Rate-Optimal Contextual Ranking and Selection

Jianzhong Du\textsuperscript{a}, Siyang Gao\textsuperscript{b} and Chun-Hung Chen\textsuperscript{c}

\textsuperscript{a}School of Management, Fudan University, Shanghai, China, jianzhodu2-c@my.cityu.edu.hk
\textsuperscript{b}Department of Advanced Design and Systems Engineering and School of Data Science, City University of Hong Kong, Hong Kong, China, siyangao@cityu.edu.hk
\textsuperscript{c}Department of Systems Engineering & Operations Research, George Mason University, Fairfax, VA 22030, USA, cchen9@gmu.edu

Abstract

The ranking and selection (R&S) problem seeks to efficiently select the best simulated system design among a finite number of alternatives. It is a well-established problem in simulation-based optimization, and has wide applications in the production, service and operation management. In this research, we consider R&S in the presence of context (also known as the covariates, side information or auxiliary information in the literature), where the context corresponds to some input information to the simulation model and can influence the performance of each design. This is a new and emerging problem in simulation for personalized decision making. The goal is to determine the best allocation of the simulation budget among various contexts and designs so as to efficiently identify the best design for all the contexts that might possibly appear. We call it contextual ranking and selection (CR&S). We utilize the OCBA approach in R&S, and solve the problem by developing appropriate objective measures, identifying the rate-optimal budget allocation rule and analyzing the convergence of the selection algorithm. We numerically test the performance of the proposed algorithm via a set of abstract and real-world problems, and show the superiority of the algorithm in solving these problems and obtaining real-time decisions.

Keywords: simulation optimization, contextual ranking and selection, OCBA, personalized decision making, convergence rate.
1 Introduction

Simulation is a powerful modeling tool for analyzing complex stochastic systems and evaluating decision problems, since these systems and problems usually do not satisfy the assumptions of analytical models. Simulation makes it possible to accurately describe a system through the use of logically complex, and often nonmathematical models. Detailed dynamics of the system can therefore be modeled (Chen and Lee 2011). Examples such as inventory control, buffer allocation, portfolio management, queueing network design, healthcare management and power grid optimization all fall into the applicable areas of simulation.

When conducting simulation experiments, designs (decisions) are simulated for multiple times and their estimated performances, typically in terms of sample means, are compared. It gives rise to the research in ranking and selection (R&S): techniques that determine the number of simulation replications for each system design so as to efficiently select the best design among a finite set of alternatives. Nowadays R&S has become one of the main research questions in simulation-based optimization. For a comprehensive review of this field, see Branke et al. (2007), Xu et al. (2015).

In the R&S methods, the standard practice is to first fix the context of the problem under consideration and then repeat experiments on the simulation model with various designs to select the best one for the given context. The context here refers to some input information to the problem which influences the performance of each design, and is also known as the covariates, side information or auxiliary information in the literature. For example, doctors use simulation to determine the best treatment method for the cancer patients (Kim et al. 2011) and the diabetic patients (Bertsimas et al. 2017) under the context of the patients’ biometric characteristics.

However, a notable issue with this fashion of conducting simulation is the computational burden. It is well known that the simulation experiments are time-consuming (Law and Kelton 2000). For example, it can take one and half hours to simulate the 1000-second dynamics of a 12000-node-single-bottleneck computer network, and two hours to simulate the 24-hour dynamics of a transportation network with 20 intersections (Ho et al. 2007). As a result, if we start the simulation after the problem is set up with the context input, the time for obtaining the best design is usually very long. This issue is even more evident in online optimization, in which the optimal decision tailored to the given context is expected soon after this context (e.g., user preference) is revealed.
In the meantime, the explosion in the availability and accessibility of data these days enables us to have a clear picture of the context space of the simulation model. It opens the door for a new way of conducting simulation-based decision making, which is to run simulation experiments on different designs and contexts in advance, and prescribe a design in anticipation of the future given context. Then, when a certain context is revealed, we can check our existing computational results and provide the best design immediately. For example, the doctor can provide a personalized treatment for a diabetic patient immediately upon his/her arrival by checking the R&S results under the same biometric characteristics (context) of this patient (Bertsimas et al. 2017). This new way of using simulation avoids the major issue of the long decision time, and, more importantly, makes it possible for the R&S approaches to be applied to problems for which R&S was hardly a solution methodology before, e.g., the real-time optimization.

In this research, we consider the R&S problem with this new perspective, namely contextual ranking and selection (CR&S). Same as our treatment to designs, we assume that there are a finite number of possible contexts (rationale of this assumption will be provided in Section 3).

Though related to R&S, CR&S is significantly different from R&S in the following two aspects. First, R&S only needs to determine the number of simulation replications for each design, while CR&S determines the number of simulation replications for each pair of context and design. The entities that receive simulation replications in CR&S are much more, and different in structure from those in R&S. Second, for CR&S, there is not even a measure like the probability of correct selection (PCS) in R&S (Chen et al. 2000, Kim and Nelson 2001) that is readily available to characterize what we want to optimize. Since there are multiple contexts in CR&S, each associated with a traditional R&S problem, the measure to be optimized should reasonably summarize the evidence of correct selection under each context.

Our contribution in this research is four-fold. First, we introduce three measures for evaluating the evidence of correct selection over the entire context space. These measures are variants of the PCS used in the traditional R&S, and are capable of depicting the quality of the selected design of the R&S problem associated with each context that might possibly appear. Moreover, we show that the three measures are asymptotically equivalent, in the sense that they have the same rate function. Then, instead of considering the three measures separately, we can solve them once and for all by directly optimizing this rate function.
Second, we establish the first budget allocation formulation for CR&S. This is an optimal computing budget allocation (OCBA)-like formulation \cite{Chen2011}, with the goal optimizing the rate function under a simulation budget constraint. In addition, we develop the rate-optimal budget allocation rule for this formulation, and design a selection algorithm that can be easily implemented in practice.

Third, we prove that the proposed selection algorithm converges to the rate-optimal budget allocation rule. That is, the algorithm is able to recover the optimal convergence rate of the three measures in the limit. In most R&S-related problems (including this study), the conditions that asymptotically maximize the probability of correct selection do not lead to closed-form budget allocation rules, which substantially adds to the difficulty of analyzing the theoretical properties of their corresponding selection algorithms due to the complex behavior of some measure estimators in the algorithms. The OCBA methods tend to handle this difficulty by making approximations on these conditions, assuming the number of simulation replications of the targeted design(s) is significantly larger than that of each rest design \cite{Chen2000, Lee2010, Lee2012}, to turn the budget allocation rules into closed form. However, the algorithms designed based on them do not converge to the rate-optimal budget allocation rules any more due the approximation that has been made. In this research, we remain the optimality conditions unchanged. A bound is identified for the difference between the outputted budget allocation rule of the algorithm and the rate-optimal one, and we show that the bound vanishes as the algorithm proceeds.

Last, we conduct extensive numerical experiments to assess the performance of our algorithm. We first show the convergence patterns of the three objective measures we proposed. The results are in line with our theoretical characterization of their converging behavior. In addition, we apply our proposed selection algorithm to a set of benchmark test functions and real-world applications, including a production line optimization, an assemble-to-order (ATO) problem and a personalized cancer treatment problem. They demonstrate the superiority and effectiveness of our method in solving different types of decision problems.
2 Related Literature

There are two primary streams of literature in R&S. The first stream assumes that a simulation budget is given, and studies how to allocate the simulation budget to the designs so that the probability of correct selection (PCS) for the best design can be maximized. Common methods along this stream include the optimal computing budget allocation (OCBA), which aims to find the asymptotic optimal budget allocation rule based on some approximation of PCS (Chen et al. 2000, Fu et al. 2007, Chen et al. 2008, Lee et al. 2012, Gao and Chen 2016, Gao et al. 2017a), and the value of information procedure (VIP), which employs a Bayesian framework and sequentially allocates simulation samples using predictive distributions of further samples in order to maximize a certain acquisition function such as the expected improvement or knowledge gradient (Frazier et al. 2008, Chick et al. 2010, Ryzhov 2016). Due to the lack of an analytical expression for PCS, the solution that directly maximizes PCS is generally unknown, and most of the research studies in this stream try to solve the problem asymptotically, i.e., finding the budget allocation rule that asymptotically maximizes PCS.

The other stream focuses on providing a guarantee on the PCS of the design selected. An important method in this regard is the indifference-zone (IZ) procedure. It assumes that the mean performance of the best design is at least \( \delta^* \) better than each alternative, where \( \delta^* \) is the minimum difference worth detecting, and keeps increasing the simulation replications allocated to the designs until the PCS is guaranteed to reach a pre-specified level (Dudewicz and Dalal 1975, Rinott 1978, Kim and Nelson 2001, Nelson et al. 2001).

Compared to R&S, CR&S is relatively new, and the literature on it is sparse. Shen et al. (2017) is perhaps the first research effort along this direction, which applies the IZ method for ranking and selection in the presence of covariates (context). That paper seeks to provide a quality guarantee on the selection over the covariate space instead of investigating the convergence rate. Consequently, the main research questions Shen et al. (2017) and this paper consider and the budget allocation rules and algorithms the two papers propose to develop differ in the very beginning. In addition, the two papers are also substantially different in the setting of the covariate space and the solution technique used. Li et al. (2018) further extended the result in Shen et al. (2017) to high-dimensional covariates and general dependence between the mean performance of a design and the covariates.
CR&S is also related to the research in multi-arm bandit (MAB) with covariates. MAB is a classic reinforcement learning problem which can be traced back to Robbins (1952). MAB sequentially allocates limited resources to a set of competing alternatives in order to maximize the expected gain, where the gain of each alternative is unknown at first, and can better understood by allocating more resources to it. Although similar in structure, MAB differs from R&S in the objective function. The R&S problems deal with offline objective functions (i.e., the measurements are associated with only the final estimates), whereas the MAB problems concern with online objectives (i.e., the measurements are associated with the outcome of each allocation). MAB with covariates, or contextual MAB, observes a context before allocating the resources to the alternatives, and the gain of the alternative receiving resources depends on this context. Zhou (2015) provides a nice review of this field.

The rest of the paper is organized as follows. Section 3 introduces the basic notation and assumptions of this research. Section 4 develops three objective measures for the CR&S problem and studies their rate functions. Section 5 formulates the selection problem and derives the optimality conditions of it. In Section 6, a corresponding sequential selection algorithm is developed for implementation, and the convergence of it is proved. Numerical examples and computational results are provided in Section 7 followed by conclusions and discussion in Section 8.

3 Preliminaries

Suppose there are $k$ different designs. The performance of each design depends on $\mathbf{X} = (X_1, \ldots, X_d)\top$, a vector of random contexts with support $\mathcal{X} \subseteq \mathbb{R}^d$. For each design $i = 1, 2, \ldots, k$, let $Y_{il}(\mathbf{x})$ be the $l$th simulation sample from design $i$ and context $\mathbf{x}$, and $y_i(\mathbf{x})$ be the mean of this design. We have $Y_{il}(\mathbf{x}) = y_i(\mathbf{x}) + \epsilon_{il}(\mathbf{x})$, where $\epsilon_{il}(\mathbf{x})$ is the random noise incurred in the simulation. Denote $n_i(\mathbf{x})$ as the number of simulation replications for design $i$ and context $\mathbf{x}$. The sample mean $\bar{Y}_i(\mathbf{x}) = \frac{1}{n_i(\mathbf{x})} \sum_{l=1}^{n_i(\mathbf{x})} Y_{il}(\mathbf{x})$. Without loss of generality, we let the best design $i^*(\mathbf{x})$ under context $\mathbf{x}$ be the design with the smallest mean performance.

Throughout the paper, we assume that $\mathcal{X}$ contains a finite number of $m$ possible contexts $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$. This setting aligns with context space that is finite in nature. For infinite context space (continuous or discrete and unbounded, e.g., the body mass index (BMI) of the patients in
the diabetes example \cite{Bertsimas:17}), we usually do not need to find the best design for each value of them; instead, a common practice is to classify the possible values of these contexts into a number of categories/levels (e.g., classify the BMI into categories <18.5 (underweight), 18.5-24.9 (normal weight), 25.0-29.9 (overweight) and $\geq$30.0 (obesity), or more accurately, into levels <18, 18, 19, ..., 29, 30 and >30). To this end, this setting provides great flexibility in the level of contextual discrepancy we want to distinguish when formulating the problem. More importantly, this setting maintains the basic feature of a R&S problem, which is to have a finite number of entities receiving simulation replications and study how to allocate the simulation budget to them.

Suppose $n$ is our total simulation budget (number of simulation replications), and $n_{i,j}$ is the number of simulation replications we allocate to design $i$ and context $x_j$. $\alpha_{i,j} = n_{i,j}/n$ and $\alpha$ is the vector of $\alpha_{i,j}$. Let the random vector $Z_{i,j}^{(n)} = (Y_i(x_j), \tilde{Y}_{i'}(x_j))$ for $i \neq i'$. Denote $\Lambda_{i,j}(\theta) = \log \mathbb{E} [\exp(\theta Y_{il}(x_j))]$, $\Lambda_{i,j}^{(n)}(\theta) = \log \mathbb{E} [\exp(\theta \tilde{Y}_{i'}(x_j))]$ and $\Upsilon_{i,j}^{(n)}(\theta) = \log \mathbb{E} [\exp(\theta, Z_{i,j}^{(n)})]$ as the log-moment generating functions of $Y_{il}(x_j)$, $\tilde{Y}_{i'}(x_j)$ and $Z_{i,j}^{(n)}$ respectively, where $\langle \cdot, \cdot \rangle$ is the inner product, $\theta \in \mathbb{R}^2$, $i, i' \in \{1, 2, ..., k\}$ and $j \in \{1, 2, ..., m\}$. The effective domain of $\Lambda_{i,j}$ is $D_{\Lambda_{i,j}} = \{\theta \in \mathbb{R} : \Lambda_{i,j}(\theta) < \infty\}$. For any set $A$, let $A^\circ$ be its interior. Define $F_{i,j} = \{\Lambda_{i,j}^{(\gamma)}(\theta) : \theta \in D_{\Lambda_{i,j}}^\circ\}$, where $\Lambda_{i,j}^{(\gamma)}(\theta)$ is the derivative of $\Lambda_{i,j}(\theta)$.

We make the following technical assumptions in our analysis.

**ASSUMPTION 1.** The best design $i^*(x)$ is unique for all $x \in X$.

**ASSUMPTION 2.** $Y_{il}(x)$’s are independent across different $i$, $l$ and $x$.

**ASSUMPTION 3.** Function $\Lambda_{i,j}(\theta)$ is finite for any $\theta \in \mathbb{R}$, $i = 1, 2, ..., k$ and $j \in \{1, 2, ..., m\}$, and the interval $[y_{i^*(x)}(x), y_{\max}(x)] \subset \bigcap_{i=1}^k F_{i,x}^\circ$, where $y_{\max}(x) = \max_{i \in \{1, 2, ..., k\}} y_i(x)$.

Assumption 1 assumes that the best design associated with each of the $m$ contexts is unique, because two designs with the same mean cannot be distinguished. The independence between simulation samples in Assumption 2 is a standard assumption in R&S problems and makes the theoretical development tractable. Assumption 3 ensures that $Y_i(x)$ can take any value in $[y_{i^*(x)}(x), y_{\max}(x)]$ and the probability $\mathbb{P}(\tilde{Y}_i(x) < Y_{i^*(x)}(x)) > 0$ for any $x \in X$ and $i \neq i^*(x)$. In addition, Assumption 3 makes sure that the rate functions

$$I_{i,j}(\gamma) = \sup_{\theta \in \mathbb{R}} (\theta \gamma - \Lambda_{i,j}(\theta)),$$
\[ I_{i,i',j}(\gamma) = \sup_{\theta \in \mathbb{R}^2} \left( \langle \theta, \gamma \rangle - \frac{1}{n} \gamma_i^{(n)}(\theta) \right), \]

are well defined. It allows us to use the large-deviations approach to solve CR&S. Denote set \( B = \{(z_1, z_2) : z_1 \geq z_2 \text{ and } z_1, z_2 \in \mathbb{R}\} \) and function

\[ G_{i,i',j}(\alpha_{i,j}, \alpha_{i',j}) = \inf_{\gamma \in B} I_{i,i',j}(\gamma). \]

According to the Gärtner-Ellis Theorem (Dembo and Zeitouni 1998), for \( i, i' \in \{1, 2, ..., k\} \) with \( y_i(x_j) < y_i'(x_j) \) and \( y_i(x_j) > \gamma \), \( I_{i,j}(\gamma) \) and \( G_{i,i',j}(\alpha_{i,j}, \alpha_{i',j}) \) describe the rates

\[
- \alpha_{i,j} I_{i,j}(\gamma) = \lim_{n \to \infty} -\frac{1}{n} \log P(\bar{Y}_i(x_j) \leq \gamma),
- G_{i,i',j}(\alpha_{i,j}, \alpha_{i',j}) = \lim_{n \to \infty} -\frac{1}{n} \log P(\bar{Y}_i(x_j) \geq \bar{Y}_{i'}(x_j)).
\]

It was further shown in Glynn and Juneja (2004) that,

\[ G_{i,i',j}(\alpha_{i,j}, \alpha_{i',j}) = \alpha_{i,j} I_{i,j}(\gamma(\alpha_{i,j}, \alpha_{i',j})) + \alpha_{i',j} I_{i',j}(\gamma(\alpha_{i,j}, \alpha_{i',j})), \]

where \( \gamma(\alpha_{i,j}, \alpha_{i',j}) \) is the unique solution that solves \( \alpha_{i,j} I_{i,j}(\gamma) + \alpha_{i',j} I_{i',j}(\gamma) = 0 \).

4 Development of the Objective Measures

In this section, we introduce three objective measures for CR&S. Next, we analyze the rate functions of the three measures and establish their equivalence.

For context \( x \), a correct selection happens when the estimated best design \( \hat{i}^*(x) \) is identical to the real best design \( i^*(x) \). However, the correct selection can never be guaranteed in practice with a finite simulation budget. Under a fixed context \( x \), traditional R&S assesses the quality of the selection for the best design by the probability of correct selection (PCS)

\[ \text{PCS}(x) = P(\hat{i}^*(x) = i^*(x)) = P\left( \bigcap_{i=1, i \neq i^*(x)}^{k} \left( \bar{Y}_{i^*(x)}(x) < \bar{Y}_i(x) \right) \right), \]

and seeks to either maximize this probability or guarantee a pre-specified level for it. The proba-
bility here is taken with respect to the random noise of the samples from the simulation.

In CR&S, each context \(x\) is associated to a R&S problem. We want to provide the best design for all the \(x\) that might possibly appear, and therefore need measures for evaluating the quality of the selection over the entire context space \(X\). To fulfill this need, we propose the following three measures based on PCS:

\[
\text{PCS}_E = \mathbb{E}[\text{PCS}(X)] = \sum_{j=1}^{m} p_j \text{PCS}(x_j),
\]

\[
\text{PCS}_M = \min_{x \in X} \text{PCS}(x),
\]

\[
\text{PCS}_A = \mathbb{P} \left( \bigcap_{j=1}^{m} \bigcap_{i=1, i \neq i^*}^{k} \left( \bar{Y}_{i^*}(x_j) < \bar{Y}_i(x_j) \right) \right).
\]

In \(\text{PCS}_E\), \(p_j\) is the probability of \(X = x_j\), \(j = 1, 2, ..., m\). It describes the expected probability of correct selection over \(X\), where the expectation is taken with respect to the randomness of \(X\). \(\text{PCS}_M\) shows the worst-case performance of \(\text{PCS}(x)\) over \(X\). This measure is, in some sense, similar to the worst-case performance in robust optimization (Bertsimas et al. 2011) and R&S with input uncertainty (Fan et al. 2019, Gao et al. 2017). Note that \(\text{PCS}_E\) and \(\text{PCS}_M\) have also been discussed and used in Shen et al. (2017) as measures for R&S with covariates.

\(\text{PCS}_A\) is defined in a way a little bit different from the two introduced above. It is not based on \(\text{PCS}(x)\); instead, it requires correctness for all the comparisons of interest, i.e., comparisons between the estimated best design and all the alternatives under all the possible contexts. \(\text{PCS}_A\) sets the highest standard for the quality of the selection among the three, and is appropriate to be adopted by highly conservative decision makers. It is obvious to see that \(\text{PCS}_A \leq \text{PCS}_M \leq \text{PCS}_E\).

Due to the lack of analytical expressions of \(\text{PCS}_E\), \(\text{PCS}_M\) and \(\text{PCS}_A\), in this research, we follow the first stream of literature in R&S, and allocate the simulation budget \(n\) to asymptotically maximize the three measures, i.e., maximizing the rates at which they converge to 1. To do so, we first characterize the rate functions of the three measures.

**THEOREM 1.** Define probabilities of false selection \(\text{PFS}_E = 1 - \text{PCS}_E\), \(\text{PFS}_M = 1 - \text{PCS}_M\) and \(\text{PFS}_A = 1 - \text{PCS}_A\) respectively. Under Assumptions 4.2, the three measures \(\text{PFS}_E\), \(\text{PFS}_M\) and
PFS converge exponentially and have the same rate function $R(\alpha)$. That is,

$$
\lim_{n \to \infty} \frac{1}{n} \log \text{PFS}_E = \lim_{n \to \infty} \frac{1}{n} \log \text{PFS}_M = \lim_{n \to \infty} \frac{1}{n} \log \text{PFS}_A = -R(\alpha).
$$

Moreover, it can be shown that

$$
R(\alpha) = \min_{j \in \{1,2,\ldots,m\}} \min_{i \in \{1,\ldots,k\}, i \neq i^*} G_{i^*(x_j),i,j}(\alpha_{i^*(x_j),j}, \alpha_{i,j}).
$$

To interpret Theorem 1, we pick $i_o$ and $j_o$ such that

$$(i_o, j_o) \in \arg \min_{i \in \{1,\ldots,k\}, i \neq i^*} G_{i^*(x_j),i,j}(\alpha_{i^*(x_j),j}, \alpha_{i,j}).$$

The theorem shows that the three measures, though defined from different perspectives, converge in the same way as the simulation budget $n$ increases, all characterized by the most difficult comparison among the comparisons between the best design and each non-best design under each context, i.e., the comparison of the sample means between designs $i^*(x_{j_o})$ and $i_o$ under context $j_o$. The reason for this effect is that, the most difficult comparison leads to the slowest convergence rate, which dominates the convergence rates of the other comparisons, and thus represents the rate these measures converge at. The result of Theorem 1 lays the foundation of this paper: instead of considering the three measures separately, we can solve them once and for all by directly optimizing the rate function $R(\alpha)$.

5 Rate-Optimal Budget Allocation Rule

In this section, we consider the optimization of $R(\alpha)$. It can be formulated by

$$
\min \quad -R(\alpha) \\
\text{s.t.} \quad \sum_{i=1}^{k} \sum_{j=1}^{m} \alpha_{i,j} = 1, \\
\alpha_{i,j} \geq 0, \quad i = 1,2,\ldots,k, \quad j = 1,2,\ldots,m.
$$

(1)
Note that this is an OCBA-like formulation (Chen et al. 2000), which finds a simulation budget allocation strategy to optimize the measure of interest, i.e., the rate function $R(\alpha)$ in our problem (in the traditional OCBA, it is typically the PCS). The simulation budget constraint \( \sum_{i=1}^{k} \sum_{j=1}^{m} \alpha_{i,j} = 1 \) is equivalent to \( \sum_{i=1}^{k} \sum_{j=1}^{m} n_{i,j} = n \). In this research, we ignore the minor technicalities associated with $n_{i,j}$’s not being integer. In implementation, they are rounded to the nearest integers. More comprehensive discussion of this setting can be found in Chen and Lee (2011).

An equivalent formulation of problem (1) is given by

\[
\text{max } z \\
\text{s.t. } G_{i^*(x_j),i,j}(\alpha_{i^*(x_j),j},\alpha_{i,j}) \geq z, \ i = 1, 2, ..., k \text{ and } i \neq i^*(x_j), j = 1, 2, ..., m, \\
\sum_{i=1}^{k} \sum_{j=1}^{m} \alpha_{i,j} = 1, \\
\alpha_{i,j} \geq 0, \ i = 1, 2, ..., k, \ j = 1, 2, ..., m. \tag{2}
\]

As discussed in Glynn and Juneja (2004), $G_{i^*(x_j),i,j}(\alpha_{i^*(x_j),j},\alpha_{i,j})$ is a concave function, so $G_{i^*(x_j),i,j}(\alpha_{i^*(x_j),j},\alpha_{i,j}) \geq z$ forms a convex set. Therefore, problem (2) is a convex optimization model. We can investigate the KKT conditions (Boyd and Vandenberghe 2004) of this model to solve it.

**THEOREM 2.** The optimal solution of (2) satisfies

\[
\sum_{i=1, i \neq i^*(x_j)}^{k} \frac{I_{i^*(x_j),j}(\gamma(\alpha_{i^*(x_j),j},\alpha_{i,j}))}{I_{i,j}(\gamma(\alpha_{i^*(x_j),j},\alpha_{i,j}))} = 1, \quad j = 1, 2, ..., m, \tag{3}
\]

\[
G_{i^*(x_j),i,j}(\alpha_{i^*(x_j),j},\alpha_{i,j}) = G_{i^*(x_{j'}),i',j'}(\alpha_{i^*(x_{j'}),j'},\alpha_{i',j'}), \ j = 1, ..., m, \ i, i' = 1, ..., k \text{ and } i \neq i' \neq i^*(x_j), \tag{4}
\]

\[
G_{i^*(x_{j'},i,j}(\alpha_{i^*(x_{j'}),j},\alpha_{i,j}) = G_{i^*(x_{j'}),i',j'}(\alpha_{i^*(x_{j'}),j'},\alpha_{i',j'}), \ j, j' = 1, ..., m, \ i, i' = 1, ..., k, \ i \neq i^*(x_j) \text{ and } i' \neq i^*(x_{j'}). \tag{5}
\]

Theorem 2 indicates that the solution satisfying conditions (3)-(5) provides the rate-optimal budget allocation rule for the CR&S problem. Condition (3) concerns rate functions $I_{i^*(x_j),j}(\gamma(\alpha_{i^*(x_j),j},\alpha_{i,j}))$
and $I_{i,j}(\gamma(\alpha_i(x), \alpha_j))$ for $i \neq i(x)$. It establishes a certain balance between $\alpha_i(x)$ and $\alpha_j$’s, i.e., the simulation samples allocated to the best design and non-best designs for each context $x_j$. Conditions (4) and (5) further adjust the ratio of the simulation samples allocated to two non-best designs under the same context and across different contexts.

In simulation-based optimization, it is a common practice to assume that the simulation samples follow normal underlying distributions. It is a reasonable assumption because the designs are often evaluated by an average performance or batch means, so that the Central Limit Theorem effects usually hold. In this research, we also explore the budget allocation rule with normal underlying distributions.

**Theorem 3.** Under Assumptions (1)-(3) and when $Y_{il}(x)$ is normally distributed as $N(y_{i}(x), \sigma_{i}(x))$ for all $i$, $l$ and $x$, the three measures $PFS_E$, $PFS_M$ and $PFS_A$ achieve the optimal convergence rate by the following budget allocation rule

$$\frac{\sigma_{i}(x)}{\sigma_{i}(x)} = \sum_{i=1, i \neq i(x)}^{k} \frac{\alpha_{i,j}}{\sigma_{i}(x)}, \quad j = 1, 2, \ldots, m, \quad (6)$$

$$\frac{(y_{i}(x) - y_{i}(x))(x_j)^2}{\sigma_{i}(x)/\alpha_{i,x}(x) + \sigma_{i}(x)/\alpha_{i,j}} = \frac{(y_{i}(x) - y_{i}(x))(x_j)^2}{\sigma_{i}(x)/\alpha_{i,x}(x) + \sigma_{i}(x)/\alpha_{i,x'}}, \quad j = 1, \ldots, m, \quad (7)$$

$$\frac{(y_{i}(x) - y_{i}(x))(x_j)^2}{\sigma_{i}(x)/\alpha_{i,x}(x) + \sigma_{i}(x)/\alpha_{i,j}} = \frac{(y_{i}(x) - y_{i}(x))(x_{j'})^2}{\sigma_{i}(x)/\alpha_{i,x}(x) + \sigma_{i}(x)/\alpha_{i,x'}}, \quad j, j' = 1, \ldots, m, \quad (8)$$

**Remark 1.** Theorem 2 is a relatively general result which applies to not only the normal distribution, but also some other commonly-encountered light-tailed distributions. For example, equations (9)-(11) give the optimality conditions when $Y_{il}(x)$ has an exponential distribution with rate parameter $\lambda_{i,j}$.

$$\sum_{i=1, i \neq i(x)}^{k} \frac{\alpha_{i,x}(x)\lambda_{i,x}(x) + \alpha_{i,j}\lambda_{i,j}(x)}{\alpha_{i,x}(x)\lambda_{i,x}(x) + \alpha_{i,j}\lambda_{i,j}} - 1 - \log\left(\frac{\alpha_{i,x}(x)\lambda_{i,x}(x) + \alpha_{i,j}\lambda_{i,j}(x)}{\alpha_{i,x}(x)\lambda_{i,x}(x) + \alpha_{i,j}\lambda_{i,j}}\right) = 1, \quad j = 1, 2, \ldots, m, \quad (9)$$
and equations (12)-(14) give the optimality conditions when $Y_{ij}(x_j)$ has a Bernoulli distribution with success probability $q_{ij}$, $h_{ij} = q_{ij}/(1-q_{ij})$ for $i = 1, 2, \ldots, k$ and $p_{ij} = \alpha_{ij}/(\alpha_{ij} + \alpha_{ij})$ for $i = 1, 2, \ldots, k$ and $i \neq i^*(x_j)$,

\begin{align*}
j = 1, \ldots, m, \quad i, i' = 1, \ldots, k \quad \text{and} \quad i \neq i^*(x_j) \quad \text{and} \quad i' \neq i^*(x_j),
\end{align*}

\begin{align*}
&\sum_{i=1, i \neq i^*(x_j)}^k \frac{h^{i(1-p_{ij})}_{i^*(x_j), j} h_{i^*(x_j), j}^{i^*(x_j), j}}{1 + h^{i(1-p_{ij})}_{i^*(x_j), j} h_{i^*(x_j), j}^{i^*(x_j), j}} \log \frac{h^{i(1-p_{ij})}_{i^*(x_j), j} h_{i^*(x_j), j}^{i^*(x_j), j}}{1 + h^{i(1-p_{ij})}_{i^*(x_j), j} h_{i^*(x_j), j}^{i^*(x_j), j}} + \frac{1}{1 + h^{i(1-p_{ij})}_{i^*(x_j), j} h_{i^*(x_j), j}^{i^*(x_j), j}} \log \frac{1}{1 + h^{i(1-p_{ij})}_{i^*(x_j), j} h_{i^*(x_j), j}^{i^*(x_j), j}} = 1, \\
&j = 1, 2, \ldots, m,
\end{align*}

\begin{align*}
&- \alpha_{i^*(x_j), j} + \alpha_{ij}, \log (1 - q_{i^*(x_j), j})^{1-p_{ij}} (1 - q_{ij})^{p_{ij}} + q_{i^*(x_j), j}^{1-p_{ij}} q_{i^*(x_j), j}^{p_{ij}} = 0, \\
&j = 1, \ldots, m, \quad i, i' = 1, \ldots, k \quad \text{and} \quad i \neq i' \neq i^*(x_j),
\end{align*}

\begin{align*}
&\sum_{i=1, i \neq i^*(x_j)}^k \frac{h^{i(1-p_{i'j})}_{i^*(x_j), j} h_{i^*(x_j), j}^{i^*(x_j), j}}{1 + h^{i(1-p_{i'j})}_{i^*(x_j), j} h_{i^*(x_j), j}^{i^*(x_j), j}} \log \frac{h^{i(1-p_{i'j})}_{i^*(x_j), j} h_{i^*(x_j), j}^{i^*(x_j), j}}{1 + h^{i(1-p_{i'j})}_{i^*(x_j), j} h_{i^*(x_j), j}^{i^*(x_j), j}} + \frac{1}{1 + h^{i(1-p_{i'j})}_{i^*(x_j), j} h_{i^*(x_j), j}^{i^*(x_j), j}} \log \frac{1}{1 + h^{i(1-p_{i'j})}_{i^*(x_j), j} h_{i^*(x_j), j}^{i^*(x_j), j}} = 1, \\
&j = 1, 2, \ldots, m, \quad i, i' = 1, \ldots, k \quad \text{and} \quad i \neq i' \neq i^*(x_j),
\end{align*}

Despite the generality of Theorem 2, it incurs computational issues when designing selection algorithms based on it. The terms $\gamma(\alpha_{ij}, \alpha_{i'j})$ and $\mathcal{G}_{i,i'j}(\alpha_{ij}, \alpha_{i'j})$ for $i, i' \in \{1, 2, \ldots, k\}$ and $j \in \{1, 2, \ldots, m\}$ in (3)-(5) are not in closed form, which requires iterative algorithms to compute them for every time they are called in each iteration of the selection algorithm, which considerably
increases the algorithm complexity. More importantly, it was noted in Glynn and Juneja (2011) that the empirical quantities that arise in estimating rate functions $I_{i,j}(\gamma)$ and $G_{i,i',j}(\alpha_{i,j}, \alpha_{i',j})$ tend to be heavy-tailed, increasing the chance of large estimation errors. The estimation errors can be propagated to the calculation of numbers of simulation replications and the determination of the best design, and therefore compromise the performance of the selection algorithm.

To avoid these issues caused by computing rate functions, in this research, we will take full advantage of the information of the underlying distributions when designing the selection algorithm instead of treating the distributions as unknown. Particularly, we will use the result in Theorem 3 rather than Theorem 2 for algorithm development and analysis.

### 6 Selection Algorithm and its Convergence

In this section, we develop a selection algorithm based on the optimality conditions (6)-(8) for the implementation purpose, and then analyze its convergence.

Define

$$U_i^b = \frac{\alpha_{i^*(x_j),j}^2}{\sigma_{i^*(x_j)}(x_j)}, \quad j = 1, 2, ..., m,$$

$$U_{j}^{\text{non}} = \sum_{i=1,i \neq i^*(x_j)}^{k} \frac{\alpha_{i,j}^2}{\sigma_{i}^2(x_j)}, \quad j = 1, 2, ..., m,$$

$$V_{i,j} = \frac{(y_i(x_j) - y_{i^*(x_j)}(x_j))^2}{\sigma_{i^*(x_j)}^2(x_j)/\alpha_{i^*(x_j),j} + \sigma_i^2(x_j)/\alpha_{i,j}}, \quad j = 1, 2, ..., m, \quad i = 1, 2, ..., k \text{ and } i \neq i^*(x_j).$$

Conditions (6)-(8) can be re-written as

$$U_j^b = U_j^{\text{non}}, \quad j = 1, 2, ..., m, \quad (15)$$

$$V_{i,j} = V_{i',j}, \quad j = 1, ..., m, \quad i, i' = 1, ..., k \text{ and } i \neq i' \neq i^*(x_j), \quad (16)$$

$$V_{i,j} = V_{i',j'}, \quad j, j' = 1, ..., m, \quad i, i' = 1, ..., k, \quad i \neq i^*(x_j) \text{ and } i' \neq i^*(x_{j'}). \quad (17)$$

Let

$$(i_*, j_*) \in \arg \min_{j \in \{1,2,\ldots,m\}, i \in \{1,\ldots,k\} \setminus \{i^*(x_j)\}} V_{i,j}.$$
Note that

\[
\frac{dU^b_i}{\alpha_i(x_j)} = \frac{2\alpha_i(x_j)}{\sigma^2_i(x_j)} > 0, \quad j = 1, 2, \ldots, m;
\]

\[
\frac{\partial U^{non}_i}{\partial \alpha_{i,j}} = \frac{2\alpha_{i,j}}{\sigma^2_i(x_j)} > 0, \quad j = 1, 2, \ldots, m, \quad i = 1, 2, \ldots, k \text{ and } i \neq i^*(x_j);
\]

\[
\frac{\partial V_{i,j}}{\partial \alpha_i(x_j)} = \frac{(y_i(x_j) - y_{i^*(x_j)}(x_j))^2}{(\sigma^2_i(x_j)/\alpha_i(x_j)) + \sigma^2_i(x_j)/\alpha_{i,j}} > 0, \quad j = 1, 2, \ldots, m, \quad i = 1, 2, \ldots, k \text{ and } i \neq i^*(x_j).
\]

That is, we can choose to increase the values of $U^b_i$ and $U^{non}_i$ by allocating additional replications to design $i^*(x_j)$ under context $x_j$ and design $i$ for any $i \neq i^*(x_j)$ under context $x_j$, respectively. We can also choose to increase the value of $V_{i,j}$ by allocating additional replications to either design $i^*(x_j)$ or design $i$ under context $x_j$, $i \in \{1, 2, \ldots, k\}$ and $i \neq i^*(x_j)$.

To design a selection algorithm based on (15)-(17), suppose for a budget allocation, (16) or (17) cannot be fulfilled. To fix it, we will provide a small incremental budget to improve $V_{i,j}$ so that the gap between $\min_{j \in \{1, 2, \ldots, m\}} \min_{i \in \{1, \ldots, k\} \setminus \{i^*(x_j)\}} V_{i,j}$ and $\max_{j \in \{1, 2, \ldots, m\}} \max_{i \in \{1, \ldots, k\} \setminus \{i^*(x_j)\}} V_{i,j}$ can be reduced. As discussed above, allocating additional replications to design $i^*(x_{j^*})$ or $i^*$ under context $x_{j^*}$ both achieves this goal. To further decide which of designs $i^*(x_{j^*})$ and $i^*$ receives the incremental budget, we check condition (15). If $U^b_{j^*} < U^{non}_{j^*}$, the additional replications should be allocated to the best design $i^*(x_{j^*})$ in order to balance the equation; otherwise, the additional replications should be allocated to the non-best design $i^*$. This idea is summarized in the algorithm below.

**CR&S Algorithm**

1. Specify the number of contexts $m$, number of designs $k$, total simulation budget $n$ and initial number of simulation replications $n_0$. Iteration counter $r \leftarrow 0$.

2. Perform $n_0$ replications on design $i$ under context $x_j$, $i = 1, 2, \ldots, k$ and $j = 1, 2, \ldots, m$, and calculate the sample mean $\hat{Y}_i(x_j)$ and sample variance $\hat{\sigma}^2_i(x_j)$. $\hat{n}_{i,j} = n_0$, $n^{(r)} = \sum_{j=1}^{m} \sum_{i=1}^{k} \hat{n}_{i,j}$ and $\hat{\alpha}_{i,j} = \hat{n}_{i,j}/n^{(r)}$.

3. If $n^{(r)} = n$, stop. Otherwise,
a. Obtain \( \hat{\mathcal{U}}^b_j, \hat{\mathcal{U}}^{\text{non}}_j, \hat{V}_{i,j} \) and \( \hat{i}^*(x_j) \) for \( i = 1, 2, ..., k, i \neq i^*(x_j) \) and \( j = 1, 2, ..., m \). Let \( (\hat{i}^*, \hat{j}^*) \in \arg\min_{j \in \{1, 2, ..., m\}, i \in \{1, ..., k\} \setminus \{i^*(x_j)\}} \hat{V}_{i,j} \).

b. If \( \hat{\mathcal{U}}^b_j < \hat{\mathcal{U}}^{\text{non}}_j \), \( i^r = \hat{i}^*(x_{jr}) \); otherwise \( i^r = \hat{i}^* \). Provide one more replication to design \( i^r \) under context \( x_{jr} \). Update \( \hat{Y}_{i^r}(x_{jr}) \) and \( \hat{\sigma}^2_{i^r}(x_{jr}) \).

c. Update \( \hat{n}_{i,j} \), \( n^{(r+1)} \) and \( \hat{\alpha}_{i,j} \), \( r \leftarrow r + 1 \).

At the beginning of the algorithm, we simulate each pair of context and design for the same number of replications and acquire initial estimates for their means and variances. In each of the subsequent iterations, we sample more on a certain pair of context and design, determined by \( \hat{\mathcal{U}}^b_j, \hat{\mathcal{U}}^{\text{non}}_j \) and \( \hat{V}_{i,j} \), and update its sample mean and sample variance. Although we have set the incremental budget \( \Delta n = 1 \) in this generic algorithm, in practice, \( \Delta n \) can be larger than 1, say 10 or 20, to reduce the number of iterations. The algorithm terminates when the total simulation budget is exhausted.

This heuristic for designing the algorithm does not involve solving a set of nonlinear equations and is thus cost-effective; more importantly, it is a faithful reflection of conditions 15)-(17), which can be shown in the following theorem.

**THEOREM 4.** For \( \hat{\alpha}_{i,j} \) generated by the CR&S Algorithm, \( i = 1, 2, ..., k \) and \( j = 1, 2, ..., m \), we have

\[
\frac{\hat{\alpha}_{i,j}^2(x_j)}{\sigma^2_{i^*(x_j)}(x_j)} - \sum_{i = 1, i \neq i^*(x_j)}^k \frac{\hat{\alpha}_{i,j}^2}{\sigma^2_{i^*(x_j)}(x_j)} \xrightarrow{a.s.} 0, \quad j = 1, 2, ..., m, \tag{18}
\]

\[
\frac{(y_i(x_j) - y_{i^*(x_j)}(x_j))^2}{\sigma^2_{i^*(x_j)}(x_j)/\hat{\alpha}_{i^*(x_j), j} + \sigma^2_{i^*(x_j)}(x_j)/\hat{\alpha}_{i,j}} - \frac{(y_{i'}(x_j) - y_{i^*(x_j)}(x_j))^2}{\sigma^2_{i^*(x_j)}(x_j)/\hat{\alpha}_{i^*(x_j), j} + \sigma^2_{i^*(x_j)}(x_j)/\hat{\alpha}_{i', j}} \xrightarrow{a.s.} 0, \quad j = 1, ..., m, \ i, i' = 1, ..., k \text{ and } i \neq i' \neq i^*(x_j), \tag{19}
\]

\[
\frac{(y_i(x_{jr}) - y_{i^*(x_{jr})}(x_{jr}))^2}{\sigma^2_{i^*(x_{jr})}(x_{jr})/\hat{\alpha}_{i^*(x_{jr}), j} + \sigma^2_{i^*(x_{jr})}(x_{jr})/\hat{\alpha}_{i,j}} - \frac{(y_{i'}(x_{jr}) - y_{i^*(x_{jr})}(x_{jr}))^2}{\sigma^2_{i^*(x_{jr})}(x_{jr})/\hat{\alpha}_{i^*(x_{jr}), j} + \sigma^2_{i^*(x_{jr})}(x_{jr})/\hat{\alpha}_{i', j'}} \xrightarrow{a.s.} 0, \quad j, j' = 1, ..., m, \ i, i' = 1, ..., k, \ i \neq i^*(x_j) \text{ and } i' \neq i^*(x_{jr}). \tag{20}
\]

This theorem shows that \( \hat{\alpha}_{i,j} \) converges almost surely to the \( \alpha_{i,j} \) that satisfies the optimality conditions 15)-(17) (or 8)-(11). That is, the CR&S Algorithm can recover the rate-optimal budget allocation rule when improving the three objective measures. Since the original OCBA for selecting
the single best design (Chen et al. 2000) is a special case of contextual R&S, Theorem 4 can also be seen as a formalization of the theoretical behavior of the OCBA-type selection algorithms, in addition to their superior empirical performances broadly seen in numerical testing (Branke et al. 2007).

**REMARK 2.** Two byproducts can be obtained from the proof of Theorem 4. The first is the consistency of the CR&S Algorithm, a direct observation from Lemma 5 in the online supplement, where the consistency here is in the sense that

\[ \lim_{r \to \infty} \hat{n}_{i,j} = \infty, \quad i = 1, \ldots, k, \quad j = 1, \ldots, m. \]

That is, the number of simulation replications allocated to each pair of context and design by the algorithm will go to infinity as the total budget \( n \) goes to infinity. It ensures that all the estimators in this algorithm, such as \( \hat{Y}_i(x_j), \hat{\sigma}^2_i(x_j), \hat{i}^*(x_j) \), etc., will converge to their true values. The second byproduct is the convergence rate of the CR&S Algorithm. By rate of the CR&S Algorithm, we refer to the rates of the three error terms on the left-hand sides of (18)-(20) converging to 0 (distinguished from the convergence rate of the three objective measures we seek to optimize in this study). It indicates how fast the difference between the budget allocation rule outputted from the CR&S Algorithm and the rate-optimal rule vanishes. From equations (62) and (138) in the online supplement, we have

\[
\frac{\hat{\alpha}^2_i(x_j),j}{\hat{\sigma}^2_i(x_j)(x_j)} - \sum_{i=1,i \neq \hat{i}^*(x_j)}^{k} \frac{\hat{\alpha}^2_i,j}{\hat{\sigma}^2_i(x_j)} \leq O\left(\frac{\sqrt{\log \log n(r)}}{(n(r))^{\kappa_1-\frac{1}{2}}}, \quad j = 1, 2, \ldots, m, \right.
\]

\[
\left| \frac{(y_i(x_j) - y_{\hat{i}^*(x_j)}(x_j))^2}{\hat{\sigma}^2_i(x_j)} + \hat{\sigma}^2_i(x_j) + \hat{\alpha}_{i,j} \right| \leq O\left(\frac{\sqrt{\log \log n(r)}}{(n(r))^{\kappa_2-\frac{1}{2}}} \right),
\]

\[
 j = 1, \ldots, m, \quad i, \hat{i}' = 1, \ldots, k \quad \text{and} \quad \hat{i} \neq \hat{i}'(x_j),
\]

\[
\left| \frac{(y_j(x_j) - y_{\hat{i}^*(x_j)}(x_j))^2}{\hat{\sigma}^2_i(x_j)} + \hat{\sigma}^2_i(x_j) + \hat{\alpha}_{i,j} \right| \leq O\left(\frac{\sqrt{\log \log n(r)}}{(n(r))^{\kappa_2-\frac{1}{2}}} \right),
\]

\[
 j, \hat{j}' = 1, \ldots, m, \quad i, \hat{i}' = 1, \ldots, k, \quad \hat{i} \neq \hat{i}'(x_j) \quad \text{and} \quad \hat{i}' \neq \hat{i}'(x_{j'}),
\]

where \( n(r) \) is the number of simulation samples used up to iteration \( r \) of the CR&S Algorithm, and \( \kappa_1 = \kappa_2 = \frac{3}{4} \). More generally, we can set \( \kappa_i \) to any value such that \( \frac{3}{4} \leq \kappa_i < 1 \) and \( \frac{1}{1-\kappa_i} \) is an
integer, \(i = 1, 2\), an implication from Lemma 11 in online supplement.

REMARK 3. The overall proof idea of Theorem 4 is partly adapted from the Section 6 of Chen and Ryzhov (2018) which provides many useful theoretical techniques. However, our proof distinguishes from the reference in two key aspects. a) The reference relies on the knowledge of variance to make the proof more accessible which is however unrealistic for most simulation scenarios. In this proof, we are able to tackle the uncertainty in variance with extra efforts due to the simplicity of the CR&S Algorithm. b) This proof provides a framework for theoretical analysis under the existence of context and is thus more complicated but also more general for applications.

7 Numerical Experiments

In this section, we conduct three sets of numerical experiments. The first set tests the convergence rates of PFS\(_E\), PFS\(_M\) and PFS\(_A\) under different budget allocation strategies. This is a numerical presentation and verification of the theoretical result obtained in Theorem 4. The second set compares the performance the proposed selection algorithm with some existing competitors on a series of benchmark functions, and the last set compares these algorithms on three real-world applications, namely the production line optimization, the assemble-to-order problem and the personalized cancer treatment.

7.1 Test on the Convergence Rate

From Theorem 4 we have

\[
\lim_{n \to \infty} \frac{1}{n} \log \text{PFS}_E = \lim_{n \to \infty} \frac{1}{n} \log \text{PFS}_M = \lim_{n \to \infty} \frac{1}{n} \log \text{PFS}_A = -\mathcal{R}(\alpha).
\]

It suggests that

\[
\begin{align*}
\log \text{PFS}_E &= -\mathcal{R}(\alpha)n + c_E(n), \quad \text{where } \lim_{n \to \infty} \frac{c_E(n)}{n} = 0; \\
\log \text{PFS}_M &= -\mathcal{R}(\alpha)n + c_M(n), \quad \text{where } \lim_{n \to \infty} \frac{c_M(n)}{n} = 0; \\
\log \text{PFS}_A &= -\mathcal{R}(\alpha)n + c_A(n), \quad \text{where } \lim_{n \to \infty} \frac{c_A(n)}{n} = 0.
\end{align*}
\]
That is, under a fixed budget allocation strategy $\alpha$, the three log PFS terms should approximately demonstrate the same linear relationship with $n$, especially when $n$ is large. In this experiment, we numerically present this relationship. Test examples are adapted from common benchmark functions and the contexts correspond to certain parameters in the test examples. Let $x$ denote the context and $z$ denote the design (solution). We consider the following benchmark functions.

1. Rastrigin function:

   $$ Y(z, x) = f(z, x) + \epsilon(z, x) = 10d + \sum_{l=1}^{d} ((z_l - x_l)^2 - 10 \cos(2\pi(z_l - x_l))) + \epsilon(z, x). $$

   The global minimum of $f(z, x)$ is 0 obtained at $z = x$. We consider the one dimensional case ($d = 1$) of this problem with 6 contexts $x \in \{-0.75, -0.45, -0.15, 0.15, 0.45, 0.75\}$ and 10 discrete designs $z \in \{-0.90, -0.70, -0.50, \cdots, 0.90\}$. $\epsilon(z, x)$ follows the normal distribution $N(0, 121)$.

2. Sphere function:

   $$ Y(z, x) = f(z, x) + \epsilon(z, x) = \sum_{l=1}^{d} (z_l - x_l)^2 + \epsilon(z, x). $$

   The global minimum of $f(z, x)$ is 0 obtained at $z = x$. We consider the one dimensional case ($d = 1$) of this problem with 4 contexts $x \in \{-0.45, -0.15, 0.15, 0.45\}$ and 11 discrete designs $z \in \{-1.25, -1.00, -0.75, \cdots, 1.25\}$. $\epsilon(z, x)$ follows the normal distribution $N(0, 0.05)$.

3. Rosenbrock function:

   $$ Y(z, x) = f(z, x) + \epsilon(z, x) = \sum_{l=1}^{d-1} \left[ 100((z_{l+1} - x_{l+1}) - (z_l - x_l))^2 + (1 - (z_l - x_l))^2 \right] + \epsilon(z, x). $$

   The global minimum of $f(z, x)$ is 0 obtained at $z_l = x_{l+1}$, $l = 1, 2, \cdots, d$, $d > 1$. We consider the two dimensional case ($d = 2$) of this problem with 25 contexts $x \in \{-0.30, -0.15, 0, 0.15, 0.30\} \times \{-0.30, -0.15, 0, 0.15, 0.30\}$ and 9 discrete designs $z \in \{0, 0.75, 1.5\} \times \{0, 0.75, 1.5\}$. $\epsilon(z, x)$ follows the normal distribution $N(0, 2.25)$. 

19
4. McCormick function:

\[
Y(z, x) = f(z, x) + \epsilon(z, x) = \sin ((z_1 - x_1) + (z_2 - x_2)) + \left((z_1 - x_1) - (z_2 - x_2)\right)^2 \\
- 1.5(z_1 - x_1) + 2.5(z_2 - x_2) + 1 + \epsilon(z, x).
\]

The global minimum of \( f(z, x) \) is -1.9133 obtained at \( z_1 \approx x_1 - 0.55 \) and \( z_2 \approx x_2 - 1.55 \) when \(-1.5 \leq z_1 - x_1 \leq 4 \) and \(-3 \leq z_2 - x_2 \leq 4 \). We consider 3 contexts \( x \in \{(-1.2, 0), (0, 1.2), (1.2, 0)\} \) and 49 discrete designs which are the mesh grids in the space \([-1.5, 1.5] \times [-3, 0]\). \( \epsilon(z, x) \) follows the normal distribution \( N(0, 0.49) \).

Note that in these examples, we have modified the original benchmark functions \( f(z) \) to \( f(z - x) \) to incorporate context \( x \), and added a stochastic noise \( \epsilon(z, x) \) for generating random samples.

From Theorem 4, we can see that the convergence rate of \( \log \text{PFS} \) relies on the budget allocation strategy \( \alpha \). In this test, we adopt four different allocation strategies: the equal allocation and three unequal allocations. Let \( n_{i,j} \) be the number of simulation replications allocated to design \( i \) of context \( x_j \). These allocation strategies are described in the online supplement.
Figure 2: Sphere function

Figure 3: Rosenbrock function
We run $10^5$ macro-replications for each of the allocation strategies on each of the test examples and observe the performance of log PFS, shown in Figures 1, 2, 3 and 4. It is clear that log PFS_E, log PFS_M and log PFS_A are approximately linear with the same slope when $n$ is large enough for all the allocation strategies and test examples. This is in line with the conclusion in Theorem 1 that the three objective measures have the same rate function.

In addition, it can be observed that the convergence rates of the log PFS terms vary a lot among different allocation strategies on the same test example. For example, in Figure 1 the equal allocation drives log PFS_E, log PFS_M and log PFS_A to converge about two times faster than the unequal allocation strategy 2. In general, the equal allocation performs the best in the first three test examples while the unequal allocation strategy 1 performs the best in the last test example. It indicates the necessity of this research, to identify the rate-optimal allocation strategy for different structures of selection problems.
7.2 Performance Comparison on the Benchmark Functions

In this test, we compare the performance of the CR&S Algorithm with some common budget allocation strategies in the literature. We use the same four benchmark functions as in Section 7.1, and consider the following methods for comparison:

- **Equal allocation.** This method allocates the same number of simulation replications to any pair of design and context (has been used in Section 7.1).

- **Proportion to variance** (PTV). This procedure was proposed in [Rinott (1978)] for a single context setting. In this test, we extend it for multiple contexts by

  \[
  \frac{n_{i_1,j_1}}{\sigma_{i_1}^2(x_{j_1})} = \frac{n_{i_2,j_2}}{\sigma_{i_2}^2(x_{j_2})}, \quad i_1, i_2 = 1, 2, \cdots, k, \quad j_1, j_2 = 1, 2, \cdots, m.
  \]

- **Optimal computing budget allocation with equal allocation among contexts** (Equal OCBA). The original OCBA was designed for a single context and has been shown to be highly efficient for R&S problems. In this test, we make use of OCBA for designs under the same context while equally distributing the simulation budget among different contexts.

  \[
  \frac{n_{i_1,j}}{n_{i_2,j}} = \frac{\sigma_{i_1}^2(x_j)}{\sigma_{i_2}^2(x_j)} \left( y_{i_1}(x_j) - y_{i_2}(x_j) \right)^2, \quad i_1, i_2 \in \{1, 2, \cdots, k\} \setminus \{i^*(x_j)\}, \quad j = 1, 2, \cdots, m,
  \]

  \[
  n_{i^*(x_j), j} = \sigma_{i^*(x_j)}(x_j) \left( \sum_{i=1, i \neq i^*(x_j)}^{k} \left( \frac{n_{i,j}}{\sigma_i(x_j)} \right)^2 \right), \quad j = 1, 2, \cdots, m,
  \]

  \[
  \sum_{i=1}^{k} n_{i,j_1} = \sum_{i=1}^{k} n_{i,j_2}, \quad j_1, j_2 = 1, 2, \cdots, m.
  \]

We run $10^4$ macro-replications to obtain the performances of the four compared methods, and the comparison results are reported in Figures 5, 6, 7 and 8 for the Rastrigin function, sphere function, Rosenbrock function and McCormick function respectively. It can be observed that CR&S has the best performance in all the examples tested and for all the three objective measures. The convergence rate of the equal OCBA is slower than CR&S but higher than the equal allocation and PTV. On average, CR&S achieves 26.36% reduction in the simulation budget compared to the equal OCBA when reaching the empirical performance of PFS $\leq 1\%$. We notice that the
Figure 5: Comparison on the Rastrigin function

Figure 6: Comparison on the sphere function

Figure 7: Comparison on the Rosenbrock function

Figure 8: Comparison on the McCormick function
efficiency advantage of CR&S over the compared methods becomes stable when the number of initial replications is sufficiently large, e.g., when \( n_0 \geq 50 \) or the initial budget takes no less than 20% of the total budget.

7.3 Performance Comparison on the Real-World Applications

We test the CR&S Algorithm and the compared methods on a set of real-world problems.

7.3.1 Optimization of the Production Line.

This test problem can be found in the Simulation Optimization Library (http://www.simopt.org/wiki/index.php?title=Optimization_of_a_Production_Line), adapted from Buchholz and Thümmel (2005). We consider a production line consisting of \( N \) successive service queues. Each queue has a single server providing service based on the first-come-first-serve discipline. The service time of server \( l \) is exponentially distributed with rate \( \mu_l \). After being processed by a server, parts will be sent immediately to the queue of the next server unless that queue is full with \( K \) parts. Parts become completed products after going through all the \( N \) servers. Arrivals at the first queue accord to a Poisson process with rate \( x \). A design of the production line corresponds to the vector \( \mu = (\mu_1, \mu_2, \ldots, \mu_N) \). The throughput of the production line \( \varrho(\mu, x) \) is unknown but can be estimated from simulation. We would like to maximize the revenue function

\[
R(\mu, x) = \frac{h \varrho(\mu, x)}{c_0 + c^\top \mu} - c_1,
\]

where \( h, c_0, c_1 \) and \( c \) are some pre-specified parameters.

We let \( K = 10 \), and consider 3 possible contexts \( x \in \{0.3, 0.5, 0.8\} \) and 36 designs which are the mesh grids in \([0.1, 1.1] \times [0.1, 1.1] \). To determine the best design under each context, we conduct \( 4 \times 10^4 \) replications for every pair of design and context, and treat the estimated best designs as the real ones. Each of the four compared selection algorithms is performed for 2000 macro-replications with a total simulation budget of 10000 and an initial number of simulation replications \( n_0 = 8 \).

The comparison result is reported in Figure 9. The CR&S Algorithm performs the best, with the fastest convergence rate among the four selection algorithms for the three objective measures. It is worthwhile to point out that when the total simulation budget exceeds 7000, the convergence
rate of CR&S becomes a little unstable and starts to have some degree of degradation. This is caused by the relatively inaccurate estimates for the mean of some pairs of (bad) design and context when PFS is small, and can be alleviated by increasing the value of $n_0$.

7.3.2 Assemble-to-Order Problem.

Assemble-to-order (ATO) is an important problem in operations management. In a typical ATO, we have $p$ products made up from $q$ different items. Demand for each product arrives independently and follows a Poisson process with a certain arrival rate. Items are classified as key items or non-key items. A product order will be lost if there are any key items out of stock; otherwise, the product will be assembled from all the key items and available non-key items. Each product sold brings a profit, and the inventory of the items incurs a holding cost. The production time for each item is normally distributed. We want to decide the inventory levels of the items to maximize the expected total profit per unit time. In this example, we let $p = 5$ and $q = 7$, and consider 5 arrival rate contexts:

\[(1.2, 1.8, 2.4, 3.0, 3.6), \quad (3.6, 3.0, 2.4, 1.8, 1.2), \quad (2.4, 2.4, 2.4, 2.4, 2.4)\]
\[(1.8, 1.8, 1.8, 1.8, 1.8), \quad (3.0, 3.0, 3.0, 3.0, 3.0)\]

and 7 inventory level design:

\[(20, 20, 20, 20, 20), \quad (10, 10, 10, 10, 10, 10, 10), \quad (20, 20, 10, 10, 5, 5, 2, 2),\]
\[(2, 2, 2, 2, 2, 2, 2, 2), \quad (5, 5, 5, 5, 5, 5, 5, 5), \quad (2, 2, 5, 5, 10, 10, 20, 20), \quad (20, 10, 5, 2, 2, 5, 10, 20).\]
The rest of the parameter setting is identical to that in the ATO example of the Simulation Optimization Library (http://simopt.org/wiki/index.php?title=Assemble_to_order).

To determine optimal inventory levels for each context, we conduct $4 \times 10^4$ replications for every pair of design and context and treat the estimated best designs as the real ones. Each of the four compared selection algorithms is performed for 2000 macro-replications with a total simulation budget of 10000 and an initial number of simulation replications $n_0 = 100$.

The comparison result is shown in Figure 10. The CR&S Algorithm still performs the best with a significant efficiency advantage over the compared methods under the three objective measures.

### 7.3.3 Personalized Treatment for Cancer Prevention.

In this test, we compare the selection algorithms on a personalized cancer treatment problem, and focus on the esophageal adenocarcinoma (EAC), which is a main sub-type of esophageal cancer. This example has also been considered in Shen et al. (2017). Nowadays, endoscopic surveillance, aspirin chemoprevention with endoscopic surveillance and statin chemoprevention with endoscopic surveillance are three prevailing alternatives to prevent the progress from Barrett’s esophagus (BE) to EAC and death. In the aspirin and statin chemopreventions, the drug effect on individuals varies depending on the patients’ characteristics. In this example, we use a Markov chain model to simulate the state transition from BE to EAC. Parameters in the probability transition matrix are set based on Hur et al. (2004) and Choi et al. (2014).

We denote the patients’ characteristics (context) by $X = (X_1, X_2, X_3, X_4)^T$, where $X_1$ is the starting age of a treatment (an older age has a higher death rate from all-cause mortality), $X_2$ is the risk (i.e. annual progression rate of BE to EAC without chemoprevention), and the $X_3$ and
X_4 are the drug effect (i.e. progression reduction effect) of the aspirin and statin respectively. The performance of a treatment alternative is measured by quality-adjusted life years. The quality of life will have a 50% discount after the development into cancer and a 3% extra discount after surgery.

According to Hur et al. (2004), X_2 is most likely to be in [0.03, 0.07], and X_3 and X_4 in [0.3, 0.7]. We set the possible values of X_i as x_1 \in \{60, 63, 66, 69, 72\}, x_2 \in \{0.03, 0.05, 0.07\}, x_3 \in \{0.3, 0.5, 0.7\} and x_4 \in \{0.3, 0.5, 0.7\}. Moreover, we exclude the cases when x_3 = x_4 for which the only difference in effect between aspirin and statin is the complication rate, and a lot of computing resources are required to distinguish this detail. The exclusion makes sense because we can easily know which chemoprevention is better by statistics of the drugs’ complication rate for such cases instead of through simulation. Then, there are a total of 90 contexts and 3 treatment methods (designs). This is different in structure from the production line optimization and assemble-to-order problem, in which there are more designs than contexts.

We run 5 \times 10^5 replications for every pair of treatment method and context to estimate the true best design of each context. Each of the four compared selection algorithms is performed for 2000 macro-replications with a total simulation budget of 6.135 \times 10^5. The initial number of simulation replications n_0 is set to be 100. The performance comparison is shown in Figure 11. Still, the CR&S Algorithm outperforms the equal allocation, PTV and equal OCBA under all the three objective measures. The relative performances of the four compared algorithms are similar in the three real-world examples tested.

Figure 11: Comparison on the cancer treatment problem

28
8 Conclusions and Discussion

In this study, we have considered ranking and selection in the presence of context, called contextual ranking and selection (CR&S). It proposes to make use of the computing time before the problem is set up with the context input, and tries to understand the best design for each context that might appear to reduce the decision time. Different from the traditional R&S, it aims to identify a budget allocation strategy over the entire context space instead of for some particular context. To solve this problem, we start by introducing three measures for evaluating the evidence of correct selection and showing that these measures have the same convergence rate function. We next identify the rate-optimal budget allocation rule that optimizes this rate function, and develop a cost-effective selection algorithm for implementation. Last, we analyze the convergence property of the selection algorithm. A series of numerical experiments demonstrate the superior empirical performance of this algorithm.

Two potential tools can be utilized to further advance the selection efficiency for CR&S. The first is the parallel computing. CR&S is well suited for the parallel computing environment. In order for their performance to be evaluated, each pair of design and context will receive multiple independent simulation replications. This process can be easily implemented in a parallel scheme by distributing the simulation computation to independent processors, without requiring any synchronization among different processors. Main research questions along this direction include theoretical issues such as the loss of i.i.d. property of the simulation samples and the dependence among the sample sizes and sample means, and implementation issues such as the master processor being a bottleneck due to the overwhelming messages it has to process (Luo et al. 2015, Ni et al. 2017).

The other tool is the common random numbers (CRN), which is a broadly-used method for variance reduction. A design selection is based on a number of pair-wise comparisons between the sample means of different designs. If we use the same stream of random numbers to generate the simulation samples for different designs, a positive correlation will be induced between these samples, and the variance will be reduced when we compare the sample means. Since CRN will change the convergence rate of the objective measures (Fu et al. 2007), we need to identify the new rate functions and the solutions that optimize them, in order to take advantage of this tool.
A Proof of Theorem \[1\]

We first present two lemmas and then show that the three measures have the same convergence rate.

**Lemma 1.** (Principle of the slowest term (Ganesh et al. [2004])) Consider positive sequences \(a_j(n), j = 1, 2, ..., m\). If \(\lim_{n \to \infty} \frac{1}{n} \log a_j(n)\) exists for all \(j\), then \(\lim_{n \to \infty} \frac{1}{n} \log(\sum_{j=1}^{m} a_j(n)) = \max_{j \in \{1, ..., m\}} (\lim_{n \to \infty} \frac{1}{n} \log a_j(n)).\)

**Lemma 2.** (Hunter and Pasupathy [2013]) Consider positive sequences \(a_i(n), i = 1, 2, ..., k\). If \(\lim_{n \to \infty} \frac{1}{n} \log a_i(n)\) exists for all \(i\), then \(\max_{i \in \{1, ..., k\}} (\lim_{n \to \infty} \frac{1}{n} \log a_i(n)) = \lim_{n \to \infty} \frac{1}{n} \log(\max_{i \in \{1, ..., k\}} a_i(n)).\)

We first consider \(\text{PFS}_E\).

\[
\lim_{n \to \infty} \frac{1}{n} \log \text{PFS}_E = \lim_{n \to \infty} \frac{1}{n} \log \left(\sum_{j=1}^{m} p_j \text{PFS}(x_j)\right)
\]

\[
= \lim_{n \to \infty} \frac{1}{n} \log \left(\sum_{j=1}^{m} p_j \mathbb{P}\left(\bigcup_{i=1, i \neq i^*(x_j)}^{k} \left(\bar{Y}_{i^*(x_j)}(x_j) \geq \bar{Y}_i(x_j)\right)\right)\right)
\]

\[
= \max_{j \in \{1, 2, ..., m\}} \lim_{n \to \infty} \frac{1}{n} \log \mathbb{P}\left(\bigcup_{i=1, i \neq i^*(x_j)}^{k} \left(\bar{Y}_{i^*(x_j)}(x_j) \geq \bar{Y}_i(x_j)\right)\right).
\]

The penultimate step is from Lemma 1. For the term \(\lim_{n \to \infty} \frac{1}{n} \log \mathbb{P}(\bigcup_{i=1, i \neq i^*(x_j)}^{k} \bar{Y}_{i^*(x_j)}(x_j) \geq \bar{Y}_i(x_j))\) in (21), note that

\[
\max_{i \in \{1, ..., k\}, i \neq i^*(x_j)} \mathbb{P}\left(\bar{Y}_{i^*(x_j)}(x_j) \geq \bar{Y}_i(x_j)\right) \leq \mathbb{P}\left(\bigcup_{i=1, i \neq i^*(x_j)}^{k} \left(\bar{Y}_{i^*(x_j)}(x_j) \geq \bar{Y}_i(x_j)\right)\right)
\]

\[
\leq (k - 1) \max_{i \in \{1, ..., k\}, i \neq i^*(x_j)} \mathbb{P}\left(\bar{Y}_{i^*(x_j)}(x_j) \geq \bar{Y}_i(x_j)\right),
\]

and that

\[
\lim_{n \to \infty} \frac{1}{n} \log \left((k - 1) \max_{i \in \{1, ..., k\}, i \neq i^*(x_j)} \mathbb{P}\left(\bar{Y}_{i^*(x_j)}(x_j) \geq \bar{Y}_i(x_j)\right)\right)
\]
\[
\begin{align*}
&= \lim_{n \to \infty} \frac{1}{n} \log \max_{i \in \{1, \ldots, k\}, i \neq i^*} \mathbb{P}\left( \bar{Y}_{i^*}(x_j) \geq \bar{Y}_i(x_j) \right) \\
&= \max_{i \in \{1, \ldots, k\}, i \neq i^*} \lim_{n \to \infty} \frac{1}{n} \log \mathbb{P}\left( \bar{Y}_{i^*}(x_j) \geq \bar{Y}_i(x_j) \right) \\
&= \max_{i \in \{1, \ldots, k\}, i \neq i^*} - \mathcal{G}_{i^*}(x_j), i, j (\alpha_{i^*}(x_j), j, \alpha_{i, j}) \\
&= - \min_{i \in \{1, \ldots, k\}, i \neq i^*} \mathcal{G}_{i^*}(x_j), i, j (\alpha_{i^*}(x_j), j, \alpha_{i, j}),
\end{align*}
\]

and then
\[
\lim_{n \to \infty} \frac{1}{n} \log \mathbb{P}\left( \bigcup_{i=1, i \neq i^*}^{k} \left( \bar{Y}_{i^*}(x_j) \geq \bar{Y}_i(x_j) \right) \right) = - \min_{i \in \{1, \ldots, k\}, i \neq i^*} \mathcal{G}_{i^*}(x_j), i, j (\alpha_{i^*}(x_j), j, \alpha_{i, j}).
\]

Combine (21) and (22), and we have
\[
\lim_{n \to \infty} \frac{1}{n} \log \text{PFS}_E = - \min_{j \in \{1, 2, \ldots, m\}} \min_{i \in \{1, \ldots, k\}, i \neq i^*} \mathcal{G}_{i^*}(x_j), i, j (\alpha_{i^*}(x_j), j, \alpha_{i, j}).
\]

We next consider PFS$_M$.
\[
\lim_{n \to \infty} \frac{1}{n} \log \text{PFS}_M = \lim_{n \to \infty} \frac{1}{n} \log \left( \max_{j \in \{1, 2, \ldots, m\}} \text{PFS}(x_j) \right) \\
= \lim_{n \to \infty} \frac{1}{n} \log \left( \max_{j \in \{1, 2, \ldots, m\}} \mathbb{P}\left( \bigcup_{i=1, i \neq i^*}^{k} \left( \bar{Y}_{i^*}(x_j) \geq \bar{Y}_i(x_j) \right) \right) \right) \\
= \max_{j \in \{1, 2, \ldots, m\}} \lim_{n \to \infty} \frac{1}{n} \log \mathbb{P}\left( \bigcup_{i=1, i \neq i^*}^{k} \left( \bar{Y}_{i^*}(x_j) \geq \bar{Y}_i(x_j) \right) \right) \\
= \max_{j \in \{1, 2, \ldots, m\}} \left( - \min_{i \in \{1, \ldots, k\}, i \neq i^*} \mathcal{G}_{i^*}(x_j), i, j (\alpha_{i^*}(x_j), j, \alpha_{i, j}) \right) \\
= - \min_{j \in \{1, 2, \ldots, m\}} \min_{i \in \{1, \ldots, k\}, i \neq i^*} \mathcal{G}_{i^*}(x_j), i, j (\alpha_{i^*}(x_j), j, \alpha_{i, j}),
\]

where the second step is from Lemma 2 and the third step is from (22).

Last, we consider PFS$_A$, given by
\[
\text{PFS}_A = \mathbb{P}\left( \bigcup_{j=1}^{m} \bigcup_{i=1, i \neq i^*}^{k} \left( \bar{Y}_{i^*}(x_j) \geq \bar{Y}_i(x_j) \right) \right).
\]
Similarly as in the analysis for PFS_E,

\[
\max_{j \in \{1, 2, \ldots, m\}} \max_{i \in \{1, \ldots, k\}, i \neq i^*(x_j)} P\left(\tilde{Y}^*_i(x_j) \geq \tilde{Y}^*_i(x_j)\right) \leq P\left(\bigcup_{j=1}^{m} \bigcup_{i=1, j \neq i^*(x_j)} \left(\tilde{Y}^*_i(x_j) \geq \tilde{Y}^*_i(x_j)\right)\right)
\]

\[
\leq m(k-1) \max_{j \in \{1, 2, \ldots, m\}} \max_{i \in \{1, \ldots, k\}, i \neq i^*(x_j)} P\left(\tilde{Y}^*_i(x_j) \geq \tilde{Y}^*_i(x_j)\right).
\]

Note that

\[
\lim_{n \to \infty} \frac{1}{n} \log \left( \max_{j \in \{1, 2, \ldots, m\}} \max_{i \in \{1, \ldots, k\}, i \neq i^*(x_j)} P\left(\tilde{Y}^*_i(x_j) \geq \tilde{Y}^*_i(x_j)\right) \right)
\]

\[
= \lim_{n \to \infty} \frac{1}{n} \log \left( \max_{j \in \{1, 2, \ldots, m\}} \max_{i \in \{1, \ldots, k\}, i \neq i^*(x_j)} P\left(\tilde{Y}^*_i(x_j) \geq \tilde{Y}^*_i(x_j)\right) \right)
\]

\[
= \max_{j \in \{1, 2, \ldots, m\}} \max_{i \in \{1, \ldots, k\}, i \neq i^*(x_j)} \lim_{n \to \infty} \frac{1}{n} \log P\left(\tilde{Y}^*_i(x_j) \geq \tilde{Y}^*_i(x_j)\right)
\]

\[
= \max_{j \in \{1, 2, \ldots, m\}} \max_{i \in \{1, \ldots, k\}, i \neq i^*(x_j)} \left( -G_{i^*, (j), i, j}(\alpha_{i^*(x_j), j}, \alpha_{i, j}) \right)
\]

\[
= - \min_{j \in \{1, 2, \ldots, m\}} \min_{i \in \{1, \ldots, k\}, i \neq i^*(x_j)} G_{i^*, (j), i, j}(\alpha_{i^*(x_j), j}, \alpha_{i, j}),
\]

where the second step is from Lemma 2. Then,

\[
\lim_{n \to \infty} \frac{1}{n} \log P\left(\bigcup_{j=1}^{m} \bigcup_{i=1, j \neq i^*(x_j)} \left(\tilde{Y}^*_i(x_j) \geq \tilde{Y}^*_i(x_j)\right)\right) = - \min_{j \in \{1, 2, \ldots, m\}} \min_{i \in \{1, \ldots, k\}, i \neq i^*(x_j)} G_{i^*, (j), i, j}(\alpha_{i^*(x_j), j}, \alpha_{i, j}).
\]

(24)

B Proof of Theorem 2

According to the KKT conditions, there exist constants \(\theta\) and \(\lambda_{i,j}\) for \(j = 1, 2, \ldots, m, i = 1, 2, \ldots, k\) and \(i \neq i^*(x_j)\) such that

\[
1 - \sum_{j=1}^{m} \sum_{i=1, i \neq i^*(x_j)}^{k} \lambda_{i,j} = 0,
\]

(25)

\[
\sum_{i=1, i \neq i^*(x_j)}^{k} \lambda_{i,j} \frac{\partial G_{i^*, (j), i, j}(\alpha_{i^*(x_j), j}, \alpha_{i, j})}{\partial \alpha_{i^*(x_j), j}} = \theta, \quad j = 1, 2, \ldots, m,
\]

(26)

\[
\lambda_{i,j} \frac{\partial G_{i^*, (j), i, j}(\alpha_{i^*(x_j), j}, \alpha_{i, j})}{\partial \alpha_{i, j}} = \theta, \quad j = 1, 2, \ldots, m, i = 1, 2, \ldots, k \text{ and } i \neq i^*(x_j),
\]

(27)
\[ \lambda_{i,j}(\gamma_{i^*(x_j)}) = 0, \quad j = 1, 2, ..., m, \quad i = 1, 2, ..., k \text{ and } i \neq i^*(x_j). \] (28)

From [26], all the \( \lambda_{i,j} \)'s are non-positive or non-negative at the same time, and from [25], \( \lambda_{i,j} \geq 0 \) for \( j = 1, 2, ..., m, \quad i = 1, 2, ..., k \) and \( i \neq i^*(x_j) \). If we assume that there exists some \( j \in \{1, 2, ..., m\} \) and \( i \in \{1, 2, ..., k\} \setminus \{i^*(x_j)\} \) such that \( \lambda_{i,j} = 0 \), from [27], \( \theta = 0 \), and then all the \( \lambda_{i,j} \)'s are equal to 0. This is a contradiction to [25]. As a result, \( \lambda_{i,j} > 0 \) for \( j = 1, 2, ..., m, \quad i = 1, 2, ..., k \) and \( i \neq i^*(x_j) \). From [28], \( \gamma_{i^*(x_j),i,j}(\alpha_{i^*(x_j)}, \alpha_{i,j}) = z \), and results (4) and (5) can be concluded.

Next, from [27], \( \lambda_{i,j} = \frac{\sigma_i^2}{\partial \gamma_{i^*(x_j),i,j}(\alpha_{i^*(x_j)}, \alpha_{i,j})/\partial \alpha_{i,j}} \). Substitute it into [26],

\[ \sum_{i = 1, i \neq i^*(x_j)}^k \frac{\partial \gamma_{i^*(x_j),i,j}(\alpha_{i^*(x_j)}, \alpha_{i,j})}{\partial \gamma_{i^*(x_j),i,j}(\alpha_{i^*(x_j)}, \alpha_{i,j})/\partial \alpha_{i,j}} = 1, \quad j = 1, 2, ..., m. \] (29)

Since

\[ \frac{\partial \gamma_{i^*(x_j),i,j}(\alpha_{i^*(x_j)}, \alpha_{i,j})}{\partial \alpha_{i^*(x_j),j}} = \gamma_{i^*(x_j),j}(\gamma(\alpha_{i^*(x_j)}, \alpha_{i,j})) + \alpha_{i^*(x_j),j}\gamma_{i^*(x_j),j}(\gamma(\alpha_{i^*(x_j)}, \alpha_{i,j})) \frac{\partial \gamma(\alpha_{i^*(x_j)}, \alpha_{i,j})}{\partial \alpha_{i^*(x_j),j}} \]

\[ + \alpha_{i,j}\gamma_{i,j}(\gamma(\alpha_{i^*(x_j)}, \alpha_{i,j})) \frac{\partial \gamma(\alpha_{i^*(x_j)}, \alpha_{i,j})}{\partial \alpha_{i^*(x_j),j}} \]

\[ = \gamma_{i^*(x_j),j}(\gamma(\alpha_{i^*(x_j)}, \alpha_{i,j})), \]

\[ \frac{\partial \gamma_{i^*(x_j),i,j}(\alpha_{i^*(x_j)}, \alpha_{i,j})}{\partial \alpha_{i,j}} = \gamma_{i,j}(\gamma(\alpha_{i^*(x_j)}, \alpha_{i,j})) + \gamma_{i,j}(\gamma(\alpha_{i^*(x_j)}, \alpha_{i,j})) \frac{\partial \gamma(\alpha_{i^*(x_j)}, \alpha_{i,j})}{\partial \alpha_{i,j}} \]

\[ + \alpha_{i,j}\gamma_{i,j}(\gamma(\alpha_{i^*(x_j)}, \alpha_{i,j})) \frac{\partial \gamma(\alpha_{i^*(x_j)}, \alpha_{i,j})}{\partial \alpha_{i,j}} \]

\[ = \gamma_{i,j}(\gamma(\alpha_{i^*(x_j)}, \alpha_{i,j})), \]

result (3) follows from (29).

C Proof of Theorem 3

It was shown in Dembo and Zeitouni (1998) that for the normal distribution,

\[ \mathcal{I}_{i,j}(\gamma) = \frac{(\gamma - \gamma_i(x_j))^2}{2\sigma_i^2(x_j)} \]
Then, it can be derived that

\[ G_{t^*(x_j),i,j}(\alpha_{t^*(x_j),j}, \alpha_{i,j}) = \frac{(y_i(x_j) - y_{t^*(x_j)}(x_j))^2}{2(\sigma_t^2(x_j)/\alpha_{t^*(x_j),j} + \sigma_i^2(x_j)/\alpha_{i,j})}, \quad (30) \]

\[ I_{t^*(x_j),j}(\gamma(\alpha_{t^*(x_j),j}, \alpha_{i,j})) = \frac{(y_i(x_j) - y_{t^*(x_j)}(x_j))^2\sigma_t^2(x_j)}{2\alpha_{i,j}^2(\sigma_t^2(x_j)/\alpha_{t^*(x_j),j} + \sigma_i^2(x_j)/\alpha_{i,j})^2}, \quad (31) \]

\[ I_{i,j}(\gamma(\alpha_{t^*(x_j),j}, \alpha_{i,j})) = \frac{(y_i(x_j) - y_{t^*(x_j)}(x_j))^2\sigma_i^2(x_j)}{2\alpha_{i,j}^2(\sigma_i^2(x_j)/\alpha_{t^*(x_j),j} + \sigma_i^2(x_j)/\alpha_{i,j})^2}. \quad (32) \]

Take (31) and (32) into (3),

\[ \sum_{i=1, i \neq t^*(x_j)}^{k} \frac{\sigma_i^2(x_j)x_j^2}{\sigma_t^2(x_j)} = 1, \quad j = 1, 2, ..., m. \]

That is,

\[ \frac{\sigma_i^2(x_j)x_j^2}{\sigma_t^2(x_j)} = \sum_{i=1, i \neq t^*(x_j)}^{k} \frac{\alpha_i^2}{\sigma_t^2(x_j)}, \quad j = 1, 2, ..., m \]

as in (6). Take (30) into (4) and (5), and (7) and (8) follow.

**D Proof of the Statements in Remark 1**

It has been shown in Glynn and Juneja (2004) that for exponential underlying distributions,

\[ I_{i,j}(\gamma) = \lambda_{i,j}\gamma - 1 - \log(\lambda_{i,j}\gamma), \quad (33) \]

and thus,

\[ \gamma(\alpha_{t^*(x_j),j}, \alpha_{i,j}) = \frac{\alpha_{t^*(x_j),j} + \alpha_{i,j}}{\alpha_{t^*(x_j),j} + \alpha_{i,j} \lambda_{t^*(x_j),j},} \quad (34) \]

\[ G_{i,j}(\alpha_{t^*(x_j),j}, \alpha_{i,j}) = -\alpha_{t^*(x_j),j} \log \frac{\lambda_{t^*(x_j),j}(\alpha_{t^*(x_j),j} + \alpha_{i,j})}{\alpha_{t^*(x_j),j} + \alpha_{i,j} \lambda_{t^*(x_j),j}} - \alpha_{i,j} \log \frac{\lambda_{i,j}(\alpha_{t^*(x_j),j} + \alpha_{i,j})}{\alpha_{t^*(x_j),j} + \alpha_{i,j} \lambda_{t^*(x_j),j}}, \quad (35) \]

Take (33)-(35) into (3)-(5), and the results in (9)-(11) follow.
Similarly for Bernoulli underlying distributions,

\[ I_{i,j}(\gamma) = \gamma \log \frac{\gamma}{q_{i,j}} + (1 - \gamma) \log \frac{1 - \gamma}{1 - q_{i,j}}, \tag{36} \]

and thus,

\[ \gamma(\alpha_{i^*(x_j),j}, \gamma_{i,j}) = \frac{h_{i^*(x_j),i}^*(\gamma_{i,j})}{1 + h_{i^*(x_j),i}^*(\gamma_{i,j})}, \tag{37} \]

\[ G_{i^*(x_j),i,j}(\alpha_{i^*(x_j),j}, \gamma_{i,j}) = -(\alpha_{i^*(x_j),j} + \gamma_{i,j}) \log \left( (1 - q_{i^*(x_j),j})(1 - q_{i,j})\rho_{i,j} + q_{i^*(x_j),j}^*(\gamma_{i,j}) \right). \tag{38} \]

Take (36)-(38) into (3)-(5), and the results in (12)-(14) follow.

**E Proof of Theorem 4**

Before introducing the proof, a subtle change in notation should be aware of. We will bracket \(i\) and \(j\) into \((i,j)\) and append another subscript \(r\), to indicate the iteration where notations come from. For example, \(\hat{\alpha}_{i,j} \rightarrow \hat{\alpha}_{(i,j),r}\). However, without influence of understanding, we omit the \(r\) occasionally for sake of simplicity.

Let \(I_{r}^{(i,j)}\) be the indicator function of design \(i\) of \(x_j\) at iteration \(r\). That is to say,

\[ I_{r}^{(i,j)} = \begin{cases} 1, & (i,j) \text{ is sampled at iteration } r, \\ 0, & (i,j) \text{ is not sampled at iteration } r. \end{cases} \]

Denote the mean and variance estimator by \(\tilde{Y}_{i}(x_j)\), \(\tilde{\sigma}^2_{i}(x_j)\) respectively.

Notations frequently appeared are listed as follows.

\[ \hat{i}_{r}^*(x_j) = \min_{i=1,2,\ldots,k} \tilde{Y}_{i}(x_j), \quad \hat{U}^{b}_{(j),r} = \frac{\hat{\alpha}_{i^*(x_j),j,r}^2}{\hat{\sigma}^2_{i^*(x_j)}(x_j)}, \quad j = 1, 2, \ldots, m; \]

\[ \hat{U}^{nom}_{(j),r} = \sum_{i \neq i^*(x_j)} \frac{\hat{\alpha}_{i,r}^2}{\hat{\sigma}^2_{i}(x_j)}, \quad S^{b}_{(j),r} = \frac{\hat{\sigma}^2_{i^*(x_j)}(x_j)}{\hat{n}_{i^*(x_j),j,r}}, \quad j = 1, 2, \ldots, m; \]

\[ \hat{\delta}_{(i,j)}^r = (\tilde{Y}_{i}(x_j) - \tilde{Y}_{i^*(x_j)}(x_j))^2, \quad S_{(i,j),r} = \tilde{\sigma}^2_{i}(x_j)/\hat{n}_{(i,j),r}, \]

35
\[
\hat{\tau}_{(i,j),r} = \frac{\hat{\delta}_{(i,j)}}{S^{b}_{(j),r} + S_{(i,j),r}}, \quad \hat{V}_{(i,j),r} = \frac{\hat{\tau}_{(i,j),r}}{n^{(r)}}, \quad j = 1, 2, \cdots, m, \quad i = 1, 2, \cdots, k, i \neq \hat{i}^{*}(x_{j}).
\]

Next, we will dedicate to prove the asymptotic property of the algorithm. In process of derivation, we got a lot of inspirations from [Chen and Ryzhov (2018)]. We will separate the conclusion of Theorem 4 into two parts as follows.

**THEOREM 5.** Under our algorithm,

\[
\lim_{r \to \infty} \left[ \left( \frac{\hat{\alpha}_{(i^{*}(x_{j}),j)}}{\sigma_{i^{*}(x_{j})}(x_{j})} \right)^{2} - \sum_{i \neq i^{*}(x_{j})} \left( \frac{\hat{\alpha}_{(i,j)}}{\sigma_{i}(x_{j})} \right)^{2} \right] = 0 \tag{39}
\]

almost surely, for any \( j = 1, 2, \cdots, m \).

**THEOREM 6.** Under our algorithm,

\[
\lim_{r \to \infty} \left[ \frac{(y_{i_{1}}(x_{j_{1}}) - y_{i^{*}(x_{j_{1}})}(x_{j_{1}}))^{2}}{\sigma^{2}_{i^{*}(x_{j_{1}})}(x_{j_{1}})/\hat{\alpha}_{(i^{*}(x_{j_{1}}),j_{1})}} + \sigma^{2}_{i_{1}}(x_{j_{1}})/\hat{\alpha}_{(i_{1},j_{1})}} - \frac{(y_{i_{2}}(x_{j_{2}}) - y_{i^{*}(x_{j_{2}})}(x_{j_{2}}))^{2}}{\sigma^{2}_{i^{*}(x_{j_{2}})}(x_{j_{2}})/\hat{\alpha}_{(i^{*}(x_{j_{2}}),j_{2})}} + \sigma^{2}_{i_{2}}(x_{j_{2}})/\hat{\alpha}_{(i_{2},j_{2})}} \right] = 0
\]

almost surely, for any \( j_{1}, j_{2} = 1, 2, \cdots, m \) and \( i_{1}, i_{2} = 1, 2, \cdots, k, i_{1} \neq i^{*}(x_{j_{1}}), i_{2} \neq i^{*}(x_{j_{2}}) \).

Moreover, without loss of generality, we assume \( \Delta n = 1 \). This makes sense because the samples generated in one iteration can be viewed as one sample with less variance when \( \Delta n > 1 \).

Among all of our conclusions, lemma 3, 4 and 5 are simple but lay foundations for the rest derivation. Lemma 6, 7 and 10 deal with the relative relationship of \( \hat{n}_{(i,j),r} \) inside the subsequence where \( x_{j} \) is considered. Meanwhile, lemma 8, 9 and 13 focus on how the algorithm would allocate computing resources among different contexts \( x_{j} \). Lemma 11 illustrates a property of variance estimator. Then, Theorem 5 comes out naturally.

Lemma 14 and 16 are preparation for 15 and 17 respectively. Combining lemma 15 and 17 with 12, which indicates the increasing property of \( \hat{V}_{(i,j),r} \), we can get Theorem 6.

**LEMMA 3.** Let \( n^{(r)}_{(j)} = \sum_{i=1}^{k} \hat{n}_{(i,j),r}, j = 1, 2, \cdots, m \). For context \( x_{j} \), if \( n^{(r)}_{(j)} \to \infty \), we have \( \hat{\tau}_{(i,j)} \to \infty \) almost surely for all \( i = 1, 2, \cdots, k \) as \( r \to \infty \).

**Proof.** For context \( x_{j} \) and a fixed sample path \( \omega \), define \( A = \{ i | \hat{n}_{(i,j),r} \to \infty \} \). Because \( n^{(r)}_{(j)} \to \infty \), it is obvious that \( A \) is non-empty. Suppose \( A^{c} \) is also non-empty and \( i_{1} \in A^{c} \), denote \( K_{1} \) to be the last time that \( (i_{1}, j) \) is sampled. It means \( K_{1} = \sup \{ r | i_{1}^{r} = i_{1} \} \).

36
Since $\hat{\sigma}_i^2(x_j)$ is almost surely convergent, we could find an upper bound $b_{\text{var}}$ such that $\hat{\sigma}_i^2(x_j) < b_{\text{var}}$ for all $i = 1, 2, \cdots, k$. Then, as $n^{(r)}_{(j)}$ increases, there must exist a finite time $K_2 > K_1$ such that

$$\frac{\hat{n}^2_{(i_1,j),r}}{\hat{\sigma}_i^2(x_j)} < \frac{1}{b_{\text{var}}} \sum_{i \neq i_1(x_j)} \hat{n}^2_{(i,j),r} < \frac{1}{b_{\text{var}}} \sum_{i \neq i_1(x_j)} \hat{n}^2_{(i,j),r} \hat{\sigma}_i^2(x_j)$$

holds when $r > K_2$. For all $r > K_2$, we have $\hat{\tau}_r(x_j) \neq i_1$. If this was not the case, since $n^{(r)}_{(j)} \to \infty$, we would be able to find some iterations $r > K_2$ where context $x_j$ is considered and $(i^r, j^r) = (i_1, j)$. This contradicts the definition of $K_1$.

The analysis above applies to any $i_1 \in A^c$. Because $A^c$ contains just a finite number of element, thus $K_3$ could be found such that both $i^r \in A$ and $\hat{\tau}_r(x_j) \in A$ for all $r > K_3$, where context $x_j$ is considered.

By the definition of $A$, we have

$$\hat{\tau}_{(i,j)} = \frac{(\hat{Y}_i(x_j) - \hat{Y}_{i^*}(x_j)(x_j))^2}{\hat{\sigma}_i^2(x_j)/\hat{n}^2_{(i^*,j),r} + \hat{\sigma}_i^2(x_j)/\hat{n}_{(i,j),r}} \to \infty,$$

because $\hat{n}_{(i^*_r(x_j),j),r} \to \infty$.  

**Lemma 4.** For all $x_j, j = 1, 2, \cdots, m$, we have

$$\lim_{r \to \infty} n^{(r)}_{(j)} = \infty$$

almost surely.

**Proof.** Fix a sample path $\omega$ and define $B = \{j | \lim_{r \to \infty} n^{(r)}_{(j)} = \infty\}$. Obviously, $B$ is non-empty. If $B^c$ is also non-empty, for every $j_1 \in B^c$, let

$$L_1 = \sup\{r | j^r = j_1\}.$$

Then context $x_{j_1}$ would not be considered when $r > L_1$. Thus, $\hat{\tau}_{(i,j_1),r}$ are same for $r > L_1, i = 1, 2, \cdots, k$.

Because $B^c$ has finite number of elements, we could find $L_2 \geq L_1$ such that, for any $j \in B^c$, contexts $x_j$ would not be considered when $r > L_2$. Then, a positive constant $b_u$ could be found
such that
\[
\max_{j \in A, i = 1, 2, \ldots, k} \tilde{\tau}(i, j) < b^a.
\]

However, by lemma \[3\] \(\tilde{\tau}(i, j) \to \infty\) as \(r \to \infty\) if \(j \in B\) and \(i = 1, 2, \ldots, k\). From here, it is straightforward to show that we could find \(L_3 > L_2\) such that
\[
\min_{j \in A, i = 1, 2, \ldots, k} \tilde{\tau}(i, j) > \max_{j \in A, i = 1, 2, \ldots, k} \tilde{\tau}(i, j),
\]
for \(r > L_3\). By then, we have to consider the contexts \(x_j\) where \(j \in B^c\). This contradicts the definition of \(L_2\). We conclude that \(B^c = \phi\).

\[\Box\]

**Lemma 5.** For all \(x_j, j = 1, 2, \ldots, m\) and \(i = 1, 2, \ldots, k\), we have
\[
\lim_{r \to \infty} \hat{n}(i, j)_r = \infty
\]
almost surely.

**Proof.** For context \(x_j\), fix a sample path and define \(D = \{i| \lim_{r \to \infty} \hat{n}(i, j)_r = \infty\}\). Suppose \(D^c\) is non-empty, \(i_1 \in D^c\) and \(i_0 \in D\). Similarly, we could find \(H_1\) such that any \(i_1 \in D^c\) would not be sampled when \(r \geq H_1\). Then, when \(r \geq H_1\), we have
\[
\tilde{\tau}(i_0, j) - \tilde{\tau}(i_1, j) = \frac{\hat{\delta}(i_0, j)}{S^b(j)_r + S(i_0, j)_r} - \frac{\hat{\delta}(i_1, j)}{S^b(j)_r + S(i_1, j)_r} = \frac{(\hat{\delta}(i_0, j) - \hat{\delta}(i_1, j)) \hat{\delta}^2 v^2(x_j)(x_j) / \hat{n}(i_0, j)_r - \hat{\delta}(i_1, j) \hat{\delta}^2 v^2(x_j) / \hat{n}(i_1, j)_r}{(S^b(j)_r + S(i_0, j)_r)(S^b(j)_r + S(i_1, j)_r)} > \frac{b_{v_1} / \hat{n}(i_0, j)_r + b_{v_2} / \hat{n}(i_1, j)_r - b_{v_3} / \hat{n}(i_0, j)_r}{(S^b(j)_r + S(i_0, j)_r)(S^b(j)_r + S(i_1, j)_r)},
\]
where \(b_{v_1}, b_{v_2}\) and \(b_{v_3}\) are lower bound of \((\hat{\delta}(i_0, j) - \hat{\delta}(i_1, j)) \hat{\delta}^2 v^2(x_j)(x_j), \hat{\delta}(i_0, j) \hat{\delta}^2 v^2(x_j)\) and upper bound of \(\hat{\delta}(i_1, j) \hat{\delta}^2 v^2(x_j)\) respectively. Moreover, \(b_{v_2} > 0\) and \(b_{v_3} > 0\).

From proof of lemma \[3\] and definition of \(i_0\), we know that both \(\hat{n}(i_0, j)_r\) and \(\hat{n}(i_1, j)_r\) go to \(\infty\)

38
as \( r \) increases. Then, \( H_2 > H_1 \) could be found such that, when \( r > H_2 \),

\[
\frac{|b_{v_1}|}{\hat{n}(\hat{i}_i^*(x_j),r)} < \frac{b_{v_2}}{2\hat{n}(i,j),H_1} = \frac{b_{v_2}}{2\hat{n}(i,j),r}, \quad \frac{b_{v_3}}{\hat{n}(i_0,j),r} < \frac{b_{v_2}}{2\hat{n}(i_0,j),H_1} = \frac{b_{v_2}}{2\hat{n}(i_0,j),r}.
\]

It means \( \hat{r}_{(i_0,j)} - \hat{r}_{(i_1,j)} > 0 \) when \( r > H_2 \). Because of the finite number of designs, we could find \( H_3 \) such that when \( r > H_3 \),

\[
\max_{i_1 \in D^c} \hat{r}_{(i_1,j)} < \min_{i_0 \in D} \hat{r}_{(i_0,j)}.
\]

By lemma 4, \( \hat{n}_{(j)}(r) \to \infty \). That is to say that context \( x_j \) will be considered for infinitely many times. However, based on the criteria of our algorithm, when \( r > H_3 \), context \( x_j \) would be considered only because of \( \hat{r}_{(i_1,j)} \), \( i_1 \in D^c \). Since \( i_1 \in D^c \), \( \hat{i}_i^*(x_j) \) will always be sampled when \( r > H_3 \) and context \( x_j \) is considered.

This will lead to contradiction because as \( \hat{U}_j^b \) increases with \( \hat{n}_{(i,j),r} \) (this is because \( \hat{a}_{(i_1,j)} = \hat{n}_{(i,j),r}/n(r) \)), we would fail to meet condition

\[
\hat{U}_j^b < \hat{U}_j^{non}
\]

and some \( i_1 \in D^c \) would be sampled. So we have proved the lemma. \qed

**REMARK 4.** Since \( \hat{n}_{(i,j),r} \to \infty \) as \( r \to \infty \) for all \( x_j, j = 1, 2, \ldots, m \) and \( i = 1, 2, \ldots, k \), we will always have \( \hat{i}_i^*(x_j) = i^*(x_j) \) on almost every sample path when \( r \) is large enough. So, without loss of generality, when we talk about iteration \( r \) from here to the end of this paper, we always assume the \( r \) is large enough so that the estimate \( \hat{Y}_i(x_j) \) is close to the real value \( y_i(x_j) \) and \( \hat{i}_i^*(x_j) = i^*(x_j) = 1 \) for all \( j = 1, 2, \ldots, m \) and \( i = 1, 2, \ldots, k \).

**LEMMA 6.** For any context \( x_j \) and any two designs \( i_1, i_2 \neq i^*(x_j) \), \( \lim_{r \to \infty} \frac{\hat{a}_{(i_1,j)}}{\hat{a}_{(i_2,j)}} > 0 \) almost surely.

**Proof.** Similar to the method used in proof of lemma 4.2 in Chen and Ryzhov (2018), We will prove this lemma by contradiction when \( r \) is large enough and \( \hat{i}_i^*(x_j) = i^*(x_j) = 1 \).

By lemma 4, \( \delta_{(i_1,j)}, \delta_{(i_2,j)}, \hat{\sigma}_1^2(x_j), \hat{\sigma}_1(x_j), \hat{\sigma}_2^2(x_j) \) and \( \hat{\sigma}_2(x_j) \) converge as \( r \to \infty \). So, positive constants
\[ b_U, b_L, b_{v,i_1}, b_{s,1} \text{ and } b_{s,i_2} \text{ could be found such that} \]
\[ \hat{\delta}_{(i_1,j)} < b_U, \quad \hat{\delta}_{(i_2,j)} > b_L, \quad b_U > b_L; \quad \hat{\sigma}_{i_1}^2(x_j) > b_{v,i_1}, \quad \hat{\sigma}_1(x_j) < b_{s,1}, \quad \hat{\sigma}_2(x_j) < b_{s,i_2}. \]

when \( r > r_1 \). Suppose that there exist \( i_1, i_2 > 1 \) and

\[ \liminf_{r \to \infty} \frac{\hat{\alpha}_{(i_1,j)}}{\hat{\alpha}_{(i_2,j)}} = 0. \]  \( \text{(40)} \)

Then, because of equation \((40)\), we could find a large enough iteration \( r_2 > r_1 \) such that

\[ \frac{\hat{\alpha}_{(i_1,j),r_2}}{\hat{\alpha}_{(i_2,j),r_2}} < (b_U - b_L)b_{v,i_1} + b_Ub_{s,i_2} + 1 \quad \triangleq \quad c < \frac{b_U b_{v,i_1}}{b_L b_{v,i_2}} + 1, \]  \( \text{(41)} \)

and we will sample alternative \( i_2 \) of \( x_j \) to make

\[ \hat{\alpha}_{(i_1,j),r_2+1}^{i_2} < \hat{\alpha}_{(i_2,j),r_2+1}^{i_2}. \]  \( \text{However, at this iteration} \ r_2, \)

\[ \hat{V}_{(i_1,j)} - \hat{V}_{(i_2,j)} = \frac{\hat{\delta}_{(i_1,j)}}{(S_{(1,j)} + S_{(i_1,j),r_2}) n_{(r_2)}} - \frac{\hat{\delta}_{(i_2,j)}}{(S_{(1,j)} + S_{(i_2,j),r_2}) n_{(r_2)}} \]
\[ < \frac{b_U b_{v,i_1} b_{v,i_2}}{b_L b_{v,i_2}} + b_U b_{s,i_2} + b_L b_{v,i_1} \quad \frac{\hat{\alpha}_{(i_1,j)}}{\hat{\alpha}_{(i_2,j)}} - b_L b_{v,i_2} \quad \frac{\hat{\alpha}_{(i_1,j)}}{\hat{\alpha}_{(i_2,j)}} \]
\[ = \frac{(b_U - b_L) \hat{\delta}_1^2(x_j) + b_U \hat{\delta}_2^2(x_j) - b_L \hat{\delta}_{i_1}^2(x_j)}{(S_{(1,j),r} + S_{(i_1,j),r})(S_{(1,j),r} + S_{(i_2,j),r}) n_{(r_2)}} \frac{1}{\hat{\alpha}_{(i_2,j)}} < 0. \]

\[ \text{(43)} \]

\[ \text{(44)} \]

\[ \text{(42)} \]

\[ \text{(43)} \]

\[ \text{(44)} \]

For the nominator, we have

\[ (b_U - b_L) \frac{\hat{\delta}_1^2(x_j)}{\hat{\alpha}_{(i_1,j)}} + b_U \frac{\hat{\delta}_2^2(x_j)}{\hat{\alpha}_{(i_2,j)}} - b_L \frac{\hat{\delta}_{i_1}^2(x_j)}{\hat{\alpha}_{(i_1,j)}} \]
\[ \leq \ (b_U - b_L) \frac{\hat{\delta}_1(x_j) \hat{\delta}_{i_2}(x_j)}{\hat{\alpha}_{(i_1,j)}} + b_U \frac{\hat{\delta}_2^2(x_j)}{\hat{\alpha}_{(i_2,j)}} - b_L \frac{\hat{\delta}_{i_1}^2(x_j)}{\hat{\alpha}_{(i_1,j)}} \]
\[ \leq \ (b_U - b_L) \frac{\hat{\delta}_1(x_j) \hat{\delta}_{i_2}(x_j)}{\hat{\alpha}_{(i_2,j)}} + b_U \frac{\hat{\delta}_2^2(x_j)}{\hat{\alpha}_{(i_2,j)}} - b_L \frac{\hat{\delta}_{i_1}^2(x_j)}{\hat{\alpha}_{(i_2,j)}} \]
\[ = \left( (b_U - b_L) \hat{\delta}_1(x_j) \hat{\delta}_{i_2}(x_j) + b_U \hat{\delta}_2^2(x_j) - b_L \frac{\hat{\delta}_{i_1}^2(x_j)}{c} \right) \frac{1}{\hat{\alpha}_{(i_2,j)}} < 0. \]

\[ \text{(42)} \]

\[ \text{holds because an suboptimal alternative is sampled which means} \]

\[ \frac{\hat{\alpha}_{(1,j),r}^2}{\hat{\sigma}_1^2(x_j)} \leq \sum_{i \neq 1} \frac{\hat{\alpha}_{(i,j),r}^2}{\hat{\sigma}_1^2(x_j)}. \]
(43) and (44) holds because of (41). Thus, \(\hat{\mathcal{V}}_{(i_1,j)} - \hat{\mathcal{V}}_{(i_2,j)} < 0\) and \(i_2\) would not be sampled based on our algorithm. We finish the proof of this lemma by contradiction.

\[\Box\]

**Remark 5.** By symmetry, we have \(\limsup_{r \to \infty} \hat{\alpha}_{(i_1,j)} < \infty\) almost surely, for \(i_1, i_2 > 1\).

**Lemma 7.** For any context \(x_j\),

\[
\limsup_{r \to \infty} \frac{\hat{\alpha}_{(i^*(x_j),j)}}{\hat{\alpha}_{(i,j)}} < \infty, \quad \liminf_{r \to \infty} \frac{\hat{\alpha}_{(i^*(x_j),j)}}{\hat{\alpha}_{(i,j)}} > 0.
\]

almost surely, for \(i \neq i^*(x_j)\).

**Proof.** For a fixed sample path, we have \(\hat{i}_r(x_j) = i^*(x_j) = 1\) when \(r\) is large enough. So, it is sufficient to prove

\[
\limsup_{r \to \infty} \frac{\hat{\alpha}_{(1,j)}}{\hat{\alpha}_{(i,j)}} < \infty, \quad i \neq 1.
\]

Because \(\hat{\sigma}_i^2(x_j)\) converges as iteration \(r\) increases, \(b_{vL} > 0\) and \(b_{vU} > 0\) could be found such that

\[
0 < b_{vL} < \hat{\sigma}_i^2(x_j) < b_{vU}, \quad \text{for all } i = 1, 2, \ldots, k. \tag{45}
\]

By lemma 6 we have

\[
\hat{\alpha}_{(i,j)} < c_i \hat{\alpha}_{(i_0,j)}, \quad i > 1.
\]

Suppose there exists a suboptimal design \(i_0 \neq 1\) of context \(x_j\) which has

\[
\limsup_{r \to \infty} \frac{\hat{\alpha}_{(1,j)}}{\hat{\alpha}_{(i_0,j)}} = \infty.
\]

Then, we could find a iteration \(r\) where \((i^r, j^r) = (1, j)\) and

\[
\left(\frac{\hat{\alpha}_{(1,j)}}{\hat{\alpha}_{(i_0,j)}}\right)^2 > \frac{b_{vU}}{b_{vL}} \sum_{i \neq 1} c_i^2 + 1 \triangleq c. \tag{46}
\]

However, at iteration \(r\),

\[
\frac{(\hat{\alpha}_{(1,j)}/\hat{\sigma}_1(x_j))^2}{\sum_{i \neq 1} (\hat{\alpha}_{(i,j)}/\hat{\sigma}_i(x_j))^2} > \frac{\hat{\alpha}_{(1,j)}^2/b_{vU}}{\sum_{i \neq 1} \hat{\alpha}_{(i,j)}^2/b_{vL}} > \frac{\hat{\alpha}_{(1,j)}^2/b_{vU}}{\hat{\alpha}_{(i_0,j)}^2 \sum_{i \neq 1} c_i^2/b_{vL}} > 1.
\]
The last inequality holds because of (46). So, base on our algorithm, \((i^r, j^r) \neq (1, j)\), which leads to contradiction.

Similarly, it is easy to show that \(\liminf_{r \to \infty} \frac{\hat{\alpha}_{(i, j)}}{\hat{\alpha}_{(i_j, j)}} > 0\).

**Proof.** Suppose that \(i_1, i_2, j_1, j_2\) satisfy that

\[
\limsup_{r \to \infty} \frac{\hat{\alpha}_{(i_1, j_1)}}{\hat{\alpha}_{(i_2, j_2)}} = \infty. \tag{47}
\]

From lemma \([7]\) \(H_1\) and \(H_2\) could be found such that \(\hat{\alpha}_{(i_1, j_1)} > H_1 \hat{\alpha}_{(i_1, j_1)}\) and \(\hat{\alpha}_{(i_2, j_2)} < H_2 \hat{\alpha}_{(i_2, j_2)}\).

By similar argument to (45), there must exist a large enough iteration \(r_1\) such that

\[
0 < b_{vL} < \hat{\sigma}^2_i(x_j) < b_{vU}, \quad \forall \ i = 1, 2, \ldots, k \text{ and } j = j_1, j_2: \quad \hat{\delta}_{(i, j_1)} > b_{dL} > 0, \quad 0 < \hat{\delta}_{(i, j_2)} < b_{dU}.
\]

when \(r > r_1\). Then, because of equation (47), we could find a iteration \(r_2 > r_1\) such that

\[
\frac{\hat{\alpha}_{(i_1, j_1)}}{\hat{\alpha}_{(i_2, j_2)}} > b_{dU} \frac{b_{vU} / H_1 + b_{vU} + 1}{b_{dL} (b_{vL} / H_2 + b_{vL})} \triangleq c > b_{dU} \frac{\hat{\sigma}^2_i(x_j) / H_1 + \hat{\sigma}^2_i(x_j) + 1}{b_{dL} (\hat{\sigma}^2_j(x_j) / H_2 + \hat{\sigma}^2_j(x_j))}, \tag{48}
\]

and \((i^{r_2}, j^{r_2}) = (i_1, j_1)\). Then, we have

\[
\hat{V}_{(i_1, j_1)} - \hat{V}_{(i_2, j_2)} \geq \frac{\hat{\delta}_{(i_1, j_1)}}{(S_{(1, j_1), r_2} + S_{(i_1, j_1), r_2}) n(r_2)} - \frac{\hat{\delta}_{(i_2, j_2)}}{(S_{(1, j_2), r_2} + S_{(i_2, j_2), r_2}) n(r_2)}
\]

\[
> \frac{\hat{\delta}_{(i_1, j_1)}}{(S_{(1, j_1), r_2} + S_{(i_1, j_1), r_2}) n(r_2)} - \frac{b_{dL} \hat{\sigma}^2_i(x_j) / H_1 + \hat{\sigma}^2_i(x_j) / H_2 + \hat{\sigma}^2_j(x_j)}{(S_{(1, j_1), r_2} + S_{(i_2, j_2), r_2}) n(r_2)}
\]

\[
> \frac{\hat{\sigma}^2_i(x_j) / H_1 + \hat{\sigma}^2_j(x_j)}{(S_{(1, j_1), r_2} + S_{(i_1, j_1), r_2}) n(r_2)} - \frac{b_{dU} \hat{\sigma}^2_i(x_j) / H_1 + \hat{\sigma}^2_i(x_j) / H_2 + \hat{\sigma}^2_j(x_j)}{(S_{(1, j_1), r_2} + S_{(i_2, j_2), r_2}) n(r_2)}
\]

\[
= \frac{b_{dL} \hat{\sigma}^2_i(x_j) / H_1 + \hat{\sigma}^2_j(x_j)}{(S_{(1, j_1), r_2} + S_{(i_1, j_1), r_2}) n(r_2)} - \frac{b_{dU} \hat{\sigma}^2_i(x_j) / H_1 + \hat{\sigma}^2_i(x_j) / H_2 + \hat{\sigma}^2_j(x_j)}{(S_{(1, j_1), r_2} + S_{(i_2, j_2), r_2}) n(r_2)}
\]

\[
> 0.
\]

The last inequality holds because of (48). So, \(\hat{V}_{(i_1, j_1)} - \hat{V}_{(i_2, j_2)} > 0\). We could not sample design \(i_1\) of context \(x_{j_1}\) based on our algorithm. Now, we can prove this lemma by contradiction.
LEMMA 9. For any $x_{j_1}, x_{j_2}$ and any two designs $i_1, i_2$, \( \liminf_{r \to \infty} \frac{\hat{\alpha}_{(i_1,j_1)}}{\alpha_{(i_2,j_2)}} > 0 \) almost surely.

Proof. This lemma is straightforward from lemma 6, 7 and 8.

This lemma implies that $\hat{\sigma}_{i_1}^2(x_{j_1}) = \Theta(\hat{\sigma}_{i_2}^2(x_{j_2}))$ and $\hat{n}_{(i_1,j_1), r} = \Theta(n^{(r)})$ almost surely, for any $i_1, i_2 = 1, \ldots, k$ and $j_1, j_2 = 1, \ldots, m$.

Pleased be noted that we will use $\hat{\sigma}_{r, l}^2(x_j)$ to indicate which iteration $\hat{\sigma}_{r}^2(x_j)$ comes from.

LEMMA 10. For every context $x_j$, consider the subsequence where $x_j$ is considered. In this subsequence, between two samples of the optimal design, the number of samples that could be allocated to any suboptimal designs is $O\left(\sqrt{n^{(r)} \log \log n^{(r)}}\right)$ almost surely; symmetrically, between two samples assigned to any suboptimal designs, the number of samples that could be allocated to the optimal design is $O\left(\sqrt{n^{(r)} \log \log n^{(r)}}\right)$ almost surely.

Proof. Fix a sample path. Since $n^{(r)}(j) \to \infty$ almost surely, let $\{r_l, l = 1, 2, \ldots\}$ be the sequences where $x_j$ is considered and assume $i^*(x_j) = 1$ when $r_l$ is large enough. We focus on

$$
\Lambda_{r_l, j} = \left(\frac{\hat{n}_{(1,j), r_l}}{\hat{\sigma}_{1,r_l}(x_j)}\right)^2 - \sum_{i \neq 1} \left(\frac{\hat{n}_{(i,j), r_l}}{\hat{\sigma}_{i,r_l}(x_j)}\right)^2, \; l = 1, 2, \ldots.
$$

(CASE 1: First, we prove that between two samples assigned to the optimal design of context $x_j$, total number of samples that could be allocated to any suboptimal designs of $x_j$ is $O\left(\sqrt{n^{(r)} \log \log n^{(r)}}\right)$ almost surely. Suppose $(i^r, j^r) = (1, j)$ and $(i^{r+u}, j^{r+u}) = (i_u, j)$, $i_u \neq 1$, $1 \leq u \leq e$. Then, by our algorithm,

$$
\Lambda_{r_l, j} < 0, \; \Lambda_{r_l+1, j} \geq 0, \; \Lambda_{r_l+e, j} \geq 0, \; \Lambda_{r_l+e+1, j} < 0.
$$

Since $\Lambda_{r_l, j} < 0$, we have

$$
\left(\frac{\hat{n}_{(1,j), r_l}}{\hat{\sigma}_{1,r_l}(x_j)}\right)^2 < \sum_{i \neq 1} \left(\frac{\hat{n}_{(i,j), r_l}}{\hat{\sigma}_{i,r_l}(x_j)}\right)^2.
$$

(50)
Since $\Lambda_{r_{l+1},j} > 0$, we have

$$
\left( \frac{\hat{n}_{(i,j), r_l} + 1}{\hat{\sigma}_{1,r_{l+1}}(x_j)} \right)^2 = \left( \frac{\hat{n}_{(i,j), r_{l+1}}}{\hat{\sigma}_{1,r_{l+1}}(x_j)} \right)^2 \geq \sum_{i \neq 1} \left( \frac{\hat{n}_{(i,j), r_l}}{\hat{\sigma}_{1,r_l}(x_j)} \right)^2 = \sum_{i \neq 1} \left( \frac{\hat{n}_{(i,j), r_l}}{\hat{\sigma}_{1,r_l}(x_j)} \right)^2.
$$

(51)

By (50) and (51), we have

$$
\left( \frac{\hat{n}_{(i,j), r_l} + 1}{\hat{\sigma}_{1,r_{l+1}}(x_j)} \right)^2 \leq \sum_{i \neq 1} \left( \frac{\hat{n}_{(i,j), r_l}}{\hat{\sigma}_{1,r_l}(x_j)} \right)^2 + \left( \frac{\hat{n}_{(i,j), r_{l+1}}}{\hat{\sigma}_{1,r_{l+1}}(x_j)} \right)^2 - \left( \frac{\hat{n}_{(i,j), r_l}}{\hat{\sigma}_{1,r_l}(x_j)} \right)^2.
$$

(52)

At iteration $r_{l+e}$, $\Lambda_{r_{l+e},j} \geq 0$ means that

$$
\left( \frac{\hat{n}_{(i,j), r_{l+e}}}{\hat{\sigma}_{1,r_{l+e}}(x_j)} \right)^2 = \left( \frac{\hat{n}_{(i,j), r_l} + 1}{\hat{\sigma}_{1,r_l}(x_j)} \right)^2 \geq \sum_{i \neq 1} \left( \frac{\hat{n}_{(i,j), r_l}}{\hat{\sigma}_{1,r_l}(x_j)} \right)^2 = \sum_{i \neq 1} \left( \frac{\hat{n}_{(i,j), r_l}}{\hat{\sigma}_{1,r_l}(x_j)} \right)^2.
$$

(53)

By (52), we must have

$$
\sum_{i \neq 1} \left( \frac{\hat{n}_{(i,j), r_l} + e_i}{\hat{\sigma}_{1,r_l}(x_j)} \right)^2 \leq \sum_{i \neq 1} \left( \frac{\hat{n}_{(i,j), r_l}}{\hat{\sigma}_{1,r_l}(x_j)} \right)^2 + \left( \frac{\hat{n}_{(i,j), r_{l+1}}}{\hat{\sigma}_{1,r_{l+1}}(x_j)} \right)^2 - \left( \frac{\hat{n}_{(i,j), r_l}}{\hat{\sigma}_{1,r_l}(x_j)} \right)^2.
$$

By multiplying $\prod_{h=1}^{l} \hat{\sigma}_{h,r_{l+e}}^2(x_j)$ on both sides, the above equation is equivalent to

$$
\sum_{i \neq 1} \left( \frac{\hat{n}_{(i,j), r_l} + e_i}{\hat{\sigma}_{1,r_l}(x_j)} \right)^2 - \frac{\hat{\sigma}_{1,r_{l+e}}^2(x_j)}{\hat{n}_{(i,j), r_l}} \sum_{h \neq i} \hat{\sigma}_{h,r_{l+e}}^2(x_j) \leq \left( \frac{\hat{n}_{(i,j), r_{l+1}}}{\hat{\sigma}_{1,r_{l+1}}(x_j)} \right)^2 - \frac{\hat{\sigma}_{1,r_{l+1}}^2(x_j)}{\hat{n}_{(i,j), r_l}} \sum_{h \neq i} \hat{\sigma}_{h,r_{l+1}}^2(x_j).
$$

(53)

(53) holds because $\hat{\sigma}_{1,r_{l+e}}^2(x_j) = \hat{\sigma}_{1,r_{l+1}}^2(x_j)$.

By property of variance estimator,

$$
\hat{\sigma}^2 = \frac{1}{N-1} \sum_{q=1}^{N} (W_q - \bar{W})^2 = \frac{1}{N-1} \sum_{q=1}^{N-1} Z_q^2,
$$

where $W_q, q = 1, 2, \cdots, N$ are $N$ observations and $Z_q = \frac{1}{\sqrt{q(q+1)}} \sum_{s=1}^{q} W_s - \frac{q}{\sqrt{q(q+1)}} W_{q+1}$. So, $Z_q \sim \mathcal{N}(0, \sigma^2)$ and $Z_q$ are independent of $Z_p, q, p = 1, 2, \cdots, N-1, q \neq p$. By the law of iterated
Then, we must have
\[ \left| \frac{\hat{\sigma}_{i,r+l+e}^2(x_j)}{\hat{\sigma}_{i,r}^2(x_j)} - 1 \right| = \left| \frac{\hat{\sigma}_{i,r+l+e}^2(x_j) - \hat{\sigma}_{i,r}^2(x_j)}{\hat{\sigma}_{i,r}^2(x_j)} \right| = \left| \frac{\hat{\sigma}_{i,r+l+e}^2(x_j) - \sigma_i^2(x_j) + \sigma_i^2(x_j) - \hat{\sigma}_{i,r}^2(x_j)}{\hat{\sigma}_{i,r}^2(x_j)} \right| = O\left( \frac{\log \log \hat{n(i,j),r_l}}{\hat{n(i,j),r_l}} \right). \] (54)

Therefore, taking (54) into (53), we have
\[
\sum_{i \neq 1} \left( 2\hat{n}_{i,i},r_l e_i + e_i^2 - c_{g_i} \sqrt{\hat{n}_{i,i},r_l} \log \log \hat{n}_{i,i},r_l \hat{n}_{i(i),r_l} \right) \prod_{h \neq i} \delta_{h,r_l+e}^2(x_j) < \left( 2\hat{n}_{i,i},r_l \right)^2 + 1 + c_{g_i} \sqrt{\hat{n}_{i,i},r_l} \log \log \hat{n}_{i(i),r_l} \hat{n}_{i(i),r_l} \prod_{h \neq i} \delta_{h,r_l+e}^2(x_j),
\] (55)

where \( c_{g_i}, i = 1, 2, \ldots, k \) are positive constants. By (55), for any \( i_0 \neq 1 \), we must have
\[
\left( 2\hat{n}_{i_0,i_0},r_l e_{i_0} + e_{i_0}^2 - c_{g_i} \sqrt{\hat{n}_{i_0,i_0},r_l} \log \log \hat{n}_{i_0,i_0},r_l \hat{n}_{i(i),r_l} \right) \prod_{h \neq i_0} \delta_{h,r_l+e}^2(x_j) < \left( 2\hat{n}_{i_0,i_0},r_l \right)^2 + 1 + c_{g_i} \sqrt{\hat{n}_{i_0,i_0},r_l} \log \log \hat{n}_{i(i),r_l} \hat{n}_{i(i),r_l} \prod_{h \neq i} \delta_{h,r_l+e}^2(x_j)
\]
\[- \sum_{i \neq 1, i \neq i_0} \left( 2\hat{n}_{i,i},r_l e_i + e_i^2 - c_{g_i} \sqrt{\hat{n}_{i,i},r_l} \log \log \hat{n}_{i(i),r_l} \hat{n}_{i(i),r_l} \right) \prod_{h \neq i} \delta_{h,r_l+e}^2(x_j)
\]
\[
\left( 2\hat{n}_{i_0,i_0},r_l \right)^2 + 1 + c_{g_i} \sqrt{\hat{n}_{i_0,i_0},r_l} \log \log \hat{n}_{i(i),r_l} \hat{n}_{i(i),r_l} \prod_{h \neq i} \delta_{h,r_l+e}^2(x_j)
\]
\[
+ \sum_{i \neq 1, i \neq i_0} c_{g_i} \sqrt{\hat{n}_{i(i),r_l} \log \log \hat{n}_{i(i),r_l} \hat{n}_{i(i),r_l} \prod_{h \neq i} \delta_{h,r_l+e}^2(x_j)}
\]
\[
< \sum_{i \neq i_0} c'_{g_i} \sqrt{\hat{n}_{i(i),r_l} \log \log \hat{n}_{i(i),r_l} \hat{n}_{i(i),r_l} \prod_{h \neq i} \delta_{h,r_l+e}^2(x_j)}
\] (56)

Reorganize (56), we have
\[
2\hat{n}_{i_0,i_0},r_l e_{i_0} + e_{i_0}^2 < \sum_{i = 1}^k c'_{g_i} \sqrt{\hat{n}_{i(i),r_l} \log \log \hat{n}_{i(i),r_l} \hat{n}_{i(i),r_l} r_l}.
\]

Then, we must have
\[
e_{i_0} < \sum_{i = 1}^k c'_{g_i} \sqrt{\hat{n}_{i(i),r_l} \log \log \hat{n}_{i(i),r_l} \hat{n}_{i(i),r_l} r_l} = O\left( \sqrt{n(r_l) \log \log n(r_l)} \right).
\]
By (57), we must have \( \hat{n}_{(i,j),r_l} \leq \Theta(1) \).

**CASE 2:** Next, we prove that between two samples assigned to any suboptimal designs of context \( x_j \), the number of samples that could be allocated to the optimal design of \( x_j \) is \( O(\sqrt{n(r) \log \log n(r)}) \) almost surely.

Suppose \( (i^u, j^u) = (i_0, j) \), where \( i_0 > 1 \), and \( (i^{r+u}, j^{r+u}) = (1, j) \), \( 1 \leq u \leq e \). Then, by our algorithm, \( \Lambda_{r_1,j} \geq 0, \quad \Lambda_{r_1+1,j} < 0, \quad \Lambda_{r_1+e,j} < 0, \quad \Lambda_{r_1+e+1,j} \geq 0. \)

Similar to (50), (51) and (52), since \( \Lambda_{r_1,j} \geq 0 \) and \( \Lambda_{r_1+1,j} < 0 \), we have

\[
\left( \frac{\hat{n}_{(1,j),r_1}}{\sigma_{1,r_1}(x_j)} \right)^2 \geq \sum_{i \neq 1} \left( \frac{\hat{n}_{(i,j),r_1}}{\sigma_{i,r_1}(x_j)} \right)^2,
\]

\[
\left( \frac{\hat{n}_{(1,j),r_1}}{\sigma_{1,r_1}(x_j)} \right)^2 < \sum_{i \neq 1} \left( \frac{\hat{n}_{(i,j),r_1+1}}{\sigma_{i,r_1+1}(x_j)} \right)^2 = \sum_{i \neq 1, i \neq 0} \left( \frac{\hat{n}_{(i,j),r_1}}{\sigma_{i,r_1}(x_j)} \right)^2 + \left( \frac{\hat{n}_{(i_0,j),r_1} + 1}{\sigma_{i_0,r_1+1}(x_j)} \right)^2,
\]

\[
\sum_{i \neq 1} \left( \frac{\hat{n}_{(i,j),r_1+1}}{\sigma_{i,r_1+1}(x_j)} \right)^2 < \left( \frac{\hat{n}_{(1,j),r_1}}{\sigma_{1,r_1}(x_j)} \right)^2 - \left( \frac{\hat{n}_{(i_0,j),r_1}}{\sigma_{i_0,r_1}(x_j)} \right)^2 + \left( \frac{\hat{n}_{(i_0,j),r_1} + 1}{\sigma_{i_0,r_1+1}(x_j)} \right)^2.
\]

(57)

At iteration \( r_{t+e} \), \( \Lambda_{r_{t+e},j} < 0 \) means that

\[
\left( \frac{\hat{n}_{(1,j),r_1} + e - 1}{\sigma_{1,r_1+e}(x_j)} \right)^2 = \left( \frac{\hat{n}_{(1,j),r_1}}{\sigma_{1,r_1}(x_j)} \right)^2 - \sum_{i \neq 1} \left( \frac{\hat{n}_{(i,j),r_1+e}}{\sigma_{i,r_1+e}(x_j)} \right)^2 = \sum_{i \neq 1} \left( \frac{\hat{n}_{(i,j),r_1+1}}{\sigma_{i,r_1+1}(x_j)} \right)^2.
\]

By (57), we must have

\[
\left( \frac{\hat{n}_{(1,j),r_1} + e - 1}{\sigma_{1,r_1+e}(x_j)} \right)^2 < \left( \frac{\hat{n}_{(1,j),r_1}}{\sigma_{1,r_1}(x_j)} \right)^2 - \left( \frac{\hat{n}_{(i_0,j),r_1}}{\sigma_{i_0,r_1}(x_j)} \right)^2 + \left( \frac{\hat{n}_{(i_0,j),r_1} + 1}{\sigma_{i_0,r_1+1}(x_j)} \right)^2,
\]

that is, equivalently,

\[
2(e - 1)\hat{n}_{(1,j),r_1} + (e - 1)^2 < \frac{\sigma^2_{1,r_1+e}(x_j)}{\sigma^2_{1,r_1}(x_j)} \hat{n}^2_{(1,j),r_1} - \hat{n}^2_{(i_0,j),r_1} + \frac{\sigma^2_{1,r_1+e}(x_j)}{\sigma^2_{i_0,r_1}(x_j)} \hat{n}^2_{(i_0,j),r_1} + \frac{\sigma^2_{i_0,r_1+1}(x_j)}{\sigma^2_{i_0,r_1}(x_j)} \hat{n}^2_{(i_0,j),r_1} + \frac{\sigma^2_{i_0,r_1+1}(x_j)}{\sigma^2_{i_0,r_1}(x_j)} \hat{n}^2_{(i_0,j),r_1} + \frac{c_{g1}}{\sqrt{\hat{n}_{(1,j),r_1} \log \log \hat{n}_{(1,j),r_1} \hat{n}_{(1,j),r_1}}} + c_{var} \left( \frac{\sqrt{\hat{n}_{(i_0,j),r_1} \log \log \hat{n}_{(i_0,j),r_1} \hat{n}_{(i_0,j),r_1}} + (2\hat{n}_{(i_0,j),r_1} + 1)\sigma^2_{i_0,r_1}(x_j) \right).
\]
\[ < c_{g1} \sqrt{\hat{n}_{(1,j),r_l}} \log \log \hat{n}_{(1,j),r_l} \hat{n}_{(1,j),r_l} \hat{n}_{(1,j),r_l} c''_{g10} \sqrt{\hat{n}_{(i_0,j),r_l} \log \log \hat{n}_{(i_0,j),r_l} \hat{n}_{(i_0,j),r_l}} \]

where \( c_{g1} \) and \( c'_{g10} \) are positive constants. Then, we must have

\[ 2e\hat{n}_{(1,j),r_l} \leq c''_{g1} \sqrt{\hat{n}_{(1,j),r_l} \log \log \hat{n}_{(1,j),r_l} \hat{n}_{(1,j),r_l} + c''_{g10} \sqrt{\hat{n}_{(i_0,j),r_l} \log \log \hat{n}_{(i_0,j),r_l} \hat{n}_{(i_0,j),r_l}}} \]

Therefore, \( e = O\left(\sqrt{n(r_l)} \log \log n(r_l)\right) \).

\[ \square \]

**Lemma 11.** Suppose \((\hat{\mu}_N, \hat{\sigma}_N^2)\) is mean and variance estimate of a Gaussian distribution \(N(\mu, \sigma^2)\).

We have

\[ N^{\frac{1}{2}} |\hat{\sigma}_{N+1}^2 - \hat{\sigma}_N^2| \to 0, \quad \text{as } N \to \infty \]

almost surely.

**Proof.** Let \( W_q \) denote the \( q \)-th sample. Because \( \hat{\sigma}_N^2 = \frac{1}{N-1} \sum_{q=1}^{N} (W_q - \hat{\mu}_N)^2 \), we have

\[ \hat{\sigma}_{N+1}^2 - \hat{\sigma}_N^2 = \frac{\hat{\sigma}_N^2}{N} - \frac{1}{N+1} (W_{N+1} - \hat{\mu}_N)^2. \]

By strong law of large number, \( |\hat{\sigma}_N^2 - \sigma^2| \to 0 \) almost surely which implies that \( \frac{\hat{\sigma}_N^2}{N^{1/2}} \to 0 \) a.s. Therefore, the lemma will be proved if \( \frac{(W_{N+1} - \hat{\mu}_N)^2}{N^{1/2}} \to 0 \) as \( N \to \infty \).

\[ \frac{(W_{N+1} - \hat{\mu}_N)^2}{N^{1/2}} = \frac{(W_{N+1} - \mu + \mu - \hat{\mu}_N)^2}{N^{1/2}} = \frac{(W_{N+1} - \mu)^2}{N^{1/2}} + \frac{(\mu - \hat{\mu}_N)^2}{N^{1/2}} + \frac{2(W_{N+1} - \mu)(\mu - \hat{\mu}_N)}{N^{1/2}}. \]

Again, by strong law of large number, \( |\mu - \hat{\mu}_N| \to 0 \) almost surely. To prove this lemma, it is sufficient to have

\[ \frac{(W_{N+1} - \mu)^2}{N^{1/2}} \to 0, \quad \frac{2|W_{N+1} - \mu|}{N^{1/2}} \to 0 \]

almost surely as \( N \to \infty \).

The following is similar to proof of Lemma 4.5 in Chen and Ryzhov (2018). By Markov’s
inequality, for all $\varepsilon > 0$,

$$P\left(\frac{(W_{N+1} - \mu)^2}{N^{1/4}} \geq \varepsilon \right) \leq \mathbb{E}\left(\frac{(W_{N+1} - \mu)^{16}}{N^2\varepsilon^8}\right) \leq \frac{u_{16}}{N^2\varepsilon^8}$$

The last inequality holds because that any order central moment of Gaussian distribution exists.

By the Borel-Cantelli lemma, since

$$\sum N P\left(\frac{(W_{N+1} - \mu)^2}{N^{1/4}} \geq \varepsilon \right) \leq \sum N \frac{u_{16}}{N^2\varepsilon^8} < \infty,$$

we have $(W_{N+1} - \mu)^2 \to 0$ almost surely. Similarly, we can also prove $\frac{2|W_{N+1} - \mu|}{N^{1/4}} \to 0$. Therefore, $(W_{N+1} - \mu N)^2 \to 0$ as $N \to \infty$ and the proof is completed.

\[ \square \]

Then, we can prove Theorem 5 now.

**Proof of Theorem 5.** For a given context $x_j$, define

$$\Delta_r = \left(\frac{\hat{\alpha}_{(1,j),r}}{\sigma_1(x_j)}\right)^2 - \sum_{i \neq 1} \left(\frac{\hat{\alpha}_{(i,j),r}}{\sigma_i(x_j)}\right)^2,$$

$$\hat{\Delta}_r = \left(\frac{\hat{\alpha}_{(1,j),r}}{\hat{\sigma}_{1,r}(x_j)}\right)^2 - \sum_{i \neq 1} \left(\frac{\hat{\alpha}_{(i,j),r}}{\hat{\sigma}_{i,r}(x_j)}\right)^2.$$

Since $\hat{n}_{(i,j),r} \to \infty$, $i = 1, 2, \cdots, k$, to have $\lim_{r \to \infty} \Delta_r = 0$, it is sufficient to prove $\lim_{r \to \infty} \hat{\Delta}_r = 0$.

If $(i^r, j^r) = (1, j)$ at iteration $r$, then $\hat{\Delta}_r < 0$ and

$$\hat{\Delta}_{r+1} - \hat{\Delta}_r = \left(\frac{\hat{\alpha}_{(1,j),r+1}}{\sigma_{1,r+1}(x_j)}\right)^2 - \sum_{i \neq 1} \left(\frac{\hat{\alpha}_{(i,j),r+1}}{\sigma_{i,r+1}(x_j)}\right)^2 - \left(\frac{\hat{\alpha}_{(1,j),r}}{\sigma_{1,r}(x_j)}\right)^2 + \sum_{i \neq 1} \left(\frac{\hat{\alpha}_{(i,j),r}}{\sigma_{i,r}(x_j)}\right)^2$$

$$= \left(\frac{\hat{n}_{(1,j),r} + 1}{(n^{(r)} + 1)\hat{\sigma}_{1,r}(x_j)}\right)^2 - \left(\frac{\hat{n}_{(1,j),r}}{n^{(r)}\hat{\sigma}_{1,r}(x_j)}\right)^2 - \sum_{i \neq 1} \left(\frac{\hat{n}_{(i,j),r}}{(n^{(r)} + 1)\hat{\sigma}_{i,r}(x_j)}\right)^2 - \sum_{i \neq 1} \left(\frac{\hat{n}_{(i,j),r}}{n^{(r)}\hat{\sigma}_{i,r}(x_j)}\right)^2.$$

\[ (58) \]
Consider the elements in (58) separately.

\[
\left| \left( \frac{\hat{n}_{(1,j),r}}{(n(r) + 1)\hat{\sigma}_{1,r+1}(x_j)} \right)^2 - \left( \frac{\hat{n}_{(1,j),r}}{n(r)\hat{\sigma}_{1,r}(x_j)} \right)^2 \right|
\]

\[
= \left| \left( \frac{\hat{n}_{(1,j),r}}{(n(r) + 1)\hat{\sigma}_{1,r+1}(x_j)} \right)^2 - \left( \frac{\hat{n}_{(1,j),r}}{n(r) + 1} \right)^2 \right|
\]

\[
\leq b_1(\hat{\sigma}_{1,r}^2(x_j) - \hat{\sigma}_{1,r+1}^2(x_j)) + b_2 \frac{1}{n(r)}
\]

\[
= O((n(r))^{-3/4}),
\]

(59)

where (59) comes from lemma 9 which implies that \( \hat{n}_{(1,j),r}/n(r) = \Theta(1) \). (60) comes from lemma 11.

Meanwhile, for the other element in (58),

\[
\left| \sum_{i \neq 1} \left( \frac{\hat{n}_{(i,j),r}}{(n(r) + 1)\hat{\sigma}_{i,r}(x_j)} \right)^2 - \sum_{i \neq 1} \left( \frac{\hat{n}_{(i,j),r}}{n(r)\hat{\sigma}_{i,r}(x_j)} \right)^2 \right|
\]

\[
= \frac{2n(r) + 1}{n(r)^2(n(r) + 1)^2} \sum_{i \neq 1} \left( \frac{\hat{n}_{(i,j),r}}{\hat{\sigma}_{i,r}(x_j)} \right)^2 = O\left( \frac{1}{n(r)} \right).
\]

(61)

Therefore, we have

\[
|\hat{\Delta}_{r+1} - \hat{\Delta}_r| = O((n(r))^{-3/4}).
\]

Let \( \{r_l, l = 1, 2, \ldots \} \) be the subsequence where \( x_j \) is considered. Suppose \( (i'^{l}, j'^{l}) = (i, j), \)

\( i \neq 1, (i'^{l+1}, j'^{l+1}) = (1, j) \) and define

\[ e = \sup \{ u \geq 1 : (i'^{l+u}, j'^{l+u}) = (1, j) \}. \]

By lemma 10, \( e = O(\sqrt{n(r_l)} \log \log n(r_l)) \). Also note that \( |\hat{\Delta}_{r+1}| = \frac{(n(r))^2}{(n(r) + 1)^2} |\Delta_r| \) when \( j^r \neq j \). Since \( \hat{\Delta}_{i^r} > 0 \) and \( \hat{\Delta} \Delta_{i^r+} < 0 \), We have

\[
|\hat{\Delta}_{r+1}| - |\hat{\Delta}_{r+1} - \hat{\Delta}_r| \leq a|\hat{\Delta}_{r+1} - \hat{\Delta}_r| \leq e|\hat{\Delta}_{r+1} - \hat{\Delta}_r| = O\left( \frac{\sqrt{n(r_l)} \log \log n(r_l)}{(n(r_l))^{3/4}} \right), \quad 1 \leq a \leq e, \quad (62)
\]

whence \( \liminf_{r \to \infty} \hat{\Delta}_r = 0. \)

Similarly, \( \limsup_{r \to \infty} \hat{\Delta}_r = 0. \) Therefore, \( \lim_{r \to \infty} \hat{\Delta}_r = 0 \), completing the proof.

\[ \square \]

**Lemma 12.** For \( j = 1, 2, \ldots, m \) and \( i = 1, 2, \ldots, k \), \( i \neq i^*(x_j) \), let \( \Gamma_{(i_1,j_1),(i_2,j_2)} = \hat{V}_{(i_1,j_1)} - \)
\[ \hat{v}_{(i_2,j_2)}: \text{Under our algorithm,} \]
\[ \left| \Gamma_{(i_1,j_1),(i_2,j_2)}(r + 1) - \Gamma_{(i_1,j_1),(i_2,j_2)}(r) \right| < C(n(r))^{-3/4} \]

holds almost surely, for any \( j_1,j_2 = 1,2,\ldots,m \) and \( i_1,i_2 = 1,2,\ldots,k \), \( i_1 \neq \hat{i}_r(x_{j_1}), i_2 \neq \hat{i}_r(x_{j_2}) \).

**Proof.** Without loss of generality, we assume \( \hat{i}_r(x_j) = i^*(x_j) = 1 \) for any \( j = 1,2,\ldots,m \), when \( r \) is large enough.

Now, we discuss on \( \hat{v}_{(i,j),r+1} - \hat{v}_{(i,j),r} \), which is the increment of \( \hat{v}_{(i,j)} \) after the \((r+1)\)th iteration takes place.

If \((i^*,j^*) = (i,j)\), \( i \neq 1 \), which means that design \( i \) of context \( x_j \) is sampled at iteration \( r \), then
\[ \begin{align*}
\frac{1}{\hat{\alpha}_{(i,j),r+1}} &= \frac{n^{(r)} + 1}{\hat{n}_{(i,j),r+1}} = \frac{n^{(r)}}{\hat{n}_{(i,j),r}} \left( 1 - \frac{n^{(r)} - \hat{n}_{(i,j),r}}{n^{(r)} + n^{(r)}\hat{n}_{(i,j),r}} \right) = \frac{1}{\hat{\alpha}_{(i,j),r}} \left( 1 - \frac{n^{(r)} - \hat{n}_{(i,j),r}}{n^{(r)} + n^{(r)}\hat{n}_{(i,j),r}} \right), \\
\frac{1}{\hat{\alpha}_{(i,j),r+1}} &= \frac{n^{(r)} + 1}{\hat{n}_{(i,j),r+1}} = \frac{1}{\hat{\alpha}_{(i,j),r}} + \frac{1}{\hat{n}_{(i,j),r}}.
\end{align*} \]

We have:
\[ \dot{\hat{V}}_{(i,j),r+1} - \dot{\hat{V}}_{(i,j),r} = \frac{\dot{\hat{\alpha}}_{(i,j),r+1}}{\hat{\alpha}_{(i,j),r+1}} - \frac{\dot{\hat{\alpha}}_{(i,j),r}}{\hat{\alpha}_{(i,j),r}} < b_{o1} \left( \dot{\hat{\sigma}}_{(i,j),r+1} - \dot{\hat{\sigma}}_{(i,j),r} \right) \]
\[ = b_{o1} \left( \dot{\hat{\sigma}}_{(i,j),r+1} - \dot{\hat{\sigma}}_{(i,j),r} \right) \]
\[ = \frac{\dot{\hat{\alpha}}_{(i,j),r+1}}{\hat{\alpha}_{(i,j),r+1}} - \frac{\dot{\hat{\alpha}}_{(i,j),r}}{\hat{\alpha}_{(i,j),r}} < b_{o1} \left( \dot{\hat{\sigma}}_{(i,j),r+1} - \dot{\hat{\sigma}}_{(i,j),r} \right) \]
\[ = \frac{\dot{\hat{\alpha}}_{(i,j),r+1}}{\hat{\alpha}_{(i,j),r+1}} - \frac{\dot{\hat{\alpha}}_{(i,j),r}}{\hat{\alpha}_{(i,j),r}} < b_{o1} \left( \dot{\hat{\sigma}}_{(i,j),r+1} - \dot{\hat{\sigma}}_{(i,j),r} \right) \]
\[ = b_{o2} \left( \dot{\hat{\alpha}}_{(i,j),r+1} - \dot{\hat{\alpha}}_{(i,j),r} \right) + b_{o3} \frac{1}{n^{(r)}}. \]

\[ (65) \]
\[ (66) \]
\[ (67) \]

(65) holds because we sample a suboptimal design of \( x_j \) and (66) comes from (63). The last equation holds because of lemma 9. By lemma 4.5 in [Chen and Ryzhov (2018)] which implies that \((n^{(r)})^{3/4}\left| \hat{\delta}_{(i,j),r+1} - \hat{\delta}_{(i,j),r} \right| \rightarrow 0 \) as \( r \rightarrow \infty \), we have \( \left| \hat{\delta}_{(i,j),r+1} - \hat{\delta}_{(i,j),r} \right| \leq b_{o4}(n^{(r)})^{-3/4} \) when \( r \) is large enough.
Therefore, it follows from (67) that

$$\left| \frac{\hat{\delta}_{i,j}^{r+1} \hat{\sigma}_{i,r}^2(x_j)}{\hat{\alpha}_{(i,j),r}} - \frac{\hat{\delta}_{i,j}^{r} \hat{\sigma}_{i,r+1}^2(x_j)}{\hat{\alpha}_{(i,j),r+1}} \right| \leq b_{65}(n^{(r)})^{-3/4}. \quad (68)$$

Next, we discuss another part of RHS of (64).

$$\begin{align*}
&\frac{\hat{\delta}_{i,j}^{r+1} \hat{\sigma}_{i,r}^2(x_j)}{\hat{\alpha}_{(i,j),r}} - \frac{\hat{\delta}_{i,j}^{r} \hat{\sigma}_{i,r+1}^2(x_j)}{\hat{\alpha}_{(i,j),r+1}} \\
&= \frac{\hat{\delta}_{i,j}^{r+1} \hat{\sigma}_{i,r}^2(x_j)}{\hat{\alpha}_{(i,j),r}} - \frac{\hat{\delta}_{i,j}^{r} \hat{\sigma}_{i,r+1}^2(x_j)}{\hat{\alpha}_{(i,j),r+1}} + \frac{\hat{\delta}_{i,j}^{r+1} \hat{\sigma}_{i,r+1}^2(x_j)}{\hat{\alpha}_{(i,j),r+1}} - \frac{\hat{\delta}_{i,j}^{r} \hat{\sigma}_{i,r+1}^2(x_j)}{\hat{\alpha}_{(i,j),r+1}} \\
&= \left( \frac{\hat{\delta}_{i,j}^{r+1} \hat{\sigma}_{i,r}^2(x_j)}{\hat{\alpha}_{(i,j),r}} - \frac{\hat{\delta}_{i,j}^{r} \hat{\sigma}_{i,r+1}^2(x_j)}{\hat{\alpha}_{(i,j),r+1}} \right) + \left( \frac{\hat{\delta}_{i,j}^{r+1} \hat{\sigma}_{i,r+1}^2(x_j)}{\hat{\alpha}_{(i,j),r+1}} - \frac{\hat{\delta}_{i,j}^{r} \hat{\sigma}_{i,r+1}^2(x_j)}{\hat{\alpha}_{(i,j),r+1}} \right) \\
&\leq b_{661}(\hat{\sigma}_{i,r}^2(x_j) - \hat{\sigma}_{i,r+1}^2(x_j)) + b_{662}(n^{(r)})^{-1} + b_{663}(n^{(r)})^{-3/4},
\end{align*}$$

(69)

where (69) comes from first equation of (63). By lemma 11

$$\left| \hat{\sigma}_{i,r}^2(x_j) - \hat{\sigma}_{i,r+1}^2(x_j) \right| \leq b_{67}(n^{(r)})^{-3/4},$$

Thus, we have

$$\left| \frac{\hat{\delta}_{i,j}^{r+1} \hat{\sigma}_{i,r}^2(x_j)}{\hat{\alpha}_{(i,j),r}} - \frac{\hat{\delta}_{i,j}^{r} \hat{\sigma}_{i,r+1}^2(x_j)}{\hat{\alpha}_{(i,j),r+1}} \right| \leq b_{68}(n^{(r)})^{-3/4}. \quad (70)$$

Combine (68) and (70), we can find a large enough $P_1$ such that the increment

$$|\hat{\delta}_{(i,j),r+1} - \hat{\delta}_{(i,j),r}| < b_{69}(n^{(r)})^{-3/4}$$

when $r > P_1$.

Similarly, if $(i', j') = (1, j)$, we can find a large enough $P_2$ such that the increment $|\hat{\delta}_{(i,j),r+1} - \hat{\delta}_{(i,j),r}| < b_{70}(n^{(r)})^{-3/4}$ when $r > P_2$.

If $(i', j') \neq (1, j)$ and $(i', j') \neq (i, j)$, we have:

$$\begin{align*}
\hat{\delta}_{(i,j),r+1} - \hat{\delta}_{(i,j),r} &= \hat{\delta}_{(i,j)}^{r+1} \hat{\sigma}_{i,r+1}^2(x_j) \frac{\hat{\sigma}_{i,r+1}^2(x_j)}{\hat{\alpha}_{(i,j),r+1}} + \hat{\delta}_{(i,j)}^{r} \hat{\sigma}_{i,r+1}^2(x_j) \frac{\hat{\sigma}_{i,r+1}^2(x_j)}{\hat{\alpha}_{(i,j),r+1}} \\
&= \hat{\delta}_{(i,j)}^{r+1} \hat{\sigma}_{i,r+1}^2(x_j) \frac{\hat{\sigma}_{i,r+1}^2(x_j)}{\hat{\alpha}_{(i,j),r+1}} + \hat{\delta}_{(i,j)}^{r} \hat{\sigma}_{i,r+1}^2(x_j) \frac{\hat{\sigma}_{i,r+1}^2(x_j)}{\hat{\alpha}_{(i,j),r+1}} \\
&= -b_{71} \frac{1}{n^{(r)}},
\end{align*}$$

(71)
where $b_{o11}$ is a positive constant and (71) holds because none estimate in $\hat{V}_{(i,j)}$ is changed from iteration $r$ to $r+1$. In summary, we could find a large enough $P_3$ such that

$$|\hat{V}_{(i,j),r+1} - \hat{V}_{(i,j),r}| \leq b_{o12}(n(r))^{-3/4}$$ (72)

for all large enough $r$.

Now, we discuss on $\hat{\Gamma}_{(i_1,j_1),(i_2,j_2)} = |\hat{V}_{(i_1,j_1)} - \hat{V}_{(i_2,j_2)}|$ when $r$ is large enough.

$$\left|\hat{\Gamma}_{(i_1,j_1),(i_2,j_2)}(r + 1) - \hat{\Gamma}_{(i_1,j_1),(i_2,j_2)}(r)\right|$$

$$= \left|\hat{V}_{(i_1,j_1),r+1} - \hat{V}_{(i_2,j_2),r+1} - (\hat{V}_{(i_1,j_1),r} - \hat{V}_{(i_2,j_2),r})\right|$$

$$\leq \left|\hat{V}_{(i_1,j_1),r+1} - \hat{V}_{(i_1,j_1),r}\right| + \left|\hat{V}_{(i_2,j_2),r+1} - \hat{V}_{(i_2,j_2),r}\right|$$

$$\leq 2b_{o12}(n(r))^{-3/4}.$$

Let $b = 2b_{o12}$ and the lemma is proved.

\[\square\]

**LEMMA 13.** For each two contexts $x_{j_1}$ and $x_{j_2}$, between any two samples for any designs of $x_{j_1}$, the number of samples allocated to $x_{j_2}$ is $O(\sqrt{n(r)} \log \log n(r))$ almost surely.

**Proof.** Fix a sample path and assume $\hat{i}_r^*(x_j) = i^*(x_j) = 1$ for $j = j_1, j_2$ when $r$ is large enough. Let $\{r_g, g = 1, 2, \cdots\}$ be the subsequence where $x_{j_1}$ and $x_{j_2}$ are considered.

Suppose we sample a design of $x_{j_1}$ at iteration $r_g$ and assign a sample to some design of $x_{j_2}$ at iteration $r_{g+1}$. Let

$$d = \inf\{l > 0 : i^{(i_{j_1})}_{r_{g+l}} = 1, \forall i \in \{1, 2, \cdots, k\}\}.$$ 

By definition, we would not consider $x_{j_1}$ but $x_{j_2}$ at iteration $r_{g+l}$ when $0 < l < d$. We study on the upper bound of $d$.

Since we consider context $x_{j_1}$ at iteration $r_g$, there exists a suboptimal design $i_1$ of $x_{j_1}$ that

$$\hat{\tau}_{(i_1,j_1)} = \arg \min_{j \in \{1, 2, \cdots, m\}, i \in \{1, 2, \cdots, k\} \setminus i^*_g(x_j)} \hat{\tau}_{(i,j)}.$$
Therefore, at iteration \( r_g \), we have

\[
\frac{\delta_{(i_1,j_1)}^{r_g}}{S_{(j_1),r_g}^b + S_{(i_1,j_1),r_g}^b} < \frac{\delta_{(i_1,j_2)}^{r_g}}{S_{(j_2),r_g}^b + S_{(i_1,j_2),r_g}^b}, \quad \forall i \in \{2, \cdots, k\}. \tag{73}
\]

Since we consider \( x_{j_2} \) at iteration \( r_{g+1} \), then there exists some suboptimal design \( i_2 \) of \( x_{j_2} \) satisfy

\[
\frac{\delta_{(i_1,j_1)}^{r_{g+1}}}{S_{(j_1),r_{g+1}}^b + S_{(i_1,j_1),r_{g+1}}^b} > \frac{\delta_{(i_2,j_2)}^{r_{g+1}}}{S_{(j_2),r_{g+1}}^b + S_{(i_2,j_2),r_{g+1}}^b} = \frac{\delta_{(i_2,j_2)}^{r_g}}{S_{(j_2),r_g}^b + S_{(i_2,j_2),r_g}^b}. \tag{74}
\]

By \((73), (74)\), it is easy to show that

\[
\frac{\delta_{(i_1,j_1)}^{r_{g+1}}}{S_{(j_1),r_{g+1}}^b + S_{(i_1,j_1),r_{g+1}}^b} < \frac{\delta_{(i_1,j_1)}^{r_g}}{S_{(j_1),r_g}^b + S_{(i_1,j_1),r_g}^b} \leq b_{w1}(n^{r_g})^{1/4},
\]

where \( b_{w1} \) is a positive constant. Therefore,

\[
\frac{\delta_{(i_1,j_1)}^{r_{g+1}}}{S_{(j_1),r_{g+1}}^b + S_{(i_1,j_1),r_{g+1}}^b} \leq \frac{\delta_{(i_2,j_2)}^{r_g}}{S_{(j_2),r_g}^b + S_{(i_2,j_2),r_g}^b} + b_{w1}(n^{r_g})^{1/4}. \tag{75}
\]

Let \( p = \sup\{1 \leq l < d : \hat{\tau}_{(i_2,j_2)} = \arg \min_{j \in \{1,2,\cdots,m\}, i \in \{1,2,\cdots,k\}\setminus \{i_{g+1}(x_j) \}} \hat{\tau}_{(i,j)} \} \). Then, \( r_{g+p} \) is the last iteration before \( r_{g+d} \) that \( x_{j_2} \) is considered because of \( \hat{\tau}_{(i_2,j_2)} \). At iteration \( r_{g+p} \), we must have

\[
\frac{\delta_{(i_1,j_1)}^{r_{g+1}}}{S_{(j_1),r_{g+1}}^b + S_{(i_1,j_1),r_{g+1}}^b} = \frac{\delta_{(i_1,j_1)}^{r_{g+p}}}{S_{(j_1),r_{g+p}}^b + S_{(i_1,j_1),r_{g+p}}^b} \geq \frac{\delta_{(i_2,j_2)}^{r_{g+p}}}{S_{(j_2),r_{g+p}}^b + S_{(i_2,j_2),r_{g+p}}^b}.
\]

By \((75)\), it is necessary to have

\[
\frac{\delta_{(i_2,j_2)}^{r_{g+p}}}{S_{(j_2),r_{g+p}}^b + S_{(i_2,j_2),r_{g+p}}^b} - \frac{\delta_{(i_2,j_2)}^{r_g}}{S_{(j_2),r_g}^b + S_{(i_2,j_2),r_g}^b} \leq b_{w1}(n^{r_g})^{1/4},
\]

53
that is, equivalently,
\[
\delta_{(i_2,j_2)}^{r_{g+p}} S^b_{(j_2),r_g} - \delta_{(i_2,j_2)}^{r_g} S^b_{(j_2),r_g+p} + \delta_{(i_2,j_2)}^{r_{g+p}} S_{(i_2,j_2),r_g} - \delta_{(i_2,j_2)}^{r_g} S_{(i_2,j_2),r_g+p}
\leq b_w \frac{n(r_g)^{1/4}}{n(i_2,j_2, r_g)}
\]  
(76)

Let \(h_i = \hat{n}_{(i,j_2),r_g+p} - \hat{n}_{(i,j_2),r_g}, \ i = 1, i_2\). We will discuss the left hand side of (76) by piece.

\[
\delta_{(i_2,j_2)}^{r_{g+p}} S^b_{(j_2),r_g} - \delta_{(i_2,j_2)}^{r_g} S^b_{(j_2),r_g+p}
\]

\[
= \delta_{(i_2,j_2)}^{r_{g+p}} S^b_{(j_2),r_g} - \delta_{(i_2,j_2)}^{r_g} S^b_{(j_2),r_g+p} + \delta_{(i_2,j_2)}^{r_{g+p}} S_{(i_2,j_2),r_g} - \delta_{(i_2,j_2)}^{r_g} S_{(i_2,j_2),r_g+p}
\]

(77)

For the first part,

\[
\delta_{(i_2,j_2)}^{r_{g+p}} S^b_{(j_2),r_g} - \delta_{(i_2,j_2)}^{r_g} S^b_{(j_2),r_g+p}
\]

\[
= \delta_{(i_2,j_2)}^{r_{g+p}} \left( \frac{\sigma^2_{1,r_g}(x_{j_2})}{\hat{n}_{(i,j_2),r_g}} - \frac{\sigma^2_{1,r_g}(x_{j_2})}{\hat{n}_{(i,j_2),r_g+p}} + \frac{\sigma^2_{1,r_g}(x_{j_2})}{\hat{n}_{(i,j_2),r_g+p}} - \frac{\sigma^2_{1,r_g+p}(x_{j_2})}{\hat{n}_{(i,j_2),r_g+p}} \right)
\]

\[
= \delta_{(i_2,j_2)}^{r_{g+p}} \left( \frac{\sigma^2_{1,r_g}(x_{j_2})}{\hat{n}_{(i,j_2),r_g}} + \frac{1}{\hat{n}_{(i,j_2),r_g+p}} \left( \sigma^2_{1,r_g}(x_{j_2}) - \sigma^2_{1,r_g+p}(x_{j_2}) \right) \right)
\]

\[
\geq h_1 \frac{b_{w2}}{n_{(i,j_2),r_g}} - \frac{b_{w3}}{n_{(i,j_2),r_g+p}} \sqrt{\log \log \hat{n}_{(i,j_2),r_g}}
\]

(78)

where \(b_{w2}, b_{w3} > 0\). At the same time,

\[
\delta_{(i_2,j_2)}^{r_{g+p}} S^b_{(j_2),r_g} - \delta_{(i_2,j_2)}^{r_g} S^b_{(j_2),r_g+p}
\]

\[
= (\delta_{(i_2,j_2)}^{r_{g+p}} - \delta_{(i_2,j_2)}^{r_g}) S^b_{(j_2),r_g+p} \geq -b_{w4} \frac{1}{n_{(i,j_2),r_g+p}} \sqrt{\log \log \hat{n}_{(i,j_2),r_g}}
\]

\[
\geq - b_{w5} \frac{1}{n_{(i,j_2),r_g+p}} \sqrt{\log \log \hat{n}_{(i,j_2),r_g}}
\]

(79)

The last equation holds because of lemma 9. Take (79) and (78) into (77), we get

\[
\delta_{(i_2,j_2)}^{r_{g+p}} S^b_{(j_2),r_g} - \delta_{(i_2,j_2)}^{r_g} S^b_{(j_2),r_g+p} \geq h_1 \frac{b_{w2}}{n_{(i,j_2),r_g}} - \frac{b_{w3} + b_{w5}}{n_{(i,j_2),r_g+p}} \sqrt{\log \log \hat{n}_{(i,j_2),r_g}}
\]

54
Therefore,

\[
\frac{\delta_{r_{g+d}} - \delta_{r_{g}}}{S_{(j_2),r_{g+d}} + S_{(i_2,j_2),r_{g}} + S_{(i_2,j_2),r_{g+d}}}
\]

\[
\geq h_1 n_{(1,j_2),r_{g+d}} n_{(1,j_2),r_{g}}
\]

\[
= b_{w8} h_1 = b_{w9} \sqrt{n(r_g) \log \log n(r_g)}
\]

(80)

(81)

where \(b_{w6}\) and \(b_{w7}\) are positive constants and (80) is got by multiply \(\hat{n}_{(1,j_2),r_{g}} \hat{n}_{(1,j_2),r_{g+d}}\) on both nominator and denominator. (81) holds because of lemma [9].

By the same way, we can show that

\[
\frac{\delta_{r_{g+d}} - \delta_{r_{g}}}{S_{(j_2),r_{g+d}} + S_{(i_2,j_2),r_{g}} + S_{(i_2,j_2),r_{g+d}}}
\]

\[
\geq b_{w8} h_2 = b_{w9} \sqrt{n(r_g) \log \log n(r_g)},
\]

(82)

where \(b_{w8}, b_{w9} > 0\). Take (81) and (82) into (76), we have

\[
b_{w6} h_1 + b_{w8} h_2 - b_{w10} \sqrt{n(r_g) \log \log n(r_g)} \leq b_{w1} (n(r_g))^{1/4},
\]

where \(b_{w6}, b_{w8}, b_{w10} > 0\). It means that

\[
b_{w2} h_1 + b_{w3} h_2 = O(\sqrt{n(r_g) \log \log n(r_g)}).
\]

Therefore, \(h_1 = O(\sqrt{n(r_g) \log \log n(r_g)})\) and \(h_2 = O(\sqrt{n(r_g) \log \log n(r_g)})\). Then,

\[
\hat{n}_{(i_2,j_2),r_{g+d}} - \hat{n}_{(i_2,j_2),r_{g}} = h_2 \quad \text{or} \quad h_2 + 1,
\]

which is of order \(O(\sqrt{n(r_g) \log \log n(r_g)})\). By discussing at the \(r_{g+d-1}\)th iteration, we have

\[
\hat{n}_{(1,j_2),r_{g+d-1}} - \hat{n}_{(1,j_2),r_{g}} = O(\sqrt{n(r_g) \log \log n(r_g)}),
\]

55
which means \( \hat{n}(1,j_2),r_{g+d} - \hat{n}(1,j_2),r_g = O\left( \sqrt{n^{(r_g)} \log \log n^{(r_g)}} \right) \). Since the number of design is finite, it is easy to show that

\[
d - 1 = \sum_{i=1}^{k} \left( \hat{n}(i,j_2),r_{g+d} - \hat{n}(i,j_2),r_g \right) = O\left( \sqrt{n^{(r_g)} \log \log n^{(r_g)}} \right).
\]

The lemma is proved.

Let \( k_{(i,j),(r,r+u)} \triangleq \hat{n}_{(i,j),r+u} - \hat{n}_{(i,j),r}, \ i = 1, 2, \ldots, l, j = 1, 2, \ldots, m \). For notation convenience, we abbreviate \( k_{(i,j),(r,r+u)} \) by \( k_{(i,j)}^{(r,u)} \).

**Lemma 14.** Suppose that some suboptimal design \( i_1 \) of context \( x_j \) is sampled at iteration \( r \geq 3 \). Define

\[
t \triangleq \inf \{ l > 0 : I_{r+l}^{(i_1,j)} = 1 \}, \quad s' \triangleq \sup \{ l < t : I_{r+l}^{(1,j)} = 1 \}, \quad s \triangleq \sup \{ l < s' : I_{r+l}^{(i,j)} = 1, \forall i > 1 \}.
\]

For any positive constant \( c_1 \), if there exists a sufficiently large positive constant \( c_2 \) (dependent on \( c_1 \), but independent of \( r \)) for which \( k_s^{(1,j)} \geq c_2 \sqrt{n^{(r)}} \log \log n^{(r)} \) holds. Then, there exists a suboptimal design \( i_2 \neq i_1 \) of context \( x_j \) and a iteration \( r + u \), where \( u \leq s \), such that \( i_2 \) of \( x_j \) is sampled at iteration \( r + u \) and

\[
(1 + c_1 \sqrt{\frac{\log \log n^{(r)}}{n^{(r)}}}) \frac{\hat{n}_{(i_2,j),r}}{\hat{n}_{(1,j),r}} < \frac{\hat{n}_{(i_2,j),r} + k_{(i_2,j)}^{(r,u)}}{\hat{n}_{(1,j),r} + k_{(1,j)}^{(r,u)}} \leq \frac{\hat{n}_{(i_2,j),r} + k_{(i_2,j)}^{(r,u)}}{\hat{n}_{(1,j),r} + k_{(1,j)}^{(r,u)}}
\]

holds almost surely.

**Proof.** This proof is partly similar to proof of lemma 6.1 in Chen and Ryzhov (2018). However, we derive it with unknown, rather than known, variance assumption. Moreover, we consider the subsequence where \( x_j \) is sampled, rather than the whole sequence.

It is clear that \( t = O(n^{(r)}) \). Otherwise, we will have contradiction because \( \frac{\hat{n}_{(i_1,j),r}}{n^{(r)}} = \Theta(1) \). Here, we assume that \( k_s^{(1,j)} < t < c_{up} n^{(r)} \) where \( c_{up} \) is a fixed positive constant.
By definition of $s$, $r$ is the last iteration before $b$ where
\[ r = \infty \]
From the definition, $r + s'$ is the last iteration before $r + t$ that $(1, j)$ is sampled and $r + s$ is the last iteration before $r + s'$ that a suboptimal design (assume it is $i_3$) of $x_j$ is sampled. Let
\[ s_1 = \inf\{l > s : I_{r+l}^{(1,j)} = 1\} \]
which means $r + s_1$ is the first iteration after $r + s$ that $(1, j)$ is sampled.

By definition of $s$ and $s_1$, there is no sample of $x_j$ between $r + s + 1$ and $r + s_1 - 1$. By lemma 13, 
\[ \frac{n^{(r+s_1)}}{n^{(r+s)}} = 1 + \frac{n^{(r+s_1)} - n^{(r+s)}}{n^{(r+s)}} = 1 + \frac{O(\sqrt{n^{(r+s)} \log \log n^{(r+s)}})}{n^{(r+s)}} \to 1, \quad \text{as } r \to \infty. \]

Then, we have
\[
\Delta_{r+s_1,j} = \left( \frac{\hat{n}^{(1,j),i,j}/\hat{\sigma}_{1,r+s_1}(x_j)}{n^{(r+s_1)}} \right)^2 - \sum_{i\neq 1} \left( \frac{\hat{n}^{(i,j),i,j}/\hat{\sigma}_{i,r}(x_j)}{n^{(r+s_1)}} \right)^2 
\]
\[
= \left( \frac{\hat{n}^{(1,j),i,j}/\hat{\sigma}_{1,r+s}(x_j)}{n^{(r+s_1)}} \right)^2 - \sum_{i=2,i\neq i_3} M \left( \frac{\hat{n}^{(i,j),i,j}/\hat{\sigma}_{i,r+s}(x_j)}{n^{(r+s_1)}} \right)^2 
\]
\[
= \left( \frac{\hat{n}^{(1,j),i,j}/\hat{\sigma}_{1,r+s}(x_j)}{n^{(r+s_1)}} \right)^2 \Delta_{r+s_1,j} + \frac{\hat{n}^{2}_{(1,j),r+s}_x}{\hat{\sigma}^2_{1,r+s}(x_j)(n^{(r+s_1)})^2} - \frac{\hat{n}^{2}_{(1,j),r+s_1}_x}{\hat{\sigma}^2_{1,r+s_1}(x_j)(n^{(r+s_1)})^2} 
\]
\[
= \left( \frac{\hat{n}^{(r+s)}_x}{n^{(r+s_1)}} \right)^2 \Delta_{r+s_1,j} + \frac{\hat{n}^{2}_{(1,j),r+s_1}_x}{(n^{(r+s_1)})^2} \frac{\hat{\sigma}^2_{1,r+s}(x_j)}{\hat{\sigma}^2_{1,r+s_1}(x_j)} - \frac{\hat{n}^{2}_{(1,j),r+s_1}_x}{\hat{\sigma}^2_{1,r+s_1}(x_j)} 
\]
\[
\geq \left( \frac{\hat{n}^{(r+s)}_x}{n^{(r+s_1)}} \right)^2 \Delta_{r+s_1,j} - \frac{b_{s_1}}{(n^{(r+s_1)})^{3/4}}, \quad \text{(87)}
\]
where $b_{s_1} > 0$. \(86\) holds because $\hat{n}^{(1,j),r+s} = \hat{n}^{(i_3,j),r+s_1} - 1$ and equation \(87\) comes from lemma 11 and 9. Therefore, by \(85\),
\[
\Delta_{r+s_1,j} \leq \left( \frac{n^{(r+s_1)}}{n^{(r+s)}} \right)^2 \Delta_{r+s_1,j} + \frac{b_{s_1} n^{(r+s_1)}}{(n^{(r+s)})^2} \leq \frac{b_{s_2}}{(n^{(r+s)})^{3/4}},
\]
57
where \( b_{s2} > 0 \). Note that \( k_s^{(i_1,j)} = 1 \), whence

\[
\sum_{i > 2, i \neq i_1} \frac{\left( \frac{\hat{n}_{(i,j),r} + k_s^{(i,j)}}{\hat{n}_{(1,j),r} + k_s^{(1,j)}} \right)^2}{\hat{n}_{(1,j),r} + k_s^{(1,j)}} + \frac{b_{s2}}{(n(r+s))^{3/4}} \left( \frac{n(r+s)}{\hat{n}_{(1,j),r} + k_s^{(1,j)}} \right)^2 \geq 1.
\]

From lemma 9, which implies \( \hat{n}_{(1,j),r} = \Theta(n(r)) \), we could find some positive constant \( b_{s3} \) such that

\[
b_{s2} \left( \frac{n(r+s)}{\hat{n}_{(1,j),r} + k_s^{(1,j)}} \right)^2 \leq b_{s3},
\]

whence

\[
\sum_{i > 2, i \neq i_1} \frac{\left( \frac{\hat{n}_{(i,j),r} + k_s^{(i,j)}}{\hat{n}_{(1,j),r} + k_s^{(1,j)}} \right)^2}{\hat{n}_{(1,j),r} + k_s^{(1,j)}} \geq 1 - \frac{\left( \frac{\hat{n}_{(i_1,j),r} + 1}{\hat{n}_{(1,j),r} + k_s^{(1,j)}} \right)^2}{(n(r+s))^{3/4}}.
\]

Therefore,

\[
\sum_{i > 2, i \neq i_1} \left[ \frac{\left( \frac{\hat{n}_{(i,j),r} + k_s^{(i,j)}}{\hat{n}_{(1,j),r} + k_s^{(1,j)}} \right)^2}{\hat{n}_{(1,j),r} + k_s^{(1,j)}} \right] \geq \frac{\left( \frac{\hat{n}_{(i_1,j),r} + 1}{\hat{n}_{(1,j),r} + k_s^{(1,j)}} \right)^2}{(n(r+s))^{3/4}} - \frac{b_{s3}}{(n(r+s))^{3/4}}.
\]

(88)
By law of the iterated logarithm, we have

\[
\hat{\sigma}^2_{i,r,s}(x_j) \hat{n}^2_{i,j,r+s}(x_j) \left( \frac{2\hat{n}_{i,j,r} k^{(1,j)}}{2} + \frac{1}{2}(k^{(1,j)})^2 \right) - \frac{b_{s3}}{(n^{(r+s)})^{3/4}},
\]

where (88) comes from (84). (89) holds because \( k^{(1,j)} > c_2 \sqrt{n^{(r)} \log \log n^{(r)}} \) and \( \hat{n}_{i,j,r} = \Theta(\hat{n}_{i,j,r}) \).

By law of the iterated logarithm, we have

\[
\left( \frac{\hat{n}_{i,j,r}}{\hat{n}_{i,j,r+s}} \right)^2 \frac{\hat{\sigma}^2_{i,r,s}(x_j)}{\hat{\sigma}^2_{i,r,s}(x_j)} - \frac{b_{s4}}{(n^{(r+s)})^{3/4}} > -b_{s4} \sqrt{\frac{\log \log n^{(r)}}{n^{(r)}}},
\]

for some positive constant \( b_{s4} \). Meanwhile,

\[
\frac{\hat{\sigma}^2_{i,r,s}(x_j)}{\hat{\sigma}^2_{i,r,s}(x_j)} \hat{n}^2_{i,j,r+s}(x_j) \left( \frac{2\hat{n}_{i,j,r} k^{(1,j)}}{2} + \frac{1}{2}(k^{(1,j)})^2 \right) - \frac{b_{s5}}{(n^{(r+s)})^{3/4}} \geq \frac{b_{s5}}{(\hat{n}_{i,j,r} + k^{(1,j)})^2} \left( 2\hat{n}_{i,j,r} k^{(1,j)} + (k^{(1,j)})^2 - \frac{b_{s6}(\hat{n}_{i,j,r} + k^{(1,j)})^2}{(n^{(r+s)})^{3/4}} \right)
\]

\[
\geq \frac{b_{s5}}{(\hat{n}_{i,j,r} + k^{(1,j)})^2} \left( 2\hat{n}_{i,j,r} k^{(1,j)} + (k^{(1,j)})^2 - \frac{b_{s6}(\hat{n}_{i,j,r} + \log n^{(r)})^2}{(n^{(r)})^{3/4}} \right)
\]

\[
\geq \frac{b_{s5}}{(\hat{n}_{i,j,r} + k^{(1,j)})^2} \left( 2\hat{n}_{i,j,r} k^{(1,j)} - 2b_{s7}\hat{n}_{i,j,r}^{5/4} \right)
\]

\[
= \frac{b_{s5}}{(\hat{n}_{i,j,r} + k^{(1,j)})^2} \left( 2\hat{n}_{i,j,r} k^{(1,j)} - b_{s7}\hat{n}_{i,j,r}^{1/4} \right)
\]

\[
\geq \frac{b_{s8}}{\hat{n}_{i,j,r}} (k^{(1,j)} - b_{s7}\hat{n}_{i,j,r}^{1/4})
\]

\[
\geq \frac{b_{s8}}{\hat{n}_{i,j,r}} (c_2 \sqrt{\frac{n^{(r)} \log \log n^{(r)}}{n^{(r)}}} - b_{s7}\hat{n}_{i,j,r}^{1/4})
\]

\[
\geq b_{s9} c_2 \sqrt{\frac{\log \log n^{(r)}}{n^{(r)}}},
\]

where (91) and (95) come from lemma 9. (93) and (94) holds because \( k^{(1,j)} < c_{up} n^{(r)} \). Note that \( b_{s9} \) is a positive constant.

Therefore, by (90) and (95), We could select \( c_2 > \frac{b_{s4}}{b_{s9}} + 1 \) such that

\[
\sum_{i > 2, j \neq i} \left[ \left( \frac{\hat{n}_{i,j,r} + k^{(1,j)}_{i,j}}{\hat{n}_{i,j,r + k^{(1,j)}}} \right)^2 \left( \frac{\hat{n}_{i,j,r}}{\hat{n}_{i,j,r + k^{(1,j)}}} \right)^2 \right]
\]
The last inequality holds because
\[ |\hat{\sigma}_{2,r+s}(x_j) - \hat{\sigma}_{2,r}(x_j)| = O(\sqrt{\log \log n(r)}) \] and \( \frac{\hat{n}_{(i_2,j),r}}{\hat{n}_{(1,j),r}} = \Theta(1) \). Then, we could further select \( c_2 > \frac{b_{s12}(m-2)}{b_{s9}b_{s11}} + \frac{b_{s4}}{b_{s9}} + 1 \) and

\[
\left( \frac{\hat{n}_{(i_2,j),r} + k_s^{(i_2,j)}}{\hat{n}_{(1,j),r} + k_s^{(1,j)}} \right)^2 - \left( \frac{\hat{n}_{(i_2,j),r}}{\hat{n}_{(1,j),r}} \right)^2 > \frac{b_{s11}b_{s10}}{m-2} \sqrt{\frac{\log \log n(r)}{n(r)}} + \frac{\hat{\sigma}_{2,r+s}^2(x_j)}{\hat{\sigma}_{2,r}^2(x_j) - \hat{\sigma}_{1,r+s}^2(x_j) - 1}
\]

where \( b_{s10} \) is a positive constant. So there must be a suboptimal design \( i_2 \) such that

\[
\left( \frac{\hat{n}_{(i_2,j),r} + k_s^{(i_2,j)}}{\hat{n}_{(1,j),r} + k_s^{(1,j)}} \right)^2 - \left( \frac{\hat{n}_{(i_2,j),r}}{\hat{n}_{(1,j),r}} \right)^2 > \frac{b_{s11}b_{s10}}{m-2} \sqrt{\frac{\log \log n(r)}{n(r)}},
\]

that is, equivalently,

\[
\left( \frac{\hat{n}_{(i_2,j),r} + k_s^{(i_2,j)}}{\hat{n}_{(1,j),r} + k_s^{(1,j)}} \right)^2 - \left( \frac{\hat{n}_{(i_2,j),r}}{\hat{n}_{(1,j),r}} \right)^2 > \frac{b_{s11}b_{s10}}{m-2} \sqrt{\frac{\log \log n(r)}{n(r)}} - b_{s12} \sqrt{\frac{\log \log n(r)}{n(r)}}.
\]

The last inequality holds because \( |\hat{\sigma}_{2,r+s}^2(x_j) - \hat{\sigma}_{2,r}^2(x_j)| = O(\sqrt{\log \log n(r)}) \) and \( \frac{\hat{n}_{(i_2,j),r}}{\hat{n}_{(1,j),r}} = \Theta(1) \). Then, we could further select \( c_2 > \frac{b_{s12}(m-2)}{b_{s9}b_{s11}} + \frac{b_{s4}}{b_{s9}} + 1 \) and

\[
\left( \frac{\hat{n}_{(i_2,j),r} + k_s^{(i_2,j)}}{\hat{n}_{(1,j),r} + k_s^{(1,j)}} \right)^2 - \left( \frac{\hat{n}_{(i_2,j),r}}{\hat{n}_{(1,j),r}} \right)^2 > \frac{b_{s11}b_{s10}}{m-2} \sqrt{\frac{\log \log n(r)}{n(r)}} + \frac{b_{s12}}{m-2} \sqrt{\frac{\log \log n(r)}{n(r)}}
\]

where \( b_{s13} > 0 \) is a constant. Then, because \( \left( \frac{\hat{n}_{(i_2,j),r}}{\hat{n}_{(1,j),r}} \right)^2 \) is \( \Theta(1) \),

\[
\left( \frac{\hat{n}_{(i_2,j),r} + k_s^{(i_2,j)}}{\hat{n}_{(1,j),r} + k_s^{(1,j)}} \right)^2 > 1 + b_{s13} \sqrt{\frac{\log \log n(r)}{n(r)}}.
\]

Let \( b_{s15} = b_{s14}b_{s13}/4 \), for all large enough \( r \), we have

\[
\left( \frac{\hat{n}_{(i_2,j),r} + k_s^{(i_2,j)}}{\hat{n}_{(1,j),r} + k_s^{(1,j)}} \right)^2 > 1 + b_{s15} \sqrt{\frac{\log \log n(r)}{n(r)}},
\]

60
whence
\[ \frac{\hat{n}_{(i_2,j),r} + k_s^{(i_2,j)}}{\hat{n}_{(1,j),r} + k_s^{(1,j)}} > \left(1 + b_{s15} \sqrt{\frac{\log \log n^{(r)}}{n^{(r)}}}\right) \frac{\hat{n}_{(i_2,j),r}}{\hat{n}_{(1,j),r}}. \]

Let \( u \triangleq \sup \{l \leq s : l^{(i_2,j)}_{r+l} = 1\} \). Then, \( r + u \) is the last iteration before \( r + s \) that design \( i_2 \) of \( x_j \) is considered. Since \( k_l^{(i,j)} \) increases monotonically with \( l \) for all \( i = 1, 2 \ldots, k \), we have
\[
\frac{\hat{n}_{(i_2,j),r} + k_u^{(i_2,j)}}{\hat{n}_{(1,j),r} + k_u^{(1,j)}} = \frac{\hat{n}_{(i_2,j),r} + k_u^{(i_2,j)} - 1}{\hat{n}_{(1,j),r} + k_u^{(1,j)}}
\[
= \left(1 - \frac{1}{\hat{n}_{(i_2,j),r} + k_u^{(i_2,j)}}\right) \frac{\hat{n}_{(i_2,j),r} + k_u^{(i_2,j)}}{\hat{n}_{(1,j),r} + k_u^{(1,j)}}
\[
> \left(1 - \frac{1}{\hat{n}_{(i_2,j),r} + k_u^{(i_2,j)}}\right) \left(1 + b_{s15} \sqrt{\frac{\log \log n^{(r)}}{n^{(r)}}}\right) \frac{\hat{n}_{(i_2,j),r}}{\hat{n}_{(1,j),r}}
\]

Since \( \frac{\hat{n}_{(i_2,j),r}}{n^{(r)}} = \Theta(1) \), for all large enough \( r \),
\[
\left(1 - \frac{1}{\hat{n}_{(i_2,j),r} + k_u^{(i_2,j)}}\right) \left(1 + b_{s15} \sqrt{\frac{\log \log n^{(r)}}{n^{(r)}}}\right) \frac{\hat{n}_{(i_2,j),r}}{\hat{n}_{(1,j),r}}
\[
> \left(1 - \frac{b_{s16}}{n^{(r)}}\right) \left(1 + b_{s15} \sqrt{\frac{\log \log n^{(r)}}{n^{(r)}}}\right) \frac{\hat{n}_{(i_2,j),r}}{\hat{n}_{(1,j),r}}
\[
= \left(1 + b_{s15} \sqrt{\frac{\log \log n^{(r)}}{n^{(r)}}}\right) \frac{\hat{n}_{(i_2,j),r}}{\hat{n}_{(1,j),r}} - b_{s16} b_{s15} \sqrt{\log \log n^{(r)}} \frac{\sqrt{n^{(r)}}}{n^{(r)^2}} \frac{\hat{n}_{(i_2,j),r}}{\hat{n}_{(1,j),r}}
\[
\geq \left(1 + b_{s15} \sqrt{\frac{\log \log n^{(r)}}{n^{(r)}}}\right) \frac{\hat{n}_{(i_2,j),r}}{\hat{n}_{(1,j),r}}
\]

By definition,
\[
\frac{b_{s15}}{2} = \frac{b_{s14} b_{s13}}{8} = \frac{b_{s14}}{8} \left(\frac{b_{s11} b_{s10}}{m-2} - b_{s12}\right) = \frac{b_{s14}}{8} \left(\frac{b_{s11} (b_{s9} c_2 - b_{s4})}{m-2} - b_{s12}\right),
\]

where \( b_{s4}, b_{s9}, b_{s11}, b_{s12}, b_{s14} \) are positive constant, which do not depend on \( c_2 \).
Therefore, for all large enough \( r \), if
\[
c_2 > \frac{m - 2}{b_{s9}b_{s11}} \left( \frac{8c_1}{b_{s14}} + b_{s12} \right) + \frac{b_{s4}}{b_{s9}},
\]
then \( \frac{b_{s9}}{b_{s4}} > c_1 \) and we have
\[
\frac{\hat{n}_{(i_2,j),r} + k_u^{(i_2,j)}}{\hat{n}_{(1,j),r} + k_s^{(1,j)}} > \left( 1 + c_1 \sqrt{\frac{\log \log n(r)}{n(r)}} \right) \frac{\hat{n}_{(i_2,j),r}}{\hat{n}_{(1,j),r}},
\]
completing the proof.

\[\square\]

**Lemma 15.** Suppose that some suboptimal design \( i_1 \neq 1 \) of context \( x_j \) is sampled at iteration \( r \geq 3 \), define
\[
t \triangleq \inf \{ l > 0 : I_{r+l}^{(i_1,j)} = 1 \}, \quad s' \triangleq \sup \{ l < t : I_{r+l}^{(1,j)} = 1 \}, \quad s \triangleq \sup \{ l < s' : I_{r+l}^{(i,j)} = 1, \forall i > 1 \}.
\]

Then, \( k_t^{(1,j)} = O(\sqrt{n(r) \log \log n(r)}) \) almost surely.

**Proof.** We will prove the result by contradiction.

Based on our algorithm, at iteration \( r \), we have:
\[
0 < \tilde{\tau}_{(i,j),r} = \frac{\tilde{\delta}_{(i_1,j)}^{r}}{S_{(j),r} + S_{(i_1,j),r}} < \frac{\tilde{\delta}_{(i_1,j)}^{r}}{S_{(j),r} + S_{(i,j),r}} = \hat{\tau}_{(i,j),r},
\]
for any \( i > 1 \) and \( i \neq i_1 \). It means that
\[
\tilde{\delta}_{(i_1,j)}^{r} \left( S_{(j),r} + S_{(i_1,j),r} \right) < \tilde{\delta}_{(i,j)}^{r} \left( S_{(j),r} + S_{(i,j),r} \right).
\]

If the conclusion of this lemma does not hold, for any positive constant \( c'_2 \), we could find an iteration \( r \) where \( (i^r, j^r) = (i_1, j) \) such that
\[
k_t^{(1,j)} > c'_2 \sqrt{n(r) \log \log n(r)}.
\]

By lemma 10, \( k_t^{(1,j)} - k_s^{(1,j)} = O(\sqrt{n(r) \log \log n(r)}) \). Thus, for any constant \( c_2 > 0 \), we could find
iteration \( r \) such that
\[
k_{s}^{(1,j)}(r) > c_2 \sqrt{n(r) \log \log n(r)}.
\] (97)

By lemma 14, since the corresponding \( c_2 \) could be found for any constant \( c_1 \), we could have an iteration \( r + u \) and a suboptimal design \( i_2 \neq i_1 \) of \( x_j \) such that \((i^{(r+u)}, j^{(r+u)}) = (i_2, j)\), \( u \leq s \) and (83) holds.

To prove this lemma by contradiction, we claims that design \( i_2 \) of context \( x_j \) could not be sampled at iteration \( r + u \) when the corresponding \( c_2 \) is large enough. To prove this claim, we will show that
\[
\hat{\delta}^{r+u}(i_1, j)\left(S_{(j), r+u}^{b} + S_{(i_2, j), r+u} \right) < \hat{\delta}^{r+u}(i_2, j)\left(S_{(j), r+u}^{b} + S_{(i_1, j), r+u} \right).
\] (98)

**CASE 1:** If \( \delta_{(i_2, j)}/\delta_{(i_1, j)} > 1 \), we have \( \hat{\delta}^{r}(i_2, j)/\hat{\delta}^{r}(i_1, j) > 1 \) when \( r \) is large enough. By equation (96),
\[
\hat{\delta}^{r+u}(i_1, j)S_{(i_2, j), r+u} = \frac{\hat{\delta}^{r+u}(i_1, j)S_{(i_2, j), r}}{\hat{\delta}^{r}(i_1, j)} \frac{S_{(i_2, j), r+u}}{S_{(i_2, j), r}} < \left(\frac{\hat{\delta}^{r+u}(i_2, j)S_{(j), r}}{\hat{\delta}^{r}(i_1, j)} + \hat{\delta}^{r+u}(i_2, j)S_{(i_1, j), r} \right) S_{(i_2, j), r+u}.
\]

Therefore, to satisfy (98), it is sufficient to show that
\[
\frac{\hat{\delta}^{r+u}(i_1, j)}{\hat{\delta}^{r}(i_1, j)} \left(\hat{\delta}^{r}(i_2, j) - \hat{\delta}^{r}(i_1, j)\right)S_{(j), r+u} + \hat{\delta}^{r}(i_1, j)S_{(i_1, j), r} \right) S_{(i_2, j), r+u} < \left(\frac{\hat{\delta}^{r+u}(i_2, j)S_{(j), r}}{\hat{\delta}^{r}(i_1, j)} + \hat{\delta}^{r+u}(i_2, j)S_{(i_1, j), r} \right) S_{(i_2, j), r+u}.
\] (99)

To prove (99), it is sufficient to have
\[
\frac{\hat{\delta}^{r+u}(i_1, j)}{\hat{\delta}^{r}(i_1, j)} \left(\hat{\delta}^{r}(i_2, j) - \hat{\delta}^{r}(i_1, j)\right)S_{(j), r+u} < \left(\frac{\hat{\delta}^{r+u}(i_2, j)S_{(j), r}}{\hat{\delta}^{r}(i_1, j)} \right) S_{(i_2, j), r+u},
\]
\[
\frac{\hat{\delta}^{r+u}(i_1, j)}{\hat{\delta}^{r}(i_1, j)} \frac{\hat{\delta}^{r}(i_2, j)S_{(i_1, j), r+u} \hat{\delta}^{r+u}(i_1, j)}{\hat{\delta}^{r}(i_1, j)} S_{(i_2, j), r+u} < \frac{\hat{\delta}^{r+u}(i_2, j)S_{(i_1, j), r+u} \hat{\delta}^{r+u}(i_2, j)}{\hat{\delta}^{r}(i_1, j)} S_{(i_2, j), r+u},
\]
that is, equivalently,
\[
\frac{\hat{\delta}^{r+u}(i_1, j)}{\hat{\delta}^{r}(i_1, j)} \left(\hat{\delta}^{r}(i_2, j) - \hat{\delta}^{r}(i_1, j)\right) \frac{\hat{\delta}^{r+u}(i_1, j)}{\hat{\delta}^{r}(i_1, j)} S_{(j), r+u} \hat{\delta}^{r+u}(i_1, j) \hat{\delta}^{r+u}(i_2, j) S_{(i_2, j), r+u} < \hat{\delta}^{r+u}(i_2, j) S_{(i_2, j), r+u} + k_u^{(i_2, j)},
\] (100)
\[
\frac{\hat{\delta}_{(i_1,j)}^{r+u}}{\hat{\delta}_{(i_1,j)}^r} - \frac{\hat{\delta}_{(i_1,j)}^{r+u} \hat{\sigma}_{i_1,r}(x_j)}{\hat{\delta}_{(i_1,j)}^r \hat{\sigma}_{i_1,r+u}(x_j)} \frac{\hat{n}_{(i_1,j),r+u}(x_j)}{\hat{n}_{(i_1,j),r}} - \frac{\hat{n}_{(i_2,j),r}}{\hat{n}_{(i_2,j),r} + k_{(i_2,j)}} < 1.
\] (101)

Let \(d_{(i_1,j)}^{r} = \bar{Y}_{i_1,r}(x_j) - y_{i_1}(x_j)\). For large enough \(r\) and any design \(i_1 \neq 1\), we have

\[
\left|\hat{\delta}_{(i_1,j)}^{r+u} - \hat{\delta}_{(i_1,j)}^r\right| = (\bar{Y}_{i_1,r+u}(x_j) - \bar{Y}_{1,r+u}(x_j))^2 - (\bar{Y}_{i_1,r}(x_j) - \bar{Y}_{1,r}(x_j))^2
\]
\[
= |\bar{Y}_{i_1,r+u}(x_j) - \bar{Y}_{1,r+u}(x_j)| \cdot |\bar{Y}_{i_1,r+u}(x_j) - \bar{Y}_{1,r+u}(x_j)|
\]
\[
\leq |\bar{Y}_{i_1,r+u}(x_j) - \bar{Y}_{1,r+u}(x_j)| + |\bar{Y}_{i_1,r}(x_j) - \bar{Y}_{1,r}(x_j)|
\]
\[
\cdot \left( |d_{(i_1,j)}^{r+u}| + |d_{(i_2,j)}^{r+u}| + |d_{(i_1,j)}^r| + |d_{(i_2,j)}^r| \right)
\]
\[
= O\left( \sqrt{\frac{\log \log \hat{n}_{(i_1,j),r}}{\hat{n}_{(i_1,j),r}}} \right) + O\left( \sqrt{\frac{\log \log \hat{n}_{(1,j),r}}{\hat{n}_{(1,j),r}}} \right)
\] (102)
\[
= O\left( \frac{\log \log n^{(r)}}{n^{(r)}} \right). \] (103)

Equation (102) holds with probability 1 because of law of iterated logarithm. Equation (103) holds due to lemma \[9\]. Then, we have

\[
\frac{\hat{\delta}_{(i_1,j)}^{r+u}}{\hat{\delta}_{(i_1,j)}^r} = \frac{\hat{\delta}_{(i_1,j)}^{r+u} \hat{\sigma}_{i_1,r}(x_j)}{\hat{\delta}_{(i_1,j)}^r \hat{\sigma}_{i_1,r+u}(x_j)} \frac{\hat{n}_{(i_1,j),r+u}(x_j)}{\hat{n}_{(i_1,j),r}} = 1 + O\left( \sqrt{\frac{\log \log n^{(r)}}{n^{(r)}}} \right).
\]

Similarly, \(\frac{\hat{\delta}_{(i_2,j)}^{r+u}}{\hat{\delta}_{(i_2,j)}^r} = 1 + O\left( \sqrt{\frac{\log \log n^{(r)}}{n^{(r)}}} \right)\) and \(\frac{\hat{\delta}_{(i_1,j)}^{r+u}}{\hat{\delta}_{(i_1,j)}^r} = 1 + O\left( \sqrt{\frac{\log \log n^{(r)}}{n^{(r)}}} \right)\). Moreover, by \[54\],

\[
\frac{\hat{\sigma}_{i_1,r}(x_j)}{\hat{\sigma}_{i_1,r+u}(x_j)} = 1 + O\left( \sqrt{\frac{\log \log n^{(r)}}{n^{(r)}}} \right), i = 1, i_1, i_2.
\]
For inequality (101),

\[
\begin{align*}
&\frac{\hat{\delta}_{i(1,j)}^{r+u}(\hat{\delta}_{i(2,j)}^{r}) - \hat{\delta}_{i(1,j)}^{r}}{\hat{\delta}_{i(1,j)}^{r}} \leq \frac{\sigma_{i(1,r)}^{2}(x_j) \sigma_{i(2,r+u)}^{2}(x_j) \hat{n}_{(1,j),r+u}}{\hat{n}_{(1,j),r}} \hat{n}_{(2,j),r} - \hat{n}_{(2,j),r} \\
&= \frac{\hat{\delta}_{i(1,j)}^{r+u}(\hat{\delta}_{i(2,j)}^{r}) - \hat{\delta}_{i(1,j)}^{r}}{\hat{\delta}_{i(1,j)}^{r}} \leq \frac{\sigma_{i(1,r)}^{2}(x_j) \sigma_{i(2,r+u)}^{2}(x_j) \hat{n}_{(1,j),r+u}}{\hat{n}_{(1,j),r}} \hat{n}_{(2,j),r} - \hat{n}_{(2,j),r} \\
&= \left(1 + O\left(\sqrt{\frac{\log \log n(r)}{n(r)}}\right)\right) \left(1 + O\left(\frac{1}{n(r)}\right)\right) \hat{n}_{(2,j),r} - \hat{n}_{(2,j),r} \\
&\leq b_{u1} \sqrt{n(r) \log n(r)}.
\end{align*}
\]

Then, to satisfy inequality (101), it is sufficient to have

\[
k^{(i2,j)}_{u} > b_{u1} \sqrt{n(r) \log n(r)}. \tag{104}
\]

For inequality (100),

\[
\begin{align*}
&\frac{\hat{\delta}_{i(1,j)}^{r+u}(\hat{\delta}_{i(2,j)}^{r}) - \hat{\delta}_{i(1,j)}^{r}}{\hat{\delta}_{i(1,j)}^{r}} \leq \frac{\sigma_{i(1,r)}^{2}(x_j) \sigma_{i(2,r+u)}^{2}(x_j) \hat{n}_{(1,j),r+u}}{\hat{n}_{(1,j),r}} \hat{n}_{(2,j),r} + k^{(i2,j)}_{u} \\
&= \left(1 + O\left(\sqrt{\frac{\log \log n(r)}{n(r)}}\right)\right) \frac{\hat{n}_{(2,j),r}}{\hat{n}_{(1,j),r}} \frac{\hat{n}_{(1,j),r+u}}{\hat{n}_{(1,j),r+u}} \hat{n}_{(2,j),r} + k^{(i2,j)}_{u} \\
&\leq \left(1 + b_{u2} \sqrt{\frac{\log \log n(r)}{n(r)}}\right) \frac{\hat{n}_{(2,j),r}}{\hat{n}_{(1,j),r}} \frac{\hat{n}_{(1,j),r+u}}{\hat{n}_{(1,j),r+u}} \hat{n}_{(2,j),r} + k^{(i2,j)}_{u} \triangleq I_r.
\end{align*}
\]

It is sufficient to have \( I_r < 1 \). By lemma 14,

\[
\begin{align*}
&\left(1 + b_{u2} \sqrt{\frac{\log \log n(r)}{n(r)}}\right) \hat{n}_{(2,j),r} < \frac{\hat{n}_{(2,j),r}}{\hat{n}_{(1,j),r}} \frac{\hat{n}_{(1,j),r+u}}{\hat{n}_{(1,j),r+u}} \hat{n}_{(2,j),r} + k^{(i2,j)}_{u} \leq \frac{\hat{n}_{(2,j),r}}{\hat{n}_{(1,j),r}} \frac{\hat{n}_{(1,j),r+u}}{\hat{n}_{(1,j),r+u}} \hat{n}_{(2,j),r} + k^{(i2,j)}_{u} \tag{105}
\end{align*}
\]

will hold if the corresponding \( b_{u3} \) could be found such that

\[
k^{(1,j)}_{s} \geq b_{u3} \sqrt{n(r) \log n(r)}. \tag{106}
\]

65
On the other hand, when (105) holds,

\[
\frac{\hat{n}_{(i_2,j),r} + k_u^{(i_2,j)}}{\hat{n}_{(1,j),r} + k_s^{(1,j)}} > \left(1 + b_u2\sqrt{\log \log n(r)}\right) \frac{\hat{n}_{(i_2,j),r}}{\hat{n}_{(1,j),r}}.
\]

By lemma 9, we have

\[k_u^{(i_2,j)} > k_s^{(1,j)}\frac{\hat{n}_{(i_2,j),r}}{\hat{n}_{(1,j),r}} \geq b_{(i_2,j)}b_u3\sqrt{n(r)\log \log n(r)}. \quad (107)\]

Let \(c_2 = \max\left(\frac{b_{(i_2,j)}}{\hat{n}_{(i_2,j),r}}, b_u3\right)\). We will always find \(r\) such that \(k_s^{(1,j)} > c_2\sqrt{n(r)\log \log n(r)}\), due to (97).

Then, (104) and (106) are satisfied, which is sufficient to prove (100) and (101). Therefore, we have \(\hat{r}_{(i_2,j)}(r + u)/\hat{r}_{(i_1,j)}(r + u) > 1\). It means that design \(i_2\) of context \(x_j\) can not be sampled at iteration \(r + u\), which leads to contradiction.

**CASE 2:** If \(\delta_{(i_2,j)}/\delta_{(i_1,j)} < 1\), we have \(\hat{\delta}_{(i_2,j)}/\hat{\delta}_{(i_1,j)} < 1\) when \(r\) is large enough. By equation (96),

\[
\frac{\hat{\delta}_{(i_2,j)}^{r+u}S_{(i_2,j),r+u}}{\hat{\delta}_{(i_2,j)}^rS_{(i_2,j),r}} = \frac{\hat{\delta}_{(i_2,j)}^{r+u}S_{(i_2,j),r+u}}{\hat{\delta}_{(i_1,j)}^rS_{(i_1,j),r}} \quad (108)
\]

Similarly, to have contradiction, we need to show that

\[\frac{\hat{\delta}_{(i_2,j)}^{r+u}S_{(i_2,j),r+u}}{\hat{\delta}_{(i_2,j)}^rS_{(i_2,j),r}} > \frac{\hat{\delta}_{(i_1,j)}^{r+u}S_{(i_1,j),r+u}}{\hat{\delta}_{(i_1,j)}^rS_{(i_1,j),r}}. \quad (109)\]

It means that

\[
\frac{\hat{\delta}_{(i_2,j)}^{r}S_{(i_2,j)}}{\hat{\delta}_{(i_2,j)}^{r+u}S_{(i_2,j),r+u}} - \frac{\hat{\delta}_{(i_1,j)}^{r}S_{(i_1,j)}}{\hat{\delta}_{(i_1,j)}^{r+u}S_{(i_1,j),r+u}} \quad (108)
\]

\[
\frac{\hat{\delta}_{(i_2,j)}^{r+u}S_{(i_2,j),r+u}}{\hat{\delta}_{(i_2,j)}^{r}S_{(i_2,j),r}} - \frac{\hat{\delta}_{(i_1,j)}^{r+u}S_{(i_1,j),r+u}}{\hat{\delta}_{(i_1,j)}^{r}S_{(i_1,j),r}} < 1. \quad (109)\]
For (108), we have
\[
\begin{align*}
&\frac{\hat{\delta}^{r} (i_{2,j})}{\hat{\delta}^{r+u} (i_{1,j})} \frac{\hat{\sigma}^2_{i_{1},r}(x_j)}{\hat{\sigma}^2_{i_{2},r}(x_j)} \frac{\hat{n}_{(i_{1,j}),r+u}(x_j)}{\hat{n}_{(i_{2,j}),r}} \hat{n}_{(i_{2,j}),r} - \hat{n}_{(i_{2,j}),r} \\
&= \frac{\hat{\delta}^{r} (i_{2,j})}{\hat{\delta}^{r+u} (i_{1,j})} \frac{\hat{\sigma}^2_{i_{1},r}(x_j)}{\hat{\sigma}^2_{i_{2},r}(x_j)} \frac{\hat{n}_{(i_{1,j}),r+u}(x_j)}{\hat{n}_{(i_{1,j}),r}} \hat{n}_{(i_{1,j}),r} - \hat{n}_{(i_{1,j}),r} \\
&= \left( 1 + O\left(\sqrt{\frac{\log \log n(r)}{n(r)}}\right) \right) \left( 1 + O\left(\frac{1}{n(r)}\right) \right) \hat{n}_{(i_{2,j}),r} - \hat{n}_{(i_{2,j}),r} \\
&\leq b_{u4} \sqrt{n(r) \log \log n(r)}.
\end{align*}
\]

For (109), we have
\[
\begin{align*}
&\frac{\hat{\delta}^{r} (i_{2,j})}{\hat{\delta}^{r+u} (i_{1,j})} \frac{\hat{\sigma}^2_{i_{1},r}(x_j)}{\hat{\sigma}^2_{i_{2},r}(x_j)} \frac{\hat{n}_{(i_{1,j}),r+u}(x_j)}{\hat{n}_{(i_{1,j}),r}} \hat{n}_{(i_{1,j}),r} - \hat{n}_{(i_{1,j}),r} \\
&= \left( 1 + O\left(\sqrt{\frac{\log \log n(r)}{n(r)}}\right) \right) \left( 1 + O\left(\frac{1}{n(r)}\right) \right) \hat{n}_{(i_{1,j}),r} - \hat{n}_{(i_{1,j}),r} \\
&\leq b_{u5} \sqrt{n(r) \log \log n(r)}
\end{align*}
\]

So, to satisfy (108) and (109), it is sufficient to have
\[
\begin{align*}
k_{u}^{(i_{2,j})} &> b_{u4} \sqrt{n(r) \log \log n(r)}, \quad (110) \\
k_{u}^{(i_{1,j})} &> b_{u5} \sqrt{n(r) \log \log n(r)}.
\end{align*}
\]

Let \( k_{u}^{(i_{1,j})} > c_{3} \sqrt{n(r) \log \log n(r)} \) where \( c_{3} \) is a positive constant to be specified. By lemma 14, the follow inequality would hold
\[
\begin{align*}
\left( 1 + b_{u5} \sqrt{\frac{\log \log n(r)}{n(r)}} \right) \frac{\hat{n}_{(i_{2,j}),r}}{\hat{n}_{(i_{1,j}),r}} < \frac{\hat{n}_{(i_{2,j}),r} + k_{u}^{(i_{2,j})}}{\hat{n}_{(i_{1,j}),r}} + k_{u}^{(i_{1,j})}
\end{align*}
\]
if the corresponding constant \( b_{u6} \) could be found such that
\[
\begin{align*}
k_{u}^{(i_{1,j})} &> c_{3} \sqrt{n(r) \log \log n(r)} > b_{u6} \sqrt{n(r) \log \log n(r)}.
\end{align*}
\]
When \((112)\) holds, \(k_{(i,j)}^{(1,j)} > \hat{n}_{(i,j),r}/n_{(1,j),r}\). By lemma \(9\), \(\hat{n}_{(i,j),r} > b_{(i,j)} > 0\). So, we have

\[
k_u^{(i,j)} > \frac{\hat{n}_{(i,j),r}}{n_{(1,j),r}} \geq b_{(i,j)}k_s^{(1,j)} \geq b_{(i,j)}c_3\sqrt{n(r) \log \log n(r)}.
\]

(114)

To satisfy \((110)\), we should have \(c_3 > b_u/2b_{(i,j)}\). Meanwhile, to satisfy \((111)\), we should have \(c_3 > b_u/3\).

Now, let \(c_3 = \max\{b_u, b_u/2b_{(i,j)}, b_u/3\} + 1\). Then both \((110)\) and \((111)\) are satisfied at iteration \(r + 1\), which means that \(\tau_{(i,j),r+1} > \tau_{(i,j),r+1} > 1\). So, based on our algorithm, design \(i_2\) of context \(\mathbf{x}_j\) could not be sampled. Again, we have the contradiction.

\[\Box\]

**Lemma 16.** Given a context \(\mathbf{x}_j\) and an iteration stage \(r\), let \(k_q^{(i,j)} = \hat{n}_{(i,j),r+q} - \hat{n}_{(i,j),r}\), \(r \geq 3\), \(0 < q < n(r)\), \(c_{up} \geq 1\) is a fixed positive constant. For any positive constant \(c_4\), if there exists a sufficiently large positive constant \(c_3\) (dependent on \(c_4\), but independent of \(r\)) for which \(c_3\sqrt{n(r) \log n(r)} \leq k_q^{(1,j)} < c_{up}n(r)\), then there exists a suboptimal design \(i_1\) of \(\mathbf{x}_j\) such that \(c_4\sqrt{n(r) \log n(r)} < k_q^{(i,j)} < c_3n(r)\). Symmetrically, for any positive constant \(c_7\), if there exists a suboptimal design \(i_2\) of \(\mathbf{x}_j\) and a sufficiently large positive constant \(c_6\) (dependent on \(c_7\), but independent of \(r\)) for which \(c_6\sqrt{n(r) \log n(r)} < k_q^{(i,j)} < c_{up}n(r)\), then \(c_7\sqrt{n(r) \log n(r)} < k_q^{(1,j)} < c_8n(r)\)

*Proof.* We will prove that if there exists a sufficiently large positive constant \(c_3\) such that \(c_3\sqrt{n(r) \log n(r)} < k_q^{(1,j)} < c_{up}n(r)\), then we could find a suboptimal design \(i_1\) of \(\mathbf{x}_j\) satisfying \(c_4\sqrt{n(r) \log n(r)} < k_q^{(i,j)} < c_5n(r)\) almost surely.

First, we show that there exists a positive constant \(c_5\) such that \(k_q^{(i,j)} < c_5n(r)\) for all \(i \neq 1\) of \(\mathbf{x}_j\). If \(k_q^{(i,j)} < cn(r)\) for some suboptimal design \(i \neq 1\) of \(\mathbf{x}_j\) and every \(c > 0\), then

\[
\hat{n}_{(i,j),r+q}/n_{(i,j),r+q} \geq \hat{n}_{(i,j),r}/n_{(i,j),r} = \frac{\hat{n}_{(i,j),r}/n(r) + k_q^{(1,j)} / n(r)}{\hat{n}_{(i,j),r}/n(r) + k_q^{(1,j)} / n(r)}
\]

\[
\begin{align*}
&\geq \frac{b_{(i,j)} + k_q^{(i,j)} / n(r)}{b_{(i,j)} + k_q^{(1,j)} / n(r)} \\
&\geq \frac{b_{(i,j)} + c}{b_{(i,j)} + c_{up}},
\end{align*}
\]

(115)

(116)

where \((115)\) holds because of lemma \(9\) which indicates that \(\hat{n}_{(i,j),r} = \Theta(n(r))\), and \((116)\) holds
because \( k_q^{(i,j)} \geq cn(r) \) and \( k_q^{(1,j)} \leq c_{up} n(r) \).

Therefore, \( \limsup_{r \to \infty} \frac{\hat{n}_{(i,j),r}}{\hat{n}_{(1,j),r}} = \infty \), which contradicts the conclusion of lemma 10. So, there exists a positive constant \( c_3 \) such that \( k_q^{(i,j)} \leq c_3 n(r) \) for all \( i \neq 1 \) of \( x_j \).

Next, we will show a suboptimal design \( i_1 \) of \( x_j \) can be found such that \( k_q^{(i_1,j)} \geq c_4 \sqrt{n(r)} \log n(r) \)
where \( c_4 \) is a given positive constant. Lemma 10 implies that between two samples assigned to any suboptimal designs of \( x_j \), the number of samples allocated to any suboptimal designs of \( (1, j) \) is \( O(\sqrt{n(r)} \log \log n(r)) \) almost surely. Because \( k_q^{(1,j)} \geq c_3 \sqrt{n(r)} \log n(r) \) and \( c_3 \) is a large enough constant to be specified, there must exist some suboptimal designs of \( x_j \) that are sampled from iteration \( r \) to \( r + q \).

Let \( u_1 = \inf \{ l \geq 0 : j^{(i,j)}_{r+l} = 1, i \neq 1 \} \) and \( h_{1,j}^{(i,j)} = \hat{n}_{(i,j),r+l} - \hat{n}_{(i,j),r+u_1}, l > u_1 \). Because iteration \( r + u_1 \) is the first time after \( r - 1 \) that a suboptimal design of \( x_j \) is sampled, it is clear that

\[
\hat{n}_{(i,j),r+u_1} = \hat{n}_{(i,j),r}, \quad i \neq 1.
\]

Because of lemma 10 we have \( \hat{n}_{(1,j),r+u_1} - \hat{n}_{(1,j),r} = O(\sqrt{n(r)} \log n(r)) \).

If \( c_3 \sqrt{n(r)} \log n(r) \leq k_q^{(1,j)} \leq c_{up} n(r) \), we could find \( c'_3 > 0 \) such that

\[
c'_3 \sqrt{n(r)} \log n(r) \leq h_q^{(1,j)} = k_q^{(1,j)} - (\hat{n}_{(1,j),r+u_1} - \hat{n}_{(1,j),r}) \leq c_{up} n(r).
\]

Moreover, since a suboptimal design is sampled at iteration \( r + u_1 \), we have

\[
\frac{\hat{n}_{(1,j),r+u_1}^2}{\hat{\sigma}_{1,r+u_1}^2(x_j)} > \sum_{i \neq 1} \frac{\hat{n}_{(i,j),r+u_1}^2}{\hat{\sigma}_{i,r+u_1}^2(x_j)} = \sum_{i \neq 1} \frac{\hat{n}_{(i,j),r}^2}{\hat{\sigma}_{i,r}^2(x_j)}.
\]

Let \( u_2 = \sup \{ l < q : j^{(1,j)}_{r+l} = 1 \} \). Because iteration \( r + u_2 \) is the last time before \( r + q \) that \( (1, j) \) is sampled, it is clear that \( c''_3 \sqrt{n(r)} \log n(r) \leq h_{u_2}^{(1,j)} = h_q^{(1,j)} - 1 \leq c_{up} n(r) \) where \( c''_3 = c_3 - b_{w0} \) and \( b_{w0} > 0 \) is independent of \( c_3 \). At iteration \( r + u_2 \), we have

\[
\frac{\hat{n}_{(1,j),r+u_2}^2}{\hat{\sigma}_{1,r+u_2}^2(x_j)} < \sum_{i \neq 1} \frac{\hat{n}_{(i,j),r+u_2}^2}{\hat{\sigma}_{i,r+u_2}^2(x_j)}.
\]
that is, equivalently,

\[
\left( \frac{\hat{n}_{(1,j),r+u_1} + h_{u_2}^{(1,j)}}{\hat{\sigma}_{1,r+u_2}^2(x_j)} \right)^2 < \sum_{i \neq 1} \left( \frac{\hat{n}_{(i,j),r+u_1} + h_{u_2}^{(i,j)}}{\hat{\sigma}_{i,r+u_2}^2(x_j)} \right)^2.
\]  

(119)

We can rewritten (119) as

\[
\frac{\hat{n}_{(1,j),r+u_1}^2}{\hat{\sigma}_{1,r+u_1}^2(x_j)} \frac{\hat{\sigma}_{1,r+u_1}^2(x_j)}{\hat{\sigma}_{1,r+u_2}^2(x_j)} + \frac{(h_{u_2}^{(1,j)})^2 + 2\hat{n}_{(1,j),r+u_1} h_{u_2}^{(1,j)}}{\hat{\sigma}_{1,r+u_2}^2(x_j)} < \sum_{i \neq 1} \left( \frac{\hat{n}_{(i,j),r+u_1}^2}{\hat{\sigma}_{i,r+u_1}^2(x_j)} \frac{\hat{\sigma}_{i,r+u_1}^2(x_j)}{\hat{\sigma}_{i,r+u_2}^2(x_j)} + \frac{(h_{u_2}^{(i,j)})^2 + 2\hat{n}_{(i,j),r+u_1} h_{u_2}^{(i,j)}}{\hat{\sigma}_{i,r+u_2}^2(x_j)} \right)
\]

(120)

\[
\frac{\hat{n}_{(1,j),r+u_1}^2}{\hat{\sigma}_{1,r+u_1}^2(x_j)} + \sum_{i \neq 1} \frac{\hat{n}_{(i,j),r+u_1}^2}{\hat{\sigma}_{i,r+u_1}^2(x_j)} - \frac{\hat{n}_{(1,j),r+u_1}^2}{\hat{\sigma}_{1,r+u_1}^2(x_j)} \left( \frac{\hat{\sigma}_{1,r+u_1}^2(x_j)}{\hat{\sigma}_{1,r+u_2}^2(x_j)} - 1 \right) + \sum_{i \neq 1} \frac{(h_{u_2}^{(i,j)})^2 + 2\hat{n}_{(i,j),r+u_1} h_{u_2}^{(i,j)}}{\hat{\sigma}_{i,r+u_2}^2(x_j)}
\]  

(121)

where (120) holds because of (118). (121) comes from (117). Similar to (54),

\[
\left| \frac{\hat{\sigma}_{1,r+u_1}^2(x_j)}{\hat{\sigma}_{1,r+u_2}^2(x_j)} - 1 \right| = O\left( \sqrt{\frac{\log \log n^r}{n^r}} \right).
\]

So, the left hand side of (121) satisfies

\[
\geq \frac{\hat{n}_{(1,j),r+u_1}^2}{\hat{\sigma}_{1,r+u_1}^2(x_j)} - b_{u_1} \frac{\hat{n}_{(1,j),r+u_1}^2}{\hat{\sigma}_{1,r+u_1}^2(x_j)} \sqrt{\frac{\log \log n^r}{n^r}} \left( \frac{(h_{u_2}^{(1,j)})^2 + 2\hat{n}_{(1,j),r+u_1} h_{u_2}^{(1,j)}}{\hat{\sigma}_{1,r+u_2}^2(x_j)} \right)
\]

(122)

\[
\geq \frac{\hat{n}_{(1,j),r+u_1}^2}{\hat{\sigma}_{1,r+u_1}^2(x_j)} - b_{u_1} \frac{\hat{n}_{(1,j),r+u_1}^2}{\hat{\sigma}_{1,r+u_1}^2(x_j)} \sqrt{n^r \log \log n^r} + \frac{(h_{u_2}^{(1,j)})^2 + 2\hat{n}_{(1,j),r+u_1} h_{u_2}^{(1,j)}}{\hat{\sigma}_{1,r+u_2}^2(x_j)} \sqrt{\frac{\log \log n^r}{n^r}}
\]

(123)

\[
\geq \frac{\hat{n}_{(1,j),r+u_1}^2}{\hat{\sigma}_{1,r+u_1}^2(x_j)} - b_{u_2} n^r \log \log \log n^r + \frac{(h_{u_2}^{(1,j)})^2 + 2\hat{n}_{(1,j),r+u_1} h_{u_2}^{(1,j)}}{\hat{\sigma}_{1,r+u_2}^2(x_j)} \sqrt{n^r \log \log n^r}
\]

(124)
where \( b_{w_1}, b_{w_2}, b_{w_3}, b_{w_4} \) are positive constant and independent from \( c''_3 \). \([122]\) comes from lemma \([10]\) which indicates that \( k_{w_1}^{(1,j)} = O(\sqrt{n(r) \log n(r)}) \). \([123]\) holds because of lemma \([4]\) which implies that \( \hat{n}_{(1,j),r} = \Theta(n(r)) \). \([124]\) holds because \( c''_3 \sqrt{n(r) \log n(r)} \leq h_{u_2}^{(1,j)} \leq c_u n(r) \). Meanwhile, the right hand side of \([121]\) satisfies

\[
\begin{align*}
\hat{n}_{(1,j),r+u_1}^2 & \leq \frac{n_{(1,j),r+u_1}^2}{\hat{\sigma}_{1,r+u_1}^2(x_j)} + \frac{\hat{n}_{(i,j),r}^2}{\hat{\sigma}_{i,r}^2(x_j)} \left( \frac{\hat{\sigma}_{i,r+u_2}^2(x_j)}{\hat{\sigma}_{i,r+u_1}^2(x_j)} - 1 \right) + \sum_{i \neq 1} \left( h_{u_2}^{(i,j)} \right)^2 + 2 \hat{n}_{(i,j),r}^2 h_{u_2}^{(i,j)} \\
& \leq \frac{n_{(1,j),r+u_1}^2}{\hat{\sigma}_{1,r+u_1}^2(x_j)} + b_{w_5} n(r) \sqrt{n(r) \log n(r)} + \sum_{i \neq 1} \left( h_{u_2}^{(i,j)} \right)^2 + 2 \hat{n}_{(i,j),r}^2 h_{u_2}^{(i,j)} \\
& \leq \frac{n_{(1,j),r+u_1}^2}{\hat{\sigma}_{1,r+u_1}^2(x_j)} + b_{w_5} n(r) \sqrt{n(r) \log n(r)} + \sum_{i \neq 1} b_i h_{u_2}^{(i,j)} n(r),
\end{align*}
\]

where \( b_{w_5} \) and \( b_i, i \neq 1 \) are positive constant. \([125]\) holds because \( h_{u_2}^{(i,j)} \leq c_5 n(r), i \neq 1 \), where \( c_5 > 0 \). Take \([124]\) and \([125]\) into \([119]\), we must have

\[
\begin{align*}
\frac{n_{(1,j),r+u_1}^2}{\hat{\sigma}_{1,r+u_1}^2(x_j)} + (b_{w_4} c''_3 - b_{w_3}) n(r) \sqrt{n(r) \log n(r)} & < \frac{n_{(1,j),r+u_1}^2}{\hat{\sigma}_{1,r+u_1}^2(x_j)} + b_{w_5} n(r) \sqrt{n(r) \log n(r)} + \sum_{i \neq 1} b_i h_{u_2}^{(i,j)} n(r),
\end{align*}
\]

that means,

\[
(b_{w_4} c''_3 - b_{w_3} - b_{w_5}) \sqrt{n(r) \log n(r)} \leq b_{w_5} \sqrt{n(r) \log n(r)} < \sum_{i \neq 1} b_i h_{u_2}^{(i,j)},
\]

Then, there must be some suboptimal design \( i_1 \) of \( x_j \) such that

\[
k_q^{(i_1,j)} = h_{i_1,j}^{(i_1,j)} \geq h_{u_2}^{(i_1,j)} > \frac{b_{w_6}}{b_{w_1} (m - 1)} \sqrt{n(r) \log n(r)}.
\]

Note that \( b_i, i \neq 1, b_{w_0}, b_{w_3}, b_{w_4}, b_{w_5} \) are independent of \( c_3 \). Therefore, if we have

\[
c_3 > (m - 1) \left( (c_4 + b_{w_3} + b_{w_5} + 1)/b_{w_4} + b_{w_0} \right) \max b_i
\]

to make \( \frac{b_{w_6}}{b_{w_1} (m - 1)} > c_4 \), then \( k_q^{(i_1,j)} > c_4 \sqrt{n(r) \log n(r)} \) which proves the first part of this lemma. By symmetry, the second claim could be proved in a similar way.
LEMMA 17. Suppose that some suboptimal design \( i_1 \) of context \( x_{j_1} \) is sampled at iteration \( r \geq 3 \). Define \( t = \inf \{ l > 0 : f_{r+l}^{(i_1,j_1)} = 1 \} \). Then, we have \( k_l^{(i_2,j_2)} = O(\sqrt{n(r) \log n(r)}) \), \( i_2 = 1, 2, \cdots, k \), \( j_2 = 1, 2, \cdots, m \), \( j_2 \neq j_1 \), almost surely.

Proof. It is clear that \( t = O(n(r)) \) by lemma 9. Since we sample \((i_1, j_1)\) at iteration \( r \), we have

\[
\hat{\tau}_{(i_1,j_1),r}/\hat{\tau}_{(i',j'),r} = \frac{\hat{d}_{r}^{(i_1,j_1)}(S_{r}^{(i'},j',r) + S_{r}^{(i,j),r})}{\hat{d}_{r}^{(i',j')}\left(S_{r}^{(i_1,j_1),r} + S_{r}^{(i_1,j_1),r}\right)} < 1, \quad i' = 1, 2, \cdots, k, \quad j' = 1, 2, \cdots, m, \quad j' \neq j_1,
\]

that is, equivalently,

\[
\frac{\hat{d}_{r}^{2}(\hat{\tau}_{(i_1,j_1),r})(x_{j_1})}{\hat{n}_{(i_1,j_1),r}} + \frac{\hat{d}_{r}^{2}(\hat{\tau}_{(i_1,j_1),r})(x_{j_1})}{\hat{n}_{(i_1,j_1),r}} < \frac{\hat{d}_{r}^{2}(\hat{\tau}_{(i',j'),r})(x_{j_1})}{\hat{n}_{(i_1,j_1),r}} + \frac{\hat{d}_{r}^{2}(\hat{\tau}_{(i',j'),r})(x_{j_1})}{\hat{n}_{(i_1,j_1),r}}.
\]

(126)

CASE 1: In this case, suppose we could find a \( x_{j_2} \) such that, for any given \( b_{ja} > 0 \),

\[
b_{ja}\sqrt{n(r) \log n(r)} < k_l^{(i_1,j_2)} < b_{ja}n(r).
\]

(127)

By lemma 16, we could have a suboptimal design \( i_2 \) of \( x_{j_2} \) and an iteration stage \( r + u \) such that

\[
u = \sup \{ l < t : f_{r+l}^{(i_2,j_2)} = 1 \}
\]

and \( k_{u}^{(i_2,j_2)} > b_{ja}^{2}n(r) \log n(r) \), where \( b_{ja}^{2} \) is a sufficient large positive constant to be specified. We will show that

\[
\hat{\tau}_{(i_1,j_1),r}(r + u)/\hat{\tau}_{(i_1,j_1),r}(r + u) < 1,
\]

(128)

which means that \( f_{r+u}^{(i_2,j_2)} = 0 \) and leads to contradiction. To have (128), we need to show

\[
\frac{\hat{d}_{r}^{2}(\hat{\tau}_{(i_1,j_1),r})(x_{j_1})}{\hat{n}_{(i_1,j_1),r} + \hat{b}_{jb}\sqrt{n(r) \log n(r)}} + \frac{\hat{d}_{r}^{2}(\hat{\tau}_{(i_1,j_1),r})(x_{j_1})}{\hat{n}_{(i_1,j_1),r} + \hat{b}_{jb}\sqrt{n(r) \log n(r)}} \leq \frac{\hat{d}_{r}^{2}(\hat{\tau}_{(i_1,j_1),r})(x_{j_1})}{\hat{n}_{(i_1,j_1),r} + \hat{b}_{jb}\sqrt{n(r) \log n(r)}}.
\]

(129)

By lemma 15, \( b_{jb} > 1/\sqrt{n(r) \log n(r)} \) could be found such that \( \hat{n}_{(i_1,j_1),r} < \hat{n}_{(i_1,j_1),r} + b_{jb}\sqrt{n(r) \log n(r)} \). Furthermore, \( \hat{n}_{(i_1,j_1),r} = \hat{n}_{(i_1,j_1),r} + 1 \). Then, to prove (129), we need to have

\[
\frac{\hat{d}_{r}^{2}(\hat{\tau}_{(i_1,j_1),r})(x_{j_1})}{\hat{n}_{(i_1,j_1),r} + \hat{b}_{jb}\sqrt{n(r) \log n(r)}} + \frac{\hat{d}_{r}^{2}(\hat{\tau}_{(i_1,j_1),r})(x_{j_1})}{\hat{n}_{(i_1,j_1),r} + \hat{b}_{jb}\sqrt{n(r) \log n(r)}} > \frac{\hat{d}_{r}^{2}(\hat{\tau}_{(i_1,j_1),r})(x_{j_1})}{\hat{n}_{(i_1,j_1),r} + \hat{b}_{jb}\sqrt{n(r) \log n(r)}} + \frac{\hat{d}_{r}^{2}(\hat{\tau}_{(i_1,j_1),r})(x_{j_1})}{\hat{n}_{(i_1,j_1),r} + \hat{b}_{jb}\sqrt{n(r) \log n(r)}}.
\]

72
to satisfy (129), it is sufficient to have

\[
\frac{\hat{\delta}_{r+u}^{i_1,j_2} \hat{\sigma}_{1,r+u}^2(x_{j_1}) \hat{\delta}_{r+u}^{i_1,j_2} \hat{\sigma}_{1,r}^2(x_{j_1})}{\hat{\delta}_{r}^{i_1,j_2} \hat{\sigma}_{1,r}^2(x_{j_1})} \hat{n}(1,j_1,r) > \frac{\hat{n}(1,j_1,r)}{\frac{n(1,j_1,r)}{\log \log n(r)}} \left(1 - b_j \sqrt{\frac{\log \log n(r)}{n(r)}}\right) + \frac{b_j \sqrt{n(r)} \log \log n(r)}{\hat{n}(1,j_1,r)}
\]

(130) comes from (126). The last inequality holds because of law of the iterated logarithm. Thus, to satisfy (129), it is sufficient to have

\[
\frac{\hat{\delta}_{r+u}^{i_1,j_2} \hat{\sigma}_{1,r+u}^2(x_{j_1}) \hat{\delta}_{r+u}^{i_1,j_2} \hat{\sigma}_{1,r}^2(x_{j_1})}{\hat{\delta}_{r}^{i_1,j_2} \hat{\sigma}_{1,r}^2(x_{j_1})} \hat{n}(1,j_1,r) \frac{\hat{n}(1,j_1,r)}{\frac{n(1,j_1,r)}{\log \log n(r)}} > \frac{\hat{n}(1,j_1,r)}{\frac{n(1,j_1,r)}{\log \log n(r)}} \left(1 - b_j \sqrt{\frac{\log \log n(r)}{n(r)}}\right) + \frac{b_j \sqrt{n(r)} \log \log n(r)}{\hat{n}(1,j_1,r)}
\]

(131)

By multiplying LHS of (131) by \(1/\hat{\delta}_{r+u}^{i_1,j_1} \hat{\sigma}_{1,r+u}^2(x_{j_2})\), we have

\[
\frac{\hat{\delta}_{r+u}^{i_1,j_2} \hat{\sigma}_{1,r+u}^2(x_{j_1}) \hat{\delta}_{r+u}^{i_1,j_2} \hat{\sigma}_{1,r}^2(x_{j_1})}{\hat{\delta}_{r}^{i_1,j_2} \hat{\sigma}_{1,r}^2(x_{j_1})} \hat{n}(1,j_1,r) \frac{\hat{n}(1,j_1,r)}{\frac{n(1,j_1,r)}{\log \log n(r)}} + b_j \sqrt{n(r)} \log \log n(r) \hat{n}(1,j_1,r) \frac{1}{\hat{n}(1,j_1,r)}
\]
where, or, equivalently, sufficient to have almost surely. By Taylor’s expansion, we have (134) from (133). Thus, to satisfy (131), it is

\[
\left| \frac{\hat{\delta}_{i,j}^{u}}{\delta_{i,j}^{u}} - 1 \right| = O\left( \sqrt{\log n(r) n(r)} \right), \quad \left| \frac{\hat{\sigma}_{i,j}^{u}}{\sigma_{i,j}^{u}} - 1 \right| = O\left( \sqrt{\log n(r) n(r)} \right).
\]

At the same time, \(\hat{\delta}_{i,j}^{u}, \hat{\sigma}_{i,j}^{u}, \hat{\sigma}_{i,j}^{u}(x_{j1})\) and \(\hat{\sigma}_{i,j}^{u}(x_{j2})\) have positive upper and lower bounds almost surely. By Taylor’s expansion, we have (134) from (133). Thus, to satisfy (131), it is sufficient to have

\[
\left( 1 - b_{jh}' \sqrt{\log n(r) n(r)} \right) \frac{1}{\hat{n}_{i,j}^{r+u}} > \frac{1}{\hat{n}_{i,j}^{r+u}},
\]

or, equivalently,

\[
\frac{1}{1 - b_{jh}' \sqrt{\log n(r) n(r)}} \hat{n}_{i,j}^{r+u} - \hat{n}_{i,j}^{r+u} \leq b_{jh} \sqrt{n(r) \log n(r)} \leq k_{i,j}^{r+u},
\]

where \(b_{jh}\) is a fixed positive constant. By multiplying LHS of (132) by \(1/\hat{\delta}_{i,j}^{u} \hat{\sigma}_{i,j}^{u}(x_{j2})\), we have

\[
\frac{\hat{\delta}_{i,j}^{u}}{\delta_{i,j}^{u}} \frac{\hat{\sigma}_{i,j}^{u}}{\sigma_{i,j}^{u}} \frac{\hat{\delta}_{i,j}^{u}(x_{j1})}{\delta_{i,j}^{u}(x_{j1})} \frac{\hat{\sigma}_{i,j}^{u}(x_{j2})}{\sigma_{i,j}^{u}(x_{j2})} \frac{1}{\hat{n}_{i,j}^{r+u}} \hat{n}_{i,j}^{r+u} + b_{jh} \sqrt{n(r) \log n(r)}
\]

\[
= \frac{\hat{\delta}_{i,j}^{u}}{\delta_{i,j}^{u}} \frac{\hat{\sigma}_{i,j}^{u}}{\sigma_{i,j}^{u}} \frac{\hat{\delta}_{i,j}^{u}(x_{j1})}{\delta_{i,j}^{u}(x_{j1})} \frac{\hat{\sigma}_{i,j}^{u}(x_{j2})}{\sigma_{i,j}^{u}(x_{j2})} \frac{1}{\hat{n}_{i,j}^{r+u}} \hat{n}_{i,j}^{r+u} + b_{jh} \sqrt{n(r) \log n(r)}
\]

\[
= \frac{\hat{\delta}_{i,j}^{u}}{\delta_{i,j}^{u}} \frac{\hat{\sigma}_{i,j}^{u}}{\sigma_{i,j}^{u}} \frac{\hat{\delta}_{i,j}^{u}(x_{j1})}{\delta_{i,j}^{u}(x_{j1})} \frac{\hat{\sigma}_{i,j}^{u}(x_{j2})}{\sigma_{i,j}^{u}(x_{j2})} \frac{1}{\hat{n}_{i,j}^{r+u}} \hat{n}_{i,j}^{r+u} + b_{jh} \sqrt{n(r) \log n(r)}
\]

74
\[
\left(1 - b_{jj} \sqrt{\frac{\log \log n(r)}{n(r)}}\right) \frac{1}{\hat{n}_{(1,j_2),r}} \geq \frac{1}{\hat{n}_{(1,j_2),r}^{1,2,\cdots} + u, r+u},
\]

Thus, to satisfy (132), it is sufficient to have

\[
\left(1 - b_{jj} \sqrt{\frac{\log \log n(r)}{n(r)}}\right) \frac{1}{\hat{n}_{(1,j_2),r}} > \frac{1}{\hat{n}_{(1,j_2),r}^{1,2,\cdots} + u, r+u},
\]

or, equivalently,

\[
\frac{1}{1 - b_{jj} \sqrt{\frac{\log \log n(r)}{n(r)}}} \hat{n}_{(1,j_2),r} - \hat{n}_{(1,j_2),r} \leq b_{jj}' \sqrt{n(r) \log n(r)} < k_u^{(1,j_2)}.
\]

Because \( t = O(n(r)) \), we have \( k_t^{(i_2,j_2)} < b_{jU} n(r) \) where \( b_{jU} \) is a fixed constant. By the second part of lemma 13 and lemma 16 would hold if \( k_u^{(i_2,j_2)} \geq b_{jj} \sqrt{n(r) \log n(r)} \) where \( b_{jj} \) is a sufficiently large positive constant dependent on \( b_{jj}' \). Let \( b_{ja}' = \max(b_{jj}, b_{jh}) \). Since \( u = \sup\{t < t : T_r^{(i_2,j_2)} = 1\}, \)
\( k_u^{(i_2,j_2)} \geq b_{ja}' \sqrt{n(r) \log n(r)} \) holds if \( k_t^{(i_2,j_2)} > (b_{ja}' + 1) \sqrt{n(r) \log n(r)} \). By the first part of lemma 16 we can have \( k_t^{(i_2,j_2)} > (b_{ja}' + 1) \sqrt{n(r) \log n(r)} \) because of (127).

Therefore, \( \hat{\tau}_{(i_1,j_1),r+u}/\hat{\tau}_{(i_2,j_2),r+u} < 1 \) at iteration \( r + u \) and \( (i_2, j_2) \) can not be sampled. Now, we have the desired contradiction.

**CASE 2:** Similarly, if we could find a suboptimal design \( i_2 \) of context \( x_j \) such that, for any \( b_{jj} > 0, \)
\[
b_{jj} \sqrt{n(r) \log n(r)} < k_t^{(i_2,j_2)} < t < b_{jU} n(r),
\]
the contradiction also exists.

\[
\iffalse
\[
\hat{\tau}_{(i_1,j_1),r+u}/\hat{\tau}_{(i_2,j_2),r+u} < 1
\]
\fi
\]

Combining lemmas 15 and 17 with lemma 12 we could prove Theorem 6 in a similar way as Theorem 5.

**Proof of Theorem 6.** Fix a sample path \( \omega \) and let \( A_{(i,j)} = \{r_l, l = 1, 2, \cdots\} \) be the subsequence where design \( i \) of context \( x_j \) is sampled. By lemma 9, \( A_{(i,j)} \) is infinite for any \( i = 1, 2, \cdots, k, \)
\( i \neq i^*(x_j) \) and \( j = 1, 2, \cdots, m. \) Without loss of generality, we assume again that \( r \) is large enough and \( \hat{i}_t^*(x_j) = i^*(x_j) = 1. \)
Let $\Gamma_{(i_1,j_1),(i_2,j_2)}(r) = \hat{Y}_{(i_1,j_1),r} - \hat{Y}_{(i_2,j_2),r}$ and suppose $r_1 \in A_{(i_1,j_1)}$. Then, by our algorithm, we have $\Gamma_{(i_1,j_1),(i_2,j_2)}(r_1) < 0$ and $\Gamma_{(i_1,j_1),(i_2,j_2)}(r_{1+1}) < 0$.

When $r_1 < r < r_{1+1}$, $\Gamma_{(i_1,j_1),(i_2,j_2)} > 0$ may occur. And when $\Gamma_{(i_1,j_1),(i_2,j_2)} > 0$, $\Gamma_{(i_1,j_1),(i_2,j_2)}$ may increase only because $(1, j_1)$, $(1, j_2)$ and $(i_2, j_2)$ are sampled. Otherwise, $\Gamma_{(i_1,j_1),(i_2,j_2)}$ would decrease. In lemma 12, we get the upper bound of the increment,

$$\left| \Gamma_{(i_1,j_1),(i_2,j_2)}(r + 1) - \Gamma_{(i_1,j_1),(i_2,j_2)}(r) \right| < C(n(r))^{-3/4}.$$ 

Therefore, for any $r_1 < r < r_{1+1}$,

$$\Gamma_{(i_1,j_1),(i_2,j_2)}(r) < \Gamma_{(i_1,j_1),(i_2,j_2)}(r - 1) + \frac{C}{(n(r_1))^{3/4}}$$

$$< \Gamma_{(i_1,j_1),(i_2,j_2)}(r_1) + \left(1 + k_{(1,j_1),(r_1,r)} + k_{(1,j_2),(r_1,r)} + k_{(i_2,j_2),(r_1,r)} \right) \frac{C}{(n(r_1))^{3/4}}$$

$$< \left(1 + k_{(1,j_1),(r_1,r)} + k_{(1,j_2),(r_1,r)} + k_{(i_2,j_2),(r_1,r)} \right) \frac{C}{(n(r_1))^{3/4}},$$

where $k_{(i,j),(r_1,r)} = \hat{n}_{(i,j),r} - \hat{n}_{(i,j),r_1}$. By lemma 15, we have

$$k_{(1,j_1),(r_1,r)} \leq k_{(1,j_1),(r_1,r_{1+1})} = O(\sqrt{n(r_1) \log n(r_1)}).$$

Meanwhile, by lemma 17, we have

$$k_{(i_2,j_2),(r_1,r)} \leq k_{(i_2,j_2),(r_1,r_{1+1})} = O(\sqrt{n(r_1) \log n(r_1)}),$$

$$k_{(1,j_2),(r_1,r)} \leq k_{(1,j_2),(r_1,r_{1+1})} = O(\sqrt{n(r_1) \log n(r_1)}).$$

It follows that

$$\Gamma_{(i_1,j_1),(i_2,j_2)}(r) = O\left(\sqrt{n(r_1) \log n(r_1)} \frac{n(r_1)}{(n(r_1))^{3/4}}\right), \quad r_1 < r < r_{1+1},$$

whence $\limsup_{r \to \infty} \Gamma_{(i_1,j_1),(i_2,j_2)}(r) = 0$. By symmetry,

$$\liminf_{r \to \infty} \Gamma_{(i_1,j_1),(i_2,j_2)}(r) = \limsup_{r \to \infty} \Gamma_{(i_1,j_1),(i_2,j_2)}(r) = 0,$$

whence $\lim_{r \to \infty} \Gamma_{(i_1,j_1),(i_2,j_2)}(r) = 0$. 

76
On the other hand, by lemma 9, \( \hat{n}_{(i_1,j_1),r} \to \infty \) and \( \hat{n}_{(i_2,j_2),r} \to \infty \) as \( r \to \infty \). Furthermore, \( \hat{i}_r^*(x_j) = i^*(x_j) \) when \( r \) is large enough. So, we have

\[
\lim_{r \to \infty} \hat{V}_{(i,j),r} = \lim_{r \to \infty} \frac{(\hat{Y}_i(x_j) - \hat{Y}_{i^*}^*(x_j)(x_j))^2}{\hat{\sigma}_i^2(x_j)/\hat{\alpha}_{i,i} + \hat{\sigma}_{i^*}^2(x_j)/\hat{\alpha}_{i,j}} = \frac{(y_i(x_j) - y_{i^*}^*(x_j)(x_j))^2}{\sigma_i^2(x_j)/\alpha_{i,i} + \sigma_{i^*}^2(x_j)/\alpha_{i,j}},
\]

for any \( i = 1, 2, \ldots, k \) and \( j = 1, 2, \ldots, m \). Since \( \lim_{r \to \infty} \Gamma_{(i_1,j_1),(i_2,j_2)}(r) = 0 \), we have

\[
\lim_{r \to \infty} \left[ \frac{(y_{i_1}(x_{j_1}) - y_{i_1^*}^*(x_{j_1})(x_{j_1}))^2}{\sigma_i^2(x_{j_1})/\alpha_{i_1,i_1} + \sigma_{i_1}^2(x_{j_1})/\alpha_{i_1,j_1}} - \frac{(y_{i_2}(x_{j_2}) - y_{i_2^*}^*(x_{j_2})(x_{j_2}))^2}{\sigma_i^2(x_{j_1})/\alpha_{i_2,i_2} + \sigma_{i_2}^2(x_{j_2})/\alpha_{i_2,j_2}} \right] = 0,
\]

and Theorem is proved.

\[\square\]

References

Bertsimas, D., D. B. Brown, C. Caramanis. 2011. Theory and applications of robust optimization. *SIAM Review*, 53, 464–501.

Bertsimas, D., N. Kallus, A. M. Weinstein, Y. D. Zhuo. 2017. Personalized diabetes management using electronic medical records. *Diabetes Care*, 40(2), 210–217.

Boyd, S., L. Vandenberghe. 2004. *Convex Optimization*. Cambridge University Press, New York.

Branke, J., S. E. Chick, C. Schmidt. 2007. Selecting a selection procedure. *Management Science*, 53(11), 1916–1932.

Buchholz, P., A. Thümmeler. 2005. Enhancing evolutionary algorithms with statistical selection procedures for simulation optimization. In *Proceedings of the 2005 Winter Simulation Conference*, 842–852. Orlando, FL.

Chen, C. H., D. He, M. Fu, L. H. Lee. 2008. Efficient simulation budget allocation for selecting an optimal subset. *INFORMS Journal on Computing*, 20(4), 579–595.

Chen, C. H., L. H. Lee. 2011. *Stochastic Simulation Optimization: An Optimal Computing Budget Allocation*. World Scientific Publishing, Singapore.

Chen, C. H., J. Lin, E. Yücesan, S. E. Chick. 2000. Simulation budget allocation for further enhancing the efficiency of ordinal optimization. *Discrete Event Dynamic Systems*, 10, 251–270.

Chen, Y., I. O. Ryzhov. 2018. Complete expected improvement converges to an optimal budget allocation.
Chick, S. E., J. Branke, C. Schmidt. 2010. Sequential sampling to myopically maximize the expected value of information. INFORMS Journal on Computing, 22(1), 71–80.

Choi, S. E., E. P. Katherine, A. C. Tramontano, C. Y. Kong, C. Hur. 2014. Statins and aspirin for chemoprevention in barrett’s esophagus: Results of a cost-effectiveness analysis. Cancer Prevention Research, 7(3), 341–350.

Dembo, A., O. Zeitouni. 1998. Large Deviations Techniques and Applications. 2nd edition. Springer, New York.

Dudewicz, E. J., S. R. Dalal. 1975. Allocation of observations in ranking and selection with unequal variances. Sankhya, 37B, 28–78.

Fan, W., L. J. Hong, X. Zhang. 2019. Distributionally robust selection of the best. Management Science. In press.

Frazier, P. I., W. B. Powell, S. Dayanik. 2008. A knowledge-gradient policy for sequential information collection. SIAM Journal on Control and Optimization, 47(5), 2410–2439.

Fu, M., J. Hu, C.-H. Chen, X. Xiong. 2007. Simulation allocation for determining the best design in the presence of correlated sampling. INFORMS Journal on Computing, 19, 101–111.

Ganesh, A. J., N. O’Connell, D. J. Wischik. 2004. Big Queues, Lecture Notes in Mathematics, volume 1838. Springer, New York.

Gao, S., W. Chen. 2016. A new budget allocation framework for selecting top simulated designs. IIE Transactions, 48, 855–863.

Gao, S., W. Chen, L. Shi. 2017a. A new budget allocation framework for the expected opportunity cost. Operations Research, 65, 787–803.

Gao, S., H. Xiao, E. Zhou, W. Chen. 2017b. Robust ranking and selection with optimal computing budget allocation. Automatica, 81, 30–36.

Glynn, P., S. Juneja. 2004. A large deviations perspective on ordinal optimization. In Proceedings of the 2004 Winter Simulation Conference, 577–585.

Glynn, P., S. Juneja. 2011. Ordinal optimization: A nonparametric framework. In Proceedings of the 2011 Winter Simulation Conference, 4062–4069.

Ho, Y., Q. Zhao, Q. Jia. 2007. Ordinal Optimization: Soft Optimization for Hard Problems. Springer, New York.
Hunter, S. R., R. Pasupathy. 2013. Optimal sampling laws for stochastically constrained simulation optimization on finite sets. *INFORMS Journal on Computing*, 25(3), 527–542.

Hur, C., N. S. Nishioka, G. S. Gazelle. 2004. Cost-effectiveness of aspirin chemoprevention for barrett’s esophagus. *Journal of the National Cancer Institute*, 96(4), 316–325.

Kim, E. S., R. S. Herbst, I. I. Wistuba, J. J. Lee, G. R. Blumenschein, A. Tsao, D. J. Stewart, M. E. Hicks, J. Erasmus, S. Gupta, et al. 2011. The battle trial: personalizing therapy for lung cancer. *Cancer discovery*, 14, 44–53.

Kim, S. H., B. L. Nelson. 2001. A fully sequential procedure for indifference-zone selection in simulation. *ACM Transactions on Modeling and Computer Simulation*, 11(3), 251–273.

Law, A. M., W. D. Kelton. 2000. *Simulation Modeling and Analysis*. 3rd edition. McGraw-Hill, New York.

Lee, L. H., E. P. Chew, S. Y. Teng, D. Goldsman. 2010. Finding the non-dominated pareto set for multi-objective simulation models. *IIE Transactions*, 42(9), 656–674.

Lee, L. H., N. A. Pujowidianto, L. Li, C. H. Chen, C. M. Yap. 2012. Approximate simulation budget allocation for selecting the best design in the presence of stochastic constraints. *IEEE Transactions on Automatic Control*, 57(11), 2940–2945.

Li, X., X. Zhang, Z. Zheng. 2018. Data-driven ranking and selection: high-dimensional covariates and general dependence. In *Proceedings of the 2018 Winter Simulation Conference*, 1933–1944. Institute of Electrical and Electronics Engineers, Inc., Piscataway, New Jersey.

Luo, J., L. J. Hong, B. L. Nelson, Y. Wu. 2015. Fully sequential procedures for large-scale ranking-and-selection problems in parallel computing environments. *Operations Research*, 63(5), 1177–1194.

Nelson, B. L., J. Swann, D. Goldsman, W. Song. 2001. Simple procedures for selecting the best simulated system when the number of alternatives is large. *Operations Research*, 49(6), 950–963.

Ni, E. C., D. F. Ciocan, S. G. Henderson, S. R. Hunter. 2017. Efficient ranking and selection in parallel computing environments. *Operations Research*, 65(3), 821–836.

Rinott, Y. 1978. On two-stage selection procedures and related probability inequalities. *Communications in Statistics*, A7(8), 799–811.

Robbins, H. 1952. Some aspects of the sequential design of experiments. *Bulletin of the American Mathematical Society*, 58(5), 527–535.

Ryzhov, I. O. 2016. On the convergence rates of expected improvement methods. *Operations Research*, 64(6), 1515–1528.

Shen, H., L. J. Hong, X. Zhang. 2017. Ranking and selection with covariates for personalized decision making. URL [https://arxiv.org/abs/1710.02642](https://arxiv.org/abs/1710.02642). Under review.
Xu, J., E. Huang, C.-H. Chen, L. H. Lee. 2015. Simulation optimization: a review and exploration in the new era of cloud computing and big data. *Asia-Pacific Journal of Operational Research*, 32(3), 1–34.

Zhou, L. 2015. A survey on contextual multiarmed bandits. URL [https://arxiv.org/pdf/1508.03326](https://arxiv.org/pdf/1508.03326)