

# Multifidelity Monte Carlo write-up

---

## Abstract

We investigate the Grad-Shafranov free boundary problem in Tokamak fusion reactors under the influence of parameter uncertainties. Using both Monte Carlo and multi-fidelity Monte Carlo sampling approaches, we quantify the impact of these uncertainties on model predictions, emphasizing the statistical characterization of solution variability across diverse parameter regimes. Our numerical results reveals that the multi-fidelity Monte Carlo estimator achieves statistical accuracy comparable to the Monte Carlo approach. However, the multi-fidelity method demonstrates superior computational efficiency, achieving a cost reduction by a factor of ..., while preserving fidelity in representing plasma boundary dynamics and geometric parameters. This work underscores the efficiency of multi-fidelity frameworks in addressing the computational demands of uncertainty quantification in complex fusion reactor models, offering a robust pathway for enhancing predictive capabilities in plasma physics.

---

## 1. Model problem

To characterize the uncertainty random parameters of our model, we introduce a  $d$ -dimensional random variable  $\omega = [\omega_1, \dots, \omega_d]$ , whose  $i$ -th component corresponds to the  $i$ -th parameter, each component is mutually independent random variables with joint distribution  $\pi(\omega)$ . The parameter space is  $W$ .

We have a high fidelity model denoted as  $f_1 : W \rightarrow Z$  we desire, and several low fidelity models (surrogates)  $f_k$  for  $k = 2, \dots, K$ . Our objective is to approximate

$$\mathbb{E}(f_1(\omega)) = \int_W f_1(r, z, \omega) \pi(\omega) d\omega.$$

Note for each  $f_k(\omega)$ , its variance and Pearson product-moment correlation coefficient are

$$\sigma_k^2 = \mathbb{V}(f_k(\omega)), \quad \rho_{k,j} = \frac{\text{Cov}(f_k(\omega), f_j(\omega))}{\sigma_k \sigma_j}, \quad k, j = 1, \dots, K,$$

where  $\mathbb{V}(f) := \mathbb{E}(\|f - \mathbb{E}(f)\|_Z^2)$  and  $\text{Cov}(f_k, f_j) := \mathbb{E}((f_k - \mathbb{E}(f_k), f_j - \mathbb{E}(f_j))_Z)$ . Note that  $\rho_{k,k} = 1$ .

## 2. Monte Carlo estimator

Let  $u_h$  be the discretized solution to (??). The Monte Carlo Finite-Element estimator  $A_N^{\text{MC}}$  is defined as the sample mean of  $N$  independent and identically distributed (i.i.d.) realizations  $\omega_1, \dots, \omega_N$

$$A_N^{\text{MC}} := \frac{1}{N} \sum_{i=1}^N u_h(\omega_i), \tag{1}$$

where  $\mathbb{E}(A_N^{\text{MC}}) = \mathbb{E}(u_h)$ ,  $\mathbb{V}(A_N^{\text{MC}}) = \mathbb{V}(u_h)/N$ .

We define the normalized *mean squared error* (nMSE) to quantify the error between  $\mathbb{E}(u)$  and estimator  $A$ , denoted as  $\mathcal{E}_A^2$ , for an estimator  $A$  with a normalizing factor  $\|\mathbb{E}(u)\|_Z^2$  as

$$\mathcal{E}_A^2 := \frac{\mathbb{E} \left[ \|\mathbb{E}(u) - A\|_Z^2 \right]}{\|\mathbb{E}(u)\|_Z^2}.$$

The nMSE for Monte Carlo estimator becomes

$$\mathcal{E}_{A_N^{\text{MC}}}^2 = \frac{\|\mathbb{E}(u) - \mathbb{E}(u_h)\|_Z^2 + \mathbb{E} \left[ \|\mathbb{E}(u_h) - A_N^{\text{MC}}\|_Z^2 \right]}{\|\mathbb{E}(u)\|_Z^2} = \frac{\|\mathbb{E}(u) - \mathbb{E}(u_h)\|_Z^2}{\|\mathbb{E}(u)\|_Z^2} + \frac{\mathbb{V}(u_h)}{N \|\mathbb{E}(u)\|_Z^2} = \mathcal{E}_{\text{Bias}}^2 + \mathcal{E}_{\text{Stat}}^2,$$

Given a target tolerance  $\epsilon^2$  for the nMSE, the required sample size  $N$  is estimated as

$$N = \frac{\sigma_1^2}{\epsilon^2 \|\mathbb{E}(u)\|_Z^2} \simeq \epsilon^{-2}.$$

Let  $C$  denote the average evaluation cost per sample for  $u_h$ . The total sampling cost using  $N$  samples with the Monte Carlo method for  $\mathbb{E}(u_h(\omega))$  is

$$\mathcal{W}_{\text{MC}} = CN = \frac{C\sigma_1^2}{\epsilon^2 \|\mathbb{E}(u)\|_Z^2}.$$

### 3. Multifidelity Monte Carlo

The Multi-fidelity Monte Carlo (MFMC) estimator is defined as

$$A^{\text{MF}} := A_{1,N_1}^{\text{MC}} + \sum_{k=2}^K \alpha_k (A_{k,N_k}^{\text{MC}} - A_{k,N_{k-1}}^{\text{MC}}), \quad (2)$$

where  $\alpha_k$  are coefficients weighting the correction terms. In each correction term, the two Monte Carlo estimators are dependent, as  $A_{k,N_{k-1}}^{\text{MC}}$  reuses the first  $N_k$  samples of  $A_{k,N_k}^{\text{MC}}$ , requiring  $N_{k-1} \leq N_k$  for  $k = 2, \dots, K$ . Exploiting this dependence, we can separate  $N_k$  samples into a disjoint union of  $N_{k-1}$  and  $N_k - N_{k-1}$  samples and rewrite (2) as

$$A^{\text{MF}} = A_{1,N_1}^{\text{MC}} + \sum_{k=2}^K \alpha_k \left[ \left( \frac{N_{k-1}}{N_k} - 1 \right) A_{k,N_{k-1}}^{\text{MC}} + \left( 1 - \frac{N_{k-1}}{N_k} \right) A_{k,N_k \setminus N_{k-1}}^{\text{MC}} \right], \quad (3)$$

where the weights in the correction terms now depend on the sample ratio. This reformulation is significant because we partition the  $N_k$  samples into two disjoint sets such that the samples used to estimate  $A_{k,N_{k-1}}^{\text{MC}}$  and  $A_{k,N_k \setminus N_{k-1}}^{\text{MC}}$  are independent. If we define the following notations

$$Y_1 := A_{1,N_1}^{\text{MC}}, \quad Y_k := \left( \frac{N_{k-1}}{N_k} - 1 \right) (A_{k,N_{k-1}}^{\text{MC}} - A_{k,N_k \setminus N_{k-1}}^{\text{MC}}), \quad k = 2, \dots, K,$$

then the multifidelity Monte Carlo estimator in (3) boils down to

$$A^{\text{MF}} = Y_1 + \sum_{k=2}^K \alpha_k Y_k,$$

where  $\mathbb{E}(Y_k) = 0$  for  $k \geq 2$ ,  $\mathbb{E}(A^{\text{MF}}) = \mathbb{E}(u_{h,1})$  and

$$\mathbb{V}(Y_1) = \frac{\sigma_1^2}{N_1}, \quad \mathbb{V}(Y_k) = \left(\frac{N_{k-1}}{N_k} - 1\right)^2 \left(\frac{\sigma_k^2}{N_{k-1}} + \frac{\sigma_k^2}{N_k - N_{k-1}}\right) = \left(\frac{1}{N_{k-1}} - \frac{1}{N_k}\right) \sigma_k^2.$$

Even though the samples in  $Y_k$  and  $Y_j$  for  $k \neq j, k, j = 2, \dots, K$  are dependent, it can be shown that  $Y_k$  and  $Y_j$  are uncorrelated using the fact that samples in the two disjoint groups are independent. However,  $Y_1$  may not necessarily be uncorrelated with  $Y_k$ . Using this fact and the covariance relation established in [1, Lemma 3.2], the variance of the multi-fidelity Monte Carlo estimator can be found as

$$\begin{aligned} \mathbb{V}(A^{\text{MF}}) &= \mathbb{V}(Y_1) + \mathbb{V}\left(\sum_{k=2}^K \alpha_k Y_k\right) + 2 \text{Cov}\left(Y_1, \sum_{k=2}^K \alpha_k Y_k\right), \\ &= \mathbb{V}(Y_1) + \sum_{k=2}^K \alpha_k^2 \mathbb{V}(Y_k) + 2 \sum_{2 \leq k < j \leq K} \alpha_k \alpha_j \text{Cov}(Y_k, Y_j) + 2 \sum_{k=2}^K \alpha_k \text{Cov}(Y_1, Y_k), \\ &= \frac{\sigma_1^2}{N_1} + \sum_{k=2}^K \left(\frac{1}{N_{k-1}} - \frac{1}{N_k}\right) (\alpha_k^2 \sigma_k^2 - 2\alpha_k \rho_{1,k} \sigma_1 \sigma_k). \end{aligned} \quad (4)$$

The nMSE error to approximate (??) for the multi-fidelity Monte Carlo estimator is

$$\mathcal{E}_{A^{\text{MF}}}^2 = \frac{\|\mathbb{E}(u) - \mathbb{E}(A^{\text{MF}})\|_Z^2 + \mathbb{E}\left[\|\mathbb{E}(A^{\text{MF}}) - A^{\text{MF}}\|_Z^2\right]}{\|\mathbb{E}(u)\|_Z^2} = \frac{\|\mathbb{E}(u) - \mathbb{E}(A^{\text{MF}})\|_Z^2}{\|\mathbb{E}(u)\|_Z^2} + \frac{\mathbb{V}(A^{\text{MF}})}{\|\mathbb{E}(u)\|_Z^2} = \mathcal{E}_{\text{Bias}}^2 + \mathcal{E}_{\text{Stat}}^2,$$

The total sampling cost for the MFMC estimator is

$$\mathcal{W}^{\text{MF}} = \sum_{k=1}^K C_k N_k.$$

Our next goal is to determine the sample sizes  $N_k$  and weights  $\alpha_k$  such that the MFMC estimator satisfies a user-specified accuracy requirement,  $\mathcal{E}_{A^{\text{MFMC}}}^2 = \epsilon^2$ . Towards this goal, we formulate a constrained optimization problem aimed at minimize the total sampling cost while ensuring that the variance of the MFMC estimator remains within the target threshold. Additionally, the constraints enforce the ordering of sample sizes  $N_{k-1} \leq N_k$  for  $k = 2, \dots, K$ , and that all sample sizes are non-negative. The resulting optimization problem is formulated as

$$\begin{aligned} \min_{N_1, \dots, N_K \in \mathbb{R}, \alpha_2, \dots, \alpha_K \in \mathbb{R}} \quad & \sum_{k=1}^K C_k N_k, \\ \text{s.t.} \quad & \mathbb{V}(A^{\text{MF}}) - \|\mathbb{E}(f_1)\|_Z^2 \epsilon^2 = 0, \\ & N_{k-1} - N_k \leq 0, \quad k = 2, \dots, K, \\ & -N_1 \leq 0. \end{aligned} \quad (5)$$

Theorem 1 gives the solution to (5).

**Theorem 1.** *Let  $f_k$  be  $K$  models that satisfy the following conditions*

$$(i) \quad |\rho_{1,1}| > \dots > |\rho_{1,K}| \quad (ii) \quad \frac{C_{k-1}}{C_k} > \frac{\rho_{1,k-1}^2 - \rho_{1,k}^2}{\rho_{1,k}^2 - \rho_{1,k+1}^2}, \quad k = 2, \dots, K.$$

Then for  $k = 1 \dots, K$ , the global minimizer to (5) is

$$\alpha_k^* = \frac{\rho_{1,k}\sigma_1}{\sigma_k}, \quad (6)$$

$$N_k^* = \frac{\sigma_1^2}{\|\mathbb{E}(f_1)\|_Z^2 \epsilon^2} \sqrt{\frac{\rho_{1,k}^2 - \rho_{1,k+1}^2}{C_k}} \sum_{j=1}^K \sqrt{C_j (\rho_{1,j}^2 - \rho_{1,j+1}^2)}, \quad \rho_{1,K+1} = 0. \quad (7)$$

*Proof.* Consider the auxiliary Lagrangian function  $L$  with multipliers  $\lambda_0, \dots, \lambda_K$  and its partial derivatives with respect to  $\alpha_k, N_k$

$$\begin{aligned} L &= \sum_{k=1}^K C_k N_k + \lambda_0 \left( \frac{\sigma_1^2}{N_1} + \sum_{k=2}^K \left( \frac{1}{N_{k-1}} - \frac{1}{N_k} \right) (\alpha_k^2 \sigma_k^2 - 2\alpha_k \rho_{1,k} \sigma_1 \sigma_k) \right) - \lambda_1 N_1 + \sum_{k=2}^K \lambda_k (N_{k-1} - N_k), \\ \frac{\partial L}{\partial \alpha_k} &= \lambda_0 \left( \frac{1}{N_{k-1}} - \frac{1}{N_k} \right) (2\alpha_k \sigma_k^2 - 2\rho_{1,k} \sigma_1 \sigma_k), \quad k = 2, \dots, K, \\ \frac{\partial L}{\partial N_1} &= C_1 + \lambda_0 \left( -\frac{\sigma_1^2}{N_1^2} - \frac{\alpha_2^2 \sigma_2^2 - 2\alpha_2 \rho_{1,2} \sigma_1 \sigma_2}{N_1^2} \right) - \lambda_1 + \lambda_2, \\ \frac{\partial L}{\partial N_k} &= C_k + \lambda_0 \left( \frac{\alpha_k^2 \sigma_k^2 - 2\alpha_k \rho_{1,k} \sigma_1 \sigma_k}{N_k^2} - \frac{\alpha_{k+1}^2 \sigma_{k+1}^2 - 2\alpha_{k+1} \rho_{1,k+1} \sigma_1 \sigma_{k+1}}{N_k^2} \right) - \lambda_k + \lambda_{k+1}, \quad k = 2, \dots, K-1, \\ \frac{\partial L}{\partial N_K} &= C_K + \lambda_0 \left( \frac{\alpha_K^2 \sigma_K^2 - 2\alpha_K \rho_{1,K} \sigma_1 \sigma_K}{N_K^2} \right) - \lambda_K. \end{aligned}$$

We can see that  $\alpha_k^* = (\rho_{1,k}\sigma_1)/\sigma_k$  satisfy  $\partial L/\partial \alpha_k = 0$ . Substitute  $\alpha_k^*$  into  $\partial L/\partial N_k = 0$ , we have

$$C_k = \frac{\lambda_0 \sigma_1^2}{N_k^2} (\rho_{1,k}^2 - \rho_{1,k+1}^2) + \lambda_k - \lambda_{k+1}, \quad \text{for } k = 1, \dots, K-1, \quad C_K = \frac{\lambda_0 \sigma_1^2}{N_K^2} \rho_{1,K}^2 + \lambda_K.$$

Karush-Kuhn-Tucker (KKT) conditions

$$\begin{aligned} \frac{\partial L}{\partial \alpha_j} &= 0, \quad \frac{\partial L}{\partial N_k} = 0, \quad j = 2 \dots, K, \quad k = 1 \dots, K, \\ \mathbb{V}(A^{\text{MF}}) - \|\mathbb{E}(f_1)\|_Z^2 \epsilon^2 &= 0, \\ N_{k-1} - N_k &\leq 0, \quad k = 2 \dots, K, \\ -N_1 &\leq 0, \\ \lambda_1, \dots, \lambda_K &\geq 0, \\ \lambda_k (N_{k-1} - N_k) &= 0, \quad k = 2 \dots, K, \\ \lambda_1 N_1 &= 0. \end{aligned}$$

Reference [1] has shown that the global minimizer is achieved with the constraints enforce the ordering of sample sizes  $N_{k-1} < N_k$  for  $k = 2, \dots, K$ , this indicates that the inequality constraints are inactive ( $\lambda_k=0$ ,  $k = 1, \dots, K$ ) in the complementary slackness condition and

$$N_k = \sigma_1 \sqrt{\lambda_0} \sqrt{\frac{\rho_{1,k}^2 - \rho_{1,k+1}^2}{C_k}}, \quad \text{with } \rho_{1,K+1} = 0 \quad k = 1, \dots, K. \quad (8)$$

Therefore, the variance of multifidelity Monte Carlo estimator (4) becomes

$$\begin{aligned}\mathbb{V}(A^{\text{MF}}) &= \frac{\sigma_1^2}{N_1} + \sum_{k=2}^K \left( \frac{1}{N_{k-1}} - \frac{1}{N_k} \right) (\alpha_k^2 \sigma_k^2 - 2\alpha_k \rho_{1,k} \sigma_1 \sigma_k), \\ &= \sigma_1^2 \sum_{k=1}^K \frac{\rho_{1,k}^2 - \rho_{1,k+1}^2}{N_k} = \frac{\sigma_1}{\sqrt{\lambda_0}} \sum_{k=1}^K \sqrt{C_k (\rho_{1,k}^2 - \rho_{1,k+1}^2)}, \quad \text{with } \rho_{K+1} = 0.\end{aligned}\quad (9)$$

If the normalized mean square error achieves a given tolerance  $\epsilon^2$ , we can solve for  $\sqrt{\lambda_0}$  as

$$\sqrt{\lambda_0} = \frac{\sigma_1}{\|\mathbb{E}(f_1)\|_Z^2 \epsilon^2} \sum_{k=1}^K \sqrt{C_k (\rho_{1,k}^2 - \rho_{1,k+1}^2)},$$

substitute  $\sqrt{\lambda_0}$  into (8) so the sample size estimation  $N_k$  can be estimated as

$$N_k = \frac{\sigma_1^2}{\|\mathbb{E}(f_1)\|_Z^2 \epsilon^2} \sqrt{\frac{\rho_{1,k}^2 - \rho_{1,k+1}^2}{C_k}} \sum_{j=1}^K \sqrt{C_j (\rho_{1,j}^2 - \rho_{1,j+1}^2)}.$$

Note that by requiring condition (ii), we can guarantee that  $N_k$  is strictly increasing as  $k$  increases.  $\square$

Using sample size  $N_k \in \mathbb{R}$  derived in Theorem (5), the total sampling cost becomes

$$\mathcal{W}^{\text{MF}} = \sum_{k=1}^K C_k N_k = \frac{\sigma_1^2}{\|\mathbb{E}(f_1)\|_Z^2 \epsilon^2} \left( \sum_{k=1}^K \sqrt{C_k (\rho_{1,k}^2 - \rho_{1,k+1}^2)} \right)^2. \quad (10)$$

In reality, the correlation parameters and cost in the sample size estimation (7) are unknown and we will approximate these quantities by the sample estimates. Moreover, the sample sizes should be round to integers. We will use the ceiling of  $N_k$ , denoted as  $\lceil N_k \rceil$ , as the sample size estimation. This leads to

$$\sum_{k=1}^K C_k N_k \leq \sum_{k=1}^K C_k \lceil N_k \rceil < \sum_{k=1}^K C_k N_k + \sum_{k=1}^K C_k, \quad (11)$$

where the term  $\sum_{k=1}^K C_k$  results from the fact that  $N_k \leq \lceil N_k \rceil < N_k + 1$ . Let  $B_k = C_k (\rho_{1,k}^2 - \rho_{1,k+1}^2)$ ,  $k = 1, \dots, K$ . In general, the behavior of  $B_k$  depends on the product of  $C_k$  and  $(\rho_{1,k}^2 - \rho_{1,k+1}^2)$ .  $B_k$  can decay, grow and stay the same as  $k$  increases. The total sampling cost (10) can be rewritten as

$$\mathcal{W}^{\text{MF}} = \sum_{k=1}^K C_k N_k = \frac{\sigma_1^2}{\|\mathbb{E}(f_1)\|_Z^2 \epsilon^2} \left( \sum_{k=1}^K \sqrt{B_k} \right)^2.$$

Then condition (ii) indicates that

$$\frac{\sqrt{B_k}}{\sqrt{B_{k-1}}} > \frac{C_k}{C_{k-1}}, \quad k = 2, \dots, K. \quad (12)$$

In the asymptotic regime when  $K$  is large, the behavior of  $\sqrt{B_k}$  can decay, grow, or stay the same. If  $\sqrt{B_k}$  decays as  $k$  increases, the decay of  $\sqrt{B_k}$  is slower than the decay of  $C_k$ . This indicates that in all

scenarios, as  $k$  grows, we should expect the impact of  $\sum_{k=1}^K C_k$  in (11) is negligible as  $\sum_{k=1}^K C_k N_k$ . As a result, the sampling cost using the integer sample size should behave like (10). Thus, using the optimal sample estimation  $\lceil N_k \rceil$ , we obtain the total sampling cost of multifidelity estimator (value of the objective function) as

$$\mathcal{W}^{\text{MF}} = \frac{\sigma_1^2}{\|\mathbb{E}(f_1)\|_Z^2 \epsilon^2} \left( \sum_{k=1}^K \sqrt{C_k (\rho_{1,k}^2 - \rho_{1,k+1}^2)} \right)^2. \quad (13)$$

Our surrogate functions are built via stochastic collocation approach with the same sparse grid nodes on a family of spatial meshes  $\{\mathcal{T}_k\}$  of decreasing resolution, with a decreasing number of grid nodes  $\{M_k\}_{0 \leq k \leq K}$  that, we will assume, satisfy

$$M_k = s M_{k-1} \quad \text{for } s > 1. \quad (14)$$

We will also assume that the sample-wise discretization error is bounded as

$$\|u(\omega^{(i)}) - u_h(\omega^{(i)})\|_Z \leq c_m(\omega^{(i)}) M^{-\alpha}, \quad (15a)$$

where  $\alpha$  is the order of the sample-wise discretization error; the constants  $c_m(\omega^{(i)})$  depends only on the problem geometry and the particular realization  $\omega^{(i)}$ . We remark that in what follows, when describing the different approaches to surrogate-building, the associated values of the constants  $c_m$  and  $\alpha$  may vary from section to section.

To ensure that the discretization error falls below  $\theta\epsilon$ , we estimate the number of points on the finest grid  $M_K$  and the required spatial grid level  $K$  as

$$M_K = M_1 s^{-K} \geq \left( \frac{\theta\epsilon}{c_m} \right)^{-\frac{1}{\alpha}} \quad \text{and} \quad K = \left\lceil \frac{1}{\alpha} \log_s \left( \frac{c_m M_1^\alpha}{\theta\epsilon} \right) \right\rceil, \quad (16)$$

where  $\lceil \cdot \rceil$  denotes the ceiling function.

**Theorem 2.** *Let  $f_k$  be  $K$  models that satisfy the following conditions Suppose there exist positive constants  $\alpha, \beta, \gamma$  such that*

- (i)  $|\rho_{1,1}| > \dots > |\rho_{1,K}|$ ,
- (ii)  $\frac{C_{k-1}}{C_k} > \frac{\rho_{1,k-1}^2 - \rho_{1,k}^2}{\rho_{1,k}^2 - \rho_{1,k+1}^2}, \quad k = 2, \dots, K, \quad \rho_{1,K+1} = 0$ ,
- (iii)  $\|\mathbb{E}(u - u_k)\|_Z = \mathcal{O}(M_k^{-\alpha})$ ,
- (iv)  $C_k = \mathcal{O}(M_k^{-\gamma})$ ,
- (v)  $\rho_{1,k-1}^2 - \rho_{1,k}^2 = \mathcal{O}(M_k^\beta)$ .

*Then for  $0 < \epsilon$  there exist spatial grid level  $K$  and sample size  $N_\ell$  for which the multi-fidelity estimator  $A^{\text{MF}}$  satisfies*

$$\|\mathbb{E}(u) - A^{\text{MF}}\|_{L^2(\mathbf{W}, Z)} < \epsilon \|\mathbb{E}(u)\|_{L^2(\mathbf{W}, Z)},$$

*with total sampling work bounded as*

$$\mathcal{W}^{\text{MF}} \simeq$$

*Proof.*

□

The total sampling cost efficiency of the multifidelity Monte Carlo (MFMC) estimator relative to the standard Monte Carlo (MC) estimator is

$$\mu = \frac{\mathcal{W}_{\text{MFMC}}}{\mathcal{W}_{\text{MC}}} = \frac{1}{C_1} \left( \sum_{k=1}^K \sqrt{C_k (\rho_{1,k}^2 - \rho_{1,k+1}^2)} \right)^2$$

This ratio also quantifies the variance reduction achieved by the MFMC estimator given the same computational budget. The quantity of gamma is determined by the cost per sample for various models and the correlation parameters. The smaller the value of  $\gamma$ , the more effective the MFMC estimator.

### 3.1. Model selection

In order to maximize the capability of multifidelity Monte Carlo estimator compared to the Monte Carlo estimator, we need to select the models from the available set such that the ratio  $\gamma$  is as small as possible. Let  $S = \{1, \dots, K\}$  be the indices of  $K$  available models. We seek a subset  $S_1 = \{i_1, i_2, \dots, i_{K^*}\} \subseteq S$  ( $K^* \leq K$ ) of indices that minimizes the sampling cost of multifidelity Monte Carlo estimator. Note that  $S_1$  is non-empty and  $i_1 = 1$  since the high fidelity model must be included. This lead to the following optimization problem to determine the index set  $S_1$  for the selected models. We will follow the exhaustive algorithm in [1, Algorithm 1] for  $2^{K-1}$  subsets of  $S$  but with the target to minimize the sampling cost  $\mathcal{W}_{\text{MFMC}}$ . This algorithm gives the optimal  $K^*$  that satisfies (??).

$$\begin{aligned} \min_{k \in S} \quad & \mathcal{W}^{\text{MF}}, \\ \text{s.t.} \quad & |\rho_{1,1}| > \dots > |\rho_{1,K}|, \\ & \frac{C_{k-1}}{C_k} > \frac{\rho_{1,k-1}^2 - \rho_{1,k}^2}{\rho_{1,k}^2 - \rho_{1,k+1}^2}, \quad k = 2, \dots, K, \quad \rho_{1,K+1} = 0, \end{aligned}$$

---

#### Algorithm 1: Multi-fidelity Model Selection

---

**Input:**  $K$  models  $f_k$ , initial sample size  $N_0$ .

**Output:** Number of models  $K^*$  in  $\mathcal{S}^*$ , model  $f_i$  in  $\mathcal{S}^*$ , coefficients  $\rho_i$ ,  $\alpha_i$  and  $C_i$  for each model  $f_i$ .

---

- 1 Estimate  $\rho_{1,k}$  and  $C_k$  for each model  $f_k$  using  $N_0$  samples.
  - 2 Sort  $f_k$  by decreasing  $\rho_{1,k}$  to create  $\mathcal{S} = \{f_k\}_{k=1}^K$ .
  - 3 Initialize  $w^* = C_1$ ,  $\mathcal{S}^* = \{f_1\}$ . Let  $\widehat{\mathcal{S}}$  be all  $2^{K-1}$  ordered subsets of  $\mathcal{S}$ , each containing  $f_1$ .
  - 4 **for** each subset  $\widehat{\mathcal{S}}_j$  **do**
  - 5     **if** condition (ii) from Theorem 1 is satisfied **then**
  - 6         Compute the objective function value  $w$  using (??).
  - 7         **if**  $w < w^*$  **then**
  - 8             Update  $\mathcal{S}^* = \widehat{\mathcal{S}}_j$  and  $w^* = w$ .
  - 9      $j = j + 1$ .
  - 10 Compute  $\alpha_i$  for  $\mathcal{S}^*$ ,  $i = 2, \dots, K^*$  by (6).
-

---

**Algorithm 2:** Multifidelity Monte Carlo

---

**Input:** Models  $f_k$  in  $\mathcal{S}^*$ , parameters  $\rho_k$ ,  $\alpha_k$  and  $C_k$  for each  $f_k$  in  $\mathcal{S}^*$ , tolerance  $\epsilon$ .

**Output:** Sample sizes  $N_k$  for  $K^*$  models, expectation estimate  $A^{\text{MFMC}}$ .

---

- 1 Compute the sample size  $N_k$  for  $1 \leq k \leq K^*$  by (7).
  - 2 Evaluate  $f_1$  to obtain  $f_1(\omega^i)$  for  $i = 1, \dots, N_1$  and compute  $A_{1,N_1}^{\text{MC}}$  by (1). Store  $N_1$  samples.
  - 3 **for**  $k = 2, \dots, K^*$  **do**
  - 4     Evaluate  $f_k$  to obtain  $f_k(\omega^i)$  for  $i = 1, \dots, N_{k-1}$  and compute  $A_{k,N_{k-1}}^{\text{MC}}$  by (1).
  - 5     Evaluate  $f_k$  to obtain  $f_k(\omega^i)$  for  $i = 1, \dots, N_k - N_{k-1}$  and compute  $A_{k,N_k \setminus N_{k-1}}^{\text{MC}}$  by (1).
  - 6     Store  $N_{k-1}$  and  $N_k - N_{k-1}$  samples as  $N_k$  samples.
  - 7 Compute  $A^{\text{MF}}$  by (3).
- 

Dof	1934365	484080	120697	30449	8019	2685
Model $k$	$f_1$	$f_2$	$f_3$	$f_4$	$f_5$	$f_6$
$C_k$ direct solve	1.2029e+02	2.6478e+01	5.3710e+00	1.1269e+00	2.9300e-01	9.6419e-02
$C_k$ surrog evaluation(24 nodes)	1.2595e-01	2.9694e-02	9.1085e-03	3.0580e-03	1.1869e-03	2.5127e-04

Table 1: Number of spatial grid points  $M_\ell$  at increasing spatial grid level  $\ell = 0$  to 5.

## 4. Numerical experiments

## 5. Appendix

## References

- [1] B. Peherstorfer, K. Willcox, and M. Gunzburger. Optimal model management for multifidelity Monte Carlo estimation. SIAM J. Sci. Comput., 38(5):A3163–A3194, 2016.



Dof	1934365	484080	120697	30449	8019	2685
Model $k$	$f_1$	$f_2$	$f_3$	$f_4$	$f_5$	$f_6$
$\rho_{1,k}$ (24 nodes), ref l=3			0.9709	0.9702	0.9696	0.9482
$\sigma_k$			1.1466e-04	1.1842e-04	1.3128e-04	8.7847e-05
Covariance			1.0525e-04	1.3157e-04	1.2503e-04	1.2312e-04

Table 2: High fidelity model: finite element solution on mesh with 30449 grid nodes.  $\sigma_1 = 1.4026e - 04$ . The data are estimated using 500 samples.

Dof	1934365	484080	120697	30449	8019	2685
Model $k$	$f_1$	$f_2$	$f_3$	$f_4$	$f_5$	$f_6$
$\rho_{1,k}$ (24 nodes), ref l=3			0.9984	0.9976	0.9958	0.9802
$\sigma_k$			1.2016e-04	1.2405e-04	1.374e-04	9.3826e-05
Covariance			1.2457e-04	1.2646e-04	1.3283e-04	1.0807e-04

Table 3: High fidelity model: finite element solution on mesh with 120697 grid nodes.  $\sigma_1 = 1.2955e - 04$ . The data are estimated using 500 samples.

$\epsilon$	Level $\ell$					
	0	1	2	3	4	5
$8 \times 10^{-3}$			5			
$6 \times 10^{-3}$			7			
$4 \times 10^{-3}$				22		
$2 \times 10^{-3}$				83		
$10^{-3}$					322	
$8 \times 10^{-4}$					527	
$6 \times 10^{-4}$					869	
$4 \times 10^{-4}$					1980	
$2 \times 10^{-4}$						8000*

$\epsilon$	Level $\ell$					
	0	1	2	3	4	5
$8 \times 10^{-3}$	10	2	2			
$6 \times 10^{-3}$	12	3	2			
$4 \times 10^{-3}$	32	5	2	2		
$2 \times 10^{-3}$	152	26	4	2		
$10^{-3}$	691	109	18	4	2	
$8 \times 10^{-4}$	841	129	23	3	2	
$6 \times 10^{-4}$	1610	231	40	8	2	
$4 \times 10^{-4}$	3791	589	104	15	3	
$2 \times 10^{-4}$	15859	2344	375	62	13	2

$\epsilon$	Level $\ell$					
	0	1	2	3	4	5
$8 \times 10^{-3}$						
$6 \times 10^{-3}$						
$4 \times 10^{-3}$						
$2 \times 10^{-3}$						
$10^{-3}$						
$8 \times 10^{-4}$						
$6 \times 10^{-4}$						
$4 \times 10^{-4}$						
$2 \times 10^{-4}$						

Table 4: The optimal sample size estimation for MC-FE (left), uniform MLMC-FE (middle), and MFMC-FE (right). The simulations were conducted for a variety of choices of  $\epsilon$ . The computational cost associated with a tolerance of  $\epsilon = 2 \times 10^{-4}$  for Monte Carlo was prohibitive; the entry in the table for this tolerance (with an asterisk) is an estimate.