IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 28, 2019, accepted October 17, 2019, date of publication October 22, 2019,
date of current version November 4, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2948980

An Improved Budget Allocation Procedure for Selecting the Best-Simulated Design in the Presence of Large Stochastic Noise

SEON HAN CHOI[1], (Member, IEEE), CHANGBEOM CHOI[2], (Member, IEEE), AND TAG GON KIM[3], (Senior Member, IEEE)

[1]Department of IT Convergence and Application Engineering, Pukyong National University, Busan 48513, South Korea
[2]School of Global Entrepreneurship and ICT, Handong Global University, Pohang 37554, South Korea
[3]School of Electrical Engineering, Korea Advanced Institute of Science and Technology, Daejeon 34141, South Korea

Corresponding author: Changbeom Choi (cbchoi@handong.edu)

This work was supported by Electronics and Telecommunications Research Institute (ETRI) grant funded by the Korean government [19ZH1100, Distributed Intelligence Core Technology of Hyper-Connected Space].

ABSTRACT Motivated by the increasing practical needs for simulation optimization of modern industrial systems, this paper proposes an efficient ranking and selection (R&S) procedure for selecting the best-simulated design from a finite set of alternatives in the presence of large stochastic noise. To obtain the correct selection under a limited simulation budget, the proposed procedure sequentially allocates the budget to minimize the evaluated uncertainty values of the selection through a two-step process based on the existing uncertainty evaluation (UE) procedure. This two-step process reduces the inefficiency of the underlying UE procedure while keeping its high robustness to noise, thereby achieving improved the efficiency for the proposed procedure in a noisy environment. This improved efficiency is demonstrated in comparative experiments with other R&S procedures on several benchmark problems. In particular, the experimental results of three practical optimization problems emphasize the necessity of the proposed procedure.

INDEX TERMS Discrete-event system, high robustness to noise, optimization, ranking and selection, stochastic simulation.

I. INTRODUCTION

Discrete-event system simulation plays a key role in evaluating, analyzing, and optimizing the performance of modern industrial systems such as the telecommunication [1], military [2], manufacturing [3], smart grid [4], and transportation [5] systems, which scarcely meet the assumptions of closed-form analytic models [6]. However, conducting a simulation experiment can be expensive and time-consuming. Let a configuration of decision variables in the system be referred to as a design. Since simulations typically involve a stochastic noise to reflect the randomness and uncertainty of the real-world, many replicated simulations for one design are necessary to reduce noise and estimate the mean performance of the design. Suppose N simulation replications are required to obtain a precise estimation (i.e., the sample mean) of the performance for each design, and suppose k designs are given. Then, the total number of simulation replications T required for this experiment is kN. If the number of feasible designs is large (i.e., k is large) and the stochastic noise is large (i.e., N should be large to meet the precision requirement), T can be very large; this may easily result in a prohibitively high cost for the simulation experiment and preclude the use of simulation [7].

There are a variety of ways to reduce the cost of simulation experiments, but if k is less than a few thousand, it falls into the area of statistical ranking and selection (R&S) [8]. Instead of assigning the same number of simulation replications to all designs blindly, R&S procedures make an intelligent allocation. That is, based on the statistical inferences for the current simulation results of designs, R&S procedures allocate the replications to each design differently in order to select the best one correctly, thereby reducing the necessary cost dramatically.

The associate editor coordinating the review of this manuscript and approving it for publication was Fanbiao Li.

Thus far, various R&S procedures have been proposed. Kim and Nelson [9] proposed the two-stage indifference-zone (IZ) procedure to achieve a prespecified lower bound guarantee of the probability of correct selection $P\{\mathrm{CS}\}$ from the frequentist perspective. On the other hand, Chick and Wu [10] presented another IZ procedure that provides an upper bound guarantee of the expected opportunity cost E[OC], of a potentially incorrect selection. In addition, Chick and Inoue [11], [12] suggested the two-stage Bayesian procedure to minimize E[OC] based on the IZ procedure. Chick *et al.* [13] eliminated several approximations used in the previous Bayesian procedure and proposed a small-sample procedure that allocates a few replications sequentially. Chen *et al.* [14] presented the well-known optimal computing budget allocation (OCBA) procedure to maximize $P\{\mathrm{CS}\}$ under a simulation budget constraint. It allocates further replications sequentially according to the optimal rule that asymptotically maximizes an approximated lower bound of $P\{\mathrm{CS}\}$. Instead of maximizing $P\{\mathrm{CS}\}$, He *et al.* [15] proposed an OCBA-EOC procedure that sequentially minimizes E[OC] by using the OCBA approach to selection. (cf. [16] for extensive empirical comparisons of these R&S procedures.)

Recently, Choi and Kim [17] proposed the uncertainty evaluation (UE) procedure to resolve practical simulation optimization problems involving large stochastic noise efficiently. Similar to the OCBA procedure, the UE procedure also aims to maximize $P\{\mathrm{CS}\}$ under a limited number of replications. Applying the statistical hypothesis test and the p-value, it evaluates a measurement for each design called uncertainty that indicates the degree to which the current simulation results are significant evidence for verifying the correct selection. Then, it sequentially allocates further replications according to the evaluation results to maximize $P\{\mathrm{CS}\}$. In contrast with the OCBA procedure, which uses only the sample mean and variance values of the design for the allocation, the UE procedure considers the precision of the sample mean by using the number of replications assigned so far for each design additionally. Consequently, in the presence of large stochastic noise, the UE procedure can prevent the waste of replications caused by a poor value of the sample mean and improve $P\{\mathrm{CS}\}$; this has been demonstrated in several practical case studies [18]–[20].

While the UE procedure has high robustness to noise, it still suffers from the drawback of inefficiency caused by heuristics in the uncertainty evaluation for each design. As a direct example to show this, consider the simple problem of selecting the better of two design alternatives under limited replications. The UE procedure allocates the replications equally to both designs, like as the most inefficient equal allocation. The reason is that the evaluations of the uncertainty of the two designs are the same regardless of their current simulation results (see Section III-A for details). Theoretically, the optimal allocation that maximizes $P\{\mathrm{CS}\}$ for this problem is to allocate the replications depending on the variance of each design [8]. That is, the fact that the UE allocation does not converge to the theoretical optimal allocation in this example implies that there is room to further improve the efficiency of the UE procedure.

Recently, the increasing stochastic noise due to the higher complexity of systems has made simulation optimization more expensive. To deal with this problem efficiently, we propose an improved R&S procedure based on the UE procedure. The proposed procedure increases efficiency by improving the inefficiency of the previous procedure through a two-step sequential allocation process. Here, this two-step process differs from the two-stage IZ procedure in that it does not include elimination [9]. Compared with the UE procedure, the proposed procedure converges to the theoretical optimal allocation in the two-design case and is more efficient in the presence of large stochastic noise. We present experimental results of a comparison with other R&S procedures through practical simulation optimization problems as well as a series of numerical benchmarks.

The rest of the paper is organized as follows: Section II defines the problem in detail, and Section III proposes the R&S procedure. Section IV exhibits the experimental results. Finally, the conclusion is given in Section V.

II. PROBLEM DEFINITION

We use the following notations in this paper, where a bold typeface notation represents a vector.

- T — Total number of simulation replications.
- k — Number of design alternatives.
- $\boldsymbol{x}_i$ — Design (i.e., simulation input), where the subscript i represents the design index, $i \in \{1, \ldots, k\}$.
- Θ — Set of feasible designs, $\Theta = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k\}$.
- Y_{ij} — Simulation output of $\boldsymbol{x}_i$ in the jth replication, where $Y_{ij} \sim \mathcal{N}\left(\mu_i, \sigma_i^2\right)$.
- μ_i — Mean of Y_{ij} (i.e., the performance of $\boldsymbol{x}_i$), $\mu_i = \mathrm{E}\left[Y_{ij}\right]$.
- σ_i^2 — Variance of Y_{ij}, $\sigma_i^2 = \mathrm{Var}\left[Y_{ij}\right]$.
- N_i — Number of output samples (i.e., number of allocated replications at $\boldsymbol{x}_i$).
- $\bar{\mu}_i$ — Sample mean of $Y_{i1}, \ldots, Y_{iN_i}$ (i.e., the estimated performance of $\boldsymbol{x}_i$), $\bar{\mu}_i = 1/N_i \sum_{j=1}^{N_i} Y_{ij} \sim N\left(\mu_i, \sigma_i^2/N_i\right)$.
- s_i^2 — Sample variance of $Y_{i1}, \ldots, Y_{iN_i}$, $s_i^2 = \sum_{j=1}^{N_i} \left(Y_{ij} - \bar{\mu}_i\right)^2 \Big/ \left(N_i - 1\right)$.

We assume that the distribution of simulation output follows a normal distribution and that every output is independent across different replications and designs. This normality assumption is reasonable in practice because the output is typically obtained as an average value or batch means in stochastic simulations and thus the central limit theorem holds [8]. In addition, in contrast to the assumptions commonly used in R&S literature, we suppose that the variance of the output is unknown. That is, the proposed procedure can be effective in practical situations where no prior knowledge of the mean and variance of each design is given.

Assuming that a lower value is better in the performance of a design, the simulation optimization problem is defined as follows:

$$\boldsymbol{x}_b = \arg\min_{\boldsymbol{x}_i \in \Theta} \mu_i, \tag{1}$$

where $\boldsymbol{x}_b$ is the best design, which we have to select from Θ. This assumption can be made is without loss of generality since the proposed procedure is equally applicable to maximization problems, as are many R&S procedures that start from a minimization problem (see the practical case studies in Section IV-B).

However, in practice, the exact value of μ_i in (1) cannot be obtained; thus, we inevitably select the estimated best design $\boldsymbol{x}_e$ based on the sample mean $\bar{\mu}_i$,

$$\boldsymbol{x}_e = \arg\min_{\boldsymbol{x}_i \in \Theta} \bar{\mu}_i. \tag{2}$$

Then, the accuracy of this selection, $P\{\mathrm{CS}\}$ is defined as

$$\begin{aligned} P\{\mathrm{CS}\} &= P\{\boldsymbol{x}_e = \boldsymbol{x}_b\} \\ &= P\{\mu_e < \mu_i,\ i \neq e,\ \text{and } i = 1, \ldots, k\}. \end{aligned} \tag{3}$$

As the precision of $\bar{\mu}_i$ increases through many simulation replications, $P\{\mathrm{CS}\}$ can increase to 1; however, our objective is to maximize $P\{\mathrm{CS}\}$ under a limited simulation budget T in order to improve the efficiency of simulation. The R&S problem for achieving this objective is given as

$$\begin{aligned} &\arg\max_{N_1,\ldots,N_k} P\{CS\} \\ &\text{such that } \sum_{i=1}^{k} N_i = T \quad \text{and } N_i \geq 0. \end{aligned} \tag{4}$$

Here, the constraint $\sum_{i=1}^{k} N_i = T$ implicitly assumes that the simulation cost of each replication is roughly the same across designs.

III. PROPOSED PROCEDURE

As mentioned previously, this paper proposes an improved R&S procedure based on the UE procedure to solve the problem defined in (4). The first subsection briefly introduces the underlying UE procedure and the drawback of its inefficiency. Then, the improved procedure will be proposed in the second subsection.

A. THE UE PROCEDURE

A statistical hypothesis test statistically verifies the comparative relationship between unknown values μ_i and μ_j via the p-value calculated using the observed statistical data such as $\bar{\mu}_i$, $\bar{\mu}_j$, s_i, s_j, N_i, and N_j. For example, the test for verifying $\mu_i > \mu_j$ is as follows:

$$H_0 : \mu_i \leq \mu_j, \quad \underline{H_A : \mu_i > \mu_j}. \tag{5}$$

H_0 is the null hypothesis, and H_A is the alternative hypothesis, which is the relationship to be verified. The p-value for this test, $\varepsilon_{i,j}$ can be calculated as

$$\varepsilon_{i,j} = F_\nu\left(\left(\bar{\mu}_j - \bar{\mu}_i\right)\Big/ s_{i,j}\right) \quad \text{where } s_{i,j} = \sqrt{s_i^2 \Big/ N_i + s_j^2 \Big/ N_j}. \tag{6}$$

The function $F_\nu(\cdot)$ is the cumulative distribution function (CDF) of the t-distribution with ν degrees of freedom, where ν can be defined as

$$\nu = \left\lfloor \frac{s_{i,j}^4}{\left(s_i^2/N_i\right)^2 \Big/ (N_i - 1) + \left(s_j^2/N_j\right)^2 \Big/ (N_j - 1)} \right\rfloor. \tag{7}$$

The p-value $\varepsilon_{i,j}$ is defined as the probability of obtaining the currently observed data or more extreme data than the current data under the assumption that H_0 is true. Here, extreme means less likely to be obtained under this assumption. In the framework of the hypothesis test, $\varepsilon_{i,j}$ indicates the degree to which the observed data can be considered significant evidence to verify $\mu_i > \mu_j$. That is, a small value of $\varepsilon_{i,j}$ close to zero means that it is very unlikely (or almost impossible) for the current data to be observed when H_0 is true; thus, this actually observed data can be regarded as significant evidence to reject H_0 and accept H_A as true. On the other hand, a relatively large value of $\varepsilon_{i,j}$ close to 0.5 indicates that the data cannot be considered evidence to verify $\mu_i > \mu_j$, as it is probable for such data to be obtained when H_0 is true. In this case, it is uncertain whether $\mu_i > \mu_j$ is true [21].

Using such features of the hypothesis test and p-value, the UE procedure evaluates the uncertainty of each design as the criterion for allocating further replications. The uncertainty of a design is defined as the maximum p-value for the comparative relationships between that design and the others that must be verified for the correct selection of the design. Similar to the p-value, the uncertainty indicates the degree to which the simulation data of the design that has been obtained can be statistically significant evidence that the selection of the design is correct (cf. [17]).

For example, supposing that $\boldsymbol{x}_1$ is selected as the estimated best design of the three designs: $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, and $\boldsymbol{x}_3$, the following comparative relationships should be verified for the correct selection of $\boldsymbol{x}_1$: $\mu_2 > \mu_1$ and $\mu_3 > \mu_1$. Thus, the uncertainty of $\boldsymbol{x}_1$ is defined as the maximum value of $\varepsilon_{2,1}$ and $\varepsilon_{3,1}$ for verifying each relationship. A small value close to zero indicates that the observed data of $\boldsymbol{x}_1$ can be considered significant evidence to statistically confirm that $\boldsymbol{x}_1$ is the best design since the data are significant evidence for verifying both relationships. On the other hand, a relatively large value close to 0.5 means that at least one relationship is uncertain, based on the data. That is, the selection of $\boldsymbol{x}_1$ is also uncertain, so more replications should be allocated to update the data of $\boldsymbol{x}_1$ so that it can be considered significant evidence.

However, the uncertainty evaluation of the UE procedure involves heuristics that may result in inefficient allocation, which is to evaluate a single design using the p-value in (6) that has been calculated with the observed data of both designs. In other words, the evaluated uncertainty of the design indicates not only the significance degree of this design's data, but also that of another design's data. When the evaluated value is small, it does not have a significant impact on the allocation, but as the value becomes higher it may

lead to inefficient allocation. In the example above, suppose that $\boldsymbol{x}_1$ and $\boldsymbol{x}_3$ have already been allocated many replications, so that $\bar{\mu}_1$ and $\bar{\mu}_3$ are quite accurate, whereas $\boldsymbol{x}_2$ has not been allocated replications and hence its observed data are insufficient. Due to the inaccurate value of $\bar{\mu}_2$, the uncertainty of $\boldsymbol{x}_1$ (i.e., $\varepsilon_{2,1}$) is assigned a high value, and thus $\boldsymbol{x}_1$ is allocated further replications due to its high uncertainty. However, this additional assignment to $\boldsymbol{x}_1$ is excessive allocation that wastes the limited replications that should be allocated to $\boldsymbol{x}_2$. In short, the UE procedure does not fully exploit the nature of the p-value. The inefficient even allocation in the two-design case mentioned in the introduction also results from this.

B. THE IMPROVED R&S PROCEDURE

To improve such inefficiency, the proposed procedure allocates further replications through a two-step process. Applying the UE procedure, the first step assigns further replications depending on the evaluated uncertainty for each design. However, it allocates the replications to the pairs of designs involved in the uncertainty to fully exploit the p-value. Then, in the second step, the allocated replications at each pair are finally distributed to the two designs to minimize the uncertainty.

1) THE FIRST STEP

Depending on the definition of uncertainty, the uncertainty of each design is evaluated as follows. For $k-1$ designs that are not selected as the estimated best design $\boldsymbol{x}_e$, $\mu_i > \mu_e$ should be verified for the correct selection of each $\boldsymbol{x}_i$ (i.e., $\forall i \in \{1,\ldots,k\}$ and $i \neq e$: $\boldsymbol{x}_i \neq \boldsymbol{x}_b \Leftarrow \mu_i > \mu_e$); thus, the uncertainty of $\boldsymbol{x}_i$ is evaluated as $\varepsilon_{i,e}$. On the other hand, for the estimated best design $\boldsymbol{x}_e$, $\mu_i > \mu_e$ should be verified for the remaining $k-1$ designs for the correct selection of $\boldsymbol{x}_e$ (i.e., $\boldsymbol{x}_e = \boldsymbol{x}_b \Leftarrow (\mu_1 > \mu_e) \wedge \cdots \wedge (\mu_i > \mu_e) \wedge \cdots \wedge (\mu_k > \mu_e)$, $i \neq e$); thus, the uncertainty of $\boldsymbol{x}_e$ is evaluated as the maximum among the values of $\varepsilon_{i,e}$ for each $\mu_i > \mu_e$ (i.e., $\max\left(\varepsilon_{1,e},\ldots,\varepsilon_{i,e},\ldots,\varepsilon_{k,e}\right)$).

The first step in applying the UE procedure allocates further replications to each design depending on its evaluated uncertainty. That is, this step allocates more replications to designs with relatively larger uncertainty based on the meaning of the uncertainty. However, in contrast with the UE procedure, the allocated replications at each design in the first step are considered to be temporarily assigned to a pair of designs involved in its uncertainty. As mentioned before, the evaluated uncertainty of a design indicates the significance degree not only of this design's data but of both designs' data, based on the nature of the p-value in (6). To make the most of the characteristic of the p-value, the first step temporarily allocates additional replications to the pair of $\boldsymbol{x}_i$ and $\boldsymbol{x}_e$ rather than allocating additional replications to $\boldsymbol{x}_i$ depending on $\varepsilon_{i,e}$, the uncertainty of $\boldsymbol{x}_i$ calculated with the data of both $\boldsymbol{x}_i$ and $\boldsymbol{x}_e$. These allocated replications will finally be distributed to $\boldsymbol{x}_i$ and $\boldsymbol{x}_e$ to minimize $\varepsilon_{i,e}$ in the second step, thereby improving the inefficiency of the UE procedure.

Given the number of further replications Δ, the number of allocated replications, $a_{i,e}$, for each pair of $\boldsymbol{x}_i$ and $\boldsymbol{x}_e$ depending on the value of $\varepsilon_{i,e}$ in the first step, is calculated as follows [17]:

$$a_{i,e}/a_{j,e} = \varepsilon_{i,e}/\varepsilon_{j,e}$$

where

$$i,j \in \{1,\ldots,k\}, \quad i \neq j \neq e, \text{ and } \Delta = \sum_{i=1,i\neq e}^{k} a_{i,e}. \tag{8}$$

Here, Equation (8) involves a small adjustment that excludes the uncertainty of $\boldsymbol{x}_e$ from the allocation of Δ, in contrast with the UE procedure. Since the uncertainty of $\boldsymbol{x}_e$ is defined as $\max\left(\forall i \neq e : \varepsilon_{i,e}\right)$, the assignment to $\boldsymbol{x}_e$ depending on this uncertainty is considered a redundant allocation to one pair of $\boldsymbol{x}_i$ and $\boldsymbol{x}_e$ in the first step that results in inefficiency. Thus, the uncertainty of $\boldsymbol{x}_e$ is not used in (8). Because of this, even if $\boldsymbol{x}_e$ is not allocated further replications in the first step, it can be allocated sufficient replications to minimize its uncertainty through the second step.

2) THE SECOND STEP

Using Bonferroni's inequality, an approximated lower bound of $P\{\mathrm{CS}\}$ (APCS) can be defined as follows [14]:

$$\begin{aligned} P\{\mathrm{CS}\} &= P\left\{\bigcap_{i=1,i\neq e}^{k} (\tilde{\mu}_i > \tilde{\mu}_e)\right\} \\ &\geq 1 - \sum_{i=1,i\neq e}^{k} \left[1 - P\{\tilde{\mu}_i > \tilde{\mu}_e\}\right] \equiv \mathrm{APCS}. \end{aligned} \tag{9}$$

Here, $P\{\tilde{\mu}_i > \tilde{\mu}_e\}$ is the posterior probability of $\mu_i > \mu_e$ based on the observed data. Under the noninformative prior assumption mentioned in Section II, the results of Bayesian and frequentist inference have the same value [22]. That is, the value of $1 - P\{\tilde{\mu}_i > \tilde{\mu}_e\}$ is equal to the value of $\varepsilon_{i,e}$; thus, the APCS in (9) can be rewritten as follows:

$$P\{\mathrm{CS}\} \geq 1 - \sum_{i=1,i\neq e}^{k} \varepsilon_{i,e}. \tag{10}$$

As a result, $P\{\mathrm{CS}\}$ can be maximized when the uncertainty of each $\boldsymbol{x}_i$ other then $\boldsymbol{x}_e$ is minimized to zero. In the second step, the additional replications temporarily allocated to each pair $\boldsymbol{x}_i$ and $\boldsymbol{x}_e$ in the previous step, $a_{i,e}$ are optimally distributed between both designs to minimize each $\varepsilon_{i,e}$.

Let a_i and a_e be the number of further replications distributed to $\boldsymbol{x}_i$ and $\boldsymbol{x}_e$, respectively, in the second step (i.e., $a_i + a_e = a_{i,e}$). Then, the minimization problem of the second step is defined as follows:

$$\begin{aligned} &\arg\min_{a_i,a_e} \left(\varepsilon_{i,e}\big|_{a_i,a_e}\right) \\ &\text{such that } a_i + a_e = a_{i,e} \quad \text{and } a_i, a_e \geq 0, \end{aligned} \tag{11}$$

where $\varepsilon_{i,e}|_{a_i,a_e}$ is the change of $\varepsilon_{i,e}$ according to the further allocation of a_i and a_e. Unfortunately, a closed-form expression for $\varepsilon_{i,e}|_{a_i,a_e}$ cannot be obtained since the simulation outputs of $\boldsymbol{x}_i$ and $\boldsymbol{x}_e$ for the additional allocation of a_i and a_e

cannot be predicted. However, if a_i and a_e are not too large compared to N_i and N_e, the number of allocated replications to $\boldsymbol{x}_i$ and $\boldsymbol{x}_e$ so far (i.e., $a_i \ll N_i$ and $a_e \ll N_e$), it can be assumed that the statistical data of both designs will not change too much after conducting further simulations. That is, the observed data after further simulations of a_i and a_e for each design can be reasonably expected to $\bar{\mu}_i$, $\bar{\mu}_e$, s_i, s_e, $N_i + a_i$, and $N_e + a_e$; thus, $\varepsilon_{i,e}|_{a_i,a_e}$ can be approximated using (6) and (7) as follows:

$$\varepsilon_{i,e}\,|_{a_i,a_e} \approx F_\nu\left(\frac{\bar{\mu}_e - \bar{\mu}_i}{\sqrt{s_i^2/(N_i+a_i) + s_e^2/(N_e+a_e)}}\right)$$

where

$$\nu = \left\lfloor \frac{\left(s_i^2/(N_i+a_i) + s_e^2/(N_e+a_e)\right)^2}{\frac{\left(s_i^2/(N_i+a_i)\right)^2}{(N_i+a_i-1)} + \frac{\left(s_e^2/(N_e+a_e)\right)^2}{(N_e+a_e-1)}} \right\rfloor. \tag{12}$$

Based on (12), the problem given in (11) can be approximated as follows:

$$\begin{aligned} &\arg\min_{a_i,a_e} \; F_\nu\left(\frac{\bar{\mu}_e - \bar{\mu}_i}{\sqrt{s_i^2/(N_i+a_i) + s_e^2/(N_e+a_e)}}\right) \\ &\text{such that } a_i + a_e = a_{i,e} \quad \text{and } a_i, a_e \geq 0. \end{aligned} \tag{13}$$

A lower bound for the degree of freedom ν in (12) is defined as $\min(N_i + a_i - 1, N_e + a_e - 1)$ [23]. Due to the nature of the R&S procedure, every design is allocated initial replications of n_0 to obtain the minimum data for the additional allocations, where n_0 has a value of 10 or more by default (see the next subsection in detail). That is, N_i and N_e are greater than 10; thus, it can be supposed that ν is at least as large as 10 depending on its lower bound. The CDF of the t-distribution with $\nu = 10$ is almost the same as that of the function with $\nu \to \infty$ (i.e., the CDF of the standard normal distribution $\Phi(\cdot)$). Actually, the cosine similarity between the two functions is 0.9987. That is, the impact of varying ν based on a_i and a_e on $F_\nu(\cdot)$ is insignificant. As a result, we neglect this to simplify the problem. Since $F_\nu(\cdot)$ is a monotonically increasing function and $\bar{\mu}_i$ is always larger than $\bar{\mu}_e$ (i.e., $\bar{\mu}_i > \bar{\mu}_e$), the problem in (13) can be simplified as follows:

$$\begin{aligned} &\arg\min_{a_i,a_e} \; \frac{s_i^2}{N_i+a_i} + \frac{s_e^2}{N_e+a_e} \\ &\text{such that } a_i + a_e = a_{i,e} \quad \text{and } a_i, a_e \geq 0. \end{aligned} \tag{14}$$

Plugging the constraint $a_e = a_{i,e} - a_i$ into (14), we have,

$$\arg\min_{a_i} \frac{s_i^2}{N_i+a_i} + \frac{s_e^2}{N_e+a_{i,e}-a_i} \quad \text{such that } 0 \leq a_i \leq a_{i,e}. \tag{15}$$

Let $s_i^2/(N_i+a_i) + s_e^2/(N_e+a_{i,e}-a_i)$ in (15) be the function $g(a_i)$. Since $g(a_i)$ is twice continuously differentiable and its second derivative is nonnegative in the interval $(-N_i, N_e+a_{i,e})$ involving the constraint of a_i, $g(a_i)$ is a strictly convex function on this interval; thus, a solution c that minimizes $g(a_i)$ exists within this interval and can be obtained as follows:

$$c = \frac{s_i\left(N_e + a_{i,e}\right) - s_e N_i}{s_i + s_e}. \tag{16}$$

Hence, depending on the constraint of a_i in (15), an optimal solution of (15) can be defined as follows:

$$a_i = \begin{cases} 0 & c \leq 0 \\ c & 0 < c < a_{i,e} \\ a_{i,e} & a_{i,e} \leq c. \end{cases} \tag{17}$$

The results are summarized as the optimal allocation rule of the second step in the following theorem.

Theorem 1: For the further replications $a_{i,e}$ allocated to the pair $\boldsymbol{x}_i$ and $\boldsymbol{x}_e$, the uncertainty of each $\boldsymbol{x}_i$, $\varepsilon_{i,e}$ can be approximately minimized by distributing $a_{i,e}$ to $\boldsymbol{x}_i$ and $\boldsymbol{x}_e$ as follows:

$$\begin{cases} a_i = 0 \text{ and } a_e = a_{i,e} & c \leq 0 \\ (N_i+a_i)/(N_e+a_e) = s_i/s_e & 0 < c < a_{i,e} \\ a_i = a_{i,e} \text{ and } a_e = 0 & a_{i,e} \leq c. \end{cases} \tag{18}$$

In summary, the number of given further replications Δ is allocated to each design through applying (8) in the first step and (18) in the second step sequentially. Let δ_i denote the number of additional replications assigned to a design $\boldsymbol{x}_i$ via the two-step process (i.e., $\Delta = \sum_{i=1}^{k} \delta_i$). For $k-1$ designs that are not selected as the estimated best design $\boldsymbol{x}_e$ (i.e., $i \neq e$), δ_i is calculated as follows according to (8) and (18):

$$\delta_i = \min\left(\max\left(\left\lceil \frac{\left(a_{i,e} + N_i + N_e\right) s_i}{s_i + s_e} \right\rceil - N_i, 0\right), a_{i,e}\right),$$

where

$$i \neq e \quad \text{and } a_{i,e} = \Delta \cdot \varepsilon_{i,e} \Big/ \sum_{j=1, j\neq e}^{k} \varepsilon_{j,e}. \tag{19}$$

On the other hand, for the estimated best design $\boldsymbol{x}_e$ (i.e., $i = e$), $\boldsymbol{x}_e$ is allocated further replications simultaneously from every pair $\boldsymbol{x}_i$ and $\boldsymbol{x}_e$ in the second step because it is involved in all $\varepsilon_{i,e}$s. These further replications allocated from each pair are independent from each other; thus, we heuristically define δ_e, the number of further replications allocated to $\boldsymbol{x}_e$, as their sum:

$$\begin{aligned} \delta_e &= \sum_{i=1, i\neq e}^{k} \min\left(\max\left(\left\lceil \frac{\left(a_{i,e} + N_i + N_e\right) s_e}{s_i + s_e} \right\rceil - N_e, 0\right), a_{i,e}\right) \\ &= \Delta - \sum_{i=1, i\neq e}^{k} \delta_i. \end{aligned} \tag{20}$$

Since the observed data of $\boldsymbol{x}_e$ is used to calculate every $\varepsilon_{i,e}$, this heuristic method can additionally reduce the uncertainties by allocating more replications to $\boldsymbol{x}_e$. That is, each $\varepsilon_{i,e}$ might be further reduced due to the $\delta_e - a_e$ additionally allocated to $\boldsymbol{x}_e$ along with the $a_{i,e}$ from the perspective of each pair.

Remark 1: In the special case where $k = 2$ and the variance of the two designs is known, the theoretically optimal allocation for given replications T can be obtained as [8]:

$$N_1/N_2 = \sigma_1/\sigma_2 \quad \text{where } N_1 + N_2 = T. \tag{21}$$

In this case, the allocation obtained through (19) and (20) is identical to this theoretical optimum. Since there is only one pair $\boldsymbol{x}_i$ and $\boldsymbol{x}_e$ when $k = 2$, the first step becomes meaningless, and thus further replications are allocated to the two designs depending on Theorem 1, where the sample variance in (18) is replaced with the known variance. In the initial situation where no replication is allocated to both designs (i.e., $N_i = N_e = 0$), the value of c in (16) is always within the interval $(0, a_{i,e})$; thus, according to the equation in the middle of (18), the given replications $a_{i,e}$ are distributed depending on the known variance of both design, as in (21).

3) SEQUENTIAL PROCEDURE

Similar to general R&S procedures, the proposed procedure divides the given limited replications T into smaller units Δ and allocates them sequentially. That is, until T is depleted, the procedure iterates the allocations of Δ in accordance with (19) and (20) and updates the observed data of the designs with these additionally allocated replications. This sequential allocation enables the procedure to use T efficiently and provides robustness to noise [8]. Algorithm 1 represents the sequential procedure.

Algorithm 1 Select the Best Design out of k Alternatives

Control parameters: T ($\geq kn_0$), n_0, and Δ

Procedures:

1: **simulate** n_0 times for each $\boldsymbol{x}_i$, $i \in \{1, \ldots, k\}$
2: **update** $\bar{\mu}_i$, s_i, and N_i for $\forall i$
3: **select** $\boldsymbol{x}_e$ with (2)
4: **while** $\sum_{i=1}^{k} N_i < T$ **do**
5: **set** $\Delta \leftarrow \min\left(T - \sum_{i=1}^{k} N_i, \Delta\right)$
6: **calculate** δ_i for $\forall \boldsymbol{x}_i \in \Theta, \boldsymbol{x}_i \neq \boldsymbol{x}_e$ with (19)
7: **calculate** δ_e for $\boldsymbol{x}_e$ with (20)
8[a]: **simulate** round(δ_i) times for each $\boldsymbol{x}_i$, $i \neq e$
9[a]: **simulate** $\Delta - \sum_{i=1, i\neq e}^{k}$ round (δ_i) times for $\boldsymbol{x}_e$
10: **update** $\bar{\mu}_i$, s_i, and N_i for $\forall i$
11: **select** $\boldsymbol{x}_e$ with (2)
12: **end while**
13: **return** $\boldsymbol{x}_e$

The control parameter n_0 is the number of initial replications allocated to every design to obtain the required minimum data for further allocation before beginning the iteration. Thus, if the value of n_0 is too small, poor initial data can be obtained, which results in bad allocations in the iteration and decreases $P\{\text{CS}\}$. In this paper, we recommend a value of at least 10 based on the literature [24], [25] and empirical studies [17], [18]. As n_0 increases, $P\{\text{CS}\}$ can increase due to the adequate further allocations based on the more precise initial data. However, an excessively large value of n_0 compared to a given value of T reduces $P\{\text{CS}\}$ by wasting some of the replications required for the iteration. Fig. 1 illustrates the trend of $P\{\text{CS}\}$ with respect to the value of n_0. As a result, n_0 should be set to a value of 10 or greater in consideration of T.

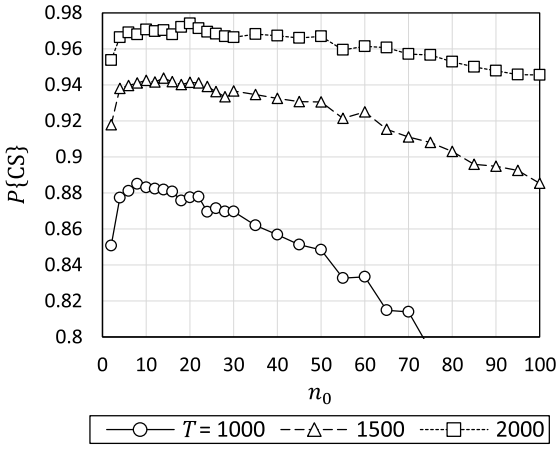

FIGURE 1. The graph illustrates the estimated value of $P\{\text{CS}\}$ versus n_0 for the specific value of T when Δ is fixed at 20. The equal variance (EV) benchmark problem in Table 1 is used, and $P\{\text{CS}\}$ was estimated over 10,000 independent repeated experiments.

On the other hand, the other parameter Δ is the number of further replications to be allocated via (19) and (20) in each iteration. It also should be set to a proper value by taking k and T into consideration. A value that is too small for a given k keeps the proposed procedure from allocating enough replications to reduce the uncertainties. A value that is too large for a given T reduces the advantages of the sequential procedure. That is, both extreme cases decrease $P\{\text{CS}\}$, as shown in Fig. 2. However, unless Δ has a very extreme value for the given k or T, its effect on $P\{\text{CS}\}$ is relatively insignificant compared to that of n_0 due to the iterative allocations of the sequential procedure. In this paper, we suggest a value between $0.5k$ and $0.05T$ as the proper value of Δ, based on empirical studies [17], [18].

Meanwhile, allocating Δ in the proposed procedure can be processed by calculating (19) $k-1$ times and (20) once. That is, the computational complexity of the additional allocation is $O(k)$, which is the same as in other R&S procedures such as the UE and OCBA procedures. However, when calculating (19), it is necessary to determine the value of the t-distribution's CDF (i.e., $F_\nu\left(\left(\bar{\mu}_j - \bar{\mu}_i\right)/s_{i,j}\right)$) to evaluate the uncertainty $\varepsilon_{i,e}$. As a result, depending on the implementation, the actual computation time of the proposed procedure may be slightly longer than that of the OCBA procedure, which uses only arithmetic operations for calculating the

[a]We use the round function to convert the real numbers δ_i and δ_e to integers and added the rounding errors to $\boldsymbol{x}_e$, which is usually allocated the largest portion of Δ. Due to the robustness of the sequential allocation, any reasonable conversion method, such as the ceiling or the floor function, has no significant impact on maximizing $P\{\text{CS}\}$.

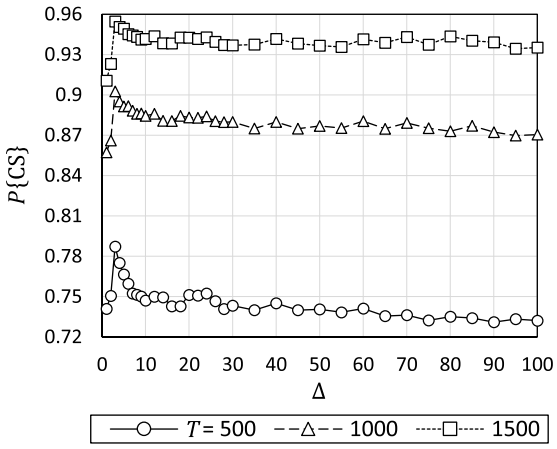

FIGURE 2. The graph illustrates the estimated value of $P\{CS\}$ versus Δ for the specific value of T when n_0 is fixed at 10. The equal variance (EV) benchmark problem in Table 1 is used, and $P\{CS\}$ was estimated over 10,000 independent repeated experiments.

allocation of Δ. In addition, the computation time of the proposed procedure is somewhat longer than that of the UE procedure due to the added second step. However, as mentioned previously, due to the higher complexity of modern industrial systems, the cost per simulation replication is relatively expensive and gradually increasing; thus, this slightly increase in the computation time of the proposed procedure becomes negligible compared to the total cost of the simulation experiment. Actually, while it took approximately 24 hours to solve the taxicab surplus problem in Section IV-B by applying the proposed procedure, the computation time of the procedure itself was only a few seconds, which is meaningless. Most of the time was spent running approximately 43,000 replications of the agent-based taxicab allocation simulator in Fig. 5.

IV. EXPERIMENTS

A. NUMERICAL EXPERIMENTS

In this subsection, we exhibit experimental results that demonstrate the improved efficiency of the proposed procedure in comparison with other R&S procedures developed to solve (4), such as the UE procedure [17], the OCBA procedure [14], the proportional-to-variance (PTV) procedure [8], and the equal allocation procedure. We used three benchmark problems that were applied to evaluate these procedures: the equal variance (EV), increasing variance (IV) and decreasing variance (DV) problems. Table 1 summarizes these problems. They each have 10 design alternatives, of which the best design to be selected is $\boldsymbol{x}_1$. In the EV problem, all designs have the same variance, whereas in the IV and DV problems the designs have an increasing or decreasing variance, respectively, as the design index decreases. Each problem has two versions: the large-noise version has relatively larger stochastic noise compared with the original version because it reduces the gap in μ_i while keeping the same value of σ_i.

TABLE 1. Three numerical benchmark problems.

Prob.	k	Original version		Large-noise version (L)	
		μ_i	σ_i	μ_i	σ_i
EV	10	$\mu_i = i$	$\sigma_i = 10$	$\mu_i = \mathbf{0.25} \times i$	$\sigma_i = 10$
IV	10	$\mu_i = i$	$\sigma_i = 5 + i$	$\mu_i = \mathbf{0.25} \times i$	$\sigma_i = 5 + i$
DV	10	$\mu_i = i$	$\sigma_i = 16 - i$	$\mu_i = \mathbf{0.25} \times i$	$\sigma_i = 16 - i$

TABLE 2. Value of T required to achieve $P\{CS\}$ of 0.99 for benchmark problems.

Problems		Proposed	UE	OCBA	PTV	Equal
EV	T req.	2,600	2,840	3,600	9,000	10,800
	Ratio[a]	-	**1.09**	**1.38**	**3.46**	**4.15**
EV(L)	T req.	43,000	50,500	72,000	145,700	171,500
	Ratio	-	**1.17**	**1.67**	**3.39**	**3.99**
IV	T req.	1,360	1,440	1,600	8,400	4,600
	Ratio	-	**1.06**	**1.18**	**6.18**	**3.38**
IV(L)	T req.	22,000	26,200	39,400	120,000	70,000
	Ratio	-	**1.19**	**1.79**	**5.45**	**3.18**
DV	T req.	5,420	6,120	7,180	12,600	24,100
	Ratio	-	**1.13**	**1.32**	**2.32**	**4.45**
DV(L)	T req.	84,500	103,200	128,500	188,500	345,000
	Ratio	-	**1.22**	**1.52**	**2.23**	**4.08**

[a] This denotes the ratio of the value of T required in the proposed procedure and that required in the other procedure (i.e., $T_{\text{other}}/T_{\text{proposed}}$).

For a fair comparison, the values of n_0 and Δ in every procedure were set to 10 and 20, respectively, according to [8], [14], and [17]. We estimated $P\{CS\}$ for each procedure with 10,000 independent repeated experiments while varying T; the results are shown in Fig. 3. In addition, to numerically verify the improved efficiency of the proposed procedure, Table 2 exhibits the value of T required to select the best design correctly (i.e., to achieve $P\{CS\}$ of 0.99) for each problem. The experimental results shown in Fig. 3 and Table 2 clearly indicate the improved efficiency of the proposed procedure in comparison with the UE, OCBA, PTV, and equal allocation procedures. In particular, the proposed procedure demonstrated relatively higher efficiency in the presence of large stochastic noise, as shown in the large-noise version of the benchmark problems (see the shaded areas in Fig. 3 and Table 2).

The relatively high robustness to noise of the proposed procedure can primarily be attributed to the allocation according to the uncertainty of the UE procedure in the first step, which takes into account the precision of the sample mean. For example, when the observed value of $\bar{\mu}_1$ is abnormally high in the EV problem (actually, $\boldsymbol{x}_1$ is the best design with the lowest μ_1), the OCBA procedure, which does not consider the precision of $\bar{\mu}_1$, cannot allocate further replications to $\boldsymbol{x}_1$ based on the poor value of $\bar{\mu}_1$. Thus, this erroneous value of $\bar{\mu}_1$ is not improved, leading to a decrease in $P\{CS\}$. If the additional allocation continues indefinitely, the $\boldsymbol{x}_1$ will eventually be allocated further replications due to changes in the simulation results of the other designs; however, numerous replications are wasted on other marginal designs in the meantime (refer

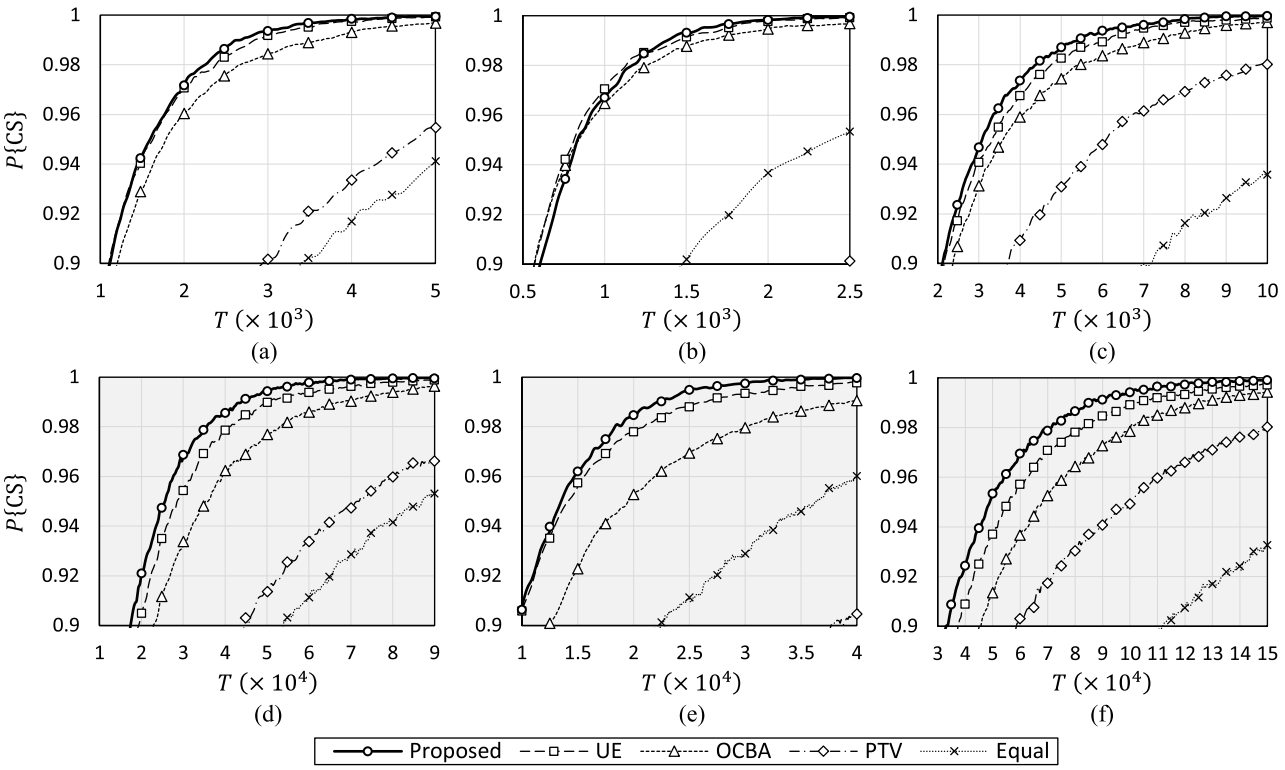

FIGURE 3. Graphs (a)-(f) illustrate the estimated value of $P\{CS\}$ versus T for the three benchmark problems, where the shaded graphs represent the results for the large-noise version of each benchmark: (a) EV, (b) IV, (c) DV, (d) EV(L), (e) IV(L), and (f) DV(L).

to [17] for more details). However, the proposed procedure considers the precision of the sample mean via the uncertainty based on the p-value calculated with the sample mean and the standard error, as shown in (6). Thus, it can allocate further replications to $\boldsymbol{x}_1$ for which the precision of $\bar{\mu}_1$ is relatively low, thereby quickly lowering the abnormally high value of $\bar{\mu}_1$ and making the correct selection. As the stochastic noise increases, such a poor value of the sample mean is obtained frequently; therefore, the proposed procedure becomes relatively more efficient (compare the results of the large-noise version with the original one in Fig. 3 and Table 2).

Furthermore, the proposed procedure achieves improved efficiency over the UE procedure via the second-step allocation. The UE procedure simply allocates further replications to each design depending on its evaluated uncertainty, whereas the proposed procedure optimally redistributes these allocated replications to each design based on the uncertainty of that design in the second step. That is, by fully exploiting the nature of the uncertainty based on the p-value in (6), the second step distributes the replications assigned at the first step to the two designs involved in the uncertainty so that this uncertainty is minimized. Therefore, the proposed procedure can further increase $P\{\mathrm{CS}\}$ compared to the UE procedure in the presence of large stochastic noise.

B. PRACTICAL CASE STUDIES

As mentioned earlier, we proposed this improved procedure to efficiently resolve practical optimization problems with large stochastic noise due to the high complexity of modern industrial systems. To demonstrate the necessity of the proposed procedure, we have applied it to three practical problems. A brief description of these problems is as follows.

1) MILES GEAR DESIGN PROBLEM

The multiple integrated laser engagement system (MILES) provides gear for military training [26]. MILES gear consists of a laser generator attached to an actual weapon and multiple sensors attached to a trainee's body. Since real bullets are not used, a hit is considered to have occurred when the generated laser beam is sensed by the sensor; thus, to give an immersive experience to trainees, we need to design MILES gear to maximize the hit rate, as with an actual weapon. There are two design factors: beam width w and sensing angle a, as shown in Fig. 4. The beam width refers to the diameter of the beam area that has an energy level greater than the sensor's threshold at 250 m. The sensing angle means the limit of the incidence angle of the beam that can be detected by the sensor. Depending on the combination of these two factors, a set of 195 feasible designs for MILES gear was given (i.e., $\Theta = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{195}\}$, where $\boldsymbol{x}_i = [w, a]$). The hit rate of the MILES gear for each design is estimated with many independent repeated simulations using the optical engineering (OE) simulator [27] in Fig. 4. Consequently, the MILES gear design problem boils down to the simulation optimization problem for selecting the best design with the highest hit rate out of 195 alternatives.

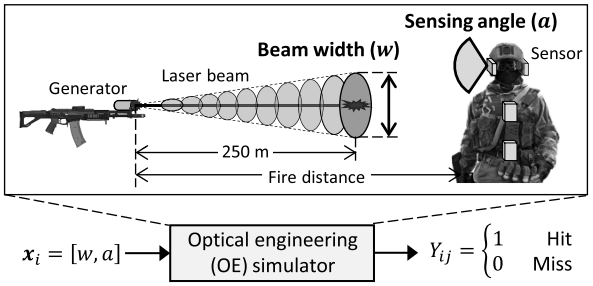

FIGURE 4. The OE simulator computes whether a beam with width of w launched randomly toward a target is sensed by the sensors attached to the target, where the sensing angle of the sensors is less than or equal to the value of a.

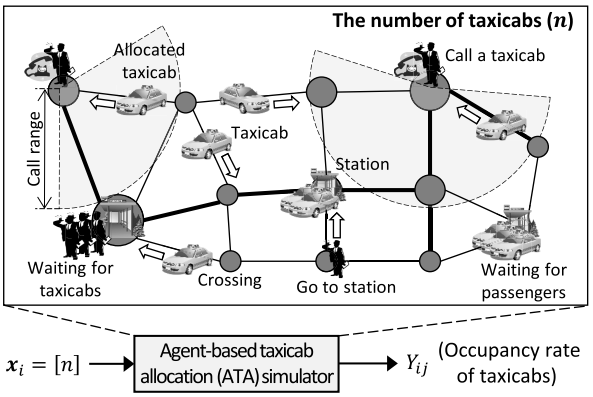

FIGURE 5. The ATA simulator provides the occupancy rate of taxicabs using rule-based taxicab agents that move on the graph representing the traffic environment of Daejeon. Here, passengers and their destinations are randomly generated at each vertex based on the statistics in the government report.

2) TAXICAB SURPLUS PROBLEM IN DAEJEON, SOUTH KOREA

Recently, the demand for taxicabs in Daejeon, South Korea, has gradually decreased due to the development of various means of transportation such as car-sharing services and personal mobility. As a result, the occupancy rate of taxicabs continues to decrease, which leads to a vicious cycle in which the income of taxi drivers is reduced, the quality of service is reduced, and consequently, the demand is further reduced. In addition, surplus taxicabs cause unnecessary traffic congestion and sometimes accidents due to competition for passengers. Although various efforts have been made to increase the occupancy rate of taxicabs, reducing the number of taxicabs is the most direct and effective way to accomplish this. Unfortunately, since it is impossible to use real experiments to find the optimal number of taxicabs to maximize the occupancy rate, this problem becomes a simulation optimization problem. For this scenario, 100 design alternatives were used, with each design containing the number of taxicabs n discretized in 100-unit intervals from 100 to 10000 (i.e., $\Theta = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{100}\}$, where $\boldsymbol{x}_i = [n]$). Many replications of the agent-based taxicab allocation (ATA) simulator [28] in Fig. 5 can be used to estimate the occupancy rate for each design. As a result, the second problem is defined as selecting the best design with the highest occupancy rate of taxicabs, out of 100 alternatives.

3) TACTICAL EVASION MANEUVER OPTIMIZATION PROBLEM

In modern warfare, domination of the air is directly linked to victory in war, so it is crucial to increase the survival rate of fighter aircraft, even if the increase is just one percent. Missiles, which are the largest threat to survival, are typically much faster than fighters and thus have that drawback of not being able to rapidly change their direction. The tactical evasion maneuver (TEM) shown in Fig. 6 allows a fighter to avoid an approaching missile by taking advantage of this drawback. When a missile comes within 20 km of a fighter, the fighter descends while maintaining the leading angle. Then, if the missile comes within the soaring distance, the fighter radically soars up using the maximum G factor, and the missile misses the fighter due to its inability to change direction rapidly. To increase the survival rate of fighter aircraft, we need to find the optimal values of the leading angle l and the soaring distance s that maximize the success rate of the TEM. Depending on the combination of these two factors of the TEM, a set of 50 feasible designs was given (i.e., $\Theta = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{50}\}$, where $\boldsymbol{x}_i = [l, s]$). The success rate for each TEM design is estimated with many independent repeated simulations using the TEM simulator [29] in Fig. 6. Thus, the third problem is defined as selecting the best design with the highest success rate out of 50 alternatives.

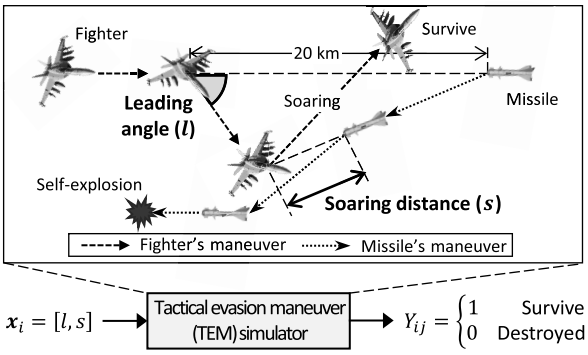

FIGURE 6. The TEM simulator determines whether a fighter can survive an attack from an approaching missile using the TEM with leading angle of l and soaring distance of s. Due to the randomness involved in the representation of control errors and environmental disturbances in the simulation, each replication gives different results.

We applied the proposed procedure to solve these practical problems efficiently. However, it is often difficult to determine T in practice. In addition, continuously allocating further replications until the given T is exhausted can waste the limited budget if the correct selection has already been made. To prevent this, the following stop condition using the APCS defined in (9) is added before line 12 of Algorithm 1: **if** APCS $\geq \tau$ then **break**. That is, the iteration for further allocation of Δ is terminated when the calculated APCS is above the target value τ, where τ was set to 0.99 for the

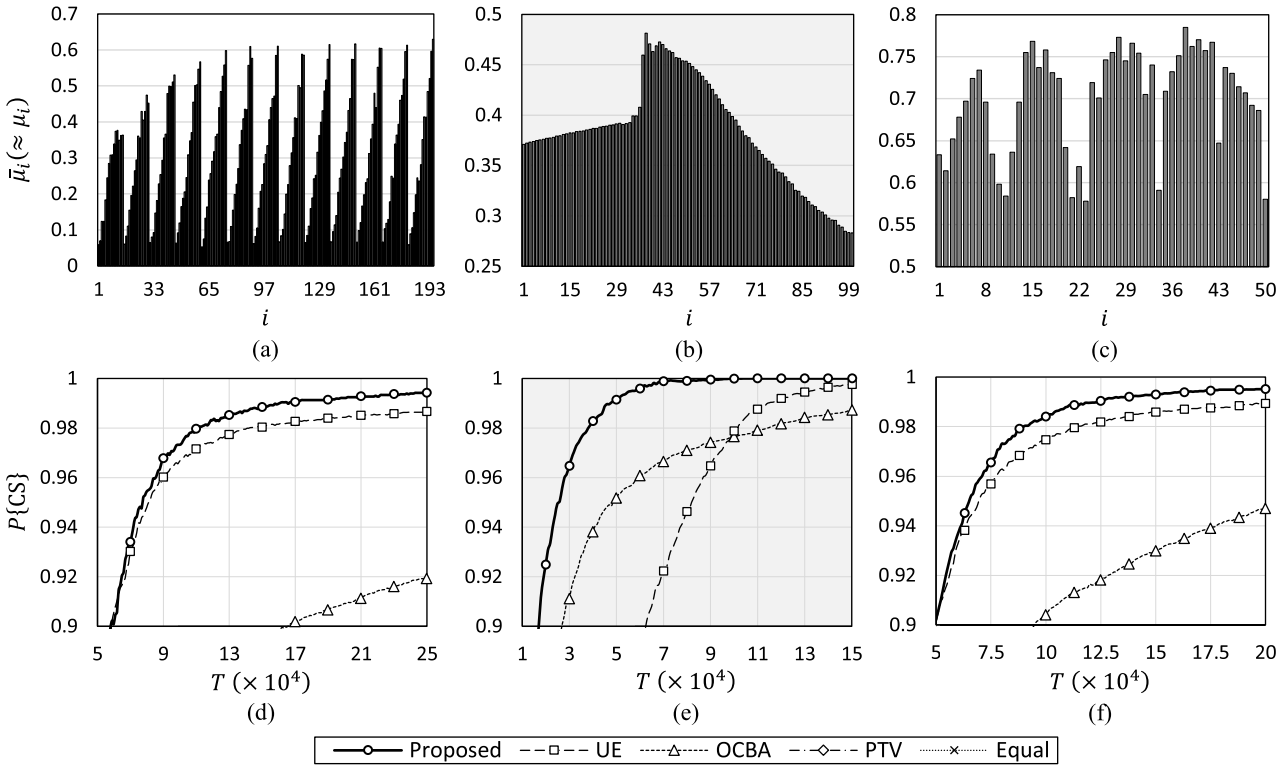

FIGURE 7. Graphs (a)-(c) illustrate the precisely estimated performance of each design for the three practical problems: (a) the hit rate of MILES gear for the 195 designs, (b) the occupancy rate of taxicabs for the 100 designs, and (c) the success rate of TEM for the 50 designs. Graphs (d)-(e) show the estimated value of $P\{CS\}$ versus T for these problems: (d) the MILES gear design problem, (e) the taxicab surplus problem in Daejeon, and (f) the TEM optimization problem.

TABLE 3. Selected best design for practical problems.

Problem	Selected best design
MILES gear design problem	$\boldsymbol{x}_{195} = [100\text{ cm}, 90\text{ °}]$
Taxicab surplus problem	$\boldsymbol{x}_{38} = [3800]$
TEM optimization problem	$\boldsymbol{x}_{38} = [70\text{ °}, 9000\text{ m}]$

correct selection. Based on the guidelines in [17], the values of n_0 and Δ were set to 20 and 100, respectively, in the first and second problems and to 20 and 50, respectively, in the third problem. As a result, we could solve these problems efficiently using the proposed procedure; Table 3 shows the selected best design for each problem.

To verify the efficiency of the proposed procedure quantitatively, we applied the other procedures used in the previous numerical experiments. In each procedure, the stop condition using the APCS was applied, and the same configuration of n_0 and Δ was used for a fair comparison. Table 4 illustrates the average value of T used with each procedure to achieve APCS of 0.99 for the three problems, where the average values were estimated through 1,000 independent repeated experiments. In addition, we estimated $P\{CS\}$ for each procedure while varying T, as shown in Figs. 7(d)-(f).

As shown in Figs. 7(a)-(c), these practical problems have large stochastic noise that is caused by the many near-optimal designs and relatively large variances. That is, numerous simulation replications will be required to obtain the correct

TABLE 4. Average value of T used to achieve APCS of 0.99 for practical problems.

Problem		Proposed	UE	OCBA	PTV	Equal
MILES gear design	$P\{CS\}$[a]	1	1	0.999	0.9023[b]	0.8529
	Avg. T used	152,719	168,496	306,657	10^6	10^6
	Ratio[c]	-	**1.10**	**2.01**	**6.55**	**6.55**
Taxicab surplus problem	$P\{CS\}$	1	0.996	0.996	1	0.9205
	Avg. T used	42,848	72,829	74,906	733,529	10^6
	Ratio	-	**1.70**	**1.75**	**17.12**	**23.34**
TEM optim. problem	$P\{CS\}$	1	0.999	0.996	0.997	0.999
	Avg. T used	116,618	121,547	142,512	980,031	856,430
	Ratio	-	**1.04**	**1.22**	**8.40**	**7.34**

[a] The estimated value of $P\{CS\}$ when the procedure achieves APCS of 0.99. Since the APCS is a lower bound of $P\{CS\}$, the estimated value of $P\{CS\}$ is greater than 0.99.

[b] Even though 1,000,000 replications were allocated, the procedures in the shaded area could not achieve APCS of 0.99; thus, the estimated value of $P\{CS\}$ is lower than 0.99.

[c] This denotes the ratio of the average value of T used in the proposed procedure and that used in the other procedure (i.e., $T_{other}/T_{proposed}$).

selection in these problems. For example, the taxicab surplus problem requires T of more than 10^6 to achieve $P\{CS\}$ of 0.99. This takes approximately 550 hours or more since the ATA simulator is a large-scale simulation involving many agents (i.e., about two seconds per replication of the ATA simulator). The comparison results in Table 4 and Figs. 7(d)-(f) demonstrate the improved efficiency of the proposed procedure in the presence of such large noise. While the UE

procedure took approximately 40 hours, on average, to obtain the correct selection in the taxicab surplus problem, the proposed procedure reduced this to less than one day. It also saved half of the simulation budget that the OCBA procedure used for the MILES gear design problem. This superior efficiency, along with its high robustness to noise and easy implementation, emphasizes the necessity of the proposed procedure for practical optimization problems of modern complex systems.

V. CONCLUSION

This paper proposed an improved R&S procedure based on the UE procedure to efficiently resolve practical optimization problems involving large stochastic noise. To select the best design correctly from a finite set of alternatives, the proposed procedure sequentially assigns a limited number of simulation replications through a two-step allocation process. The first step in applying the UE procedure temporarily allocates further replications to all pairs consisting of a design and the estimated best design depending on the uncertainty of each design. Then, in the second step, the replications allocated to each pair are optimally distributed between the two designs to minimize the uncertainty. This two-step process enables the proposed procedure to fully exploit the nature of the uncertainty based on the p-value, thereby increasing $P\{\mathrm{CS}\}$ while maintaining the high robustness to noise of the underlying UE procedure. The experimental results for benchmark and practical problems demonstrate this improved efficiency in the presence of large stochastic noise in comparison to existing R&S procedures. It is expected that the proposed procedure will allow simulation to be applied more effectively in the fourth industrial revolution.

ACKNOWLEDGMENT

The authors would like to thank the Editor, the Associate Editor, and anonymous reviewers for their valuable comments and constructive suggestions.

REFERENCES

[1] L. Wang, R. Iida, and A. M. Wyglinski, "Vehicular network simulation environment via discrete event system modeling," *IEEE Access*, vol. 7, pp. 87246–87264, 2019.

[2] B. G. Kang, K.-M. Seo, and T. G. Kim, "Model-based design of defense cyber-physical systems to analyze mission effectiveness and network performance," *IEEE Access*, vol. 7, pp. 42063–42080, 2019.

[3] Y. Liu, K. Wang, L. Ge, L. Ye, and J. Cheng, "Adaptive evaluation of virtual machine placement and migration scheduling algorithms using stochastic Petri nets," *IEEE Access*, vol. 7, pp. 79810–79824, 2019.

[4] S. H. Mousavi, M. Davari, and H. J. Marquez, "An innovative event-based filtering scheme using H_∞ performance for stochastic LTI systems considering a practical application in smart modernized microgrids," *IEEE Access*, vol. 7, pp. 48138–48150, 2019.

[5] S. Yang, K. Song, and G. Zhu, "Stochastic process and simulation of traction load for high speed railways," *IEEE Access*, vol. 7, pp. 76049–76060, 2019.

[6] L. H. Lee, E. P. Chew, P. I. Frazier, Q.-S. Jia, and C.-H. Chen, "Advances in simulation optimization and its applications," *IIE Trans.*, vol. 45, no. 7, pp. 683–684, 2013.

[7] C.-H. Chen, E. Yücesan, L. Dai, and H.-C. Chen, "Optimal budget allocation for discrete-event simulation experiments," *IIE Trans.*, vol. 42, no. 1, pp. 60–70, 2010.

[8] C.-H. Chen and L. H. Lee, *Stochastic Simulation Optimization: An Optimal Computing Budget Allocation*, vol. 1. Singapore: World Science, 2011.

[9] S.-H. Kim and B. L. Nelson, "A fully sequential procedure for indifference-zone selection in simulation," *ACM Trans. Model. Comput. Simul.*, vol. 11, no. 3, pp. 251–273, 2001.

[10] S. E. Chick and Y. Z. Wu, "Selection procedures with frequentist expected opportunity cost bounds," *Oper. Res.*, vol. 53, no. 5, pp. 867–878, 2005.

[11] S. E. Chick and K. Inoue, "New procedures for identifying the best simulated system using common random numbers," *Manag. Sci.*, vol. 47, no. 8, pp. 1133–1149, 2001.

[12] S. E. Chick and K. Inoue, "New two-stage and sequential procedures for selecting the best simulated system," *Oper. Res.*, vol. 49, no. 5, pp. 732–743, 2001.

[13] S. E. Chick, J. Branke, and C. Schmidt, "Sequential sampling to myopically maximize the expected value of information," *INFORMS J. Comp.*, vol. 22, no. 1, pp. 71–80, 2010.

[14] C.-H. Chen, J. Lin, E. Yücesan, and S. E. Chick, "Simulation budget allocation for further enhancing the efficiency of ordinal optimization," *Discrete Event Dyn. Syst.*, vol. 10, no. 3, pp. 251–270, Jul. 2000.

[15] D. He, S. E. Chick, and C.-H. Chen, "Opportunity cost and OCBA selection procedures in ordinal optimization for a fixed number of alternative systems," *IEEE Trans. Syst., Man, Cybern., C (Appl. Rev.)*, vol. 37, no. 5, pp. 951–961, Sep. 2007.

[16] J. Branke, S. E. Chick, and C. Schmidt, "Selecting a selection procedure," *Manage. Sci.*, vol. 53, no. 12, pp. 1916–1932, 2007.

[17] S. H. Choi and T. G. Kim, "Efficient ranking and selection for stochastic simulation model based on hypothesis test," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 48, no. 9, pp. 1555–1565, Sep. 2018.

[18] S. H. Choi and T. G. Kim, "Pareto set selection for multiobjective stochastic simulation model," *IEEE Trans. Syst., Man, Cybern. Syst.*, to be published. doi: 10.1109/TSMC.2018.2846680.

[19] S. H. Choi and T. G. Kim, "A heuristic approach for selecting best-subset including ranking within the subset," *IEEE Trans. Syst., Man, Cybern. Syst.*, to be published. doi: 10.1109/TSMC.2018.2870408.

[20] S. H. Choi, B. G. Kang, and T. G. Kim, "A heuristic procedure for selecting the best feasible design in the presence of stochastic constraints," *IEEE Trans. Syst., Man, Cybern. Syst.*, to be published. doi: 10.1109/TSMC.2019.2894523.

[21] A. Hayter, *Probability and Statistics for Engineer and Scientists*, 4th ed. Boston, MA, USA: Cengage Learning, 2012.

[22] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*, 4th ed. Cambridge, MA, USA: MIT Press, 2012.

[23] D. C. Howell, "Hypothesis tests applied to means," in *Statistical Methods for Psychology*, 7th ed. Belmont, CA, USA: Wadsworth, 2010, pp. 179–223.

[24] A. M. Law, *Simulation Modeling and Analysis*, 4th ed. New York, NY, USA: McGraw-Hill, 2007.

[25] R. E. Bechhofer, T. J. Santner, and D. M. Goldsman, *Design and analysis of Experiments for Statistical Selection, Screening, and Multiple Comparisons*. New York, NY, USA: Wiley, 1995.

[26] G. D. Jones, B. Huang, Q. Bian, C. A. Linson, P. M. Wallrich, W. W. Price, J. Decker, and E. R. Davis, "Simulated firearm having a multiple integrated laser engagement system," WO Patent 2008 115 624, Sep. 25, 2008. [Online]. Available: https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2008115624

[27] S. H. Choi, C.-B. Choi, and T. G. Kim, "Optical engineering simulator development for efficient design of MILES gear," in *Proc. 12th I3M Multiconf.*, Bergeggi, Italy, Sep. 2015, pp. 56–64.

[28] S. H. Choi, K.-M. Seo, and T. G. Kim, "Accelerated simulation of discrete event dynamic systems via a multi-fidelity modeling framework," *Appl. Sci.*, vol. 7, no. 10, p. 1056, 2017.

[29] S. H. Choi, J. H. Lee, S. H. Lee, H. D. Yoo, J. Koo, and T. G. Kim, "6 dof aircraft simulation model capable of handling maneuver events (WIP)," in *Proc. Summer Comput. Simulation Conf.*, 2016, Art. no. 54.

SEON HAN CHOI (M'18) received the B.S., M.S., and Ph.D. degrees in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea, in 2012, 2014, and 2018, respectively.

He was a Teaching Assistant in computer programming with the School of Electrical Engineering, from 2013 to 2017. In 2018, he was a Postdoctoral Researcher with the Information and Electronics Research Institute, KAIST, where he was a Senior Researcher, from 2018 to 2019. In 2019, he joined the Department of IT Convergence and Application Engineering, Pukyong National University, Busan, South Korea, as an Assistant Professor. His current research interests include modeling and simulation of discrete-event systems, efficient simulation optimization under stochastic noise, evolutionary computing, and machine learning.

CHANGBEOM CHOI (M'18) received the B.S. degree in computer engineering from Kyunghee University, Suwon, South Korea, in 2005, the M.S. degree in computer science and the Ph.D. degree in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea, in 2007 and 2014, respectively.

He has been an Assistant Professor with Handong Global University, Pohang, South Korea, since 2014. His research interests include discrete-event systems modeling and simulation, experiment management, verification, validation, and accreditation, and agent-based simulation in various domains.

TAG GON KIM (SM'95) received the Ph.D. degree in computer engineering with a specialization in systems modeling and simulation from the University of Arizona, Tucson, in 1988.

From 1989 to 1991, he was an Assistant Professor with the Department of Electrical and Computer Engineering, University of Kansas, Lawrence. In 1991, he joined the Department of Electrical Engineering, KAIST, Daejeon, South Korea, as an Assistant Professor, where he has been a Full Professor, since 1998. He is the coauthor of the book *Theory of Modeling and Simulation* (Burlington, MA, USA: Academic, 2000). He has published more than 200 articles on modeling and simulation (M&S) theory and practice in international journals and conference proceedings. He is very active in research and education concerning defense M&S in South Korea. He has been a Technical Advisor of the defense M&S area at various Korean Government organizations, including the Ministry of Defense, the Defense Agency for Technology and Quality, the Korea Institute for Defense Analysis, and the Agency for Defense Development. He has developed DEVSimHLA, which is a tool set for the development of high-level architecture compliant war game simulators that have been used in the development of three war game simulators for the navy, air force, and marines in South Korea. Since 2004, the three simulators have been interoperated with U.S. war game simulators at the U.S.–South Korea combined exercises such as Ulchi Focus Lens.

Dr. Kim is a member of the Eta Kappa Nu and a Fellow of the SCS. He was the President of the Korea Society for Simulation. He was the Editor-in-Chief of *SIMULATION: Transactions of the Society for Modeling and Simulation: International* (SCS).