

# Multi-fidelity Monte Carlo for uncertainty quantification in the free boundary Grad-Shafranov equation

Matthias Heinkenschloss<sup>a</sup>, Jiaxing Liang<sup>b</sup>

<sup>a</sup>*Department of Computational Applied Mathematics & Operations Research, Rice University.*

<sup>b</sup>*Department of Computational Applied Mathematics & Operations Research, Rice University.*

---

## Abstract

We investigate the Grad-Shafranov free boundary problem in Tokamak fusion reactors under the influence of parameter uncertainties. Using both traditional 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.

**Keywords:** Multi-fidelity Monte Carlo Finite-Element, Parametric expectation,, Sparse Grid Stochastic Collocation, Uncertainty Quantification, Free Boundary Grad-Shafranov Problem.

**2020 MSC:**

---

## 1. Introduction

The pursuit of controlled nuclear fusion as a clean and virtually limitless energy source has spurred extensive research into the physics of magnetic confinement in fusion reactors. At the core of this effort lies the Grad-Shafranov free-boundary problem, which governs the equilibrium state of plasma in axially symmetric geometries, such as those found in Tokamaks. The governing equation encapsulates the intricate interplay between magnetic fields and plasma pressure, which determines critical confinement and stability properties essential for efficient plasma performance. However, the predictive accuracy of these models is significantly challenged by uncertainties arising from measurement limitations, model assumptions, and operational variabilities. Addressing these uncertainties effectively requires advanced computational frameworks capable of robust statistical analysis, enabling accurate predictions of the plasma equilibrium response under diverse scenarios and ensuring reliable assessments of reactor designs and operations.

The Monte Carlo (MC) method is widely used to quantify uncertainty and estimate statistical properties of input random variables. However, its application is computationally prohibitive due to its notoriously slow convergence rate, which requires a high-fidelity numerical approximation of the nonlinear partial differential equations for a large number of sample realizations. To mitigate this challenge, low-fidelity

---

*Email addresses:* heinken@rice.edu (Matthias Heinkenschloss), j1508@rice.edu (Jiaxing Liang)

models have been proposed as computationally efficient alternatives. For instance, in [3], a low-fidelity model constructed using stochastic collocation was used to facilitate Monte Carlo sampling. In [4], several low-fidelity models, based on coarse spatial grids, were developed as surrogate models for multilevel Monte Carlo (MLMC) sampling. Additionally, other work (e.g., [ ]) has explored generating low-fidelity models using stochastic collocation on coarse grids and integrating them with MLMC sampling. While these approaches reduce computational costs, the use of simplified surrogates introduces the risk of compromising accuracy. The trade-off between efficiency and accuracy is therefore critical and warrants careful investigation to balance computational savings with the reliability of results.

In this paper, we consider a multi-fidelity Monte Carlo (MFMC) method that uses the control variate approach to exploit correlations between the computationally expensive high-fidelity model and a series of appropriately chosen low-fidelity models. This method aims to further enhance sampling efficiency while preserving accuracy. Our goal is to demonstrate that this multi-fidelity approach can significantly accelerate sampling process while maintaining statistical fidelity, offering a practical and robust solution to the challenges posed by uncertainty quantification in plasma equilibrium modeling.

The paper is structured as follows. Section 2 introduces the Grad-Shafranov free boundary problem. Section 3 provides an overview of the sparse grid stochastic collocation technique used to construct low-fidelity models. Section 4 and Section 5 explore the surrogate-based Monte Carlo Finite Element and multilevel Monte Carlo Finite Element methods. Finally, Section 6 presents numerical experimental results that evaluate the efficiency and accuracy of these strategies.

## 2. The Grad-Shafranov free boundary problem with uncertainty

In magnetic confinement fusion reactors, atoms with light nuclei such as deuterium and tritium are injected into a chamber and heated to high temperature, ionizing to form a hot plasma. To maintain the conditions required for nuclear fusion, magnetic fields generated by external coils are used to confine the plasma and isolate it from the reactor wall. The confinement is governed by a balance between hydrostatic pressure and magnetic pressure generated from the external coils and internal plasma current. The equilibrium state, coupled with Maxwell's equations, is mathematically described by the Grad-Shafranov equation [6, 11, 16]. In axially symmetric tokamaks configurations, the problem simplifies to a two-dimensional formulation in the meridional  $r$ - $z$  plane using cylindrical coordinates  $(r, z, \varphi)$ . The equation determines the poloidal flux function  $u(r, z)$  from

$$-\nabla \cdot \left( \frac{1}{\mu r} \nabla u \right) = \begin{cases} r \frac{d}{du} p(u) + \frac{1}{2\mu r} \frac{d}{du} g^2(u) & \text{in } \Omega_p(u) \\ I_k / S_k & \text{in } \Omega_{C_k} \\ 0 & \text{elsewhere,} \end{cases} \quad (1a)$$

where  $\nabla$  and  $\nabla \cdot$  denote the Cartesian gradient and divergence operators in two dimensions; the magnetic permeability  $\mu$  is a constant  $\mu_0$  in vacuum but varies as  $\mu = \mu(|\nabla u|^2/r^2)$  within ferromagnetic materials; the regions  $\Omega_p$  and  $\Omega_{C_k}$  correspond to the plasma confinement zone and the meridian cross-sections of external coils, respectively, with the latter carrying currents  $I_k$  over areas  $S_k$ ; the source term in  $\Omega_p$  represents the toroidal component of the plasma current density, which depends non-linearly on the poloidal flux through the hydrostatic pressure  $p(u)$  and toroidal magnetic field  $g(u)$ . We adopt experimentally motivated formulations for  $p(u)$  and  $g(u)$  proposed in [12] as

$$\frac{d}{du} p(u) = \frac{\beta}{r_0} \left( 1 - u_N^{\alpha_1} \right)^{\alpha_2}, \quad \frac{1}{2} \frac{d}{du} g^2(u) = \mu_0 r_0 (1 - \beta) \left( 1 - u_N^{\alpha_1} \right)^{\alpha_2}, \quad (1b)$$

where  $u_N \in [0, 1]$  is the normalized flux, scaled by the difference between its values at the *magnetic axis* and the plasma boundary; the parameter  $r_0$  denotes the outer radius of the vacuum chamber,  $\alpha_1$  and  $\alpha_2$  dictate the sharpness of the current peaks near the magnetic axis, and  $\beta$ , the *poloidal beta*, measures the ratio of *plasma pressure* to *magnetic pressure*. The plasma boundary  $\Omega_p$ , delineated by the last closed streamline, is dependent on  $u$  and unknown a priori, introducing a *free boundary problem*. The nonlinear dependence of the poloidal flux, combined with the magnetic permeability and the nonlinearity of  $p(u)$  and  $g(u)$ , significantly contributes to the mathematical complexity of the Grad-Shafranov equation.

Let  $\Omega$  be a bounded, Lipschitz domain enclosing the confinement region  $\Omega_p$ , the external coils  $\Omega_{c_i}$ , and all other structural components of the reactor. The solution space  $Z$  for (1), as defined in [7], is

$$Z := \left\{ u : \Omega \rightarrow \mathbb{R} \mid \int_{\Omega} u^2 r dr dz < \infty; \int_{\Omega} \frac{|\nabla u|^2}{r} dr dz < \infty; u(0, z) = 0 \right\} \cap C^0(\bar{\Omega}), \quad (2)$$

ensuring finite energy and continuity over  $\bar{\Omega}$ . The corresponding inner product and energy norm are

$$\langle u, v \rangle_Z := \int_{\Omega} \frac{1}{r} \nabla u \cdot \nabla v dr dz, \quad \|u\|_Z := \left( \int_{\Omega} \frac{|\nabla u|^2}{r} dr dz \right)^{1/2}.$$

In practice, the equilibrium state of the plasma is vulnerable to uncertainties arising from sources such as measurement inaccuracies and operational fluctuations. These uncertainties manifest themselves as stochastic variations in the solution and its derived quantities, influencing the stability and performance of plasma confinement. This paper investigates the impact of parametric uncertainties on two key factors: the current intensities  $I_k$  of the external coils and the parameters governing the source term (1b). To model these uncertainties, we define a  $d$ -dimensional random variable  $\omega := (\omega_1, \dots, \omega_d)$ , where each component  $\omega_i$  is treated as an independent random variable characterized by a probability density function  $\pi_k$ . The baseline values for these parameters are denoted by the vector  $\tilde{\omega} = (\tilde{\omega}_1, \dots, \tilde{\omega}_d)$ . Assuming a uniform distribution for each parameter, centered around its corresponding baseline value  $\tilde{\omega}_i$  and perturbed by a relative magnitude  $\tau$ , the joint probability density function  $\pi(\tilde{\omega})$  and the associated  $d$ -dimensional parameter space  $W$  are

$$\pi(\omega) = \prod_{k=1}^d \pi_k(\omega_k) = \prod_{k=1}^d \frac{1}{2\tau|\tilde{\omega}_k|}, \quad W := \prod_{k=1}^d [\tilde{\omega}_k - \tau|\tilde{\omega}_k|, \tilde{\omega}_k + \tau|\tilde{\omega}_k|]. \quad (3)$$

Incorporating stochasticity into the current intensities and the source function requires solving for a solution operator  $u(\cdot, \omega) : W \rightarrow Z$ , which maps the realizations of the random variable  $\omega$  to the corresponding solutions of the free-boundary problem (1). To address the variability introduced by stochastic parameters, we adopt the *weighted Bochner space* framework, which provides a rigorous way to quantify combined spatial and parametric variability. For a stochastic function  $u : W \rightarrow Z$ , we consider function space  $L^2(W, Z)$ , consisting of functions with finite second moments, defined as

$$L^2(W, Z) = \{u : W \rightarrow Z \mid u \text{ is strongly measurable and } \int_W \|u(\cdot, \omega)\|_Z^2 \pi(\omega) d\omega < +\infty\},$$

and the associated norm is defined as

$$\|u\|_{L^2(W, Z)} = \left( \int_W \|u(\cdot, \omega)\|_Z^2 \pi(\omega) d\omega \right)^{1/2} = \left( \mathbb{E}[\|u(\cdot, \omega)\|_Z^2] \right)^{1/2}.$$

The goal of this paper is to explore the propagation of uncertainty and efficiently approximate the parametric expectation of  $u(\cdot, \omega)$

$$\mathbb{E}(u(\cdot, \omega)) = \int_W u(\cdot, \omega) \pi(\omega) d\omega, \quad (4)$$

and to compute some derived quantities from (4), such as the plasma boundary and features of solutions such as the locations of x-points.

To achieve computational efficiency, we use sampling techniques including Monte Carlo and multi-fidelity Monte Carlo methods. The high-fidelity model in the multi-fidelity Monte Carlo framework is implemented using the finite element-based solver FEEQS.m [8] – a lightweight Matlab implementation of the CEDRES++ code [5, 9], developed by Holger Heumann and collaborators; it implements a piecewise linear Finite Element discretization of a weak formulation of (1) and incorporates a globalized variant of Newton’s method to resolve its inherent nonlinearity. Hereafter, we refer to this solver as the *direct solver*. Moreover, the multi-fidelity framework relies on a sequence of low-fidelity models to mitigate the computational cost associated with the high-fidelity model. To construct these low-fidelity models for  $u(\cdot, \omega)$ , we use the sparse grid stochastic collocation method [1, 10, 13, 17].

### 3. Sparse grid stochastic collocation

This section provides a brief overview of the sparse grid stochastic collocation approach [1, 10, 13, 17], demonstrated using a generic solution  $u$ .

The method starts by defining a univariate set of  $m_i$  collocation nodes  $X^i = \{x_1^i, \dots, x_{m_i}^i\}$  over the interval  $[-1, 1]$ . Using these nodes, a univariate interpolation operator is built as  $I_{X^i}[u] := \sum_{j=1}^{m_i} u(\cdot, x_j^i) \phi_j$ , where the basis function  $\phi_k(x_j^i)$  is the Kronecker delta, evaluating to 1 when  $k = j$  and 0 otherwise. To extend this approach to high-dimensional parameter space, the method combines univariate nodes across multiple dimensions using a tensor product. Instead of using a full tensor grid – which grows exponentially with the number of dimensions  $d$  – the sparse grid framework selects fewer nodes  $m_i$  per dimension, drastically reducing computational costs while maintaining accuracy. The *sparse grid nodes* for a domain of dimension  $d$  and *level*  $q$  (where  $q \geq d$ ) are defined as

$$H(q, d) = \bigcup_{q-d+1 \leq |\mathbf{i}| \leq q} (X^{i_1} \times \cdots \times X^{i_d}) \in [-1, 1]^d,$$

where  $|\mathbf{i}| = i_1 + \dots + i_d$  specifies the refinement rule. These nodes form a sparse representation of the domain, capturing the essential features of  $u$  with fewer evaluations than a full grid.

For our problem, the collocation nodes are selected as the extrema of the Chebyshev polynomials [1, 2], with the  $j$ -th node calculated as  $x_j^i = -\cos(\pi(j-1)/(m_i-1))$  for  $j = 1, \dots, m_i$ . The number of nodes is determined so that  $m_1 = 1$  and  $m_i = 2^{i-1} + 1$  for  $i \geq 2$ . This specific choice ensures that the univariate nodal sets  $X^i$  exhibit a *nested* structure, satisfying  $X^i \subset X^{i+1}$ . Consequently, the multidimensional sparse grid nodes inherit this nesting property, resulting in

$$H(q, d) \subset H(q+1, d), \quad \text{and} \quad H(q, d) = \bigcup_{|\mathbf{i}|=q} (X^{i_1} \times \cdots \times X^{i_d}). \quad (5)$$

This hierarchical nesting allows for the reuse of function evaluations at coarser levels when constructing finer levels, further enhancing computational efficiency compared to non-nested grids. Interpolation over the hierarchical sparse grid nodes  $H(q, d)$  is performed using the *Smolyak quadrature formula*, which combines univariate interpolation operators across dimensions as

$$S_{q,d}[u] = \sum_{q+1 \leq |\mathbf{i}| \leq q+d} (-1)^{q+d-|\mathbf{i}|} \binom{d-1}{q+d-|\mathbf{i}|} \cdot (I_{X^{i_1}} \otimes \cdots \otimes I_{X^{i_d}})[u]. \quad (6)$$

This method efficiently computes high-dimensional interpolations by exploiting the sparsity and hierarchical structure of the grid, providing an effectively balances between computational efficiency and precision, making it particularly suited for solving deterministic parametrized problems involving high-dimensional uncertainties.

As discussed in [14, 18], consider the function  $u \in C^0(W, Z)$ , where the parameter space  $W$  and the solution space  $Z$  are defined in (3) and (2) respectively. Let the interval in the  $k$ -th dimension be defined as  $W_k = [\tilde{\omega}_k - \tau |\tilde{\omega}_k|, \tilde{\omega}_k + \tau |\tilde{\omega}_k|]$ . The complementary multi-dimensional parameter space that excludes the  $k$ -th dimension is

$$W_k^c = \prod_{i=1, i \neq k}^d W_i.$$

Now, for any fixed element  $\omega_k^c \in W_k^c$ , and for each  $\omega_k \in W_k$ , we assume the function  $u(\cdot, \omega_k, \omega_k^c) : W_k \rightarrow C^0(W_k^c; Z)$  admits an analytic extension  $u(\cdot, z, \omega_k^c)$  in the complex plane, specifically in the region

$$W_k^* := \{z \in \mathbb{C} : \text{dist}(z, W_k) \leq \iota_k \text{ for some } \iota_k > 0\},$$

where  $\iota_k$  denotes the proximity of the analytic extension to the real interval  $W_k$ . Under these assumptions, the interpolation error associated with the sparse grid method demonstrates an algebraic convergence rate

$$\|u - \mathcal{S}_{q,d}(u)\|_\infty = CP^{-\mu}, \quad (7)$$

where  $P$  denotes the sparse grid node count,  $C$  is a constant dependent on dimension  $d$  and analytic extension proximity to the interval  $W_k$ , and  $\mu$  is related to the dimension of parameter space and function's analytic extension in the complex plane.

#### 4. Monte Carlo estimator

To estimate the expectation in (4), Monte Carlo sampling is typically used. Let  $u_h(\cdot, \omega)$  denote a discrete approximation of  $u(\cdot, \omega)$  for the stochastic version of (1), obtained using a spatial discretization characterized by the mesh parameter  $h$  and consisting of  $M$  nodes. For simplicity, we write  $u$  and  $u_h$  as  $u(\cdot, \omega)$  and  $u_h(\cdot, \omega)$  respectively. 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  $\omega^{(1)}, \dots, \omega^{(N)}$

$$A_N^{\text{MC}} := \frac{1}{N} \sum_{i=1}^N u_h(\cdot, \omega^{(i)}), \quad (8)$$

where  $\mathbb{E}(A_N^{\text{MC}}) = \mathbb{E}(u_h)$ ,  $\mathbb{V}(A_N^{\text{MC}}) = \mathbb{V}(u_h)/N$  and  $\mathbb{V}(u) := \mathbb{E}(\|u - \mathbb{E}(u)\|_Z^2)$ . The central limit theorem ensures that as  $N$  approaches infinity, the estimator (8) converges in distribution to  $\mathbb{E}(u)$ . While this guarantees asymptotic behavior, the normalized *mean squared error* (nMSE) offers a quantitative measure of how much the estimator deviates from the true value in the mean-square sense. Normalized by  $\|\mathbb{E}(u)\|_Z^2$ , the nMSE is defined as

$$\mathcal{E}_{A_N^{\text{MC}}}^2 := \frac{\mathbb{E}[\|\mathbb{E}(u) - A_N^{\text{MC}}\|_Z^2]}{\|\mathbb{E}(u)\|_Z^2}.$$

The nMSE for the Monte Carlo estimator can be decomposed into two components: a bias term, reflecting the discretization of  $u$  using  $u_h$ , and the statistical error, arising from finite sampling. The decomposition is

expressed as

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

For the bias error, we assume that the sample-wise discretization error satisfies

$$\|u(\cdot, \omega^{(i)}) - u_h(\cdot, \omega^{(i)})\|_Z \leq C_m(\omega^{(i)})M^{-\alpha},$$

where  $C_m(\omega^{(i)})$  is a constant depending only on the geometry of the spatial domain and the particular realization  $\omega^{(i)}$ ,  $\alpha > 0$  is the convergence rate of spatial discretization, and  $M$  is the number of spatial grid nodes. Given a user-specified threshold  $\epsilon^2$  for the nMSE, we introduce a *splitting ratio*  $\theta \in (0, 1)$ , which allocates the total error budget  $\epsilon^2$  between bias and statistical components requiring that

$$\mathcal{E}_{\text{Bias}}^2 = \|u - u_h\|_{L^2(W, Z)} \leq C_m M^{-\alpha} = \theta \epsilon^2, \quad \mathcal{E}_{\text{Stat}}^2 = \frac{\sigma_1^2}{N\|\mathbb{E}(u)\|_Z^2} = (1 - \theta) \epsilon^2, \quad (9)$$

where  $C_m$  is independent of the sample and  $\sigma_1^2 = \mathbb{V}(u_h)$ . To satisfy the error constraints, the required number of spatial nodes  $M$  and sample size  $N$  must satisfy

$$M \geq \left(\frac{\theta \epsilon^2}{C_m}\right)^{-\frac{1}{\alpha}}, \quad N \geq \frac{\sigma_1^2}{\epsilon^2(1 - \theta)\|\mathbb{E}(u)\|_Z^2}, \quad (10)$$

Since  $M$  and  $N$  are integers, we round them up to the smallest integers that satisfy (10). Assume the average cost of evaluating  $u_h$  for a single sample is  $C$ . The total computational cost of estimating  $\mathbb{E}(u_h)$  with  $N$  samples is

$$\mathcal{W}_{\text{MC}} = CN = \frac{C\sigma_1^2}{\epsilon^2(1 - \theta)\|\mathbb{E}(u)\|_Z^2}.$$

## 5. Multi-fidelity Monte Carlo

The Monte Carlo finite element estimator is often computationally expensive due to the extensive sampling on a fixed finite element discretization. This can become prohibitively costly, particularly under stringent accuracy requirements. To mitigate this challenge, we turn to multi-fidelity Monte Carlo sampling [15], an approach that combines high-fidelity models – characterized by their accurate approximation of system behavior – and low-fidelity models, which are computationally efficient albeit less detailed. The core of this approach is rooted in the *control variate* technique, which reduces the variance of the estimator by exploiting the statistical correlations between high- and low-fidelity models. By optimally allocating computational effort across these models, the multi-fidelity framework significantly decreases the reliance on expensive high-fidelity evaluations, ultimately lowering computational costs. Importantly, this strategy preserves the accuracy and robustness of the resulting estimator, offering a practical and efficient alternative to traditional high-fidelity-only Monte Carlo sampling. Below, we provide a brief overview of this method.

To approximate (4), the multi-fidelity Monte Carlo combines the high-fidelity model  $u_{h,1}$  with a sequence of progressively less detailed but computationally cheaper low-fidelity models  $u_{h,k} : W \rightarrow Z$  for  $k = 2, \dots, K$ . The fidelity of the models decreases as  $k$  increases, with  $u_{h,1}$  being the most accurate. For

each model  $u_{h,k}$ , we define the variance and Pearson correlation coefficient between two models  $u_{h,k}$  and  $u_{h,j}$  as

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

where  $\text{Cov}(u_{h,k}, u_{h,j}) := \mathbb{E}[\langle u_{h,k} - \mathbb{E}(u_{h,k}), u_{h,j} - \mathbb{E}(u_{h,j}) \rangle_Z]$  and  $\rho_{k,k} = 1$ . The multi-fidelity Monte Carlo Finite-Element estimator  $A^{\text{MF}}$  incorporates weighted corrections derived from low-fidelity models into the Monte Carlo estimator of the high-fidelity model, defined as

$$A^{\text{MF}} := A_{1,N_1}^{\text{MC}} + \sum_{k=2}^K \alpha_k (\bar{A}_{k,N_k} - \bar{A}_{k,N_{k-1}}), \quad (11)$$

where  $A_{1,N_1}^{\text{MC}}$  is the Monte Carlo estimator for the high-fidelity model using  $N_1$  samples,  $\alpha_k$  are the weights for the correction terms, and  $\bar{A}_{k,N_k}$  denotes the sample average of the  $k$ -th model based on  $N_k$  samples. The sample average term  $\bar{A}_{k,N_{k-1}}$  in each correction reuses the first  $N_k$  samples from  $\bar{A}_{k,N_k}$ , requiring the condition  $N_{k-1} \leq N_k$  for  $k = 2, \dots, K$ . By partitioning the  $N_k$  samples into two disjoint sets, one of size  $N_{k-1}$  and the other of size  $N_k - N_{k-1}$ , we can reformulate the MFMC estimator (11) as

$$A^{\text{MF}} = A_{1,N_1}^{\text{MC}} + \sum_{k=2}^K \alpha_k \left( \frac{N_{k-1}}{N_k} - 1 \right) (A_{k,N_{k-1}}^{\text{MC}} - A_{k,N_k \setminus N_{k-1}}^{\text{MC}}), \quad (12)$$

The weights in the correction terms of this reformulation are now functions of the sample ratio between successive fidelity levels. This reformulation is crucial because it ensures that the samples used in the two components  $A_{k,N_{k-1}}^{\text{MC}}$  and  $A_{k,N_k \setminus N_{k-1}}^{\text{MC}}$  of the low-fidelity corrections are independent. The MFMC estimator can be compactly written as

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

where  $Y_k$  are defined as

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

It can be shown that  $\mathbb{E}(Y_k) = 0$  for  $k \geq 2$ , which implies  $\mathbb{E}(A^{\text{MF}}) = \mathbb{E}(u_{h,1})$ , ensuring the unbiasedness of the estimator. Since the samples in two disjoint components of  $Y_k$  are independent, the variances of  $Y_k$  can be computed as

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

While  $Y_k$  and  $Y_j$  for  $k \neq j, k, j = 2, \dots, K$  are dependent due to overlapping sample sets, it can be shown that  $Y_k$  and  $Y_j$  are uncorrelated. However,  $Y_k$  is correlated with  $Y_1$  since the samples used in  $Y_k$  reuse those of  $Y_1$ . Using the covariance relation from [15, Lemma 3.2], we express the covariance between  $Y_1$  and  $Y_k$  as

$$\text{Cov}(Y_1, Y_k) = - \left( \frac{1}{N_{k-1}} - \frac{1}{N_k} \right) \rho_{1,k} \sigma_1 \sigma_k.$$

The variance of the multi-fidelity Monte Carlo estimator is

$$\begin{aligned}
\mathbb{V}(A^{\text{MF}}) &= \mathbb{V}(Y_1) + \mathbb{V}\left(\sum_{k=2}^K \alpha_k Y_k\right) + 2 \operatorname{Cov}\left(Y_1, \sum_{k=2}^K \alpha_k Y_k\right), \\
&= \mathbb{V}(Y_1) + \sum_{k=2}^K \alpha_k^2 \mathbb{V}(Y_k) + 2 \sum_{2 \leq k < j \leq K} \alpha_k \alpha_j \operatorname{Cov}(Y_k, Y_j) + 2 \sum_{k=2}^K \alpha_k \operatorname{Cov}(Y_1, Y_k), \\
&= \frac{\sigma_1^2}{N_1} + \sum_{k=2}^K \left( \frac{1}{N_{k-1}} - \frac{1}{N_k} \right) (\alpha_k^2 \sigma_k^2 - 2 \alpha_k \rho_{1,k} \sigma_1 \sigma_k).
\end{aligned} \tag{13}$$

Using the variance of the MFMC estimator, the nMSE for the multi-fidelity Monte Carlo estimator can be split as bias and statistical errors as

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

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

$$M_L = M_0 s^{-L} \geq \left( \frac{\theta\epsilon}{C_m} \right)^{-\frac{1}{\alpha}} \quad \text{and} \quad L = \left\lceil \frac{1}{\alpha} \log_s \left( \frac{c_m M_0^\alpha}{\theta\epsilon} \right) \right\rceil. \tag{14}$$

Let  $C_k$  denote the cost of generating a single sample from the  $k$ -th low-fidelity model, and  $N_k$  represent the number of samples taken from that model. The total computational cost for the MFMC estimator is

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

Using the same splitting ratio  $\theta$  between bias and statistical errors as introduced in the discussion of the Monte Carlo Finite-Element approach, our next goal is to identify the optimal sample sizes  $N_k$  and weights  $\alpha_k$  for the MFMC estimator (12). Towards this goal, we formulate an optimization problem such that the total sampling cost,  $\mathcal{W}^{\text{MF}}$ , is minimized while ensuring that the normalized statistical error of the MFMC estimator is controlled to achieve the required accuracy  $\mathcal{E}_{\text{Stat}}^2 = (1-\theta)\epsilon^2$ . Moreover, the optimization must enforce hierarchical ordering of the sample sizes,  $N_{k-1} \leq N_k$  for  $k = 2, \dots, K$ , and non-negativity of all sample sizes. This leads to the constrained optimization problem

$$\begin{aligned}
&\min_{\substack{N_1, \dots, N_K \in \mathbb{R} \\ \alpha_2, \dots, \alpha_K \in \mathbb{R}}} \quad \sum_{k=1}^K C_k N_k, \\
&\text{subject to} \quad \mathbb{V}(A^{\text{MF}}) - \|\mathbb{E}(u)\|_Z^2 (1-\theta)\epsilon^2 = 0, \\
&\quad -N_1 \leq 0, \\
&\quad N_{k-1} - N_k \leq 0, \quad k = 2, \dots, K.
\end{aligned} \tag{15}$$

The solution to (15), which provides the real-valued optimal sample sizes and weights, is explicit presented in Theorem 1.

**Theorem 1.** Consider a set of  $K$  models,  $u_k$  for  $k = 1, \dots, K$ , where each model is characterized by the standard deviation of its output  $\sigma_k$ , the correlation coefficient  $\rho_{1,k}$  between the highest-fidelity model  $u_{h,1}$  and the  $k$ -th low-fidelity model, and the computational cost per sample  $C_k$ . Assume the following conditions are satisfied

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

Under these assumptions, the optimal weights  $\alpha_k^*$  and sample sizes  $N_k^*$ , for  $k = 1, \dots, K$  to the optimization problem (15) is

$$\alpha_k^* = \frac{\rho_{1,k}\sigma_1}{\sigma_k}, \quad (16)$$

$$N_k^* = \frac{\sigma_1^2}{\|\mathbb{E}(u)\|_Z^2(1-\theta)\epsilon^2} \sqrt{\frac{\rho_{1,k}^2 - \rho_{1,k+1}^2}{C_k}} \sum_{j=1}^K \sqrt{C_j(\rho_{1,j}^2 - \rho_{1,j+1}^2)}, \quad \rho_{1,K+1} = 0. \quad (17)$$

*Proof.* The optimization problem is approached with the method of Lagrange multipliers, where the auxiliary Lagrangian function  $L$  incorporates multipliers  $\lambda_0, \dots, \lambda_K$  to enforce the constraints on the variance and sample sizes

$$L = \sum_{k=1}^K C_k N_k + \lambda_0 \left( \frac{\sigma_1^2}{N_1} + \sum_{k=2}^K \left( \frac{1}{N_{k-1}} - \frac{1}{N_k} \right) (\alpha_k^2 \sigma_k^2 - 2\alpha_k \rho_{1,k} \sigma_1 \sigma_k) \right) - \lambda_1 N_1 + \sum_{k=2}^K \lambda_k (N_{k-1} - N_k),$$

To solve this, we apply the Karush-Kuhn-Tucker (KKT) conditions for (15), leading to a system of equations for the partial derivatives of the Lagrangian with respect to the optimization variables  $\alpha_k$  and  $N_k$ , in addition to primal feasibility, dual feasibility, and complementary slackness conditions

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

The partial derivatives of the Lagrangian with respect to  $\alpha_k$  and  $N_k$  are computed as

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

By solving the equation  $\partial L / \partial \alpha_k = 0$ , the optimal weights  $\alpha_k^*$  are determined as

$$\alpha_k^* = \frac{\rho_{1,k}\sigma_1}{\sigma_k}.$$

Substituting  $\alpha_k^*$  into  $\partial L / \partial N_k = 0$  yields a representation for  $C_k$  in terms of the sample sizes  $N_k$ , variance contributions, correlation coefficients, and the Lagrangian multipliers

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

As shown by [15], the global minimizer is achieved when the inequality constraints are inactive ( $\lambda_k = 0$ ,  $k = 1, \dots, K$ ) in the complementary slackness conditions, indicating that the sample sizes strictly increase with  $k$  ( $N_{k-1} < N_k$  for  $k = 2, \dots, K$ ). This results in the expression for the optimal sample sizes

$$N_k = \sigma_1 \sqrt{\lambda_0} \sqrt{\frac{\rho_{1,k}^2 - \rho_{1,k+1}^2}{C_k}}, \quad \text{with } \rho_{1,K+1} = 0 \quad k = 1, \dots, K. \quad (18)$$

Using the optimal coefficients  $\alpha_k^*$  and sample size estimations  $N_k$  in (18), the variance (13) of the multi-fidelity Monte Carlo estimator becomes

$$\mathbb{V}(A^{\text{MF}}) = \frac{\sigma_1}{\sqrt{\lambda_0}} \sum_{k=1}^K \sqrt{C_k (\rho_{1,k}^2 - \rho_{1,k+1}^2)}, \quad \text{with } \rho_{K+1} = 0.$$

To satisfy the equality constraint for the variance, we solve for  $\sqrt{\lambda_0}$

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

Substituting  $\sqrt{\lambda_0}$  into the equation for (18), we obtain an estimate for the optimal sample sizes

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

Note that by ensuring the condition (ii) is satisfied, we can guarantee that  $N_k$  increases strictly as  $k$  grows.  $\square$

The total sampling cost  $\mathcal{W}^{\text{MF}}$  is derived from the sample size estimates  $N_k^* \in \mathbb{R}$  obtained in Theorem 1 as

$$\mathcal{W}^{\text{MF}} = \sum_{k=1}^K C_k N_k^* = \frac{\sigma_1^2}{\|\mathbb{E}(u)\|_Z^2 (1-\theta)\epsilon^2} \left( \sum_{k=1}^K \sqrt{C_k (\rho_{1,k}^2 - \rho_{1,k+1}^2)} \right)^2. \quad (19)$$

However, since the correlation parameters and costs are unknown in practice, these quantities must be estimated based on sample data. Additionally, sample sizes should be rounded to the nearest integers for implementation. Thus, we use the ceiling function  $\lceil N_k \rceil$  to represent the sample size, which leads to the lower and upper bounds on the sampling cost

$$\sum_{k=1}^K C_k N_k \leq \sum_{k=1}^K C_k \lceil N_k \rceil < \sum_{k=1}^K C_k N_k + \sum_{k=1}^K C_k, \quad (20)$$

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 (19) can be rewritten as

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

Then condition (ii) indicates that

$$\frac{\sqrt{B_k}}{\sqrt{B_{k-1}}} > \frac{C_k}{C_{k-1}}, \quad k = 2, \dots, K. \quad (21)$$

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 (20) 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 (19). 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

$$\mathcal{W}^{\text{MF}} = \frac{\sigma_1^2}{\|\mathbb{E}(u)\|_Z^2 (1-\theta)\epsilon^2} \left( \sum_{k=1}^K \sqrt{C_k (\rho_{1,k}^2 - \rho_{1,k+1}^2)} \right)^2. \quad (22)$$

Our surrogate functions are built via a 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

$$M_k = sM_{k-1} \quad \text{for } s > 1. \quad (23)$$

**Theorem 2.** *Let  $u_{h,k}$  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)  $\|\mathbb{E}(u - u_{h,k})\|_Z = O(M_k^{-\alpha})$ ,
- (iv)  $C_k = O(M_k^{-\gamma})$ ,
- (v)  $\rho_{1,k-1}^2 - \rho_{1,k}^2 = O(M_k^\beta)$ .

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

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

with total sampling work bounded as

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

*Proof.* □

The total sampling cost efficiency of the multi-fidelity Monte Carlo estimator relative to the standard Monte Carlo estimator is

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

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.

### 5.1. Model selection

In order to maximize the capability of multi-fidelity 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 multi-fidelity Monte Carlo estimator. Note that  $S_1$  is nonempty and  $i_1 = 1$  since the high-fidelity model must be included. This leads to the following optimization problem to determine the index set  $S_1$  for the selected models. We will follow the exhaustive algorithm in [15, Algorithm 1] for  $2^{K-1}$  subsets of  $S$  but with the target to minimize the sampling cost  $\mathcal{W}^{\text{MF}}$ . This algorithm gives the optimal  $K^*$ .

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

---

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

---

**Input:**  $K$  models  $u_{h,k}$ , initial sample size  $N_0$ .

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

---

- 1 Estimate  $\rho_{1,k}$  and  $C_k$  for each model  $u_{h,k}$  using  $N_0$  samples.
- 2 Sort  $u_{h,k}$  by decreasing  $\rho_{1,k}$  to create  $\mathcal{S} = \{u_{h,k}\}_{k=1}^K$ .
- 3 Initialize  $w^* = C_1$ ,  $\mathcal{S}^* = \{u_{h,1}\}$ . Let  $\widehat{\mathcal{S}}$  be all  $2^{K-1}$  ordered subsets of  $\mathcal{S}$ , each containing  $u_{h,1}$ .
- 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 (22).
- 7     **if**  $w < w^*$  **then**
- 8       Update  $\mathcal{S}^* = \widehat{\mathcal{S}}_j$  and  $w^* = w$ .
- 9
- 10
- 11
- 12
- 13
- 14
- 15
- 16
- 17
- 18
- 19
- 20
- 21
- 22
- 23
- 24
- 25
- 26
- 27
- 28
- 29
- 30
- 31
- 32
- 33
- 34
- 35
- 36
- 37
- 38
- 39
- 40
- 41
- 42
- 43
- 44
- 45
- 46
- 47
- 48
- 49
- 50
- 51
- 52
- 53
- 54
- 55
- 56
- 57
- 58
- 59
- 60
- 61
- 62
- 63
- 64
- 65
- 66
- 67
- 68
- 69
- 70
- 71
- 72
- 73
- 74
- 75
- 76
- 77
- 78
- 79
- 80
- 81
- 82
- 83
- 84
- 85
- 86
- 87
- 88
- 89
- 90
- 91
- 92
- 93
- 94
- 95
- 96
- 97
- 98
- 99
- 100
- 101
- 102
- 103
- 104
- 105
- 106
- 107
- 108
- 109
- 110
- 111
- 112
- 113
- 114
- 115
- 116
- 117
- 118
- 119
- 120
- 121
- 122
- 123
- 124
- 125
- 126
- 127
- 128
- 129
- 130
- 131
- 132
- 133
- 134
- 135
- 136
- 137
- 138
- 139
- 140
- 141
- 142
- 143
- 144
- 145
- 146
- 147
- 148
- 149
- 150
- 151
- 152
- 153
- 154
- 155
- 156
- 157
- 158
- 159
- 160
- 161
- 162
- 163
- 164
- 165
- 166
- 167
- 168
- 169
- 170
- 171
- 172
- 173
- 174
- 175
- 176
- 177
- 178
- 179
- 180
- 181
- 182
- 183
- 184
- 185
- 186
- 187
- 188
- 189
- 190
- 191
- 192
- 193
- 194
- 195
- 196
- 197
- 198
- 199
- 200
- 201
- 202
- 203
- 204
- 205
- 206
- 207
- 208
- 209
- 210
- 211
- 212
- 213
- 214
- 215
- 216
- 217
- 218
- 219
- 220
- 221
- 222
- 223
- 224
- 225
- 226
- 227
- 228
- 229
- 230
- 231
- 232
- 233
- 234
- 235
- 236
- 237
- 238
- 239
- 240
- 241
- 242
- 243
- 244
- 245
- 246
- 247
- 248
- 249
- 250
- 251
- 252
- 253
- 254
- 255
- 256
- 257
- 258
- 259
- 260
- 261
- 262
- 263
- 264
- 265
- 266
- 267
- 268
- 269
- 270
- 271
- 272
- 273
- 274
- 275
- 276
- 277
- 278
- 279
- 280
- 281
- 282
- 283
- 284
- 285
- 286
- 287
- 288
- 289
- 290
- 291
- 292
- 293
- 294
- 295
- 296
- 297
- 298
- 299
- 300
- 301
- 302
- 303
- 304
- 305
- 306
- 307
- 308
- 309
- 310
- 311
- 312
- 313
- 314
- 315
- 316
- 317
- 318
- 319
- 320
- 321
- 322
- 323
- 324
- 325
- 326
- 327
- 328
- 329
- 330
- 331
- 332
- 333
- 334
- 335
- 336
- 337
- 338
- 339
- 340
- 341
- 342
- 343
- 344
- 345
- 346
- 347
- 348
- 349
- 350
- 351
- 352
- 353
- 354
- 355
- 356
- 357
- 358
- 359
- 360
- 361
- 362
- 363
- 364
- 365
- 366
- 367
- 368
- 369
- 370
- 371
- 372
- 373
- 374
- 375
- 376
- 377
- 378
- 379
- 380
- 381
- 382
- 383
- 384
- 385
- 386
- 387
- 388
- 389
- 390
- 391
- 392
- 393
- 394
- 395
- 396
- 397
- 398
- 399
- 400
- 401
- 402
- 403
- 404
- 405
- 406
- 407
- 408
- 409
- 410
- 411
- 412
- 413
- 414
- 415
- 416
- 417
- 418
- 419
- 420
- 421
- 422
- 423
- 424
- 425
- 426
- 427
- 428
- 429
- 430
- 431
- 432
- 433
- 434
- 435
- 436
- 437
- 438
- 439
- 440
- 441
- 442
- 443
- 444
- 445
- 446
- 447
- 448
- 449
- 450
- 451
- 452
- 453
- 454
- 455
- 456
- 457
- 458
- 459
- 460
- 461
- 462
- 463
- 464
- 465
- 466
- 467
- 468
- 469
- 470
- 471
- 472
- 473
- 474
- 475
- 476
- 477
- 478
- 479
- 480
- 481
- 482
- 483
- 484
- 485
- 486
- 487
- 488
- 489
- 490
- 491
- 492
- 493
- 494
- 495
- 496
- 497
- 498
- 499
- 500
- 501
- 502
- 503
- 504
- 505
- 506
- 507
- 508
- 509
- 510
- 511
- 512
- 513
- 514
- 515
- 516
- 517
- 518
- 519
- 520
- 521
- 522
- 523
- 524
- 525
- 526
- 527
- 528
- 529
- 530
- 531
- 532
- 533
- 534
- 535
- 536
- 537
- 538
- 539
- 540
- 541
- 542
- 543
- 544
- 545
- 546
- 547
- 548
- 549
- 550
- 551
- 552
- 553
- 554
- 555
- 556
- 557
- 558
- 559
- 560
- 561
- 562
- 563
- 564
- 565
- 566
- 567
- 568
- 569
- 570
- 571
- 572
- 573
- 574
- 575
- 576
- 577
- 578
- 579
- 580
- 581
- 582
- 583
- 584
- 585
- 586
- 587
- 588
- 589
- 590
- 591
- 592
- 593
- 594
- 595
- 596
- 597
- 598
- 599
- 600
- 601
- 602
- 603
- 604
- 605
- 606
- 607
- 608
- 609
- 610
- 611
- 612
- 613
- 614
- 615
- 616
- 617
- 618
- 619
- 620
- 621
- 622
- 623
- 624
- 625
- 626
- 627
- 628
- 629
- 630
- 631
- 632
- 633
- 634
- 635
- 636
- 637
- 638
- 639
- 640
- 641
- 642
- 643
- 644
- 645
- 646
- 647
- 648
- 649
- 650
- 651
- 652
- 653
- 654
- 655
- 656
- 657
- 658
- 659
- 660
- 661
- 662
- 663
- 664
- 665
- 666
- 667
- 668
- 669
- 670
- 671
- 672
- 673
- 674
- 675
- 676
- 677
- 678
- 679
- 680
- 681
- 682
- 683
- 684
- 685
- 686
- 687
- 688
- 689
- 690
- 691
- 692
- 693
- 694
- 695
- 696
- 697
- 698
- 699
- 700
- 701
- 702
- 703
- 704
- 705
- 706
- 707
- 708
- 709
- 710
- 711
- 712
- 713
- 714
- 715
- 716
- 717
- 718
- 719
- 720
- 721
- 722
- 723
- 724
- 725
- 726
- 727
- 728
- 729
- 730
- 731
- 732
- 733
- 734
- 735
- 736
- 737
- 738
- 739
- 740
- 741
- 742
- 743
- 744
- 745
- 746
- 747
- 748
- 749
- 750
- 751
- 752
- 753
- 754
- 755
- 756
- 757
- 758
- 759
- 760
- 761
- 762
- 763
- 764
- 765
- 766
- 767
- 768
- 769
- 770
- 771
- 772
- 773
- 774
- 775
- 776
- 777
- 778
- 779
- 780
- 781
- 782
- 783
- 784
- 785
- 786
- 787
- 788
- 789
- 790
- 791
- 792
- 793
- 794
- 795
- 796
- 797
- 798
- 799
- 800
- 801
- 802
- 803
- 804
- 805
- 806
- 807
- 808
- 809
- 810
- 811
- 812
- 813
- 814
- 815
- 816
- 817
- 818
- 819
- 820
- 821
- 822
- 823
- 824
- 825
- 826
- 827
- 828
- 829
- 830
- 831
- 832
- 833
- 834
- 835
- 836
- 837
- 838
- 839
- 840
- 841
- 842
- 843
- 844
- 845
- 846
- 847
- 848
- 849
- 850
- 851
- 852
- 853
- 854
- 855
- 856
- 857
- 858
- 859
- 860
- 861
- 862
- 863
- 864
- 865
- 866
- 867
- 868
- 869
- 870
- 871
- 872
- 873
- 874
- 875
- 876
- 877
- 878
- 879
- 880
- 881
- 882
- 883
- 884
- 885
- 886
- 887
- 888
- 889
- 890
- 891
- 892
- 893
- 894
- 895
- 896
- 897
- 898
- 899
- 900
- 901
- 902
- 903
- 904
- 905
- 906
- 907
- 908
- 909
- 910
- 911
- 912
- 913
- 914
- 915
- 916
- 917
- 918
- 919
- 920
- 921
- 922
- 923
- 924
- 925
- 926
- 927
- 928
- 929
- 930
- 931
- 932
- 933
- 934
- 935
- 936
- 937
- 938
- 939
- 940
- 941
- 942
- 943
- 944
- 945
- 946
- 947
- 948
- 949
- 950
- 951
- 952
- 953
- 954
- 955
- 956
- 957
- 958
- 959
- 960
- 961
- 962
- 963
- 964
- 965
- 966
- 967
- 968
- 969
- 970
- 971
- 972
- 973
- 974
- 975
- 976
- 977
- 978
- 979
- 980
- 981
- 982
- 983
- 984
- 985
- 986
- 987
- 988
- 989
- 990
- 991
- 992
- 993
- 994
- 995
- 996
- 997
- 998
<li

---

**Algorithm 2:** Multifidelity Monte Carlo

---

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

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

---

- 1 Compute the sample size  $N_k$  for  $1 \leq k \leq K^*$  by (17).
  - 2 Evaluate  $u_{h,1}$  to obtain  $u_{h,1}(\omega^i)$  for  $i = 1, \dots, N_1$  and compute  $A_{1,N_1}^{\text{MC}}$  by (8). Store  $N_1$  samples.
  - 3 **for**  $k = 2, \dots, K^*$  **do**
  - 4   Evaluate  $u_{h,k}$  to obtain  $u_{h,k}(\omega^i)$  for  $i = 1, \dots, N_{k-1}$  and compute  $A_{k,N_{k-1}}^{\text{MC}}$  by (8).
  - 5   Evaluate  $u_{h,k}$  to obtain  $u_{h,k}(\omega^i)$  for  $i = 1, \dots, N_k - N_{k-1}$  and compute  $A_{k,N_k \setminus N_{k-1}}^{\text{MC}}$  by (8).
  - 6   Store  $N_{k-1}$  and  $N_k - N_{k-1}$  samples as  $N_k$  samples.
  - 7 Compute  $A^{\text{MF}}$  by (12).
- 

Following [5], we consider an ITER geometry with 12 coils and a “baseline” vector of target current intensities  $\mathbf{I}$  given by

$$\begin{aligned} I_1 &= -1.4 \times 10^6 A, & I_2 &= -9.5 \times 10^6 A, & I_3 &= -2.0388 \times 10^7 A, & I_4 &= -2.0388 \times 10^7 A, \\ I_5 &= -9 \times 10^6 A, & I_6 &= 3.564 \times 10^6 A, & I_7 &= 5.469 \times 10^6 A, & I_8 &= -2.266 \times 10^6 A, \\ I_9 &= -6.426 \times 10^6 A, & I_{10} &= -4.82 \times 10^6 A, & I_{11} &= -7.504 \times 10^6 A, & I_{12} &= 1.724 \times 10^7 A. \end{aligned} \quad (24)$$

We will refer to these values as the *reference currents*. The profiles for  $p$  and  $g$  on the right hand side of (1) follow the form given in (1b) with the specific values  $r_0 = 6.2m$ ,  $\beta = 0.5978$ ,  $\alpha_1 = 2$ , and  $\alpha_2 = 1.395$ , and  $j_0 = 1.3655 \times 10^6 A/m^2$ . The reactor and coil array geometries follow the ones described in [? ]. In our experiments, we will take the vector of current intensities to be subject to uncertainty modeled as a uniformly distributed perturbation of magnitude  $\tau = 2\%$  centered around the reference values above.

The high-fidelity model is the finite element approximation of (1). The low-fidelity models are constructed using sparse grid stochastic collocation with 25 sparse grid nodes on a sequence of spatial mesh  $\{\mathcal{T}_k\}$ , with the number of spatial grid nodes shown in Table 1. We calculate the CPU time of 500 samples to estimate the cost per sample using both direct computation and surrogate evaluation. Experiments were conducted using MATLAB R2024a on the batch-scheduled HPC/HTC cluster, NOTSx, which operates on the Rice Big Research Data (BiRD) cloud infrastructure. The system comprises 298 dual-socket compute blades housed in HPE s6500, HPE Apollo 2000, and Dell PowerEdge C6400 chassis. All nodes are interconnected via a high-speed network featuring 10 or 25 Gigabit Ethernet.

We use Welford’s algorithm to update sample mean, variance and covariance using the formula

$$\begin{aligned} m_i &= m_{i-1} + \frac{u(\omega^i) - m_{i-1}}{i}, & s_i &= s_{i-1} + (u(\omega^i) - m_{i-1}, u(\omega^i) - m_i) \\ v_i &= v_{i-1} + (u^{\text{high}}(\omega^i) - m_{i-1}^{\text{high}}, u^{\text{low}}(\omega^i) - m_{i-1}^{\text{low}}) \end{aligned}$$

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.

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.

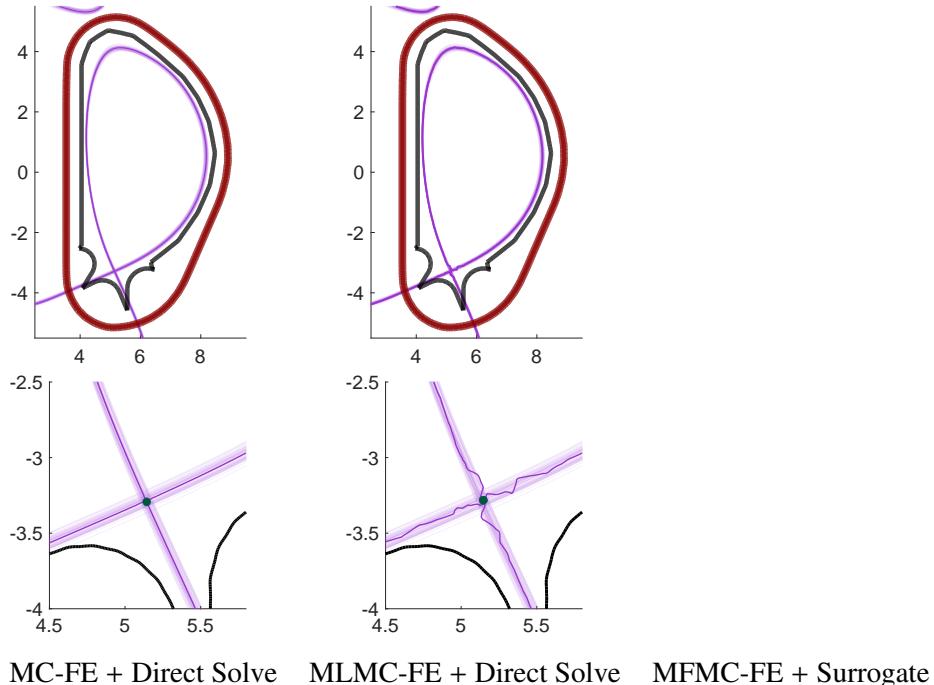


Figure 1: The overlayed plasma boundaries of 50 random realizations are displayed in the top row as violet curves (interpolated to the neighboring finer mesh). The solid violet line is the plasma boundary of the expected poloidal flux generated with tolerance  $\epsilon = 4 \times 10^{-4}$ . The inner and outer walls of the reactor are displayed in solid black and dark red respectively. The bottom row shows the regions close to the x-points in more detail. The dark green dots are the x-points of the expected solution. The columns from left to right correspond to simulations using the MC-FE approach with the direct solver and surrogate, MLMC-FE with direct solver and surrogate. All simulations were performed using the discretization level  $\ell = 5$  on geometry-conforming uniform meshes.

## 7. Acknowledgment

This work was supported in part by the Big-Data Private-Cloud Research Cyberinfrastructure MRI-award funded by NSF under grant CNS-1338099 and by Rice University's Center for Research Computing.

$\epsilon$	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.

## References

- [1] V. Barthelmann, E. Novak, and K. Ritter. High dimensional polynomial interpolation on sparse grids. *Advances in Computational Mathematics*, 12:273–288, 2000.
- [2] C. W. Clenshaw and A. R. Curtis. A method for numerical integration on an automatic computer. *Numer. Math.*, 2:197–205, 1960.
- [3] H. C. Elman, J. Liang, and T. Sánchez-Vizuet. Surrogate approximation of the Grad-Shafranov free boundary problem via stochastic collocation on sparse grids. *Journal of Computational Physics*, 448:110699, 20, 2022.
- [4] H. C. Elman, J. Liang, and T. Sánchez-Vizuet. Multilevel Monte Carlo methods for the Grad-Shafranov free boundary problem. *Computer Physics Communications*, 298:109099, 2024.
- [5] B. Faugeras and H. Heumann. FEM-BEM coupling methods for Tokamak plasma axisymmetric free-boundary equilibrium computations in unbounded domains. *Journal of Computational Physics*, 343:201 – 216, 2017.
- [6] H. Grad and H. Rubin. Hydromagnetic equilibria and force-free fields. In *Proc. Second international conference on the peaceful uses of atomic energy, Geneva*, volume 31,190, New York, Oct 1958. United Nations.
- [7] V. Grandgirard. Modélisation de l'équilibre d'un plasma de tokamak. Technical report, Université de Franche-Comté, 1999.
- [8] H. Heumann. FEEQS.M. <http://www-sop.inria.fr/members/Holger.Heumann/Software.html>.
- [9] H. Heumann, J. Blum, C. Boulbe, B. Faugeras, G. Selig, J.-M. Ané, S. Brémond, V. Grandgirard, P. Hertout, E. Nardon, and et al. Quasi-static free-boundary equilibrium of toroidal plasma with CEDRES: Computational methods and applications. *Journal of Plasma Physics*, 81(3):905810301, 2015.
- [10] A. Klimke and B. Wohlmuth. Algorithm 847: spinterp: piecewise multilinear hierarchical sparse grid interpolation in MATLAB. *Association for Computing Machinery. Transactions on Mathematical Software*, 31(4):561–579, 2005.
- [11] R. Lüst and A. Schlüter. Axialsymmetrische magnetohydrodynamische Gleichgewichtskonfigurationen. *Z. Naturf.*, 12a:850–854, 1957.
- [12] J. Luxon and B. Brown. Magnetic analysis of non-circular cross-section tokamaks. *Nuclear Fusion*, 22(6):813–821, jun 1982.
- [13] X. Ma and N. Zabaras. An adaptive hierarchical sparse grid collocation algorithm for the solution of stochastic differential equations. *Journal of Computational Physics*, 228(8):3084–3113, 2009.
- [14] F. Nobile, R. Tempone, and C. G. Webster. A sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 46(5):2309–2345, 2008.
- [15] B. Peherstorfer, K. Willcox, and M. Gunzburger. Optimal model management for multifidelity Monte Carlo estimation. *SIAM J. Sci. Comput.*, 38(5):A3163–A3194, 2016.
- [16] V. D. Shafranov. On magnetohydrodynamical equilibrium configurations. *Soviet Physics JETP*, 6:545–554, 1958.
- [17] S. A. Smolyak. Quadrature and interpolation formulae on tensor products of certain function classes. *Doklady Akademii Nauk SSSR*, 148:1042–1045, 1963.
- [18] A. L. Teckentrup, P. Jantsch, C. G. Webster, and M. Gunzburger. A multilevel stochastic collocation method for partial differential equations with random input data. *SIAM/ASA Journal on Uncertainty Quantification*, 3(1):1046–1074, 2015.