Rate-optimality of complete expected improvement

Ye Chen Ilya O. Ryzhov

March 22, 2017

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 achieve the optimal asymptotic convergence rate under Gaussian observations. This is the first strong rate-optimality result for any EI-type method.

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 $\mu^{(j)} \in \mathbb{R}$. We wish to identify the best alternative $j^* = \arg\max_j \mu^{(j)}$, which is assumed to be unique. For any j, we have the ability to collect noisy samples of the form $W^{(j)} \sim \mathcal{N}\left(\mu^{(j)}, \left(\lambda^{(j)}\right)^2\right)$, 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).

The *optimal* allocation (with regard to probability of correct selection) for R&S with independent normal samples was characterized by Glynn & Juneja (2004). We state this characterization here, as it is directly relevant to what follows. 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 $\alpha_*^{(j)}$ satisfy two conditions:

- Proportion assigned to alternative j^*:

$$\left(\frac{\alpha_*^{(j^*)}}{\lambda^{(j^*)}}\right)^2 = \sum_{j \neq j^*} \left(\frac{\alpha_*^{(j)}}{\lambda^{(j)}}\right)^2 \tag{1}$$

- Proportions assigned to arbitrary $i, j \neq j^*$:

$$\frac{\left(\mu^{(i)} - \mu^{(j^*)}\right)^2}{\frac{\left(\lambda^{(i)}\right)^2}{\alpha_*^{(i)}} + \frac{\left(\lambda^{(j^*)}\right)^2}{\alpha_*^{(j^*)}}} = \frac{\left(\mu^{(j)} - \mu^{(j^*)}\right)^2}{\frac{\left(\lambda^{(j)}\right)^2}{\alpha_*^{(j)}} + \frac{\left(\lambda^{(j^*)}\right)^2}{\alpha_*^{(j^*)}}} \tag{2}$$

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.

The main contribution of this paper is a new rate-optimality proof for CEI. We show 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 optimality results for any R&S heuristic to date. To compare it with the state of the art, Russo (2017) presents a class of heuristics, based on Thompson sampling, which can also attain the optimal convergence rate, but only when a tuning parameter is set optimally. By contrast, CEI requires no tuning whatsoever. Another recent 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 describe the Bayesian statistical model used by EI-type methods. Let $\mu^{(j)} \sim \mathcal{N}\left(\theta_0^{(j)}, \left(\sigma_0^{(j)}\right)^2\right)$, where $\theta_0^{(j)}$ and $\sigma_0^{(j)}$ are pre-specified prior parameters, and assume $\mu^{(i)}, \mu^{(j)}$ are independent for all $i \neq j$. Let $\{j_n\}_{n=0}^{\infty}$ be a sequence of alternatives chosen for sampling. For each j_n, we observe $W_{n+1}^{(j_n)} \sim \mathcal{N}\left(\mu^{(j_n)}, \left(\lambda^{(j_n)}\right)^2\right)$ where $\lambda^{(j)} > 0$ is assumed to be known for all j. We let $\mathcal{F}_n$ be the sigma-algebra generated by $j_0, W_1^{(j_0)}, ..., j_{n-1}, W_n^{(j_{n-1})}$. This definition allows the sampling decisions to be adapted to the incoming information; under the allocation rules studied in this paper, j_n will in fact be $\mathcal{F}_n$-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 $n = 1, 2, ...$

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 calculated using recursive Bayesian updating. For simplicity, we assume the non-informative prior $\sigma_0^{(j)} = \infty$, whence

$$\theta_n^{(j)} = \frac{1}{N_n^{(j)}} \sum_{m=0}^{n-1} I_m^{(j)} W_{m+1}^{(j)}, \tag{3}$$

$$\left(\sigma_n^{(j)}\right)^2 = \frac{\left(\lambda^{(j)}\right)^2}{N_n^{(j)}}, \tag{4}$$

matching the usual frequentist notions of the sample mean and its variance. Thus, Bayesian arguments are not central to the model, though they will be used to derive the computational forms of the algorithms discussed in this paper.

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 adaptive rule used to allocate the samples. Glynn & Juneja (2004) considers a frequentist setting (in which j^* is fixed) and proves that, if every alternative is assigned a non-zero proportion of the budget asymptotically, the convergence rate of PCS can be expressed in terms of the limit

$$\Gamma^J = -\lim_{n\to\infty} \frac{1}{n} \log P\left(j_n^* \neq j^*\right),$$

where $J = \{j_n\}_{n=0}^{\infty}$ denotes the allocation rule. That is, the probability of *incorrect* selection converges to zero at an exponential rate where the exponent includes a constant Γ^J that depends on the allocation rule J. Equations (1)-(2) characterize the proportions that optimize the rate (maximize Γ^J) under the assumption of independent normal samples.

Since the remainder of the paper will focus on the class of expected improvement methods, we give a brief summary of the EI concept. 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

$$\begin{aligned} v_n^{(j)} &= \mathbb{E}\left(\max\left\{\mu^{(j)} - \theta_n^{(j_n^*)}, 0\right\} \mid \mathcal{F}_n, j_n = j\right) \\ &= \sigma_n^{(j)} f\left(-\frac{\left|\theta_n^{(j)} - \theta_n^{(j_n^*)}\right|}{\sigma_n^{(j)}}\right), \end{aligned} \tag{5}$$

and $f(z) = z\Phi(z) + \phi(z)$ with ϕ, Φ being the standard Gaussian pdf and cdf, respectively. We can view (5) as a measure of the potential that the true value of j will improve upon the current-best estimate $\theta_n^{j_n^*}$. 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

$$\lim_{n\to\infty} \frac{N_n^{(j^*)}}{n} = 1, \tag{6}$$

$$\lim_{n\to\infty} \frac{N_n^{(i)}}{N_n^{(j)}} = \left(\frac{\lambda^{(i)}\left|\mu^{(j)} - \mu^{(j^*)}\right|}{\lambda^{(j)}\left|\mu^{(i)} - \mu^{(j^*)}\right|}\right)^2, \qquad i, j \neq j^*, \tag{7}$$

where the limits hold almost surely. Clearly, (6)-(7) do not match (1)-(2) except in the limiting case where $\alpha_*^{(j^*)} \to 1$. Pasupathy et al. (2014) showed that this case arises when the number of alternatives $M \to \infty$. However, since $\frac{N^{(j)}}{n} \to 0$ 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 (5) with

$$v_n^{(j)} = \mathbb{E}\left(\max\left\{\mu^{(j)} - \mu^{(j_n^*)}, 0\right\} \mid \mathcal{F}_n, j_n = j\right), \tag{8}$$

Let $n = 0$ and repeat the following:

1: Check whether

$$\left(\frac{N_n^{(j_n^*)}}{\lambda^{(j_n^*)}}\right)^2 < \sum_{j \neq j_n^*} \left(\frac{N_n^{(j)}}{\lambda^{(j)}}\right)^2. \tag{10}$$

If (10) holds, assign $j_n = j_n^*$. If (10) does not hold, assign $j_n = \arg\max_{j \neq j_n^*} v_n^{(j)}$, where $v_n^{(j)}$ is given by (9).

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

Figure 1: Modified CEI (mCEI) algorithm for R&S.

which can be written in closed form as

$$v_n^{(j)} = \sqrt{\left(\sigma_n^{(j)}\right)^2 + \left(\sigma_n^{(j_n^*)}\right)^2} f\left(-\frac{\left|\theta_n^{(j)} - \theta_n^{(j_n^*)}\right|}{\sqrt{\left(\sigma_n^{(j)}\right)^2 + \left(\sigma_n^{(j_n^*)}\right)^2}}\right) \tag{9}$$

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 (9), which translates the CEI concept to our R&S model.

From (8), 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, this is a challenging question even outside the EI literature. For example, Russo (2017) observes that the popular Thompson sampling algorithm (Russo & Van Roy, 2014) will sample j_n^* too often, and addresses this problem by essentially assigning a fixed proportion β of samples to j_n^*, while using Thompson sampling to choose between the other alternatives. The optimal rate of convergence can be achieved if β is tuned correctly, but the optimal choice of β is problem-dependent and generally difficult to find.

Based on these considerations, we give a modified CEI procedure in Figure 1. The modification adds condition (10), 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 straightforward to show that mCEI samples every alternative infinitely often as $n \to \infty$, so we use this fact without proof.

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 $n \to \infty$.

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*

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

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

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

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

4 Proofs of main results

Before delving into the proofs, we explain two simplifications used in the analysis. First, we assume in the proofs that the true values $\mu^{(j)}$ are fixed constants, though unknown to the algorithm. This is consistent with all previous theoretical treatments of convergence rates in R&S, such as Glynn & Juneja (2004). Although Bayesian arguments were used to derive the computational form of (9), the resulting algorithm may be implemented and studied in a non-Bayesian setting. Since we prove almost sure convergence (that is, convergence on a fixed sample path), this is not a major issue. We then assume that $j^* = 1$ is the unique optimal alternative.

Second, by the arguments in Section 4.1 of Ryzhov (2016), it is sufficient to prove Theorems 3.1 and 3.2 for a simplified version of mCEI with (9) replaced by

$$v_n^{(j)} = \sqrt{\frac{\left(\lambda^{(j)}\right)^2}{N_n^{(j)}} + \frac{\left(\lambda^{(1)}\right)^2}{N_n^{(1)}}}\, f\left(-\frac{\left|\mu^{(j)} - \mu^{(1)}\right|}{\sqrt{\frac{\left(\lambda^{(j)}\right)^2}{N_n^{(j)}} + \frac{\left(\lambda^{(1)}\right)^2}{N_n^{(1)}}}}\right). \quad (11)$$

and (10) replaced by

$$\left(\frac{N_n^{(1)}}{\lambda^{(1)}}\right)^2 < \sum_{j>1}\left(\frac{N_n^{(j)}}{\lambda^{(j)}}\right)^2. \tag{12}$$

That is, we first apply (4) to the posterior variance, and then replace the posterior means by their limits. As discussed in Ryzhov (2016), the error in the sample mean does not substantially affect the declining behaviour of f, so it suffices to study the convergence rates of the purely deterministic criterion in (11). We also replace j_n^* in both (9) and (10) by its limit $j^* = 1$. This version of mCEI does not depend on the sampled values, and thus we do not have to consider or explicitly denote dependence on a sample path. Aside from this simplification, however, the version of mCEI given by (11)-(12) requires substantial new technical content over Ryzhov (2016), since $v_n^{(j)}$ now explicitly depends on both $N_n^{(j)}$ and $N_n^{(1)}$.

Our results can still be achieved if the noise parameters $\lambda^{(j)}$ are unknown. The standard way of handling this issue (see, e.g., Jones et al., 1998 or Salemi et al., 2014) is to replace $\lambda^{(j)}$ in (4) by the corresponding sample standard deviation, then simply plug these estimates into (9). As the estimates will converge to $\lambda^{(j)}$ a.s., the limiting allocation will not be affected.

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 the deterministic version of mCEI with (11)-(12) replacing (9)-(10).

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,

$$\begin{aligned}
&\Delta_{n+1} - \Delta_n \\
=\ &\left(\frac{\left(N_n^{(1)}+1\right)/\lambda^{(1)}}{n+1}\right)^2 - \sum_{j=2}^{M}\left(\frac{N_n^{(j)}/\lambda^{(j)}}{n+1}\right)^2 - \left(\left(\frac{N_n^{(1)}/\lambda^{(1)}}{n}\right)^2 - \sum_{j=2}^{M}\left(\frac{N_n^{(j)}/\lambda^{(j)}}{n}\right)^2\right)
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{\left(\lambda^{(1)}\right)^2}\left(\left(\left(\frac{\left(N_n^{(1)}+1\right)}{n+1}\right)^2 - \left(\frac{N_n^{(1)}}{n}\right)^2\right) + \left(\sum_{j=2}^{M}\left(\frac{N_n^{(j)}/\lambda^{(j)}}{n}\right)^2 - \sum_{j=2}^{M}\left(\frac{N_n^{(j)}/\lambda^{(j)}}{n+1}\right)^2\right)\right) \\
&> 0.
\end{aligned}
$$

If alternative $j' > 1$ is sampled, then $\Delta_n \geq 0$ and

$$
\begin{aligned}
\Delta_{n+1} - \Delta_n &= \left(\frac{N_n^{(1)}/\lambda^{(1)}}{n+1}\right)^2 - \sum_{j\neq j'}\left(\frac{N_n^{(j)}/\lambda^{(j)}}{n+1}\right)^2 - \left(\frac{(N_n^{(j')}+1)/\lambda^{(j')}}{n+1}\right)^2 \\
&\quad - \left(\left(\frac{N_n^{(1)}/\lambda^{(1)}}{n}\right)^2 - \sum_{j=2}^{M}\left(\frac{N_n^{(j)}/\lambda^{(j)}}{n}\right)^2\right) \\
&= \left(\frac{N_n^{(1)}/\lambda^{(1)}}{n+1}\right)^2 - \sum_{j=2}^{M}\left(\frac{N_n^{(j)}/\lambda^{(j)}}{n+1}\right)^2 - \frac{2N_n^{(j')}+1}{\left(\lambda^{(j')}(n+1)\right)^2} \\
&\quad - \left(\left(\frac{N_n^{(1)}/\lambda^{(1)}}{n}\right)^2 - \sum_{j=2}^{M}\left(\frac{N_n^{(j)}/\lambda^{(j)}}{n}\right)^2\right) \\
&= \left(\frac{n^2}{(n+1)^2} - 1\right)\Delta_n - \frac{2N_n^{(j')}+1}{\left(\lambda^{(j')}(n+1)\right)^2} \\
&< 0,
\end{aligned}
$$

which completes the proof. □

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 $n_1 > \frac{2}{\ell^2\varepsilon} - 1$. 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,

$$
\begin{aligned}
\Delta_{n+1} &= \left(\frac{\left(N_n^{(1)}+1\right)/\lambda^{(1)}}{n+1}\right)^2 - \sum_{j=2}^{M}\left(\frac{N_n^{(j)}/\lambda^{(j)}}{n+1}\right)^2 \\
&= \Delta_n + \frac{2N_n^{(1)}+1}{\left(\lambda^{(1)}(n+1)\right)^2} \\
&< \frac{2n+2}{\left(\lambda^{(1)}(n+1)\right)^2} \\
&\leq \frac{2}{\left(\lambda^{(1)}\right)^2(n_1+1)} \\
&< \frac{\ell^2}{\left(\lambda^{(1)}\right)^2}\varepsilon \\
&\leq \varepsilon.
\end{aligned}
$$

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. Furthermore, using similar arguments as before, we find

$$\begin{aligned}
\Delta_{n+1} &= \left(\frac{N_n^{(1)}/\lambda^{(1)}}{n+1}\right)^2 - \sum_{j=2}^{M}\left(\frac{N_n^{(j)}/\lambda^{(j)}}{n+1}\right)^2 - \frac{2N_n^{(j')}+1}{\left(\lambda^{(j')}(n+1)\right)^2} \\
&= \Delta_n - \frac{2N_n^{(j')}+1}{\left(\lambda^{(j')}(n+1)\right)^2} \\
&\geq -\frac{2n+2}{\left(\lambda^{(j')}(n+1)\right)^2} \\
&\geq -\varepsilon.
\end{aligned}$$

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 $d^{(j)} \triangleq \left|\mu^{(j)} - \mu^{(1)}\right|$ and $\delta^{(j)} = \left(d^{(j)}\right)^2$ for all $j > 1$.

The first technical lemma implies that, for any two alternatives i and j, $N_n^{(i)} = O(N_n^{(j)})$ and $N_n^{(i)} = O(n)$.

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

In the following arguments, we refer to "a large enough n" to indicate that n is chosen to make $N_n^{(j)}$ large enough for some desired condition to hold. For any j and any positive integer m, define

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

as the number of samples allocated to alternative j from stage n to stage $n+m-1$. We introduce the next lemma, which establishes a technical relationship between $k^{(1)}_{(n,n+s)}$ and samples assigned to suboptimal alternatives.

Lemma 4.3. *Let C_1 be any positive constant, and fix $i > 1$. Suppose that i is sampled at at some stage n and define*

$$m \triangleq \inf\left\{l : k^{(i)}_{(n,n+l)} = 2\right\} - 2, \qquad s \triangleq \sup\left\{l < m : I^{(1)}_{n+l} = 0\right\}.$$

Suppose also that, for large enough n, we have

$$C_2 \log n \leq k^{(1)}_{(n,n+s)} \leq (C_2+1)\log n$$

for some large enough positive constant C_2. Then, there must exist a suboptimal alternative $j \neq i$ and a stage $n+u$, where $u \leq s$, such that j is sampled at stage $n+u$ and

$$\left(1 + C_1 \frac{\log n}{n}\right) \frac{N^{(j)}_n}{N^{(1)}_n} < \frac{N^{(j)}_n + k^{(j)}_{(n,n+u)}}{N^{(1)}_n + k^{(1)}_{(n,n+s)}} \leq \frac{N^{(j)}_n + k^{(j)}_{(n,n+u)}}{N^{(1)}_n + k^{(1)}_{(n,n+u)}}.$$

Now let

$$\begin{aligned} z^{(j)}_n &\triangleq \frac{d^{(j)}}{\sqrt{\frac{\left(\lambda^{(j)}\right)^2}{N^{(j)}_n} + \frac{\left(\lambda^{(1)}\right)^2}{N^{(1)}_n}}}, \\ t^{(j)}_n &\triangleq \left(z^{(j)}_n\right)^2 = \frac{\delta^{(j)}}{\frac{\left(\lambda^{(j)}\right)^2}{N^{(j)}_n} + \frac{\left(\lambda^{(1)}\right)^2}{N^{(1)}_n}}. \end{aligned}$$

For any j, both $z^{(j)}_n$ and $t^{(j)}_n$ increase to infinity as $n \to \infty$. We apply an expansion of the Mills ratio (Ruben, 1962) to $v^{(j)}_n$. For large enough n,

$$\begin{aligned} v^{(j)}_n &= \frac{d^{(j)}_n}{z^{(j)}_n} f\left(-z^{(j)}_n\right) \\ &= \frac{d^{(j)}_n}{z^{(j)}_n} \phi\left(z^{(j)}_n\right) \left(-z^{(j)}_n \frac{1 - \Phi\left(z^{(j)}_n\right)}{\phi\left(z^{(j)}_n\right)} + 1\right) \\ &= \frac{d^{(j)}_n}{z^{(j)}_n} \phi\left(z^{(j)}_n\right) \left(-z^{(j)}_n \frac{1}{z^{(j)}_n} \left(1 - \frac{1}{\left(z^{(j)}_n\right)^2} + O\left(\frac{1}{\left(z^{(j)}_n\right)^4}\right)\right) + 1\right) \end{aligned} \tag{13}$$

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

where (13) comes from the Mills ratio. Then,

$$\begin{aligned}
2\log\left(v_n^{(j)}\right) &= 2\log d_n^{(j)} - 6\log z_n^{(j)} + 2\log\phi\left(z_n^{(j)}\right) + 2\log\left(1+O\left(\frac{1}{\left(z_n^{(j)}\right)^2}\right)\right) \\
&= \log\delta^{(j)} - 3\log t_n^{(j)} - \log(2\pi) - t_n^{(j)} + 2\log\left(1+O\left(\frac{1}{t_n^{(j)}}\right)\right) \\
&= -t_n^{(j)}\left(1+O\left(\frac{\log t_n^{(j)}}{t_n^{(j)}}\right)\right).
\end{aligned}$$

For any two suboptimal alternatives i and j, define

$$\begin{aligned}
r_n^{(i,j)} &\triangleq \frac{2\log\left(v_n^{(i)}\right)}{2\log\left(v_n^{(j)}\right)} \\
&= \frac{t_n^{(i)}}{t_n^{(j)}}\frac{1+O\left(\frac{\log t_n^{(i)}}{t_n^{(i)}}\right)}{1+O\left(\frac{\log t_n^{(j)}}{t_n^{(j)}}\right)}, \qquad (14)
\end{aligned}$$

and note that both $1+O\left(\frac{\log t_n^{(i)}}{t_n^{(i)}}\right)$ and $1+O\left(\frac{\log t_n^{(j)}}{t_n^{(j)}}\right)$ converge to 1 as $n\to\infty$. We will show that $r_n^{(i,j)}\to 1$ for any suboptimal i and j; then, (14) will yield $\frac{t_n^{(i)}}{t_n^{(j)}}\to 1$, completing the proof of Theorem 3.2. This requires one final technical lemma, showing that $k_{(n,n+m)}^{(1)}$ is $\mathcal{O}(\log n)$, which is proved (in the Appendix) using Lemma 4.3.

Lemma 4.4. *If $i>1$ is sampled at stage n, then $k_{(n,n+m)}^{(1)} = O(\log n)$ for*

$$m \triangleq \inf\left\{l>0 : I_{n+l}^{(i)} = 1\right\}.$$

Let $i,j>1$ and suppose that i is sampled at stage n. We will first place an $O\left(\frac{1}{n}\right)$ bound on the increment $r_{n+1}^{(i,j)} - r_n^{(i,j)}$. We will then place a bound of $O\left(\frac{\log n}{n}\right)$ on the growth of $\left(r_n^{(i,j)}\right)$ in between two samples of i (note that, by definition, $r_n^{(i,j)}\le 1$ at any stage n when i is sampled). As this bound vanishes to zero as $n\to\infty$, it will follow (as explained later) that $r_n^{(i,j)}\to 1$.

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

$$\begin{aligned}
r_{n+1}^{(i,j)} - r_n^{(i,j)} &= \frac{1}{2\left|\log\left(v_n^{(j)}\right)\right|}\left\{\left(\frac{\delta^{(i)}}{\frac{\left(\lambda^{(i)}\right)^2}{N_n^{(i)}+1}+\frac{\left(\lambda^{(1)}\right)^2}{N_n^{(1)}}} - \frac{\delta^{(i)}}{\frac{\left(\lambda^{(i)}\right)^2}{N_n^{(i)}}+\frac{\left(\lambda^{(1)}\right)^2}{N_n^{(1)}}}\right)\right. \\
&\quad +3\left(\log\frac{\delta^{(i)}}{\frac{\left(\lambda^{(i)}\right)^2}{N_n^{(i)}+1}+\frac{\left(\lambda^{(1)}\right)^2}{N_n^{(1)}}} - \log\frac{\delta^{(i)}}{\frac{\left(\lambda^{(i)}\right)^2}{N_n^{(i)}}+\frac{\left(\lambda^{(1)}\right)^2}{N_n^{(1)}}}\right) \\
&\quad \left.-2\left[\log\left(1+O\left(\frac{\left(\lambda^{(i)}\right)^2}{N_n^{(i)}+1}+\frac{\left(\lambda^{(1)}\right)^2}{N_n^{(1)}}\right)\right) - \log\left(1+O\left(\frac{\left(\lambda^{(i)}\right)^2}{N_n^{(i)}}+\frac{\left(\lambda^{(1)}\right)^2}{N_n^{(1)}}\right)\right)\right]\right\} \qquad (15)
\end{aligned}$$

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

$$\begin{aligned}
2\left|\log\left(v_n^{(j)}\right)\right| &= t_n^{(j)}\left(1+O\left(\frac{\log t_n^{(j)}}{t_n^{(j)}}\right)\right) \\
&> \frac{1}{2}\frac{\delta^{(j)}}{\frac{\left(\lambda^{(j)}\right)^2}{N_n^{(j)}}+\frac{\left(\lambda^{(1)}\right)^2}{N_n^{(1)}}} \\
&\geq C_1 n.
\end{aligned}$$

On the other hand, there also exists a positive constant C_2 such that

$$\frac{\delta^{(i)}}{\frac{\left(\lambda^{(i)}\right)^2}{N_n^{(i)}+1}+\frac{\left(\lambda^{(1)}\right)^2}{N_n^{(1)}}} - \frac{\delta^{(i)}}{\frac{\left(\lambda^{(i)}\right)^2}{N_n^{(i)}}+\frac{\left(\lambda^{(1)}\right)^2}{N_n^{(1)}}} \leq C_2$$

and for large enough n, we have

$$\begin{aligned}
& 3\left(\log\frac{\delta^{(i)}}{\frac{\left(\lambda^{(i)}\right)^2}{N_n^{(i)}+1}+\frac{\left(\lambda^{(1)}\right)^2}{N_n^{(1)}}} - \log\frac{\delta^{(i)}}{\frac{\left(\lambda^{(i)}\right)^2}{N_n^{(i)}}+\frac{\left(\lambda^{(1)}\right)^2}{N_n^{(1)}}}\right) \\
\leq\ & 3\left(\frac{\delta^{(i)}}{\frac{\left(\lambda^{(i)}\right)^2}{N_n^{(i)}+1}+\frac{\left(\lambda^{(1)}\right)^2}{N_n^{(1)}}} - \frac{\delta^{(i)}}{\frac{\left(\lambda^{(i)}\right)^2}{N_n^{(i)}}+\frac{\left(\lambda^{(1)}\right)^2}{N_n^{(1)}}}\right) \\
\leq\ & 3C_2
\end{aligned}$$

and

$$\left|2\left[\log\left(1+O\left(\frac{\left(\lambda^{(i)}\right)^2}{N_n^{(i)}+1}+\frac{\left(\lambda^{(1)}\right)^2}{N_n^{(1)}}\right)\right) - \log\left(1+O\left(\frac{\left(\lambda^{(i)}\right)^2}{N_n^{(i)}}+\frac{\left(\lambda^{(1)}\right)^2}{N_n^{(1)}}\right)\right)\right]\right|$$

$$\leq 2\left[\left|\log\left(1+O\left(\frac{\left(\lambda^{(i)}\right)^2}{N_n^{(i)}+1}+\frac{\left(\lambda^{(1)}\right)^2}{N_n^{(1)}}\right)\right)\right|+\left|\log\left(1+O\left(\frac{\left(\lambda^{(i)}\right)^2}{N_n^{(i)}}+\frac{\left(\lambda^{(1)}\right)^2}{N_n^{(1)}}\right)\right)\right|\right]$$
$$\leq C_2.$$

We have now bounded all three terms in (15). Therefore, for large enough n, we have

$$r_{n+1}^{(i,j)} - r_n^{(i,j)} \leq \frac{5C_2/C_1}{n},$$

and

$$r_{n+1}^{(i,j)} - 1 \leq r_n^{(i,j)} - 1 + \frac{5C_2/C_1}{n} \leq \frac{5C_2/C_1}{n}.$$

Thus, we have established a bound on the growth of $r_n^{(i,j)}$ due to having sampled i at time n.

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

$$m \triangleq \inf\left\{l > 0 : I_{n+l}^{(i)} = 1\right\}$$

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 $r_{n+s+1}^{(i,j)} > r_{n+s}^{(i,j)}$ can only hold if the optimal alternative is sampled at stage $n+s$. When this happens, we have

$$\begin{aligned}
r_{n+s+1}^{(i,j)} - r_{n+s}^{(i,j)} &= \frac{\log\left(v_{n+s+1}^{(i)}\right)}{\log\left(v_{n+s+1}^{(j)}\right)} - \frac{\log\left(v_{n+s}^{(i)}\right)}{\log\left(v_{n+s}^{(j)}\right)} \\
&\leq \frac{\log\left(v_{n+s+1}^{(i)}\right)}{\log\left(v_{n+s}^{(j)}\right)} - \frac{\log\left(v_{n+s}^{(i)}\right)}{\log\left(v_{n+s}^{(j)}\right)} \qquad (16) \\
&= \frac{1}{2\left|\log\left(v_{n+s}^{(j)}\right)\right|}\left\{\left(\frac{\delta^{(i)}}{\frac{\left(\lambda^{(i)}\right)^2}{N_{n+s}^{(i)}}+\frac{\left(\lambda^{(1)}\right)^2}{N_{n+s}^{(1)}+1}} - \frac{\delta^{(i)}}{\frac{\left(\lambda^{(i)}\right)^2}{N_{n+s}^{(i)}}+\frac{\left(\lambda^{(1)}\right)^2}{N_{n+s}^{(1)}}}\right)\right. \\
&\quad +3\left(\log\frac{\delta^{(i)}}{\frac{\left(\lambda^{(i)}\right)^2}{N_{n+s}^{(i)}}+\frac{\left(\lambda^{(1)}\right)^2}{N_{n+s}^{(1)}+1}} - \log\frac{\delta^{(i)}}{\frac{\left(\lambda^{(i)}\right)^2}{N_{n+s}^{(i)}}+\frac{\left(\lambda^{(1)}\right)^2}{N_{n+s}^{(1)}}}\right) \\
&\quad \left.-2\left[\log\left(1+O\left(\frac{\left(\lambda^{(i)}\right)^2}{N_{n+s}^{(i)}}+\frac{\left(\lambda^{(1)}\right)^2}{N_{n+s}^{(1)}+1}\right)\right) - \log\left(1+O\left(\frac{\left(\lambda^{(i)}\right)^2}{N_{n+s}^{(i)}}+\frac{\left(\lambda^{(1)}\right)^2}{N_{n+s}^{(1)}}\right)\right)\right]\right\},
\end{aligned}$$

where (16) holds because $v_n^{(j)}$ is a monotone decreasing function of n. Then, by Lemma 4.2, there exists a positive constant C_3 such that, for large enough n,

$$\begin{aligned}
2\left|\log\left(v_{n+s}^{(j)}\right)\right| &= t_{n+s}^{(j)}\left(1+O\left(\frac{\log t_{n+s}^{(j)}}{t_{n+s}^{(j)}}\right)\right) \\
&> \frac{1}{2}\frac{\delta^{(j)}}{\frac{\left(\lambda^{(j)}\right)^2}{N_{n+s}^{(j)}}+\frac{\left(\lambda^{(1)}\right)^2}{N_{n+s}^{(1)}}} \\
&\geq C_3(n+s) \\
&\geq C_3 n.
\end{aligned}$$

On the other hand, there also exists a positive constant C_4 such that

$$\frac{\delta^{(i)}}{\frac{\left(\lambda^{(i)}\right)^2}{N_{n+s}^{(i)}}+\frac{\left(\lambda^{(1)}\right)^2}{N_{n+s}^{(1)}+1}}-\frac{\delta^{(i)}}{\frac{\left(\lambda^{(i)}\right)^2}{N_{n+s}^{(i)}}+\frac{\left(\lambda^{(1)}\right)^2}{N_{n+s}^{(1)}}}\leq C_4,$$

and, for large enough n, we have

$$\begin{aligned}
& 3\left(\log\frac{\delta^{(i)}}{\frac{\left(\lambda^{(i)}\right)^2}{N_{n+s}^{(i)}}+\frac{\left(\lambda^{(1)}\right)^2}{N_{n+s}^{(1)}+1}}-\log\frac{\delta^{(i)}}{\frac{\left(\lambda^{(i)}\right)^2}{N_{n+s}^{(i)}}+\frac{\left(\lambda^{(1)}\right)^2}{N_{n+s}^{(1)}}}\right) \\
\leq\ & 3\left(\frac{\delta^{(i)}}{\frac{\left(\lambda^{(i)}\right)^2}{N_{n+s}^{(i)}}+\frac{\left(\lambda^{(1)}\right)^2}{N_{n+s}^{(1)}+1}}-\frac{\delta^{(i)}}{\frac{\left(\lambda^{(i)}\right)^2}{N_{n+s}^{(i)}}+\frac{\left(\lambda^{(1)}\right)^2}{N_{n+s}^{(1)}}}\right) \\
\leq\ & 3C_4,
\end{aligned}$$

and

$$\begin{aligned}
& \left|2\left[\log\left(1+O\left(\frac{\left(\lambda^{(i)}\right)^2}{N_{n+s}^{(i)}}+\frac{\left(\lambda^{(1)}\right)^2}{N_{n+s}^{(1)}+1}\right)\right)-\log\left(1+O\left(\frac{\left(\lambda^{(i)}\right)^2}{N_{n+s}^{(i)}}+\frac{\left(\lambda^{(1)}\right)^2}{N_{n+s}^{(1)}}\right)\right)\right]\right| \\
\leq\ & 2\left[\left|\log\left(1+O\left(\frac{\left(\lambda^{(i)}\right)^2}{N_{n+s}^{(i)}}+\frac{\left(\lambda^{(1)}\right)^2}{N_{n+s}^{(1)}+1}\right)\right)\right|+\left|\log\left(1+O\left(\frac{\left(\lambda^{(i)}\right)^2}{N_{n+s}^{(i)}}+\frac{\left(\lambda^{(1)}\right)^2}{N_{n+s}^{(1)}}\right)\right)\right|\right] \\
\leq\ & C_4.
\end{aligned}$$

Therefore, for large enough n, we have

$$r_{n+s+1}^{(i,j)}-r_{n+s}^{(i,j)}\leq\frac{5C_4/C_3}{n},$$

whence

$$r_{n+s+1}^{(i,j)} - 1 \leq r_{n+s}^{(i,j)} - 1 + \frac{5C_4/C_3}{n} \leq r_n^{(i,j)} - 1 + k_s^{(1)} \frac{5C_4/C_3}{n} \leq k_s^{(1)} \frac{5C_4/C_3}{n}.$$

However, from Lemma 4.4, we have $k_s^{(1)} \leq k_m^{(1)} = O(\log n)$ for all $0 < s \leq m$, implying that

$$r_{n+s+1}^{(i,j)} - 1 \leq k_m^{(1)} \frac{5C_4/C_3}{n} = O\left(\frac{\log n}{n}\right),$$

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 rates, or match the optimality conditions only in special limiting cases). The optimality conditions are achieved without 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

Bechhofer, R. (1954), 'A single-sample multiple decision procedure for ranking means of normal populations with known variances', *The Annals of Mathematical Statistics* **25**(1), 16–39.

Branke, J., Chick, S. E. & Schmidt, C. (2007), 'Selecting a selection procedure', *Management Science* **53**(12), 1916–1932.

Chau, M., Fu, M. C., Qu, H. & Ryzhov, I. O. (2014), Simulation optimization: a tutorial overview and recent developments in gradient-based methods, *in* A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley & J. A. Miller, eds, 'Proceedings of the 2014 Winter Simulation Conference', pp. 21–35.

Chen, C.-H., Lin, J., Yücesan, E. & Chick, S. E. (2000), 'Simulation Budget Allocation for Further Enhancing the Efficiency of Ordinal Optimization', *Discrete Event Dynamic Systems* **10**(3), 251–270.

Chick, S. E., Branke, J. & Schmidt, C. (2010), 'Sequential Sampling to Myopically Maximize the Expected Value of Information', *INFORMS Journal on Computing* **22**(1), 71–80.

DeGroot, M. H. (1970), *Optimal Statistical Decisions*, John Wiley and Sons.

Gittins, J. C., Glazebrook, K. D. & Weber, R. (2011), *Multi-armed bandit allocation indices (2nd ed.)*, John Wiley and Sons.

Glynn, P. W. & Juneja, S. (2004), A large deviations perspective on ordinal optimization, *in* R. Ingalls, M. D. Rossetti, J. S. Smith & B. A. Peters, eds, 'Proceedings of the 2004 Winter Simulation Conference', pp. 577–585.

Han, B., Ryzhov, I. O. & Defourny, B. (2016), 'Optimal learning in linear regression with combinatorial feature selection', *INFORMS Journal on Computing* **28**(4), 721–735.

Hong, L. J. & Nelson, B. L. (2009), A Brief Introduction To Optimization Via Simulation, *in* M. Rosetti, R. Hill, B. Johansson, A. Dunkin & R. Ingalls, eds, 'Proceedings of the 2009 Winter Simulation Conference', pp. 75–85.

Jones, D., Schonlau, M. & Welch, W. (1998), 'Efficient global optimization of expensive black-box functions', *Journal of Global Optimization* **13**(4), 455–492.

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

Pasupathy, R., Hunter, S. R., Pujowidianto, N. A., Lee, L. H. & Chen, C.-H. (2014), 'Stochastically constrained ranking and selection via SCORE', *ACM Transactions on Modeling and Computer Simulation* **25**(1), 1:1–1:26.

Peng, Y. & Fu, M. C. (2017), 'Myopic allocation policy with asymptotically optimal sampling rate', *IEEE Transactions on Automatic Control (to appear)*.

Powell, W. B. & Ryzhov, I. O. (2012), *Optimal Learning*, John Wiley and Sons.

Ruben, H. (1962), 'A new asymptotic expansion for the normal probability integral and Mill's ratio', *Journal of the Royal Statistical Society* **B24**(1), 177–179.

Russo, D. (2017), 'Simple Bayesian algorithms for best arm identification', *arXiv preprint arXiv:1602.08448*.

Russo, D. & Van Roy, B. (2014), 'Learning to optimize via posterior sampling', *Mathematics of Operations Research* **39**(4), 1221–1243.

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

Salemi, P., Nelson, B. L. & Staum, J. (2014), Discrete optimization via simulation using Gaussian Markov random fields, *in* A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley & J. A. Miller, eds, 'Proceedings of the 2014 Winter Simulation Conference', pp. 3809–3820.

Scott, W. R., Powell, W. B. & Simão, H. P. (2010), Calibrating simulation models using the knowledge gradient with continuous parameters, *in* B. Johansson, S. Jain, J. Montoya-Torres, J. Hugan & E. Yücesan, eds, 'Proceedings of the 2010 Winter Simulation Conference', pp. 1099–1109.

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 $\limsup_{n\to\infty} \frac{N_n^{(i)}}{N_n^{(j)}} = \infty$. Then, there must exist a stage m such that

$$\frac{N_m^{(i)}}{N_m^{(j)}} > \max\left\{\frac{\delta^{(j)}}{\delta^{(i)}}, 1\right\} \frac{\left(\lambda^{(i)}\right)^2 + \lambda^{(1)}\lambda^{(i)}}{\left(\lambda^{(j)}\right)^2},$$

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,

$$
\begin{aligned}
v_m^{(i)} &= \sqrt{\frac{\left(\lambda^{(i)}\right)^2}{N_m^{(i)}} + \frac{\left(\lambda^{(1)}\right)^2}{N_m^{(1)}}} f\left(-\frac{d^{(i)}}{\sqrt{\frac{\left(\lambda^{(i)}\right)^2}{N_m^{(i)}} + \frac{\left(\lambda^{(1)}\right)^2}{N_m^{(1)}}}}\right) \\
&\leq \sqrt{\frac{\left(\lambda^{(i)}\right)^2}{N_m^{(i)}} + \frac{\lambda^{(1)}\lambda^{(i)}}{N_m^{(i)}}} f\left(-\frac{d^{(i)}}{\sqrt{\frac{\left(\lambda^{(i)}\right)^2}{N_m^{(i)}} + \frac{\lambda^{(1)}\lambda^{(i)}}{N_m^{(i)}}}}\right) \qquad (17) \\
&= \sqrt{\frac{\left(\lambda^{(i)}\right)^2 + \lambda^{(1)}\lambda^{(i)}}{N_m^{(i)}}} f\left(-\frac{d^{(i)}}{\sqrt{\frac{\left(\lambda^{(i)}\right)^2 + \lambda^{(1)}\lambda^{(i)}}{N_m^{(i)}}}}\right) \\
&< \sqrt{\frac{\left(\lambda^{(j)}\right)^2}{N_m^{(j)}}} f\left(-\frac{d^{(j)}}{\sqrt{\frac{\left(\lambda^{(j)}\right)^2}{N_m^{(j)}}}}\right) \\
&< \sqrt{\frac{\left(\lambda^{(j)}\right)^2}{N_m^{(j)}} + \frac{\left(\lambda^{(1)}\right)^2}{N_m^{(1)}}} f\left(-\frac{d^{(j)}}{\sqrt{\frac{\left(\lambda^{(j)}\right)^2}{N_m^{(j)}} + \frac{\left(\lambda^{(1)}\right)^2}{N_m^{(1)}}}}\right) \\
&= v_m^{(j)}, \qquad (18)
\end{aligned}
$$

where (17) holds because a suboptimal alternative is sampled at stage m. From the definition of the mCEI algorithm, (18) implies that we cannot sample i at stage m. We conclude that $\limsup_{n\to\infty} \frac{N_n^{(i)}}{N_n^{(j)}} < \infty$ 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

For convenience, we abbreviate $k^{(j)}_{(n,n+m)}$ by the notation $k^{(j)}_m$ for all j. At stage n, since we sample a suboptimal i by assumption, we must have

$$\left(N_n^{(1)}/\lambda^{(1)}\right)^2 \geq \sum_{j=2}^{M}\left(N_n^{(j)}/\lambda^{(j)}\right)^2. \tag{19}$$

Let $s \leq m$ be a positive integer and suppose that some suboptimal $j \neq i$ is sampled at stage $n+s$. Recall that, by assumption, we must have $k^{(1)}_{(n,n+m)} \geq C_2 \log n$ for some positive constant C_2. Repeating the arguments in the proof of Theorem 3.1, we obtain

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

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

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

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

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

whence

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

and

$$\sum_{j=2,j\neq i}^{M}\left[\left(\frac{\left(N_n^{(j)}+k_s^{(j)}\right)/\lambda^{(j)}}{\left(N_n^{(1)}+k_s^{(1)}\right)/\lambda^{(1)}}\right)^2 - \left(\frac{N_n^{(j)}/\lambda^{(j)}}{N_n^{(1)}/\lambda^{(1)}}\right)^2\right]$$

$$
\begin{aligned}
&\geq \quad 1 - \sum_{j=2, j\neq i}^{M} \left(\frac{N_n^{(j)}/\lambda^{(j)}}{N_n^{(1)}/\lambda^{(1)}} \right)^2 - \left(\frac{\left(N_n^{(i)} + 1\right)/\lambda^{(i)}}{\left(N_n^{(1)} + k_s^{(1)}\right)/\lambda^{(1)}} \right)^2 - \frac{C_4}{n} \\
&\geq \quad \left(\frac{N_n^{(i)}/\lambda^{(i)}}{N_n^{(1)}/\lambda^{(1)}} \right)^2 - \left(\frac{\left(N_n^{(i)} + 1\right)/\lambda^{(i)}}{\left(N_n^{(1)} + k_s^{(1)}\right)/\lambda^{(1)}} \right)^2 - \frac{C_4}{n} \qquad (20) \\
&> \quad \left(\frac{\lambda^{(1)}}{\lambda^{(i)}} \right)^2 \frac{1}{\left(N_n^{(1)} + k_s^{(1)}\right)^2} \left(\frac{N_n^{(i)}}{N_n^{(1)}} \right)^2 \left(2N_n^{(1)} k_s^{(1)} + \left(k_s^{(1)}\right)^2 \right) - \frac{C_4}{n},
\end{aligned}
$$

where (20) holds because of (19). Since $\liminf_{n\to\infty} \frac{N_n^{(i)}}{N_n^{(1)}} > 0$ and $\liminf_{n\to\infty} \frac{N_n^{(1)}}{n} > 0$, there must exist positive constants C_5, C_6, C_7, C_8 and C_9 such that, for large enough n, we have

$$
\begin{aligned}
&\left(\frac{\lambda^{(1)}}{\lambda^{(i)}} \right)^2 \frac{1}{\left(N_n^{(1)} + k_s^{(1)}\right)^2} \left(\frac{N_n^{(i)}}{N_n^{(1)}} \right)^2 \left(2N_n^{(1)} k_s^{(1)} + \left(k_s^{(1)}\right)^2 \right) - \frac{C_4}{n} \\
\geq &\quad C_5 \frac{1}{\left(N_n^{(1)} + k_s^{(1)}\right)^2} \left(2N_n^{(1)} k_s^{(1)} + \left(k_s^{(1)}\right)^2 \right) - \frac{C_4}{N_n^{(1)}} \\
= &\quad \frac{C_5}{\left(N_n^{(1)} + k_s^{(1)}\right)^2} \left(2N_n^{(1)} k_s^{(1)} + \left(k_s^{(1)}\right)^2 - C_6 \frac{\left(N_n^{(1)} + k_s^{(1)}\right)^2}{N_n^{(1)}} \right) \\
\geq &\quad \frac{C_5}{\left(N_n^{(1)} + k_s^{(1)}\right)^2} \left(2N_n^{(1)} k_s^{(1)} + \left(k_s^{(1)}\right)^2 - 2C_7 N_n^{(1)} \right) \qquad (21) \\
\geq &\quad \frac{C_5}{\left(N_n^{(1)} + k_s^{(1)}\right)^2} 2 \left(k_s^{(1)} - C_7 \right) N_n^{(1)} \\
\geq &\quad \frac{C_8 \left(k_s^{(1)} - C_7 \right)}{N_n^{(1)}} \\
\geq &\quad \frac{C_8 \left(C_2 \log n - C_7 \right)}{n} \\
\geq &\quad \frac{C_9 C_2 \log n}{n},
\end{aligned}
$$

where (21) holds since $k_s^{(1)} = O(\log n)$. Then,

$$
\sum_{j=2, j\neq i}^{M} \left[\left(\frac{\left(N_n^{(j)} + k_s^{(j)}\right)/\lambda^{(j)}}{\left(N_n^{(1)} + k_s^{(1)}\right)/\lambda^{(1)}} \right)^2 - \left(\frac{N_n^{(j)}/\lambda^{(j)}}{N_n^{(1)}/\lambda^{(1)}} \right)^2 \right] > \frac{C_9 C_2 \log n}{n},
$$

so there must be some suboptimal j such that

$$\left(\frac{\left(N_n^{(j)}+k_s^{(j)}\right)/\lambda^{(j)}}{\left(N_n^{(1)}+k_s^{(1)}\right)/\lambda^{(1)}}\right)^2-\left(\frac{N_n^{(j)}/\lambda^{(j)}}{N_n^{(1)}/\lambda^{(1)}}\right)^2>\frac{1}{M-2}\frac{C_9C_2\log n}{n}.$$

Let $C_{10}=\frac{C_9}{M-2}$ and $C_{11}=\sqrt{C_{10}C_2}-1$. Then,

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

and, for large enough n, we have

$$\frac{\left(N_n^{(j)}+k_s^{(j)}\right)/\left(N_n^{(1)}+k_s^{(1)}\right)}{N_n^{(j)}/N_n^{(1)}}>1+\frac{C_{11}\log n}{n}, \tag{22}$$

whence

$$\frac{N_n^{(j)}+k_s^{(j)}}{N_n^{(1)}+k_s^{(1)}}>\left(1+C_{11}\frac{\log n}{n}\right)\frac{N_n^{(j)}}{N_n^{(1)}}. \tag{23}$$

For the alternative j that satisfies (23), let

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

Since $k_s^{(j)}$ is monotone increasing in s, we have

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

where the last line follows from (22). By Lemma 4.2, there must exist a positive constant C_{12} such that, for large enough n,

$$\left(1-\frac{1}{N_n^{(j)}+k_s^{(j)}}\right)\left(1+C_{11}\frac{\log n}{n}\right)\frac{N_n^{(j)}}{N_n^{(1)}}$$

$$\geq \left(1-\frac{C_{12}}{n}\right)\left(1+C_{11}\frac{\log n}{n}\right)\frac{N_n^{(j)}}{N_n^{(1)}}$$
$$= \left(1+C_{11}\frac{\log n}{n}-\frac{C_{12}}{n}-C_{12}C_{11}\frac{\log n}{n^2}\right)\frac{N_n^{(j)}}{N_n^{(1)}}$$
$$\geq \left(1+\frac{C_{11}}{2}\frac{\log n}{n}\right)\frac{N_n^{(j)}}{N_n^{(1)}}$$
$$= \left(1+C_{13}\frac{\log n}{n}\right)\frac{N_n^{(j)}}{N_n^{(1)}},$$

where $C_{13} = \frac{C_{11}}{2} = \frac{\sqrt{C_{10}C_2}-1}{2}$. Thus, if we take C_2 to be large enough to make $C_{13} \geq C_1$, then

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

which completes the proof.

6.3 Proof of Lemma 4.4

First, from Theorem 3.1 and Lemma 4.2, it is trivial to see that, between two samples of suboptimal alternatives, 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 the suboptimal alternatives is at most equal to some fixed constant B_2. (Note that this is different from bounding the number of samples of alternative 1 in between two samples of the *same* suboptimal alternative i; that bound is treated in Lemma 4.3).

For convenience, we abbreviate $k_{(n,n+l)}^{(j)}$ 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, $\frac{k_m^{(1)}}{\log n}$ can be arbitrarily large. Since we sample $i > 1$ at stage n, then for any other suboptimal alternative $j \neq i$, we have

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

Then, there must exist positive constants C_1 and C_2 such that, for large enough n,

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

that is, equivalently,

$$\frac{\delta^{(i)}\left(\lambda^{(j)}\right)^2}{N_n^{(j)}}+\frac{\delta^{(i)}\left(\lambda^{(1)}\right)^2}{N_n^{(1)}}\leq\frac{\delta^{(j)}\left(\lambda^{(i)}\right)^2\left(1+C_2\frac{\log n}{n}\right)}{N_n^{(i)}}+\frac{\delta^{(j)}\left(\lambda^{(1)}\right)^2\left(1+C_2\frac{\log n}{n}\right)}{N_n^{(1)}}. \tag{24}$$

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

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

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

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

or, equivalently,

$$\frac{\delta^{(i)}\left(\lambda^{(j)}\right)^2}{N_n^{(j)}+k_u^{(j)}}+\frac{\delta^{(i)}\left(\lambda^{(1)}\right)^2}{N_n^{(1)}+k_u^{(1)}}\leq\frac{\delta^{(j)}\left(\lambda^{(i)}\right)^2\left(1-C_4\frac{\log n}{n}\right)}{N_n^{(i)}+k_u^{(i)}}+\frac{\delta^{(j)}\left(\lambda^{(1)}\right)^2\left(1-C_4\frac{\log n}{n}\right)}{N_n^{(1)}+k_u^{(1)}}. \tag{25}$$

Note that $k_u^{(i)}=1$. Now let n be large enough such that

$$\begin{aligned}\left(\delta^{(j)}-\delta^{(i)}\right)\left(\delta^{(j)}\left(1+C_2\frac{\log n}{n}\right)-\delta^{(i)}\right) &> 0,\\ \left(\delta^{(j)}-\delta^{(i)}\right)\left(\delta^{(j)}\left(1-C_4\frac{\log n}{n}\right)-\delta^{(i)}\right) &> 0.\end{aligned}$$

If $\delta^{(j)}>\delta^{(i)}$, then by (24) we have

$$\begin{aligned}\frac{\delta^{(i)}\left(\lambda^{(j)}\right)^2}{N_n^{(j)}+k_u^{(j)}} &= \frac{\delta^{(i)}\left(\lambda^{(j)}\right)^2}{N_n^{(j)}}\frac{N_n^{(j)}}{N_n^{(j)}+k_u^{(j)}}\\ &\leq \frac{\delta^{(j)}\left(\lambda^{(i)}\right)^2\left(1+C_2\frac{\log n}{n}\right)}{N_n^{(i)}}\frac{N_n^{(j)}}{N_n^{(j)}+k_u^{(j)}}\\ &\quad+\frac{\delta^{(j)}\left(\lambda^{(1)}\right)^2\left(1+C_2\frac{\log n}{n}\right)-\delta^{(i)}\left(\lambda^{(1)}\right)^2}{N_n^{(1)}}\frac{N_n^{(j)}}{N_n^{(j)}+k_u^{(j)}}\\ &= \frac{\delta^{(j)}\left(\lambda^{(i)}\right)^2\left(1-C_4\frac{\log n}{n}\right)}{N_n^{(i)}+1}\frac{\left(1+C_2\frac{\log n}{n}\right)}{\left(1-C_4\frac{\log n}{n}\right)}\frac{N_n^{(i)}+1}{N_n^{(i)}}\frac{N_n^{(j)}}{N_n^{(j)}+k_u^{(j)}}\end{aligned}$$

$$+\frac{\delta^{(j)}\left(\lambda^{(1)}\right)^2\left(1-C_4\frac{\log n}{n}\right)-\delta^{(i)}\left(\lambda^{(1)}\right)^2}{N_n^{(1)}+k_u^{(1)}}\frac{\delta^{(j)}\left(1+C_2\frac{\log n}{n}\right)-\delta^{(i)}}{\delta^{(j)}\left(1-C_4\frac{\log n}{n}\right)-\delta^{(i)}}\frac{N_n^{(1)}+k_u^{(1)}}{N_n^{(1)}}\frac{N_n^{(j)}}{N_n^{(j)}+k_u^{(j)}}.$$

It follows that there must exist a positive constant C_5 such that

$$\begin{aligned}\frac{\delta^{(i)}\left(\lambda^{(j)}\right)^2}{N_n^{(j)}+k_u^{(j)}} \quad\leq\quad & \frac{\delta^{(j)}\left(\lambda^{(i)}\right)^2\left(1-C_4\frac{\log n}{n}\right)}{N_n^{(i)}+1}\left(1+C_5\frac{\log n}{n}\right)\frac{N_n^{(i)}+1}{N_n^{(i)}}\frac{N_n^{(j)}}{N_n^{(j)}+k_u^{(j)}}\\ &+\frac{\delta^{(j)}\left(\lambda^{(1)}\right)^2\left(1-C_4\frac{\log n}{n}\right)-\delta^{(i)}\left(\lambda^{(1)}\right)^2}{N_n^{(1)}+k_u^{(1)}}\left(1+C_5\frac{\log n}{n}\right)\frac{N_n^{(1)}+k_u^{(1)}}{N_n^{(1)}}\frac{N_n^{(j)}}{N_n^{(j)}+k_u^{(j)}}.\end{aligned}$$

Thus, to satisfy (25), it is sufficient to have

$$\left(1+C_5\frac{\log n}{n}\right)\frac{N_n^{(i)}+1}{N_n^{(i)}}\frac{N_n^{(j)}}{N_n^{(j)}+k_u^{(j)}} \quad\leq\quad 1, \tag{26}$$

$$\left(1+C_5\frac{\log n}{n}\right)\frac{N_n^{(1)}+k_u^{(1)}}{N_n^{(1)}}\frac{N_n^{(j)}}{N_n^{(j)}+k_u^{(j)}} \quad\leq\quad 1. \tag{27}$$

By Lemma 4.2, there exists a positive constant C_6 such that, for large enough n,

$$\left(1+C_5\frac{\log n}{n}\right)\frac{N_n^{(i)}+1}{N_n^{(i)}}N_n^{(j)}-N_n^{(j)}\leq C_6\log n.$$

Therefore, to satisfy (26), it is sufficient to have

$$C_6\log n\leq k_u^{(j)}. \tag{28}$$

Now define

$$s\triangleq\sup\left\{l<m:I_{n+l}^{(1)}=0\right\}. \tag{29}$$

Since $\frac{k_m^{(1)}}{\log n}$ can be arbitrarily large, we can suppose that $k_s^{(1)}>C_7\log n$, where C_7 is a positive constant to be specified. By Lemma 4.3, since C_5 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

$$\left(1+C_5\frac{\log n}{n}\right)\frac{N_n^{(j)}}{N_n^{(1)}}<\frac{N_n^{(j)}+k_u^{(j)}}{N_n^{(1)}+k_s^{(1)}}\leq\frac{N_n^{(j)}+k_u^{(j)}}{N_n^{(1)}+k_u^{(1)}}.$$

Then, (27) will hold at stage $n+u$. At the same time, since

$$\frac{N_n^{(j)}+k_u^{(j)}}{N_n^{(1)}+k_s^{(1)}} > \left(1+C_5\frac{\log n}{n}\right)\frac{N_n^{(j)}}{N_n^{(1)}} \geq \frac{N_n^{(j)}}{N_n^{(1)}},$$

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

$$k_u^{(j)} \geq C_9 k_s^{(1)} \geq C_9 C_7 \log n.$$

Now let $C_7 = \max\left\{C_8, \frac{C_6}{C_9}\right\}$. Then, both (27) and (28) are satisfied at stage $n+u$, so (25) is satisfied, which means

$$r_{n+u}^{(i,j)} < 1 \qquad \Rightarrow \qquad v_{n+u}^{(i)} > v_{n+u}^{(j)}.$$

But the alternative j is sampled at stage $n+u$, which means $v_{n+u}^{(i)} \leq v_{n+u}^{(j)}$. The desired contradiction follows.

Now, consider the other case where $\delta^{(j)} < \delta^{(i)}$. By (24), we have

$$\begin{aligned}
\frac{\delta^{(j)}\left(\lambda^{(i)}\right)^2\left(1-C_4\frac{\log n}{n}\right)}{N_n^{(i)}+1} &= \frac{\delta^{(j)}\left(\lambda^{(i)}\right)^2\left(1+C_2\frac{\log n}{n}\right)}{N_n^{(i)}}\frac{1-C_4\frac{\log n}{n}}{1+C_2\frac{\log n}{n}}\frac{N_n^{(i)}}{N_n^{(i)}+1} \\
&\geq \frac{\delta^{(i)}\left(\lambda^{(j)}\right)^2}{N_n^{(j)}}\frac{1-C_4\frac{\log n}{n}}{1+C_2\frac{\log n}{n}}\frac{N_n^{(i)}}{N_n^{(i)}+1} \\
&\quad + \frac{\delta^{(i)}\left(\lambda^{(1)}\right)^2-\delta^{(j)}\left(\lambda^{(1)}\right)^2\left(1+C_2\frac{\log n}{n}\right)}{N_n^{(1)}}\frac{1-C_4\frac{\log n}{n}}{1+C_2\frac{\log n}{n}}\frac{N_n^{(i)}}{N_n^{(i)}+1}.
\end{aligned}$$

Then, there must exist a positive constant C_{10} such that, for large enough n,

$$\begin{aligned}
\frac{\delta^{(j)}\left(\lambda^{(i)}\right)^2\left(1-C_4\frac{\log n}{n}\right)}{N_n^{(i)}+1} &\geq \frac{\delta^{(i)}\left(\lambda^{(j)}\right)^2}{N_n^{(j)}}\frac{1}{1+C_{10}\frac{\log n}{n}}\frac{N_n^{(i)}}{N_n^{(i)}+1} \\
&\quad + \frac{\delta^{(i)}\left(\lambda^{(1)}\right)^2-\delta^{(j)}\left(\lambda^{(1)}\right)^2\left(1+C_2\frac{\log n}{n}\right)}{N_n^{(1)}}\frac{1}{1+C_{10}\frac{\log n}{n}}\frac{N_n^{(i)}}{N_n^{(i)}+1}.
\end{aligned}$$

Thus, to satisfy (25), it is sufficient to have

$$\frac{1}{N_n^{(j)}}\frac{1}{1+C_{10}\frac{\log n}{n}}\frac{N_n^{(i)}}{N_n^{(i)}+1} \quad \geq \quad \frac{1}{N_n^{(j)}+k_u^{(j)}},$$

$$\frac{\delta^{(i)} - \delta^{(j)}\left(1 + C_2 \frac{\log n}{n}\right)}{N_n^{(1)}} \frac{1}{1 + C_{10}\frac{\log n}{n}} \frac{N_n^{(i)}}{N_n^{(i)} + 1} \quad \geq \quad \frac{\delta^{(i)} - \delta^{(j)}\left(1 - C_4 \frac{\log n}{n}\right)}{N_n^{(1)} + k_u^{(1)}},$$

which can equivalently be rewritten as

$$k_u^{(j)} \quad \geq \quad \left(1 + C_{10}\frac{\log n}{n}\right) \frac{N_n^{(i)} + 1}{N_n^{(i)}} N_n^{(j)} - N_n^{(j)}, \tag{30}$$

$$k_u^{(1)} \quad \geq \quad \left(1 + C_{10}\frac{\log n}{n}\right) \frac{\delta^{(i)} - \delta^{(j)}\left(1 - C_4 \frac{\log n}{n}\right)}{\delta^{(i)} - \delta^{(j)}\left(1 + C_2 \frac{\log n}{n}\right)} \frac{N_n^{(i)} + 1}{N_n^{(i)}} N_n^{(1)} - N_n^{(1)}. \tag{31}$$

By Lemma 4.2, there exist positive constants C_{11}, C_{12}, C_{13} and C_{14} such that, for large enough n,

$$\left(1 + C_{10}\frac{\log n}{n}\right) \frac{N_n^{(i)} + 1}{N_n^{(i)}} N_n^{(j)} - N_n^{(j)} \quad \leq \quad C_{11} \log n,$$

and

$$\begin{aligned} & \left(1 + C_{10}\frac{\log n}{n}\right) \frac{\delta^{(i)} - \delta^{(j)}\left(1 - C_4 \frac{\log n}{n}\right)}{\delta^{(i)} - \delta^{(j)}\left(1 + C_2 \frac{\log n}{n}\right)} \frac{N_n^{(i)} + 1}{N_n^{(i)}} N_n^{(1)} - N_n^{(1)} \\ \leq \quad & \left(1 + C_{10}\frac{\log n}{n}\right)\left(1 + C_{12}\frac{\log n}{n}\right) \frac{N_n^{(i)} + 1}{N_n^{(i)}} N_n^{(1)} - N_n^{(1)} \\ \leq \quad & \left(1 + C_{13}\frac{\log n}{n}\right) \frac{N_n^{(i)} + 1}{N_n^{(i)}} N_n^{(1)} - N_n^{(1)} \\ \leq \quad & C_{14} \log n. \end{aligned}$$

Therefore, to satisfy (30) and (31), it is sufficient to have

$$k_u^{(j)} \quad \geq \quad C_{11} \log n, \tag{32}$$

$$k_u^{(1)} \quad \geq \quad C_{14} \log n. \tag{33}$$

Again, define s as in (29). Since $\frac{k_m^{(1)}}{\log n}$ can be arbitrarily large, we can suppose that $k_s^{(1)} > C_{15} \log n$, where C_{15} is a positive constant to be specified. By Lemma 4.3, 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

$$\left(1 + C_{11}\frac{\log n}{n}\right) \frac{N_n^{(j)}}{N_n^{(1)}} < \frac{N_n^{(j)} + k_u^{(j)}}{N_n^{(1)} + k_s^{(1)}} \leq \frac{N_n^{(j)} + k_u^{(j)}}{N_n^{(1)} + k_u^{(1)}},$$

whence

$$\frac{N_n^{(j)} + k_u^{(j)}}{N_n^{(1)} + k_s^{(1)}} > \left(1 + C_5 \frac{\log n}{n}\right) \frac{N_n^{(j)}}{N_n^{(1)}} \geq \frac{N_n^{(j)}}{N_n^{(1)}}.$$

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

$$k_u^{(j)} \geq C_{17} k_s^{(1)} \geq C_{17} C_{15} \log n.$$

At the same time, for large enough n, we also have

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

Now, let $C_{15} = \max\left\{C_{16}, \frac{C_{11}}{C_{17}}, \frac{2B_2 C_{14}}{C_{17}}\right\}$. Then both (32) and (33) are satisfied at stage $n+u$, so (25) is satisfied, which means that

$$r_{n+u}^{(i,j)} < 1 \qquad \Rightarrow \qquad v_{n+u}^{(i)} > v_{n+u}^{(j)}.$$

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.