MULTIFIDELITY MODELING OF IRRADIATED PARTICLE-LADEN TURBULENCE SUBJECT TO UNCERTAINTY

L. Jofre,[1,2,*] *M. Papadakis,*[3] *P. T. Roy,*[4] *A. Aiken,*[3] *& G. Iaccarino*[1]

[1]*Center for Turbulence Research, Stanford University, Stanford, CA 94305, USA*

[2]*Dept. Fluid Mechanics, Technical University of Catalonia - BarcelonaTech, Barcelona 08019, Spain*

[3]*Stanford Computer Science, Stanford University, Stanford, CA 94305, USA*

[4]*Bongfish, Am Eisernen Tor 3, 8010 Graz, Austria*

Address all correspondence to: L. Jofre, Dept. Fluid Mechanics, Technical University of Catalonia - BarcelonaTech, Barcelona 08019, Spain, E-mail: lluis.jofre@upc.edu

The study of thermal radiation interacting with particle-laden turbulence is of great importance in a wide range of scientific and engineering applications. The computational study of such systems is challenging as a result of the large number of thermo-fluid mechanisms governing the underlying physics. 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 infeasible for expensive simulation models featuring a large number of uncertain inputs and highly non-linear behavior. In this regard, multifidelity methods have become increasingly popular in the past years as acceleration strategies to reduce the computational cost. These methods are based on a hierarchy of generalized numerical resolutions, or model fidelities, and attempt to leverage the correlation between high- and low-fidelity models to obtain a more accurate statistical estimator with a relatively small number of high-fidelity calculations. In this work, the performance of a collection of different multifidelity strategies and modeling approaches are assessed to propagate the uncertainties encountered in the simulation of irradiated particle-laden turbulence relevant to volumetric solar energy receivers. The results obtained indicate that multifidelity methods provide speedups in the order of $10^1 - 10^3 \times$ with respect to straightforward Monte Carlo approaches, resulting in remarkable reductions in computational cost.

KEY WORDS: *Multifidelity Monte Carlo; Particle-laden flow; Thermal radiation; Turbulence; Uncertainty quantification*

1. INTRODUCTION

The number of uncertainties involved in the study of 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 is typically not conclusive and neither truly 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. In this regard, the field of uncertainty quantification (UQ) 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 configurations and operating conditions of volumetric solar receivers.

1.1 Irradiated particle-laden turbulence

Turbulent flow laden with inertial particles exposed to thermal radiation is encountered in a wide range of natural phenomena and industrial applications. For instance, it is well established that turbulence-driven particle inhomogeneities play a fundamental role in determining the rate of droplet coalescence and evaporation in atmospheric clouds [3]. This problem is also important in fires due to the strong interaction between turbulence, soot particles, and radiation that result in complex heat transfer mechanisms [4]. From an industrial viewpoint, important applications include the atomization of liquid fuels in combustors [5,6], soot formation in rocket engines [7], and, more recently, volumetric particle-based concentrated solar power (CSP) systems [8].

Even in the most elementary configuration, i.e., homogeneous isotropic turbulence (HIT), particle-laden turbulent flow is known to exhibit complex interactions between the carrier and dispersed phases in the form of preferential concentration and turbulence modulation. Preferential concentration is the mechanism by which heavy particles tend to avoid intense vortical motions and accumulate in regions of high strain rate. On the other hand, turbulence modulation refers to the alteration of fluid flow characteristics in the near-field region of particle clusters as a result of two-way coupling effects, e.g., enhanced dissipation, kinetic energy transfer, or formation of wakes. The physical complexity is further increased by the simple addition of solid walls as turbophoresis [9], i.e., tendency of particles to migrate towards regions of decreasing turbulence levels, becomes an important mechanism for augmenting the inhomogeneity in the spatial distribution of the dispersed phase by driving particle accumulation at the walls.

Consequently, the analysis and characterization of particle-laden turbulent flow is a challenging endeavor; many experimental and computational research studies have been devoted to this objective over the past decades, e.g., [9–11]. In addition to particle-flow coupling, the problem studied in this work involves an additional layer of complexity by considering the heat transfer from the particles to the fluid through radiation absorption. The engineering application motivating the study of these phenomena is the improvement of energy harvesting in volumetric particle-based solar receivers. This innovative technology is expected to increase the performance of CSP plants by avoiding the necessity of heat-exchanging stages. However, the physical mechanisms governing these systems are still not fully comprehended. Examples of recent work focusing on this problem include the interaction between radiation, particles and buoyancy in HIT by Zamansky et al. [12], the impact of heating on the settling of particles by Frankel et al. [13], the effect of Stokes number and polydispersity on particle-gas heat transfer rates by Pouransari & Mani [14] and Rahmani et al. [15], and the quantification of uncertainties and sensitivity/dimensional analysis of complex irradiated particle-laden turbulent flow by Jofre & Fairbanks et al. [16–18].

1.2 Objectives and organization of the work

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 [15] and thermal radiative properties, variability of the incident radiation and its complex interaction with boundaries, and model-form incertitude [19,20]. In addition, accurate representation of the underlying physical phenomena mandates the utilization of expensive high-fidelity (HF) numerical simulations. 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 work, therefore, is to investigate the utilization of multifidelity UQ strategies to efficiently characterize irradiated particle-laden turbulent flow subject to uncertainty.

The paper is organized as follows. The accelerated sampling strategies considered are presented first in Section 2. Next, the physics modeling and computational approach utilized to study irradiated particle-laden turbulence are

described in Section 3, and the system of interest is detailed in terms of setup, uncertainties, QoIs, and models designed in Section 4. Then, in Section 5, the performances of the accelerated estimators are analyzed. Finally, conclusions and future work are discussed in Section 6.

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 multifidelity (MF) 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 MF UQ strategies exist in the literature; see, for example, the exhaustive review by Peherstorfer et al. [21]. 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 Monte Carlo (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 $\mathbb{V}[Q]$ 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 $\sqrt{\mathbb{V}[Q]/N}$, 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.

2.1 Multilevel Monte Carlo

One of the most popular acceleration strategies is the multilevel (ML) method [22]. 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 multilevel Monte Carlo (MLMC), or simply ML, and the resulting estimator has a variance equal to $\mathbb{V}\left[\hat{Q}^{\mathrm{ML}}\right] = \sum_{\ell=0}^{L} N_\ell^{-1} \mathbb{V}[Y_\ell]$. Consequently, if the level definition is such that $Q_\ell \to Q$ in mean square sense, then $\mathbb{V}[Y_\ell] \to 0$ 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 ML estimator equal to $\varepsilon^2/2$ and resulting in [22]

$$N_\ell = \frac{\sum_{k=0}^{L} \sqrt{\mathcal{C}_k \mathbb{V}[Y_k]}}{\varepsilon^2/2} \sqrt{\frac{\mathbb{V}[Y_\ell]}{\mathcal{C}_\ell}}, \tag{2}$$

where the computational cost of the individual Y_ℓ evaluations is denoted by $\mathcal{C}_\ell$, and ε^2 represents the mean square error (MSE) of the estimator. It is important to note that the variance decay can be proven to be satisfied only for levels based on a numerical discretization (spatial/temporal meshes) and not for general hierarchies of models, such as 2-D versus 1-D, large-eddy simulation versus Reynolds-Averaged Navier Stokes, etc.

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 [21,23,24]. In statistics, the control variates approach replaces a generic quantity q by $q + \alpha(\mathbb{E}\,[g] - g)$, where g is a function chosen for its high correlation with q and for which the value of $\mathbb{E}\,[g]$ is readily available. However, in the problem of interest here the LF correlations and expected values are not available *a priori*, and consequently need to be established during the computations along with the HF calculations. As a consequence, the expected values of the LF models are generally approximated by means of MC estimators requiring a set of additional (independent) LF computations. In mathematical form, if N_0 and N_f are, respectively, the number of HF and LF $f = 1, \ldots F$ model evaluations, with $0 < N_0 \leq N_1 \leq \cdots \leq N_F$ and associated computational costs $\mathbf{c} = [\mathcal{C}_0, \mathcal{C}_1 \ldots, \mathcal{C}_F]^\intercal$, then the control variates MC estimator (CV) is defined as

$$\hat{Q}^{\mathrm{CV}} = \hat{Q}^{\mathrm{MC}}_{\mathrm{HF}} + \sum_{f=1}^{F} \alpha_f \left(\mathbb{E}\,[Q_f] - \hat{Q}^{\mathrm{MC}}_f \right), \tag{3}$$

where $\hat{Q}^{\mathrm{MC}}_{\mathrm{HF}} = N_0^{-1} \sum_{i=1}^{N_0} Q^{(i)}_{\mathrm{HF}}$, $\mathbb{E}\,[Q_f] \approx (N_f - N_0 + 1)^{-1} \sum_{i=N_0+1}^{N_f} Q^{(i)}_f$, $\hat{Q}^{\mathrm{MC}}_f = N_0^{-1} \sum_{i=1}^{N_0} Q^{(i)}_f$, and $\alpha_f = \mathrm{Cov}\,[Q_{\mathrm{HF}}, Q_f] / \mathbb{V}\,[Q_f]$ are the control variate coefficients chosen to minimize the variance of $\hat{Q}^{\mathrm{CV}}$. The optimal α_f selection leads to

$$\mathbb{V}\left[\hat{Q}^{\mathrm{CV}}\right] = \frac{1}{N_0} \mathbb{V}\,[Q_{\mathrm{HF}}] + \sum_{f=1}^{F} \left(\frac{1}{N_0} - \frac{1}{N_f - N_0 + 1} \right) \left(\alpha_f^2 \mathbb{V}\,[Q_f] - 2\alpha_f \rho_f \sqrt{\mathbb{V}\,[Q_{\mathrm{HF}}] \,\mathbb{V}\,[Q_f]} \right), \tag{4}$$

with $-1 \leq \rho_f = \mathrm{Cov}\,[Q_{\mathrm{HF}}, Q_f] / \sqrt{\mathbb{V}\,[Q_{\mathrm{HF}}] \,\mathbb{V}\,[Q_f]} \leq 1$ the Pearson correlation coefficient between the HF and LF models, resulting from considering the optimal number of evaluations $N_0^*, N_f^*, \ldots N_F^*$ calculated as

$$N_0^* = \frac{\gamma}{\mathbf{c}^\intercal \mathbf{r}} \quad \text{and} \quad N_f^* = r_f N_0^* \quad \text{for} \quad f = 1, \ldots F, \tag{5}$$

where γ is the total computational budget, and the components of the vector $\mathbf{r} = [1, r_1, \ldots r_F]^\intercal$ are given by

$$r_f = \sqrt{\frac{\mathcal{C}_0(\rho_f^2 - \rho_{f+1}^2)}{\mathcal{C}_f(1 - \rho_f^2)}} \quad \text{for} \quad f = 1, \ldots F, \tag{6}$$

with $\rho_{F+1} = 0$. A comprehensive description of the control variates MC estimator and derivation of optimal coefficients and number of samples per fidelity is detailed in Peherstorfer et al. [21].

2.3 Multilevel multifidelity 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 multilevel multifidelity (MLMF), e.g., [24,25]. For example, this work considers the MLMF estimator introduced in [24], 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 statistical error, 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

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

where the vector $r_\ell = \sqrt{\dfrac{\rho_\ell^2}{1 - \rho_\ell^2} \dfrac{\mathcal{C}_\ell^{\mathrm{RF}}}{\mathcal{C}_\ell^{\mathrm{CV}}}} - 1$ is used to calculate the optimal supplementary $r_\ell N_\ell^{\mathrm{RF}}$ CV realizations per level ℓ, and the optimal parameters $\alpha_\ell = -\rho_\ell \sqrt{\mathbb{V}\left[Y_\ell^{\mathrm{RF}}\right]} / \sqrt{\mathbb{V}\left[Y_\ell^{\mathrm{CV}}\right]}$ are obtained similarly to the standard control variates approach with ρ_ℓ (a level-based Pearson correlation coefficient) defined as

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

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

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

where the optimal number of RF samples per level is obtained as [24]

$$N_\ell^{\mathrm{RF}} = \frac{2}{\varepsilon^2} \left[\sum_{k=0}^{L} \sqrt{\frac{\mathbb{V}\left[Y_k^{\mathrm{RF}}\right] \mathcal{C}_k^{\mathrm{RF}}}{1 - \rho_\ell^2}} \left(1 - \frac{r_\ell}{1 + r_\ell} \rho_\ell^2\right)\right] \sqrt{\left(1 - \rho_\ell^2\right) \frac{\mathbb{V}\left[Y_\ell^{\mathrm{RF}}\right]}{\mathcal{C}_\ell^{\mathrm{RF}}}}. \tag{10}$$

3. MODELING OF IRRADIATED PARTICLE-LADEN TURBULENT FLOW

The mathematical representation of dispersed multiphase flow can be separated into two main formulations based on the strategy chosen to describe the particle and fluid phases: Eulerian-Eulerian (EE) and Lagrangian-Eulerian (LE). The EE type is a random field approach in which both dispersed and carrier phases are represented as random fields in the Eulerian frame, whereas the LE class is a point-based formulation in which the dispersed phase is represented as a stochastic process in a Lagrangian setting and the carrier phase is represented as a random field in the Eulerian frame. The scope of this work is restricted to LE schemes as some of its properties are advantageous for the application of interest, e.g., applicability to a wide range of mass loading fractions and Stokes numbers, and ability to accurately represent particle-particle/wall collisions. The LE approach can be further decomposed into different subcategories (see the review by Subramaniam [26] for a detailed exposition) corresponding to: (i) direct numerical simulation (DNS) with fully-resolved particles, (ii) point particle (PP) DNS with physical particles, (iii) PP-DNS with stochastic (parcel) particles, (iv) large-eddy simulation (LES) with physical or stochastic particles, and (v) Reynolds-averaged Navier-Stokes (RANS).

3.1 Point particle direct numerical simulation

Within the realm of LE formulations, DNS with fully-resolved particles is the modeling approach with highest possible fidelity as all the turbulent flow features in the bulk of the fluid, i.e., integral to Kolmogorov scales, and near the particles, e.g., boundary layers and wakes, are resolved. This extreme level of resolution, however, comes at the cost of exceedingly large computing resources, and therefore simulations are restricted typically to few hundreds/thousands of particles. In the case of particle sizes being smaller than the smallest flow scales, a practical alternative is to represent them as infinitesimally small Lagrangian points, i.e., PP-DNS, with masses m_p and time-dependent positions $\mathbf{x}_p$ and velocities $\mathbf{v}_p$. This mesoscopic description of particles significantly reduces the computational cost of simulating dispersed multiphase flow, but requires assumed models to represent interphase transfer terms such as particle acceleration and fluid-particle heat exchange. The conservation equations describing irradiated, dispersed multiphase flow following a PP-DNS representation are introduced in the following subsections.

3.1.1 Eulerian description of turbulent flow

At low-pressure conditions, the carrier fluid is well approximated by the ideal-gas equation of state (EoS) $P_{th} = \rho_f R_f T_f$, where P_{th} is the thermodynamic pressure, ρ_f is the fluid density, $R_f = C_{P,f} - C_{v,f}$ is the specific gas constant of the fluid with $C_{P,f}$ and $C_{v,f}$ the isobaric and isochoric specific heat capacities, respectively, and T_f is the fluid temperature. As indicated by the EoS, density varies with temperature and pressure. However, the Mach number of the flow, $Ma = u/c$ with u the local flow velocity and c the speed of sound of the medium, is small for the range of thermodynamic conditions considered, i.e., $Ma \ll 1$. As a result, the low-Mach-number approximation is utilized to separate the hydrodynamic part, $p \ll P_{th}$, from the total pressure, $P = P_{th} + p$. This decomposition results in the following equations of fluid motion

$$\frac{\partial \rho_f}{\partial t} + \nabla \cdot (\rho_f \mathbf{u}_f) = 0, \tag{11}$$

$$\frac{\partial (\rho_f \mathbf{u}_f)}{\partial t} + \nabla \cdot (\rho_f \mathbf{u}_f \otimes \mathbf{u}_f) = -\nabla p + \nabla \cdot \left[\mu_f \left(\nabla \mathbf{u}_f + \nabla \mathbf{u}_f^{\mathsf{T}} \right) - \frac{2}{3} \mu_f (\nabla \cdot \mathbf{u}_f) \mathbf{I} \right] + \mathbf{f}_{TWC}, \tag{12}$$

$$\frac{\partial (\rho_f C_{v,f} T_f)}{\partial t} + \nabla \cdot (\rho_f C_{P,f} T_f \mathbf{u}_f) = \nabla \cdot (\lambda_f \nabla T_f) + S_{TWC}, \tag{13}$$

where $\mathbf{u}_f$ is the fluid velocity, $\mathbf{I}$ is the identity matrix, μ_f and λ_f are the dynamic viscosity and thermal conductivity, and $\mathbf{f}_{TWC}$ and S_{TWC} are two-way coupling terms representing the effect of particles on the fluid and defined as

$$\mathbf{f}_{TWC} = \sum_p m_p \frac{\mathbf{v}_p - \mathbf{u}_p}{\tau_p} \delta (\mathbf{x} - \mathbf{x}_p) \quad \text{and} \quad S_{TWC} = \sum_p \pi d_p^2 h (T_p - T_f) \delta (\mathbf{x} - \mathbf{x}_p), \tag{14}$$

with $\mathbf{u}_p$ the fluid velocity at the particle location, $\tau_p = \rho_p d_p^2 / (18 \mu_f)$ the particle relaxation time, d_p and T_p the particle diameter and temperature, respectively, and $\delta (\mathbf{x} - \mathbf{x}_p)$ the Dirac delta function concentrated at the particle position. In this work, particles are assumed to be isothermal as their Biot number, $Bi = h d_p / \lambda_p$ with h the fluid-particle convection coefficient and λ_p the thermal conductivity of particles, is small, i.e., $Bi \ll 1$.

3.1.2 Lagrangian representation of particles

The particle sizes are several orders of magnitude smaller than the smallest significant flow scale, and the density ratio between particles and fluid is $\rho_p / \rho_f \gg 1$. As a result, particles are modeled following a Lagrangian PP approach in which Stokes' drag is the most important force [27]. Their description in terms of position, velocity and temperature is given by

$$\frac{d\mathbf{x}_p}{dt} = \mathbf{v}_p, \tag{15}$$

$$\frac{d\mathbf{v}_p}{dt} = \frac{\mathbf{u}_p - \mathbf{v}_p}{\tau_p}, \tag{16}$$

$$\frac{d (m_p C_{v,p} T_p)}{dt} = \frac{\pi d_p^2}{4} \epsilon_p I_0 - \pi d_p^2 h (T_p - T_f), \tag{17}$$

where $C_{v,p}$ and ϵ_p are the particle specific isochoric heat capacity and radiation absorption coefficient (an optically thin medium has been assumed), and I_0 is the intensity irradiating the system. In the conservation equation for particle temperature (17), the first term on the right-hand-side accounts for the amount of radiation absorbed by a particle, while the second term represents the heat transferred to its surrounding fluid.

3.2 Point particle parcel modeling

The MF strategies described in Section 2 rely on the construction of levels with different ratios of fidelity and computational cost to accelerate the process of reducing the variance of the estimators. Diverse approaches can be pursued

to develop such cheaper models, with the only requirement that these lower-fidelity representations need to be orders of magnitude faster to compute while maintaining some degree of interdependence, the higher the better, for the QoI calculated at the different levels. In this work, the method chosen to build coarser representations of the dispersed Lagrangian field is based on parcel particles [28], also referred to as computational particles.

The particle parcels method represents the ensemble of N_p physical particles by N_s parcel particles. To reduce the computational cost, N_s is chosen to be smaller than N_p by ratios that can range several orders of magnitude. The average number of physical particles represented by a parcel particle is denoted by the statistical weight W_s, and their summation over all the parcel particles equals the total number of physical particles, i.e., $\sum_{s=1}^{N_s} W_s = N_p$. Similar to the physical particles, the N_s parcel particles are represented also in a Lagrangian frame with time-dependent positions $\mathbf{x}_s$ and velocities $\mathbf{v}_s$ described by transport equations. The parcel particles do not have to evolve identically to their physical counterparts because the equivalence of the parcel and physical systems is in a (weak) statistical sense; viz. sample trajectories do not have to coincide, ideally only the statistical moments, e.g., mean and variance. However, if the statistical weighting for all N_s is uniform in time and space, i.e., $W_s = N_p/N_s \ \forall s$, then the evolution of the parcel particles can be described with the same set of Lagrangian equations used for the physical particles, Eqs. (15)-(17), with the only modification of multiplying the physics coupling terms by W_s as

$$\mathbf{f}_{TWC} = \sum_p W_s m_p \frac{\mathbf{v}_p - \mathbf{u}_p}{\tau_p} \delta \left(\mathbf{x} - \mathbf{x}_p \right) \quad \text{and} \quad S_{TWC} = \sum_p W_s \pi d_p^2 h \left(T_p - T_f \right) \delta \left(\mathbf{x} - \mathbf{x}_p \right). \tag{18}$$

3.3 Computational approach

The equations of fluid motion (11-13) are solved following an Eulerian discretization implemented in an in-house solver [29] that is second-order accurate in space and suitable to non-uniform Cartesian meshes. A fourth-order Runge-Kutta scheme is used for integrating the equations in time, together with a fractional-step method for imposing conservation of mass. Integration in time of the Lagrangian position, velocity, and temperature of particles (15-17) is fully coupled with the advancement of the flow equations to ensure fourth-order accuracy.

4. DESCRIPTION OF THE IRRADIATED PARTICLE-LADEN TURBULENCE SYSTEM

The problem considered is inspired by the study of thermo-fluid mechanisms in volumetric particle-based solar energy receivers. As illustrated in Figure 1, the analysis of this type of systems involves the interaction of particles, turbulent flow and radiative heat transfer. The instantaneous snapshot, extracted from a PP-DNS with physical particles, corresponds to the normalized temperature increment, $\Delta T/T_0 = (T - T_0)/T_0$, of fluid and particles at the xy-plane along the streamwise direction. A complete description of the problem setup, uncertainties and QoIs, and fidelities designed is presented in the subsections below.
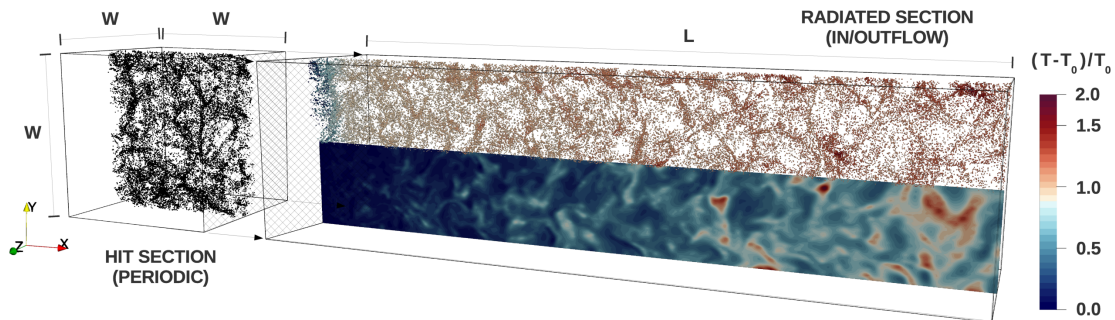

Figure 1: Setup of the irradiated particle-laden turbulence problem. The forced HIT (left) domain is utilized to provide fluid-particle inflow conditions to the radiated (right) section. The color scheme indicates normalized instantaneous temperature increment, $\Delta T/T_0$, of the dispersed (top) and carrier (bottom) phases on the streamwise xy-plane.

4.1 Problem setup

The study of particle-laden turbulence in an irradiated environment is performed by considering the two domains depicted in Figure 1. The cubic domain of size W is an isothermal, particle-laden HIT flow generator in which the fluid phase (initial density $\rho_{0,f}$ and temperature $T_{0,f}$, dynamic viscosity μ_f) is volumetrically forced [30]. The dispersed phase is initialized at the same time and temperature as the fluid with $N_{p,0}$ monodisperse particles (density ρ_p and diameter d_p) randomly distributed in the volume. The turbulence forcing scheme is targeted to produce an averaged turbulent kinetic energy, $k_{\infty,f}$, such that the ratio between domain size and Kolmogorov length scale η, is $W/\eta \sim \mathcal{O}(10^2)$, and therefore the small-scale features of the flow are not affected by the triply periodic boundaries. This first domain is designed to provide turbulent steady-state fluid-particle inflow conditions to the (second) radiated section.

The rectangular radiated section is of size $L \times W \times W$ in the streamwise (x, in/outflow boundaries), spanwise (y, periodic boundaries) and crossflow (z, periodic boundaries) directions, respectively. The turbulent fluid-particle mixture is sampled in time from a yz-plane in the HIT volume and introduced to this second domain by adding a bulk velocity U_0 to the streamwise velocity component. To achieve similar turbulent behavior as in wall-bounded flows, the ratio between root-mean-square (rms) velocity fluctuations, u_{rms}, in the HIT domain and U_0 is selected to be $u_{rms}/U_0 \sim \mathcal{O}(10^{-1})$, and the gravitational acceleration is not considered as its effects are negligible compared to the inertia of the bulk flow. As the fluid-particle mixture flows through the domain, it is irradiated with uniform intensity I_0. The result is that particles (isochoric heat capacity $C_{v,p}$ and absorption coefficient ϵ_p) absorb thermal radiation, increasing their temperature, T_p, and subsequently transferring energy to the surrounding fluid (thermal conductivity λ_f and isobaric heat capacity $C_{P,f}$) by thermal exchange (fluid-particle heat convection coefficient h).

The fluid-particle mixture in this problem is optically thin, allowing us to model the radiation absorption by the particles with the algebraic model described in Section 3, viz. all particles receive the same amount of radiation intensity. As a result of the particles heating and transferring thermal energy to the carrier fluid, the average fluid temperature $\overline{T}_f$ increases along the streamwise direction. This deposition of energy accelerates the flow by means of thermal expansion due to a decrease in fluid density ρ_f.

4.2 Uncertainties and quantities of interest

The analysis of the ML, CV and MLMF strategies is carried out by considering the 16 stochastic variables listed in Table 1, which are assumed to be uniform and independently distributed. As a theoretical research exercise, this set of uncertainties has been characterized by assuming a $\pm 10\%$ variation with respect to the following nominal values: inflow-outflow domain width $W = 0.04$ m and length $L = 0.16$ m; initial mixture temperature $T_0 = 300$ K; fluctuating $u_{rms} = 0.1$ m/s and bulk $U_0 = 1.0$ m/s velocities; initial fluid density $\rho_{f,0} = 1.2$ kg/m^3, dynamic viscosity $\mu_f = 1.95 \cdot 10^{-5}$ Pa$\cdot$s, thermal conductivity $\lambda_f = 2.75 \cdot 10^{-2}$ W/(m$\cdot$K), and isobaric heat capacity $C_{P,f} = 1000$ J/(kg$\cdot$K); fluid-particle convection coefficient $h = 2000$ W/(m$^2 \cdot$K); initial number of particles $N_{p,0} = 8 \cdot 10^5$; particles density $\rho_p = 8900$ kg/m^3, diameter $d_p = 2.75 \cdot 10^{-5}$ m, isochoric heat capacity $C_{v,p} = 450$ J/(kg$\cdot$K), and absorption coefficient $\epsilon_p = 0.4$; and radiation intensity $I_0 = 1.0 \cdot 10^6$ W/m^2.

The thermo-fluid response of the system is one of the principal functions to maximize in volumetric particle-based solar receivers. Consequently, the time-averaged fluid temperature increment (normalized), i.e., $Q = (\overline{T}_f - T_0)/T_0$, at the outlet of the radiated domain is chosen as the QoI studied. Data is collected by running HF and LF calculations for different values of the input parameters following a design of experiment (DoE) based on the KDOE approach [31], which is an iterative method that introduces stochasticity in the sampling process by means of a variable kernel density estimation to optimize the uniformity of the DoE. The range of values for W, L and u_{rms} are utilized to estimate the mesh resolution required to perform the HF PP-DNS calculations such that all the significant turbulent flow scales are captured, and as a result the contribution of the deterministic bias (i.e., discretization/approximation error) to the MSE of the estimators is negligible. The resulting Eulerian meshes for the HIT and radiated domains correspond to uniform Cartesian grids of sizes $128 \times 128 \times 128$ and $512 \times 128 \times 128$, respectively. The time-averaging is computed by taking the ensemble average of 5 flow through times (FTT) after the transient period is surpassed (approximately 10 FTTs), defined as FTT $\approx L/U_0$, on yz-planes since the solution is symmetric in the transverse directions.

Table 1: List of uncertainties with the corresponding ranges.

Uncertainty	Interval	Uncertainty	Interval
W	$[0.036 : 0.044]$ m	$C_{P,f}/C_{v,f}$	1.4 (diatomic ideal gases)
L	$[0.144 : 0.176]$ m	h	$[1800 : 2200]$ W/(m$^2 \cdot$ K)
T_0	$[270 : 330]$ K	$N_{p,0}$	$[7.2 \cdot 10^5 : 8.8 \cdot 10^5]$
u_{rms}	$[0.09 : 0.11]$ m/s	ρ_p	$[8010 : 9790]$ kg/m^3
U_0	$[0.9 : 1.1]$ m/s	d_p	$[2.48 \cdot 10^{-5} : 3.03 \cdot 10^{-5}]$ m
$\rho_{f,0}$	$[1.08 : 1.32]$ kg/m^3	$C_{v,p}$	$[405 : 495]$ J/(kg $\cdot$ K)
μ_f	$[1.76 \cdot 10^{-5} : 2.15 \cdot 10^{-5}]$ Pa $\cdot$ s	ϵ_p	$[0.36 : 0.44]$
λ_f	$[2.45 \cdot 10^{-2} : 3.03 \cdot 10^{-2}]$ W/(m $\cdot$ K)	I_0	$[9.0 \cdot 10^5 : 1.1 \cdot 10^6]$ W/m^2
$C_{P,f}$	$[900 : 1100]$ J/(kg $\cdot$ K)		

4.3 Fidelity levels/models designed

The performance of a selected set of different levels, or models, within a MF framework is assessed by considering 5 configuration parameters and their combinations. These parameters correspond to the Eulerian flow resolution in the (1) streamwise and (2) transverse directions, the (3) weight of particle parcels and their (4) one-way coupling with the flow, and (5) the total number of FTTs integrated in time to complete the calculations. The Hindu-Arabic numerals (1, 2, 3, 4, 5) indicate which low-fidelity parameters are activated for each level/model with respect to the full HF configuration. For instance, the notation LF-1 specifies that the flow resolution in the streamwise direction has been coarsened, while the other parameters are kept at the HF resolution/configuration. The notation utilized also denotes if a combination of LF parameters has been considered, like for example LF-12 which indicates that the flow resolution has been coarsed in both streamwise and transverse directions. Following this notation, 31 different LF levels/models have been constructed corresponding to LF-1, LF-2, LF-3, LF-4, LF-5, LF-12, LF-13 . . . LF-12345. The discrete set of HF and LF values for the different parameters are: (1) 512 or 64 grid points; (2) 128 or 16 grid points; (3) $W_s = 1$ or $W_s = 500$ parcel weights; (4) one-way coupling or fixed particles; (5) a total of 5 or 1 FTTs. To estimate the cost of the different fidelity levels/models, the idealized cost expression

$$C_{\text{FL}} = \text{FTT} \times \left\{ \text{ITER} \times \left[W_f \times \left(N_{f,yz}^3 + N_{f,x} \times N_{f,yz}^2 \right) + W_p \times \frac{N_{p,0}}{W_s} \right] \right\} \tag{19}$$

is utilized, where FTT $= 5, 1$ is the number of flow through times, ITER $= 6400, 3200$ is the number of time-step iterations per FTT, $W_f = 1.0$ and $W_p = 0.25, 0.0$ are the flow and particles sub-solver relative costs (viz. the cost of solving 4 particles is equivalent to solving 1 flow cell), and $N_{f,x} = 512, 64$, $N_{f,yz} = 128, 16$, $N_p = 8.0 \cdot 10^5, 1.6 \cdot 10^3$ are the number of grid points in the streamwise and transverse directions and number of physical particles, respectively. The cost of a HF sample following the expression above is $C_{\text{HF}} = 3.4 \cdot 10^{11}$ time-units. To facilitate the comparison of computational costs between levels within the ML, CV and MLMF approaches, the cost values per sample are normalized as $1 \geq \mathcal{C}_k = C_{\text{FL}}/C_{\text{HF}} > 0$, and the computational solver is assumed to scale linearly within the number of compute nodes and problem sizes considered. The normalized costs $\mathcal{C}_k$ of the different fidelity levels/models considered in this work are given in Table 2.

5. RESULTS AND DISCUSSION

This section presents and analyzes the data acquired by computing a set of 32 pilot samples per fidelity (resulting in a total of 1024 data points) as described in Section 4, and provides a discussion of the results obtained from the methodologies introduced in Section 2. A concise statistical characterization of the thermal-fluid QoI for the different fidelities is performed first in Section 5.1. Next, in Section 5.2, a quasi-optimal set of fidelities for the ML and CV estimators are selected based on balancing variance reduction / correlation and computational costs. Finally, the performance of the ML and CV estimators proposed and their MLMF hybridization is quantitatively analyzed in Section 5.3.

Table 2: Normalized cost of the different fidelity levels/models. The HF cost is taken as reference.

Fidelity	$\mathcal{C}_k$	Fidelity	$\mathcal{C}_k$	Fidelity	$\mathcal{C}_k$	Fidelity	$\mathcal{C}_k$
HF	1.00	LF-12	$1.03 \cdot 10^{-2}$	LF-123	$9.77 \cdot 10^{-4}$	LF-1234	$9.57 \cdot 10^{-4}$
LF-1	$1.57 \cdot 10^{-1}$	LF-13	$1.47 \cdot 10^{-1}$	LF-124	$9.58 \cdot 10^{-4}$	LF-1235	$1.95 \cdot 10^{-4}$
LF-2	$1.57 \cdot 10^{-2}$	LF-14	$1.46 \cdot 10^{-1}$	LF-125	$2.06 \cdot 10^{-3}$	LF-1245	$1.92 \cdot 10^{-4}$
LF-3	$9.81 \cdot 10^{-1}$	LF-15	$3.13 \cdot 10^{-2}$	LF-134	$1.45 \cdot 10^{-1}$	LF-1345	$2.92 \cdot 10^{-2}$
LF-4	$9.80 \cdot 10^{-1}$	LF-23	$6.34 \cdot 10^{-3}$	LF-135	$2.94 \cdot 10^{-2}$	LF-2345	$1.25 \cdot 10^{-3}$
LF-5	$2.00 \cdot 10^{-1}$	LF-24	$6.32 \cdot 10^{-3}$	LF-145	$2.93 \cdot 10^{-2}$	—	—
—	—	LF-25	$3.14 \cdot 10^{-3}$	LF-234	$6.31 \cdot 10^{-3}$	—	—
—	—	LF-34	$9.79 \cdot 10^{-1}$	LF-235	$1.27 \cdot 10^{-3}$	—	—
—	—	LF-35	$1.96 \cdot 10^{-1}$	LF-245	$1.26 \cdot 10^{-3}$	—	—
—	—	LF-45	$1.95 \cdot 10^{-1}$	LF-345	$1.94 \cdot 10^{-1}$	LF-12345	$1.91 \cdot 10^{-4}$

5.1 Statistical distribution of the quantity of interest

As a first step to construct efficient MF estimators, data have been collected by running the same 32 pilot samples for the 32 fidelities designed. For example, instantaneous thermal fluid increments in the radiated domain for the first sample generated by the HF, LF-1, LF-2, LF-3, LF-4 and LF-5 fidelities are depicted in Fig. 2. It can be inferred from the snapshots that with respect to the HF: (i) LF-1 qualitatively captures the local deposition of thermal energy and turbulent mixing on the carrier phase in the transverse direction, but is notably underresolved in the streamwise direction; (ii) the coarsening of the y and z flow resolutions largely modifies the patterns of the thermal-fluid structures in all three dimensions for LF-2, which become significantly broader and more diffused; (iii) the fluid temperature distribution is also underresolved in LF-3, but in this case the transfer of thermal energy is located around more concentrated and partitioned regions; (iv) the suppression of the particles one-way coupling with the flow in LF-4 results in a solution that resembles the diffusion of energy by forced convection; (v) the thermal-fluid patterns are well captured by LF-5, but the quantitative values are considerably smaller. Conclusive arguments, however, cannot be obtained by solely analyzing Fig. 2 as, first, the QoI targeted in this work is the normalized averaged increment of fluid temperature, $Q \equiv (\overline{T} - T_0)/T_0$, at the outlet of the radiated domain, and second the performance of the MF estimators is not directly related (to a first-order approximation) to the accuracy of the LF data.

All the data generated for the different samples and fidelities are collected and visually summarized in Fig. 3 by means of boxplots. The fidelities are sorted in decreasing $\mathcal{C}_k$ order starting from HF ($\mathcal{C}_k = 1$), finishing with LF-12345 ($\mathcal{C}_k \sim 10^{-4}$), and spanning a total of 5 decades. It can be seen that, in general, the cost of the fidelities reduces by orders of magnitude as more LF options are combined. For example, the cost of LF-1 is within the $\mathcal{C}_k \sim 10^{-1}$ decade, and when combined with LF options 2, 25 and 235 it reduces to $\mathcal{C}_k \sim 10^{-2}$, $\mathcal{C}_k \sim 10^{-3}$ and $\mathcal{C}_k \sim 10^{-4}$, respectively. The cost of the fidelities has been estimated by the idealized expression given in Eq. (19). Since these values are approximated, the cost of the fidelities within a decade is assumed to be the same in the following subsections, and consequently the ML and CV estimators will be constructed by selecting one fidelity per $\mathcal{C}_k$ decade.

As emphasized in the paragraph above, the accuracy of the LF data with respect to the HF results is not the principal component for the performance of the ML and CV estimators as they are, respectively, formulated in terms of variance reduction and correlation. Nevertheless, it is instructive to analyze the relations between the HF and LF data from a statistical perspective to gain insight and facilitate the effective construction of the MF estimators. The distributions in Fig. 3 show that the data are organized in 3 main blocks across the $\mathcal{C}_k$ decades: (1) fidelities HF, LF-3, LF-45, LF-345, LF-1, LF-13, LF-145, LF-1345, LF-2, LF-12, LF-23, LF-245, LF-2345, LF-1245, LF-12345 generate data distributed around $Q \approx 0.5$ and displaying moderate variances; (2) fidelities involving the fourth LF option LF-4, LF-34, LF-14, LF-134, LF-24, LF-234, LF-124, LF-1234 overestimate Q by approximately $2\times$ with significantly larger variability; and (3) the outcomes of fidelities LF-5, LF-35, LF-15, LF-135, LF-25, LF-125, LF-235, LF-1235, which consider 1 FTT only, are biased toward small Q values, but present variabilities similar to the HF data.
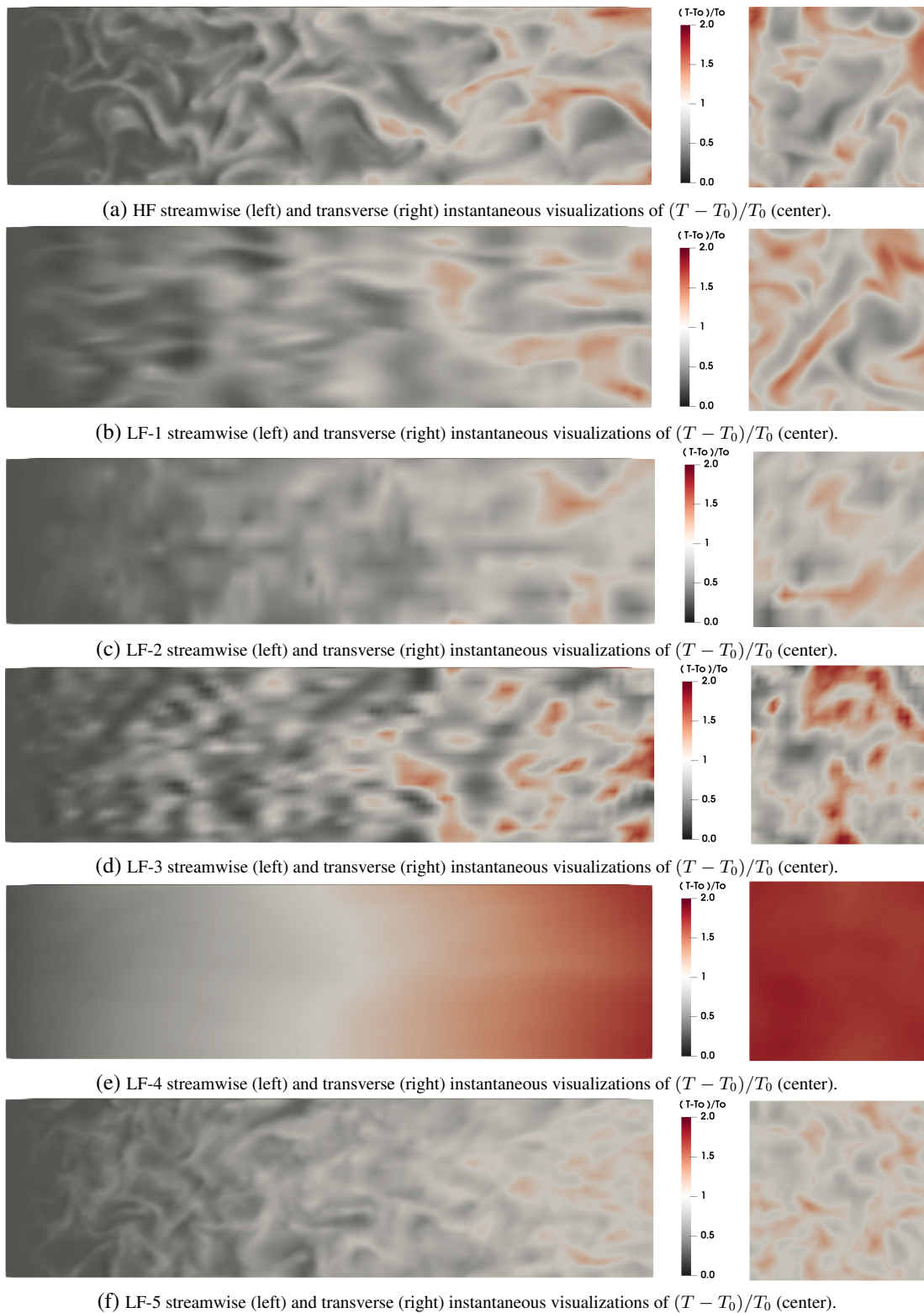
(a) HF streamwise (left) and transverse (right) instantaneous visualizations of $(T - T_0)/T_0$ (center).

(b) LF-1 streamwise (left) and transverse (right) instantaneous visualizations of $(T - T_0)/T_0$ (center).

(c) LF-2 streamwise (left) and transverse (right) instantaneous visualizations of $(T - T_0)/T_0$ (center).

(d) LF-3 streamwise (left) and transverse (right) instantaneous visualizations of $(T - T_0)/T_0$ (center).

(e) LF-4 streamwise (left) and transverse (right) instantaneous visualizations of $(T - T_0)/T_0$ (center).

(f) LF-5 streamwise (left) and transverse (right) instantaneous visualizations of $(T - T_0)/T_0$ (center).

Figure 2: Snapshots of normalized instantaneous temperature increment, $(T - T_0)/T_0$, of the fluid phase on the streamwise xy-plane and transverse outlet yz-plane for the HF (a), LF-1 (b), LF-2 (c), LF-3 (d), LF-4 (e) and LF-5 (f) fidelities.
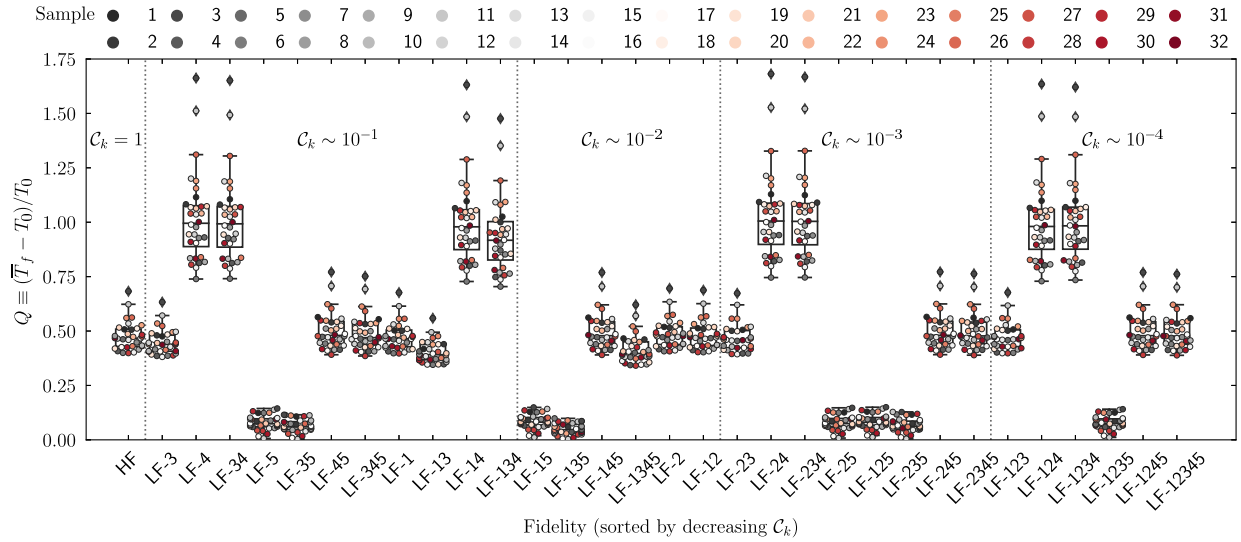
Figure 3: Distributions of normalized averaged temperature increment, $Q \equiv (\overline{T} - T_0)/T_0$, data at the outlet of the radiated domain for the different samples and fidelities sorted by decreasing $\mathcal{C}_k$. Boxplots display the minimum (small horizontal lines at Q1-1.5×IQR), maximum (small horizontal lines at Q3+1.5×IQR), whiskers (vertical lines), interquartile range (boxes spanning IQR=Q3-Q1), median (large horizontal lines), outliers (diamonds), and data points (colored circles) of the distributions.

5.2 Pragmatic selection of fidelities

The pilot data analyzed previously is utilized in this subsection to propose quasi-optimal MF estimators. The methodologies followed to select the set of fidelities for the ML, CV and MLMF estimators are based on optimizing the variance reduction / correlation with respect to the costs $\mathcal{C}_k$, as represented in Fig. 4, and their resulting performances are depicted in Figs. 5 and 6.

The quasi-optimal ML estimator is designed with 5 levels (one per $\mathcal{C}_k$ decade) and constructed by utilizing the fidelities highlighted in Fig. 4(a) with the solid grey line. These correspond to HF, LF-1, LF-12, LF-245 and LF-12345, which are combined to generate the levels $Y_0 = Q_{\text{LF-12345}}$, $Y_1 = Q_{\text{LF-245}} - Q_{\text{LF-12345}}$, $Y_2 = Q_{\text{LF-12}} - Q_{\text{LF-245}}$, $Y_3 = Q_{\text{LF-1}} - Q_{\text{LF-12}}$ and $Y_4 = Q_{\text{HF}} - Q_{\text{LF-1}}$. Since $\mathbb{V}\left[\hat{Q}^{\text{ML}}\right]$ is the sum of the Y_ℓ variances divided by the number of samples per level, this set of fidelities is selected by following the minimum of $\mathbb{V}[Y_\ell]$ across levels between potential fidelity pairs starting from the most expensive HF level. These fidelities pertain to block 1, which is characterized by boxplots presenting small biases and moderate variances, and as a result provide small $\mathbb{V}[Y_\ell]$ values. As shown in Fig. 5, the resulting ML estimator efficiently reduces its variance by computing many (cheap) samples at the coarse level, where $\mathbb{V}[Y_\ell]$ is large, and few samples at the expensive Y_4 level, where $\mathbb{V}[Y_\ell]$ is by construction very small.

Following a similar procedure, the quasi-optimal CV estimator is composed of the 5 fidelities (one per $\mathcal{C}_k$ decade) indicated by the solid grey line in Fig. 4(b). These correspond to HF, LF-134, LF-2, LF-2345 and LF-1245, and their optimal ratios of LF-to-HF number of samples are $r_f \approx 1$, 2, 5, 15, 100, respectively. Since the rate of reduction of $\mathbb{V}\left[\hat{Q}^{\text{CV}}\right]$ is proportional to ρ_f divided by N_f, this set of fidelities is chosen by considering the maximum ρ_f path along the 5 different $\mathcal{C}_k$ decades, which only decays from $\rho_f = 1$ to $\rho_f \approx 0.8$ while achieving $\mathcal{C}_k$ reductions of $\sim 10^4 \times$. An interesting observation is that fidelity LF-134, as well as some of the fidelities discarded (violet circles) close to the quasi-optimal path, does not pertain to block 1, indicating that CV estimators can be constructed with modeling approaches that present significant biases with respect to HF. This is a very important result as it exemplifies the higher robustness of CV estimators with respect to ML strategies. However, as demonstrated in Fig. 5, if optimal LF models can be designed, ML estimators typically outperform CV approaches.
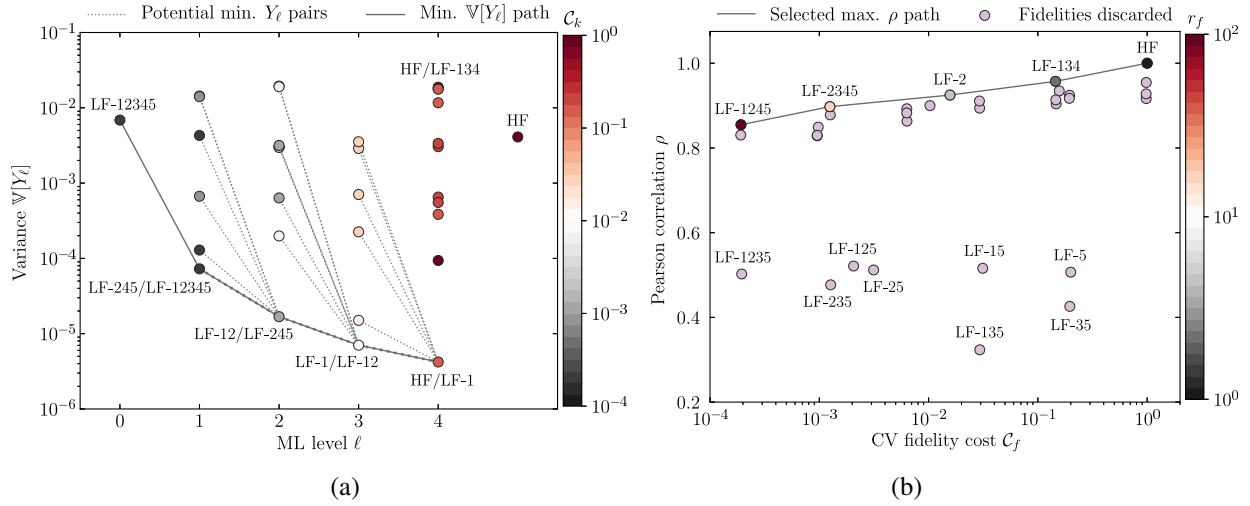
Figure 4: Variance of Y_ℓ (a) across levels for a set of potential pairs (dashed lines) and selected combination (solid line), with color indicating fidelity cost $\mathcal{C}_k$. Pearson correlation coefficient (b) versus fidelity cost for all the HF-LF combinations and selected maximum ρ path (solid line). The color of the circles indicates the r_f optimal ratio between HF and LF samples.

5.3 Performance of the multifidelity estimators

Finally, the performance of the quasi-optimal ML and CV multifidelity estimators constructed in the previous subsection are reported in Fig. 5(a). The horizontal logarithmic-scale axis corresponds to the total cost $C = \sum_{\ell=0}^{L} N_\ell \mathcal{C}_\ell$ for ML and computational budget γ for CV normalized by the cost of a HF sample. On the vertical logarithmic-scale axis, target estimators' standard deviations, ε, normalized by the MC value, ε_0^{MC}, are shown for MC and ML estimators through evaluating Eq. (2), and for the CV estimator utilizing the expression [21]

$$\frac{\varepsilon}{\varepsilon^{MC}} = \sqrt{1 - \rho_1^2} + \sum_{f=1}^{F} \sqrt{\frac{\mathcal{C}_f}{\mathcal{C}_{HF}} \left(\rho_f^2 + \rho_{f+1}^2 \right)}. \tag{20}$$

The results show that the MF speedups with respect to a straightforward MC correspond to $\sim 10^3$ and $\sim 10^1$ for the ML and CV quasi-optimal estimators, respectively. For example, to reduce ε by an order of magnitude with respect to ε_0^{MC}, the MC approach requires computing 3200 HF runs, while the ML (composed by 8 HF, 26 LF-1, 157 LF-12, 931 LF-245 and 2324 LF-12345) and CV (composed by 444 HF, 992 LF-134, 2081 LF-2, 7772 LF-2345 and 52733 LF-1245) demand only 20 and 640 HF equivalent runs, respectively.

The results depicted in Fig. 5(a) highlight the better performance of the ML with respect to the CV estimator in the case of LF models/levels that present small bias and moderate variances. However, in a more general problem involving complex, non-linear flow phenomena, like for example boundary layers or shocks, in which such "good" LF models are more challenging to design and/or discover, the CV approach may be a more robust option as shown in Fig. 5(b), where the MF estimators have been constructed by taking the cheapest fidelity for each $\mathcal{C}_k$ decade corresponding to HF, LF-134, LF-12, LF-2345 and LF-12345. In that case, since $\mathbb{V}[Y_\ell]$ of the 4th (most expensive) level is larger than for HF, the performance of the ML estimator is significantly degraded to the point in which it performs approximately $10\times$ worse than the straightforward MC. Instead, the performance of the CV estimator is barely impacted, and therefore maintains a speedup of order $10\times$ with respect to the plain MC. This observation indicates that hybridization strategies, like for example the bi-fidelity approximation [17,25] and the multilevel multifidelity [24,32], are promising extensions of the standard ML and CV methods for accelerating the convergence rate of statistical estimators in challenging flow problems.

The MLMF hybridization strategy is analyzed in Fig. 6 for the case of constructing MF estimators with two
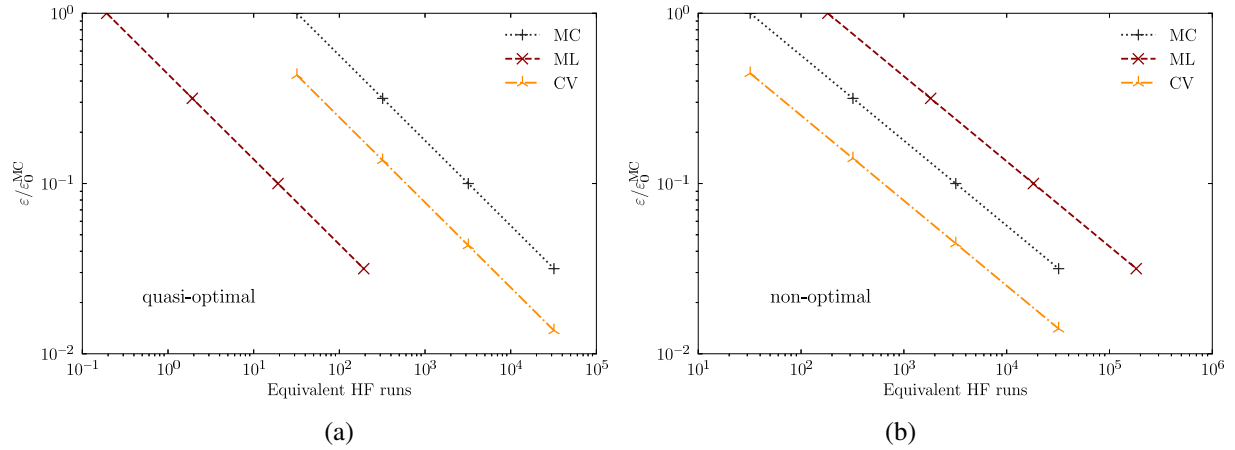
Figure 5: Extrapolated MSE of the MC, ML and CV estimators (normalized by the MC value), $\varepsilon/\varepsilon_0^{\mathrm{MC}}$, for quasi-optimal (a) and non-optimal (b) configurations as function of the overall computational cost in terms of equivalent number of HF runs.

levels. The estimators studied correspond to: (i) ML using the HF and LF-1 models, (ii) CV between fidelities LF-1 and LF-1245, and (iii) MLMF resulting from the hybridization of i and ii. The variance of the Y_ℓ levels is depicted in Fig. 6(a), where the fine level (HF/LF-1) presents a very small value, and consequently few samples are required at this expensive level, whereas the coarse level (LF-1) produces a large variance at a significantly high $\mathcal{C}_k$ cost. The panel also provides the correlation value $\rho = 0.85$ between LF-1 and LF-1245, indicating that the MLMF hybridization will (potentially) improve the performance of the estimator with respect to ML. The performances of the two-level ML, CV and MLMF estimators constructed are reported in Fig. 6(b). Three main observations can be extracted from the plot. First, the performance of the ML estimator with respect to the previous five-level design is reduced by almost $\sim 100\times$, which denotes that smooth $\mathbb{V}[Y_\ell]$ transitions between levels is necessary to obtain large speedups. Second, once again, the CV estimator is not significantly impacted by modifying its structure if models with "good" balances between correlation and cost are available. Finally, the third observation is that the performance of ML can be largely improved (an order of magnitude in this case) by easily adding a CV at the coarsest level.

6. SUMMARY, CONCLUSIONS AND FUTURE WORK

Performing uncertainty quantification studies of large-scale, multiphysics flow applications is challenging due to the expensive high-fidelity calculations usually required and the large number of uncertainties typically encountered. For instance, the utilization of a straightforward MC approach for the problem studied would require in the order of billion core-hours on some of the most advanced supercomputers, while, as shown in Section 5, equivalent predictions can be obtained at significant cost saving ratios by means of multifidelity strategies.

On the basis of the problem of interest and methods considered, the multilevel Monte Carlo performs better than the control variates due to the small bias and moderate variability of the low-fidelity data generated. However, it is also shown that the speedup achieved by the control variates approach is more robust to situations in which designing, or discovering, "good" low-fidelity models is challenging as a result of complex flow phenomena. An interesting observation is that the hybridization of these two approaches, based on the multilevel multifidelity, results in a significant speedup improvement by simply adding a control variates at the coarse levels.

Ongoing work is focused on leveraging the speedup obtained from the acceleration strategies analyzed to characterize thermal-fluid mechanisms subject to uncertainty for this problem. Future work will explore Exascale-ready ensemble-based computational strategies to effectively perform multifidelity calculations on hybrid (CPU/GPU) computing facilities [33], and the utilization of turbulence models (LES & RANS) for the design of LF levels.
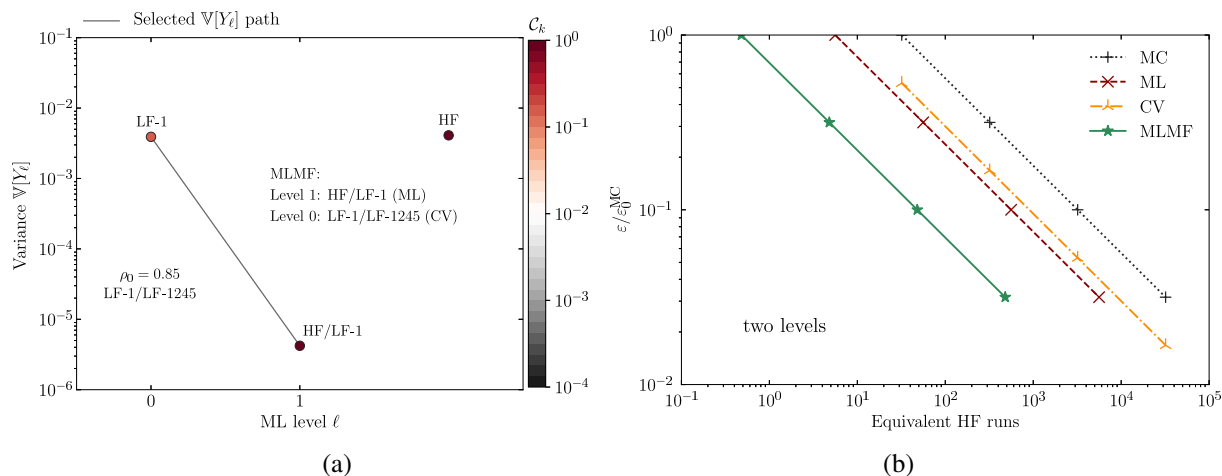
Figure 6: Variance of Y_ℓ (a) across the two levels of the MLMF estimator (solid line), with color indicating fidelity cost $\mathcal{C}_k$. Extrapolated MSE (b) of the MC and two-level ML, CV and MLMF estimators (normalized by the MC value), $\varepsilon / \varepsilon_0^{\mathrm{MC}}$, as function of equivalent number of HF runs.

ACKNOWLEDGMENTS

REFERENCES

1. Najm, H.N., Uncertainty Quantification and Polynomial Chaos Techniques in Computational Fluid Dynamics, *Ann. Rev. Fluid Mech.*, vol. **41**, pp. 35-52, 2009.

2. Exascale Computing Engineering Center. Predictive Science Academic Alliance Program (PSAAP) II, Stanford University, 2019. Accessed from https://exascale.stanford.edu/

3. Shaw, R.A., Particle-turbulence Interactions in Atmospheric Clouds, *Ann. Rev. Fluid Mech.*, vol. **35**, pp. 183-227, 2003.

4. Tieszen, S.R., On the Fluid Mechanics of Fires, *Ann. Rev. Fluid Mech.*, vol. **33**, pp. 67-92, 2001.

5. Lasheras, J.C. and Hopfinger, E.J., Liquid Jet Instability and Atomization in a Coaxial Gas Stream, *Ann. Rev. Fluid Mech.*, vol. **32**, pp. 275-308, 2000.

6. Dodd, M.S. and Jofre, L., Small-scale Flow Topologies in Decaying Isotropic Turbulence Laden with Finite-Size Droplets, *Phys. Rev. Fluids*, vol. **4**, pp. 064303, 2019.

7. Raman, V. and Fox, R.O., Modeling of Fine-Particle Formation in Turbulent Flames, *Annu. Rev. Fluid Mech.*, vol. **48**, pp. 159-190, 2016.

8. Ho, C.K., Advances in Central Receivers for Concentrating Solar Applications, *Sol. Energy*, vol. **152**, pp. 38-56, 2017.

9. Caporaloni, M., Tampieri, F., Trombetti, F. and Vittori, O., Transfer of Particles in Nonisotropic Air Turbulence, *J. Atmos. Sci.*, vol. **32**, pp. 565-568, 1975.

10. Balachandar, S. and Eaton, J.K., Turbulent Dispersed Multiphase Flow, *Ann. Rev. Fluid Mech.*, vol. **42**, pp. 111-133, 2010.

11. Dunton, A.M., Jofre, L., Doostan, A. and Iaccarino, G., Pass-efficient Methods For

Compression of High-Dimensional Turbulent Flow Data, *CTR Ann. Res. Briefs*, pp. 313-325, 2017.

12. Zamansky, R., Coletti, F., Massot, M. and Mani, A., Radiation Induces Turbulence in Particle-Laden Fluids, *Phys. Fluids*, vol. **26**, pp. 071701, 2014.

13. Frankel, A., Pouransari, H., Coletti, F. and Mani, A., Settling of Heated Particles in Homogeneous Turbulence, *J. Fluid Mech.*, vol. **792**, pp. 869-893, 2016.

14. Pouransari, H. and Mani, A., Effects of Preferential Concentration on Heat Transfer in Particle-Based Solar Receivers, *J. Sol. Energy Eng.*, vol. **139**, pp. 021008, 2017.

15. Rahmani, M., Geraci, G., Iaccarino, G. and Mani, A., Effects of Particle Polydispersity on Radiative Heat Transfer in Particle-Laden Turbulent Flows, *Int. J. Multiph. Flow*, vol. **104**, pp. 42-59, 2018.

16. Jofre, L., Geraci, G., Fairbanks, H.R., Doostan, A. and Iaccarino, G., Multi-fidelity Uncertainty Quantification of Irradiated Particle-Laden Turbulence, *CTR Ann. Res. Briefs*, pp. 21-34, 2017.

17. Fairbanks, H.R., Jofre, L., Geraci, G., Iaccarino, G., and Doostan, A., Bi-fidelity Approximation for Uncertainty Quantification and Sensitivity Analysis of Irradiated Particle-Laden Turbulence, *J. Comput. Phys.*, vol. **402**, pp. 108996, 2020.

18. Jofre, L., del Rosario, Z.R., and Iaccarino, G., Data-driven Dimensional Analysis of Heat Transfer in Irradiated Particle-Laden Turbulent Flow, *Int. J. Multiph. Flow*, vol. **125**, pp. 103198, 2020.

19. Jofre, L., Domino, S.P., and Iaccarino, G., A Framework for Characterizing Structural Uncertainty in Large-Eddy Simulation Closures, *Flow Turbul. Combust.*, vol. **100**, no. 2, pp. 341-363, 2018.

20. Jofre, L., Domino, S.P. and Iaccarino, G., Eigensensitivity Analysis of Subgrid-Scale Stresses in Large-Eddy Simulation of a Turbulent Axisymmetric Jet, *Int. J. Heat Fluid Fl.*, vol. **7**, pp. 314-335, 2019.

21. Peherstorfer, B., Willcox, K. and Gunzburger, M., Survey of Multifidelity Methods in Uncertainty Propagation, Inference, and Optimization, *SIAM Rev.*, vol. **60**, no. 3, pp. 550-591, 2018.

22. Giles, M.B., Multi-level Monte Carlo Path Simulation, *Oper. Res.*, vol. **56**, pp. 607-617, 2008.

23. Pasupathy, R., Taaffe, M., Schmeiser, B.W. and Wang, W., Control-variate Estimation Using Estimated Control Means, *IIE Trans.*, vol. **44**, pp. 381-385, 2014.

24. Geraci, G., Eldred, M., and Iaccarino, G., A Multifidelity Multilevel Monte Carlo Method for Uncertainty Propagation in Aerospace Applications, *19th AIAA Non-Deterministic Approaches Conference*, pp. 2017-1951, 2017.

25. Fairbanks, H.R., Doostan, A., Ketelsen, C. and Iaccarino, G., A Low-rank Control Variate For Multilevel Monte Carlo Simulation Of High-Dimensional Uncertain Systems, *J. Comput. Phys.*, vol. **341**, pp. 121-139, 2017.

26. Subramaniam, S., Lagrangian-Eulerian Methods for Multiphase Flows, *Prog. Energy Combust. Sci.*, vol. **39**, pp. 215-245, 2013.

27. Maxey, M.R. and Riley, J.J., Equation of Motion for a Small Rigid Sphere in a Nonuniform Flow, *Phys. Fluids*, vol. **26**, pp. 883-889, 1983.

28. Amsden, A.A., O'Rourke, P.J. and Butler, T.D., KIVA-II: a Computer Program for Chemically Reactive Flows with Sprays, Tech. Rep. LA-12503-MS, LANL, 1989.

29. Esmaily, M., Jofre, L., Mani, A. and Iaccarino, G., A Scalable Geometric Multigrid Solver for Nonsymmetric Elliptic Systems with Application to Variable-Density Flows, *J. Comput. Phys.*, vol. **357**, pp. 142-158, 2018.

30. Bassenne, M., Urzay, J., Park, G.I. and Moin, P., Constant-energetics Physical-Space Forcing Methods for Improved Convergence to Homogeneous-Isotropic Turbulence with Applications to Particle-Laden Flows, *Phys. Fluids*, vol. **28**, pp. 035114, 2016.

31. Roy, P.T., Jofre, L., Jouhaud, J.C. and Cuenot, B., Versatile sequential sampling algorithm using kernel density estimation, *Euro. J. Oper. Res.*, vol. 284, no. 1, 2020. DOI: 10.1016/j.ejor.2019.11.070

32. Gorodetsky, A.A., Geraci, G., Eldred, M.S. and Jakeman, J.D., A Generalized Approximate Control Variate Framework for Multifidelity Uncertainty Quantification, *arXiv preprint*, *arXiv:1811.04988:[stat.CO]*, 2019.

33. Jofre, L., Papadakis, M., Aiken, A. and Iaccarino, G., Multifidelity Ensemble-based Prediction Of Turbulent Flows at the Exascale, *72nd Annual Meeting of the American Physical Society - Division of Fluid Dynamics*, pp. DFD19–2019–000161, 2019.