Proceedings of the 2023 Winter Simulation Conference
C. G. Corlu, S. R. Hunter, H. Lam, B. S. Onggo, J. Shortle, and B. Biller, eds.

ROBUST IMPORTANCE SAMPLING FOR STOCHASTIC SIMULATIONS WITH UNCERTAIN PARAMETRIC INPUT MODEL

Seung Min Baik
Young Myoung Ko

Industrial and Management Engineering
Pohang University of Science and Technology
77, Cheongam-ro, Nam-gu, Pohang-si
Gyeongsangbuk-do, 37673, REPUBLIC OF KOREA

Eunshin Byon

Industrial and Operations Engineering
University of Michigan
1205 Beal Avenue
Ann Arbor, MI 48109-2117, USA

ABSTRACT

In stochastic simulations, input model uncertainty may significantly impact output estimation accuracy. Although variance reduction techniques alleviate the computational burden, input model uncertainty remains unaddressed. Among several variance reduction techniques, we propose a robust version of the importance sampling method. We formulate a min-max optimization problem for finding a robust sampling density for simulation inputs considering a parametric uncertainty set that represents candidates of the true input distribution. We utilize the Bayesian optimization framework for solving the outer problem and the barrier method for tackling the inner problem. By incorporating input model uncertainty in the sampling stage, our method effectively allocates simulation effort to improve estimation robustness. Numerical experiments demonstrate the advantages of the proposed method over a benchmark model assuming a precisely known input model. Our approach produces more accurate output estimation (i.e., an estimator with lower variance), highlighting its robustness and potential applicability in a variety of situations.

1 INTRODUCTION

Stochastic simulation is widely used for analyzing system performance and guiding decision-making across various applications. When a stochastic computer model is employed in stochastic simulation, the computation of simulation outputs involves two inherent sources of randomness: one in sampling simulation input from a probability distribution, and another in repeatedly executing the computer models to obtain stochastic simulation outputs. The accuracy of the input model is crucial for producing reliable results, as it can significantly affect output estimation. However, in practice, input models are often uncertain due to limited or noisy data.

A particular domain where input model uncertainty could significantly affect output estimation is wind turbine reliability analysis. To design reliable wind turbines capable of withstanding varying wind conditions, researchers employ aeroelastic computer models, such as TurbSim (Jonkman 2009) and FAST (Jonkman and Buhl 2005), developed by the National Renewable Energy Laboratory of the U.S. Department of Energy, to generate load responses given the wind condition. Assuming a known wind condition, variance reduction techniques have been proposed to effectively estimate the failure probability that a load response exceeds a threshold under a limited computational budget (Choe et al. 2016; Liu et al. 2022; Ko and Byon 2022). However, in real-world situations, the true wind condition may not be available or may be estimated with limited data. In such instances, the estimated failure probability could be highly inaccurate.

Importance sampling (IS), one of the powerful variance reduction techniques, has been widely utilized to estimate various quantities of interest such as failure probability, or rare event probability, in stochastic simulations. In this paper, we propose a robust version of the IS method to address input model uncertainty. Specifically, this study focuses on estimating the expected value of interest for a single uncertain input model. Our objective is to determine an input sampling strategy that is efficient in terms of estimation accuracy while also being resilient to the unknown true input model's realization. To achieve this, we model the true input distribution to lie within a set of plausible distributions of a specific parametric family. The parametric input model is commonly adopted in both real-world problems and recent studies (Wu et al. 2018; Song and Nelson 2019). We formulate the min-max optimization problem to find a robust sampling density and provide a solution approach utilizing a probabilistic framework to solve it.

We demonstrate the effectiveness of the proposed methodology through numerical experiments. Our sampling strategy reduces the estimator variance when the true model realization deviates from the initial belief, compared to the benchmark model that does not account for input model uncertainty. The proposed robust approach has the potential to be applied across diverse applications and industries by enhancing estimation accuracy. Below we summarize the contribution of our study.

- We propose a distributionally robust importance sampling method for stochastic simulation models, namely DR-SIS, to cope with input model uncertainty when determining the input samples for simulation. This study is, to the best of our knowledge, the first to propose the concept of input uncertainty for IS strategies in stochastic simulations.
- We provide a formulation of the min-max problem for finding the robust input sampling density in consideration of a parametric uncertainty set as candidates for the unknown true input distribution. A solution method to tackle the problem is provided by utilizing a probabilistic optimization framework, Bayesian optimization (BO), for the outer problem and the barrier method for the inner problem.
- Our numerical experiments demonstrate the robustness of the proposed methodology in the presence of input uncertainty. Even with imprecise information about the input model, the DR-SIS method produces an estimator with lower variance with respect to the true model, compared to the benchmark model. This showcases the potential of our method across diverse applications and industries which consider parametric input models.

The remainder of this paper is organized as follows. Section 2 reviews the related works and elucidates the connections of our work to them. Section 3 describes our problem by illustrating input model uncertainty and variance reduction techniques. Section 4 introduces the DR-SIS method and the solution method for obtaining the robust IS density. Section 5 presents the results of numerical experiments and conducts a discussion on the performance of the DR-SIS method in comparison to a benchmark model. Finally, Section 6 concludes the paper, summarizing the contributions of our work and suggesting future research directions.

2 RELATED WORKS

Our study is closely related to the two research fields: studies on the input uncertainty and variance reduction methods for stochastic simulations. This section reviews relevant studies and highlights the similarities and differences between our approach and existing research.

2.1 Studies on Input Uncertainty

Studies on input uncertainty can be broadly categorized into two areas, including stochastic simulations under input uncertainty and distributionally robust optimization (DRO). First, addressing input model uncertainty in stochastic simulations has recently drawn the attention of researchers. Related works include quantification of the impact of input uncertainty on simulation optimization, analysis of simulation output under input uncertainty, and optimization of the simulation model itself. Among these, we present previous works that are particularly relevant to our study. Chick (2001) developed Bayesian model averaging methods using the Bayesian approaches for the parametric distribution models. Xie et al. (2014) employed the delta method and measured the overall uncertainty as a posterior of the Gaussian process (GP) concerning the uncertainty about the input model. Gao et al. (2017) proposed a robust optimal computing budget allocation that approximately maximizes the expected opportunity cost. Fan et al. (2020) investigated a robust selection of the best problem using the indifference-zone method with the concept of ambiguity set. More extensive reviews are provided in Barton (2012), Lam (2016), and Song and Nelson (2017).

Next, as our approach seeks a robust solution with the lowest worst-case estimator variance, we make use of uncertainty sets as in DRO literature. Mottet and Lam (2017) computed the extremal performance measures with monotonicity-based uncertainty set considering moment constraints based on Choquet's theory. Lam and Mottet (2017) explored the worst-case objective value when estimating the interest measure of tail-related quantity. Blanchet and Murthy (2019) derived strong duality results for DRO formed via optimal transport discrepancy, and computed a bound for the expected value of interest. Blanchet et al. (2019) investigated an estimation of a performance measure utilizing the uncertainty set and Kullback-Leibler divergence in analyzing the insurance risk. Bai et al. (2022) proposed a method for calibrating the model parameters and quantified an output error with confidence bounds provided regarding the IS method. Zhou and Wu (2017) summarized the relevant works, and Rahimian and Mehrotra (2019) provided the overall survey about DRO, including its main concept, contribution, and relationship to other studies.

Overall, this study shares similarities with stochastic simulation and DRO studies using parametric models for uncertain input distribution. However, our approach differs from existing studies in that it employs an input sampling strategy to lower estimator variance. The following section will review the research on variance reduction.

2.2 Studies on Variance Reduction

Simulations with stochastic computer models often take the nested simulation structure (Choe et al. 2018), where the first level of input sampling can significantly impact the performance of the estimator, potentially leading to high variance along with poorly selected samples. The crude Monte Carlo is the most naive method that samples the inputs from the original distribution. As an alternative, variance reduction techniques are exploited to enhance the precision of estimates (Glasserman 2003; Asmussen and Glynn 2007). Despite the differences in implementation, they commonly share the fundamental idea of effectively drawing inputs in a statistically more efficient manner to improve estimation accuracy under a fixed computational budget.

Importance sampling is one of the widely used variance reduction methods, seeking the sampling density encouraging more on *important* inputs. Many related studies have been conducted to explore the advantages of IS and develop algorithms based on its concept, both for deterministic and stochastic computer models. Choe et al. (2015) provided theoretically optimal sampling densities in the nested simulation framework with different numbers of simulation replications for each input and demonstrated substantial improvements in estimator performance. Kurtz and Song (2013) and Cao and Choe (2019) investigated parametric input sampling density, which closely resembles the optimal density by minimizing the Kullback-Leibler distance via a tractable expectation-maximization algorithm for deterministic and stochastic simulation models, respectively. Pan et al. (2020) proposed an adaptive method that iteratively runs the simulation and exploits the obtained simulation outcomes to refine the IS density while estimating the unknown quantile value. Li et al. (2021) investigated nonparametric IS in consideration of multiple environmental factors as simulation input. Ko and Byon (2022) addressed the budget allocation problem for stochastic simulations with IS and explored the optimal balance between exploration and exploitation.

It is essential to note that previous approaches are mostly designed for situations where the input model is precisely known. However, in practical applications, the true input model might differ from the empirical distribution fitted from the data. Sometimes, we may have to utilize a predictive model to characterize the input model that remains unknown until it is actually revealed. That is, the existence of input model uncertainty inevitably causes conventional approaches to have intrinsic limitations; a particular input sampling strategy optimal for a specific input model may become suboptimal and yield estimators with high variance for another input model. This potential issue motivates the development of a more robust estimator that can efficiently handle the plausible variations in the input model.

3 PROBLEM DESCRIPTION

This study considers a nested simulation problem involving a stochastic black box computer model. The random variables (or vectors) $\mathbf{X}$ and Y denote the input and output of the simulation model. The input $\mathbf{X} \in \mathbb{R}^p$ follows a distribution F and the computer model produces a stochastic output $Y(\mathbf{X})$, given $\mathbf{X}$. General reliability analysis defines $Z \in \mathbb{R}$ as a function of Y and seeks to estimate the expectation of Z, $\mathbb{E}[Z]$. The mean $\mathbb{E}[Z]$ then can be estimated by a two-level simulation framework using the law of total expectation as $\mathbb{E}[Z] = \mathbb{E}_{\mathbf{X}}[\mathbb{E}_Y[Z|\mathbf{X}]]$. The first level involves sampling the input data $\mathbf{X}$, and given the sampled $\mathbf{X}$, the second level entails conducting a simulation run to obtain Y (or Z). For example, in our numerical experiments in Section 5, we aim to estimate the failure probability $\mathbb{P}(Y(\mathbf{X}) > l)$, which represents the probability of the simulation output exceeding a threshold l. To estimate the failure probability, we can set $Z = \mathbb{1}(Y > l)$, which makes $\mathbb{E}[Z]$ equivalent to the probability $\mathbb{P}(Y(\mathbf{X}) > l)$ (Choe et al. 2018).

In response to the need to account for input model uncertainty, we propose the distributionally robust importance sampling method for the stochastic simulation model, namely DR-SIS. This method is designed to account for input uncertainty and provide a robust and efficient estimator, even when the true input model is uncertain. The true distribution F^c, followed by a random variable $\mathbf{X}^c$, specifies the true input model and is assumed to be unknown. Superscript c denotes the *true* (or *correct*) information. As the true distribution is unknown, we adopt the concept of an uncertainty set used in DRO studies and consider a set of plausible distributions (i.e., each element F of the uncertainty set $\mathscr{F}$ is a likely candidate for the true model). Information at hand in the simulation planning stage can be utilized to establish the nominal distribution $\bar{F}$ (e.g., empirical distribution or a predictive distribution), which acts as a representative element or the center of the uncertainty set. The elements of the set are selected as the distributions *close* to the nominal one in some sense (e.g., 1-Wasserstein distance).

Let us consider the following estimator, referred to as stochastic importance sampling (SIS) estimator (Chen and Choe 2019; Ko and Byon 2022).

$$\hat{P}_{\text{SIS}} = \frac{1}{N_T}\sum_{i=1}^{N_T} Z(Y(\mathbf{X}_i))\frac{f(\mathbf{X}_i)}{q(\mathbf{X}_i)}, \tag{1}$$

where f is the probability density function of $\mathbf{X}$ with the cumulative distribution function F. This estimator is calculated by drawing a total of N_T inputs from the IS density q, running the second-level simulation once for obtaining $Y(\mathbf{X}_i)$ for each input $\mathbf{X}_i$, and then aggregating the results by calculating the weighted sum of Z values of $Y(\mathbf{X}_i)$. The weight for each input is the ratio of density functions $f(\mathbf{X}_i)/q(\mathbf{X}_i)$. It corrects the bias and makes the estimator unbiased. It has been known that the following density, denoted by q_{SIS}, is optimal and minimizes the estimator variance when the input model F (or f) is known.

$$q_{\text{SIS}}(\mathbf{x}) \propto \sqrt{\mathbb{E}_Y[Z^2|\mathbf{X}=\mathbf{x}]}f(\mathbf{x}). \tag{2}$$

Yet, when the true model F^c (or f^c) is unknown, neither the estimator nor q_{SIS} can be calculated. That is, we are facing a situation where we do not know the density required for calculating the objective value to be minimized. Under input uncertainty, the performance measure of the estimator itself becomes vague. To resolve the issue, we consider a maximum value of plausible estimator variances, $\max_{F\in\mathscr{F}} \text{Var}[\hat{P}_{\text{DR-SIS}}(F)]$ (i.e., estimator variances with respect to plausible distributions $F \in \mathscr{F}$). Unlike the conventional SIS estimator, we consider a robust estimator $\hat{P}_{\text{DR-SIS}}$ which also takes an argument of F. The input density f is assumed to be known in equation (1), however, in our approach, it is rather treated as a variable under uncertainty. We defer the rigorous details to Section 4. Our goal is to obtain a robust input sampling strategy by solving the following min-max optimization problem:

$$\min_q \max_{F\in\mathscr{F}} \text{Var}\left[\hat{P}_{\text{DR-SIS}}(F;q)\right]. \tag{3}$$

The inner maximization finds the worst-case estimator variance among the plausible input distributions within the uncertainty set. The outer minimization seeks the best sampling density that minimizes the worst-case estimator variance.

4 ROBUST IMPORTANCE SAMPLING METHOD FOR STOCHASTIC SIMULATION

This section presents a detailed approach to the proposed DR-SIS method. We first describe our robust estimator and the evaluation of its performance regarding the worst-case estimator variance, followed by the mathematical modeling of the input model uncertainty, and the formulation of the DR-SIS optimization problem. Finally, we discuss the solution method for deriving the optimal DR-SIS sampling density.

4.1 Modeling the Uncertainty of Input Model

To estimate the expected value of interest $\mathbb{E}[Z] = \mathbb{E}_{\mathbf{X}}[\mathbb{E}_Y[Z|\mathbf{X}]] = \int \mathbb{E}_Y[Z|\mathbf{X}=\mathbf{x}] f(\mathbf{x})d\mathbf{x}$ where $\mathbf{X}$ obeys the input distribution F, we define our estimator $\hat{P}_{\text{DR-SIS}}(F;q)$, namely DR-SIS estimator, as follows:

$$\hat{P}_{\text{DR-SIS}}(F;q) = \frac{1}{N_T}\sum_{i=1}^{N_T} Z(Y(\mathbf{X}_i))\frac{f(\mathbf{X}_i)}{q(\mathbf{X}_i)},$$

where $\{\mathbf{X}_i\}_{i\in\{1,2,\ldots,N_T\}}$ are independent and identically distributed samples drawn from a density q. The right-hand side of this definition is equivalent to the estimator in equation (1) as we use the same input sampling and estimator calculation procedures. The only difference is that our estimator considers a distribution F (or f) for evaluation. While being sounds trivial, we redefine the estimator here to emphasize the difference in robust estimation when the true input distribution is unknown. In the conventional IS approach that ignores the input uncertainty, the goal was to optimize the sampling density q when the distribution F is precisely known, that is, $F = F^c$. In our context, F^c is uncertain and F can be any plausible distribution within the uncertainty set.

Generally, the performance of the estimator is comprehensively evaluated by its bias and variance. If the estimator is unbiased, the lower the variance the more efficient the estimation becomes. In Chen and Choe (2019), $\hat{P}_{\text{SIS}}$ in equation (1) was proven to be unbiased if $q(\mathbf{x}) = 0$ implies $Z(Y(\mathbf{x}))f(\mathbf{x}) = 0$ for any $\mathbf{x}$. Such a condition is easily satisfiable by restricting the domain of q to be $\{\mathbf{x}|f(\mathbf{x}) > 0\}$ even when Y is unknown, and this also naturally leads to the unbiasedness of $\hat{P}_{\text{DR-SIS}}$. With $s_1(\mathbf{x}) = \mathbb{E}_Y[Z|\mathbf{X}=\mathbf{x}]$ and $s_2(\mathbf{x}) = \mathbb{E}_Y\left[Z^2|\mathbf{X}=\mathbf{x}\right]$, the variance of $\hat{P}_{\text{DR-SIS}}(F;q)$ is given by (Chen and Choe 2019; Ko and Byon 2022)

$$\begin{aligned}\text{Var}\left[\hat{P}_{\text{DR-SIS}}(F;q)\right] &= \frac{1}{N_T}\left(\mathbb{E}_q\left[s_2(\mathbf{x})\frac{f(\mathbf{x})^2}{q(\mathbf{x})^2}\right] - \left(\mathbb{E}_q\left[s_1(\mathbf{x})\frac{f(\mathbf{x})}{q(\mathbf{x})}\right]\right)^2\right) \\ &= \frac{1}{N_T}\left(\int s_2(\mathbf{x})\frac{f(\mathbf{x})^2}{q(\mathbf{x})}d\mathbf{x} - \left(\int s_1(\mathbf{x})f(\mathbf{x})d\mathbf{x}\right)^2\right).\end{aligned} \tag{4}$$

Assuming the true input distribution is known, the optimal IS density q_{SIS} in equation (2) minimizes the estimator variance. With unknown F^c, we instead take into account a set of plausible distributions $F \in \mathscr{F}$ and calculate the worst-case estimator variance. We, therefore, consider the performance measure of our robust estimator as $\max_{F\in\mathscr{F}} \text{Var}\left[\hat{P}_{\text{DR-SIS}}(F;q)\right]$. Please recall that the inner problem in equation (3) aims to calculate such a measure. Also, it is noteworthy that both estimators, $\hat{P}_{\text{SIS}}$ and $\hat{P}_{\text{DR-SIS}}$, are unbiased under any given distribution F. However, they may still either overestimate or underestimate the true quantity under F^c if the true distribution F^c is never revealed, due to a lack of information.

Then, the next question is how to construct the uncertainty set. When seeking a robust input sampling density, the uncertainty set plays a crucial role as it defines the search space for the inner maximization problem. Previous studies on DRO have considered a wide variety of uncertainty set types, such as discrepancy-based and moment-based sets (Rahimian and Mehrotra 2019). The set design can be chosen based on several aspects, including the characteristics of the problem at hand, the tractability of the solution method, and prior knowledge of the uncertain input model. In this study, we focus on a set type that is particularly useful for stochastic simulations which incorporate the parametric input models. Such an input model has been handled in many earlier studies on input uncertainty (Wu et al. 2018; Song and Nelson 2019) and input sampling strategy, especially for wind turbine analysis (Cao and Choe 2019). Still, we would like to note that another set design could also be adopted for our framework, although difficulty in coping with the inner problem might arise. Under different uncertainty sets, different constraints should be engaged within the inner problem, and thus a distinct solution method is required. We leave the consideration of another uncertainty set and the development of a tractable solution method regarding the set as a future research topic.

Consider a plausible distribution $F(\cdot;\theta)$ belonging to a certain parametric family prescribed by the parameter θ and its plausible range Θ. For example, θ could denote the mean vector and covariance matrix of Gaussian distribution or the Rayleigh scale parameter for a Rayleigh distribution. We define the uncertainty set as a collection of distributions with all admissible parameters as follows:

$$\mathscr{F} = \{F(\cdot;\theta) | \theta \in \Theta\}.$$

The candidate parameter set Θ is usually constructed by setting about its nominal value to represent the distribution most likely to approximate the true distribution. In practice, the fitted empirical distribution or the predicted model is often utilized. The size of the set, which relates to the magnitude of Θ, affects the estimation performance in terms of robustness and efficiency. A smaller uncertainty set may lead to higher efficiency (i.e., smaller worst-case estimator variance), but lower robustness. A larger set may provide better robustness in an attempt to be resilient to a wider range of true model realizations, but at the expense of decreased efficiency. Therefore, it is important to strike a balance by selecting an appropriate set size based on one's belief about the uncertainty. Sensitivity analysis in our numerical experiments demonstrates that the proposed methodology remains robust even when the actual uncertainty during variance evaluation is larger than the coverage of the uncertainty set used for determining $q_{\text{DR-SIS}}$. More details will be discussed in Section 5.

4.2 Solving DR-SIS Problem for Obtaining Robust Sampling Density

With the aforementioned estimation framework and model uncertainty, we formulate the min-max optimization model, referred to as the DR-SIS problem, for finding a robust input sampling density as follows:

$$\min_{q} \max_{F \in \mathscr{F}} \int s_2(\mathbf{x}) \frac{f(\mathbf{x})^2}{q(\mathbf{x})} d\mathbf{x} - \left(\int s_1(\mathbf{x}) f(\mathbf{x}) d\mathbf{x} \right)^2 . \tag{5}$$

Please note that the constant value $1/N_T$ in equation (4) is dropped out in the objective function as it does not affect the solution. We refer to the optimal density which minimizes the worst-case estimator variance, denoted by $q_{\text{DR-SIS}}$, as a DR-SIS density.

Unfortunately, the optimization problem in equation (5) cannot be solved using commonly adopted approaches to typical DRO problems, which are called distributionally robust stochastic problems (DRSP). In most cases, DRSP has an objective function expressed as an expectation with respect to the random variable ξ following the distribution F. The min-max optimization problem for determining the robust solution is expressed as $\min_{\mathbf{x}} \max_{F \in \mathscr{F}} \mathbb{E}_F [h(\mathbf{x},\xi)]$ where the objective is linear in F (i.e., $\mathbb{E}_F [h(\mathbf{x},\xi)] = \int h(\mathbf{x},\xi) dF(\xi)$). Cutting-surface and dual methods are the two most widely used techniques for solving the problem (Rahimian and Mehrotra 2019). The cutting-surface method requires the ability to solve a relaxation of the original problem with a finite number of plausible distributions. The dual method requires the guarantee for conic duality to formulate the inner problem's dual with zero duality gap and solve the integrated single minimization problem. Neither of them applies to our problem because our objective function has an unconventional form such as the integral of f^2 minus the square of the integral of f. Moreover, the decision variable q of the outer problem is a function, while most DRSP considers a vector $\mathbf{x}$. Taken together, these factors make it difficult to address our problem with conventional approaches.

To address the challenges, our solution method employs the probabilistic optimization approach using Bayesian optimization (BO) to find the best sampling density q. Given q, we evaluate the inner problem's objective value by solving it via the barrier method. The obtained objective value, denoting the worst-case estimator variance, is considered a fitness value of q. Then, we model the relationship between the sampling density and worst-case variance by the GP. Under the BO framework, we repeat the search process by updating the input sampling density q (outer minimization) and evaluating the fitness of the density (inner maximization) iteratively.

We now elaborate on the procedure in more detail. As shown in equation (5), the objective function of the inner maximization problem is calculated through the distribution F, which is an element of the uncertainty set, or its density f. Considering the parametric distribution, the objective value becomes a computationally tractable function of the input parameter θ for a given q. Hence, we can express the equivalent representation of the inner problem as a constrained optimization problem with the decision variable θ, as follows:

$$\begin{aligned} \max_{\theta}\ & g(\theta) := \int s_2(\mathbf{x}) \frac{f(\mathbf{x};\theta)^2}{q(\mathbf{x})} d\mathbf{x} - \left(\int s_1(\mathbf{x}) f(\mathbf{x};\theta) d\mathbf{x} \right)^2, \\ s.t.\ & \theta \in \Theta \end{aligned} \tag{6}$$

where $f(\cdot;\theta)$ is a density function of the parametric distribution $F(\cdot;\theta)$ with parameter θ.

It is notable that the objective function is differentiable with respect to θ. To solve the inner problem in equation (6), we employ the barrier method (Nocedal and Wright 1999) in conjunction with the gradient-based algorithm. Specifically, we use a logarithmic barrier function to penalize the constraints, and we solve the optimization problem using the L-BFGS-B algorithm (Zhu et al. 1997). Due to the non-concavity of g, the method poses a risk of converging to local optimality. To mitigate this risk, we run the algorithm multiple times with different initial parameters. Since the search space of θ is low-dimensional for commonly used input distributions, this approach can practically find the near-worst-case estimator variance.

With the ability to evaluate the inner problem, we now shift our focus to solving the outer minimization problem. Searching for the best sampling density q is not straightforward, as the domain of the decision variable is a function space, and the near-worst-case estimator variance, given q, does not have an analytical expression that could be explicitly derived. Therefore, we investigate the parametric IS density when q is restricted to being a parametric distribution. Using a specific parametric distribution, on the other hand, may not provide enough flexibility.

To circumvent this, we consider a mixture of multiple parametric densities, specifically the Gaussian mixture model (GMM) (Kurtz and Song 2013; Wang and Zhou 2015; Cao and Choe 2019; Geyer et al. 2019). The procedure for solving the outer problem is then reduced to finding the best GMM parameter vector, including the weight and Gaussian parameters of each component, $\theta^{\mathrm{BO}} = \left(\alpha_1^{\mathrm{BO}}, \alpha_2^{\mathrm{BO}}, \ldots, \alpha_k^{\mathrm{BO}}, \mu_1^{\mathrm{BO}}, \mu_2^{\mathrm{BO}}, \ldots, \mu_k^{\mathrm{BO}}, \Sigma_1^{\mathrm{BO}}, \Sigma_2^{\mathrm{BO}}, \ldots, \Sigma_k^{\mathrm{BO}}\right)$. When the true input density is known, the expectation-maximization algorithm can be used to find the parameter vector of GMM that approximates the optimal density q_{SIS} in equation (2). However, it cannot be applied to our case due to the presence of uncertainty. Instead, we obtain the GMM parameter vector to form a robust IS density using the BO framework. The BO establishes a GP surrogate model that approximates the objective value of the inner problem at the GMM parameter vector.

The implementation of the algorithm consists of several key steps. We start by initializing a set of GMM parameter vectors Θ^{BO} as candidate solutions. These initial candidates can be selected randomly or using an experimental design such as the Latin hypercube design. In our experiment, we begin with the GMM parameter vector that best approximates the theoretically optimal sampling density q_{SIS} without uncertainty. We then calculate the fitness of these candidates, $\mathscr{V}^{\mathrm{BO}} = \left\{ v(\theta^{\mathrm{BO}}),\ \forall \theta^{\mathrm{BO}} \in \Theta^{\mathrm{BO}} \right\}$, with each fitness value corresponding to the objective value of the inner problem for a given q defined by the GMM

Algorithm 1 Bayesian Optimization Algorithm for Solving DR-SIS Problem

Require: An oracle for solving the inner maximization problem given the GMM parameter vector.

Input: Set of initial GMM parameter vectors Θ^{BO} and corresponding objective values of the inner maximization problems $\mathscr{V}^{\text{BO}} = \left\{ v(\theta^{\text{BO}}),\ \forall \theta^{\text{BO}} \in \Theta^{\text{BO}} \right\}$.

Output: The best GMM parameter vector $\theta^{\text{BO}}_{best}$ which minimizes the objective of the inner problem.

1: **Initialization:** Find the best initial GMM parameter vector, $\theta^{\text{BO}}_{best} \leftarrow \underset{\theta^{\text{BO}} \in \Theta^{\text{BO}}}{\operatorname{argmin}}\ v\left(\theta^{\text{BO}}\right)$.
2: **while** Stopping criterion is not satisfied **do**
3: Fit the Gaussian process using dataset $\left(\theta^{\text{BO}}, \mathscr{V}^{\text{BO}}\right)$.
4: Obtain the new GMM parameter vector θ^{BO}_{new} that maximizes the expected improvement acquisition function, $\theta^{\text{BO}}_{new} \leftarrow \underset{\theta^{\text{BO}}}{\operatorname{argmax}}\ \mathbb{E}\left[\max\left(v\left(\theta^{\text{BO}}_{best}\right) - v\left(\theta^{\text{BO}}\right), 0\right)\right]$.
5: Update $\Theta^{\text{BO}} \leftarrow \Theta^{\text{BO}} \cup \theta^{\text{BO}}_{new}$, $\mathscr{V}^{\text{BO}} \leftarrow \mathscr{V}^{\text{BO}} \cup v\left(\theta^{\text{BO}}_{new}\right)$.
6: **if** $v\left(\theta^{\text{BO}}_{new}\right) < v\left(\theta^{\text{BO}}_{best}\right)$ **then**
7: $\theta^{\text{BO}}_{best} \leftarrow \theta^{\text{BO}}_{new}$.
8: **end if**
9: **end while**
10: **return** $\theta^{\text{BO}}_{best}$

parameter vector. Next, the following steps are iteratively performed; Using dataset $\left\{\Theta^{\text{BO}}, \mathscr{V}^{\text{BO}}\right\}$, we fit a GP surrogate model. We identify a new candidate GMM parameter vector θ^{BO}_{new} that maximizes the acquisition function (here, we utilize an expected improvement that denotes the possible gain in terms of variance reduction of the worst case). We update the dataset by adding the new parameter and its fitness value $v\left(\theta^{\text{BO}}_{new}\right)$. The algorithm continues iterating through these steps until a stopping criterion is met, such as a predefined number of iterations or convergence condition of the objective function (i.e., fitness decrease). Once the algorithm stops, the best GMM parameter vector obtained during the search is returned as the solution for the outer problem. Algorithm 1 summarizes the described procedure. Numerical experiments in Section 5 show that this algorithm efficiently searches for a sampling density while avoiding exhaustive search.

5 NUMERICAL EXPERIMENTS

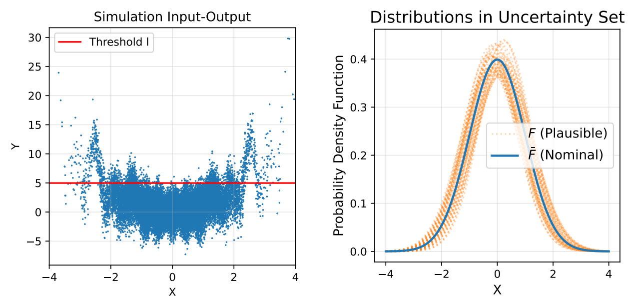
We conduct numerical experiments to evaluate the performance of the DR-SIS method. As a benchmark model, we adopt q_{SIS} in equation (2), which is the theoretically optimal IS density for stochastic simulations when the true input model is known. The benchmark model treats the nominal distribution $\bar{F}$ as if it is the same as the true distribution, neglecting its uncertainty. On the contrary, in the DR-SIS, the nominal distribution acts as the representative of the uncertainty set.

Following our motivating examples discussed earlier in Section 1, we consider the estimation of failure probability where $Z = \mathbb{1}\left(Y(\mathbf{X}) > l\right)$. However, please note that the proposed methodology remains applicable to other cases with a more general Z, even though we only handle a specific problem here due to the page limitations. We consider a numerical example in Choe et al. (2015) and Cao and Choe (2019) that aims to estimate the tail probability $\mathbb{P}(Y(X) > l)$ with one-dimensional input X. Given the input X, the computer model produces a stochastic output as follows:

$$\begin{gathered} Y|\{X = x\} \sim \mathscr{N}\left(\mu_Y(x), \sigma_Y(x)\right), \\ \mu_Y(x) = 0.95x^2(1 + 0.5\cos(5x) + 0.5\cos(10x)), \quad \sigma_Y(x) = 1 + 0.7|x| + 0.4\cos(x) + 0.3\cos(14x), \end{gathered} \tag{7}$$

where the nominal input distribution is $X \sim \mathscr{N}(0,1)$. The threshold l is set to be 4.98. With this threshold, the failure probability with the nominal input model is 0.05. With $Z = \mathbb{1}\left(Y(X) > l\right)$, the values $s_1(x)$ and $s_2(x)$ become identical as the conditional output mean $s(x) = \mathbb{P}(Y > l|X = x)$, which can be obtained from the data generating structure in equation (7). In practice, $s(x)$ can be estimated using data generated from the pilot stage simulation. Since the focus of this study is to investigate the influence of input uncertainty on output estimation, we leave the analysis of the effect of estimation accuracy for $s(x)$ as a future research topic.

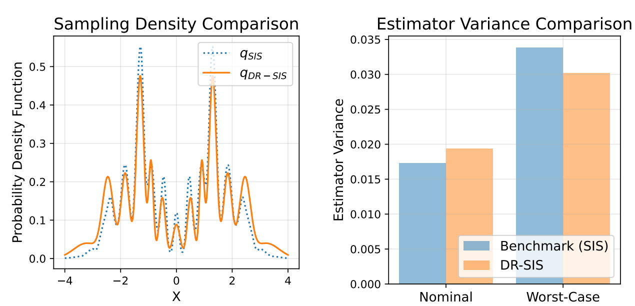
In implementing DR-SIS, we slightly modify the input model setting to account for its uncertainty. The set of plausible input distributions is modeled with a family of normal distributions, following a parametric uncertainty set. To be specific, we consider the uncertainty set $\mathscr{F} = \{\mathscr{N}(\mu_X, \sigma_X) | \, |\mu_X| < 0.3, |\sigma_X - 1| < 0.1\}$, with mean μ_X and standard deviation σ_X. We construct $q_{\text{DR-SIS}}$ with the GMM density. The number of components in GMM can be decided using the information criterion suggested in Cao and Choe (2019). In this study, we use 13 components. Figure 1 illustrates the experimental setup.



(a) Simulation input-output. (b) Input distributions.

Figure 1: Experimental setup: simulation model and plausible input distributions in uncertainty set.

Now, we evaluate the performance of the DR-SIS and the benchmark methods. Figure 2 compares the results from the two approaches. First, Figure 2(a) presents the robust IS density $q_{\text{DR-SIS}}$ from DR-SIS and the IS density q_{SIS} from the benchmark approach. Overall, they show similar patterns, sampling more inputs in the regions that are considered to be *important* (i.e., they show spikes at similar input regions). We, however, observe that $q_{\text{DR-SIS}}$ allocates more simulation efforts to larger $|X|$. This region has a significant conditional failure probability, as $s(x) = \mathbb{P}(Y > l | X = x)$ is high. This greater allocation in large $|X|$ enables us to obtain a better estimate when the true input model has a higher probability at extreme X than expected in the nominal distribution. The benchmark model, on the other hand, neglects this region, because the nominal distribution has a low probability at large $|X|$.

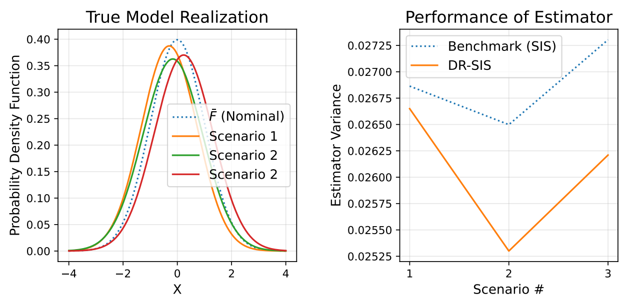


(a) Sampling density. (b) Estimator variance.

Figure 2: Comparison results between DR-SIS and benchmark model.

These results align with the fact that the worst-case distribution within the uncertainty set occurs when the true input distribution is $X \sim \mathscr{N}(-0.3, 1.1)$ or $X \sim \mathscr{N}(0.3, 1.1)$. These tilted and widely spread input distributions result in greater probabilities at larger $|X|$. The benchmark approach yields the highest estimator variance in these worst-case input scenarios, as shown in Figure 2(b). Our robust sampling density is resilient even in those situations, lowering the estimator variance by approximately 11% over the benchmark approach. On the contrary, when the true F^c is identical to our initial belief (i.e., the nominal distribution $\bar{F}$), q_{SIS} produces a lower estimator variance compared to $q_{\text{DR-SIS}}$.

It should be worthwhile to mention that DR-SIS has benefits even when the input model moderately deviates from our expectations. Figure 3 shows three other scenarios of true model realization and their corresponding estimator variances, when $X \sim \mathscr{N}(-0.3, 1.033)$, $X \sim \mathscr{N}(-0.167, 1.1)$, and $X \sim \mathscr{N}(0.233, 1.067)$, respectively. With these moderate deviations from the nominal distribution, DR-SIS still outperforms the benchmark method.



(a) Realization scenarios. (b) Estimator variance.

Figure 3: Comparison results when the true distribution moderately deviates from the nominal distribution.

Finally, we conduct a sensitivity analysis to investigate the effect of the magnitude of uncertainty set on the performance of the DR-SIS method. The true input distribution could deviate from the nominal distribution either less or more than we initially expected. We aim to understand how the range of such realization affects the estimator variance. Suppose we obtain the robust IS density with the pre-specified uncertainty set $\mathscr{F}$, as shown in Figure 2(a). Then, we consider uncertainty sets of different sizes, $\mathscr{F}_\alpha = \{\mathscr{N}(\mu_X, \sigma_X)|\ |\mu_X| < 0.3\alpha, |\sigma_X - 1| < 0.1\alpha\}$ with $0 \leq \alpha \leq 2$. Here, the parameter α regulates the set size. When α is larger than 1, it implies that the presumed uncertainty set is too small and that the worst-case scenario could happen outside of it. With $\alpha < 1$, a relatively conservative uncertainty set is chosen when obtaining $q_{\text{DR-SIS}}$. Figure 4 presents the result of the sensitivity analysis.

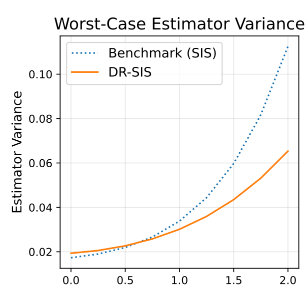


Figure 4: Sensitivity analysis of DR-SIS and benchmark model under different uncertainty set sizes.

The proposed method demonstrates a smaller worst-case estimator variance compared to the benchmark model for $\alpha \geq 0.612$. Even when the true input distribution lies outside the uncertainty set (i.e., $\alpha > 1$), DR-SIS can maintain robustness. It is also noticeable that the worst-case estimator variance increases faster for larger α. This result suggests that when the input uncertainty is large, employing a conservative uncertainty set is preferable. On the other hand, the benchmark model slightly performs better for α smaller than 0.612 because the nominal distribution is close to the true distribution. In such a case, a tighter uncertainty set can be employed. However, the difference between the two approaches is not substantial. Collectively, these findings show that being conservative in determining the uncertainty set size may be advantageous in most circumstances, particularly when the input model is highly uncertain.

6 CONCLUSION

In this study, we introduce a new approach for addressing input model uncertainty when devising the IS density in stochastic simulations. By incorporating input model uncertainty, we can allocate simulation efforts robustly when the true input distribution deviates from the presumed nominal distribution, which is often estimated with empirical data. Numerical experiments, including sensitivity analysis, demonstrate the effectiveness of the proposed method, highlighting its advantages over a benchmark model assuming the precisely known input model. The proposed robust approach has the potential to be applied across diverse applications to enhance the reliability of simulation-based estimations. Future research directions include investigating a solution approach for the proposed robust IS framework with different uncertainty set designs (e.g., discrepancy-based set with 1-Wasserstein distance) and exploring adaptive algorithms for dynamic adjustment of uncertainty set sizes.

ACKNOWLEDGMENTS

This work was supported in part by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2021R1A2C1094699 and NRF-2021R1A4A1031019) and in part by U.S. National Science Foundation (CMMI- 2226348, IIS-1741166).

REFERENCES

Asmussen, S., and P. W. Glynn. 2007. *Variance-Reduction Methods*, 126–157. New York: Springer.

Bai, Y., Z. Huang, and H. Lam. 2022. "Model Calibration via Distributionally Robust Optimization: On the NASA Langley Uncertainty Quantification Challenge". *Mechanical Systems and Signal Processing* 164:108211.

Barton, R. R. 2012. "Tutorial: Input Uncertainty in Outout Analysis". In *Proceedings of the 2012 Winter Simulation Conference*, edited by C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A. M. Uhrmacher, 1–12. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Blanchet, J., H. Lam, Q. Tang, and Z. Yuan. 2019. "Robust Actuarial Risk Analysis". *North American Actuarial Journal* 23(1):33–63.

Blanchet, J., and K. Murthy. 2019. "Quantifying Distributional Model Risk Via Optimal Transport". *Mathematics of Operations Research* 44(2):565–600.

Cao, Q. D., and Y. Choe. 2019. "Cross-Entropy Based Importance Sampling for Stochastic Simulation Models". *Reliability Engineering & System Safety* 191:106526.

Chen, Y.-C., and Y. Choe. 2019. "Importance Sampling and its Optimality for Stochastic Simulation Models". *Electronic Journal of Statistics* 13(2):3386 – 3423.

Chick, S. E. 2001. "Input Distribution Selection for Simulation Experiments: Accounting for Input Uncertainty". *Operations Research* 49(5):744–758.

Choe, Y., E. Byon, and N. Chen. 2015. "Importance Sampling for Reliability Evaluation with Stochastic Simulation Models". *Technometrics* 57(3):351–361.

Choe, Y., H. Lam, and E. Byon. 2018. "Uncertainty Quantification of Stochastic Simulation for Black-Box Computer Experiments". *Methodology and Computing in Applied Probability* 20:1155–1172.

Choe, Y., Q. Pan, and E. Byon. 2016. "Computationally Efficient Uncertainty Minimization in Wind Turbine Extreme Load Assessments". *Journal of Solar Energy Engineering* 138(4):041012.

Fan, W., L. J. Hong, and X. Zhang. 2020. "Distributionally Robust Selection of the Best". *Management Science* 66(1):190–208.

Gao, S., W. Chen, and L. Shi. 2017. "A New Budget Allocation Framework for the Expected Opportunity Cost". *Operations Research* 65(3):787–803.

Geyer, S., I. Papaioannou, and D. Straub. 2019. "Cross Entropy-Based Importance Sampling Using Gaussian Densities Revisited". *Structural Safety* 76:15–27.

Glasserman, P. 2003. *Variance Reduction Techniques*, 185–279. New York: Springer.

Jonkman, B. J. 2009. "TurbSim User's Guide: Version 1.50". Technical Report NREL/TP-500-46198, National Renewable Energy Laboratory, Golden, Colorado.

Jonkman, J. M., and M. L. Buhl, Jr. 2005. "FAST User's Guide". Technical Report NREL/EL-500-38230, National Renewable Energy Laboratory, Golden, Colorado.

Ko, Y. M., and E. Byon. 2022. "Optimal Budget Allocation for Stochastic Simulation with Importance Sampling: Exploration vs. Replication". *IISE Transactions* 54(9):881–893.

Kurtz, N., and J. Song. 2013. "Cross-Entropy-Based Adaptive Importance Sampling using Gaussian Mixture". *Structural Safety* 42:35–44.

Lam, H. 2016. "Advanced Tutorial: Input Uncertainty and Robust Analysis in Stochastic Simulation". In *Proceedings of the 2016 Winter Simulation Conference*, edited by T. M. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, 178–192. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Lam, H., and C. Mottet. 2017. "Tail Analysis without Parametric Models: A Worst-Case Perspective". *Operations Research* 65(6):1696–1711.

Li, S., Y. M. Ko, and E. Byon. 2021. "Nonparametric Importance Sampling for Wind Turbine Reliability Analysis with Stochastic Computer Models". *The Annals of Applied Statistics* 15(4):1850–1871.

Liu, B., X. Yue, E. Byon, and R. A. Kontar. 2022. "Parameter Calibration in Wake Effect Simulation Model with Stochastic Gradient Descent and Stratified Sampling". *The Annals of Applied Statistics* 16(3):1795–1821.

Mottet, C., and H. Lam. 2017. "On Optimization over Tail Distributions". *arXiv preprint arXiv:1711.00573*.

Nocedal, J., and S. J. Wright. 1999. *Numerical Optimization*. New York: Springer.

Pan, Q., E. Byon, Y. M. Ko, and H. Lam. 2020. "Adaptive Importance Sampling for Extreme Quantile Estimation with Stochastic Black Box Computer Models". *Naval Research Logistics* 67(7):524–547.

Rahimian, H., and S. Mehrotra. 2019. "Distributionally Robust Optimization: A Review". *arXiv preprint arXiv:1908.05659*.

Song, E., and B. L. Nelson. 2017. "Input Model Risk". In *Advances in Modeling and Simulation*, edited by A. Tolk, J. Fowler, G. Shao, and E. Yücesan, 63–80. Cham: Springer.

Song, E., and B. L. Nelson. 2019. "Input–Output Uncertainty Comparisons for Discrete Optimization via Simulation". *Operations Research* 67(2):562–576.

Wang, H., and X. Zhou. 2015. "A Cross-Entropy Scheme for Mixtures". *ACM Transactions on Modeling and Computer Simulation* 25(1):1–20.

Wu, D., H. Zhu, and E. Zhou. 2018. "A Bayesian Risk Approach to Data-Driven Stochastic Optimization: Formulations and Asymptotics". *SIAM Journal on Optimization* 28(2):1588–1612.

Xie, W., B. L. Nelson, and R. R. Barton. 2014. "A Bayesian Framework for Quantifying Uncertainty in Stochastic Simulation". *Operations Research* 62(6):1439–1452.

Zhou, E., and D. Wu. 2017. "Simulation Optimization under Input Model Uncertainty". In *Advances in Modeling and Simulation*, edited by A. Tolk, J. Fowler, G. Shao, and E. Yücesan, 219–247. Cham: Springer.

Zhu, C., R. H. Byrd, P. Lu, and J. Nocedal. 1997. "Algorithm 778: L-BFGS-B: Fortran Subroutines for Large-Scale Bound-Constrained Optimization". *ACM Transactions on Mathematical Software* 23(4):550–560.

AUTHOR BIOGRAPHIES

SEUNG MIN BAIK received the B.S. degree in industrial and management engineering from the Pohang University of Science and Technology (POSTECH), Pohang, South Korea, in 2016, where he is currently pursuing the Ph.D. degree in industrial and management engineering. His research interests include robust decision-making in stochastic systems and reliability engineering in stochastic simulation models. His email address is seungmin.baik@lstlab.org.

YOUNG MYOUNG KO is an Associate Professor in the Department of Industrial and Management Engineering at Pohang University of Science and Technology (POSTECH), Pohang, South Korea. He received the B.S. and M.S. degrees in industrial engineering from Seoul National University, Seoul, South Korea, and the Ph.D. degree in industrial engineering from Texas A&M University, College Station, TX, USA, in 1998, 2000, and 2011, respectively. His research interests include, but are not limited to, simulation and optimization of stochastic systems, such as telecommunication networks, ICT infrastructure, and renewable energy systems. His email address is youngko@postech.ac.kr and his homepage is https://www.lstlab.org/.

EUNSHIN BYON is an Associate Professor in the Department of Industrial and Operations Engineering at the University of Michigan. Her research interests include reliability evaluation, fault diagnosis/condition monitoring, predictive modeling and data analytics, and operations and maintenance decision-making for stochastic systems. Her recent research focuses on uncertainty quantification of stochastic systems using stochastic simulations, reliability analysis, and improvement of large-scale, interconnected systems with applications to renewable power systems and manufacturing processes. She is a member of IIE, INFORMS, and IEEE. Her email address is ebyon@umich.edu and her homepage is https://ebyon.engin.umich.edu/.