

ANALYSIS OF THE CHALLENGES IN DEVELOPING SAMPLE-BASED MULTI-FIDELITY ESTIMATORS FOR NON-DETERMINISTIC MODELS

Bryan W. Reuter, Gianluca Geraci, & Timothy Wildey*

Sandia National Laboratories, Albuquerque, NM, 87185

*Address all correspondence to: Bryan W. Reuter, E-mail: bwreute@sandia.gov

Original Manuscript Submitted: mm/dd/yyyy; Final Draft Received: mm/dd/yyyy

Multifidelity (MF) Uncertainty Quantification (UQ) seeks to leverage and fuse information from a collection of models to achieve greater statistical accuracy with respect to a single-fidelity counterpart, while maintaining an efficient use of computational resources. Despite many recent advancements in MF UQ, several challenges remain and these often limit its practical impact in certain application areas. In this manuscript, we focus on the challenges introduced by non-deterministic models to sampling MF UQ estimators. Non-deterministic models produce different responses for the same inputs, which means their outputs are effectively noisy. MF UQ is complicated by this noise since many state-of-the-art approaches rely on statistics, e.g., the correlation among models, to optimally fuse information and allocate computational resources. We demonstrate how the statistics of the quantities of interest, which impact the design, effectiveness, and use of existing MF UQ techniques, change as functions of the noise. With this in hand, we extend the unifying Approximate Control Variate framework to account for non-determinism, providing for the first time a rigorous means of comparing the effect of non-determinism on different multifidelity estimators and analyzing their performance with respect to one another. Numerical examples are presented throughout the manuscript to illustrate and discuss the consequences of the presented theoretical results.

KEY WORDS: *Uncertainty Quantification, Multifidelity, Non-deterministic/stochastic models, Monte Carlo*

1. INTRODUCTION

Uncertainty quantification (UQ) is widely regarded as a cornerstone of computational science. While tremendous progress has been made in developing new algorithms to perform UQ, the majority of this effort has focused on deterministic models where repeated model evaluations at the same input parameter location give the same response. In contrast, a non-deterministic model is one whose output response will vary even when evaluated for the same inputs due to an intrinsic and uncontrollable source of variability. Non-deterministic models can be broken into two classes based on the source of this intrinsic variability. The first class includes models where the outputs are corrupted by non-determinism directly – for example, emulotics for cybersecurity [1–3], flow through an uncertain porous medium [4,5], and stochastic modeling of material microstructures [6]. For such models, the intrinsic noise can be categorized as aleatoric. The second class includes models where the quantities of interest (QoIs) are necessarily statistics with respect to a non-deterministic response – for example, Monte Carlo methods for radiation transport [7] which average over a finite number of particle trajectories, particle-in-cell methods for plasmas [8], and turbulent flow simulations [9] which require ensemble- or time-averaging. For these models, the intrinsic noise can be categorized as epistemic since the noise can be reduced by increasing the computational effort. Formally, these scenarios can be treated equivalently if the QoIs are interpreted as noisy functions of the controllable inputs.

The accurate computation of integrals which depend on a high-fidelity computational model, as required by tasks such as uncertainty quantification, can easily become intractable due to the model's cost and the required number of

samples. One path to alleviating this issue is surrogate construction, which has recently been examined in a single-fidelity setting for UQ of a non-deterministic model [10–13]. Another path is sampling-based multifidelity (MF) approaches for uncertainty quantification, which are explored here. MF approaches aim to improve the reliability of statistics of high-fidelity computational models by leveraging information from different sources that vary in both accuracy and cost. There exists a growing body of literature focusing on MF techniques (see, e.g., [14–23]); however, at the time of writing this manuscript, little attention has been given to these methods in the context of non-deterministic models, which introduce challenges that impact the optimal performance of MF techniques. MF methods exploit correlations between models to estimate statistics of quantities of interest with respect to the uncertain input space in a more computationally efficient manner. For deterministic models, the resource allocation optimization problem in MF UQ determines how to distribute samples among a model set given a constraint on accuracy or cost of the estimator. This step may also include finding the optimal subset of models and the groups which share UQ samples (i.e., the optimal model graph) given some constraints on the class of estimator; see e.g., [24,25]. Both the correlations among quantities of interest and the relative computational costs among models are key factors in the design and performance of MF UQ estimators. In the context of non-deterministic models, these factors are directly impacted by the noise in the QoIs. The noise in each model manifests due to an independent source of variability, so the correlations among models' responses are always degraded. At least in principle the detrimental effect of noise on correlation can be mitigated by acquiring more samples for the same, fixed input; however, this directly increases the computational cost. Moreover, due to the noise the relationships among models cannot be determined *a priori* as done in the deterministic context. This is because the optimal model graph expressing the ordering and dependencies among models is directly affected by the noise level of each model's response. Without loss of generality, we simply refer to hierarchical and non-hierarchical estimators according to their need to order the models following a strict model sequence (former) or not (latter). The interplay among the noise levels, correlations, and computational cost directly impacts the relationships among models and represents the main challenge characterizing the use of MF UQ for non-deterministic models. These aspects have not been extensively studied in literature. Preliminary contributions of the authors [13,26–30] have explored some of these aspects either in the single fidelity case, to identify the corruption of statistics like variance by the models' noise, or in a simplified multifidelity setting in which only a single low-fidelity model is employed. In the present work the formulation is generalized to ensemble of models, i.e., more than one low-fidelity model, which allows for the study of how several existing estimators are impacted by non-determinism.

For all the reasons introduced above, the presence of non-determinism in the models' responses further emphasizes the need for flexible MF UQ estimators that are capable to encode non-hierarchical relationships among models. The Approximate Control Variate (ACV) framework [22] was introduced to support the design and analysis of non-hierarchical estimators while providing a unifying framework that encompasses other existing hierarchical estimators like multilevel Monte Carlo [14] and multifidelity Monte Carlo [18]. In this paper we adopt the ACV framework to analyze the effect of non-determinism on the performances of several MF UQ estimators. In particular, we note that ACV has been formulated and analyzed only for deterministic models and the role of non-determinism is studied and discussed here for the first time.

An ACV estimator is obtained by augmenting a single fidelity Monte Carlo estimator based on the high-fidelity model with low-fidelity control variates with estimated means. In practice, since we deal with low-fidelity models with unknown statistics, sampling strategies that allow to simultaneously estimate low- and high-fidelity statistics need to be employed. Herein, we focus on the sampling approach, named ACV-MF (introduced later in the text), which can be obtained from the same samples of MFMC. Indeed, the acronym ACV-MF has been introduced in [22] to underscore the similarity with MFMC. The choice to use ACV-MF in the present manuscript is dictated by the need to provide a straightforward comparison among estimators. We note that the theory we develop applies to any other flavor of ACV (see [22] for additional details) and, due to the relationship between ACV and other approaches (e.g., the so-called multilevel Best Linear Unbiased Estimators (ML BLUE) developed in [23,31,32]), we are putting forth a framework that will aid with the analysis of virtually any MF UQ approach based on control variates.

The novel contributions of this paper are to:

- introduce the relationships between noisy and non-noisy statistics that determine the design and performance of MF UQ estimators;

- extend the ACV framework [22] to the case of non-deterministic models by deriving how the effects of non-determinism propagate to estimator performance;
- leverage the introduced framework to analyze and compare instances of ACV estimators based on more than one low-fidelity model – namely multilevel Monte Carlo, multifidelity Monte Carlo, and ACV-MF;
- illustrate through numerical examples how several estimators perform against each other and with respect to single fidelity estimators as a function of the models' non-determinism.

The remainder of this paper is organized as follows. Necessary preliminary material on non-deterministic models, the forward UQ task, and key statistics for MF UQ are included in Section 2 along with relationships which show how the statistics degrade as a function of noise. The single-fidelity Monte Carlo estimator, which constitutes the building block of all sampling-based MF UQ approaches, is discussed in Section 3. Section 4 briefly introduces the ACV framework from [22] and extends it to the case of non-deterministic models. Moreover, three instances of ACV estimators (multilevel Monte Carlo, multifidelity Monte Carlo, and ACV-MF) are explicitly derived and the effect of noise is discussed. Numerical experiments shown in Section 5 compare various estimators and examine the effect of non-determinism on MF UQ in different application scenarios. Finally, Section 6 summarizes and highlights future work.

2. PRELIMINARIES

Consider the task of sample-based forward propagation of uncertainty through a model f with inputs x which, without loss of generality, are uncertain. The model output $y = f(x)$ is a random variable or possibly a vector of random variables, e.g., the vector of solution values over a mesh. We consider a generic functional of the solution as our quantities of interest, i.e., $Q(y(x))$, which we simply denote as $Q(x)$. The goal is computing statistics of $Q(x)$ over the uncertain input space x with a set of realizations $\{Q(x^{(i)})\}$. In this context, a *non-deterministic model* is one with an additional and uncontrollable source of variability such that for a given input sample $x^{(i)}$, the model output is still a random variable. The presence of non-determinism in the models makes forward propagation more challenging as its additional variability pollutes the statistics of the QoIs – that is, larger statistical errors are incurred compared to forward propagation through the same model without the additional variability.

2.1 Non-deterministic models and quantities of interest

Let $f_0 : X \times \Omega_0 \rightarrow \mathcal{Y}$ be the mapping from the joint input space $X \times \Omega_0$ to the output space \mathcal{Y} which defines a non-deterministic, high-fidelity computational model:

$$y_0 = f_0(x, \omega_0). \quad (1)$$

The vector of input parameters $x \in X$ are explicitly specified for each model evaluation and, therefore, are deemed *accessible* and *controllable*. In contrast, ω_0 is a vector of *inaccessible* and *uncontrollable* random variables which represent the intrinsic stochasticity of f_0 and cause the model output to vary for a fixed $x^{(i)}$. Formally, we associate x with the probability space $\mathcal{P}_X = \{X, \mathcal{F}_X, \mathbb{P}_X\}$ where \mathcal{F}_X is the event space and \mathbb{P}_X the probability measure. Similarly, $\mathcal{P}_{\Omega_0} = \{\Omega_0, \mathcal{F}_{\Omega_0}, \mathbb{P}_{\Omega_0}\}$ is the probability space for ω_0 . The relevant probability space for f_0 is associated with the product measure, e.g., the space $\mathcal{P}_{X \times \Omega_0} = \{X \times \Omega_0, \mathcal{F}_X \times \mathcal{F}_{\Omega_0}, \mathbb{P}_X \times \mathbb{P}_{\Omega_0}\}$.

For this work, we define our QoI as the conditional expectation

$$Q_0(x) = \mathbb{E}_{\Omega_0} [f_0(x, \omega_0)] \equiv \mathbb{E} [f_0(\bar{x}, \omega_0) | x = \bar{x}], \quad (2)$$

whose definition also establishes our notation for conditional statistics, e.g., $\mathbb{E}_A [f(a, b)]$ indicates marginalizing over A . $\mathbb{E}_{\Omega_0} [f_0]$ recovers the underlying deterministic response of f_0 with respect to x . Q_0 is usually estimated in a Monte

Carlo fashion from a finite number of *replicas*. The Monte Carlo approximation of $Q_0(x)$ is defined as

$$\mathbb{E}_{\Omega_0} [f_0(x, \omega_0)] \approx \tilde{Q}_0(x; N_{\omega_0}) \equiv \frac{1}{N_{\omega_0}} \sum_{j=1}^{N_{\omega_0}} f_0(x, \omega_0^{(j)}) \quad (3)$$

where N_{ω_0} replicas of f_0 , obtained by solving the model multiple times at the same location x , are used; see Figure 1 for an illustrative example. Due to the finite number of replicas that is possible to afford in a practical application,

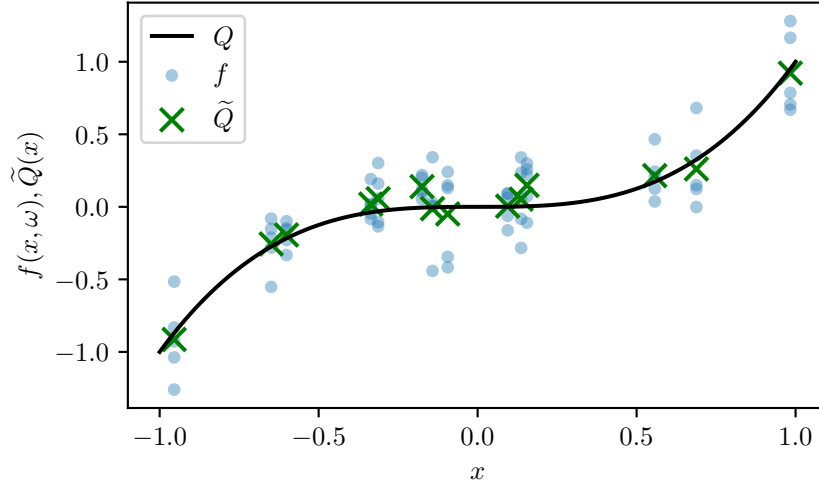


FIG. 1: Monte Carlo sampling of a non-deterministic model f with $N_{\omega} = 5$ replicas. True conditional expectation conditioned on x shown in black.

this procedure results in *noisy* observations of the QoI. In our context, *noisy* refers to variability incurred due to non-determinism, i.e., $|\tilde{Q}_0(x; N_{\omega_0}) - Q_0(x)| \sim \mathcal{N}(0, \sigma_{\omega_0}^2/N_{\omega_0})$ where $\sigma_{\omega_0}^2 \equiv \mathbb{V}_{\Omega_0} [f_0(x, \omega_0)]$ is the variance of the response for a fixed x . The *non-noisy* QoI Q_0 is recovered when $N_{\omega_0} \rightarrow \infty$. We use $\tilde{\cdot}$ to distinguish noisy quantities from their non-noisy counterparts, with a small number of exceptions, herein. To aid the reader, Table A.7 summarizes the nomenclature and key relationships between many of these quantities defined throughout this work.

As mentioned above, the focus of the UQ task is to estimate statistics over the uncertain parameters x of QoIs derived from the high-fidelity model. For deterministic models, this task can be performed more efficiently by leveraging multifidelity techniques. Our target will be computing the mean $\mathbb{E}_X [Q_0(x)]$ using multifidelity techniques in a non-deterministic setting. With this in mind, consider a set of L lower fidelity non-deterministic models f_1, \dots, f_L^* such that f_ℓ for $\ell = 1, \dots, L$ is a map from the input space $X \times \Omega_\ell$ to \mathcal{Y} with associated probability space $\mathcal{P}_{X \times \Omega_\ell}$, in direct analogy to the high-fidelity model. The use of Ω_ℓ implies the stochastic dependence need not be the same for each model. In any case, due to the uncontrollable nature of the ω_ℓ , realizations from two different models are assumed to be independent with respect to their stochastic variability. On the other hand, the accessible input space X is shared between all models (including the high-fidelity model) so they can be sampled at the same location $x \in X$. As in Equations (2) and (3), N_{ω_ℓ} replicas of f_ℓ generate noisy observations \tilde{Q}_ℓ of the non-noisy QoI $Q_\ell \equiv \mathbb{E}_{\Omega_\ell} [f_\ell]$. We have overloaded the expectation operator and will overload the variance operator with the understanding that the appropriate probability space can be inferred from their arguments – e.g., $\mathbb{V} [f_\ell]$ is an integral defined with respect to $\mathcal{P}_{X \times \Omega_\ell}$.

Let W_ℓ be the cost of a single evaluation of $f_\ell(x, \omega_\ell)$. We assume W_ℓ is independent of x and ω_ℓ , e.g., the cost of the evaluation $f_\ell(x^{(i)}, \omega_\ell^{(j)})$ is the same as the cost of the evaluation $f_\ell(x^{(k)}, \omega_\ell^{(m)})$ for $i \neq k, j \neq m$. Additionally,

*Note that no assumptions are made about the ordering of the models.

we assume that for a given model N_{ω_ℓ} is fixed, i.e., the same number of replicas are used for each sample $x^{(i)}$. Note that each model may use a different number of replicas, i.e., $N_{\omega_m} \neq N_{\omega_n}$ for $m \neq n$. Hence, the effective model cost, or the cost of obtaining one evaluation of \tilde{Q}_ℓ , is $\tilde{W}_\ell \equiv N_{\omega_\ell} W_\ell$ and the cost of N_ℓ QoI evaluations is $N_\ell \tilde{W}_\ell$. In other words, the computational cost is linear in both the number of replicas and the number of QoI evaluations. The results provided in this manuscript leverage this assumption; extensions to this cost model are provided in [33].

In this work, our goal is to study and discuss the effect of non-determinism on MF estimator variance. In order to reach this goal, the first step is to understand the relationship between statistics of the noisy QoIs and the corresponding statistics of the non-noisy QoIs. In the next subsection, we summarize a series of results presented in [26,28,29,34], extending them where appropriate for our multifidelity, non-deterministic context.

2.2 Relationships between noisy and non-noisy statistics

The following propositions highlight relationships between the noisy statistics, which determine multifidelity estimator design and performance, and their non-noisy counterparts. These are the variance of a noisy QoI and covariance and correlation among noisy QoIs.

Proposition 1. *The total variance of the noisy QoI \tilde{Q}_ℓ is denoted $\mathbb{V}[\tilde{Q}_\ell]$ and can be decomposed as*

$$\mathbb{V}[\tilde{Q}_\ell(x; N_{\omega_\ell})] = \mathbb{V}_X[Q_\ell(x)] + \frac{\mathbb{E}_X[\sigma_{\omega_\ell}^2]}{N_{\omega_\ell}} \quad (4)$$

where $\mathbb{V}_X[Q_\ell(x)]$ is the parametric variance and $\sigma_{\omega_\ell}^2 \equiv \mathbb{V}_{\Omega_\ell}[f_\ell]$ quantifies the level of noise from the stochastic response.

Proof. Equation (4) follows from the law of total variance (see APPENDIX B.1). \square

The parametric variance $\mathbb{V}_X[Q_\ell]$, or the variance with respect to the accessible parameters, is not directly observable. Instead, given N_ℓ independent, identically-distributed (i.i.d.) samples of x , $\mathbf{x} = \{x^{(i)}\}$, and N_{ω_ℓ} replicas for each sample, the evaluations of f_ℓ can be used to estimate

1. $\mathbb{E}_X[\sigma_{\omega_\ell}^2]$, by first estimating $\sigma_{\omega_\ell}^2(x^{(i)})$ with the N_{ω_ℓ} replicas generated at each location $x^{(i)}$ and then averaging w.r.t. x ;
2. $\mathbb{V}[\tilde{Q}_\ell]$, by taking the variance of the N_ℓ available samples of \tilde{Q}_ℓ at the $x^{(i)}$.

Then $\mathbb{V}_X[Q_\ell]$ can be computed from Equation (4). The procedure of breaking the total noisy variance into contributions from the parametric and stochastic variability has been introduced in several contributions [26,28,29,34] and denoted *variance deconvolution*.

Proposition 2. *The total covariance between two noisy QoIs \tilde{Q}_m and \tilde{Q}_n for $m \neq n$, denoted $\mathbb{C}[\tilde{Q}_m, \tilde{Q}_n]$, is equal to the non-noisy parametric covariance:*

$$\mathbb{C}[\tilde{Q}_m(x; N_{\omega_m}), \tilde{Q}_n(x; N_{\omega_n})] = \mathbb{C}_X[Q_m(x), Q_n(x)]. \quad (5)$$

Proof. By assumption, the stochastic responses of two different models are independent. Hence, any shared variability between models is due to the parametric response. See APPENDIX B.2 for details of the proof and the precise definition of the total covariance. \square

As discussed in Section 4, it is convenient to express the existing multifidelity methods in the ACV [22] formalism since many estimators can be derived directly as instances of it (or directly related to it, e.g., in the case of MLBLUE [23]). The main quantities needed to formulate an ACV estimator are the matrix of covariances among the low-fidelity QoIs and the vector of covariances between low- and high-fidelity QoIs. Let $\tilde{C} = \tilde{C}(N_{\omega}^{LF})$ be the $L \times L$

matrix of noisy covariances $\mathbb{C} [\tilde{Q}_m(x; N_{\omega_m}), \tilde{Q}_n(x; N_{\omega_n})] \equiv \tilde{C}_{m,n}$ between the low-fidelity models and let $\tilde{\mathbf{c}} = \tilde{\mathbf{c}}(\mathbf{N}_\omega) = [\tilde{C}_{0,1}, \dots, \tilde{C}_{0,L}]^T$ be the vector of noisy covariances between the low-fidelity models and the high-fidelity model. Since $\tilde{\mathbf{C}}$ and $\tilde{\mathbf{c}}$ represent extensions of the non-noisy quantities used in ACV, we derive the expression linking them in the following corollary; these expressions are obtained by identifying and separating the deterministic and non-deterministic contributions.

Corollary 1. $\tilde{\mathbf{C}} = \mathbf{C} + \delta\tilde{\mathbf{C}}$ where \mathbf{C} is the $L \times L$ matrix of non-noisy covariances $\mathbb{C}_X [Q_m, Q_n]$ between low-fidelity models and $\delta\tilde{\mathbf{C}}$ is a diagonal matrix depending only the noise and number of replicas:

$$\delta\tilde{C}_{mn} = \begin{cases} \frac{\mathbb{E}_X [\sigma_{\omega_m}^2]}{N_{\omega_m}} & \text{if } m = n \\ 0 & \text{else} \end{cases}. \quad (6)$$

Furthermore, $\tilde{\mathbf{c}} = \mathbf{c}$, where \mathbf{c} is the vector of non-noisy covariances between the low-fidelity models and the high-fidelity model.

A simple relationship between the correlation of two non-noisy QoIs and the correlation of the corresponding noisy QoIs follows from Propositions 1 and 2.

Corollary 2. Define the (Pearson's) correlation of two non-noisy QoIs as

$$\rho_{m,n} \equiv \frac{\mathbb{C}_X [Q_m(x), Q_n(x)]}{\sqrt{\mathbb{V}_X [Q_m(x)] \mathbb{V}_X [Q_n(x)]}} \quad (7)$$

and the correlation of two noisy QoIs as

$$\tilde{\rho}_{m,n}(N_{\omega_m}, N_{\omega_n}) \equiv \frac{\mathbb{C} [\tilde{Q}_m(x, N_{\omega_m}), \tilde{Q}_n(x, N_{\omega_n})]}{\sqrt{\mathbb{V} [\tilde{Q}_m(x; N_{\omega_m})] \mathbb{V} [\tilde{Q}_n(x; N_{\omega_n})]}}. \quad (8)$$

Let $\tau_\ell = \sqrt{\mathbb{V}_X [Q_\ell] / \mathbb{V}_X [Q_0]}$ and $p_\ell^2 = \mathbb{E}_X [\sigma_{\omega_\ell}^2] / \mathbb{V}_X [Q_0]$. Then

$$\tilde{\rho}_{m,n} = \rho_{m,n} \frac{\tau_m \tau_n}{\sqrt{\left(\tau_m^2 + \frac{p_m^2}{N_{\omega_m}}\right) \left(\tau_n^2 + \frac{p_n^2}{N_{\omega_n}}\right)}}.$$

It is easy to see the noisy statistics converge to their non-noisy counterparts in the limit of large numbers of replicas; this is illustrated in the following corollary.

Corollary 3. As $N_{\omega_\ell} \rightarrow \infty$, $\mathbb{V} [\tilde{Q}_\ell] \rightarrow \mathbb{V}_X [Q_\ell]$ and $\tilde{C}_{\ell\ell} \rightarrow C_{\ell\ell}$. Similarly, if both $N_{\omega_m}, N_{\omega_n} \rightarrow \infty$ then $\tilde{\rho}_{m,n} \rightarrow \rho_{m,n}$.

3. SINGLE-FIDELITY MONTE CARLO ESTIMATOR

All the MF UQ estimators based on linear control variates [15,35–38] are built from a single-fidelity Monte Carlo (SFMC) estimator based on high-fidelity model evaluations. Afterwards, unbiased terms based on low-fidelity model evaluations are added (and weighted) to reduce its variance while maintaining the bias of the high-fidelity model (see Section 4). In this section we are able to leverage the concepts introduced in Section 2 and establish additional definitions and notation for the SFMC. These developments will be fundamental for the formulation of multifidelity estimators, as discussed in Section 4, and will allow the comparison of MF UQ estimators with their single-fidelity counterpart, which should be considered the baseline in all sampling-based UQ approaches.

3.1 Definition and estimator variance

Consider the MC estimator for $\mathbb{E}_X [Q_0(x)]$ for which we use Equation (3) to write

$$\mathbb{E}_X [Q_0(x)] \approx \langle Q \rangle_0(\mathbf{x}; N_{\omega_0}) \equiv \frac{1}{N_0} \sum_{i=1}^{N_0} \tilde{Q}_0(x^{(i)}; N_{\omega_0}) = \frac{1}{N_0} \sum_{i=1}^{N_0} \left[\frac{1}{N_{\omega_0}} \sum_{j=1}^{N_{\omega_0}} f_0(x^{(i)}, \omega_0^{i,(j)}) \right] \quad (9)$$

where \mathbf{x} is a set of i.i.d. samples of x of size N_0 . For each sample $x^{(i)}$, f_0 is evaluated N_{ω_0} times, producing realizations $f_0(x^{(i)}, \omega_0^{i,(j)})$ for $j = 1, \dots, N_{\omega_0}$. The total cost of the estimator is $W^{MC} = N_0 N_{\omega_0} W = N_0 \tilde{W}_0$. Given the nested nature of the estimator and the linearity of the expectation operator, it is easy to see that $\langle Q \rangle_0$ is unbiased.

With N_0 samples of the noisy QoI, the estimator variance is $\mathbb{V}[\langle Q \rangle_0(\mathbf{x}; N_{\omega_0})] = N_0^{-1} \mathbb{V}[\tilde{Q}_0(x, N_{\omega_0})]$ which, using Equation (4), can be decomposed as

$$\mathbb{V}[\langle Q \rangle_0(\mathbf{x}; N_{\omega_0})] = \frac{1}{N_0} \left(\mathbb{V}_X [Q_0] + \frac{\mathbb{E}_X [\sigma_{\omega_0}^2]}{N_{\omega_0}} \right). \quad (10)$$

Akin to the discussion in Section 2.2, let

$$\hat{Q}_0(\mathbf{x}) = \lim_{N_{\omega_0} \rightarrow \infty} \langle Q \rangle_0(\mathbf{x}; N_{\omega_0}) \quad (11)$$

be the related non-noisy MC estimator which uses the same parametric samples \mathbf{x} but evaluates the QoI exactly and achieves the minimal possible variance, i.e., $\mathbb{V}[\hat{Q}_0] = N_0^{-1} \mathbb{V}_X [Q_0]$. Then Equation (10) can be written as

$$\mathbb{V}[\langle Q \rangle_0(\mathbf{x}; N_{\omega_0})] = \frac{\mathbb{V}_X [Q_0]}{N_0} \left(1 + \frac{p_0^2}{N_{\omega_0}} \right) = \mathbb{V}_X [\hat{Q}_0] \left(1 + \frac{p_0^2}{N_{\omega_0}} \right). \quad (12)$$

This type of splitting provides one useful way of highlighting the effect of non-determinism and will be extended in Section 4. The variance in Equation (10) expresses the variability of the mean estimator $\langle Q \rangle_0(\mathbf{x}; N_{\omega_0})$ that would be observed over independent realizations of x and ω_0 . It is important to note that the noise increases the variance of the non-deterministic SFMC estimator compared to its deterministic counterpart with the same N_0 . To maintain the same level of precision requires either a larger number of QoI evaluations N_0 or an increased number of replicas N_{ω_0} , which effectively reduces the noise introduced by non-determinism. Practically, with a constraint on the total computational budget, these two parameters are optimized to recover the smallest possible estimator variance.

3.2 Optimal estimator under a budget constraint

Under the assumption on the cost introduced in the previous section (the cost of evaluating a model is independent from the realization of x and ω), the SFMC estimator with the lowest variance for a given budget can be obtained with a single replica, (i.e., $N_{\omega_0} = 1$), as demonstrated in the following proposition; see also [13].

Proposition 3. *Assuming the model cost is independent of x and ω_0 (linear cost model) and a fixed computational budget W_{tot} , the optimal (lowest variance) single-fidelity Monte Carlo estimator is the one with $N_{\omega_0} = 1$.*

Proof. Given the budget W_{tot} , $N_{\omega_0} = W_{tot}/(N_0 W_0)$. Substituting into Equation (10) gives

$$\mathbb{V}[\langle Q \rangle_0] = \frac{\mathbb{V}_X [Q_0]}{N_0} + \frac{W_0}{W_{tot}} \mathbb{E}_X [\sigma_{\omega_0}^2].$$

The variance $\mathbb{V}[\langle Q \rangle_0]$ is minimized for the maximum number of evaluations $N_0 = W_{tot}/W$, which corresponds to $N_{\omega_0} = 1$. \square

Hence, Equation (10) shows that, in the single-fidelity context and under the linear cost assumption, the lowest variance is obtained with one replica per each sample $x^{(i)}$. Indeed, additional replicas ($N_{\omega_0} > 1$) only serve to reduce the noise, but additional samples for x (larger N_0) drive down the contribution of both terms. This optimal approach is equivalent to the so-called “direct” sampling strategy [13] explored in the context of single-fidelity surrogate construction (Polynomial Chaos) for radiation transport models. We note once again that more complex/realistic cost models would lead to different conclusions; see for instance [33,39].

4. MULTIFIDELITY ESTIMATORS

In this section, we introduce multifidelity estimators which aim at reducing estimator variance by leveraging realizations from cheaper low-fidelity models. Indeed, if high-fidelity evaluations are expensive, the single fidelity estimator introduced in the previous section, even with a single replica, could still produce a variance that is too large for practical purposes. A larger number of low-fidelity simulations can be usually obtained for a fraction of the high-fidelity model’s cost and, if these models are opportunistically correlated, multifidelity estimators can achieve a *variance reduction* with respect to SFMC. Unfortunately, non-deterministic models carry uncorrelated stochastic noise which diminishes their potential for variance reduction. In this section, the approximate control variate framework for multifidelity uncertainty quantification [22] is extended for non-deterministic models. As discussed in [22], several existing multifidelity approaches are instances of ACV estimators. The goal here is to incorporate the effect of noisy statistics (due to non-determinism) into ACV and then explore this effect on several instances of ACV estimators (multilevel Monte Carlo, multifidelity Monte Carlo, and ACV-MF).

4.1 The non-deterministic Approximate Control Variate framework

We first establish the notion of a non-deterministic ACV estimator and identify its variance, then we show how noise in the QoIs affects variance reduction. We note here that all the developments that follow, which involve the noisy statistics introduced in the previous sections, are novel since the original ACV framework [22] only dealt with deterministic models. Nevertheless, our presentation and notation is intentionally as close as possible to the original work for two reasons. First, the reader familiar with the original ACV can easily parse the role played by the models’ noise and recover the original ACV expressions, if no noise is present. Second, by explicitly writing the effect of the noise as additional terms or corrections with respect to the deterministic ACV baseline, we can easily propagate the effect of the noisy quantities into the estimator’s performance, e.g., from the degradation of the correlations to the variance reduction.

4.1.1 Definitions

Let $\langle Q \rangle_\ell(\mathbf{x}_\ell; N_{\omega_\ell})$ be the nested Monte Carlo estimator for model ℓ generated with a batch of samples \mathbf{x}_ℓ , directly akin to Equation (9). For each sample $x_\ell^{(i)}$, a vector of realizations of the inaccessible variables ω_ℓ^i of size N_{ω_ℓ} is generated during replication. An ACV estimator for non-deterministic model sets is an unbiased MC estimator of the form

$$\begin{aligned} \langle Q \rangle^{ACV}(\tilde{\boldsymbol{\alpha}}, \underline{\mathbf{x}}; \mathbf{N}_\omega) &= \langle Q \rangle_0(\mathbf{x}_0; N_{\omega_0}) + \sum_{\ell=1}^L \tilde{\alpha}_\ell (\langle Q \rangle_\ell(\mathbf{x}_\ell^1; N_{\omega_\ell}) - \langle Q \rangle_\ell(\mathbf{x}_\ell^2; N_{\omega_\ell})) \\ &= \langle Q \rangle_0(\mathbf{x}_0; N_{\omega_0}) + \tilde{\boldsymbol{\alpha}}^T \tilde{\boldsymbol{\Delta}}(\underline{\mathbf{x}}^{LF}; \mathbf{N}_\omega^{LF}), \end{aligned} \quad (13)$$

where $\tilde{\Delta}_\ell = \tilde{\Delta}_\ell(\mathbf{x}_\ell^1, \mathbf{x}_\ell^2; N_{\omega_\ell}) = \langle Q \rangle_\ell(\mathbf{x}_\ell^1; N_{\omega_\ell}) - \langle Q \rangle_\ell(\mathbf{x}_\ell^2; N_{\omega_\ell})$ for $\ell = 1, \dots, L$ are differences of estimators obtained with possibly overlapping sets \mathbf{x}_ℓ^1 and \mathbf{x}_ℓ^2 , while $\tilde{\boldsymbol{\alpha}}$ are a set of weights (which can be optimized). The number of replicas employed for each model are indicated with $\mathbf{N}_\omega = \{N_{\omega_0}, \dots, N_{\omega_L}\}$, whereas the ordered set of all the samples is denoted as $\underline{\mathbf{x}} = (\mathbf{x}_0, \mathbf{x}_1^1, \mathbf{x}_1^2, \dots, \mathbf{x}_L^1, \mathbf{x}_L^2)$. To simplify the following expressions, we also introduce the low-fidelity only quantities $\underline{\mathbf{x}}^{LF} = (\mathbf{x}_1^1, \mathbf{x}_1^2, \dots, \mathbf{x}_L^1, \mathbf{x}_L^2)$ and $\mathbf{N}_\omega^{LF} = \{N_{\omega_1}, \dots, N_{\omega_L}\} = \mathbf{N}_\omega \setminus \{N_{\omega_0}\}$. Letting N_0 be the size of \mathbf{x}_0 and N_ℓ be the size of $\mathbf{x}_\ell^1 \cup \mathbf{x}_\ell^2$ (i.e., the total number of noisy QoI evaluations of model ℓ), we

introduce the commonly employed sample ratio $\tilde{\mathbf{r}}$ where $\tilde{r}_\ell \equiv N_\ell/N_0$ for $\ell = 1, \dots, L$; see, e.g., [18]. We assume the QoIs are such that $|\tilde{\rho}_{0,\ell}| > |\tilde{\rho}_{0,\ell-1}|$, i.e., the first low-fidelity QoI has the largest noisy correlation with the HF QoI, and so on. While not necessary for a generic ACV estimator, this is easily done in practice, is often found in the MF literature, and has advantages when optimizing estimators under constraints (more detail in Section 4.4). Following [22] closely, the following lemmas express the ACV estimator's variance.

Lemma 1. (*Variance of the noisy ACV estimator*). *The variance of the noisy ACV estimator in Equation (13) for a vector of weights $\tilde{\boldsymbol{\alpha}} \in \mathbb{R}^L$ is*

$$\mathbb{V} \left[\langle Q \rangle^{ACV} (\tilde{\boldsymbol{\alpha}}; \mathbf{N}_\omega) \right] = \mathbb{V} \left[\langle Q \rangle_0 \right] \left(1 - \tilde{R}^2 \right), \quad (14)$$

$$\text{where } \tilde{R}^2 = \tilde{R}^2 (\tilde{\boldsymbol{\alpha}}; \mathbf{N}_\omega) = - \left(\tilde{\boldsymbol{\alpha}}^T \frac{\mathbb{C} \left[\tilde{\boldsymbol{\Delta}}, \tilde{\boldsymbol{\Delta}} \right]}{\mathbb{V} \left[\langle Q \rangle_0 \right]} \tilde{\boldsymbol{\alpha}} + 2 \tilde{\boldsymbol{\alpha}}^T \frac{\mathbb{C} \left[\tilde{\boldsymbol{\Delta}}, \langle Q \rangle_0 \right]}{\mathbb{V} \left[\langle Q \rangle_0 \right]} \right). \quad (15)$$

Proof. The proof is provided in [22], however, in the present manuscript, the notation is adapted to our context, i.e., $\tilde{\cdot}$ indicates the noisy quantities (see Section 2.2). \square

Lemma 2. (*Optimal noisy ACV weights*). *The noisy ACV variance can be minimized by using the weights $\tilde{\boldsymbol{\alpha}}^{opt}$ given by (see Proposition 2 and Eq. 11 in [22])*

$$\tilde{\boldsymbol{\alpha}}^{opt} = -\mathbb{C} \left[\tilde{\boldsymbol{\Delta}}, \tilde{\boldsymbol{\Delta}} \right]^{-1} \mathbb{C} \left[\tilde{\boldsymbol{\Delta}}, \langle Q \rangle_0 \right], \quad (16)$$

which lead to

$$\tilde{R}^2 (\tilde{\boldsymbol{\alpha}}^{opt}; \mathbf{N}_\omega) = \mathbb{C} \left[\tilde{\boldsymbol{\Delta}}, \langle Q \rangle_0 \right]^T \frac{\mathbb{C} \left[\tilde{\boldsymbol{\Delta}}, \tilde{\boldsymbol{\Delta}} \right]^{-1}}{\mathbb{V} \left[\langle Q \rangle_0 \right]} \mathbb{C} \left[\tilde{\boldsymbol{\Delta}}, \langle Q \rangle_0 \right]. \quad (17)$$

Proof. For the proof see Proposition 2 in [22]. \square

We have introduced the *noisy variance reduction term* \tilde{R}^2 , which represents the variance reduction achieved by the multifidelity strategy compared to its single-fidelity counterpart, specifically the one corresponding to the high-fidelity estimator with the same N_0 and N_ω . The *noisy variance reduction ratio*, defined as

$$\tilde{\gamma} \equiv \frac{\mathbb{V} \left[\langle Q \rangle^{ACV} \right]}{\mathbb{V} \left[\langle Q \rangle_0 \right]} = 1 - \tilde{R}^2, \quad (18)$$

can then be used to measure the effectiveness of the ACV estimator.

Finally, different instances of ACV estimators vary by the relationship between the different samples sets that comprise $\underline{\mathbf{x}}$. As stated in [22,25], the statistics that impact ACV estimator performance can be written explicitly in terms of the sampling strategy as

$$\mathbb{C} \left[\tilde{\boldsymbol{\Delta}}, \langle Q \rangle_0 \right] = \frac{1}{N_0} \text{diag} \left(\mathbf{F}(\tilde{\mathbf{r}}) \right) \circ \tilde{\mathbf{c}}(\mathbf{N}_\omega) \quad (19)$$

$$\text{and } \mathbb{C} \left[\tilde{\boldsymbol{\Delta}}, \tilde{\boldsymbol{\Delta}} \right] = \frac{1}{N_0} \mathbf{F}(\tilde{\mathbf{r}}) \circ \tilde{\mathbf{C}}(\mathbf{N}_\omega^{LF}), \quad (20)$$

extended here for non-determinism. We indicate with \circ the Hadamard (element-wise) product and $\mathbf{F}(\tilde{\mathbf{r}})$ is the matrix which depends only on $\tilde{\mathbf{r}}$ and encodes the aforementioned parametric sampling strategy (e.g., the overlap between the different sample sets in $\underline{\mathbf{x}}$).

4.1.2 How does noise impact ACV performance?

The previous lemmas have adapted the results of [22] to the non-deterministic model case. Going forward, we gain additional insights by separating and identifying the effect of noise on the estimator's performance. This is illustrated in the subsequent results, which culminate in Theorem 1.

First, we link the noisy and the non-noisy variance reduction in the following lemma.

Lemma 3. (*Relationship between the noisy \tilde{R}^2 and deterministic R^2 variance reduction*). Given the noisy variance reduction \tilde{R}^2 , introduced in Lemma 1, the deterministic variance reduction R^2 (see [22]) is obtained in the limit of infinite replicas from all models, i.e.,

$$R^2 = \lim_{N_{\omega} \rightarrow \infty} \tilde{R}^2, \quad (21)$$

where both

$$R^2 = \begin{cases} - \left(\boldsymbol{\alpha}^T \frac{\mathbb{C}_X [\boldsymbol{\Delta}, \boldsymbol{\Delta}]}{\mathbb{V}_X [\hat{Q}_0]} \boldsymbol{\alpha} + 2 \boldsymbol{\alpha}^T \frac{\mathbb{C}_X [\boldsymbol{\Delta}, \hat{Q}_0]}{\mathbb{V}_X [\hat{Q}_0]} \right) & \boldsymbol{\alpha} \text{ arbitrary} \\ \mathbb{C}_X [\boldsymbol{\Delta}, \hat{Q}_0]^T \frac{\mathbb{C}_X [\boldsymbol{\Delta}, \boldsymbol{\Delta}]^{-1}}{\mathbb{V}_X [\hat{Q}_0]} \mathbb{C}_X [\boldsymbol{\Delta}, \hat{Q}_0] & \boldsymbol{\alpha} = \boldsymbol{\alpha}^{opt}, \end{cases} \quad (22)$$

and $\boldsymbol{\alpha}^{opt} = -\mathbb{C}_X [\boldsymbol{\Delta}, \boldsymbol{\Delta}]^{-1} \mathbb{C}_X [\boldsymbol{\Delta}, \hat{Q}_0]$ are defined as in [22].

Proof. This result follows from the definition of \tilde{R}^2 provided in Equation (15) and its optimal value (Equation (17)), the use of Equation (11), and

$$\boldsymbol{\Delta} = \lim_{N_{\omega}^{LF} \rightarrow \infty} \tilde{\boldsymbol{\Delta}}, \quad (23)$$

where we employed

$$\hat{Q}_\ell(\mathbf{x}_\ell^i) = \lim_{N_{\omega}^{LF} \rightarrow \infty} \langle Q_\ell \rangle (\mathbf{x}_\ell^i; N_{\omega_\ell}), \quad \text{for } i = 1, 2. \quad (24)$$

□

In the next lemma, we recognize the high-fidelity model's noise affects \tilde{R}^2 only through the normalization term $\mathbb{V} [\langle Q \rangle_0]$.

Lemma 4. (*Influence of high-fidelity noise on \tilde{R}^2*). \tilde{R}^2 only depends on the high-fidelity model's non-determinism through the term $\mathbb{V} [\langle Q \rangle_0]$.

Proof. Starting with Equation (19) and employing Corollary 1, it immediately follows that

$$\mathbb{C} [\tilde{\boldsymbol{\Delta}}, \langle Q \rangle_0] = \mathbb{C}_X [\boldsymbol{\Delta}, \hat{Q}_0], \quad (25)$$

i.e., the deterministic term is recovered. Again using Corollary 1, Equation (20) is simplified:

$$\mathbb{C} [\tilde{\boldsymbol{\Delta}}, \tilde{\boldsymbol{\Delta}}] = \frac{1}{N_0} \mathbf{F}(\tilde{\mathbf{r}}) \circ \left(\mathbf{C} + \delta \tilde{\mathbf{C}}(N_{\omega}^{LF}) \right) = \mathbb{C}_X [\boldsymbol{\Delta}, \boldsymbol{\Delta}] + \frac{1}{N_0} \mathbf{F}(\tilde{\mathbf{r}}) \circ \delta \tilde{\mathbf{C}}(N_{\omega}^{LF}), \quad (26)$$

which shows $\mathbb{C} [\tilde{\boldsymbol{\Delta}}, \tilde{\boldsymbol{\Delta}}]$ only depends on the LF QoIs and their noise. Hence, the dependence of \tilde{R}^2 (Equations (15) and (17)) on HF non-determinism is only through the denominator $\mathbb{V} [\langle Q \rangle_0]$. □

Corollary 4. The product $\left(1 + \frac{p_0^2}{N_{\omega_0}} \right) \tilde{R}^2$ has no dependence on the high-fidelity noise or N_{ω_0} .

Proof. Since $\mathbb{V}[\langle Q \rangle_0]$ is the denominator of \tilde{R}^2 , the product $\mathbb{V}[\langle Q \rangle_0] \tilde{R}^2$ will be independent of the high-fidelity noise. From Equation (12), this product can be rewritten as $\mathbb{V}_X[\hat{Q}_0] (1 + p_0^2/N_{\omega_0}) \tilde{R}^2$. $\mathbb{V}_X[\hat{Q}_0]$ is a deterministic quantity; hence, $(1 + p_0^2/N_{\omega_0}) \tilde{R}^2$ will also be independent of the high-fidelity noise. \square

We are now ready to identify the models' noise effect (from all the high- and low-fidelity models) on the noisy ACV variance as a degradation with respect to the deterministic variance reduction R^2 . This is illustrated in the following theorem.

Theorem 1. (Relationship between the non-deterministic ACV variance and its deterministic counterpart). *The variance of the non-deterministic ACV estimator in Equation (13) for a vector of weights $\tilde{\alpha} \in \mathbb{R}^L$ can be expressed as*

$$\mathbb{V}[\langle Q \rangle^{ACV}] = \mathbb{V}_X[\hat{Q}_0] \left(1 - R^2 + \delta\tilde{R}_{HF}^2(N_{\omega_0}) + \delta\tilde{R}_{LF}^2(\mathbf{N}_{\omega}^{LF}) \right), \quad (27)$$

where the degradation in variance reduction introduced by the high- and low-fidelity models' noise is represented by $\delta\tilde{R}_{HF}^2(N_{\omega_0})$ and $\delta\tilde{R}_{LF}^2(\mathbf{N}_{\omega}^{LF})$, defined as

$$\delta\tilde{R}_{HF}^2(N_{\omega_0}) = \frac{p_0^2}{N_{\omega_0}} \quad (28)$$

$$\delta\tilde{R}_{LF}^2(\mathbf{N}_{\omega}^{LF}) = R^2 - \left(1 + \frac{p_0^2}{N_{\omega_0}} \right) \tilde{R}^2, \quad (29)$$

respectively.

Proof. This proof starts from the use of Equation (12) in the noisy ACV variance expression (Equation (14)), which leads to

$$\mathbb{V}[\langle Q \rangle^{ACV}] = \mathbb{V}_X[\hat{Q}_0] \left(1 + \frac{p_0^2}{N_{\omega_0}} - \left(1 + \frac{p_0^2}{N_{\omega_0}} \right) \tilde{R}^2 \right). \quad (30)$$

By definition, p_0^2/N_{ω_0} depends only on the non-determinism of the HF model, so this can be identified as $\delta\tilde{R}_{HF}^2$ as long as the product $(1 + p_0^2/N_{\omega_0}) \tilde{R}^2$ is independent of the non-determinism of the HF model. This is precisely what is demonstrated in Corollary 4. Hence, the result follows from addition and subtraction of R^2 . \square

Theorem 1 is central in understanding the role played by the models' noise. Indeed, it shows that the deterministic variance reduction is reduced (i.e., the effectiveness of the estimator with respect to the single-fidelity) by two independent contributions that separately depend on the high- and low-fidelity noise. As a consequence, it is possible to obtain an alternative expression for the noisy variance reduction, as shown in the following corollary.

Corollary 5. (Noisy variance reduction ratio). *From Equation (27), the noisy reduction ratio is expressed by*

$$\tilde{\gamma} = 1 - \frac{R^2 - \delta\tilde{R}_{LF}^2}{1 + \delta\tilde{R}_{HF}^2}. \quad (31)$$

From the previous corollary it follows that higher noise in the high-fidelity model inhibits the potential for variance reduction, diminishing the effectiveness of the low-fidelity models even if $\delta\tilde{R}_{LF}^2$ is small.

Finally, $\delta\tilde{R}_{LF}^2$ can be further rewritten to emphasize the effect of noisy statistics, as summarized in the following proposition.

Proposition 4. *The impact on variance reduction from the non-determinism of the low-fidelity models $\delta\tilde{R}_{LF}^2$ is*

$$\delta R_{LF}^2 = \begin{cases} \frac{1}{N_0 \mathbb{V}_X[\hat{Q}_0]} \boldsymbol{\alpha}^T [\mathbf{F} \circ \delta\tilde{\mathbf{C}}] \boldsymbol{\alpha} & \boldsymbol{\alpha} \text{ arbitrary} \\ \frac{1}{\mathbb{V}_X[\hat{Q}_0]} \mathbb{C}_X[\boldsymbol{\Delta}, \hat{Q}_0]^T \left(\mathbb{C}_X[\boldsymbol{\Delta}, \boldsymbol{\Delta}]^{-1} - \mathbb{C}[\tilde{\boldsymbol{\Delta}}, \tilde{\boldsymbol{\Delta}}]^{-1} \right) \mathbb{C}_X[\boldsymbol{\Delta}, \hat{Q}_0] & \boldsymbol{\alpha}^{opt}, \tilde{\boldsymbol{\alpha}}^{opt} \text{ employed.} \end{cases} \quad (32)$$

Proof. The result follows from Equations (25), (26) and (29), and the definitions of R^2 and \tilde{R}^2 . Details in APPENDIX B.3. \square

4.2 Examples of ACV estimators

In this section, we discuss three instances of ACV estimators, namely multilevel Monte Carlo (MLMC) [14], multifidelity Monte Carlo (MFMC) [18], and ACV-MF [22], in the light of the results introduced in the previous section.

4.2.1 Multilevel Monte Carlo

As shown in [22], this estimator can be obtained within the ACV framework (when the number of levels is prescribed):

$$\langle Q \rangle^{ML}(\boldsymbol{\alpha}, \boldsymbol{x}; \mathbf{N}_\omega) = \langle Q \rangle_0(\mathbf{x}_0; N_{\omega_0}) - \sum_{\ell=1}^L \left(\langle Q \rangle_\ell(\mathbf{x}_{\ell-1}^2; N_{\omega_\ell}) - \langle Q \rangle_\ell(\mathbf{x}_\ell^2; N_{\omega_\ell}) \right) \quad (33)$$

where $\mathbf{x}_{\ell-1}^2 \cap \mathbf{x}_\ell^2 = \emptyset$ and \mathbf{x}_ℓ^2 has size $N_{\ell,2}$. This means $\mathbf{x}_\ell^1 = \mathbf{x}_{\ell-1}^2$, $\mathbf{x}_0^2 = \mathbf{x}_0$ and $\alpha_\ell = -1 \forall \ell$ (see Equation (13)).

Proposition 5. Let $\eta_\ell = N_{\ell,2}/N_0$ and recall $\tau_\ell = \sqrt{\mathbb{V}_X[Q_\ell]/\mathbb{V}_X[Q_0]}$, $p_\ell^2 = \mathbb{E}_X[\sigma_{\omega_\ell}^2]/\mathbb{V}_X[Q_0]$. The variance of the MLMC estimator (Equation (33)) can be written in the form of Equation (27) where

$$R_{ML}^2 = -\tau_1^2 + 2\rho_{0,1}\tau_1 - \frac{\tau_L^2}{\eta_L} - \sum_{\ell=2}^L \frac{1}{\eta_{\ell-1}} (\tau_\ell^2 + \tau_{\ell-1}^2 - 2\rho_{\ell,\ell-1}\tau_\ell\tau_{\ell-1}) \quad (34)$$

$$\text{and } \delta \tilde{R}_{ML,LF}^2 = \sum_{\ell=1}^L \frac{\eta_\ell + \eta_{\ell-1}}{\eta_\ell \eta_{\ell-1}} \frac{p_\ell^2}{N_{\omega_\ell}}. \quad (35)$$

Proof. Included in APPENDIX B.4. The result follows Equation (32) and the specific partitioning for ML, \mathbf{F}_{ML} . \square

As expected, R_{ML}^2 is the deterministic result from Eq. (24) in [22] if $\alpha_\ell = -1 \forall \ell^\dagger$. The following proposition gives a bound on the achievable variance reduction ratio $\tilde{\gamma}$ for MLMC.

Proposition 6. In the limit of infinite evaluations of the low-fidelity QoIs \tilde{Q}_ℓ , $\ell \geq 1$, the MLMC estimator (Equation (33)) has a minimal attainable variance reduction ratio corresponding to

$$\tilde{\gamma}_{ML}^{min} = 1 - \frac{-\tau_1^2 + 2\rho_{0,1}\tau_1 - \frac{p_1^2}{N_{\omega_1}}}{1 + \frac{p_0^2}{N_{\omega_0}}}. \quad (36)$$

Proof. Using Equation (31), this follows since $\delta \tilde{R}_{ML,LF}^2 \rightarrow p_1^2/N_{\omega_1}$ and $R^2 \rightarrow -\tau_1^2 + 2\rho_{0,1}\tau_1$ as $N_\ell \rightarrow \infty$. \square

Note that $\mathbb{V}[\langle Q \rangle^{ML}]$ and, in turn, $\tilde{\gamma}_{ML}^{min}$ are suboptimal since the weights are fixed and, in general, differ from the optimal weights (Equation (16)).

Corollary 6. In the limit of infinite evaluations of the low-fidelity QoIs, it is possible to obtain a lower variance with the MLMC estimator than the single-fidelity MC, i.e., $\tilde{\gamma}_{ML}^{min} < 1$, when the low-fidelity noise, in the first model, satisfies $p_1^2/N_{\omega_1} < -\tau_1^2 + 2\rho_{0,1}\tau_1$.

The above proposition illustrates how the presence of noise can have a detrimental effect on the performance of a multilevel estimator, even in the limit of infinite LF QoI evaluations. MFMC and ACV-MF are partially able to counteract this negative impact of the noise, however, their performance is also impacted by the models' noise as illustrated in the next sections.

\dagger There are typos in Eq. 24 in [22]: τ_M should read τ_M^2 and the $\tau_{i-1}^2 \tau_i^2$ term should read $\alpha_{i-1}^2 \tau_{i-1}^2$.

4.2.2 Multifidelity Monte Carlo

As with MLMC, the multifidelity Monte Carlo strategy can be obtained as part of the ACV framework [22]. The multifidelity Monte Carlo estimator [18] partitions samples in the following manner: with $0 < N_0 \leq N_1 \leq \dots \leq N_L$, one set of i.i.d. samples of size N_L , \mathbf{x}_L , is drawn and the first N_ℓ are used to form the partition \mathbf{x}_ℓ^\ddagger . In the ACV framework, these choices correspond to $\mathbf{x}_\ell^1 = \mathbf{x}_{\ell-1}$ and $\mathbf{x}_\ell^2 = \mathbf{x}_\ell$, which allow to write Equation (13) as

$$\langle Q \rangle^{MF}(\boldsymbol{\alpha}, \underline{\mathbf{x}}; \mathbf{N}_\omega) = \langle Q \rangle_0(\mathbf{x}_0; N_{\omega_0}) + \sum_{\ell=1}^L \alpha_\ell \left(\langle Q \rangle_\ell(\mathbf{x}_{\ell-1}; N_{\omega_\ell}) - \langle Q \rangle_\ell(\mathbf{x}_\ell; N_{\omega_\ell}) \right). \quad (37)$$

Due to the absence of off-diagonal terms in $\mathbb{C}_X \left[\tilde{\boldsymbol{\Delta}}, \tilde{\boldsymbol{\Delta}} \right]^{-1}$, it is possible to obtain the MFMC optimal weights in closed-form as

$$\tilde{\alpha}_\ell = -\frac{\mathbb{C}_X [Q_0, Q_\ell]}{\mathbb{V} [\tilde{Q}_\ell]} = -\frac{\rho_{0,\ell} \tau_\ell}{\tau_\ell^2 + \frac{p_\ell^2}{N_{\omega_\ell}}}, \quad (38)$$

which results in the estimator variance outlined in the following proposition.

Proposition 7. *Using the standard convention $r_0 = 1$, the variance of the MFMC estimator (Equation (37)) can be written in the form of Equation (27) where*

$$R_{MF}^2 = \sum_{\ell=1}^L \frac{r_\ell - r_{\ell-1}}{r_\ell r_{\ell-1}} \rho_{0,\ell}^2 \quad (39)$$

$$\text{and } \delta \tilde{R}_{MF,LF}^2 = \sum_{\ell=1}^L \frac{r_\ell - r_{\ell-1}}{r_\ell r_{\ell-1}} \rho_{0,\ell}^2 \frac{\frac{p_\ell^2}{N_{\omega_\ell}}}{\tau_\ell^2 + \frac{p_\ell^2}{N_{\omega_\ell}}}. \quad (40)$$

Proof. Included in APPENDIX B.5. The result follows Equation (32) and the specific partitioning for MFMC, \mathbf{F}_{MF} . \square

As done for the MLMC estimator, in the following proposition, we quantify the effect of the noise on the achievable variance reduction ratio $\tilde{\gamma}$ of MFMC, under the assumption of infinite low-fidelity evaluations.

Proposition 8. *In the limit of infinite evaluations of the low-fidelity QoIs \tilde{Q}_ℓ , $\ell \geq 1$, the MFMC estimator (Equation (37)) has a minimal attainable variance reduction ratio corresponding to*

$$\tilde{\gamma}_{MF}^{min} = 1 - \rho_{0,1}^2 \frac{\tau_1^2}{\left(1 + \frac{p_0^2}{N_{\omega_0}}\right) \left(\tau_1^2 + \frac{p_1^2}{N_{\omega_1}}\right)} = 1 - \tilde{\rho}_{0,1}^2. \quad (41)$$

Proof. Using Equation (31), this follows since $\delta \tilde{R}_{MF,LF}^2 \rightarrow \rho_{0,1}^2 (p_1^2/N_{\omega_1}) / (\tau_1^2 + p_1^2/N_{\omega_1})$ and $R^2 \rightarrow \rho_{0,1}^2$ as $N_\ell \rightarrow \infty$. \square

This bound is the same as that achieved by the optimal control variate estimator which employs only the first ($\ell = 1$), most highly correlated low-fidelity model (OCV-1). The following corollary illustrates how MFMC is able to always achieve a variance reduction in the presence of noise if enough low-fidelity evaluations are available.

Corollary 7. *In the limit of infinite evaluations of the low-fidelity QoIs, the MFMC estimator, excluding the trivial case with no correlation among high- and low-fidelity models, always achieves a variance reduction, i.e., $\tilde{\gamma}_{MF}^{min} < 1$, since $\tilde{\rho}_{0,1}^2 > 0$.*

\ddagger Each subsequent partition is a superset of the previous.

The above results illustrate a key difference between MLMC and MFMC. While the MLMC variance reduction is strictly dependent on the low-fidelity models' noise, the MFMC estimator can always achieve a variance reduction if enough low-fidelity noisy evaluations are provided.

4.2.3 ACV-MF

In [22], a few *convergent* multifidelity estimators are introduced. These are ACV estimators which converge to the optimal control variate estimator (OCV) in the limit of infinite samples of the low-fidelity QoIs (see section 2.1 in [22]). An OCV has a minimal variance reduction ratio

$$\tilde{\gamma}_{OCV}^{min} = 1 - \frac{\bar{\mathbf{c}}^T \tilde{\mathbf{C}}^{-1} \bar{\mathbf{c}}}{1 + \frac{p_0^2}{N_{\omega_0}}} \quad (42)$$

where $\bar{\mathbf{c}} = \mathbf{c} / \sqrt{\mathbb{V}_X [Q_0]}$. This last expression, unlike its counterparts for MFMC and MLMC, includes statistics of the high-fidelity and all the low-fidelity models, rather than only the first low-fidelity model. One convergent estimator is ACV-MF which, as with MFMC, generates a set of samples size N_L (\mathbf{x}_L) and uses the first N_ℓ to form $\mathbf{x}_\ell^2 = \mathbf{x}_\ell$ but instead always selects $\mathbf{x}_\ell^1 = \mathbf{x}_0$ (rather than $\mathbf{x}_{\ell-1}$). The optimal weights (Equation (16)) are used. In general, $\tilde{\alpha}$, \tilde{R}^2 , etc. need be computed numerically since $\mathbb{C} [\tilde{\Delta}, \tilde{\Delta}]$ is a full matrix and analytic inversion is not possible.

Proposition 9. *If $\tilde{\mathbf{C}}$ is full rank, the ACV-MF estimator with infinite low-fidelity QoI evaluations always achieves variance reduction.*

Proof. This is a property of all convergent estimators. Since $\tilde{\mathbf{C}}$ is full rank, it is symmetric positive definite. Hence, $\bar{\mathbf{c}}^T \tilde{\mathbf{C}}^{-1} \bar{\mathbf{c}} > 0$ and $\tilde{\gamma}_{OCV}^{min} < 1$. \square

4.3 Example: comparing optimal variance reduction

In this section we illustrate via a numerical example the effects of non-deterministic models described in the previous sections. We augment the monomial example from Section 2.5 in [22], which computes the variance reduction ratio for a given five model set as a function of low-fidelity QoI evaluations for different ACV estimators, by introducing centered Gaussian noise with variance $\sigma_{\omega,\ell}^2$ to the models. In the notation adopted in this paper, we can define the ℓ th QoI[§] as

$$\tilde{Q}_\ell(x, \omega_\ell) = x^{5-\ell} + \frac{1}{N_{\omega_\ell}} \sum_{j=1}^{N_{\omega_\ell}} \mathcal{N}^j(0, \sigma_{\omega,\ell}^2) \quad (43)$$

where $x \sim \mathcal{U}(0, 1)$. Table 5 in [22] provides the correlation matrix and all necessary statistics follow from the models' definition in Equation (43).

We study the variance reduction ratio $\tilde{\gamma}$ for MLMC, MFMC, and ACV-MF in the two following scenarios:

1. Firstly, as a function of N_ℓ for $\ell \geq 1$, the number of evaluations of the low-fidelity QoIs \tilde{Q}_ℓ , with a given amount of effective noise p_ℓ^2/N_{ω_ℓ} for all models (including HF);
2. Then, in the limit of infinite low-fidelity sample evaluations (N_ℓ for $\ell \geq 1$) and infinite low-fidelity replicas (N_{ω_ℓ} for $\ell \geq 1$), such that we effectively sample the non-noisy low-fidelity QoIs Q_ℓ and extract their maximum potential for variance reduction while the high-fidelity model remains noisy.

Furthermore, it is important to also note that for this example we are interested in the variance reduction ratio without assuming a constraint on the computational budget. Cases with budget constraints are discussed in Sections 4.4 and 5.

[§]Before reordering, if necessary.

With a sample ratio $\tilde{r}_\ell(k) = 2^{\ell+k}$, Figure 2 shows $\tilde{\gamma}$ as a function of low-fidelity QoI evaluations under different stochastic noise scenarios. For simplicity, p_ℓ^2/N_{ω_ℓ} (which measures the effective amount of stochastic noise with respect to the high-fidelity parametric variance) is taken to be the same for all models. We include the deterministic, non-noisy case, i.e., $p_\ell^2/N_{\omega_\ell} = 0 \forall \ell$, for reference. As expected, for large $k = \log_2(\tilde{r}_\ell) - \ell$, $\tilde{\gamma}$ nears its optimal value $\tilde{\gamma}^{min}$ for all three strategies because this corresponds to a large number of samples of each model. For the deterministic case, OCV provides more than four orders of magnitude reduction in variance compared to the single fidelity MC (SFMC), whereas OCV-1, which corresponds to the use of a single most correlated low-fidelity model, can attain two orders of magnitude reduction with respect to MC. Note that MLMC is unable to asymptotically reach OCV-1 due to the fixed coefficients, i.e., $\alpha_\ell = -1$, as predicted by the theory. The introduction of noise increasingly decreases the maximum attainable variance reduction while, for large values of k , all methods still converge to their maximum variance reduction. In particular, OCV provides just ~ 1.2 times more variance reduction than OCV-1 and the gap between the multifidelity strategies and SFMC is much smaller, $\tilde{\gamma}_{OCV}^{min} = .0226, .1455, .2448$ for $p_\ell^2/N_{\omega_\ell} = .01, .1, .2$, respectively. Hence, there is still significant potential for variance reduction compared to a single-fidelity strategy, but no longer significant differences between an OCV with four LF models or one using \tilde{Q}_1 alone. The limiting values $\tilde{\gamma}^{min}$ for OCV and OCV-1 are summarized in Table 1. Finally, note as the effective noise level grows, fewer number of LF QoI evaluations are needed before the methods converge to their maximum variance reduction since their performance is dominated by the amount of stochastic noise present.

Now we move to the second scenario and allow $N_{\omega_\ell} \rightarrow \infty$ for all low-fidelity models ($\ell \geq 1$) such that $\tilde{Q}_\ell \rightarrow Q_\ell$. The maximal potential for variance reduction increases for both OCV and OCV-1, as the effective LF noise is removed (Table 1). However, as the effective noise in the high-fidelity model increases, the gap in variance reduction between OCV and OCV-1 again decreases. This indicates the three additional low-fidelity models for OCV are becoming less beneficial even with all noise removed in the limit of infinite low-fidelity samples. Ultimately, the HF noise limits all correlations $\tilde{\rho}_{0,\ell}$ and, in turn, the efficacy of these MF techniques in this scenario.

$\frac{p_\ell^2}{N_{\omega_\ell}}$	$\tilde{\gamma}_{OCV}^{min}$, All models noisy	$\tilde{\gamma}_{OCV}^{min}$, Only HF model noisy	$\tilde{\gamma}_{OCV-1}^{min}$, All models noisy	$\tilde{\gamma}_{OCV-1}^{min}$, Only HF model noisy
0	2.54×10^{-5}	-	9.98×10^{-3}	-
.01	.0226	.0099	.0284	.0198
.1	.1455	.0909	.1734	.1000
.2	.2448	.1667	.2993	.1750

TABLE 1: Minimal variance reduction ratio for various amounts of effective stochastic noise, monomial problem (Section 4.3).

4.4 Optimal ACV estimators under a budget constraint

In a practical scenario, large numbers of LF QoI evaluations with many replicas are not possible, but rather the available computational budget would need to be allocated between evaluations in the uncertainty space and replicas. More formally, the aim is to minimize the ACV estimator variance for a given budget W_{tot} :

$$\operatorname{argmin}_{N_0, \tilde{\mathbf{r}}, \mathbf{N}_\omega} \left(\mathbb{V} \left[\langle Q \rangle^{ACV} (N_0, \tilde{\mathbf{r}}; \mathbf{N}_\omega) \right] \right) \quad \text{s.t.} \quad W^{ACV} \leq W_{tot}, \quad (44)$$

where $W^{ACV} = W^{ACV}(N_0, \tilde{\mathbf{r}}; \mathbf{N}_\omega) = N_0 \left(\tilde{W}_0 + \sum_{\ell=1}^L \tilde{r}_\ell \tilde{W}_\ell \right)$ is the achieved cost of the ACV estimator. The explicit dependence of α is dropped with the understanding that either the coefficients are fixed or $\tilde{\alpha}^{opt}$ is used depending on the estimator.

For simplicity, Equation (44) is solved via a nested approach in this work. The nested approach has an outer loop

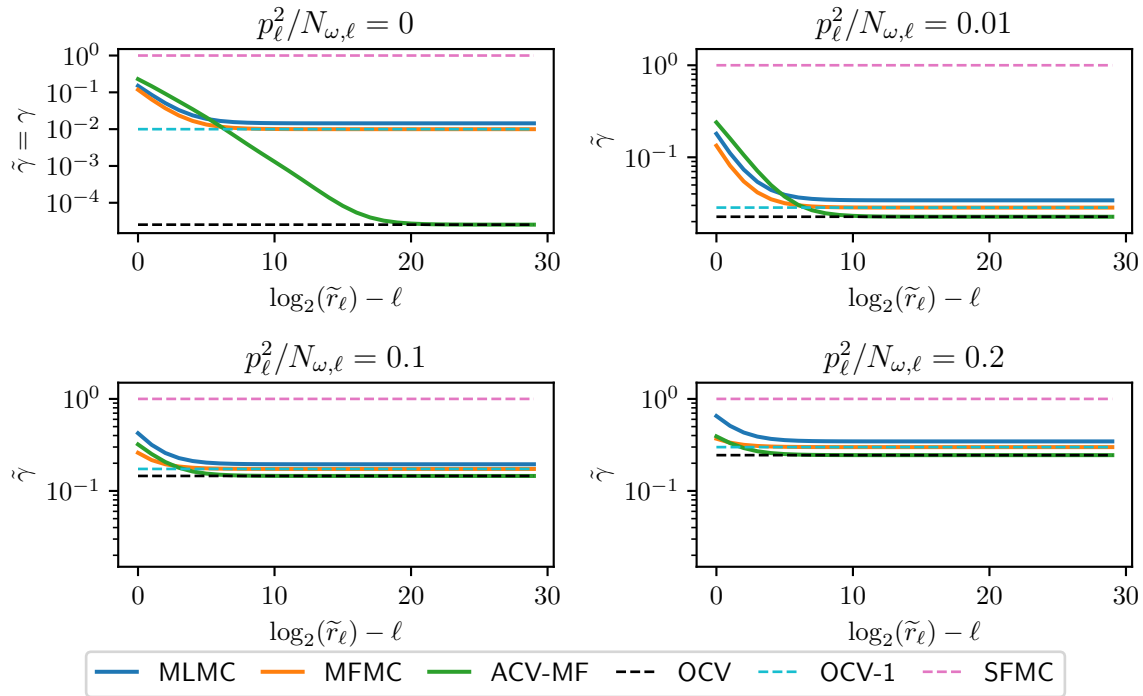


FIG. 2: Variance reduction ratio $\tilde{\gamma}$ versus number of LF QoI evaluations (measured by r_ℓ) for the example in Section 4.3 under four different effective stochastic noise scenarios. The noise measure $p_\ell^2/N_{\omega,\ell}$ is the same for all models in each scenario. The asymptotic $\tilde{\gamma}^{min}$ are shown for the OCV (employing all models), the OCV-1 (using only 1 LF model), and SFMC (HF only). Note the three cases with noise share the same axis scaling, while the non-noisy case (upper left) does not.

over replicate profiles N_ω in some search space \mathcal{N}^\sharp . In the inner loop, the standard ACV optimization problem

$$\operatorname{argmin}_{N_0, \tilde{\mathbf{r}}} \left(\mathbb{V} \left[\langle Q \rangle^{ACV} (N_0, \tilde{\mathbf{r}}; N_\omega) \right] \right) \quad \text{s.t.} \quad W^{ACV} \leq W_{tot}, \quad (45)$$

can be solved since N_ω fixes \tilde{W}_ℓ and \tilde{C} . This is attractive since existing solvers for Equation (45) can be used with these modified costs and covariances.

For a given N_ω , an analytic solution to the inner loop problem (Equation (45)) always exists for MLMC (see APPENDIX C). For MFMC, an analytic solution exists if certain constraints are met (see Theorem 3.4 in [18]):

$$1 = |\tilde{\rho}_{0,0}| > |\tilde{\rho}_{0,1}| > \cdots > |\tilde{\rho}_{0,L}| \quad \text{and} \\ \frac{\tilde{W}_{\ell-1}}{\tilde{W}_\ell} > \frac{\tilde{\rho}_{0,\ell-1}^2 - \tilde{\rho}_{0,\ell}^2}{\tilde{\rho}_{0,\ell}^2 - \tilde{\rho}_{0,\ell+1}^2} \quad \text{for } \ell = 1, \dots, L \quad (46)$$

where, for notational convenience, $\tilde{\rho}_{0,L+1} = 0$. This, as well as Equation (41), motivates reordering the QoIs based on decreasing $|\tilde{\rho}_{0,\ell}|$, as has been assumed herein. This guarantees the first constraint in Equation (46) is always met for MFMC. Reordering can also improve performance for MLMC and robustness for estimators requiring a numerical solution of Equation (45); hence, it is employed for all estimators tested here. Note the reordering depends on N_ω and may differ while searching for the optimal N_0 , $\tilde{\mathbf{r}}$, and N_ω . A numerical solution of MFMC which does not impose the constraints can also be used. When converged, the numerical solution should match the analytical solution

[‡]For example, \mathcal{N} might be the Cartesian product $(1, 2, \dots, N_\omega^{max})^L$ for a given maximum number of replicas N_ω^{max} .

(to numerical tolerances), if the latter exists. However, the solution of Equation (44) may differ for the analytic and numerical versions of MFMC since the optimal N_ω may not permit an analytic solution even with the aforementioned reordering due to the second constraint on effective QoI costs.

4.5 Optimal number of HF replicas for analytic MFMC

The analytic MFMC is still useful since Equation (45) can be solved at basically no cost and may give comparable results to its numerical counterpart. Additionally, it is useful for theoretical advancements such as the observation that, under the assumptions laid forth in [18], $N_{\omega_0} = 1$ is optimal for MFMC. For N_ω fixed, Equation (45) becomes the same optimization problem as in [18] whose analytic solution, under certain constraints on the model correlations and costs (Equation (46)), is

$$N_0 = \frac{W_{tot}}{\sum_{\ell=0}^L \tilde{r}_\ell \tilde{W}_\ell}, \quad \tilde{r}_\ell = \sqrt{\frac{\tilde{W}_0 (\tilde{\rho}_{0,\ell}^2 - \tilde{\rho}_{0,\ell+1}^2)}{\tilde{W}_\ell (1 - \tilde{\rho}_{0,1}^2)}}. \quad (47)$$

The minimized variance is

$$\mathbb{V} [\langle Q \rangle^{MF} (N_\omega)] = \frac{\mathbb{V} [\tilde{Q}_0]}{W_{tot}} \left(\sum_{\ell=0}^L \sqrt{\tilde{W}_\ell (\tilde{\rho}_{0,\ell}^2 - \tilde{\rho}_{0,\ell+1}^2)} \right)^2. \quad (48)$$

Proposition 10. *Assuming the conditions of (46) hold, $|\rho_{0,1}| < 1$, a linear cost model, and fixed budget W_{tot} , $N_{\omega_0} = 1$ is optimal for minimizing the MFMC variance, Equation (48), with respect to the number of high-fidelity replicas.*

Proof. Included in APPENDIX B.6. It is shown $\partial \mathbb{V} [\langle Q \rangle^{MF}] / \partial N_{\omega_0} > 0 \forall N_{\omega_0}$. □

Recall the optimal single-fidelity estimator (Proposition 3) also has $N_{\omega_0} = 1$.

5. NUMERICAL EXPERIMENTS

In this section, we explore the performance and characteristics of the MF UQ methods from Section 4 under a budget constraint. As stated in Section 4.4, a nested approach is used to solve Equation (44) since our goal is simply to investigate the performance of various sampling-based MF UQ estimators in a variety of scenarios defined by the noise levels associated with the non-deterministic models. An approach focused on obtaining the optimal resource allocation while limiting the exploration of the search space will be the subject of future studies. Finally, we note that non-noisy statistics can be obtained directly from their noisy counterparts by employing the relationships derived in Section 2.2, which effectively provide an avenue for the efficient solution of this problem by only requiring a set of pilot runs; see, e.g., [30]. To solve the inner loop problem the optimization approach from [22] is used when numerical solutions are required. For a given inner loop iteration (i.e., N_ω is fixed), the analytic MLMC and MFMC (if it exists) solutions are used to generate two initial guesses for the optimization of ACV-MF and the numerical MFMC (without cost/correlation constraints, labelled here as MFMC-NUM). This process generates two solutions for each of these estimators and the solution with the lowest variance is then selected. Note that in the following examples the model covariances are known, not estimated. Additionally, the noise-to-variance ratios p_ℓ^2 are also specified, not estimated.

When imposing a budget constraint W_{tot} , it is most appropriate to compare estimators optimized for the same total cost. Hence, rather than $\tilde{\gamma}$ for which N_0 and N_{ω_0} are fixed between the SFMC and MF estimators, in this section the ratio of the optimized ACV estimator variance to the optimized SFMC variance for the given W_{tot} is reported to describe variance reduction. We will use $\tilde{\gamma}^*$ to distinguish this quantity. For purposes of comparison, γ^* will indicate the ratio of deterministic, non-noisy MF and SFMC estimators for the same effective cost.

5.1 Monomial problem

We first revisit the monomial exemplar from Section 4.3, now imposing a budget constraint. Letting $W_0 = 1$, Equation (44) is solved for $W_{tot} = 10^3$ with two different cost profiles:

1. $W_\ell = 10^{-\ell}$ for $\ell \geq 1$ (smaller cost gap) – Figure 3;
2. $W_\ell = 10^{-\ell-1}$ for $\ell \geq 1$ (larger cost gap) – Figure 4.

For the outer loop, a simple grid search over $N_\omega \in [1, 10]^5$ is employed. Figures 3 and 4 plot variance reduction ratio $\tilde{\gamma}^*$ against stochastic noise level for a budget of $W_{tot} = 10^3$. For each noise scenario, p_ℓ^2 is the same for the four models. The single fidelity MC (SFMC) is implicitly indicated by $\tilde{\gamma}^* = 1$.

In all cases, the multifidelity strategies are preferable to the single-fidelity estimator. In the deterministic, non-noisy case (included in Table 2 for reference), ACV-MF performs best, achieving slightly less than a two order-of-magnitude reduction in variance compared to SFMC for cost profile 1 and slightly more than a two order-of-magnitude improvement for cost profile 2. For the lower noise levels shown in Figures 3 and 4, ACV-MF provides closer to one order-of-magnitude improvement over SFMC in these non-deterministic scenarios. ACV-MF achieves ~ 4 times more variance reduction than the second most performant MF strategy for both cost profiles in the deterministic case. For $p_\ell \geq 10^{-2}$, all the MF estimators tested except MLMC, i.e., ACV-MF/MFMC/MFMC-NUM, perform similarly. This indicates that the noise’s negative impact on correlation effectively prevents the efficient use of low-fidelity evaluations. Finally, as predicted by the theory, MLMC is less robust to noise as its performance degrades more rapidly than ACV-MF and MFMC with increasing noise for cost profile 1 (Figure 3) and cost profile 2 (Figure 4). One way to see this is by looking at the point for which $\tilde{\gamma}^* \geq 1$ (SFMC is preferred to the MF strategy). Of the estimators shown, MLMC crosses this threshold first for both cost profiles.

Tables 3 and 4 detail the optimal N_ω^{opt} for MFMC, ACV-MF, MFMC-NUM, and MLMC for various p_ℓ^2 for the two cost profiles. Additionally, the percentage of invalid optimizations out of the 10^5 outer loop iterations are listed (not applicable for MLMC). We note here that the number of invalid optimizations is high for MFMC (e.g., $\sim 9\%$ and 27% for $p_\ell^2 = .1$ and $.2$) because we consider MFMC to be invalid when the second constraint of Equation (46) is not met, e.g., because the correlation/cost trade-off does not follow the second of Equation (46). For the estimators requiring numerical optimizations (ACV-MF and MFMC-NUM), invalid optimizations (that is, when the optimizer does not converge) are infrequent: 0.006% at most, for all shown p_ℓ^2 . Despite this, the optimized variances for MFMC and MFMC-NUM differ at most by $\sim 2.5\%$ for $p_\ell^2 \leq .2$, indicating the optimal MFMC solution is nearby the optimal MFMC-NUM solution. As the noise level grows further, we observe that it is increasingly more difficult to satisfy the MFMC constraint and a larger discrepancy appears between MFMC and MFMC-NUM, especially for cost profile 1. Finally, note N_ω^{opt} varies from estimator to estimator, as well as for different p_ℓ^2 and for the two cost profiles. However, $N_{\omega_0}^{opt}$, the number of high-fidelity replicas, is always 1 for ACV-MF, MF-NUM, and MLMC. For cost profile 1 and $p_\ell^2 = 1$ and 2, $N_{\omega_0}^{opt}$ is 2 for MFMC, presumably because the second constraint of Equation (46) cannot be satisfied for $N_{\omega_0} = 1$.

Estimator	γ^*	γ^*
	smaller cost gap	larger cost gap
SFMC	1	1
MLMC	0.048	0.022
MFMC	0.046	0.019
MFMC-NUM	0.047	0.019
ACV-MF	0.015	0.005

TABLE 2: Ratio of MF estimator variance to SFMC variance for the same cost, non-noisy estimators. Monomial example with budget constraint: $W_{tot} = 10^3$. MFMC-NUM is the numerical MFMC solution.

As demonstrated in Section 4.3, high noise scenarios can lessen the gap between an OCV which employs all LF models and one which uses only one LF model. Hence, we perform an additional study where Equation (44) is

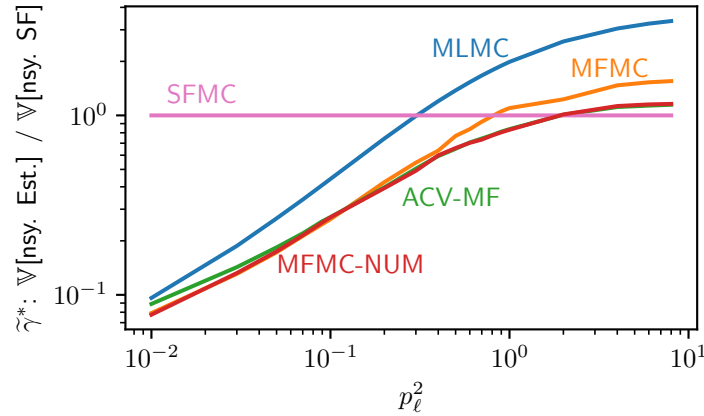


FIG. 3: Ratio of MF estimator variance to SFMC variance for same effective cost versus stochastic noise level, monomial example with budget constraint: $W_0 = 1$, $W_\ell = 10^{-\ell}$ for $\ell \geq 1$ and $W_{tot} = 10^3$. The noise measure p_ℓ^2 is the same for all models in each scenario. MFMC–NUM is the numerical MFMC solution.

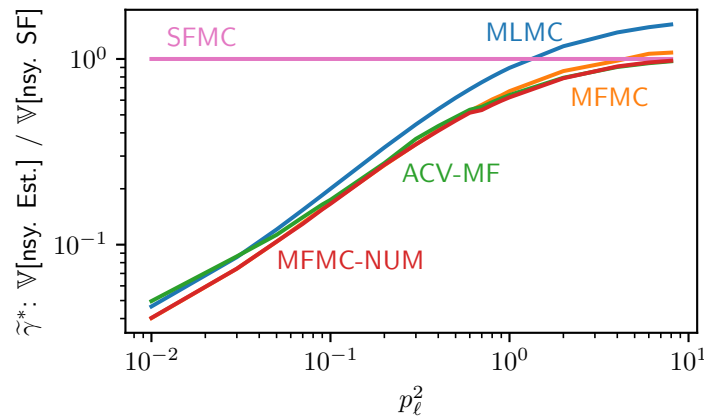


FIG. 4: Ratio of MF estimator variance to SFMC variance for same effective cost versus stochastic noise level, monomial example with budget constraint: $W_0 = 1$, $W_\ell = 10^{-\ell-1}$ for $\ell \geq 1$ and $W_{tot} = 10^3$. The noise measure p_ℓ^2 is the same for all models in each scenario. MFMC–NUM is the numerical MFMC solution.

p_ℓ^2	N_{ω}^{opt} , ACV-MF	Inv. Opt. ACV-MF	N_{ω}^{opt} , MFMC	Inv. Opt., MFMC	N_{ω}^{opt} , MFMC-NUM	Inv. Opt., MFMC-NUM	N_{ω}^{opt} , MLMC
.01	[1, 1, 1, 1, 7]	0.001%	[1, 1, 1, 1, 1]	0.000%	[1, 1, 1, 1, 1]	0.000%	[1, 1, 1, 1, 1]
.1	[1, 1, 1, 4, 10]	0.000%	[1, 1, 1, 1, 1]	9.200%	[1, 1, 1, 1, 1]	0.000%	[1, 2, 3, 2, 1]
.2	[1, 1, 1, 3, 9]	0.000%	[1, 2, 1, 2, 1]	27.200%	[1, 1, 4, 4, 1]	0.006%	[1, 2, 3, 3, 2]
1	[1, 1, 2, 2, 6]	0.000%	[2, 2, 1, 1, 2]	76.536%	[1, 1, 2, 10, 8]	0.000%	[1, 3, 6, 8, 5]
2	[1, 1, 2, 2, 4]	0.000%	[2, 2, 1, 1, 1]	89.106%	[1, 1, 2, 8, 5]	0.000%	[1, 3, 7, 10, 8]

TABLE 3: Optimal replicate profiles and percentage of invalid optimizations (out of 10^5 candidate profiles) for various amounts of stochastic noise, monomial problem under a cost budget of $W_{tot} = 1000$ and cost profile 1 (Section 5.1). MFMC results in an invalid optimization when the constraints (Equation (46)) are not met.

p_ℓ^2	N_{ω}^{opt} , ACV-MF	Inv. Opt. ACV-MF	N_{ω}^{opt} , MFMC	Inv. Opt., MFMC	N_{ω}^{opt} , MFMC-NUM	Inv. Opt., MFMC-NUM	N_{ω}^{opt} , MLMC
.01	[1, 1, 7, 10, 2]	0.004%	[1, 2, 1, 1, 1]	0.000%	[1, 2, 1, 1, 1]	0.000%	[1, 2, 1, 1, 1]
.1	[1, 2, 5, 6, 10]	0.000%	[1, 4, 3, 2, 1]	5.000%	[1, 4, 4, 2, 1]	0.000%	[1, 5, 5, 3, 1]
.2	[1, 2, 3, 5, 10]	0.000%	[1, 3, 6, 3, 1]	11.620%	[1, 4, 5, 4, 1]	0.000%	[1, 6, 7, 5, 2]
1	[1, 1, 4, 7, 3]	0.000%	[1, 7, 4, 8, 4]	57.280%	[1, 1, 10, 10, 4]	0.000%	[1, 8, 10, 10, 6]
2	[1, 1, 6, 9, 4]	0.000%	[1, 6, 3, 3, 9]	73.593%	[1, 1, 10, 10, 5]	0.000%	[1, 8, 10, 10, 8]

TABLE 4: Optimal replicate profiles and percentage of invalid optimizations (out of 10^5 candidate profiles) for various amounts of stochastic noise, monomial problem under a cost budget of $W_{tot} = 1000$ and cost profile 1 (Section 5.1). MFMC results in an invalid optimization when the constraints (Equation (46)) are not met.

solved using only the HF model and a single LF model. Note that, numerical MFMC and ACV-MF are equivalent in these two model cases. Figure 5 shows $\tilde{\gamma}^*$ versus p_ℓ^2 for ACV-MF for the four possible HF/LF pairings and cost profile 1. Figure 6 shows the same for cost profile 2. In both cases, the ACV-MF solution with all models and SFMC are included for reference. For cost profile 2, where the raw model costs are cheaper and more QoI evaluations are possible, LF2, LF3, and LF4 collapse to the performance of all models for the higher noise levels. For cost profile 1, this collapse is not seen. Figures 5 and 6 also show as the noise level changes so does the relative performance of the estimators. The results here include only a simple subset of the possible estimators that can be formed through model selection (i.e., subselecting from the set of L low-fidelity models) but highlight how variable performance can be as a function of W_ℓ , $\rho_{m,n}$, and p_ℓ^2 once computational cost must be spread between replication and QoI evaluation.

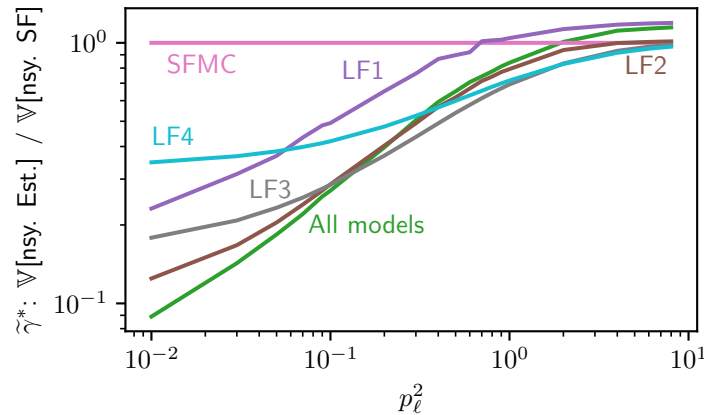


FIG. 5: Ratio of two-model ACV-MF estimator variance to SFMC variance for same effective cost versus stochastic noise level, monomial example with budget constraint: $W_0 = 1$, $W_\ell = 10^{-\ell}$ for $\ell \geq 1$ and $W_{tot} = 10^3$. The noise measure p_ℓ^2 is the same for all models in each scenario. $LF\ell$ indicates \tilde{Q}_ℓ is used as the LF QoI. The solution using all models is included for reference (All models).

5.2 Turbulent signals

This exemplar is designed to mimic models whose outputs are dependent on unknown input parameters and fluctuate in time about a mean as in statistically-stationary turbulent flow. For this case, the QoIs \tilde{Q}_ℓ are obtained by averaging over N_{ω_ℓ} equally-spaced *snapshots* of the time-dependent process (see Figure 7). We consider models which are time-discretized Ornstein-Uhlenbeck processes with x -dependent means $\mu_\ell(x)$ and parameters θ_ℓ and s_ℓ :

$$g_\ell(x, t_{N+1}) = g_\ell(x, t_N) + \theta_\ell \Delta t (\mu_\ell(x) - g_\ell(x, t_N)) + s_\ell \sqrt{\Delta t} \xi_\ell(t_N) \quad (49)$$

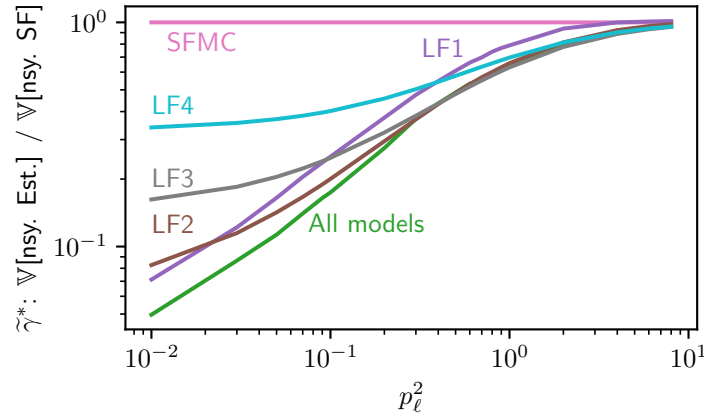


FIG. 6: Ratio of two-model ACV-MF estimator variance to SFMC variance for same effective cost versus stochastic noise level, monomial example with budget constraint: $W_0 = 1$, $W_\ell = 10^{-\ell-1}$ for $\ell \geq 1$ and $W_{tot} = 10^3$. The noise measure p_ℓ^2 is the same for all models in each scenario. $LF\ell$ indicates \bar{Q}_ℓ is used as the LF QoI. The solution using all models is included for reference (All models).

where $\xi_\ell \sim \mathcal{N}(0, 1)$. For simplicity, all models share the same uniform timestep $\Delta t = t_{N+1} - t_N$ and the same window between snapshots $t_{window} = K\Delta t$, where K is an integer. Recognizing Equation (49) is a first-order autoregressive process, the mean and variance are easily computed to be $\mu_\ell(x)$ and $\Psi_\ell^2 \equiv s_\ell^2 \Delta t / (2\theta_\ell \Delta t - \theta_\ell^2 \Delta t^2)$, respectively. To ensure stationarity, the initial condition is taken such that $g_\ell(x, t_0) \sim \mathcal{N}(\mu_\ell, \Psi_\ell^2)$. The snapshots can be viewed as samples of a random variable $f_\ell(x, \omega_\ell)$ with the same mean and variance, $\mathbb{E}_{\Omega_\ell} [f_\ell] = \mu_\ell(x)$, $\mathbb{V}_{\Omega_\ell} [f_\ell] = \Psi_\ell^2$, as the time-varying process. Hence, the theory introduced in this work can be directly applied.

The means

$$\mu_\ell(x) = \mu_\ell(x_1, x_2) = 2 + (2x_1^5 + 2x_2^5) \mathcal{I}_\ell^1 + 3x_1x_2 + (x_1^2 + x_2^2 + 5x_1^2x_2^2) \mathcal{I}_\ell^2 + .5x_1 + .5x_2 \quad (50)$$

are adopted from an example in [40]. The uncertain parameters are assumed to be $x_i \sim \mathcal{U}(-1, 1)$ and independent and the \mathcal{I}_ℓ^i are set to 1 or 0 to create a four model set (summarized in Table 5). As in the previous example, p_ℓ^2 is set to be the same for all models. The model costs are consistent with cost profile 1 above: $W_\ell = 10^{-\ell}$ and $W_{tot} = 10^3$. For this example, we solve the ACV optimization problem (Equation (44)) with a grid search over $N_\omega \in [1, 10]^4$, as well as a search over $N_\omega \in [5, 10]^4$. The second case reflects a scenario often encountered in practice: the noisy and non-noisy statistics are only approximately known in practice and will be computed with the samples used to construct the mean estimator. Hence, a minimum threshold in the number of replicas is necessary. The value of 5 is chosen for purposes of example, but realistically the threshold is problem-dependent and may also vary from model to model.

As a baseline, ACV-MF ($\gamma^* = .075$) achieves more than five times the variance reduction of MFMC-NUM and nearly eight times the variance reduction of MLMC in the non-noisy, deterministic case (Table 6). All non-noisy estimators are more performant than SFMC. Figure 8 plots variance reduction ratio $\tilde{\gamma}^*$ against p_ℓ where the grid search over $N_\omega \in [1, 10]^4$ is employed. Figure 9 shows the same quantity but for a search over $N_\omega \in [5, 10]^4$. The trends in each case are similar: ACV-MF achieves 2 – 3 times more variance reduction than MFMC-NUM for small noise levels ($p_\ell^2 < .1$), ACV-MF and MFMC-NUM perform similarly for larger noise levels ($p_\ell^2 \geq 1$), MLMC lags in performance compared to the other MF estimators. Comparing the two cases, ACV-MF and MFMC-NUM outperform SFMC for higher noise levels (up to $p_\ell^2 \sim 4$) for the $N_\omega \in [5, 10]^4$ case, compared to $p_\ell^2 \sim .7$ for the case without a lower threshold. For MLMC, the breaking points are $p_\ell^2 \sim .7$ and $p_\ell^2 \sim .1$. We remark here that differences among estimators are expected in the low-noise regime since in this regime the estimators approach their non-noisy counterparts. On the contrary, for high-noise regimes the degradation in the correlations is so significant

that effectively no estimator is capable of exploiting low-fidelity information; as a consequence their performance is similarly poor. This clearly manifests in the estimators attaining a higher variance than SFMC.

ℓ	\mathcal{I}_ℓ^1	\mathcal{I}_ℓ^2	$\rho_{0,\ell}$
0	1	1	1
1	0	1	.907
2	1	0	.792
3	0	0	.679

TABLE 5: Values of \mathcal{I}_ℓ^i which define the means of the four Ornstein-Uhlenbeck processes. The non-noisy correlation between the resulting four models and the HF model also shown.

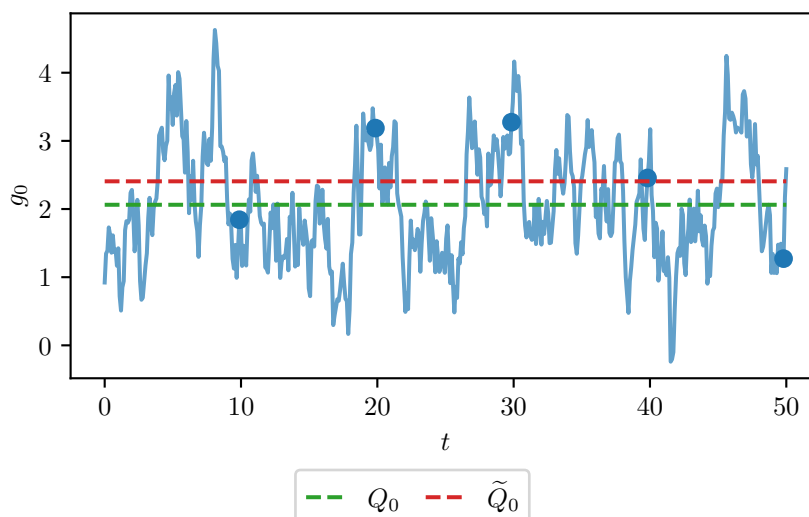


FIG. 7: One trajectory of the discrete high-fidelity Ornstein-Uhlenbeck process ($x_1 = .5, x_2 = .5, \sigma_0 = 1.25, \theta_0 = 1, \Delta t = .1$). $N_{\omega_0} = 5$ snapshots are indicated with filled circles. The exact and computed time averages, Q_0 and \tilde{Q}_0 , are shown as dashed lines.

Estimator	γ^*
SFMC	1
MLMC	0.580
MFMC-NUM	0.387
ACV-MF	0.075

TABLE 6: Ratio of MF estimator variance to SFMC variance for same effective cost, non-noisy estimators. Turbulent signal example with budget constraint: $W_0 = 1, W_\ell = 10^{-\ell}$ for $\ell \geq 1$ and $W_{tot} = 10^3$. MFMC-NUM is the numerical MFMC solution.

6. CONCLUSION

Non-deterministic models have received far less attention than their deterministic counterparts from a UQ perspective. As a consequence, MF UQ approaches in literature are not as well-developed to efficiently serve this class of applications/models. In this contribution, we extended the ACV framework introduced in [22] to the case of non-deterministic

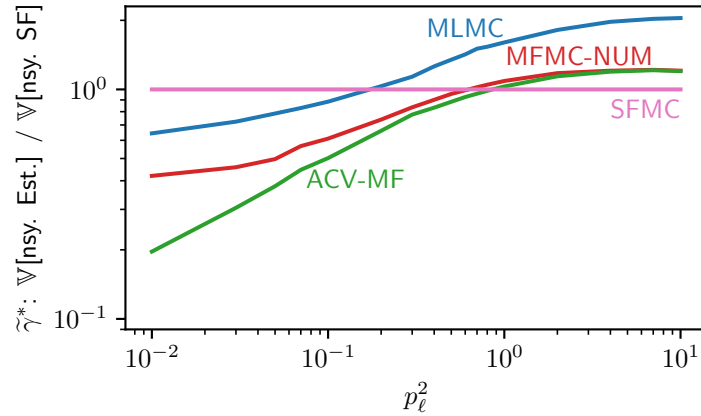


FIG. 8: Ratio of MF estimator variance to SFMC variance for same effective cost, turbulent signal example with budget constraint: $W_0 = 1$, $W_\ell = 10^{-\ell}$ for $\ell \geq 1$ and $W_{tot} = 10^3$. $\mathbf{N}\omega \in [1, 10]^4$ is explored to find the optimal solution. The noise measure p_ℓ^2 is the same for all models in each scenario. MFMC-NUM is the numerical MFMC solution.

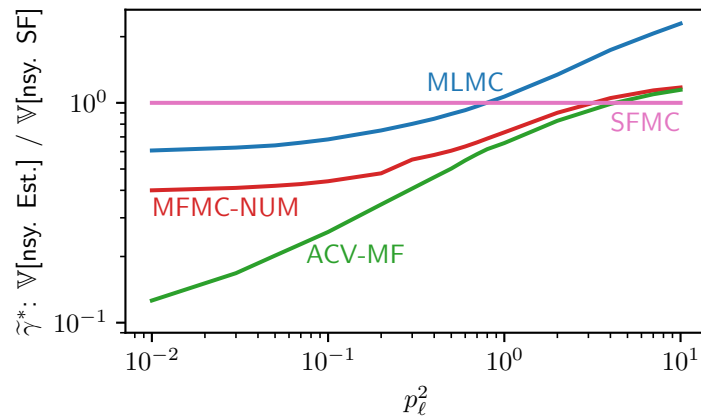


FIG. 9: Ratio of MF estimator variance to SFMC variance for same effective cost, turbulent signal example with budget constraint: $W_0 = 1$, $W_\ell = 10^{-\ell}$ for $\ell \geq 1$ and $W_{tot} = 10^3$. $\mathbf{N}\omega \in [5, 10]^4$ is explored to find the optimal solution. The noise measure p_ℓ^2 is the same for all models in each scenario. MFMC-NUM is the numerical MFMC solution.

models. Several theoretical results are presented. First, the effect of noise introduced by the high- and low-fidelity models in the ACV variance was shown to manifest as separate, additive degradations to the non-deterministic variance reduction. Second, we analyzed a few instances of ACV estimators, namely MLMC, MFMC, and ACV-MF and demonstrated that MLMC is intrinsically more susceptible to the presence of noise. MFMC and ACV-MF are more equipped to handle noise, but their performance can severely degrade, compared to the deterministic case, due to lower correlations and the need for replication. Finally, it was shown that, under a budget constraint, only one HF replica is optimal for MFMC.

Numerical results were provided to illustrate the challenges in this space. Two examples were considered, namely a verification test problem extracted from [22], extended to include non-determinism, and a turbulent signal model problem. Noise was shown to prevent the most efficient use of LF QoI evaluations. ACV-MF, which is most performant in the non-deterministic case for the examples shown, performs more on par with MFMC (i.e., gap between

OCV and OCV-1 is much smaller) for large levels of stochastic noise. Additionally, by examining a few different estimators created by subselecting models from the available set, the variability of estimator performance as a function of the noise appeared. Finally, we note the use of a numerical solution for MFMC, which circumvents the analytic cost constraint was robust to changes in N_ω and was able to find a more optimal solution than the analytic MFMC for large noise levels.

The goal of this manuscript was not necessarily to provide a numerically efficient approach to deal with the challenges of non-determinism in multifidelity estimators, but rather to illustrate the main features of this endeavor and to provide a starting point for subsequent theoretical and algorithmic development. Clearly, selecting the optimal model graph (see, e.g., [25]) is crucial as carrying additional LF models means introducing additional sources of noise, which may be detrimental. This pairs with the trade-off between replication and sampling the UQ space, which can be viewed as a model tuning problem with hyperparameters N_ω that set the cost and correlation of the QoIs. An advantage here is these costs and correlations (as well as other statistics) are known functions of N_ω . This advantage could be leveraged when developing approaches to solve the combined optimization over samples and replicas (inner and outer loops) [30], potentially building on recent work in this space [24,41,42]. A few key assumptions could be relaxed in future work. For instance, we assumed a perfect knowledge of statistics that govern the behavior of the estimators; however, in a realistic deployment these statistics will need to be estimated. Hence, there exists the challenge of balancing computational resources between exploration of the model ensemble via pilot runs (to accurately compute the statistics) and the exploitation phase (obtaining samples for the optimized estimator). In the context of non-deterministic models, this challenge is exacerbated by the need to obtain replicas for each model to estimate their noise, i.e., the exploration cost is augmented by the need for replication. Strategies focused on tailoring this exploration to the problem and models at hand, e.g., by targeting a level of precision for noise estimation rather than employing a fixed number of replicas per model, should be devised to make the approach viable for real-world applications. Other cost models, such as those explored in [30,33], should be considered to reflect the fact that, in many applications, replication is much cheaper than evaluating a model at a new location in the UQ space. Finally, strategies which let the number of replicas varying over the UQ space should be investigated to improve the overall computational efficiency of the approaches based on the presented framework.

ACKNOWLEDGMENTS

This article has been authored by an employee of National Technology & Engineering Solutions of Sandia, LLC under Contract No. DE-NA0003525 with the U.S. Department of Energy (DOE). The employee owns all right, title and interest in and to the article and is solely responsible for its contents. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this article or allow others to do so, for United States Government purposes. The DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan

<https://www.energy.gov/downloads/doe-public-access-plan>. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government. B. Reuter and G. Geraci were partially supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Scientific Discovery through Advanced Computing (SciDAC) Program through the FASTMath Institute. T. Wildey was supported by the CHaRMNET MMICC Center funded by the Advanced Scientific Computing Research program of the U.S. Department of Energy.

APPENDIX A. SUMMARY OF NOTATION AND RELATIONSHIPS BETWEEN NOISY AND NON-NOISY QUANTITIES

Quantity	Noisy	Non-Noisy	Important relationships
QoI	\tilde{Q}_ℓ	Q_ℓ	$Q_\ell = \lim_{N_{\omega_\ell} \rightarrow \infty} \tilde{Q}_\ell$
Cost	\tilde{W}_ℓ	W_ℓ	$\tilde{W}_\ell = N_{\omega_\ell} W_\ell$
QoI variance	$\mathbb{V}[\tilde{Q}_\ell]$	$\mathbb{V}_X[Q_\ell]$	$\mathbb{V}[\tilde{Q}_\ell] = \mathbb{V}_X[Q_\ell] + \mathbb{E}_X[\sigma_{\omega_\ell}^2]/N_{\omega_\ell},$ $\mathbb{V}_X[Q_\ell] = \lim_{N_{\omega_\ell} \rightarrow \infty} \mathbb{V}[\tilde{Q}_\ell]$
QoI covariance	$\mathbb{C}[\tilde{Q}_m, \tilde{Q}_n]$	$\mathbb{C}_X[Q_m, Q_n]$	$\mathbb{C}[\tilde{Q}_m, \tilde{Q}_n] = \mathbb{C}_X[Q_m, Q_n], m \neq n$
QoI correlation	$\tilde{\rho}_{m,n}$	$\rho_{m,n}$	$\tilde{\rho}_{m,n} = \frac{\rho_{m,n} \tau_m \tau_n}{\sqrt{(\tau_m^2 + p_m^2/N_{\omega_m})(\tau_n^2 + p_n^2/N_{\omega_n})}},$ $\rho_{m,n} = \lim_{N_{\omega_m}, N_{\omega_n} \rightarrow \infty} \tilde{\rho}_{m,n}$
MC estimator (w.r.t. x)	$\langle Q \rangle_\ell$	\hat{Q}_ℓ	$\hat{Q}_\ell = \lim_{N_{\omega_\ell} \rightarrow \infty} \langle Q \rangle_\ell$
MC variance	$\mathbb{V}[\langle Q \rangle_\ell]$	$\mathbb{V}_X[\hat{Q}_\ell]$	$\mathbb{V}[\langle Q \rangle_\ell] = 1/N_0 (\mathbb{V}_X[Q_\ell] + \mathbb{E}_X[\sigma_{\omega_\ell}^2]/N_{\omega_\ell}),$ $\mathbb{V}_X[\hat{Q}_\ell] = \lim_{N_{\omega_\ell} \rightarrow \infty} \mathbb{V}[\langle Q \rangle_\ell]$
ACV variance reduction	\tilde{R}^2	R^2	$R^2 = \lim_{N_{\omega_\ell} \rightarrow \infty} \tilde{R}^2,$ $\mathbb{V}[\langle Q \rangle^{ACV}] = \mathbb{V}[\langle Q \rangle_0] (1 - \tilde{R}^2) =$ $\mathbb{V}_X[\hat{Q}_0] (1 - R^2 + \delta \tilde{R}_{HF}^2 + \delta \tilde{R}_{LF}^2)$
ACV variance reduction ratio	$\tilde{\gamma}$	γ	$\gamma = \lim_{N_{\omega_\ell} \rightarrow \infty} \tilde{\gamma},$ $\gamma = 1 - R^2,$ $\tilde{\gamma} = 1 - (R^2 - \delta \tilde{R}_{LF}^2) / (1 + \delta \tilde{R}_{HF}^2)$

TABLE A.7: Nomenclature and important relationships for various quantities needed to describe and analyze non-deterministic MF estimators. Recall $\tau_\ell = \sqrt{\mathbb{V}_X[Q_\ell]}/\sqrt{\mathbb{V}_X[Q_0]}$, $p_\ell^2 = \mathbb{E}_X[\sigma_{\omega_\ell}^2]/\mathbb{V}_X[Q_0]$, $\delta \tilde{R}_{HF}^2 = p_0^2/N_{\omega_0}$, $\delta \tilde{R}_{LF}^2 = R^2 - (1 + p_0^2/N_{\omega_0}) \tilde{R}^2$.

APPENDIX B. PROOFS

APPENDIX B.1 Variance of noisy QoI

By the law of total variance and the definition of the noisy QoI (Equation (3)), it follows

$$\begin{aligned}
\mathbb{V}[\tilde{Q}_\ell] &= \mathbb{V}_X[\mathbb{E}_{\Omega_\ell}[\tilde{Q}_\ell]] + \mathbb{E}_X[\mathbb{V}_{\Omega_\ell}[\tilde{Q}_\ell]] \\
&= \mathbb{V}_X[\mathbb{E}_{\Omega_\ell}[f_\ell]] + \mathbb{E}_X\left[\frac{\mathbb{V}_{\Omega_\ell}[f_\ell]}{N_{\omega_\ell}}\right] \\
&= \mathbb{V}_X[Q_\ell] + \frac{\mathbb{E}_X[\sigma_{\omega_\ell}^2]}{N_{\omega_\ell}}.
\end{aligned} \tag{B.1}$$

APPENDIX B.2 Covariance of noisy QoI

We would like to demonstrate $\mathbb{C} [\tilde{Q}_m, \tilde{Q}_n] = \mathbb{C}_X [Q_m, Q_n]$ for two different models $f_m(x, \omega_m)$ and $f_n(x, \omega_n)$. To do so it is necessary to define the product probability space $\mathcal{P}_{X \times \Omega_m \times \Omega_n}$ and trivially extend f_m and f_n to $X \times \Omega_m \times \Omega_n$ in Ω_n and Ω_m , respectively. The covariance $\mathbb{C} [f_m, f_n]$ is then well-defined on $\mathcal{P}_{X \times \Omega_m \times \Omega_n}$. Let $f_m^i \equiv f_m(x, \omega_m^{(i)})$ be the i^{th} replica of the m^{th} model for $x \in X$. Then

$$\begin{aligned} \mathbb{C} [\tilde{Q}_m, \tilde{Q}_n] &= \mathbb{C} \left[\frac{1}{N_{\omega_m}} \sum_{i=1}^{N_{\omega_m}} f_m^i, \frac{1}{N_{\omega_n}} \sum_{j=1}^{N_{\omega_n}} f_n^j \right] \\ &= \sum_{i=1}^{N_{\omega_m}} \sum_{j=1}^{N_{\omega_n}} \mathbb{C} \left[\frac{1}{N_{\omega_m}} f_m^i, \frac{1}{N_{\omega_n}} f_n^j \right] \\ &= \sum_{i=1}^{N_{\omega_m}} \sum_{j=1}^{N_{\omega_n}} \frac{1}{N_{\omega_m} N_{\omega_n}} \mathbb{C} [f_m^i, f_n^j]. \end{aligned} \quad (\text{B.2})$$

By the law of total covariance

$$\begin{aligned} \mathbb{C} [f_m^i, f_n^j] &= \mathbb{C}_X [\mathbb{E}_{\Omega_m \times \Omega_n} [f_m^i], \mathbb{E}_{\Omega_m \times \Omega_n} [f_n^j]] + \mathbb{E}_X [\mathbb{C}_{\Omega_m \times \Omega_n} [f_m^i, f_n^j]] \\ &= \mathbb{C}_X [Q_m, Q_n] \end{aligned} \quad (\text{B.3})$$

where the second term on the first line is zero since ω_m and ω_n are assumed independent and, since f_m has no ω_n dependence, $\mathbb{E}_{\Omega_m \times \Omega_n} [f_m^i] = \mathbb{E}_{\Omega_m} [f_m^i] = Q_m$. Finally,

$$\begin{aligned} \sum_{i=1}^{N_{\omega_m}} \sum_{j=1}^{N_{\omega_n}} \frac{1}{N_{\omega_m} N_{\omega_n}} \mathbb{C} [f_m^i, f_n^j] &= \sum_{i=1}^{N_{\omega_m}} \sum_{j=1}^{N_{\omega_n}} \frac{1}{N_{\omega_m} N_{\omega_n}} \mathbb{C}_X [Q_m, Q_n] \\ &= \mathbb{C}_X [Q_m, Q_n] \end{aligned} \quad (\text{B.4})$$

which proves the result.

APPENDIX B.3 Low-fidelity noise's impact on variance reduction

For arbitrary coefficients, start with Equation (15) and Equations (25) and (26) to find

$$\begin{aligned} \left(1 + \frac{p_0^2}{N_{\omega_0}}\right) \tilde{R}^2 &= -\frac{1}{\mathbb{V}_X [\hat{Q}_0]} \left(\frac{1}{N_0} \alpha^T [\mathbf{F} \circ (\mathbf{C} + \delta \tilde{\mathbf{C}})] \alpha + 2\alpha^T \mathbb{C}_X [\Delta, \hat{Q}_0] \right) \\ &= -\frac{1}{\mathbb{V}_X [\hat{Q}_0]} \left(\frac{1}{N_0} \alpha^T [\mathbf{F} \circ \delta \tilde{\mathbf{C}}] \alpha + \alpha^T \mathbb{C}_X [\Delta, \Delta] \alpha + 2\alpha^T \mathbb{C}_X [\Delta, \hat{Q}_0] \right) \\ &= -\frac{1}{\mathbb{V}_X [\hat{Q}_0]} \frac{1}{N_0} \alpha^T [\mathbf{F} \circ \delta \tilde{\mathbf{C}}] \alpha + R^2. \end{aligned} \quad (\text{B.5})$$

The result directly follows.

If the optimal coefficients α^{opt} are used for the non-noisy estimator and $\tilde{\alpha}^{\text{opt}}$ are used for the noisy estimator, then

$$\begin{aligned} \left(1 + \frac{p_0^2}{N_{\omega_0}}\right) \tilde{R}^2 &= \mathbb{C}_X [\Delta, \hat{Q}_0]^T \frac{\mathbb{C}_X [\tilde{\Delta}, \tilde{\Delta}]^{-1}}{\mathbb{V}_X [\hat{Q}_0]} \mathbb{C}_X [\Delta, \hat{Q}_0] \\ \implies R^2 - \left(1 + \frac{p_0^2}{N_{\omega_0}}\right) \tilde{R}^2 &= \mathbb{C}_X [\Delta, \hat{Q}_0]^T \left(\mathbb{C}_X [\Delta, \Delta]^{-1} - \mathbb{C} [\tilde{\Delta}, \tilde{\Delta}]^{-1} \right) \mathbb{C}_X [\Delta, \hat{Q}_0] \end{aligned} \quad (\text{B.6})$$

APPENDIX B.4 Variance of MLMC

Let $\mathbf{1} = -\alpha$ be a length- L vector of ones and $\eta_\ell = N_{\ell,2}/N_0$ such that $\tilde{r}_\ell = \eta_{\ell-1} + \eta_\ell$. Starting with Equation (32) and

$$\mathbf{F}_{ML} = \begin{bmatrix} \ddots & \ddots & 0 & \dots & \dots & \dots & 0 \\ sym & \ddots & \ddots & 0 & \dots & \dots & \vdots \\ \vdots & sym & \frac{\eta_\ell + \eta_{\ell-1}}{\eta_\ell \eta_{\ell-1}} & -\frac{1}{\eta_\ell} & 0 & \dots & \vdots \\ \vdots & \dots & sym & \ddots & \ddots & 0 & \vdots \\ sym & \dots & \dots & sym & \ddots & \ddots & 0 \end{bmatrix}, \quad (\text{B.7})$$

$N_0^{-1} \mathbf{1}^T [\mathbf{F}_{ML} \circ \delta \tilde{\mathbf{C}}] \mathbf{1}$ is simply the sum of the entries of $\mathbf{F}_{ML} \circ \delta \tilde{\mathbf{C}}$ (which is diagonal due to $\delta \tilde{\mathbf{C}}$) scaled by $1/N_0$. This means

$$\mathbb{V}_X [\hat{Q}_0] \delta R_{ML,LF}^2 = \frac{1}{N_0} \sum_{\ell=1}^L \frac{\eta_\ell + \eta_{\ell-1}}{\eta_\ell \eta_{\ell-1}} \frac{\mathbb{E}_X [\sigma_{\omega_\ell}^2]}{N_{\omega_\ell}}, \quad (\text{B.8})$$

which is easily rearranged to get the final result.

APPENDIX B.5 Variance of MFMC

Recognizing that \mathbf{F}_{MF} is diagonal with entries $(\tilde{r}_\ell - \tilde{r}_{\ell-1})/(\tilde{r}_\ell \tilde{r}_{\ell-1})$, it follows that

$$\begin{aligned} \mathbb{C}_X [\Delta, \hat{Q}_0]^T \left(\mathbb{C}_X [\Delta, \Delta]^{-1} - \mathbb{C} [\tilde{\Delta}, \tilde{\Delta}]^{-1} \right) \mathbb{C}_X [\Delta, \hat{Q}_0] &= \\ \frac{1}{N_0} \sum_{\ell=1}^L \left(\frac{\tilde{r}_\ell - \tilde{r}_{\ell-1}}{\tilde{r}_\ell \tilde{r}_{\ell-1}} \right) c_\ell \left(\frac{\tilde{r}_\ell \tilde{r}_{\ell-1}}{\tilde{r}_\ell - \tilde{r}_{\ell-1}} \right) \left[\frac{1}{C_{\ell\ell}} - \frac{1}{\tilde{C}_{\ell\ell}} \right] \left(\frac{\tilde{r}_\ell - \tilde{r}_{\ell-1}}{\tilde{r}_\ell \tilde{r}_{\ell-1}} \right) c_\ell & \\ = \frac{1}{N_0} \sum_{\ell=1}^L \left(\frac{\tilde{r}_\ell - \tilde{r}_{\ell-1}}{\tilde{r}_\ell \tilde{r}_{\ell-1}} \right) c_\ell^2 \frac{\delta \tilde{C}_{\ell\ell}}{C_{\ell\ell} \tilde{C}_{\ell\ell}}. & \end{aligned} \quad (\text{B.9})$$

Recalling $\delta \tilde{C}_{\ell\ell} = \mathbb{E}_X [\sigma_{\omega_\ell}^2] / N_{\omega_\ell}$ and employing Equation (32) gives

$$\begin{aligned} \delta R_{MF,LF}^2 &= \frac{N_0}{\mathbb{V}_X [Q_0]} \frac{1}{N_0} \sum_{\ell=1}^L \left(\frac{\tilde{r}_\ell - \tilde{r}_{\ell-1}}{\tilde{r}_\ell \tilde{r}_{\ell-1}} \right) c_\ell^2 \frac{\mathbb{E}_X [\sigma_{\omega_\ell}^2]}{N_{\omega_\ell}} \frac{\mathbb{V}_X [Q_0]}{\mathbb{V}_X [Q_\ell] \mathbb{V} [\tilde{Q}_\ell] \mathbb{V}_X [Q_0]} \\ &= \sum_{\ell=1}^L \left(\frac{\tilde{r}_\ell - \tilde{r}_{\ell-1}}{\tilde{r}_\ell \tilde{r}_{\ell-1}} \right) \rho_{0,\ell}^2 \frac{p_\ell^2}{\tau_\ell^2 + \frac{p_\ell^2}{N_{\omega_\ell}}}. \end{aligned} \quad (\text{B.10})$$

APPENDIX B.6 Optimal number of high-fidelity replicas

Starting with the expression for the optimized MFMC variance given N_ω and a budget W_{tot} (Equation (48)), we would like to show $\partial \mathbb{V} [\langle Q \rangle^{MF}] / \partial N_{\omega_0} > 0 \forall N_{\omega_0}$. By assumption, $1 > |\tilde{\rho}_{0,1}| > \dots > |\tilde{\rho}_{0,L}|$ (see Theorem 3.4 in

[18]) and $|\rho_{0,1}| < 1$. First, recognize

$$\frac{\partial \tilde{\rho}_{0,\ell}^2}{\partial N_{\omega_0}} = \frac{\mathbb{E}_X [\sigma_{\omega_0}^2]}{N_{\omega_0}^2 \mathbb{V} [\tilde{Q}_0]} \tilde{\rho}_{0,\ell}^2 \quad \text{for } \ell \neq 1. \quad (\text{B.11})$$

Differentiating Equation (48) then gives

$$\begin{aligned} \frac{\partial \mathbb{V} [\langle Q \rangle^{MF}]}{\partial N_{\omega_0}} &= -\frac{\mathbb{E}_X [\sigma_{\omega_0}^2]}{W_{tot} N_{\omega_0}^2} \times \left\{ \sum_{\ell=0}^L \sqrt{\tilde{W}_\ell (\tilde{\rho}_{0,\ell}^2 - \tilde{\rho}_{0,\ell+1}^2)} \right\}^2 \\ &+ \frac{\mathbb{V} [\tilde{Q}_0]}{W_{tot}} \times \left\{ \sum_{\ell=0}^L \sqrt{\tilde{W}_\ell (\tilde{\rho}_{0,\ell}^2 - \tilde{\rho}_{0,\ell+1}^2)} \right\} \times \\ &\left\{ \frac{W_0 (1 - \tilde{\rho}_{0,1}^2) - \tilde{W}_0 \frac{\mathbb{E}_X [\sigma_{\omega_0}^2]}{N_{\omega_0}^2 \mathbb{V} [\tilde{Q}_0]} \tilde{\rho}_{0,1}^2}{\sqrt{\tilde{W}_0 (1 - \tilde{\rho}_{0,1}^2)}} + \frac{\mathbb{E}_X [\sigma_{\omega_0}^2]}{N_{\omega_0}^2 \mathbb{V} [\tilde{Q}_0]} \sum_{\ell=1}^L \sqrt{\tilde{W}_\ell (\tilde{\rho}_{0,\ell}^2 - \tilde{\rho}_{0,\ell+1}^2)} \right\}. \end{aligned} \quad (\text{B.12})$$

Let

$$\zeta = \frac{\sum_{\ell=0}^L \sqrt{\tilde{W}_\ell (\tilde{\rho}_{0,\ell}^2 - \tilde{\rho}_{0,\ell+1}^2)}}{W_{tot} \sqrt{\tilde{W}_0 (1 - \tilde{\rho}_{0,1}^2)}}$$

then Equation (B.12) becomes

$$\begin{aligned} &\zeta \mathbb{V} [\tilde{Q}_0] W_0 (1 - \tilde{\rho}_{0,1}^2) - \zeta W_0 \frac{\mathbb{E}_X [\sigma_{\omega_0}^2]}{N_{\omega_0}} \tilde{\rho}_{0,1}^2 \\ &+ \zeta \frac{\mathbb{E}_X [\sigma_{\omega_0}^2]}{N_{\omega_0}^2} \sqrt{\tilde{W}_1 (1 - \tilde{\rho}_{0,1}^2)} \times \left\{ \sum_{\ell=1}^L \sqrt{\tilde{W}_\ell (\tilde{\rho}_{0,\ell}^2 - \tilde{\rho}_{0,\ell+1}^2)} \right\} \\ &- \zeta \frac{\mathbb{E}_X [\sigma_{\omega_0}^2]}{N_{\omega_0}^2} \sqrt{\tilde{W}_0 (1 - \tilde{\rho}_{0,1}^2)} \times \left\{ \sum_{\ell=0}^L \sqrt{\tilde{W}_\ell (\tilde{\rho}_{0,\ell}^2 - \tilde{\rho}_{0,\ell+1}^2)} \right\} \\ &= \zeta \mathbb{V} [\tilde{Q}_0] W_0 (1 - \tilde{\rho}_{0,1}^2) - \zeta W_0 \frac{\mathbb{E}_X [\sigma_{\omega_0}^2]}{N_{\omega_0}} \tilde{\rho}_{0,1}^2 \\ &\quad - \zeta \frac{\mathbb{E}_X [\sigma_{\omega_0}^2]}{N_{\omega_0}^2} \sqrt{\tilde{W}_0 (1 - \tilde{\rho}_{0,1}^2)} \sqrt{\tilde{W}_0 (1 - \tilde{\rho}_{0,1}^2)} \\ &= \zeta \mathbb{V}_X [\tilde{Q}_0] W_0 (1 - \tilde{\rho}_{0,1}^2) - \zeta W_0 \frac{\mathbb{E}_X [\sigma_{\omega_0}^2]}{N_{\omega_0}} \tilde{\rho}_{0,1}^2 - \zeta \frac{\mathbb{E}_X [\sigma_{\omega_0}^2]}{N_{\omega_0}} W_0 (1 - \tilde{\rho}_{0,1}^2) \\ &= \zeta W_0 \left[\mathbb{V}_X [Q_0] (1 - \tilde{\rho}_{0,1}^2) - \frac{\mathbb{E}_X [\sigma_{\omega_0}^2]}{N_{\omega_0}} \tilde{\rho}_{0,1}^2 \right] \end{aligned} \quad (\text{B.13})$$

Note when $|\tilde{\rho}_{0,\ell}| < 1$ then $\zeta > 0$, so it is sufficient to show the bracketed quantity is non-negative. Then

$$\begin{aligned}
\mathbb{V}_X [Q_0] (1 - \tilde{\rho}_{0,1}^2) - \frac{\mathbb{E}_X [\sigma_{\omega_0}^2]}{N_{\omega_0}} \tilde{\rho}_{0,1}^2 &= \mathbb{V}_X [Q_0] - \mathbb{V} [\tilde{Q}_0] \tilde{\rho}_{0,1}^2 \\
&= \mathbb{V}_X [Q_0] - \frac{\mathbb{C}_X [Q_0, Q_1]^2}{\mathbb{V} [\tilde{Q}_1]} \\
&> \mathbb{V}_X [Q_0] - \frac{\mathbb{V}_X [Q_0] \mathbb{V}_X [Q_1]}{\mathbb{V} [\tilde{Q}_1]} \\
&> \mathbb{V}_X [Q_0] - \frac{\mathbb{V}_X [Q_0] \mathbb{V}_X [Q_1]}{\mathbb{V}_X [Q_1]} = 0.
\end{aligned} \tag{B.14}$$

Since the trivial case where models 0 and 1 are perfectly correlated is ignored, the inequality is strict and the derivative is always positive. Hence, $N_{\omega_0} = 1$ will be optimal.

APPENDIX C. MLMC ANALYTIC SOLUTION TO INNER LOOP OPTIMIZATION PROBLEM

In [14] an analytic solution to a similar optimization problem to Equation (45) is given but in their problem, a target variance is the constraint rather than the budget. For completeness, the solution to Equation (45) is included here. The solution procedure parallels that in [14]. For simplicity, the solution is expressed in terms of $N_{\ell,2}$ and, for convenience, $Q_{L+1} = 0, W_{L+1} = 0$:

$$\begin{aligned}
N_{\ell,2} &= \frac{W_{tot}}{S} \sqrt{\frac{\mathbb{V} [\tilde{Q}_\ell] + \mathbb{V} [\tilde{Q}_{\ell+1}] - 2\mathbb{C}_X [Q_\ell, Q_{\ell+1}]}{\tilde{W}_\ell^*}} \\
\text{where } S &= \sum_{\ell=0}^L \sqrt{(\mathbb{V} [\tilde{Q}_\ell] + \mathbb{V} [\tilde{Q}_{\ell+1}] - 2\mathbb{C}_X [Q_\ell, Q_{\ell+1}]) \tilde{W}_\ell^*} \\
&\text{and } \tilde{W}_\ell^* = N_{\omega_\ell} W_\ell + N_{\omega_{\ell+1}} W_{\ell+1}.
\end{aligned} \tag{C.1}$$

Note that $N_{0,2} = N_0$ and $N_\ell = N_{\ell-1,2} + N_{\ell,2}$ for $\ell \geq 1$ so $\tilde{r}_\ell = N_\ell/N_0$ is easily computed from Equation (C.1).

REFERENCES

1. Tarman, T., Rollins, T., Swiler, L., Cruz, J., Vugrin, E., Huang, H., Sahu, A., Wlazlo, P., Goulart, A., and Davis, K., Comparing reproduced cyber experimentation studies across different emulation testbeds, In *Proceedings of the 14th Cyber Security Experimentation and Test Workshop*, CSET '21, p. 63–71, New York, NY, USA, 2021. Association for Computing Machinery.
2. Crussell, J., Kroeger, T.M., Brown, A., and Phillips, C., Virtually the same: Comparing physical and virtual testbeds, In *2019 International Conference on Computing, Networking and Communications (ICNC)*. IEEE, 2019.
3. minimega developers. minimega: a distributed vm management tool, 2019.
4. Bear, J., *Modeling phenomena of flow and transport in porous media*, Vol. 1, Springer, 2018.
5. Grigoriu, M., Reduced order models for random functions. application to stochastic problems, *Applied Mathematical Modelling*, 33(1):161–175, 2009.
6. Gupta, A., Cecen, A., Goyal, S., Singh, A.K., and Kalidindi, S.R., Structure–property linkages using a data science approach: application to a non-metallic inclusion/steel composite system, *Acta Materialia*, 91:239–254, 2015.
7. Lewis, E.E. and Miller, W.F., *Computational methods of neutron transport*, John Wiley and Sons, Inc., New York, NY, 1984.
8. Tskhakaya, D., Matyash, K., Schneider, R., and Taccogna, F., The particle-in-cell method, *Contributions to Plasma Physics*, 47(8-9):563–594, 2007.

9. Moin, P. and Mahesh, K., Direct numerical simulation: a tool in turbulence research, *Annual review of fluid mechanics*, 30(1):539–578, 1998.
10. Zhu, X. and Sudret, B., Replication-based emulation of the response distribution of stochastic simulators using generalized lambda distributions, *International Journal for Uncertainty Quantification*, 10(3), 2020.
11. Zhu, X. and Sudret, B., Emulation of stochastic simulators using generalized lambda models, *SIAM/ASA Journal on Uncertainty Quantification*, 9(4):1345–1380, 2021.
12. Zhu, X. and Sudret, B., Stochastic polynomial chaos expansions to emulate stochastic simulators, *International Journal for Uncertainty Quantification*, 2022.
13. Geraci, G. and Olson, A.J., Impact of sampling strategies in the polynomial chaos surrogate construction for Monte Carlo transport applications, In *Proceedings of the American Nuclear Society M&C 2021*, pp. 76–86, 2021.
14. Giles, M.B., Multilevel Monte Carlo Path Simulation, *Operations Research*, 56(3):607–617, June 2008, publisher: INFORMS.
15. Pasupathy, R., Schmeiser, B.W., Taaffe, M.R., and Wang, J., Control-variate estimation using estimated control means, *IIE Transactions*, 44(5):381–385, 2012.
16. Ng, L.W. and Willcox, K.E., Multifidelity approaches for optimization under uncertainty, *International Journal for numerical methods in Engineering*, 100(10):746–772, 2014.
17. Giles, M.B., Multilevel Monte Carlo methods, *Acta numerica*, 24:259–328, 2015.
18. Peherstorfer, B., Willcox, K., and Gunzburger, M., Optimal Model Management for Multifidelity Monte Carlo Estimation, *SIAM Journal on Scientific Computing*, 38(5):A3163–A3194, January 2016, publisher: Society for Industrial and Applied Mathematics.
19. Haji-Ali, A.L., Nobile, F., and Tempone, R., Multi-index Monte Carlo: when sparsity meets sampling, *Numerische Mathematik*, 132(4):767–806, 2016.
20. Geraci, G., Eldred, M.S., and Iaccarino, G., A multifidelity multilevel Monte Carlo method for uncertainty propagation in aerospace applications, In *19th AIAA non-deterministic approaches conference*, pp. 19–51, 2017.
21. Peherstorfer, B., Willcox, K., and Gunzburger, M., Survey of multifidelity methods in uncertainty propagation, inference, and optimization, *Siam Review*, 60(3):550–591, 2018.
22. Gorodetsky, A.A., Geraci, G., Eldred, M.S., and Jakeman, J.D., A generalized approximate control variate framework for multifidelity uncertainty quantification, *Journal of Computational Physics*, 408:109257, May 2020.
23. Schaden, D. and Ullmann, E., On multilevel best linear unbiased estimators, *SIAM/ASA Journal on Uncertainty Quantification*, 8(2):601–635, 2020.
24. Adams, B.M., Eldred, M.S., Geraci, G., Portone, T., Ridgway, E.M., Stephens, J.A., and Wildey, T.M., Deployment of multifidelity uncertainty quantification for thermal battery assessment Part I: Algorithms and single cell results, 9 2022.
25. Bomarito, G., Leser, P., Warner, J., and Leser, W., On the optimization of approximate control variates with parametrically defined estimators, *Journal of Computational Physics*, 451:110882, February 2022.
26. Olson, A., Calculation of parametric variance using variance deconvolution., Technical report SAND2019-2691C, Sandia National Laboratories, 2019.
27. Geraci, G., Swiler, L., and Debusschere, B., Multifidelity UQ sampling for stochastic simulations, *16th U.S. National Congress on Computational Mechanics*, 2021, SAND2021-8907C, <https://doi.org/10.2172/1889573>.
28. Clements, K.B., Geraci, G., and Olson, A.J., Numerical investigation on the performance of a variance deconvolution estimator, *Trans. Am. Nucl. Soc.*, 126:344–347, 2022.
29. Clements, K.B., Geraci, G., Olson, A.J., and Palmer, T., A variance deconvolution estimator for efficient uncertainty quantification in Monte Carlo radiation transport applications. (Under Review), *Journal of Quantitative Spectroscopy and Radiative Transfer*, 2023.
30. Geraci, G., Reuter, B., Olson, A., and Clements, K., Multifidelity UQ methods for Monte Carlo radiation applications and stochastic media, *17th U.S. National Congress on Computational Mechanics*, 2023, SAND2023-07000C.
31. Schaden, D. and Ullmann, E., Asymptotic analysis of multilevel best linear unbiased estimators, *SIAM/ASA Journal on Uncertainty Quantification*, 9(3):953–978, 2021.
32. Croci, M., Willcox, K., and Wright, S., Multi-output multilevel best linear unbiased estimators via semidefinite programming, *Computer Methods in Applied Mechanics and Engineering*, 413:116130, 2023.

33. Geraci, G. and Olson, A., Multifidelity uncertainty quantification in stochastic media transport problems, *Transactions American Nuclear Society*, 127(1):446–449, 2022.
34. Clements, K.C., Geraci, G., and Olson, A.J., A variance deconvolution approach to sampling uncertainty quantification for Monte Carlo radiation transport solvers, In *Computer Science Research Institute Summer Proceedings 2021*, pp. 293–307, 2021. Technical Report SAND2022-0653R, <https://cs.sandia.gov/summerproceedings/CCR2021.html>.
35. Lavenberg, S.S. and Welch, P.D., A perspective on the use of control variables to increase the efficiency of monte carlo simulations, *Management Science*, 27(3):322–335, 1981.
36. Hesterberg, T., Control variates and importance sampling for efficient bootstrap simulations, *Statistics and Computing*, 6:147–157, 1996.
37. Lavenberg, S., Moeller, T., and Welch, P., *Statistical results on multiple control variables with application to variance reduction in queueing network simulation*, IBM Thomas J. Watson Research Division, 1978.
38. Lavenberg, S.S., Moeller, T.L., and Welch, P.D., Statistical results on control variables with application to queueing network simulation, *Operations Research*, 30(1):182–202, 1982.
39. Geraci, G., Clements, K., and Olson, A.J., A polynomial chaos approach for uncertainty quantification of Monte Carlo transport codes, In *Proceedings of the American Nuclear Society M&C 2023*, 2023.
40. Gorodetsky, A.A., Jakeman, J.D., and Geraci, G., Mfnets: data efficient all-at-once learning of multifidelity surrogates as directed networks of information sources, *Computational Mechanics*, 68(4):741–758, 2021.
41. Bomarito, G., Geraci, G., Warner, J., Leser, P., Leser, W., Eldred, M.S., Jakeman, J., and Gorodetsky, A., Improving multi-model trajectory simulation estimators using model selection and tuning, In *AIAA SCITECH 2022 Forum*, 2022. AIAA 2022-1099.
42. Thompson, M., Geraci, G., Bomarito, G., Warner, J., Leser, P., Leser, W.P., Eldred, M.S., Jakeman, J., and Gorodetsky, A., Strategies for automation of model tuning in multi-fidelity trajectory uncertainty propagation, In *AIAA SCITECH 2023 Forum*, p. 1481, 2023.