
DATA-DRIVEN OPTIMAL COMPUTING BUDGET ALLOCATION UNDER INPUT UNCERTAINTY

Yuhao Wang

School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, GA 30332
yuhaowang@isye.gatech.edu

Enlu Zhou

School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, GA 30332
enlu.zhou@isye.gatech.edu

September 27, 2022

ABSTRACT

In a fixed budget ranking and Selection (R&S) problem, one aims to identify the best design among a finite number of candidates by efficiently allocating the given simulation budget to evaluate design performance. Classical methods for fixed budget R&S usually assume known input distributions, which are the distributions that model the randomness in the system and drive the simulation. In this paper, we consider the practical scenario where the input distribution is unknown but can be estimated from streaming input data that arrive in batches over time. We model the R&S problem in this dynamic setting as a multi-stage problem, where the input distribution estimate is updated at each stage and a stage-wise optimization problem is formulated to allocate the simulation budget. We characterize the optimality conditions for the stage-wise budget allocation problem by applying the large deviations theory to maximize the decay rate of probability of false selection. Based on the optimality conditions and combined with the updating of input distribution estimates, we design two sequential budget allocation procedures for R&S under streaming input data. We theoretically guarantee the consistency and asymptotic optimality of the two proposed procedures. We also demonstrate the practical efficiency through numerical experiments in comparison with the equal allocation policy and two extensions of the Optimal Computing Budget Allocation (OCBA) algorithm.

Keywords: ranking and selection, input uncertainty, large deviations theory, optimal computing budget allocation

1 Introduction

Stochastic simulation has been a powerful tool for designing and analyzing modern complex systems arising in many industrial areas. In building a good simulation model, one needs to capture not only the system's internal logic but also the random factors that affect the system's performance. These random factors, such as customer demand, traveling time, and service time, are often modeled by probability distributions, which are commonly referred to as "input distributions". A simulator generates samples from input distributions to simulate real-world scenarios and evaluate the system's performance on these scenarios.

Evaluation via simulation brings simulation error, which mainly comes from two sources. The first is known as the input uncertainty (IU), which refers to the estimation error of the input distribution from input data observed from the real world. Another source is the stochastic uncertainty (SU), which is induced by random samples generated in the simulation process. The evaluation accuracy can be severely impacted if either type of uncertainty remains large and is ignored. Therefore, it is important for a simulation analysis procedure to take into account both sources of error/uncertainty.

SU can be reduced by carrying out more simulation replications, while IU is determined by the amount of input data. In many application problems, input data are often collected in batches over time to reduce IU and refine the simulation model, which is then used to test and compare potential system designs or strategies. More specifically, each time a new

batch of data come, the simulator updates the estimate of the input distributions and run additional simulation based on the new input distributions. The simulation budget, which is the amount of simulation effort can be implemented, between arrivals of two successive batches of input data is usually limited and determined by external factors such as the inter-arrival time of input data and simulation cost. The goal is to identify the best design/strategy through smart allocation of the limited given simulation budget to evaluate different designs/strategies.

This motivates us to consider a fixed budget ranking and selection (R&S) problem with streaming data where new input data arrive over time in batches of possibly varying sizes. A procedure for fixed budget R&S aims to achieve a probability of correct selection (PCS) as high as possible with a given simulation budget. Our proposed procedure for the streaming data setting is built on the Optimal Computing Budget Allocation (OCBA) algorithm, which is one of the most widely applied and studied algorithms for fixed budget R&S with a fixed input distribution. It computes the budget allocation rule by repeatedly maximizing an approximate PCS objective in each iteration with plug-in estimators of design performances. OCBA was first proposed in [5] and was shown to converge asymptotically to the optimal allocation rule in [32]. The statistical validity of OCBA crucially relies on the stationarity of the underlying input distribution, which implies that the simulation outputs are independent and identically distributed (i.i.d.) samples even though they are generated in different iterations of the algorithm. Hence, the performance estimation error diminishes as more samples are generated over iterations, leading to convergence of the allocation policy to the optimal policy. OCBA have been extended in the past years to various problems such as subset selection ([4] and [18]), contextual R&S ([20] and [27]), multi-objective ([31]), finding simplest good designs ([26] and [45]), maximizing opportunity cost ([19]), robust R&S under input uncertainty ([21]), and many others. All of these works either assume a known input distributions or consider an empirical input distribution estimated with a fixed set of input data, where the performance samples for the same design follow the same input distribution over iterations.

The setting of streaming data considered in this paper is more challenging than the setting of fixed input distribution in the aforementioned OCBA works. Most notably, the underlying input distribution is no longer fixed but is updated with new input data at each time stage, and hence it breaks the i.i.d. condition of the simulation outputs for each fixed design. Moreover, since simulation is often time consuming compared to the inter-arrival time of data, only a limited (small) number of simulation replications can be carried out at each stage, i.e., in between arrivals of input data. To control the variance of performance estimates, it is necessary to reuse simulation outputs generated under heterogeneous input distributions over part stages. To efficiently aggregate simulation outputs from different stages as well as maintaining statistical validity, we first approximate the input distribution by an empirical distribution on a finite support that is fixed over time, and each point in the finite support is called an “input realization”. The central idea is that by breaking down the entire input distribution into input realizations, the simulation output conducted for each design at a fixed input realization (called a design-input pair) is i.i.d. over time. More specifically, the fixed finite support allows us to update the estimated input distribution through probability mass function (pmf) and enables us to define the performance estimate of a design as the weighted sum of the design-input performance estimates weighted by the empirical pmf over all input realizations. The estimation error of the performance then can be split into two parts, the error from IU reflected by the estimation of input pmf and the error from SU reflected by the sample average of the design-input performance. The convergence of the budget allocation policy relies on the convergence of both the empirical input distribution and the design-input performance estimate. We note that the design-input pairwise simulation is similar as contextual R&S where simulation is run on each design-context pair (see [20]). However, our setting with input uncertainty is different from contextual R&S in nature: the goal of R&S with input uncertainty is to identify one single optimal design with the best performance measure under the true input distribution, while contextual R&S aims to choose one optimal design for each possible context and typically considers a fixed input distribution. Another distinction is that [20] applied the same large deviations rate (LDR) function for i.i.d Gaussian random variables as in [22] to derive the optimal allocation rule, but we cannot directly apply this result due to our different definition of the performance estimate, which contains samples under different input realizations. We recalculate the rate function of our performance estimate with input distribution and derive the optimality conditions for stage-wise budget allocation policy. Based on the stage-wise optimality conditions, we develop two data-driven procedures for the streaming data setting. We show that the two allocation policies converge to the optimal allocation policy under the true input distribution as the time stage goes to infinity.

We summarize the contributions of this paper as follows.

1. This paper along with our earlier conference version, [37], are the first to consider streaming input data in fixed budget R&S problems and design a data-driven approach. Compared with the conference version, this paper develops a new procedure IU-OCBA-balance and extends the problem to the more general setting of unknown variance in the evaluation of design performance.
2. To aggregate simulation outputs generated under heterogeneous input distributions over time, we define a new performance estimator as the weighted sum of design-input sample means over all input realizations, which contains

IU in the weights from the input distribution estimate and SU in the design-input sample means. We apply the Gartner-Ellis theorem (see [13]) to calculate the rate function of this new performance estimator. Unlike many other works derived from [22] where they directly use the same rate function, we need to recalculate it due to samples from different distributions in the performance estimator. We formulate a stage-wise rate maximization problem using the recalculated rate function and derive the corresponding optimality conditions. Compared with optimality conditions in [22], we obtain an additional ‘‘Derivative Balance’’ condition that characterizes the allocation rule among different input realizations, showing the impact of different input distributions on the budget allocation.

3. We develop two fully sequential procedures, called IU-OCBA-approx and IU-OCBA-balance, based on different approaches to solve the optimality conditions. IU-OCBA-approx takes an OCBA-type approach to approximate the optimality conditions and compute the approximately optimal solution in an explicit form; IU-OCBA-balance adapts the approach in [6] to avoid directly computing the solution but instead try to balance the two sides of the optimality equations each time. We prove the statistical consistency and asymptotic optimality of both procedures, where the major analytic challenge is to show that the procedures asymptotically satisfy the additional ‘‘Derivative Balance’’ condition and the much more complicated ‘‘Local Balance’’ condition (compared to the case without IU). Our proof indicates the importance of balancing IU and SU, or in other words, having sufficient input data in accordance with the simulation budget to ensure the convergence of allocation policy.

Next we briefly review the relevant literature with an emphasis on the relation with our work.

1.1 Literature Review

The research on R&S largely falls into two related yet different categories. The fixed confidence R&S procedures aim to achieve a pre-specified probability of correct selection (PCS) using the least possible amount of simulation effort, whereas the fixed budget R&S procedures typically tend to attain a PCS as high as possible with a given simulation budget. For fixed confidence, a large body of literature goes to the indifference zone (IZ) formulation. An IZ procedure guarantees selecting the best design with at least a pre-specified confidence level, given that the difference between the top-two designs is sufficiently large. Existing IZ procedures in the R&S literature include but are not limited to the KN procedure in [28], the KVP and UVP procedures in [25], and the BIZ procedure in [17]. We refer the reader to [29] for a comprehensive review of IZ formulations. In addition, the Bayesian approaches in [8] and the probably approximately correct (PAC) selection in [34] have also been studied in this stream of works.

In this paper, we focused on the fixed budget R&S. As discussed in Section 1, OCBA is originally derived under a normality assumption and an approximate PCS objective. The allocation rule of OCBA can be justified from a rigorous perspective of the large deviations theory in [22], after which lots of works followed this large deviations formulation. For instance, [6] designed a fully sequential budget allocation algorithm for general distributions using the optimality conditions in [22]; [24] and [35] applied the large deviations theory to constrained R&S; [20] and [3] extended the large deviations approach to contextual R&S; [21] computed the LDR function with respect to a worst-case performance estimator. In this paper, we compute the LDR function of a performance estimator aggregating samples across different input distributions. Other well-known fixed budget R&S procedures include the expected value of information (EVI) approach proposed by [9] and the knowledge-gradient (KG) approach proposed by [16], where EVI is derived by asymptotically minimizing a bound of the expected loss and KG determines the optimal sampling allocation policy by maximizing the so-called acquisition function. We refer the reader to [23] for a recent overview of the R&S literature.

There is a sophisticated literature on quantifying the impact of IU on simulation output, which includes but is not limited to the frequentist methods ([1]), the meta-assisted methods such as [2], [42, 43], the empirical likelihood method ([30]), single-run method ([33]), and bayesian methods ([7, 46, 47, 42]). We refer the reader to [10] for a recent review on the topic of input uncertainty.

Despite the extensive study on IU quantification and classical R&S, works on R&S with IU have only been studies in recent years. [11, 12] aimed to eliminate inferior designs as many as possible and return a subset of superior designs with fixed amount of input data. [39] formulated a fixed budget problem under OCBA framework to simultaneously allocate the effort to carry out stimulation and the effort to obtain input data; this work is followed by [44], which proposed a general framework that integrates input data collection and simulation in which the data collection and simulation costs themselves can be random. [21], [41], [40] took a fixed budget formulation with a robust approach, aiming to select a design with the best worst-case performance over an uncertainty set of finite distributions that contains the true input distribution; [15] also used this worst-case criteria but took an IZ formulation. [36] derived confidence bands to account for both IU and SU in R&S. All of the above works considered a fixed set of input data, but more recently [38] considered R&S with streaming input data, the same setting as this paper, but used fixed confidence formulation. They proposed a moving average performance estimator to aggregate simulation outputs from different

input distributions over time stages and designed sequential elimination procedures to screen out the inferior designs until one is left to be the optimal one with at least a pre-specified confidence level.

The rest of the paper is organized as follows. We describe the problem setting and present the overall framework of the proposed data-driven procedures in Section 2. In Section 3 we explicitly solve the stage-wise budget allocation problem by applying the large deviations theory to calculate the rate function for the performance estimator and characterize the stage-wise optimal allocation policy. In Section 4 we propose sequential procedures for the R&S problem under streaming data. We show the statistical consistency and asymptotic optimality of the two procedures in Section 5. We present numerical results in Section 6 and conclude in Section 7.

2 Problem Statement

We first lay down some basic notations. Suppose we have a set of finite number of designs $\mathcal{I} = \{1, 2, \dots, K\}$, and the goal is to find the design with the highest expected performance. The performance of each design $i \in \mathcal{I}$ is evaluated through repeated simulation runs. The simulation budget, which is the total number of simulation replications we can run on all designs, is often limited by computational time or expense. The core of the fixed budget R&S problem is to devise procedures that maximize the probability of correct selection (PCS) of the optimal design when exhausting the simulation budget.

In classical R&S, the input distributions, $\{F_i\}$, that capture various sources of system randomness are assumed to be known. However, in practice the input distributions are seldom known and need to be estimated from input data, which are a finite amount of real-world observations. Sources of system randomness, such as demand load in an inventory-production problem, are often shared among all designs. Therefore, throughout the paper we assume that all the designs share the same input distribution F (thus, dropping the subscript i) and consequently common input data from these distributions. Note that in addition to the common input distribution, there are possibly other design-specific distributions that drive the simulation. For example, for a service network where we want to select the best design (network configuration, service rates), the customer demand is unknown and common to all designs, whereas the service distribution is known and determined by the design.

Let ζ denote the input random vector, which follows the unknown input distribution F . Let $X_i(\tilde{\zeta})$ denote the sample performance, i.e., the simulation output of design i under the input realization $\zeta = \tilde{\zeta}$. Denote by $\varepsilon_i(\tilde{\zeta})$ the randomness in simulation output under input realization $\tilde{\zeta}$, which is caused by simulating the design-specific distributions, such as the service distribution mentioned above. We assume $X_i(\tilde{\zeta})$ has the following form:

$$X_i(\tilde{\zeta}) = \mu_i(\tilde{\zeta}) + \varepsilon_i(\tilde{\zeta}),$$

where $\varepsilon_i(\tilde{\zeta})$ has a known distribution with mean 0 and $\mu_i(\tilde{\zeta}) = \mathbf{E}[X_i(\tilde{\zeta})|\tilde{\zeta}]$ is the expected performance of design i under the input realization $\tilde{\zeta}$. Let $\bar{\mu}_i = \mathbf{E}[\mu_i(\zeta)]$ denote the expected performance of design i , where the expectation is taken with respect to (w.r.t.) the input distribution ζ . The goal is to select the design with the largest expected performance, i.e., the design b (for best) such that

$$b = \arg \max_{1 \leq i \leq K} \bar{\mu}_i. \quad (1)$$

Without loss of generality, we assume design 1 is the unique optimal design, i.e., $b = 1$.

A common approach of simulating one replication for design i , as in the classical R&S, is to first generate a sample $\tilde{\zeta}$ of ζ and then generate a sample performance $X_i(\tilde{\zeta})$. However, a major challenge of taking this approach with streaming input data is that the estimated input distribution under which the simulation is conducted varies over time. As a result, the simulation output is no more i.i.d., which makes it analytically difficult to apply methodologies in classical R&S. Instead of generating samples from different input distributions, we conduct the simulation on each design-input pair, that is, we simulate a design i under a specific input realization ζ . Then the simulation outputs of the design under the same input realization are i.i.d. samples. To have a finite number of input realizations, we make the following assumption on the input distribution.

Assumption 1 *The true (unknown) input distribution F_ζ has a finite support $\{\zeta_1, \zeta_2, \dots, \zeta_D\}$, with probability mass function (pmf) $\mathbf{P}(\zeta = \zeta_j) = p_j > 0$, $j = 1, \dots, D$.*

Assumption 1 can be satisfied by discretizing the support of the input distribution if its support is continuous. Specifically, Suppose the input distribution has cumulative distribution function (CDF) F with the support U . We discretize the support by a partition of U , $\{U_j\}_{j=1}^D$, such that $U = \bigcup_{j=1}^D U_j$. The probability mass function can then be defined as the probability of each subdomain U_j , i.e., $p_j = \mathbf{P}(\zeta \in U_j) = \int_{U_j} dF(x)$.

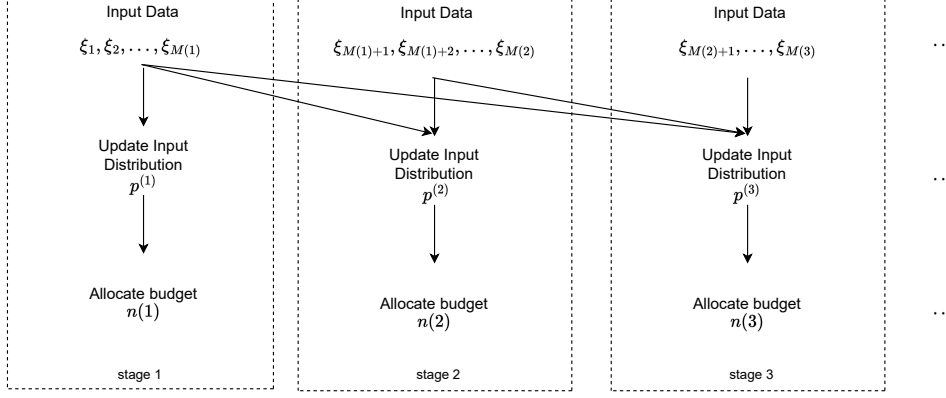


Figure 1: Illustration of budget allocation with streaming input data

With this finite support assumption, we simplify the notations of design-input performance, expected performance and simulation error respectively as $X_{i,j} := X_i(\zeta_j)$, $\mu_{i,j} := \mu_i(\zeta_j)$ and $\varepsilon_{i,j} := \varepsilon_i(\zeta_j)$. Denote by $X_{i,j}^{(l)}$ the l th replication for $X_{i,j}$. We make the following assumption on the simulation output.

Assumption 2

1. The simulation error $\varepsilon_{i,j}$ follows a normal distribution $\mathbf{N}(0, \sigma_{\varepsilon_{i,j}}^2)$ where the variance is unknown.
2. The simulation output $\{X_{i,j}^{(l)}\}$ are independent for all i, j and l .

Assumption 2.1 models the simulation error as Gaussian noise with both mean and variance being unknown. It is common in the R&S literature to assume Gaussian randomness with known variance (see, for example, [5], [6]), but we develop the procedures and analysis in the more general setting of unknown variance. Assumption 2.2 can be guaranteed since now we simulate on fixed input realization, where the randomness only comes from simulation error, which is often independent across designs, input realizations, and iterations. It is worth noting here we assume the simulator can choose a specific input realization to run simulation in finding the best design, while for implementation on the real system, the decision maker does not know the input realization before making a decision; this is the key difference from contextual R&S, where the decision maker first observes a realization of the context and then makes a decision.

Next, we describe the overall framework of our proposed data-driven budget allocation procedures. Recall that our goal is to select the design $b = \arg \max_{1 \leq i \leq K} \bar{\mu}_i$, where $\bar{\mu}_i = \sum_{j=1}^D p_j \mu_{i,j}$ under Assumption 1. The true input distribution $\{p_j\}$ needs to be estimated sequentially via streaming data. More specifically, at time stage t , new input data of batch size $m(t)$ can be obtained and used to update the estimate of the input distribution, and then we allocate simulation budget $n(t)$ to design-input pairs to maximize the PCS with respect to the current estimated input distribution. We assume both $n(t)$ and $m(t)$ are given. This process is illustrated in Figure 1, where $M(t) = \sum_{\tau=1}^t m(\tau)$ is the total amount of input data collected up to stage t . For input data, we make the following assumption.

Assumption 3 The input data, $\{\xi_s\}_{s=1}^{\infty}$, are identically and independently distributed.

With Assumption 3, at stage t the input distribution can be estimated by the empirical distribution consisting of the observed data up to stage t :

$$p_j^{(t)} = \frac{\sum_{s=1}^{M(t)} \mathbb{1}\{\xi_s = \zeta_j\}}{M(t)}.$$

To find the budget allocation rule for each stage, we apply the large deviations theory to formulate an optimization problem under the current estimated input distribution and characterize its optimality condition to derive the stage-wise optimal budget allocation rule in Section 3. Then combined with the updating of estimated input distributions, we develop two data-driven budget allocation procedures for the multi-stage setting in 4.

3 Stage-wise Optimal Budget Allocation

In this section, we formulate and solve a static optimal budget allocation problem to guide allocating the stage-wise budget to design-input pairs under the current estimated input distribution, estimated expected performance, and estimated simulation variance. Let $n_{i,j}$ denote the simulation budget allocated to design i under input realization ζ_j . Let $\alpha_i = (\alpha_{i,1}, \dots, \alpha_{i,D})^\top$ be the ratio of budget allocated to design i and input realization j . Denote by

$$\hat{\mu}_{i,j}(\alpha_{i,j}n) = \frac{1}{\alpha_{i,j}n} \sum_{s=1}^{\alpha_{i,j}n} X_{i,j}^{(s)} \quad \text{and} \quad \hat{\mu}_i(\alpha_i, n) = \sum_{j=1}^D p_j^{(t)} \hat{\mu}_{i,j}(\alpha_{i,j}n),$$

the estimated performance for the (i, j) (design-input) pair and the estimated performance for design i , respectively. Ignoring the minor issue of $\alpha_{i,j}n$ not being an integer, we formulate an optimization problem from the large deviations perspective as in [22] and define the rate function

$$\mathbf{G}_i(\alpha_1, \alpha_i) := - \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbf{P} \left(\hat{\mu}_1(\alpha_1, n) \leq \hat{\mu}_i(\alpha_i, n) \right).$$

That is, $n\mathbf{G}_i(\alpha_1, \alpha_i)$ is the exponential rate of the probability of the event $\{\hat{\mu}_1(\alpha_1, n) \leq \hat{\mu}_i(\alpha_i, n)\}$ going to zero as n goes to infinity. Since the probability of false selection (PFS), which is defined as

$$PFS = \mathbf{P} \left(\hat{\mu}_1(\alpha_1, n) \leq \max_{2 \leq i \leq K} \hat{\mu}_i(\alpha_i, n) \right),$$

can be bounded by

$$\max_{2 \leq i \leq K} \mathbf{P} \left(\hat{\mu}_1(\alpha_1, n) \leq \hat{\mu}_i(\alpha_i, n) \right) \leq PFS \leq (K-1) \max_{2 \leq i \leq K} \mathbf{P} \left(\hat{\mu}_1(\alpha_1, n) \leq \hat{\mu}_i(\alpha_i, n) \right),$$

we have

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log PFS = - \min_{2 \leq i \leq K} \mathbf{G}_i(\alpha_1, \alpha_i). \quad (2)$$

That is, $\min_{2 \leq i \leq K} \mathbf{G}_i(\alpha_1, \alpha_i)$ is the asymptotically exponential decay rate of PFS. To maximize this decay rate of PFS, we consider the following optimization problem:

$$\begin{aligned} & \max_{\alpha_{i,j}, 1 \leq i \leq K, 1 \leq j \leq D} z \\ & \text{s.t.} \quad \mathbf{G}_i(\alpha_1, \alpha_i) - z \geq 0 \quad 2 \leq i \leq K \\ & \quad \sum_{i=1}^K \sum_{j=1}^D \alpha_{i,j} = 1 \\ & \quad \alpha_{i,j} \geq 0 \quad 1 \leq i \leq K, 1 \leq j \leq D, \end{aligned} \quad (3)$$

To solve (3), we need the true input distribution, the expected performance $\bar{\mu}_i$, and the variance of simulation error $\sigma_{i,j}^2$, all of which are unknown. Therefore, we replace the unknown parameters with their estimates at the current stage and solve (3) to guide the budget allocation.

Optimizing the large deviations rate of PFS is first studied in [22], where IU is not considered. Formulation without IU can be seen as a special case of $D = 1$ in our setting. Our approach is an extension of [22], which incorporates simulation samples under different input realizations in calculating the rate function. Due to this difference, we cannot directly apply their result. Instead, we take a similar approach using the Gartner-Ellis Theorem (see [13]), and the detailed calculation is shown in the next section.

3.1 Calculation of the Rate Function

In this section we give the explicit form of the rate function $\mathbf{G}_i(\alpha_1, \alpha_i)$. The rate function is computed with full knowledge of involved distributions. That is, all the parameters such as p_j , $\bar{\mu}_i$ and $\sigma_{i,j}$ that will be used are assumed known in deriving the rate function.

For a fixed i , let $\Lambda_{i,j}(\cdot)$ denote the log-moment generating function of $X_{i,j}$ and $\Lambda_n(\cdot, \cdot)$ denote the log-moment generating function of $Z_n = (\hat{\mu}_1(\alpha_1, n), \hat{\mu}_i(\alpha_i, n))$. That is,

$$\begin{aligned}\Lambda_n(\lambda_1, \lambda_i) &= \log \mathbf{E} \left[e^{\lambda_1 \hat{\mu}_1(\alpha_1, n) + \lambda_i \hat{\mu}_i(\alpha_i, n)} \right] \\ &= \log \mathbf{E} \left[\exp \left(\lambda_1 \sum_{j=1}^D p_j \sum_{s=1}^{n\alpha_{1,j}} \frac{X_{1,j}^{(s)}}{n\alpha_{1,j}} + \lambda_i \sum_{j=1}^D p_j \sum_{s=1}^{n\alpha_{i,j}} \frac{X_{i,j}^{(s)}}{n\alpha_{i,j}} \right) \right] \\ &= \sum_{j=1}^D n\alpha_{1,j} \log \mathbf{E} \left[\exp \left(\frac{\lambda_1 p_j}{n\alpha_{1,j}} X_{1,j} \right) \right] + \sum_{j=1}^D n\alpha_{i,j} \log \mathbf{E} \left[\exp \left(\frac{\lambda_i p_j}{n\alpha_{i,j}} X_{i,j} \right) \right] \\ &= \sum_{j=1}^D n\alpha_{1,j} \Lambda_{1,j} \left(\frac{\lambda_1 p_j}{n\alpha_{1,j}} \right) + \sum_{j=1}^D n\alpha_{i,j} \Lambda_{i,j} \left(\frac{\lambda_i p_j}{n\alpha_{i,j}} \right).\end{aligned}$$

The third equality follows from Assumption 2.2 that all simulation outputs are independent across designs and simulation outputs are identically distributed for the same design-input pair. Then, substituting λ_1 and λ_i with $n\lambda_1$ and $n\lambda_i$, respectively, we obtain

$$\frac{1}{n} \Lambda_n(n\lambda_1, n\lambda_i) = \sum_{j=1}^D \alpha_{1,j} \Lambda_{1,j} \left(\frac{\lambda_1 p_j}{\alpha_{1,j}} \right) + \sum_{j=1}^D \alpha_{i,j} \Lambda_{i,j} \left(\frac{\lambda_i p_j}{\alpha_{i,j}} \right).$$

Since $X_{i,j}$ follows a normal distribution with mean $\mu_{i,j}$ and variance $\sigma_{i,j}^2$, $\Lambda_{i,j}(\lambda) = \lambda \mu_{i,j} + \frac{1}{2} \lambda^2 \sigma_{i,j}^2$. Let $I(x_1, x_i)$ be the Fenchel-Legendre transform of Λ_n . Then,

$$\begin{aligned}I(x_1, x_i) &= \sup_{\lambda_1, \lambda_i} \left\{ \lambda_1 x_1 + \lambda_i x_i - \sum_{j=1}^D \alpha_{1,j} \Lambda_{1,j} \left(\frac{\lambda_1 p_j}{\alpha_{1,j}} \right) - \sum_{j=1}^D \alpha_{i,j} \Lambda_{i,j} \left(\frac{\lambda_i p_j}{\alpha_{i,j}} \right) \right\} \\ &= \sup_{\lambda_1} \left\{ \lambda_1 x_1 - \sum_{j=1}^D \alpha_{1,j} \Lambda_{1,j} \left(\frac{\lambda_1 p_j}{\alpha_{1,j}} \right) \right\} + \sup_{\lambda_i} \left\{ \lambda_i x_i - \sum_{j=1}^D \alpha_{i,j} \Lambda_{i,j} \left(\frac{\lambda_i p_j}{\alpha_{i,j}} \right) \right\} \\ &= \sup_{\lambda_1} \left\{ \lambda_1 x_1 - \sum_{j=1}^D \left(\lambda_1 p_j \mu_{1,j} + \frac{1}{2} \frac{\sigma_{1,j}^2 \lambda_1^2 p_j^2}{\alpha_{1,j}} \right) \right\} + \sup_{\lambda_i} \left\{ \lambda_i x_i - \sum_{j=1}^D \left(\lambda_i p_j \mu_{i,j} + \frac{1}{2} \frac{\sigma_{i,j}^2 \lambda_i^2 p_j^2}{\alpha_{i,j}} \right) \right\} \\ &= \underbrace{\frac{1}{2} \frac{(x_1 - \bar{\mu}_1)^2}{\sum_{j=1}^D \frac{\sigma_{1,j}^2 p_j^2}{\alpha_{1,j}}}}_{= I_1} + \underbrace{\frac{1}{2} \frac{(x_i - \bar{\mu}_i)^2}{\sum_{j=1}^D \frac{\sigma_{i,j}^2 p_j^2}{\alpha_{i,j}}}}_{= I_i}.\end{aligned}$$

By the Gartner-Ellis Theorem, $\mathbf{G}_i(\alpha_1, \alpha_i) = \inf_{x_1 \leq x_i} I(x_1, x_i)$. It is easy to see that I_1 is decreasing for $x_1 \leq \bar{\mu}_1$ and increasing for $x_1 \geq \bar{\mu}_1$, and I_i is decreasing for $x_i \leq \bar{\mu}_i$ and increasing for $x_i \geq \bar{\mu}_i$. Since $\bar{\mu}_1 > \bar{\mu}_i$, we must have

$$\mathbf{G}_i(\alpha_1, \alpha_i) = \inf_{\bar{\mu}_i \leq x \leq \bar{\mu}_1} I(x, x) = \frac{(\bar{\mu}_1 - \bar{\mu}_i)^2}{2 \left(\sum_{j=1}^D \frac{\sigma_{1,j}^2 p_j^2}{\alpha_{1,j}} + \sum_{j=1}^D \frac{\sigma_{i,j}^2 p_j^2}{\alpha_{i,j}} \right)}. \quad (4)$$

When $D = 1$, we recover exactly the same rate function as in [22]. The following lemma summarizes some important properties of $\mathbf{G}_i(\alpha_1, \alpha_i)$.

Lemma 1 *Suppose Assumption 2 holds. Then,*

1. $\mathbf{G}_i(\alpha_1, \alpha_i)$ is strictly increasing in $\alpha_{1,j}$ and $\alpha_{i,j}$ for $\alpha_{1,j}, \alpha_{i,j} > 0, j = 1, 2, \dots, D$. Moreover, $\mathbf{G}_i(\alpha_1, \alpha_i) = 0$ if there exists j_0 such that $\min(\alpha_{1,j_0}, \alpha_{i,j_0}) = 0$.
2. $\mathbf{G}_i(\alpha_1, \alpha_i)$ is concave in (α_1, α_i) for $(\alpha_1, \alpha_i) > 0$.

Lemma 1.1 implies that any design-input pair must be allocated with a positive ratio of the simulation budget; otherwise, the rate will be zero. Lemma 1.2 claims the concavity of \mathbf{G}_i , which guarantees the optimality with the Karush–Kuhn–Tucker (KKT) condition for the optimization problem (3) in the following section.

3.2 Optimal Allocation Policy

In this section we derive the optimality conditions for problem (3), shown in the following theorem.

Theorem 1 *Suppose Assumption 2 holds. Let $\alpha \geq 0$ be a feasible allocation policy. Then α is the optimal solution to (3) if and only if the following three conditions hold:*

$$1. (\text{Derivative Balance}) \quad \frac{\partial \mathbf{G}_i(\alpha_1, \alpha_i)}{\partial \alpha_{i,j}} = \frac{\partial \mathbf{G}_i(\alpha_1, \alpha_i)}{\partial \alpha_{i,j'}} \quad 2 \leq i \leq K \text{ and } 1 \leq j < j' \leq D, \quad \text{if } p_j, p_{j'} > 0; \quad (5)$$

$$2. (\text{Total Balance}) \quad \sum_{i=2}^K \frac{\partial \mathbf{G}_i(\alpha_1, \alpha_i) / \partial \alpha_{1,j}}{\partial \mathbf{G}_i(\alpha_1, \alpha_i) / \partial \alpha_{i,j}} = 1 \quad 1 \leq j \leq D, \quad \text{if } p_j > 0; \quad (6)$$

$$3. (\text{Local Balance}) \quad \mathbf{G}_i(\alpha_1, \alpha_i) = \mathbf{G}_{i'}(\alpha_1, \alpha_{i'}) \quad 2 \leq i < i' \leq K. \quad (7)$$

Or equivalently in the explicit form:

$$1. \quad \frac{\alpha_{i,j}}{\sigma_{i,j} p_j} = \frac{\alpha_{i,j'}}{\sigma_{i,j'} p_{j'}} \quad 2 \leq i \leq K, 1 \leq j \leq D, \quad (8)$$

$$2. \quad \left(\frac{\alpha_{1,j}}{\sigma_{1,j}} \right)^2 = \sum_{i=2}^K \left(\frac{\alpha_{i,j}}{\sigma_{i,j}} \right)^2 \quad 1 \leq j \leq D, \quad (9)$$

$$3. \quad \frac{(\bar{\mu}_1 - \bar{\mu}_i)^2}{\sum_{j=1}^D \frac{\sigma_{1,j}^2 p_j^2}{\alpha_{1,j}} + \sum_{j=1}^D \frac{\sigma_{i,j}^2 p_j^2}{\alpha_{i,j}}} = \frac{(\bar{\mu}_1 - \bar{\mu}_{i'})^2}{\sum_{j=1}^D \frac{\sigma_{1,j}^2 p_j^2}{\alpha_{1,j}} + \sum_{j=1}^D \frac{\sigma_{i',j}^2 p_j^2}{\alpha_{i',j}}} \quad 2 \leq i < i' \leq K. \quad (10)$$

Furthermore, the optimal solution α^* to (3) is unique.

Proof. We first show the existence of α . The existence follows from the continuity of G_i with respect to $\alpha \in \Delta^{KB-1}$, where Δ^n denotes the n -dimensional simplex. Furthermore, by Lemma 1.1, G_i is strictly increasing in $\alpha_{1,j'}$ and $\alpha_{i,j'}$. Since $\alpha_{i,j} = \frac{1}{KB} \forall i, j$ is a feasible solution and the corresponding objective value is strictly positive, the optimal solution α must satisfy $\alpha_{i,j} > 0$ for all i, j .

Now we show the necessity of the three optimality conditions. By Lemma 1.2, the optimization problem (3) is a concave maximization problem, and therefore the KKT conditions are both sufficient and necessary for the optimality. With α strictly positive, the KKT conditions can be written as

$$1 - \sum_{i=1}^K \lambda_i = 0, \quad (11)$$

$$\lambda_i \frac{\partial G_i}{\partial \alpha_{i,j}}(\alpha_1, \alpha_i) = \gamma \quad 2 \leq i \leq K, 1 \leq j \leq D, \quad (12)$$

$$\sum_{i=2}^K \lambda_i \frac{\partial G_i}{\partial \alpha_{1,j}}(\alpha_1, \alpha_i) = \gamma \quad 1 \leq j \leq D, \quad (13)$$

$$\lambda_i (G_i(\alpha_1, \alpha_i) - z) = 0 \quad 2 \leq i \leq K, \quad (14)$$

for some γ and $\lambda_i \geq 0$, $2 \leq i \leq K$. By (11) there exists at least one i_0 such that $\lambda_{i_0} > 0$. Then since G_i is increasing in $\alpha_{i,j}$, we have $\frac{\partial G_{i_0}}{\partial \alpha_{i_0,j}}(\alpha_1, \alpha_{i_0}) > 0$. This implies $\gamma > 0$ by (12). Hence, we must have $\lambda_i > 0$ for all $2 \leq i \leq K$. Then we have $\frac{\partial G_i(\alpha_1, \alpha_i)}{\partial \alpha_{i,j}} = \frac{\gamma}{\lambda_i}$, $1 \leq j \leq D$, $2 \leq i \leq K$, which proves (5). Since $\lambda_i > 0$, $G_i(\alpha_1, \alpha_i) = z$, $2 \leq i \leq K$ by (14). Hence, (7) holds. To see why (6) holds, solving for $\lambda_i = \frac{\gamma}{\frac{\partial G_i}{\partial \alpha_{i,j}}(\alpha_1, \alpha_i)}$ in (12) and substituting λ_i in (13), we get the desired result.

For sufficiency, first let $\lambda_i = \frac{1}{\frac{\partial G_i(\alpha_1, \alpha_i)}{\partial \alpha_{i,j}} / (\sum_{k=2}^K \frac{1}{\frac{\partial G_k(\alpha_1, \alpha_k)}{\partial \alpha_{k,j}}})}$ for $i \geq 2$. Notice that $\lambda_i > 0$ and does not depend on the choice of j by (5). Moreover, $\{\lambda_i\}_{i \geq 2}$ satisfy condition (11). Further let $\gamma = (\sum_{k=2}^K \frac{1}{\frac{\partial G_k(\alpha_1, \alpha_k)}{\partial \alpha_{k,j}}})^{-1}$, which is also independent of j . We can easily verify that both (12) and (13) hold. (11) also holds by setting $z = G_i(\alpha_1, \alpha_i)$, which is independent of i by (7).

Now we are only left to show the uniqueness of α . First notice that from (8) and (9), we have

$$\frac{\alpha_{1,j}}{\sigma_{1,j}p_j} = \sqrt{\sum_{i=2}^D \left(\frac{\alpha_{i,j}}{\sigma_{i,j}p_j}\right)^2} = \sqrt{\sum_{i=2}^D \left(\frac{\alpha_{i,j'}}{\sigma_{i,j'}p_{j'}}\right)^2} = \frac{\alpha_{1,j'}}{\sigma_{1,j'}p_{j'}} \quad 1 \leq j < j' \leq D.$$

Letting $\beta_i = \frac{\alpha_{i,j}}{\sigma_{i,j}p_j}$ which is independent of j , we can write $\alpha_{i,j} = p_j\sigma_{i,j}\beta_i$ for all $i = 1, 2, \dots, K$ and $j = 1, 2, \dots, D$. Since α and $\beta = (\beta_1, \dots, \beta_K)$ are bijective, it is sufficient to show the uniqueness of β . Plugging $\alpha_{i,j}$ into (10) and (9), we have

$$\frac{(\bar{\mu}_1 - \bar{\mu}_i)^2}{\frac{\sum_{j=1}^D \sigma_{1,j}p_j}{\beta_1} + \frac{\sum_{j=1}^D \sigma_{i,j}p_j}{\beta_i}} = \frac{(\bar{\mu}_1 - \bar{\mu}_{i'})^2}{\frac{\sum_{j=1}^D \sigma_{1,j}p_j}{\beta_1} + \frac{\sum_{j=1}^D \sigma_{i',j}p_j}{\beta_{i'}}} \quad 2 \leq i < i' \leq K$$

with $\beta_1^2 = \sum_{i=2}^K \beta_i^2$. Let $\eta = \frac{\beta}{\beta_1}$. Then η satisfies

$$\frac{(\bar{\mu}_1 - \bar{\mu}_i)^2}{\sum_{j=1}^D \sigma_{1,j}p_j + \frac{\sum_{j=1}^D \sigma_{i,j}p_j}{\eta_i}} = \frac{(\bar{\mu}_1 - \bar{\mu}_{i'})^2}{\sum_{j=1}^D \sigma_{1,j}p_j + \frac{\sum_{j=1}^D \sigma_{i',j}p_j}{\eta_{i'}}} \quad 2 \leq i < i' \leq K \quad (15)$$

with $1 = \sum_{i=2}^K \eta_i^2$. If there exists $\eta' \neq \eta$ satisfying these two conditions, then there must be $i \neq k \neq 1$ such that $\eta_i < \eta'_i$ and $\eta_k > \eta'_k$. Then, we have

$$\frac{(\bar{\mu}_1 - \bar{\mu}_i)^2}{\sum_{j=1}^D \sigma_{1,j}p_j + \frac{\sum_{j=1}^D \sigma_{i,j}p_j}{\eta_i}} < \frac{(\bar{\mu}_1 - \bar{\mu}_i)^2}{\sum_{j=1}^D \sigma_{1,j}p_j + \frac{\sum_{j=1}^D \sigma_{i,j}p_j}{\eta'_i}} = \frac{(\bar{\mu}_1 - \bar{\mu}_k)^2}{\sum_{j=1}^D \sigma_{1,j}p_j + \frac{\sum_{j=1}^D \sigma_{k,j}p_j}{\eta'_k}} < \frac{(\bar{\mu}_1 - \bar{\mu}_k)^2}{\sum_{j=1}^D \sigma_{1,j}p_j + \frac{\sum_{j=1}^D \sigma_{k,j}p_j}{\eta_k}},$$

which contradicts (15). Hence, η is unique, which implies $\beta = C * \eta$ for some constant C . Then if there exists $\beta' \neq \beta$ and both are optimal, we have $\beta > (<)\beta'$. This implies the corresponding $\alpha > (<)\alpha'$, which contradicts $\sum_{i,j} \alpha_{i,j} = \sum_{i,j} \alpha'_{i,j} = 1$. ■

Compared with the optimality condition in [22], in addition to the ‘‘total balance’’ that characterizes the relation between the optimal design and the non-optimal designs and the ‘‘local balance’’ conditions that characterizes the relation between two non-optimal designs, here we have the additional optimality condition (5), the local ‘‘derivative balance’’ condition. It states that within the allocation for a certain design i , the partial derivative of the rate function G_i with respect to $\alpha_{i,j}$ is the same for all j 's. That is, simulation for each fixed input realization should provide the same improvement to identify that design 1 is better than i . Furthermore, with normally distributed simulation errors, equation (8) indicates that for a fixed design i the optimal allocation ratio $\alpha_{i,j}$ should be proportional to the input probability mass p_j and the standard deviation $\sigma_{i,j}$, which quantitatively characterizes how input uncertainty affects the optimal allocation policy. Also notice for fixed i , (8) only depends on i , which means the relative allocation ratios among different input realizations for a fixed design do not depend on other designs. On the other hand, (10) indicates that the relative allocation ratios among designs under the same input realization j are affected by all p_j 's, which implies directly applying OCBA to designs under a fixed input realization j may perform poorly since it does not take information from other design-input pairs into consideration. Moreover, notice that the three optimality conditions (5)-(7) not only hold for Gaussian simulation noise but also hold as long as the rate function G_i has the properties shown in Lemma 1.

4 Sequential Procedure with Streaming Input Data

In deriving Theorem 1 above, we assume full knowledge of the input distribution and simulation error distribution. In this section, by trying to satisfy the optimality conditions in Theorem 1 with the current input distribution estimate as ‘‘plug-in’’ estimate at each time stage, we develop two fully sequential procedures, named as IU-OCBA-approx and IU-OCBA-balance, for simulation budget allocation in the multi-stage setting with streaming input data. The two procedures mainly differ in how to satisfy the optimality conditions: IU-OCBA-approx solves the optimality conditions approximately, while IU-OCBA-balance tries to balance the two sides of the optimality equations.

A major difficulty of solving the optimality conditions is that the optimality equations (8)-(10) do not have closed-form solutions, and it is usually computationally expensive to solve them using numerical methods such as gradient decent. To improve the computational efficiency, we design the two procedures tackling the optimality conditions in different ways. The IU-OCBA-approx procedure tries to directly solve the optimization problem (3) at each iteration but approximating (10) by assuming that a weighted ratio of allocation budget assigned to the optimal design is much larger than that assigned to other designs, which is a similar assumption taken by [5]. This approximation enables us to compute the solution in a much simpler way. Alternatively, by taking a similar approach in [6], the IU-OCBA-balance

procedure avoids directly solving the optimality equations and instead balances the two sides of the equations, i.e., reduces the difference between two sides of the equations in each iteration when allocating the budget. Plausibly, IU-OCBA-approx is expected to converge faster since we solve the equations every time, while IU-OCBA-balance may converge slower since we only balance instead of solving the equations. However, IU-OCBA-approx approximates the optimality conditions, meaning that the ‘‘optimal solution’’ we get may not be really optimal in the original problem. IU-OCBA-balance, instead, targets at the original problem and will eventually converge to the true optimal solution as more and more data are collected. The empirical comparison of these two methods will be carried out numerically in Section 6.

4.1 IU-OCBA-approx

In this section we derive the IU-OCBA-approx procedure. Let $\beta_i = \frac{\alpha_{i,j}}{p_j \sigma_{i,j}}$, $1 \leq i \leq K$, which is independent of j by (8). Plugging $\alpha_{i,j} = \beta_i \sigma_{i,j} p_j$ into (10), we have

$$\frac{(\bar{\mu}_1 - \bar{\mu}_i)^2}{\frac{\sum_{j=1}^D \sigma_{1,j} p_j}{\beta_1} + \frac{\sum_{j=1}^D \sigma_{i,j} p_j}{\beta_i}} = \frac{(\bar{\mu}_1 - \bar{\mu}_{i'})^2}{\frac{\sum_{j=1}^D \sigma_{1,j} p_j}{\beta_1} + \frac{\sum_{j=1}^D \sigma_{i',j} p_j}{\beta_{i'}}}, \quad 2 \leq i < i' \leq K.$$

Assume $\beta_1 \gg \beta_i$, $\forall i \neq 1$, i.e., $\frac{\alpha_{1,j}}{\sigma_{1,j}} \gg \frac{\alpha_{i,j}}{\sigma_{i,j}} \forall i \neq 1, \forall j$, the simulation budget assigned to the optimal design-input pair divided by its standard deviation is much larger than that of other designs. Then we have $\frac{\beta_i}{\beta_{i'}} \approx \frac{\sum_{j=1}^D \sigma_{i,j} p_j / (\bar{\mu}_1 - \bar{\mu}_i)^2}{\sum_{j=1}^D \sigma_{i',j} p_j / (\bar{\mu}_1 - \bar{\mu}_{i'})^2} \forall i \neq i' \neq 1$. Plugging $\beta_i = \frac{\alpha_{i,j}}{p_j \sigma_{i,j}}$ back with this approximation, we have

$$\frac{\alpha_{i,j}}{\alpha_{i',j'}} = \frac{p_j \sigma_{i,j} \sum_{k=1}^D \sigma_{i,k} p_k / (\bar{\mu}_1 - \bar{\mu}_i)^2}{p_{j'} \hat{\sigma}_{i',j'}^{(l)} \sum_{k=1}^D \sigma_{i',k} p_k / (\bar{\mu}_1 - \bar{\mu}_{i'})^2}, \quad i, i' \neq 1. \quad (16)$$

Furthermore, with (9) and $\sum_{i=1}^K \sum_{j=1}^D \alpha_{i,j} = 1$, we can calculate $\alpha_{i,j}$ explicitly.

IU-OCBA-approx

1. **Input.** Number of designs K , input distribution support $\{\zeta_1, \zeta_2, \dots, \zeta_B\}$, initial sample size n_0 , total simulation budget n , input data batch size $\{m(t)\}_{t=1}^\infty$ and stage-wise simulation budget $\{n(t)\}_{t=1}^\infty$.
2. **Initialization.** Time stage counter $t \leftarrow 0$, replication counter $l \leftarrow 0$, total input data $M(t) \leftarrow 0$. Collect n_0 initial samples for each design-input pair (i, j) . Set $N_{i,j}^{(l)} = n_0$. Compute the initial sample mean $\hat{\mu}_{i,j}^{(l)} = \frac{1}{N_{i,j}^{(l)}} \sum_{s=1}^{N_{i,j}^{(l)}} X_{i,j}^{(s)}$, and sample standard deviation $\hat{\sigma}_{i,j}^{(l)} = \sqrt{\frac{1}{N_{i,j}^{(l)} - 1} \sum_{s=1}^{N_{i,j}^{(l)}} (X_{i,j}^{(s)} - \hat{\mu}_{i,j}^{(l)})^2}$.
3. **WHILE** $\sum_{i=1}^K \sum_{j=1}^D N_{i,j}^{(l)} < n$ **DO**
4. $t \leftarrow t + 1$, given input data of batch size $m(t)$, let $M(t) = \sum_{\tau=1}^t m(\tau)$ and update $p_j^{(t)} = \frac{\sum_{s=1}^{M(t)} \mathbb{1}\{\xi_s = \zeta_j\}}{M(t)}$ for $j = 1, 2, \dots, D$. Compute $\hat{\mu}_i^{(l)} = \sum_{j=1}^D p_j^{(t)} \hat{\mu}_{i,j}^{(l)}$.
5. **REPEAT** $n(t)$ **TIMES**
6. $\hat{b}^{(l)} \leftarrow \arg \max_i \hat{\mu}_i^{(l)}$.
7. Update $\hat{\alpha}_{i,j}^{(l)}$ using (16), (9) and $\sum_{i,j} \hat{\alpha}_{i,j} = 1$, with $p_j, \bar{\mu}_i, \sigma_{i,j}$ replaced by $p_j^{(t)}, \hat{\mu}_i^{(l)}$ and $\hat{\sigma}_{i,j}^{(l)}$, respectively. Calculate $\hat{N}_{i,j}^{(l)} = \hat{\alpha}_{i,j}^{(l)} \left(1 + \sum_{i=1}^K \sum_{j=1}^D N_{i,j}^{(l)} \right)$, $\forall 1 \leq i \leq K, 1 \leq j \leq D$.
8. Find the design-input pair index $(I, J) = \arg \max_{i,j} \left(\hat{N}_{i,j}^{(l)} - N_{i,j}^{(l)} \right)$. Simulate the pair (I, J) once. Update $\hat{\mu}_{I,J}^{(l+1)}, \hat{\sigma}_{I,J}^{(l)}$ and $\hat{\mu}_I^{(l+1)}$ using the new simulation output, and set $\hat{\mu}_{i,j}^{(l+1)} = \hat{\mu}_{i,j}^{(l)}, \hat{\sigma}_{i,j}^{(l+1)} = \hat{\sigma}_{i,j}^{(l)}$ and $\hat{\mu}_i^{(l+1)} = \hat{\mu}_i^{(l)}$ for $i \neq I, j \neq J$. Let $N_{I,J}^{(l+1)} = N_{I,J}^{(l)} + 1$ and $N_{i,j}^{(l+1)} = N_{i,j}^{(l)}$ for all $i \neq I, j \neq J$.
9. $l \leftarrow l + 1$.
10. **END REPEAT**
11. **END WHILE**
12. **Output:** Output $i_b = \arg \max_i \hat{\mu}_i^{(l)}$ as the best design.

4.2 IU-OCBA-balance

Unlike IU-OCBA-approx where we try to directly solve for the optimal solutions, IU-OCBA-balance only requires to evaluate both sides of the three optimality equations given the current number of replications for each design-input pair. The procedures select a design-input pair each time to reduce the difference (balance) of at least one of the optimality equations. In particular, at each iteration, the procedure will first decide whether to simulate the estimated best design or one of the non-optimal designs to balance the ‘‘total balance’’ conditions. If the estimated best design is not selected, then the procedure selects a non-optimal design to balance the ‘‘total balance’’ conditions. After selecting the design, an input realization is chosen by balancing the ‘‘derivative balance’’ conditions. Notice that although in (8) the ‘‘derivative balance’’ conditions are only for non-optimal designs, (8) also holds for $i = b$ by (9) with $\frac{\alpha_{i,j}}{\sigma_{i,j}}$ replaced by $\frac{\alpha_{i,j'} p_j}{\sigma_{i,j'} p_{j'}} , \forall j \neq j'$. The balancing approach utilizes the monotonicity of both sides of all optimality equations in terms of the allocation policy $\alpha_{i,j}$. For example, if we have in one of the equations in (8) violated by $\frac{\alpha_{i,j}}{\sigma_{i,j} p_j} < \frac{\alpha_{i,j'}}{\sigma_{i,j'} p_{j'}}$, then we may want to simulate the design-input pair (i, j) to make the left hand side larger. The IU-OCBA-balance procedure is presented as follows:

IU-OCBA-balance

1. **Input.** Number of designs K , input distribution support $\{\zeta_1, \zeta_2, \dots, \zeta_B\}$, initial sample size n_0 , total simulation budget n , input data batch size $\{m(t)\}_{t=1}^\infty$, and stage-wise simulation budget $\{n(t)\}_{t=1}^\infty$.
2. **Initialization.** Time stage counter $t \leftarrow 0$, replication counter $l \leftarrow 0$, total input data $M(t) \leftarrow 0$. Collect n_0 initial samples for each design-input pair (i, j) . Set $N_{i,j}^{(l)} = n_0$. Compute the initial sample mean $\hat{\mu}_{i,j}^{(l)} = \frac{1}{N_{i,j}^{(l)}} \sum_{s=1}^{N_{i,j}^{(l)}} X_{i,j}^{(s)}$, and sample standard deviation $\hat{\sigma}_{i,j}^{(l)} = \sqrt{\frac{1}{N_{i,j}^{(l)} - 1} \sum_{s=1}^{N_{i,j}^{(l)}} (X_{i,j}^{(s)} - \hat{\mu}_{i,j}^{(l)})^2}$.
3. **WHILE** $\sum_{i=1}^K \sum_{j=1}^D N_{i,j}^{(l)} < n$ **DO**
4. $t \leftarrow t + 1$, given input data of batch size $m(t)$, let $M(t) = \sum_{\tau=1}^t m(\tau)$ and update $p_j^{(t)} = \frac{\sum_{s=1}^{M(t)} \mathbf{1}\{\xi_s = \zeta_j\}}{M(t)}$ for $j = 1, 2, \dots, D$. Compute $\hat{\mu}_i^{(l)} = \sum_{j=1}^D p_j^{(t)} \hat{\mu}_{i,j}^{(l)}$.
5. **REPEAT** $n(t)$ **TIMES**
6. $\hat{b}^{(l)} \leftarrow \arg \max_i \hat{\mu}_i^{(l)}$.
7. Let $\hat{j}^* = \arg \min_j \left\{ \left(\frac{N_{\hat{b}^{(l)},j}^{(l)}}{\hat{\sigma}_{\hat{b}^{(l)},j}^{(l)}} \right)^2 - \sum_{i \neq \hat{b}^{(l)}} \left(\frac{N_{i,j}^{(l)}}{\hat{\sigma}_{i,j}^{(l)}} \right)^2 \right\}$.
8. **IF** $\left(\frac{N_{\hat{b}^{(l)},\hat{j}^*}^{(l)}}{\hat{\sigma}_{\hat{b}^{(l)},\hat{j}^*}^{(l)}} \right)^2 - \sum_{i \neq \hat{b}^{(l)}} \left(\frac{N_{i,\hat{j}^*}^{(l)}}{\hat{\sigma}_{i,\hat{j}^*}^{(l)}} \right)^2 < 0$, set $I_l = \hat{b}^{(l)}$, $J_l = \arg \min_j \frac{N_{\hat{b}^{(l)},j}^{(l)}}{\hat{\sigma}_{\hat{b}^{(l)},j}^{(l)} p_j^{(t)}}$.
9. **ELSE** set $I_l = \arg \min_{i \neq \hat{b}^{(l)}} \frac{(\hat{\mu}_{\hat{b}^{(l)}}^{(l)} - \hat{\mu}_i^{(l)})^2}{\sum_{j=1}^D \frac{(\hat{\sigma}_{\hat{b}^{(l)},j}^{(l)})^2 (p_j^{(t)})^2}{N_{\hat{b}^{(l)},j}^{(l)}} + \sum_{j=1}^D \frac{(\hat{\sigma}_{i,j}^{(l)})^2 (p_j^{(t)})^2}{N_{i,j}^{(l)}}$, $J_l = \arg \min_j \frac{N_{I_l,j}^{(l)}}{\hat{\sigma}_{I_l,j}^{(l)} p_j^{(t)}}$.
10. **END IF**
11. Simulate the pair (I, J) once. Update $\hat{\mu}_{I,J}^{(l+1)}$, $\hat{\sigma}_{I,J}^{(l)}$ and $\hat{\mu}_I^{(l+1)}$ using the new simulation output, and set $\hat{\mu}_{i,j}^{(l+1)} = \hat{\mu}_{i,j}^{(l)}$, $\hat{\sigma}_{i,j}^{(l+1)} = \hat{\sigma}_{i,j}^{(l)}$ and $\hat{\mu}_i^{(l+1)} = \hat{\mu}_i^{(l)}$ for $i \neq I, j \neq J$. Let $N_{I,J}^{(l+1)} = N_{I,J}^{(l)} + 1$ and $N_{i,j}^{(l+1)} = N_{i,j}^{(l)}$ for all $i \neq I, j \neq J$.
12. $l \leftarrow l + 1$.
13. **END REPEAT**
14. **END WHILE**

5 CONSISTENCY AND ASYMPTOTIC OPTIMALITY

In this section we show consistency and asymptotic optimality of IU-OCBA-approx and IU-OCBA-balance. The consistency of classical R&S procedures, in a nutshell, is usually guaranteed by the Strong Law of Large Number

(SLLN) as long as we simulate each design infinitely many times. However, here we also need the convergence of input distribution estimate $\{p_j^{(t)}\}_{1 \leq j \leq D}$ to ensure we have the correct estimate of the true expected performance. For this purpose, we make the following assumption about the input data batch size and simulation budget in each stage for the IU-OCBA-approx procedure.

Assumption 4 *At stage t , the input data batch size $m(t)$ and simulation budget $n(t)$ satisfy*

$$\lim_{T \rightarrow \infty} \sum_{t=1}^T n(t) = \infty, \quad \lim_{T \rightarrow \infty} \sum_{t=1}^T m(t) = \infty.$$

Assumption 4 ensures that both the total amount of input data and total simulation replications go to infinity as time stage T goes to infinity, which helps guarantee the consistency of both the input estimate and the design-input performance estimate. Recall that t denotes the time stage and l denotes the amount of simulation budget that has been allocated so far. The following theorem shows the consistency (with respect to t) and asymptotic optimality (with respect to l) of IU-OCBA-approx.

Theorem 2 *Suppose Assumption 1, 2, 3 and 4 hold and the total simulation budget $n = \infty$. Then,*

1. (Consistency) IU-OCBA-approx selects the optimal design almost surely as $t \rightarrow \infty$.
2. (Asymptotic optimality) $\lim_{l \rightarrow \infty} \frac{N_{i,j}^{(l)}}{N^{(l)}} = \alpha_{i,j}^*$ almost surely, $1 \leq i \leq K, 1 \leq j \leq D$, where $\alpha_{i,j}^*$ satisfies the optimality conditions (16) and (9).

Proof Since $\sum_{s=1}^t n(s) \rightarrow \infty$ as $t \rightarrow \infty$, we know the empirical distribution $p_j^{(t)} \rightarrow p_j$ almost surely by Glivenko-Cantelli Theorem. Hence, if we can show $N_{i,j}^{(l)} \rightarrow \infty$ almost surely for all i, j , we then have $\hat{\mu}_i^{(l)} \rightarrow \bar{\mu}_i$ and $\hat{\sigma}_{i,j}^{(l)} \rightarrow \sigma_{i,j}$ almost surely by the SLLN and the fact that $p_j^{(t)} \rightarrow p_j$ almost surely. Hence, to prove 1 and 2, it suffices to show $N_{i,j}^{(l)} \rightarrow \infty$ almost surely. Denote by $A = \{(i, j) | N_{i,j}^{(l)} \rightarrow \infty\}$. Clearly $A \neq \emptyset$.

Proof of 1. Denote by ω any sample path of one simulation process. We fix a sample path ω in the following proof. Prove by contradiction. Suppose there exists $(i_0, j_0) \notin A$. Notice $\hat{\mu}_{i_0, j_0}^{(l)}$ and $\hat{\sigma}_{i_0, j_0}^{(l)}$ will converge almost surely, no matter whether $N_{i_0, j_0}^{(l)}$ tend to infinity. This is because if $N_{i_0, j_0}^{(l)}$ is at most finite, then $\hat{\mu}_{i_0, j_0}^{(l)}$ and $\hat{\sigma}_{i_0, j_0}^{(l)}$ will remain unchanged after finite iterations. Since $p_j^{(t)}$ converges almost surely, $\hat{\mu}_{i_0, j_0}^{(l)}$ will also converge almost surely. Denote by $N^{(l)} = \sum_{i,j} N_{i,j}^{(l)}$.

Then there exists an allocation policy $\{\tilde{\alpha}_{i,j}\}$ satisfying $\lim_{l \rightarrow \infty} \frac{\tilde{N}_{i_0, j_0}^{(l)}}{N^{(l)}} = \tilde{\alpha}_{i_0, j_0}$. Since (i_0, j_0) can be sampled for at most finitely many times, it must hold for l large enough, $\tilde{N}_{i_0, j_0}^{(l)} - N_{i_0, j_0}^{(l)} \leq \tilde{N}_{i_0, j_0}^{(l)} - N_{i_0, j_0}^{(l)}$ for any $(i, j) \in A$. Then we have

$$\liminf_{l \rightarrow \infty} \frac{\tilde{N}_{i_0, j_0}^{(l)} - N_{i_0, j_0}^{(l)}}{N^{(l)}} \leq \liminf_{l \rightarrow \infty} \frac{\tilde{N}_{i,j}^{(l)} - N_{i,j}^{(l)}}{N^{(l)}}. \quad (17)$$

The left hand side (LHS) of (17) = $\lim_{l \rightarrow \infty} \frac{\tilde{N}_{i_0, j_0}^{(l)}}{N^{(l)}} = \tilde{\alpha}_{i_0, j_0} > 0$, where positiveness comes from the fact that $\alpha_{i,j} > 0$ by proof of Theorem 1, where we replace the true $\bar{\mu}$ with the limit of $\hat{\mu}^{(l)}$ and the true $\sigma_{i,j}$ with the limit of $\hat{\sigma}_{i,j}^{(l)}$. On the other hand, the right hand side (RHS) of (17) = $\lim_{l \rightarrow \infty} \frac{\tilde{N}_{i,j}^{(l)}}{N^{(l)}} - \limsup_{l \rightarrow \infty} \frac{N_{i,j}^{(l)}}{N^{(l)}} = \tilde{\alpha}_{i,j} - \limsup_{l \rightarrow \infty} \frac{N_{i,j}^{(l)}}{N^{(l)}}$. Hence, we have $\limsup_{l \rightarrow \infty} \frac{N_{i,j}^{(l)}}{N^{(l)}} \leq \tilde{\alpha}_{i,j} - \tilde{\alpha}_{i_0, j_0} < \tilde{\alpha}_{i,j}$. Since this holds for all $(i, j) \in A$, we have

$$1 = \sum_{i,j} \frac{N_{i,j}^{(l)}}{N^{(l)}} = \limsup_{l \rightarrow \infty} \sum_{i,j} \frac{N_{i,j}^{(l)}}{N^{(l)}} \leq \sum_{i,j} \limsup_{l \rightarrow \infty} \frac{N_{i,j}^{(l)}}{N^{(l)}} = \sum_{(i,j) \in A} \limsup_{l \rightarrow \infty} \frac{N_{i,j}^{(l)}}{N^{(l)}} < \sum_{(i,j) \in A} \tilde{\alpha}_{i,j} \leq 1 - \tilde{\alpha}_{i_0, j_0} < 1,$$

a contradiction.

Proof of 2. Again we fix some sample path ω . By Assumption 4, we have $l \rightarrow \infty \iff t \rightarrow \infty$. Since $N_{i,j}^{(l)} \rightarrow \infty$ and $p_j^{(t)} \rightarrow p_j$ as $l \rightarrow \infty$, we have $\hat{\alpha}_{i,j}^{(l)} \rightarrow \alpha_{i,j}^*$ as $l \rightarrow \infty$. Then, $\forall \varepsilon > 0, \exists \tilde{L}$ such that $|\hat{\alpha}_{i,j}^{(l)} - \alpha_{i,j}^*| < \varepsilon$ and $\frac{1}{N^{(l)}} < \varepsilon$ for all $l \geq \tilde{L}$ and $i \geq 1, j \geq 1$. Let $\tilde{L}_{i,j} = \min\{l > \tilde{L} : N_{i,j}^{(l)} = N_{i,j}^{(\tilde{L})} + 1\} \forall (i, j) \in A$, the first time (i, j) is sampled after \tilde{L} . Let $L = \max_{(i,j) \in A} \tilde{L}_{i,j} < \infty$ by the definition of A . Then for any $l > L$, let $D_l = \{(i, j) : \hat{\alpha}_{i,j}^{(l)} - \frac{N_{i,j}^{(l)}}{N^{(l)}} < 0\}$. Then

if $(i, j) \in D_l$, let $L_{i,j} = \max\{s < l : N_{i,j}^{(s)} = N_{i,j}^{(l)} - 1\}$. Then we have $L_{i,j} \geq \tilde{L}$ by the definition of L . Furthermore,

$$\hat{\alpha}_{i,j}^{(l)} - \frac{N_{i,j}^{(l)}}{N^{(l)}} \geq \hat{\alpha}_{i,j}^{(l)} - \frac{N_{i,j}^{(L_{i,j}+1)}}{N^{(L_{i,j}+1)}} = \underbrace{\left[\hat{\alpha}_{i,j}^{(L_{i,j})} - \frac{N_{i,j}^{(L_{i,j})}}{N^{(L_{i,j})}} \right]}_{= E_1} + \underbrace{\left[\hat{\alpha}_{i,j}^{(l)} - \hat{\alpha}_{i,j}^{(L_{i,j})} \right]}_{= E_2} + \underbrace{\left[\frac{N_{i,j}^{(L_{i,j})}}{N^{(L_{i,j})}} - \frac{N_{i,j}^{(L_{i,j}+1)}}{N^{(L_{i,j}+1)}} \right]}_{= E_3}, \quad (18)$$

where the second inequality follows from $N_{i,j}^{(l)} = N_{i,j}^{(L_{i,j}+1)}$ and $N^{(L_{i,j}+1)} < N^{(l)}$ by the definition of $L_{i,j}$. Since (i, j) is sampled at $L_{i,j}$, we must have $E_1 \geq 0$. Further since $l, L_{i,j} \geq \tilde{L}$, we have $E_2 \geq -|\alpha_{i,j}^* - \hat{\alpha}_{i,j}^{(l)}| - |\alpha_{i,j}^* - \hat{\alpha}_{i,j}^{(L_{i,j})}| > -2\varepsilon$ and $E_3 = \frac{N_{i,j}^{(L_{i,j})}}{N^{(L_{i,j})}} - \frac{N_{i,j}^{(L_{i,j}+1)}}{N^{(L_{i,j}+1)}} = \frac{N_{i,j}^{(L_{i,j})} - N_{i,j}^{(L_{i,j}+1)}}{N^{(L_{i,j})}(N^{(L_{i,j}+1)})} > \frac{-N^{(L_{i,j})}}{N^{(L_{i,j})}(N^{(L_{i,j}+1)})} = \frac{-1}{N^{(L_{i,j}+1)}} > -\varepsilon$. Hence, we have $\hat{\alpha}_{i,j}^{(l)} - \frac{N_{i,j}^{(l)}}{N^{(l)}} \geq -3\varepsilon$. As a result,

$$0 = \sum_{i,j} \hat{\alpha}_{i,j}^{(l)} - \sum_{i,j} \frac{N_{i,j}^{(l)}}{N^{(l)}} = \sum_{(i,j) \in D_l} \left(\hat{\alpha}_{i,j}^{(l)} - \frac{N_{i,j}^{(l)}}{N^{(l)}} \right) + \sum_{(i,j) \in D_l^c} \left(\hat{\alpha}_{i,j}^{(l)} - \frac{N_{i,j}^{(l)}}{N^{(l)}} \right) \geq -3|D_l|\varepsilon + \sum_{(i,j) \in D_l^c} \left(\hat{\alpha}_{i,j}^{(l)} - \frac{N_{i,j}^{(l)}}{N^{(l)}} \right)$$

Hence, $0 \leq \max_{i,j} \left\{ \hat{\alpha}_{i,j}^{(l)} - \frac{N_{i,j}^{(l)}}{N^{(l)}} \right\} \leq \sum_{(i,j) \in D_l^c} \left(\hat{\alpha}_{i,j}^{(l)} - \frac{N_{i,j}^{(l)}}{N^{(l)}} \right) \leq 3|D_l|\varepsilon \leq 3KB\varepsilon$. By arbitrary $\varepsilon > 0$, we get $\lim_{l \rightarrow \infty} \max_{i,j} \left(\hat{\alpha}_{i,j}^{(l)} - \frac{N_{i,j}^{(l)}}{N^{(l)}} \right) = 0$. Since for any (i_0, j_0) we have

$$\max_{i,j} \left\{ \hat{\alpha}_{i,j}^{(l)} - \frac{N_{i,j}^{(l)}}{N^{(l)}} \right\} \geq \hat{\alpha}_{i_0, j_0}^{(l)} - \frac{N_{i_0, j_0}^{(l)}}{N^{(l)}} = - \sum_{i \neq i_0, j \neq j_0} \left(\hat{\alpha}_{i,j}^{(l)} - \frac{N_{i,j}^{(l)}}{N^{(l)}} \right) \geq -(KB - 1) \max_{i,j} \left\{ \hat{\alpha}_{i,j}^{(l)} - \frac{N_{i,j}^{(l)}}{N^{(l)}} \right\},$$

we obtain $\lim_{l \rightarrow \infty} \hat{\alpha}_{i_0, j_0}^{(l)} - \frac{N_{i_0, j_0}^{(l)}}{N^{(l)}} = 0 = \alpha_{i_0, j_0}^* - \lim_{l \rightarrow \infty} \frac{N_{i_0, j_0}^{(l)}}{N^{(l)}}$ as desired. \blacksquare

For the IU-OCBA-balance procedure, we need more regularity conditions on the input data batch size and stage-wise simulation budget to obtain the asymptotic optimality, as shown in the following assumption.

Assumption 5 *At stage t , the input data batch size $m(t)$ and simulation budget $n(t)$ satisfy*

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T n(t) = \bar{n}, \quad \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T m(t) = \bar{m}$$

for some $\bar{n}, \bar{m} > 0$.

Assumption 5 ensures that both the total amount of input data and total simulation replications go to infinity at the same rate $O(T)$ as time stage T goes to infinity. The same linear increasing rate allows the usage of law of iterated logarithm (LII) on both the design-input performance estimator $\{\hat{\mu}_{i,j}^{(l)}\}_{i,j}$ and empirical input pmf $\{p_j^{(t)}\}_j$ and obtain almost sure concentration bounds of order $O\left(\sqrt{\frac{\log \log t}{t}}\right)$, $O\left(\sqrt{\frac{\log \log t}{t}}\right)$. The convergence of the allocation policy relies on this uniform bound on errors from IU and SU. The following theorem shows the consistency and asymptotic optimality of IU-OCBA-balance.

Theorem 3 *Suppose Assumption 1, 2, 3 and 5 hold and the total simulation budget $n = \infty$. Then,*

1. (Consistency) IU-OCBA-balance selects the optimal design almost surely as $t \rightarrow \infty$.

2. (Asymptotic optimality) Define $\alpha_{ij}^{(l)} = \frac{N_{i,j}^{(l)}}{N^{(l)}} \forall i, j$. Then $\alpha_{ij}^{(l)}$ asymptotically satisfies the optimality condition (8), (9) and (10). That is, as $l \rightarrow \infty$,

$$i. \left| \frac{\alpha_{i,j}^{(l)}}{\sigma_{i,j} p_j} - \frac{\alpha_{i,j'}^{(l)}}{\sigma_{i,j'} p_{j'}} \right| \rightarrow 0 \quad a.s. \quad 2 \leq i \leq K, 1 \leq j < j' \leq D, \quad (19)$$

$$ii. \left| \left(\frac{\alpha_{1,j}^{(l)}}{\sigma_{1,j}} \right)^2 - \sum_{i=2}^D \left(\frac{\alpha_{i,j}^{(l)}}{\sigma_{i,j}} \right)^2 \right| \rightarrow 0 \quad a.s. \quad 1 \leq j \leq D, \quad (20)$$

$$iii. \left| \frac{(\bar{\mu}_1 - \bar{\mu}_i)^2}{\sum_{j=1}^D \frac{\sigma_{1,j}^2 p_j^2}{\alpha_{1,j}^{(l)}} + \sum_{j=1}^D \frac{\sigma_{i,j}^2 p_j^2}{\alpha_{i,j}^{(l)}}} - \frac{(\bar{\mu}_1 - \bar{\mu}_{i'})^2}{\sum_{j=1}^D \frac{\sigma_{1,j}^2 p_j^2}{\alpha_{1,j}^{(l)}} + \sum_{j=1}^D \frac{\sigma_{i',j}^2 p_j^2}{\alpha_{i',j}^{(l)}}} \right| \rightarrow 0 \quad a.s. \quad 2 \leq i < i' \leq K. \quad (21)$$

Proof of 1. Again denote by ω any sample path of one simulation process and $A = \{(i, j) | N_{i,j}^{(l)} \rightarrow \infty\}$. Fix the sample path ω . We first prove $(\hat{b}, j) \in A$ for some j , where $\hat{b} = \lim_{l \rightarrow \infty} \hat{b}^{(l)}$. We prove by contradiction. Suppose $(\hat{b}, j) \notin A, \forall j$. Then there exists $L > 0$, such that \hat{b} will not be sampled after L th sample. Take any $(i', j') \in A$,

then $\left(\frac{N_{\hat{b},j'}^{(l)}}{\hat{\sigma}_{\hat{b},j'}^{(l)}} \right)^2 - \sum_{i \neq \hat{b}} \left(\frac{N_{i,j'}^{(l)}}{\hat{\sigma}_{i,j'}^{(l)}} \right)^2 \leq \left(\frac{N_{\hat{b},j'}^{(l)}}{\hat{\sigma}_{\hat{b},j'}^{(l)}} \right)^2 - \left(\frac{N_{i',j'}^{(l)}}{\hat{\sigma}_{i',j'}^{(l)}} \right)^2 < 0$ for l large enough, contradicting to $\hat{b}^{(l)} = \hat{b}$

will not be sampled for $l > L$. Hence, there exists j_0 such that $(\hat{b}, j_0) \in A$. Second, We prove $(\hat{b}, j) \in A, \forall j$. Since

$\frac{N_{\hat{b},j_0}^{(l)}}{\hat{\sigma}_{\hat{b},j_0}^{(l)} p_{j_0}^{(t)}} \rightarrow \infty$ as $l \rightarrow \infty$, then by Line 8 in IU-OCBA-balance, we have for each j , $\frac{N_{\hat{b},j}^{(l)}}{\hat{\sigma}_{\hat{b},j}^{(l)} p_j^{(t)}} \rightarrow \infty$ as $l \rightarrow \infty$. Then,

since $p_j^{(t)} \rightarrow p_j > 0$, and the fact that $\sigma_{\hat{b},j}^{(l)}$ converges almost surely, no matter whether $N_{\hat{b},j}^{(l)}$ goes to infinity, we have $N_{\hat{b},j}^{(l)} \rightarrow \infty, \forall j$. The third step is to prove for each j , there exists an $i \neq \hat{b}$ such that $N_{i,j}^{(l)} \rightarrow \infty$. Otherwise if for some

j , all (i, j) pairs, $i \neq \hat{b}$ will only be simulated finitely many times, then $\left(\frac{N_{\hat{b},j}^{(l)}}{\hat{\sigma}_{\hat{b},j}^{(l)}} \right)^2 - \sum_{i \neq \hat{b}(l)} \left(\frac{N_{i,j}^{(l)}}{\hat{\sigma}_{i,j}^{(l)}} \right)^2 > 0$ for all l

large enough, a contradiction to (\hat{b}, j) will be sampled infinitely many times. Consequently, fix any j_0 , there exists $i_0 \neq \hat{b}$ such that $N_{i_0,j_0}^{(l)} \rightarrow \infty$. For this specific i_0 , by Line 9 in IU-OCBA-balance, we have $(i_0, j) \in A \forall j$, since

$\frac{N_{i_0,j_0}^{(l)}}{\hat{\sigma}_{i_0,j_0}^{(l)} p_j^{(t)}} \rightarrow \infty$. Finally, we show $N_{i,j}^{(l)} \rightarrow \infty, \forall i \neq \hat{b}, \forall j$. Since $\frac{(\hat{\mu}_{\hat{b}(l)}^{(l)} - \hat{\mu}_{i_0}^{(l)})^2}{\sum_{j=1}^D \frac{(\hat{\sigma}_{\hat{b}(l),j}^{(l)})^2 (p_j^{(t)})^2}{N_{\hat{b}(l),j}^{(l)}} + \sum_{j=1}^D \frac{(\hat{\sigma}_{i_0,j}^{(l)})^2 (p_j^{(t)})^2}{N_{i_0,j}^{(l)}}} \rightarrow \infty$, we

must have $\forall i \neq \hat{b}, \frac{(\hat{\mu}_{\hat{b}(l)}^{(l)} - \hat{\mu}_i^{(l)})^2}{\sum_{j=1}^D \frac{(\hat{\sigma}_{\hat{b}(l),j}^{(l)})^2 (p_j^{(t)})^2}{N_{\hat{b}(l),j}^{(l)}} + \sum_{j=1}^D \frac{(\hat{\sigma}_{i,j}^{(l)})^2 (p_j^{(t)})^2}{N_{i,j}^{(l)}}} \rightarrow \infty$. This implies that $N_{i,j}^{(l)} \rightarrow \infty, \forall i, j$.

Sketch of proof of 2. The proof of the asymptotic optimality of IU-OCBA-balance is quite technical. So, we only give a ketch of the proof here and leave the formal proof in the appendix. The proof takes a similar approach as [14] and [6]. We will show that between any two successive samples of a certain design-input pair, the number of replications that can be assigned to any other design is $O(\sqrt{l \log \log l})$, where l is the iteration at which the first of the two successive samples is sampled. Furthermore, for each sample, the fluctuation of both terms in (19), (20) and (21) is $O(\frac{1}{l})$. Then,

we can bound (19), (20) and (21) by $O\left(\sqrt{\frac{\log \log l}{l}}\right)$, which proves the result. \blacksquare

6 NUMERICAL EXPERIMENT

We test the performance of IU-OCBA-approx and IU-OCBA-balance by comparing with (i) Equal Allocation, which allocates equal simulation budget to all design-input pairs; (ii) Equal-OCBA-approx/Equal-OCBA-balance, which

assigns equal budget to input realizations and but uses OCBA-approx or OCBA-balance to allocate budget to different designs for each fixed input realization. We test the procedures on two different problems: a simple quadratic problem with Gaussian simulation error, and a portfolio optimization problem with non-Gaussian simulation noise.

6.1 Quadratic Problem

In the quadratic problem, we want to minimize the expected value of a quadratic function $X_i = (-0.5 + 0.5 * i - \zeta)^2 + \varepsilon_i(\zeta)$, where $i \in \mathcal{I} = [0, 10] \cap \mathbf{Z}$, ζ takes value in $\{0, 1, \dots, 4\}$ with probability p_j , $j = 0, 1, \dots, 4$, and $\varepsilon_{i,j}$ follows the normal distribution with mean 0 and stand deviation $\sigma_{i,j}$. We set $p_j \propto j + 5$, and $\sigma_{i,j} = 1 + \frac{1}{i+j+1}$. The true best design is $b = \arg \min_{i \in \mathcal{I}} \sum_{j=0}^5 (-0.5 + 0.5 * i - j)^2 p_j = 5$.

Experiment Results

We first test with constant data batch size with different initial number of replications, $n(0)$, which affect the initial estimation of parameter $\{\bar{\mu}_i\}_i$ and $\{\sigma_{i,j}\}_{i,j}$. We also test different stage-wise simulation budgets, $n(t)$, which together with $m(t)$ affect the impact from SU and IU on the empirical PCS. We set the batch size of input data $m(t) = 50$, initial input data batch size $m(0) = 50$, and total simulation budget $\sum_{t=1}^T n(t) = 20000$. For different choices of $n(0)$, we set $n(t) = 50$; for different choices of $n(t)$, we set $n(0) = 200$. In each scenario, we run 200 macro-replications to compute the empirical PCS, i.e., the ratio of macro-replications where the best design is correctly selected. Figure 2 and 3 show the empirical PCS with respect to the total simulation budget.

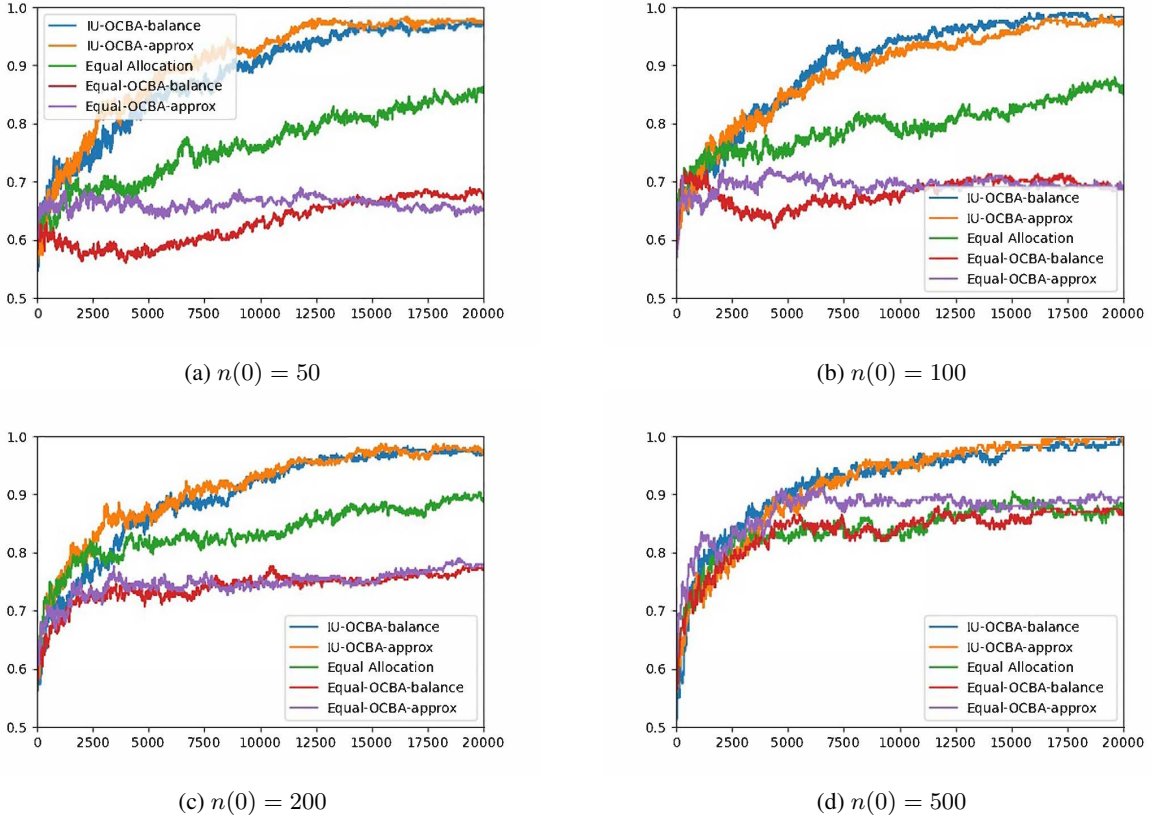


Figure 2: Quadratic example with constant batch size and different choices of $n(0)$

The observations from Figure 2 and 3 are summarized as follows:

1. For all choices of $n(0)$ and $n(t)$, the two proposed procedures IU-OCBA-approx and IU-OCBA-balance outperform the other three procedures, showing great efficiency in achieving much higher PCS with the same

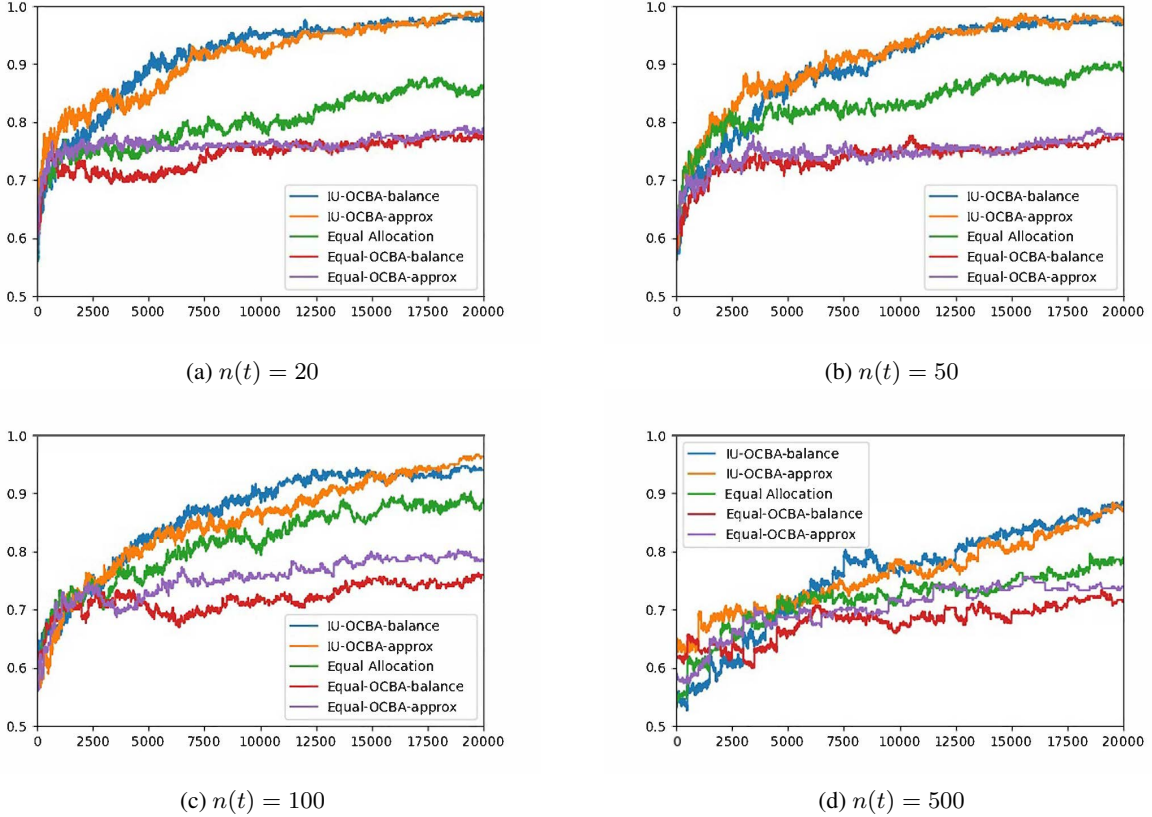


Figure 3: Quadratic example with constant batch size and different choices of $n(t)$

simulation budget. In addition, IU-OCBA-balance and IU-OCBA-approx have similar performance, implying that (16) is a proper approximation of local balance condition (10) in this example.

2. For the different choices of $n(0)$, Figure 2 indicates both of the “balance” procedures perform worse than the corresponding “approx” procedures when $n(0)$ is small. The reason is that IU-OCBA-balance and Equal-OCBA-balance are more sensitive to the estimation error, especially the estimation error in variance, which affects how fast it will converge to the optimal policy.
3. For the different choices of $n(t)$, we see the performance of five procedures become more and more similar as the stage-wise budget $n(t)$ becomes larger. The reason is that as $n(t)$ becomes larger, SU reduces and IU becomes dominant in the estimation error. Figure 3d indicates all five procedures perform poorly to achieve a high PCS, which implies balancing the effort on reducing IU and SU is important to achieve a high PCS.
4. It may be surprising that the Equal Allocation procedure outperforms Equal-OCBA-balance and Equal-OCBA-approx in most scenarios, even though it is the simplest one and does not utilize any information from the simulation outputs. The reason is that ranking of the design performances conditioned on an input realization can be drastically different from ranking of the unconditional expected performances. Both Equal-OCBA-approx and Equal-OCBA-balance tend to allocate more simulation budget to the design-input pair (i, j) where i is optimal under the input realization ζ_j but can be suboptimal when averaging with respect to the entire input distribution. These simulations contribute little to finding the optimal design, making these procedures even worse than equal allocation since they prohibit the chance of simulating the best design b under the same input realization ζ_j .

The previous experiments are carried out with data constant batch size in each stage, and next we test with random batch sizes, where the stage-wise budget and input data batch size are taken as $m(t) = \tilde{m} * Z$ and $n(t) = \tilde{n} * Z$ with Z being a random variable equally distributed among $\{1, 2, 3, 4, 5\}$. We set $\tilde{m} = 20$ for testing different \tilde{n} and $n(0)$. The results of empirical PCS with respect to the total number of simulation replications are shown in Figure 4 and 5.

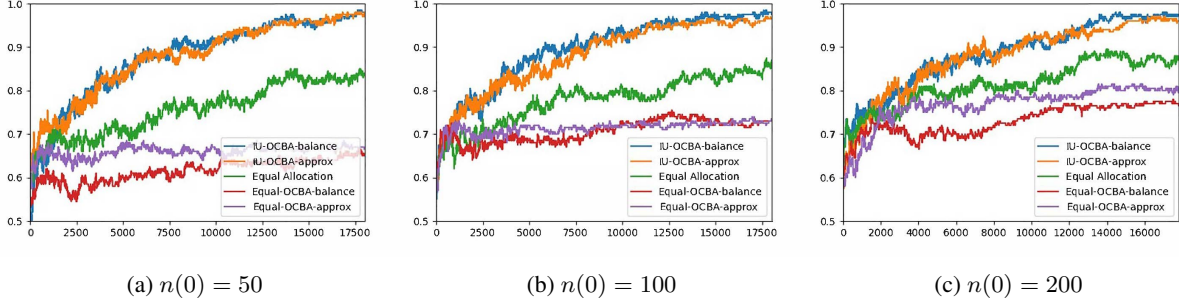
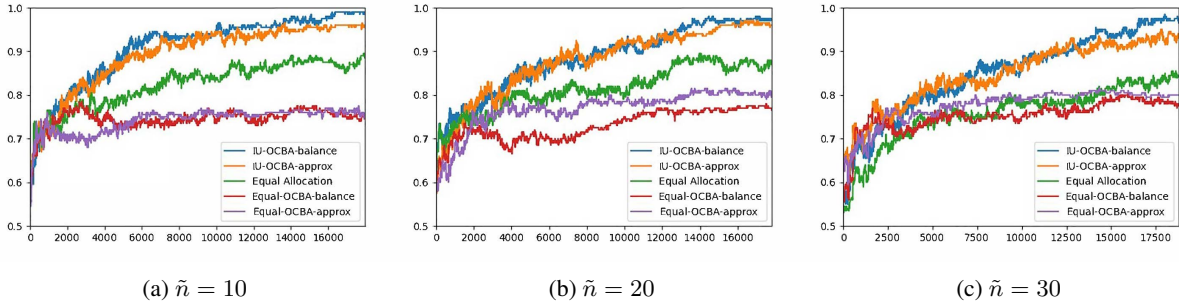

Figure 4: Quadratic example with random batch size: different choices of $n(0)$

Figure 5: Quadratic example with random batch size: different choices of \tilde{n}

Figure 4 and 5 show that both our proposed procedures work well in every scenario with random batch sizes. Other similar observations can be drawn as in the case of constant batch size.

6.2 Portfolio Optimization

We also demonstrate the performance of our proposed procedures on a more general problem of portfolio optimization, where the simulation error does not follow a normal distribution. Although the procedures are derived under the normal assumption, the numerical results show they still perform well in this example. We consider a portfolio optimization problem where an investor invests a certain amount of capital in a riskless asset with interest rate r and a risky asset, whose price per share at time t is denoted as S_t . Suppose $\{S_t\}$ follows a Geometric Brownian motion with initial price S_0 , which admits the following expression for any fixed t :

$$S_t = S_0 \exp \left[\left(\mu - \frac{\sigma^2}{2} \right) t + \sigma B(t) \right],$$

where σ is the volatility parameter, μ is the drift, and $\{B(t) : t \geq 0\}$ is a standard Brownian motion. At time 0, the investor makes a one-time decision $x \in [0, 1]$, which is the proportion of investment in the risky asset. Then, the total wealth at time t , denoted by W_t , is

$$W_t = xW_0 \exp \left[\left(\mu - \frac{\sigma^2}{2} \right) t + \sigma B(t) \right] + (1-x)W_0 e^{rt}.$$

A risk-averse investor aims to maximize the mean-variance of the total asset after T length of time with a risk-averse parameter ρ , as follows:

$$F(x) = \mathbf{E} \left\{ xW_0 \exp \left[\left(\mu - \frac{\sigma^2}{2} \right) T + \sigma B(T) \right] + (1-x)W_0 e^{rT} \right\} - \rho x^2 W_0^2 e^{2(\mu - \frac{\sigma^2}{2})T} (e^{\sigma^2 T} - 1) e^{\sigma^2 T}.$$

Here the variance term is calculated explicitly using the distribution of log-normal random variable, and the expectation term needs to be estimated. We set the volatility parameter $\sigma = 0.1$, time length $T = 5$, interest rate of riskless asset $r = 0.1$, risk-averse parameter $\rho = 0.2$, and initial asset price $S_0 = 1$. The drift μ takes values randomly in $\mathbf{J} = \{0.2 + 0.1 * j : j = 0, 1, \dots, 8\}$ with unknown pmf $p_j \propto j + |\mathbf{J}|/2$, which needs to be estimated via streaming

i.i.d. input data. The i th candidate design $x_i = 0.1 * i$, $i = 0, 1, \dots, 10$. For simulation under (x_i, μ_j) , we generate a sample z of $B(t)$, (i.e., from a normal distribution with mean 0 and variance t) and compute

$$X_{i,j} = x_i S_0 \exp \left[\left(\mu_j - \frac{\sigma^2}{2} \right) t + \sigma z \right] + (1 - x_i) S_0 e^{rt} - \rho x_i^2 S_0^2 e^{2(\mu_j - \frac{\sigma^2}{2})t} (e^{\sigma^2 t} - 1) e^{\sigma^2 t}$$

as a single simulation output on design-input pair $i-j$.

Experiment Result

As in the quadratic example, we also test with both constant batch size and random batch size, and vary initial number of samples $n(0)$ and stage-wise budget $n(t)$. We set the initial input data batch size $m(0) = 50$, constant input data batch size $m(t) = 50$, random input data batch size $\tilde{m} = 20$ and $m(t) = \tilde{m} * Z$, $n(t) = \tilde{n} * Z$, where Z is uniformly distribution in $\{1, 2, 3, 4, 5\}$. For varying $n(0)$, we set $n(t) = 50$ for constant batch size and $\tilde{n} = 20$ for random batch size. For varying $n(t)$ (or equivalently, \tilde{n}), we set $n(0) = 300$. Figure 6-9 show the results of empirical PCS with respect to total simulation budget allocated.

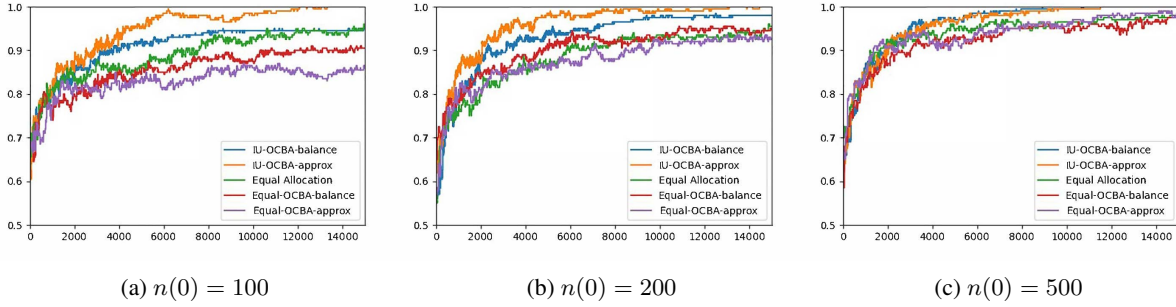


Figure 6: Portfolio example with constant batch size and different choices of $n(0)$

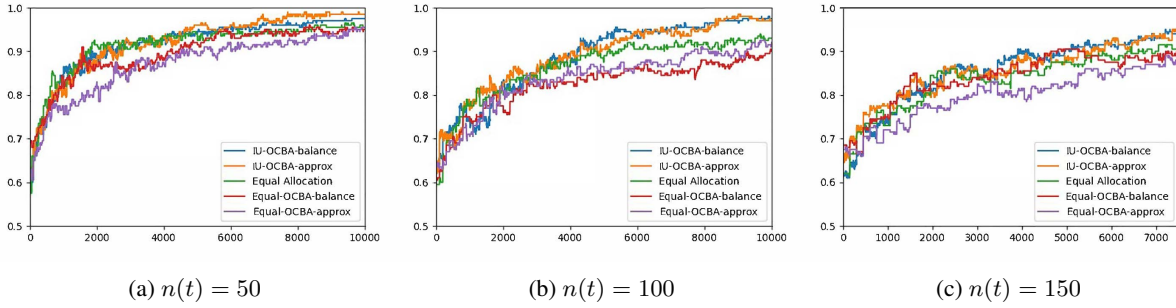


Figure 7: Portfolio example with constant batch size and different choices of $n(t)$

We have the following observations from Figure 6-9.

1. In scenarios with constant batch sizes, even with non-Gaussian simulation error, the two proposed procedures still achieve the best two performance among all. This demonstrates the practical value of the two procedures since they can be applied to problems with general simulation error and still gain robust performance. However, they do not outperform the other three procedures as much as in the case of Gaussian simulation error, since the optimality conditions (8) - (10) do not hold for non-Gaussian noise.
2. Comparing IU-OCBA-balance and IU-OCBA-approx, IU-OCBA-balance gradually catches up or even outperforms IU-OCBA-approx as $n(0)$ becomes larger, which is seen in the quadratic example as well. This phenomenon is more evident in the portfolio optimization problem, since the true variance here is larger and leads to a larger estimation error.

For the scenarios with random batch size, we obtain similar results as for those with constant batch size and our proposed procedures still achieve the highest empirical PCS. The numerical results are shown in Figure 8 and 9.

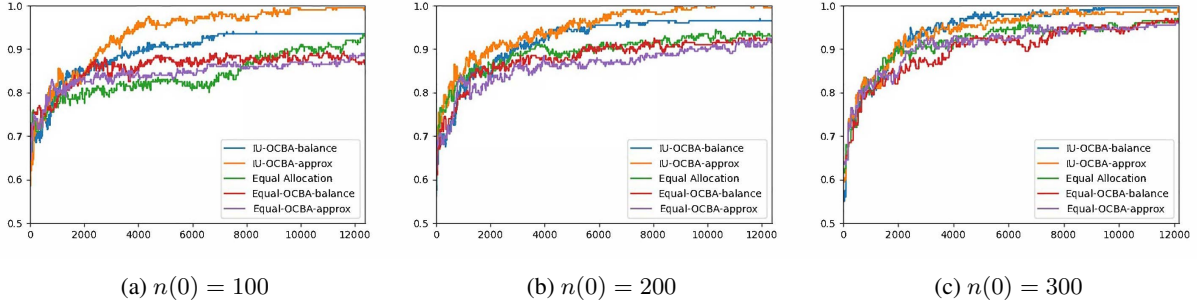


Figure 8: Portfolio example with random batch size: different choices of $n(0)$

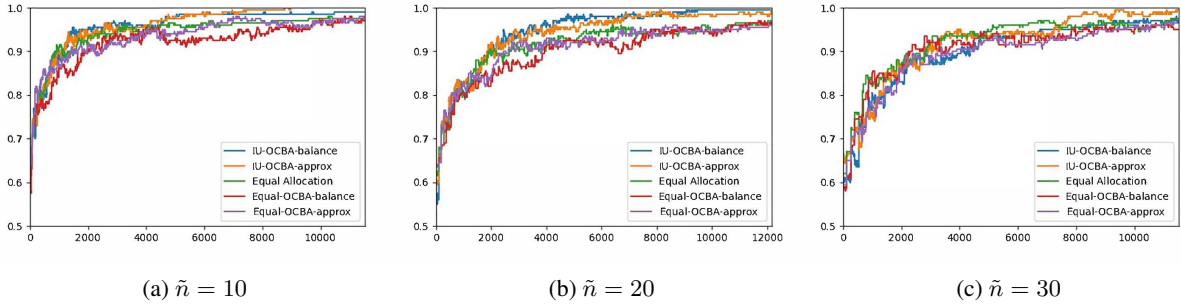


Figure 9: Portfolio example with random batch size and different choices of \tilde{n}

7 Conclusion

In this paper we consider a fixed budget ranking and selection (R&S) problem, where the common input distribution across designs is unknown but can be estimated with streaming input data that come sequentially in time. As the input distribution estimate is updated with the new batch of data at each time stage, the stage-wise simulation budget needs to be allocated to carry out new simulations for evaluating design performances. We apply the large deviations theory to obtain the optimal stage-wise budget allocation policy for design-input pairs. Then, combined with the updating of the input distribution estimate, we design two fully sequential procedures for the streaming data setting. Both procedures are shown to have consistency (i.e., select the best design with probability 1 as times goes to infinity) and asymptotic optimality (i.e., converge to the optimal budget allocation policy under the true input distribution). Our numerical experiments demonstrate that the proposed procedures have a great advantage over the equal allocation rule and two extensions of OCBA, when dealing with unknown input distributions with streaming input data.

ACKNOWLEDGMENT

The authors gratefully acknowledge the support by the Air Force Office of Scientific Research under Grants FA9550-19-1-0283 and FA9550-22-1-0244 and the National Science Foundation under Grant NSF-DMS2053489.

References

- [1] Russell R Barton, Barry L Nelson, and Wei Xie. Quantifying input uncertainty via simulation confidence intervals. *INFORMS Journal on Computing*, 26(1):74–87, 2013.
- [2] Russell R Barton, Barry L Nelson, and Wei Xie. Quantifying input uncertainty via simulation confidence intervals. *INFORMS journal on computing*, 26(1):74–87, 2014.
- [3] Sait Cakmak, Siyang Gao, and Enlu Zhou. Contextual ranking and selection with gaussian processes. *arXiv preprint arXiv:2201.07782*, 2022.
- [4] Chun-Hung Chen, Donghai He, Michael Fu, and Loo Hay Lee. Efficient simulation budget allocation for selecting an optimal subset. *INFORMS Journal on Computing*, 20(4):579–595, 2008.

- [5] Chun-Hung Chen, Jianwu Lin, Enver Yücesan, and Stephen E Chick. Simulation budget allocation for further enhancing the efficiency of ordinal optimization. Discrete Event Dynamic Systems, 10(3):251–270, 2000.
- [6] Ye Chen and Ilya O Ryzhov. Complete expected improvement converges to an optimal budget allocation. Advances in Applied Probability, 51(1):209–235, 2019.
- [7] Stephen E Chick. Input distribution selection for simulation experiments: accounting for input uncertainty. Operations Research, 49(5):744–758, 2001.
- [8] Stephen E Chick and Peter Frazier. Sequential sampling with economics of selection procedures. Management Science, 58(3):550–569, 2012.
- [9] Stephen E Chick and Koichiro Inoue. New two-stage and sequential procedures for selecting the best simulated system. Operations Research, 49(5):732–743, 2001.
- [10] Canan G Corlu, Alp Akcay, and Wei Xie. Stochastic simulation under input uncertainty: A review. Operations Research Perspectives, page 100162, 2020.
- [11] Canan G Corlu and Bahar Biller. A subset selection procedure under input parameter uncertainty. In Proceedings of the 2013 Winter Simulation Conference, pages 463–473, 2013.
- [12] Canan G Corlu and Bahar Biller. Subset selection for simulations accounting for input uncertainty. In Proceedings of the 2015 Winter Simulation Conference, pages 437–446, 2015.
- [13] Amir Dembo, Ofer Zeitouni, and IH Dinwoodie. Large deviations techniques and applications. SIAM Review, 36(2):303–303, 1994.
- [14] Jianzhong Du, Siyang Gao, and Chun-Hung Chen. Rate-optimal contextual ranking and selection. arXiv preprint arXiv:2206.12640, 2022.
- [15] Weiwei Fan, L Jeff Hong, and Xiaowei Zhang. Distributionally robust selection of the best. Management Science, 66(1):190–208, 2020.
- [16] Peter Frazier, Warren Powell, and Savas Dayanik. The knowledge-gradient policy for correlated normal beliefs. INFORMS Journal on Computing, 21(4):599–613, 2009.
- [17] Peter I Frazier. A fully sequential elimination procedure for indifference-zone ranking and selection with tight bounds on probability of correct selection. Operations Research, 62(4):926–942, 2014.
- [18] Siyang Gao and Weiwei Chen. Efficient subset selection for the expected opportunity cost. Automatica, 59:19–26, 2015.
- [19] Siyang Gao, Weiwei Chen, and Leyuan Shi. A new budget allocation framework for the expected opportunity cost. Operations Research, 65(3):787–803, 2017.
- [20] Siyang Gao, Jianzhong Du, and Chun-Hung Chen. Selecting the optimal system design under covariates. In 2019 IEEE 15th International Conference on Automation Science and Engineering (CASE), 2019.
- [21] Siyang Gao, Hui Xiao, Enlu Zhou, and Weiwei Chen. Robust ranking and selection with optimal computing budget allocation. Automatica, 81:30–36, 2017.
- [22] Peter Glynn and Sandeep Juneja. A large deviations perspective on ordinal optimization. In Proceedings of the 2004 Winter Simulation Conference, pages 489–494, 2004.
- [23] L Jeff Hong, Weiwei Fan, and Jun Luo. Review on ranking and selection: A new perspective. Frontiers of Engineering Management, 8(3):321–343, 2021.
- [24] Susan R Hunter and Raghu Pasupathy. Optimal sampling laws for stochastically constrained simulation optimization on finite sets. INFORMS Journal on Computing, 25(3):527–542, 2013.
- [25] L Jeff Hong. Fully sequential indifference-zone selection procedures with variance-dependent sampling. Naval Research Logistics (NRL), 53(5):464–476, 2006.
- [26] Qing-Shan Jia, Enlu Zhou, and Chun-Hung Chen. Efficient computing budget allocation for finding simplest good designs. IIE Transactions, 45(7):736–750, 2013.
- [27] Xiao Jin, Haobin Li, and Loo Hay Lee. Optimal budget allocation in simulation analytics. In 2019 IEEE 15th International Conference on Automation Science and Engineering (CASE), 2019.
- [28] Seong-Hee Kim and Barry L Nelson. A fully sequential procedure for indifference-zone selection in simulation. ACM Transactions on Modeling and Computer Simulation (TOMACS), 11(3):251–273, 2001.
- [29] Seong-Hee Kim and Barry L Nelson. Recent advances in ranking and selection. In Proceedings of the 2007 Winter Simulation Conference, pages 162–172, 2007.

- [30] Henry Lam and Enlu Zhou. The empirical likelihood approach to quantifying uncertainty in sample average approximation. *Operations Research Letters*, 45(4):301–307, 2017.
- [31] Loo Hay Lee, Ek Peng Chew, Suyan Teng, and David Goldsman. Optimal computing budget allocation for multi-objective simulation models. In *Proceedings of the 2004 Winter Simulation Conference*, pages 594–602, 2004.
- [32] Yanwen Li and Siyang Gao. On the convergence of optimal computing budget allocation algorithms. In *Proceedings of the 2021 Winter Simulation Conference*, pages 1–12, 2021.
- [33] Y Lin, Eunhye Song, and Barry L Nelson. Single-experiment input uncertainty. *Journal of Simulation*, 9(3):249–259, 2015.
- [34] Sijia Ma and Shane G Henderson. An efficient fully sequential selection procedure guaranteeing probably approximately correct selection. In *Proceedings of the 2017 Winter Simulation Conference*, pages 2225–2236, 2017.
- [35] Raghu Pasupathy, Susan R Hunter, Nugroho A Pujowidianto, Loo Hay Lee, and Chun-Hung Chen. Stochastically constrained ranking and selection via score. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 25(1):1–26, 2014.
- [36] Eunhye Song and Barry L Nelson. Input–output uncertainty comparisons for discrete optimization via simulation. *Operations Research*, 67(2):562–576, 2019.
- [37] Yuhao Wang and Enlu Zhou. Fixed budget ranking and selection with streaming input data. In *Accepted by the 2022 Winter Simulation Conference*, 2022.
- [38] Di Wu, Yuhao Wang, and Enlu Zhou. Data-driven ranking and selection under input uncertainty. *Accepted by Operations Research*, 2022.
- [39] Di Wu and Enlu Zhou. Ranking and selection under input uncertainty: a budget allocation formulation. In *Proceedings of the 2017 Winter Simulation Conference*, pages 2245–2256, 2017.
- [40] Hui Xiao, Fei Gao, and Loo Hay Lee. Optimal computing budget allocation for complete ranking with input uncertainty. *IIE Transactions*, 52(5):489–499, 2020.
- [41] Hui Xiao and Siyang Gao. Simulation budget allocation for selecting the top- m designs with input uncertainty. *IEEE Transactions on Automatic Control*, 63(9):3127–3134, 2018.
- [42] Wei Xie, Barry L Nelson, and Russell R Barton. A bayesian framework for quantifying uncertainty in stochastic simulation. *Operations Research*, 62(6):1439–1452, 2014.
- [43] Wei Xie, Barry L Nelson, and Russell R Barton. Multivariate input uncertainty in output analysis for stochastic simulation. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 27(1):1–22, 2016.
- [44] Jingxu Xu, Zeyu Zheng, and Peter W Glynn. Joint resource allocation for input data collection and simulation. In *2020 Winter Simulation Conference (WSC)*, 2020.
- [45] Shen Yan, Enlu Zhou, and Chun-Hung Chen. Efficient selection of a set of good enough designs with complexity preference. *Automation Science and Engineering, IEEE Transactions on*, 9:596–606, 07 2012.
- [46] Faker Zouaoui and James R Wilson. Accounting for parameter uncertainty in simulation input modeling. *Iie Transactions*, 35(9):781–792, 2003.
- [47] Faker Zouaoui and James R Wilson. Accounting for input-model and input-parameter uncertainties in simulation. *IIE Transactions*, 36(11):1135–1151, 2004.

Technical Proof

Proof of Lemma 1 Lemma 1.1 is easily seen from (4). To prove Lemma 1.2, it suffices to show the concavity of the function for $x > 0$ with form $f(x) = 1/(\sum_{i=1}^n \frac{a_i}{x_i})$, where $a_i > 0$ for $i = 1, 2, \dots, n$. We prove the concavity of the multivariate function by proving the concavity along all lines. For any $y \in \mathbb{R}^n$, let $g(t) = f(x + ty)$ where $t \in \mathbb{R}$ such that $x + ty > 0$. We have

$$g''(t) = \frac{2}{(\sum_{i=1}^n \frac{a_i}{x_i + ty_i})^3} \left\{ \left[\sum_{i=1}^n \frac{a_i y_i}{(x_i + ty_i)^2} \right]^2 - \sum_{i=1}^n \frac{a_i y_i^2}{(x_i + ty_i)^3} \sum_{i=1}^n f \frac{a_i}{x_i + ty_i} \right\} \leq 0,$$

where the inequality uses the Cauchy inequality. Hence, f is concave in $x > 0$. ■

Proof of Theorem 3.2 Let t_l be the time stage s.t. $\sum_{t=1}^{t_l} n(t) < l \leq \sum_{t=1}^{t_l+1} n(t)$. Denote by $p_j^l = p_j^{(t_l)}$. Let $\mathbf{1}_{(i,j)}^{(l)}$ be the indicator function that $\mathbf{1}_{(i,j)}^{(l)} = 1$ if (i, j) is sampled at iteration l . Several lemmas are needed to complete the proof. The first Lemma 2 ensures the convergence rate of estimated input distribution.

Lemma 2 $|p_j^l - p_j| = O(\sqrt{\frac{\log \log l}{l}})$ almost surely .

Proof By assumption 3, the input data $\{\xi_s\}_{s=1}^\infty$ are i.i.d., which means $\{\mathbf{1}\{\xi_s = \zeta_j\}\}_{s=1}^\infty$ are also i.i.d.. Hence, we have $\limsup_{t \rightarrow \infty} \frac{|p_j^{(t)} - p_j|}{\sqrt{\frac{\log \log M(t)}{M(t)}}} = \limsup_{t \rightarrow \infty} \frac{|\sum_{s=1}^M(t) \mathbf{1}\{\xi_s = \zeta_j\} - p_j|}{\sqrt{M(t) \log \log M(t)}} = p_j(1 - p_j)$ by law of the iterated logarithm (LIL).

Hence, $|p_j^{(t)} - p_j| = O(\sqrt{\frac{\log \log M(t)}{M(t)}})$. By Assumption 5, $\lim_{t \rightarrow \infty} \frac{M(t)}{t} = \bar{m}$ almost surely .

$$|p_j^l - p_j| = |p_j^{(t_l)} - p_j| = O(\sqrt{\frac{\log \log M(t_l)}{M(t_l)}}) = O(\sqrt{\frac{\log \log t_l}{t_l}}) \text{ almost surely .}$$

Further since

$$\lim_{l \rightarrow \infty} \frac{1}{t_l} \sum_{t=1}^{t_l} n(t) < \lim_{l \rightarrow \infty} \frac{l}{t_l} \leq \lim_{l \rightarrow \infty} \frac{1}{t_l} \sum_{t=1}^{t_l+1} n(t) \text{ almost surely ,}$$

by Assumption 5 we get $\lim_{l \rightarrow \infty} \frac{l}{t_l} = \bar{n}$ almost surely . Hence $|p_j^l - p_j| = O(\sqrt{\frac{\log \log t_l}{t_l}}) = O(\sqrt{\frac{\log \log l}{l}})$ almost surely . \blacksquare

Lemma 3-6 ensure that a positive ratio of budget will be assigned to all design-input pairs.

Lemma 3

$\liminf_{l \rightarrow \infty} \frac{\alpha_{i,j}^{(l)}}{\alpha_{i,j'}^{(l)}} > 0, \forall i, \forall j \neq j'$ almost surely.

Proof Prove by contradiction. Suppose $\liminf_{l \rightarrow \infty} \frac{\alpha_{i,j}^{(l)}}{\alpha_{i,j'}^{(l)}} = 0$, then there exists a subsequence $\{l_k\}_{k=1}^\infty$ such that

$\frac{\alpha_{i,j}^{(l_k)}}{\alpha_{i,j'}^{(l_k)}} \rightarrow 0$. Further more since $\frac{\alpha_{i,j}^{(l)}}{\alpha_{i,j'}^{(l)}}$ increases only if (i, j) is sampled at l and decreases only if (i, j') is sampled at l .

The subsequence $\{l_k\}$ can be chosen such that (i, j') is sampled at $l_k \forall k$. Then since $p_j^{(l)} \rightarrow p_j$ and $\hat{\sigma}_{i,j}^{(l)} \rightarrow \sigma_{i,j}^{(l)}, \forall i, j$, we have

$$\frac{\frac{\alpha_{i,j}^{(l_k)}}{\hat{\sigma}_{i,j}^{(l_k)} p_j^{(l_k)}}}{\frac{\alpha_{i,j'}^{(l_k)}}{\hat{\sigma}_{i,j'}^{(l_k)} p_{j'}^{(l_k)}}} = \frac{\alpha_{i,j}^{(l_k)} \hat{\sigma}_{i,j'}^{(l_k)} p_{j'}^{(l_k)}}{\alpha_{i,j'}^{(l_k)} \hat{\sigma}_{i,j}^{(l_k)} p_j^{(l_k)}} \rightarrow 0 \text{ as } l \rightarrow \infty.$$

Hence for k large enough, (i, j') cannot be sampled at l_k , a contradiction. \blacksquare

Lemma 4

$\liminf_{l \rightarrow \infty} \frac{\alpha_{i,j}^{(l)}}{\alpha_{i',j}^{(l)}} > 0, \forall i \neq i' \neq b, \forall j$ almost surely.

Proof Prove by contradiction. Suppose there exists j_0 such that $\liminf_{l \rightarrow \infty} \frac{\alpha_{i,j}^{(l)}}{\alpha_{i',j}^{(l)}} = 0$. By similar arguments as in the

proof of Lemma 3, for any positive constant $\varepsilon > 0$, we can find a sufficiently large l such that (i', j_0) is sampled at l and $\frac{\alpha_{i,j_0}^{(l)}}{\alpha_{i',j_0}^{(l)}} \leq \varepsilon$. By Lemma 2 and Theorem 3.1, we have $p_j^l \rightarrow p_j$ and $N_{i,j}^{(l)} \rightarrow \infty$ almost surely $\forall i, j$. Hence, $\bar{\mu}_b^{(l)}, \bar{\mu}_i^{(l)}$

and $\bar{\mu}_{i'}^{(l)}$ all will converge to the true value a.s. and $\hat{b}^{(l)} = b$ for l sufficiently large. Then there exists constants a, e and $U > L > 0$, such that for l sufficiently large, $0 < a < p_j^l < e, 0 < r \leq \hat{\sigma}_{i,j}^{(l)} \leq h$ and $U > (\bar{\mu}_b^{(l)} - \bar{\mu}_i^{(l)})^2$ and

$L < (\bar{\mu}_b^{(l)} - \bar{\mu}_{i'}^{(l)})^2$. Then

$$\begin{aligned}
& \frac{(\hat{\mu}_b^{(l)} - \hat{\mu}_i^{(l)})^2}{\sum_{j=1}^D \frac{(\hat{\sigma}_{b,j}^{(l)})^2 (p_j^l)^2}{\alpha_{b,j}^{(l)}} + \sum_{j=1}^D \frac{(\hat{\sigma}_{i,j}^{(l)})^2 (p_j^l)^2}{\alpha_{i,j}^{(l)}}} - \frac{(\hat{\mu}_b^{(l)} - \hat{\mu}_{i'}^{(l)})^2}{\sum_{j=1}^D \frac{(\hat{\sigma}_{b,j}^{(l)})^2 (p_j^l)^2}{\alpha_{b,j}^{(l)}} + \sum_{j=1}^D \frac{(\hat{\sigma}_{i',j}^{(l)})^2 (p_j^l)^2}{\alpha_{i',j}^{(l)}}} \\
& < \frac{D}{\sum_{j=1}^D \frac{(\hat{\sigma}_{b,j}^{(l)})^2 (p_j^l)^2}{\alpha_{b,j}^{(l)}} + \sum_{j=1}^D \frac{(\hat{\sigma}_{i,j}^{(l)})^2 (p_j^l)^2}{\alpha_{i,j}^{(l)}}} - \frac{D}{\sum_{j=1}^D \frac{(\hat{\sigma}_{b,j}^{(l)})^2 (p_j^l)^2}{\alpha_{b,j}^{(l)}} + \sum_{j=1}^D \frac{(\hat{\sigma}_{i',j}^{(l)})^2 (p_j^l)^2}{\alpha_{i',j}^{(l)}}} \\
& = \frac{\sum_{j=1}^D \frac{(\hat{\sigma}_{b,j}^{(l)})^2 (p_j^l)^2}{\alpha_{b,j}^{(l)}} (U - L) + U \sum_{j=1}^D \frac{(\hat{\sigma}_{i',j}^{(l)})^2 (p_j^l)^2}{\alpha_{i',j}^{(l)}} - L \sum_{j=1}^D \frac{(\hat{\sigma}_{i,j}^{(l)})^2 (p_j^l)^2}{\alpha_{i,j}^{(l)}}}{\left(\sum_{j=1}^D \frac{(\hat{\sigma}_{b,j}^{(l)})^2 (p_j^l)^2}{\alpha_{b,j}^{(l)}} + \sum_{j=1}^D \frac{(\hat{\sigma}_{i,j}^{(l)})^2 (p_j^l)^2}{\alpha_{i,j}^{(l)}} \right) \left(\sum_{j=1}^D \frac{(\hat{\sigma}_{b,j}^{(l)})^2 (p_j^l)^2}{\alpha_{b,j}^{(l)}} + \sum_{j=1}^D \frac{(\hat{\sigma}_{i',j}^{(l)})^2 (p_j^l)^2}{\alpha_{i',j}^{(l)}} \right)} \quad (22)
\end{aligned}$$

By Lemma 3, there exist c_1, c_2 such that $\alpha_{i,j}^{(l)} \leq \frac{1}{c_1} \alpha_{i,j_0}^{(l)}$ and $\alpha_{i',j}^{(l)} \geq \frac{1}{c_2} \alpha_{i',j_0}^{(l)}$. Then, the numerator of (22) is upper bounded by

$$\begin{aligned}
& \sum_{j=1}^D \frac{(\hat{\sigma}_{b,j}^{(l)})^2 e^2}{\alpha_{b,j}^{(l)}} (U - L) + \frac{U}{\alpha_{i',j_0}^{(l)}} c_2 \sum_{j=1}^D h^2 e^2 - \frac{L}{\alpha_{i,j_0}^{(l)}} c_1 \sum_{j=1}^D r^2 a^2 \\
& \leq \sum_{j=1}^D \frac{\hat{\sigma}_{b,j}^{(l)} \hat{\sigma}_{i',j}^{(l)} e^2}{\alpha_{i',j}^{(l)}} (U - L) + \frac{U}{\alpha_{i',j_0}^{(l)}} c_2 \sum_{j=1}^D h^2 e^2 - \frac{L}{\alpha_{i,j_0}^{(l)}} c_1 \sum_{j=1}^D r^2 a^2 \quad (23)
\end{aligned}$$

$$\begin{aligned}
& \leq \frac{(U - L)}{\alpha_{i',j_0}^{(l)}} c_2 D h^2 e^2 + \frac{U}{\alpha_{i',j_0}^{(l)}} c_2 D^2 h^2 e^2 - \frac{L}{\varepsilon \alpha_{i',j_0}^{(l)}} c_1 B r^2 a^2 \\
& = \frac{1}{\alpha_{i',j_0}^{(l)}} \left\{ D \left[c_2 h^2 e^2 (U - L) + c_2 U h^2 e^2 - \frac{L}{\varepsilon} c_1 B r^2 a^2 \right] \right\} \quad (24)
\end{aligned}$$

(23) holds because a non-best design is sampled at l which implies $\left(\frac{\alpha_{b,j}^{(l)}}{\sigma_{b,j}^{(l)}} \right)^2 \geq \sum_i \left(\frac{\alpha_{i,j}^{(l)}}{\sigma_{i,j}^{(l)}} \right)^2 \geq \left(\frac{\alpha_{i',j}^{(l)}}{\sigma_{i',j}^{(l)}} \right)^2 \forall j$. Then if

we choose ε that makes (24) < 0 . We obtain
$$\frac{(\hat{\mu}_b^{(l)} - \hat{\mu}_i^{(l)})^2}{\sum_{j=1}^D \frac{(\hat{\sigma}_{b,j}^{(l)})^2 (p_j^l)^2}{\alpha_{b,j}^{(l)}} + \sum_{j=1}^D \frac{(\hat{\sigma}_{i,j}^{(l)})^2 (p_j^l)^2}{\alpha_{i,j}^{(l)}}} - \frac{(\hat{\mu}_b^{(l)} - \hat{\mu}_{i'}^{(l)})^2}{\sum_{j=1}^D \frac{(\hat{\sigma}_{b,j}^{(l)})^2 (p_j^l)^2}{\alpha_{b,j}^{(l)}} + \sum_{j=1}^D \frac{(\hat{\sigma}_{i',j}^{(l)})^2 (p_j^l)^2}{\alpha_{i',j}^{(l)}}} <$$

0, which implies (i', j_0) cannot be sampled at l , a contradiction.

Lemma 5 (i) $\liminf_{l \rightarrow \infty} \frac{\alpha_{b,j}^{(l)}}{\alpha_{i,j}^{(l)}} > 0 \forall i \neq b, \forall j$ almost surely ; (ii) $\liminf_{l \rightarrow \infty} \frac{\alpha_{i,j}^{(l)}}{\alpha_{b,j}^{(l)}} > 0, \forall i \neq b, \forall j$ almost surely.

proof of (i) Prove by contradiction. By Lemma 4, there exists a positive constant $c > 0$ such that $\frac{\alpha_{k,j}^{(l)}}{\alpha_{i,j}^{(l)}} \geq c$ for all $k \neq b$

and l sufficiently large. Since $\liminf_{l \rightarrow \infty} \frac{\alpha_{b,j}^{(l)}}{\alpha_{i,j}^{(l)}} = 0$, there exists a sufficiently large l such that (i, j) is sampled at l and

$\frac{1}{2} \sigma_{i,j} \leq \hat{\sigma}_{i,j}^{(l)} \leq 2 \sigma_{i,j} \forall i, j$ and $\frac{\alpha_{b,j}^{(l)}}{\alpha_{i,j}^{(l)}} \leq d$, where d satisfies $\frac{d^2}{\sigma_{b,j}^2} < \frac{1}{16} c \left(\sum_{k \neq b} \frac{1}{\sigma_{k,j}^2} \right)$. Then we have

$$\frac{\left(\frac{\alpha_{b,j}^{(l)}}{\hat{\sigma}_{b,j}^{(l)}} \right)^2}{\sum_{k \neq b} \left(\frac{\alpha_{k,j}^{(l)}}{\hat{\sigma}_{k,j}^{(l)}} \right)^2} \leq \frac{\left(\frac{\alpha_{b,j}^{(l)}}{\hat{\sigma}_{b,j}^{(l)}} \right)^2}{c \left(\alpha_{i,j}^{(l)} \right)^2 \sum_{k \neq b} \left(\frac{1}{\hat{\sigma}_{i,j}^{(l)}} \right)^2} \leq \frac{d^2}{(\sigma_{b,j})^2} \frac{16}{c \sum_{k \neq b} \frac{1}{(\sigma_{k,j})^2}} < 1.$$

Hence we have (i, j) cannot be sampled at l , a contradiction.

Proof of (ii) (ii) can be proved in a similar way. ■

Lemma 6 (i) $\liminf_{l \rightarrow \infty} \frac{\alpha_{i_1, j_1}^{(l)}}{\alpha_{i_2, j_2}^{(l)}} > 0 \forall i_1, i_2, j_1, j_2$ almost surely ; (ii) $\liminf_{l \rightarrow \infty} \alpha_{i,j}^{(l)} > 0, \forall i, j$ almost surely.

Proof of (i) By Lemma 3, there exists $c_1 > 0$ such that $\alpha_{i_1, j_1}^{(l)} \geq c_1 \alpha_{i_1, j_2}^{(l)}$. If $b \notin \{i_1, i_2\}$, then by Lemma 4 there exists $c_2 > 0$ such that $\alpha_{i_1, j_2}^{(l)} \geq c_2 \alpha_{i_2, j_2}^{(l)}$. Else if $b \in \{i_1, i_2\}$, by Lemma 5 there exists $c_3 > 0$ such that $\alpha_{i_1, j_2}^{(l)} \geq c_3 \alpha_{i_2, j_2}^{(l)}$. In both cases there exists $c_4 > 0$ such that $\frac{\alpha_{i_1, j_1}^{(l)}}{\alpha_{i_2, j_2}^{(l)}} \geq c_4 > 0$, which proves (i).

Proof of (ii) This is a direct result of (i). By (i), for any fixed (i, j) there exists $c > 0$ such that $\alpha_{i, j}^{(l)} > c \alpha_{i', j'}^{(l)}, \forall i', j'$. Hence $\alpha_{i, j}^{(l)} = \frac{\alpha_{i, j}^{(l)}}{\sum_{i', j'} \alpha_{i', j'}^{(l)}} > \frac{c}{KB} > 0$. ■

Lemma 7 and 8 guarantee the convergence rate of the estimated expected performance and estimated variance, respectively.

Lemma 7 $|\hat{\mu}_i^{(l)} - \bar{\mu}_i| = O(\sqrt{\frac{\log \log l}{l}})$ almost surely.

Proof Notice that $\hat{\mu}_i^{(l)} - \bar{\mu}_i = \sum_{j=1}^D (p_j^l \hat{\mu}_{i, j}^{(l)} - p_j \mu_{i, j})$. It is sufficient to show

$$|p_j^l \hat{\mu}_{i, j}^{(l)} - p_j \mu_{i, j}| = O(\sqrt{\frac{\log \log l}{l}}), \quad j = 1, 2, \dots, D.$$

By Lemma 6, there exists $C_{i, j} > 0$, such that $N_{i, j}^{(l)} \geq C_{i, j} l$. By LIL, $|\hat{\mu}_{i, j}^{(l)} - \mu_{i, j}| = O(\sqrt{\frac{\log \log N_{i, j}^{(l)}}{N_{i, j}^{(l)}}}) = O(\sqrt{\frac{\log \log l}{l}})$. Then by Lemma 2,

$$|p_j^l \hat{\mu}_{i, j}^{(l)} - p_j \mu_{i, j}| = |p_j^l (\hat{\mu}_{i, j}^{(l)} - \mu_{i, j}) + (p_j^l - p_j) \mu_{i, j}| \leq p_j^l |\hat{\mu}_{i, j}^{(l)} - \mu_{i, j}| + |p_j^l - p_j| |\mu_{i, j}| = O(\sqrt{\frac{\log \log l}{l}}). \quad \blacksquare$$

Lemma 8 $|(\hat{\sigma}_{i, j}^{(l)})^2 - (\sigma_{i, j})^2| = O(\sqrt{\frac{\log \log l}{l}})$ almost surely. As a result, $|\hat{\sigma}_{i, j}^{(l)} - \sigma_{i, j}| = O(\sqrt{\frac{\log \log l}{l}})$ almost surely.

Proof Since

$$\sum_{s=1}^{N_{i, j}^{(l)}} (X_{i, j}^{(s)} - \hat{\mu}_{i, j}^{(l)})^2 = \sum_{s=1}^{N_{i, j}^{(l)}} (X_{i, j}^{(s)} - \mu_{i, j})^2 - N_{i, j}^{(l)} (\hat{\mu}_{i, j}^{(l)} - \mu_{i, j})^2,$$

We have

$$(\hat{\sigma}_{i, j}^{(l)})^2 - (\sigma_{i, j})^2 = \frac{1}{N_{i, j}^{(l)} - 1} \sum_{s=1}^{N_{i, j}^{(l)}} [(X_{i, j}^{(s)} - \mu_{i, j})^2 - \sigma_{i, j}^2] - \frac{N_{i, j}^{(l)}}{N_{i, j}^{(l)} - 1} (\hat{\mu}_{i, j}^{(l)} - \mu_{i, j})^2 + \frac{\sigma_{i, j}^2}{N_{i, j}^{(l)} - 1}.$$

Since $(X_{i, j}^{(s)} - \mu_{i, j})^2 - \sigma_{i, j}^2$ are i.i.d. with mean 0, by LIL, we have with probability 1,

$$\left| \frac{1}{N_{i, j}^{(l)} - 1} \sum_{s=1}^{N_{i, j}^{(l)}} [(X_{i, j}^{(s)} - \mu_{i, j})^2 - \sigma_{i, j}^2] \right| = O\left(\sqrt{\frac{\log \log N_{i, j}^{(l)}}{N_{i, j}^{(l)}}}\right) = O\left(\sqrt{\frac{\log \log l}{l}}\right),$$

where the second equality comes from Lemma 6. Further since $|\hat{\mu}_{i, j}^{(l)} - \mu_{i, j}| = O(\sqrt{\frac{\log \log l}{l}})$ and $\left| \frac{\sigma_{i, j}^2}{N_{i, j}^{(l)} - 1} \right| = O\left(\sqrt{\frac{\log \log l}{l}}\right)$, we get the desired result. ■

Lemma 9 is a simple but useful result which we will use frequently in the following proof.

Lemma 9 Let (i, j) be a fixed design-input pair. Suppose (i, j) is sampled at iteration r . Let $t_r = \inf\{l > 0 : \mathbf{1}_{(i, j)}^{(r+l)} = 1\}$. Hence $r + t_r$ is the next iteration (i, j) will be sampled after r . Then we have $r < r + t_r = O(r)$ almost surely.

Proof Prove by contradiction. Suppose $\forall C_0 > 0$, there exists an iteration r such that $t_r > C_0 r$. We have

$$\alpha_{i,j}^{(r+t_r)} = \frac{N_{i,j}^{(r+t_r)}}{N^{(r+t_r)}} = \frac{N_{i,j}^{(r)} + 1}{N^{(r)} + t_r} < \frac{2(n_0 + r)}{KBn_0 + (C_0 + 1)r} < \frac{3}{C_0}$$

for large r . The first inequality holds since $N(r) = KBn_0 + r$ and $N_{i,j}^{(0)} = n_0$. By the arbitrariness of C_0 and the fact that if $C_0 \rightarrow \infty$, the iteration r that satisfy $t_r > C_0 r$ must also go to ∞ . We have $\liminf_{l \rightarrow \infty} \alpha_{i,j}^{(l)} = 0$, contradicting Lemma 6.(ii). ■

The following Lemma 10 bounds the amount of budget allocated to design-input pair (i, j) between two successive samples of another design-input pair (i, j') with the same design but different input realizations.

Lemma 10 *Let i denote any fixed design and j a fixed input realization. Suppose (i, j) is sampled at iteration r and $t_r := \inf\{l > 0 : \mathbf{1}_{(i,j)}^{(r+l)} = 1\}$. $r + t_r$ is the next iteration where (i, j) will be sampled again. Then between the two samples of (i, j) , the number of samples that can be allocated to (i, j') , $j' \neq j$ is $O(\sqrt{r \log \log r})$ almost surely.*

Proof For any $j' \neq j$, let $s_r = \sup\{l < t_r : \mathbf{1}_{(i,j')}^{(r+l)} = 1\}$. $r + s_r$ is the last time before $r + t_r$ that (i, j') is sampled. If $s_r < 0$, then the lemma holds true, otherwise we assume $s_r > 0$. Then at iteration $r + s_r$, we have $\exists C_1, C_2, C_3$ independent of r ,

$$\begin{aligned} \frac{N_{i,j'}^{(r+s_r)}}{\sigma_{i,j'}^{(r+s_r)} p_{j'}^{(r+s_r)}} &\leq \frac{N_{i,j}^{(r+s_r)}}{\sigma_{i,j}^{(r+s_r)} p_j^{r+s_r}} \\ &= \frac{N_{i,j}^{(r)} + 1}{\sigma_{i,j}^{(r)} p_j^{r+s_r}} \\ &\leq \frac{N_{i,j}^{(r)}}{\sigma_{i,j}^{(r)} p_j^{r+s_r}} + C_1 \\ &\leq \frac{N_{i,j}^{(r)}}{\sigma_{i,j}^{(r)} p_j^r} + C_2 \sqrt{r \log \log r} \end{aligned} \tag{25}$$

$$\leq \frac{N_{i,j'}^{(r)}}{\sigma_{i,j'}^{(r)} p_{j'}^r} + C_2 \sqrt{r \log \log r} \tag{26}$$

$$\leq \frac{N_{i,j'}^{(r)}}{\sigma_{i,j'}^{(s_r+r)} p_{j'}^{s_r+r}} + C_3 \sqrt{r \log \log r} \tag{27}$$

where (25) and (27) are by Lemma 2 and Lemma 8 and (26) is because (i, j) is sampled at r . Hence, we obtain $N_{i,j'}^{(r+t_r)} - 1 - N_{i,j'}^{(r)} = N_{i,j'}^{(r+s_r)} - N_{i,j'}^{(r)} \leq C_3 \sigma_{i,j'}^{(r+s_r)} p_{j'}^{r+s_r} \sqrt{r \log \log r} \leq C_4 \sqrt{r \log \log r}$ for some C_4 independent of r . The proof is complete. ■

The following Lemma 11 proved the ‘‘Derivative Balance’’ optimality condition.

Lemma 11 $\left| \frac{N_{i,j}^{(r)}}{\sigma_{i,j}^{(r)} p_j^r} - \frac{N_{i,j'}^{(r)}}{\sigma_{i,j'}^{(r)} p_{j'}^r} \right| = O(\sqrt{r \log \log r})$, $\forall i, \forall j \neq j'$ almost surely.

Proof Without loss of generality, suppose $\frac{N_{i,j}^{(r)}}{\sigma_{i,j}^{(r)} p_j^r} \leq \frac{N_{i,j'}^{(r)}}{\sigma_{i,j'}^{(r)} p_{j'}^r}$. Let $u := \inf\{l > r : \mathbf{1}_{(i,j')}^{(r+l)} = 1\}$ be the next iteration at which (i, j') is sampled. Then by Lemma 10 we have $N_{i,j}^{(u)} - N_{i,j}^{(r)} = O(\sqrt{r \log \log r})$. Then there exists $C_1, C_2, C_2', C_3 > 0$ such that

$$\frac{N_{i,j}^{(u)}}{\sigma_{i,j}^{(u)} p_j^u} \leq \frac{N_{i,j}^{(r)} + C_1 \sqrt{r \log \log r}}{(\sigma_{i,j}^{(r)} - C_2' \sqrt{\frac{\log \log r}{r}})(p_j^r - C_2 \sqrt{\frac{\log \log r}{r}})} \leq \frac{N_{i,j}^{(r)}}{\sigma_{i,j}^{(r)} p_j^r} + C_3 \sqrt{r \log \log r}.$$

Furthermore, since $N_{i,j'}^{(u)} = N_{i,j'}^{(r)} + 1$, $\exists C_4, C_5 > 0$, such that

$$\frac{N_{i,j'}^{(u)}}{\sigma_{i,j'}^{(u)} p_j^u} = \frac{N_{i,j'}^{(r)} + 1}{\sigma_{i,j'}^{(u)} p_j^u} \leq \frac{N_{i,j'}^{(r)}}{\sigma_{i,j'}^{(u)} p_j^u} + C_4 \leq \frac{N_{i,j'}^{(r)}}{\sigma_{i,j'}^{(r)} p_j^r} + C_5 \sqrt{r \log \log r}.$$

Then,

$$\begin{aligned} 0 &\leq \frac{N_{i,j}^{(r)}}{\sigma_{i,j}^{(r)} p_j^r} - \frac{N_{i,j'}^{(r)}}{\sigma_{i,j'}^{(r)} p_{j'}^r} \\ &\leq \frac{N_{i,j}^{(u)}}{\sigma_{i,j}^{(u)} p_j^u} - \frac{N_{i,j'}^{(u)}}{\sigma_{i,j'}^{(u)} p_{j'}^u} + (C_3 + C_5) \sqrt{r \log \log r} \\ &\leq (C_3 + C_5) \sqrt{r \log \log r}. \end{aligned}$$

This implies $\left| \frac{N_{i,j}^{(r)}}{\sigma_{i,j}^{(r)} p_j^r} - \frac{N_{i,j'}^{(r)}}{\sigma_{i,j'}^{(r)} p_{j'}^r} \right| = O(\sqrt{r \log \log r})$.

Proof of (19) in Theorem 3 Proof The result follows from Lemma 11. Since $N^{(l)} = l + KBn_0$. We have almost surely,

$$\begin{aligned} &\left| \frac{\alpha_{i,j}^{(l)}}{\sigma_{i,j} p_j} - \frac{\alpha_{i,j'}^{(l)}}{\sigma_{i,j'} p_{j'}} \right| \\ &= \left| \frac{\alpha_{i,j}^{(l)}}{\hat{\sigma}_{i,j}^{(l)} p_j^l} - \frac{\alpha_{i,j'}^{(l)}}{\hat{\sigma}_{i,j'}^{(l)} p_{j'}^l} \right| + O\left(\sqrt{\frac{\log \log l}{l}}\right) \\ &= \left| \frac{N_{i,j}^{(l)}/N^{(l)}}{\hat{\sigma}_{i,j}^{(l)} p_j^l} - \frac{N_{i,j'}^{(l)}/N^{(l)}}{\hat{\sigma}_{i,j'}^{(l)} p_{j'}^l} \right| + O\left(\sqrt{\frac{\log \log l}{l}}\right) \\ &= O\left(\sqrt{\frac{\log \log l}{l}}\right) \rightarrow 0 \quad \text{as } l \rightarrow \infty. \end{aligned}$$

Hence, we obtain $\left| \frac{\alpha_{i,j}^{(l)}}{\sigma_{i,j} p_j} - \frac{\alpha_{i,j'}^{(l)}}{\sigma_{i,j'} p_{j'}} \right| \rightarrow 0$. ■

The following Lemma 10 bounds the amount of budget allocated to a non-optimal design-input pair (i, j) between two successive samples of the best design-input pair (b, j) under the same input realization.

Lemma 12 *Under a fixed input realization j , Suppose (b, j) is sampled at iteration r . Let $t_r = \inf\{l > 0 : \mathbf{1}_{(b,j)}^{(r+l)} = 1\}$. $r + t_r$ is the next iteration at which (b, j) is sampled. Then between the two samples of (b, j) , the number of samples that can be allocated to (i, j) , $i \neq b$ is at most $O(\sqrt{r \log \log r})$ almost surely.*

Proof Fix a non-optimal design i . Let $s_r = \sup\{l < t_r : \mathbf{1}_{(i,j)}^{(r+l)} = 1\}$. $r + s_r$ is the last time before $r + t_r$ at which (i, j) is sampled. If $s_r < 0$, then the lemma holds true, otherwise assume $s_r > 0$. Since (b, j) is sampled at r , there exists j_0 such that $\left(\frac{N_{b,j_0}^{(r)}}{\sigma_{b,j_0}^{(r)} p_{j_0}^r}\right)^2 \leq \sum_{k \neq b} \left(\frac{N_{k,j_0}^{(r)}}{\sigma_{k,j_0}^{(r)} p_{j_0}^r}\right)^2$. By Lemma 11, there exists $C_1 > 0$ such that $\left| \frac{N_{i,j}^{(r)}}{\sigma_{i,j}^{(r)} p_j^r} - \frac{N_{i,j'}^{(r)}}{\sigma_{i,j'}^{(r)} p_{j'}^r} \right| \leq C_1 \sqrt{r \log \log r}$, $\forall i$. Then

$$\left| \left(\frac{N_{i,j}^{(r)}}{\sigma_{i,j}^{(r)} p_j^r}\right)^2 - \left(\frac{N_{i,j'}^{(r)}}{\sigma_{i,j'}^{(r)} p_{j'}^r}\right)^2 \right| = \left| \frac{N_{i,j}^{(r)}}{\sigma_{i,j}^{(r)} p_j^r} - \frac{N_{i,j'}^{(r)}}{\sigma_{i,j'}^{(r)} p_{j'}^r} \right| \left(\frac{N_{i,j}^{(r)}}{\sigma_{i,j}^{(r)} p_j^r} + \frac{N_{i,j'}^{(r)}}{\sigma_{i,j'}^{(r)} p_{j'}^r} \right) \leq C_1 C_2 r \sqrt{r \log \log r}. \quad (28)$$

The last inequality holds for some $C_2 > 0$ since $|\frac{N_{i,j}^{(r)}}{\sigma_{i,j}^{(r)} p_j^r}| = O(r)$, $\forall i, j$. Hence, there exist $C_3, C_4 > 0$ independent of r ,

$$\begin{aligned}
0 &\leq \left(\frac{N_{b,j}^{(r+s_r)}}{\sigma_{b,j}^{(r+s_r)} p_j^{r+s_r}} \right)^2 - \sum_{k \neq b} \left(\frac{N_{k,j}^{(r+s_r)}}{\sigma_{k,j}^{(r+s_r)} p_j^{r+s_r}} \right)^2 \\
&= \left(\frac{N_{b,j}^{(r)} + 1}{\sigma_{b,j}^{(r+s_r)} p_j^{r+s_r}} \right)^2 - \sum_{k \neq b} \left(\frac{N_{k,j}^{(r+s_r)}}{\sigma_{k,j}^{(r+s_r)} p_j^{r+s_r}} \right)^2 \\
&\leq \left(\frac{N_{b,j}^{(r)}}{\sigma_{b,j}^{(r+s_r)} p_j^{r+s_r}} \right)^2 - \sum_{k \neq b} \left(\frac{N_{k,j}^{(r+s_r)}}{\sigma_{k,j}^{(r+s_r)} p_j^{r+s_r}} \right)^2 + C_3 r \\
&\leq \left(\frac{N_{b,j_0}^{(r)}}{\sigma_{b,j_0}^{(r+s_r)} p_{j_0}^{r+s_r}} \right)^2 - \sum_{k \neq b} \left(\frac{N_{k,j}^{(r+s_r)}}{\sigma_{k,j}^{(r+s_r)} p_j^{r+s_r}} \right)^2 + C_3 r + C_1 C_2 r \sqrt{r \log \log r} \tag{29}
\end{aligned}$$

$$\begin{aligned}
&\leq \sum_{k \neq b} \left(\frac{N_{k,j_0}^{(r)}}{\sigma_{k,j_0}^{(r)} p_{j_0}^r} \right)^2 - \sum_{k \neq b} \left(\frac{N_{k,j}^{(r+s_r)}}{\sigma_{k,j}^{(r+s_r)} p_j^{r+s_r}} \right)^2 + C_3 r + C_1 C_2 r \sqrt{r \log \log r} \\
&\leq \sum_{k \neq b} \left(\frac{N_{k,j}^{(r)}}{\sigma_{k,j}^{(r)} p_j^r} \right)^2 - \sum_{k \neq b} \left(\frac{N_{k,j}^{(r+s_r)}}{\sigma_{k,j}^{(r+s_r)} p_j^{r+s_r}} \right)^2 + C_3 r + K C_1 C_2 r \sqrt{r \log \log r} \tag{30}
\end{aligned}$$

$$\leq \sum_{k \neq b} \left(\frac{N_{k,j}^{(r)}}{\sigma_{k,j}^{(r)} p_j^r} \right)^2 - \sum_{k \neq b} \left(\frac{N_{k,j}^{(r+s_r)}}{\sigma_{k,j}^{(r)} p_j^r} \right)^2 + C_4 r \sqrt{r \log \log r} \tag{31}$$

(29) and (30) hold by (28). (31) holds for some $C_4 > C_3 + K C_1 C_2$. Then, since for each $k \neq b$, $\left(\frac{N_{k,j}^{(r+s_r)}}{\sigma_{k,j}^{(r)} p_j^r} \right)^2 - \left(\frac{N_{k,j}^{(r)}}{\sigma_{k,j}^{(r)} p_j^r} \right)^2 \geq 0$. We obtain

$$\left(\frac{N_{i,j}^{(r+s_r)}}{\sigma_{i,j}^{(r+s_r)} p_j^{r+s_r}} \right)^2 - \left(\frac{N_{i,j}^{(r)}}{\sigma_{i,j}^{(r)} p_j^r} \right)^2 \leq C_4 r \sqrt{r \log \log r}.$$

Since there exists $C_5 > C_4 \sigma_{k,j}^{(r)} p_j^r \forall k, r$, we obtain

$$C_5 r \sqrt{r \log \log r} \geq (N_{i,j}^{(r+s_r)})^2 - (N_{i,j}^{(r)})^2 = (N_{i,j}^{(r+s_r)} - N_{i,j}^{(r)})(N_{i,j}^{(r+s_r)} + N_{i,j}^{(r)}) \geq 2N_{i,j}^{(r)}(N_{i,j}^{(r+s_r)} - N_{i,j}^{(r)}) \tag{32}$$

By Lemma 6.(ii), there exists $C_6 > 0$, such that $\alpha_{i,j}^{(r)} > C_6$ for r sufficiently large. This implies $N_{i,j}^{(r)} > C_6 r$. Then (32) $\geq 2C_6 r(N_{i,j}^{(r+s_r)} - N_{i,j}^{(r)})$, which implies $N_{i,j}^{(r+s_r)} - N_{i,j}^{(r)} \leq \frac{C_5}{2C_6} \sqrt{r \log \log r}$. Hence $N_{i,j}^{(r+t_r)} - N_{i,j}^{(r)} = O(\sqrt{r \log \log r})$. ■

Conversely, the following Lemma 10 bounds the amount of budget allocated to the optimal design-input pair (b, j) between two successive samples of any two non-optimal design-input pair under the same input realization.

Lemma 13 *Under a fixed input realization $j = 1, 2, \dots, D$, Suppose at iteration r a non-optimal design i_1 is sampled. Let $t_r = \inf\{l > 0 : \exists i \neq b, \mathbf{1}_{(i,j)}^{(r+l)} = 1\}$. $r + t_r$ is the next iteration at which a non-optimal design is sampled. Then between iteration r and $r + t_r$, the number of samples that can be allocated to (b, j) is $O(\sqrt{r \log \log r})$ almost surely.*

Proof Define $s_r = \sup\{l < t_r : \mathbf{1}_{(b,j)} = 1\}$. $r + s_r$ is the last time before $r + t_r$ the optimal design is sampled. If $s_r < 0$, then the lemma holds. Otherwise assume $s_r > 0$. Since (b, j) is sampled at $r + s_r$, there exists j_0 such that

$$0 \geq \left(\frac{N_{b,j_0}^{(r+s_r)}}{\sigma_{b,j_0}^{(r+s_r)} p_{j_0}^{r+s_r}} \right)^2 - \sum_{k \neq b} \left(\frac{N_{k,j_0}^{(r+s_r)}}{\sigma_{k,j_0}^{(r+s_r)} p_{j_0}^{r+s_r}} \right)^2. \text{ Further by Lemma 9, we have}$$

$$s_r < t_r = O(r). \tag{33}$$

Then, there exists $C_1, C_2, C_3 > 0$, such that

$$\begin{aligned} 0 &\geq \left(\frac{N_{b,j_0}^{(r+s_r)}}{\sigma_{b,j_0}^{(r+s_r)} p_{j_0}^{r+s_r}} \right)^2 - \sum_{k \neq b} \left(\frac{N_{k,j_0}^{(r+s_r)}}{\sigma_{k,j_0}^{(r+s_r)} p_{j_0}^{r+s_r}} \right)^2 \\ &\geq \left(\frac{N_{b,j}^{(r+s_r)}}{\sigma_{b,j}^{(r+s_r)} p_j^{r+s_r}} \right)^2 - \sum_{k \neq b} \left(\frac{N_{k,j}^{(r+s_r)}}{\sigma_{k,j}^{(r+s_r)} p_j^{r+s_r}} \right)^2 - C_1 r \sqrt{r \log \log r} \end{aligned} \quad (34)$$

$$\begin{aligned} &= \left(\frac{N_{b,j}^{(r+s_r)}}{\sigma_{b,j}^{(r+s_r)} p_j^{r+s_r}} \right)^2 - \sum_{k \neq i_1 \neq b} \left(\frac{N_{k,j}^{(r)}}{\sigma_{k,j}^{(r+s_r)} p_j^{r+s_r}} \right)^2 - \left(\frac{N_{i_1,j}^{(r)} + 1}{\sigma_{i_1,j}^{(r+s_r)} p_{j_0}^{r+s_r}} \right)^2 - C_1 r \sqrt{r \log \log r} \\ &\geq \left(\frac{N_{b,j}^{(r+s_r)}}{\sigma_{b,j}^{(r+s_r)} p_j^{r+s_r}} \right)^2 - \sum_{k \neq i_1 \neq b} \left(\frac{N_{k,j}^{(r)}}{\sigma_{k,j}^{(r+s_r)} p_j^{r+s_r}} \right)^2 - \left(\frac{N_{i_1,j}^{(r)}}{\sigma_{i_1,j}^{(r+s_r)} p_j^{r+s_r}} \right)^2 - C_2 r \sqrt{r \log \log r} \\ &\geq \left(\frac{N_{b,j}^{(r+s_r)}}{\sigma_{b,j}^{(r+s_r)} p_j^r} \right)^2 - \sum_{k \neq b} \left(\frac{N_{k,j}^{(r)}}{\sigma_{k,j}^{(r)} p_j^r} \right)^2 - C_3 r \sqrt{r \log \log r} \end{aligned} \quad (35)$$

$$\geq \left(\frac{N_{b,j}^{(r+s_r)}}{\sigma_{b,j}^{(r)} p_j^r} \right)^2 - \left(\frac{N_{b,j}^{(r)}}{\sigma_{b,j}^{(r)} p_j^r} \right)^2 - C_3 r \sqrt{r \log \log r} \quad (36)$$

(34) holds because of Lemma 11 and (33); (35) holds due to Lemma 2, Lemma 8 and (33); (36) is because at iteration r (i_1, j) is sampled. Hence, there exists $C_4 > 0$, such that

$$0 \leq \left(N_{b,j}^{(r+s_r)} \right)^2 - \left(N_{b,j}^{(r)} \right)^2 = \left(N_{b,j}^{(r+s_r)} - N_{b,j}^{(r)} \right) \left(N_{b,j}^{(r+s_r)} + N_{b,j}^{(r)} \right) \leq C_3 \sigma_{b,j}^{(r)} p_j^r r \sqrt{r \log \log r} \leq C_4 r \sqrt{r \log \log r}.$$

By Lemma 6, there exists $C_5 > 0$, such that $N_{b,j}^{(r)} \geq C_5 r$. Then

$$N_{b,j}^{(r+s_r)} - N_{b,j}^{(r)} \leq \frac{C_4 r \sqrt{r \log \log r}}{\left(N_{b,j}^{(r+s_r)} + N_{b,j}^{(r)} \right)} \leq \frac{C_4 \sqrt{r \log \log r}}{2C_5}.$$

The proof is complete. ■

We prove the ‘‘Total Balance’’ condition in the following Lemma 14.

Lemma 14 $\left| \left(\frac{\alpha_{1,j}^{(r)}}{\sigma_{1,j}^{(r)}} \right)^2 - \sum_{i \neq b} \left(\frac{\alpha_{i,j}^{(r)}}{\sigma_{i,j}^{(r)}} \right)^2 \right| = O\left(\sqrt{\frac{\log \log r}{r}}\right) \forall j \quad a.s..$

Proof Let $\{l_k\}_{k=1}^\infty$ denote the subsequence where input realization ζ_j is sampled. We first show $|\Delta_j^{(k)}| :=$

$$\left| \left(\frac{\alpha_{b,j}^{(l_k)}}{\sigma_{b,j}^{(l_k)}} \right)^2 - \sum_{i \neq b} \left(\frac{\alpha_{i,j}^{(l_k)}}{\sigma_{i,j}^{(l_k)}} \right)^2 \right| = O\left(\sqrt{\frac{\log \log l_k}{l_k}}\right). \text{ Fix an iteration } k_0.$$

Case 1. $\Delta_j^{(k_0)} < 0$, let $s = \sup\{k < k_0 : \exists i \neq b, \mathbf{1}_{(i,j)}^{(k)} = 1\}$. Then l_s the last time before l_k that a non-optimal design, denote by i_0 , is sampled under input realization ζ_j . Then, there exists $C_1 > 0$,

$$\begin{aligned} 0 &> \left(\frac{N_{b,j}^{(l_{k_0})}}{\sigma_{b,j}^{(l_{k_0})}} \right)^2 - \sum_{i \neq b} \left(\frac{N_{i,j}^{(l_{k_0})}}{\sigma_{i,j}^{(l_{k_0})}} \right)^2 \\ &\geq \left(\frac{N_{b,j}^{(l_s)}}{\sigma_{b,j}^{(l_{k_0})}} \right)^2 - \sum_{i \neq i_0 \neq b} \left(\frac{N_{i,j}^{(l_s)}}{\sigma_{i,j}^{(l_{k_0})}} \right)^2 - \left(\frac{N_{i_0,j}^{(l_s)} + 1}{\sigma_{i_0,j}^{(l_{k_0})}} \right)^2 \\ &\geq \left(\frac{N_{b,j}^{(l_s)}}{\sigma_{b,j}^{(l_{k_0})}} \right)^2 - \sum_{i \neq i_0} \left(\frac{N_{i,j}^{(l_s)}}{\sigma_{i,j}^{(l_{k_0})}} \right)^2 - C_1 l_s \end{aligned}$$

Divide both sides by $(N^{(l_s)})^2$ and notice that by Lemma 9 there exists $C_2 > 0$ independent of k_0, s , such that $l_{k_0} < C_2 l_s$. We obtain

$$\begin{aligned} 0 &> \left(\frac{\alpha_{b,j}^{(l_s)}}{\sigma_{b,j}^{(l_{k_0})}} \right)^2 - \sum_{i \neq i_0} \left(\frac{\alpha_{i,j}^{(l_s)}}{\sigma_{i,j}^{(l_{k_0})}} \right)^2 - \frac{C_1}{l_s} \\ &\geq \left(\frac{\alpha_{b,j}^{(l_s)}}{\sigma_{b,j}^{(l_{k_0})}} \right)^2 - \sum_{i \neq i_0} \left(\frac{\alpha_{i,j}^{(l_s)}}{\sigma_{i,j}^{(l_{k_0})}} \right)^2 - \frac{C_1 C_2}{l_{k_0}} \\ &\geq \left(\frac{\alpha_{b,j}^{(l_s)}}{\sigma_{b,j}^{(l_s)}} \right)^2 - \sum_{i \neq i_0} \left(\frac{\alpha_{i,j}^{(l_s)}}{\sigma_{i,j}^{(l_s)}} \right)^2 - C_3 \sqrt{\frac{\log \log l_{k_0}}{l_{k_0}}} \end{aligned}$$

for some C_3 by Lemma 8.

Case 2. $\Delta_j^{(k_0)} \geq 0$. Let $s' = \sup\{k < k_0 : \mathbf{1}_{(b,j)}^{(k)} = 1\}$. $l_{s'}$ is the last time before l_{k_0} at which (b, j) is sampled. Then there exists j' such that $\Delta_{j'}^{(l_s)} \leq 0$. Furthermore by Lemma 9, $l_{k_0} = O(l_s)$ since (b, j) is not sampled between l_s and l_{k_0} ; by Lemma 12, $N_{i,j}^{(l_k)} - N_{i,j}^{(l_s)} = O(\sqrt{l_s \log \log l_s})$, $\forall i \neq b$. Then, there exists C_3, C_4 and $C_5 > 0$,

$$\begin{aligned} 0 &\leq \left(\frac{N_{b,j}^{(l_{k_0})}}{\sigma_{b,j}^{(l_{k_0})}} \right)^2 - \sum_{i \neq i_0} \left(\frac{N_{i,j}^{(l_{k_0})}}{\sigma_{i,j}^{(l_{k_0})}} \right)^2 \\ &\leq \left(\frac{N_{b,j}^{(l_s)} + 1}{\sigma_{b,j}^{(l_{k_0})}} \right)^2 - \sum_{i \neq i_0} \left(\frac{N_{i,j}^{(l_s)}}{\sigma_{i,j}^{(l_{k_0})}} \right)^2 + C_3 l_{k_0} \sqrt{l_{k_0} \log \log l_{k_0}} \\ &\leq \left(\frac{N_{b,j}^{(l_s)}}{\sigma_{b,j}^{(l_s)}} \right)^2 - \sum_{i \neq i_0} \left(\frac{N_{i,j}^{(l_s)}}{\sigma_{i,j}^{(l_s)}} \right)^2 + C_4 l_{k_0} \sqrt{l_{k_0} \log \log l_{k_0}} \\ &\leq \left(\frac{N_{b,j'}^{(l_s)}}{\sigma_{b,j'}^{(l_s)}} \right)^2 - \sum_{i \neq i_0} \left(\frac{N_{i,j'}^{(l_s)}}{\sigma_{i,j'}^{(l_s)}} \right)^2 + C_5 l_{k_0} \sqrt{l_{k_0} \log \log l_{k_0}} \\ &\leq C_5 l_{k_0} \sqrt{l_{k_0} \log \log l_{k_0}} \end{aligned} \tag{37}$$

(37) holds because of Lemma 11. Divide both sides by $(N^{(l_k)})^2$, we obtain $0 \leq \Delta_j^{(k_0)} \leq C_5 \sqrt{\frac{\log \log l_{k_0}}{l_{k_0}}}$.

We now show for $\forall r$, the lemma holds. Let $l_k \leq r < l_{k+1}$. Then

$$\begin{aligned} &\left| \left(\frac{\alpha_{b,j}^{(r)}}{\sigma_{b,j}^{(r)}} \right)^2 - \sum_{i \neq b} \left(\frac{\alpha_{i,j}^{(r)}}{\sigma_{i,j}^{(r)}} \right)^2 \right| \\ &= \left(\frac{N^{(l_k)}}{N^{(r)}} \right)^2 \left| \left(\frac{\alpha_{b,j}^{(l_k)}}{\hat{\sigma}_{b,j}^{(l_k)}} \right)^2 - \sum_{i \neq b} \left(\frac{\alpha_{i,j}^{(l_k)}}{\hat{\sigma}_{i,j}^{(l_k)}} \right)^2 + O\left(\sqrt{\frac{\log \log l_k}{l_k}} \right) \right| \\ &= \left(\frac{N^{(l_k)}}{N^{(r)}} \right)^2 O\left(\sqrt{\frac{\log \log l_k}{l_k}} \right) \\ &= O\left(\sqrt{\frac{\log \log r}{r}} \right), \end{aligned}$$

where the last equality holds since $r = O(l_k)$ by Lemma 9. ■

Proof of (20) in Theorem 3

By Lemma 8 and Lemma 14,

$$\begin{aligned} \left| \left(\frac{\alpha_{1,j}^{(r)}}{\sigma_{1,j}} \right)^2 - \sum_{i \neq b} \left(\frac{\alpha_{i,j}^{(r)}}{\sigma_{i,j}} \right)^2 \right| &= \left| \left(\frac{\alpha_{1,j}^{(r)}}{\sigma_{1,j}^{(r)}} \right)^2 - \sum_{i \neq b} \left(\frac{\alpha_{i,j}^{(r)}}{\sigma_{i,j}^{(r)}} \right)^2 \right| + O\left(\sqrt{\frac{\log \log r}{r}}\right) \\ &= O\left(\sqrt{\frac{\log \log r}{r}}\right) \end{aligned}$$

■

The next Lemma 15 is a little technical, which is used to bound the amount of budget allocated to a non-optimal design-input pair (b, j) between two successive samples of a non-optimal design-input pair (i, j) under the same input realization, as shown in Lemma 16.

Lemma 15 *Under a fixed input realization j , suppose a non-optimal design (k, j) is sampled at iteration r . Define*

$$\begin{cases} t_r := \inf_l \{l > 0 : \mathbf{1}_{(k,j)}^{(r+l)} = 1\} \\ s'_r := \sup_l \{l < t_r : \mathbf{1}_{(b,j)}^{(r+l)} = 1\} \\ s_r := \sup_l \{l < s'_r : \mathbf{1}_{(i,j)}^{(r+l)} = 1 \text{ for some } i \neq b\} \\ d_{i,j}^{(r,q)} = N_{i,j}^{(r+q)} - N_{i,j}^{(r)} \end{cases}$$

For all $C_1 > 0$, if there exists C_2 sufficiently large (depend on C_1 but not on r), such that $C_2\sqrt{r \log \log r} \leq d_{b,j}^{(r,s_r)}$ holds for infinitely many r 's, then for such sufficiently large r , there exists another sub-optimal design $i \neq k \neq b$ and $u \leq s_r$, i is sampled at $r + u$ and

$$\left(1 + C_1\sqrt{\frac{\log \log r}{r}}\right) \frac{N_{i,j}^{(r)}}{N_{b,j}^{(l)}} \leq \frac{N_{i,j}^{(r+u)}}{N_{b,j}^{(r+s_r)}} \leq \frac{N_{i,j}^{(r+u)}}{N_{b,j}^{(r+u)}} \quad (38)$$

holds almost surely.

Proof By Lemma 9, $t_r = O(r)$, which implies there exists $C_0 > 0$, $d_{b,j}^{(r,s_r)} \leq C_0 r$. Hence, for any fixed C_1 , there exists C_2 such that for infinitely many r 's $C_2\sqrt{r \log \log r} \leq d_{b,j}^{(r,s_r)} \leq C_0 r$. Let $\Delta_j^{(r)} = \left(\frac{\alpha_{b,j}^{(r)}}{\sigma_{b,j}^{(r)}}\right)^2 - \sum_{i \neq b} \left(\frac{\alpha_{i,j}^{(r)}}{\sigma_{i,j}^{(r)}}\right)^2$. By the definition of s_r , $\Delta_j^{(r+s_r)} \geq 0$ and $\Delta_j^{(r+s_r+1)} = \Delta_j^{(r+s'_r)} + O\left(\sqrt{\frac{\log \log(r+s_r)}{r+s_r}}\right)$. Since (b, j) is sampled at $r + s'_r$, there exists j' such that $\Delta_{j'}^{(r+s'_r)} < 0$. By lemma 11, there exists $C_3, C'_3, C''_3 > 0$,

$$\Delta_j^{(r+s_r+1)} \leq \Delta_j^{(r+s'_r)} + C''_3\sqrt{\frac{\log \log(r+s_r)}{r+s_r}} \leq \Delta_{j'}^{(r+s'_r)} + C'_3\sqrt{\frac{\log \log(r+s'_r)}{r+s'_r}} \leq C'_3\sqrt{\frac{\log \log(r+s'_r)}{r+s'_r}} \leq C_3\sqrt{\frac{\log \log r}{r}}.$$

The last inequality holds since $s'_r \leq t_r = O(r)$. Then one can choose $C_4 > C_3$,

$$0 \leq \Delta_j^{(r+s_r)} \leq C_4\sqrt{\frac{\log \log r}{r}} \quad (39)$$

Since $N_{k,j}^{(r+s_r)} - N_{k,j}^{(r)} = 1$, (39) can be expressed as

$$\left(\frac{N_{b,j}^{(r+s_r)}/\sigma_{b,j}^{(r+s_r)}}{N^{(r+s_r)}}\right)^2 - \sum_{i \neq b \neq k} \left(\frac{N_{i,j}^{(r+s_r)}/\sigma_{i,j}^{(r+s_r)}}{N^{(r+s_r)}}\right)^2 - \left(\frac{(N_{k,j}^{(r)} + 1)/\sigma_{k,j}^{(r+s_r)}}{N^{(r+s_r)}}\right)^2 \leq C_4\sqrt{\frac{\log \log r}{r}}.$$

By some simple algebraic calculation we get

$$\sum_{i \neq b \neq k} \left(\frac{N_{i,j}^{(r+s_r)}/\sigma_{i,j}^{(r+s_r)}}{N_{b,j}^{(r+s_r)}/\sigma_{b,j}^{(r+s_r)}}\right)^2 + \left(\frac{(N_{k,j}^{(r)} + 1)/\sigma_{k,j}^{(r+s_r)}}{N_{b,j}^{(r+s_r)}/\sigma_{b,j}^{(r+s_r)}}\right)^2 + C_4\sqrt{\frac{\log \log r}{r}} \left(\frac{N^{(r+s_r)}}{N_{b,j}^{(r+s_r)}/\sigma_{b,j}^{(r+s_r)}}\right)^2 \geq 1.$$

By Lemma 5, there exists $C_5 > 0$, $C_5 N_{b,j}^{(r)}/\sigma_{b,j}^{(r)} > N^{(r)}$ for all large r . Hence,

$$\sum_{i \neq b \neq k} \left(\frac{N_{i,j}^{(r+s_r)}/\sigma_{i,j}^{(r+s_r)}}{N_{b,j}^{(r+s_r)}/\sigma_{b,j}^{(r+s_r)}} \right)^2 \geq 1 - \left(\frac{(N_{k,j}^{(r+1)})/\sigma_{k,j}^{(r+s_r)}}{N_{b,j}^{(r+s_r)}/\sigma_{b,j}^{(r+s_r)}} \right)^2 - C_4 C_5^2 \sqrt{\frac{\log \log r}{r}}$$

Furthermore, there exists $C_6, C_7 > 0$, such that

$$\begin{aligned} & \sum_{i \neq b \neq k} \left(\frac{N_{i,j}^{(r+s_r)}/\sigma_{i,j}^{(r+s_r)}}{N_{b,j}^{(r+s_r)}/\sigma_{b,j}^{(r+s_r)}} \right)^2 - \sum_{i \neq b \neq k} \left(\frac{N_{i,j}^{(r)}/\sigma_{i,j}^{(r+s_r)}}{N_{b,j}^{(r)}/\sigma_{b,j}^{(r+s_r)}} \right)^2 \\ & \geq 1 - \sum_{i \neq b \neq k} \left(\frac{N_{i,j}^{(r)}/\sigma_{i,j}^{(r+s_r)}}{N_{b,j}^{(r)}/\sigma_{b,j}^{(r+s_r)}} \right)^2 - \left(\frac{(N_{k,j}^{(r)} + 1)/\sigma_{k,j}^{(r+s_r)}}{N_{b,j}^{(r+s_r)}/\sigma_{b,j}^{(r+s_r)}} \right)^2 - C_4 C_5^2 \sqrt{\frac{\log \log r}{r}} \\ & \geq 1 - \sum_{i \neq b \neq k} \left(\frac{N_{i,j}^{(r)}/\sigma_{i,j}^{(r)}}{N_{b,j}^{(r)}/\sigma_{b,j}^{(r)}} \right)^2 - \left(\frac{(N_{k,j}^{(r)} + 1)/\sigma_{k,j}^{(r+s_r)}}{N_{b,j}^{(r+s_r)}/\sigma_{b,j}^{(r+s_r)}} \right)^2 - (C_4 C_5^2 + C_6) \sqrt{\frac{\log \log r}{r}} \\ & \geq \left(\frac{N_{k,j}^{(r)}/\sigma_{k,j}^{(r)}}{N_{b,j}^{(r)}/\sigma_{b,j}^{(r)}} \right)^2 - \left(\frac{(N_{k,j}^{(r)} + 1)/\sigma_{k,j}^{(r+s_r)}}{N_{b,j}^{(r+s_r)}/\sigma_{b,j}^{(r+s_r)}} \right)^2 - (C_4 C_5^2 + C_6) \sqrt{\frac{\log \log r}{r}} \quad (40) \\ & \geq \left(\frac{N_{k,j}^{(r)}/\sigma_{k,j}^{(r+s_r)}}{N_{b,j}^{(r)}/\sigma_{b,j}^{(r+s_r)}} \right)^2 - \left(\frac{(N_{k,j}^{(r)} + 1)/\sigma_{k,j}^{(r+s_r)}}{N_{b,j}^{(r+s_r)}/\sigma_{b,j}^{(r+s_r)}} \right)^2 - (C_4 C_5^2 + C_7) \sqrt{\frac{\log \log r}{r}} \\ & = \left(\frac{\sigma_{b,j}^{(r+s_r)}}{\sigma_{k,j}^{(r+s_r)}} \right)^2 \left[\frac{(N_{k,j}^{(r)})^2 (N_{b,j}^{(r+s_r)})^2 - (N_{k,j}^{(r)} + 1)^2 (N_{b,j}^{(r)})^2}{(N_{b,j}^{(r)})^2 (N_{b,j}^{(r+s_r)})^2} \right] - (C_4 C_5^2 + C_7) \sqrt{\frac{\log \log r}{r}} \\ & = \left(\frac{\sigma_{b,j}^{(r+s_r)}}{\sigma_{k,j}^{(r+s_r)}} \right)^2 \left[\frac{(N_{k,j}^{(r)})^2 (N_{b,j}^{(r)} + d_{b,j}^{(r,s_r)})^2 - (N_{k,j}^{(r)} + 1)^2 (N_{b,j}^{(r)})^2}{(N_{b,j}^{(r)})^2 (N_{b,j}^{(r+s_r)})^2} \right] - (C_4 C_5^2 + C_7) \sqrt{\frac{\log \log r}{r}} \\ & = \left(\frac{\sigma_{b,j}^{(r+s_r)}}{\sigma_{k,j}^{(r+s_r)}} \right)^2 \left[\frac{(2N_{b,j}^{(r)} N_{k,j}^{(r)} + N_{k,j}^{(r)} d_{b,j}^{(r,s_r)} + N_{b,j}^{(r)}) (N_{k,j}^{(r)} d_{b,j}^{(r,s_r)} - N_{b,j}^{(r)})}{(N_{b,j}^{(r)})^2 (N_{b,j}^{(r+s_r)})^2} \right] - (C_4 C_5^2 + C_7) \sqrt{\frac{\log \log r}{r}}. \quad (41) \end{aligned}$$

(40) holds since (k, j) is sampled at r . By Lemma 5 and $s_r = O(r)$, there exists $C_8, C_9, C_{10} > 0$, $\frac{N_{k,j}^{(r)}}{N_{b,j}^{(r)}} \geq C_8$, $N_{b,j}^{(r)} \geq C_9 r$ and $N_{b,j}^{(r+s_r)} \leq C_{10} r$. We have (41) is lower bounded by

$$\begin{aligned} & \left(\frac{\sigma_{b,j}^{(r+s_r)}}{\sigma_{k,j}^{(r+s_r)}} \right)^2 \left[\frac{C_8^2 (2C_9 r + d_{b,j}^{(r,s_r)} + 1/C_8) (d_{b,j}^{(r,s_r)} - 1/C_8)}{C_{10}^2 r^2} \right] - (C_4 C_5^2 + C_7) \sqrt{\frac{\log \log r}{r}} \\ & = \left(\frac{\sigma_{b,j}^{(r+s_r)}}{\sigma_{k,j}^{(r+s_r)}} \right)^2 \left(\frac{C_8}{C_{10}} \right)^2 \left[\frac{2C_9 d_{b,j}^{(r,s_r)}}{r} + \left(\frac{d_{b,j}^{(r,s_r)}}{r} \right)^2 - \frac{2C_9}{C_8 r} - \frac{1}{C_8^2 r^2} \right] - (C_4 C_5^2 + C_7) \sqrt{\frac{\log \log r}{r}} \\ & \geq \left(\frac{\sigma_{b,j}^{(r+s_r)}}{\sigma_{k,j}^{(r+s_r)}} \right)^2 \left(\frac{C_8}{C_{10}} \right)^2 \left[\frac{d_{b,j}^{(r,s_r)}}{r} - \frac{1}{C_8 r} - \frac{1}{2C_9 C_8^2 r^2} \right] - (C_4 C_5^2 + C_7) \sqrt{\frac{\log \log r}{r}} \quad (42) \end{aligned}$$

There exists $C_{11} > 0$, $\frac{1}{C_8 r} + \frac{1}{2C_9 C_8^2 r^2} \leq C_{11} \sqrt{\frac{\log \log r}{r}}$. Choose $0 < C_{12} \leq \left(\frac{\sigma_{b,j}^{(r+s_r)}}{\sigma_{k,j}^{(r+s_r)}}\right)^2 \left(\frac{C_8}{C_{10}}\right)^2 2C_9$ for all large r . Furthermore since $d_{b,j}^{(r,s_r)} \geq C_2 \sqrt{r \log \log r}$, there exists $C_{13} > 0$

$$\begin{aligned}
(42) &\geq C_{12} \left[(C_2 - C_{11}) \sqrt{\frac{\log \log r}{r}} \right] - (C_4 C_5^2 + C_7) \sqrt{\frac{\log \log r}{r}} \\
&= [C_{12}(C_2 - C_{11}) - (C_4 C_5^2 + C_7)] \sqrt{\frac{\log \log r}{r}} \\
&\geq C_{13} C_2 \sqrt{\frac{\log \log r}{r}}
\end{aligned} \tag{43}$$

(43) holds for C_2 large enough (but not depends on r). For example, we can choose $C_{13} = C_{12}/2$ and then (43) holds for all $C_2 \geq 2C_{11} + \frac{2C_4 C_5^2 + C_7}{C_{12}}$. Since C_3, C_4, \dots, C_{13} are all independent of r , the C_2 here is also independent of r . Then,

$$\sum_{i \neq b \neq k} \left(\frac{N_{i,j}^{(r+s_r)} / \sigma_{i,j}^{(r+s_r)}}{N_{b,j}^{(r+s_r)} / \sigma_{b,j}^{(r+s_r)}} \right)^2 - \sum_{i \neq b \neq k} \left(\frac{N_{i,j}^{(r)} / \sigma_{i,j}^{(r+s_r)}}{N_{b,j}^{(r)} / \sigma_{b,j}^{(r+s_r)}} \right)^2 \geq C_{13} C_2 \sqrt{\frac{\log \log r}{r}}$$

There exists a non-optimal design $h \neq k \neq b$, such that

$$\left(\frac{N_{h,j}^{(r+s_r)} / \sigma_{h,j}^{(r+s_r)}}{N_{b,j}^{(r+s_r)} / \sigma_{b,j}^{(r+s_r)}} \right)^2 - \left(\frac{N_{h,j}^{(r)} / \sigma_{h,j}^{(r+s_r)}}{N_{b,j}^{(r)} / \sigma_{b,j}^{(r+s_r)}} \right)^2 \geq \frac{1}{K-2} C_{13} C_2 \sqrt{\frac{\log \log r}{r}}.$$

Or equivalently,

$$\left(\frac{N_{h,j}^{(r+s_r)} / N_{b,j}^{(r+s_r)}}{N_{h,j}^{(r)} / N_{b,j}^{(r)}} \right)^2 \geq 1 + \frac{C_{13} C_2}{K-2} \left(\frac{\sigma_{h,j}^{(r+s_r)}}{\sigma_{b,j}^{(r+s_r)}} \right)^2 \left(\frac{N_{b,j}^{(r)}}{N_{h,j}^{(r)}} \right)^2 \sqrt{\frac{\log \log r}{r}}.$$

By Lemma 5 and the convergence of the sample variance, there exists $0 < C_{14} < \frac{C_{13}}{K-2} \left(\frac{\sigma_{h,j}^{(r+s_r)}}{\sigma_{b,j}^{(r+s_r)}}\right)^2 \left(\frac{N_{b,j}^{(r)}}{N_{h,j}^{(r)}}\right)^2$ for all large r . We obtain

$$\left(\frac{N_{h,j}^{(r+s_r)} / N_{b,j}^{(r+s_r)}}{N_{h,j}^{(r)} / N_{b,j}^{(r)}} \right)^2 \geq 1 + C_{14} C_2 \sqrt{\frac{\log \log r}{r}}$$

Hence,

$$\frac{N_{h,j}^{(r+s_r)}}{N_{b,j}^{(r+s_r)}} \geq \sqrt{1 + C_{14} C_2 \sqrt{\frac{\log \log r}{r}}} \frac{N_{h,j}^{(r)}}{N_{b,j}^{(r)}} \geq \left(1 + \frac{C_{14} C_2}{4} \sqrt{\frac{\log \log r}{r}}\right) \frac{N_{h,j}^{(r)}}{N_{b,j}^{(r)}} \tag{44}$$

for all large r by Taylor Expansion. Let $v = \sup\{l \leq s_r : \mathbf{1}_{h,j}^{(r+l)} = 1\}$. $r + v$ is the last time before $r + s_r$ at which h is sampled. We then have

$$\begin{aligned}
\frac{N_{h,j}^{(r+v)}}{N_{b,j}^{(r+v)}} &\geq \frac{N_{h,j}^{(r+v)}}{N_{b,j}^{(r+s_r)}} \\
&= \frac{N_{h,j}^{(r+s_r)} - 1}{N_{b,j}^{(r+s_r)}} \\
&= \left(1 - \frac{1}{N_{b,j}^{(r+s_r)}}\right) \frac{N_{h,j}^{(r+s_r)}}{N_{b,j}^{(r+s_r)}} \\
&\geq \left(1 - \frac{1}{N_{b,j}^{(r)}}\right) \left(1 + \frac{C_{14}C_2}{4} \sqrt{\frac{\log \log r}{r}}\right) \frac{N_{h,j}^{(r)}}{N_{b,j}^{(r)}} \\
&\geq \left(1 - \frac{1}{C_9 r}\right) \left(1 + \frac{C_{14}C_2}{4} \sqrt{\frac{\log \log r}{r}}\right) \frac{N_{h,j}^{(r)}}{N_{b,j}^{(r)}} \\
&\geq \left(1 + \frac{C_{14}C_2}{8} \sqrt{\frac{\log \log r}{r}}\right) \frac{N_{h,j}^{(r)}}{N_{b,j}^{(r)}}
\end{aligned}$$

for all large r . Then, given any C_1 , there exists $C_2 \geq \max\{\frac{8C_1}{C_{14}}, 2C_{11} + \frac{2C_4C_5^2 + C_7}{C_{12}}\}$. The Lemma holds for true for the C_1, C_2 . \blacksquare

Lemma 16 For a fixed input realization j_0 and sub-optimal design k , between two samples of (k, j_0) . Suppose (k, j_0) is sampled at r and let $t_r = \inf\{l > 0 : \mathbf{1}_{(k, j_0)} = 1\}$. $r + t_r$ is the next time (k, j_0) being sampled. Then the number of samples that can be allocated to (b, j_0) between r and $r + t_r$ is $O(\sqrt{r \log \log r})$ almost surely.

Proof We use the same notation of $s'_r, s_r, d_{i,j}^{(r,q)}$ as in Lemma 15. Since $d_{b,j_0}^{(r,t_r)} - d_{b,j_0}^{(r,s_r)} = d_{b,j_0}^{(r,s'_r)} - d_{b,j_0}^{(r,s_r)} + 1 = O(\sqrt{r \log \log r})$ by Lemma 13, it is sufficient to prove $d_{b,j_0}^{(r,s_r)} = O(\sqrt{r \log \log r})$. Prove by contradiction. Suppose the statement does not hold. Then $\forall C_2 > 0$, there exists r such that $d_{b,j_0}^{(r,s_r)} \geq C_2 \sqrt{r \log \log r}$. By Lemma 15, $\forall C_1 > 0$ (remain to be specified), there exists a iteration r at which (k, j_0) is sampled, an another non-optimal design $h \neq k$ and a iteration $v < s_r$, such that (h, j_0) is sampled at v and

$$\frac{\alpha_{h,j_0}^{(r+v)}}{\alpha_{b,j_0}^{(r+s_r)}} \geq \left(1 + C_1 \sqrt{\frac{\log \log r}{r}}\right) \frac{\alpha_{h,j_0}^{(r)}}{\alpha_{b,j_0}^{(r)}}$$

holds. We aim to show (h, j_0) cannot be sampled at $r + v$ for a contradiction. It is sufficient to show

$$\frac{(\hat{\mu}_b^{(r+v)} - \hat{\mu}_h^{(r+v)})^2}{\sum_{j=1}^D \frac{(\sigma_{b,j}^{(r+v)})^2 (p_j^{r+v})^2}{\alpha_{b,j}^{(r+v)}} + \sum_{j=1}^D \frac{(\sigma_{h,j}^{(r+v)})^2 (p_j^{r+v})^2}{\alpha_{h,j}^{(r+v)}}} > \frac{(\hat{\mu}_b^{(r+v)} - \hat{\mu}_k^{(r+v)})^2}{\sum_{j=1}^D \frac{(\sigma_{b,j}^{(r+v)})^2 (p_j^{r+v})^2}{\alpha_{b,j}^{(r+v)}} + \sum_{j=1}^D \frac{(\sigma_{k,j}^{(r+v)})^2 (p_j^{r+v})^2}{\alpha_{k,j}^{(r+v)}}}. \quad (45)$$

Denote by $\delta_i^{(l)} = (\hat{\mu}_b^{(l)} - \hat{\mu}_i^{(l)})^2$. It is equivalent to show

$$\delta_h^{(r+v)} \left(\sum_{j=1}^D \frac{(\sigma_{b,j}^{(r+v)})^2 (p_j^{r+v})^2}{\alpha_{b,j}^{(r+v)}} + \sum_{j=1}^D \frac{(\sigma_{k,j}^{(r+v)})^2 (p_j^{r+v})^2}{\alpha_{k,j}^{(r+v)}} \right) > \delta_k^{(r+v)} \left(\sum_{j=1}^D \frac{(\sigma_{b,j}^{(r+v)})^2 (p_j^{r+v})^2}{\alpha_{b,j}^{(r+v)}} + \sum_{j=1}^D \frac{(\sigma_{h,j}^{(r+v)})^2 (p_j^{r+v})^2}{\alpha_{h,j}^{(r+v)}} \right) \quad (46)$$

By Lemma 11, we have $\left| \frac{\alpha_{i,j}^{(r)}}{\sigma_{i,j}^{(r)} p_j^r} - \frac{\alpha_{i,j'}^{(r)}}{\sigma_{i,j'}^{(r)} p_{j'}^r} \right| = O(\sqrt{\frac{\log \log r}{r}})$. Hence, $\left| \frac{\sigma_{i,j}^{(r)} p_j^r}{\alpha_{i,j}^{(r)}} - \frac{\sigma_{i,j'}^{(r)} p_{j'}^r}{\alpha_{i,j'}^{(r)}} \right| = O(\sqrt{\frac{\log \log r}{r}})$. Then, there exist $C_3 > 0$

$$\text{LHS of (46)} \geq \delta_h^{(r+v)} \left(\frac{(\sigma_{b,j}^{(r+v)})^2 p_j^{r+v}}{\alpha_{b,j}^{(r+v)}} \sum_{q=1}^D \sigma_{b,q}^{(r+v)} p_q^{r+v} + \frac{(\sigma_{k,j}^{(r+v)})^2 p_j^{r+v}}{\alpha_{k,j}^{(r+v)}} \sum_{q=1}^D \sigma_{k,q}^{(r+v)} p_q^{r+v} \right) - C_3 \sqrt{\frac{\log \log r}{r}}$$

$$\text{RHS of (46)} \leq \delta_k^{(r+v)} \left(\frac{\sigma_{b,j}^{(r+v)} p_j^{r+v}}{\alpha_{b,j}^{(r+v)}} \sum_{q=1}^D \sigma_{b,q}^{(r+v)} p_q^{r+v} + \frac{\sigma_{h,j}^{(r+v)} p_j^{r+v}}{\alpha_{k,j}^{(r+v)}} \sum_{q=1}^D \sigma_{h,q}^{(r+v)} p_q^{r+v} \right) + C_3 \sqrt{\frac{\log \log r}{r}}$$

Denote by $\gamma_i^{(l)} = \sum_{q=1}^D \hat{\sigma}_{i,q}^{(l)} p_q^l$. It is equivalent to show

$$\begin{aligned} & \delta_h^{(r+v)} \left(\frac{\sigma_{b,j}^{(r+v)} p_j^{r+v}}{\alpha_{b,j}^{(r+v)}} \gamma_b^{(r+v)} + \frac{\sigma_{k,j}^{(r+v)} p_j^{r+v}}{\alpha_{k,j}^{(r+v)}} \gamma_k^{(r+v)} \right) - C_3 \sqrt{\frac{\log \log r}{r}} \\ & > \delta_k^{(r+v)} \left(\frac{\sigma_{b,j}^{(r+v)} p_j^{r+v}}{\alpha_{b,j}^{(r+v)}} \gamma_b^{(r+v)} + \frac{\sigma_{h,j}^{(r+v)} p_j^{r+v}}{\alpha_{k,j}^{(r+v)}} \gamma_h^{(r+v)} \right) + C_3 \sqrt{\frac{\log \log r}{r}} \end{aligned} \quad (47)$$

Since (k, j_0) is sampled at r ,

$$\delta_h^{(r)} \left(\sum_{j=1}^D \frac{(\sigma_{b,j}^{(r)})^2 (p_j^r)^2}{\alpha_{b,j}^{(r)}} + \sum_{j=1}^D \frac{(\sigma_{k,j}^{(r)})^2 (p_j^r)^2}{\alpha_{k,j}^{(r)}} \right) \geq \delta_k^{(r)} \left(\sum_{j=1}^D \frac{(\sigma_{b,j}^{(r)})^2 (p_j^r)^2}{\alpha_{b,j}^{(r)}} + \sum_{j=1}^D \frac{(\sigma_{h,j}^{(r)})^2 (p_j^r)^2}{\alpha_{h,j}^{(r)}} \right).$$

Again by Lemma 11, there exists $C_4 > 0$, such that

$$\delta_h^{(r)} \left(\frac{\sigma_{b,j_0}^{(r)} p_{j_0}^r}{\alpha_{b,j_0}^{(r)}} \gamma_b^{(r)} + \frac{\sigma_{k,j_0}^{(r)} p_{j_0}^r}{\alpha_{k,j_0}^{(r)}} \gamma_k^{(r)} \right) + C_4 \sqrt{\frac{\log \log r}{r}} \geq \delta_k^{(r)} \left(\frac{\sigma_{b,j_0}^{(r)} p_{j_0}^r}{\alpha_{b,j_0}^{(r)}} \gamma_b^{(r)} + \frac{\sigma_{h,j_0}^{(r)} p_{j_0}^r}{\alpha_{k,j_0}^{(r)}} \gamma_h^{(r)} \right) - C_4 \sqrt{\frac{\log \log r}{r}}. \quad (48)$$

Then,

$$\begin{aligned} \text{LHS of (47)} &= \delta_h^{(r)} \frac{\sigma_{b,j_0}^{(r)} p_{j_0}^r}{\alpha_{b,j_0}^{(r)}} \gamma_b^{(r)} \cdot \frac{\delta_h^{(r+v)} \sigma_{b,j_0}^{(r+v)} \alpha_{b,j_0}^{(r)} \gamma_b^{(r+v)} p_{j_0}^{r+v}}{\delta_h^{(r)} \sigma_{b,j_0}^{(r)} \alpha_{b,j_0}^{(r+v)} \gamma_b^{(r)} p_{j_0}^r} \\ &+ \delta_h^{(r)} \frac{\sigma_{k,j_0}^{(r)} p_{j_0}^r}{\alpha_{k,j_0}^{(r)}} \gamma_k^{(r)} \cdot \frac{\delta_h^{(r+v)} \sigma_{k,j_0}^{(r+v)} \alpha_{k,j_0}^{(r)} \gamma_k^{(r+v)} p_{j_0}^{r+v}}{\delta_h^{(r)} \sigma_{k,j_0}^{(r)} \alpha_{k,j_0}^{(r+v)} \gamma_k^{(r)} p_{j_0}^r} - C_3 \sqrt{\frac{\log \log r}{r}}. \end{aligned} \quad (49)$$

and

$$\begin{aligned} \text{RHS of (47)} &= \delta_k^{(r)} \frac{\sigma_{b,j_0}^{(r)} p_{j_0}^r}{\alpha_{b,j_0}^{(r)}} \gamma_b^{(r)} \cdot \frac{\delta_k^{(r+v)} \sigma_{b,j_0}^{(r+v)} \alpha_{b,j_0}^{(r)} \gamma_b^{(r+v)} p_{j_0}^{r+v}}{\delta_k^{(r)} \sigma_{b,j_0}^{(r)} \alpha_{b,j_0}^{(r+v)} \gamma_b^{(r)} p_{j_0}^r} \\ &+ \delta_k^{(r)} \frac{\sigma_{h,j_0}^{(r)} p_{j_0}^r}{\alpha_{h,j_0}^{(r)}} \gamma_h^{(r)} \cdot \frac{\delta_k^{(r+v)} \sigma_{h,j_0}^{(r+v)} \alpha_{h,j_0}^{(r)} \gamma_h^{(r+v)} p_{j_0}^{r+v}}{\delta_k^{(r)} \sigma_{h,j_0}^{(r)} \alpha_{h,j_0}^{(r+v)} \gamma_h^{(r)} p_{j_0}^r} + C_3 \sqrt{\frac{\log \log r}{r}}. \end{aligned} \quad (50)$$

$\forall 1 \leq i \leq K$, by Lemma 7, $\frac{\delta_i^{(r+v)}}{\delta_i^{(r)}} = 1 + O(\sqrt{\frac{\log \log r}{r}})$; by Lemma 2 and 8, $\frac{p_{j_0}^{r+v}}{p_{j_0}^r} = 1 + O(\sqrt{\frac{\log \log r}{r}})$, $\frac{\sigma_{i,j_0}^{(r+v)}}{\sigma_{i,j_0}^{(r)}} = 1 + O(\sqrt{\frac{\log \log r}{r}})$ and $\frac{\gamma_i^{(r+v)}}{\gamma_i^{(r)}} = 1 + O(\sqrt{\frac{\log \log r}{r}})$; Furthermore $N_{k,j_0}^{(r+v)} = N_{k,j_0}^{(r)} + 1$. Hence, there exist $C_5 > 0$, such that

$$(49) \geq \delta_h^{(r)} \frac{\sigma_{b,j_0}^{(r)} p_{j_0}^r}{\alpha_{b,j_0}^{(r)}} \gamma_b^{(r)} \frac{\alpha_{b,j_0}^{(r)}}{\alpha_{b,j_0}^{(r+v)}} (1 - C_5 \sqrt{\frac{\log \log r}{r}}) + \delta_h^{(r)} \frac{\sigma_{k,j_0}^{(r)} p_{j_0}^r}{\alpha_{k,j_0}^{(r)}} \gamma_h^{(r)} (1 - C_5 \sqrt{\frac{\log \log r}{r}}) - C_3 \sqrt{\frac{\log \log r}{r}} \quad (51)$$

and

$$(50) \leq \delta_k^{(r)} \frac{\sigma_{b,j_0}^{(r)} p_{j_0}^r}{\alpha_{b,j_0}^{(r)}} \gamma_b^{(r)} \frac{\alpha_{b,j_0}^{(r)}}{\alpha_{b,j_0}^{(r+v)}} (1 + C_5 \sqrt{\frac{\log \log r}{r}}) + \delta_k^{(r)} \frac{\sigma_{h,j_0}^{(r)} p_{j_0}^r}{\alpha_{h,j_0}^{(r)}} \gamma_h^{(r)} \frac{\alpha_{h,j_0}^{(r)}}{\alpha_{h,j_0}^{(r+v)}} (1 + C_5 \sqrt{\frac{\log \log r}{r}}) + C_3 \sqrt{\frac{\log \log r}{r}} \quad (52)$$

Divide by $\frac{\alpha_{b,j_0}^{(r)}}{\alpha_{b,j_0}^{(r+v)}}$ by both sides and notice $1 \leq \frac{\alpha_{b,j_0}^{(r+v)}}{\alpha_{b,j_0}^{(r)}} \leq \frac{1}{\alpha_{b,j_0}^{(r)}} \leq C_7$ for some $C_7 > 0$ by Lemma 6. Then (51) > (52) can be implied by

$$\begin{aligned} & \delta_h^{(r)} \frac{\sigma_{b,j_0}^{(r)} p_{j_0}^r}{\alpha_{b,j_0}^{(r)}} \gamma_b + \delta_h^{(r)} \frac{\sigma_{k,j_0}^{(r)} p_{j_0}^r}{\alpha_{k,j_0}^{(r)}} \gamma_h - C_8 \sqrt{\frac{\log \log r}{r}} \\ & > \delta_k^{(r)} \frac{\sigma_{b,j_0}^{(r)} p_{j_0}^r}{\alpha_{b,j_0}^{(r)}} \gamma_b + \delta_k^{(r)} \frac{\sigma_{h,j_0}^{(r)} p_{j_0}^r}{\alpha_{h,j_0}^{(r)}} \gamma_h \left(\frac{\alpha_{h,j_0}^{(r)}}{\alpha_{b,j_0}^{(r)}} / \frac{\alpha_{h,j_0}^{(r+v)}}{\alpha_{b,j_0}^{(r+v)}} \right) (1 + C_5 \sqrt{\frac{\log \log r}{r}}) + C_8 \sqrt{\frac{\log \log r}{r}} \end{aligned}$$

for some C_8 independent of r . Since $\left(\frac{\alpha_{h,j_0}^{(r)}}{\alpha_{b,j_0}^{(r)}} / \frac{\alpha_{h,j_0}^{(r+v)}}{\alpha_{b,j_0}^{(r+v)}} \right) < \frac{1}{1+C_1 \sqrt{\frac{\log \log r}{r}}} < 1 - \frac{C_1}{2} \sqrt{\frac{\log \log r}{r}}$ for all large r 's, (53) can be further implied by

$$\begin{aligned} & \delta_h^{(r)} \frac{\sigma_{b,j_0}^{(r)} p_{j_0}^r}{\alpha_{b,j_0}^{(r)}} \gamma_b + \delta_h^{(r)} \frac{\sigma_{k,j_0}^{(r)} p_{j_0}^r}{\alpha_{k,j_0}^{(r)}} \gamma_h - C_8 \sqrt{\frac{\log \log r}{r}} \\ & > \delta_k^{(r)} \frac{\sigma_{b,j_0}^{(r)} p_{j_0}^r}{\alpha_{b,j_0}^{(r)}} \gamma_b + \delta_k^{(r)} \frac{\sigma_{h,j_0}^{(r)} p_{j_0}^r}{\alpha_{h,j_0}^{(r)}} \gamma_h \left(1 - \frac{C_1}{2} \sqrt{\frac{\log \log r}{r}}\right) (1 + C_5 \sqrt{\frac{\log \log r}{r}}) + C_8 \sqrt{\frac{\log \log r}{r}} \quad (53) \\ & = \delta_k^{(r)} \frac{\sigma_{b,j_0}^{(r)} p_{j_0}^r}{\alpha_{b,j_0}^{(r)}} \gamma_b + \delta_k^{(r)} \frac{\sigma_{h,j_0}^{(r)} p_{j_0}^r}{\alpha_{h,j_0}^{(r)}} \gamma_h + \left[C_8 + C_5 \delta_k^{(r)} \frac{\sigma_{h,j_0}^{(r)} p_{j_0}^r}{\alpha_{h,j_0}^{(r)}} \gamma_h - \frac{C_1}{2} \delta_k^{(r)} \frac{\sigma_{h,j_0}^{(r)} p_{j_0}^r}{\alpha_{h,j_0}^{(r)}} \gamma_h \right] \sqrt{\frac{\log \log r}{r}} \end{aligned}$$

Let C_9, C_{10} be two constants such that $C_9 \leq \delta_k^{(r)} \frac{\sigma_{h,j_0}^{(r)} p_{j_0}^r}{\alpha_{h,j_0}^{(r)}} \gamma_h \leq C_{10}$. Then by (48), (53) is implied by

$$\left[2C_8 + 2C_4 + C_5 C_{10} - \frac{C_1}{2} C_9 \right] \sqrt{\frac{\log \log r}{r}} < 0. \quad (54)$$

Since C_3, C_4, \dots, C_{10} are all independent of r , we can choose C_1 to be large enough such that (54) holds, which gives us the contradiction that (h, j_0) cannot be sampled at $r + v$.

We can now finally prove the ‘‘Local Balance’’ optimality condition.

Proof of (21) in Theorem 3. Denote by $\delta_i^{(l)} = (\hat{\mu}_b^{(l)} - \hat{\mu}_i^{(l)})^2$, $\gamma_i^{(l)} = \sum_{j=1}^D \sigma_{i,j}^{(r)} p_j^l$ and $\Theta_i^{(l)} =$

$$\frac{\delta_i^{(l)}}{\sum_{j=1}^D \frac{(\sigma_{i,j}^{(r)})^2 (p_j^l)^2}{\alpha_{i,j}^{(l)}} + \sum_{j=1}^D \frac{(\sigma_{i,j}^{(r)})^2 (p_j^l)^2}{\alpha_{i,j}^{(l)}}}. \text{ For any two non-optimal designs } i \neq k \text{ and an iteration } r. \text{ Without loss of generality, assume } \Theta_i^{(r)} \leq \Theta_k^{(r)}. \text{ Let } u = \sup\{l < r : \sum_{j=1}^D \mathbf{1}_{k,j}^{(l)} = 1\} \text{ be the last time } k \text{ is sampled before } r. \text{ Suppose } (k, j_0)$$

is sampled at u . Then $r = O(u)$ by Lemma 9. By Lemma 11 and the proof of Lemma 16, there exists $C_1 > 0$,

$$\begin{aligned} 0 & \leq \Theta_k^{(r)} - \Theta_i^{(r)} \\ & \leq \frac{\delta_k^{(r)}}{\frac{\sigma_{b,j_0}^{(r)} p_{j_0}^r}{\alpha_{b,j_0}^{(r)}} \gamma_b + \frac{\sigma_{k,j_0}^{(r)} p_{j_0}^r}{\alpha_{k,j_0}^{(r)}} \gamma_k} - \frac{\delta_i^{(r)}}{\frac{\sigma_{b,j_0}^{(r)} p_{j_0}^r}{\alpha_{b,j_0}^{(r)}} \gamma_b + \frac{\sigma_{i,j_0}^{(r)} p_{j_0}^r}{\alpha_{i,j_0}^{(r)}} \gamma_i} + C_1 \sqrt{\frac{\log \log r}{r}} \quad (55) \end{aligned}$$

Since by Lemma 6, all $\frac{\sigma_{i',j'}^{(r)} p_{j'}^r}{\alpha_{i',j'}^{(r)}} \gamma_{i'}$ are upper and lower bounded by positive constants almost surely. Then there exists $C_2 > 0$,

$$(55) \leq \frac{\delta_k^{(r)}}{\frac{\sigma_{b,j_0}^{(r)} p_{j_0}^r}{\alpha_{b,j_0}^{(r)}} \gamma_b + \frac{\sigma_{k,j_0}^{(r)} p_{j_0}^r}{\alpha_{k,j_0}^{(r)}} \gamma_k} - \frac{\delta_i^{(r)}}{\frac{\sigma_{b,j_0}^{(r)} p_{j_0}^r}{\alpha_{b,j_0}^{(r)}} \gamma_b + \frac{\sigma_{i,j_0}^{(r)} p_{j_0}^r}{\alpha_{i,j_0}^{(r)}} \gamma_i} + C_2 \sqrt{\frac{\log \log r}{r}} \quad (56)$$

By Lemma (7), $\delta_{i'}^{(r)} = \delta_{i'}^{(u)} + O(\sqrt{\frac{\log \log u}{u}}) = \delta_{i'}^{(u)} + O(\sqrt{\frac{\log \log r}{r}})$, $\forall i'$; by Lemma 2 and Lemma 8, $p_j^r = p_j^u + O(\sqrt{\frac{\log \log r}{r}})$, $\hat{\sigma}_{i,j}^{(r)} = \hat{\sigma}_{i,j}^{(u)} + O(\sqrt{\frac{\log \log r}{r}})$ and hence $\gamma_{i'}^{(r)} = \gamma_{i'}^{(u)} + O(\sqrt{\frac{\log \log r}{r}})$, $\forall i'$. Furthermore $\alpha_{i',j'}^{(l)}$ is upper and lower bounded by positive constants $\forall i', j'$. There exists $C_3 > C_2$,

$$(56) \leq \frac{\delta_k^{(u)}}{\frac{\sigma_{b,j_0}^{(u)} p_{j_0}^u}{\alpha_{b,j_0}^{(r)}} \gamma_b^{(u)} + \frac{\sigma_{k,j_0}^{(u)} p_{j_0}^u}{\alpha_{k,j_0}^{(r)}} \gamma_k^{(u)}} - \frac{\delta_i^{(u)}}{\frac{\sigma_{b,j_0}^{(u)} p_{j_0}^u}{\alpha_{b,j_0}^{(r)}} \gamma_b^{(u)} + \frac{\sigma_{i,j_0}^{(u)} p_{j_0}^u}{\alpha_{i,j_0}^{(r)}} \gamma_i^{(u)}} + C_3 \sqrt{\frac{\log \log r}{r}}. \quad (57)$$

By Lemma 16, $N_{b,j_0}^{(r)} = N_{b,j}^{(u)} + O(\sqrt{u \log \log u}) = N_{b,j}^{(u)} + O(\sqrt{r \log \log r})$. Hence, there exists $C_4 > 0$,

$$(57) \leq \frac{1}{N^{(r)}} \frac{\delta_k^{(u)}}{\frac{\sigma_{b,j_0}^{(u)} p_{j_0}^u}{N_{b,j_0}^{(u)} + C_4 \sqrt{r \log \log r}} \gamma_b^{(u)} + \frac{\sigma_{k,j_0}^{(u)} p_{j_0}^u}{N_{k,j_0}^{(u)} + 1} \gamma_k^{(u)}} - \frac{1}{N^{(r)}} \frac{\delta_i^{(u)}}{\frac{\sigma_{b,j_0}^{(u)} p_{j_0}^u}{N_{b,j_0}^{(u)}} \gamma_b^{(u)} + \frac{\sigma_{i,j_0}^{(u)} p_{j_0}^u}{N_{i,j_0}^{(u)}} \gamma_i^{(u)}} + C_3 \sqrt{\frac{\log \log r}{r}} \quad (58)$$

By Lemma 6 and $l = O(u)$, $\liminf_{r \rightarrow \infty} \frac{N_{i,j}^{(u)}}{N^{(r)}} > 0 \forall i, j$. Then, there exists $C_5 > 0$,

$$(58) \leq \frac{1}{N^{(r)}} \frac{\delta_k^{(u)}}{\frac{\sigma_{b,j_0}^{(u)} p_{j_0}^u}{N_{b,j_0}^{(u)}} \gamma_b^{(u)} + \frac{\sigma_{k,j_0}^{(u)} p_{j_0}^u}{N_{k,j_0}^{(u)}} \gamma_k^{(u)}} - \frac{1}{N^{(r)}} \frac{\delta_i^{(u)}}{\frac{\sigma_{b,j_0}^{(u)} p_{j_0}^u}{N_{b,j_0}^{(u)}} \gamma_b^{(u)} + \frac{\sigma_{i,j_0}^{(u)} p_{j_0}^u}{N_{i,j_0}^{(u)}} \gamma_i^{(u)}} + C_5 \sqrt{\frac{\log \log r}{r}} \\ \leq C_5 \sqrt{\frac{\log \log r}{r}}, \quad (59)$$

which proves the desired result. ■