

# Multifidelity uncertainty quantification with models based on dissimilar parameters

Xiaoshu Zeng<sup>1</sup>, Gianluca Geraci<sup>2</sup>, Michael S. Eldred<sup>2</sup>, John D. Jakeman<sup>2</sup>, Alex A. Gorodetsky<sup>3</sup>, and Roger Ghanem<sup>1</sup>

<sup>1</sup>University of Southern California, Los Angeles, CA

<sup>2</sup>Sandia National Laboratories, Albuquerque, NM

<sup>3</sup>University of Michigan, Ann Arbor, MI

June 27, 2023

## Abstract

Multifidelity uncertainty quantification (MF UQ) sampling approaches have been shown to significantly reduce the variance of statistical estimators while preserving the bias of the highest-fidelity model, provided that the low-fidelity models are well correlated. However, maintaining a high level of correlation can be challenging, especially when models depend on different input uncertain parameters, which drastically reduces the correlation. Existing MF UQ approaches do not adequately address this issue. In this work, we propose a new sampling strategy that exploits a shared space to improve the correlation among models with dissimilar parameterization. We achieve this by transforming the original coordinates onto an auxiliary manifold using the adaptive basis (AB) method [64]. The AB method has two main benefits: (1) it provides an effective tool to identify the low-dimensional manifold on which each model can be represented, and (2) it enables easy transformation of polynomial chaos representations from high- to low-dimensional spaces. This latter feature is used to identify a shared manifold among models without requiring additional evaluations. We present two algorithmic flavors of the new estimator to cover different analysis scenarios, including those with legacy and non-legacy high-fidelity (HF) data. We provide numerical results for analytical examples, a direct field acoustic test, and a finite element model of a nuclear fuel assembly. For all examples, we compare the proposed strategy against both single-fidelity and MF estimators based on the original model parameterization.

## Nomenclature

AB	adaptive basis	MSE	mean squared error
ACV	Approximate Control Variate	PCE	polynomial chaos expansions
ANOVA	Analysis of Variance	QOI	quantity of interest
AS	active subspace	UQ	uncertainty quantification
HF	high-fidelity	$\mathcal{N}$	normal distribution
LF	low-fidelity	$\mathcal{U}$	uniform distribution
MC	Monte Carlo	$\mathbf{A}$	rotation matrix of obtained by the AB method
MF	multifidelity	$N$	number of samples of the underlying model
MFAB	multifidelity adaptive basis	$Q$	quantity of interest
MFMC	multifidelity Monte Carlo	$d$	dimension of the underlying model in the original space
MIMC	multi index Monte Carlo	$e_i$	multi-index with one at the $i$ -th location and zeros elsewhere
MLMC	multilevel Monte Carlo	$r$	reduced dimension of the underlying model in the adapted space
MLMF	multilevel multifidelity		

$\alpha, \beta, \gamma$	multi-index	$\varpi$	difference in the first-order PCE — the second bias estimator
$\alpha$	weight coefficient in the ACV estimator	$\text{MSE}_{PCE}$	MSE of the adapted PCE — the first bias estimator
$\mu$	mean value of the underlying random variable	$\mathcal{J}_p$	set of $d$ -dimensional multi-index of a PCE with maximum order of $p$
$\xi$	set of random variable in the original space	$\mathcal{J}_p^r$	set of $r$ -dimensional multi-index of a PCE with maximum order of $p$
$\xi^r$	set of transformed variables from the $r$ -dimensional adapted space	$\psi_\alpha$	polynomial basis associated with multi-index $\alpha$
$\eta$	set of random variable in the adapted space	$c_\alpha$	PCE coefficient associated with multi-index $\alpha$
$\delta$	bias of the HF model	SUBSCRIPT “ $H$ ”	indicates the underlying symbol is associated with the HF model
$\sigma^2$	variance of the HF model	SUBSCRIPT “ $L$ ”	indicates the underlying symbol is associated with the LF model
$\rho$	Pearson correlation coefficient	SUBSCRIPT “ $r$ ”, “ $s$ ”	indicates the first $r$ or $s$ components of the underlying symbol
$\varrho$	set of cumulative to total difference ratios associated with different reduced dimensions		
$\vartheta$	threshold of the cumulative to total difference ratio		
$v$	oversampling ratio of the LF model with respect to the HF model		
$\chi$	indicator of the two bias estimators, 1 or 2		

## 1 Introduction

Model-based predictions of complex phenomena are inherently subject to uncertainty. In this work, we are interested in the so-called forward propagation within the Uncertainty Quantification (UQ) framework, which consists of evaluating the statistics of the system’s response given the uncertain distributions of the inputs. To address this, a wide range of UQ algorithms have been developed and successfully applied to various applications [68, 6, 48, 46, 12]. However, most of these methods rely on evaluations of a single highly accurate numerical model, which can become computationally intractable for models that are expensive to evaluate. To mitigate this computational cost, multifidelity (MF) methods have been developed [25, 55, 27]. These methods combine limited high-fidelity (HF) numerical simulations with less accurate, but cheaper to evaluate, lower-fidelity models. Low-fidelity (LF) models typically arise from coarse spatial/temporal discretizations and/or simplifying physics assumptions. Despite their reduced accuracy, LF models are often sufficiently correlated with the highest-fidelity model, allowing for extensive sampling to reduce the cost of estimating HF statistics.

MF methods can be classified into two main categories: sampling-based methods, which are derived from Monte Carlo (MC) [25, 7, 52, 55, 31, 16, 21, 27, 60], and surrogate-based approaches [62, 30, 35, 41, 58, 29, 28, 26]. The main idea introduced in this paper applies to both categories; however, the focus is on sampling-based methods.

Sampling MF methods use one or more LF models to reduce the error in statistics of the quantity of interest (QoI), e.g., central moments. Their mean squared error (MSE) is the sum of two contributions: variance and bias. If the highest-fidelity is unbiased, the MSE can only be reduced by decreasing the estimator variance. Control variate approaches have been introduced to decrease the estimator variance by leveraging lower-fidelity models with known means; for instance, [43]. However, for LF models used in science and engineering, the knowledge of their means is an unrealistic assumption. To address this limitation, MF approaches have been introduced that evaluate the means of the LF models along with the HF statistics. A non-exhaustive list of references includes [25, 7, 55, 31, 56, 60, 52, 16, 21, 27]. For all these methods, the variance reduction directly depends on the correlation between the HF and LF models and the relative computational expense. Coarsening the spatial and/or temporal discretizations is the most straightforward technique to generate correlated models, particularly for partial differential equations, and has been relied upon extensively in literature, such as [25, 7, 31]. Simplified modeling assumptions are often necessary in physical models, making it difficult to preserve the original parameterization. This can lead to more parameters being introduced in LF models or *vice versa*. In the current MF practice, non-shared parameters are generally ignored, although engineers and scientists usually embed physical knowledge into the LF models’ construction to avoid this situation. However, this is impossible to do for complex multi-physics and/or multi-discipline applications. Consequently, the construction of LF models often gives rise to dissimilar parameterizations, wherein the number or choice of parameters varies across models. This can considerably reduce the correlations among the models, rendering MF methods less efficient or even inapplicable.

The work presented in this paper addresses a crucial gap in the field by introducing a novel contribution. Our pivotal idea

is to abandon the original parameterization of models and instead rely on a shared manifold among models, which can be lower-dimensional. This shared manifold is obtained by transforming the original coordinates of the models. For each model, we propose a mapping from the original input parameters to a more convenient *adapted space* where variables can be ordered according to their importance in representing the model’s output. This allows the adapted space to be truncated to include only the important variables while quantifying the error committed during this truncation. The truncated adapted space is referred to as the *auxiliary space*, with the adapted coordinates that span it referred to as *important directions*. By defining a common *shared space* among all models based on their auxiliary spaces, we can demonstrate that the performance and generality of MF UQ estimators can be enhanced by sampling the models over this shared space rather than their original coordinates.

Our approach is inspired by the conjecture that the system response often depends on a limited number of variables that can be obtained as a function of the original input parameters. Our method is agnostic of the problem’s physics and relies only on data, making it applicable to a large class of numerical problems, even those whose parameters may not have a strict physical meaning, such as the coefficients in a Reynolds-Averaged Navier Stokes turbulence model. Dimension reduction strategies have previously been used to identify dimensionless groups responsible for the most variability in a system’s response, e.g., [39]. We postulate that if a shared manifold exists, it captures the common underlying causes of a model’s response, thus maximizing the correlation among models. Any dimension reduction strategy can be used, but we initially designed our approach by exploiting the Active Subspace (AS) method proposed by Constantine and collaborators [10]. We have previously presented our original idea in a series of contributions [17, 18, 20] but only considered a simplified context where each model was fully captured by a one-dimensional auxiliary space and, as a consequence, the shared space was also one-dimensional. Despite this strong assumption, the method was still effective in realistic application scenarios, as demonstrated in previous works [10, 20].

The present contribution addresses the theoretical and algorithmic challenges presented by situations where more than one important direction is required to represent a model effectively. We integrate the Adaptive Basis (AB) method [64, 69, 67], which has useful features for its integration within a MF UQ approach. AB can identify the auxiliary space by relying on a set of random realizations for the model’s QoI, and a first-order polynomial regression can be built to provide the necessary information, avoiding the need for derivative information. This inaccurate polynomial chaos expansion (PCE) is only needed to build the transformation among variables and is never required to be accurate. Our approach can be applied to problems that are usually not amenable to surrogate-based methods, such as high-dimensional problems and problems with poor solution regularity. Moreover, a crude PCE is sufficient to guide the construction of the best shared manifold without requiring additional models’ evaluations, taking into account the cost limitations associated with realistic problems.

The original contributions of this work are:

- I. A framework for designing MF sampling estimators with models based on dissimilar parameterizations is introduced and analyzed.
- II. The AB dimension reduction method is integrated in the MF sampling algorithm and we derive estimators for the statistics that enable this coupling without extensive re-sampling of the models.
- III. Two practical algorithms are designed to enable the application of this framework in contexts with different computational requirements, e.g., legacy or non-legacy HF data, various degrees of relative computational expense between HF and LF models, *etc.*
- IV. Several numerical examples verify the methodology and gauge its effectiveness against both single- and MF approaches based on the original models’ parameterizations.

The remainder of the paper is organized as follows. In Section 2, we provide an example to illustrate the relevance of dissimilar parameterization in the practice of MF UQ. In Section 3, we introduce the mathematical background for our new framework. We discuss the MF sampling strategy in Section 3.1 and the AB approach in Section 3.2. Our novel framework, which integrates AB within MF UQ, is discussed in Section 4. In Section 5, we present the practical implementation of the approach, where we offer two different algorithmic variants. Numerical results are presented in Section 6, where we explore the properties and performance of our new estimators with respect to both single- and multi-fidelity approaches based on the models’ original parameterizations. Finally, conclusions and perspectives are provided in Section 7.

## 2 Dissimilar parameterization in MF UQ: a constructive example

In this section, we illustrate how dissimilar parameterizations can naturally arise in MF UQ analyses. We show that this situation can arise either because the number of parameters is different or because different parameters are used across models. Our strategy addresses both scenarios. For simplicity, we resort to a textbook compressible fluid dynamics problem to illustrate the idea. Consider an isentropic supersonic flow through the divergent portion of a converging-diverging de Laval nozzle [2]

that results in supersonic exit flow and an exit pressure greater than the ambient pressure (under-expanded flow). Isentropic relationships describe the flow quantities and the exit pressure  $P_e$  can be expressed as

$$P_e = P_0 \left( 1 + \frac{\gamma-1}{2} M_e^2 \right)^{-\frac{\gamma}{\gamma-1}}, \quad (1)$$

in which  $P_0$  indicates the stagnation (reservoir) pressure,  $\gamma$  the ratio of polytropic coefficients (equal to 1.4 for air modeled as diatomic gas), and  $M_e$  the Mach number of the fluid velocity at the exit of the duct. The Mach number  $M_e$  can be obtained by solving the non-linear equation for the isentropic flows with a section area variation such that

$$\frac{1}{M_e} \left[ \frac{2}{\gamma+1} \left( 1 + \frac{\gamma-1}{2} M_e^2 \right) \right]^{\frac{\gamma+1}{2(\gamma-1)}} - \frac{A_e}{A^*} = 0, \quad (2)$$

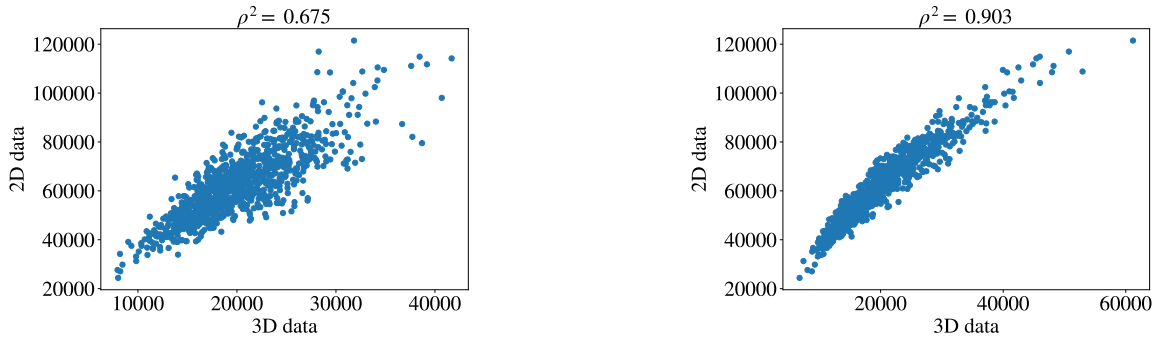
where  $A_e$  and  $A^*$  indicate the area of the exit section and the throat (at which sonic conditions exist), respectively. We consider uncertainty in the reservoir pressure  $P_0$  and in the geometry of the duct, e.g., a nozzle with a circular throat section with radius  $r_t$  and an elliptical exit section with axes  $r_a$  and  $r_b$ . The exit pressure is only affected by the area ratio

$$\frac{A_e}{A^*} = \frac{\pi r_a r_b}{\pi r_t^2} = \frac{r_a r_b}{r_t^2}. \quad (3)$$

The simplest LF model can be obtained by considering (planar) two-dimensional flow, i.e., all derivatives in the transversal direction are zero. The model is agnostic of the area variation due to the transversal dimension  $r_b$ ; therefore the area ratio is linear in  $r_a/r_t$

$$\left. \frac{A_e}{A^*} \right|_{2D} = \frac{r_a}{r_t}. \quad (4)$$

Even in a simple problem like this, preserving the models' parameterization does not follow naturally. In Figure 1a, we show the scatter plot for the exit pressure  $P_e$  as function of the uncertain stagnation pressure  $P_0 \sim \mathcal{N}(658612.5, 118212.5)$  Pa and the nozzle geometry with parameters  $r_a/r_t$  and  $r_b/r_t$  for the 3D model and  $r_a/r_t$  for the 2D model, which we assume to be distributed as  $\mathcal{N}(2, 1/6)$ . The scatter plot is a graphical representation of the linear correlation between the models and the



(a) Exit pressure  $P_e$  [Pa] for the elliptical geometry case

(b) Exit pressure  $P_e$  [Pa] for the circular geometry case

Figure 1: Scatter plot for the nozzle flow in the elliptical (a) and circular (b) geometry cases. The data represent the exit pressure  $P_e$  [Pa] with uncertain stagnation pressure  $P_0 \sim \mathcal{N}(658612.5, 118212.5)$  and exit area geometry. In the elliptical case, the 3D model has uncertain parameters  $r_a$  and  $r_b$ , while the 2D is only described by  $r_a$ . In the circular section case, the geometrical parameter is  $r_a$  for both models. We consider the geometrical parameters distributed as:  $r_a/r_t, r_b/r_t \sim \mathcal{N}(2, 1/6)$ .

spread of the data in Figure 1a shows how dissimilar parameterization negatively impacts the correlation; the effect of the extra parameter in the 3D model de-correlates the models, even if models' variability along the other (common) variable is correlated.

A more subtle case is observed in the presence of a common shared parameter with a different physical meaning in distinct models. For instance, if the above target model has a circular exit section, rather than an elliptical section, we would have  $r_a = r_b$  and the area ratio would become

$$\frac{A_e}{A^*} = \frac{\pi r_a^2}{\pi r_t^2} = \frac{r_a^2}{r_t^2}. \quad (5)$$

Although this parameter is nominally the same in both models, in the 3D configuration the area would depend on its squared value; on the other hand, the LF model depends linearly on it. In Figure 1b, we show how the scatter plot is impacted in this scenario; although the spread of data is reduced compared to the previous case, the models are still not perfectly correlated.

Understanding the role played by the models' parameters is easy in simple situations and it can be used to define well-correlated LF models. However, in complex multi-physics and/or multi-discipline applications this would be impossible. The method presented in this paper has been designed to aid in both the situations described above; in particular, it leverages an optimally constructed mapping among models (and their variables) without requiring the definition of new LF based on a physics-informed parameterization.

### 3 Mathematical background

Let  $(\Omega, \mathcal{F}, \mathcal{P})$  be a probability space and  $\boldsymbol{\xi} = (\xi_1, \dots, \xi_d)^T \in \mathbb{R}^d$  be a vector of random variables defined on this space. The focus of this paper is the accurate estimation of moments, e.g., the expected value of a QoI  $Q := Q(\boldsymbol{\xi}) \in \mathbb{R}$ . Various numerical quadrature methods can also be used to compute such expectations, e.g., sparse grids [22, 49], cubature rules [38, 42], PCE [65, 37], and Gaussian processes [53]. However, typically, the performance of these methods degrades with dimensionality and when applied to functions with limited smoothness (e.g., the presence of high gradients and discontinuities). Moreover, these methods can require structured samples, which limits their use in the presence of solver failures at certain samples. In contrast, while MC-based methods have the slowest rate of convergence with respect to the number of evaluations, compared to the methods described above for low-dimensional smooth functions, their convergence rate is independent of dimension and smoothness. This paper presents a method that ameliorates the slow convergence rate of the MC methods whenever surrogate methods cannot be easily employed.

#### 3.1 Monte Carlo-based MF estimation of statistics

Given  $N$  samples  $\{\boldsymbol{\xi}^{(i)}\}_{i=1}^N$  drawn from the joint-distribution  $p(\boldsymbol{\xi})$ , the MC estimator of the mean  $\mathbb{E}_{\boldsymbol{\xi}}[Q]$  is

$$\hat{Q} = \frac{1}{N} \sum_{i=1}^N Q^{(i)}, \quad \text{where} \quad Q^{(i)} := Q(\boldsymbol{\xi}^{(i)}). \quad (6)$$

The MC estimator is itself a random variable, i.e., different realizations corresponds to distinct set of  $N$  samples for  $\boldsymbol{\xi}$ . However, MC is unbiased, i.e.,  $\mathbb{E}[\hat{Q}] = \mathbb{E}_{\boldsymbol{\xi}}[Q]$ , and converges almost surely as  $N \rightarrow \infty$  with an error  $|\hat{Q} - \mathbb{E}_{\boldsymbol{\xi}}[Q]|$  that follows (by the law of large numbers) a normal distribution with zero mean and standard deviation  $\sqrt{\text{Var}[Q]/N}$ .

Multi-fidelity methods that exploit features like the variance decay or correlation among models have been proposed to reduce the error of MC estimators, for example, multilevel Monte Carlo (MLMC) [25], Multi index Monte Carlo (MIMC) [31], MF Monte Carlo (MFMC) [55] and Multilevel MF (MLMF) [52, 16, 21]. Among the most versatile strategies to improve MC estimation we can find *control variate* methods; see, e.g., [54]. Control variate methods [54] rely on the use of an additional random variable with known mean to reduce the variance of an MC estimator, like the one in Eq (6), while preserving the estimator bias due to the high-fidelity model. These methods exploit the correlation among models and can be extended to multiple information sources, as described in, e.g., [43, 44]. Unfortunately, in computational problems, low-fidelity models are usually simplified models for which the expected value is not known *a priori*, but rather it needs to be approximated. In this work, we focus on Approximate Control Variate (ACV) estimators [27], which are linear control variate approaches that can generalize many existing MF sampling techniques, including MLMC and MFMC. A generic ACV estimator is written as,

$$\hat{Q}^{ACV} = \hat{Q} + \sum_{i=1}^M \alpha_i (\hat{Q}_i - \hat{\mu}_i), \quad (7)$$

where  $M$  is the number of LF models;  $\hat{Q}_i$  and  $\hat{\mu}_i$  are two MC approximations of the expected value of the  $i$ th LF model obtained using two different sets of samples (see also Eq (6));  $\alpha_i \in \mathbb{R}$  are the weights (to be optimized) for the differences  $\hat{Q}_i - \hat{\mu}_i$ . In order to obtain a non-zero (but still unbiased) contribution for each low-fidelity model, the two approximations  $\hat{Q}_i$  and  $\hat{\mu}_i$  need to be distinct and  $\hat{\mu}_i$  is chosen to be more accurate than  $\hat{Q}_i$  by employing a larger set of samples. Different ACV estimators can be obtained from this framework by varying the structure of the sets of samples used by  $\hat{Q}_i$  and  $\hat{\mu}_i$  as explained in [27]. The best sampling scheme is generally problem-dependent; in this work we focus on the use of a single LF model such that  $M = 1$  and

$$\hat{Q}^{MF}(\alpha, \boldsymbol{\xi}) = \hat{Q}_H(\boldsymbol{\xi}_H) + \alpha (\hat{Q}_L(\boldsymbol{\xi}_L^1) - \hat{\mu}_L(\boldsymbol{\xi}_L^2)), \quad (8)$$

where the subscripts  $H$  and  $L$  denote the HF and LF models and the different sample sets used to evaluate each model. We indicate the (ordered) collection of samples for the HF model as  $\underline{\xi}_H = \{\xi^{(i)}\}_{i=1}^N$ , the (ordered) collection of samples for the LF model as  $\underline{\xi}_L^1 = \{\xi^{1,(i)}\}_{i=1}^N$  and  $\underline{\xi}_L^2 = \{\xi^{2,(i)}\}_{i=1}^{N_L}$ , respectively, and the ordered collection of all samples as  $\underline{\xi} = (\underline{\xi}_H, \underline{\xi}_L^1, \underline{\xi}_L^2) = (\underline{\xi}_H, \underline{\xi}_L)$ , where  $\underline{\xi}_L = (\underline{\xi}_L^1, \underline{\xi}_L^2)$ . For ease of discussion and without loss of generality, we focus on the sample configuration that results in the MFMC estimator introduced in [51, 55]. In this estimator the ordered low-fidelity set  $\underline{\xi}_L$  includes the high-fidelity samples  $\underline{\xi}_H$  and, specifically, the first  $N$  samples of the  $N_L$  LF samples are exactly  $\underline{\xi}_H$ , i.e.,  $\underline{\xi}_H \cap \underline{\xi}_L = \underline{\xi}_H = \underline{\xi}_L^1$ . An additional set of  $N_L - N$  independent samples are used for the remaining low-fidelity samples  $\underline{\xi}_L^2$ , such that  $(\underline{\xi}_L^2 \setminus \underline{\xi}_L^1) \cap \underline{\xi}_H = \emptyset$ . As proposed in [51], it is customary to define the ratio between the number of low-fidelity samples and high-fidelity samples by resorting to the  $\mathbb{R} \ni v > 1$ , such that we can write  $N_L = vN$ . As a consequence, the low-fidelity estimators in Eq (8) are evaluated from  $\underline{\xi}_L^1 = \underline{\xi}_H$  and  $\underline{\xi}_L^2 = \underline{\xi}_L$ , i.e.,  $\hat{Q}_L = \hat{Q}_L(\underline{\xi}_L^1 = \underline{\xi}_H)$  and  $\hat{\mu}_L = \hat{\mu}_L(\underline{\xi}_L)$ .

The variance of the MFMC estimator is [51]

$$\mathbb{V}ar[\hat{Q}^{MF}] = \frac{\mathbb{V}ar[Q]}{N} \left( 1 - \frac{v-1}{v} \rho^2 \right), \quad (9)$$

where  $\rho$  indicates the Pearson's correlation coefficient between the HF and LF model, defined as

$$\rho = \frac{\text{Cov}[Q_H, Q_L]}{\sigma_H \sigma_L}, \quad (10)$$

where  $\sigma_H$  and  $\sigma_L$  are the standard deviation of the HF and LF, respectively. The need for the evaluation of this coefficient is the main reason for requiring an overlapping set of samples, between HF and LF. In the case of dissimilar parametrization, the overlapping set of samples is obtained in the shared coordinates, while, in the original coordinates, these samples can be still be distinct, as described in Section 4. The variance in (9) is minimized when the coefficient  $\alpha$  and the ratio  $v$  are equal to

$$\alpha = -\rho \frac{\mathbb{V}ar^{1/2}[Q_H]}{\mathbb{V}ar^{1/2}[Q_L]} \quad \text{and} \quad v = \sqrt{\frac{\rho^2 \mathcal{C}_H}{1 - \rho^2 \mathcal{C}_L}}, \quad (11)$$

where  $\mathcal{C}_H$  and  $\mathcal{C}_L$  represent the computational cost of obtaining a single HF and LF model, respectively. The computational cost of a model refers to the amount of computational resources, such as time, memory and/or number of floating point operations, required to evaluate a single realization of the model. Assuming a fixed hardware configuration, hereinafter, we have assumed the computational cost to be proportional to the wallclock time and used this latter quantity for its quantification since we are interested in the ratio between the cost of the two models employed in the MF estimator. From Eq (9) it follows that, if a specified estimator variance  $\varepsilon^2$  is given, the number of HF model evaluations is chosen as

$$N = \frac{\mathbb{V}ar[Q]}{\varepsilon^2} \left( 1 - \frac{v-1}{v} \rho^2 \right). \quad (12)$$

Similarly, one could equivalently minimize the variance for a prescribed computational cost.

In the reminder of the paper we refer to the estimator obtained using Eq (8), as the classical MF estimator or simply the MF estimator in the original coordinates. In this framework, the HF model is considered the truth and not an approximation of a HF process, i.e., the mean squared error of the estimator is equal to the estimator's variance. The variance reduction depends on the correlation between the HF and LF model and their relative cost. Therefore, it is necessary to increase the correlation between the HF and LF models to improve the efficiency of a MF estimator for a fixed computational cost of the LF model. In the following sections we will show that this can be accomplished by sampling the models in a shared space, which alleviates the degradation of the correlation due to the presence of non-shared parameters (as described in the previous section).

### 3.2 Identifying important input features using adaptive basis

The AB method was introduced as a dimension reduction strategy for UQ in [64, 69]. Our goal in dimension reduction is to identify the most significant input parameters of a model for a certain QoI. Two relevant classes of methods are those that identify important parameters among the original coordinates (e.g., Analysis of variance (ANOVA) [59, 11]) and those that seek transformations of the original coordinates to express the model with a different set of parameters (e.g., Principal Component Analysis [40], AS [10, 9], and AB). We will focus on linear approaches as they seek a linear combination of the original parameters, defining an *adapted space*. If the number of important directions is lower than the original model's dimensionality, these variables effectively describe a low-dimensional manifold (the adapted space). We will use the adapted space of each available model to define a shared manifold that enhances the correlation among models.

In this work, we focus on AB for two reasons. Firstly, AB, like the AS method, can be implemented from a random set of model evaluations. By solving a linear regression from this set of realizations, we can obtain the first-order polynomial coefficients. Unlike the straightforward version of the AS method, this knowledge is sufficient to identify more than one important direction. Secondly, the PCE construction within AB can be used to estimate several statistical properties for designing the estimator, without requiring additional model runs.

### 3.2.1 Fundamentals of polynomial chaos expansions and adaptive basis

Given a set of independent Gaussian variables  $\xi$  and the Hilbert space  $\mathcal{H}$  generated by the linear span of  $\xi$ , we can approximate any QoI  $Q \in L^2(\Omega, \mathcal{F}(\mathcal{H}), \mathcal{P})$  by the PCE [23], which can be truncated to a total degree  $p$

$$Q(\xi) = \sum_{\alpha \in \mathcal{J}} c_{\alpha} \psi_{\alpha}(\xi) \approx \sum_{\alpha \in \mathcal{J}_p} c_{\alpha} \psi_{\alpha}(\xi). \quad (13)$$

Here  $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathcal{J} := (\mathbb{N}_0)^d$  is a  $d$ -dimensional multi-index,  $\{\psi_{\alpha}(\xi) : \alpha \in \mathcal{J}\}$  are orthonormal Hermite polynomials, which are a complete basis for  $\mathcal{H}$ ,  $\mathcal{J}_p = \{\alpha \in \mathcal{J} : |\alpha| \leq p\}$ , and  $|\alpha| = \sum_{i=1}^d \alpha_i$ . The PCE coefficients can be expressed by projection as

$$c_{\alpha} = \frac{\langle Q, \psi_{\alpha} \rangle}{\langle \psi_{\alpha}^2 \rangle} = \langle Q, \psi_{\alpha} \rangle, \quad \alpha \in \mathcal{J}_p, \quad (14)$$

where  $\langle \cdot, \cdot \rangle$  is the  $L^2$  inner product which is defined as the expectation of the product with respect to the Gaussian measure. Eq (14) can be understood as a multi-dimensional integral involving the underlying Gaussian measure. The PCE converges to  $Q$  in the mean squared sense as the number of terms in the expansion tends to infinity, i.e.,  $\text{card}(\mathcal{J}_p) \rightarrow \infty$ . Without loss of generality, we assume a Gaussian distribution for the uncertain parameters  $\xi$ ; however, one could rely on generalized PCE [65] for distributions belonging to the Askey-Wiener scheme or on transformations, e.g., the Nataf transformation [47] for arbitrary distributions. Furthermore, any set of independent random variables can be transformed into a set of independent and identically distributed standard Gaussian variables. Such transformations can be applied separately to variables of dissimilar models. AB can then be applied to the composition of the variable transformation and the model response. These variable transformations add a non-linearity that can reduce the correlation among models. However, we have demonstrated in practice that variable transformations are viable when applying Active Subspaces, as shown in [17, 20], to high-dimensional models. Given the commonality between AB and Active Subspaces, we expect variable transformations to work equally well with AB. Several approaches have been proposed to obtain the PCE coefficients; see, e.g., [45, 57, 32, 50, 3, 36]. In this work we rely on regression-based strategies to leverage available samples generated via random sampling, which is natural in the MF UQ sampling strategy. Therefore, we will work with Ordinary Least-Square (OLS) regression to obtain the coefficients  $c_{\alpha}$ .

The AB method efficiently identifies the important directions in the parameter space by constructing an isometry (rotation matrix) that transforms the Gaussian inputs into a new basis. Specifically, the AB method constructs a rotation matrix  $\mathbf{A}$  on  $\mathbb{R}^{d \times d}$  such that  $\mathbf{A}\mathbf{A}^T = \mathbf{I}$ . This matrix defines the so-called adapted variables  $\eta$  as

$$\eta = \eta(\xi) = \mathbf{A}\xi. \quad (15)$$

Because  $\xi$  are independent Gaussian variables and (15) is a linear transformation,  $\eta$  are also independent Gaussian variables. The PCE expansion of  $Q$  can be expressed in the adapted variables as

$$Q^{\mathbf{A}}(\eta) = \sum_{\beta \in \mathcal{J}_p} c_{\beta}^{\mathbf{A}} \psi_{\beta}(\eta), \quad (16)$$

where the superscript  $\mathbf{A}$  on  $Q$  and  $c_{\beta}^{\mathbf{A}}$  denotes that the expansion is in terms of new variables generated by the rotation matrix  $\mathbf{A}$ . Since  $\xi$  and  $\eta$  are both sets of independent Gaussian variables,  $\{\psi_{\alpha}(\xi) : \alpha \in \mathcal{J}_p\}$  and  $\{\psi_{\beta}(\eta) : \beta \in \mathcal{J}_p\}$  span the same space; thus, we have two equivalent expansions  $Q^{\mathbf{A}}(\eta(\xi)) \equiv Q(\xi)$  with “ $\equiv$ ” denoting equivalence. Furthermore, by defining  $\psi_{\beta}^{\mathbf{A}}(\xi) := \psi_{\beta}(\eta) = \psi_{\beta}(\mathbf{A}\xi)$ , the new PCE coefficients can be computed from the knowledge of the expansion in the original coordinates

$$c_{\beta}^{\mathbf{A}} = \sum_{\alpha \in \mathcal{J}_p} c_{\alpha} \langle \psi_{\alpha}, \psi_{\beta}^{\mathbf{A}} \rangle, \quad \beta \in \mathcal{J}_p. \quad (17)$$

In summary, the AB method requires the evaluation of a rotation matrix  $\mathbf{A}$ , which can be used to derive a PCE that is distinct from the PCE expressed in the original coordinates. The flexibility of this approach arises from the fact that the rotation matrix can be determined using only the first-order coefficients, as outlined in Algorithm 1. To obtain the first-order PCE coefficients, we can use OLS to fit a model to the available samples of the QoI (Step 1). We obtain the first row of  $\mathbf{A}$  using these coefficients (Step 2), while all subsequent rows are constructed in Step 4 after rearranging the coefficients in descending order of importance (Step 3). The algorithm is concluded by applying the Gram-Schmidt procedure (Step 5) to ensure that  $\mathbf{A}\mathbf{A}^T = \mathbf{I}$ .

---

**Algorithm 1:** Construction of rotation matrix by Gaussian adaptation [67]

---

- 1 Use least squares regression to estimate the coefficients  $c_{e_0}$  and  $\{c_{e_i}\}_{i=1}^d$  of the first-order PCE

$$Q(\xi) = \sum_{\alpha \in \mathcal{J}_1} c_{\alpha} \psi_{\alpha}(\xi) = c_{e_0} + \sum_{i=1}^d c_{e_i} \xi_i, \quad (18)$$

with  $e_0$  the multi-index of all zeros, and  $e_i$  the multi-index with one at  $i$ -th location and zeros elsewhere.

- 2 Construct the first row of the rotation matrix  $\mathbf{A} \in \mathbb{R}^{d \times d}$  such that

$$\eta_1 = \sum_{i=1}^d A_{1i} \xi_i = \sum_{i=1}^d c_{e_i} \xi_i \implies A_{1i} = c_{e_i} \quad \text{for } i = 1, \dots, d. \quad (19)$$

- 3 Rank the first-order coefficients  $\{c_{e_i}\}_{i=1}^d$  by absolute value in descending order and record their indices in the original coordinates as  $\{\kappa_j\}_{j=1}^d$ ;  
4 For  $j = 2, \dots, d$  construct the  $j^{\text{th}}$  row of  $\mathbf{A}$  such that

$$\eta_j = \xi_{\kappa_{j-1}}. \quad (20)$$

- 5 Perform Gram-Schmidt procedure on  $\mathbf{A}$  to make it an isometry (rotation matrix).
- 

The algorithm requires only a limited number of QoI evaluations because only the coefficients of a first-order PCE are necessary to define the rotation matrix. In regression-based strategies, the required number of random samples is in the order of  $N \sim \mathcal{O}(P \ln P)$  or, in some cases,  $N \sim \mathcal{O}(P^2 \ln P)$  [8], where  $P$  is the total number of terms in the PCE. We are using a first-order PCE so that  $P = d + 1$ . Furthermore, we note that, in practice, the cost of evaluating the models is significantly higher compared to the computational cost required to obtain the rotation matrix through the evaluation of the PCE coefficients.

AB enables the recovery of all  $d$  dimensions of  $\boldsymbol{\eta}$  without incurring additional computational expense. Note that in the scenario where the first-order derivatives to the original coordinates are zero, the aforementioned algorithm may be unable to identify the important directions. However, this issue has been addressed in literature [64] by devising a high-order adaptation, which can be used without modifications in our context. We also note that the first-order PCE here is only used to construct the rotation matrix. For low-dimensional manifold representation of the QoI, a higher-order PCE can be constructed using the adapted variables. This captures the nonlinearity of the problem with higher-order PCE terms, allowing for adjustment of the PCE order to improve accuracy. The AB method has been successfully applied to various applications for reducing representation of nonlinear QoIs [24, 69, 67].

The importance of the variables  $\boldsymbol{\eta}$  is captured in the order of the rows in the matrix  $\mathbf{A}$ . Then, a dimension reduction can be performed by partitioning the adapted variables  $\boldsymbol{\eta}$  as

$$\boldsymbol{\eta} = \begin{pmatrix} \mathbf{A}_r \boldsymbol{\xi} \\ \mathbf{A}_{-r} \boldsymbol{\xi} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\eta}_r \\ \boldsymbol{\eta}_{-r} \end{pmatrix}, \quad (21)$$

where  $\boldsymbol{\eta}_r$  are defined as the first  $r$  important directions. The PCE  $Q^{\mathbf{A}_r}$ , defined on the first  $r$  important directions, satisfies

$$Q^{\mathbf{A}_r}(\boldsymbol{\eta}_r) = \sum_{\gamma \in \mathcal{J}_p^r} c_{\gamma}^{\mathbf{A}_r} \psi_{\gamma}(\boldsymbol{\eta}_r) \approx Q(\boldsymbol{\xi}), \quad (22)$$

where  $\gamma = (\gamma_1, \dots, \gamma_r)$  is a  $r$ -dimensional multi-index and  $\mathcal{J}_p^r := \{\alpha \in \mathcal{J} : |\alpha| \leq p\}$  such that  $\{\psi_{\gamma}(\boldsymbol{\eta}_r) : \gamma \in \mathcal{J}_p^r\}$  is a basis of order up to  $p$  defined on  $\boldsymbol{\eta}_r$ . Then, by defining  $\psi_{\gamma}^{\mathbf{A}_r}(\boldsymbol{\xi}) := \psi_{\gamma}(\boldsymbol{\eta}_r) = \psi_{\gamma}(\mathbf{A}_r \boldsymbol{\xi})$  and assuming that  $Q^{\mathbf{A}_r}(\boldsymbol{\eta}_r(\boldsymbol{\xi}))$  and  $Q(\boldsymbol{\xi})$  are equivalent expansions of the QoI, that is,  $Q^{\mathbf{A}_r}(\boldsymbol{\eta}_r(\boldsymbol{\xi})) \equiv Q(\boldsymbol{\xi})$ , we can obtain the PCE coefficients by projection as

$$c_{\gamma}^{\mathbf{A}_r} = \langle Q, \psi_{\gamma}^{\mathbf{A}_r} \rangle, \quad \gamma \in \mathcal{J}_p^r. \quad (23)$$

The estimation of the PCE coefficients in Eq (22) (with  $r \leq d$ ) requires function evaluations as training data; samples generated in the  $\boldsymbol{\eta}_r$  space can be mapped to  $\boldsymbol{\xi}$  to allow for the evaluation of the model. We introduce  $\boldsymbol{\eta}_{-r} = \mathbf{0}$  to regularize this transformation<sup>1</sup>

$$\boldsymbol{\xi}^r = \boldsymbol{\xi}^r(\boldsymbol{\eta}_r) = \mathbf{A}^T \begin{pmatrix} \boldsymbol{\eta}_r \\ \mathbf{0} \end{pmatrix} = \mathbf{A}_r^T \boldsymbol{\eta}_r. \quad (24)$$

---

<sup>1</sup>This is required to preserve a bijective mapping



The superscript  $r$  on  $\xi$  signifies that the full space coordinates  $\xi^r \in \mathbb{R}^d$  are approximated using a transformation of the  $r$ -dimensional adapted coordinates  $\eta_r \in \mathbb{R}^r$ . The accuracy of the adapted PCE in Eq (22) can be affected by the transformation in (24), especially when the variation of  $Q$  along the directions  $\eta_{\neg r}$  is significant. If  $Q$  does not change significantly in these directions, the error introduced by the transformation is negligible when computing the adapted PCE.

Note that most AB literature focuses on improving the accuracy of the PCE surrogate of a single model. This is accomplished by sequentially increasing its value until the difference between two successive adaptations, denoted as  $(r-1)$  and  $r$ , is within the specified tolerance. Please refer to [24, 69, 67] for more information on this topic. As we increase  $r$ , the error of the transformation (24) decreases, leading to an increased accuracy of the adapted model. When  $r = d$ , the adaptation recovers the full-dimensional PCE. However, when using AB for constructing MF estimators, the choice of  $r$  must be determined based on its impact on the error of the MF estimator. We will discuss procedures for choosing  $r$  in the next section.

## 4 MF estimation via embedded adaptive basis

The first step for obtaining an MF estimator with embedded AB, which we denote with MFAB, is to reformulate the MF estimator (Eq (8)) by introducing the change of coordinates explained in the previous section

$$\xi_H = \mathbf{A}_H^T \eta_H, \quad \xi_L^1 = \mathbf{A}_L^T \eta_L^1, \quad \text{and} \quad \xi_L^2 = \mathbf{A}_L^T \eta_L^2, \quad (25)$$

which leads to

$$\widehat{Q}^{MFAB}(\underline{\eta}; \alpha) = \widehat{Q}_H(\mathbf{A}_H^T \underline{\eta}_H) + \alpha \left( \widehat{Q}_L(\mathbf{A}_L^T \underline{\eta}_L^1) - \widehat{\mu}_L(\mathbf{A}_L^T \underline{\eta}_L^2) \right), \quad (26)$$

where we defined the collection of samples as  $\underline{\eta}_H = \{\eta^{(i)}\}_{i=1}^N$ ,  $\underline{\eta}_L^1 = \{\eta_L^{(i),1}\}_{i=1}^N$ ,  $\underline{\eta}_L^2 = \{\eta_L^{(i),2}\}_{i=1}^{vN}$ , and  $\underline{\eta} = (\underline{\eta}_H, \underline{\eta}_L^1, \underline{\eta}_L^2)$ , as done in Section 3.1. The estimators in Eq. (26) are defined as

$$\begin{aligned} \widehat{Q}_H(\mathbf{A}_H^T \underline{\eta}_H) &= \frac{1}{N} \sum_{i=1}^N Q_H(\xi_H^{(i)}) = \frac{1}{N} \sum_{i=1}^N Q_H(\mathbf{A}_H^T \eta_H^{(i)}) \\ \widehat{Q}_L(\mathbf{A}_L^T \underline{\eta}_L^1) &= \frac{1}{N} \sum_{i=1}^N Q_L(\xi_L^{1,(i)}) = \frac{1}{N} \sum_{i=1}^N Q_L(\mathbf{A}_L^T \eta_L^{1,(i)}) \\ \widehat{\mu}_L(\mathbf{A}_L^T \underline{\eta}_L^2) &= \frac{1}{vN} \sum_{i=1}^{vN} Q_L(\xi_L^{2,(i)}) = \frac{1}{vN} \sum_{i=1}^{vN} Q_L(\mathbf{A}_L^T \eta_L^{2,(i)}), \end{aligned} \quad (27)$$

where, as discussed in Section 3.1, we have  $\xi_H = \xi_L^1$  and  $(\xi_L^2 \setminus \xi_L^1) \cap \xi_L^1 = \emptyset$ .

Up to this point, we have only changed the coordinates for each model. However, it is not guaranteed that  $\eta_H$  and  $\eta_L$  (as well as  $\xi_H$  and  $\xi_L$ ) will have the same cardinality, since  $d_H := \text{card}(\eta_H) = \text{card}(\xi_H)$  is generally different from  $d_L := \text{card}(\eta_L) = \text{card}(\xi_L)$ . Moreover, we want to take advantage of the dimension reduction offered by AB. In the previous section, we demonstrated how to truncate the model by selecting the first  $r$  important directions. Here, we consider a truncation independent for each model, i.e.,  $r_H \leq d_H$  and  $r_L \leq d_L$ . The values of  $r_H$  and  $r_L$  can be chosen to control the MFAB performance, as described later. Assuming that the HF and LF adapted coordinates  $\eta_H$  and  $\eta_L$  are arranged in decreasing order of importance (see Algorithm 1), we can define a vector of shared (important) directions as

$$\eta_s = [\eta_{s,1}, \dots, \eta_{s,r_s}]^T \in \mathbb{R}^{r_s}, \quad (28)$$

where  $r_s = \max(r_H, r_L)$  and

$$\eta_{s,i} = \begin{cases} \eta_{H,i} & \text{if } i \leq \min(r_H, r_L) \\ \eta_{H,i} & \text{if } r_L < i \leq r_H \\ \eta_{L,i} & \text{if } r_H < i \leq r_L. \end{cases} \quad (29)$$

To satisfy the above definition, the HF and LF realizations must be associated with the same samples of  $\eta_s$  in the adapted space. We achieve this by proposing two novel sampling algorithms, which we will discuss later in Section 5. The current section is devoted to how a MF estimator is constructed with embedded AB.

Starting from Eq (24) we can determine the map from the shared variables  $\eta_s$  to original coordinates corresponding to the truncation, either  $r_H$  or  $r_L$ , for the HF and LF model, respectively

$$\begin{aligned} \xi_H^{r_H} &= \mathbf{A}_{H,r_H}^T \eta_{H,r_H} = [\mathbf{A}_{H,r_H}^T \quad \mathbf{0}_{d_H \times (r_s - r_H)}] \eta_s = \mathcal{A}_{H,r_H}^T \eta_s \\ \xi_L^{r_L} &= \mathbf{A}_{L,r_L}^T \eta_{L,r_L} = [\mathbf{A}_{L,r_L}^T \quad \mathbf{0}_{d_L \times (r_s - r_L)}] \eta_s = \mathcal{A}_{L,r_L}^T \eta_s. \end{aligned} \quad (30)$$

These transformations allow for re-writing the estimators in Eqs. (27) as function of the shared coordinates<sup>2</sup>

$$\begin{aligned}\widehat{Q}_H(\mathbf{A}_H^T \underline{\boldsymbol{\eta}}_H) &= \frac{1}{N} \sum_{i=1}^N Q_H(\boldsymbol{\xi}_H^{(i)}) \approx \frac{1}{N} \sum_{i=1}^N Q_H(\boldsymbol{\xi}_H^{r_H, (i)}) \approx \frac{1}{N} \sum_{i=1}^N Q_H(\mathcal{A}_{H, r_H}^T \boldsymbol{\eta}_s^{(i)}) =: \widehat{Q}_H^{r_H}(r_H, \underline{\boldsymbol{\eta}}_s) \\ \widehat{Q}_L(\mathbf{A}_L^T \underline{\boldsymbol{\eta}}_L^1) &= \frac{1}{N} \sum_{i=1}^N Q_L(\boldsymbol{\xi}_L^{1, (i)}) \approx \frac{1}{N} \sum_{i=1}^N Q_L(\boldsymbol{\xi}_L^{r_L, (i)}) \approx \frac{1}{N} \sum_{i=1}^N Q_L(\mathcal{A}_{L, r_L}^T \boldsymbol{\eta}_s^{(i)}) =: \widehat{Q}_L^{r_L}(r_L, \underline{\boldsymbol{\eta}}_s) \\ \widehat{\mu}_L(\mathbf{A}_L^T \underline{\boldsymbol{\eta}}_L^2) &= \frac{1}{vN} \sum_{i=1}^{vN} Q_L(\boldsymbol{\xi}_L^{2, (i)}) \approx \frac{1}{vN} \sum_{i=1}^{vN} Q_L(\boldsymbol{\xi}_L^{r_L, (i)}) \approx \frac{1}{vN} \sum_{i=1}^{vN} Q_L(\mathcal{A}_{L, r_L}^T \boldsymbol{\eta}_s^{2, (i)}) =: \widehat{\mu}_L^{r_L}(r_L, \underline{\boldsymbol{\eta}}_s^2),\end{aligned}\tag{31}$$

where  $\underline{\boldsymbol{\eta}}_s = \{\boldsymbol{\eta}_s^{(i)}\}_{i=1}^N$  and  $\underline{\boldsymbol{\eta}}_s^2 = \{\boldsymbol{\eta}_s^{2, (i)}\}_{i=1}^{vN}$ .

Finally, the MFAB estimator (26) can be written as a function of the samples in the shared coordinates as

$$\widehat{Q}^{MFAB}(\underline{\boldsymbol{\eta}}_s, \underline{\boldsymbol{\eta}}_s^2; \alpha, r_H, r_L) = \widehat{Q}_H^{r_H}(r_H, \underline{\boldsymbol{\eta}}_s) + \alpha \left( \widehat{Q}_L^{r_L}(r_L, \underline{\boldsymbol{\eta}}_s) - \widehat{\mu}_L^{r_L}(r_L, \underline{\boldsymbol{\eta}}_s^2) \right).\tag{32}$$

We reiterate that even though the MFAB estimator necessitates samples solely from the shared coordinates  $\underline{\boldsymbol{\eta}}_s$  and  $\underline{\boldsymbol{\eta}}_s^2$ , each model employs a separate rotation matrix (refer to Eq (30)). This aspect is crucial for the method, especially in the context of dissimilar parametrization. Whenever common samples are needed for both models, they are only required in  $\underline{\boldsymbol{\eta}}_s$ , while they can still correspond to different samples in the original coordinates. The following proposition illustrates the relationship between the original coordinates of the HF and LF models.

**Proposition 4.1** (Original coordinate mapping through common shared space). *Assume  $r_H = r_L$  and that the samples are generated on the common shared space  $\boldsymbol{\eta}_s$ , then the original coordinates, obtained as  $\boldsymbol{\xi}_H^{r_H} = \mathbf{A}_{H, r_H}^T \boldsymbol{\eta}_{H, r_H}$  and  $\boldsymbol{\xi}_L^{r_L} = \mathbf{A}_{L, r_L}^T \boldsymbol{\eta}_{L, r_L}$  (see Eq (30)), are related through*

$$\boldsymbol{\xi}_H^{r_H} = \mathbf{A}_{H, r_H}^T \mathbf{A}_{L, r_L} \boldsymbol{\xi}_L^{r_L}, \quad \boldsymbol{\xi}_L^{r_L} = \mathbf{A}_{L, r_L}^T \mathbf{A}_{H, r_H} \boldsymbol{\xi}_H^{r_H}.\tag{33}$$

*Proof.* The proof is provided in Appendix A. □

The properties of the MFAB estimator (32) depend on the numbers of important directions  $r_H$  and  $r_L$ , which affect both correlation between the truncated models and the estimator's MSE. The following proposition capture the effect of  $r_H$  and  $r_L$  in the MSE.

**Proposition 4.2** (MSE error for an MFAB estimator as function of  $r_H$  and  $r_L$ ).

$$\text{MSE} \left[ \widehat{Q}^{MFAB} \right] = \text{Var} \left[ \widehat{Q}^{MFAB} \right] + \left( \mathbb{E} \left[ Q_H(\mathcal{A}_{H, r_H}^T \boldsymbol{\eta}_s) \right] - \mathbb{E} \left[ Q_H(\boldsymbol{\xi}_H) \right] \right)^2 = \sigma^2(r_H, r_L) + \delta^2(r_H),\tag{34}$$

where

$$\sigma^2(r_H, r_L) = \text{Var} \left[ \widehat{Q}_H^{r_H}(r_H, \underline{\boldsymbol{\eta}}_s) \right] \left( 1 - \frac{v-1}{v} \rho_{r_L, r_H}^2 \right),\tag{35}$$

$\rho_{r_L, r_H}^2$  is the correlation between the models' samples in the adapted coordinates  $Q_L(\mathcal{A}_{L, r_L}^T \boldsymbol{\eta}_s)$  and  $Q_H(\mathcal{A}_{H, r_H}^T \boldsymbol{\eta}_s)$ , i.e.,

$$\rho_{r_L, r_H}^2 = \frac{\text{Cov}^2[Q_H(\mathcal{A}_{H, r_H}^T \boldsymbol{\eta}_s), Q_L(\mathcal{A}_{L, r_L}^T \boldsymbol{\eta}_s)]}{\text{Var} \left[ Q_H(\mathcal{A}_{H, r_H}^T \boldsymbol{\eta}_s) \right] \text{Var} \left[ Q_L(\mathcal{A}_{L, r_L}^T \boldsymbol{\eta}_s) \right]},\tag{36}$$

and the estimator's squared bias is

$$\delta^2(r_H) = \left( \mathbb{E} \left[ Q_H(\mathcal{A}_{H, r_H}^T \boldsymbol{\eta}_s) \right] - \mathbb{E} \left[ Q_H(\boldsymbol{\xi}_H) \right] \right)^2.\tag{37}$$

*Proof.* The proof is provided in Appendix B. □

We see that the estimator's bias is introduced due to the truncation of HF in the AB method, while it is not present in the MF estimator obtained by sampling the models in the original coordinates. It is easy to show that for  $r_H \rightarrow d_H$  the term  $\delta^2(r_H) \rightarrow 0$ ; however, this choice has a less intuitive effect on the estimator's variance. For an efficient design of the MFAB estimator it is necessary to estimate the impact that  $r_H$  and  $r_L$  have on the correlation and MSE of the estimator. This can be done by leveraging the PCE construction within AB, as discussed in the next section.

<sup>2</sup>The symbol “ $\Leftarrow$ ” is the reverse of “ $\Rightarrow$ ”, denoting that the right-hand side is defined as the left-hand side.

## 4.1 Leveraging the AB framework to efficiently estimate MFAB properties without re-sampling

We will describe in this section how to estimate bias and correlation between models based on  $r_H$  and  $r_L$ , with minimal computational cost. This knowledge will be fundamental for constructing optimal MFAB estimators.

### 4.1.1 Bias quantification

We aim to develop bias estimators that require only a small number of pilot samples. The approach is straightforward: first, we estimate the significance of adapted coordinates in contributing to the model's bias. Then, we use a reduced space spanned by the important directions to obtain additional samples and estimate the bias due to truncation. We present two strategies for achieving this goal.

**Bias estimator 1: MSE of adapted PCEs.** The first bias estimator is based on the mean square error (MSE) of the projected PCE onto the adapted space. To obtain this estimator, we first evaluate a set of  $N_p$  pilot samples  $Q_H^{(i)}_{i=1}^{N_p}$  of the HF model in the original coordinates  $\xi_H$ . Using Algorithm 1, we construct the rotation matrix  $A_H$ , and then, without the need for additional samples, we project the pilot samples onto the adapted space to obtain the corresponding adapted variables  $\eta_{H,r_H} = A_{H,r_H} \xi_H$ . Next, we project  $\{Q_H^{(i)}\}_{i=1}^{N_p}$  to the adapted PCE

$$Q_H(\eta_{H,r_H}) = \sum_{\gamma \in \mathcal{J}_{p_H}^{r_H}} c_{H,\gamma} \psi_\gamma(\eta_{H,r_H}). \quad (38)$$

The expansion is the same as Eq (22), but specific to the HF.  $\mathcal{J}_{p_H}^{r_H}$  represents the sets of PCE multi-indices now for only  $r_H$  adapted coordinates. An estimator for the MSE of the PCE can be defined as

$$\text{MSE}_{PCE} \approx \frac{1}{N_p} \sum_{i=1}^{N_p} \left( Q_H(\xi_H^{(i)}) - Q_H(\eta_{H,r_H}^{(i)}) \right)^2. \quad (39)$$

This estimator captures the bias convergence, which will be used as a target for the variance allocation in the estimator, as shown in the numerical results section.

**Bias estimator 2: difference in first-order PCEs.** The second estimator uses simpler information and is quantified by the difference of the first-order PCEs in the original space. In Eq (24), we explained that the assumption  $\eta_{H,-r} = \mathbf{0}$  introduces errors in the adapted PCEs. Writing the transformation  $\xi_H \mapsto \eta_{H,r_H} \mapsto \xi_H^{r_H}$  as  $\xi_H^{r_H} = A_{H,r_H}^T A_{H,r_H} \xi_H$ , the error is embedded in the difference between  $\xi_H^{r_H}$  and  $\xi_H$ . Since these two quantities are in the original space, we can use the PCE in this space to quantify the differences without requiring additional model evaluations. Although the high-order PCE in the original space is not available, we have constructed a first-order pilot PCE during the construction of  $A_H$ . Therefore, we can use the difference in the first-order PCEs to quantify the difference.

Suppose the first-order coefficients of the pilot PCE are  $(c_{H,e_1}, \dots, c_{H,e_d})$ , the first-order components with respect to  $\xi_H$  and  $\xi_H^{r_H}$  are  $(c_{H,e_1} \xi_{H,1}, \dots, c_{H,e_d} \xi_{H,d})$  and  $(c_{H,e_1} \xi_{H,1}^{r_H}, \dots, c_{H,e_d} \xi_{H,d}^{r_H})$ , respectively. The difference between them can be quantified by the  $l^2$ -norm  $\left\| (c_{H,e_1} \xi_{H,1}, \dots, c_{H,e_d} \xi_{H,d}) - (c_{H,e_1} \xi_{H,1}^{r_H}, \dots, c_{H,e_d} \xi_{H,d}^{r_H}) \right\|_2$ . Similarly to the previous estimator, we only use this quantity to estimate the convergence rather than the MSE's magnitude. We introduce the following error estimator

$$\varpi = \frac{1}{N_p} \sum_{i=1}^{N_p} \left\| \mathbf{w} \odot \xi_H^{r_H,(i)} - \mathbf{w} \odot \xi_H^{(i)} \right\|_2, \quad (40)$$

where " $\odot$ " denotes the element-wise product and  $\mathbf{w}$  indicates the normalized first-order coefficients, i.e.,  $\mathbf{w} := (c_{H,e_1}, \dots, c_{H,e_d}) / \|(c_{H,e_1}, \dots, c_{H,e_d})\|_2$  (obtained from the first row of the rotation matrix  $A_H$ ). The estimator  $\varpi$  is the mean of the weighted  $l^2$ -norm of the difference between  $\{\xi_H^{r_H,(i)}\}_{i=1}^{N_p}$  and  $\{\xi_H^{(i)}\}_{i=1}^{N_p}$ . This approach has the promise of more efficiency than the previous one since the first-order coefficients can be obtained with limited model evaluations.

Algorithm 2 summarizes the numerical procedures designed to obtain both the estimators presented above. After the evaluation of the MSE, a truncation of the adapted space is obtained (according to a user-specified tolerance). It is worth noting that this procedure does not require additional evaluations since the absolute value of the MSE is not computed, but rather the convergence is compared with respect to a threshold level specified as input.

---

**Algorithm 2:** Determination of the adapted HF dimension based on the bias estimators

---

**Input:**  $\{\xi_H^{(i)}\}_{i=1}^{N_p}$ ,  $\{Q_H(\xi_H^{(i)})\}_{i=1}^{N_p}$ : pilot input and corresponding HF samples;  
 $\mathbf{A}_H$ : rotation matrix of the HF model;  
 $\chi$ : estimator indicator ( $\chi = 1$  for estimator 1 and  $\chi = 2$  for estimator 2);  
 $\vartheta$ : threshold of cumulative difference ratio ( $0 < \vartheta < 1$ ).

```

1 if  $\chi = 1$  then
2   for  $r_H = 1, \dots, d_H$  do
3     Compute  $\eta_{H,r_H}^{(i)} = \mathbf{A}_{H,r_H} \xi_H^{(i)}$ , for  $i = 1, \dots, N_p$ ;
4     Compute the adapted PCE coefficients  $c_{H,\gamma}$  in (38) by least square regression;
5     Evaluate  $Q_H(\eta_{H,r_H}^{(i)})$  by the constructed PCE;
6     Compute the MSE of the PCE surface,  $\text{MSE}_{PCE}$ , of Eq (39).
7 else if  $\chi = 2$  then
8   Let  $\omega$  be the first row of  $\mathbf{A}_H$ ;
9   for  $r_H = 1, \dots, d_H$  do
10    Compute  $\xi_H^{r_H,(i)} = \mathbf{A}_{H,r_H}^T \mathbf{A}_{H,r_H} \xi_H^{(i)}$ , for  $i = 1, \dots, N_p$ , and  $\varpi$  by Eq (40).
11 Denote the estimator with respect to  $r_H$  as  $\theta = \{\theta_{r_H}\}_{r_H=1}^{d_H}$  without distinguishing the two estimators;
12 Compute the bias improvement as the adjacent difference of  $\theta$  as  $\theta_{\text{diff}} = \{0, \theta_1 - \theta_2, \theta_2 - \theta_3, \dots, \theta_{d_H-1} - \theta_{d_H}\}$ ;
13 Compute the cumulative bias improvement as the cumulative difference  $\theta_{\text{cdiff}} = \{0, \theta_1 - \theta_2, \theta_1 - \theta_3, \dots, \theta_1 - \theta_{d_H}\}$ ;
14 Compute the cumulative to total difference ratio as

```

$$\varrho = \{\varrho_1, \dots, \varrho_{d_H}\} = \{\theta_{\text{cdiff},r_H} / (\theta_1 - \theta_{d_H})\}_{r_H=1}^{d_H} = \frac{\theta_{\text{cdiff}}}{\theta_1 - \theta_{d_H}}; \quad (41)$$

15 Choose  $r_H$  as the adapted dimension for HF where  $\varrho_{r_H}$  is the first element in  $\varrho$  such that  $\varrho_{r_H} \geq \vartheta$ .

**Output:**  $r_H$ : the adapted dimension of the HF model.

---

#### 4.1.2 Correlation quantification

In this section, we develop a procedure to estimate the variance of the MFAB estimator, which depends on both  $r_H$  and  $r_L$ , as shown in Proposition 4.1. We extend a simplified approach presented in [17], which avoids re-evaluating the models by using the PCE in AB.

Similar to Eq (38), an adapted PCE of LF can be built using the pilot samples as

$$Q_L(\eta_{L,r_L}) = \sum_{\gamma \in \mathcal{J}_{p_L}^{r_L}} c_{L,\gamma} \psi_\gamma(\eta_{L,r_L}), \quad (42)$$

where  $\eta_{L,r_L} = \mathbf{A}_{L,r_L} \xi_{p_L}^{r_L}$ , while  $\mathcal{J}_{p_L}^{r_L}$  represents the sets of PCE multi-indices corresponding to the truncated space for the LF. The expansion is the same as Eq (22), but specific to the LF. The correlation between the HF and LF models can be estimated via the PCE as illustrated in the following proposition.

**Proposition 4.3** (Correlation between models as function of the number of important directions). *Let two models be represented by their PCEs, Eqs (38) and (42), their correlation is expressed by*

$$\rho(Q_H(\eta_{H,r_H}), Q_L(\eta_{L,r_L})) = \frac{\sum_{\gamma \in (\mathcal{J}_{p_H}^{r_H} \cap \mathcal{J}_{p_L}^{r_L}) \setminus \mathbf{e}_0} c_{H,\gamma} c_{L,\gamma}}{\sqrt{\left(\sum_{\gamma \in \mathcal{J}_{p_H}^{r_H} \setminus \mathbf{e}_0} c_{H,\gamma}^2\right) \left(\sum_{\gamma \in \mathcal{J}_{p_L}^{r_L} \setminus \mathbf{e}_0} c_{L,\gamma}^2\right)}}. \quad (43)$$

*Proof.* The proof is provided in Appendix C. □

From Proposition 4.3, the following corollary can be obtained.

**Corollary 4.1** (The correlation between two PCEs is maximized if each model is represented with only the shared multi-indices). *For two PCEs with multi-indices  $\mathcal{J}_{p_H}^{r_H}$  and  $\mathcal{J}_{p_L}^{r_L}$ , their correlation  $\rho$  is maximized if both expansions use the shared multi-indices given by  $\mathcal{J}_{p_H}^{r_H} \cap \mathcal{J}_{p_L}^{r_L}$ , that is, with abuse of notation,*

$$\arg \max_{(\mathcal{J}_{p_H}^{r_H}, \mathcal{J}_{p_L}^{r_L})} \rho^2(Q_H(\mathcal{J}_{p_H}^{r_H}), Q_L(\mathcal{J}_{p_L}^{r_L})) = \left(\mathcal{J}_{p_H}^{r_H} \cap \mathcal{J}_{p_L}^{r_L}, \mathcal{J}_{p_H}^{r_H} \cap \mathcal{J}_{p_L}^{r_L}\right), \quad (44)$$

where  $\rho^2(Q_H(\mathcal{J}_{p_H}^{r_H}), Q_L(\mathcal{J}_{p_L}^{r_L}))$  is obtained by Eq (43). If we assume that the PCE expansion for the HF model includes all multi-indices from the LF model, we can further write  $\mathcal{J}_{p_H}^{r_H} \cap \mathcal{J}_{p_L}^{r_L} = \mathcal{J}_{p_L}^{r_L}$ .

*Proof.* The proof for the later case is provided in Appendix D.  $\square$

Corollary 4.1 demonstrates that increasing the dimension of one model in the adapted space beyond the dimension of the other model reduces the correlation between the HF and LF models.

We observe that the PCE framework provides a straightforward method to estimate the correlation, as explained in Proposition 4.3. Additionally, the framework enables us to comprehend the benefits of sampling on the shared space over sampling in the original space. This proposition is presented below to illustrate this advantage.

**Proposition 4.4** (Correlation in the adapted space is greater than correlation in the original space). *Consider PCEs of the HF and LF of Eqs (38) and (42), but with polynomial order one, i.e.,  $p_H = p_L = 1$ , and with dimensions  $r_H = r_L = d_H = d_L = d$ , implying  $\eta = \eta_s = \eta_{H,r_H} = \eta_{L,r_L}$  for the shared adapted variables. If we*

- I. Assume that the first-order PCE coefficients of HF and LF are greater than or equal to zero;
- II. The rotation matrices on HF and LF only permute the original variables  $\xi$  to rearrange them in decreasing order as in  $\eta$ ;

*then the squared correlation between the two models obtained by sampling in the important directions is greater than the correlation in the physical variables*

$$\rho^2(Q_H(\eta), Q_L(\eta)) \geq \rho^2(Q_H(\xi), Q_L(\xi)). \quad (45)$$

*Proof.* The proof is provided in Appendix E.  $\square$

This proposition demonstrates that if the variables in both the HF and LF models are ranked by their importance and a shared shape is constructed based on this ranking, their correlation can be increased. Although the construction of the rotation in the proposition is simplified compared to AB, the key assumption of rearranging the variables in descending order of importance is still present. Therefore, this result also applies to AB.

Proposition 4.4 states that the correlation between the HF and LF models in the adapted space is at least as good as the correlation in the original space. The AB method arranges the rotated variables in descending order, aligning the dominant polynomial chaos expansion (PCE) terms that have high magnitude coefficients in both the HF and LF models. This key observation is utilized in the proof of Proposition 4.4, presented in Appendix E, to demonstrate that the element-wise product of PCE coefficients for the HF and LF models, with positive values, is maximized when the two coefficient vectors are ranked by their magnitudes. In cases where models do not share a common underlying manifold, the correlation in the rotated space may show negligible improvement. However, our empirical examples demonstrate that for models with shared physics or common causes of variability, the AB method substantially enhances the correlation in the rotated space compared to the original correlation, regardless of whether the models have the same or dissimilar parameterizations.

The propositions presented in this section aim to provide intuitions and are valid under specific assumptions. For instance, they do not consider practical scenarios where truncation can introduce rotation errors. Therefore, it is necessary to develop a numerical strategy for selecting the truncation dimensions and incorporating it into the construction of MFAB, as discussed in the next section.

## 5 Practical implementation of MFAB

In Section 4, we demonstrated how the correlation between models could be maximized by limiting the number of terms in their AB representations. On the other hand, the MSE of the MFAB estimator depends only on the bias of the AB representation of the HF model, which decreases as  $r_H$  approaches  $d_H$ . Moreover, the optimal truncation for the models needs to be determined from a limited number of available pilot samples in the original coordinates. In this section, we introduce the design of MFAB estimators based on these considerations. We provide two MFAB estimators. The first one balances its variance with the residual bias of the HF model in the truncated AB space, aiming to maximize the correlation among the models by minimizing the shared space as much as possible. The second estimator maintains a full representation of the HF model, which does not introduce any bias, but results in a lower correlation between models. These estimators have different data requirements and can be useful in different scenarios, as discussed later.

## 5.1 Bias-variance balanced MFAB estimator

The first MFAB estimator aims to maximize the correlation among the HF and LF models by leveraging the smallest possible important directions  $r_H$  (see Proposition 4.1). As this choice affects the MFAB estimator's MSE (see Proposition 4.2), we need to determine the acceptable estimator bias and then use its value as the target variance in the sample allocation step. By relying on the tools presented in the previous section, this approach is computationally straightforward and is summarized in Algorithm 4, which uses Algorithm 3 as a subroutine.

The approach assumes that additional HF and LF model evaluations are allowed, and we refer to this as the *non-legacy dataset* case. To construct the estimator, we start with collecting pilot samples for both models corresponding to the same coordinates in the original space. Algorithm 1 can then be used for the HF and LF models to obtain their rotation matrices  $A_H$  and  $A_L$ , independently. We use the HF rotation matrix as input (along with other user-defined quantities) for Algorithm 2 to determine the truncation dimension  $r_H$ . Then, we perform a grid search on the truncation  $r_L$  to evaluate the maximum correlation between the two models. Maximizing the correlation is equivalent to maximizing the use of LF information, which, in turn, corresponds to minimizing the estimator's variance for a fixed cost. The shared space can be defined according to Eq (28), which allows us to evaluate additional samples for both models, if needed.

Since Algorithm 4 balances bias and variance, we must estimate a target variance for the MFAB estimator. The two proposed bias estimators introduced in the previous section only capture the convergence trend, so we also need to estimate the bias associated with  $r_H$  in HF. This can be done by evaluating an additional  $N_{r_H}$  HF samples in the shared space, as shown in Eq (46). The MFAB estimator is designed for problems with a small  $r_H$ , so the additional number of  $N_{r_H}$  samples required is also small. Once the shared space has been defined, converging the sample allocation for the MFAB estimator is no different from traditional multilevel/MF estimators, such as MLMC, MPMC, or ACV. This task involves converging on the values for the oversampling ratio  $v$  and the number of HF runs  $N$ , where the total number of LF samples is  $\lceil vN \rceil$ . The detailed process to compute the MFAB estimator is presented in Algorithm 3, which specifies options for non-legacy datasets and bias-variance balancing.

---

### Algorithm 3: MFAB estimator with given adapted HF and LF models

---

**Input:**  $\{\xi_H^{(i)}\}_{i=1}^{N_p}, \{Q_H(\xi_H^{(i)})\}_{i=1}^{N_p}, A_H, A_L, r_H, r_L$ ;  
 Legacy\_flag: using legacy dataset (True) or allow for additional HF runs (False);  
 Bias\_flag: select if the bias-variance balanced (True) estimator is desired  
 $\sigma^2$ : target variance when non-legacy dataset is used.

```

1 if Legacy_flag == True then
2   Project  $\xi_H$  to  $\eta_s$  by Eq (28), map them to  $\xi_L^{r_L}$  by Eq (30), and evaluate  $Q_L(\xi_L^{r_L})$ ;
3   Compute  $v$  by (11);
4 else
5   Specify the target variance  $\sigma^2$ ;
6   if Bias_flag == True then
7     Generate  $N_p$  samples of  $\eta_s$ , map them to  $\xi_H^{r_H}$  and  $\xi_L^{r_L}$  by Eq (30), and evaluate  $Q_H(\xi_H^{r_H})$  and  $Q_L(\xi_L^{r_L})$ ;
8   else
9     Generate  $N_p$  samples of  $\xi_H$ , project them to  $\eta_s$ , map  $\eta_s$  to  $\xi_L^{r_L}$ , and evaluate  $Q_H(\xi_H)$  and  $Q_L(\xi_L^{r_L})$ ;
10  while True do
11    Compute  $v$  by (11) and the required HF samples  $N$  by (12);
12    Compute additional HF samples  $\Delta N = N - N_p$ ;
13    if  $\Delta N = 0$  then
14      Break;
15    else
16      if Bias_flag == True then
17        Generate  $\Delta N$  samples of  $\eta_s$ , map them to  $\xi_H^{r_H}$  and  $\xi_L^{r_L}$ , and evaluate  $Q_H(\xi_H^{r_H})$  and  $Q_L(\xi_L^{r_L})$ ;
18      else
19        Generate  $\Delta N$  samples of  $\xi_H$ , map to  $\eta_s$  and then to  $\xi_L^{r_L}$ , and evaluate  $Q_H(\xi_H)$  and  $Q_L(\xi_L^{r_L})$ ;
20      Append the new HF and LF samples to the existing samples and update  $N_p = N$ ;
21  Generate  $\lceil (v-1)N \rceil$  samples of  $\eta_{L,r_L}$ , map them to  $\xi_L^{r_L}$ , and evaluate  $Q_L(\xi_L^{r_L})$ ;
22  Append LF samples to the existing samples and compute MFAB estimator by (32) and its variance.
```

**Output:**  $\hat{Q}^{MFAB}$ , MFAB estimator.

---

---

**Algorithm 4:** Bias-variance balanced MFAB estimator.

---

**Input:**  $\{\xi_H^{(i)}\}_{i=1}^{N_p}$ ,  $\{Q_H(\xi_H^{(i)})\}_{i=1}^{N_p}$ ,  $\{\xi_L^{(i)}\}_{i=1}^{N_p}$ ,  $\{Q_L(\xi_L^{(i)})\}_{i=1}^{N_p}$ ;

$\chi$ : bias estimator indicator (choose from  $\{1, 2\}$ );

$\vartheta$ : threshold of cumulative to total difference ratio ( $0 < \vartheta < 1$ );

$N_{r_H}$ : number of additional samples to quantify the bias of the  $r_H$ -d adapted HF model.

1 Construct rotation matrices of  $\mathbf{A}_H$  and  $\mathbf{A}_L$  by Algorithm 1 from pilot samples;

2 Determine  $r_H$  from  $\mathbf{A}_H, \chi, \vartheta$  by Algorithm 2;

3 **Choose the adapted LF dimension with maximized correlation**

4   **for**  $r_L = 1$  to  $d_L$  **do**

5     Compute the correlation of  $r_H$ -d adapted HF and  $r_L$ -d adapted LF by (43);

6   Choose the LF dimension  $r_L$  that maximizes the correlation (43);

7 **Estimate squared bias**  $\delta^2(r_H)$

8   Generate addition  $N_{r_H}$  samples of  $\eta_s$ ,

$$\delta(r_H) \approx \hat{\delta}(r_H) := \frac{1}{N_{r_H}} \sum_{i=1}^{N_{r_H}} Q_H(\mathbf{A}_{H,r_H}^T \eta_s^{(i)}) - \frac{1}{N_p} \sum_{i=1}^{N_p} Q(\xi_H^{(i)}). \quad (46)$$

9 Compute MFAB estimator by Algorithm 3 with `Legacy_flag=False`, `Bias_flag=True`, and target variance  $\sigma^2 = \hat{\delta}^2(r_H)$ .

**Output:**  $\hat{Q}^{MFAB}$ , MFAB estimator.

---

## 5.2 Unbiased MFAB estimator for *legacy* and *non-legacy* HF data

As shown in the previous section, constructing the bias-variance balanced MFAB estimator always requires evaluating additional HF samples by sampling the shared space  $\eta_s$ . However, it's not always possible to obtain additional evaluations for the HF model, such as when using a *legacy dataset*. In this case, an MFAB estimator can still be obtained by sampling the LF model to determine the optimal sample allocation, while the number of HF evaluations  $N$  remains fixed. This corresponds to using a full (non-truncated) representation of the HF model, while the LF model can still be arbitrarily reduced.

One key difference from the bias-variance balanced strategy is that the estimator is unbiased because the HF model is non-truncated. The shared space spanned by  $\eta_s$  can still be constructed and used to maximize correlation. However, the shared samples are not sampled directly but obtained by mapping the available HF realizations to this space through the rotation matrix  $\mathbf{A}_H$ .

A numerical procedure to build this MFAB estimator is presented in Algorithm 5, where the subroutine Algorithm 3 is used with the `Legacy_flag` set to `True`. First, we search for the HF important directions that can represent the model accurately by Algorithm 2. This step does not require re-evaluating the HF model to determine  $\delta^2$ , i.e.,  $\delta^2 = 0$ , given that HF samples are evaluated on the full space. Second, we map the set of HF samples from its original space to its auxiliary space by

$$\eta_{H,r_H} = \mathbf{A}_{H,r_H} \xi_H. \quad (47)$$

For an assigned number of important directions  $r_L$  for the LF, we can assemble the shared space  $\eta_s$  by Eq (28) and then proceed by sampling the LF given the mapping  $\xi_L^{r_L} = \mathbf{A}_{L,r_L}^T \eta_s$ . For the last step of the algorithm, we need to converge the oversampling ratio  $v$ ; by doing so we minimize the variance for the estimator with an assigned number of HF realizations.

The remaining question is how to select the optimal number of important directions for the LF model. Conceptually, this can be done similarly to the previous case (Algorithm 4) by performing a grid search over  $r_L$ . However, there is a notable difference: we cannot rely on the correlation expression introduced in Proposition 4.3, as the HF model is not represented by its AB counterpart. Hence, it is necessary here to re-evaluate the LF model for each  $1 \leq r_L \leq d_L$  and explicitly compute the correlation between the models as

$$\rho^2(Q_H(\xi_H), Q_L(\xi_L^{r_L})) \approx \frac{\left( \sum_{i=1}^{N_p} \left( Q_H(\xi_H^{(i)}) - \hat{Q}_H(\xi_H^{(i)}) \right) \left( Q_L(\xi_L^{r_L,(i)}) - \hat{Q}_L(\xi_L^{r_L,(i)}) \right) \right)^2}{\sum_{i=1}^{N_p} \left( Q_H(\xi_H^{(i)}) - \hat{Q}_H(\xi_H^{(i)}) \right)^2 \sum_{i=1}^{N_p} \left( Q_L(\xi_L^{r_L,(i)}) - \hat{Q}_L(\xi_L^{r_L,(i)}) \right)^2}. \quad (48)$$

This process incurs the additional cost of discarding LF evaluations, which, depending on the application and the computational cost of the model, can be more or less feasible. To alleviate the cost of this grid search, we limit it to the space  $1 \leq r_L \leq r_H$  (knowing that increasing it would not help, see Corollary 4.1).

---

**Algorithm 5:** Unbiased MFAB estimator for legacy and non-legacy HF dataset.

---

**Input:**  $\{\xi_H^{(i)}\}_{i=1}^{N_p}$ ,  $\{Q_H(\xi_H^{(i)})\}_{i=1}^{N_p}$ ,  $\{\xi_L^{(i)}\}_{i=1}^{N_p}$ ,  $\{Q_L(\xi_L^{(i)})\}_{i=1}^{N_p}$ ;

$\chi$ : bias estimator indicator (choose from  $\{1, 2\}$ );

$\vartheta$ : threshold of cumulative to total difference ratio ( $0 < \vartheta < 1$ );

Legacy\_flag: using legacy dataset if True else using non-legacy dataset with target variance.

- 1 Construct rotation matrices of  $\mathbf{A}_H$  and  $\mathbf{A}_L$  by Algorithm 1 from pilot samples;
- 2 Determine  $r_H$  from  $\mathbf{A}_H, \chi, \vartheta$  by Algorithm 2;
- 3 **Choose the adapted LF dimension with maximized correlation**
- 4   **for**  $r_L = 1$  to  $r_H$  **do**
- 5     Project  $\xi_H$  to  $\eta_{H, r_H}$  by (47), and assemble  $\eta_s$  by (28);
- 6     Map  $\eta_s$  to  $\xi_L^{r_L}$  by the second equation in Eq (30) and evaluate  $Q_L(\xi_L^{r_L})$ ;
- 7     Compute the correlation of  $r_H$ -d projected HF and  $r_L$ -d adapted LF by (48);
- 8   Choose the LF dimension  $r_L$  that maximizes the correlation expressed by Eq (48);
- 9 Compute MFAB estimator by Algorithm 3 with Bias\_flag=False.

**Output:**  $\hat{Q}^{MFAB}$ , MFAB estimator.

---

Due to the projection error of the HF model, the correlation in this case is smaller than when both models are generating input samples in the reduced space (see Proposition 4.3). Since this contribution would depend on the effect introduced by the truncated  $(d_H - r_H)$  variables, it is expected to decay rapidly once the model's response has been captured by the first  $r_H$  important variables.

In principle, the approach with the legacy HF dataset can be easily extended to the non-legacy HF dataset, which is also presented in Algorithms 5 and 3 by using the *non-legacy dataset* option. The only modification required would happen in the resource allocation step of the Algorithm 3. With the unbiased option, the HF model must be sampled in its original space to obtain the optimal  $N$ . The resulting algorithm would be an unbiased MFAB estimator with prescribed target variance, which could provide an alternative to Algorithm 4. We will discuss and compare these two algorithms in the next section.

## 6 Numerical examples

In this section, we present several numerical results to demonstrate the features of the MFAB approach. First, we adopt an analytical test problem in Section 6.1 to discuss a wide range of results, given the availability of both an exact solution and the low computational cost of the models. Afterward, we illustrate two more challenging computational problems. In Section 6.2, we consider an acoustic problem inspired by a direct field acoustic test, while in Section 6.3, we use a finite element model for a realistic nuclear fuel assembly.

Before considering the numerical examples, let us return to the nozzle example introduced in Section 2 to demonstrate the increase in correlation in both MF scenarios. The purpose of this simple demonstration is to support our intuition that relying on a shared space to correlate models with dissimilar parameterizations can be understood from the physical understanding of the problem<sup>3</sup>. For the nozzle example, a shared space with a single important direction is sufficient to correlate the elliptical and circular high-fidelity cases, in which the elliptical case has 3 parameters for the HF model and 2 parameters for the LF model. In both cases, the LF model is a 2D approximation with only two parameters: one for the geometry, *in lieu* of the two geometrical parameters of the HF model, and one for the total pressure, which is shared by the models. In Figure 2 shows the scatter plots for these two scenarios; this figure augments Figure 1 with the datapoints obtained by sampling the models at the same shared locations. We note that, in both scenarios, the correlation obtained by sampling the models in the shared space is significantly increased compared to the correlation obtained by sampling the original coordinates. As described in the previous sections, the correlation is the main parameter controlling the variance of the MF estimator (along with the model costs). Therefore, we can easily link the improvement in correlation to the improvement in MFAB efficiency. We will discuss the performance of the MFAB estimator in detail for the numerical examples in the following sections.

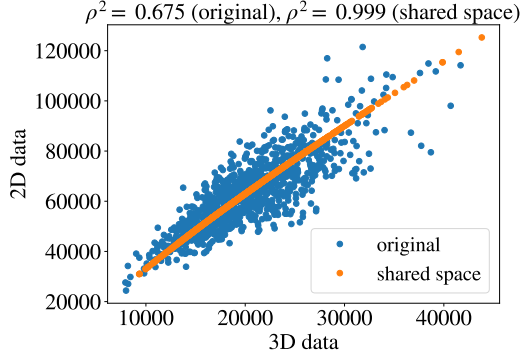
### 6.1 Analytical test problem

The first example we consider is an analytical test case, which, by providing closed-form solutions, is helpful to illustrate and verify the algorithm. Consider two models: a HF model  $f = f(\mathbf{x})$ , characterized by 10 uncertain parameters, i.e.,  $\mathbf{x} \in \mathbb{R}^{10}$ ,

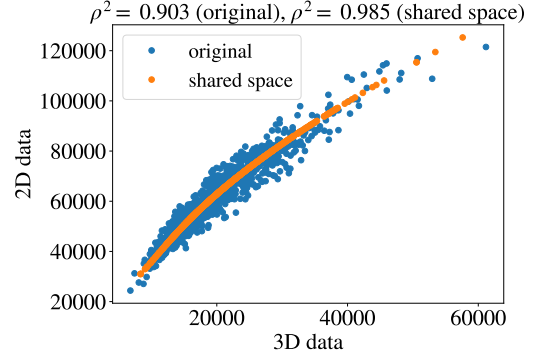
---

<sup>3</sup>similar interpretations for more complex problems, like the ones presented in the following sections, are hampered by the complexity of the underlying physical models





(a) Exit pressure  $P_e$  [Pa] for the elliptical geometry section



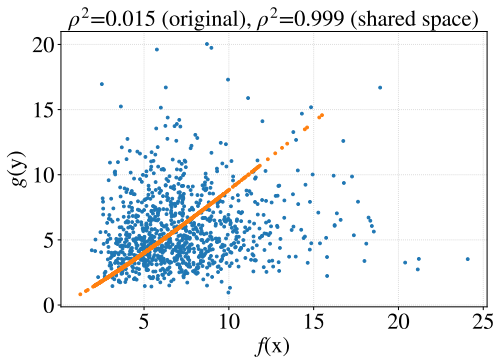
(b) Exit pressure  $P_e$  [Pa] for the circular geometry section

Figure 2: Scatter plot for the nozzle flow in the elliptical (a) and circular (b) geometry case. This figure augments Figure 1 by overlaying the scatter plot of the model sampled in the shared space (orange dots) with the one corresponding to the models sampled in the original space (blue dots). In both cases, the correlation is greatly increased, demonstrating how the variability of the models can be well-captured by resorting to a 1D shared space even if the models are described by different parameters.

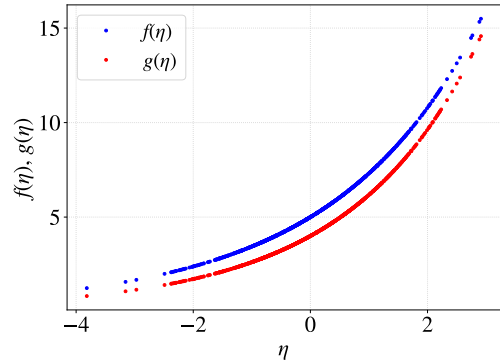
and a LF model  $g = g(\mathbf{y})$  with 8 uncertain inputs, i.e.,  $\mathbf{y} \in \mathbb{R}^8$ , defined by

$$\begin{aligned} f(\mathbf{x}) &= \exp(x_1 + 0.05x_2) + \exp(0.8x_3) + \exp(0.8x_4 + 0.05x_5 + 0.05x_6) \\ &\quad + \exp(0.8x_7 + 0.05x_8) + \exp(0.08x_9 + 0.05x_{10}), \\ g(\mathbf{y}) &= \exp(0.1y_1 + y_2) + \exp(0.1y_3 + 0.01y_4 + 0.8y_5) + \exp(0.8y_6 + 0.1y_7) + \exp(0.8y_8). \end{aligned} \quad (49)$$

We assume  $x_1, \dots, x_{10} \sim \mathcal{N}(0, 1)$  and  $y_1, \dots, y_8 \sim \mathcal{N}(0, 1)$ . To illustrate the characteristics of the models, in Figure 3 we report the scatter plots (1000 samples) for the models sampled in the original and shared coordinates. Hereinafter, we refer to the correlation in the original space as the one obtained by sampling the variables in the original coordinates, e.g.,  $\mathbf{x}$  and  $\mathbf{y}$  here. In contrast, the correlation in the adapted variables is obtained by first sampling the shared variables  $\boldsymbol{\eta}_s$  and then evaluating the computational models at their inputs obtained after the transformation, i.e.,  $f(\mathbf{x}(\boldsymbol{\eta}_s))$  and  $g(\mathbf{y}(\boldsymbol{\eta}_s))$  here. As explained in the previous sections, the two correlations are indeed different since the rotation matrices are distinct for the two models. Moreover, since the transformation from the shared variables to the models' original variables depends on the number of directions retained in the rotation matrix, the correlation in the adapted variables is a function of the value  $r$  chosen for both the HF and LF; see 4.1.2. Here, we used a single shared variable such that the models' response over this variable can be easily represented (see Figure 3(b)). As visible from the plot, the correlation among the models is greatly increased by resorting to the shared variables.



(a) Scatter plots



(b) Models' response over the first shared variable

Figure 3: (a) scatter plots for the models sampled in the original and shared spaces with a single shared variable and (b) their responses over the shared variable.

The MFAB algorithms described in Algorithms 4 and 5 rely on estimating and controlling the bias of the MFAB estimator.

In our application, we use  $N_p = 40$  pilot samples, but we also want to estimate statistics in the full dimensional spaces, where  $d_H = 10$  and  $d_L = 8$ , so we use a larger dataset of  $N'_p = 100$  samples. The bias estimators introduced in Section 4.1.1 (with  $\chi = 1, 2$  in Algorithm 2) are both computed and reported in Figure 4. From the figure, we can see that using five variables significantly decreases the model's bias, while the remaining five only have a marginal effect.

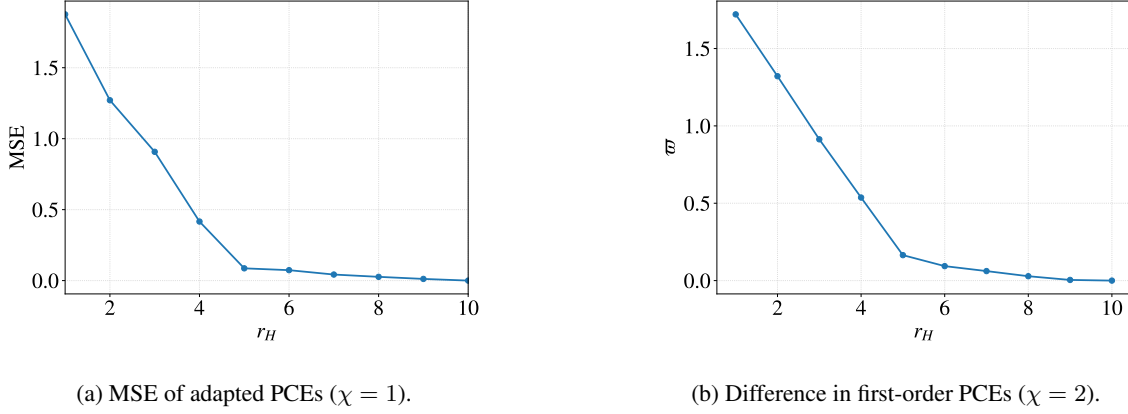


Figure 4: Comparison of the bias estimators introduced in Section 4.1.1 for the analytical test problem.

The cumulative to total difference ratio of the two metrics can be computed following Algorithm 2 with results reported in Figure 5. We use  $\vartheta = 1$  to show all achievable levels as a function of the HF dimension.

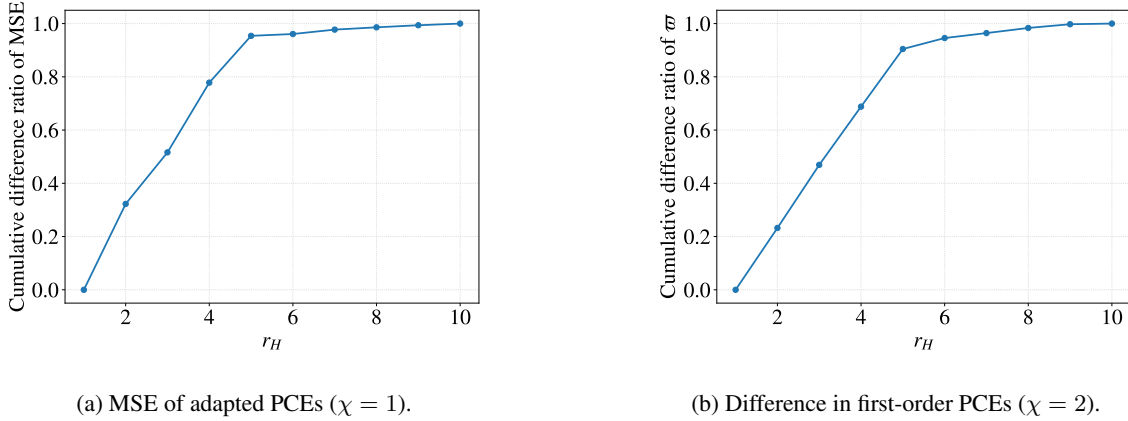


Figure 5: Adaptive selection of the HF model dimension based on the bias estimators introduced in Section 4.1.1 and Algorithm 2 with  $\vartheta = 1$  for the analytical test problem.

### 6.1.1 Bias-variance balanced estimator

We start by presenting the bias-variance balanced estimator described in Algorithm 4. For illustration purposes, in Figure 6, we report the correlation (Proposition 4.3) among models, for different truncation dimensions  $r_H$  and  $r_L$ , by employing  $N'_p = 100$  samples.

To estimate the bias of the HF model with user-specified  $r_H$  in step 8 of Algorithm 4, we sample the model on the shared space with additional  $N_{r_H}$  samples and evaluate the bias using Eq (46) in Algorithm 4. The required number of additional HF samples,  $N_{r_H}$ , is linked to  $r_H$ . For  $r_H = 1$ , we choose  $N_{r_H} = 10$  and estimate the squared bias to be approximately  $\hat{\delta}^2 = 3.1$ . For  $r_H = 3$ , we choose  $N_{r_H} = 20$  and obtain  $\hat{\delta}^2 = 0.5$ , and for  $r_H = 5$ , we choose  $N_{r_H} = 30$  and obtain  $\hat{\delta}^2 = 0.04$ . Once the HF dimension is fixed, the LF dimension is chosen to maximize correlation, which can be determined through a grid search as shown in Figure 6.

Once the dimensions of adapted HF and adapted LF are fixed, one can specify the desired level of variance of the MF estimator and run the optimization problem (Algorithm 3 with *non-legacy dataset* and bias-variance balanced options) to find parameters  $N, v, \alpha$  by Eq (32). Here, to balance the two contributions of the MSE, we require  $\sigma^2 = \hat{\delta}^2$ .

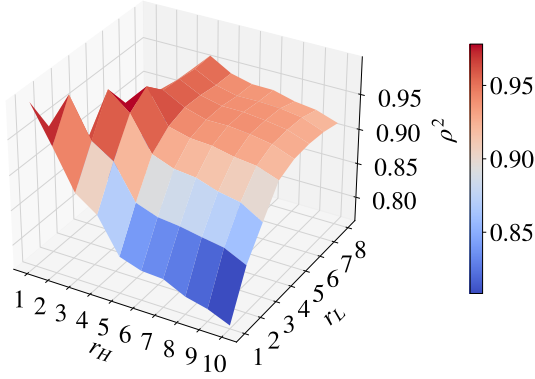


Figure 6: Estimated correlation for the two models in the analytical test problem following Proposition 4.3 as function of the truncated dimensions  $r_H$  and  $r_L$ .

To demonstrate the performance of the MFAB estimator, we conduct multiple repetitions based on a pilot set containing 40 samples. We run 500 repetitions of the estimators to obtain their probability density functions (PDFs), as shown in Figure 7a. We use the MFAB estimator cost to compute the equivalent single-fidelity MC and MF estimators, both of which are obtained by sampling the models in their original coordinates. The mean value of the MC estimator is represented by a vertical solid line in Figure 7a.

The squared correlation between HF and LF for the MFAB estimator has significantly increased to 0.996 from about 0.017. As a result, the MFAB estimator has reduced the variance from 0.0515 to 0.007 compared to the MC estimator. With high correlation between HF and LF, the variance of the MFAB estimator is already less than the target variance,  $\hat{\delta}^2 = 3.1$ , with only 40 HF samples. However, it is also observed that, although the variance of the MFAB estimator has decreased significantly, it has a substantial bias due to the model reduction of the HF model. In this case, the MSE can be further reduced by increasing  $r_H$ . For instance, for  $r_H = 3$ , an optimal value of  $r_L = 3$  can be selected, as shown in Figure 6. The 500 repetitions for this case are reported in Figure 7b for the three estimators. The correlation of LF and HF in MFAB has dropped to 0.877 compared to the previous case; however, it is still a substantial increase from 0.031. Moreover, the optimal MFAB estimator can achieve a variance of 0.0498, which is again a significant improvement from the variance of the MC estimator of 0.199. As expected, the bias of the MFAB estimator has decreased compared to the case of  $r_H = 1$ .

Similar results can be obtained for  $r_H = 5$  and  $r_L = 5$ , where  $\hat{\delta}^2 = 0.04$ , as shown in Figure 7c. The correlation of HF and LF in the MFAB estimator decreases with respect to  $(r_H, r_L) = (3, 3)$ ; however, it is still much larger than that in the MF estimator obtained by sampling the models in their original space. The optimal estimator to obtain an estimator with a variance less than 0.04 has a variance of 0.0333, substantially less than the MC and classical MF estimator with the same cost. Moreover, although the correlation of HF and LF has dropped slightly, the bias of the MFAB estimator in this case is much smaller than in the previous one. A virtually unbiased estimator can be obtained by increasing  $r_H$  even more, e.g.,  $r_H = 6$ . The results are reported in Figure 7d; in this case, the optimal correlation is obtained with  $r_L = 5$ .

### 6.1.2 Unbiased MFAB estimator

In this case, we still assume non-legacy data. Algorithm 5 can be used to construct the MFAB estimator with `bias_flag=False` in the subroutine Algorithm 3.

The first two steps of Algorithm 5 are identical to those in Algorithm 4. The adapted HF model is converged using  $r_H = 5$ . In the next step of Algorithm 5, we compute the correlation between HF and LF using Eq (48), with LF re-evaluated on the shared space for  $r_L = 1, \dots, r_H$ . The results are shown in Figure 8a, where we expand the correlation calculation to full dimensions for both HF and LF for demonstration purposes only. When  $r_H = 5$ , the value of  $r_L = 5$  maximizes their correlation. However,  $r_H \times N_p$  evaluations of the LF model are required, which is the additional cost of the unbiased strategy. In this case, the correlation cannot be estimated without re-running the LF model.

We can specify a desired level of variance and compute the optimal MF estimator by finding the parameters  $N$ ,  $v$ ,  $\alpha$  that satisfy Eq (32). To compare the results with those obtained with the biased strategy in the previous section, we choose  $\sigma^2 = 0.04$ . We calculate three estimators 500 times, and their PDFs are presented in Figure 8b. The corresponding (equivalent total cost) MC and MF estimators are also reported for comparison. We observe that the correlation between HF and LF for the MFAB estimator is 0.785, which is a significant improvement from the correlation of 0.021 in the original space. As a

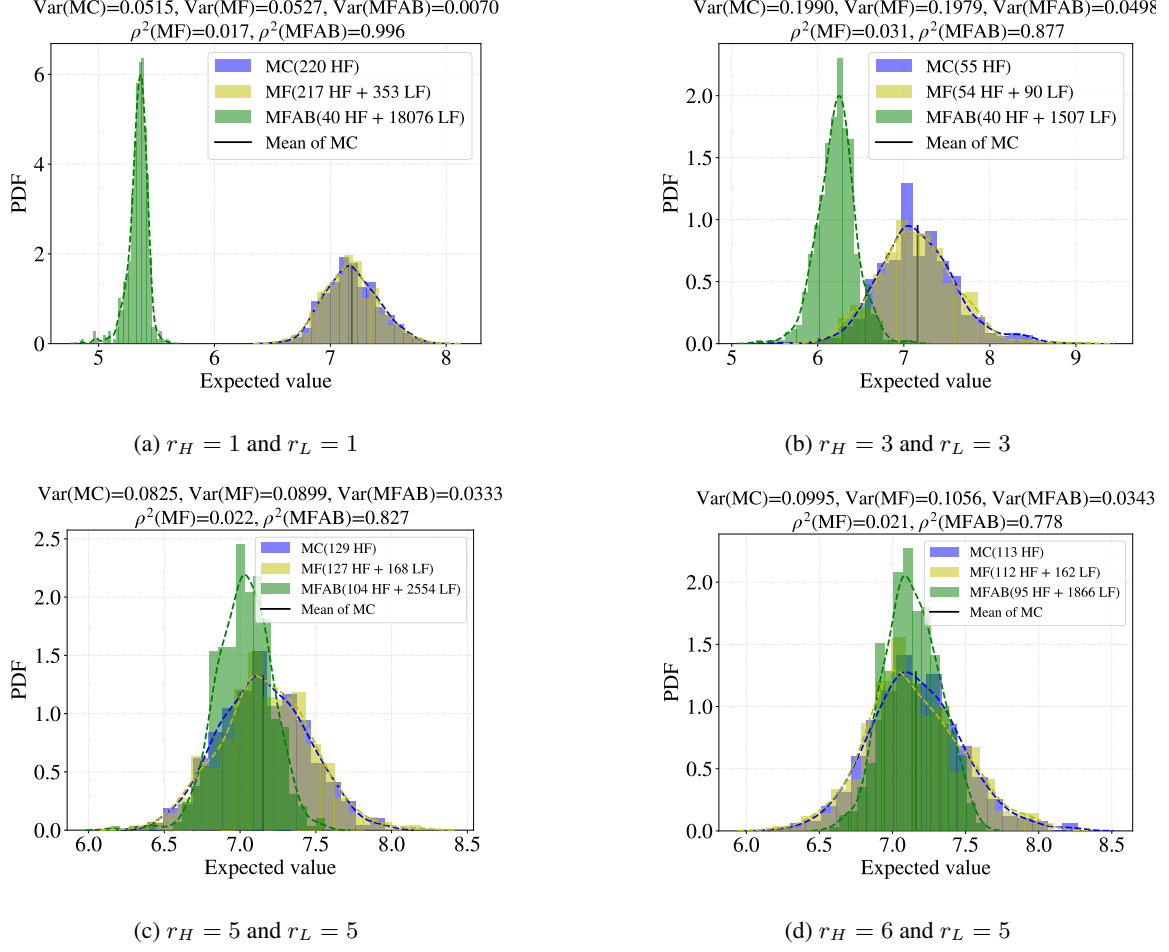


Figure 7: Probability density functions for 500 realizations of the MFAB, MF, and MC estimators for the analytical test problem with different choices of  $r_H$  and  $r_L$ . Increasing the number of dimensions  $r_H$  contributes to reducing the bias, as predicted by the theory, while the number of LF dimensions  $r_L$  is determined to maximize the correlation between the models.

consequence, the variance of the MFAB estimator is reduced to 0.0324 from the value of 0.0943 for the MC estimator, which represents a gain of approximately 65%. As in the previous case, the MF estimator, which samples the models in their original coordinates, cannot achieve better results than MC since the correlation between them is too low.

## 6.2 A direct field acoustic testing application

We will now introduce a MF benchmark case, which is included in the PyApprox software [34] and is inspired by a direct field acoustic test described in [61]. The problem involves a circular scatterer with a circular air inclusion placed in the center of an octagonal domain, where each side is a speaker with its cabinet (see Figure 9(a)). The acoustic pressure  $u$  in the octagonal domain  $D$  is governed by the real Helmholtz equation under the assumptions of the scatterer being a dense fluid and the absence of impedance for the speakers' cabinets. For a fixed angular frequency  $\omega = 2\pi\nu$ , the equation is given by

$$\begin{aligned} \Delta u + \kappa^2 u &= 0 \quad \text{in } D \\ \frac{\partial u}{\partial n} &= \rho_0 \omega \sum_{j=1}^s h_j \chi_j \quad \text{on } \partial D, \end{aligned} \quad (50)$$

where  $\kappa = \omega/c$  is the wavenumber,  $c$  is the local speed of sound,  $\rho_0$  is the fluid density, and  $\chi_j : \partial D \rightarrow 0, 1$  is the characteristic function of the  $j$ th speaker, which oscillates with velocity  $h_j \cos(\omega t)$ . The QoI in this problem is the average sound pressure, which is measured in the red rectangle area induced by eight loudspeakers located at the center of each side of an octagonal speaker cabinet, as shown in Figure 9(a). Our goal is to estimate the expected value of this QoI.

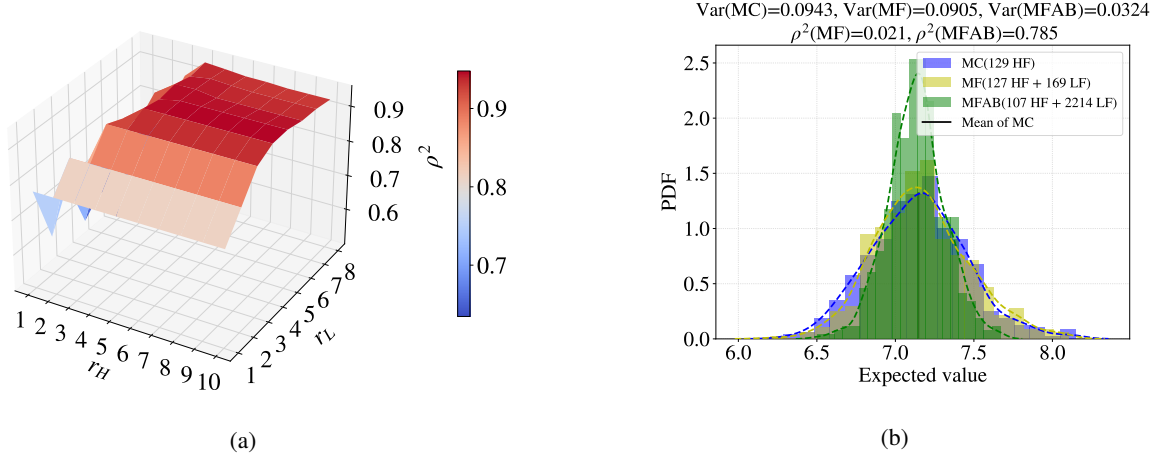


Figure 8: (a) Estimated correlation for the two models in the analytical test problem with the unbiased strategy, and (b) probability density functions for 500 realizations of the MFAB, MF, and MC estimators, where the unbiased strategy is used in the MFAB estimator with  $r_H = 5$  and  $r_L = 5$ .

In this problem, the pressure is obtained by activating the speakers, whose amplitudes follow a uniform distribution. The scatterer is considered to be made of aluminum, for which the sound speed varies as  $\mathcal{U}(6000, 6600)$  m/s in aluminum, and the density of air is distributed as  $\mathcal{U}(0.8, 1.6)$  Kg/m<sup>3</sup>. The liner frequency is  $\nu = 400$ Hz. The complete list of random variables for this problem is summarized in Table 1.

Uncertain Parameter	Distribution	Units
Amplitude oscillation for Speakers 1 and 2 ( $\theta_1$ )	$\mathcal{U}[1, 9]$	m/s
Amplitude oscillation for Speakers 3 and 4 ( $\theta_2$ )	$\mathcal{U}[1, 9]$	m/s
Sound speed of aluminum ( $\theta_3$ )	$\mathcal{U}[6000, 6600]$	m/s
Air density ( $\theta_4$ )	$\mathcal{U}[0.8, 1.6]$	Kg/m <sup>3</sup>
Amplitude oscillation for Speakers 5 and 6 ( $\theta_5$ )	$\mathcal{U}[1, 9]$	m/s
Amplitude oscillation for Speakers 7 and 8 ( $\theta_6$ )	$\mathcal{U}[1, 9]$	m/s

Table 1: Uncertain parameters and distribution for the Acoustic problem. Speakers are ordered counterclockwise, with the first speaker located on the right vertical edge of the octagon.

The Helmholtz equation is solved using a finite element (FE) method, which involves linear superposition of basis functions. The HF model has a dimension of 6 and is denoted as  $f(\mathbf{x})$ , where  $\mathbf{x} = [\theta_1, \dots, \theta_6]^T$ , and the mesh size for the FE, denoting the number of element in one side of the speaker cabinet, is 21. The LF model, denoted as  $g(\mathbf{y})$ , has a dimension of 4 and includes only the first four activated speakers, with their distributions being the same as in the HF model. The sound speed in aluminum and atmospheric pressure are assumed to follow the same distributions as in the HF model. The mesh size for the LF model is 11, which is coarser than the HF model, resulting in a cost of only 10% of the HF model. Figure 9(b) presents scatter plots (500 samples) for the models sampled in the original and shared coordinates. For demonstration purposes, a single shared variable is used. The scatter plots show that the correlation among the models significantly increases by using shared variables.

To implement Algorithms 4 and 5, the first two steps involve estimating and controlling the estimator bias. We use  $N_p = 40$  pilot samples in this application. However, for demonstration purposes, we employ a larger dataset of  $N'_p = 100$  samples to explore the full space for quantities like bias and correlation, as we did in the previous section. Algorithm 2 with  $\chi = 1, 2$  is used to compute the two bias metrics, which are presented in Figure 10. We notice a rapid decrease in bias as the dimension of the adapted HF increases from 1 to 3, after which the bias remains relatively stable, indicating that the adapted HF will converge to the full-dimensional model with a dimension of 3. Furthermore, we compute the cumulative-to-total difference ratio of the two bias metrics using Algorithm 2, and the results are shown in Figure 11 for all achievable levels as a function of the HF dimension with  $\vartheta = 1$ .

### 6.2.1 Bias-variance balanced estimator

The correlations estimated for different truncation dimensions  $r_H$  and  $r_L$  are presented in Figure 12. The bias is estimated

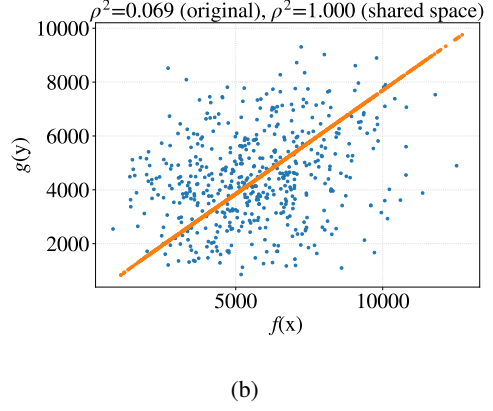
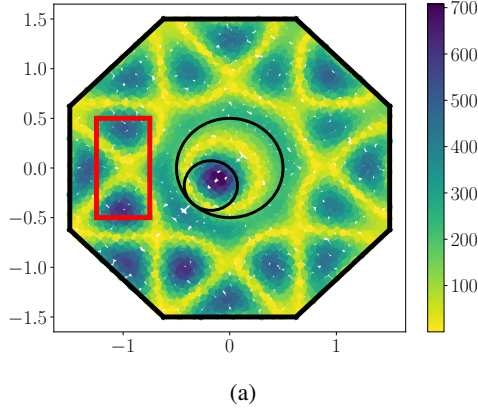
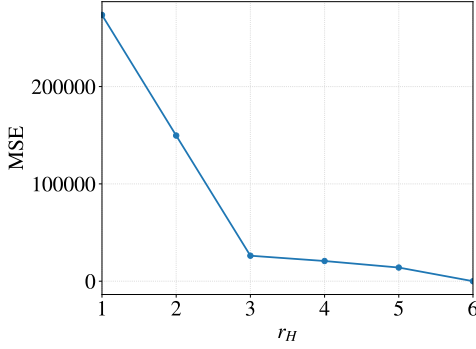
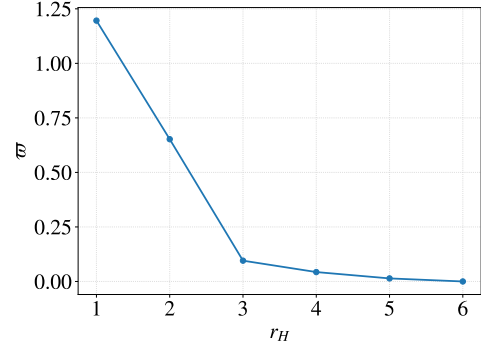


Figure 9: (a) Example pressure plot of the speaker cabinet. The QoI is the average pressure inside the red rectangle. (b) Scatter plot of HF and LF in the original and shared spaces with a single shared variable of the acoustic application.

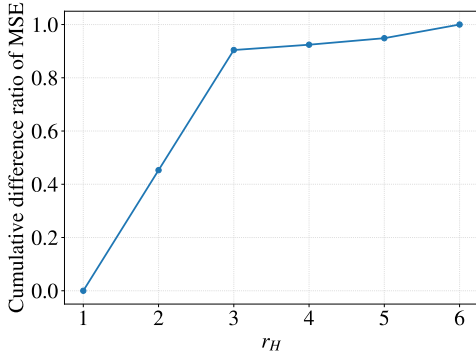


(a) MSE of adapted PCEs ( $\chi = 1$ ).

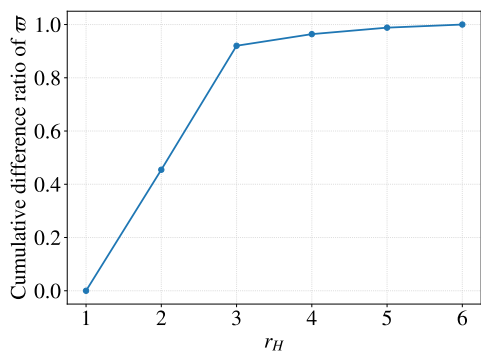


(b) Difference in the first-order PCEs ( $\chi = 2$ ).

Figure 10: Comparison of the bias estimators introduced in Section 4.1.1 for the acoustic application.



(a) MSE of adapted PCEs ( $\chi = 1$ ).



(b) Difference in first-order PCEs ( $\chi = 2$ ).

Figure 11: Adaptive selection of the HF model dimension based on the bias estimators introduced in Section 4.1.1 and Algorithm 2 with  $\vartheta = 1$  for the acoustic application.

using an additional  $N_{r_H}$  samples. For  $r_H = 1$ , we choose  $N_{r_H} = 10$ , and the estimated squared bias is  $\hat{\delta}^2 = 29864.0$ . For  $r_H = 2$ , we choose  $N_{r_H} = 20$ , and obtain  $\hat{\delta}^2 = 20201.4$ . For  $r_H = 3$ , we choose  $N_{r_H} = 30$ , and obtain  $\hat{\delta}^2 = 2496.3$ . As with the previous example, we consider different truncations. For  $r_H = r_L = 1$ ,  $\sigma^2 = \hat{\delta}^2 = 29864.0$ . Figure 13a shows the results for the three estimators (MC, MF, and MFAB) for 500 repetitions obtained with  $N_p = 40$ . From Figure 13a, we can see that compared to the MF estimator, the squared correlation between HF and LF of the MFAB estimator has increased from 0.088 to



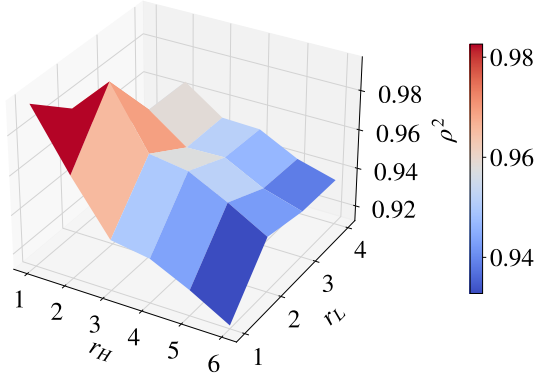


Figure 12: Estimated correlation for the two models in the acoustic application following Proposition 4.3 as function of the truncated dimensions  $r_H$  and  $r_L$ .

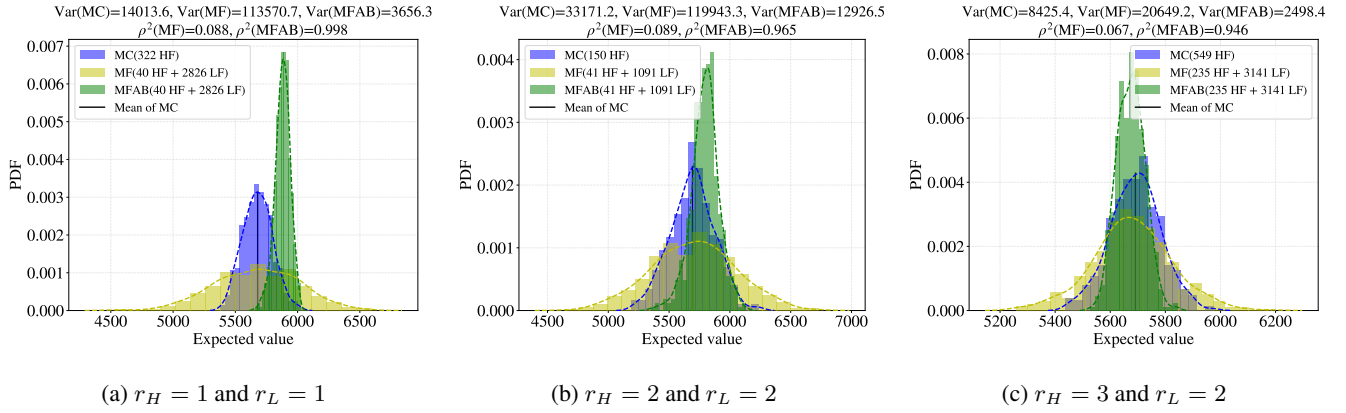


Figure 13: Probability density functions for 500 realizations of the MFAB, MF, and MC estimators for the acoustic application with different choices of  $r_H$  and  $r_L$ . Increasing the number of dimensions  $r_H$  contributes to reducing the bias, as predicted by the theory, while the number of LF dimensions  $r_L$  is determined to maximize the correlation between the models.

0.998. As a result, the MFAB estimator has a much-reduced variance compared to the MC estimator, and the former is about 26% of the latter. Additionally, due to the high correlation between HF and LF, the variance of the MFAB estimator is already less than the target variance,  $\hat{\sigma}^2 = 29864.0$ , with 40 HF samples. However, we can also see a significant bias introduced in the estimator. A bias reduction can be obtained for  $r_H = 2$ , for which the optimal truncation  $r_L = 2$  is obtained; see Figure 12. For this case, the 500 repetitions for the estimators are reported in Figure 13b. Compared to the previous case, the squared correlation has dropped to 0.965, but it is still a significant improvement from sampling in the original coordinates, which only achieves 0.089. Therefore, the variance of the MFAB estimator is only 39% of the MC estimator. The MFAB estimator's bias has also decreased from the previous case. A virtually unbiased estimator can be obtained using  $r_H = 3$  and  $r_L = 2$ , for which a target variance of  $\hat{\sigma}^2 = 2496.3$  is obtained. The 500 repetitions for the three estimators are reported in Figure 13c. The squared correlation between HF and LF has decreased slightly from 0.965 to 0.946 compared to the case when  $(r_H, r_L) = (2, 2)$ , but the variance of the MFAB estimator is only 30% of its MC counterpart. As by design, the bias of the MFAB estimator is almost negligible.

### 6.2.2 Unbiased MFAB estimator

In this application, we can still afford to re-run the HF model and, as a consequence, we assume non-legacy data options in the algorithm. Figure 14a illustrates the correlations for the complete range of choices for  $r_H$  and  $r_L$ . We observe that the maximum correlation is obtained when  $r_H = 3$  and  $r_L = 2$ . At this step,  $r_H \times N_p$  LF evaluations are performed, which is the additional cost of the unbiased strategy. To ensure fair comparison, we use the same target variance as in the previous case, i.e.,  $\sigma^2 = 2496.3$ . Figure 14b presents the 500 repetitions for the three estimators obtained using 40 pilot samples. The squared

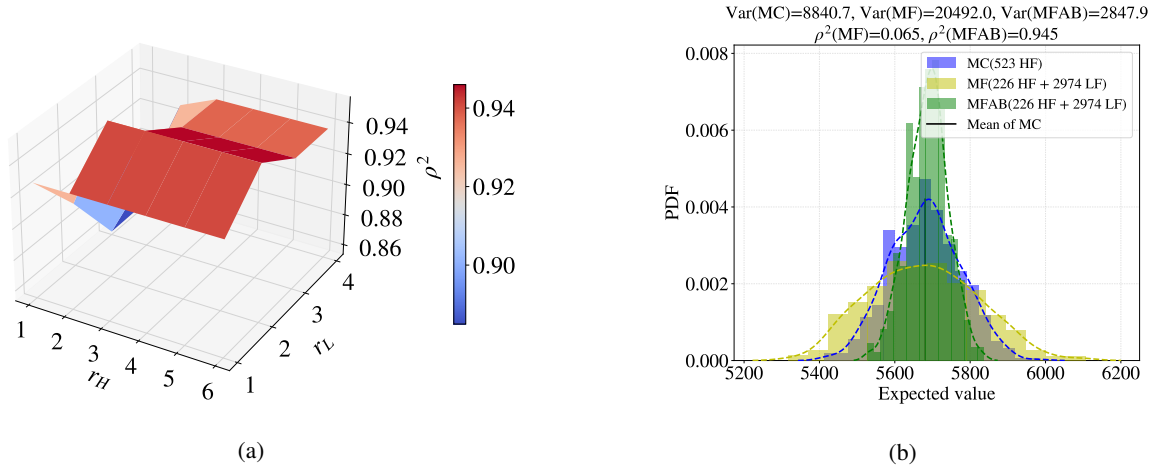


Figure 14: (a) Estimated correlation for the two models in the acoustic application with unbiased strategy, and (b) probability density functions for 500 realizations of the MFAB, MF, and MC estimators, where the unbiased strategy is used in the MFAB estimator with  $r_H = 3$  and  $r_L = 2$ .

correlation has increased from 0.065 to 0.945, leading to a substantial decrease in the variance of the MFAB estimator, which is only 32% of the MC estimator. Moreover, as expected, the MFAB estimator is unbiased. In this application, the oversampling ratio  $v$  for the MF estimator is assumed to be the same as the MFAB estimator since the ratio  $v$  computed by the pilot samples is less than 1, implying that leveraging the LF model in the MF estimator has no gain when sampling the original parameters space. Since we are also forcing the estimators to have the same cost, the number of HF samples in the MF estimator is less than in the MC estimator. Since the additional LF samples in the MF estimator does not help to improve the performance, the MF estimator has a greater variance than the MC estimator.

### 6.3 Nuclear spent fuel assembly model

We will next demonstrate the application of the developed methodologies to a nuclear fuel assembly model. The fuel assembly model is based on the General Electric GE14 design [33] and is utilized in boiling water reactors. The model is a slender structure with dimensions of 4517.7 mm in length and 140.2 mm in width. The fuel rods are arranged in a  $10 \times 10$  array, held together by eight uniformly spaced spacer grids, and the upper and lower tie plates. Two water rods are inserted after removing eight fuel rods. These components are placed in a slender squared tube known as the channel. In the finite element (FE) model, linear springs are used to model the connections between the fuel rods and spacer grids, as well as between the spacer grids and the channel. The channel is connected to the lower tie plate by finger springs and is attached to a post of the upper tie plate. The FE model assumes that the channel is clamped to the lower tie plate and the post of the upper tie plate. Figure 15a illustrates local views of the FE model.

The FE model consists of  $N = 1,912,506$  degrees-of-freedom (DOFs). The analysis focuses on vibration in the frequency range  $\mathcal{B} = 2\pi \times [0, 1000]$  rad/s. The lower tie plate's tip node is fully constrained, and the normal translation DOFs of two nodes near the top for each of the four channel faces are also constrained. The QoI of interest is the frequency response function (FRF) of acceleration excited at a node on one face of the lower tie plate and observed at a node on the adjacent face, both in the normal direction. The excitation and observation points are located in the red region of the lower tie plate shown in Figure 15a. The QoI is the average value of the eight largest peaks in the frequency range  $\mathcal{B}$ . Accurate estimations can be obtained by using the Craig-Bampton sub-structuring method described in [15], where all the modes in the frequency band  $\mathcal{B}_c = 2\pi \times [0, 1200]$  rad/s are considered. The modes in  $2\pi \times [1000, 1200]$  rad/s are included to consider the spread contribution to the response within 1000 Hz. To reduce the computational cost, the shift-invert Lanczos method is used to solve the eigenvalue problem [15, 13]. The model can be solved by 50 computer nodes (16-core Intel(R) Xeon(R) CPU E5-2640 v3 @ 2.60 GHz) in about 5 minutes. The connections between different structural levels, rod-to-grid, and grid-channel, are critical to computing the QoI and are intrinsically uncertain due to the modeling process. Therefore, these two types of connections are randomized by Beta distributions with mean values of 0.15 kN/mm and 1.5 kN/mm, respectively, and coefficients of variation of 20%. The connections associated with a spacer grid share the same random germ. Therefore, the problem has eight random variables. Choosing the parameters to be their mean values and solving the dynamic problem yields the FRF, as shown in Figure 16a. In the figure the peaks of the FRF are marked by blue circles.

A much simpler LF model has also been constructed, which is a  $10 \times 10$  array of fuel rods held by eight spacer grids, an



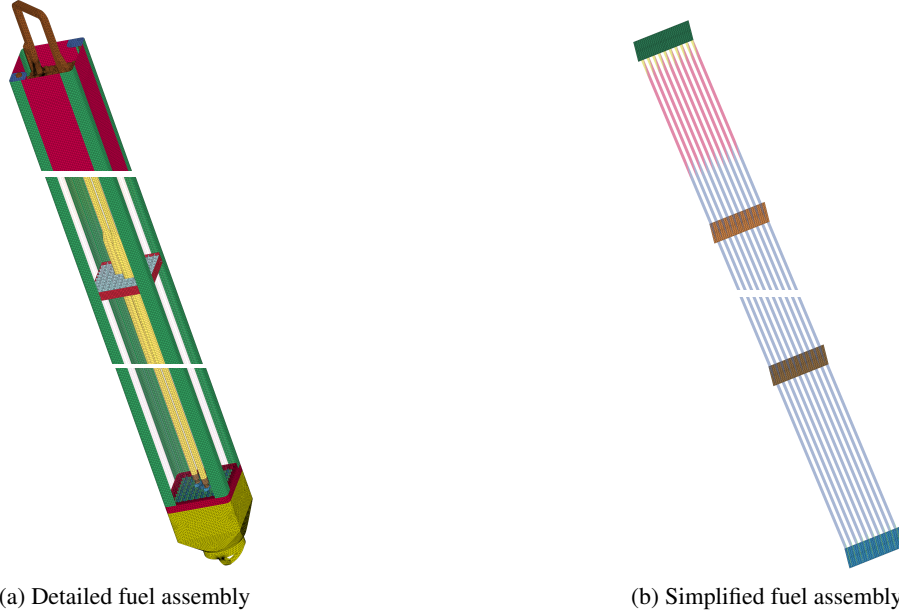


Figure 15: Finite element model of the detailed (a) / simplified (b) fuel assembly: view of the top (top), the middle span (middle), and the bottom (bottom) without the channel thin panel.

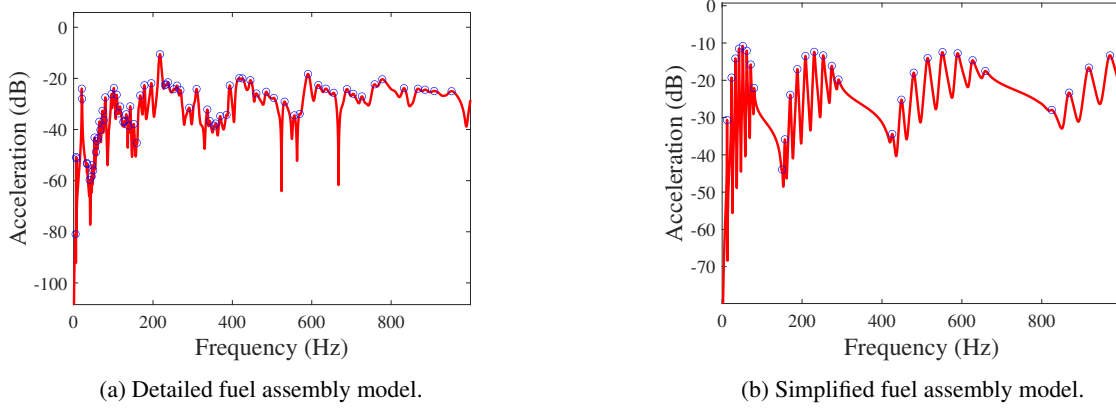


Figure 16: Frequency response function on mean values for the detailed fuel assembly model (a) and simplified (b) model.

upper tie plate, and a lower tie plate. The upper and lower tie plates have been simplified as transverse shell elements, and the spacer grids have also been modeled using transverse shell elements, as shown in the FE model in Figure 15b. The simplified model does not include any water rods or channel structure, resulting in  $N = 368,892$  DOFs. The upper and lower tie plates are constrained to simulate the same boundary conditions as the HF model. In this model, the FRF of acceleration is excited at a node on one side of the first spacer grid from the bottom and observed at a node on the adjacent side, both along the normal direction. Although the computational cost of the LF model by mode superposition has already been significantly reduced from the HF model, we further reduce the cost by applying the global reduced-order model (ROM) introduced in [14]. The cost to obtain the FRF is about 3 minutes using one computer node. Therefore, the HF and LF models have a cost ratio of about 80.

The simplified fuel assembly model employs shared nodes to connect the fuel rods to the spacer grids. In addition, we attempt to introduce randomness by modeling the Young's modulus of the spacer grids as Beta distributions with a coefficient of variation of 20% and a mean value of  $70.0 \text{ kN/mm}^2$ . Each spacer grid contains one random variable, resulting in a UQ dimension of 8. By setting the parameters to their mean values and solving the dynamic problem using global ROM, we obtain the FRF, which is shown in Figure 16b. The FRF of the simplified fuel assembly differs significantly from that of the detailed fuel assembly in terms of the shape and magnitude of the response. This is not unexpected, given that the simplified version is much simpler in several aspects and that the excitation and observation locations differ from those of the detailed model. The differences in modeling result in significant differences in the models' parameterizations.

We will now use the MFAB method to estimate the expected value of the QoI and compare it with the MC and MF methods.

Since the HF model is expensive, it is unrealistic to obtain repeated realizations of the estimators and then compare them. Therefore, we will use the unbiased approach with the *legacy dataset* options. For this application, we will generate  $N_p = 200$  pilot samples of the HF model, which will be reused to obtain different estimators for comparison. Although the HF samples are fixed, we can generate as many LF samples. The squared correlation of the HF and LF models on the original coordinates for the pilot samples is 0.068, as shown in Figure 17a. The two models are almost uncorrelated. Thus, the gain of leveraging

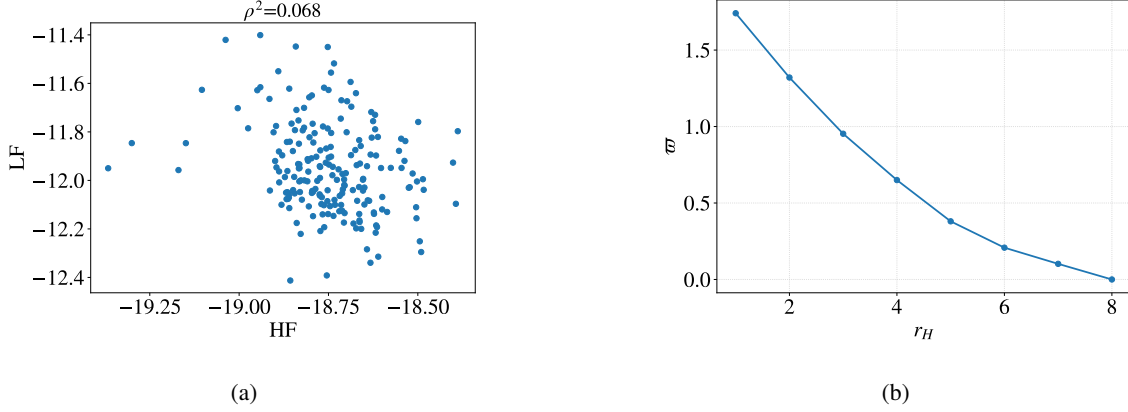


Figure 17: (a) Scatter plot of the HF and LF models on the original coordinates, and (b) bias estimator based on first order information for the fuel assembly application ( $\chi = 2$ ).

LF models is limited when using the MF estimator, which samples the models in their original space. Following the first two steps of Algorithm 5, we can compute the bias estimators of the HF model. Here, we will use the estimator based on first-order information for presentation, as shown in Figure 17b. There is no clear convergence trend, implying that all adapted directions are important and cannot be reduced, making correlation enhancement more challenging. We highlight that this feature reflects that the problem response cannot be satisfactorily represented on a lower dimensional manifold. However, this is an intrinsic property of this application and does not depend on the proposed strategy.

Next, we need to estimate the correlations for the unbiased estimator using Eq.(48). The results for various dimensions are shown in Figure 18a. When  $r_H = r_L = 2$ , we observe that the squared correlation is 0.423, which is the highest among all

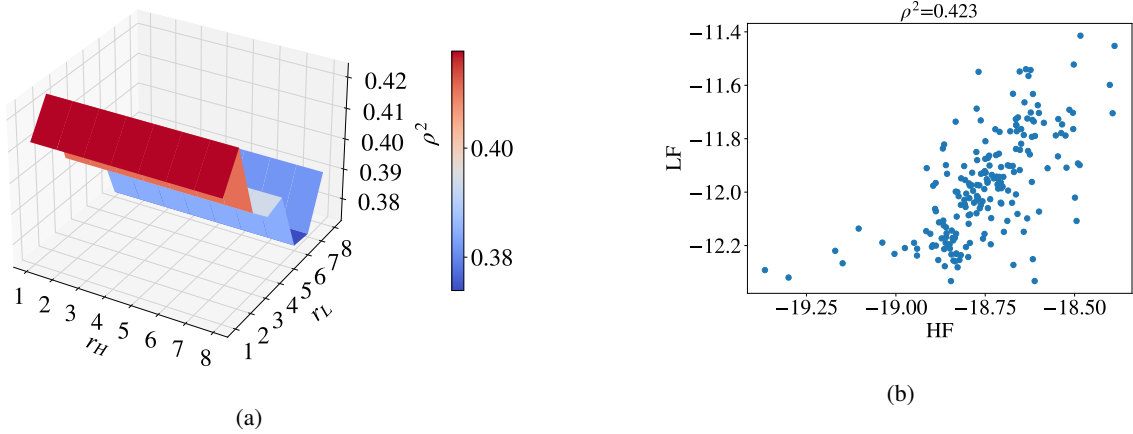


Figure 18: (a) Estimated correlation for the two models in the fuel assembly application with the unbiased strategy, and (b) scatter plot of the pilot HF and LF samples after rerun the LF model in the shared space.

dimensions. Although the value is lower than the previous two examples, it still represents a significant increase from 0.068. Subsequently, with  $r_H = r_L = 2$ , we project the pilot HF samples onto the 2-dimensional adapted space, map these samples to the original coordinates of the LF model, and reevaluate the LF model on these samples. The scatter plot of the pilot HF samples and the new pilot LF samples is presented in Figure 18b. Once again, we observe an increase in the correlation between HF and LF.

Following Algorithm 3, we estimate the oversampling ratio  $v$  of the LF model to be  $v = 7.66$ , and compute the number of

independent LF samples as  $\lceil (v - 1)N_p \rceil = 1332$ . Once the unbiased sampling strategy generates these 1332 independent LF samples, we use them to compute the MFAB estimator and its variance. Unlike the previous two applications, where multiple estimators could be computed, in this application, only one estimator is computed using the *legacy dataset* option. To provide a graphical representation of the estimators' performance, we plot a normal distribution with the estimator mean and variance in Figure 19. We also compare the MFAB estimator with the MC estimator, the optimal MF estimator, and the MF estimator with the same HF and LF samples as the MFAB estimator. For the MC estimator, we first use the 200 pilot HF samples to estimate the mean and variance, and then scale the variance to match the cost of the MFAB estimator. The optimal MF estimator is obtained using the pilot samples sampled in the original coordinates. Due to the low correlation in the original space, the oversampling ratio is calculated as  $v = 2.42$ , which is much smaller than that of the MFAB estimator. Figure 19 presents a comparison of these estimators. The variances of the two MF estimators are close to that of the MC estimator, while the

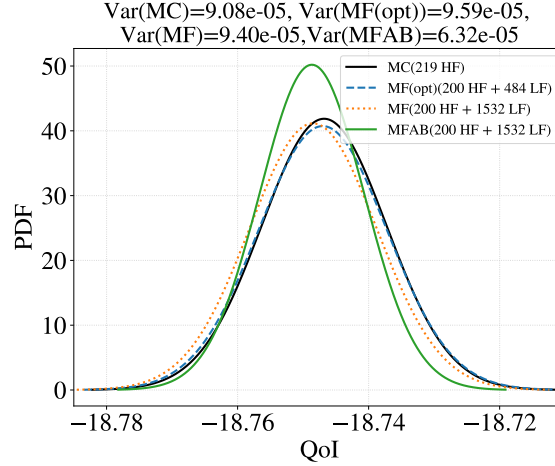


Figure 19: Scatter plot of the pilot HF and LF samples after rerun the LF model for the fuel assembly application.

variance of the MFAB estimator is 30% lower than that of the MC estimator. Comparing the two MF estimators, we observe that increasing the number of LF samples barely changes the MF estimator's performance since the correlation between the models in the original spaces is negligible.

The fuel assembly application presents significant challenges because the LF model is simplified to such an extent that its physics differs dramatically from the HF model. Additionally, as shown in Figure 17b, the HF model cannot be accurately represented on a lower-dimensional space, which further complicates the challenge of correlation enhancement. Generally, as the number of required dimensions for accurate representation in the adapted space increases, the correlation enhancement decreases. However, the MFAB approach can still increase the correlation between the models, thereby improving the performance of the MF estimator.

## 7 Concluding remarks

Multifidelity uncertainty quantification (MF UQ) approaches play a crucial role in deploying UQ analyses for realistic scientific and engineering problems. However, the effectiveness of the various methods developed in the literature is often greatly hindered by the presence of models with dissimilar parameterization. For example, a subset of the parameters may not be shared among models, which reduces the correlation between models by introducing an independent source of variability. Therefore, the dissimilar parameterization poses a significant challenge to MF UQ approaches since their performance is highly dependent on model correlations.

In this work, we addressed this challenge by relying on dimension reduction strategies to identify a shared set of parameters that does not correspond to the original model parameterization but can be obtained from it. We extended the previous work [17, 18] by integrating the Adaptive Basis (AB) [64, 69] dimension reduction strategy with a MF estimator corresponding to the MF Monte Carlo estimator with a single low-fidelity model. AB allows for easy and efficient identification of lower-dimensional manifolds in the parameter space that can be used to sample the models and improve their correlations. We presented a general discussion about the design of such estimators and introduced several strategies to support the practical construction of these estimators from available pilot samples. As a result, we proposed two novel strategies to embed AB into MF estimators. One strategy leads to bias-variance balanced MF estimators, enabling the flexibility of controlling bias and variance. In contrast,

the other strategy leads to unbiased estimators with variance control. Thanks to AB, both strategies have enhanced correlation among the models and consequently improve the performance and applicability of the MFUQ approaches.

Note that this paper primarily focuses on the use of a single LF model to demonstrate the effectiveness of the AB method in improving multi-fidelity estimators when dealing with models that have dissimilar parameterizations. Extensions to multiple LF models are possible, but they are not currently within the scope of our ongoing work. However, we want to emphasize that when truncating multiple LF models ( $r_L$ ), the selection must consider the impact of all rotated LF models on the performance of the MF estimator. The optimal approach may vary for different ACV estimators, such as MLMC and MFMC. In future work, we plan to leverage advancements in multi-fidelity model tuning [4, 19, 1, 5, 63], which involve selecting the numerical discretization of LF models (with the same parameters) to reduce the error of the MF estimator within a given budget. This will enable us to develop algorithms that optimally determine  $r_L$  for multiple LF models.

Finally, we presented an array of numerical test cases to illustrate the features of the methods in different practical scenarios, such as in the presence of legacy and non-legacy high-fidelity datasets. For all numerical tests presented, the novel estimator was able to make use of low-fidelity models that would otherwise be too poorly correlated to be exploited within existing MF methods. Our current work focuses on integrating this strategy with multiple low-fidelity models based on the Approximate Control Variate [27] method. Moreover, we are focusing on using this approach in the context of MF surrogate construction, such as MF networks [28]. Preliminary results for this latter case are presented in [66].

## Acknowledgements

The authors were partially supported by the Laboratory Directed Research Development (LDRD) program at Sandia National Laboratories and DOE SciDAC FASTMath institute. Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC (NTESS), a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy’s National Nuclear Security Administration (DOE/NNSA) under contract DE-NA0003525. This written work is authored by an employee of NTESS. The employee, not NTESS, owns the right, title and interest in and to the written work and is responsible for its contents. Any subjective views or opinions that might be expressed in the written work do not necessarily represent the views of the U.S. Government. The publisher acknowledges that the U.S. Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this written work or allow others to do so, for U.S. Government purposes. The DOE will provide public access to results of federally sponsored research in accordance with the DOE Public Access Plan.

## References

- [1] Brian M. Adams, Michael S. Eldred, Gianluca Geraci, Teresa Portone, Elliott M. Ridgway, John Adam Stephens, and Timothy M. Wildey. Deployment of multifidelity uncertainty quantification for thermal battery assessment part i: Algorithms and single cell results. 9 2022.
- [2] John D. Anderson. *Modern Compressible Flow: With Historical Perspective*. McGraw-Hill, New York, 2nd edition, 1990.
- [3] Géraud Blatman and Bruno Sudret. Adaptive sparse polynomial chaos expansion based on least angle regression. *Journal of Computational Physics*, 230(6):2345–2367, 2011.
- [4] Patrick J. Blonigan, Gianluca Geraci, Francesco Rizzi, and Michael S. Eldred. Towards an integrated and efficient framework for leveraging reduced order models for multifidelity uncertainty quantification. In *AIAA Scitech 2020 Forum*, 2020.
- [5] Geoffrey Bomarito, Gianluca Geraci, James Warner, Patrick Leser, William Leser, Michael S. Eldred, John Jakeman, and Alex Gorodetsky. Improving multi-model trajectory simulation estimators using model selection and tuning. In *AIAA SCITECH 2022 Forum*, 2022. AIAA 2022-1099.
- [6] Jianbing Chen, Xiaoshu Zeng, and Yongbo Peng. Probabilistic analysis of wind-induced vibration mitigation of structures by fluid viscous dampers. *Journal of Sound and Vibration*, 409:287–305, 2017.
- [7] K. A. Cliffe, M. B. Giles, R. Scheichl, and A. L. Teckentrup. Multilevel monte carlo methods and applications to elliptic pdes with random coefficients. *Computing and Visualization in Science*, 14(1):3, 2011.
- [8] Albert Cohen and Giovanni Migliorati. Optimal weighted least-squares methods. *The SMAI journal of computational mathematics*, 3:181–203, 2017.
- [9] Paul G. Constantine. *Active Subspaces*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2015.

- [10] Paul G Constantine, Michael Emory, Johan Larsson, and Gianluca Iaccarino. Exploiting active subspaces to quantify uncertainty in the numerical simulation of the hyshot ii scramjet. *Journal of Computational Physics*, 302:1–20, 2015.
- [11] Thierry Crestaux, Olivier Le Maitre, and Jean-Marc Martinez. Polynomial chaos expansion for sensitivity analysis. *Reliability Engineering & System Safety*, 94(7):1161 – 1172, 2009.
- [12] Michael S Eldred and Howard C Elman. Design under uncertainty employing stochastic expansion methods. *International Journal for Uncertainty Quantification*, 1(2), 2011.
- [13] Olivier Ezvan, X Zeng, R Ghanem, and B Gencturk. Dominant substructural vibration modes for fully-loaded spent nuclear fuel canisters. *Computational Mechanics*, 67(1):365–384, 2021.
- [14] Olivier Ezvan, Xiaoshu Zeng, Roger Ghanem, and Bora Gencturk. Dominant vibration modes for broadband frequency analysis of multiscale structures with numerous local vibration modes. *International Journal for Numerical Methods in Engineering*, 117(6):644–692, 2019.
- [15] Olivier Ezvan, Xiaoshu Zeng, Roger Ghanem, and Bora Gencturk. Multiscale modal analysis of fully-loaded spent nuclear fuel canisters. *Computer Methods in Applied Mechanics and Engineering*, 367:113072, 2020.
- [16] H.R. Fairbanks, A. Doostan, C. Ketelsen, and G. Iaccarino. A low-rank control variate for multilevel monte carlo simulation of high-dimensional uncertain systems. *Journal of Computational Physics*, 341:121–139, 2017.
- [17] G. Geraci and M.S. Eldred. Leveraging intrinsic principal directions for multifidelity uncertainty quantification. *SAND2018-10817*, 2018.
- [18] G. Geraci, M.S. Eldred, A.A. Gorodetsky, and J.D. Jakeman. Leveraging active directions for efficient multifidelity uncertainty quantification. In *6th European Conference on Computational Mechanics (ECCM 6)*, pages 2735–2746, 2018.
- [19] Gianluca Geraci, Jonathan Crussell, Laura P. Swiler, and Bert J. Deusschere. Exploration of multifidelity uq sampling strategies for computer network applications. *International Journal for Uncertainty Quantification*, 11(1):93–118, 2021.
- [20] Gianluca Geraci, Michael S Eldred, Alex Gorodetsky, and John Jakeman. Recent advancements in multilevel-multifidelity techniques for forward uq in the darpa sequoia project. In *AIAA Scitech 2019 Forum*, page 0722, 2019.
- [21] Gianluca Geraci, Michael S. Eldred, and Gianluca Iaccarino. A multifidelity multilevel Monte Carlo method for uncertainty propagation in aerospace applications. In *19th AIAA Non-Deterministic Approaches Conference*. AIAA, jan 2017.
- [22] T. Gerstner and M. Griebel. Numerical integration using sparse grids. *Numerical Algorithms*, 18(3-4):209–232, 1998.
- [23] Roger G Ghanem and Pol D Spanos. *Stochastic finite elements: a spectral approach*. Courier Corporation, 2003.
- [24] Ziad G Ghauch, Venkat Aitharaju, William R Rodgers, Praveen Pasupuleti, Arnaud Dereims, and Roger G Ghanem. Integrated stochastic analysis of fiber composites manufacturing using adapted polynomial chaos expansions. *Composites Part A: Applied Science and Manufacturing*, 118:179–193, 2019.
- [25] Michael B Giles. Multilevel Monte Carlo path simulation. *Operations Research*, 56(3):607–617, 2008.
- [26] A.A. Gorodetsky, J.D. Jakeman, and G. Geraci. MFNets: data efficient all-at-once learning of multifidelity surrogates as directed networks of information sources. *Computational Mechanics*, (68):741–758, 2021.
- [27] Alex A. Gorodetsky, Gianluca Geraci, Michael S. Eldred, and J.D. Jakeman. A generalized approximate control variate framework for multifidelity uncertainty quantification. *Journal of Computational Physics*, 408:109257, 2020.
- [28] Alex A. Gorodetsky, J.D. Jakeman, Gianluca Geraci, and Michael S. Eldred. MFNets: Multi-fidelity data-driven networks for Bayesian learning and prediction. *International Journal for Uncertainty Quantification*, 10(6):595–622, 2020.
- [29] Loic Le Gratiet and Josselin Garnier. Recursive co-kriging model for design of experiments with multiple levels of fidelity. *International Journal for Uncertainty Quantification*, 4(5):365–386, 2014.
- [30] A. Haji-Ali, F. Nobile, L. Tamellini, and R. Tempone. Multi-index stochastic collocation for random pdes. *Computer Methods in Applied Mechanics and Engineering*, 306:95 – 122, 2016.

- [31] Abdul-Lateef Haji-Ali, Fabio Nobile, and Raúl Tempone. Multi-index Monte Carlo: when sparsity meets sampling. *Numerische Mathematik*, 132(4):767–806, Apr 2016.
- [32] Jerrad Hampton and Alireza Doostan. Coherence motivated sampling and convergence analysis of least squares polynomial chaos regression. *Computer Methods in Applied Mechanics and Engineering*, 290:73–97, 2015.
- [33] HITACHI. *GE14-BWR Nuclear Fuel*, 2020. <https://nuclear.gepower.com/fuel-a-plant/products/ge14>.
- [34] J. D. Jakeman. Pyapprox: Approximation and probabilistic analysis of data. <https://sandialabs.github.io/pyapprox/index.html>, 2021.
- [35] J.D. Jakeman, M.S. Eldred, G. Geraci, and A. Gorodetsky. Adaptive multi-index collocation for uncertainty quantification and sensitivity analysis. *International Journal for Numerical Methods in Engineering*, 2019.
- [36] J.D. Jakeman, M.S. Eldred, and K. Sargsyan. Enhancing  $\ell_1$ -minimization estimates of polynomial chaos expansions using basis selection. *Journal of Computational Physics*, 289(0):18 – 34, 2015.
- [37] J.D. Jakeman, Fabian Franzelin, Akil Narayan, Michael Eldred, and Dirk Plfuger. Polynomial chaos expansions for dependent random variables. *Computer Methods in Applied Mechanics and Engineering*, 351:643 – 666, 2019.
- [38] J.D. Jakeman and A. Narayan. Generation and application of multivariate polynomial quadrature rules. *Computer Methods in Applied Mechanics and Engineering*, 338:134 – 161, 2018.
- [39] Lluís Jofre, Zachary R. del Rosario, and Gianluca Iaccarino. Data-driven dimensional analysis of heat transfer in irradiated particle-laden turbulent flow. *International Journal of Multiphase Flow*, 125:103198, 2020.
- [40] Ian T Jolliffe and Jorge Cadima. Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065):20150202, 2016.
- [41] M. C Kennedy and A. O’Hagan. Predicting the output from a complex computer code when fast approximations are available. *Biometrika*, 87(1):1–13, 2000.
- [42] Vahid Keshavarzzadeh, Robert M. Kirby, and Akil Narayan. Numerical integration in multiple dimensions with designed quadrature. *SIAM Journal on Scientific Computing*, 40(4):A2033–A2061, 2018.
- [43] S. S. Lavenberg and P. D. Welch. A perspective on the use of control variables to increase the efficiency of monte carlo simulations. *Management Science*, 27(3):322–335, 1981.
- [44] Stephen S Lavenberg, Thomas L Moeller, and Peter D Welch. Statistical results on control variables with application to queueing network simulation. *Operations Research*, 30(1):182–202, 1982.
- [45] Olivier P Le Maître, Matthew T Reagan, Habib N Najm, Roger G Ghanem, and Omar M Knio. A stochastic projection method for fluid flow: Ii. random process. *Journal of computational Physics*, 181(1):9–44, 2002.
- [46] Guotu Li, Mohamed Iskandarani, Matthieu Le Hénaff, Justin Winokur, Olivier P Le Maître, and Omar M Knio. Quantifying initial and wind forcing uncertainties in the gulf of mexico. *Computational Geosciences*, 20(5):1133–1153, 2016.
- [47] Pei-Ling Liu and Armen Der Kiureghian. Multivariate distribution models with prescribed marginals and covariances. *Probabilistic Engineering Mechanics*, 1(2):105–112, 1986.
- [48] Habib N Najm. Uncertainty quantification and polynomial chaos techniques in computational fluid dynamics. *Annual review of fluid mechanics*, 41:35–52, 2009.
- [49] A. Narayan and J.D. Jakeman. Adaptive Leja sparse grid constructions for stochastic collocation and high-dimensional approximation. *SIAM Journal on Scientific Computing*, 36(6):A2952–A2983, 2014.
- [50] A. Narayan, J.D. Jakeman, and T. Zhou. A Christoffel function weighted least squares algorithm for collocation approximations. *Mathematics of Computation*, 86:1913–1947, 2017.
- [51] Leo WT Ng and Karen E Willcox. Multifidelity approaches for optimization under uncertainty. *International Journal for numerical methods in Engineering*, 100(10):746–772, 2014.
- [52] Fabio Nobile and Francesco Tesei. A multi level monte carlo method with control variate for elliptic pdes with log-normal coefficients. *Stochastic Partial Differential Equations: Analysis and Computations*, 3(3):398–444, Sep 2015.

- [53] A. O’Hagan. Bayes–hermite quadrature. *Journal of Statistical Planning and Inference*, 29(3):245–260, 1991.
- [54] Art B. Owen. *Monte Carlo theory, methods and examples*. 2013.
- [55] Benjamin Peherstorfer, Karen Willcox, and Max Gunzburger. Optimal model management for multifidelity Monte Carlo estimation. *SIAM Journal on Scientific Computing*, 38(5):A3163–A3194, 2016.
- [56] Benjamin Peherstorfer, Karen Willcox, and Max Gunzburger. Survey of multifidelity methods in uncertainty propagation, inference, and optimization. *SIAM Review*, 60(3):550–591, 2018.
- [57] Matthew T Reagan, Habib N Najm, Roger G Ghanem, and Omar M Knio. Uncertainty quantification in reacting-flow simulations through non-intrusive spectral projection. *Combustion and Flame*, 132(3):545–555, 2003.
- [58] Markus P. Rumpfkeil and Philip S. Beran. Multifidelity sparse polynomial chaos surrogate models applied to flutter databases. *AIAA Journal*, 58(3):1292–1303, 2020.
- [59] A. Saltelli and *et al.* *Global Sensitivity Analysis. The primer*. John Wiley and Sons, 2008.
- [60] Daniel Schaden and Elisabeth Ullmann. On multilevel best linear unbiased estimators. *SIAM/ASA Journal on Uncertainty Quantification*, 8(2):601–635, 2020.
- [61] E. C. Stasiunas, R. A. Schultz, and Ross M. R. Performing direct-field acoustic test environments on a Sandia flight system to provide data for finite element simulation. *Rotating Machinery, Hybrid Test Methods, Vibro-Acoustics & Laser Vibrometry*, 8:267–279, 2016.
- [62] A. Teckentrup, P. Jantsch, C. Webster, and M. Gunzburger. A multilevel stochastic collocation method for partial differential equations with random input data. *SIAM/ASA Journal on Uncertainty Quantification*, 3(1):1046–1074, 2015.
- [63] Marten Thompson, Gianluca Geraci, Geoffrey Bomarito, James Warner, Patrick Leser, William P. Leser, Michael S. Eldred, John Jakeman, and Alex Gorodetsky. Strategies for automation of model tuning in multi-fidelity trajectory uncertainty propagation. In *AIAA SCITECH 2023 Forum*, 2022. AIAA 2023-1481.
- [64] Ramakrishna Tipireddy and Roger Ghanem. Basis adaptation in homogeneous chaos spaces. *Journal of Computational Physics*, 259:304–317, 2014.
- [65] D. Xiu and G.E. Karniadakis. The Wiener-Askey Polynomial Chaos for stochastic differential equations. *SIAM J. Sci. Comput.*, 24(2):619–644, 2002.
- [66] Xiaoshu Zeng, Gianluca Geraci, Alex Gorodetsky, John Jakeman, Michael S Eldred, and Roger Ghanem. Improving bayesian networks multifidelity surrogate construction with basis adaptation. In *AIAA Scitech 2023 Forum*, pages AIAA 2023–0917, 2023.
- [67] Xiaoshu Zeng and Roger Ghanem. Projection pursuit adaptation on polynomial chaos expansions. *Computer Methods in Applied Mechanics and Engineering*, 405:115845, 2023.
- [68] Xiaoshu Zeng, Yongbo Peng, and Jianbing Chen. Serviceability-based damping optimization of randomly wind-excited high-rise buildings. *The structural design of tall and special buildings*, 26(11):e1371, 2017.
- [69] Xiaoshu Zeng, John Red-Horse, and Roger Ghanem. Accelerated basis adaptation in homogeneous chaos spaces. *Computer Methods in Applied Mechanics and Engineering*, 386:114109, 2021.

## Appendix A Proof of Proposition 4.1

*Proof.* Given the assumptions,  $r_s = \max(r_H, r_L) = r_H = r_L$ . It follows that

$$\boldsymbol{\eta}_s = \boldsymbol{\eta}_{H,r_H} = \boldsymbol{\eta}_{L,r_L}.$$

Then, from Eq (30),  $\boldsymbol{\xi}_L^{r_L} = \mathbf{A}_{L,r_L}^T \boldsymbol{\eta}_{L,r_L} = \mathbf{A}_{L,r_L}^T \boldsymbol{\eta}_s$ , implying  $\boldsymbol{\eta}_{H,r_H} = \boldsymbol{\eta}_s = \mathbf{A}_{L,r_L} \boldsymbol{\xi}_L^{r_L}$ . It follows that

$$\boldsymbol{\xi}_H^{r_H} = \mathbf{A}_{H,r_H}^T \boldsymbol{\eta}_{H,r_H} = \mathbf{A}_{H,r_H}^T \mathbf{A}_{L,r_L} \boldsymbol{\xi}_L^{r_L}.$$

Similarly, we can prove the second equation of (33). □

## Appendix B Proof of Proposition 4.2

*Proof.*

$$\begin{aligned}
\text{MSE}(\hat{Q}^{\text{MFAB}}) &= \mathbb{E} \left[ \left( \hat{Q}^{\text{MFAB}} - \mathbb{E}[Q_H] \right)^2 \right] \\
&= \mathbb{E} \left[ \left( \hat{Q}^{\text{MFAB}} - \mathbb{E}[\hat{Q}^{\text{MFAB}}] + \mathbb{E}[\hat{Q}^{\text{MFAB}}] - \mathbb{E}[Q_H] \right)^2 \right] \\
&= \mathbb{E} \left[ \left( \hat{Q}^{\text{MFAB}} - \mathbb{E}[\hat{Q}^{\text{MFAB}}] \right)^2 \right] + \left( \mathbb{E}[\hat{Q}^{\text{MFAB}}] - \mathbb{E}[Q_H] \right)^2 \\
&= \text{Var}[\hat{Q}^{\text{MFAB}}] + \left( \mathbb{E}[Q_H(\mathcal{A}_{H,r_H}^T \boldsymbol{\eta}_s)] - \mathbb{E}[Q_H(\boldsymbol{\xi}_H)] \right)^2.
\end{aligned}$$

And, by Eq (9)

$$\text{Var}[\hat{Q}^{\text{MFAB}}] = \text{Var}[\hat{Q}_H^{r_H}(r_H, \underline{\boldsymbol{\eta}}_s)] \left( 1 - \frac{v-1}{v} \rho_{r_L, r_H}^2 \right).$$

□

## Appendix C Proof of Proposition 4.3

*Proof.* The expected values for the two expansions are

$$\begin{aligned}
\mathbb{E}[Q_H(\boldsymbol{\eta}_{H,r_H})] &= \mathbb{E} \left[ \sum_{\gamma \in \mathcal{J}_{p_H}^{r_H}} c_{H,\gamma} \psi_\gamma(\boldsymbol{\eta}_{H,r_H}) \right] = c_{H,e_0} + \sum_{\gamma \in \mathcal{J}_{p_H}^{r_H} \setminus e_0} c_{H,\gamma} \mathbb{E}[\psi_\gamma(\boldsymbol{\eta}_{H,r_H})] = c_{H,e_0} \\
\mathbb{E}[Q_L(\boldsymbol{\eta}_{L,r_L})] &= \mathbb{E} \left[ \sum_{\gamma \in \mathcal{J}_{p_L}^{r_L}} c_{L,\gamma} \psi_\gamma(\boldsymbol{\eta}_{L,r_L}) \right] = c_{L,e_0} + \sum_{\gamma \in \mathcal{J}_{p_L}^{r_L} \setminus e_0} c_{L,\gamma} \mathbb{E}[\psi_\gamma(\boldsymbol{\eta}_{L,r_L})] = c_{L,e_0},
\end{aligned}$$

by properties of Hermite polynomial chaos. In the above equation,  $c_{H,e_0}$  and  $c_{L,e_0}$  are zero order coefficients of the two models. It follows that

$$\begin{aligned}
\text{Cov}[Q_H(\boldsymbol{\eta}_{H,r_H}), Q_L(\boldsymbol{\eta}_{L,r_L})] &= \mathbb{E}[(Q_H(\boldsymbol{\eta}_{H,r_H}) - \mathbb{E}[Q_H(\boldsymbol{\eta}_{H,r_H})])(Q_L(\boldsymbol{\eta}_{L,r_L}) - \mathbb{E}[Q_L(\boldsymbol{\eta}_{L,r_L})])] \\
&= \mathbb{E} \left[ \left( \sum_{\gamma \in \mathcal{J}_{p_H}^{r_H} \setminus e_0} c_{H,\gamma} \psi_\gamma(\boldsymbol{\eta}_{H,r_H}) \right) \left( \sum_{\gamma \in \mathcal{J}_{p_L}^{r_L} \setminus e_0} c_{L,\gamma} \psi_\gamma(\boldsymbol{\eta}_{L,r_L}) \right) \right].
\end{aligned}$$

Since  $\boldsymbol{\eta}_{H,r_H}$  and  $\boldsymbol{\eta}_{L,r_L}$  are standard Gaussian variables, they share the same probability measure. Then, by the orthonormality of Hermite polynomial chaos, the terms in the above equation are nonzero only when the multi-indices  $\gamma$  are identical for both expansions. Then, by noting  $\langle \psi_\gamma, \psi_\gamma \rangle = 1$ , it follows that

$$\text{Cov}[Q_H(\boldsymbol{\eta}_{H,r_H}), Q_L(\boldsymbol{\eta}_{L,r_L})] = \sum_{\gamma \in (\mathcal{J}_{p_H}^{r_H} \cap \mathcal{J}_{p_L}^{r_L}) \setminus e_0} c_{H,\gamma} c_{L,\gamma}.$$

By the orthonormality of Hermite polynomial chaos, we can also get

$$\begin{aligned}
\text{Var}[Q_H(\boldsymbol{\eta}_{H,r_H})] &= \mathbb{E}[(Q_H(\boldsymbol{\eta}_{H,r_H}) - \mathbb{E}[Q_H(\boldsymbol{\eta}_{H,r_H})])^2] = \mathbb{E} \left[ \left( \sum_{\gamma \in \mathcal{J}_{p_H}^{r_H} \setminus e_0} c_{H,\gamma} \psi_\gamma(\boldsymbol{\eta}_{H,r_H}) \right)^2 \right] = \sum_{\gamma \in \mathcal{J}_{p_H}^{r_H} \setminus e_0} c_{H,\gamma}^2 \\
\text{Var}[Q_L(\boldsymbol{\eta}_{L,r_L})] &= \mathbb{E}[(Q_L(\boldsymbol{\eta}_{L,r_L}) - \mathbb{E}[Q_L(\boldsymbol{\eta}_{L,r_L})])^2] = \mathbb{E} \left[ \left( \sum_{\gamma \in \mathcal{J}_{p_L}^{r_L} \setminus e_0} c_{L,\gamma} \psi_\gamma(\boldsymbol{\eta}_{L,r_L}) \right)^2 \right] = \sum_{\gamma \in \mathcal{J}_{p_L}^{r_L} \setminus e_0} c_{L,\gamma}^2
\end{aligned}$$

Then, by definition, the correlation of  $Q_H(\boldsymbol{\eta}_{H,r_H})$  and  $Q_L(\boldsymbol{\eta}_{L,r_L})$  has the expression of (43).

□



## Appendix D Proof of Corollary 4.1

*Proof.* By the definition of correlation we can write

$$\begin{aligned}\rho^2 &= \frac{(\text{Cov}[Q_H, Q_L])^2}{\text{Var}[Q_H] \text{Var}[Q_L]} = \frac{\left(\sum_{\gamma \in \mathcal{J}_{p_L}^{r_L} \setminus e_0} c_{H,\gamma} c_{L,\gamma} \mathbb{E}[\psi_\gamma^2]\right)^2}{\left(\sum_{\gamma \in \mathcal{J}_{p_H}^{r_H} \setminus e_0} c_{H,\gamma}^2 \mathbb{E}[\psi_\gamma^2]\right) \left(\sum_{\gamma \in \mathcal{J}_{p_L}^{r_L} \setminus e_0} c_{L,\gamma}^2 \mathbb{E}[\psi_\gamma^2]\right)} \\ &= \frac{\left(\sum_{\gamma \in \mathcal{J}_{p_L}^{r_L} \setminus e_0} c_{H,\gamma} c_{L,\gamma} \mathbb{E}[\psi_\gamma^2]\right)^2}{\left(\sum_{\gamma \in \mathcal{J}_{p_L}^{r_L} \setminus e_0} c_{H,\gamma}^2 \mathbb{E}[\psi_\gamma^2] + \sum_{\gamma \in \mathcal{J}_\Delta \setminus e_0} c_{H,\gamma}^2 \mathbb{E}[\psi_\gamma^2]\right) \left(\sum_{\gamma \in \mathcal{J}_{p_L}^{r_L} \setminus e_0} c_{L,\gamma}^2 \mathbb{E}[\psi_\gamma^2]\right)} \\ &\leq \frac{\left(\sum_{\gamma \in \mathcal{J}_{p_L}^{r_L} \setminus e_0} c_{H,\gamma} c_{L,\gamma} \mathbb{E}[\psi_\gamma^2]\right)^2}{\left(\sum_{\gamma \in \mathcal{J}_{p_L}^{r_L} \setminus e_0} c_{H,\gamma}^2 \mathbb{E}[\psi_\gamma^2]\right) \left(\sum_{\gamma \in \mathcal{J}_{p_L}^{r_L} \setminus e_0} c_{L,\gamma}^2 \mathbb{E}[\psi_\gamma^2]\right)} = \rho_{\mathcal{J}_{p_L}^{r_L}}^2\end{aligned}$$

where  $\mathcal{J}_\Delta = \mathcal{J}_{p_H}^{r_H} \setminus \mathcal{J}_{p_L}^{r_L}$ . The last inequality is obtained because all terms corresponding to multi-indices belonging to the set  $\mathcal{J}_\Delta$  are contributing to the HF variance only, thus decreasing the correlation squared.  $\square$

## Appendix E Proof of Proposition 4.4

*Proof.* For simplified notation, we write the two first-order PCEs in the adapted spaces as

$$Q_H(\boldsymbol{\eta}) = c_{H,0} + \sum_{k=1}^d c_{H,k} \eta_k, \quad Q_L(\boldsymbol{\eta}) = c_{L,0} + \sum_{k=1}^d c_{L,k} \eta_k.$$

By assumption, we have  $c_{H,k} \geq 0$  and  $c_{L,k} \geq 0$ , for  $k = 1, \dots, d$ , and

$$c_{H,1} \geq \dots \geq c_{H,d} \quad \text{and} \quad c_{L,1} \geq \dots \geq c_{L,d}.$$

The correlations in the adapted space and the original space can be written as

$$\begin{aligned}\rho(Q_H(\boldsymbol{\eta}), Q_L(\boldsymbol{\eta})) &= \frac{\sum_{k=1}^d c_{H,k} c_{L,k}}{\sqrt{(\sum_{k=1}^d c_{H,k}^2)(\sum_{k=1}^d c_{L,k}^2)}}, \\ \rho(Q_H(\boldsymbol{\xi}), Q_L(\boldsymbol{\xi})) &= \frac{\sum_{k=1}^d c_{H,\sigma(k)} c_{L,k}}{\sqrt{(\sum_{k=1}^d c_{H,k}^2)(\sum_{k=1}^d c_{L,k}^2)}},\end{aligned}$$

respectively, where  $\sigma_{(1)}, \dots, \sigma_{(d)}$  is a permutation such that  $c_{H,\sigma(k)}$  and  $c_{L,k}$  are PCE coefficients associated with a same basis in the original space (i.e.  $\boldsymbol{\xi}$  space). The rearrangement inequality in mathematics states that

$$x_n y_1 + \dots + x_1 y_n \leq x_{\sigma(1)} y_1 + \dots + x_{\sigma(n)} y_n \leq x_1 y_1 + \dots + x_n y_n$$

for all choices of real numbers

$$x_1 \leq \dots \leq x_n \quad \text{and} \quad y_1 \leq \dots \leq y_n$$

and any permutation

$$x_{\sigma(1)}, \dots, x_{\sigma(n)}$$

of  $x_1, \dots, x_n$ . Apply the rearrangement inequality to the correlations we immediately get

$$\rho(Q_H(\boldsymbol{\eta}), Q_L(\boldsymbol{\eta})) \geq \rho(Q_H(\boldsymbol{\xi}), Q_L(\boldsymbol{\xi})).$$

Since we also assume that the PCE coefficients are greater than or equal to 0,

$$\rho^2(Q_H(\boldsymbol{\eta}), Q_L(\boldsymbol{\eta})) \geq \rho^2(Q_H(\boldsymbol{\xi}), Q_L(\boldsymbol{\xi})).$$

$\square$