SAND2019-15494C

Higher moment multilevel estimators for optimization under uncertainty applied to wind plant design

Friedrich Menhorn*
Technical University of Munich, 85748 Garching, Germany

Gianluca Geraci† and Daniel T. Seidl‡
Sandia National Laboratories, Albuquerque, New Mexico 87123, USA

Michael S. Eldred§
Sandia National Laboratories, Albuquerque, New Mexico 87123, USA

Ryan N. King¶
National Renewable Laboratory, Golden, Colorado 80401, USA.

Hans-Joachim Bungartz‖
Technical University of Munich, 85748 Garching, Germany

Youssef M. Marzouk**
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

The design of wind power plants is a complex engineering task that requires accurate numerical simulations that include both computational fluid and structural dynamics. The accurate prediction of the global performance of the plant is therefore based on the possibility to capture the evolving dynamics for several fluid scales, from local eddies affecting the loading on the turbine blades up to large eddies that form the interacting wakes downstream with respect to each rotor. The presence of these complex flow structures and their interactions requires the use of high-fidelity tools and high resolution grids. At the same time the operative conditions of the plant are intrinsically stochastic in their nature and uncertainty quantification techniques are needed to both characterize the sources of uncertainty and propagate them through the numerical codes. The high computational burden of this task is very often prohibitive for a single plant configuration and it is exacerbated in the case of a design process in which an uncertainty quantification propagation is required for each design iteration. In this work we present and discuss the integration of multilevel uncertainty quantification and design strategies that have the potential to drastically reduce the overall cost of an optimization under uncertainty study. The main objective of the multilevel strategy is to combine computational tools with different accuracy and computational cost such that information from a hierarchy of resolutions can be efficiently fused to decrease the overall cost without compromising the overall accuracy. We focus on algorithmic advancements for both the forward uncertainty quantification step and the optimization step as well. We present multilevel estimators for higher moments, particularly the variance which is employed in robust optimization problems and show the advantage compared to standard multilevel estimators. We use simple model problems to describe the different algorithmic components and their features and performance. As a final demonstration we consider a wind plant design problem involving two turbines for which we combine fluid dynamics tools with different numerical accuracy based on Reynolds

*PhD Candidate, Department of Informatics.
†Senior Member of Technical Staff, Optimization and Uncertainty Estimation, Sandia National Laboratory. Member AIAA.
‡Senior Member of Technical Staff, Optimization and Uncertainty Estimation, Sandia National Laboratory.
§Distinguished Member of Technical Staff, Optimization and Uncertainty Estimation, Sandia National Laboratory. Associate Fellow AIAA.
¶Senior Scientist, Computational Science Center, National Renewable Laboratory. Member AIAA.
‖Professor, Department of Informatics.
**Associate Professor, Department of Aeronautics and Astronautics.

Averaged Navier-Stokes equations. Finally, we compare the performance of the newly developed multilevel strategies with their single fidelity counterpart.

I. Introduction

The numerical design of wind power plants requires to perform a series of challenging computational tasks that includes accurate turbulent computational fluid dynamics (CFD) simulations and structural dynamics analyses. Moreover, wind power plants operate in an environment that is stochastic in nature and thus the accurate quantification of their performance also requires the characterization and propagation of uncertainty sources through the numerical codes. This latter tasks concur to define the so called uncertainty quantification (UQ) workflow. It is well known that UQ requires multiple realizations of the numerical code which correspond to the response of the system for different combinations of input/uncertain parameters. This process, despite the recent algorithmic advancements, is still computational expensive especially since it requires a number of system realizations that roughly scales with the number of input parameters. It is evident that a similar workflow is prohibitively expensive when the system, as in our case, requires high-fidelity simulations for its accurate characterization. This cost is greatly augmented when the UQ step is encapsulated in an external loop as for instance the one required by optimization. To overcome this difficulty in recent year the concept of multilevel and multifidelity UQ has been introduced in a series of papers[1–9] mimicking earlier developments presented in the optimization literature [10]. These approaches are all based on a similar key concept: whenever several models with increasing accuracy (and computational cost) are available, it is possible to optimally and efficiently fuse information from all of them in order to minimize the overall computational cost while preserving the accuracy requirements.

In this work we are interested in demonstrating how multilevel approaches can be introduced and adapted for the UQ step in order to greatly reduce the overall computational cost of the optimization under uncertainty (OUU) for wind plants. In particular, we focus our attention on multifidelity sampling-based approaches as the multilevel Monte Carlo (MLMC) [2, 3] to accelerate the UQ propagation step. For the optimization, we base our work on the derivative-free optimization library SNOWPAC [11]. Our algorithmic contributions are focused on the coupling between SNOWPAC and the MLMC implementation in Dakota extending what we have presented in [12]. Specifically, we present new multilevel estimators for higher order moments, namely the variance extending work from [13, 14]. By coupling those estimators with SNOPWAC, we extend the single fidelity approach to a more efficient multilevel version. We present numerical tests on a simplified problem to highlight and demonstrate the features of the proposed strategy compared to its standard single level counterpart.

As a final demonstration of the overall multilevel OUU workflow we focus on a wind power plant design problem for which we combine computational fluid dynamics tools with increasing computational complexity as a result of a hierarchy of grid resolutions of Reynolds Averaged Navier Stokes (RANS) equations. The remainder of the paper is organized as it follows. The wind plant design problem and computational fluid dynamics tools used is described in Section II. The optimization and uncertainty quantification approaches are briefly described in Section III. Results for the model and the wind problem are discussed in Section IV. Concluding remarks end the paper in Section V.

II. Application description

Wind power plant design can benefit from different strategies in order to maximize the power extraction. However, for a wind plant that have been already designed and installed, control techniques can be adopted to improve their performance. One effective controls technique is wake steering in which the turbines are intentionally yawed away from a perpendicular alignment with the incoming wind. When doing so, the turbines act like flow control devices and can direct the wake away from downstream rotors and also entrain more momentum from aloft. Obtaining effective wake steering strategies is complicated by uncertainty in the incoming wind direction, wind intensity, turbulence levels, *etc.* Furthermore, noise in wind vane sensors and errors in the yaw control system introduce additional uncertainty in obtaining optimal wake steering strategies. The complex turbulent flow in the turbine wake, with high-vorticity regions interacting with other wakes and downstream rotors makes the sensor measurements subject to a significant uncertainty that obfuscates the actual yaw angle of each turbine.

In this work we are interested in formulating and solving an OUU problem for finding the optimal wake steering strategy inspired by the work presented in [15]. In this previous work, the authors formulated an OUU problem with the expected annual energy production (AEP) as the quantity of interest. The authors focused on a two-turbine case with different level of the uncertainty for the yaw angle as well as the Princess Amalia wind farm. For both cases the authors

were able to show the advantage of a OUU formulation with respect to a deterministic optimization (which assumes perfect knowledge of the yaw misalignment).

We study a similar OUU problem to maximize the total power production of two turbines subject to a set of box constraints on yaw angles. However, we consider a situation with partial wake overlap that commonly occurs within a wind farm. Futhermore, we demonstrate that better results can be achieved by employing a multilevel OUU strategy leveraging a hierarchy of resolutions of RANS equations while the costs are equal to a standard Monte Carlo strategy.

Regarding the underlying flow model and solver we are using Reynolds-averaged Navier Stokes (RANS) codes to solve for ensemble averaged flow fields with a turbulence closure for the Reynolds stress and which provide a mid-fidelity representation of wake interactions and intra-farm turbulent flow. WindSE [16] is a RANS code with actuator disk turbine representation and automatic differentiation for high dimensional gradients that is intended for optimization and UQ studies. Additionally, it offers automatic and parametrized meshing which allows the construction of different resolutions levels. Finally, computational cost for a high resolution mesh is large enough to show the effectiveness of multilevel approaches while still being feasible for a test environment.

III. Multilevel Monte Carlo for OUU

In this section we briefly describe the main components of our OUU approach: the library SNOWPAC for the derivative-free optimization solver and the multilevel Monte Carlo capabilities implemented in DAKOTA which is also tasked with managing the overall UQ workflow.

A. OUU formulation

In the following we summarize the OUU formulation for a generic quantity of interest that we want to solve. Given a stochastic optimization problem with objective function f and inequality constraints c

$$\begin{aligned} \min_{x} \quad & f(x,\theta), \\ \text{s.t.} \quad & c(x,\theta) \leq 0 \end{aligned} \tag{1}$$

where $x \in \mathbb{R}^d$ is the design variable and $\theta := (\theta_1, ..., \theta_r) : (\Omega, F, P) \to (\Theta, B(\Theta), \mu)$ is a stochastic parameters mapping from probability space (Ω, F, P) to $(\Theta, B(\Theta), \mu), \Theta \subset \mathbb{R}^m$. Here, $B(\Theta)$ denotes the standard Borel σ-field. The uncertain parameters θ represents the uncertainty in measured quantities or may model a lack of knowledge about process parameters.

Next, we reformulate (1) into a robust optimization problem using measures of robustness and risk $\mathcal{R}$

$$\begin{aligned} \min_{x} \quad & \mathcal{R}^f(x,\theta), \\ \text{s.t.} \quad & \mathcal{R}^c(x,\theta) \leq 0 \end{aligned} \tag{2}$$

and we are interested in optimal designs x which are robust with respect to the introduced uncertainty by θ. Typical examples for those measures given a general function $b \in \{f, c\}$ are expected value, $\mathcal{R}^b(x,\theta) := \mathbb{E}[b(x,\theta)]$, a linear combination with the standard deviation, $\mathcal{R}^b(x,\theta) := \mathbb{E}[b(x,\theta)] + \alpha\mathbb{V}[b(x,\theta)]^{\frac{1}{2}}$, or—more specifically—chance constraints, $\mathcal{R}^c(x,\theta) := \mathbb{E}[\mathbb{1}(c(x,\theta) \geq 0)] - (1-\beta)$, given a probability level $\beta \in]0, 1[$.

B. (S)NOWPAC

We use the derivative-free constrained stochastic optimization method SNOWPAC [11] for solving the resulting robust optimization problem given in Eq. (2). The method is an extension of the deterministic derivative-free method NOWPAC [17] which employs an inner boundary path to ensure feasible trial points. Furthermore, it uses a trust region approach to approximate the objective f and constraint c locally around the current optimal design employing minimum Frobenius norm surrogate models (see [18]) m^f and m^c, respectively. Next, classical quadratic optimization algorithms can be used on the surrogate to find the best optimization step. For NOWPAC, the authors in [17] show that the method is globally convergent to a feasible local optimum.

SNOWPAC extends NOWPAC for robust optimization problems which requires multiple adaptations to the algorithm. It introduces Monte Carlo sampling estimators to evaluate the robustness measures $R^b = \mathcal{R}^b + \varepsilon_b$. This, however, introduces an error ε_b restricting the approximation quality of the surrogates m^f and m^c in the trust region. The authors

show that the minimal possible trust region radius ρ_k at a given optimization step k is restricted by the maximal noise $\varepsilon^k_{\max} = \max_{b\in\{f,c\}} \varepsilon^k_b$,

$$\varepsilon^k_{max}\rho_k^{-2} \leq \lambda_t^{-2}, \quad \text{resp.} \qquad \rho_k \geq \lambda_t\sqrt{\varepsilon^k_{max}} = \max_i \lambda_t\sqrt{\varepsilon^k_i}, \tag{3}$$

where λ_t is a safety parameter. Thus, progress of the algorithm is only achieved if the noise can be reduced.

Therefore, Gaussian process surrogates (see [19] for an introduction) are utilized additionally to bias the evaluations and to reduce the noise. They are built over all previous evaluations using a squared exponential kernel and maximum likelihood hyperparameter optimization. Smoothened evaluations $\tilde{R}^b_k$ and noise estimates $\tilde{\varepsilon}^b_k$ are constructed via a convex sum

$$\begin{aligned} \tilde{R}^b_k &= \gamma_k \mathcal{G}^k_b(x_k) + (1-\gamma_k)R^b, \\ \tilde{\varepsilon}^b_k &= \gamma_k 2\sigma_b(x_k) + (1-\gamma_k)\varepsilon^b_k, \end{aligned} \tag{4}$$

where $\mathcal{G}^k_b(x_k)$ denotes the Gaussian process estimate and $\sigma_b(x_k)$ denotes the standard deviation of $\mathcal{G}^k_b$ at point x_k. The weight factor $\gamma_k := e^{-\sigma_b(x_k)}$ is chosen to approach 1 following the approximation quality of the Gaussian process. The corrected evaluations $\tilde{R}^b_k$ at the local interpolation points are then used to build local surrogates and the associated reduced noise level $\tilde{\varepsilon}^b$ allows a reduction in the trust region radius ρ_k.

In summary, SNOWPAC takes two sources of error into account with the use of those two surrogate models. While the Gaussian process surrogate is built over a larger domain and, therefore, holds more global information, the minimum Frobenius norm surrogate models are built locally in the trust region. Through the combination following Eq. (4) we balance the error in the surrogate model via the lower bound on the trust region with the error in the Gaussian process model represented by its standard deviation estimate. With an increasing number of evaluations we gain confidence in the Gaussian process model, are able to decrease the noise and, finally, the trust region.

SNOWPAC showed promising results as described in [11], outperforming optimization methods like COBYLA [20], NOMAD [21] or cBO [22] on a collection of benchmark problems; it is also available in the optimization and uncertainty quantification framework DAKOTA (from version 6.7 [23]) where it can be used a stand-alone solver or an approximate subproblem solver. With its derivative-free approach it offers the flexibility and applicability to a wide range of problems. Thus, it is our method of choice. For a more elaborate introduction to the method we refer the interested reader to [11] and will focus on the most recent algorithmic developments next.

C. Multilevel Monte Carlo for the expected value

As stated before, SNOWPAC employs Monte Carlo sampling estimators R to evaluate the measures of risk and robustness $\mathcal{R}$ resulting from the OUU formulation. By having access to a hierarchy of models/fidelities we can speed up the computation by employing newly developed multifidelity and multilevel approaches (see [2, 3]).

Let us first introduce some notation. Given a mapping $f : \mathbb{R}\times\Omega \to \mathbb{R}$ where $\theta \in \Omega$ is a random variable as in Section III.A, we use the shorthand $Q := f(\cdot,\theta)$ or $Q^{(\ell)} := f^{(\ell)}(\cdot,\theta)$ when multiple levels are available. A realization (or sample) is then written as $Q_i := f(\cdot,\theta_i)$ or $Q^{(\ell)}_i := f^{(\ell)}(\cdot,\theta_i)$ where N_ℓ samples are used for level ℓ, such that: $\{Q^{(\ell)}_1,\dots,Q^{(\ell)}_{N_\ell}\} = \{f^{(\ell)}(\cdot,\theta_1),\dots,f^{(\ell)}(\cdot,\theta_{N_\ell})\}$. We employ $\mu_0[Q] := \mathbb{E}[Q]$ for the expected value, while $\mu_i := \mathbb{E}[(Q-\mu[Q])^i], i > 0$, is used for the i-th central moment. If obvious from context, we also allow ourselves to omit the random variable we integrate over, e.g. $\mu_i := \mu_i[Q]$. Additionally, the hat symbol stands for a sampling approximation of the quantity, e.g. $\widehat{\mu_0} \approx \mu_0$ while a multilevel estimator is equipped with a superscript, e.g. $\widehat{\mu_2^{\text{ML}}}$. Furthermore, for the multilevel notation, the current level is shown as superscript, e.g. $\mu^{(\ell)}_4$, while $\mu^{(\ell-1)}_{i,\ell}$ means we approximate $\mu_i, i \geq 0$, on level $(\ell-1)$ using samples from ℓ, e.g. $\widehat{\mu^{(\ell-1)}_{0,\ell}} = \frac{1}{N_\ell}\sum_{i=1}^{N_\ell} f^{(\ell-1)}(\cdot,\theta^{(\ell)}_i) = \frac{1}{N_\ell}\sum_{i=1}^{N_\ell} Q^{(\ell-1)}_{i,\ell}$ and similarly for higher order moments. Finally, all described estimators are unbiased, e.g. $\mu = \mathbb{E}[\widehat{\mu_0}]$, if not explicitly noted otherwise.

Having introduced the notation, the generic multilevel Monte Carlo (MLMC) estimator for a QoI Q at a given highest resolution level L ($\ell = 0$ being the coarsest level) is computed as

$$\mathbb{E}[Q_L] = \mu_0^{\text{ML}}[Q_L] \approx \widehat{\mu_0^{\text{ML}}}[Q_L] = \sum_{\ell=0}^{L}\widehat{\mu_0}[Q^{(\ell)} - Q^{(\ell-1)}] = \sum_{\ell=0}^{L}\frac{1}{N_\ell}\sum_{i=1}^{N_\ell}(Q^{(\ell)}_i - Q^{(\ell-1)}_{i,\ell}), \quad Q^{(-1)}_{i,0} := 0, \tag{5}$$

and its variance is given by

$$\mathbb{V}[\mu_0^{\mathrm{ML}}] = \sum_{\ell=0}^{L} \frac{1}{N_\ell} \mathbb{V}[Q_\ell - Q_{\ell-1}] \approx \mathbb{V}[\widehat{\mu_0^{\mathrm{ML}}}] = \sum_{\ell=0}^{L} \mathbb{V}[\widehat{\mu_0^{(\ell)}} - \widehat{\mu_{0,\ell}^{(\ell-1)}}]. \tag{6}$$

Therefore, for a sequence of levels for which $\mathbb{V}[Q_\ell - Q_{\ell-1}] \to 0$ with $\ell \to L$, it is possible to redistribute the computational load toward the coarser level in order to reach a desired accuracy. Moreover, the optimal allocation of samples $N_\ell^{\mathbb{E}}$ across levels can be obtained in closed form once the variance on each level $\mathbb{V}[Q_\ell - Q_{\ell-1}]$ is estimated. The optimization that we need to solve for the optimal sample allocation as described in [1, 2] is given as

$$\begin{aligned} &\min_{N_\ell^{\mathbb{E}}} \sum_{\ell=0}^{L} C_\ell N_\ell^{\mathbb{E}}, \\ &\text{s.t. } \mathbb{V}[\widehat{\mu_0^{\mathrm{ML}}}] = \epsilon^2. \end{aligned} \tag{7}$$

Here, we minimize the cost while targeting a certain error tolerance ϵ^2. Since the constraint $\mathbb{V}[\widehat{\mu_0^{\mathrm{ML}}}] = \epsilon^2$ is based on the variance of the MLMC estimator for the expected value Eq. (6)*, we further on also refer to solving Eq. (7) as "targeting the mean".This is further symbolized by the superscript $\mathbb{E}$ for the sample allocation $N_\ell^{\mathbb{E}}$. It is known (see e.g. [2]) that the optimization problem has an analytic solution where the method of Lagrange multipliers is used to find the multiplier

$$\lambda = \epsilon^{-2} \sum_{\ell=0}^{L} \sqrt{\mathbb{V}[Q_\ell - Q_{\ell-1}] C_\ell}, \tag{8}$$

and finally the optimal sample allocation for each level

$$N_\ell^{\mathbb{E}} = \left\lceil \lambda \sqrt{\frac{\mathbb{V}[Q_\ell - Q_{\ell-1}]}{C_\ell}} \right\rceil. \tag{9}$$

D. Multilevel Monte Carlo for the variance

In general, for UQ problems we are not only interested in the expected value but also in higher order moments. Specifically in the field of OUU the variance and standard deviation play an important role as mentioned in the section III.A. When optimizing, e.g., over the standard deviation we decrease the variation in our optimal design. Additionally, often times a linear combination of mean and standard deviation is used as probabilistic constraints to ensure a robust solution under uncertainty. Hence, we present here our contribution to find the optimal sample allocation for higher order moments, specifically the variance. To find the optimal sample allocation for the variance $N_\ell^{\mathbb{V}}$ the optimization problem that we have to solve is

$$\begin{aligned} &\min_{N_\ell^{\mathbb{V}}} \sum_{\ell=0}^{L} C_\ell N_\ell^{\mathbb{V}}, \\ &\text{s.t. } \mathbb{V}[\widehat{\mu_2^{\mathrm{ML}}}] = \epsilon^2. \end{aligned} \tag{10}$$

Hence, we need to compute the variance (or mean squared error) for the MLMC estimator of the variance instead of the expected value in Eq. (7). Therefore, we refer to the solution of Eq. (10) as "targeting the variance". This is similarly symbolized by the superscript $\mathbb{V}$ in the sample allocation $N_\ell^{\mathbb{V}}$. This problem is not analytically traceable any longer. Thus, we give closed form solutions for the different terms of the optimization problem in the following and finally revert to solving it numerically to get the sample allocation $N_\ell^{\mathbb{V}}$.

*which is equal to the mean squared error of $\widehat{\mu_0^{\mathrm{ML}}}$ since its an unbiased estimator

First, we introduce the MLMC estimator for the variance by again employing the telescopic sum

$$\begin{aligned}\mathbb{V}[Q_L] \approx \widehat{\mu_2^{\text{ML}}}[Q_L] &= \sum_{\ell=0}^{L} \widehat{\mu_2}[Q^{(\ell)}] - \widehat{\mu_2}[Q^{(\ell-1)}] \\ &= \sum_{\ell=0}^{L} \frac{1}{N_\ell - 1}\Big(\sum_{i=1}^{N_\ell} (Q_i^{(\ell)} - \widehat{\mu_0^{(\ell)}})^2 - (Q_{i,\ell}^{(\ell-1)} - \widehat{\mu_{0,\ell}^{(\ell-1)}})^2 \Big) \\ &= \sum_{\ell=0}^{L} (\widehat{\mu_2^{(\ell)}} - \widehat{\mu_{2,\ell}^{(\ell-1)}}).\end{aligned} \tag{11}$$

Next, using this estimator we decompose the constraint of Eq. (10) in its single terms

$$\mathbb{V}[\widehat{\mu_2^{\text{ML}}}] = \mathbb{V}[\sum_{\ell=0}^{L} (\widehat{\mu_2^{(\ell)}} - \widehat{\mu_{2,\ell}^{(\ell-1)}})] = \sum_{\ell=0}^{L} \mathbb{V}[\widehat{\mu_2^{(\ell)}} - \widehat{\mu_{2,\ell}^{(\ell-1)}}] = \sum_{\ell=0}^{L} \mathbb{V}[\widehat{\mu_2^{(\ell)}}] + \mathbb{V}[\widehat{\mu_{2,\ell}^{(\ell-1)}}] - 2\text{Cov}[\widehat{\mu_2^{(\ell)}}, \widehat{\mu_{2,\ell}^{(\ell-1)}}] \tag{12}$$

where we use independence of the samples over the different levels. Note again the notation $\widehat{\mu_{2,\ell}^{(\ell-1)}}$ where we evaluate the moment of interest on level $(\ell - 1)$ but use samples on level ℓ which results in a dependence expressed by the covariance term.

In the following, we derive the single terms in Eq. (12). The variance of variance is a well-know expression and can be for example found in [24]. The authors, however, only describe a biased estimator and we derive an unbiased estimator for the expression as:

$$\mathbb{V}[\mu_2] \approx \mathbb{V}[\widehat{\mu_2}] = \frac{(N-1)}{N^2 - 2N + 3}\left(\widehat{\mu_4} - \frac{N-3}{N-1}\widehat{\mu_2}^2 \right). \tag{13}$$

This also includes the fourth central moment for which we derive an unbiased single level estimator

$$\widehat{\mu_4} = \frac{1}{(N^2 - 3N + 3) - \frac{(6N-9)(N^2-N)}{N(N^2-2N+3)}}\left(\frac{N^3}{N-1}\widehat{\mu_{4,\text{biased}}} - \frac{(6N-9)(N^2-N)}{N^2-2N+3}\widehat{\mu_2}^2 \right), \tag{14}$$

where

$$\widehat{\mu_{4,\text{biased}}} = \frac{1}{N}\sum_{i=1}^{N} (Q_i - \widehat{\mu_0})^4. \tag{15}$$

Thus, we use single level expressions Eq. (13) and Eq. (14) to evaluate the variance terms of Eq. (12).

Finally, we derive the covariance term as

$$\text{Cov}[\widehat{\mu_2^{(\ell)}}, \widehat{\mu_{2,\ell}^{(\ell-1)}}] = \frac{1}{N_\ell}\mathbb{E}[\widehat{\mu_2^{(\ell)}}\widehat{\mu_{2,\ell}^{(\ell-1)}}] + \frac{1}{N_\ell(N_\ell - 1)}(\widehat{\mu_0^{(\ell)}}[Q^{(\ell)}Q^{(\ell-1)}] - \widehat{\mu_0^{(\ell)}}\widehat{\mu_{0,\ell}^{(\ell-1)}})^2 \tag{16}$$

where

$$\begin{aligned}\mathbb{E}[\widehat{\mu_2^{(\ell)}}\widehat{\mu_{2,\ell}^{(\ell-1)}}] = {} & \widehat{\mu_0^{(\ell)}}\left[{Q^{(\ell)}}^2{Q^{(\ell-1)}}^2\right] - 2\widehat{\mu_0^{(\ell)}}\left[{Q^{(\ell)}}^2 Q^{(\ell-1)}\right]\widehat{\mu_0^{(\ell)}}\left[Q^{(\ell-1)}\right] \\ & + 2\widehat{\mu_0^{(\ell)}}\left[Q^{(\ell-1)}\right]^2\widehat{\mu_0^{(\ell)}}\left[{Q^{(\ell)}}^2\right] - 2\widehat{\mu_0^{(\ell)}}\left[Q^{(\ell)}\right]\widehat{\mu_0^{(\ell)}}\left[Q^{(\ell)}{Q^{(\ell-1)}}^2\right] \\ & + 4\widehat{\mu_0^{(\ell)}}\left[Q^{(\ell-1)}\right]\widehat{\mu_0^{(\ell)}}\left[Q^{(\ell)}\right]\widehat{\mu_0^{(\ell)}}\left[Q^{(\ell)}Q^{(\ell-1)}\right] + 2\widehat{\mu_0^{(\ell)}}\left[Q^{(\ell)}\right]^2\widehat{\mu_0^{(\ell)}}\left[{Q^{(\ell-1)}}^2\right] \\ & - 4\widehat{\mu_0^{(\ell)}}\left[Q^{(\ell)}\right]^2\widehat{\mu_0^{(\ell)}}\left[Q^{(\ell-1)}\right]^2 - \widehat{\mu_0^{(\ell)}}\left[{Q^{(\ell)}}^2\right]\widehat{\mu_0^{(\ell)}}\left[{Q^{(\ell-1)}}^2\right].\end{aligned} \tag{17}$$

Note that the product of (multiple) mean estimators is again a biased estimator. Therefore, we also derived unbiased estimators for double, triple and quadruple products of expected values in Eq. (17) which, however, we do not list explicitly for the sake of compactness of this presentation.

Finally, we close this section with a brief review about work in this field and putting our contribution into context. First work regarding MLMC estimators known to us was published in [13]. While they present estimators for higher order central moments, the authors employ biased estimators. More recent work by [14] leverages h-statistics and symbolic computations to find unbiased closed form solutions for the higher order moments; however, they only approximate the underlying optimization problem for the sample allocation, thereby solving an approximate analytic problem while we use numerical optimization instead.

E. Coupling MLMC, SNOWPAC and DAKOTA

Additionally, in order to couple a given UQ strategy with SNOWPAC, for a generic objective function $\mathbb{E}[Q] \pm \alpha\sigma[Q]$ it is necessary to provide an estimation of the standard error SE for $\mathbb{E}$, $\mathbb{V}$ and σ as estimate for ε^b in Eq. (3) and Eq. (4). While the standard error for the sample mean $\widehat{\mu_0}$ is simply obtained as $\sqrt{\mathbb{V}[\widehat{\mu_0}]} \approx \sqrt{\frac{\widehat{\mu_2}}{N}}$, we can now use the derived quantities from the previous chapter to also compute the standard error for the variance as given in Eq. (12) for the multilevel estimator and Eq. (13) for a single level.

Finally, the standard error for the sample standard deviation $\sigma[Q_L] \approx \sqrt{\widehat{\mu_2}^{(L)}}$ is approximated by

$$SE(\sigma[Q_L]) \approx \frac{1}{2\sqrt{\mu_2^{(L)}}}\sqrt{\mathbb{V}[\mu_2^{(L)}]} \tag{18}$$

for a single level and for $\sigma[Q_L] \approx \sqrt{\widehat{\mu_2^{\mathrm{ML}}}}$ by

$$SE(\sigma[Q_L]) \approx \frac{1}{2\sqrt{\widehat{\mu_2^{\mathrm{ML}}}}}\sqrt{\mathbb{V}[\widehat{\mu_2^{\mathrm{ML}}}]} \tag{19}$$

by employing the so-called Delta Method. This enables us to approximate the probability distribution of a function (the root mean square for us) of an asymptotically normal estimator, i.e. the sample variance in our case. Of course, in our case the sample variance is not an asymptotically normal estimator, therefore this is only an approximation.

The developed MLMC algorithms and estimators are implemented in DAKOTA [23], a software that offers state-of-the-art research and robust, usable algorithms for optimization and UQ. It furthermore includes SNOWPAC as an external solver. By its modular design, a number of surrogate models can be employed with the optimization algorithm of the user's choice. Applications can be easily coupled, e.g., through the exchange of input and output files. This offers a powerful tool to utilize and exchange a multitude of algorithms in a straight-forward manner.

IV. OUU results

After having introduced the new algorithmic developments regarding higher moment MLMC estimators, we show their computational efficiency on two problems. In Section IV.A we work with an analytic OUU example which also serves as a verification case for the approach. Afterwards, we continue to the two turbine wind problem in Section IV.B to show the effectiveness of the method for a real word problem.

A. Analytical test problem

Before approaching the complex wind application, we perform a series of tests to verify both the implementation and the performance of the OUU algorithm. The test is based on problem 18 of a collection of optimization problems given in [25] and has been modified by adding a constraint. Given the following deterministic functions

$$f(x) = \begin{cases} (x-2)^2 & \text{if } x \le 3 \\ 2\log(x-2)+1 & \text{if } x > 3 \end{cases}, \tag{20}$$

$$g(x) = \tfrac{2.\log(1.5)}{2.5}(x-1)+1 \tag{21}$$

which we also illustrate in the left figure of Figure 1 we define the deterministic optimization problem:

$$\begin{aligned} \min_x \quad & f(x), \\ \text{s.t.} \quad & c(x) := g(x) - f(x) \le 0. \end{aligned} \tag{22}$$

Next, we add a uniform random variable ξ to the constraint function c based on a parameter A to obtain a traceable multilevel OUU formulation:

$$\begin{aligned} c_H(x,\xi) &:= g_H(x,\xi) - f(x) = g(x) + \xi^3 - f(x), \\ c_L(x,\xi) &:= g_L(x,\xi) - f(x) = g(x) + A\xi^3 - f(x), \xi \sim \mathcal{U}(-0.5, 0.5), \end{aligned} \tag{23}$$

where H describes the fine and L describes the coarse level, respectively. A is a control parameter that correlates the two levels which we modify in order to obtain an interesting problem with desired features: this gives us a verification case, not only with a control parameter for the correlation over the two levels but also an analytical solution since $\mathbb{E}[c_{H/L}(x,\xi)] = g(x) - f(x)$, $\mathbb{V}[c_H] = \frac{0.5^6}{7}$ and $\mathbb{V}[c_L] = A^2\frac{0.5^6}{7}$.

Finally, we define two robust optimization problems to see the effect of the newly developed MLMC estimators

- Mean:

$$\begin{aligned} \min_x \quad & f(x), \\ \text{s.t.} \quad & \mathcal{R}_{\text{Mean}} := \mathbb{E}[c_H(x,\xi)] \leq 0. \end{aligned} \tag{24}$$

- Mean plus push back:

$$\begin{aligned} \min_x \quad & f(x), \\ \text{s.t.} \quad & \mathcal{R}_{\text{Pback}} := \mathbb{E}[c_H(x,\xi)] + 3\sigma[c_H(x,\xi)] \leq 0. \end{aligned} \tag{25}$$

Here, we evaluate the robustness measure using either Monte Carlo sampling on c_H as reference or MLMC sampling with an appropriate combination of samples on c_H and c_L.

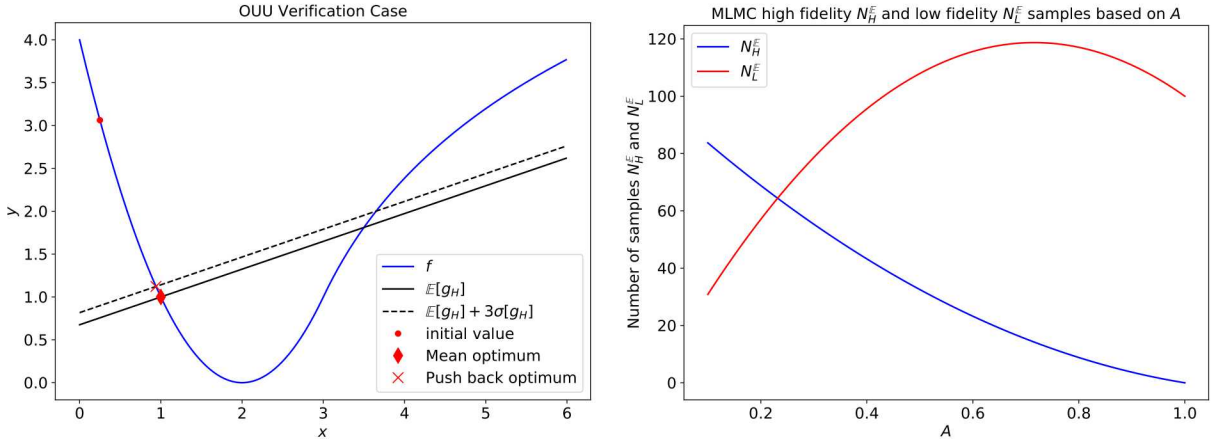

Fig. 1 Test functions f and g in the interval $x \in [0,6]$ (left). MLMC profile for the number of samples N_H and N_L given parameter A. (right)

To find the optimal MLMC profile we first solve problem reported in Eq. (7) where we use a Monte Carlo estimator on c_H with $N_H = 100$ samples as reference case and a cost ratio of 10 between the two levels. Therefore, we set $\epsilon^2 = \mathbb{V}[c_H]/100$ and solve the optimization problem which returns a sample allocation dependent on A. In the right plot of Figure 1, we plot the number of samples N_H and N_L in relation to A. Based on this profile we pick $A = 0.7$ which results in a sample allocation of $N_H^{\mathbb{E}} = 16$ and $N_L^{\mathbb{E}} = 119$. Additionally, we compute a second sample profile by solving Eq. (10) numerically since we not only have to estimate the expected value but also the standard deviation for Eq. (25). This returns the sample allocation $N_H^{\mathbb{V}} = 38$ and $N_L^{\mathbb{V}} = 108$ for $A = 0.7$ which tells us to redistribute samples to the fine level to properly resolve the variance. Though the cost are theoretical here, this results in total cost reduction of 70.5% for $N^{\mathbb{E}}$ and 47.4% for $N^{\mathbb{V}}$ compared to $N_H = 100$ samples of the Monte Carlo reference solution.

To compare the two samples allocations in a first verification test we sample the stochastic moments $\mathbb{E}$, $\mathbb{V}$ and σ 5000 times over ξ at location $x = 3$ and compare the MC and MLMC estimation given the sample allocations. The resulting histograms are visualized in Figure 2. While the MLMC estimator (red) for the expected value (left) matches the Monte Carlo estimator (blue) we see a discrepancy for the variance (center) and standard deviation (right). This is reasonable since the sample profile was chosen to match the expected value of the Monte Carlo estimator following Eq. (7). However, this does not ensure that also the variance or standard deviation are matching. We can clearly see a better match for the variance and standard deviation for the sample allocation targeting Eq. (10) (green). However, we now over-resolve for the expected value due to the higher sample sizes on the fine grid.

If we now compare the two optimization benchmarks given in Eq. (24) and Eq. (25) we can see how the sample allocation influences the optimization. For this test we run 100 optimizations for Eq. (24) and Eq. (25) starting from the

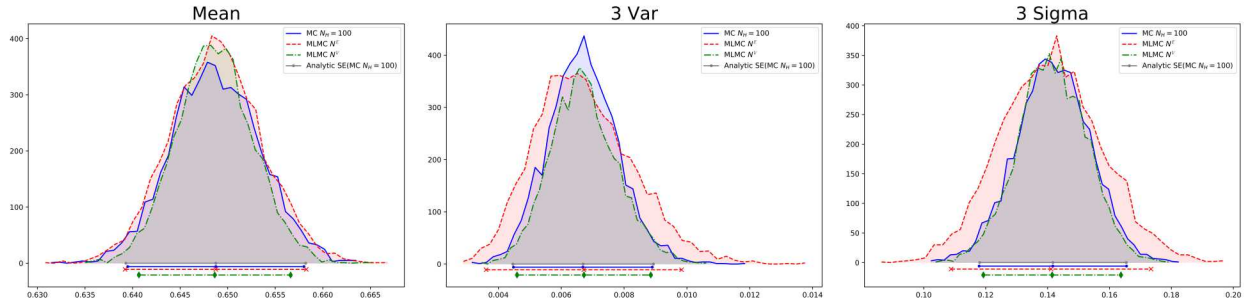

Fig. 2 Histogram plot over 5000 test runs evaluating the expected value $\mathbb{E}$ (left), the variance $\mathbb{V}$ and standard deviation σ of the stochastic constraint. The Monte Carlo estimator using 100 samples is shown in (solid) blue. The MLMC estimator optimized following Eq. (7) using samples $N_H^{\mathbb{E}}$ and $N_L^{\mathbb{E}}$ is plotted in (dashed) red. The MLMC estimator when targeting the variance as in Eq. (10) employing the sample allocation $N_H^{\mathbb{V}}$ and $N_L^{\mathbb{V}}$ is shown in (dot dashed) green. The flat lines on the bottom show two times the standard error as well as the mean of the histogram while the grey top most line reflect the analytic error.

same initial value $x_0 = 0.25$, different random seed and comparing MC and MLMC estimators. In Figure 3 we plot the final designs obtained for the optimization runs. Here, we show that the discrepancy when higher order moments are evaluated using the MLMC estimator based on Eq. (7) is visible. While we observe the final designs obtained for Eq. (24) matching for Monte Carlo (red diamond) and MLMC (blue cross), MLMC performs poorly for Eq. (25) compared to the designs found by Monte Carlo. Here, although we would like to see the MLMC designs match the Monte Carlo designs, they are much more infeasible when only targeting Eq. (7). Using the sample allocation obtained from Eq. (10) (green plus) we see an improved performance in the right plot of Figure 3 where the designs are more feasible and located around the analytic optimal solution. Therefore, we conclude that an appropriate sample allocation is necessary when including higher order moments—like the standard deviation—in the OUU problem formulation, still leveraging a cost reduction compared to classical Monte Carlo.

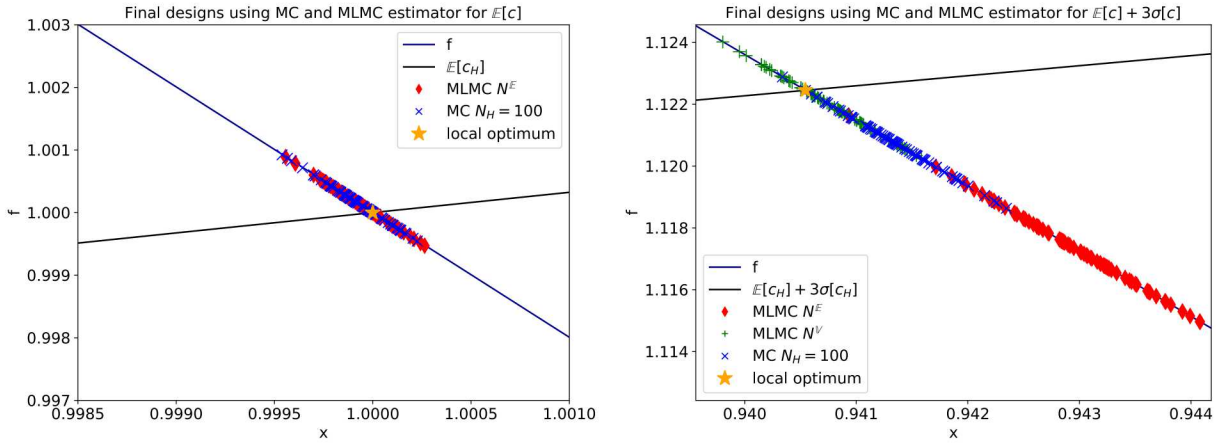

Fig. 3 Final designs obtained after 100 optimization runs starting from the same initial design $x_0 = 0.25$ for formulation (24) (left) and (25) (right). The blue line shows the deterministic objective function f, the black line shows the analytic solution of the stochastic constraint. The blue crosses show the final designs using a classic Monte Carlo estimator with 100 samples on the final level. The red diamonds show the final designs using a MLMC estimator following (7). The MLMC estimator targeting the variance (10) is shown as green pluses. The orange star shows the respective optimal solution.

B. OUU for Wake Steering

After showing the general validity of the approach on the test problem we are returning to the wind application. For this application, we are considering a scenario involving two turbines arranged such that the upwind turbine partially wakes the downwind turbine. The turbines are represented as the NREL 5 MW [26] machine, are spaced apart 600 m in the streamwise direction, and offset by half a rotor diameter in the spanwise direction. We assumed an average 8 m/s hub height velocity with the turbines operating at a fixed axial induction factor of 1/3. This partial wake overlap situation commonly occurs in wind farms, and is know to have a clear wake steering strategy in the deterministic case. The partial offset introduces an asymmetry in the power response to yaw motions that avoids a bimodal situation that would otherwise occur when using the non-rotating actuator disks in WindSE in cases without spanwise offsets. Furthermore, we treat the prescribed yaw angles γ_i as the optimization variables, and assume uncertainty in the incoming velocity θ_u as well as random additive noise at each turbine θ_{γ_i} as done in previous studies. The noise in yaw angles is attributed to errors in the wind vane sensors, yaw positioning system, and unresolved turbulent fluctuations.

We consider two different robust optimization problems where we are able to control the yaw angle of the turbines while we model uncertainty in the yaw angle sensors as well as the velocity measurements of the inflowing wind. Finally, we are interested in the combined power production of the two turbines as described in Section II. Similar to Section IV.A we maximize for the expected value of the power production only; or, maximize for its expected value while trying to minimize its variance. The two optimization problems are defined as follows:

- Mean:

$$\max_{\gamma_1,\gamma_2} \mathcal{R}_{\text{Mean}} := \max_{\gamma_1,\gamma_2} \mathbb{E}[f_{\text{power}}\left(\gamma_1, \gamma_2, \theta_u, \theta_{\gamma_1}, \theta_{\gamma_2}\right)]. \tag{26}$$

- Mean plus push back:

$$\max_{\gamma_1,\gamma_2} \mathcal{R}_{\text{Pback}} := \max_{\gamma_1,\gamma_2} \mathbb{E}[f_{\text{power}}\left(\gamma_1, \gamma_2, \theta_u, \theta_{\gamma_1}, \theta_{\gamma_2}\right)] - 3\sigma[f_{\text{power}}\left(\gamma_1, \gamma_2, \theta_u, \theta_{\gamma_1}, \theta_{\gamma_2}\right)]. \tag{27}$$

Here, $\gamma_i \in [-45°, 45°]$, $i = 1, 2$ are box constraints for the design variables and $\theta_u \sim \mathcal{U}(7.6\frac{m}{s}, 8.4\frac{m}{s}), \theta_{\gamma_i} \sim \mathcal{U}(-5°, 5°), i = 1, 2$ the uniform distributions for the uncertain parameters.

We employ a hierarchy of three different grid resolutions COARSE, MEDIUM and FINE for the multilevel case. The different grids and a velocity field snapshot for each of the grids are visualized in Figure 4 using angles of the initial design of $\gamma_1 = 0°$ and $\gamma_2 = 0°$. Starting from different base resolutions, each of the grids uses additional refinement along a channel following the flow around the turbines and is most refined around the turbines themselves to capture the flow field and wake. The degrees of freedom of each of the grids and the absolute and relative cost are given in Figure 1. Here, the runtimes were averaged over 20 pilot runs over the uncertain space at the centered value $\gamma_i = 0°, i = 1, 2$.

	COARSE	MEDIUM	FINE
Degrees of freedom	12548	86364	242788
Absolute cost [s]	17.288	133.41	531.567
Approximate relative cost	1	8	31

Table 1 Degrees of freedom, averaged absolute cost in seconds and relative cost with respect to COARSE for the three grid employed grid resolutions COARSE, MEDIUM and FINE for a single evaluation.

Regarding the sample allocation we solve the optimization problems Eq. (7) and Eq. (10) with respect to a Monte Carlo reference solution with $N_{\text{FINE}} = 150$ samples. Hence, we search for an MLMC estimator matching the same accuracy but at lower computational cost. Using the previously mentioned 20 pilot runs we compute the sample allocations $N^{\mathbb{E}} = [223, 7, 1]$ for Eq. (7). For this case we consider the final profile of $N^{\mathbb{E}} = [223, 7, 5]$ to account for a realistic number (five) of pilot samples at the fine level. When targeting the variance we compute the sample allocation $N^{\mathbb{V}} = [396, 46, 20]$ numerically solving Eq. (10). As we can observe, to target the variance we again need more samples. While the Monte Carlo reference using $N_{\text{FINE}} = 150$ samples has a total cost of $C_{N_{\text{FINE}}} = 4650$ for a single optimization step, the MLMC estimators reduce it to $C_{N^{\mathbb{E}}} = 481$ and $C_{N^{\mathbb{V}}} = 1590$, respectively. Hence, we again leverage the resolution hierarchy for a more efficient computational performance.

We present the optimization results in Figure 5. We plot here the optimization progress of the objective function that we maximize (right) as well as the change of the design parameters $\gamma_i, i = 1, 2$ (left). We compare results using a MLMC estimator in the optimization against a Monte Carlo reference estimator employing 150 samples on FINE; additionally,

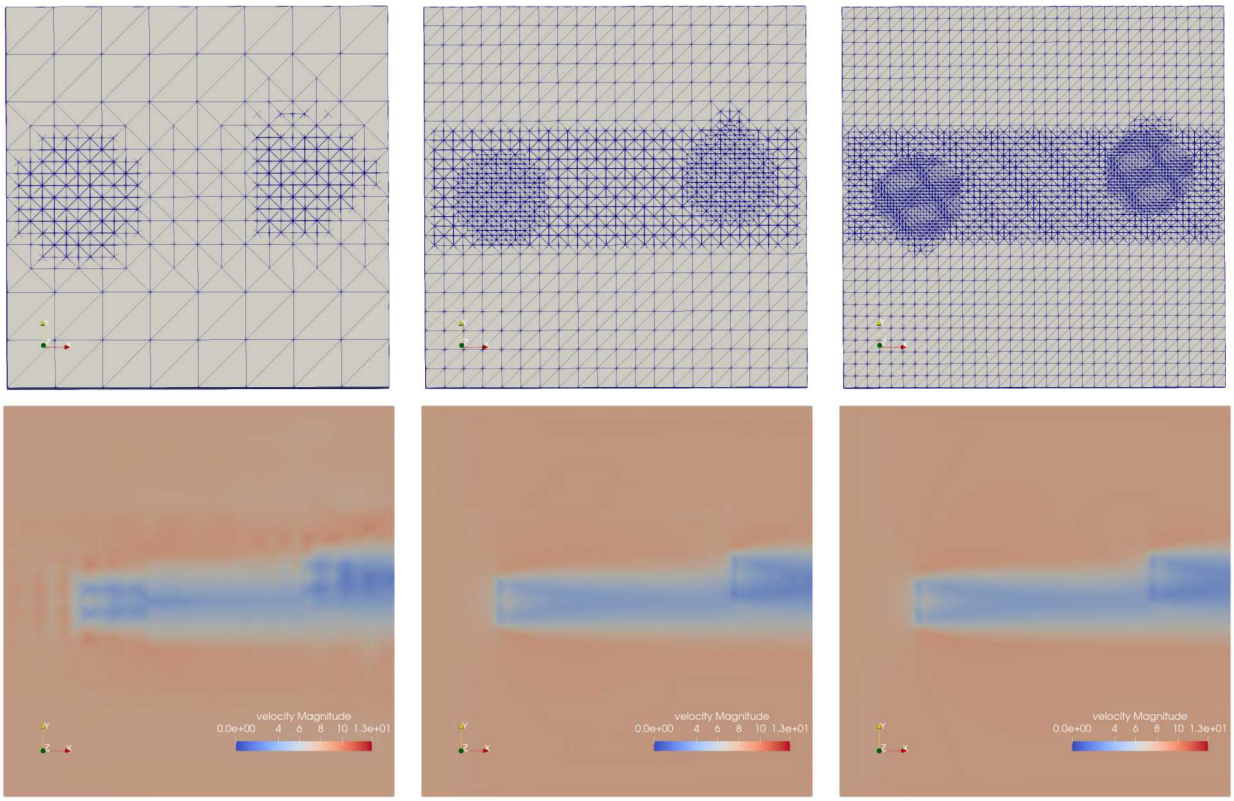

Fig. 4 Grid resolutions employed for the multilevel case (top row) and example velocity fields for initial design $\gamma_1 = 0°$ and $\gamma_2 = 0°$ (bottom row) visualized on the three different resolutions from left to right: COARSE, MEDIUM and FINE

we compare to the deterministic problem using the nominal values of the uncertain parameters, i.e. $\theta_{\gamma_i} = 0°, i = 1, 2$ and $\theta_u = 8\frac{m}{s}$. In the following plots, the gray lines extend the converged the deterministic solution as well as the Monte Carlo reference solutions that stopped because of limited computational resources.

When we consider the first OUU formulation Eq. (26) we can see that the Monte Carlo reference solution (blue) finds a similar optimal design as the deterministic solution (black) in Figure 5. The MLMC estimator (red) finds a similar objective value but at a different design with additional optimization steps though the solution does not seem fully converged yet. However, we deduce a similar performance of the Monte Carlo and MLMC case when targeting the mean. This is reasonable since they are designed for the same accuracy (neglecting the extra samples for numerical stability).

Regarding the second formulation Eq. (27) we observe the effect of the added standard deviation term in Figure 6. Now, the MLMC case targeting the mean (red) clearly has problems and is underresolved. Looking at the MLMC estimator targeting the variance (green), however, we clearly see a better match with the Monte Carlo reference solution (blue) in the design as well as the objective value. This again, as in the previous sections, is due to the design of the sample allocation and shows the importance of employing Eq. (10) when a variance or standard deviation has to be estimated. Therefore, similar to the previous test problem in Figure IV.A, the appropriate sample allocation ensures a better performance during the optimization despite a slight additional cost compared to $N^{\mathbb{E}}$; the cost reduction to standard Monte Carlo is more significant.

To further investigate the final designs found we sample each of the solutions over the uncertain space. In Figure 7 we plot histograms for 200 sample runs. We compare the final distributions for formulations reported in Eq. (26) (left) and Eq. (27) (right) where each plot shows the distributions at the initial design (gray), the optimal deterministic solution (black) as well the optimal solutions found using Monte Carlo (blue) and MLMC (red for $N^{\mathbb{E}}$, green for $N^{\mathbb{V}}$). Additionally, the vertical lines show the mean value while the horizontal lines show the standard deviation for the respective solution. Note, that we plot the histogram value of f_{power}, not the objective value.

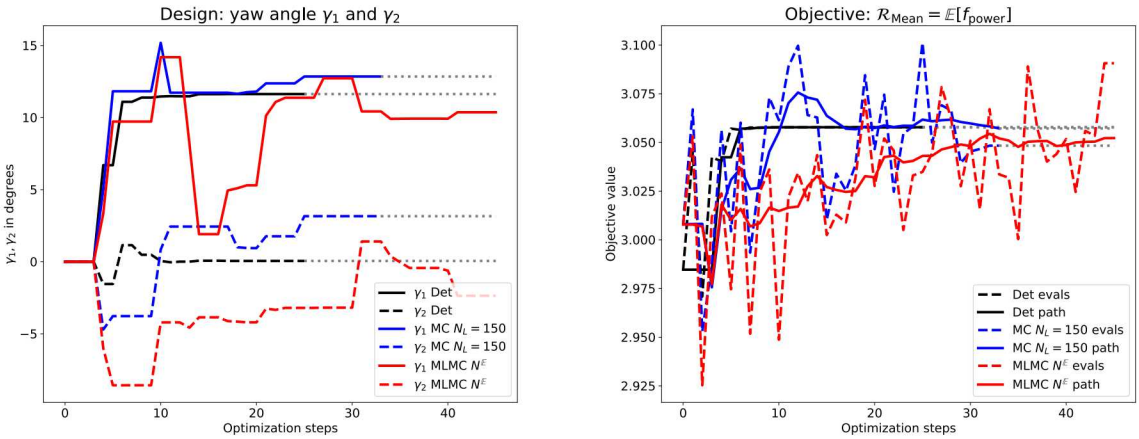

Fig. 5 Optimization path (left) and optimal design path (right) using Monte Carlo and MLMC estimator for formulation (26)

First, we observe that all designs show approximately a uniform distribution in the quantity of interest. Additionally, we clearly see the improvement in the expected value of the solution due to the optimization. Furthermore, all optimal designs result in a similar final distribution and expectation value. Moreover, also the variance is approximately equal. Thus, we cannot observe a significant effect from adding the standard deviation term in formulation Eq. (27). It seems that the variance over the design does not change considerably to enforce different optimal robust solution. Therefore, the solution of Eq. (27) returns to the optimal solution of Eq. (26).

Apart from observing the performance of the optimizer qualitatively we state the final design values for each of the approaches in Table 2. Additionally, we state the approximated expected value as well as the standard deviation of f_{power} at the corresponding designs. These values are computed from the 200 samples also used for the histograms. We see that all approaches find a similar objective value in expectation. Additionally, we notice a similar standard deviation value. This further strengthens the argument that the standard deviation seems constant over the design space and explains the similar results between the formulations reported in Eq. (26) and Eq. (27).

			$\mathcal{R}_{\text{Mean}}$		$\mathcal{R}_{\text{Pback}}$	
	Initial	Det	MC	MLMC	MC	MLMC $N^{\mathbb{V}}$
γ_1	0.0	11.6341	12.8512	10.3720	17.6596	12.9430
γ_2	0.0	0.0570	3.1616	-2.3608	-2.8550	-1.8230
$\mathbb{E}[f_{\text{power}}(\gamma_1, \gamma_2, \theta_u, \theta_{\gamma_1}, \theta_{\gamma_2})]$	2.9851	3.0558	3.0506	3.0529	3.0297	3.0534
$\sigma[f_{\text{power}}(\gamma_1, \gamma_2, \theta_u, \theta_{\gamma_1}, \theta_{\gamma_2})]$	0.2490	0.2514	0.2512	0.2513	0.2496	0.2511
Objective	2.9851	3.0558	3.0506	3.0529	2.2808	2.3001

Table 2 Optimization results for the different methods compared to the initial condition. All values are rounded to four digits.

Finally, we show three flow fields for the velocity magnitude computed for the final design at the nominal values for the deterministic (left), mean (center) and mean plus push back OUU formulation (right). The results are visualized in Figure 8. Again, there is no major difference visible and all cases try to steer mainly the wake of the first turbine to optimize the total power production.

We summarize this section by the following three observations and conclusions: First, we see an advantage of employing MLMC to leverage a hierarchy of available resolutions compared to standard Monte Carlo and therefore reducing the computational cost. Second, when employing higher order moment in the problem formulation, an adaptation of the MLMC estimator following the description in Section III.D. Third, for this application we do not see a direct advantage of using a OUU formulation spending the additional computational resources while the deterministic

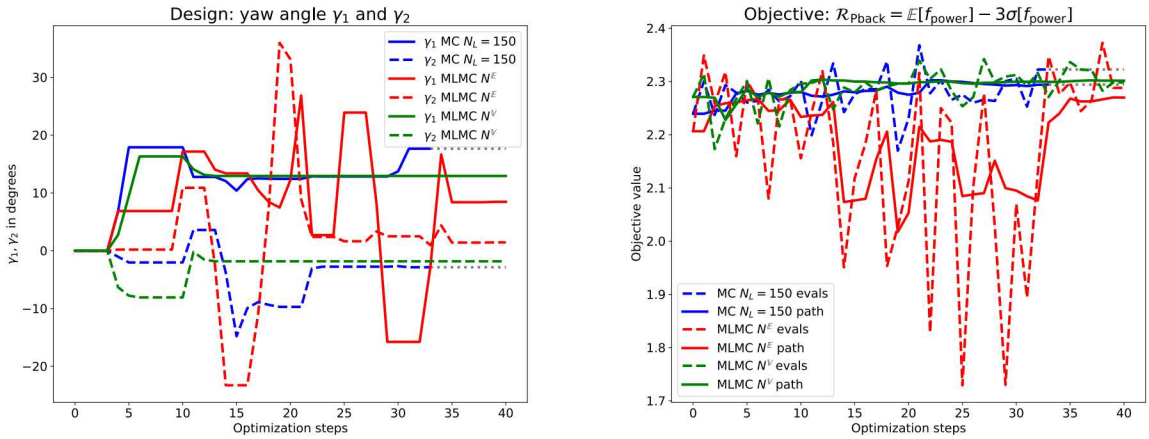

Fig. 6 Optimization path (left) and optimal design path (right) using Monte Carlo and MLMC estimator for formulation (27)

solution already seems quite promising. Therefore, we envision to move to more challenging applications, e.g. adding more turbines to the problem and, thus, enlarging the design as well as the uncertain space or add additional uncertain as well as design parameters and nonlinear constraints.

V. Conclusions

Wind plant design and control is a formidable engineering problem that requires the use of accurate and computationally expensive high-fidelity simulations to capture the interactions among the wakes and the rotors. In this work we will focus on the wake steering control approach that enables to increase the power extraction of a wind plant already installed by just modify the yaw angle of each turbine. By changing the yaw angle of a turbine its wake can be tilted as well in order to minimize the loss of power caused the wake impinging on downstream turbines. The difficulty in measuring accurately the yaw angle of each turbine by using sensors mounted on the nacelle motivates the use of an optimization under uncertainty approach. In particular, we propose an efficient multilevel strategy that combines efficient multilevel forward propagation with the derivative-free stochastic optimization method SNOWPAC in Dakota that enables us to either accelerate this workflow by requiring only a limited number of high-fidelity simulations complemented by a large number of lower fidelity system realizations or improve accuracy while having similar computational cost compared to standard Monte Carlo forward propagation. In this paper we have reported the performance of the proposed approach with respect to its single fidelity counterpart based on MC sampling. We employed both verification test cases as well as a more challenging, albeit quite simplified, wind turbine test case based on RANS presented in [15]. For this particular optimization formulation we have observed an improvement of the performance for the optimized configuration with respect to the nominal configuration. However, we have not observed a substantial difference between the deterministic and OUU formulation. In the future we plan to consider more challenging test cases, *e.g.* wind plant configurations with a higher number of turbines, for which we expect a greater non-linearity of the response and therefore a significant difference between the deterministic and OUU formulations.

Acknowledgments

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525. The views expressed in the article do not necessarily represent the views of the U.S. Department Of Energy or the United States Government. This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Wind Energy

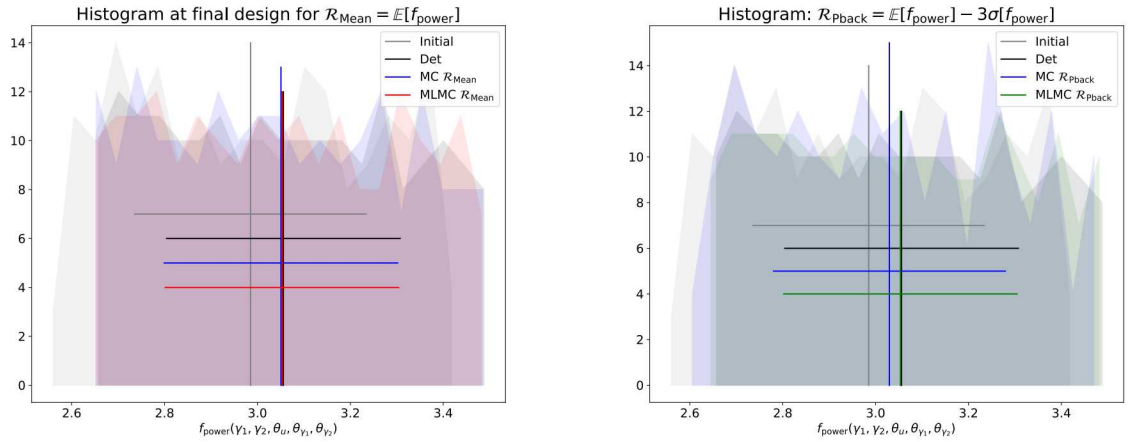

Fig. 7 Histogram for 200 sample runs over the uncertain space at the respective final design for the formulations reported in Eq. (26) (left) and Eq. (27) (right).

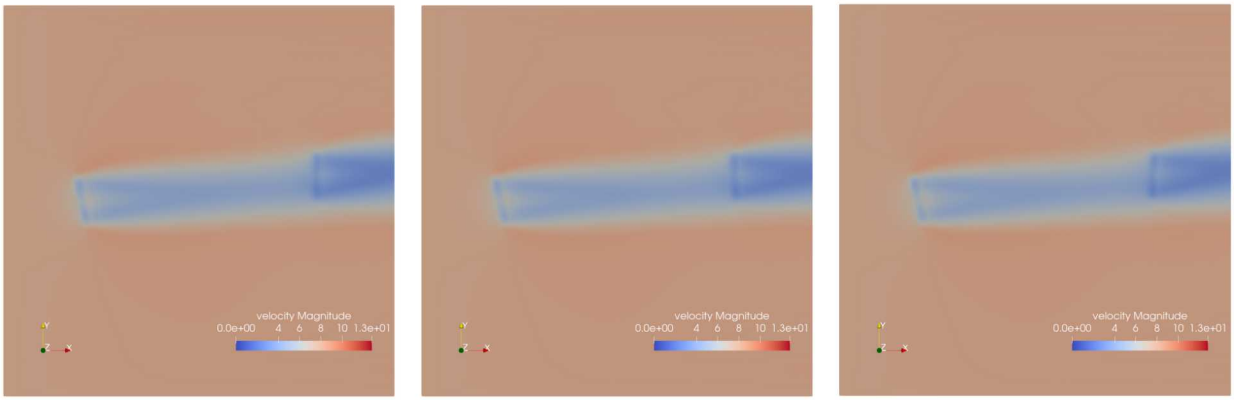

Fig. 8 Flowfield for velocity magnitude for final design found for deterministic (left), mean (26) (center) and pushback (27) (right) formulation.

Technologies Office. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

References

[1] Heinrich, S., "Multilevel Monte Carlo Methods," *Large-Scale Scientific Computing*, edited by S. Margenov, J. Waśniewski, and P. Yalamov, Springer Berlin Heidelberg, Berlin, Heidelberg, 2001, pp. 58–67.

[2] Giles, M. B., "Multilevel Monte Carlo Path Simulation," *Operations Research*, Vol. 56, No. 3, 2008, pp. 607–617. doi: 10.1287/opre.1070.0496, URL http://pubsonline.informs.org/doi/abs/10.1287/opre.1070.0496.

[3] Giles, M. B., "Multilevel Monte Carlo methods," *Acta Numerica*, Vol. 24, 2015, p. 259–328. doi:10.1017/S096249291500001X.

[4] Pasupathy, R., Schmeiser, B. W., Taaffe, M. R., and Wang, J., "Control-variate estimation using estimated control means," *IIE Transactions*, Vol. 44, No. 5, 2012, pp. 381–385.

[5] Ng, L., and Willcox, K., "Multifidelity approaches for optimization under uncertainty," *Int. J. Numer. Meth. Engng.*, Vol. 100, No. 10, 2014, pp. 746–772.

[6] Peherstorfer, B., Willcox, K., and Gunzburger, M., "Optimal model management for multifidelity Monte Carlo estimation," *SIAM Journal on Scientific Computing*, Vol. 38, No. 5, 2016, pp. A3163–A3194.

[7] Geraci, G., Iaccarino, G., and Eldred, M. S., "A multi fidelity control variate approach for the multilevel Monte Carlo technique." *CTR Annual Research Briefs 2015*, 2015, pp. 169–181.

[8] Geraci, G., Eldred, M. S., and Iaccarino, G., "A multifidelity multilevel Monte Carlo method for uncertainty propagation in aerospace applications," *19th AIAA Non-Deterministic Approaches Conference*, 2017, p. 1951.

[9] Fairbanks, H., Doostan, A., Ketelsen, C., and Iaccarino, G., "A low-rank control variate for multilevel Monte Carlo simulation of high-dimensional uncertain systems," *Journal of Computational Physics*, Vol. 341, 2017, pp. 121–139.

[10] Eldred, M., Giunta, A., and Collis, S., "Second-order corrections for surrogate-based optimization with model hierarchies," *10th AIAA/ISSMO multidisciplinary analysis and optimization conference*, 2004, p. 4457.

[11] Augustin, F., and Marzouk, Y. M., "A Trust-Region Method for Derivative-Free Nonlinear Constrained Stochastic Optimization," *arXiv:1703.04156*, 2017.

[12] Geraci, G., Menhorn, F., Huan, X., Safta, C., Marzouk, Y., Najm, H., and Eldred, M., "Progress in Scramjet Design Optimization Under Uncertainty Using Simulations of the HIFiRE Configuration," 2019. doi:10.2514/6.2019-0725.

[13] Bierig, C., and Chernov, A., "Estimation of arbitrary order central statistical moments by the multilevel Monte Carlo method," *Stochastics and Partial Differential Equations Analysis and Computations*, Vol. 4, No. 1, 2016, pp. 3–40. doi:10.1007/s40072-015-0063-9, URL https://doi.org/10.1007/s40072-015-0063-9.

[14] Pisaroni, M., Krumscheid, S., and Nobile, F., "MATHICSE Technical Report : Quantifying uncertain system outputs via the multilevel Monte Carlo method - Part I: Central moment estimation," 2017, p. 29. doi:10.5075/epfl-MATHICSE-263564, URL http://infoscience.epfl.ch/record/263564, mATHICSE Technical Report Nr. 23.2017 October 2017.

[15] Quick, J., Annoni, J., King, R., Dykes, K., Fleming, P., and Ning, A., "Optimization Under Uncertainty for Wake Steering Strategies," *Journal of Physics: Conference Series*, Vol. 854, 2017, p. 012036. doi:10.1088/1742-6596/854/1/012036, URL https://doi.org/10.1088%2F1742-6596%2F854%2F1%2F012036.

[16] King, R. N., Dykes, K., Graf, P., and Hamlington, P. E., "Optimization of wind plant layouts using an adjoint approach," *Wind Energy Science*, Vol. 2, No. 1, 2017, pp. 115–131. doi:10.5194/wes-2-115-2017, URL http://www.wind-energ-sci.net/2/115/2017/.

[17] Augustin, F., and Marzouk, Y. M., "A Path-Augmented Constraint Handling Approach for Nonlinear Derivative-Free Optimization," *arXiv:1403.1931v3*, 2014.

[18] Kannan, A., and Wild, S. M., "Obtaining Quadratic Models of Noisy Functions," Tech. rep., Argonne National Laboratory, Illinois, 2012.

[19] Rasmussen, C. E., and Williams, C. K. I., *Gaussian Processes for Machine Learning*, MIT Press, 2006.

[20] J. D. Powell, M., "A View of Algorithms for Optimization Without Derivatives," *Mathematics TODAY*, Vol. 43, No. 5, 2007.

[21] Le Digabel, S., "NOMAD: Nonlinear optimization with the MADS algorithm," *ACM Trans. Math. Softw.*, Vol. 37, 2011, p. 44.

[22] Gardner, J. R., Kusner, M. J., Xu, Z., Weinberger, K. Q., and Cunningham, J. P., "Bayesian Optimization with Inequality Constraints," *Proceedings of the 31st International Conference on International Conference on Machine Learning - Volume 32*, JMLR.org, 2014, pp. II–937–II–945. URL http://dl.acm.org/citation.cfm?id=3044805.3044997.

[23] Adams, B., Bauman, L., Bohnhoff, W., Dalbey, K., M.S., E., J.P., E., M.S., E., Hough, P., Hu, K., Jakeman, J., Stephens, J., Swiler, L., Vigil, D., and Wildey, T., "DAKOTA, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis - Version 6.6 Users Manual," , July 2014. Updated May, 2017.

[24] Mood, A., Graybill, F., and Boes, D., *Introduction to the Theory of Statistics*, International Student edition, McGraw-Hill, 1974. URL https://books.google.com/books?id=Viu2AAAAIAAJ.

[25] Gavana, A., "Infinity 77," http://infinity77.net/global_optimization/test_functions_1d.html, 2013. [Online; accessed June 15th, 2019].

[26] Jonkman, J., Butterfield, S., Musial, W., and Scott, G., "Definition of a 5-MW Reference Wind Turbine for Offshore System Development," Technical Report NREL/TP-500-38060, National Renewable Energy Laboratory, Feb. 2009. URL http://www.osti.gov/bridge/servlets/purl/947422-nhrlni/.