Multi-fidelity Monte Carlo for Uncertainty Quantification in the Free-Boundary Grad–Shafranov Equation

Jiaxing Liang

Department of Computational Applied Mathematics and Operations Research, Rice University

September 19, 2025

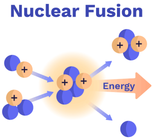
Outline

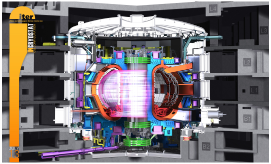
- Physical Model: Grad-Shafranov Equation
- Uncertainty Quantification
- Objectives & Approaches
- Multi-fidelity Monte Carlo
- Parameter Estimation
- Numerical Results

- ▷ **Physical Model: Grad-Shafranov Equation**
- ▷ Uncertainty Quantification
- ▷ Objectives & Approaches
- ▷ Multi-fidelity Monte Carlo
- ▷ Parameter Estimation
- ▷ Numerical Results

Fusion & Magnetic Confinement

- **Fusion reactor:** Focus on Tokamak – doughnut-shaped vacuum chamber.
- **Fusion mechanism:** Inject fuel gas (hydrogen isotopes) into the chamber $\rightarrow$ heated with microwave $\rightarrow$ form a plasma $\rightarrow$ particles overcome coulomb repulsion $\rightarrow$ fuse to produce energy.





- During fusion, the hot plasma tends to expand, to prevent it from contacting the reactor wall, it must be confined.
- **Magnetic confinement:** Confinement is achieved through magnetic fields generated by currents running through external coils surrounding the reactor.
- Our project focuses on studying the **impact of variations in current** on the physical properties of the confinement field. These **variations** may come from power supply, temperature fluctuations, and material impurities in the conduction wire.

Source: https://www.iter.org/mach/Tokamak

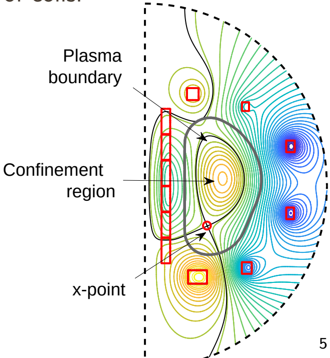
The Grad-Shafranov Free Boundary Problem

$$-\nabla \cdot \left(\frac{1}{\mu x} \nabla \psi \right) = \begin{cases} x \frac{d}{d\psi} p(\psi) + \frac{1}{2\,\mu x} \frac{d}{d\psi} g^2(\psi) & \text{in } \underline{\Omega_p(\psi)} \\ \boldsymbol{I_i}/S_i & \text{in } \Omega_{C_i} \\ 0 & \text{elsewhere} \end{cases}$$

$$\psi(0, y) = 0; \qquad \lim_{\|(x,y)\| \to \infty} \psi(x, y) = 0.$$

ψ: poloidal flux. $\mu(\psi)$: magnetic permeability. $p(\psi)$: hydrodynamic pressure. $g(\psi)$: toroidal field function. $\Omega_p(\psi)$: confinement region. I_i: current intensity. S_i: cross section area of the coils. Ω_{c_i}: locations of coils.

- **Non-linear** free boundary problem with the source of uncertainty arises from the **current intensities**.
- **Knowns:** $p(\psi), g(\psi), I_i, \Omega_{c_i}$, and μ. **Unknowns:** ψ and $\Omega_p(\psi)$.
- **Software:** MATLAB finite element package `FEEQS.M` – H. Heumann and his collaborators.



- Physical Model: Grad-Shafranov Equation
- # Uncertainty Quantification
- Objectives & Approaches
- Multi-fidelity Monte Carlo
- Parameter Estimation
- Numerical Results

Uncertain Coil Currents

- **Source of uncertainty:** Investigate the impact of uncertainties in the current intensities on the confinement properties of a plasma.
- **Baseline current intensities:**

$$
\begin{array}{lll}
I_1: -1.40\times 10^6\mathrm{A} & I_5: -9.00\times 10^6\mathrm{A} & I_9: -6.43\times 10^6\mathrm{A} \\
I_2: -9.50\times 10^6\mathrm{A} & I_6: 3.56\times 10^6\mathrm{A} & I_{10}: -4.82\times 10^6\mathrm{A} \\
I_3: -2.04\times 10^7\mathrm{A} & I_7: 5.47\times 10^6\mathrm{A} & I_{11}: -7.50\times 10^6\mathrm{A} \\
I_4: -2.04\times 10^7\mathrm{A} & I_8: -2.27\times 10^6\mathrm{A} & I_{12}: 1.72\times 10^7\mathrm{A}
\end{array}
$$

- **Uncertianty Quantification UQ:** Variations in the baseline current intensities. Each current in the array is considered to be a parameter, so we have 12-dimensional parameter space.
- **Currents uniformly distributed:**
  τ: perturbation level ($\tau = 1\%$ & 2%). I_k: current intensity in the k-th coil.
  - Joint density function: $\pi(\boldsymbol{\omega}) = \prod_{k=1}^{d} \pi_k(\omega_k) = \prod_{k=1}^{d} \frac{1}{2\tau |I_k|}$.
  - 12-d parameter space: $W := \prod_{k=1}^{d} [I_k - \tau|I_k|, I_k + \tau|I_k|]$.

▷ Physical Model: Grad-Shafranov Equation

▷ Uncertainty Quantification

▷ Objectives & Approaches

▷ Multi-fidelity Monte Carlo

▷ Parameter Estimation

▷ Numerical Results

Objectives & approaches

- **Uncertainty Quantification in Plasma Equilibria**
  Quantify the **impact of variations in current intensity** on determining
  - **Expectation:** $\mathbb{E}\left[\psi(x, y, \boldsymbol{\omega})\right] = \int_W \psi(x, y, \boldsymbol{\omega})\pi(\boldsymbol{\omega})d\boldsymbol{\omega}$.
  - **Key plasma features:** presence and locations of x-points, contact points, strike points & separatrices.
  - **Geometry parameters:** elongation, triangularity etc.

- **Enhancing Sampling Efficiency**
  Use **sampling methods** to approximate the expectation efficiently
  - Multi-fidelity Monte Carlo
    Blending high- and low-fidelity models, where high-fidelity models provide accuracy and low-fidelity surrogates reduce computational cost.

- **New Contributions**
  - Optimization-based formulation for MFMC sample allocation.
  - Dynamic pilot sampling strategy for correlation estimation.

- Physical Model: Grad-Shafranov Equation
- Uncertainty Quantification
- Objectives & Approaches
- **Multi-fidelity Monte Carlo**
- Parameter Estimation
- Numerical Results

Monte Carlo Sampling

- **Monte Carlo sampling (MC)**
  Estimate $\mathbb{E}[u]$ by a finite number of i.i.d. sample average.

$$A_N^{\mathrm{MC}} := \frac{1}{N}\sum_{i=1}^{N} u_h^{(i)}.$$

- **MC estimator A^{MC} for $\mathbb{E}[u]$**
  - Unbiased $\mathbb{E}[A_N^{\mathrm{MC}}] = \mathbb{E}[u_h]$.
  - Variance decays slowly: $\mathcal{O}(N^{-1/2})$. $\mathbb{V}[A_N^{\mathrm{MC}}] = \mathbb{V}[u_h]/N$, where $\mathbb{V}[u_h] = \frac{1}{N-1}\left(\sum_{i=1}^{N}\left\|u_h^{(i)}\right\|_U^2 - \frac{1}{N}\left\|\sum_{i=1}^{N} u_h^{(i)}\right\|_U^2\right).$

Multi-fidelity Monte Carlo

▷ Multi-fidelity Monte Carlo (MFMC)

A variance reduction technique that efficiently estimates statistical quantities by using low fidelity (LF) models to reduce variance of high fidelity (HF) model. HF models are used to maintaining accuracy.

B. Peherstorfer, K. Willcox, and M. Gunzburger. Optimal Model Management for Multifidelity Monte Carlo Estimation. SIAM J. SCI. COMPUT, Vol. 38, No. 5, pp. A3163–A3194.

▷ MFMC estimator A^{MF} for $\mathbb{E}[u]$

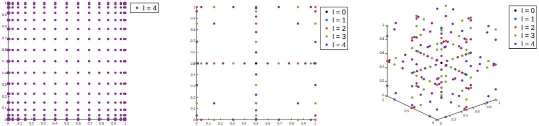
$$A^{\mathrm{MF}} = A^{\mathrm{MC}}_{1,N_1} + \sum_{k=2}^{K} \alpha_k \left(\overline{A}_{k,N_k} - \overline{A}_{k,N_{k-1}} \right)$$

- $\overline{A}_{k,N_{k-1}}$ reuses the first N_{k-1} samples from $\overline{A}_{k,N_k}$, $N_k \geq N_{k-1}$.
- Unbiased $\mathbb{E}[A^{\mathrm{MF}}] = \mathbb{E}[u_h]$.
- Let $\Delta_k = \rho^2_{1,k} - \rho^2_{1,k+1}$ for $k = 1, \ldots, K$, with $\rho_{1,K+1} = 0$, $\mathbb{V}[A^{\mathrm{MF}}] = \sigma_1^2 \sum_{k=1}^{K} \frac{\Delta_k}{N_k}$.
- Requires estimating optimal weights α_k and sample sizes N_k.
- Efficiency depends on correlation $\rho_{1,k}$ between HF and LF.

Multi-fidelity Monte Carlo

- **HF and LF models**
  - **High fidelity models:** Expensive but accurate
    (e.g., finite element solution of Grad-Shafranov equation).
  - **Low fidelity models:** Cheap but less accurate, correlated with HF as control variates.
    (e.g., sparse grid stochastic collocation + coarser meshes).

- Build surrogate via **Sparse Grid Stochastic Collocation**.
  - **Idea:** Select a special set of current $\{\boldsymbol{\omega}^{(k)}\}_{1\le k\le n_{sc}}$ in the parameter space. For each $\boldsymbol{\omega}^{(k)}$, evaluate the nonlinear solver to yield the realization $\{\psi_h^{(k)}\}$. An interpolation (surrogate) is then constructed to mimic the original nonlinear function with
$$\widehat{\psi_h}(\cdot,\boldsymbol{\omega}) = \sum_k \psi_h^{(k)}(\cdot) L_{\boldsymbol{\omega}^{(k)}}(\boldsymbol{\omega}).$$
  - **Sparse grids**:



Left to right: Full tensor grid 2d, level 4. Chebyshev sparse grids for 2d and 3d from level 0 to level 4.

Theorem: Interpolation error bound in L_∞ norm
$$\|f - \mathscr{S}(f)\|_\infty = \mathscr{O}\left(N^{-k}\cdot|\log N|^{(k+2)(d-1)+1}\right)$$
N: # of collocation nodes, d: dimension, k: measure of smoothness of ψ wrt $\boldsymbol{\omega}$.

  - **Software:** MATLAB SPINTERP package – A.Klimke.

    SPINTERP, piecewise multilinear hierarchical sparse grid interpolation, http://people.math.sc.edu/burkardt/msrc/spinterp/spinterp.html, 2007.

Multi-fidelity Monte Carlo

Let $\Delta_k = \rho_{1,k}^2 - \rho_{1,k+1}^2$ for $k = 1, \ldots, K$, with $\rho_{1,K+1} = 0$.

- **The original optimization problem:**

$$
\begin{aligned}
\min_{\alpha_2,\ldots,\alpha_K \in \mathbb{R}} \quad & \mathbb{V}\left[A^{\text{MF}}\right] = \sigma_1^2 \textstyle\sum_{k=1}^K \frac{\Delta_k}{N_k}, \\
\text{subject to} \quad & \sum_{k=1}^K C_k N_k = p, \\
& -N_1 \leq 0, \quad N_{k-1} - N_k \leq 0, \quad k = 2 \ldots, K, \\
& N_1, \ldots, N_K \in \mathbb{R}.
\end{aligned}
$$

- **Ours:**

$$
\begin{aligned}
\min_{\alpha_2,\ldots,\alpha_K \in \mathbb{R}} \quad & \sum_{k=1}^K C_k N_k, \\
\text{subject to} \quad & \mathbb{V}\left[A^{\text{MF}}\right] = \epsilon_{\text{tar}}^2, \\
& -N_1 \leq 0, \quad N_{k-1} - N_k \leq 0, \quad k = 2 \ldots, K, \\
& N_1, \ldots, N_K \in \mathbb{R}.
\end{aligned}
$$

MFMC Optimal Sample Allocation

Theorem (Original problem – Simplified sample size)

Consider an ensemble of K models $\{u_{h,k}\}_{k=1}^{K}$ each characterized by the standard deviation σ_k of its output, the correlation coefficient $\rho_{1,k}$ with the highest-fidelity model $u_{h,1}$, and the computational cost per sample evaluation C_k. Define $\Delta_k = \rho_{1,k}^2 - \rho_{1,k+1}^2$ for $k = 1, \ldots, K$, with $\rho_{1,K+1} = 0$. Assume the following conditions hold

$$(i) \ \ \textit{Correlation monotonicity}: \quad |\rho_{1,1}| > \cdots > |\rho_{1,K}|,$$

$$(ii) \ \ \textit{Cost-correlation ratio}: \qquad \frac{\Delta_k}{C_k} > \frac{\Delta_{k-1}}{C_{k-1}}, \quad k = 2, \ldots, K.$$

Under these assumptions, the solution to the original optimization problem yields optimal weights $\alpha_k^* = \frac{\rho_{1,k}\sigma_1}{\sigma_k}$. define an intermediate vector $\boldsymbol{r}^* = [r_1, \ldots, r_k]^T$, then $\boldsymbol{r}^*$ and the sample sizes N_k^*

$$r_k^* = \sqrt{\frac{C_1 \Delta_k}{C_k \Delta_1}}, \quad N_1^* = \frac{p}{\sum_{k=1}^{K} C_k r_k^*}, \quad N_k^* = N_1^* r_k^*, \quad \Rightarrow \quad N_k^* = \sqrt{\frac{\Delta_k}{C_k}} \frac{p}{\sum_{j=1}^{K} \sqrt{C_j \Delta_j}}.$$

The resulting MFMC estimator achieves a variance of

$$\mathbb{V}\left[A^{MF}\right] = \frac{\sigma_1^2}{p} \left(\sum_{k=1}^{K} \sqrt{C_k \Delta_k} \right)^2 .$$

MFMC Optimal Sample Allocation

Theorem (Our problem)

Consider an ensemble of K models $\{u_{h,k}\}_{k=1}^{K}$ each characterized by the standard deviation σ_k of its output, the correlation coefficient $\rho_{1,k}$ with the highest-fidelity model $u_{h,1}$, and the computational cost per sample evaluation C_k. Define $\Delta_k = \rho_{1,k}^2 - \rho_{1,k+1}^2$ for $k = 1, \ldots, K$, with $\rho_{1,K+1} = 0$. Assume the following conditions hold

$$\begin{aligned} &(i) \;\; \textit{Correlation monotonicity}: && |\rho_{1,1}| > \cdots > |\rho_{1,K}|, \\ &(ii) \;\; \textit{Cost-correlation ratio}: && \frac{\Delta_k}{C_k} > \frac{\Delta_{k-1}}{C_{k-1}}, \quad k = 2, \ldots, K. \end{aligned}$$

Under these assumptions, the solution to the original optimization problem yields optimal weights $\alpha_k^* = \frac{\rho_{1,k}\sigma_1}{\sigma_k}$. The sample sizes N_k^* and total computational cost are

$$N_k^* = \frac{\sigma_1^2}{\epsilon_{tar}^2}\sqrt{\frac{\Delta_k}{C_k}}\sum_{j=1}^{K}\sqrt{C_j\Delta_j}, \qquad \mathcal{W}^{MF} = \frac{\sigma_1^2}{\epsilon_{tar}^2}\left(\sum_{k=1}^{K}\sqrt{C_k\Delta_k}\right)^2.$$

- Convex problem $\rightarrow$ unique solution.
- Sample sizes N_k^* increase monotonically with fidelity level.
- Global optimality established via KKT and block-structure analysis.

Rounding & Example

- Closed-form N_k^* are real $\rightarrow$ need integer rounding.
- Parameters

Model index	1	2	3	4	5
Correlation coeff $\rho_{1,k}$	1	9.9977e-01	9.9925e-01	9.9728e-01	9.8390e-01
Standard deviation σ_k	1.0840e-02	1.0838e-02	1.1001e-02	1.1549e-02	9.5720e-03
Cost	73	7.0318e-03	1.4018e-03	5.0613e-04	2.6803e-04

- Sample size

	Sample size	Total cost p	$\mathbb{V}\left[A^{\mathrm{MF}}\right]$
Real valued	[8.8130e-01, 1.3499e+02, 5.8811e+02, 2.5409e+03, 2.1100e+04]	blue!40 73.05	yellow! 6.9633e-08
Real valued, floor	[0, 134, 588, 2540, 21100]	blue!20 8.7075	∞
Modified MFMC	[1, 1, 3, 13, 114]	blue!2073.0484	1.5677e-06
Integer program	[1, 2, 3, 11, 97]	blue!2073.0498	1.7250e-06
Real valued, ceil	[1, 135, 589, 2541, 21101]	81.7167	yellow!306.2353e-08

- Ours: Simple rounding technique preserves variance constraint, adds negligible overhead to computational cost.

Modified MFMC

Algorithm Modified MFMC

Input: Parameters $\rho_{1,k}$ and C_k for each $\widehat{u}_{h,k}$, total cost p. Ensure the K models $\widehat{u}_{h,k}$ after the model selection satisfies $p \geq \sum_i C_i$.

Output: Sample sizes N_k for K models.

Set the weights $\alpha_k = \frac{\rho_{1,k}\sigma_1}{\sigma_k}, \quad 1 \leq k \leq K$.

while there is $1 \leq j \leq K - 1$ such that $N_j < 1$ **do**

  Set j equal to the first such index.

  $N_j \leftarrow 1$.

  Set the k-st component of N_k,

$$N_k = \sqrt{\frac{\Delta_k}{C_k}} \frac{p - \sum_{i=1}^{j} C_i}{\sum_{i=j+1}^{K} \sqrt{C_i \Delta_i}}, \qquad j + 1 \leq k \leq K.$$

end while

Summary for MFMC formulations

	Original	Modified MFMC (due to decimal sample size btw $0 \sim 1$)	Ours
$\mathbb{V}\left[A^{\mathrm{MF}}\right]$	$\frac{\sigma_1^2}{p}\left(\sum_{k=1}^{K}\sqrt{C_k\Delta_k}\right)^2$	$\sigma_1^2\sum_{k=1}^{K}\frac{\Delta_k}{N_k}$	$\epsilon_{\mathrm{tar}}^2$
$\mathcal{W}^{\mathrm{MF}}$	p	$\leq p$	$\frac{\sigma_1^2}{\epsilon_{\mathrm{tar}}^2}\left(\sum_{k=1}^{K}\sqrt{C_k\Delta_k}\right)^2$
N_k	$\sqrt{\frac{\Delta_k}{C_k}}\frac{p}{\sum_{j=1}^{K}\sqrt{C_j\Delta_j}}$	While there is $1 \leq j \leq K-1$ such that $N_j < 1$, set $N_j \leftarrow 1$ and update $\sqrt{\frac{\Delta_k}{C_k}}\frac{p-\sum_{i=1}^{j}C_i}{\sum_{i=j+1}^{K}\sqrt{C_i\Delta_i}}, \quad j+1 \leq k \leq K.$	$\frac{\sigma_1^2}{\epsilon_{\mathrm{tar}}^2}\sqrt{\frac{\Delta_k}{C_k}}\sum_{j=1}^{K}\sqrt{C_j\Delta_j}$
Rounding	Floor	Floor	Ceil

- ▷ Physical Model: Grad-Shafranov Equation
- ▷ Uncertainty Quantification
- ▷ Objectives & Approaches
- ▷ Multi-fidelity Monte Carlo
- ▷ **Parameter Estimation**
- ▷ Numerical Results

Parameter Estimation

- Parameters to estimate: C_k, σ_k, $\rho_{1,k}$.
- Sample standard deviation and cost: unbiased, converge quickly.
- **Sample correlation coefficient**: biased, skewed distribution mean $|\rho_{1,k}| \to 1$. Requires careful pilot sampling to avoid efficiency loss.

$$\widehat{\rho}_{1,k} = \frac{\sum_{i=1}^{Q} \left\langle u_1^{(i)} - A_{1,Q}^{\mathrm{MC}},\ u_k^{(i)} - A_{k,Q}^{\mathrm{MC}} \right\rangle_U}{\sqrt{\sum_{i=1}^{Q} \left\| u_1^{(i)} - A_{1,Q}^{\mathrm{MC}} \right\|_U^2} \sqrt{\sum_{i=1}^{Q} \left\| u_k^{(i)} - A_{k,Q}^{\mathrm{MC}} \right\|_U^2}},$$

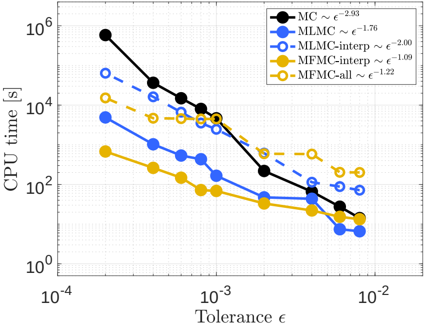
  Q: number of pilot samples,
  $u_k^{(i)}$: i-th realization of the k-th model,
- **Dynamic Pilot Sampling Strategy**
  - Sequentially increase pilot sample size until the relative cost efficiency update falls below certain threshold using sensitivity of cost efficiency.
  - Avoids unnecessary large pilot runs while ensuring reliability. Achieves accuracy comparable to fixed-Q but at lower cost.

- Physical Model: Grad-Shafranov Equation
- Uncertainty Quantification
- Objectives & Approaches
- Multi-fidelity Monte Carlo
- Parameter Estimation

Numerical Results

Numerical results

▷ **Efficiency:** Multi-fidelity Monte Carlo reduces the computational cost of Monte Carlo with direct non-linear solve by up to $\mathbf{10^3}$ ($\sim \epsilon^{-1}$). However, when accounting for **upfront costs** such as surrogate construction and parameter estimation, the overall gain is limited to a factor of **50**.



Theorem (Asymptotic Cost for MFMC-SC Estimator)

Consider a hierarchy of models $\{u_{\ell(k),k}\}_{k=1}^{K_c}$, where $u_{L,1}$ denotes the high-fidelity model on the finest mesh $\mathcal{T}_L$ with M_L degrees of freedom, and the low-fidelity models $\{u_{\ell(k),k}\}_{k=2}^{K_c}$ are constructed via sparse grid stochastic collocation on coarser meshes $\mathcal{T}_{\ell(k)}$ with $\ell(k) = L+1-k$. Let $\mathcal{I} = \{i_k \mid k = 1,\ldots,K\}$ denote the index set of selected models, ordered by decreasing correlation ρ_{1,i_k} and associated costs C_{i_k}. For the family of high-fidelity models $\{u_{L,1}\}_{L\leq L_{\max}}$ with M_L degrees of freedom, assume there exist positive constants α, β, γ such that

$$(i) \quad \|\mathbb{E}[u] - \mathbb{E}[u_{L,1}]\|_U \simeq M_L^{-\alpha}, \qquad (ii) \quad 1-\rho_{1,i_2}^2 \simeq M_L^{-\beta}, \qquad (iii) \quad C_1 \simeq M_L^{\gamma},$$

For each low-fidelity model $u_{\ell(i_k),i_k}$ with $k = 2,\ldots,K$, assume there exist positive constants β_1, γ_1 such that the correlation decay and cost growth follow

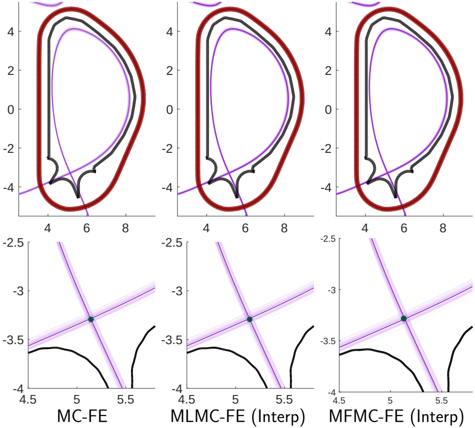
$$(iv) \quad \rho_{1,i_k}^2 - \rho_{1,i_{k+1}}^2 \simeq M_{\ell(i_k)}^{-\beta_1}, \qquad (v) \quad C_{i_k} \simeq M_{\ell(i_k)}^{\gamma_1},$$

with $\rho_{1,i_{K^*+1}}^2 := 0$. Then, for any tolerance $\epsilon \in (0, e^{-1})$, there exists a finest mesh level $L \leq L_{\max}$ and sample allocation $\{N_{i_k}\}_{i_k\in\mathcal{I}^*}$ such that the multifidelity estimator A^{MF} achieves an bounded nMSE with ϵ^2. In the regime where the high-fidelity contribution dominates, the total cost reduces to

$$\mathcal{W}^{\mathrm{MF}} \lesssim \epsilon^{-2-\frac{\gamma-\beta}{\alpha}}.$$

Numerical results

▷ **Accuracy:** Plasma boundary and geometric descriptors from MFMC sampling exhibit behavior consistent with those from MC and MLMC when interpolated onto a common mesh.



Thank You |¬_¬|