# Multifidelity write-up


**Abstract**

We derived the explicit sample size estimation in terms of the desired accuracy requirement. We make an improvement on the efficiency gain estimate and model selection approach compared to [1].


## 1. Model problem

We have a high fidelity model denoted as $f_1 : \Omega \to Z$ we desire, and several low fidelity models (surrogates) $f_k$ for $k \in \mathbb{N}$. Our objective is to approximate

$$\mathbb{E}\left(f_1(\omega)\right).$$

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

$$\sigma_k^2 = \mathbb{V}\left(f_k(\omega)\right), \qquad \rho_{k,j} = \frac{\mathrm{Cov}\left(f_k(\omega), f_j(\omega)\right)}{\sigma_k \sigma_j}, \quad k, j = 1, \ldots, K,$$

where $\mathbb{V}(f) := \mathbb{E}\left(\|f - \mathbb{E}(f)\|_Z^2\right)$. Note that $\rho_{k,k} = 1$.

## 2. Monte Carlo estimator

The Monte Carlo estimator for the expectation of each $f_k$ is defined as the sample mean of $N_k$ i.i.d realizations $\omega_1, \ldots, \omega_N$

$$A^{\mathrm{MC}}_{k,N_k} := \frac{1}{N_k} \sum_{i=1}^{N_k} f_k(\omega_i), \quad \forall k = 1, \ldots, K, \tag{1}$$

where $\mathbb{E}(A^{\mathrm{MC}}_{k,N_k}) = \mathbb{E}(f_k)$, $\mathbb{V}(A^{\mathrm{MC}}_{k,N_k}) = \mathbb{V}(f_k)/N_k$. Let $C_k$ denote the average evaluation cost per sample for $f_k$, then the total sampling cost for each Monte Carlo estimator $A^{\mathrm{MC}}_{k,N_k}$ is

$$\mathcal{W}^k_{\mathrm{MC}} = C_k N_k.$$

We define the *mean squared error* (nMSE), denoted as $\mathcal{E}_A^2$, with normalizing factor $\|\mathbb{E}(f)\|_Z^2$ for estimator $A$ as

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

If we use a Monte Carlo estimator to estimate $\mathbb{E}\left(f_1(\omega)\right)$,

$$\mathcal{E}^2_{A^{\mathrm{MC}}_{1,N_1}} = \frac{\mathbb{E}\left[\left\|\mathbb{E}(f_1) - A^{\mathrm{MC}}_{1,N_1}\right\|_Z^2\right]}{\|\mathbb{E}(f_1)\|_Z^2} = \frac{\left\|\mathbb{E}(f_1) - \mathbb{E}(A^{\mathrm{MC}}_{1,N_1})\right\|_Z^2 + \mathbb{E}\left[\left\|\mathbb{E}(A^{\mathrm{MC}}_{1,N_1}) - A^{\mathrm{MC}}_{1,N_1}\right\|_Z^2\right]}{\|\mathbb{E}(f_1)\|_Z^2} = \frac{\mathbb{V}\left(A^{\mathrm{MC}}\right)}{\|\mathbb{E}(f_1)\|_Z^2} = \frac{\mathbb{V}\left(f_1\right)}{N_1 \|\mathbb{E}(f_1)\|_Z^2}.$$

Given a target tolerance $\epsilon^2$ for the nMSE, the sample size $N_1$ is estimated as

$$N_1 = \frac{\sigma_1^2}{\epsilon^2 \|\mathbb{E}(f_1)\|_Z^2} \simeq \epsilon^{-2}.$$

So the total expense of sampling with $N_1$ samples using Monte Carlo method for $\mathbb{E}(f_1(\omega))$ is

$$\mathcal{W}_{\mathrm{MC}} = C_1 N_1 = \frac{C_1 \sigma_1^2}{\epsilon^2 \|\mathbb{E}(f_1)\|_Z^2}.$$

## 3. Multifidelity Monte Carlo

The Multifidelity Monte Carlo (MFMC) estimator is defined as

$$A^{\mathrm{MFMC}} := A_{1,N_1}^{\mathrm{MC}} + \sum_{k=2}^{K} \alpha_k \left( A_{k,N_k}^{\mathrm{MC}} - A_{k,N_{k-1}}^{\mathrm{MC}} \right), \tag{2}$$

where $\alpha_k$ are the coefficients to weight the correction term. In each correction term, the two Monte Carlo estimators are dependent in the sense that $A_{k,N_{k-1}}^{\mathrm{MC}}$ recycles the first $N_{k-1}$ samples of $A_{k,N_k}^{\mathrm{MC}}$ so we require $N_{k-1} \leq N_k$ for $k = 2, \ldots, K$. Using this property and (1), we can remove the dependent samples and rewrite (2) as

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

Now, the weights in the new correction terms involve a sample ratio and the importance of this reformulation is that the samples to estimate $A_{k,N_{k-1}}^{\mathrm{MC}}$ and $A_{k,N_k-N_{k-1}}^{\mathrm{MC}}$ are independent with each other with the same model $f_k$. Define $Y_1 := A_{1,N_1}^{\mathrm{MC}}$, $Y_k := (\frac{N_{k-1}}{N_k} - 1)(A_{k,N_{k-1}}^{\mathrm{MC}} - A_{k,N_k-N_{k-1}}^{\mathrm{MC}})$ for $k = 2 \ldots, K$, but $Y_1$ and $Y_k$ may not necessarily uncorrelated with each other. Note that $Y_k$ are independent of each other for $k = 2, \ldots, K$. Using the new notation, (3) can be simplified as

$$A^{\mathrm{MFMC}} = Y_1 + \sum_{k=2}^{K} \alpha_k Y_k.$$

So $\mathbb{E}(Y_k) = 0$ for $k \geq 2$ and $\mathbb{E}(A^{\mathrm{MFMC}}) = \mathbb{E}(f_1)$. For independent random variable, since each realization is uncorrelated to each other, this indicates that the sum of sample realizations and variance are interchangeable, therefore

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

So,

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

where we use the fact that $Y_k$ and $Y_j$ are independent with each other for $k \geq 2$ and [1, Lemma 3.2]

$$
\text{Cov}(Y_1, Y_k) = \text{Cov}\left(A^{\text{MC}}_{1,N_1}, A^{\text{MC}}_{k,N_k}\right) - \text{Cov}\left(A^{\text{MC}}_{1,N_1}, A^{\text{MC}}_{k,N_{k-1}}\right) = -\left(\frac{1}{N_{k-1}} - \frac{1}{N_k}\right)\rho_{1,k}\sigma_1\sigma_k.
$$

The nMSE error for the multifidelity Monte Carlo estimator is

$$
\mathcal{E}^2_{A^{\text{MFMC}}} = \frac{\mathbb{E}\left[\left\|\mathbb{E}(f_1) - A^{\text{MFMC}}\right\|_Z^2\right]}{\|\mathbb{E}(f_1)\|_Z^2} = \frac{\left\|\mathbb{E}(f_1) - \mathbb{E}(A^{\text{MFMC}})\right\|_Z^2 + \mathbb{E}\left[\left\|\mathbb{E}(A^{\text{MFMC}}) - A^{\text{MFMC}}\right\|_Z^2\right]}{\|\mathbb{E}(f_1)\|_Z^2} = \frac{\mathbb{V}\left(A^{\text{MFMC}}\right)}{\|\mathbb{E}(f_1)\|_Z^2}.
$$

Unlike multilevel Monte Carlo, the normalize mean square error splitting has a bias term representing the discretization error, for multifidelity Monte Carlo, the bias term reflects the difference between $f_1$ and multifidelity models. Since the multifidelity Monte Carlo estimator is unbiased, so the nMSE only reflects the statistical error. The total sampling cost for MFMC estimator is

$$
\sum_{k=1}^{K} C_k N_k.
$$

Our next goal is to determine the sample size $N_k$ such that the MFMC estimation satisfy the accuracy threshold $\mathcal{E}^2_{A^{\text{MFMC}}} = \epsilon^2$. We formulate an following optimization problem to minimize the sampling cost subject to a bounded variance of MFMC estimator, and solve for sample size $N_k \in \mathbb{R}$ for $k = 1\ldots, K$ and $\alpha_k \in \mathbb{R}$ for $k = 2\ldots, K$ as

$$
\begin{array}{lll}
\displaystyle\min_{N_1,\ldots N_K\in\mathbb{R},\alpha_2,\ldots,\alpha_K\in\mathbb{R}} & \sum_{k=1}^{K} C_k N_k, & \\
\text{s.t.} & \mathbb{V}\left(A^{\text{MFMC}}\right) - \|\mathbb{E}(f_1)\|_Z^2\, \epsilon^2 = 0, & \\
& N_{k-1} - N_k \leq 0, & k = 2\ldots, K, \\
& -N_1 \leq 0. &
\end{array} \tag{5}
$$

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

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

*Then the global minimizer to* (5) *is*

$$\alpha_k^* = \frac{\rho_{1,k}\sigma_1}{\sigma_k}, \quad k = 2\dots,K,$$
$$N_k^* = \frac{\sigma_1^2}{\|\mathbb{E}(f_1)\|_Z^2\,\epsilon^2}\sqrt{\frac{\rho_{1,k}^2-\rho_{1,k+1}^2}{C_k}}\sum_{j=1}^{K}\left(\sqrt{\frac{C_j}{\rho_{1,j}^2-\rho_{1,j+1}^2}}-\sqrt{\frac{C_{j-1}}{\rho_{1,j-1}^2-\rho_{1,j}^2}}\right)\rho_{1,j}^2, \quad k = 1\dots,K,$$

*with* $\rho_{1,0} = \infty$ *and* $\rho_{1,K+1} = 0$.

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

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

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

$$C_1 = \frac{\lambda_0\sigma_1^2}{N_1^2}\left(1-\rho_{1,2}^2\right) + \lambda_1 - \lambda_2, \quad C_k = \frac{\lambda_0\sigma_1^2}{N_k^2}\left(\rho_{1,k}^2 - \rho_{1,k+1}^2\right) + \lambda_k - \lambda_{k+1}, \quad C_K = \frac{\lambda_0\sigma_1^2}{N_K^2}\rho_{1,K}^2 + \lambda_K.$$

Karush-Kuhn-Tucker (KKT) conditions

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

If the inequality constraints are inactive ($\lambda_k$=0, $k = 1,\dots,K$) in the complementary slackness condition, then

$$N_1 = \sigma_1\sqrt{\lambda_0}\sqrt{\frac{1-\rho_{1,2}^2}{C_1}}, \quad N_k = \sigma_1\sqrt{\lambda_0}\sqrt{\frac{\rho_{1,k}^2-\rho_{1,k+1}^2}{C_k}}, \quad N_K = \sigma_1\sqrt{\lambda_0}\sqrt{\frac{\rho_{1,K}^2}{C_K}},$$

or we can simplify the notation as

$$N_k = \sigma_1 \sqrt{\lambda_0} \sqrt{\frac{\rho_{1,k}^2 - \rho_{1,k+1}^2}{C_k}}, \quad \frac{1}{N_k} = \frac{1}{\sigma_1 \sqrt{\lambda_0}} \sqrt{\frac{C_k}{\rho_{1,k}^2 - \rho_{1,k+1}^2}}, \quad k = 0, \ldots, K.$$

with $\rho_{1,0} = \infty, \rho_{1,K+1} = 0$.

substitute $\frac{1}{N_k}$ for $k = 1, \ldots, K$ into (4)

$$\begin{aligned}
\mathbb{V}\left(A^{\text{MFMC}}\right) &= \frac{\sigma_1^2}{N_1} + \sum_{k=2}^{K} \left(\frac{1}{N_{k-1}} - \frac{1}{N_k}\right)\left(\alpha_k^2 \sigma_k^2 - 2\alpha_k \rho_{1,k} \sigma_1 \sigma_k\right) \\
&= \frac{\sigma_1^2}{N_1} - \sigma_1^2 \sum_{k=2}^{K} \left(\frac{1}{N_{k-1}} - \frac{1}{N_k}\right)\rho_{1,k}^2 = \sigma_1^2 \left(\frac{\rho_{1,1}^2}{N_1} + \sum_{k=2}^{K}\left(\frac{1}{N_k} - \frac{1}{N_{k-1}}\right)\rho_{1,k}^2\right) \\
&= \sigma_1^2 \sum_{k=1}^{K} \left(\frac{1}{N_k} - \frac{1}{N_{k-1}}\right)\rho_{1,k}^2, \text{with } N_0 = 0 \\
&= \frac{\sigma_1}{\sqrt{\lambda_0}} \sum_{k=1}^{K} \left(\sqrt{\frac{C_k}{\rho_{1,k}^2 - \rho_{1,k+1}^2}} - \sqrt{\frac{C_{k-1}}{\rho_{1,k-1}^2 - \rho_{1,k}^2}}\right)\rho_{1,k}^2 = \|\mathbb{E}(f_1)\|_Z^2 \, \epsilon^2
\end{aligned}$$

solve for $\sqrt{\lambda_0}$, we have

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

Substitute $\sqrt{\lambda_0}$ into $N_k$ for $k = 1 \ldots, K$,

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

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

Next we want to show this local minimizer is global.

□

Consider the sample size estimation in Theorem 1 found from optimization problem 5. In reality, the sample size estimation should be integer. We will use the ceiling of $N_k$ as the sample size estimation. Note that

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

where the term $\sum_{k=1}^{K} C_k$ results from the fact that $N_k \leq \lceil N_k \rceil < N_k + 1$. In the asymptotic regime when $N_k$ is large, we should expect the impact of $\sum_{k=1}^{K} C_k$ is negligible compared to $\sum_{k=1}^{K} C_k N_k$ and thus the sampling cost behaves as (7). This indicates that

$$\frac{\sigma_1^2}{\|\mathbb{E}(f_1)\|_Z^2 \, \epsilon^2} \sum_{k=1}^{K} \sqrt{\left(\rho_{1,k}^2 - \rho_{1,k+1}^2\right) C_k} \sum_{k=1}^{K} \left(\sqrt{\frac{C_k}{\rho_{1,k}^2 - \rho_{1,k+1}^2}} - \sqrt{\frac{C_{k-1}}{\rho_{1,k-1}^2 - \rho_{1,k}^2}}\right)\rho_{1,k}^2 \geq \sum_{k=1}^{K} C_k. \tag{6}$$

Using the optimal sample estimation $\lceil N_k \rceil$, we obtain the total sampling cost of multifidelity estimator (value of the objective function) as

$$\mathcal{W}_{\text{MFMC}} = \sum_{k=1}^{K} C_k \lceil N_k \rceil = \sum_{k=1}^{K} C_k N_k = \frac{\sigma_1^2}{\|\mathbb{E}(f_1)\|_Z^2 \epsilon^2} \sum_{k=1}^{K} \sqrt{\left(\rho_{1,k}^2 - \rho_{1,k+1}^2\right) C_k} \sum_{k=1}^{K} \left( \sqrt{\frac{C_k}{\rho_{1,k}^2 - \rho_{1,k+1}^2}} - \sqrt{\frac{C_{k-1}}{\rho_{1,k-1}^2 - \rho_{1,k}^2}} \right) \rho_{1,k}^2. \tag{7}$$

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

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

Further more, we observe that

$$\begin{aligned}
\mathcal{W}_{\text{MC}} \mathbb{V}\left(A^{\text{MC}}\right) &= \frac{C_1 \sigma_1^2}{\|\mathbb{E}(f_1)\|_Z^2}, \\
\mathcal{W}_{\text{MFMC}} \mathbb{V}\left(A^{\text{MFMC}}\right) &= \frac{\sigma_1^2}{\|\mathbb{E}(f_1)\|_Z^2} \sum_{k=1}^{K} \sqrt{\left(\rho_{1,k}^2 - \rho_{1,k+1}^2\right) C_k} \sum_{k=1}^{K} \left( \sqrt{\frac{C_k}{\rho_{1,k}^2 - \rho_{1,k+1}^2}} - \sqrt{\frac{C_{k-1}}{\rho_{1,k-1}^2 - \rho_{1,k}^2}} \right) \rho_{1,k}^2.
\end{aligned}$$

This implies that if both Monte Carlo and multifidelity Monte Carlo have a same sampling cost, then $\gamma = \mathbb{V}\left(A^{\text{MFMC}}\right) / \mathbb{V}\left(A^{\text{MC}}\right)$. Therefore, this ratio also quantifies the variance reduction achieved by the MFMC estimator. 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 $S = \{1, \ldots, K\}$ be the indices of $K$ available models. We seek a subset $S_1 = \{i_1, i_2, \ldots, i_{K_1}\} \subseteq S\,(K_1 \leq K)$ of indices that minimizes the sampling cost of multifidelity Monte Carlo estimator. This lead to the following optimization problem to determine the index set $S_1$ for the selected models.

$$\begin{aligned}
\min_{S_1} \quad & \frac{\sigma_1^2}{\|\mathbb{E}(f_1)\|_Z^2 \epsilon^2} \sum_{k \in S_1} \sqrt{\left(\rho_{1,k}^2 - \rho_{1,k+1}^2\right) C_k} \sum_{k \in S_1} \left( \sqrt{\frac{C_k}{\rho_{1,k}^2 - \rho_{1,k+1}^2}} - \sqrt{\frac{C_{k-1}}{\rho_{1,k-1}^2 - \rho_{1,k}^2}} \right) \rho_{1,k}^2, \\
\text{s.t.} \quad & |\rho_{1,1}| > \ldots > |\rho_{1,K}|, \\
& \frac{C_{k-1}}{C_k} > \frac{\rho_{1,k-1}^2 - \rho_{1,k}^2}{\rho_{1,k}^2 - \rho_{1,k+1}^2}, \quad k = 2, \ldots, K, \\
& \frac{\sum_{k \in S_1} \sqrt{\left(\rho_{1,k}^2 - \rho_{1,k+1}^2\right) C_k}}{\sum_{k \in S_1} C_k} \sum_{k \in S_1} \left( \sqrt{\frac{C_k}{\rho_{1,k}^2 - \rho_{1,k+1}^2}} - \sqrt{\frac{C_{k-1}}{\rho_{1,k-1}^2 - \rho_{1,k}^2}} \right) \rho_{1,k}^2 \geq \frac{\|\mathbb{E}(f_1)\|_Z^2 \epsilon^2}{\sigma_1^2}, \\
& \rho_{1,0} = \infty \;\text{ and }\; \rho_{1,K+1} = 0.
\end{aligned} \tag{8}$$

Note that $S_1$ is non-empty and $i_1 = 1$ since the high fidility model must be included.

## 4. 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.