Adv. Appl. Prob. **51,** 209–235 (2019)
*doi:*10.1017/apr.2019.9

COMPLETE EXPECTED IMPROVEMENT CONVERGES TO AN OPTIMAL BUDGET ALLOCATION

YE CHEN,* *Virginia Commonwealth University*
ILYA O. RYZHOV,** *University of Maryland*

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.

Keywords: Optimal learning; ranking and selection; expected improvement; large deviations rate

2010 Mathematics Subject Classification: Primary 93E35
Secondary 68T05

1. Introduction

In the ranking and selection (R&S) problem, there are M 'alternatives' (or 'systems'), and each alternative $j \in \{1, \ldots, M\}$ 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 $W^{(j)} \sim \mathcal{N}\big(\mu^{(j)}, (\lambda^{(j)})^2\big)$, 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.

The R&S problem has a long history dating back to [1], and continues to be an active area of research; see [3] and [11]. 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 [14], optimal computing budget allocation [4], and expected improvement [13]. The related literature on multiarmed bandits [8] has contributed other approaches, such as Thompson

Received 22 March 2017; revision received 23 March 2018.
* Postal address: Statistical Sciences and Operations Research, Virginia Commonwealth University, Richmond, VA 23284, USA.
** Postal address: Robert H. Smith School of Business, University of Maryland, College Park, MD 20742, USA. Email address: iryzhov@rhsmith.umd.edu

sampling [21], although the bandit problem uses a different objective function from the R&S problem and, thus, a good method for one problem may work poorly in the other [20].

Glynn and Juneja [9] gave a rigorous foundation for the notion of *optimal budget allocation* with regard to probability of correct selection. Denote by $0 \le N^{(j)} \le N$ the number of samples assigned to alternative j (thus, $\sum_j N^{(j)} = N$), and take $N \to \infty$ while keeping the proportion $\alpha^{(j)} = N^{(j)}/N$ constant. The optimal proportions $\alpha_*^{(j)}$ (among all possible vectors $\alpha \in \mathbb{R}^M_{++}$ satisfying $\sum_j \alpha^{(j)} = 1$) satisfy the following conditions.

- Proportion assigned to alternative j^*:

$$\left(\frac{\alpha_*^{(j^*)}}{\lambda^{(j^*)}}\right)^2 = \sum_{j \neq j^*} \left(\frac{\alpha_*^{(j)}}{\lambda^{(j)}}\right)^2. \tag{1.1}$$

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

$$\frac{(\mu^{(i)} - \mu^{(j^*)})^2}{(\lambda^{(i)})^2/\alpha_*^{(i)} + (\lambda^{(j^*)})^2/\alpha_*^{(j^*)}} = \frac{(\mu^{(j)} - \mu^{(j^*)})^2}{(\lambda^{(j)})^2/\alpha_*^{(j)} + (\lambda^{(j^*)})^2/\alpha_*^{(j^*)}}. \tag{1.2}$$

Under this allocation, the probability of incorrect selection will converge to 0 at the fastest possible rate (exponential with the best possible exponent). Of course, (1.1) and (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.* [15] used large deviations theory to derive optimality conditions, analogous to (1.1) and (1.2), for a general class of simulation-based optimization problems, but advocated 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 [2], [10], [24] ever since their introduction in [13]. Expected improvement is a Bayesian approach to the R&S problem 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 [13], the knowledge gradient criterion [17], or the LL_1 criterion of [6]. Ryzhov [22] 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.* [23] proposed a new criterion called *complete expected improvement* (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.* [23] created and implemented this idea in the context of Gaussian Markov random fields, a more sophisticated Bayesian learning model than the version of the R&S problem 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 (1.1) and (1.2) is available for statistical models with a Gaussian Markov structure. In this 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 it was originally laid out in [23], this modified version of CEI achieves both (1.1) and (1.2) 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 problem heuristic to date. To compare it with state-of-the-art research, Russo [20] presented 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.* [18], 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. From this work, it can be seen that top-two methods are structured similarly to our approach, but essentially deflect the calculation of $\alpha_*^{(j^*)}$ to the decision maker, whereas we develop a simple, adaptive scheme that learns this quantity with no tuning whatsoever. Also related is the work by Peng and Fu [16], who found a way to reverse-engineer the EI calculations to optimize the rate, but this approach requires one to first solve (1.1) and (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 convergence rate theory and the more practical concerns that motivate EI methods.

2. Preliminaries

We first provide some formal background for the optimality conditions (1.1)–(1.2) derived in [9], and then give an overview of EI-type methods. It is important to note that the theoretical framework of [9], as well as the theoretical analysis developed in this paper, relies on a frequentist interpretation of the R&S problem, 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 $i = 1, \ldots, M$. Let $\{j_n\}_{n=0}^{\infty}$ be a sequence of alternatives chosen for sampling. For each j_n, we observe $W_{n+1}^{(j_n)} \sim \mathcal{N}\big(\mu^{(j_n)}, \big(\lambda^{(j_n)}\big)^2\big)$, where $\lambda^{(j)} > 0$ is assumed to be known for all j. We let $\mathcal{F}_n$ be the σ-algebra generated by $j_0, W_1^{(j_0)}, \ldots, j_{n-1}, W_n^{(j_{n-1})}$. 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 the j_n are $\mathcal{F}_0$-measurable. We define $I_n^{(j)} = \mathbf{1}_{\{j_n = j\}}$, 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 $n = 1, 2, \ldots$.

At time n, we can calculate the statistics

$$\theta_n^{(j)} = \frac{1}{N_n^{(j)}} \sum_{m=0}^{n-1} I_m^{(j)} W_{m+1}^{(j)}, \tag{2.1}$$

$$(\sigma_n^{(j)})^2 = \frac{(\lambda^{(j)})^2}{N_n^{(j)}}. \tag{2.2}$$

https://doi.org/10.1017/apr.2019.9 Published online by Cambridge University Press

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 $\mathbb{P}\left(j_n^* = j^*\right)$, depends on the rule used to allocate the samples. Glynn and Juneja [9] proved that, for any static allocation that assigns a proportion $\alpha^{(j)} > 0$ of the budget to each alternative j, the convergence rate of the PCS can be expressed in terms of the limit

$$\Gamma^{\alpha} = - \lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(j_n^* \neq j^*\right). \tag{2.3}$$

That is, the probability of *incorrect* selection converges to 0 at an exponential rate, where the exponent includes a constant Γ^{α} that depends on the vector α of proportions. Equations (1.1) and (1.2) characterize the proportions that optimize the rate (maximize Γ^{α}) under the assumption of independent normal samples. Although Glynn and Juneja [9] only considered static allocations, nonetheless, to date, (2.3) continues to be one of the strongest rate results for the R&S problem. Optimal static allocations derived through this framework can be used as guidance for the design of dynamic allocations; see, for example, [12] and [15].

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 $\mu^{(j)} \sim \mathcal{N}\left(\theta_0^{(j)}, (\sigma_0^{(j)})^2\right)$, where $\theta_0^{(j)}$ and $\sigma_0^{(j)}$ are prespecified prior parameters. It is also assumed that $\mu^{(i)}$ and $\mu^{(j)}$ are independent for all $i \neq j$. Under these assumptions, it is well known [7] that the posterior distribution of $\mu^{(j)}$ given $\mathcal{F}_n$ is $\mathcal{N}\left(\theta_n^{(j)}, (\sigma_n^{(j)})^2\right)$, where the posterior mean and variance can be computed recursively. Under the noninformative prior $\sigma_0^{(j)} = \infty$, the Bayesian posterior parameters $\theta_n^{(j)}$ and $\sigma_n^{(j)}$ are identical to the frequentist statistics defined in (2.1) and (2.2), 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.* [13]. In this version of EI, as applied to our R&S model, we take $j_n = \arg\max_j v_n^{(j)}$, where

$$v_n^{(j)} = \mathbb{E}\left(\max\left\{\mu^{(j)} - \theta_n^{(j_n^*)}, 0\right\} \middle| \mathcal{F}_n\right) = \sigma_n^{(j)} f\left(- \frac{\left|\theta_n^{(j)} - \theta_n^{(j_n^*)}\right|}{\sigma_n^{(j)}}\right), \tag{2.4}$$

and $f(z) = z\Phi(z) + \phi(z)$, with ϕ and Φ being the standard Gaussian PDF and CDF, respectively. We can view (2.4) 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 [22] gave the first convergence rate analysis of this algorithm. Under EI, we have

$$\lim_{n\to\infty} \frac{N_n^{(j^*)}}{n} = 1, \tag{2.5}$$

$$\lim_{n\to\infty} \frac{N_n^{(i)}}{N_n^{(j)}} = \left(\frac{\lambda^{(i)} \left|\mu^{(j)} - \mu^{(j^*)}\right|}{\lambda^{(j)} \left|\mu^{(i)} - \mu^{(j^*)}\right|} \right)^2, \qquad i, j \neq j^*, \tag{2.6}$$

where the limits hold almost surely. Clearly, (2.5) and (2.6) do not match (1.1) and (1.2) except in the limiting case where $\alpha_*^{(j^*)} \to 1$. Because $N^{(j)}/n \to 0$ for $j \neq j^*$, EI will not achieve an exponential convergence rate for any finite M. Ryzhov [22] also derived the limiting allocations for two other variants of EI, but they do not recover (1.1) and (1.2) either.

3. Algorithm and main results

Salemi *et al.* [23] proposed to replace (2.4) with

$$v_n^{(j)} = \mathbb{E}\big(\max\{\mu^{(j)} - \mu^{(j_n^*)}, 0\} \mid \mathcal{F}_n\big), \tag{3.1}$$

which can be written in closed form as

$$v_n^{(j)} = \sqrt{\big(\sigma_n^{(j)}\big)^2 + \big(\sigma_n^{(j_n^*)}\big)^2}\, f\left(-\frac{\big|\theta_n^{(j)} - \theta_n^{(j_n^*)}\big|}{\sqrt{\big(\sigma_n^{(j)}\big)^2 + \big(\sigma_n^{(j_n^*)}\big)^2}}\right) \tag{3.2}$$

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.* [23] considered 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 (3.2), which translates the CEI concept to our R&S model.

From (3.1), 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 [23]. In fact, many adaptive methods are unable to efficiently identify when j_n^* should be measured; thus, both the classic EI method of [13], and the popular Thompson sampling algorithm [21], will sample j_n^* too often. The class of top-two methods, first introduced in [20], addresses this problem by essentially assigning a fixed proportion β of samples to j_n^*, while using Thompson sampling or other means to choose between the other alternatives. Optimal allocations can be attained 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 (mCEI) procedure in Algorithm 3.1. The modification adds condition (3.3), which mimics (1.1) 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 was shown in [5] that the mCEI algorithm samples every alternative infinitely often as $n \to \infty$.

Algorithm 3.1. (*mCEI algorithm for the R&S problem.*) Let $n = 0$ and repeat the following.

- Check whether

$$\left(\frac{N_n^{(j_n^*)}}{\lambda^{(j_n^*)}}\right)^2 < \sum_{j \neq j_n^*}\left(\frac{N_n^{(j)}}{\lambda^{(j)}}\right)^2. \tag{3.3}$$

  If (3.3) holds, assign $j_n = j_n^*$. If (3.3) does not hold, assign $j_n = \arg\max_{j \neq j_n^*} v_n^{(j)}$, where $v_n^{(j)}$ is given by (3.2).

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

We now state our main results on the asymptotic rate optimality of mCEI. Essentially, these theorems state that conditions (1.1) and (1.2) 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* $\alpha_n^{(j)} = N_n^{(j)}/n$. *Under the mCEI algorithm,*

$$\lim_{n\to\infty} \left(\frac{\alpha_n^{(j^*)}}{\lambda^{(j^*)}}\right)^2 - \sum_{j\neq j_n^*} \left(\frac{\alpha_n^{(j)}}{\lambda^{(j)}}\right)^2 = 0 \quad \textit{almost surely.}$$

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

$$\tau_n^{(j)} = \frac{(\mu^{(j)} - \mu^{(j^*)})^2}{(\lambda^{(j)})^2/\alpha_n^{(j)} + (\lambda^{(j^*)})^2/\alpha_n^{(j^*)}},$$

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

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

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

4. Proofs of the main results

For notational convenience, we assume that $j^* = 1$ is the unique optimal alternative. Since, under the mCEI algorithm, $N_n^{(j)} \to \infty$ 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 the mCEI algorithm with (3.2) replaced by

$$v_n^{(j)} = \sqrt{\frac{(\lambda^{(j)})^2}{N_n^{(j)}} + \frac{(\lambda^{(1)})^2}{N_n^{(1)}}} f\left(-\frac{|\theta_n^{(j)} - \theta_n^{(1)}|}{\sqrt{(\lambda^{(j)})^2/N_n^{(j)} + (\lambda^{(1)})^2/N_n^{(1)}}}\right), \tag{4.1}$$

and (3.3) replaced by

$$\left(\frac{N_n^{(1)}}{\lambda^{(1)}}\right)^2 < \sum_{j>1} \left(\frac{N_n^{(j)}}{\lambda^{(j)}}\right)^2. \tag{4.2}$$

To simplify the presentation of the key arguments, we treat the noise parameters $\lambda^{(j)}$ as being known. If in (2.2) we replace $\lambda^{(j)}$ by the sample standard deviation (as recommended, e.g. in both [13] and [23]) then simply plug the resulting approximation into (3.2), the limiting allocation will not be affected. Because the rate-optimality framework of Glynn and Juneja [9] is frequentist and assumes that selection is based only on sample means, it does not make any distinction between known and unknown variances in terms of characterizing an optimal allocation.

4.1. Proof of Theorem 3.1

First, we define the quantity

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

https://doi.org/10.1017/apr.2019.9 Published online by Cambridge University Press

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 mCEI algorithm with (4.1) and (4.2) replacing (3.2) and (3.3).

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

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

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

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

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 $n_1 > 2/\ell^2\varepsilon - 1$. 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,

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

https://doi.org/10.1017/apr.2019.9 Published online by Cambridge University Press

$$\begin{aligned}
&< \frac{2n+2}{(\lambda^{(1)}(n+1))^2} \\
&\leq \frac{2}{(\lambda^{(1)})^2(n_1+1)} \\
&< \frac{\ell^2}{(\lambda^{(1)})^2}\varepsilon \\
&\leq \varepsilon.
\end{aligned}$$

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. Using similar arguments as before, we find that

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

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 that $\lim_{n\to\infty} \Delta_n = 0$ 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 the mCEI algorithm 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 that $-\varepsilon < \Delta_{n_2} < 0$. 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 Appendix A. For notational convenience, we define $d_n^{(j)} := \left|\theta_n^{(j)} - \theta_n^{(1)}\right|$ and $\delta_n^{(j)} = (d_n^{(j)})^2$ for all $j > 1$. Furthermore, for any j and any positive integer m, we define

$$k_{(n,n+m)}^{(j)} := N_{n+m}^{(j)} - N_n^{(j)}$$

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, $N_n^{(i)} = \Theta(N_n^{(j)})$ and $N_n^{(i)} = \Theta(n)$.

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

Now let

$$z_n^{(j)} := \frac{d_n^{(j)}}{\sqrt{(\lambda^{(j)})^2/N_n^{(j)} + (\lambda^{(1)})^2/N_n^{(1)}}}, \qquad t_n^{(j)} := (z_n^{(j)})^2 = \frac{\delta_n^{(j)}}{(\lambda^{(j)})^2/N_n^{(j)} + (\lambda^{(1)})^2/N_n^{(1)}}.$$

For any j, both $z_n^{(j)}$ and $t_n^{(j)}$ go to ∞ as $n \to \infty$. We apply an expansion of the Mills ratio [19] to $v_n^{(j)}$. For all large enough n,

$$\begin{aligned} v_n^{(j)} &= \frac{d_n^{(j)}}{z_n^{(j)}} f(-z_n^{(j)}) \\ &= \frac{d_n^{(j)}}{z_n^{(j)}} \phi(z_n^{(j)}) \biggl(-z_n^{(j)} \frac{1-\Phi(z_n^{(j)})}{\phi(z_n^{(j)})} + 1 \biggr) \\ &= \frac{d_n^{(j)}}{z_n^{(j)}} \phi(z_n^{(j)}) \biggl(-z_n^{(j)} \frac{1}{z_n^{(j)}} \biggl(1 - \frac{1}{(z_n^{(j)})^2} + O\biggl(\frac{1}{(z_n^{(j)})^4} \biggr) \biggr) + 1 \biggr) \qquad (4.3) \\ &= \frac{d_n^{(j)}}{(z_n^{(j)})^3} \phi(z_n^{(j)}) \biggl(1 + O\biggl(\frac{1}{(z_n^{(j)})^2} \biggr) \biggr), \end{aligned}$$

where (4.3) comes from the Mills ratio. Then

$$\begin{aligned} 2\log\bigl(v_n^{(j)}\bigr) &= 2\log d_n^{(j)} - 6\log z_n^{(j)} + 2\log\phi\bigl(z_n^{(j)}\bigr) + 2\log\biggl(1 + O\biggl(\frac{1}{(z_n^{(j)})^2}\biggr)\biggr) \\ &= \log\delta_n^{(j)} - 3\log t_n^{(j)} - \log(2\pi) - t_n^{(j)} + 2\log\biggl(1 + O\biggl(\frac{1}{t_n^{(j)}}\biggr)\biggr) \\ &= -t_n^{(j)}\biggl(1 + O\biggl(\frac{\log t_n^{(j)}}{t_n^{(j)}}\biggr)\biggr). \end{aligned}$$

For any two suboptimal alternatives i and j, define

$$r_n^{(i,j)} := \frac{2\log(v_n^{(i)})}{2\log(v_n^{(j)})} = \frac{t_n^{(i)}}{t_n^{(j)}} \frac{1 + O(\log t_n^{(i)}/t_n^{(i)})}{1 + O(\log t_n^{(j)}/t_n^{(j)})}, \qquad (4.4)$$

and note that both $1 + O(\log t_n^{(i)}/t_n^{(i)})$ and $1 + O(\log t_n^{(j)}/t_n^{(j)})$ converge to 1 as $n \to \infty$. We will show that $r_n^{(i,j)} \to 1$ for any suboptimal i and j; then (4.4) will yield $t_n^{(i)}/t_n^{(j)} \to 1$, 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 Appendix A. 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 $O(1)$ and vice versa; next, Lemma 4.4 shows that $k_{(n,n+m)}^{(1)}$ is $O\bigl(\sqrt{n\log\log n}\bigr)$; finally, Lemma 4.5 bounds $n^{3/4}\bigl|\delta_{n+1}^{(i)} - \delta_n^{(i)}\bigr|$.

Lemma 4.3. *Between two samples assigned to any suboptimal alternatives (i.e. two time stages when condition (4.2) 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.*

https://doi.org/10.1017/apr.2019.9 Published online by Cambridge University Press

Lemma 4.4. *On almost every sample path, there exists a fixed positive constant $C < \infty$ and time $n_0 \geq 3$ such that, for any $n \geq n_0$ in which some suboptimal alternative i is sampled and*

$$m := \inf\{l > 0 \colon I^{(i)}_{n+l} = 1\},$$

we have $k^{(1)}_{(n,n+m)} \leq C\sqrt{n \log\log n}$.

Lemma 4.5. *For any alternative i, $n^{3/4}\big|\delta^{(i)}_{n+1} - \delta^{(i)}_n\big| \to 0$ almost surely as $n \to \infty$.*

Let $i, j > 1$, and suppose that i is sampled at stage n. We will first place an $O(1/n^{3/4})$ bound on the increment $r^{(i,j)}_{n+1} - r^{(i,j)}_n$. We will then place a bound of $O(\sqrt{n \log\log n}/n^{3/4})$ on the growth of $(r^{(i,j)}_n)$ in between two samples of i (note that, by definition, $r^{(i,j)}_n \leq 1$ at any stage n when i is sampled). As this bound vanishes to 0 as $n \to \infty$, it will then be shown to follow that $r^{(i,j)}_n \to 1$.

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

$$\begin{aligned}
r^{(i,j)}_{n+1} - r^{(i,j)}_n &= \frac{\log\big(v^{(i)}_{n+1}\big)}{\log\big(v^{(j)}_{n+1}\big)} - \frac{\log\big(v^{(i)}_n\big)}{\log\big(v^{(j)}_n\big)} \\
&= \frac{\log\big(v^{(i)}_{n+1}\big) - \log\big(v^{(i)}_n\big)}{\log\big(v^{(j)}_n\big)} \\
&\leq \frac{\big|\log\big(v^{(i)}_{n+1}\big) - \log\big(v^{(i)}_n\big)\big|}{\big|\log\big(v^{(j)}_n\big)\big|} \\
&= \frac{n^{1/4}}{2\big|\log\big(v^{(j)}_n\big)\big|}\frac{1}{n^{1/4}}\left|\left(\frac{\delta^{(i)}_{n+1}}{\frac{(\lambda^{(i)})^2}{N^{(i)}_n+1} + \frac{(\lambda^{(1)})^2}{N^{(1)}_n}} - \frac{\delta^{(i)}_n}{\frac{(\lambda^{(i)})^2}{N^{(i)}_n} + \frac{(\lambda^{(1)})^2}{N^{(1)}_n}}\right)\right. \\
&\qquad + 3\left(\log\frac{\delta^{(i)}_{n+1}}{\frac{(\lambda^{(i)})^2}{N^{(i)}_n+1} + \frac{(\lambda^{(1)})^2}{N^{(1)}_n}} - \log\frac{\delta^{(i)}_n}{\frac{(\lambda^{(i)})^2}{N^{(i)}_n} + \frac{(\lambda^{(1)})^2}{N^{(1)}_n}}\right) \\
&\qquad - 2\left[\log\left(1 + O\left(\frac{(\lambda^{(i)})^2}{N^{(i)}_n+1} + \frac{(\lambda^{(1)})^2}{N^{(1)}_n}\right)\right)\right. \\
&\qquad \left. - \log\left(1 + O\left(\frac{(\lambda^{(i)})^2}{N^{(i)}_n} + \frac{(\lambda^{(1)})^2}{N^{(1)}_n}\right)\right)\right] \\
&\qquad \left. - \big(\log\delta^{(i)}_{n+1} - \log\delta^{(i)}_n\big)\right|.
\end{aligned} \tag{4.5}$$

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

$$\begin{aligned}
\frac{2|\log(v^{(j)}_n)|}{n^{1/4}} &= \frac{t^{(j)}_n}{n^{1/4}}\left(1 + O\left(\frac{\log t^{(j)}_n}{t^{(j)}_n}\right)\right) \\
&> \frac{1}{2n^{1/4}}\frac{\delta^{(j)}_n}{(\lambda^{(j)})^2/N^{(j)}_n + (\lambda^{(1)})^2/N^{(1)}_n}
\end{aligned}$$

https://doi.org/10.1017/apr.2019.9 Published online by Cambridge University Press

$$\geq C_1 \frac{n}{n^{1/4}}$$
$$= C_1 n^{3/4}.$$

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

$$\begin{aligned}
&\frac{1}{n^{1/4}}\left|\frac{\delta_{n+1}^{(i)}}{(\lambda^{(i)})^2/(N_n^{(i)}+1)+(\lambda^{(1)})^2/N_n^{(1)}}-\frac{\delta_n^{(i)}}{(\lambda^{(i)})^2/N_n^{(i)}+(\lambda^{(1)})^2/N_n^{(1)}}\right|\\
&\leq \frac{1}{n^{1/4}}\left(\left|\frac{\delta_{n+1}^{(i)}}{(\lambda^{(i)})^2/(N_n^{(i)}+1)+(\lambda^{(1)})^2/N_n^{(1)}}-\frac{\delta_{n+1}^{(i)}}{(\lambda^{(i)})^2/N_n^{(i)}+(\lambda^{(1)})^2/N_n^{(1)}}\right|\right.\\
&\quad\left.+\left|\frac{\delta_{n+1}^{(i)}}{(\lambda^{(i)})^2/N_n^{(i)}+(\lambda^{(1)})^2/N_n^{(1)}}-\frac{\delta_n^{(i)}}{(\lambda^{(i)})^2/N_n^{(i)}+(\lambda^{(1)})^2/N_n^{(1)}}\right|\right)\\
&= \frac{1}{n^{1/4}}\left(O(1)+O(n)\left|\delta_{n+1}^{(i)}-\delta_n^{(i)}\right|\right)\\
&= O(1)\left(\frac{1}{n^{1/4}}+n^{3/4}\left|\delta_{n+1}^{(i)}-\delta_n^{(i)}\right|\right)\\
&= O\,(1)\\
&\leq C_2,
\end{aligned}$$

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

$$\begin{aligned}
&\frac{3}{n^{1/4}}\left|\log\frac{\delta_{n+1}^{(i)}}{\left(\lambda^{(i)}\right)^2/(N_n^{(i)}+1)+\left(\lambda^{(1)}\right)^2/N_n^{(1)}}-\log\frac{\delta_n^{(i)}}{\left(\lambda^{(i)}\right)^2/N_n^{(i)}+\left(\lambda^{(1)}\right)^2/N_n^{(1)}}\right|\\
&\leq \frac{3}{n^{1/4}}\left|\frac{\delta_{n+1}^{(i)}}{\left(\lambda^{(i)}\right)^2/(N_n^{(i)}+1)+\left(\lambda^{(1)}\right)^2/N_n^{(1)}}-\frac{\delta_n^{(i)}}{\left(\lambda^{(i)}\right)^2/N_n^{(i)}+\left(\lambda^{(1)}\right)^2/N_n^{(1)}}\right|\\
&\leq 3C_2,
\end{aligned}$$

and

$$\begin{aligned}
&\left|\frac{2}{n^{1/4}}\left[\log\left(1+O\left(\frac{(\lambda^{(i)})^2}{N_n^{(i)}+1}+\frac{(\lambda^{(1)})^2}{N_n^{(1)}}\right)\right)-\log\left(1+O\left(\frac{(\lambda^{(i)})^2}{N_n^{(i)}}+\frac{(\lambda^{(1)})^2}{N_n^{(1)}}\right)\right)\right]\right|\\
&\quad+\frac{1}{n^{1/4}}\left|\log\delta_{n+1}^{(i)}-\log\delta_n^{(i)}\right|\\
&\leq \frac{2}{n^{1/4}}\left[\left|\log\left(1+O\left(\frac{(\lambda^{(i)})^2}{N_n^{(i)}+1}+\frac{(\lambda^{(1)})^2}{N_n^{(1)}}\right)\right)\right|+\left|\log\left(1+O\left(\frac{(\lambda^{(i)})^2}{N_n^{(i)}}+\frac{(\lambda^{(1)})^2}{N_n^{(1)}}\right)\right)\right|\right]\\
&\quad+\frac{1}{n^{1/4}}\left|\log\delta_{n+1}^{(i)}-\log\delta_n^{(i)}\right|\\
&\leq C_2.
\end{aligned}$$

We now have bounded all four terms in (4.5). Therefore, for all large enough n, we have

$$r_{n+1}^{(i,j)}-r_n^{(i,j)}\leq\frac{5C_2/C_1}{n^{3/4}},$$

and

$$r_{n+1}^{(i,j)} - 1 \leq r_n^{(i,j)} - 1 + \frac{5C_2/C_1}{n^{3/4}} \leq \frac{5C_2/C_1}{n^{3/4}}.$$

Thus, we have established a bound on the growth of $r_n^{(i,j)}$ 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

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

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 alternative j or the optimal alternative is sampled at stage $n+s$.

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

$$\begin{aligned}
r_{n+s+1}^{(i,j)} - r_{n+s}^{(i,j)} &= \frac{\log\left(v_{n+s+1}^{(i)}\right)}{\log\left(v_{n+s+1}^{(j)}\right)} - \frac{\log\left(v_{n+s}^{(i)}\right)}{\log\left(v_{n+s}^{(j)}\right)} \\
&= \frac{\log\left(v_{n+s}^{(i)}\right)}{\log\left(v_{n+s+1}^{(j)}\right)} - \frac{\log\left(v_{n+s}^{(i)}\right)}{\log\left(v_{n+s}^{(j)}\right)} \\
&\leq \left|\frac{\log\left(v_{n+s}^{(i)}\right)}{\log\left(v_{n+s+1}^{(j)}\right)}\right| \frac{\left|\log\left(v_{n+s+1}^{(j)}\right) - \log\left(v_{n+s}^{(j)}\right)\right|}{\left|\log\left(v_{n+s}^{(j)}\right)\right|}.
\end{aligned}$$

Using similar arguments as above, we have

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

and, by Lemma 4.2,

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

Thus, there exists a constant C_3 such that

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

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

$$\begin{aligned}
r_{n+s+1}^{(i,j)} - r_{n+s}^{(i,j)} &= \frac{\log\left(v_{n+s+1}^{(i)}\right)}{\log\left(v_{n+s+1}^{(j)}\right)} - \frac{\log\left(v_{n+s}^{(i)}\right)}{\log\left(v_{n+s}^{(j)}\right)} \\
&\leq \left|\frac{\log\left(v_{n+s+1}^{(i)}\right)}{\log\left(v_{n+s+1}^{(j)}\right)} - \frac{\log\left(v_{n+s}^{(i)}\right)}{\log\left(v_{n+s+1}^{(j)}\right)}\right| + \left|\frac{\log\left(v_{n+s}^{(i)}\right)}{\log\left(v_{n+s+1}^{(j)}\right)} - \frac{\log\left(v_{n+s}^{(i)}\right)}{\log\left(v_{n+s}^{(j)}\right)}\right| \\
&= \left|\frac{\log(v_{n+s+1}^{(i)}) - \log(v_{n+s}^{(i)})}{\log(v_{n+s+1}^{(j)})}\right| + \left|\frac{\log(v_{n+s}^{(i)})}{\log(v_{n+s+1}^{(j)})}\right| \frac{\left|\log(v_{n+s+1}^{(j)}) - \log(v_{n+s}^{(j)})\right|}{\left|\log(v_{n+s}^{(j)})\right|}.
\end{aligned}$$

https://doi.org/10.1017/apr.2019.9 Published online by Cambridge University Press

Similarly as above, we have

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

Then there exists a constant C_4 such that

$$r_{n+s+1}^{(i,j)}-r_{n+s}^{(i,j)} \le C_4 n^{-3/4}.$$

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

$$r_{n+s+1}^{(i,j)}-r_{n+s}^{(i,j)} \le \frac{C_5}{n^{3/4}},$$

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

$$\begin{aligned} r_{n+s+1}^{(i,j)}-1 &\le r_{n+s}^{(i,j)}-1+\frac{C_5}{n^{3/4}} \\ &\le r_n^{(i,j)}-1+\left(1+k_s^{(j)}+k_s^{(1)}\right)\frac{C_5}{n^{3/4}} \\ &\le \left(1+k_s^{(j)}+k_s^{(1)}\right)\frac{C_5}{n^{3/4}}. \end{aligned}$$

However, from Lemma 4.4, we have $k_s^{(1)} \le k_m^{(1)} = O\left(\sqrt{n\log\log n}\right)$ for all $0<s\le m$, 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 $k_s^{(j)} \le k_m^{(j)} \le B_2\big(k_m^{(1)}+1\big)$, whence $k_m^{(j)} = O\left(\sqrt{n\log\log n}\right)$. It follows that

$$r_{n+s+1}^{(i,j)}-1 \le \left(1+k_m^{(j)}+k_m^{(1)}\right)\frac{C_5}{n^{3/4}} = O\bigg(\frac{\sqrt{n\log\log n}}{n^{3/4}}\bigg),$$

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 [23]) achieves the rate-optimality conditions of [9] asymptotically. While convergence to the rate-optimal static allocation need not preserve the convergence rate of that allocation, nonetheless, the static framework of [9] is widely used by the simulation community as a guide for the development of dynamic algorithms, and from this viewpoint it is noteworthy that simple and efficient algorithms can learn optimal static allocations without tuning or approximations.

https://doi.org/10.1017/apr.2019.9 Published online by Cambridge University Press

This 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 [23]. 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.

Appendix A. Additional proofs

Below, we give the full proofs of some technical lemmas that were stated in the main text.

A.1. Proof of Lemma 4.2

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

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

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

$$\begin{aligned}
v_m^{(i)} &= \sqrt{\frac{(\lambda^{(i)})^2}{N_m^{(i)}} + \frac{(\lambda^{(1)})^2}{N_m^{(1)}}}\, f\left(-\frac{d_m^{(i)}}{\sqrt{(\lambda^{(i)})^2/N_m^{(i)} + (\lambda^{(1)})^2/N_m^{(1)}}}\right) \\
&\le \sqrt{\frac{\left(\lambda^{(i)}\right)^2}{N_m^{(i)}} + \frac{\lambda^{(1)}\lambda^{(i)}}{N_m^{(i)}}}\, f\left(-\frac{d_m^{(i)}}{\sqrt{(\lambda^{(i)})^2/N_m^{(i)} + \lambda^{(1)}\lambda^{(i)}/N_m^{(i)}}}\right) && \text{(A.1)} \\
&= \sqrt{\frac{\left(\lambda^{(i)}\right)^2 + \lambda^{(1)}\lambda^{(i)}}{N_m^{(i)}}}\, f\left(-\frac{d_m^{(i)}}{\sqrt{((\lambda^{(i)})^2 + \lambda^{(1)}\lambda^{(i)})/N_m^{(i)}}}\right) \\
&< \sqrt{\frac{\left(\lambda^{(j)}\right)^2}{N_m^{(j)}}}\, f\left(-\frac{d_m^{(j)}}{\sqrt{(\lambda^{(j)})^2/N_m^{(j)}}}\right) && \text{(A.2)} \\
&< \sqrt{\frac{\left(\lambda^{(j)}\right)^2}{N_m^{(j)}} + \frac{\left(\lambda^{(1)}\right)^2}{N_m^{(1)}}}\, f\left(-\frac{d_m^{(j)}}{\sqrt{(\lambda^{(j)})^2/N_m^{(j)} + (\lambda^{(1)})^2/N_m^{(1)}}}\right) \\
&= v_m^{(j)}, && \text{(A.3)}
\end{aligned}$$

where (A.1) holds because a suboptimal alternative is sampled at stage m, and (A.2) holds because $\lim_{m\to\infty} d_m^{(j)}/d_m^{(i)} = |(\mu^{(j)} - \mu^{(1)})|/|(\mu^{(i)} - \mu^{(1)})|$. From the definition of the mCEI algorithm, (A.3) implies that we cannot sample i at stage m. We conclude that $\limsup_{n\to\infty} N_n^{(i)}/N_n^{(j)} < \infty$ for any two suboptimal alternatives i and j.

From this result, we can see that, for $i, j > 1$, we have

$$0 < \liminf_{n\to\infty} \frac{N_n^{(i)}}{N_n^{(j)}} \le \limsup_{n\to\infty} \frac{N_n^{(i)}}{N_n^{(j)}} < \infty.$$

Together with Theorem 3.1, this implies that, for any $i > 1$, we have

$$0 < \liminf_{n\to\infty} \frac{N_n^{(i)}}{N_n^{(1)}} \le \limsup_{n\to\infty} \frac{N_n^{(i)}}{N_n^{(1)}} < \infty,$$

completing the proof.

A.2. Proof of Lemma 4.3

Define $Q_n := (N_n^{(1)}/\lambda^{(1)})^2 - \sum_{j=2}^{M} (N_n^{(j)}/\lambda^{(j)})^2$. Suppose that, at some stage n, $Q_n < 0$ and $Q_{n+1} \ge 0$, which means that the optimal alternative is sampled at time n and then a suboptimal alternative is sampled at time $n+1$. Let $m := \inf\{l > 0 \colon Q_{n+l} < 0\}$, 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

$$\begin{aligned}
Q_{n+1} &= \left(\frac{N_n^{(1)}+1}{\lambda^{(1)}}\right)^2 - \sum_{j=2}^{M}\left(\frac{N_n^{(j)}}{\lambda^{(j)}}\right)^2 \\
&= \left(\frac{N_n^{(1)}}{\lambda^{(1)}}\right)^2 - \sum_{j=2}^{M}\left(\frac{N_n^{(j)}}{\lambda^{(j)}}\right)^2 + \frac{2N_n^{(1)}+1}{(\lambda^{(1)})^2} \\
&= Q_n + \frac{2N_n^{(1)}+1}{(\lambda^{(1)})^2} \\
&< \frac{2N_n^{(1)}+1}{(\lambda^{(1)})^2} \\
&\le C_1 N_n^{(1)},
\end{aligned} \tag{A.4}$$

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

$$\begin{aligned}
Q_{n+s} &= \left(\frac{N_{n+s}^{(1)}}{\lambda^{(1)}}\right)^2 - \sum_{j=2}^{M}\left(\frac{N_{n+s}^{(j)}}{\lambda^{(j)}}\right)^2 \\
&= \left(\frac{N_{n+1}^{(1)}}{\lambda^{(1)}}\right)^2 - \sum_{j=2}^{M}\left(\frac{N_{n+s}^{(j)}}{\lambda^{(j)}}\right)^2 \\
&= \left(\frac{N_n^{(1)}+1}{\lambda^{(1)}}\right)^2 - \sum_{j=2}^{M}\left(\frac{N_n^{(j)}}{\lambda^{(j)}}\right)^2 - \left(\sum_{j=2}^{M}\left(\frac{N_{n+s}^{(j)}}{\lambda^{(j)}}\right)^2 - \sum_{j=2}^{M}\left(\frac{N_n^{(j)}}{\lambda^{(j)}}\right)^2\right) \\
&< C_1 N_n^{(1)} - \left(\sum_{j=2}^{M}\left(\frac{N_{n+s}^{(j)}}{\lambda^{(j)}}\right)^2 - \sum_{j=2}^{M}\left(\frac{N_n^{(j)}}{\lambda^{(j)}}\right)^2\right),
\end{aligned}$$

where the inequality holds because of (A.4). We can also see that, after stage n, the increment of $\sum_{j=2}^{M} (N_n^{(j)}/\lambda^{(j)})^2$ obtained by allocating a sample to alternative j is at least $2N_n^{(j)}/(\lambda^{(j)})^2$.

https://doi.org/10.1017/apr.2019.9 Published online by Cambridge University Press

Then, for all large enough n,

$$\sum_{j=2}^{M}\left(\frac{N_{n+s}^{(j)}}{\lambda^{(j)}}\right)^2-\sum_{j=2}^{M}\left(\frac{N_{n}^{(j)}}{\lambda^{(j)}}\right)^2 \quad\geq\quad 2s\frac{\min_{\{j>1\}} N_n^{(j)}}{\max_{\{j>1\}}(\lambda^{(j)})^2}\geq C_2 s N_n^{(1)},$$

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 that $m=O(1)$ 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.

A.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 Appendix A.5.

Lemma A.1. *Fix a sample path, and let C_1 be any positive constant. For a time stage n at which some suboptimal alternative $i>1$ is sampled, define*

$$m:=\inf\left\{l>0\colon I_{n+l}^{(i)}=1\right\},\qquad s:=\sup\left\{l<m\colon I_{n+l}^{(1)}=0\right\}.$$

Suppose that there are infinitely many time stages n in which the condition

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

holds, where C_2 is a positive constant whose value does not depend on these values of n (but may depend on C_1). Then, for such large enough n, there exists some 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

$$\left(1+C_1\frac{\sqrt{n\log\log n}}{n}\right)\frac{N_n^{(j)}}{N_n^{(1)}}<\frac{N_n^{(j)}+k_{(n,n+u)}^{(j)}}{N_n^{(1)}+k_{(n,n+s)}^{(1)}}\leq\frac{N_n^{(j)}+k_{(n,n+u)}^{(j)}}{N_n^{(1)}+k_{(n,n+u)}^{(1)}}, \tag{A.5}$$

holds.

Essentially, Lemma A.1 will be used to prove the desired result by contradiction; we will show that (A.5) cannot arise, and, therefore, $k_{(n,n+m)}^{(j)}$ must be $O(\sqrt{n\log\log n})$.

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

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

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

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

https://doi.org/10.1017/apr.2019.9 Published online by Cambridge University Press

that is, equivalently,

$$\frac{\delta_n^{(i)}(\lambda^{(j)})^2}{N_n^{(j)}} + \frac{\delta_n^{(i)}(\lambda^{(1)})^2}{N_n^{(1)}} \leq \frac{\delta_n^{(j)}(\lambda^{(i)})^2(1+C_2(\log n)/n)}{N_n^{(i)}} + \frac{\delta_n^{(j)}(\lambda^{(1)})^2(1+C_2(\log n)/n)}{N_n^{(1)}}. \tag{A.6}$$

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,

$$\begin{aligned} r_{n+u}^{(i,j)} &= \frac{t_{n+u}^{(i)}}{t_{n+u}^{(j)}} \frac{1+O(\log t_{n+u}^{(i)}/t_{n+u}^{(i)})}{1+O(\log t_{n+u}^{(j)}/t_{n+u}^{(j)})} \\ &\leq \frac{t_{n+u}^{(i)}}{t_{n+u}^{(j)}} \frac{1}{1-C_3(\log t_{n+u}^{(i)}/t_{n+u}^{(i)} + \log t_{n+u}^{(j)}/t_{n+u}^{(j)})} \\ &< \frac{t_{n+u}^{(i)}}{t_{n+u}^{(j)}} \frac{1}{1-C_4(\log n)/n}. \end{aligned}$$

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

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

or, equivalently,

$$\begin{aligned} &\frac{\delta_{n+u}^{(i)}(\lambda^{(j)})^2}{N_n^{(j)}+k_u^{(j)}} + \frac{\delta_{n+u}^{(i)}(\lambda^{(1)})^2}{N_n^{(1)}+k_u^{(1)}} \\ &\qquad \leq \frac{\delta_{n+u}^{(j)}(\lambda^{(i)})^2(1-C_4(\log n)/n)}{N_n^{(i)}+k_u^{(i)}} + \frac{\delta_{n+u}^{(j)}(\lambda^{(1)})^2(1-C_4(\log n)/n)}{N_n^{(1)}+k_u^{(1)}}. \end{aligned} \tag{A.7}$$

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

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

If $\lim_{n\to\infty} \frac{\delta_n^{(j)}}{\delta_n^{(i)}} > 1$, i.e. $\mu^{(j)} < \mu^{(i)}$, then by (A.6), we have

$$\begin{aligned} \frac{\delta_{n+u}^{(i)}(\lambda^{(j)})^2}{N_n^{(j)}+k_u^{(j)}} &= \frac{\delta_{n+u}^{(i)}}{\delta_n^{(i)}} \frac{\delta_n^{(i)}(\lambda^{(j)})^2}{N_n^{(j)}} \frac{N_n^{(j)}}{N_n^{(j)}+k_u^{(j)}} \\ &\leq \frac{\delta_{n+u}^{(i)}}{\delta_n^{(i)}} \frac{\delta_n^{(j)}(\lambda^{(i)})^2(1+C_2(\log n)/n)}{N_n^{(i)}} \frac{N_n^{(j)}}{N_n^{(j)}+k_u^{(j)}} \\ &\quad + \frac{\delta_{n+u}^{(i)}}{\delta_n^{(i)}} \frac{\delta_n^{(j)}(\lambda^{(1)})^2(1+C_2(\log n)/n) - \delta_n^{(i)}(\lambda^{(1)})^2}{N_n^{(1)}} \frac{N_n^{(j)}}{N_n^{(j)}+k_u^{(j)}} \end{aligned}$$

https://doi.org/10.1017/apr.2019.9 Published online by Cambridge University Press

$$= \frac{\delta_{n+u}^{(i)}}{\delta_n^{(i)}} \frac{\delta_n^{(j)}(\lambda^{(i)})^2(1-C_4(\log n)/n)}{N_n^{(i)}+1} \frac{(1+C_2(\log n)/n)}{(1-C_4(\log n)/n)} \frac{N_n^{(i)}+1}{N_n^{(i)}} \frac{N_n^{(j)}}{N_n^{(j)}+k_u^{(j)}}$$
$$+ \frac{\delta_{n+u}^{(i)}}{\delta_n^{(i)}} \frac{\delta_n^{(j)}(\lambda^{(1)})^2(1+C_2(\log n)/n) - \delta_n^{(i)}(\lambda^{(1)})^2}{N_n^{(1)}+k_u^{(1)}} \frac{N_n^{(1)}+k_u^{(1)}}{N_n^{(1)}} \frac{N_n^{(j)}}{N_n^{(j)}+k_u^{(j)}}.$$

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

$$\frac{\delta_{n+u}^{(i)}(\lambda^{(j)})^2}{N_n^{(j)}+k_u^{(j)}} \le \frac{\delta_{n+u}^{(i)}}{\delta_n^{(i)}} \frac{\delta_n^{(j)}(\lambda^{(i)})^2(1-C_4(\log n)/n)}{N_n^{(i)}+1} \left(1+C_5\frac{\log n}{n}\right) \frac{N_n^{(i)}+1}{N_n^{(i)}} \frac{N_n^{(j)}}{N_n^{(j)}+k_u^{(j)}}$$
$$+ \frac{\delta_{n+u}^{(i)}}{\delta_n^{(i)}} \frac{\delta_n^{(j)}(\lambda^{(1)})^2(1+C_2(\log n)/n) - \delta_n^{(i)}(\lambda^{(1)})^2}{N_n^{(1)}+k_u^{(1)}} \frac{N_n^{(1)}+k_u^{(1)}}{N_n^{(1)}} \frac{N_n^{(j)}}{N_n^{(j)}+k_u^{(j)}}.$$

Thus, to satisfy (A.7), it is sufficient to have

$$\frac{\delta_{n+u}^{(i)}}{\delta_n^{(i)}} \frac{\delta_n^{(j)}}{\delta_{n+u}^{(j)}} \left(1+C_5\frac{\log n}{n}\right) \frac{N_n^{(i)}+1}{N_n^{(i)}} \frac{N_n^{(j)}}{N_n^{(j)}+k_u^{(j)}} \le 1, \tag{A.8}$$
$$\frac{\delta_{n+u}^{(i)}}{\delta_n^{(i)}} \frac{\delta_n^{(j)}(1+C_2(\log n)/n) - \delta_n^{(i)}}{\delta_{n+u}^{(j)}(1-C_4(\log n)/n) - \delta_{n+u}^{(i)}} \frac{N_n^{(1)}+k_u^{(1)}}{N_n^{(1)}} \frac{N_n^{(j)}}{N_n^{(j)}+k_u^{(j)}} \le 1. \tag{A.9}$$

Note that, for all large enough n and any alternative $i \ne 1$, by Lemma 4.2, we have

$$\begin{aligned}
|\delta_{n+u}^{(i)} - \delta_n^{(i)}| &= \big|(d_{n+u}^{(i)})^2 - (d_n^{(i)})^2\big| \\
&= \big|(\theta_{n+u}^{(i)} - \theta_{n+u}^{(1)})^2 - (\theta_n^{(i)} - \theta_n^{(1)})^2\big| \\
&= \big|(\theta_{n+u}^{(i)} - \theta_{n+u}^{(1)}) + (\theta_n^{(i)} - \theta_n^{(1)})\big|\big|(\theta_{n+u}^{(i)} - \theta_n^{(i)}) - (\theta_{n+u}^{(1)} - \theta_n^{(1)})\big| \\
&\le \big|(\theta_{n+u}^{(i)} - \theta_{n+u}^{(1)}) + (\theta_n^{(i)} - \theta_n^{(1)})\big| \\
&\quad \cdot \big(|\theta_{n+u}^{(i)} - \mu^{(i)}| + |\theta_n^{(i)} - \mu^{(i)}| + |\theta_{n+u}^{(1)} - \mu^{(1)}| + |\theta_n^{(1)} - \mu^{(1)}|\big) \\
&= O\left(\sqrt{\frac{\log\log N_n^{(i)}}{N_n^{(i)}}}\right) + O\left(\sqrt{\frac{\log\log N_n^{(1)}}{N_n^{(1)}}}\right) \\
&= O\left(\sqrt{\frac{\log\log n}{n}}\right),
\end{aligned}$$

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

$$\frac{\delta_{n+u}^{(i)}}{\delta_n^{(i)}} = \frac{\delta_{n+u}^{(i)} - \delta_n^{(i)} + \delta_n^{(i)}}{\delta_n^{(i)}} = 1 + \frac{\delta_{n+u}^{(i)} - \delta_n^{(i)}}{\delta_n^{(i)}} = 1 + O\left(\sqrt{\frac{\log\log n}{n}}\right),$$
$$\frac{\delta_n^{(j)}}{\delta_{n+u}^{(j)}} = 1 + O\left(\sqrt{\frac{\log\log n}{n}}\right),$$

https://doi.org/10.1017/apr.2019.9 Published online by Cambridge University Press

and

$$\frac{\delta_n^{(j)}(1+C_2(\log n)/n)-\delta_n^{(i)}}{\delta_{n+u}^{(j)}(1-C_4(\log n)/n)-\delta_{n+u}^{(i)}}$$
$$=1+\frac{\delta_n^{(j)}(1+C_2(\log n)/n)-\delta_{n+u}^{(j)}(1-C_4(\log n)/n)-(\delta_n^{(i)}-\delta_{n+u}^{(i)})}{\delta_{n+u}^{(j)}(1-C_4(\log n)/n)-\delta_{n+u}^{(i)}}$$
$$\leq 1+\frac{|\delta_n^{(j)}-\delta_{n+u}^{(j)}|+|\delta_n^{(j)}C_2(\log n)/n|+|\delta_{n+u}^{(j)}C_4(\log n)/n|+|\delta_n^{(i)}-\delta_{n+u}^{(i)}|}{\delta_{n+u}^{(j)}(1-C_4(\log n)/n)-\delta_{n+u}^{(i)}}$$
$$=1+O\Big(\sqrt{\frac{\log\log n}{n}}\Big).$$

Then together with Lemma 4.2, there exists a positive constant C_6 such that, for all large enough n, the left-hand side of (A.8) satisfies

$$\frac{\delta_{n+u}^{(i)}}{\delta_n^{(i)}}\frac{\delta_n^{(j)}}{\delta_{n+u}^{(j)}}\Big(1+C_5\frac{\log n}{n}\Big)\frac{N_n^{(i)}+1}{N_n^{(i)}}N_n^{(j)}-N_n^{(j)}$$
$$=\Big(1+O\Big(\sqrt{\frac{\log\log n}{n}}\Big)\Big)\Big(1+C_5\frac{\log n}{n}\Big)\Big(1+O\Big(\frac{1}{n}\Big)\Big)N_n^{(j)}-N_n^{(j)}$$
$$\leq C_6\sqrt{n\log\log n},$$

while the left-hand side of (A.9) satisfies

$$\frac{\delta_{n+u}^{(i)}}{\delta_n^{(i)}}\frac{\delta_n^{(j)}\,(1+C_2(\log n)/n)-\delta_n^{(i)}}{\delta_{n+u}^{(j)}\,(1-C_4(\log n)/n)-\delta_{n+u}^{(i)}}\frac{N_n^{(1)}+k_u^{(1)}}{N_n^{(1)}}\frac{N_n^{(j)}}{N_n^{(j)}+k_u^{(j)}}$$
$$=\Big(1+O\Big(\sqrt{\frac{\log\log n}{n}}\Big)\Big)\frac{N_n^{(1)}+k_u^{(1)}}{N_n^{(1)}}\frac{N_n^{(j)}}{N_n^{(j)}+k_u^{(j)}}$$
$$\leq\Big(1+C_6\frac{\sqrt{n\log\log n}}{n}\Big)\frac{N_n^{(1)}+k_u^{(1)}}{N_n^{(1)}}\frac{N_n^{(j)}}{N_n^{(j)}+k_u^{(j)}}.$$

Therefore, to satisfy (A.8), it is sufficient to have

$$C_6\sqrt{n\log\log n}\leq k_u^{(j)}. \tag{A.10}$$

Now define

$$s:=\sup\big\{l<m\colon I_{n+l}^{(1)}=0\big\}. \tag{A.11}$$

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

$$\Big(1+C_6\frac{\sqrt{n\log\log n}}{n}\Big)\frac{N_n^{(j)}}{N_n^{(1)}}<\frac{N_n^{(j)}+k_u^{(j)}}{N_n^{(1)}+k_s^{(1)}}\leq\frac{N_n^{(j)}+k_u^{(j)}}{N_n^{(1)}+k_u^{(1)}}.$$

https://doi.org/10.1017/apr.2019.9 Published online by Cambridge University Press

Then, (A.9) holds at stage $n+u$. At the same time, since

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

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

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

Now let $C_7 = \max\{C_8, C_6/C_9\}$. Then, both (A.9) and (A.10) are satisfied at stage $n+u$, so (A.7) is satisfied, which means that

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

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 $\lim_{n\to\infty}\delta_n^{(j)}/\delta_n^{(i)} < 1$, i.e. $\mu^{(j)} > \mu^{(i)}$. By (A.6), we have

$$\begin{aligned}
&\frac{\delta_{n+u}^{(j)}(\lambda^{(i)})^2(1-C_4(\log n)/n)}{N_n^{(i)}+1}\\
&\quad= \frac{\delta_{n+u}^{(j)}}{\delta_n^{(j)}}\frac{\delta_n^{(j)}(\lambda^{(i)})^2(1+C_2(\log n)/n)}{N_n^{(i)}}\frac{1-C_4(\log n)/n}{1+C_2(\log n)/n}\frac{N_n^{(i)}}{N_n^{(i)}+1}\\
&\quad\geq \frac{\delta_{n+u}^{(j)}}{\delta_n^{(j)}}\frac{\delta_n^{(i)}(\lambda^{(j)})^2}{N_n^{(j)}}\frac{1-C_4(\log n)/n}{1+C_2(\log n)/n}\frac{N_n^{(i)}}{N_n^{(i)}+1}\\
&\qquad+\frac{\delta_{n+u}^{(j)}}{\delta_n^{(j)}}\frac{\delta_n^{(i)}(\lambda^{(1)})^2-\delta_n^{(j)}(\lambda^{(1)})^2(1+C_2(\log n)/n)}{N_n^{(1)}}\frac{1-C_4(\log n)/n}{1+C_2(\log n)/n}\frac{N_n^{(i)}}{N_n^{(i)}+1}.
\end{aligned}$$

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

$$\begin{aligned}
&\frac{\delta_{n+u}^{(j)}(\lambda^{(i)})^2\,(1-C_4(\log n)/n)}{N_n^{(i)}+1}\\
&\quad\geq \frac{\delta_{n+u}^{(j)}}{\delta_n^{(j)}}\frac{\delta_n^{(i)}(\lambda^{(j)})^2}{N_n^{(j)}}\frac{1}{1+C_{10}(\log n)/n}\frac{N_n^{(i)}}{N_n^{(i)}+1}\\
&\qquad+\frac{\delta_{n+u}^{(j)}}{\delta_n^{(j)}}\frac{\delta_n^{(i)}(\lambda^{(1)})^2-\delta_n^{(j)}(\lambda^{(1)})^2(1+C_2(\log n)/n)}{N_n^{(1)}}\frac{1}{1+C_{10}(\log n)/n}\frac{N_n^{(i)}}{N_n^{(i)}+1}.
\end{aligned}$$

Thus, to satisfy (A.7), for all large enough n, it is sufficient to have

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

https://doi.org/10.1017/apr.2019.9 Published online by Cambridge University Press

$$\frac{\delta_{n+u}^{(j)}}{\delta_n^{(j)}} \frac{\delta_n^{(i)} - \delta_n^{(j)}\,(1 + C_2(\log n)/n)}{N_n^{(1)}} \frac{1}{1 + C_{10}(\log n)/n} \frac{N_n^{(i)}}{N_n^{(i)} + 1}$$
$$\geq \frac{\delta_{n+u}^{(i)} - \delta_{n+u}^{(j)}\,(1 - C_4(\log n)/n)}{N_n^{(1)} + k_u^{(1)}},$$

which can equivalently be rewritten as

$$k_u^{(j)} \geq \frac{\delta_n^{(j)}}{\delta_{n+u}^{(j)}} \frac{\delta_{n+u}^{(i)}}{\delta_n^{(i)}} \left(1 + C_{10}\frac{\log n}{n}\right) \frac{N_n^{(i)} + 1}{N_n^{(i)}} N_n^{(j)} - N_n^{(j)}, \tag{A.12}$$

$$k_u^{(1)} \geq \frac{\delta_n^{(j)}}{\delta_{n+u}^{(j)}} \left(1 + C_{10}\frac{\log n}{n}\right) \frac{\delta_{n+u}^{(i)} - \delta_{n+u}^{(j)}\,(1 - C_4(\log n)/n)}{\delta_n^{(i)} - \delta_n^{(j)}\,(1 + C_2(\log n)/n)} \frac{N_n^{(i)} + 1}{N_n^{(i)}} N_n^{(1)} - N_n^{(1)}. \tag{A.13}$$

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,

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

and

$$\frac{\delta_n^{(j)}}{\delta_{n+u}^{(j)}} \left(1 + C_{10}\frac{\log n}{n}\right) \frac{\delta_{n+u}^{(i)} - \delta_{n+u}^{(j)}\,(1 - C_4(\log n)/n)}{\delta_n^{(i)} - \delta_n^{(j)}\,(1 + C_2(\log n)/n)} \frac{N_n^{(i)} + 1}{N_n^{(i)}} N_n^{(1)} - N_n^{(1)}$$
$$\leq \left(1 + C_{10}\frac{\log n}{n}\right)\left(1 + C_{12}\sqrt{\frac{\log\log n}{n}}\right) \frac{N_n^{(i)} + 1}{N_n^{(i)}} N_n^{(1)} - N_n^{(1)}$$
$$\leq \left(1 + C_{13}\sqrt{\frac{\log\log n}{n}}\right) \frac{N_n^{(i)} + 1}{N_n^{(i)}} N_n^{(1)} - N_n^{(1)}$$
$$\leq C_{14}\sqrt{n \log\log n}.$$

Therefore, to satisfy (A.12) and (A.13), it is sufficient to have

$$k_u^{(j)} \geq C_{11}\sqrt{n \log\log n}, \tag{A.14}$$
$$k_u^{(1)} \geq C_{14}\sqrt{n \log\log n}. \tag{A.15}$$

Again, define s as in (A.11). Since $k_m^{(1)}/\sqrt{n \log\log n}$ can be arbitrarily large, we can suppose that $k_s^{(1)} > C_{15}\sqrt{n \log\log n}$, where C_{15} is a positive constant to be specified. By Lemma A.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

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

whence

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

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

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

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 $C_{15} = \max\{C_{16}, C_{11}/C_{17}, 2B_2C_{14}/C_{17}\}$. Then both (A.14) and (A.15) are satisfied at stage $n+u$, so (A.7) is satisfied, which means that

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

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.

A.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 $n^{3/4}|\delta_{n+1}^{(i)} - \delta_n^{(i)}| = 0$.

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

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

where

$$\begin{aligned} n\big|\theta_{n+1}^{(i)} - \theta_n^{(i)}\big| &= n\Big|\frac{1}{N_n^{(i)}+1}\big(N_n^{(i)}\theta_n^{(i)} + W_{n+1}^{(i)}\big) - \theta_n^{(i)}\Big| \\ &\leq \frac{n}{N_n^{(i)}+1}\big|\theta_n^{(i)}\big| + \frac{n}{N_n^{(i)}+1}\big|W_{n+1}^{(i)}\big| \\ &= O(1)(1+|W_{n+1}^{(i)}|); \end{aligned}$$

thus, there exists a constant C_2 such that

$$n^{3/4}\big|\delta_{n+1}^{(i)} - \delta_n^{(i)}\big| \quad \leq \quad \frac{C_2}{n^{1/4}}\big(1+\big|W_{n+1}^{(i)}\big|\big).$$

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

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

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

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

https://doi.org/10.1017/apr.2019.9 Published online by Cambridge University Press

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

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

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

A.5. Proof of Lemma A.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 $\lim_{n\to\infty} \sqrt{n \log\log n}/n = 0$, 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 $C_2\sqrt{n \log\log n} \leq k_s^{(1)} \leq n$ for some positive constant C_2 to be specified.

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

$$\left(\frac{N_n^{(1)}}{\lambda^{(1)}}\right)^2 \geq \sum_{j=2}^{M} \left(\frac{N_n^{(j)}}{\lambda^{(j)}}\right)^2. \tag{A.16}$$

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

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

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

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

From Lemma 4.2, we know that $\liminf_{n\to\infty} N_n^{(1)}/n > 0$. 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

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

https://doi.org/10.1017/apr.2019.9 Published online by Cambridge University Press

and, for all large enough n,

$$\begin{aligned}
&\sum_{j\geq 2,\, j\neq i}\left[\left(\frac{(N_n^{(j)}+k_s^{(j)})/\lambda^{(j)}}{(N_n^{(1)}+k_s^{(1)})/\lambda^{(1)}}\right)^2-\left(\frac{N_n^{(j)}/\lambda^{(j)}}{N_n^{(1)}/\lambda^{(1)}}\right)^2\right]\\
&\quad\geq 1-\sum_{j\geq 2,\, j\neq i}\left(\frac{N_n^{(j)}/\lambda^{(j)}}{N_n^{(1)}/\lambda^{(1)}}\right)^2-\left(\frac{(N_n^{(i)}+1)/\lambda^{(i)}}{(N_n^{(1)}+k_s^{(1)})/\lambda^{(1)}}\right)^2-\frac{C_4}{n}\\
&\quad\geq \left(\frac{N_n^{(i)}/\lambda^{(i)}}{N_n^{(1)}/\lambda^{(1)}}\right)^2-\left(\frac{(N_n^{(i)}+1)/\lambda^{(i)}}{(N_n^{(1)}+k_s^{(1)})/\lambda^{(1)}}\right)^2-\frac{C_4}{n} \qquad\qquad (A.17)\\
&\quad=\left(\frac{\lambda^{(1)}}{\lambda^{(i)}}\right)^2\frac{(N_n^{(i)})^2(N_n^{(1)}+k_s^{(1)})^2-(N_n^{(1)})^2(N_n^{(i)}+1)^2}{(N_n^{(1)}+k_s^{(1)})^2(N_n^{(1)})^2}-\frac{C_4}{n}\\
&\quad=\left(\frac{\lambda^{(1)}}{\lambda^{(i)}}\right)^2\frac{(N_n^{(i)})^2(2N_n^{(1)}k_s^{(1)}+(k_s^{(1)})^2)-(N_n^{(1)})^2(2N_n^{(i)}+1)}{(N_n^{(1)}+k_s^{(1)})^2(N_n^{(1)})^2}-\frac{C_4}{n}\\
&\quad=\left(\frac{\lambda^{(1)}}{\lambda^{(i)}}\right)^2\frac{(N_n^{(i)})^2\big(2N_n^{(1)}(k_s^{(1)}-N_n^{(1)}/N_n^{(i)})+(k_s^{(1)})^2-(N_n^{(1)}/N_n^{(i)})^2\big)}{(N_n^{(1)}+k_s^{(1)})^2(N_n^{(1)})^2}-\frac{C_4}{n}\\
&\quad>\left(\frac{\lambda^{(1)}}{\lambda^{(i)}}\right)^2\frac{(N_n^{(i)})^2(2N_n^{(1)}(k_s^{(1)}/2)+(k_s^{(1)})^2/2)}{(N_n^{(1)}+k_s^{(1)})^2(N_n^{(1)})^2}-\frac{C_4}{n} \qquad\qquad (A.18)\\
&\quad=\frac{1}{2}\left(\frac{\lambda^{(1)}}{\lambda^{(i)}}\right)^2\frac{1}{(N_n^{(1)}+k_s^{(1)})^2}\left(\frac{N_n^{(i)}}{N_n^{(1)}}\right)^2(2N_n^{(1)}k_s^{(1)}+(k_s^{(1)})^2)-\frac{C_4}{n},
\end{aligned}$$

where (A.17) holds due to (A.16), while (A.18) holds since $\liminf_{n\to\infty} N_n^{(i)}/N_n^{(1)}>0$ and $k_s^{(1)}\geq C_2\sqrt{n\log\log n}$ for a positive constant C_2. Since $\liminf_{n\to\infty} N_n^{(i)}/N_n^{(1)}>0$ and $\liminf_{n\to\infty} N_n^{(1)}/n>0$, there must exist positive constants C_5, C_6, C_7, C_8, and C_9 such that, for all large enough n, we have

$$\begin{aligned}
&\frac{1}{2}\left(\frac{\lambda^{(1)}}{\lambda^{(i)}}\right)^2\frac{1}{(N_n^{(1)}+k_s^{(1)})^2}\left(\frac{N_n^{(i)}}{N_n^{(1)}}\right)^2(2N_n^{(1)}k_s^{(1)}+(k_s^{(1)})^2)-\frac{C_4}{n}\\
&\quad\geq C_5\frac{1}{(N_n^{(1)}+k_s^{(1)})^2}(2N_n^{(1)}k_s^{(1)}+(k_s^{(1)})^2)-\frac{C_4}{N_n^{(1)}}\\
&\quad=\frac{C_5}{(N_n^{(1)}+k_s^{(1)})^2}\left(2N_n^{(1)}k_s^{(1)}+(k_s^{(1)})^2-C_6\frac{(N_n^{(1)}+k_s^{(1)})^2}{N_n^{(1)}}\right)\\
&\quad\geq\frac{C_5}{(N_n^{(1)}+k_s^{(1)})^2}(2N_n^{(1)}k_s^{(1)}+(k_s^{(1)})^2-2C_7N_n^{(1)}) \qquad\qquad (A.19)\\
&\quad\geq\frac{C_5}{(N_n^{(1)}+k_s^{(1)})^2}2(k_s^{(1)}-C_7)N_n^{(1)}\\
&\quad\geq\frac{C_8(k_s^{(1)}-C_7)}{N_n^{(1)}} \qquad\qquad (A.20)
\end{aligned}$$

https://doi.org/10.1017/apr.2019.9 Published online by Cambridge University Press

$$\geq \frac{C_8(C_2\sqrt{n\log\log n}-C_7)}{n}$$
$$\geq \frac{C_9C_2\sqrt{n\log\log n}}{n},$$

where (A.19) and (A.20) hold because $k_s^{(1)} \leq n$. Then

$$\sum_{j\geq 2,\, j\neq i}\left[\left(\frac{(N_n^{(j)}+k_s^{(j)})/\lambda^{(j)}}{(N_n^{(1)}+k_s^{(1)})/\lambda^{(1)}}\right)^2 - \left(\frac{N_n^{(j)}/\lambda^{(j)}}{N_n^{(1)}/\lambda^{(1)}}\right)^2\right] > \frac{C_9C_2\sqrt{n\log\log n}}{n},$$

so there must be some suboptimal j such that

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

Let $C_{10}=C_9/(M-2)$ and $C_{11}=C_{10}C_2/4$. Then,

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

and, for all large enough n, we have

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

whence

$$\frac{N_n^{(j)}+k_s^{(j)}}{N_n^{(1)}+k_s^{(1)}} > \left(1+C_{11}\frac{\sqrt{n\log\log n}}{n}\right)\frac{N_n^{(j)}}{N_n^{(1)}}. \tag{A.21}$$

For the alternative j that satisfies (A.21), let

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

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

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

https://doi.org/10.1017/apr.2019.9 Published online by Cambridge University Press

where the last line follows from (A.21). By Lemma 4.2, there must exist a positive constant C_{12} such that, for all large enough n,

$$\begin{aligned}
&\Big(1 - \frac{1}{N_n^{(j)} + k_s^{(j)}}\Big)\Big(1 + C_{11}\frac{\sqrt{n\log\log n}}{n}\Big)\frac{N_n^{(j)}}{N_n^{(1)}} \\
&\geq \Big(1 - \frac{C_{12}}{n}\Big)\Big(1 + C_{11}\frac{\sqrt{n\log\log n}}{n}\Big)\frac{N_n^{(j)}}{N_n^{(1)}} \\
&= \Big(1 + C_{11}\frac{\sqrt{n\log\log n}}{n} - \frac{C_{12}}{n} - C_{12}C_{11}\frac{\sqrt{n\log\log n}}{n^2}\Big)\frac{N_n^{(j)}}{N_n^{(1)}} \\
&\geq \Big(1 + \frac{C_{11}}{2}\frac{\sqrt{n\log\log n}}{n}\Big)\frac{N_n^{(j)}}{N_n^{(1)}} \\
&= \Big(1 + C_{13}\frac{\sqrt{n\log\log n}}{n}\Big)\frac{N_n^{(j)}}{N_n^{(1)}},
\end{aligned}$$

where $C_{13} = C_{11}/2 = C_{10}C_2/8$. 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. $C_2 \geq 8C_1/C_{10}$, to make $C_{13} \geq C_1$, then

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

which completes the proof.

Acknowledgement

The second author's research was partially supported by the National Science Foundation (EAGER award CMMI-1745198).

References

[1] BECHHOFER, R. E. (1954). A single-sample multiple decision procedure for ranking means of normal populations with known variances. *Ann. Math. Statist.* **25,** 16–39.
[2] BRANKE, J., CHICK, S. E. AND SCHMIDT, C. (2007). Selecting a selection procedure. *Manag. Sci.* **53,** 1916–1932.
[3] CHAU, M., FU, M. C., QU, H. AND RYZHOV, I. O. (2014). Simulation optimization: A tutorial overview and recent developments in gradient-based methods. In *Proc. 2014 Winter Simulation Conference*, eds A. Tolk *et al.*, IEEE, Piscataway, NJ, pp. 21–35.
[4] CHEN, C.-H., LIN, J., YÜCESAN, E. AND CHICK, S. E. (2000). Simulation budget allocation for further enhancing the efficiency of ordinal optimization. *Discrete Event Dynamic Systems* **10,** 251–270.
[5] CHEN, Y. AND RYZHOV, I. O. (2017). Rate-optimality of the complete expected improvement criterion. In *Proc. 2017 Winter Simulation Conference*, eds W. K. V. Chan *et al.*, IEEE, Piscataway, NJ, pp. 2173–2182.
[6] CHICK, S. E., BRANKE, J. AND SCHMIDT, C. (2010). Sequential sampling to myopically maximize the expected value of information. *INFORMS J. Computing* **22,** 71–80.
[7] DEGROOT, M. H. (1970). *Optimal Statistical Decisions*. John Wiley, Hoboken.
[8] GITTINS, J., GLAZEBROOK, K. AND WEBER, R. (2011). *Multi-Armed Bandit Allocation Indices*, 2nd edn. John Wiley, Chichester.
[9] GLYNN, P. W. AND JUNEJA, S. (2004). A large deviations perspective on ordinal optimization. In *Proc. 2004 Winter Simulation Conference*, eds R. Ingalls *et al.*, IEEE, pp. 577–585.
[10] HAN, B., RYZHOV, I. O. AND DEFOURNY, B. (2016). Optimal learning in linear regression with combinatorial feature selection. *INFORMS J. Computing* **28,** 721–735.

[11] HONG, L. J. AND NELSON, B. L. (2009). A brief introduction to optimization via simulation. In *Proc. 2009 Winter Simulation Conference*, eds M. Rosetti *et al.*, IEEE, pp. 75–85.

[12] HUNTER, S. R. AND MCCLOSKY, B. (2016). Maximizing quantitative traits in the mating design problem via simulation-based Pareto estimation. *IIE Trans.* **48,** 565–578.

[13] JONES, D. R., SCHONLAU, M. AND WELCH, W. J. (1998). Efficient global optimization of expensive black-box functions. *J. Global Optimization* **13,** 455–492.

[14] KIM, S.-H. AND NELSON, B. L. (2001). A fully sequential procedure for indifference-zone selection in simulation. *ACM Trans. Model. Comput. Simul.* **11,** 251–273.

[15] PASUPATHY, R. *et al.* (2014). Stochastically constrained ranking and selection via SCORE. *ACM Trans. Model. Comput. Simul.* **25,** 26p.

[16] PENG, Y. AND FU, M. C. (2017). Myopic allocation policy with asymptotically optimal sampling rate. *IEEE Trans. Automatic Control* **62,** 2041–2047.

[17] POWELL, W. B. AND RYZHOV, I. O. (2012). *Optimal Learning*. John Wiley, Hoboken.

[18] QIN, C., KLABJAN, D. AND RUSSO, D. (2017). Improving the expected improvement algorithm. In *Advances in Neural Information Processing Systems*, Vol. 30, eds I. Guyon *et al.*, Neural Information Processing Systems, pp. 5381–5391.

[19] RUBEN, H. (1962). A new asymptotic expansion for the normal probability integral and Mill's ratio. *J. R. Statist. Soc. B* **24,** 177–179.

[20] RUSSO, D. (2017). Simple Bayesian algorithms for best arm identification. Preprint. Available at https://arxiv.org/abs/1602.08448.

[21] RUSSO, D. AND VAN ROY, B. (2014). Learning to optimize via posterior sampling. *Math. Operat. Res.* **39,** 1221–1243.

[22] RYZHOV, I. O. (2016). On the convergence rates of expected improvement methods. *Operat. Res.* **64,** 1515–1528.

[23] SALEMI, P., NELSON, B. L. AND STAUM, J. (2014). Discrete optimization via simulation using Gaussian Markov random fields. In *Proc. 2014 Winter Simulation Conference*, eds A. Tolk *et al.*, IEEE, Piscataway, NJ, pp. 3809–3820.

[24] SCOTT, W. R., POWELL, W. B. AND SIMÃO, H. P. (2010). Calibrating simulation models using the knowledge gradient with continuous parameters. In *Proc. 2010 Winter Simulation Conference*, eds B. Johansson *et al.*, IEEE, Piscataway, NJ, pp. 1099–1109.