EXPLOITING MULTI-FIDELITY STRATEGIES TO QUANTIFY UNCERTAINTY IN IRRADIATED PARTICLE-LADEN TURBULENT FLOW

L. Jofre[1,*], G. Geraci[2], H. R. Fairbanks[3], A. Doostan[4] and G. Iaccarino[5]

[1] Center for Turbulence Research, Stanford University, USA. jofre@stanford.edu

[2] Center for Computing Research, Sandia National Laboratories, USA. ggeraci@sandia.gov

[3] Applied Mathematics, University of Colorado, USA. hillary.fairbanks@colorado.edu

[4] Aerospace Engineering Sciences, University of Colorado, USA. alireza.doostan@colorado.edu

[5] Center for Turbulence Research, Stanford University, USA. jops@stanford.edu

Key words: Multi-fidelity Monte Carlo, Uncertainty quantification, Turbulent flow

Abstract. The study of complex physical systems is often based on computationally intensive high-fidelity simulations. To build confidence and improve the prediction accuracy of such simulations, the impact of uncertainties on the quantities of interest must be measured. This, however, requires a computational budget that is typically a large multiple of the cost of a single calculation, and thus may become unfeasible for expensive simulation models. In this regard, multi-fidelity methods have become increasingly popular in the past decade as acceleration strategies to reduce the computational cost. These approaches are based on a hierarchy of models and attempt to leverage the relation between high- and low-fidelity levels to obtain a more accurate statistical estimator without requiring a large number of additional high-fidelity calculations. In this work, three different multi-fidelity strategies are assessed for the numerical simulation of irradiated particle-laden turbulent flow relevant to solar receiver systems.

1 INTRODUCTION

The number of uncertainties involved in the study of complex, multiphysics turbulent flow is typically large due to (i) the modeling assumptions required to mathematically describe the different physics and their couplings, i.e., epistemic uncertainty, and (ii) the aleatoric incertitude resulting, for instance, from the lack of detailed evidence regarding the initial and boundary conditions. Therefore, numerical analyses based on a single deterministic realization for a particular set of input parameters cannot be deemed predictive. A solution to this problem is to consider the system under study stochastic and analyze the relation between input and output probability distributions by means of efficient statistical methods. The field of uncertainty quantification (UQ) applied to

computational sciences and engineering has remarkably grown over the last decades [1], and it is now extensively accepted that the potential of estimating and minimizing uncertainties, in combination with numerical verification and physics validation (V&V), is crucial for augmenting the confidence in the numerical predictions. Within this scope, the goal of the Predictive Science Academic Alliance Program (PSAAP) II at Stanford University [2] is to perform large-scale, predictive simulations of particle-laden turbulent flow in a radiation environment based on the configuration and operating conditions of an in-house experimental apparatus.

The investigation of thermal radiation interacting with particle-laden turbulent flow is of great importance in diverse scientific and engineering disciplines. For instance, phenomena reminiscent of irradiated particle-laden turbulence are ubiquitous in the fields of earth and combustion sciences, like for example, the impact of preferential concentration on the rate of droplet coalescence and evaporation in atmospheric clouds [3], and the fluid mechanics of reacting boundary layers and plumes [4]. Of particular interest to this work is the study of these physical processes in the context of volumetric particle-based solar receivers for energy harvesting [5].

Inertial particles in homogeneous isotropic turbulence (HIT) may exhibit complex interactions between the carrier and dispersed phases as a result of preferential concentration and turbulence modulation [6]. The former is the process by which heavy particles are expelled from intense vortical structures and concentrate in regions of the flow dominated by strain. This phenomenon is characterized by the Stokes number $St = \tau_p/\tau_f$, defined as the ratio between particle relaxation, τ_p, and fluid, τ_f, time scales, and it is accentuated for $St_\eta \approx 1$ when τ_f is taken as the Kolmogorov time scale τ_η. The latter corresponds to the modification of fluid flow characteristics in the near-field region of particle clusters as a result of two-way coupling effects. The addition of no-slip boundaries further complicates the problem as turbophoresis [7], i.e., tendency of particles to migrate towards regions of decreasing turbulence intensity, acts as a dominant mechanism for accumulating particles at the walls. These three mechanisms are well known and have been extensively studied by means of experimental and numerical approaches over the past decades, e.g., [8, 9]. However, the problem of interest in this work involves, in addition to particle-flow coupling, heat transfer from the particles to the fluid by means of radiation absorption, particle-to-fluid heat exchange, and thermal mixing enhanced by turbulent motions. The physical processes governing this type of flows, and their complex interactions, are not fully characterized, and therefore are still subject to intense research, e.g., [10].

Modeling of irradiated particle-laden turbulent flow and its numerical investigation and validation against experimental data are difficult tasks that intrinsically require several model assumptions, selection of coefficient and parameter values and characterization of initial and boundary conditions. These steps, even if performed carefully, result in potential sources of uncertainty that can impact the quantities of interest (QoI). Some examples encompass the incomplete description of particle diameters [11] and thermal radiative properties, variability of the incident radiation and its complex interaction with boundaries, and model-form incertitude [12]. In addition, accurate representation of the underlying physical phenomena mandates the utilization of expensive high-fidelity (HF)

numerical simulations. As a representative example, the total cost (transient + collection of statistics) of a HF calculation of this problem requires in the order of 5M core-hours per realization on the Argonne Leadership Computing Facility (ALCF, Mira supercomputer) [13]. Hence, characterization of the stochastic output by means of running hundreds, or thousands, of realizations with different input values easily exceeds the resources of the largest computing facilities available. The objective of this paper, therefore, is to investigate the utilization of multi-fidelity UQ strategies to efficiently characterize irradiated particle-laden turbulent flow subject to uncertainty.

Detailed descriptions on modeling and numerical discretization of irradiated particle-laden turbulent flow are given in Jofre et al. [14] and Esmaily et al. [15], respectively. Therefore, the remaining of the paper is organized as follows. First, the accelerated sampling strategies investigated are presented in Section 2. The system of interest is detailed next, in Section 3, in terms of physical parameters, uncertainties, QoIs, and models designed. Then, in Section 4, the performances of the accelerated estimators are analyzed. Finally, conclusions and future work are discussed in Section 5.

2 MONTE CARLO-TYPE ACCELERATED SAMPLING STRATEGIES

In computational science and engineering, multiple physical/mathematical/numerical models with different features can be constructed to characterize a system of interest. Typically, computationally expensive HF models are designed to describe the system with the degree of accuracy required by the problem under study, while low-fidelity (LF) models are formulated as less accurate, but relatively cheaper, representations. Outer-loop problems, such as inference, UQ and optimization, require large numbers of model evaluations for different input values, resulting in unaffordable computational requirements in the case of large-scale, multiphysics calculations. The objective of multi-fidelity methods, therefore, is to reduce the cost of the outer-loop problem by combining the accuracy of the HF models with the speedup achieved by the LF representations. Different multi-fidelity UQ strategies exist in the literature; see, for example, the exhaustive review by Peherstorfer et al. [16]. However, due to the high-dimensional input space and the complexity of the conservation equations involved, this study is restricted to a reduced subset of acceleration strategies appertaining to surrogate-based MC-type sampling approaches.

As its name indicates, MC-type approaches are derived from the original Monte Carlo method, in which the expectation of the QoI $Q = Q(\xi)$, as a function of the stochastic inputs ξ, is estimated via a sample average. Let $\mathbb{E}[Q]$ and $\mathbb{V}ar(Q)$ denote the mean and variance of Q. Given N independent realizations of the stochastic input, denoted $\xi^{(i)}$, the MC estimator of $\mathbb{E}[Q]$ is defined as $\hat{Q}_N^{\mathrm{MC}} = N^{-1} \sum_{i=1}^{N} Q^{(i)}$, where $Q^{(i)} = Q(\xi^{(i)})$. Although unbiased, the accuracy of $\hat{Q}_N^{\mathrm{MC}}$, measured by its standard deviation $\sqrt{\mathbb{V}ar(Q)/N}$, decays slowly as a function of N. Therefore, for a fixed computational budget ($\propto N$) a viable alternative to increase the MC accuracy is to possibly replace Q with other quantities with smaller variances.

3

2.1 Multi-level Monte Carlo

One of the most popular acceleration strategies is the multi-level (ML) method [17]. This technique, inspired by the multigrid solver idea in linear algebra, is based on evaluating realizations of Q from a hierarchy of models with different levels ℓ, $\ell = 0, \ldots, L$, with L the most accurate model, in which Q is replaced by the sum of differences $Y_\ell = Q_\ell - Q_{\ell-1}$, where by definition $Y_0 = Q_0$. As a result, the QoIs of the original and new ML problems have the same mean $\mathbb{E}[Q]$. An example of a level is the grid resolution considered for solving the system of equations, so that a LF (or HF) model can be established by simulating Q on a coarse (or fine) grid. Then, $\mathbb{E}[Q]$ can be computed using the ML QoI and an independent MC estimator on each level ℓ as

$$\hat{Q}^{\mathrm{ML}} = \sum_{\ell=0}^{L} \hat{Y}_\ell^{\mathrm{MC}} = \sum_{\ell=0}^{L} \frac{1}{N_\ell} \sum_{i=1}^{N_\ell} Y_\ell^{(i)}. \tag{1}$$

This approach is referred to as multi-level Monte Carlo (MLMC) and the resulting estimator has a variance equal to $\mathbb{V}ar\left(\hat{Q}^{\mathrm{ML}}\right) = \sum_{\ell=0}^{L} N_\ell^{-1} \mathbb{V}ar\left(Y_\ell\right)$. Consequently, if the level definition is such that $Q_\ell \to Q$ in mean square sense, then $\mathbb{V}ar\left(Y_\ell\right) \to 0$ as $\ell \to \infty$, and therefore fewer samples are required on the finer level L. In particular, it is possible to show that the optimal sample allocation across levels, N_ℓ, is obtained in closed form given a target variance of the MLMC estimator equal to $\varepsilon^2/2$ and resulting in [17]

$$N_\ell = \frac{\sum_{k=0}^{L} \sqrt{\mathcal{C}_k \mathbb{V}ar\left(Y_k\right)}}{\varepsilon^2/2} \sqrt{\frac{\mathbb{V}ar\left(Y_\ell\right)}{\mathcal{C}_\ell}}, \tag{2}$$

where the computational cost of the individual Y_ℓ evaluations is denoted by $\mathcal{C}_\ell$.

2.2 Control variates Monte Carlo

To accommodate LF representations that are not obtained directly from coarsening the HF models, a common approach is to utilize LF realizations as a control variate [18, 19]. This approach requires that a generic QoI, f, be replaced by $f + \alpha(g - \mathbb{E}[g])$, where g is a function chosen for its high correlation with f and for which the value of $\mathbb{E}[g]$ is readily available. However, in the problem of interest here the LF correlation and expected value are not available *a priori*, and consequently need to be established during the computations along with the HF calculations. As a consequence, the expected value of the LF model is generally approximated by means of an MC estimator requiring a set of additional (independent) LF computations. The control variates MC estimator (CV) is defined as

$$\hat{Q}^{\mathrm{CV}} = \hat{Q}^{\mathrm{HF,MC}} + \alpha\left(\hat{Q}^{\mathrm{LF,MC}} - \mathbb{E}\left[\hat{Q}^{\mathrm{LF}}\right]\right), \tag{3}$$

where the parameter $\alpha = -\mathrm{Cov}\left(Q^{\mathrm{HF}}, Q^{\mathrm{LF}}\right)/\mathbb{V}ar\left(Q^{\mathrm{LF}}\right)$ is chosen to minimize the variance of $\hat{Q}^{\mathrm{CV}}$ by requiring $\mathrm{d}\mathbb{V}ar\left(\hat{Q}^{\mathrm{CV}}\right)/\mathrm{d}\alpha = 0$. The optimal α selection leads to

$$\mathbb{V}ar\left(\hat{Q}^{\mathrm{CV}}\right) = \mathbb{V}ar\left(\hat{Q}^{\mathrm{HF,MC}}\right)\left(1 - \frac{r}{1+r}\rho^2\right), \tag{4}$$

where $\rho = \text{Cov}\left(Q^{\text{HF}}, Q^{\text{LF}}\right)/\sqrt{\mathbb{V}ar\left(Q^{\text{HF}}\right)\mathbb{V}ar\left(Q^{\text{LF}}\right)}$ is the Pearson correlation coefficient between the HF and the LF models, and $r = N^{\text{LF}}/N^{\text{HF}} - 1$ is used to parameterize the additional rN^{HF} LF realizations needed in order to evaluate

$$\mathbb{E}\left[\hat{Q}^{\text{LF}}\right] \approx \frac{1}{N^{\text{HF}}(1+r)} \sum_{i=1}^{N^{\text{HF}}(1+r)} Q^{\text{LF},(i)}. \tag{5}$$

As a result, the optimal control variate is obtained for a particular r value that, in turn, depends on the correlation between the two models and their cost ratio. As fully detailed in [20], the optimal r value is obtained from

$$r = -1 + \sqrt{\frac{\mathcal{C}^{\text{HF}}}{\mathcal{C}^{\text{LF}}} \frac{\rho^2}{1 - \rho^2}}, \tag{6}$$

with $\mathcal{C}^{\text{HF}}$ and $\mathcal{C}^{\text{LF}}$ the computational cost of individual HF and LF samples.

2.3 Multi-level multi-fidelity Monte Carlo

Given the flexible nature of the MC-type sampling methods, the ML and CV approaches can be hybridized in different fashions to generate a class of estimators generally referred to as multi-level multi-fidelity (MLMF), e.g., [20, 21]. For example, this work considers the MLMF estimator introduced in [20], which is constructed by taking the difference Y_ℓ to be the object of an additional control variate. Similar to the two previous estimators, the central idea is to use the HF level to minimize the deterministic bias, while the LF models are exploited to speedup the process of reducing the estimator's variance.

Before presenting the estimator, some definitions applicable to the scope of this subsection are introduced to simplify the subsequent exposition: (i) subscript ℓ refers to the different ML levels, and (ii) $Y_\ell = Q_\ell - Q_{\ell-1}$ with $Y_0 = Q_0$; (iii) at each level ℓ, the reference fidelity is identified by superscript RF, whereas CV indicates control variate realizations. Following this terminology, the MLMF is formulated as

$$\hat{Q}^{\text{MLMF}} = \sum_{\ell=0}^{L}\left[\frac{1}{N_\ell^{\text{RF}}}\sum_{i=1}^{N_\ell^{\text{RF}}}Y_\ell^{\text{RF},(i)} + \frac{\alpha_\ell}{N_\ell^{\text{RF}}\left(1+r_\ell\right)}\left(r_\ell\sum_{i=1}^{N_\ell^{\text{RF}}}Y_\ell^{\text{CV},(i)} - \sum_{j=1}^{r_\ell N_\ell^{\text{RF}}}Y_\ell^{\text{CV},(j)}\right)\right], \tag{7}$$

where the vector $r_\ell = N_\ell^{\text{CV}}/N_\ell^{\text{RF}} - 1$ is used to characterize the supplementary $r_\ell N_\ell^{\text{RF}}$ CV realizations per level ℓ, and the optimal parameters $\alpha_\ell = -\rho_\ell\sqrt{\mathbb{V}ar\left(Y_\ell^{\text{RF}}\right)}/\sqrt{\mathbb{V}ar\left(Y_\ell^{\text{CV}}\right)}$ are obtained similarly to the standard control variates approach with ρ_ℓ — a level-based Pearson correlation coefficient — defined as

$$\rho_\ell = \frac{\sum_{i=1}^{N_\ell^{\text{RF}}}\left(Y_\ell^{\text{RF},(i)} - \mathbb{E}\left[Y_\ell^{\text{RF}}\right]\right)\left(Y_\ell^{\text{CV},(i)} - \mathbb{E}\left[Y_\ell^{\text{CV}}\right]\right)}{\sqrt{\sum_{i=1}^{N_\ell^{\text{RF}}}\left(Y_\ell^{\text{RF},(i)} - \mathbb{E}\left[Y_\ell^{\text{RF}}\right]\right)^2\sum_{i=1}^{N_\ell^{\text{RF}}}\left(Y_\ell^{\text{CV},(i)} - \mathbb{E}\left[Y_\ell^{\text{CV}}\right]\right)^2}}. \tag{8}$$

Table 1: List of random inputs. All inputs are assumed to be independent and uniformly distributed.

Variable	Interval	Variable	Interval
1. Prt. rest. coeff. 1	$[0.0 : 0.6]$	8. Mass load. ratio	$[18 : 22]\%$
2. Prt. rest. coeff. 2	$[0.1 : 0.7]$	9. Prt. abs. eff.	$[0.37 : 0.41]$
3. Prt. rest. coeff. 3	$[0.2 : 0.8]$	10. Prt. scatt. eff.	$[0.69 : 0.76]$
4. Prt. rest. coeff. 4	$[0.3 : 0.9]$	11. Radiation	$[1.8 : 2.0]$ MW/m^2
5. Prt. rest. coeff. 5	$[0.4 : 1.0]$	12. Radiated wall	$[1.6 : 6.4]$ kW/m^2
6. Stokes' drag corr.	$[1.0 : 1.5]$	13. Opposite wall	$[1.2 : 4.7]$ kW/m^2
7. Prt. Nusselt num.	$[1.5 : 2.5]$	14. Side x-y walls	$[0.1 : 0.2]$ kW/m^2

The $\hat{Q}^{\mathrm{MLMF}}$ estimator is unbiased since each term is obtained from an unbiased MC estimator, and its variance, resulting from considering the contribution of each Y_ℓ term plus the covariance between the RF and CV estimators, is given by

$$\mathbb{V}ar\left(\hat{Q}^{\mathrm{MLMF}}\right) = \sum_{\ell=0}^{L} \frac{1}{N_\ell^{\mathrm{RF}}} \mathbb{V}ar\left(Y_\ell^{\mathrm{RF}}\right)\left(1 - \frac{r_\ell}{1+r_\ell}\rho_\ell^2\right). \tag{9}$$

3 DESCRIPTION OF THE SYSTEM OF INTEREST

The system studied in this work is based on a small-scale, experimental apparatus designed to represent the physical mechanisms encountered in volumetric particle-based solar receivers. The main physical parameters of the problem, and their values, are bulk Reynolds number $Re_b = 20000$, Kolmogorov Stokes numbers $5 < St_\eta = \tau_p/\tau_\eta < 20$, total mass loading ratio $\mathrm{MLR} = n_p m_p/\rho_f \approx 20\%$, and radiation power $P \approx 1000$ W. Complete details of the experiment configuration and computational setup are found in [22] and [14], respectively. The subsections below describe (i) the uncertainties and QoIs considered, and (ii) the high- and low-fidelity models constructed to perform the UQ analysis.

3.1 Uncertainties and quantities of interest

This work considers 14 stochastic variables in the UQ analysis. As shown in Table 1, the variables target experiment and model-form uncertainties. These correspond to incertitude in particle restitution coefficient for the different classes $(1 - 5)$, correction to Stokes' drag law (6), particle Nusselt number (7), mass loading ratio (8), particle absorption and scattering efficiencies $(9, 10)$, incident radiation flux (11), and heat fluxes from the walls to the fluid $(12 - 14)$. The performance of the estimators is analyzed by focusing on time-averaged, normalized temperature increments, i.e., $Q = (\langle T \rangle - T_0)/T_0 = \Delta T/T_0$, at the probe location as these quantities are available from the experiments; $T_0 = 300$ K is the temperature of the inflow particle-fluid mixture.

The intervals of the stochastic variables listed in Table 1 have been carefully characterized on the basis of information from the experiments, and by taking into consideration results and conclusions extracted from published studies; see Jofre et al. [14] for complete details.
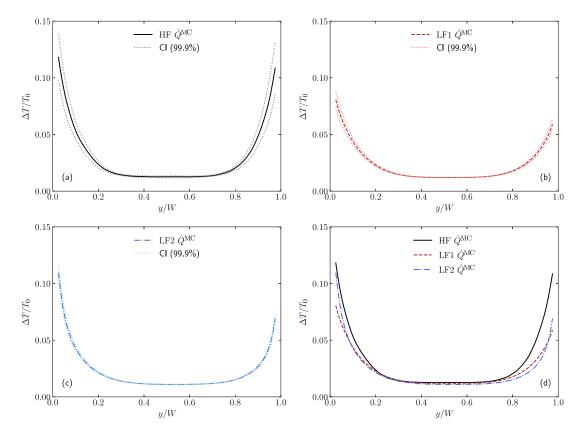
Figure 1: Prediction and CI of the individual, pilot HF (a), LF1 (b) and LF2 (c) MC estimators for $Q = \Delta T/T_0$ profile at probe location. Comparison between MC estimators (d).

3.2 High- and low-fidelity levels designed

As described in [23], the Lagrangian-Eulerian (LE) approach is flexible enough to support different degrees of particle-laden flow modeling. This feature is leveraged in this work to construct one HF model and two LF representations, denoted LF1 and LF2.

The HF level corresponds to a point particle direct numerical simulation (PP-DNS) with physical particles and sufficient resolution ($\approx$ 55M cells/section) to capture all the significant turbulent scales, while approximating the particles as Lagrangian inertial points ($\approx$ 15M particles/section). The flow grid is uniform in the streamwise direction with spacings in wall units equal to $\Delta x^+ \approx 12$, while stretched in the wall-normal directions with the first grid point at $y^+, z^+ \approx 0.5$ and with resolutions in the range $0.5 < \Delta y^+, \Delta z^+ < 6$. The radiative heat transfer equation is solved on an uniform discrete ordinates method (DOM) mesh of $270 \times 160 \times 160$ gridpoints with 350 quadrature points (discrete angles).

The LF representations are designed following the surrogate particle approach [23] together with a coarsening of the flow and radiation HF resolutions. The statistical weights (uniform) of the surrogate particles for the LF1 and LF2 levels are 5 and 10, respectively. Similarly, the flow and radiation meshes, and quadrature points are uniformly coarsen in each direction by a factor of 5 (LF1) and 10 (LF2). As a result, the LF1 and LF2 levels are approximately $170\times$ and $1300\times$ cheaper per sample than HF, respectively.
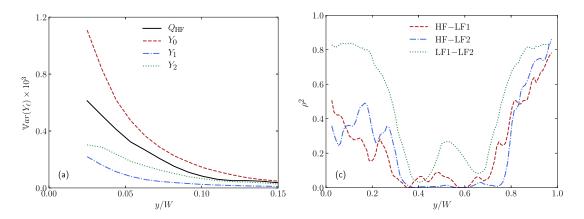
Figure 2: Variance of Y_ℓ across levels (a) and squared correlation coefficient between fidelities (b) for $Q = \Delta T/T_0$ profile at probe location based on pilot samples.

4 PERFORMANCE ANALYSIS OF THE ESTIMATORS

The first step toward efficiently investigating the impact of uncertainty on heat transfer mechanisms is to analyze the performance of the estimators in terms of speedup and quality of their predictions. This section is devoted to this objective by (i) carefully analyzing the features of the individual HF, LF1 and LF2 MC estimators and their interdependences, and (ii) calculating the optimal distribution of samples for the ML, CV and MLMF estimators and quantifying their coefficient of variation (CoV) and comparison against experimental data.

4.1 Features of the Monte Carlo estimators based on pilot samples

The calculation of the optimal allocation of samples for the ML, CV and MLMF estimators is based on the properties of $\mathbb{V}ar\,(Y_\ell)$ and ρ between the different levels. However, these quantities are not available *a priori*, and therefore the typical strategy is to evaluate them based on a small, initial set of realizations referred to as pilot samples. Following this approach, 16 HF, 128 LF1 and 256 LF2 realizations are initially computed and their features used to extrapolate the optimal number of samples. Their MC estimations for $Q = \Delta T/T_0$ profile at probe location are depicted in Figure 1.

Four main observations can be extracted from the plots in the figure. First, the shape of the temperature profile indicates that most of the radiation absorption is produced close to the walls due to the large concentration of particles in those regions as a result of turbophoretic effects. The second observation is that there is remarkably more variability toward the walls, viz. $0.2 > y/W > 0.8$, than at the centerline. A clear indication of this feature is, for example, the larger width of the confidence interval (CI) for $\hat{Q}_{HF}^{MC}$ at $y/W \approx 0$ and 1. Third, as indicated by the CI, the estimation is significantly converged for $\hat{Q}_{LF2}^{MC}$, but not for $\hat{Q}_{LF1}^{MC}$ nor (specially) $\hat{Q}_{HF}^{MC}$. Finally, the fourth observation is that the LF levels approximate reasonably well the HF prediction except for the near-wall regions where the discrepancy becomes substantial.
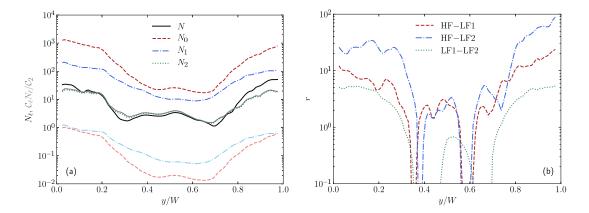
8

Figure 3: Optimal ($\epsilon \approx 0.05\,\mathbb{E}\,[Q]$) number of samples per level (opaque lines) and equivalent HF realizations (semi-opaque lines) (a), and ratio r (b) for $Q = \Delta T/T_0$ profile at probe location based on pilot samples.

4.2 Variance across levels and correlation between fidelities

The variance across levels, $\mathbb{V}ar\,(Y_\ell)$, and squared correlation coefficient, ρ^2, for $Q = \Delta T/T_0$ profile at probe location based on pilot samples are shown in Figure 2. The colored dashed lines in plot (a) correspond to $Y_0 = \mathbb{V}ar\,(Q_{\mathrm{LF2}})$, $Y_1 = \mathbb{V}ar\,(Q_{\mathrm{LF1}} - Q_{\mathrm{LF2}})$ and $Y_2 = \mathbb{V}ar\,(Q_{\mathrm{HF}} - Q_{\mathrm{LF1}})$, whereas the black solid line represents $\mathbb{V}ar\,(Q_{\mathrm{HF}})$. In plot (b), the colored dashed line indicates the squared correlation coefficient for the control variates constructed using HF and LF levels (HF $-$ LF1 and HF $-$ LF2), and utilizing LF1 and LF2 levels (LF1 $-$ LF2).

The results in plot (a) show that there is variance decay from Y_0 to Y_1, but not between Y_1 and Y_2. Even so, speedup would be still obtained with respect to the plain $\hat{Q}_{\mathrm{HF}}^{\mathrm{MC}}$ estimator by using the ML strategy as $\mathbb{V}ar\,(Q_{\mathrm{HF}})/\mathbb{V}ar\,(Y_2) \approx 2$ for the points in the profile with largest variability. In terms of correlation between the HF and LF levels, plot (b) shows that ρ^2 is very small at the centerline points, where accelerating the estimators is not important since the variability is very low. However, it increases to $\rho^2 \approx 0.5$ and $\rho^2 \approx 0.8$ for $y/W < 0.2$ and $y/W > 0.8$, respectively. With respect to the HF$-$LF results, the correlation between the LF levels is larger along the entire profile, with values $\rho^2 \approx 0.8$ for regions close to the walls, indicating that the construction of a MLMF estimator can be beneficial.

4.3 Allocation of samples for optimal performance of the estimators

The values of $\mathbb{V}ar\,(Y_\ell)$ and ρ based on the pilot samples can be used to obtain an extrapolated number of samples per level for the construction of estimators with optimal performance. The expressions to calculate these numbers are given in Eqs. 2 and 6. As detailed in Section 3.2, setting the computational cost per sample to be $\mathcal{C}_{\mathrm{HF}} = 1$, $\mathcal{C}_{\mathrm{LF1}} = 1/170$ and $\mathcal{C}_{\mathrm{LF2}} = 1/1300$, the optimal number of samples per level, N_ℓ, and ratio of additional samples, r, for $Q = \Delta T/T_0$ profile at probe location with an accuracy of 5% of the estimated QoI, i.e., $\epsilon \approx 0.05\,\mathbb{E}\,[Q]$, are shown in Figure 3. Similarly as in

9

Table 2: Optimal samples allocation computational cost for the MC, ML, CV and MLMF estimators.

Model	# HF runs	# LF1 runs	# LF2 runs	# Equiv. HF cost
MC HF	51	-	-	51
ML HF-LF1-LF2	22	199	1279	25
CV HF-LF1	22	217	-	24
CV HF-LF2	26	-	541	27
MLMF HF-LF1-LF2	21	92	528	22

Figure 2, the colored dashed lines in plot (a) indicate N_0, N_1 and N_2, while the black solid line corresponds to N for a straight $\hat{Q}_{\mathrm{HF}}^{\mathrm{MC}}$ with same ϵ. The additional semi-opaque lines represent the number of equivalent HF realizations for the N_0 LF2 (dashed red), N_1 LF1 (dashed-dotted blue) and total (solid gray) number of samples. The colored dashed lines in plot (b) indicate r for CV estimators based on HF − LF1 and HF − LF2, and a MLMF estimator utilizing HF − LF1 as a ML and LF1 − LF2 as a CV.

The results shown in the plots are summarized in Table 2 by indicating the optimal number of HF, LF1, and LF2 realizations required to construct MC, ML, CV and MLMF estimators with $\epsilon \approx 0.05\,\mathbb{E}\,[Q]$ for the point in the profile with highest computational cost. The point is different for each estimator, but, as inferred from the plots, it is located close to the walls for all cases. The table also summarizes the total computational cost, given in terms of equivalent HF realizations, for the different estimators and level combinations. As a reference, the utilization of a plain $\hat{Q}_{\mathrm{HF}}^{\mathrm{MC}}$ estimator would necessitate 51 HF realizations. Instead, using acceleration strategies, equivalent estimator predictions can be obtained at less than half the cost. It is important to note that the speedup obtained with these approaches depends on the problem and QoI. Correspondingly, the $\hat{Q}^{\mathrm{ML}}$ estimator requires 22 HF, 199 LF1 and 1279 LF2 realizations resulting in a total equivalent cost of 25 HF runs. The $\hat{Q}^{\mathrm{CV}}$ estimators demand 24 and 27 equivalent HF realizations, and the $\hat{Q}^{\mathrm{MLMF}}$ estimator obtains a similar prediction by means of 22 equivalent HF runs. From a computational cost point of view, the four acceleration strategies perform comparably with the $\hat{Q}^{\mathrm{MLMF}}$ estimator achieving the lowest value. Therefore, the subsequent analysis is based on the prediction given by the optimal $\hat{Q}^{\mathrm{MLMF}}$ estimator.

4.4 Quality of the prediction and validation against experimental data

The final step of the analysis is to verify and validate the predictions. This is portrayed in Figure 4, where the quality of the optimal $\hat{Q}^{\mathrm{MLMF}}$ estimator is shown in terms of CoV and comparison against experimental data. The solid black line in plot (a) represents the CoV of the MC estimator based on the 16 pilot HF realizations which depicts a maximum value of CoV $\approx 10\%$ at $y/W \approx 1$. In comparison, the red dashed line indicates that the maximum CoV reduces to less than 5% when utilizing the optimal $\hat{Q}^{\mathrm{MLMF}}$. Moreover, as shown in plot (b), the prediction obtained by the optimal $\hat{Q}^{\mathrm{MLMF}}$ is significantly converged as demonstrated by the collapse of the CI interval into a single line, and the experimental data is found within the distribution ($\pm$ standard deviation) of the estimated QoI.
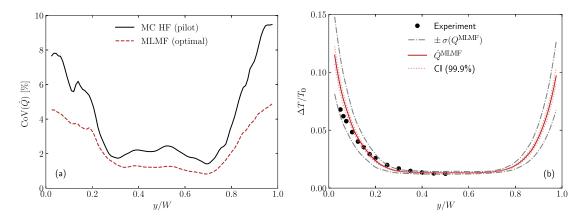
Figure 4: Quality of the MC (pilot) and MLMF (optimal) estimators measured in terms of CoV (a) and comparison against experimental data (b) for $Q = \Delta T/T_0$ profile at probe location.

5   CONCLUSIONS

Performing UQ studies of large-scale, multiphysics applications is challenging due to the expensive HF calculations required and the large number of uncertainties typically encountered. For instance, the utilization of a straightforward MC approach for the problem under study would require $\mathcal{O}(10^2)$ million core-hours in some of the most advanced supercomputers, while, as shown in Section 4, equivalent predictions can be obtained at significant cost saving ratios by means of multi-fidelity strategies.

Ongoing work is focused on leveraging the speedup obtained from the acceleration strategies to characterize heat transfer mechanisms subject to uncertainty for this problem. Future work will tackle the utilization, and further performance improvement, of this type of strategies on problems with higher particle mass and thermal loadings.

References

[1] H. N. Najm. Uncertainty quantification and polynomial chaos techniques in computational fluid dynamics. *Annu. Rev. Fluid Mech.*, 41:35–52, 2009.

[2] Exascale Computing Engineering Center. Predictive Science Academic Alliance Program (PSAAP) II, Stanford University, 2018.

[3] R. A. Shaw. Particle-turbulence interactions in atmospheric clouds. *Annu. Rev. Fluid Mech.*, 35:183–227, 2003.

[4] S. R. Tieszen. On the fluid mechanics of fires. *Annu. Rev. Fluid Mech.*, 33:67–92, 2001.

[5] C. K. Ho. Advances in central receivers for concentrating solar applications. *Sol. Energy*, 152:38–56, 2017.

[6] S. Balachandar and J. K. Eaton. Turbulent dispersed multiphase flow. *Annu. Rev. Fluid Mech.*, 42:111–133, 2010.

[7] M. Caporaloni, F. Tampieri, F. Trombetti, and O. Vittori. Transfer of particles in nonisotropic air turbulence. *J. Atmos. Sci.*, 32:565–568, 1975.

[8] K. D. Squires and J. K. Eaton. Preferential concentration of particles by turbulence. *Phys. Fluids A*, 3:1169–1178, 1991.

[9] S. Elghobashi. On predicting particle-laden turbulent flows. *Appl. Sci. Res.*, 52:309–329, 1994.

[10] A. Frankel, H. Pouransari, F. Coletti, and A. Mani. Settling of heated particles in homogeneous turbulence. *J. Fluid Mech.*, 792:869–893, 2016.

[11] M. Rahmani, G. Geraci, G. Iaccarino, and A. Mani. Effects of particle polydispersity on radiative heat transfer in particle-laden turbulent flows. *Int. J. Multiph. Flow*, Accepted, 2018.

[12] L. Jofre, S. P. Domino, and G. Iaccarino. A framework for characterizing structural uncertainty in large-eddy simulation closures. *Flow Turbul. Combust.*, 100(2):341–363, 2018.

[13] Mira, Argonne Leadership Computing Facility, 2018.

[14] L. Jofre, G. Geraci, H. R. Fairbanks, A. Doostan, and G. Iaccarino. Multi-fidelity uncertainty quantification of irradiated particle-laden turbulence. *CTR Annu. Res. Briefs*, pages 21–34, 2017.

[15] M. Esmaily, L. Jofre, A. Mani, and G. Iaccarino. A scalable geometric multigrid solver for nonsymmetric elliptic systems with application to variable-density flows. *J. Comput. Phys.*, 357:142–158, 2018.

[16] B. Peherstorfer, K. Willcox, and M. Gunzburger. Survey of multifidelity methods in uncertainty propagation, inference, and optimization. *ACDL-MIT, TR-16-1*, 2016.

[17] M. B. Giles. Multi-level Monte Carlo path simulation. *Oper. Res.*, 56:607–617, 2008.

[18] R. Pasupathy, M. Taaffe, B. W. Schmeiser, and W. Wang. Control-variate estimation using estimated control means. *IIE Trans.*, 44:381–385, 2014.

[19] L. W. T. Ng and K. Willcox. Multifidelity approaches for optimization under uncertainty. *Int. J. Numer. Meth. Eng.*, 100:746–772, 2014.

[20] G. Geraci, M. Eldred, and G. Iaccarino. A multifidelity multilevel monte carlo method for uncertainty propagation in aerospace applications. *19th AIAA Non-Deterministic Approaches Conference*, pages AIAA 2017–1951, 2017.

[21] H. R. Fairbanks, A. Doostan, C. Ketelsen, and G. Iaccarino. A low-rank control variate for multilevel Monte Carlo simulation of high-dimensional uncertain systems. *J. Comput. Phys.*, 341:121–139, 2017.

[22] L. Villafañe, A. Banko, C. Elkins, and J. K. Eaton. Gas heating by radiation absorbing inertial particles in a turbulent duct flow. *CTR Annu. Res. Briefs*, pages 35–47, 2017.

[23] S. Subramaniam. Lagrangian-Eulerian methods for multiphase flows. *Prog. Energy Combust. Sci.*, 39:215–245, 2013.