Multifidelity rare event simulation

Dayoung Kang

Contents

1 Motivation

Most engineering systems are often designed so that the probability of failure p_f is very small, which implies that there are very few number of samples in the failed region. Since the failure is estimated by the indicator function, which yields a binary outcome, the small number of failed samples leads to a large variance of the estimates. For example, if the target $p_f = 1 \times 10^{-3}$, we would expect single '1' and '0' for the rest of the samples. And as the event becomes more rare, it becomes more difficult to get failed samples. Often there may be zero number of failed samples and it would estimate p_f as zero, which is not the case. To obtain the meaningful estimates that are not zero, the Bayesian post-processing [1] is studied, which is able to estimate p_f when there are zero failed samples by combining prior information. In our studies, we propose Multifidelity (MF) Bayesian posterior mean estimators that ensures non-zero p_f estimates.

Another challenge of the p_f estimation is that the sparse occurrence of '1', which predominantly contributes to the p_f estimates, leads to high variance. To reduce the variance of p_f estimates, there have been numerous studies that leverage importance sampling. Importance sampling introduces the biasing distribution that can draw more samples in the failure region, which has been shown to reduce the variance. In addition to the importance sampling technique, the MF modeling is widely studied in a literature as a variance reduction method compared to the standard single fidelity modeling. Due to its variance reduction ability, the MF modeling has recently been applied to the reliability analysis that estimates p_f, which often introduces high variance. [2] proposed a multifidelity importance sampling (MFIS) estimator that derives the biasing distribution using a lower fidelity model and estimates p_f using a high fidelity model, which ensures the unbiasedness to the high fidelity p_f estimates. It achieves a large variance reduction compared to the standard Monte Carlo (MC) sampling, but there is still room for more variance reduction by using the control variate method. [3] uses both importance sampling and MF control variate method, which achieves the further variance reduction. However, the optimal model allocation is left for future work. In our study, we derive the optimal model allocation and propose multifidelity control variate importance sampling (MFCVIS) estimator that achieves further variance reduction.

A possible bottleneck of using the control variate approach is that it may introduce the negative estimates due to the subtraction structure of the estimator. Since the quantity of interest in reliability analysis is positive, it is possible to enforce the positivity to p_f estimates by introducing the logarithm as studied in [4]; [4] applied logarithm to the MF estimator to enforce the positivity of the covariance estimator. Our study adopts this logarithm to the Bayesian posterior mean estimator, and propose MF log Bayesian posterior mean estimator that ensures the positive p_f estimates.

In this paper, we propose four MF estimators: Multifidelity Monte Carlo (MFMC), MF Bayesian posterior mean, MF log Bayesian posterior mean, MFCVIS estimators, which all use control variates. The first three estimators are formulated based on the nominal distribution, while the MFCVIS uses the biasing distribution for the further variance reduction. Using two analytical examples, we show the MFCVIS estimator achieves the highest variance reduction compared to three standard and MFIS estimators.

2 Method

2.1 Monte Carlo estimator

We define a limit state function with respect to a high-fidelity model $f^{(1)} : \mathbb{R}^d \to \mathbb{R}^{d'}$, e.g.,

$$g^{(1)} = \beta - f^{(1)}, \tag{1}$$

where β is a problem-dependent or user-defined threshold value. In reliability analysis, $g^{(1)} \leq 0$ implies the failure and an indicator function is defined such that

$$I^{(1)} = \begin{cases} 1 & \text{if } g^{(1)} \leq 0 \\ 0 & \text{if } g^{(1)} > 0. \end{cases} \tag{2}$$

Then the probability of failure is defined by

$$p_f^{(1)} = E[I^{(1)}(X)] = \int_{\mathbb{R}^d} I^{(1)}(x)p(x)\mathrm{d}x, \tag{3}$$

where X is a random variable of the particular distribution, $x = [x_1, \cdots, x_d]^\top \in \mathbb{R}^d$ realizations of X, and $p(x)$ a probability density function (PDF) of a nominal distribution. The MC p_f estimator estimates $p_f^{(1)}$ using independently identically distributed realizations by

$$\hat{p}_{f,n_1}^{(1)} = \frac{1}{n_1} \sum_{i=1}^{n_1} I^{(1)}(x_i). \tag{4}$$

Here, n_1 is the number of samples used to estimate p_f.

2.2 Bayesian posterior mean estimator

2.2.1 Single fidelity estimator

Estimator For single fidelity $p_f^{(1)}$ estimation, we use an estimator

$$\hat{p}_{f,\text{bayes},n_1}^{(1)} = \frac{m_{n_1}^{(1)} + m'}{n_1 + n'}, \tag{5}$$

where n_1 is a total number of samples used to estimate $p_f^{(1)}$, $m_{n_1}^{(1)}$ is the number of failed samples among n_1, i.e., $m_{n_1}^{(1)} = \sum_{i=1}^{n_1} I^{(1)}$ and n', m' are the prior distribution parameters. With the maximum entropy prior, $m' = 1, n' = 2$, Jeffrey's prior, $m' = 0.5, n' = 1$, and for Haldane's prior, $m' = 0, n' = 0$.

Biasedness The $\hat{p}_{f,\text{bayes},n_1}^{(1)}$ is a biased estimator of $p_f^{(1)}$.

$$E[\hat{p}_{f,\text{bayes},n_1}^{(1)}] = \frac{E[m_{n_1}^{(1)} + m']}{n_1 + n'} = \frac{E[m_{n_1}^{(1)}] + m'}{n_1 + n'}. \tag{6}$$

Since m_{n_1} is a binomial random variable $m_{n_1} \sim \text{Binom}(n_1, p_f^{(1)})$, $E[m_{n_1}^{(1)}] = n_1 p_f^{(1)}$. Therefore,

$$E[\hat{p}_{f,\text{bayes},n_1}^{(1)}] = \frac{n_1 p_f^{(1)} + m'}{n_1 + n'}. \tag{7}$$

With Haldane's prior $(m' = 0, n' = 0)$, $\hat{p}_{f,\text{bayes},n_1}^{(1)} = \hat{p}_{f,\text{freq},n_1}^{(1)}$ is an unbiased estimator of $p_f^{(1)}$.

Bias The bias of the single fidelity estimator $b = E[\hat{p}_{f,\text{bayes},n_1}^{(1)}] - p_f^{(1)}$ is

$$b = \frac{n_1 p_f^{(1)} + m' - p_f^{(1)}(n_1 + n')}{n_1 + n'} = \frac{m' - n' p_f^{(1)}}{n_1 + n'}. \tag{8}$$

The estimator is asymptotically unbiased since

$$\lim_{n_1 \to \infty} b = \lim_{n_1 \to \infty} \frac{m' - n' p_f^{(1)}}{n_1 + n'} = 0. \tag{9}$$

Mean squared error Since the estimator is biased, we consider the bias term in addition to the variance of the estimator to compute Mean squared error (MSE).

$$\begin{aligned} \text{MSE} &= \text{Var}\left[\hat{p}_{f,\text{bayes},n_1}\right] + \text{Bias}\left(\hat{p}_{f,\text{bayes},n_1}, \hat{p}_f^{(1)}\right)^2 \\ &= \frac{n_1 \sigma_1^2}{(n_1+n')^2} + \left(\frac{m' - n' p_f^{(1)}}{n_1 + n'}\right)^2. \end{aligned} \tag{10}$$

Prior distribution parameters The Bayesian p_f estimator is a posterior mean of the beta distribution $\text{B}(m_{n_i} + m', n_i - m_{n_i} + n' - m')$ for $i = 1, \dots, k$. This posterior distribution

$$\pi_{\text{pos}} = \frac{p^{m_{n_i}+m'-1}(1-p)^{n_i - m_{n_i} + n' - m' - 1}}{\text{B}(m_{n_i} + m', n_i - m_{n_i} + n' - m')} \tag{11}$$

is the product of a prior distribution (assumed a conjugate beta distribution)

$$\pi_{\text{prior}} = \frac{p^{m'-1}(1-p)^{n'-m'-1}}{\text{B}(m', n'-m')} \tag{12}$$

and the likelihood function (binomial distribution)

$$\mathcal{L}(p|m_{n_i}, n_i) = \binom{n_i}{m_{n_i}} p^{m_{n_i}} (1-p)^{n_i - m_{n_i}}. \tag{13}$$

If $m' = 0$ and $n' = 0$, the prior information does not affect the posterior at all, so the posterior is purely affected by the likelihood (observation). Since the posterior mean is $\frac{m_{n_i}}{m_{n_i} + n_i - m_{n_i}} = \frac{m_{n_i}}{n_i}$, $\hat{p}^{(1)}_{f,\text{bayes},n_i} = \hat{p}^{(1)}_{f,\text{freq},n_i}$ for $m' = 0$ and $n' = 0$.

The prior mean $\frac{m'}{n'}$ implies the prior information or expected value of p_f. Figures 1 and 2 show the affect of changing m' (with fixed n') and n' (with fixed m').

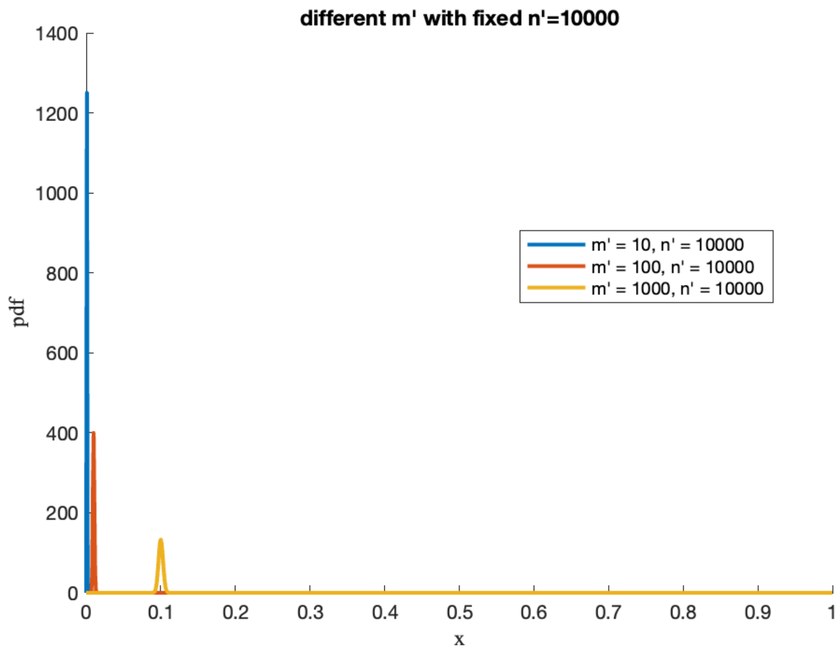


Figure 1: Beta prior distribution with fixed n'

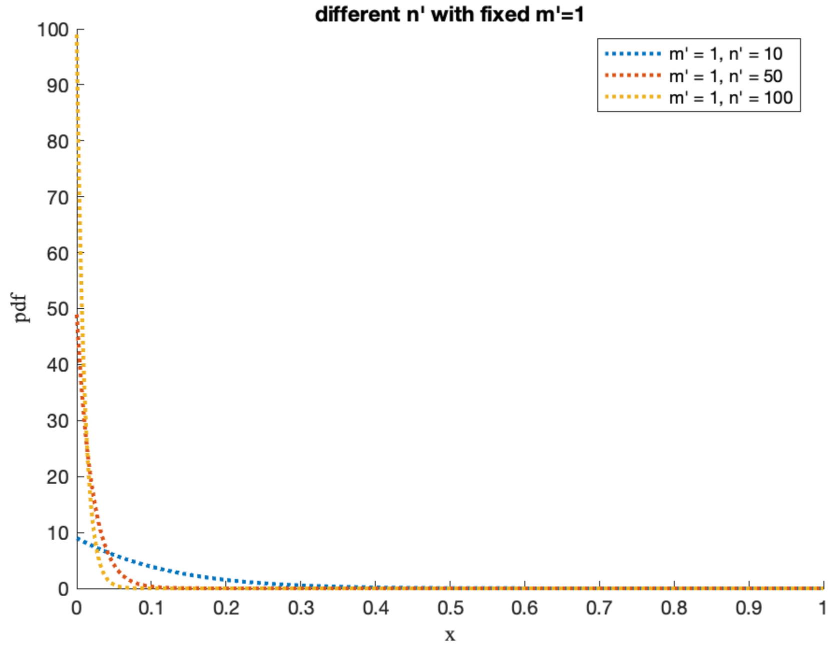


Figure 2: Beta prior distribution with fixed m'

As shown in the Figure 1, adjusting m' mainly affects to the shift of mean. Figure 2 shows that changing n' affects to the confidence on the prior information; Larger n' implies the stronger belief on the prior $p_f = \frac{m'}{n'}$ since it leads to the narrower variance. The maximum entropy prior $(m' = 1, n' = 2)$ and Jeffry's prior $(m' = 0.5, n' = 1)$ have the same mean of 0.5, but they assign different amount of confidence to the expected p_f as shown in Figure 3.

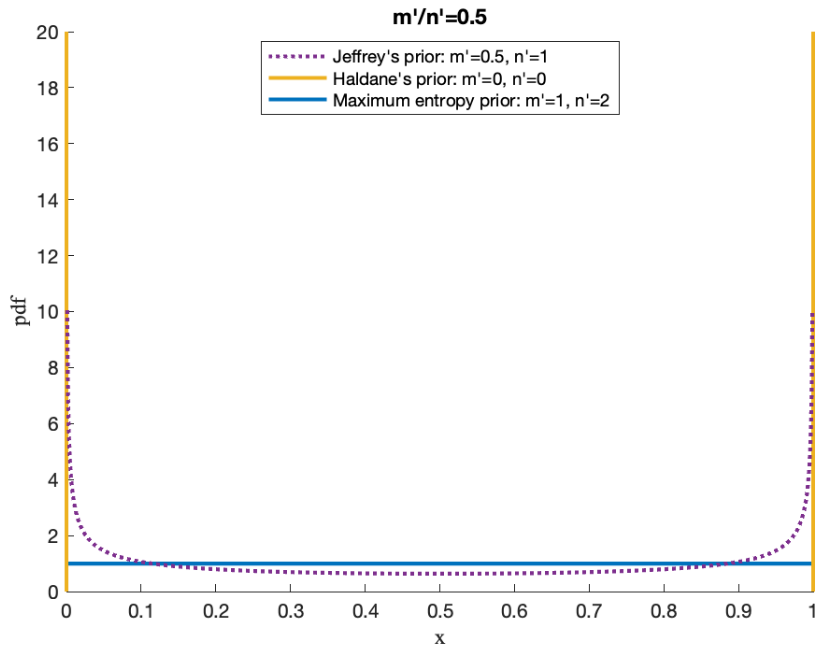


Figure 3: Choice of priors

The maximum entropy prior assigns the equal weight to the $p_f \in [0, 1]$, but the Jeffrey's prior assigns the extreme weight to both p_f of 0 and 1 rather than the middle values. Haldane's prior $(m' = n' = 0)$ assign infinity weights on p_f of 0 and 1, which doesn't assign any confidence on prior information: p_f is expected to be either 0 or 1; other than that, Haldane's prior doesn't assign any certainty on the prior information (absence of prior information). Since the prior is not informative at all, $\hat{p}^{(1)}_{f,\text{bayes},n_i} = \hat{p}^{(1)}_{f,\text{freq},n_i}$, intuitively.

2.2.2 Multifidelity estimator

Model We consider the high-fidelity model $I^{(1)}$ and correlated low fidelity models $I^{(i)} (i = 1, 2, \ldots, k)$ with respect to the Pearson correlation coefficient

$$\rho_{i,j} = \frac{\text{Cov}[I^{(i)}(X), I^{(j)}(X)]}{\sigma_i \sigma_j} \tag{14}$$

for $i, j = 1, 2, \ldots, k$ and $\sigma_i^2 = \text{Var}\left[I^{(i)}(X)\right] (i = 1, 2, \ldots, k)$, where k is the total number of high and low fidelity functions.

Estimator The multifidelity estimator is

$$\hat{p}_{f,\text{bayes}-\text{MF}} = \hat{p}^{(1)}_{f,\text{bayes},n_1} + \sum_{i=2}^{k} \alpha_i \left(\hat{p}^{(i)}_{f,\text{bayes},n_i} - \hat{p}^{(i)}_{f,\text{bayes},n_{i-1}} \right), \tag{15}$$

where $\hat{p}^{(1)}_f$ and $\hat{p}^{(i)}_f$ are the p_f evaluated by high-fidelity and low-fidelity models, respectively. From Equation 5, we have $\hat{p}^{(i)}_{f,\text{bayes},n_j} = \dfrac{m^{(i)}_{n_j} + m'}{n_j + n'}$, where $m^{(i)}_{n_j} \sim \text{Binom}(n_j, p^{(i)}_f)$.

Biasedness

$$\begin{aligned}
E\left[\hat{p}_{f,\text{bayes}-\text{MF}}\right] &= E\left[\hat{p}^{(1)}_{f,\text{bayes},n_1} + \sum_{i=2}^{k} \alpha_i \left(\hat{p}^{(i)}_{f,\text{bayes},n_i} - \hat{p}^{(i)}_{f,\text{bayes},n_{i-1}} \right)\right] \\
&= \frac{E\left[m^{(1)}_{n_1}\right] + m'}{n_1 + n'} + \sum_{i=2}^{k} \alpha_i \left(\frac{E\left[m^{(i)}_{n_i}\right] + m'}{n_i + n'} - \frac{E\left[m^{(i)}_{n_{i-1}}\right] + m'}{n_{i-1} + n'} \right) \\
&= \frac{n_1 p^{(1)}_f + m'}{n_1 + n'} + \sum_{i=2}^{k} \alpha_i \left(\frac{n_i p^{(i)}_f + m'}{n_i + n'} - \frac{n_{i-1} p^{(i)}_f + m'}{n_{i-1} + n'} \right)
\end{aligned} \tag{16}$$

Therefore, $\hat{p}_{f,\text{bayes}-\text{MF}}$ is a biased estimator of $p^{(1)}_f$. With Haldane's prior $(m' = 0, n' = 0)$, $\hat{p}_{f,\text{bayes}-\text{MF}} = \hat{p}_{f,\text{freq}-\text{MF}}$ is an unbiased estimator of $p^{(1)}_f$.

Bias The bias of the multifidelity estimator $b_{\text{MF}} = E[\hat{p}_{f,\text{bayes}-\text{MF}}] - p^{(1)}_f$ is

$$\begin{aligned}
b_{\text{MF}} &= \frac{n_1 p^{(1)}_f + m' - p^{(1)}_f (n_1 + n')}{n_1 + n'} + \sum_{i=2}^{k} \alpha_i \left(\frac{n_i p^{(i)}_f + m'}{n_i + n'} - \frac{n_{i-1} p^{(i)}_f + m'}{n_{i-1} + n'} \right) \\
&= \frac{m' - n' p^{(1)}_f}{n_1 + n'} + \sum_{i=2}^{k} \alpha_i \left(\frac{n_i p^{(i)}_f + m'}{n_i + n'} - \frac{n_{i-1} p^{(i)}_f + m'}{n_{i-1} + n'} \right).
\end{aligned} \tag{17}$$

The estimator is asymptotically unbiased since

$$\begin{aligned}
\lim_{n_1, n_i \to \infty} b_{\text{MF}} &= \lim_{n_1, n_i \to \infty} \left\{ \frac{m' - n' p^{(1)}_f}{n_1 + n'} + \sum_{i=2}^{k} \alpha_i \left(\frac{n_i p^{(i)}_f + m'}{n_i + n'} - \frac{n_{i-1} p^{(i)}_f + m'}{n_{i-1} + n'} \right) \right\} \\
&= 0 + \sum_{i=2}^{k} \alpha_i \left(p^{(i)}_f - p^{(i)}_f \right) \\
&= 0.
\end{aligned} \tag{18}$$

Mean squared error The MSE is

$$
\begin{aligned}
&\mathrm{Var}\left[\hat{p}_{f,\mathrm{bayes-MF}}\right] + \mathrm{Bias}\left(\hat{p}_{f,\mathrm{bayes-MF}}, \hat{p}_f^{(1)}\right)^2 \\
&= \mathrm{Var}\left[\hat{p}^{(1)}_{f,\mathrm{bayes},n_1} + \sum_{i=2}^{k} \alpha_i \left(\hat{p}^{(i)}_{f,\mathrm{bayes},n_i} - \hat{p}^{(i)}_{f,\mathrm{bayes},n_{i-1}}\right)\right] + \mathrm{Bias}\left(\hat{p}_{f,\mathrm{bayes-MF}}, \hat{p}_f^{(1)}\right)^2 \\
&= \mathrm{Var}\left[\hat{p}^{(1)}_{f,\mathrm{bayes},n_1}\right] + \sum_{i=2}^{k} \alpha_i^2 \left(\mathrm{Var}\left[\hat{p}^{(i)}_{f,\mathrm{bayes},n_i}\right] + \mathrm{Var}\left[\hat{p}^{(i)}_{f,\mathrm{bayes},n_{i-1}}\right]\right) \\
&+ 2\sum_{i=2}^{k} \alpha_i \left(\mathrm{Cov}\left[\hat{p}^{(1)}_{f,\mathrm{bayes},n_1}, \hat{p}^{(i)}_{f,\mathrm{bayes},n_i}\right] - \mathrm{Cov}\left[\hat{p}^{(1)}_{f,\mathrm{bayes},n_1}, \hat{p}^{(i)}_{f,\mathrm{bayes},n_{i-1}}\right]\right) \\
&- 2\sum_{i=2}^{k} \alpha_i^2 \mathrm{Cov}\left[\hat{p}^{(i)}_{f,\mathrm{bayes},n_i}, \hat{p}^{(i)}_{f,\mathrm{bayes},n_{i-1}}\right] + \left\{\frac{m' - n' p_f^{(1)}}{n_1 + n'} + \sum_{i=2}^{k} \alpha_i \left(\frac{n_i p_f^{(i)} + m'}{n_i + n'} - \frac{n_{i-1} p_f^{(i)} + m'}{n_{i-1} + n'}\right)\right\}^2
\end{aligned} \tag{19}
$$

Here,

$$
\mathrm{Var}\left[\hat{p}^{(i)}_{f,\mathrm{bayes},n_i}\right] = \frac{n_i \sigma_i^2}{(n_i + n')^2} \tag{20}
$$

by

$$
\begin{aligned}
\mathrm{Var}\left[\hat{p}^{(i)}_{f,\mathrm{bayes},n_i}\right] &= \mathrm{Var}\left[\frac{m^{(i)}_{n_i} + m'}{n_i + n'}\right] = \frac{\mathrm{Var}\left[\sum_{i'=1}^{n_i} I^{(i)}(x_{i'}) + m'\right]}{(n_i + n')^2} = \frac{\mathrm{Var}\left[\sum_{i'=1}^{n_i} I^{(i)}(x_{i'})\right]}{(n_i + n')^2} \\
&= \frac{n_i^2}{(n_i + n')^2} \mathrm{Var}\left[\frac{\sum_{i'=1}^{n_i} I^{(i)}(x_{i'})}{n_i}\right] = \frac{n_i^2}{(n_i + n')^2} \mathrm{Var}\left[\bar{I}^{(i)}_{n_i}\right] \\
&= \frac{n_i^2}{(n_i + n')^2} \frac{\sigma_i^2}{n_i} \\
&= \frac{n_i \sigma_i^2}{(n_i + n')^2}.
\end{aligned} \tag{21}
$$

The covariance terms in Equation 19 can be expressed as follows.

$$
\begin{aligned}
\mathrm{Cov}\left[\hat{p}^{(l)}_{f,\mathrm{bayes},n_i}, \hat{p}^{(t)}_{f,\mathrm{bayes},n_j}\right] &= \mathrm{Cov}\left[\frac{\sum_{i'=1}^{n_i} I^{(l)}(x_{i'}) + m'}{n_i + n'}, \frac{\sum_{j'=1}^{n_j} I^{(t)}(x_{j'}) + m'}{n_j + n'}\right] \\
&= \frac{1}{(n_i + n')(n_j + n')} \mathrm{Cov}\left[\sum_{i'=1}^{n_i} I^{(l)}(x_{i'}) + m', \sum_{j'=1}^{n_j} I^{(t)}(x_{j'}) + m'\right] \\
&= \frac{1}{(n_i + n')(n_j + n')} \sum_{i'=1}^{n_i} \sum_{j'=1}^{n_j} \mathrm{Cov}\left[I^{(l)}(x_{i'}), I^{(t)}(x_{j'})\right].
\end{aligned} \tag{22}
$$

Note that $\mathrm{Cov}\left[\sum_{i=1}^{m} a_i X_i, \sum_{j=1}^{n} b_j Y_j\right] = \sum_{i=1}^{m} \sum_{j=1}^{n} a_i b_j \mathrm{Cov}(X_i, Y_j)$. And since $x_{i'}$ and $x_{j'}$ are independent and identically distributed realizations and $x_{i'} = x_{j'}$ upto $i', j' = n_i$,

$$
\sum_{i'=1}^{n_i} \sum_{j'=1}^{n_j} \mathrm{Cov}\left[I^{(l)}(x_{i'}), I^{(t)}(x_{j'})\right] = \min\{n_i, n_j\} \mathrm{Cov}\left[I^{(l)}(X), I^{(t)}(X)\right]. \tag{23}
$$

Therefore,

$$
\begin{aligned}
\mathrm{Cov}\left[\hat{p}^{(l)}_{f,\mathrm{bayes},n_i}, \hat{p}^{(t)}_{f,\mathrm{bayes},n_j}\right] &= \frac{\min\{n_i, n_j\}}{(n_i + n')(n_j + n')} \mathrm{Cov}\left[I^{(l)}(X), I^{(t)}(X)\right] \\
&= \frac{\min\{n_i, n_j\} \rho_{l,t} \sigma_l \sigma_t}{(n_i + n')(n_j + n')}.
\end{aligned} \tag{24}
$$

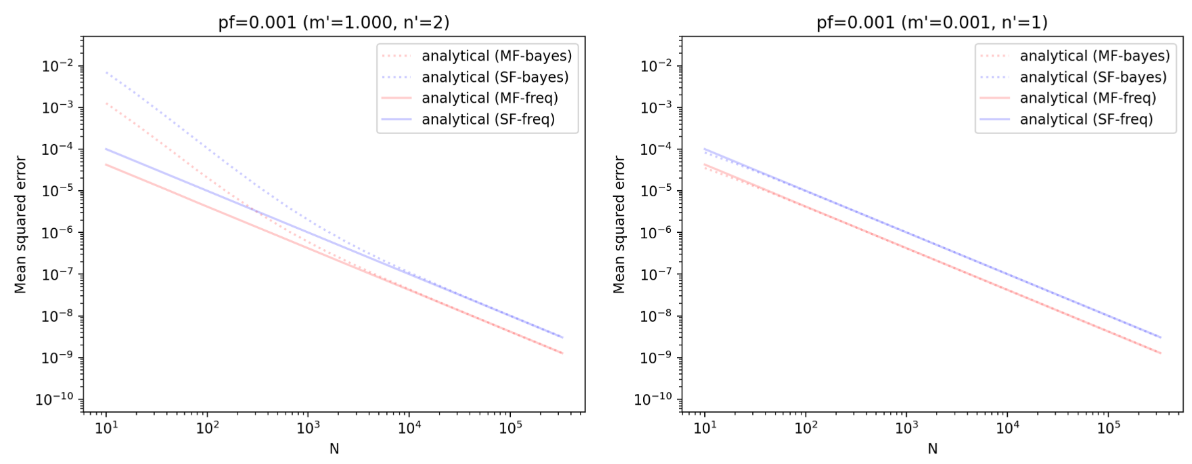
using Equation 14. With Equations 23 and 24, Equation 19 can be expressed using model statistics as follows.

$$
\begin{aligned}
&\mathrm{Var}\left[\hat{p}_{f,\mathrm{bayes-MF}}\right] + \mathrm{Bias}\left(\hat{p}_{f,\mathrm{bayes-MF}}, \hat{p}_f^{(1)}\right)^2 \\
&= \frac{n_1\sigma_1^2}{(n_1+n')^2} + \sum_{i=2}^{k} \alpha_i^2 \left\{ \frac{n_i\sigma_i^2}{(n_i+n')^2} + \frac{n_{i-1}\sigma_i^2}{(n_{i-1}+n')^2} \right\} \\
&+ 2\sum_{i=2}^{k} \alpha_i \left\{ \frac{n_1\rho_{1,i}\sigma_1\sigma_i}{(n_1+n')(n_i+n')} - \frac{n_1\rho_{1,i}\sigma_1\sigma_i}{(n_1+n')(n_{i-1}+n')} \right\} - 2\sum_{i=2}^{k} \alpha_i^2 \frac{n_{i-1}\sigma_i^2}{(n_{i-1}+n')(n_i+n')} \\
&+ \mathrm{Bias}\left(\hat{p}_{f,\mathrm{bayes-MF}}, \hat{p}_f^{(1)}\right)^2 \\
&= \frac{n_1\sigma_1^2}{(n_1+n')^2} + \sum_{i=2}^{k} \left(\frac{1}{n_{i-1}+n'} - \frac{1}{n_i+n'} \right) \left(\frac{-2\alpha_i n_1 \rho_{1,i}\sigma_1\sigma_i}{n_1+n'} \right) \\
&+ \sum_{i=2}^{k} \alpha_i^2 \left(\frac{n_i\sigma_i^2}{(n_i+n')^2} + \frac{n_{i-1}\sigma_i^2(n_i^2 - 2n_i n_{i-1} - 2n_{i-1}n' - n'^2)}{(n_{i-1}+n')^2(n_i+n')^2} \right) + \mathrm{Bias}\left(\hat{p}_{f,\mathrm{bayes-MF}}, \hat{p}_f^{(1)}\right)^2 \\
&= \frac{n_1\sigma_1^2}{(n_1+n')^2} + \sum_{i=2}^{k} \left(\frac{1}{n_{i-1}+n'} - \frac{1}{n_i+n'} \right) \left(\frac{-2\alpha_i n_1 \rho_{1,i}\sigma_1\sigma_i}{n_1+n'} \right) \\
&+ \sum_{i=2}^{k} \frac{\alpha_i^2\sigma_i^2 \left\{ n_i n_{i-1}(n_i - n_{i-1}) + n'(n_i n' - n_{i-1}n' + 2n_{i-1}n_i - 2n_{i-1}^2) \right\}}{(n_{i-1}+n')^2(n_i+n')^2} \\
&+ \left\{ \frac{m' - n'p_f^{(1)}}{n_1+n'} + \sum_{i=2}^{k} \alpha_i \left(\frac{n_i p_f^{(i)} + m'}{n_i+n'} - \frac{n_{i-1}p_f^{(i)} + m'}{n_{i-1}+n'} \right) \right\}^2 .
\end{aligned}
\tag{25}
$$

Note that if $m' = n' = 0$,

$$
\begin{aligned}
\mathrm{Var}\left[\hat{p}_{f,\mathrm{bayes-MF}}\right] &= \frac{\sigma_1^2}{n_1} + \sum_{i=2}^{k} \left(\frac{1}{n_{i-1}} - \frac{1}{n_i} \right) (\alpha_i^2\sigma_i^2 - 2\alpha_i\rho_{1,i}\sigma_1\sigma_i) \\
&= \mathrm{Var}[\hat{p}_{f,\mathrm{freq-MF}}].
\end{aligned}
\tag{26}
$$

Therefore, if we choose m' and n' to be very small near to zero, the MSE of both Frequentist and Bayesian approaches become similar as shown in Figure 4. Figure 4 compares MSE of MC and Bayesian posterior mean estimators for different prior parameters. While the maximum entropy prior introduces difference of MSE for the small budget, MSE derived by small prior parameters in Figure 4b follow the MSE of that of MC estimators.



(a) Maximum entropy prior: $m' = 1, n' = 2$

(b) Small prior parameters: $m' = 0.001, n' = 1$

Figure 4: MSE of MC and Bayesian posterior mean estimators

Theoretically, it is possible to leverage prior information to make MSE of Bayesian approach smaller than that of the Frequentist approach. Figure 5 shows the MSE derived by high confidence on prior information, i.e., high $n' = 1000$, for the case that prior guess of p_f and true p_f matches. In this case, Bayesian estimator has smaller MSE than that of the MC estimator. However, leveraging prior information with high confidence can introduce higher MSE if the prior guess is wrongly assigned by differing too much from the true p_f.

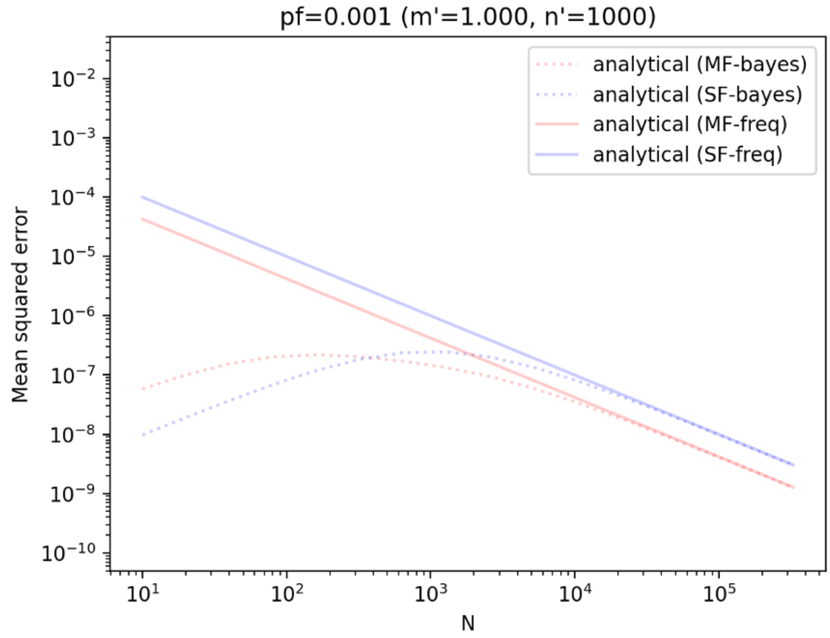


Figure 5: MSE of Bayesian posterior mean estimator for prior information with high confidence

2.2.3 Multifidelity log estimator

The multifidelity Bayesian posterior mean estimator in Equation 15 might introduce the negativity to the p_f estimates due to the subtracting structure. Since $p_f \in \mathbb{R}_+$ in practical, we introduce a logarithm into the estimator defined in Equation 15 as follows.

$$\hat{p}_{f,\log-\mathrm{bayes}-\mathrm{MF}} = \exp\left(\log \hat{p}^{(1)}_{f,\mathrm{bayes},n_1} + \sum_{i=2}^{k} \alpha_i \left(\log \hat{p}^{(i)}_{f,\mathrm{bayes},n_i} - \log \hat{p}^{(i)}_{f,\mathrm{bayes},n_{i-1}}\right)\right).$$

Note that the estimator is defined even when there is zero failed samples (except we choose Haldane's prior where $n' = m' = 0$) thanks to the Bayesian post-processing. This logarithm structure can also be applied to the multifidelity MC estimator with a caveat that the p_f estimates are undefined for the zero failed samples.

2.3 Importance sampling estimator

2.3.1 Single fidelity estimator

Estimator In rare event simulation, it is often hard to obtain the failed samples by sampling from the nominal distribution especially when the computational budget is limited. For example, we expect only one failed samples among the thousands for the target $p_f = 10^{-3}$. Importance sampling introduces a biasing distribution with a PDF q that satisfies $\mathrm{supp}(p) \subseteq \mathrm{supp}(q)$ and a random variable X' that follows the biasing distribution. We can estimate p_f in terms of q and x' (realizations of X') as follows.

$$p^{(1)}_{f,\mathrm{IS}} = \int_{\mathbb{R}^d} I^{(1)}(x)p(x)\mathrm{d}x = \int_{\mathbb{R}^d} I^{(1)}(x')\frac{p(x')}{q(x')}q(x')\mathrm{d}x'. \tag{27}$$

By defining weight $w(x') = \dfrac{p(x')}{q(x')}$, we obtain

$$p_f^{(1)} = E[I^{(1)}(X')w(X')] \tag{28}$$

and p_f estimator using importance sampling

$$\hat{p}_{f,\mathrm{IS},n_1}^{(1)} = \frac{1}{n_1}\sum_{i=1}^{n_1} I^{(1)}(x_i')w(x_i').$$

The biasing distribution is chosen to reduce the variance of p_f estimation by introducing more samples in the failure region than the nominal distribution. The variance of MC and importance sampling estimators are

$$\mathrm{Var}\left[\hat{p}_{f,n_1}^{(1)}\right] = \frac{1}{n_1}\left(p_f^{(1)} - p_f^{(1)^2}\right) \tag{29}$$

and

$$\mathrm{Var}\left[\hat{p}_{f,\mathrm{IS},n_1}^{(1)}\right] = \frac{1}{n_1}\left(E\left[\left(I^{(1)}(x')w(x')\right)^2\right] - p_f^{(1)^2}\right), \tag{30}$$

respectively. For the case $p = q$, $E\left[\left(I^{(1)}(x')w(x')\right)^2\right] = E\left[I^{(1)}(x)^2\right] = E\left[I^{(1)}(x)\right] = p_f^{(1)}$, but normally the biasing distribution is chosen such that $E\left[\left(I^{(1)}(x')w(x')\right)^2\right] < p_f^{(1)}$, i.e., $\mathrm{Var}\left[\hat{p}_{f,\mathrm{IS},n_1}^{(1)}\right] < \mathrm{Var}\left[\hat{p}_{f,n_1}^{(1)}\right]$.

2.3.2 Multifidelity estimator

In this section, we introduce three different types of multifidelity importance sampling estimators. All multifidelity estimators are the same in that they use lower fidelity model to find the biasing distribution, but they differ in a way that they estimate p_f: 1) The MFIS estimator computes p_f using high fidelity model only; 2) The MFCVIS estimator computes p_f using control variate, i.e., lower fidelity models; 3) The multifidelity approximate control variate importance sampling (MFACVIS) estimator computes p_f using approximate control variate method.

MFIS estimator To find the biasing distribution, the MFIS estimator proposed by [2] uses the expectation-maximization (EM) algorithm that fits the biasing distribution in a form of a mixture model of l Gaussian distributions ϕ_{MM} as follows.

$$\phi_{\mathrm{MM}}(z;\mu_1,\cdots,\mu_k,\Sigma_1,\cdots,\Sigma_l,\pi) = \sum_{i=1}^{l}\pi_i\phi(z;\mu_i,\Sigma_i). \tag{31}$$

Here, μ_i and Σ_i are the mean and covariance of the Gaussian distribution and π are the mixture coefficients that satisfy $\sum_{i=1}^{l}\pi_i = 1$. The EM algorithm finds μ, Σ, and π by solving the optimization problem

$$\max_{\mu,\Sigma,\pi} \Phi_{j=1}^{N_{\mathrm{fail}}}\phi_{\mathrm{MM}}(z_{\mathrm{fail}_j};\mu_1,\cdots,\mu_l,\Sigma_1,\cdots,\Sigma_l,\pi), \tag{32}$$

where N_{fail} is the number of failed samples estimated by the low fidelity function and z_{fail} are the input samples in the failure domain which were drawn by the nominal distribution. In other words, we input the failed samples, EM algorithm fits the Gaussian model to each failed samples and outputs the biasing distribution in a form of the Gaussian mixture model.

With the biasing distribution known, the MFIS estimator estimate the p_f using high-fidelity model only. Once we derived the biasing distribution, we can achieve accurate p_f estimates by drawing small number of samples since it is able to obtain more failed samples compared to drawing from the nominal distribution. In this step, we use importance sampling estimator $\hat{p}_{f,\mathrm{IS},n_i}^{(i)}$ using low fidelity model $f^{(i)}$.

MFCVIS estimator The MFCVIS estimator takes the same approach to derive the biasing distribution, but equip control variate method for p_f estimation step. In this step, we define a new multifidelity estimator as

$$\hat{p}_{f,\mathrm{MFCVIS}} = \hat{p}_{f,\mathrm{IS},n_1}^{(1)} + \sum_{i=2}^{k}\alpha_i\left(\hat{p}_{f,\mathrm{IS},n_i}^{(i)} - \hat{p}_{f,\mathrm{IS},n_{i-1}}^{(i)}\right).$$

Unbiasedness The $\hat{p}_{f,\mathrm{MFCVIS}}$ is an unbiased estimator of $p_f^{(1)}$.

Proof. We prove unbiased by showing $E\left[\hat{p}_{f,\mathrm{MFCVIS}}\right] = p_f^{(1)}$.

$$
\begin{aligned}
E\left[\hat{p}_{f,\mathrm{MFIS}}\right] &= \frac{1}{n_1}\sum_{i=1}^{n_1} E\left[I^{(1)}(x_i')w(x_i')\right] + \sum_{i=2}^{k}\alpha_i\left(\frac{1}{n_i}\sum_{i=1}^{n_i} E\left[I^{(i)}(x_i')w(x_i')\right] - \frac{1}{n_{i-1}}\sum_{i=1}^{n_{i-1}} E\left[I^{(i)}(x_i')w(x_i')\right]\right)\\
&= p_f^{(1)} + \sum_{i=2}^{k}\alpha_i\left(p_f^{(i)} - p_f^{(i)}\right)\\
&= p_f^{(1)}.
\end{aligned}
\tag{33}
$$

□

Mean squared error Since $\hat{p}_{f,\mathrm{MFCVIS}}$ is an unbiased estimator, $\mathrm{MSE}\left[\hat{p}_{f,\mathrm{MFCVIS}}\right] = \mathrm{Var}\left[\hat{p}_{f,\mathrm{MFCVIS}}\right]$.

$$
\begin{aligned}
\mathrm{MSE}\left[\hat{p}_{f,\mathrm{MFCVIS}}\right] &= \mathrm{Var}\left[\hat{p}^{(1)}_{f,\mathrm{IS},n_1} + \sum_{i=2}^{k}\alpha_i\left(\hat{p}^{(i)}_{f,\mathrm{IS},n_i} - \hat{p}^{(i)}_{f,\mathrm{IS},n_{i-1}}\right)\right]\\
&= \mathrm{Var}\left[\hat{p}^{(1)}_{f,\mathrm{IS},n_1}\right] + \sum_{i=2}^{k}\alpha_i^2\left(\mathrm{Var}\left[\hat{p}^{(i)}_{f,\mathrm{IS},n_i}\right] + \mathrm{Var}\left[\hat{p}^{(i)}_{f,\mathrm{IS},n_{i-1}}\right]\right)\\
&+ 2\sum_{i=2}^{k}\alpha_i\left(\mathrm{Cov}\left[\hat{p}^{(1)}_{f,\mathrm{IS},n_1}, \hat{p}^{(i)}_{f,\mathrm{IS},n_i}\right] - \mathrm{Cov}\left[\hat{p}^{(1)}_{f,\mathrm{IS},n_1}, \hat{p}^{(i)}_{f,\mathrm{IS},n_{i-1}}\right]\right) - 2\sum_{i=2}^{k}\alpha_i^2\left(\mathrm{Cov}\left[\hat{p}^{(i)}_{f,\mathrm{IS},n_i}, \hat{p}^{(i)}_{f,\mathrm{IS},n_{i-1}}\right]\right).
\end{aligned}
\tag{34}
$$

Here,

$$
\mathrm{Var}\left[\hat{p}^{(l)}_{f,\mathrm{IS},n_i}\right] = \frac{1}{n_i^2}\sum_{i'=1}^{n_i}\mathrm{Var}\left[I^{(l)}(x'_{i'})w(x'_{i'})\right] = \frac{1}{n_i}\mathrm{Var}\left[I^{(l)}(X')w(X')\right]
$$

and

$$
\begin{aligned}
\mathrm{Cov}\left[\hat{p}^{(l)}_{f,\mathrm{IS},n_i}, \hat{p}^{(t)}_{f,\mathrm{IS},n_j}\right] &= \mathrm{Cov}\left[\frac{1}{n_i}\sum_{i'=1}^{n_i} I^{(l)}(x'_{i'})w(x'_{i'}) + \frac{1}{n_j}\sum_{j'=1}^{n_j} I^{(t)}(x'_{j'})w(x'_{j'})\right]\\
&= \frac{1}{n_i n_j}\sum_{i'=1}^{n_i}\sum_{j'=1}^{n_j}\mathrm{Cov}\left[I^{(l)}(x'_{i'})w(x'_{i'}), I^{(t)}(x'_{j'})w(x'_{j'})\right]\\
&= \frac{\min\{n_i, n_j\}}{n_i n_j}\mathrm{Cov}\left[I^{(l)}(X')w(X'), I^{(t)}(X')w(X')\right]\\
&= \frac{1}{\max\{n_i, n_j\}}\mathrm{Cov}\left[I^{(l)}(X')w(X'), I^{(t)}(X')w(X')\right].
\end{aligned}
$$

Let $\sigma^2_{\mathrm{IS},i} = \mathrm{Var}\left[I^{(i)}(X')w(X')\right]$ and

$$
\rho_{\mathrm{IS},l,t} = \frac{\mathrm{Cov}\left[I^{(l)}(X')w(X'), I^{(t)}(X')w(X')\right]}{\sigma_{\mathrm{IS},l}\sigma_{\mathrm{IS},t}},
$$

then

$$
\mathrm{Var}\left[\hat{p}^{(l)}_{f,\mathrm{IS},n_i}\right] = \frac{\sigma^2_{\mathrm{IS},l}}{n_i}, \tag{35}
$$

$$
\mathrm{Cov}\left[\hat{p}^{(l)}_{f,\mathrm{IS},n_i}, \hat{p}^{(t)}_{f,\mathrm{IS},n_j}\right] = \frac{1}{\max\{n_i, n_j\}}\rho_{\mathrm{IS},l,t}\sigma_{\mathrm{IS},l}\sigma_{\mathrm{IS},t}. \tag{36}
$$

By substituting Equations 35 and 36 to Equation 34,

$$
\begin{aligned}
\mathrm{Var}\left[\hat{p}_{f,\mathrm{MFCVIS}}\right] &= \frac{\sigma^2_{\mathrm{IS},1}}{n_1} + \sum_{i=2}^{k} \alpha_i^2 \left(\frac{\sigma^2_{\mathrm{IS},i}}{n_i} + \frac{\sigma^2_{\mathrm{IS},i}}{n_{i-1}} \right) + 2 \sum_{i=2}^{k} \alpha_i \left(\frac{1}{n_i} \rho_{\mathrm{IS},1,i} \sigma_{\mathrm{IS},1} \sigma_{\mathrm{IS},i} - \frac{1}{n_{i-1}} \rho_{\mathrm{IS},1,i} \sigma_{\mathrm{IS},1} \sigma_{\mathrm{IS},i} \right) \\
&\quad - 2 \sum_{i=2}^{k} \alpha_i^2 \left(\frac{1}{n_i} \rho_{\mathrm{IS},i,i} \sigma_{\mathrm{IS},i} \sigma_{\mathrm{IS},i} \right) \\
&= \frac{\sigma^2_{\mathrm{IS},1}}{n_1} + \sum_{i=2}^{k} \left(\frac{1}{n_{i-1}} - \frac{1}{n_i} \right) \left(\alpha_i^2 \sigma^2_{\mathrm{IS},i} - 2 \alpha_i \rho_{\mathrm{IS},1,i} \sigma_{\mathrm{IS},1} \sigma_{\mathrm{IS},i} \right),
\end{aligned}
$$

which has the same structure with the variance of the MFMC estimator in Lemma 3.3 of [5].

Optimal control variate coefficients and model allocation To find the optimal control variate coefficients $\boldsymbol{\alpha}$ and number of function evaluations $\boldsymbol{m}$, we set up an optimization problem that minimizes the MSE $[\hat{p}_{f,\mathrm{MFIS}}]$ as follows. Here, s_i is the function evaluation cost of $I^{(i)}$, and p is the total computational budget.

$$
\begin{aligned}
&\underset{\mathbf{m}\in\mathbb{R}^{\mathbf{k}},\boldsymbol{\alpha}\in\mathbb{R}^{\mathbf{k}}}{\arg\min} \quad \mathrm{MSE}\left[\hat{p}_{f,\mathrm{MFIS}}\right] = J(\boldsymbol{m}, \boldsymbol{\alpha}) \\
&\text{subject to} \quad m_{i-1} - m_i \leq 0, \quad i = 2, \cdots, k, \\
&\qquad\qquad\quad -m_1 \leq 0, \\
&\qquad\qquad\quad \boldsymbol{s}^\top \boldsymbol{m} = p.
\end{aligned}
\tag{37}
$$

If $I^{(1)}w, \cdots, I^{(k)}w$ satisfy the two conditions

$$
|\rho_{\mathrm{IS},1,1}| > \cdots > |\rho_{\mathrm{IS},1,k}| \tag{38}
$$

and

$$
\frac{s_{i-1}}{s_i} > \frac{\rho^2_{\mathrm{IS},1,i-1} - \rho^2_{\mathrm{IS},1,i}}{\rho^2_{\mathrm{IS},1,i} - \rho^2_{\mathrm{IS},1,i+1}}, \tag{39}
$$

the global minimum of the optimization problem in Equation 37 are

$$
\alpha_i^* = \frac{\rho_{\mathrm{IS},1,i} \sigma_{\mathrm{IS},1}}{\sigma_{\mathrm{IS},i}} \tag{40}
$$

and

$$
m_i^* = m_1^* r_i^*, \tag{41}
$$

where

$$
\begin{aligned}
m_1^* &= \frac{p}{\boldsymbol{s}^\top \boldsymbol{r}^*}, \\
r_i^* &= \sqrt{\frac{s_1(\rho^2_{\mathrm{IS},1,i} - \rho^2_{\mathrm{IS},1,i+1})}{s_i(1 - \rho^2_{\mathrm{IS},1,2})}}.
\end{aligned}
$$

Proof. Refer to the proof of Theorem 3.4 in [5]. The optimization problem in Equation 37 corresponds to Equation 3.10 in [5], except our estimator defines the correlation coefficient $\rho_{\mathrm{IS},1,i}$ differently from $\rho_{1,i}$. Therefore, we follow the proof in [5] to reach the optimal solutions in Equations 40 and 41. □

MFACVIS estimator Likewise the MFIS and MFCVIS estimators, the MFACVIS estimator uses EM algorithm to derive the biasing distribution. The difference between the MFCVIS and MFACVIS is that MFACVIS uses approximate control variate method [6] that uses different sample sets for the estimator and to estimate the control variate mean.

3 Results

In this section, we present two analytic examples using linear and quadratic limit state functions for multifidelity estimators. We denote estimators derived from the nominal distribution to standard estimators, which include multifidelity MC, Bayesian posterior mean, and log Bayesian posterior mean estimators. We categorize the estimators derived from the biasing distribution as importance sampling estimators, which are MFIS, MFCVIS, and MFACVIS estimators.

3.1 Linear limit state function example

3.1.1 Problem setup

We consider the same limit state function described in the liner limit state function example of [7] for the high fidelity function

$$g^{(1)} = \beta - f^{(1)}, \quad f^{(1)} = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} z_i, \tag{42}$$

where $d = 2$ is a dimension of the random variable Z, $Z \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, and $\boldsymbol{\mu} = [\mu_1, \mu_2] = [0, 0] \in \mathbb{R}^2$, $\boldsymbol{\Sigma} = [1, 0; 0, 1] \in \mathbb{R}^{2\times 2}$. We define low fidelity models by tweaking $g^{(1)}$ with coefficients $\boldsymbol{a} = [a_2, a_3]$ and $\boldsymbol{b} = [b_2, b_3]$ such that $\rho_{1,2} = 0.85$ and $\rho_{1,3} = 0.65$:

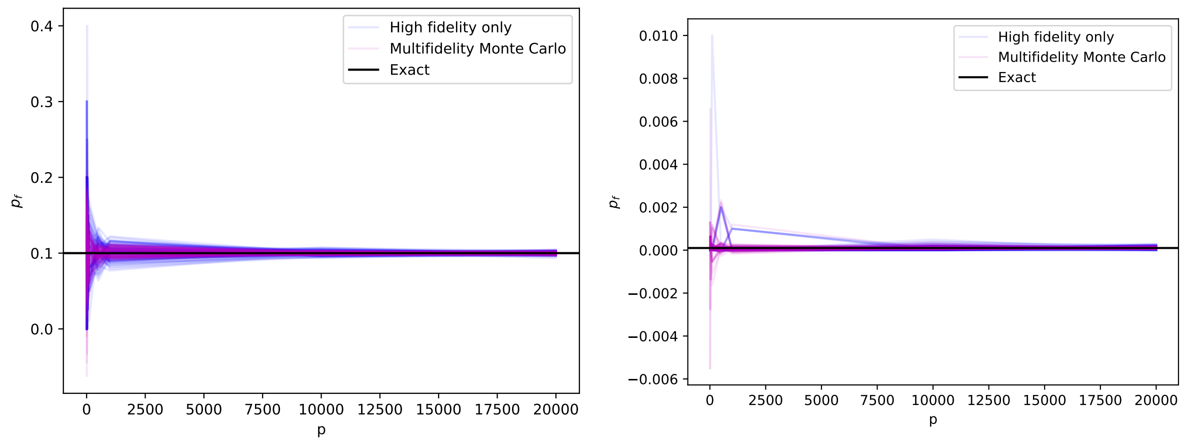
$$g^{(2)} = \beta - f^{(2)}, \quad f^{(2)} = 1.1 f^{(1)} - a_2 z_1 + b_2 z_2, \tag{43}$$

$$g^{(3)} = \beta - f^{(3)}, \quad f^{(3)} = 1.15 f^{(1)} - a_3 z_1 + b_3 z_2. \tag{44}$$

In this example, we consider lists of target $p_f = [1\times 10^{-1}, 1\times 10^{-2}, 1\times 10^{-3}, 1\times 10^{-4}]$ and $\beta = -\Phi^{-1}(p_f)$. As a result, each element of $\boldsymbol{a}$ and $\boldsymbol{b}$ are set to have four different elements depending on β to satisfy the aforementioned correlation. We show results for standard and importance sampling estimators for the computational budget $p = [10, 20, 40, 100, 500, 1000, 10000, 20000]$. We assign evaluation cost for each function to $\boldsymbol{s} = [1, 0.02, 0.001]$ and used 1×10^7 pilot samples to compute model statistics ρ and σ.

3.1.2 Standard estimators

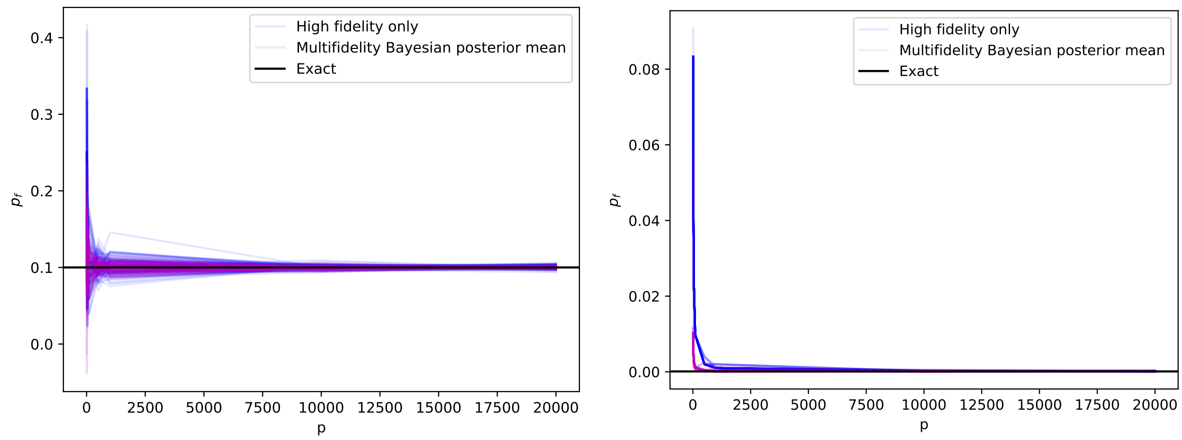
In this section, we present the performance of standard estimators compared to the single fidelity MC estimator using a high fidelity model $g^{(1)}$ along with two low fidelity functions $g^{(2)}$ and $g^{(3)}$. To evaluate the performance, we show p_f estimates and MSE of 100 sample replicates for the case of target $p_f = 0.1$ and $p_f = 0.0001$. In this estimation, maximum entropy distribution is chosen as a prior distribution for Bayesian estimators as recommended in [1]. Figures 6-8 show the p_f estimates of multifidelity Monte Carlo, Bayesian posterior mean, log Bayesian posterior mean estimators, respectively, along with the single fidelity estimates.



(a) Target $p_f = 1 \times 10^{-1}$

(b) Target $p_f = 1 \times 10^{-4}$

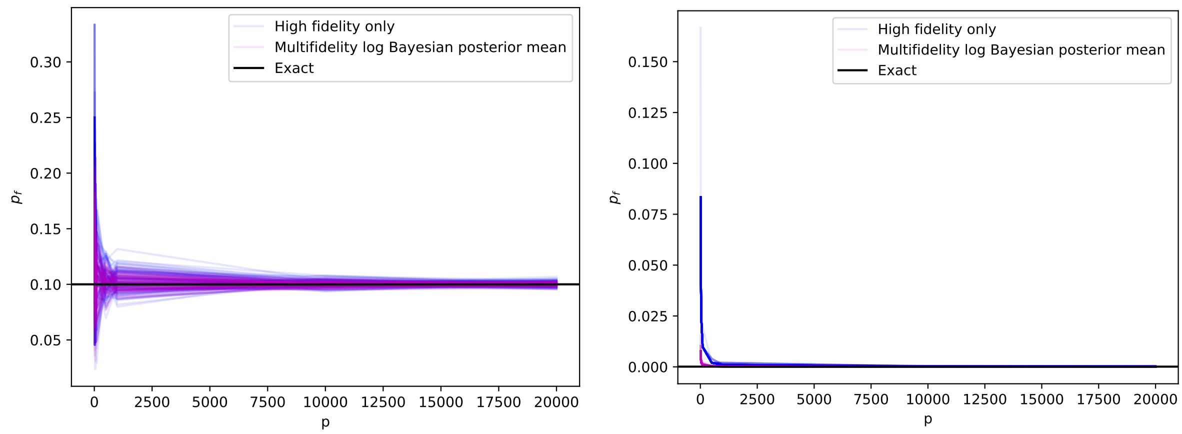
Figure 6: Sample p_f estimates using MFMC estimator



(a) Target $p_f = 1 \times 10^{-1}$

(b) Target $p_f = 1 \times 10^{-4}$

Figure 7: Sample p_f estimates using MF Bayesian posterior mean estimator



(a) Target $p_f = 1 \times 10^{-1}$

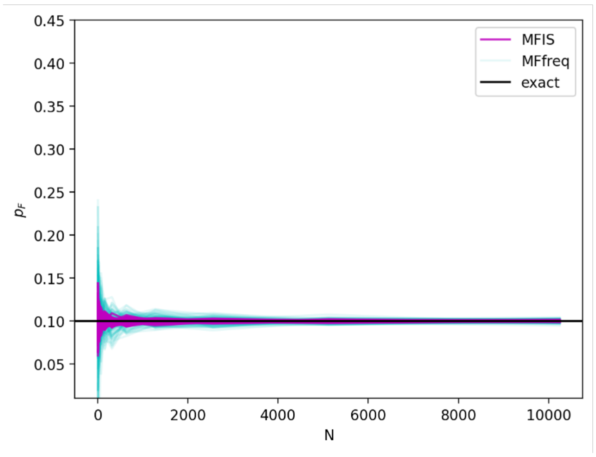
(b) Target $p_f = 1 \times 10^{-4}$

Figure 8: Sample p_f estimates using MF log Bayesian posterior mean estimator

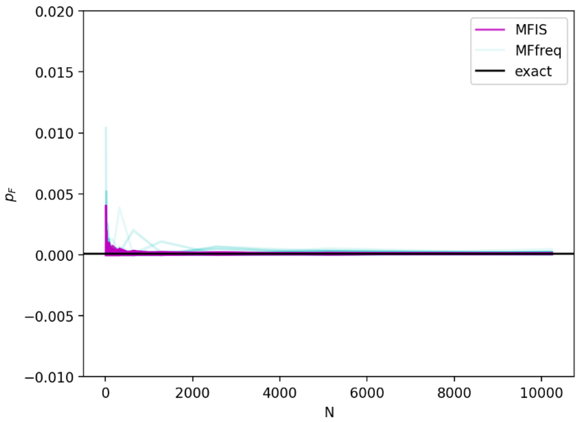
As shown in Figures 6-8, all three multifidelity estimators are unbiased and achieve lower variance than the single fidelity. As the event becomes rare, i.e., to have small p_f, the variance reduction effect becomes more prominent especially at the low budget. The single fidelity shows smaller number of lines compared to the multifidelity case, yielding unmeaningful estimates. This is because as the event becomes rare, it becomes much harder to get failed samples for a single fidelity estimator. On the other hand, the Bayesian posterior mean estimators gives meaningful estimates since Bayesian estimator is able to produce the p_f estimates without samples in the failure region. And as we expected, the multifidelity estimator performs better with smaller variance. Compared to the single fidelity estimator that p_f does not go below zero, the MFMC and Bayesian posterior mean estimators often yielded the negative p_f due to the additional term with subtraction structure. The log Bayesian estimator fixes this using logarithm and since Bayesian estimator doesn't yield zero p_f estimates, all the input samples leads to $p_f > 0$.

3.1.3 Importance sampling estimators

We now compare the standard estimators sequentially with MFIS, MFCVIS, and MFACVIS estimators. All three importance sampling estimators derive the biasing distribution using 50% of the budget and use rest of the budget to estimate p_f. Since the MFIS estimator is defined for bifidelity case, we choose one low fidelity functions either $g^{(2)}$ or $g^{(3)}$. We first present the case using the low fidelity function $g^{(2)}$, followed by $g^{(3)}$. For the bifidelity case using $g^{(2)}$, Figures 9-11 each compare p_f estimates from a hundred samples using either MFMC, Bayesian posterior mean, or log Bayesian posterior mean estimators, respectively, with those evaluated by MFIS estimator. Here, we derive the biasing distribution for each sample replicate to take into account the variance of biasing distribution search. As shown in Figures 9-11, the MFIS estimator achieves better variance reduction compared to all other three standard estimators. The performance of MFIS estimator is more prominent for the lower $p_f = 1 \times 10^{-4}$, since MFIS approach had more samples in the failure region, whereas the MFMC approach was not, which is shown by the less number of cyan lines. This result shows that the importance sampling approach can achieve higher variance reduction compared to control variate methods.

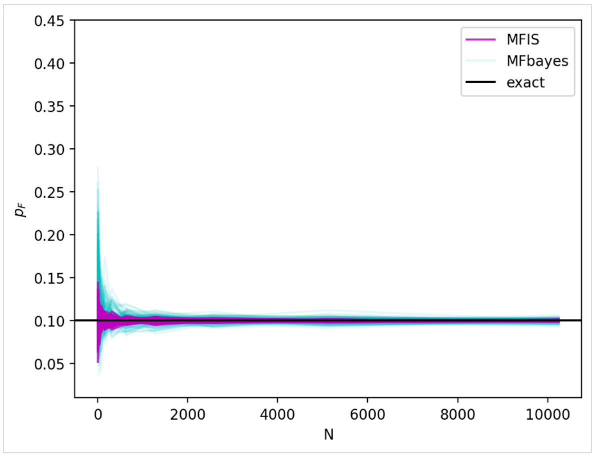


(a) Target $p_f = 1 \times 10^{-1}$

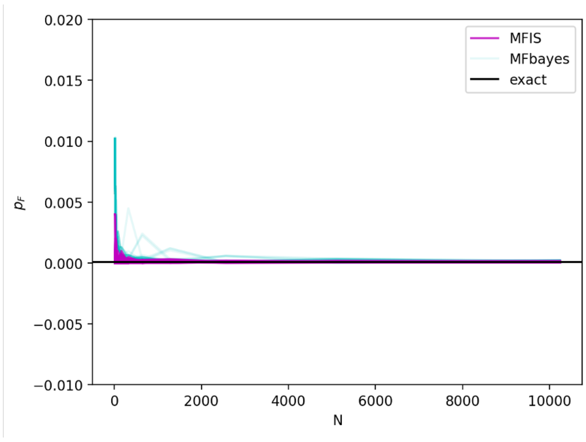


(b) Target $p_f = 1 \times 10^{-4}$

Figure 9: Comparison of sample p_f estimates for MFMC and MFIS estimators using low fidelity function $g^{(2)}$

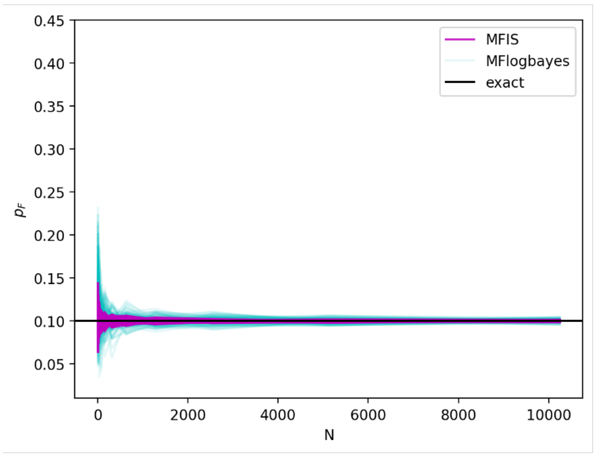


(a) Target $p_f = 1 \times 10^{-1}$

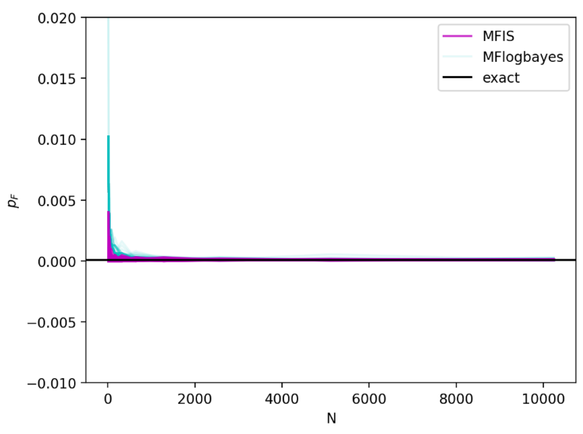


(b) Target $p_f = 1 \times 10^{-4}$

Figure 10: Comparison of sample p_f estimates for MF posterior Bayesian posterior mean and MFIS estimators using low fidelity function $g^{(2)}$



(a) Target $p_f = 1 \times 10^{-1}$



(b) Target $p_f = 1 \times 10^{-4}$

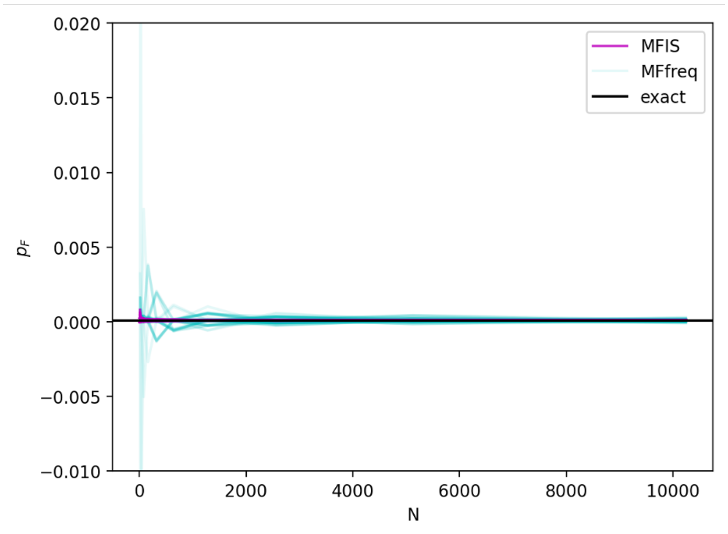
Figure 11: Comparison of sample p_f estimates for MF log Bayesian posterior mean and MFIS estimators using low fidelity function $g^{(2)}$

Table 1: Comparison using low fidelity model $g^{(2)}$

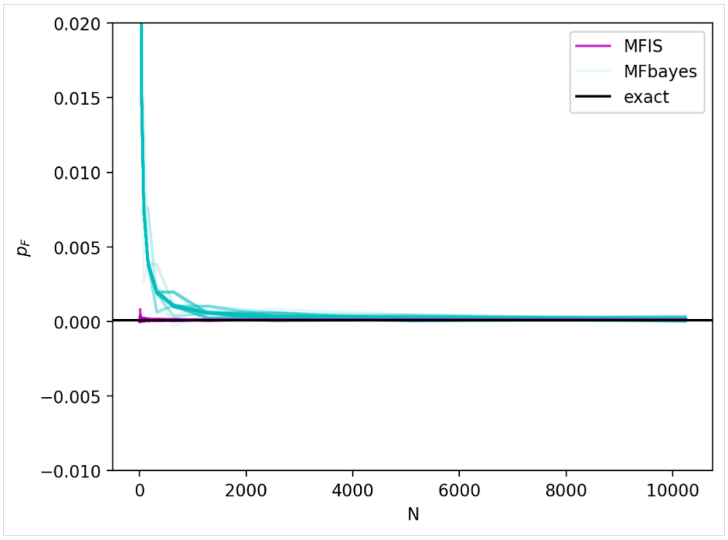
	SF	MFIS	MFMC		MFBayes		MFlog-Bayes	
	m_k	m_k	m_k	α_k	m_k	α_k	m_k	α_k
$g^{(1)}$	100	50	81	1	81	1	81	1
$g^{(2)}$	-	2500	929	0.7505	929	0.7505	929	0.7505
Total budget p	100	100	100		100		100	
MSE	9.0116×10^{-4}	3.6995×10^{-5}	3.7557×10^{-4}		3.7966×10^{-4}		4.0233×10^{-4}	

Table 1 summarizes the model allocation m and weight α (if estimator is based on control variate method) for single fidelity, multifidelity standard estimators, and MFIS estimaotr for a fixed budget $p = 100$ and target $p_f = 1 \times 10^{-1}$. The MSE is computed by 2000 samples. All standard multifidelity estimators results the same m and α since these parameters are derived from the same model statistics of $g^{(1)}$ and $g^{(2)}$. Table 1 shows that all multifidelity estimators result smaller MSE than the single fidelity case. Specifically, MFIS estimator achieves an order of magnitude smaller MSE compared to other standard estimators, which shows that the importance sampling method achieves better variance reduction compared to the control variate method.

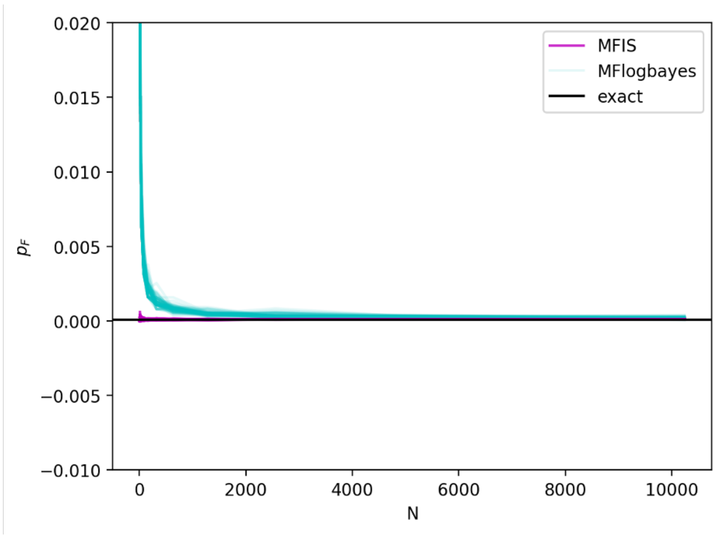
We now switch the low fidelity function to $g^{(3)}$ and compare the p_f estimates from standard and MFIS estimators for target $p_f = 1 \times 10^{-4}$. The 100 p_f sample estimates are shown in Figures 12 and the results are tabulized in Table 2. As the same as $g^{(2)}$ case, Table 2 shows averaged MSE of 2000 samples for target $p_f = 1 \times 10^{-1}$.



(a) MFMC estimates



(b) MF Bayesian posterior mean estimates



(c) MF log Bayesian posterior mean estimates

Figure 12: Comparison of sample p_f estimates for standard and MFIS estimators using low fidelity function $g^{(3)}$ for target $p_f = 1 \times 10^{-4}$

Table 2: Comparison using low fidelity model $g^{(3)}$

	SF	MFIS	MFMC		MFBayes		MFlog-Bayes	
	m_k	m_k	m_k	α_k	m_k	α_k	m_k	α_k
$g^{(1)}$	100	50	97	1	97	1	97	1
$g^{(3)}$	-	50000	2635	0.4806	2635	0.4806	2635	0.4806
Total budget p	100	100	100		100		100	
MSE	9.0344×10^{-4}	1.8361×10^{-6}	5.596×10^{-4}		5.6261×10^{-4}		6.4956×10^{-4}	

Similarly with $g^{(2)}$ case, MFIS estimator has smaller variance than all three standard estimators as shown in both Figure 12. In addition, Table 2 shows that the MSE of MFIS estimator is two order of magnitude smaller than other standard estimators. With $g^{(3)}$ having lower correlation, p_f estimates introduced higher variance for standard estimators while the variance of the MFIS estimator is reduced. Therefore, we observe higher variance reduction effect for the MFIS estimator compared to $g^{(2)}$ case. The difference of variance stands out when the budget is small. This is because we can better find biasing distribution using $m_3 = 50000$ samples which is much larger than m_2 as shown in Table 2. Once we find the biasing distribution, the MFIS estimator is able to well estimate p_f enough with small high fidelity samples. Therefore, we conclude that the importance sampling can better contribute to variance reduction compared to control variate methods if the enough budget is appropriately assigned for the low fidelity model to find the biasing distribution. Our study do not derive what is the optimal budget to be assigned for the biasing distribution search, but we choose to use 50% of budget for the search, which our numerical results is proved to show better variance reduction than the control variate methods.

References

[1] Wolfgang Betz, Iason Papaioannou, and Daniel Straub. Bayesian post-processing of monte carlo simulation in reliability analysis. *Reliability Engineering & System Safety*, 227:108731, 2022.

[2] Benjamin Peherstorfer, Tiangang Cui, Youssef Marzouk, and Karen Willcox. Multifidelity importance sampling. *Computer Methods in Applied Mechanics and Engineering*, 300:490–509, 2016.

[3] Trung Pham and Alex A Gorodetsky. Ensemble approximate control variate estimators: Applications to multifidelity importance sampling. *SIAM/ASA Journal on Uncertainty Quantification*, 10(3):1250–1292, 2022.

[4] A Maurais, T Alsup, B Peherstorfer, and Y Marzouk. Multi-fidelity covariance estimation in the log-euclidean geometry. 2023. doi: 10.48550. *arXiv preprint arXiv.2301.13749*.

[5] Benjamin Peherstorfer, Karen Willcox, and Max Gunzburger. Optimal model management for multifidelity monte carlo estimation. *SIAM Journal on Scientific Computing*, 38(5):A3163–A3194, 2016.

[6] Alex A Gorodetsky, Gianluca Geraci, Michael S Eldred, and John D Jakeman. A generalized approximate control variate framework for multifidelity uncertainty quantification. *Journal of Computational Physics*, 408:109257, 2020.

[7] Felipe Uribe, Iason Papaioannou, Youssef M Marzouk, and Daniel Straub. Cross-entropy-based importance sampling with failure-informed dimension reduction for rare event simulation. *SIAM/ASA Journal on Uncertainty Quantification*, 9(2):818–847, 2021.