Complete expected improvement converges to an optimal budget allocation

Ye Chen                     Ilya O. Ryzhov*

March 22, 2018

Abstract

The ranking and selection problem is a well-known mathematical framework for the formal study of optimal information collection. Expected improvement (EI) is a leading algorithmic approach to this problem; the practical benefits of EI have repeatedly been demonstrated in the literature, especially in the widely studied setting of Gaussian sampling distributions. However, it was recently proved that some of the most well-known EI-type methods achieve suboptimal convergence rates. We investigate a recently-proposed variant of EI (known as “complete EI”) and prove that, with some minor modifications, it can be made to converge to the rate-optimal static budget allocation without requiring any tuning.

1 Introduction

In the ranking and selection (R&S) problem, there are $M$ “alternatives” (or “systems”), and each alternative $j \in \{1, ..., M\}$ has an unknown value $\mu^{(j)} \in \mathbb{R}$ (for simplicity, suppose that $\mu^{(i)} \neq \mu^{(j)}$ for $i \neq j$). We wish to identify the unique best alternative $j^* = \arg \max_j \mu^{(j)}$. For any $j$, we have the ability to collect noisy samples of the form $W^{(j)} \sim \mathcal{N}(\mu^{(j)}, (\lambda^{(j)})^2)$, but we are limited to a total of $N$ samples that have to be allocated among the alternatives, under independence assumptions ensuring that samples of $j$ do not provide any information about $i \neq j$. After the sampling budget has been consumed, we select the alternative with the highest sample mean. We say that “correct selection” occurs if the selected alternative is identical to $j^*$. We seek to allocate the budget in a way that maximizes the probability of correct selection.

R&S has a long history dating back to Bechhofer (1954), and continues to be an active area of research; see the tutorials by Hong & Nelson (2009) and Chau et al. (2014). Most modern research on this problem considers sequential allocation strategies, in which the decision-maker may spend
part of the sampling budget, observe the results, and adjust the allocation of the remaining samples accordingly. The literature has developed various algorithmic approaches, including indifference-zone methods (Kim & Nelson, 2001), optimal computing budget allocation (or OCBA; see Chen et al., 2000), and expected improvement (Jones et al., 1998). The related literature on multi-armed bandits (Gittins et al., 2011) has contributed other approaches such as Thompson sampling (Russo & Van Roy, 2014), although the bandit problem uses a different objective function from R&S and thus a good method for one problem may work poorly in the other (Russo, 2017).

Glynn & Juneja (2004) gave a rigorous foundation for the notion of optimal budget allocation with regard to probability of correct selection. Denote by $0 \leq N^{(j)} \leq N$ the number of samples assigned to alternative $j$ (thus, $\sum_j N^{(j)} = N$), and take $N \to \infty$ while keeping the proportion $\alpha^{(j)} = \frac{N^{(j)}}{N}$ constant. The optimal proportions $\alpha^{(j)}_*$ (among all possible vectors $\alpha \in \mathbb{R}_+^M$ satisfying $\sum_j \alpha^{(j)} = 1$) satisfy two conditions:

- Proportion assigned to alternative $j^*$:
  \[
  \left(\frac{\alpha^{(j^*)}}{\lambda^{(j^*)}}\right)^2 = \sum_{j \neq j^*} \left(\frac{\alpha^*_j}{\lambda^*_j}\right)^2
  \]
  \[(1)\]

- Proportions assigned to arbitrary $i, j \neq j^*$:
  \[
  \frac{(\mu^{(i)} - \mu^{(j^*)})^2}{(\lambda^{(i)})^2 + (\lambda^{(j^*)})^2} = \frac{(\mu^{(j)} - \mu^{(j^*)})^2}{(\lambda^{(j)})^2 + (\lambda^{(j^*)})^2}
  \]
  \[(2)\]

Under this allocation, the probability of incorrect selection will converge to zero at the fastest possible rate (exponential with the best possible exponent). Of course, (1)-(2) themselves depend on the unknown performance values. A common work-around is to replace these values with plug-in estimators and repeatedly solve for the optimal proportions in a sequential manner. Even then, the optimality conditions are cumbersome to solve, which may explain why researchers and practitioners prefer suboptimal heuristics that are easier to implement. To give a recent example, Pasupathy et al. (2014) uses large deviations theory to derive optimality conditions, analogous to (1)-(2), for a general class of simulation-based optimization problems, but advocates approximating the conditions to obtain a more tractable solution.

In this paper, we focus on one particular class of heuristics, namely expected improvement (EI) methods, which have consistently demonstrated computational and practical advantages in
a wide variety of problem classes (Branke et al., 2007; Scott et al., 2010; Han et al., 2016) ever since their introduction in Jones et al. (1998). EI is a Bayesian approach to R&S that allocates samples in a purely sequential manner: each successive sample is used to update the posterior distributions of the values \( \mu^{(j)} \), and the next sample is adaptively assigned using the so-called “value of information” criterion. This notion will be formalized in Section 2; here, we simply note that there are many competing definitions, such as the classic EI criterion of Jones et al. (1998), the knowledge gradient criterion (Powell & Ryzhov, 2012), or the \( LL_1 \) criterion of Chick et al. (2010). Ryzhov (2016) showed that the seemingly minor differences between these variants produce very different asymptotic allocations, but also that all of these allocations are suboptimal.

Recently, however, Salemi et al. (2014) proposed a new criterion called “complete expected improvement” or CEI. The formal definition of CEI is given in Section 3, but the main idea is that, when we evaluate the potential of a seemingly-suboptimal alternative to improve over the current-best value, we treat both of the values in this comparison as random variables (unlike classic EI, which only uses a plug-in estimate of the best value). Salemi et al. (2014) created and implemented this idea in the context of Gaussian Markov random fields, a more sophisticated Bayesian learning model than the version of R&S with independent normal samples that we consider here. Although the Gaussian Markov model is far more scalable and practical, it also presents greater difficulties for theoretical analysis: for example, no analog of (1)-(2) is available for statistical models with Gaussian Markov structure. In the present paper, we translate the CEI criterion to our simpler model, which enables us to study its theoretical convergence rate, and ultimately leads to strong new theoretical arguments in support of the CEI method.

Our main contribution in this paper is to prove that, with a slight modification to the method as laid out in Salemi et al. (2014), this modified version of CEI achieves both (1) and (2) asymptotically as \( N \to \infty \). Not only is this a new result for EI-type methods, it is also one of the strongest guarantees for any R&S heuristic to date. To compare it with the state of the art, Russo (2017) presents a class of heuristics, called “top-two methods,” which can also achieve optimal allocations, but only when a tuning parameter is set optimally. A more recent work by Qin et al. (2017), which appeared while the present paper was under review, extended the top-two approach to use CEI calculations, but kept the requirement of a tunable parameter. By contrast, our approach requires no tuning whatsoever. A different work by Peng & Fu (2017) finds a way to reverse-engineer the EI calculations to optimize the rate, but this approach requires one to first solve (1)-(2) with plug-
in estimators, and the procedure does not have a natural interpretation as an EI criterion. By contrast, CEI requires no additional computational effort compared to classic EI, and has a very simple and intuitive interpretation. In this way, our paper bridges the gap between theoretical notions of rate-optimality and the more practical concerns that motivate EI methods.

2 Preliminaries

We first provide some formal background for the optimality conditions (1)-(2) derived in Glynn & Juneja (2004), and then give an overview of EI-type methods. It is important to note that the theoretical framework of Glynn & Juneja (2004), as well as the theoretical analysis developed in the present paper, relies on a frequentist interpretation of R&S, in which the value of alternative $i$ is treated as a fixed (though unknown) constant. On the other hand, EI methods are derived using Bayesian arguments; however, once the derivation is complete, one is free to apply and study the resulting algorithm in a frequentist setting (as we do in this paper). To avoid confusion, we first describe the frequentist model, then introduce details of the Bayesian model where necessary.

In the frequentist model, the values $\mu(i)$ are fixed for $i = 1, ..., M$. Let $\{j_n\}_{n=0}^\infty$ be a sequence of alternatives chosen for sampling. For each $j_n$, we observe $W^{(j_n)}_{n+1} \sim N\left(\mu(j_n), (\lambda(j_n))^2\right)$ where $\lambda(j) > 0$ is assumed to be known for all $j$. We let $\mathcal{F}_n$ be the sigma-algebra generated by $j_0, W_1^{(j_0)}, ..., j_{n-1}, W_n^{(j_{n-1})}$. The allocation $\{j_n\}_{n=0}^\infty$ is said to be *adaptive* if each $j_n$ is $\mathcal{F}_n$-measurable, and *static* if all $j_n$ are $\mathcal{F}_0$-measurable. We define $I_n^{(j)} = 1_{\{j_n = j\}}$ and let $N_n^{(j)} = \sum_{m=0}^{n-1} I_m^{(j)}$ be the number of times that alternative $j$ is sampled up to time index $n = 1, 2, ...$. At time $n$, we can calculate the statistics

$$\theta_n^{(j)} = \frac{1}{N_n^{(j)}} \sum_{m=0}^{n-1} I_m^{(j)} W_m^{(j)}$$

$$\left(\sigma_n^{(j)}\right)^2 = \frac{(\lambda(j))^2}{N_n^{(j)}}.$$  

If our sampling budget is limited to $n$ samples, then $j^*_n = \arg \max_j \theta_n^{(j)}$ will be the final selected alternative. Correct selection occurs at time index $n$ if $j^*_n = j^*$. The probability of correct selection (PCS), written as $P(j^*_n = j^*)$, depends on the rule used to allocate the samples. Glynn & Juneja (2004) proves that, for any static allocation that assigns a proportion $\alpha^{(j)} > 0$ of the budget to
each alternative $j$, the convergence rate of PCS can be expressed in terms of the limit

$$\Gamma^\alpha = - \lim_{n \to \infty} \frac{1}{n} \log P (j^*_n \neq j^*).$$

(5)

That is, the probability of incorrect selection converges to zero at an exponential rate where the exponent includes a constant $\Gamma^\alpha$ that depends on the vector $\alpha$ of proportions. Equations (1)-(2) characterize the proportions that optimize the rate (maximize $\Gamma^\alpha$) under the assumption of independent normal samples. Although Glynn & Juneja (2004) only considers static allocations, nonetheless, to date, (5) continues to be one of the strongest rate results for R&S. Optimal static allocations derived through this framework can be used as guidance for the design of dynamic allocations; see, for example, Pasupathy et al. (2014) and Hunter & McClosky (2016).

We now describe EI, a prominent class of adaptive methods. EI uses a Bayesian model of the learning process, which is very similar to the model presented above, but makes the additional assumption that $\mu^{(j)} \sim \mathcal{N}\left(\theta^{(j)}, \left(\sigma^{(j)}\right)^2\right)$, where $\theta^{(j)}$ and $\sigma^{(j)}$ are pre-specified prior parameters. It is also assumed that $\mu^{(i)}, \mu^{(j)}$ are independent for all $i \neq j$. Under these assumptions, it is well-known (DeGroot, 1970) that the posterior distribution of $\mu^{(j)}$ given $\mathcal{F}_n$ is $\mathcal{N}\left(\theta_n^{(j)}, \left(\sigma_n^{(j)}\right)^2\right)$ where the posterior mean and variance can be computed recursively. Under the non-informative prior $\sigma_0^{(j)} = \infty$, the Bayesian posterior parameters $\theta_n^{(j)}, \sigma_n^{(j)}$ are identical to the frequentist statistics defined in (3)-(4), and so we can use the same notation for both settings.

One of the first (and probably the best-known) EI algorithms was introduced by Jones et al. (1998). In this version of EI, as applied to our R&S model, we take $j_n = \arg \max_j v_n^{(j)}$ where

$$v_n^{(j)} = \mathbb{E} \left( \max \left\{ \mu^{(j)} - \theta_n^{(j^*)}, 0 \right\} | \mathcal{F}_n \right)$$

$$= \sigma_n^{(j)} f \left( \frac{|\theta_n^{(j)} - \theta_n^{(j^*)}|}{\sigma_n^{(j)}} \right),$$

(6)

and $f (z) = z \Phi (z) + \phi (z)$ with $\phi, \Phi$ being the standard Gaussian pdf and cdf, respectively. We can view (6) as a measure of the potential that the true value of $j$ will improve upon the current-best estimate $\theta_n^{(j^*)}$. The EI criterion $v_n^{(j)}$ may be recomputed at each time stage $n$ based on the most recent posterior parameters.

Ryzhov (2016) gave the first convergence rate analysis of this algorithm. Under EI, we have

$$\lim_{n \to \infty} \frac{N_n^{(j^*)}}{n} = 1,$$

(7)
\[
\lim_{n \to \infty} \frac{N_n^{(i)}}{N_n^{(j)}} = \left( \frac{\lambda(i)}{\lambda(j)} \frac{|\mu(j) - \mu(j^*)|}{|\mu(i) - \mu(j^*)|} \right)^2, \quad i, j \neq j^*, \quad (8)
\]

where the limits hold almost surely. Clearly, (7)-(8) do not match (1)-(2) except in the limiting case where \(\alpha_n^{(j^*)} \to 1\). Because \(\frac{N_n^{(j)}}{n} \to 0\) for \(j \neq j^*\), EI will not achieve an exponential convergence rate for any finite \(M\). Ryzhov (2016) also derives the limiting allocations for two other variants of EI, but they do not recover (1)-(2) either.

### 3 Algorithm and main results

Salemi et al. (2014) proposed to replace (6) with
\[
v_n^{(j)} = \mathbb{E} \left( \max \left\{ \mu(j) - \mu(j^*_n), 0 \right\} \mid F_n \right), \quad (9)
\]
which can be written in closed form as
\[
v_n^{(j)} = v_n^{(j)} = \sqrt{\left( \sigma_n^{(j)} \right)^2 + \left( \sigma_n^{(j_*)} \right)^2} f \left( -\frac{\theta_n^{(j)} - \theta_n^{(j_*)}}{\sqrt{\left( \sigma_n^{(j)} \right)^2 + \left( \sigma_n^{(j_*)} \right)^2}} \right) \quad (10)
\]
for any \(j \neq j^*_n\). In this way, the value of collecting information about \(j\) depends, not only on our uncertainty about \(j\), but also on our uncertainty about \(j^*_n\). Salemi et al. (2014) considers a more general Gaussian Markov model with correlated beliefs, so the original presentation of CEI included a term representing the posterior covariance between \(\mu^{(j)}\) and \(\mu^{(j^*_n)}\). In this paper we only consider independent priors, so we work with (10), which translates the CEI concept to our R&S model.

From (9), it follows that \(v_n^{(j^*_n)} = 0\) for all \(n\). Thus, we cannot simply assign \(j_n = \arg \max_j v_n^{(j)}\) because, in that case, \(j^*_n\) would never be chosen. It is necessary to modify the procedure by introducing some additional logic to handle samples assigned to \(j^*_n\). To the best of our knowledge, this issue is not explicitly discussed in Salemi et al. (2014). In fact, many adaptive methods are unable to efficiently identify when \(j^*_n\) should be measured; thus, both the classic EI method of Jones et al. (1998), and the popular Thompson sampling algorithm (Russo & Van Roy, 2014), will sample \(j^*_n\) too often. The class of top-two methods, first introduced in Russo (2017), addresses this problem by essentially assigning a fixed proportion \(\beta\) of samples to \(j^*_n\), while using Thompson sampling or other means to choose between the other alternatives. Optimal allocations can be attained if \(\beta\) is tuned correctly, but the optimal choice of \(\beta\) is problem-dependent and generally difficult to find.
Let $n = 0$ and repeat the following:

1: Check whether

$$\left( \frac{N_n^{(j^*_n)}}{\lambda(j^*_n)} \right)^2 < \sum_{j \neq j^*_n} \left( \frac{N_n^{(j)}}{\lambda(j)} \right)^2.$$  

(11)

If (11) holds, assign $j_n = j^*_n$. If (11) does not hold, assign $j_n = \arg \max_{j \neq j^*_n} v_n^{(j)}$, where $v_n^{(j)}$ is given by (10).

2: Observe $W_{n+1}^{(j_n)}$, update posterior parameters, and increment $n$ by 1.

---

Based on these considerations, we give a modified CEI procedure in Figure 1. The modification adds condition (11), which mimics (1) to decide whether $j^*_n$ should be sampled. This condition is trivial to implement, and the mCEI algorithm is completely free of tunable parameters. It is shown in Chen & Ryzhov (2017) that mCEI samples every alternative infinitely often as $n \to \infty$.

We now state our main results on the asymptotic rate-optimality of mCEI. Essentially, these theorems state that conditions (1) and (2) will hold in the limit as $n \to \infty$. Both theorems should be interpreted in the frequentist sense, that is, $\mu^{(j)}$ is a fixed but unknown constant for each $j$.

**Theorem 3.1** (Optimal alternative). Let $\alpha_n^{(j^*)} = \frac{N_n^{(j^*)}}{n}$. Under the mCEI algorithm,

$$\lim_{n \to \infty} \left( \frac{\alpha_n^{(j^*)}}{\lambda(j^*)} \right)^2 - \sum_{j \neq j^*_n} \left( \frac{\alpha_n^{(j)}}{\lambda(j)} \right)^2 = 0$$

almost surely.

**Theorem 3.2** (Suboptimal alternatives). For $j \neq j^*$, define

$$\tau_n^{(j)} = \frac{\left( \frac{\mu^{(j)}}{\alpha_n^{(j)}} \right)^2}{\left( \frac{\lambda(j)}{\alpha_n^{(j)}} \right)^2 + \left( \frac{\lambda(j^*)}{\alpha_n^{(j^*)}} \right)^2}.$$

where $\alpha_n^{(j)} = \frac{N_n^{(j)}}{n}$. Under the mCEI algorithm,

$$\lim_{n \to \infty} \frac{\tau_n^{(i)}}{\tau_n^{(j)}} = 1$$

almost surely, for any $i, j \neq j^*$.  

---
4 Proofs of main results

For notational convenience, we assume that \( j^* = 1 \) is the unique optimal alternative. Since, under mCEI, \( N_n^{(j)} \to \infty \) for all \( j \), on almost every sample path we will always have \( j_n^* = 1 \) for all large enough \( n \). It is therefore sufficient to prove Theorems 3.1 and 3.2 for a simplified version of mCEI with (10) replaced by

\[
v_n^{(j)} = \sqrt{\frac{(\lambda(j))^2}{N_n^{(j)}} + \frac{(\lambda(1))^2}{N_n^{(1)}}} f \left( -\frac{\theta_n^{(j)} - \theta_n^{(1)}}{\sqrt{\frac{(\lambda(j))^2}{N_n^{(j)}} + \frac{(\lambda(1))^2}{N_n^{(1)}}}} \right),
\]

and (11) replaced by

\[
\left( \frac{N_n^{(1)}}{\lambda(1)} \right)^2 < \sum_{j>1} \left( \frac{N_n^{(j)}}{\lambda(j)} \right)^2.
\]

To simplify the presentation of the key arguments, we treat the noise parameters \( \lambda^{(j)} \) as being known. If, in (4), we replace \( \lambda^{(j)} \) by the standard sample deviation (as recommended, e.g., by both Jones et al., 1998 and Salemi et al., 2014), then simply plug the resulting approximation into (10), the limiting allocation will not be affected. Because the rate-optimality framework of Glynn & Juneja (2004) is frequentist and assumes that selection is based only on sample means, it does not make any distinction between known and unknown variance in terms of characterizing an optimal allocation.

4.1 Proof of Theorem 3.1

First, we define the quantity

\[
\Delta_n \triangleq \left( \frac{N_n^{(1)}/\lambda(1)}{n} \right)^2 - \sum_{j=2}^{M} \left( \frac{N_n^{(j)}/\lambda(j)}{n} \right)^2
\]

and prove the following technical lemma. We remind the reader that, in this and all subsequent proofs, we assume that sampling decisions are made by mCEI with (12)-(13) replacing (10)-(11).

**Lemma 4.1.** If alternative 1 is sampled at time \( n \), then \( \Delta_{n+1} - \Delta_n > 0 \). If any other alternative is sampled at time \( n \), then \( \Delta_{n+1} - \Delta_n < 0 \).

**Proof:** Suppose that alternative 1 is sampled at time \( n \). Then,

\[
\Delta_{n+1} - \Delta_n
\]
\[
\frac{(N_n^{(1)} + 1)}{\lambda(1)} - \frac{M}{2} \left( \frac{N_n^{(j)}}{\lambda(j)} \right)^2 - \left( \frac{N_n^{(1)}}{n} \right)^2 - \left( \frac{N_n^{(1)}}{n + 1} \right)^2 - \sum_{j=2}^{M} \left( \frac{N_n^{(j)}}{\lambda(j)} \right)^2
\]
\[
= \frac{1}{\lambda(1)^2} \left( \frac{(N_n^{(1)} + 1)}{n + 1} \right)^2 - \left( \frac{N_n^{(1)}}{n} \right)^2 - \sum_{j=2}^{M} \left( \frac{N_n^{(j)}}{\lambda(j)} \right)^2 - \sum_{j=2}^{M} \left( \frac{N_n^{(j)}}{n + 1} \right)^2
\]
\[
> 0.
\]

If some alternative \( j' > 1 \) is sampled, then \( \Delta_n \geq 0 \) and

\[
\Delta_{n+1} - \Delta_n = \left( \frac{N_n^{(1)}}{n + 1} \right)^2 - \sum_{j \neq j'} \left( \frac{N_n^{(j)}}{n + 1} \right)^2 - \left( \frac{N_n^{(j')}}{\lambda(j')} \right)^2
\]
\[
= \left( \frac{N_n^{(1)}}{n + 1} \right)^2 - \sum_{j=2}^{M} \left( \frac{N_n^{(j)}}{\lambda(j)} \right)^2 - \frac{2N_n^{(j')}}{\lambda(j')(n + 1)^2}
\]
\[
= \left( \frac{n^2}{(n + 1)^2} - 1 \right) \Delta_n - \frac{2N_n^{(j')}}{(\lambda(j')(n + 1)^2}
\]
\[
< 0,
\]
which completes the proof. \( \Box \)

Let \( \ell = \min_j \lambda^{(j)} \) and recall that \( \ell > 0 \) by assumption. Now, for all \( \varepsilon > 0 \), there exists a large enough \( n_1 \) such that \( n_1 > \frac{2}{\varepsilon^2} - 1 \). Consider arbitrary \( n \geq n_1 \) and suppose that \( \Delta_n < 0 \). This means that alternative 1 is sampled at time \( n \), whence \( \Delta_{n+1} - \Delta_n > 0 \) by Lemma 4.1. Furthermore,

\[
\Delta_{n+1} = \left( \frac{(N_n^{(1)} + 1)}{\lambda(1)} \right)^2 - \sum_{j=2}^{M} \left( \frac{N_n^{(j)}}{\lambda(j)} \right)^2
\]
\[
= \Delta_n + \frac{2N_n^{(1)}}{(\lambda(1)(n + 1)^2)
\]
\[
< \frac{2n + 2}{(\lambda(1)(n + 1)^2}
\]
\[
\leq \frac{2}{(\lambda(1))^2 (n_1 + 1)}
\]
\[
\ell^2 \leq \frac{\lambda(1)^2}{\varepsilon} \leq \varepsilon.
\]

Similarly, suppose that \( \Delta_n \geq 0 \). This means that some \( j' > 1 \) is sampled, whence \( \Delta_{n+1} - \Delta_n < 0 \) by Lemma 4.1. Using similar arguments as before, we find

\[
\Delta_{n+1} = \left( \frac{N_n^{(1)} / \lambda(1)}{n + 1} \right)^2 - \sum_{j=2}^M \left( \frac{N_n^{(j')} / \lambda(j')}{(n+1)} \right)^2 - \frac{2N_n^{(j')} + 1}{(\lambda(j')(n+1))^2}
\]

\[
= \Delta_n - \frac{2N_n^{(j')} + 1}{(\lambda(j')(n+1))^2}
\]

\[
\geq -\frac{2n + 2}{(\lambda(j')(n+1))^2}
\]

\[
\geq -\varepsilon.
\]

Thus, if there exists some large enough \( n_2 \) satisfying \( n_2 \geq n_1 \) and \( -\varepsilon < \Delta_{n_2} < \varepsilon \), then it follows that, for all \( n \geq n_2 \), we have \( \Delta_n \in (-\varepsilon, \varepsilon) \), which implies \( \lim_{n \to \infty} \Delta_n = 0 \) and completes the proof of Theorem 3.1. It only remains to show the existence of such \( n_2 \).

Again, we consider two cases. First, suppose that \( \Delta_{n_1} < 0 \). Since mCEI samples every alternative infinitely often, we can let \( n_2 = \inf\{n > n_1 : \Delta_n \geq 0\} \). Since \( n_2 \) will be the first time after \( n_1 \) that any \( j' > 1 \) is sampled, we have \( \Delta_{n_2-1} < 0 \) and \( n_2 - 1 \geq n_1 \). From the previous arguments, we have \( 0 \leq \Delta_{n_2} < \varepsilon \). Similarly, in the second case where \( \Delta_{n_1} \geq 0 \), we let \( n_2 = \inf\{n > n_1 : \Delta_n < 0\} \), whence \( \Delta_{n_2-1} \geq 0 \) and \( n_2 - 1 \geq n_1 \). The previous arguments imply \( -\varepsilon < \Delta_{n_2} < 0 \). Thus, we can always find \( n_2 \geq n_1 \) satisfying \( -\varepsilon < \Delta_{n_2} < \varepsilon \), as required.

## 4.2 Proof of Theorem 3.2

The proof relies on several technical lemmas; to present the main argument more clearly, these lemmas are stated here, and the full proofs are given in the Appendix. For notational convenience, we define \( \delta_n^{(j)} \triangleq \left| \theta_n^{(j)} - \theta_n^{(1)} \right| \) and \( \delta_n^{(j)} = \left( \delta_n^{(j)} \right)^2 \) for all \( j > 1 \). Furthermore, for any \( j \) and any positive integer \( m \), we define

\[
k_{n,n+m}^{(j)} \triangleq N_{n+m}^{(j)} - N_n^{(j)}
\]

to be the number of samples allocated to alternative \( j \) from stage \( n \) to stage \( n+m-1 \).
The first technical lemma implies that, for any two alternatives \(i\) and \(j\), \(N_n^{(i)} = \Theta \left( N_n^{(j)} \right)\) and \(N_n^{(i)} = \Theta \left( n \right)\).

**Lemma 4.2.** For any two alternatives \(i\) and \(j\), \(\lim_{n \to \infty} \frac{N_n^{(i)}}{N_n^{(j)}} < \infty\).

Now let

\[
z_n^{(j)} \triangleq \frac{d_n^{(j)}}{\sqrt{\left( \frac{\lambda(j)}{N_n^{(j)}} \right)^2 + \left( \frac{\lambda(1)}{N_n^{(1)}} \right)^2}} \;
\]

\[
t_n^{(j)} \triangleq \left( \frac{z_n^{(j)}}{\lambda(j)} \right)^2 = \frac{d_n^{(j)}}{N_n^{(j)}} + \left( \frac{\lambda(1)}{N_n^{(1)}} \right)^2.
\]

For any \(j\), both \(z_n^{(j)}\) and \(t_n^{(j)}\) go to infinity as \(n \to \infty\). We apply an expansion of the Mills ratio (Ruben, 1962) to \(v_n^{(j)}\). For all large enough \(n\),

\[
v_n^{(j)} = \frac{d_n^{(j)}}{z_n^{(j)}} f \left( -z_n^{(j)} \right) = \frac{d_n^{(j)}}{z_n^{(j)}} \phi \left( z_n^{(j)} \right) \left( -z_n^{(j)} \left( 1 - \Phi \left( z_n^{(j)} \right) \right) + 1 \right)
\]

\[
= \frac{d_n^{(j)}}{z_n^{(j)}} \phi \left( z_n^{(j)} \right) \left( -z_n^{(j)} \frac{1 - \Phi \left( z_n^{(j)} \right)}{\phi \left( z_n^{(j)} \right)} + 1 \right)
\]

\[
= \frac{d_n^{(j)}}{z_n^{(j)}} \phi \left( z_n^{(j)} \right) \left( -z_n^{(j)} \frac{1 - 1}{\left( z_n^{(j)} \right)^2} + O \left( \frac{1}{\left( z_n^{(j)} \right)} \right) \right) + 1
\]

\[
= \frac{d_n^{(j)}}{\left( z_n^{(j)} \right)^3} \phi \left( z_n^{(j)} \right) \left( 1 + O \left( \frac{1}{\left( z_n^{(j)} \right)^2} \right) \right),
\]

where (14) comes from the Mills ratio. Then,

\[
2 \log \left( v_n^{(j)} \right) = 2 \log d_n^{(j)} - 6 \log z_n^{(j)} + 2 \log \phi \left( z_n^{(j)} \right) + 2 \log \left( 1 + O \left( \frac{1}{z_n^{(j)}} \right) \right)
\]

\[
= \log \delta_n^{(j)} - 3 \log t_n^{(j)} - \log (2\pi) - t_n^{(j)} + 2 \log \left( 1 + O \left( \frac{1}{t_n^{(j)}} \right) \right)
\]

\[
= -t_n^{(j)} \left( 1 + O \left( \frac{\log t_n^{(j)}}{t_n^{(j)}} \right) \right).
\]

11
For any two suboptimal alternatives \( i \) and \( j \), define
\[
    r_n^{(i,j)} \triangleq \frac{2 \log \left( v_n^{(i)} \right)}{2 \log \left( v_n^{(j)} \right)} = \frac{t_n^{(i)}}{t_n^{(j)}} \left( 1 + O \left( \frac{\log t_n^{(i)}}{t_n^{(j)}} \right) \right).
\]

(15)

and note that both \( 1 + O \left( \log t_n^{(i)} \right) \) and \( 1 + O \left( \log t_n^{(j)} \right) \) converge to 1 as \( n \to \infty \). We will show that \( r_n^{(i,j)} \to 1 \) for any suboptimal \( i \) and \( j \); then, (15) will yield \( \frac{t_n^{(i)}}{t_n^{(j)}} \to 1 \), completing the proof of Theorem 3.2.

Note that, for any \( j \), the CEI quantity \( v_n^{(j)} \) can change when either \( j \) or the optimal alternative is sampled. Thus, it is necessary to characterize the relative frequency of such samples. This requires three other technical lemmas, which are stated below and proved in the Appendix. First, Lemma 4.3 shows that the number of samples that could be allocated to the optimal alternative between two samples of any suboptimal alternatives (not necessarily the same one) is \( O \left( 1 \right) \) and vice versa; next, Lemma 4.4 shows that \( k_{(n,n+m)}^{(1)} \) is \( O \left( \sqrt{n \log \log n} \right) \); finally, Lemma 4.5 bounds
\[
    n^{3/4} \left| \delta_{n+1}^{(i)} - \delta_n^{(i)} \right|.
\]

**Lemma 4.3.** Between two samples assigned to any suboptimal alternatives (i.e., two time stages when condition (13) fails), the number of samples that could be allocated to the optimal alternative is at most equal to some fixed constant \( B_1 \); symmetrically, between two samples of alternative 1, the number of samples that could be allocated to any suboptimal alternatives is at most equal to some fixed constant \( B_2 \).

**Lemma 4.4.** If some suboptimal alternative \( i > 1 \) is sampled at stage \( n \geq 3 \), then \( k_{(n,n+m)}^{(1)} = O \left( \sqrt{n \log \log n} \right) \) for
\[
    m \triangleq \inf \left\{ l > 0 : r_{n+l}^{(i)} = 1 \right\}.
\]

**Lemma 4.5.** For any alternative \( i \), \( n^{3/4} \left| \delta_{n+1}^{(i)} - \delta_n^{(i)} \right| \to 0 \) almost surely as \( n \to \infty \).

Let \( i, j > 1 \) and suppose that \( i \) is sampled at stage \( n \). We will first place an \( O \left( \frac{1}{n^{3/4}} \right) \) bound on the increment \( r_{n+1}^{(i,j)} - r_n^{(i,j)} \). We will then place a bound of \( O \left( \frac{\sqrt{n \log \log n}}{n^{3/4}} \right) \) on the growth of \( r_n^{(i,j)} \).
in between two samples of \(i\) (note that, by definition, \(r_{n}^{(i,j)} \leq 1\) at any stage \(n\) when \(i\) is sampled).

As this bound vanishes to zero as \(n \to \infty\), it will then be shown to follow that \(r_{n}^{(i,j)} \to 1\).

If \(i\) is sampled at stage \(n\), then \(r_{n}^{(i,j)} \leq 1\) and

\[
\begin{align*}
\log \left( \frac{t_{n+1}^{(i,j)}}{t_{n}^{(i,j)}} \right) & = \log \left( \frac{t_{n+1}^{(i)}}{t_{n}^{(i)}} \right) - \log \left( \frac{t_{n+1}^{(j)}}{t_{n}^{(j)}} \right) \\
& = \frac{1}{\log \left( \frac{t_{n}^{(j)}}{t_{n}^{(i)}} \right)} \left( \log \left( \frac{t_{n+1}^{(i)}}{t_{n}^{(i)}} \right) - \log \left( \frac{t_{n+1}^{(j)}}{t_{n}^{(j)}} \right) \right) \\
& \leq \frac{1}{\log \left( \frac{t_{n}^{(j)}}{t_{n}^{(i)}} \right)} \left( \log \left( \frac{t_{n+1}^{(i)}}{t_{n}^{(i)}} \right) - \log \left( \frac{t_{n+1}^{(j)}}{t_{n}^{(j)}} \right) \right) \\
& = \frac{1}{2 \log \left( \frac{t_{n}^{(j)}}{t_{n}^{(i)}} \right)} \left( \frac{\delta_{n+1}^{(i)}}{N_{n+1}^{(i)}} + \frac{\delta_{n}^{(i)}}{N_{n}^{(i)}} \right) - \frac{\delta_{n+1}^{(i)}}{N_{n+1}^{(i)} + \frac{\delta_{n}^{(i)}}{N_{n}^{(i)}}} + \frac{\delta_{n}^{(i)}}{N_{n}^{(i)} + \frac{\delta_{n}^{(i)}}{N_{n}^{(i)}}} \right) \\
& \geq C_1 \frac{n}{n^{3/4}} = C_1 n^{3/4}.
\end{align*}
\]

By Lemma 4.2, there exists a positive constant \(C_1\) such that, for all large enough \(n\),

\[
\begin{align*}
2 \log \left( \frac{t_{n}^{(j)}}{t_{n}^{(i)}} \right) & = \frac{t_{n}^{(j)}}{n^{1/4}} \left( 1 + O \left( \frac{\log t_{n}^{(j)}}{t_{n}^{(j)}} \right) \right) \\
& \geq \frac{1}{2n^{1/4}} \left( \frac{\lambda(i)^2}{N_{n}^{(i)}} + \frac{\lambda(i)^2}{N_{n}^{(i)}} \right) \\
& \geq C_1 \frac{n}{n^{3/4}} = C_1 n^{3/4}.
\end{align*}
\]

On the other hand, for all large enough \(n\), there also exists a positive constant \(C_2\) such that

\[
\begin{align*}
\frac{1}{n^{1/4}} \left( \frac{\delta_{n+1}^{(i)}}{N_{n+1}^{(i)} + \frac{\lambda(i)^2}{N_{n}^{(i)}} + \frac{\lambda(i)^2}{N_{n}^{(i)}}} \right) - \frac{\delta_{n}^{(i)}}{N_{n}^{(i)} + \frac{\lambda(i)^2}{N_{n}^{(i)}} + \frac{\lambda(i)^2}{N_{n}^{(i)}}} \right)
\end{align*}
\]
We have now bounded all four terms in (16). Therefore, for all large enough \( n \) and for all large enough \( n \), where the first equality holds from Lemma 4.2 and the last equality holds from Lemma 4.5. Then

\[
\leq \frac{1}{n^{1/4}} \left( \left| \frac{\delta^{(i)}_{n+1}}{\frac{(\lambda^{(i)})^2}{N_n^{(i)} + 1} + \frac{(\lambda^{(1)})^2}{N_n^{(1)}}} - \frac{\delta^{(i)}_n}{\frac{(\lambda^{(i)})^2}{N_n^{(i)} + 1} + \frac{(\lambda^{(1)})^2}{N_n^{(1)}}} \right| + \left| \frac{\delta^{(i)}_{n+1}}{\frac{(\lambda^{(i)})^2}{N_n^{(i)} + 1} + \frac{(\lambda^{(1)})^2}{N_n^{(1)}}} - \frac{\delta^{(i)}_n}{\frac{(\lambda^{(i)})^2}{N_n^{(i)} + 1} + \frac{(\lambda^{(1)})^2}{N_n^{(1)}}} \right| \right)
\]

\[
= \frac{1}{n^{1/4}} \left( O(1) + O(n) \left| \delta^{(i)}_{n+1} - \delta^{(i)}_n \right| \right)
\]

\[
= O(1) \left( \frac{1}{n^{1/4}} + n^{3/4} \left| \delta^{(i)}_{n+1} - \delta^{(i)}_n \right| \right)
\]

\[
= O(1) \leq C_2,
\]

where the first equality holds from Lemma 4.2 and the last equality holds from Lemma 4.5. Then for all large enough \( n \), we have

\[
\leq \frac{3}{n^{1/4}} \left| \log \left( \frac{\delta^{(i)}_{n+1}}{\frac{(\lambda^{(i)})^2}{N_n^{(i)} + 1} + \frac{(\lambda^{(1)})^2}{N_n^{(1)}}} - \frac{\delta^{(i)}_n}{\frac{(\lambda^{(i)})^2}{N_n^{(i)} + 1} + \frac{(\lambda^{(1)})^2}{N_n^{(1)}}} \right) \right|
\]

\[
\leq \frac{3}{n^{1/4}} \left| \log \left( \frac{\delta^{(i)}_{n+1}}{\frac{(\lambda^{(i)})^2}{N_n^{(i)} + 1} + \frac{(\lambda^{(1)})^2}{N_n^{(1)}}} - \frac{\delta^{(i)}_n}{\frac{(\lambda^{(i)})^2}{N_n^{(i)} + 1} + \frac{(\lambda^{(1)})^2}{N_n^{(1)}}} \right) \right|
\]

\[
\leq 3C_2,
\]

and

\[
\leq \frac{2}{n^{1/4}} \left[ \log \left( 1 + O \left( \frac{(\lambda^{(i)})^2}{N_n^{(i)} + 1} + \frac{(\lambda^{(1)})^2}{N_n^{(1)}} \right) \right) \right] - \log \left( 1 + O \left( \frac{(\lambda^{(i)})^2}{N_n^{(i)} + 1} + \frac{(\lambda^{(1)})^2}{N_n^{(1)}} \right) \right)
\]

\[
+ \frac{1}{n^{1/4}} \left| \log \delta^{(i)}_{n+1} - \log \delta^{(i)}_n \right|
\]

\[
\leq \frac{2}{n^{1/4}} \left[ \log \left( 1 + O \left( \frac{(\lambda^{(i)})^2}{N_n^{(i)} + 1} + \frac{(\lambda^{(1)})^2}{N_n^{(1)}} \right) \right) \right] + \log \left( 1 + O \left( \frac{(\lambda^{(i)})^2}{N_n^{(i)} + 1} + \frac{(\lambda^{(1)})^2}{N_n^{(1)}} \right) \right)
\]

\[
+ \frac{1}{n^{1/4}} \left| \log \delta^{(i)}_{n+1} - \log \delta^{(i)}_n \right|
\]

\[
\leq C_2.
\]

We have now bounded all four terms in (16). Therefore, for all large enough \( n \), we have

\[
r^{(i,j)}_{n+1} - r^{(i,j)}_n \leq \frac{5C_2/C_1}{n^{3/4}},
\]

and

\[
r^{(i,j)}_{n+1} - 1 \leq r^{(i,j)}_n - 1 + \frac{5C_2/C_1}{n^{3/4}} \leq \frac{5C_2/C_1}{n^{3/4}}.
\]
Thus, we have established a bound on the growth of $r^{(i,j)}_n$ that can occur as a result of sampling $i$ at time $n$.

We now consider the growth of the ratio between stages $n$ and $n + m$, where

$$m \triangleq \inf \{ l > 0 : f^{(i)}_{n+l} = 1 \}$$

as in the statement of Lemma 4.4. In words, $n + m$ is the index of the next time after $n$ that we sample $i$. For any stage $n + s$ with $0 < s \leq m$, the inequality $r^{(i,j)}_{n+s+1} > r^{(i,j)}_{n+s}$ can only hold if alternative $j$ or the optimal alternative is sampled at stage $n + s$.

If alternative $j$ is sampled at stage $n + s$, then

$$r^{(i,j)}_{n+s+1} - r^{(i,j)}_{n+s} = \frac{\log (v^{(i)}_{n+s+1})}{\log (v^{(j)}_{n+s+1})} - \frac{\log (v^{(i)}_{n+s})}{\log (v^{(j)}_{n+s})}$$

$$= \frac{\log (v^{(i)}_{n+s}) - \log (v^{(j)}_{n+s})}{\log (v^{(j)}_{n+s+1})} \leq \frac{\log (v^{(i)}_{n+s}) - \log (v^{(j)}_{n+s})}{\log (v^{(j)}_{n+s+1})} \cdot \frac{\log (v^{(j)}_{n+s+1}) - \log (v^{(j)}_{n+s})}{\log (v^{(j)}_{n+s})}.$$  

Using similar arguments as above, we have

$$\left| \frac{\log (v^{(j)}_{n+s+1}) - \log (v^{(j)}_{n+s})}{\log (v^{(j)}_{n+s+1})} \right| = O \left( (n + s)^{-3/4} \right) = O \left( n^{-3/4} \right),$$

and, by Lemma 4.2,

$$\left| \frac{\log (v^{(i)}_{n+s})}{\log (v^{(j)}_{n+s+1})} \right| = O (1).$$

Thus, there exists a constant $C_3$ such that

$$r^{(i,j)}_{n+s+1} - r^{(i,j)}_{n+s} \leq C_3 n^{-3/4}.$$  

On the other hand, if alternative 1 is sampled at stage $n + s$, then

$$r^{(i,j)}_{n+s+1} - r^{(i,j)}_{n+s} = \frac{\log (v^{(i)}_{n+s+1})}{\log (v^{(j)}_{n+s+1})} - \frac{\log (v^{(i)}_{n+s})}{\log (v^{(j)}_{n+s})}.$$
Then, there exists a constant $C$

Similarly as above, we have

\[
\begin{align*}
\left| \frac{\log (v_{n+s}^{(i)})}{\log (v^{(j)}_{n+s+1})} - \frac{\log (v_{n+s}^{(i)})}{\log (v^{(j)}_{n+s+1})} \right| &= \frac{\log (v_{n+s}^{(i)})}{\log (v^{(j)}_{n+s+1})} - \frac{\log (v_{n+s}^{(i)})}{\log (v^{(j)}_{n+s+1})}, \\
\left| \frac{\log (v^{(i)}_{n+s})}{\log (v^{(j)}_{n+s+1})} - \frac{\log (v^{(i)}_{n+s})}{\log (v^{(j)}_{n+s+1})} \right| &= \frac{\log (v^{(i)}_{n+s})}{\log (v^{(j)}_{n+s+1})} - \frac{\log (v^{(i)}_{n+s})}{\log (v^{(j)}_{n+s+1})}.
\end{align*}
\]

Then, there exists a constant $C_4$ such that

\[
\left| r_{n+\delta}^{(i,j)} - r_{n+s}^{(i,j)} \right| \leq C_4 n^{-3/4}.
\]

Therefore, in all cases, for all large enough $n$, we have

\[
\left| r_{n+\delta}^{(i,j)} - r_{n+s}^{(i,j)} \right| \leq \frac{C_5}{n^{3/4}},
\]

where $C_5 = \max \{5C_2/C_1, C_3, C_4\}$. It follows that

\[
\begin{align*}
r_{n+\delta}^{(i,j)} - 1 &\leq r_{n+s}^{(i,j)} - 1 + \frac{C_5}{n^{3/4}} \\
&\leq r_{n}^{(i,j)} - 1 + \left(1 + k^{(j)} + k^{(j)}_s\right) \frac{C_5}{n^{3/4}} \\
&\leq \left(1 + k^{(j)} + k^{(j)}_s\right) \frac{C_5}{n^{3/4}}.
\end{align*}
\]

However, from Lemma 4.4, we have $k^{(1)}_s \leq k^{(1)}_m = O \left(\sqrt{n \log \log n} \right)$ for all $0 < s \leq m$, and at the same time, from Lemma 4.3, we know that at most $B_2$ samples could be allocated to any suboptimal alternatives between two samples of alternative 1. Then we also have $k^{(j)}_s \leq k^{(j)}_m \leq B_2 \left(k^{(1)}_m + 1\right)$, whence $k^{(j)}_m = O \left(\sqrt{n \log \log n} \right)$. It follows that

\[
r_{n+\delta}^{(i,j)} - 1 \leq \left(1 + k^{(j)}_m + k^{(1)}_m\right) \frac{C_5}{n^{3/4}} = O \left(\frac{\sqrt{n \log \log n}}{n^{3/4}}\right),
\]

16
whence \( \limsup_{n \to \infty} r_n^{(i,j)} = 1 \). By symmetry,

\[
\liminf_{n \to \infty} r_n^{(i,j)} = \limsup_{n \to \infty} r_n^{(j,i)} = 1,
\]

whence \( \lim_{n \to \infty} r_n^{(i,j)} = 1 \). This completes the proof.

5 Conclusion

We have considered a ranking and selection problem with independent normal priors and samples, and shown that an EI-type method (a modified version of the CEI method of Salemi et al., 2014) achieves the rate-optimality conditions of Glynn & Juneja (2004) asymptotically. This is the first such result available for any EI-type algorithm (previous rate results for other EI-type methods have shown that those methods achieve suboptimal allocations) that does not require any tuning.

The present paper strengthens the existing body of theoretical support for EI-type methods in general, and for the CEI method in particular. An interesting question is whether CEI would continue to perform optimally in, e.g., the more general Gaussian Markov framework of Salemi et al. (2014). However, the current theoretical understanding of such models is quite limited, and more fundamental questions (for example, how correlated Bayesian models impact the rate of convergence) should be answered before any particular algorithm can be analyzed.

References


### 6 Appendix: additional proofs

Below, we give the full proofs of some technical lemmas that were stated in the text.
6.1 Proof of Lemma 4.2

We proceed by contradiction. Suppose that \( i, j > 1 \) satisfy \( \lim_{n \to \infty} \frac{N_n^{(i)}}{N_n^{(j)}} = \infty \). Let \( c = \lim_{n \to \infty} \frac{\delta_n^{(i)}}{s_n^{(i)}} + 1 = \frac{(\mu^{(i)} - \mu^{(j)})^2}{(\mu^{(i)} - \mu^{(j)})^2} + 1 \). Then, there must exist a large enough stage \( m \) such that

\[
\frac{N_m^{(i)}}{N_m^{(j)}} > \max \{ c, 1 \} \frac{(\lambda^{(i)})^2 + \lambda^{(i)} \lambda^{(j)}}{(\lambda^{(j)})^2},
\]

and we will sample alternative \( i \) to make \( \frac{N_{m+1}^{(i)}}{N_{m+1}^{(j)}} > \frac{N_m^{(i)}}{N_m^{(j)}} \). But, at this stage \( m \),

\[
t_m^{(i)} = \sqrt{\frac{(\lambda^{(i)})^2}{N_m^{(i)}} + \frac{(\lambda^{(j)})^2}{N_m^{(j)}}} f \left( -\frac{d_m^{(i)}}{\sqrt{\frac{(\lambda^{(j)})^2}{N_m^{(j)}} + \frac{(\lambda^{(j)})^2}{N_m^{(j)}}}} \right)
\]

\[
\leq \sqrt{\frac{(\lambda^{(i)})^2}{N_m^{(i)}} + \lambda^{(i)} \lambda^{(j)} f} \left( -\frac{d_m^{(i)}}{\sqrt{\frac{(\lambda^{(j)})^2}{N_m^{(j)}} + \lambda^{(i)} \lambda^{(j)}}} \right)
\]

\[
= \sqrt{\frac{(\lambda^{(i)})^2 + \lambda^{(i)} \lambda^{(j)}}{N_m^{(i)}}} f \left( -\frac{d_m^{(i)}}{\sqrt{(\lambda^{(j)})^2 + \lambda^{(i)} \lambda^{(j)}}} \right)
\]

\[
< \sqrt{\frac{(\lambda^{(j)})^2}{N_m^{(j)}}} f \left( -\frac{d_m^{(j)}}{\sqrt{(\lambda^{(j)})^2}} \right)
\]

\[
< \sqrt{\frac{(\lambda^{(j)})^2}{N_m^{(j)}} + \frac{(\lambda^{(i)})^2}{N_m^{(j)}}} f \left( -\frac{d_m^{(j)}}{\sqrt{(\lambda^{(j)})^2 + \frac{(\lambda^{(i)})^2}{N_m^{(j)}}}} \right)
\]

\[
= n_m^{(j)},
\]

where (17) holds because a suboptimal alternative is sampled at stage \( m \), and (18) holds because \( \lim_{m \to \infty} \frac{d_m^{(i)}}{d_m^{(j)}} = \frac{\mu^{(i)} - \mu^{(j)}}{\mu^{(i)} - \mu^{(j)}} \). From the definition of the mCEI algorithm, (19) implies that we cannot sample \( i \) at stage \( m \). We conclude that \( \lim_{n \to \infty} \frac{N_n^{(i)}}{N_n^{(j)}} < \infty \) for any two suboptimal alternatives \( i \) and \( j \).

From this result, we can see that, for \( i, j > 1 \), we have

\[
0 < \liminf_{n \to \infty} \frac{N_n^{(i)}}{N_n^{(j)}} \leq \limsup_{n \to \infty} \frac{N_n^{(i)}}{N_n^{(j)}} < \infty.
\]
Together with Theorem 3.1, this implies that, for any \( i > 1 \), we have
\[
0 < \liminf_{n \to \infty} \frac{N_n^{(i)}}{N_n^{(1)}} \leq \limsup_{n \to \infty} \frac{N_n^{(i)}}{N_n^{(1)}} < \infty,
\]
completing the proof.

6.2 Proof of Lemma 4.3

Define \( Q_n \triangleq \left( \frac{N_n^{(1)}}{\lambda^{(1)}} \right)^2 - \sum_{j=2}^{M} \left( \frac{N_n^{(j)}}{\lambda^{(j)}} \right)^2 \). Suppose that, at some stage \( n \), \( Q_n < 0 \) and \( Q_{n+1} \geq 0 \), which means that the optimal alternative is sampled at time \( n \) and then a suboptimal alternative is sampled at time \( n+1 \). Let \( m \triangleq \inf \{ l > 0 : Q_{n+l} < 0 \} \), i.e., stage \( n + m \) is the first time that alternative 1 is sampled after stage \( n \). Then, in order to show that between two samples of alternative 1, the number of samples that could be allocated to suboptimal alternatives is \( O(1) \), it is sufficient to show that \( m = O(1) \).

To show this, first we can see that
\[
Q_{n+1} = \left( \frac{N_{n+1}^{(1)}}{\lambda^{(1)}} \right)^2 - \sum_{j=2}^{M} \left( \frac{N_{n+1}^{(j)}}{\lambda^{(j)}} \right)^2
\]
\[
= \left( \frac{N_{n}^{(1)}}{\lambda^{(1)}} \right)^2 - \sum_{j=2}^{M} \left( \frac{N_{n}^{(j)}}{\lambda^{(j)}} \right)^2 + \frac{2N_{n}^{(1)} + 1}{(\lambda^{(1)})^2}
\]
\[
< \frac{2N_{n}^{(1)} + 1}{(\lambda^{(1)})^2}
\]
\[
\leq C_1 N_{n}^{(1)},
\]
where \( C_1 \) is a suitable fixed positive constant and the first inequality holds because \( Q_n < 0 \). Then, for any stage \( n + s \), where \( 0 < s < m \), we have
\[
Q_{n+s} = \left( \frac{N_{n+s}^{(1)}}{\lambda^{(1)}} \right)^2 - \sum_{j=2}^{M} \left( \frac{N_{n+s}^{(j)}}{\lambda^{(j)}} \right)^2
\]
\[
= \left( \frac{N_{n+1}^{(1)}}{\lambda^{(1)}} \right)^2 - \sum_{j=2}^{M} \left( \frac{N_{n+1}^{(j)}}{\lambda^{(j)}} \right)^2
\]
\[
= \left( \frac{N_{n}^{(1)} + 1}{\lambda^{(1)}} \right)^2 - \sum_{j=2}^{M} \left( \frac{N_{n}^{(j)}}{\lambda^{(j)}} \right)^2 - \left( \sum_{j=2}^{M} \left( \frac{N_{n+s}^{(j)}}{\lambda^{(j)}} \right)^2 - \sum_{j=2}^{M} \left( \frac{N_{n}^{(j)}}{\lambda^{(j)}} \right)^2 \right)
\]

21
< \ C_1 N_n^{(1)} - \left( \sum_{j=2}^{M} \left( \frac{N_n^{(j)}}{\lambda^{(j)}} \right)^2 - \sum_{j=2}^{M} \left( \frac{N_n^{(j)}}{\lambda^{(j)}} \right)^2 \right),

where the inequality holds because of (20). We can also see that, after stage \( n \), the increment of 
\( \sum_{j=2}^{M} \left( \frac{N_n^{(j)}}{\lambda^{(j)}} \right)^2 \) obtained by allocating a sample to alternative \( j \) is at least \( 2 \frac{N_n^{(j)}}{\lambda^{(j)}} \). Then, for all large enough \( n \),

\[
\sum_{j=2}^{M} \left( \frac{N_n^{(j)}}{\lambda^{(j)}} \right)^2 - \sum_{j=2}^{M} \left( \frac{N_n^{(j)}}{\lambda^{(j)}} \right)^2 \geq 2s \frac{\min_{j>1} N_n^{(j)}}{\max_{j>1} (\lambda^{(j)})^2} \geq C_2 N_n^{(1)},
\]

where \( C_2 \) is a suitable positive constant and the last inequality follows by Lemma 4.2. Therefore, for any \( 0 < s < m \), we have 
\( Q_{n+s} < (C_1 - C_2 s) N_n^{(1)} \). But, from the definition of \( m \), for any 
\( 0 < s < m \), \( Q_{n+s} \geq 0 \) must hold. Thus, any \( 0 < s < m \) cannot be greater than \( C_1/C_2 \); in other words, we must have \( m \leq C_1/C_2 + 1 \), which implies \( m = O(1) \) for all large enough \( n \). This proves the second claim of the lemma. The first claim of the lemma can be proved in a similar way due to symmetry.

### 6.3 Proof of Lemma 4.4

We first introduce a technical lemma, which establishes a relationship between \( k_{(n,n+m)}^{(1)} \) and samples assigned to suboptimal alternatives. The lemma is proved in a separate subsection of the Appendix.

**Lemma 6.1.** Let \( C_1 \) be any positive constant, and take a large enough \( n \) such that some suboptimal alternative \( i > 1 \) is sampled at stage \( n \). Define

\[
m \triangleq \inf \left\{ l > 0 : I_{n+l}^{(i)} = 1 \right\}, \quad s \triangleq \sup \left\{ l < m : I_{n+l}^{(1)} = 0 \right\}.
\]

Suppose that there exists a sufficiently large positive constant \( C_2 \) (dependent on \( C_1 \), but independent of \( n \)) for which

\[
C_2 \sqrt{n \log \log n} \leq k_{(n,n+s)}^{(1)} \leq n
\]

holds. Then, there exists a suboptimal alternative \( j \neq i \) and a time stage \( n + u \), where \( u \leq s \), such that \( j \) is sampled at stage \( n + u \) and

\[
\left( 1 + C_1 \frac{\sqrt{n \log \log n}}{n} \right) \frac{N_n^{(j)}}{N_n^{(1)}} < \frac{N_n^{(j)}}{N_n^{(1)}} + k_{(n,n+u)}^{(j)} \leq \frac{N_n^{(j)}}{N_n^{(1)}} + k_{(n,n+s)}^{(1)},
\]

holds.
Essentially, Lemma 6.1 will be used to prove the desired result by contradiction; we will show that (21) cannot arise, and therefore $k_{n,n+m}^{(j)}$ must be $O\left(\sqrt{n\log\log n}\right)$.

For convenience, we abbreviate $k_{n,n+m}^{(j)}$ by the notation $k_l^{(j)}$. We will prove the lemma by contradiction. Suppose that the conclusion of the lemma does not hold, that is, $\frac{k_l^{(j)}}{\sqrt{n\log\log n}}$ can be arbitrarily large. Since we sample $i > 1$ at stage $n$, then for any other suboptimal alternative $j \neq i$, we have

$$r_{n}^{(i,j)} = \frac{r_{n}^{(i)} + O\left(\frac{\log r_{n}^{(i)}}{r_{n}^{(j)}}\right)}{1 + O\left(\frac{\log r_{n}^{(j)}}{r_{n}^{(j)}}\right)} \leq 1.$$ 

Then, by Lemma 4.2, there must exist positive constants $C_1$ and $C_2$ such that, for all large enough $n$,

$$\frac{\log r_{n}^{(i)}}{r_{n}^{(j)}} \leq 1 + C_1 \left(\frac{\log r_{n}^{(j)}}{r_{n}^{(i)}} + \frac{\log r_{n}^{(j)}}{r_{n}^{(i)}}\right) \leq 1 + C_2 \frac{\log n}{n},$$

that is, equivalently,

$$\frac{\delta_n^{(i)} (\lambda(j))^2}{N_n^{(j)}} + \frac{\delta_n^{(j)} (\lambda(i))^2}{N_n^{(i)}} \leq \frac{\delta_n^{(i)} (\lambda(i))^2}{1 + C_2 \log n} + \frac{\delta_n^{(j)} (\lambda(j))^2}{1 + C_2 \log n}.$$ 

Then, at stage $n+u$, where $0 < u < m$, there must exist positive constants $C_3$ and $C_4$ such that, for all large enough $n$,

$$r_{n+u}^{(i,j)} = \frac{r_{n+u}^{(i)} + O\left(\frac{\log r_{n+u}^{(i)}}{r_{n+u}^{(j)}}\right)}{1 + O\left(\frac{\log r_{n+u}^{(j)}}{r_{n+u}^{(i)}}\right)} \leq \frac{1}{r_{n+u}^{(j)} + C_3 \left(\frac{\log r_{n+u}^{(i)}}{r_{n+u}^{(j)}} + \frac{\log r_{n+u}^{(i)}}{r_{n+u}^{(j)}}\right)} < \frac{1}{C_4 \log n}.$$ 

Thus, for all large enough $n$, in order to have $r_{n+u}^{(i,j)} < 1$, it is sufficient to require

$$\frac{r_{n+u}^{(i)}}{r_{n+u}^{(j)}} \leq 1 - C_4 \frac{\log n}{n},$$

or, equivalently,

$$\frac{\delta_n^{(i)} (\lambda(j))^2}{N_n^{(j)} + k_u^{(j)}} + \frac{\delta_n^{(j)} (\lambda(i))^2}{N_n^{(i)} + k_u^{(i)}} \leq \frac{\delta_n^{(i)} (\lambda(i))^2}{1 - C_4 \frac{\log n}{n}} + \frac{\delta_n^{(j)} (\lambda(j))^2}{1 - C_4 \frac{\log n}{n}}.$$ 

(23)
Note that \( k_u^{(i)} = 1 \). By the convergence of \( \delta_n^{(i)} \) and \( \delta_n^{(j)} \), for all large enough \( n \), we have

\[
\left( \delta_n^{(j)} - \delta_n^{(i)} \right) \left( \delta_n^{(j)} \left( 1 + C_2 \frac{\log n}{n} \right) - \delta_n^{(i)} \right) > 0,
\]

\[
\left( \delta_n^{(j)} - \delta_n^{(i)} \right) \left( \delta_n^{(j)} \left( 1 - C_4 \frac{\log n}{n} \right) - \delta_n^{(i)} \right) > 0.
\]

If \( \lim_{n \to \infty} \frac{\delta_n^{(j)}}{\delta_n^{(i)}} > 1 \), i.e., \( \mu^{(j)} < \mu^{(i)} \), then by (22) we have

\[
\frac{\delta^{(i)}_{n+u} \left( \lambda^{(j)} \right)^2}{N^{(j)}_{n} + k^{(j)}_u} = \frac{\delta^{(i)}_n \delta^{(i)}_{n+u} \left( \lambda^{(j)} \right)^2}{N^{(j)}_{n} N^{(j)}_{n+u} + k^{(j)}_u} \frac{N^{(j)}_{n}}{N^{(j)}_{n+u} + k^{(j)}_u} \leq \frac{\delta^{(i)}_n \delta^{(i)}_{n+u} \left( \lambda^{(j)} \right)^2}{N^{(j)}_{n} N^{(j)}_{n+u} + k^{(j)}_u} \frac{N^{(j)}_{n}}{N^{(j)}_{n+u} + k^{(j)}_u} \frac{N^{(j)}_{n+1} + 1}{N^{(j)}_{n+1} + k^{(j)}_u} \frac{N^{(j)}_{n+1} + k^{(j)}_u}{N^{(j)}_{n+1} + k^{(j)}_u} \frac{N^{(j)}_{n}}{N^{(j)}_{n} + k^{(j)}_u} \frac{N^{(j)}_{n} + k^{(j)}_u}{N^{(j)}_{n} + k^{(j)}_u}.
\]

It follows that there must exist a positive constant \( C_5 \) such that

\[
\frac{\delta^{(i)}_{n+u} \left( \lambda^{(j)} \right)^2}{N^{(j)}_{n} + k^{(j)}_u} \leq \frac{\delta^{(i)}_{n+u} \delta^{(j)}_n \left( \lambda^{(j)} \right)^2}{N^{(j)}_{n} + k^{(j)}_u} \frac{N^{(j)}_{n+1} + 1}{N^{(j)}_{n+1} + k^{(j)}_u} \frac{N^{(j)}_{n+1} + k^{(j)}_u}{N^{(j)}_{n+1} + k^{(j)}_u} \frac{N^{(j)}_{n}}{N^{(j)}_{n} + k^{(j)}_u} \frac{N^{(j)}_{n} + k^{(j)}_u}{N^{(j)}_{n} + k^{(j)}_u}.
\]

Thus, to satisfy (23), it is sufficient to have

\[
\frac{\delta^{(i)}_{n+u} \delta^{(j)}_n \left( 1 + C_5 \frac{\log n}{n} \right) N^{(j)}_{n} + 1}{N^{(j)}_{n} + k^{(j)}_u} \frac{N^{(j)}_{n}}{N^{(j)}_{n} + k^{(j)}_u} \leq 1,
\]

(24)

\[
\frac{\delta^{(i)}_{n+u} \delta^{(j)}_n \left( 1 + C_2 \frac{\log n}{n} \right) - \delta^{(i)}_n N^{(j)}_{n} + k^{(j)}_u}{N^{(j)}_{n} + k^{(j)}_u} \frac{N^{(j)}_{n}}{N^{(j)}_{n} + k^{(j)}_u} \leq 1.
\]

(25)

Note that for all large enough \( n \) and any alternative \( i \neq 1 \), by Lemma 4.2, we have

\[
\left| \delta^{(i)}_{n+u} - \delta^{(i)}_n \right| = \left| \left( d^{(i)}_{n+u} \right)^2 - \left( d^{(i)}_n \right)^2 \right|
\]
\[
\frac{\delta_{n+u}^{(i)}}{\delta_{n}^{(i)}} = \frac{\delta_{n+u}^{(i)} - \delta_{n+u}^{(j)} + \delta_{n}^{(i)} - \delta_{n}^{(j)}}{\delta_{n}^{(i)} - \delta_{n+u}^{(i)} - \delta_{n+u}^{(j)}} = 1 + O \left( \frac{\log \log n}{n} \right),
\]

\[
\frac{\delta_{n+u}^{(j)}}{\delta_{n}^{(j)}} = 1 + O \left( \frac{\log \log n}{n} \right),
\]

and

\[
\frac{\delta_{n}^{(j)} \left( 1 + C_2 \frac{\log n}{n} \right) - \delta_{n}^{(i)}}{\delta_{n+u}^{(j)} \left( 1 - C_4 \frac{\log n}{n} \right) - \delta_{n+u}^{(i)}} = 1 + \frac{\delta_{n}^{(j)} \left( 1 + C_2 \frac{\log n}{n} \right) - \delta_{n+u}^{(j)} \left( 1 - C_4 \frac{\log n}{n} \right) - \left( \delta_{n}^{(i)} - \delta_{n+u}^{(i)} \right)}{\delta_{n+u}^{(j)} \left( 1 - C_4 \frac{\log n}{n} \right) - \delta_{n+u}^{(i)}}.
\]

\[
\leq 1 + \frac{\delta_{n}^{(j)} \left( 1 + C_2 \frac{\log n}{n} \right) - \delta_{n+u}^{(j)} \left( 1 - C_4 \frac{\log n}{n} \right) + \delta_{n+u}^{(j)} \left( 1 - C_4 \frac{\log n}{n} \right) - \left( \delta_{n}^{(i)} - \delta_{n+u}^{(i)} \right)}{\delta_{n+u}^{(j)} \left( 1 - C_4 \frac{\log n}{n} \right) - \delta_{n+u}^{(i)}}.
\]

\[
= 1 + O \left( \frac{\log \log n}{n} \right).
\]

Then together with Lemma 4.2, there exists a positive constant \(C_5\) such that, for all large enough \(n\), the LHS of (24) satisfies

\[
\frac{\delta_{n+u}^{(i)} \delta_{n+u}^{(j)} \left( 1 + C_5 \frac{\log n}{n} \right)}{\delta_{n}^{(i)} \delta_{n+u}^{(j)} \left( 1 + C_5 \frac{\log n}{n} \right)^2} \frac{N_{n+u}^{(i)} + 1}{N_{n}^{(i) - N_{n}^{(j)}}}
\]

\[
= \left( 1 + O \left( \frac{\log \log n}{n} \right) \right) \left( 1 + C_5 \frac{\log n}{n} \right) \left( 1 + O \left( \frac{1}{n} \right) \right) N_{n}^{(j)} - N_{n}^{(j)}
\]

where the fourth equality holds because of the law of the iterated logarithm, and the last equality holds by Lemma 4.2. Then for all large enough \(n\), we have

\[
\frac{\theta_{n+u}^{(i)} - \theta_{n+u}^{(1)}}{\delta_{n+u}^{(i)}} = \frac{\theta_{n+u}^{(i)} - \theta_{n+u}^{(1)} + \theta_{n}^{(i)} - \theta_{n}^{(1)}}{\delta_{n}^{(i)} - \delta_{n+u}^{(i)}} = 1 + O \left( \frac{\log \log n}{n} \right),
\]

\[
\frac{\theta_{n+u}^{(j)} - \mu^{(j)}}{\delta_{n+u}^{(j)}} = \frac{\theta_{n+u}^{(j)} - \mu^{(j)} + \theta_{n}^{(j)} - \mu^{(j)}}{\delta_{n}^{(j)} - \delta_{n+u}^{(j)}} = 1 + O \left( \frac{\log \log n}{n} \right),
\]

and

\[
\frac{\theta_{n}^{(i)} - \theta_{n}^{(1)}}{\delta_{n}^{(i)}} = \frac{\theta_{n}^{(i)} - \theta_{n}^{(1)} + \theta_{n}^{(j)} - \theta_{n}^{(1)}}{\delta_{n}^{(j)} - \delta_{n}^{(1)}} = 1 + O \left( \frac{\log \log n}{n} \right).
\]
\[ \leq C_6 \sqrt{n \log \log n}, \]

while the LHS of (25) satisfies

\[
\frac{\delta_n^{(i)}}{\delta_{n+u}^{(i)}} \frac{\delta_n^{(j)}}{\delta_{n+u}^{(j)}} \left( 1 + C_2 \frac{\log n}{n} \right) - \left( 1 - C_4 \frac{\log n}{n} \right) \frac{N_n^{(1)}}{N_n^{(1)}} \frac{N_n^{(j)}}{N_n^{(j)} + k_u^{(j)}} \\
= \left( 1 + O \left( \sqrt{\frac{\log \log n}{n}} \right) \right) \frac{N_n^{(1)} + k_u^{(1)}}{N_n^{(1)}} \frac{N_n^{(j)}}{N_n^{(j)} + k_u^{(j)}} \\
\leq \left( 1 + C_6 \sqrt{\frac{n \log \log n}{n}} \right) \frac{N_n^{(1)} + k_u^{(1)}}{N_n^{(1)}} \frac{N_n^{(j)}}{N_n^{(j)} + k_u^{(j)}}.
\]

Therefore, to satisfy (24), it is sufficient to have

\[ C_6 \sqrt{n \log \log n} \leq k_u^{(j)}. \text{ (26)} \]

Now define

\[ s \triangleq \sup \left\{ l < m : I_{n+l}^{(1)} = 0 \right\}. \text{ (27)} \]

Since \( k_{n+u}^{(1)} \sqrt{n \log \log n} \) can be arbitrarily large, we can suppose that \( k_{n+u}^{(1)} > C_7 \sqrt{n \log \log n} \), where \( C_7 \) is a positive constant to be specified. By Lemma 6.1, since \( C_6 \) is a fixed positive constant, there must exist a constant \( C_8 \) such that, if \( C_7 \geq C_8 \), there exists a suboptimal \( j \neq i \), and a stage \( n + u \) with \( u \leq s \), such that \( j \) is sampled at stage \( n + u \) and

\[ \left( 1 + C_6 \sqrt{\frac{n \log \log n}{n}} \right) \frac{N_n^{(j)}}{N_n^{(1)}} < \frac{N_n^{(j)} + k_u^{(j)}}{N_n^{(1)} + k_u^{(1)}} \leq \frac{N_n^{(j)} + k_u^{(j)}}{N_n^{(1)} + k_u^{(1)}}. \]

Then, (25) holds at stage \( n + u \). At the same time, since

\[ \frac{N_n^{(j)} + k_u^{(j)}}{N_n^{(1)} + k_u^{(1)}} \geq \left( 1 + C_6 \sqrt{\frac{n \log \log n}{n}} \right) \frac{N_n^{(j)}}{N_n^{(1)}} \geq \frac{N_n^{(j)}}{N_n^{(1)}}, \]

we have \( k_u^{(j)} \geq C_9 k_u^{(1)} \). From Lemma 4.2, there must exist a positive constant \( C_9 \) such that, for all large enough \( n \),

\[ k_u^{(j)} \geq C_9 k_u^{(1)} \geq C_9 C_7 \sqrt{n \log \log n}. \]
Now let \( C_7 = \max \{ C_8, \frac{C_9}{C_9} \} \). Then, both (25) and (26) are satisfied at stage \( n + u \), so (23) is satisfied, which means
\[
r^{(i,j)}_{n+u} < 1 \quad \Rightarrow \quad v^{(i)}_{n+u} > v^{(j)}_{n+u}.
\]
But the alternative \( j \) is sampled at stage \( n + u \), which means \( v^{(i)}_{n+u} \leq v^{(j)}_{n+u} \). The desired contradiction follows.

Now, consider the other case where \( \lim_{n \to \infty} \frac{\delta^{(j)}_{n+u}}{\delta^{(i)}_{n+u}} < 1 \), i.e., \( \mu^{(j)} > \mu^{(i)} \). By (22), we have
\[
\frac{\delta^{(j)}_{n+u} (\lambda^{(i)})^2 \left(1 - C_4 \log \frac{n}{n} \right)}{N^{(i)}_{n} + 1} = \frac{\delta^{(j)}_{n+u} \delta^{(j)}_{n} (\lambda^{(i)})^2 \left(1 + C_2 \log \frac{n}{n} \right) \left(1 - C_4 \log \frac{n}{n} \right) N^{(i)}_{n}}{N^{(j)}_{n} + 1} \geq \frac{\delta^{(j)}_{n+u} \delta^{(i)}_{n} (\lambda^{(j)})^2 \left(1 - C_4 \log \frac{n}{n} \right) N^{(i)}_{n}}{N^{(j)}_{n} + 1} + \frac{\delta^{(j)}_{n+u} \delta^{(i)}_{n} (\lambda^{(i)})^2 - \delta^{(j)}_{n+u} (\lambda^{(j)})^2 \left(1 + C_2 \log \frac{n}{n} \right) \left(1 - C_4 \log \frac{n}{n} \right) N^{(i)}_{n}}{N^{(j)}_{n} + 1} \geq \frac{1}{N^{(j)}_{n} + k^{(j)}_{n+u} N^{(i)}_{n}}.
\]
Then, there must exist a positive constant \( C_{10} \) such that, for all large enough \( n \),
\[
\frac{\delta^{(j)}_{n+u} (\lambda^{(i)})^2 \left(1 - C_4 \log \frac{n}{n} \right)}{N^{(i)}_{n} + 1} \geq \frac{\delta^{(j)}_{n+u} \delta^{(i)}_{n} (\lambda^{(j)})^2 \left(1 + C_2 \log \frac{n}{n} \right) \left(1 - C_4 \log \frac{n}{n} \right) N^{(i)}_{n}}{N^{(j)}_{n} + 1} + \frac{\delta^{(j)}_{n+u} \delta^{(i)}_{n} (\lambda^{(i)})^2 - \delta^{(j)}_{n+u} (\lambda^{(j)})^2 \left(1 + C_2 \log \frac{n}{n} \right) \left(1 - C_4 \log \frac{n}{n} \right) N^{(i)}_{n}}{N^{(j)}_{n} + 1} \geq \frac{1}{N^{(j)}_{n} + k^{(j)}_{n+u} N^{(i)}_{n}} + \frac{1}{N^{(j)}_{n} + k^{(j)}_{n+u} N^{(i)}_{n}} \geq \frac{1}{N^{(j)}_{n} + k^{(j)}_{n+u} N^{(i)}_{n}}.
\]
Thus, to satisfy (23), for all large enough \( n \), it is sufficient to have
\[
\frac{\delta^{(j)}_{n+u} \delta^{(i)}_{n} \left(1 + C_2 \log \frac{n}{n} \right) \left(1 - C_4 \log \frac{n}{n} \right) N^{(i)}_{n}}{N^{(j)}_{n} + 1} \geq \frac{1}{N^{(j)}_{n} + k^{(j)}_{n+u} N^{(i)}_{n}} + \frac{1}{N^{(j)}_{n} + k^{(j)}_{n+u} N^{(i)}_{n}} \geq \frac{1}{N^{(j)}_{n} + k^{(j)}_{n+u} N^{(i)}_{n}} + \frac{1}{N^{(j)}_{n} + k^{(j)}_{n+u} N^{(i)}_{n}}.
\]
which can equivalently be rewritten as
\[
k^{(j)}_{u} \geq \frac{\delta^{(j)}_{n+u} \delta^{(i)}_{n} \left(1 + C_2 \log \frac{n}{n} \right) \left(1 - C_4 \log \frac{n}{n} \right) N^{(i)}_{n}}{N^{(j)}_{n} + 1} \frac{N^{(i)}_{n} + 1}{N^{(j)}_{n} + k^{(j)}_{n+u} N^{(i)}_{n}} - N^{(j)}_{n},
\]
\[
k^{(1)}_{u} \geq \frac{\delta^{(j)}_{n+u} \delta^{(i)}_{n} \left(1 + C_2 \log \frac{n}{n} \right) \left(1 - C_4 \log \frac{n}{n} \right) N^{(i)}_{n}}{N^{(j)}_{n} + 1} \frac{N^{(i)}_{n} + 1}{N^{(j)}_{n} + k^{(j)}_{n+u} N^{(i)}_{n}} - N^{(j)}_{n}.
\]
Similarly as above, by Lemma 4.2, there exist positive constants $C_{11}, C_{12}, C_{13}$ and $C_{14}$ such that, for all large enough $n$,

$$\frac{\delta^{(j)}_n}{\delta^{(j)}_{n+u}} \frac{\delta^{(i)}_n}{\delta^{(i)}_{n+u}} \left( 1 + C_{10} \frac{\log n}{n} \right) \frac{N^{(i)}_n + 1}{N^{(i)}_n} N^{(j)}_n - N^{(j)}_n \leq C_{11} \sqrt{n \log \log n},$$

and

$$\frac{\delta^{(j)}_n}{\delta^{(j)}_{n+u}} \left( 1 + C_{10} \frac{\log n}{n} \right) \frac{\delta^{(i)}_n}{\delta^{(i)}_{n+u}} \left( 1 + C_{12} \frac{\log n}{n} \right) \frac{N^{(i)}_n + 1}{N^{(i)}_n} N^{(1)}_n - N^{(1)}_n \leq C_{14} \sqrt{n \log \log n}.$$ 

Therefore, to satisfy (28) and (29), it is sufficient to have

$$k^{(j)}_u \geq C_{11} \sqrt{n \log \log n}, \quad (30)$$

$$k^{(1)}_u \geq C_{14} \sqrt{n \log \log n}. \quad (31)$$

Again, define $s$ as in (27). Since $\frac{k^{(j)}_u}{\sqrt{n \log \log n}}$ can be arbitrarily large, we can suppose that $k^{(1)}_u > C_{15} \sqrt{n \log \log n}$, where $C_{15}$ is a positive constant to be specified. By Lemma 6.1, since $C_{11}$ is a fixed positive constant, there must exist a constant $C_{16}$ such that, if $C_{15} \geq C_{16}$, there exists a suboptimal alternative $j \neq i$, and a stage $n + u$ with $u \leq s$, such that $j$ is sampled at stage $n + u$ and

$$\left( 1 + C_{11} \frac{\sqrt{n \log \log n}}{n} \right) N^{(j)}_n < \frac{N^{(j)}_n + k^{(j)}_u}{N^{(1)}_n + k^{(1)}_u} \leq \frac{N^{(j)}_n + k^{(j)}_u}{N^{(1)}_n + k^{(1)}_u},$$

whence

$$\frac{N^{(j)}_n + k^{(j)}_u}{N^{(1)}_n + k^{(1)}_u} > \left( 1 + C_{11} \frac{\sqrt{n \log \log n}}{n} \right) \frac{N^{(j)}_n}{N^{(1)}_n} \geq \frac{N^{(j)}_n}{N^{(1)}_n}.$$ 

Then, we have $\frac{k^{(j)}_u}{k^{(1)}_u} \geq \frac{N^{(j)}_n}{N^{(1)}_n}$. From Lemma 4.2, there must exist a positive constant $C_{17}$ such that for all large enough $n$,

$$k^{(j)}_u \geq C_{17} k^{(1)}_u \geq C_{17} C_{15} \sqrt{n \log \log n}.$$ 

28
At the same time, by Lemma 4.3, for all large enough \( n \), we also have

\[
k_u^{(1)} \geq \frac{k_u^{(j)} + 1}{B_2} - 1 \geq \frac{C_{17}C_{15}\sqrt{n \log \log n} + 1}{B_2} - 1 \geq \frac{C_{17}C_{15}\sqrt{n \log \log n}}{2B_2}.
\]

Now, let \( C_{15} = \max \{ C_{16}, \frac{C_{11}}{C_{17}}, \frac{2B_2C_{14}}{C_{17}} \} \). Then both (30) and (31) are satisfied at stage \( n + u \), so (23) is satisfied, which means that

\[
r_{n+u}^{(i,j)} < 1 \quad \Rightarrow \quad v_{n+u}^{(i)} > v_{n+u}^{(j)}.
\]

But the alternative \( j \) is sampled at stage \( n + u \), which means that \( v_{n+u}^{(i)} \leq v_{n+u}^{(j)} \). Again, we have the desired contradiction.

### 6.4 Proof of Lemma 4.5

First, if an alternative \( j \) other than 1 or \( i \) is sampled at stage \( n \), it is obvious that \( n^{3/4} \left| \delta_{n+1}^{(i)} - \delta_n^{(i)} \right| = 0 \).

Second, if alternative \( i \) is sampled at stage \( n \), then for all large enough \( n \), there exists a constant \( C_1 \) such that

\[
n^{3/4} \left| \delta_{n+1}^{(i)} - \delta_n^{(i)} \right| = n^{3/4} \left| \left( d_{n+1}^{(i)} \right)^2 - \left( d_n^{(i)} \right)^2 \right| \\
\leq C_1 n^{3/4} \left| d_{n+1}^{(i)} - d_n^{(i)} \right| \\
= \frac{C_1}{n^{1/4}} \left| \theta_{n+1}^{(i)} - \theta_n^{(i)} \right|,
\]

where

\[
n \left| \theta_{n+1}^{(i)} - \theta_n^{(i)} \right| = n \left| \frac{1}{N_n^{(i)} + 1} \left( N_n^{(i)} \theta_n^{(i)} + W_n^{(i)} \right) - \theta_n^{(i)} \right| \\
\leq \frac{n}{N_n^{(i)} + 1} \left| \theta_n^{(i)} \right| + \frac{n}{N_n^{(i)} + 1} \left| W_n^{(i)} \right| \\
= O(1) \left( 1 + \left| W_n^{(i)} \right| \right),
\]

thus there exists a constant \( C_2 \) such that

\[
n^{3/4} \left| \delta_{n+1}^{(i)} - \delta_n^{(i)} \right| \leq \frac{C_2}{n^{1/4}} \left( 1 + \left| W_n^{(i)} \right| \right).
\]
Finally, if alternative 1 is sampled at stage $n$, then similarly as above, for all large enough $n$, there exist constants $C_3$ and $C_4$ such that

$$n^{3/4} \left| \delta_{n+1}^{(i)} - \delta_n^{(i)} \right| \leq \frac{C_3}{n^{1/4}} n \left| \theta_{n+1}^{(1)} - \theta_n^{(1)} \right| \leq \frac{C_4}{n^{1/4}} \left( 1 + \left| W_{n+1}^{(1)} \right| \right).$$

Then it is sufficient to show $\left| W_{n+1}^{(i)} \right| / n^{1/4} \rightarrow 0$ and $\left| W_{n+1}^{(1)} \right| / n^{1/4} \rightarrow 0$ almost surely. By Markov’s inequality, for all $\varepsilon > 0$,

$$P \left( \frac{\left| W_{n+1}^{(i)} \right|}{n^{1/4}} \geq \varepsilon \right) \leq \mathbb{E} \left( \frac{\left( W_{n+1}^{(i)} \right)^8}{n^2 \varepsilon^8} \right) \leq \frac{C_5}{n^2 \varepsilon^8},$$

where $C_5$ is a fixed constant, thus $\left| W_{n+1}^{(i)} \right| / n^{1/4} \rightarrow 0$ in probability. Furthermore, by the Borel-Cantelli lemma, since

$$\sum_n P \left( \frac{\left| W_{n+1}^{(i)} \right|}{n^{1/4}} \geq \varepsilon \right) \leq \sum_n \frac{C_5}{n^2 \varepsilon^8} < \infty,$$

then we have $\left| W_{n+1}^{(i)} \right| / n^{1/4} \rightarrow 0$ almost surely. Using similar arguments, we also have $\left| W_{n+1}^{(1)} \right| / n^{1/4} \rightarrow 0$ almost surely, completing the proof.

### 6.5 Proof of Lemma 6.1

For convenience, we abbreviate $k_{(n,n+m)}^{(j)}$ by the notation $k_m^{(j)}$ for all $j$. First, since $C_2$ is a constant and $\lim_{n \to \infty} \sqrt[n]{\log \log n} \leq C_2 \sqrt{n \log \log \log n} = 0$, it follows that, for all large enough $n$, we must have $C_2 \sqrt{n \log \log n} \leq n$. Intuitively, from the definition of $m$ and $s$, stage $n+m$ is the first time that alternative $i$ is sampled after stage $n$, and stage $n+s$ is the last time that a suboptimal alternative is sampled before stage $n+m$. Recall that, by assumption, we must have $C_2 \sqrt{n \log \log n} \leq k_n^{(1)} \leq n$ for some positive constant $C_2$ to be specified.

At stage $n$, since we sample a suboptimal $i$ by assumption, we must have

$$\left( N_{n+1}^{(1)} / \lambda^{(1)} \right)^2 \geq \sum_{j=2}^M \left( N_{n+1}^{(j)} / \lambda^{(j)} \right)^2.$$  \hfill (32)
At stage $n + s$, from the definition of $s$, it is also some suboptimal alternative that is sampled. Repeating the arguments in the proof of Theorem 3.1, we obtain

$$\left( \frac{N_n^{(1)} + k_s^{(1)}}{n + s} / \lambda^{(1)} \right)^2 \leq \sum_{j=2}^{M} \left( \frac{N_n^{(j)} + k_s^{(j)}}{n + s} / \lambda^{(j)} \right)^2 \leq \frac{C_3}{n}$$

for some fixed positive constant $C_3$. Note that $k_s^{(i)} = 1$, whence

$$\sum_{j \geq 2, j \neq i} \left( \frac{N_n^{(j)} + k_s^{(j)}}{N_n^{(1)} + k_s^{(1)}} / \lambda^{(1)} \right)^2 \leq \sum_{j \geq 2, j \neq i} \left( \frac{N_n^{(j)} + 1}{N_n^{(1)} + k_s^{(1)}} / \lambda^{(i)} \right)^2 + \frac{C_3}{n} \left( \frac{n + s}{N_n^{(1)} + k_s^{(1)}} / \lambda^{(1)} \right)^2 \geq 1.$$  

From Lemma 4.2, we know that $\liminf_{n \to \infty} \frac{N_n^{(1)}}{n} > 0$. Then, there must exist some constant $C_4$ such that

$$C_3 \left( \frac{n + s}{N_n^{(1)} + k_s^{(1)}} / \lambda^{(1)} \right)^2 \leq C_4,$$

whence

$$\sum_{j \geq 2, j \neq i} \left( \frac{N_n^{(j)} + k_s^{(j)}}{N_n^{(1)} + k_s^{(1)}} / \lambda^{(1)} \right)^2 \geq 1 - \left( \frac{N_n^{(i)} + 1}{N_n^{(1)} + k_s^{(1)}} / \lambda^{(i)} \right)^2 - \frac{C_4}{n},$$

and for all large enough $n$,

\[
\sum_{j \geq 2, j \neq i} \left[ \left( \frac{N_n^{(j)} + k_s^{(j)}}{N_n^{(1)} + k_s^{(1)}} / \lambda^{(1)} \right)^2 - \left( \frac{N_n^{(j)} / \lambda^{(j)}}{N_n^{(1)} / \lambda^{(1)}} \right)^2 \right] \\
\geq 1 - \sum_{j \geq 2, j \neq i} \left( \frac{N_n^{(j)} / \lambda^{(j)}}{N_n^{(1)} / \lambda^{(1)}} \right)^2 - \left( \frac{N_n^{(i)} + 1}{N_n^{(1)} + k_s^{(1)}} / \lambda^{(i)} \right)^2 - \frac{C_4}{n} \\
\geq \left( \frac{\lambda^{(1)}}{\lambda^{(i)}} \right)^2 \left( \frac{N_n^{(i)}}{N_n^{(1)}} \right)^2 \left( \frac{N_n^{(i)} + k_s^{(i)}}{N_n^{(1)}} \right)^2 - \left( \frac{N_n^{(i)} + 1}{N_n^{(1)} + k_s^{(1)}} \right)^2 - \frac{C_4}{n} \\
= \left( \frac{\lambda^{(1)}}{\lambda^{(i)}} \right)^2 \left( \frac{N_n^{(i)}}{N_n^{(1)}} \right)^2 \left( 2N_n^{(1)} k_s^{(1)} + \left( k_s^{(1)} \right)^2 \right) - \left( \frac{N_n^{(1)}}{N_n^{(1)} + k_s^{(1)}} \right)^2 \left( 2N_n^{(i)} + 1 \right) - \frac{C_4}{n}
\] (33)
where (35) and (36) hold because $k_s^{(1)} \leq n$. Then,

$$\sum_{j \geq 2, j \neq i} \left[ \left( \frac{N_n^{(j)} + k_s^{(j)}}{N_n^{(1)} + k_s^{(1)}} / \lambda^{(j)} \right)^2 - \left( \frac{N_n^{(j)} / \lambda^{(j)}}{N_n^{(1)} / \lambda^{(1)}} \right)^2 \right] > \frac{C_9 C_2 \sqrt{n \log \log n}}{n},$$
so there must be some suboptimal $j$ such that

$$\left( \frac{N_n^{(j)} + k_s^{(j)}}{N_n^{(1)} + k_s^{(1)}} \right) / \lambda^{(j)} > \frac{1}{M - 2} \frac{C_9 C_2 \sqrt{n \log \log n}}{n}. $$

Let $C_{10} = \frac{C_9}{M - 2}$ and $C_{11} = \frac{C_9 C_2}{4}$. Then,

$$\left( \frac{N_n^{(j)} + k_s^{(j)}}{N_n^{(1)} + k_s^{(1)}} \right) ^2 > 1 + \frac{C_{10} C_2 \sqrt{n \log \log n}}{n},$$

and, for all large enough $n$, we have

$$\frac{N_n^{(j)} + k_s^{(j)}}{N_n^{(1)} + k_s^{(1)}} > 1 + \frac{C_{11} \sqrt{n \log \log n}}{n}, \quad (37)$$

whence

$$\frac{N_n^{(j)} + k_s^{(j)}}{N_n^{(1)} + k_s^{(1)}} > \left( 1 + C_{11} \sqrt{n \log \log n} \right) \frac{N_n^{(j)}}{N_n^{(1)}}. \quad (38)$$

For the alternative $j$ that satisfies (38), let

$$u \triangleq \sup \{ l \leq s : I_{n+l}^{(j)} = 1 \}.$$

Then, stage $n + u$ is the last time that alternative $j$ is sampled before or at stage $n + m$. Since $k_s^{(j)}$ is monotonically increasing in $s$, we have

$$\frac{N_n^{(j)} + k_u^{(j)}}{N_n^{(1)} + k_u^{(1)}} > \frac{N_n^{(j)} + k_s^{(j)}}{N_n^{(1)} + k_s^{(1)}} \geq \frac{N_n^{(j)} + k_s^{(j)} - 1}{N_n^{(1)} + k_s^{(1)}} = \left( 1 - \frac{1}{N_n^{(j)} + k_s^{(j)}} \right) \left( N_n^{(j)} + k_s^{(j)} \right) \left( 1 + C_{11} \sqrt{n \log \log n} \right) \frac{N_n^{(j)}}{N_n^{(1)}},$$

where the last line follows from (38). By Lemma 4.2, there must exist a positive constant $C_{12}$ such that, for all large enough $n,$

$$\left( 1 - \frac{1}{N_n^{(j)} + k_s^{(j)}} \right) \left( 1 + C_{11} \sqrt{n \log \log n} \right) \frac{N_n^{(j)}}{N_n^{(1)}}
\[
\begin{align*}
\geq & \left( 1 - \frac{C_{12}}{n} \right) \left( 1 + C_{11} \frac{\sqrt{n \log \log n}}{n} \right) \frac{N_n^{(j)}}{N_n^{(i)}} \\
= & \left( 1 + C_{11} \frac{\sqrt{n \log \log n}}{n} - \frac{C_{12}}{n} - C_{12} C_{11} \frac{\sqrt{n \log \log n}}{n^2} \right) \frac{N_n^{(j)}}{N_n^{(i)}} \\
\geq & \left( 1 + \frac{C_{11}}{2} \frac{\sqrt{n \log \log n}}{n} \right) \frac{N_n^{(j)}}{N_n^{(i)}} \\
= & \left( 1 + C_{13} \frac{\sqrt{n \log \log n}}{n} \right) \frac{N_n^{(j)}}{N_n^{(i)}},
\end{align*}
\]

where \(C_{13} = \frac{C_{11}}{2} = \frac{C_{10} C_{2}}{8}\). Note that constants \(C_3\) through \(C_{10}\) are fixed and do not depend on \(C_1\) or \(C_2\). Thus, for all large enough \(n\), if we take \(C_2\) to be sufficiently large, i.e., \(C_2 \geq 8 C_1 / C_{10}\), to make \(C_{13} \geq C_1\), then

\[
\frac{N_n^{(j)} + k_u^{(j)}}{N_n^{(i)} + k_s^{(j)}} > \left( 1 + C_1 \frac{\sqrt{n \log \log n}}{n} \right) \frac{N_n^{(j)}}{N_n^{(i)}},
\]

which completes the proof.