6th European Conference on Computational Mechanics (ECCM 6)
7th European Conference on Computational Fluid Dynamics (ECFD 7)
1115 June 2018, Glasgow, UK

LEVERAGING ACTIVE DIRECTIONS FOR EFFICIENT MULTIFIDELITY UNCERTAINTY QUANTIFICATION

GIANLUCA GERACI[1], MICHAEL S. ELDRED[1], ALEX A. GORODETSKY[2], JOHN D. JAKEMAN[1]

[1] Sandia National Laboratories
PO BOX 5800, MS 1318, Albuquerque, NM 87185, USA
ggeraci@sandia.gov, mseldre@sandia.gov, jdjakem@sandia.gov

[2] University of Michigan, Dept. of Aerospace Engineering
1320 Beal Avenue, Ann Arbor, MI 48109, USA
goroda@umich.edu

Key words: Uncertainty Quantification, Sampling, Multifidelity, Active Subspaces

Abstract. In this work, we focus on sampling-based methods for forward uncertainty propagation, and in particular, on multifidelity sampling methods that employ a control variate approach. The novel component of this work is to accelerate the sampling by relying on important directions, as for instance in Active Subspaces. The idea consists of discovering the important directions for each model in order to provide a reduced dimensional link between the parameterization of each model fidelity and the relevant quantities of interest. This accomplishes two goals: *(i)* enhancing the correlation between models, and *(ii)* providing a mechanism to effectively bridge dissimilar input parameterizations. We demonstrate the performance of our approach for two model problems.

1 INTRODUCTION

Uncertainty Quantification (UQ) is a key component within modern computational science efforts, as it enhances the predictive utility of numerical simulations in support of scientific discovery and advanced engineering design. In recent years, multifidelity UQ has been demonstrated to be an effective strategy to accelerate UQ for complex problems whenever less expensive lower accuracy models can be built for the original problem. In this work, we focus intially on Monte Carlo sampling-based methods, and in particular the multifidelity control variate method [8, 7], but the extension to the multilevel Monte Carlo method [5] and the multilevel-multifidelity Monte Carlo strategy [4] is also possible and a straightforward extension. The novel idea introduced in this work is to accelerate the previous sampling techniques by relying on important directions, as for instance in Active Subspaces [1] and related techniques (ridge functions, proper orthogonal decomposition, etc.). The core of the idea is straightforward and consists in discovering the important directions while performing the usual pilot runs for the multifidelity sampling. Once the important directions (distinct for each model) are discovered, they are used to define a shared space among all model fidelities on which the sampling procedures can

be performed. This is expected to provide two advantages: *(i)* the correlation between the quantity of interest (QoI) predictions for different models can be enhanced by emphasizing active directions along which these QoI have the most variation, and *(ii)* since this approach reduces an original dimensionality to a subspace for each model that can be shared, it provides a bridging mechanism for the case where different fidelities have dissimilar input parameterizations.

In this work, we briefly present the idea and demonstrate its performance on two test problems. Additional application examples will be presented during the Conference. The paper is structured as follows. The multifidelity control variate (CV) approach is introduced in Sec. 2, and key features of the Active Subspace (AS) approach are presented in Sec. 3. The combination of CV and AS is discussed in Sec. 3.2 and some numerical examples are presented in Sec. 4. Final remarks close the paper in Sec. 5.

2 MULTIFIDELITY SAMPLING METHODS

In this section, we briefly present the control variate approach, which provides an example multifidelity technique where our strategy can be deployed. The control variate is well-known in statistics; it provides a strategy to accelerate a Monte Carlo (MC) simulation by reducing its estimator variance. In the classical CV approach, the expected value of the correlated estimate (e.g., from a low-fidelity model) is known and it is possible to obtain a reduction of the variance of the standard MC estimator, which is only a function of the correlation between the two models. However, in many practical applications, the expected value of the low-fidelity model needs to be computed at the same time; therefore we are interested in the so-called control variate with estimated control means [8, 7]. We focus our attention to a single CV from a single low-fidelity model, however the approach can be extended to multiple models as done in [9] or to latent variable networks as in [6].

We aim at computing statistical moments for a QoI $Q = Q(\boldsymbol{\xi})$, $Q : \boldsymbol{\xi} \to \mathbb{R}$ where $\boldsymbol{\xi} \in \Xi \subset \mathbb{R}^d$ is the vector of d random input parameters. We consider them to be distributed according to the joint probability distribution $p(\boldsymbol{\xi}) = \Pi_{i=1}^{d} p(\xi_i)$. For simplicity, we consider here the expected value $\mathbb{E}[Q]$ which can be approximated by a MC estimator based on N realization of Q drawn according to $p(\boldsymbol{\xi})$ as

$$\mathbb{E}[Q] = \int_{\Xi} Q(\boldsymbol{\xi}) p(\boldsymbol{\xi}) \, \mathrm{d}\boldsymbol{\xi} \approx \hat{Q}_N^{\mathrm{MC}} = \frac{1}{N} \sum_{i=1}^{N} Q(\boldsymbol{\xi}^{(i)}) = \frac{1}{N} \sum_{i=1}^{N} Q^{(i)}. \tag{1}$$

For the MC estimator in Eq. (1) it is easy to show that its variance is $\mathbb{V}ar\left(\hat{Q}_N^{\mathrm{MC}}\right) = \mathbb{V}ar\left(Q\right)/N$. Therefore, in order to obtain a more reliable estimator, the number of realizations N need to be increased; however, this can be untenable for the case of high computational expensive simulations. In the following, we use the subscripts HF and LF to indicate realizations from the high-fidelity and low-fidelity models, respectively.

Following [3, 4], the CV method with estimated control means can be written as follows

$$\mathbb{E}\left[Q^{\mathrm{HF}}\right] \approx \hat{Q}_{N_{\mathrm{HF}}}^{\mathrm{CV,HF}} = \hat{Q}_{N_{\mathrm{HF}}}^{\mathrm{MC,HF}} + \alpha \left(\hat{Q}_{N_{\mathrm{HF}}}^{\mathrm{MC,LF}} - \hat{Q}_{N_{\mathrm{LF}}}^{\mathrm{MC,LF}} \right), \tag{2}$$

where $N_{\text{LF}} = N_{\text{HF}} + \Delta_{\text{LF}} = N_{\text{HF}}(1+r)$ and the additional LF evaluations $\Delta_{\text{LF}} = rN_{\text{HF}}$ are drawn independently from the first set N_{HF}. The values for α and the additional parameter $r > 0$ are obtained by minimizing the overall computational cost under the constraint of the variance of the estimator $\mathbb{V}ar\left(\hat{Q}^{\text{CV,HF}}_{N_{\text{HF}}}\right)$ being equal to a target, namely ε^2. This constrained optimization has the solution

$$N_{\text{HF}} = \frac{\mathbb{V}ar\left(Q^{\text{HF}}\right)}{\varepsilon^2}\left(1 - \frac{r}{r+1}\rho^2\right) \quad \text{and} \quad r = -1 + \sqrt{\frac{\mathcal{C}_{\text{HF}}}{\mathcal{C}_{\text{LF}}}\frac{\rho^2}{1-\rho^2}}, \tag{3}$$

where ρ is the Pearson's correlation between HF and LF and $\mathcal{C}_{\text{HF}}$ and $\mathcal{C}_{\text{LF}}$ are the computational cost of each HF and LF, respectively. The final variance of the control variate estimator $\hat{Q}^{\text{CV}}_{N_{\text{HF}}}$ is proportional to that of the standard MC estimator according to

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

and, therefore, since $0 \leq \rho^2 \leq 1$, $\mathbb{V}ar\left(\hat{Q}^{\text{CV}}_{N_{\text{HF}}}\right) \leq \mathbb{V}ar\left(\hat{Q}^{\text{MC}}_{N_{\text{HF}}}\right)$. It follows from Eq. (4) that the performance of the control variate approach is directly related to the correlation between the models. In this work, we are seeking an efficient strategy to maximize this correlation for a given pair of HF and LF models.

3 LEVERAGING IMPORTANT DIRECTIONS FOR MULTIFIDELITY UQ

In UQ, the curse of dimensionality is a key challenge, such that considerable attention has been devoted to techniques that can reduce the dimensionality of a problem. In this work, our goal is to use dimension reduction techniques to identify the most relevant directions for each model separately, and then exploit these directions to link different fidelities. Important directions are identified as linear combinations of the original coordinates and, in the context of sampling in the reduced space, the sample coordinates can be always rotated back to the original coordinates of each model. It it reasonable to assume that, if the important directions are aligned for different models, their correlation might be higher along those directions than in the original space. Since a common shared space can be always obtained between models if the input parameters (of each model) are jointly Gaussian distributed, this strategy can also facilitate multifidelity approaches for problems where the model fidelities have dissimilar parametrizations. For instance, consider the case of performing a forward UQ analysis for a flying aircraft in which Large Eddy Simulation (LES) is used as the high-fidelity model and steady Reynolds-averaged Navier Stokes (RANS) equations are solved as a low-fidelity surrogate. In this case, the RANS model parameterization will likely include a set of parameters that are not shared with the LES code since they align more with the specific modeling approach than the physical problem to be solved. Conversely, one can easily imagine cases where a HF model contains a set of parameters that is not shared with lower fidelity models, as they derive directly from the more detailed and sophisticated physical modeling being employed.

As an example, we initially focus on a specific dimension reduction technique, namely Active Subspaces [1]. However, the concept is equally applicable to other dimension reduction strategies such as ridge approximation [2], basis adaptation and others.

3.1 The Active Subspaces example

We describe here a few concepts associated with the Active Subspace (AS) method that are central to designing our strategy. The interested reader should refer to [1] for a comprehensive discussion of the method. We consider a computational model represented by $f(\boldsymbol{\xi})$ where, as described in the previous section, $\boldsymbol{\xi} \in \Xi \subset \mathbb{R}^d$ is the vector of d random input parameters. We are interested in the forward UQ problem for computing moments of quantity of interest Q, and again we focus our interest on the expected value for simplicity. We introduce the $d \times d$ matrix $\mathbf{C}$ as the expected value of the outer product or the gradient vector $\vec{\nabla} f = [\partial f/\partial \xi_1, \cdots, \partial f/\partial \xi_d]^{\mathrm{T}}$ as

$$\mathbf{C} = \int \left(\vec{\nabla} f\right) \left(\vec{\nabla} f\right)^{\mathrm{T}} p(\boldsymbol{\xi}) \, \mathrm{d}\boldsymbol{\xi}. \tag{5}$$

Since the matrix $\mathbf{C}$ is symmetric and positive semidefinite, a real eigenvalue decomposition exists and $\mathbf{C}$ can be factorized as $\mathbf{C} = \mathbf{W}\boldsymbol{\Lambda}\mathbf{W}^{\mathrm{T}}$ where $\mathbf{W}$ is the $d \times d$ orthogonal matrix whose columns are the normalized eigenvectors and $\boldsymbol{\Lambda} = \mathrm{diag}\left\{\lambda_1, \ldots, \lambda_d\right\}$ has the corresponding eigenvalues $\lambda_1 \geq \cdots \geq \lambda_d \geq 0$ as its diagonal entries. The eigenvectors $\mathbf{W}$ define a rotation of $\mathbb{R}^d$ and the eigenvalues $\lambda_{1,\ldots,d}$ are arranged in decreasing order, enabling a separation of the rotated coordinates into two sets, namely active and inactive variable sets. We can write

$$\Lambda = \begin{bmatrix} \Lambda_A & \\ & \Lambda_I \end{bmatrix} \quad \text{and} \quad \mathbf{W} = \left[\mathbf{W}_A \mathbf{W}_I\right], \tag{6}$$

where $\Lambda_A = \mathrm{diag}\left\{\lambda_1, \ldots, \lambda_{d_A}\right\}$ and $\mathbf{W}_A$ contains the first $d_A < d$ eigenvectors.

The original coordinates can be rotated to obtain the new active $\boldsymbol{\xi}_A = \mathbf{W}_A^{\mathrm{T}}\boldsymbol{\xi} \in \mathbb{R}^{d_A}$ and inactive $\boldsymbol{\xi}_I = \mathbf{W}_I^{\mathrm{T}}\boldsymbol{\xi} \in \mathbb{R}^{(d-d_A)}$ variables. This decomposition and rotation of the original variable is effective because (as demonstrated in Lemma 3.2 [1]) the mean-squared gradients of f with respect to the the active and inactive variables ($\vec{\nabla}_A$ and $\vec{\nabla}_I$ respectively) is the sum of the eigenvalues associated to these sets. Another important property is that any $\boldsymbol{\xi}$ can be expressed in term of active and inactive variables, *i.e.* $\boldsymbol{\xi} = \mathbf{W_A}\boldsymbol{\xi}_A + \mathbf{W_I}\boldsymbol{\xi}_I$. In order to generate a shared subspace between different models, we assume the random input variables are distributed according to standard normal distributions, *i.e.* $\boldsymbol{\xi} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_d)$ (we will relax this later on). Hence, thanks to the linearity of the rotation between the original and active/inactive variables it follows that the new variable are also distributed as standard normal variable because $\mathbb{E}\left[\boldsymbol{\xi}_A\right] = \mathbf{W_A}^{\mathrm{T}}\mathbb{E}\left[\boldsymbol{\xi}\right] = \mathbf{0}$ and $\mathbb{V}ar\left(\boldsymbol{\xi}_A\right) = \mathbf{W_A}^{\mathrm{T}}\mathbb{V}ar\left(\boldsymbol{\xi}\right)\mathbf{W_A} = \mathbf{W_A}^{\mathrm{T}}\mathbf{I}_{d_A}\mathbf{W_A} = \mathbf{I}_{d_A}$ (the same applies for $\boldsymbol{\xi}_I$). It is important to note that the variance of the rotated variable will always be the unitary matrix because we consider the case of independent random input and the columns of $\mathbf{W}$ are always the normalized eigenvectors. This paves the way for obtaining a shared space between different models, even in the case of dissimilar parameterizations.

Ultimately, we are interested in the representation of the different models on the active subspaces which are the range of the eigenvectors in $\mathbf{W_A}$; hence, we want an approximation like $f_{AS}(\boldsymbol{\xi}_A) \approx f(\boldsymbol{\xi})$. The map from $\boldsymbol{\xi}$ to $\boldsymbol{\xi}_A$ is ill-posed and a regularization needs to be chosen. Here, for sake of simplicity, we use an initial approximation $f_{AS}(\boldsymbol{\xi}_A) = f(\mathbf{W_A}\boldsymbol{\xi}_A)$, which is true only if $\sum_{i=d_A+1}^{d} \lambda_i = 0$. However, since we are ultimately interested in the computation of the expected value of Q, this approximation is also equivalent to the use of a first order approximation of the conditional expected value with respect to the inactive variables. By means of the AS techniques we transformed a d dimensional integration problem in a reduced d_A integration (with $p(\boldsymbol{\xi}_A)$ as distribution).

3.2 Dimension Reduction for multifidelity UQ

In this section, we briefly describe the main idea of enhancing multifidelity UQ based on AS. We are interested in solving a forward UQ problem where the computational model is an expensive computer code that we represent as f^{HF}, which admits as input the vector of random parameters $\boldsymbol{\xi}^{\mathrm{HF}} \in \mathbb{R}^{d^{\mathrm{HF}}}$ and produces the scalar QoI Q^{HF}. We also consider a less computationally demanding approximation of the same physical system, namely f^{LF}, which admits as input the vector of random parameters $\boldsymbol{\xi}^{\mathrm{LF}} \in \mathbb{R}^{d^{\mathrm{LF}}}$ and produces the scalar QoI Q^{LF}. We propose a simple idea: instead of sampling in the original coordinates, we first search for the active directions for each model independently, and then we sample in the common shared subspace aiming at maximizing the correlation between models. Indeed it is reasonable to assume that correlation might be higher along the directions (defined by the active variables) for which the variability of the function is most pronounced. Our strategy includes the following steps:

1. For each model the AS representation is sought and the active variables are obtained
$$\begin{aligned} \boldsymbol{\xi}_A^{\mathrm{HF}} &= (\mathbf{W}_{\mathbf{A}}^{\mathrm{HF}})^{\mathrm{T}} \boldsymbol{\xi}^{\mathrm{HF}} \\ \boldsymbol{\xi}_A^{\mathrm{LF}} &= (\mathbf{W}_{\mathbf{A}}^{\mathrm{LF}})^{\mathrm{T}} \boldsymbol{\xi}^{\mathrm{LF}}; \end{aligned} \tag{7}$$
2. The shared space between the models is defined as variables $d_A^\star = \max\left(d_A^{\mathrm{HF}}, d_A^{\mathrm{LF}}\right)$. We call this vector of random variables $\boldsymbol{\xi}^\star \in \mathbb{R}^{d_A^\star}$;
3. A set of N_{pilot} pilot realizations for $\boldsymbol{\xi}^\star \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{d_A^\star})$ is generated, *i.e.* $\left\{\boldsymbol{\xi}^{\star,(i)}\right\}_{i=1}^{N_{pilot}}$;
4. The set of N_{pilot} realizations is rotated back to each set of original coordinates independently for each model
$$\begin{aligned} \boldsymbol{\xi}^{\mathrm{HF},(i)} &= \mathbf{W}_{\mathbf{A}}^{\mathrm{HF}} \begin{bmatrix} \mathbf{I}_{d_A^{\mathrm{HF}}} & \\ & \mathbf{0} \end{bmatrix} \boldsymbol{\xi}^{\star,(i)} \\ \boldsymbol{\xi}^{\mathrm{LF},(i)} &= \mathbf{W}_{\mathbf{A}}^{\mathrm{LF}} \begin{bmatrix} \mathbf{I}_{d_A^{\mathrm{LF}}} & \\ & \mathbf{0} \end{bmatrix} \boldsymbol{\xi}^{\star,(i)} \quad \text{for} \quad i = 1, \ldots, N_{pilot}. \end{aligned} \tag{8}$$
5. The quantities of interest $Q^{\mathrm{HF},(i)}$ and $Q^{\mathrm{LF},(i)}$ are obtained by invoking the models;
6. The correlation ρ between the two models is computed and the optimal sampling allocation (N_{HF} and N_{LF} from r) is obtained from Eq. (3);

7. The additional HF and LF simulations are carried out to reach the final sample allocation (steps 6 and 7 are performed until convergence is reached);

8. Statistics and the predicted variance of the estimator are evaluated.

We note here that in order to obtain a shared space between models, we considered standard normal variables for $\boldsymbol{\xi}^{\mathrm{HF}}$ and $\boldsymbol{\xi}^{\mathrm{LF}}$. However, for general distribution cases, a non-linear transformation between the actual distributions and standard normal variables can be performed. For example, if we consider a uniform random variable $\omega_i \sim \mathcal{U}(-1,1)$, it is possible to map to ω_i from the variable $\xi_i \sim \mathcal{N}(0,1)$ through the inverse of the cumulative distribution function $\omega_i = h(\xi_i) = \mathrm{erf}\left(\xi_i/\sqrt{2}\right)$. For more complex distributions, it is necessary to resort to transformations such as Nataf or Rosenblatt. The model is ultimately defined as the product of the original model and the transformation, *i.e.* $f(\boldsymbol{\xi})h(\boldsymbol{\xi})$. The drawback is that these transformations generally add non-linearity to the model, increasing complexity in the UQ process. In our context, we hope to compensate for the increase in the non-linearity of the models through the increase in their correlation. Indeed, if the non-linearity introduced by $h(\boldsymbol{\xi})$ is small relative to the global variability of $f(\boldsymbol{\xi})$ along the active directions, the correlation between models is expected to be preserved.

4 NUMERICAL TESTS

4.1 Analytical test case

We first demonstrate our approach for a simple analytical test problem, which has been constructed to highlight features of the strategy. Consider two numerical models defined as follows

$$\begin{aligned} f(x,y) &= \exp\left(0.7x + 0.3y\right) + 0.15\sin\left(2\pi x\right) \\ g(x,y) &= \exp\left(0.01x + 0.99y\right) + 0.15\sin\left(3\pi y\right). \end{aligned} \tag{9}$$

We want to compute the expected value $\mathbb{E}\left[f\right]$ and we consider the model $g(x,y)$ to be a lower accuracy representation of $f(x,y)$, therefore we want to use g as a control variate. In Fig.1, the functions f and g are reported on the domain $[-1,1]^2$. The model f exhibits a variability along both the directions x and y, whereas g is essentially varying along y. In a situation like this one the control variate approach is not effective because the correlation between the two models is expected to be very low. In Fig. 2, we take $x, y \sim \mathcal{N}(0, 1/3)$ and report the scatter plot obtained evaluating 1000 realizations of f and g using the same input values. For this set of samples, the correlation squared is approximately 0.05. In Fig. 3, the active directions are reported for the two models. It is evident that the active directions of f and g are nearly orthogonal for this case, however both functions are approximately one-dimensional. We can therefore represent each function along the direction t, where $t = \left(\mathbf{W}_{\mathbf{A}}^{f}\right)^{\mathrm{T}}\boldsymbol{\xi} = \left(\mathbf{W}_{\mathbf{A}}^{g}\right)^{\mathrm{T}}\boldsymbol{\xi}$ for standard normal $\boldsymbol{\xi} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_2)$.

Moreover, given a set of samples t it is possible to obtain the original coordinates for f and g by using $\boldsymbol{\xi} = \mathbf{W}_{\mathbf{A}} t$ (independently for f and g). In Fig. 4, we report the functions f and g along the active coordinates. It is evident that, after aligning the important directions, g essentially captures a similar trend to that of f (left plot). This can be

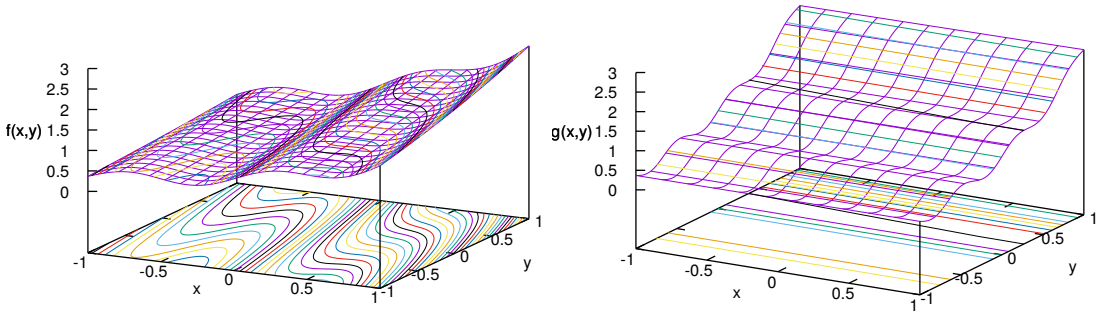

Figure 1: Test model functions.

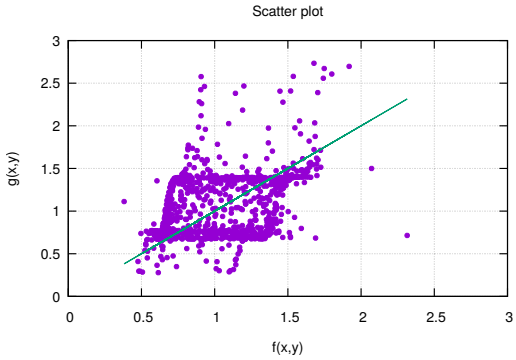

Figure 2: Scatter plot for f and g in the original coordinates.

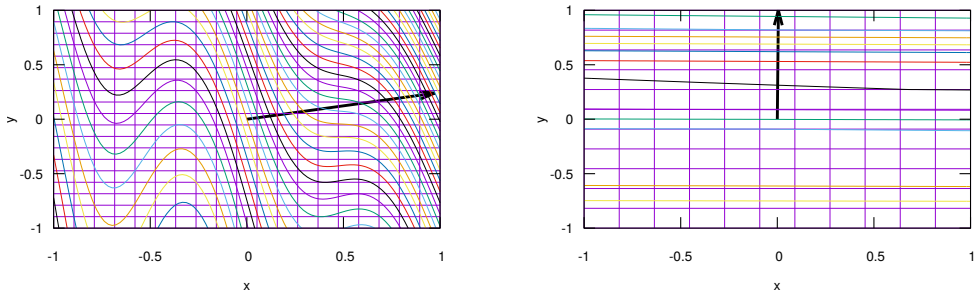

Figure 3: Test model functions.

confirmed by generating 1000 random realizations of t and then rotating back to each of the original model coordinates independently (right plot in Fig. 4). The correlation squared measured for this set of samples is approximately 0.9; therefore, we expect to observe an increase in the performance of a CV method.

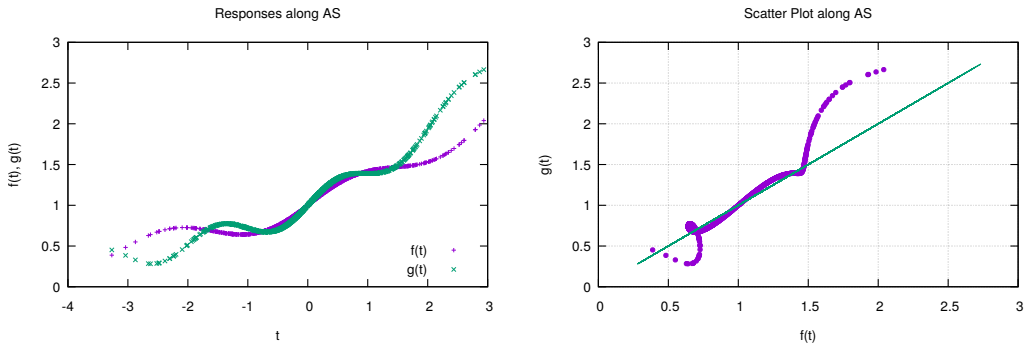

Figure 4: Functions f and g represented along the active coordinate t with $t = \left(\mathbf{W}_{\mathbf{A}}^{f}\right)^{\mathrm{T}}\boldsymbol{\xi} = \left(\mathbf{W}_{\mathbf{A}}^{g}\right)^{\mathrm{T}}\boldsymbol{\xi}$ (Left). Scatter plot obtained by sampling t and mapping back independently for each model (Right).

In order to define the forward UQ problem, we assume the cost of the low-fidelity model g to be $\mathcal{C}^{\mathrm{LF}} = 0.01\mathcal{C}^{\mathrm{HF}}$, which is representative of cost ratios observed in practical applications. We fix a computational budget of 300 HF simulations as a reference and compare four sampling estimators with equivalent cost: *(i)* the MC estimator, *(ii)* the MC estimator obtained by sampling along t (MC-AS), *(iii)* a standard CV estimator with 100 HF realizations and 20000 LF simulations (MC-MF), *(iv)* a CV estimator obtained by sampling along t and rotating the samples back to the original coordinates independently for each model (MC-MFAS). We note that, for the two estimators adopting the AS, the active directions are estimated every time in order to include the variance of this process within the overall estimator variance. In order to demonstrate the performance of the strategy, we perform 1000 repetitions of each estimator and report the normalized histograms for the estimated expected values.

Fig. 5 shows the distributions for the four mean estimators alongside the known exact solution. As expected, all of the estimators are unbiased and the distributions are approximately Gaussian. The MC and MC-AS are almost identical since the model f is essentially unidimensional in the active variable t; therefore, the variance is entirely captured by the MC-AS technique and the same behavior as MC is recovered for the same number of samples. Of course, MC is not the best choice for exploiting a lower dimensional subspace; here we include it only for demonstration purposes. The MC-MF method has been forced to redistribute the computational cost between HF and LF, irregardless of correlation. In a realistic deployment, a pilot sample could detect the low correlation and redistribute work (resulting in reliance on the HF model as in a standard MC, although penalized by the cost of the N_{pilot} samples). Once the sampling of two models are linked through the active variable $t = \left(\mathbf{W}_{\mathbf{A}}^{f}\right)^{\mathrm{T}}\boldsymbol{\xi} = \left(\mathbf{W}_{\mathbf{A}}^{g}\right)^{\mathrm{T}}\boldsymbol{\xi}$, however, the correlation is increased and improvement is evident. The MC-MFAS recovers a sharpened distribution for the expected value; i.e., its mean-squared error is much lower and the reliability of its estimates is improved.

Next, we demonstrate the effect of relaxing the requirement on normal input distribu-

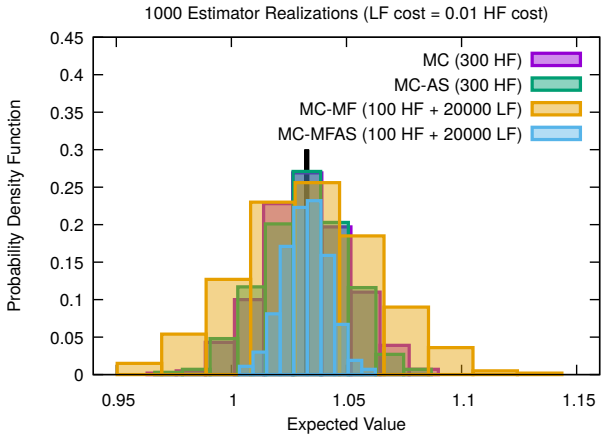

Figure 5: Normalized histograms for 1000 realizations of several expected value estimators. The exact solution is reported as a black bold vertical line.

tions and incurring a nonlinear transformation of variables. If we assume $x, y \sim \mathcal{U}(-1, 1)$, we are still able to identify the active direction t; however, the resulting model exhibits greater non-linearity as shown in Fig.6 (left plot compares the function $f(t)$ for normal and uniform cases). The scatter plot in Fig.6 (right) show that the correlation has not been degraded despite the greater non-linearity induced by the variable transformation.

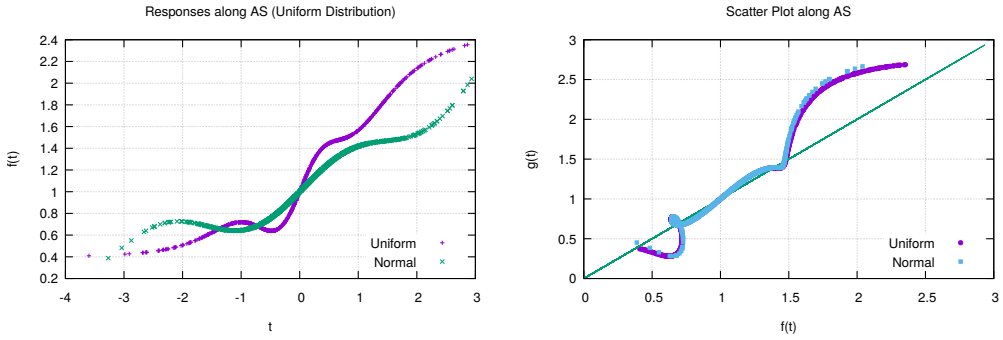

Figure 6: Comparison between normal and uniform model input. The function f represented along the active variable t is reported on the left. On the right, a comparison between scatter plots obtained for the two input distribution cases.

4.2 One-dimensional wave propagation

As a demonstration for a more complex case, we study compressional elastic wave propagation in a one-dimensional rod composed by 50 layers of two alternating materials with uncertain properties, *i.e.* uncertain non-linear stress-strain relations. We consider an uncertain initial pre-loading for the rod and we are interested in the expected value of the maximum stress in the rod after a finite time. We consider a high-fidelity model

given by an high-resolution MUSCL-Hancock scheme on a fine mesh of 801 equally spaced nodes, as opposed to a low-fidelity Godunov scheme of a very coarse mesh of 5 nodes. An approximated Riemann solver is adopted for both models. The characteristics of the two numerical models are summarized in Table 1. The problem has a total of 28 independent uniformly distributed random variables.

	N_x	N_t	Δ_t
Low-fidelity	5	50	36×10^{-4}
High-fidelity	801	600	30×10^{-5}

Table 1: HF and LF definition. The cost ratio between the model is ~ 2800.

In Fig. 7 (left), five random realizations for the two models are shown. In Fig.7 (right), we report the scatter plot which shows that the LF model already displays good correlation (approximately 0.89 squared) with respect to the HF model. Therefore, a challenge for our approach is to improve an already high correlation. In this case, the AS are discovered by using a linear regression; this is done to explore the case in which derivatives are not readily available from an adjoint solution. Extension to alternative derivative-free techniques to compute the AS is straightforward, but out of scope for this study.

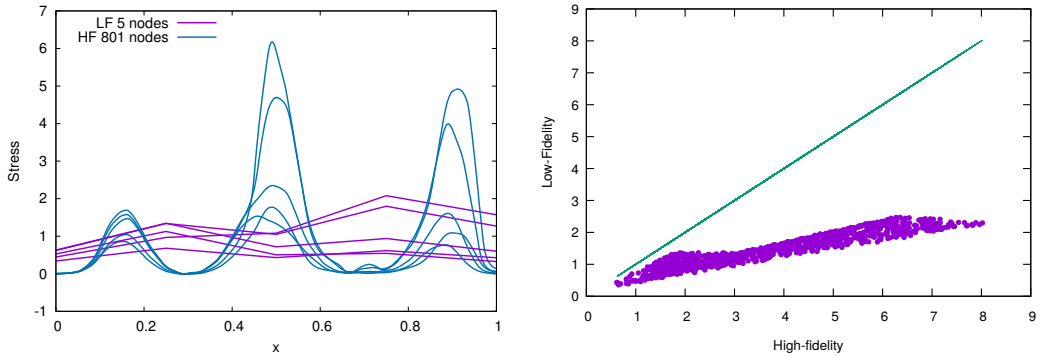

Figure 7: Five realizations of the HF and LF for the non-linear elastic propagation problem (Left). Scatter plot for several realizations of the HF and LF (Right).

In Fig.8, we display the change in the scatter plot obtained when the sampling is performed along the active variable (left plot) and the distributions for all of the estimators (right plot) based on 250 repetitions. For this case, it is again evident that we obtain a more reliable estimator by sampling in the active variable and rotating back to the original coordinates for each model. The scatter plot on the left of Fig.8 shows how the cloud of points is nearly collapsed to a line when the link between the models is enhanced using the active variables.

4.3 Preliminary results on dissimilar parameterizations

To illustrate the flexibility of this strategy, we consider again the simple analytic case of §4.1, but we include a dependence on an extra parameter z for f, *i.e.* $f(x, y, z) =$

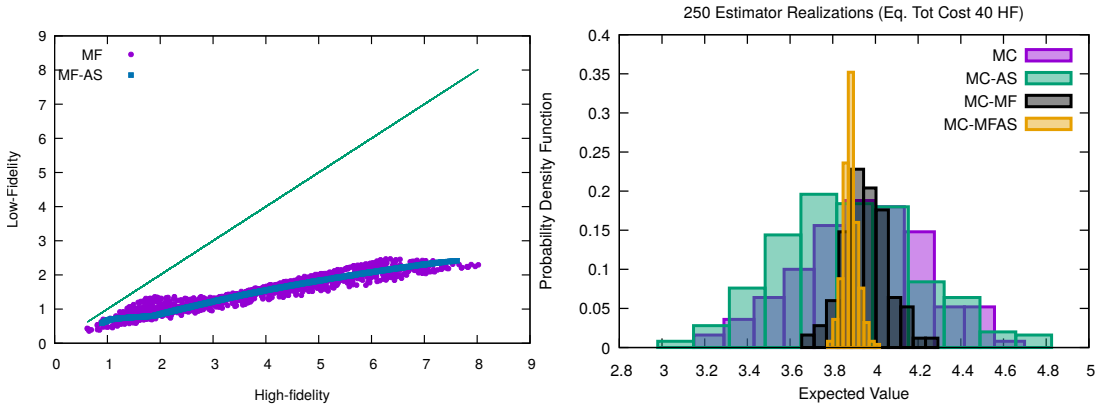

Figure 8: Five realizations of the HF and LF for the non-linear elastic propagation problem (Left). Scatter plot for several realizations of the HF and LF (Right).

$\exp(0.7x + 0.3y) + 0.15\sin(2\pi x) + 0.75z^3$, where $z \sim \mathcal{N}(0, 1/3)$. We still assume the model $g(x, y)$ to be as defined in Eq. (9). In this case, the standard control variate (MC-MF) and sampling only over the active variables (MC-AS) are expected to perform as in §4.1; therefore CV formulated by linking the two models through the active variables (MC-MFAS) and sampling over the original variables (MC) are reported in Fig. 9. It is evident that the strategy remains effective while bridging the distinct parameterizations for the two models.

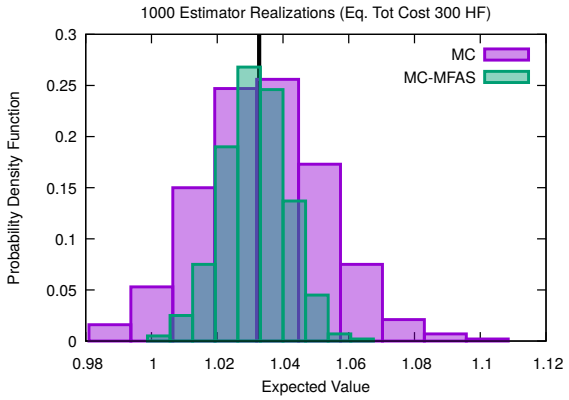

Figure 9: Normalized histograms for 1000 realizations of the MC and MC-MFAS estimators in the case of dissimilar parametrization.

5 CONCLUDING REMARKS

In this work, we propose the acceleration of multifidelity strategies (e.g., control variate Monte Carlo) by leveraging active directions separately for each model. The active directions are used to link the different models in order both to enhance their correlation

and to bridge dissimilar model parameterizations. The idea is demonstrated by means of a simple analytical test as well as a non-linear hyperbolic propagation problem in 28 random dimensions. Additional test cases will be presented during the Conference.

ACKNOWLEDGMENTS

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & 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-NA0003525. The views expressed in the article do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

REFERENCES

[1] Constantine, P., Active Subspaces. Emerging Ideas for Dimension Reduction in Parameter Studies. *SIAM Spotlights*, 2015.

[2] Constantine, P., Eftekhari, A., Hokanson, J., Rachel A.Ward. A near-stationary subspace for ridge approximation. *Comput. Method Appl. M.* (2017) Vol. 326, pp. 402-421.

[3] Geraci, G. and Eldred, M. and Iaccarino, G. A multifidelity control variate approach for the multilevel Monte Carlo technique. *CTR Annu. Res. Briefs*. (2015) 169–181

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

[5] Giles, M. B. Multi-level Monte Carlo path simulation. *Oper. Res.* (2008) **56**, pp. 607–617.

[6] Gorodetsky, A., Geraci, G., Eldred, M. and Jakeman, J. Latent Variable Networks for Multifidelity Uncertainty Quantification and Data Fusion. *7th European Conference on Computational Fluid Dynamics (ECFD 7)*, 11-15 June 2018, Glasgow, UK

[7] Ng, L.W.T., Willcox, K. Multifidelity approaches for optimization under uncertainty. *Int. J. Numer. Meth. Engng.* (2014) **100**(10), pp. 746–772.

[8] Pasupathy R., Schmeiser B. W., Taaffe and Jin Wang, M. R. Control-variate estimation using estimated control means. *IIE Transactions*, (2012) **44:5**.

[9] Peherstorfer, B., Willcox, K. and Gunzburger, M. Optimal model management for multifidelity Monte Carlo estimation. *SIAM Journal on Scientific Computing*. (2016) **38:5**, pp. A3163A3194..