Proceedings of the 2024 Winter Simulation Conference
H. Lam, E. Azar, D. Batur, S. Gao, W. Xie, S. R. Hunter, and M. D. Rossetti, eds.

SELECTING THE SAFEST DESIGN IN RARE EVENT SETTINGS

Anirban Bhattacharjee, and Sandeep Juneja

Ashoka University, Haryana, INDIA

ABSTRACT

Finitely many simulatable designs are given and we aim to identify the safest one, i.e., that with the smallest probability of catastrophic failure. We consider this problem in a ranking and selection or equivalently the multi-armed-bandit best-arm-identification framework where we aim to identify with high probability the safest design/arm with the lowest probability of failure. To illustrate the rarity structure crisply, we study the problem in an asymptotic regime where the design failure probabilities shrink to zero at varying rates. In this set-up, we consider the well known information theoretic lower bound on sample complexity, and identify the simplifications that arise due to the rarity framework. A key insight is that sample complexity is governed by the rarity of the second safest design. The proposed algorithm is guided by the lower bound, it is intuitive and asymptotically matches the lower bound.

1 INTRODUCTION

Consider simulation models of complex systems where failure in each system is rare but its consequences can be disastrous. For instance, models that capture a nuclear power plant, electrical grid for a region, insurance company cash-flow model, and so on. Our objective may be to identify a system amongst many with the smallest rare event probability at a minimum computation cost. We model this rare event problem in a ranking and selection, or equivalently, a best-arm identification problem in a multi-armed bandit learning setting. Specifically, we consider a fixed confidence, pure exploration best-arm-identification setting. Simulation model of each system can generate i.i.d. samples with failure occurring with a tiny system dependent probability. Our aim is to sequentially sample and identify the system with the lowest failure probability with correct selection probabilistic guarantee of $1-\delta$, for a pre-decided $\delta > 0$.

These problems of best arm identification with minimum sample complexity have been well studied in the learning theory literature (see e.g., Garivier and Kaufmann (2016); Even-Dar et al. (2006); Kalyanakrishnan et al. (2012)). Earlier, substantial literature existed in the simulation community, typically referred to as ranking and selection problems (see, e.g., Chen et al. (2000); Glynn and Juneja (2004); Kim and Nelson (2006)). Often in these works, the system samples are assumed to be from Gaussian distributions. See Chernoff (1959); Paulson (1964) for early literature in statistics. Best arm problems related to rare events were studied by Bekki et al. (2007) and Batur and Choobineh (2010), where quantile-based selection methods were examined. CVar-related BAI problems were considered in Agrawal et al. (2021) and Ahn and Kim (2023).

Garivier and Kaufmann (2016) use information theoretic ideas to develop lower bounds on sample complexity in the best arm selection problem when the underlying distributions belong to a single parameter exponential family. This includes Bernoulli distribution (see Agrawal et al. (2020) for extension to general distributions), and they develop algorithms that match the lower bound asymptotically as $\delta \to 0$.

Our aim in this paper is to develop structural insights into this best-arm-identification (BAI) setting when the underlying probability of failure is tiny. We do this by analyzing the problem in an asymptotic regime where the failure probabilities are indexed by a rarity parameter γ and analyze the system as $\gamma \to 0$. BAI in the rare event setting was also considered by Bhattacharjee et al. (2023) where the focus is on online advertising where the click probabilities are small but the rewards at each click could potentially be very large, and the aim is to identify the most profitable system. Structurally, their lower bound analysis as well the designed algorithms are substantially different from our safest system identification problem.

Our key insights are well explained through a somewhat extreme example. Suppose there are three systems with unknown failure probabilities 10^{-4}, 10^{-6} and 10^{-10}. Initially the simulator will keep simulating these systems equally until each system has multiples of 10^4 samples and one begins to observe failure events in system 1. Once statistically enough failures have been observed in system 1, so that the simulator is convinced that this system is not safe, it will focus effort on the other two systems. Once the samples given to the other two systems are in multiples of 10^6, and statistically adequate failure events are observed in system 2, the simulator is ready to conclude that system 3 is the safest without having to generate order 10^{10} samples. Thus, the computation order is governed by the second safest system. We also observe that computationally, each system may require a different budget on average to generate each sample. This heterogeneity results in relatively less sampling of the arm with larger computation requirement per sample, and is easily incorporated in the lower bound analysis and the algorithm.

Our key contributions are: 1) We analyze the lower bound in our rarity framework and conclude that the overall computational effort is driven by the rarity of the second safest design. 2) Through lower bound analysis, we also develop further structural insights into efficient algorithms for this problem, and propose a modification to existing algorithm tailored to the rare event setting. 3) We show how to incorporate computational effort per sample into the lower bound analysis and hence on optimal algorithms in fixed-confidence best-arm identification problems.

In Section 2, we discuss the background associated with fixed confidence best-arm identification problems and discuss the lower bound in the setting where the average computational cost per sample may differ across different designs. In Section 3, we introduce our rarity framework and discuss simplifications arising from this framework. In Section 4 we introduce our algorithm and show its performance on simple highly reliable systems. All proofs are outlined (due to space paucity) and are in the appendix.

2 BACKGROUND

Consider K simulatable systems. Simulating system i involves generating i.i.d. samples of a random element $\mathbf{X}_i$. Let A_i denote the rare failure event for system (or, equivalently in our terminology, arm) i. Our aim is to identify the system with the smallest probability $\eta_i = P(\mathbf{X}_i \in A_i)$.

Let $d(\eta,\kappa)$ denote the Kullback-Leibler divergence between two Bernoulli distributions with failure probabilities η and κ, then $d(\eta,\kappa) = \eta \log\left(\frac{\eta}{\kappa}\right) + (1-\eta)\log\left(\frac{1-\eta}{1-\kappa}\right)$.

δ-correct algorithms: We propose sequential algorithms, where the K systems are sequentially sampled. Sampling rule is adaptive and depends upon observed history. The algorithm stops at a random time τ at which point it declares the safest system. Algorithm is said to be $\delta \in (0,1)$ correct if it guarantees that the probability of correct selection is $\geq 1-\delta$.

Suppose, without loss of generality, that design 1 is the safest, that is, $\eta_1 < \min_{i\geq 2} \eta_i$. Let $\kappa = (\kappa_i : i \in [K])$ belong to the alternate conclusion set A where each $\kappa_1 \geq \min_{i\geq 2} \kappa_i$, then using data processing inequality, it is easily seen that for any δ correct algorithm that terminates at stopping time τ_δ with allocations $(N_i : i \in [K])$,

$$\sum_{i\in[K]} N_i d(\eta_i,\kappa_i) \geq \log(1/2.4\delta). \tag{1}$$

See, e.g., Kaufmann (2020). A lower bound on sample complexity $\mathbb{E}_\eta(\tau_\delta)$ then follows as a solution to the optimization problem $\min_{i\in[K]} \sum_{i\in[K]} N_i$: (1) holds for each $\kappa \in A$, where the subscript η denotes that the underlying arm sampling probability measures correspond to η. This through simplifications leads to the well known lower bound: $\mathbb{E}_\eta[\tau_\delta] \geq V^*(\eta)^{-1}\log\left(\frac{1}{\delta}\right)$ where $V^*(\eta) := \max_{w_1+\ldots+w_K=1} \min_{i\neq 1} \inf_{\eta_{1i}\in[\eta_i,\eta_1]} w_1 d(\eta_1,\eta_{1i}) + w_i d(\eta_i,\eta_{1i})$. Recall that $\mathbb{E}_\eta$ is standard notation for expectation under η.

In our setting, the computational effort for generating a sample from each arm i can be random and different. Letting c_i be the expected computational effort under η, we consider the optimization problem $\min_{i\in[K]}\sum_{i\in[K]} c_i N_i$: (1) holds for each $\kappa\in A$. This leads to the following lower bound on the computational complexity, call it, $E_\eta[\tau_{c,\delta}]$:

$$\mathbb{E}_\eta[\tau_{c,\delta}] \geq V_c^*(\eta)^{-1}\log\left(\frac{1}{2.4\delta}\right) \tag{2}$$

where $V_c^*(\eta) := \max_{w_1+\ldots+w_K=1}\min_{i\neq 1}\inf_{\eta_{1i}\in[\eta_i,\eta_1]}\left(\frac{w_1}{c_1}d(\eta_1,\eta_{1i})+\frac{w_i}{c_i}d(\eta_i,\eta_{1i})\right)$.

Let $\eta_i^* := \dfrac{c_i w_1\eta_1 + c_1 w_i \eta_i}{c_i w_1 + c_1 w_i}$ $\forall i\neq 1$ (we hide the dependence on η_1, w_1, η_i, w_i for notational simplicity). This can be seen to be the unique solution to the inner infimum in $V_c^*(\eta)$ for Bernoulli distributions. Thus, (2) simplifies to $V_c^*(\eta) = \max_{\sum_{i=1}^K w_i=1}\min_{i\neq 1}\left(\frac{w_1}{c_1}d(\eta_1,\eta_i^*)+\frac{w_i}{c_i}d(\eta_i,\eta_i^*)\right)$ (see, e.g., Kaufmann et al. (2016)). Further, through simple convex analysis, it can be seen that the optimal w_i^*, $i\in[K]$, uniquely satisfy the first order conditions (see, e.g., Agrawal et al. (2023) when all $c_i=1$).

$$\frac{w_1^*}{c_1}d(\eta_1,\eta_{12}^*)+\frac{w_2^*}{c_2}d(\eta_2,\eta_{12}^*) \;=\; \frac{w_1^*}{c_1}d(\eta_1,\eta_{1j}^*)+\frac{w_j^*}{c_j}d(\eta_j,\eta_{1j}^*)\;\forall\; j\neq 1,2 \tag{3}$$

$$\sum_{i=2}^K \frac{c_i d(\eta_1,\eta_i^*)}{c_1 d(\eta_i,\eta_i^*)} = 1.$$

3 THE RARE EVENT REGIME

We assume that each $\eta_i = f_i(\gamma)$, where each f_i is continuous and strictly increasing in γ with $f_i(0)=0$. Further, for all γ, $\gamma^b \leq f_1(\gamma) < f_2(\gamma)\leq f_3(\gamma)\leq \ldots \leq f_K(\gamma)$ for some $b>0$. We study the lower bound value and optimal allocations $w^*(\gamma) = (w_1^*(\gamma),\ldots,w_K^*(\gamma))$ as $\gamma\to 0$. For notational ease, we suppress the dependence of f_i and w_i^* on γ when it causes no confusion. Let $L_{ij} := \lim_{\gamma\to 0}\dfrac{f_i(\gamma)}{f_j(\gamma)}$, where we assume that the limit exists, hence $L_{ij}\in[0,\infty]$ for all i,j.

We partition our problem instances into three sets: $\mathbf{U}_1$ denotes the instances where $L_{12}=L_{2j}=0$ $\forall j\geq 3$, i.e., the safest arm is much safer than all other arms, and the second safest arm is much safer than the remaining arms. $\mathbf{U}_2$ denotes the instances where $L_{1j}=0$ $\forall j\geq 2$ and there exists a $2<m\leq K$ such that $L_{2j}>0$, $2\leq j\leq m$. Thus, the safest arm continues to be much safer than the rest, while there exist arms whose safety levels are similar in order of magnitude to the second safest. $\mathbf{N}$ denotes the instances where $L_{12}>0$.

Theorem 1 below offers insights on the optimal weights $\{w_i^*: i\in[K]\}$ that solve the maxmin problem given by $V_c^*(\eta)$ in the rare event framework.

Theorem 1 Let w_i^*, $i\in[K]$ solve $V_c^*(\eta)$. Then,

(a) For all j such that $L_{2j}=0$, $w_j^* = \Theta\left(\frac{f_2}{f_j\log\left(\frac{f_j}{f_2}\right)}\right)$.

(b) For all j such that $L_{2j}>0$ or $L_{1j}>0$, $w_j^*(\gamma)\to\hat{w}_j$ as $\gamma\to 0$, where these $\hat{w}_j$ uniquely solve

(i) $(c_2\hat{w}_1+c_1\hat{w}_2)\log\left(1+\frac{c_2\hat{w}_1}{c_1\hat{w}_2}\right) = c_1c_2$ and $\hat{w}_1+\hat{w}_2=1$ for $\mathbf{U}_1$,

(ii) $\hat{w}_2 f_2\log\left(\frac{c_2\hat{w}_1+c_1\hat{w}_2}{c_1\hat{w}_2}\right) = \hat{w}_i f_i \log\left(\frac{c_i\hat{w}_1+c_1\hat{w}_i}{c_1\hat{w}_i}\right)$ $\forall\; i=3,\ldots,m$ and

$$\sum_{i=2}^m \frac{\frac{c_1c_i\hat{w}_i}{c_2\hat{w}_1+c_1\hat{w}_i}}{c_1\log\left(\frac{c_i\hat{w}_1+c_1\hat{w}_j}{c_1\hat{w}_i}\right)-\frac{c_1c_i\hat{w}_1}{c_i\hat{w}_1+c_1\hat{w}_i}} = 1 \text{ for } \mathbf{U}_2, \text{ and}$$

(iii) $$\frac{c_2\hat{w}_1 f_1 \log\left(\frac{c_2\hat{w}_1 f_1 + c_1\hat{w}_2 f_1}{c_2\hat{w}_1 f_1 + c_1\hat{w}_2 f_2}\right) + c_1\hat{w}_2 f_2 \log\left(\frac{c_2\hat{w}_1 f_2 + c_1\hat{w}_2 f_2}{c_2\hat{w}_1 f_1 + c_1\hat{w}_2 f_2}\right)}{c_i\hat{w}_1 f_1 \log\left(\frac{c_i\hat{w}_1 f_1 + c_1\hat{w}_i f_1}{c_i\hat{w}_1 f_1 + c_1\hat{w}_i f_i}\right) + c_1\hat{w}_i f_i \log\left(\frac{c_i\hat{w}_1 f_i + c_1\hat{w}_i f_i}{c_i\hat{w}_1 f_1 + c_1\hat{w}_i f_i}\right)} = 1 \ \forall\ i = 3,\dots,m \text{ and}$$
$$\sum_{i=1}^{m} \frac{f_1 \log\left(\frac{c_i\hat{w}_1 f_1 + c_1\hat{w}_i f_1}{c_i\hat{w}_1 f_1 + c_1\hat{w}_i f_i}\right) - \frac{c_1\hat{w}_i(f_i - f_1)}{c_i\hat{w}_1 + c_1\hat{w}_i}}{f_i \log\left(\frac{c_i\hat{w}_1 f_i + c_1\hat{w}_i f_i}{c_i\hat{w}_1 f_1 + c_1\hat{w}_i f_i}\right) + \frac{c_i\hat{w}_1(f_i - f_1)}{c_i\hat{w}_1 + c_1\hat{w}_i}} = 1 \text{ for } \mathbf{N}.$$

The proof of the above theorem makes use of Lemmas 1 through 5. Before we state the lemmas, we need some notation.

Let

$$\tilde{d}_{1i}(w_1,w_i) := \begin{cases} \dfrac{c_1 w_i}{c_i w_1 + c_1 w_i} f_i, & L_{1i} = 0 \\ f_1 \log\left(\dfrac{c_i w_1 f_1 + c_1 w_i f_1}{c_i w_1 f_1 + c_1 w_i f_i}\right) + \dfrac{c_1 w_i}{c_i w_1 + c_1 w_i}(f_i - f_1), & L_{1i} > 0 \end{cases}$$

$$\tilde{d}_{i}(w_1,w_i) := \begin{cases} f_i \log\left(\dfrac{c_i w_1 + c_1 w_i}{c_i w_1 \frac{f_1}{f_i} + c_1 w_i}\right) - \dfrac{c_i w_1}{c_i w_1 + c_1 w_i} f_i, & L_{1i} = 0 \\ f_i \log\left(\dfrac{c_i w_1 f_i + c_1 w_i f_i}{c_i w_1 f_1 + c_1 w_i f_i}\right) - \dfrac{c_i w_1}{c_i w_1 + c_1 w_i}(f_i - f_1), & L_{1i} > 0. \end{cases}$$

$$\text{Let } \tilde{V}_c^*(\eta) = \max_{\sum_{i=1}^{K} w_i = 1} \min_{i \neq 1} \left(\frac{w_1}{c_1}\tilde{d}_{1i}(w_1,w_i) + \frac{w_i}{c_i}\tilde{d}_i(w_1,w_i)\right).$$

The $\tilde{w}_i^*$s solving this optimization problem can be seen to satisfy

$$\frac{\tilde{w}_1^*}{c_1}\tilde{d}_{12}(\tilde{w}_1^*,\tilde{w}_2^*) + \frac{\tilde{w}_2^*}{c_2}\tilde{d}_2(\tilde{w}_1^*,\tilde{w}_2^*) \ = \ \frac{\tilde{w}_1^*}{c_1}\tilde{d}_{1j}(\tilde{w}_1^*,\tilde{w}_j^*) + \frac{\tilde{w}_j^*}{c_j}\tilde{d}_j(\tilde{w}_1^*,\tilde{w}_j^*) \ \forall\ j \neq 1,2 \tag{4}$$

and

$$\sum_{i=2}^{K} \frac{c_i}{c_1} \frac{\tilde{d}_{1i}(\tilde{w}_1^*,\tilde{w}_j^*)}{\tilde{d}_i(\tilde{w}_1^*,\tilde{w}_j^*)} = 1. \tag{5}$$

Below, we suppress the dependence on γ in our notation for each $f_i(\gamma)$.

Lemma 1 There exist constants $A_{U_1}, A_{U_2}, A_{N_1}, A_{N_2}$ independent of w_i, $\forall i \in [K]$, such that

(a) for all i such that $L_{1i} = 0$,
  (i) $|d(f_1,f_i^*) - \tilde{d}_{1i}(w_1,w_i)| \leq A_{U_1} f_1$,
  (ii) $|d(f_i,f_i^*) - \tilde{d}_i(w_1,w_i)| \leq A_{U_2} f_1$; and

(b) for all i such that $L_{1i} \neq 0$,
  (i) $|d(f_1,f_i^*) - \tilde{d}_{1i}(w_1,w_i)| \leq A_{N_1} f_1^2$
  (ii) $|d(f_i,f_i^*) - \tilde{d}_i(w_1,w_i)| \leq A_{N_2} f_1^2$.

Lemma 2 There exist constants C_U and C_N such that

$$|V_c^*(\eta) - \tilde{V}_c^*(\eta)| \leq \begin{cases} C_U f_1 & \text{in } \mathbf{U}_1 \cup \mathbf{U}_2 \\ C_N f_1^2 & \text{in } \mathbf{N}. \end{cases}$$

Remark 1 In Theorem 1, we see that the maxmin optimization problem is solved when $w_j^* = \Theta\left(\frac{f_2}{f_j(\log(\gamma^{-1}))^{-1}}\right)$ for all non-rarest arms j. Given the lower bound sample complexity of $\Theta(1/f_2)$ (see Lemma 4), the expected number of failures seen in non-rarest arms turns out to be $\Theta(\log(1/\delta)^{-1})$, which is less than 1 for small δ. This suggests that the lower bound on sample complexity obtained by solving the maxmin problem need not be an accurate guide to the algorithm. In particular, more samples need to be given to sub optimal arms so that a reasonable number of failure events are observed.

The next lemma uses the approximations from Lemma 1 to arrive at an important conclusion which will be used to prove Theorem 1.

Lemma 3 $w_j^*, \tilde{w}_j^* \to 0$ as $\gamma \to 0 \; \forall \; j$ such that $L_{2j} = 0$.

The above lemma can be used to obtain the following lower bound guarantee.

Lemma 4 Let $\tau_{c,\delta}$ be the random stopping time of a δ-correct algorithm in the rare event regime. Then, for γ sufficiently small, $V_c^*(\eta), \tilde{V}_c^*(\eta) = \Theta(f_2)$. Therefore, $\mathbb{E}_\eta[\tau_{c,\delta}] \geq \Theta(1/f_2)$.

In the following result, we prove that our approximate solutions are close to the exact solutions.

Lemma 5 There exist constants S_0 and S_1 independent of w_i, $i \in [K]$, such that whenever γ lies in a sufficiently small neighborhood of 0,

(a) $|w_i^* - \tilde{w}_i^*| \leq S_0 \frac{f_1}{f_i}$, whenever $L_{1i} = 0$.
(b) $|w_i^* - \tilde{w}_i^*| \leq S_1 f_1$, whenever $L_{1i} > 0$.

4 PROPOSED ALGORITHM

We first discuss the Track and Stop (TS) paradigm, popularly used in solving best arm bandit problems. We then suggest modifications to suit our rare event setting.

TS algorithm: The lower bound suggests that each arm should be sampled in proportion to the optimal weights w^* in (2). This idea guides TS algorithms that match the lower bound asymptotically as $\delta \to 0$. Broadly, such algorithms have the following structure (see Garivier and Kaufmann (2016), Agrawal, Juneja, and Glynn (2020) for further details):

1. Arms are sampled sequentially in batches. When a total of t samples are allocated, each arm is sampled at least order $\sqrt{t}$ times (to avoid starvation).
2. Empirical estimates of Bernoulli means at stage t, $\hat{\eta}_t \in \mathbb{R}^K$ are plugged into the lower bound maxmin problem, which is then solved to estimate the prescriptive proportions $\hat{w}_t$. The algorithm then samples to closely track these proportions.
3. The algorithm stops when the generalized log-likelihood ratio (GLLR) at stage t (see, e.g, Chernoff (1959)),
$$\min_{b \neq i^*} Z_{i^*b} := N_{i^*}(t) d(\hat{\eta}_{i^*}, \eta^*_{i^*b}) + N_b(t) d(\hat{\eta}_b, \eta^*_{i^*b})$$
where i^* is the arm with empirically minimum mean, each $N_a(t)$ denotes the samples of arm a generated by stage t, and $\eta^*_{i^*b} = \frac{c_b w_{i^*} \eta_{i^*} + c_{i^*} w_b \eta_b}{c_b w_{i^*} + c_{i^*} w_b} \; \forall b \neq i^*$, exceeds a well chosen threshold $\beta(t,\delta)$ (asymptotically similar to $\log(1/\delta)$ for δ small). Typically $\beta(t,\delta) := \log\left(\frac{2t(K-1)}{\delta}\right)$.

TS algorithms in the above form are not easy to implement in the regime of rare events for two main reasons. First, the rarity structure implies that we need large number of trials to see failure events and form reliable estimates of empirical distributions. Second, even though the high computational costs of repeatedly solving the lower bound problem are usually mitigated by sampling in batches, we need to be careful so that batch sizes are neither so small that no failures are seen in the rarer systems, nor so large that the less rare systems see more failures than we require.

Taking the above two factors into consideration, we modify the TS algorithm so that it fits into the rare events framework. First, our batch sizes increase at a geometric rate until we start seeing failures in some arms. Second, once we start seeing failure instances in some arms, we eliminate them once the corresponding index Z_{i^*b} crosses the threshold $\beta(t,\delta)$. See Algorithm 1 for details. The key insight that we exploit here is that the derivative of the index Z_{i^*b} to $N_{i^*}(t)$, (assuming that the empirical distributions are fixed) can be seen to equal $\tilde{d}(\hat{\eta}_{i^*}, \eta^*_{i^*b})$. This quantity times $N_{i^*}(t)$, becomes close to zero as $N_{i^*}(t)$ becomes much larger than $N_b(t)$ (as should be in optimal allocations as per the lower bound). So by increasing $N_b(t)$ a little more than maybe optimally needed, the index exceeds the threshold $\beta(t,\delta)$, and the arm can be eliminated. Finally, amongst the remaining arms, if in all but one, enough failure events are seen, we can use the standard track and stop optimization to decide further allocations. The rarest arm may have an empirical estimate zero but that does not impact the algorithm. Also, once we see more than 20 failures in an arm, we don't sample from it again it until we reach the final stage. The number 20 is reasonable, although more analysis may help fine tune it.

Algorithm 1: Track and Stop with Elimination

Input: Confidence level δ, Set of K systems that fail with varying rarities.
Output: System k^* that fails with the least probability, correctly identified with probability at least $1-\delta$.

1: $m \leftarrow 1$, $g = 10^{-1}$, $N_i = 0\ \forall i \in [K]$.
2: $\mathscr{A} = [K]$ to store arms that will be sampled in stage 1.
3: $\mathscr{S} = [K]$ to store arms that are not eliminated in stage 1.
4: Generate $\lfloor \frac{1}{g^m} \rfloor$ samples for each system in $\mathscr{A}$.
5: Update N_i, the number of times the ith system has been sampled so far, set $t \leftarrow \sum_{i=1}^{K} N_i$.
6: Update S_i, the number of failures observed in the ith system so far.
7: Compute the empirical means $\hat{\eta} = (\hat{\eta}_i)_{i\in\mathscr{A}}$.
8: **if** $\hat{\eta}_i = 0\ \forall i \in \mathscr{A}$ **then**
9:     $m \leftarrow m+1$.
10:     Go to 2.
11: **end if**
12: $I^* \leftarrow \arg\min_{i\in\mathscr{A}} \hat{\eta}_i$. Note that the set I^* may have more than one value.
13: **while** $|\mathscr{S}| > 1$ **do**
    Stage 1
14:     Arbitrarily pick $i^* \in I^*$.
15:     For each $b \notin I^*$, compute $Z_{i^*b} := N_i^* \tilde{d}(\hat{\eta}_{i^*}, \eta^*_{i^*b}) + N_b \tilde{d}(\hat{\eta}_b, \eta^*_{i^*b})$.
16:     For each b such that $Z_{i^*b} > \beta(t,\delta)$, set $\mathscr{S} \leftarrow \mathscr{S} \backslash \{b\}$.
17:     For each b such that $Z_{i^*b} > \beta(t,\delta)$ or $S_b > 20$, set $\mathscr{A} \leftarrow \mathscr{A} \backslash \{b\}$.
18:     **if** $\min_{b\in\mathscr{A}\backslash I^*} S_b = 0$ **then**
19:         Generate$\lfloor \frac{1}{g^{m+1}} \rfloor$ samples for each $i \in \mathscr{A}$ such that $S_i = 0$.
20:         Generate $\lfloor \frac{1}{g^m} \rfloor$ samples for each $i \in \mathscr{A}$ such that $S_i > 0$.
Continues on next page
21:         Update N_i, S_i, t, $\hat{\eta}$, i^*.
22:         **Continue**.
23:     **end if**

Stage 2

Let $\hat{w}^*$ be the optimal weights satisfying $V_c^*(\hat{\eta}_{\mathscr{S}})$where $\hat{\eta}_{\mathscr{S}}$ is the vector of empirical means of the arms that did not get eliminated in Stage 1.

For each $i \in \mathscr{S}$, evaluate $Y_i = t\hat{w}_i^* - N_i$.

Generate$\lfloor \frac{1}{g^{m+1}} \rfloor$ samples for each $i \in \arg\max Y_i$

Update N_i, S_i, t, $\hat{\eta}$, i^*.

Continue.

end while

Return the remaining element of $\mathscr{S}$.

δ-correctness: The proof of δ-correctness of our algorithm employs standard machinery where it is more or less independent of arms allocation strategy. We omit the proof. It can be found in Kaufmann (2020).
Sample complexity: The proof that the sample complexity of our proposed algorithm matches the lower bound up to a constant is again a set of standard steps, which are similar to those in e.g., Kaufmann and Koolen (2021).

5 NUMERICAL EXPERIMENTS

We let $\gamma = 10^{-2}$. For the sake of simplicity, we allow all systems to have equal computation cost and let f_i be of the form $p_i\gamma^{\alpha_i}$ for each i. Each system has four independent components, and is considered to fail only if all its components fail. Let $\alpha = (\alpha_1, ..., \alpha_K)$ be a vector representing the rarities of systems 1 through K, and $p = (p_1, ..., p_K)$ be a vector representing the coefficients of γ^{α_i}. We consider $\alpha = (3,2,1)$, $\alpha = (4,3,3,2)$ and $\alpha = (3,3,2)$ - one example of each of the scenarios $\mathbf{U}_1$, $\mathbf{U}_2$ and $\mathbf{N}$. For each scenario, we list the values of δ in Table 1 and run our algorithm $1/\delta$ times. Observed average sample complexities are represented by $\hat{\tau}_\delta$, theoretical optimal weights by w^* and observed average weights by $\tilde{w}^*$. The findings are reported in the table below. In all three examples, the best arm was correctly returned in all simulations.

Table 1: Results of simulation experiments on highly reliable systems. Their respective average runtimes were 0.2s, 1.3s and 53.2s respectively.

α	p	δ	Lower Bound	$\hat{\tau}_\delta$	w^*	$\tilde{w}^*$
$(3,2,1)$	$(1,2,4)$	10^{-5}	$1.32e+05$	$2.63e+05$	$(0.62, 0.37, 1.64e-04)$	$(0.62, 0.37, 4.55e-03)$
$(3,2,2,1)$	$(1,3,4,2)$	10^{-5}	$1.32e+05$	$5.74e+05$	$(0.59, 0.34, 0.06, 3.15e-04)$	$(0.58, 0.35, 0.06, 3.14e-04)$
$(2,2,1)$	$(1,3,5)$	10^{-3}	$2.41e+06$	$6.33e+06$	$(0.513, 0.486, 7.69e-06)$	$(0.513, 0.486, 1.75e-04)$

Above, when w_b^* is small for some arm, the corresponding $\tilde{w}_b^*$ is a little larger. As we discussed, this is since the lower bound allocation w_b^* can be misleading, and require that we give insufficient samples to an arm so that even a single failure is not observed.

ACKNOWLEDGMENTS

We thank the reviewers of WSC 2024 for their insightful comments that helped us improve the manuscript. Much of this work was conducted when both the authors were in TIFR. We also acknowledge support from the Department of Atomic Energy, Government of India, under project no. 12-R&D-TFR-5.01-0500 to TIFR.

A PROOF OUTLINE OF LEMMA 1

For part (a),

$$
\begin{aligned}
d(f_1, f_i^*) &= d\left(f_1, \frac{c_i w_1 f_1 + c_1 w_i f_i}{c_i w_1 + c_1 w_i}\right) \\
&= f_1 \log\left(\frac{c_i w_1 f_1 + c_1 w_i f_1}{c_i w_1 f_1 + c_1 w_i f_i}\right) - (1-f_1)\log\left(1 - \frac{(f_i - f_1)c_1 w_i}{(c_i w_1 + c_1 w_i)(1-f_1)}\right)
\end{aligned}
$$

The following statements can easily be shown using $\log(1+x) \leq x$ along with simple algebra.

- $\log f_1 - b\log\gamma \leq \log\left(\frac{c_i w_1 f_1 + c_1 w_i f_1}{c_i w_1 f_1 + c_1 w_i f_i}\right) \leq -\log f_1 + b\log\gamma.$
- $L_{1i} \leq -(1-f_1)\log\left(1 - \frac{(f_i-f_1)c_1 w_i}{(c_i w_1 + c_1 w_i)(1-f_1)}\right) \leq U_{1i}$, where $L_{1i} := \frac{c_1 w_i}{c_i w_1 + c_1 w_i} f_i - f_1$ and $U_{1i} := \frac{c_1 w_i}{c_i w_1 + c_1 w_i} f_i + f_1$.

These inequalities together show that

$$
\begin{aligned}
&- f_1(-\log f_1 + b\log\gamma) - f_1 \leq d(f_1, f_i^*) - \frac{c_1 w_i}{c_i w_1 + c_1 w_i} f_i \leq f_1 + f_1(-\log f_1 + b\log\gamma)\frac{p_i}{p_1} p_1 \gamma^{\alpha_1} \\
&\Rightarrow \left| d(f_1, f_i^*) - \tilde{d}_{1i}(w_1, w_i) \right| \leq f_1(1 - \log f_1 + b\log\gamma) \leq f_1
\end{aligned}
$$

which in turn proves part (a)(i).

Next, $d(f_i, f_i^*) = f_i \log\left(\frac{c_i w_1 f_i + c_1 w_i f_i}{c_i w_1 f_1 + c_1 w_i f_i}\right) - (1-f_i)\log\left(1 + \frac{(f_i-f_1)c_1 w_i}{(c_i w_1 + c_1 w_i)(1-f_1)}\right)$. Similarly as part (a)(i), $L_i \leq -(1-f_i)\log\left(1 + \frac{(f_i-f_1)c_1 w_i}{(c_i w_1 + c_1 w_i)(1-f_1)}\right) \leq U_i$, where $L_i := -\frac{c_1 w_i}{c_i w_1 + c_1 w_i} f_i - f_1$ and $U_i := -\frac{c_1 w_i}{c_i w_1 + c_1 w_i} f_i + f_1$. Therefore,

$$
\begin{aligned}
&- f_1 \leq d(f_i, f_i^*) - f_i \log\left(\frac{c_i w_1 f_i + c_1 w_i f_i}{c_i w_1 f_1 + c_1 w_i f_i}\right) + \frac{c_1 w_i}{c_i w_1 + c_1 w_i} f_i \leq f_1 \\
&\Rightarrow \left| d(f_i, f_i^*) - \tilde{d}_{1i}(w_1, w_i) \right| \leq f_1
\end{aligned}
$$

This proves part (a)(ii).

For part (b),similar to part (a), we get

$$
\left| d(f_1, f_i^*) - \tilde{d}_{1i}(w_1, w_i) \right| \leq \max\left\{ f_i^2 + (f_i - f_1) f_1, \frac{(f_i - f_1)^2}{2(1-f_1)} \right\},
$$

making use of

- $d(f_1, f_i^*) = f_1 \log\left(\frac{c_i w_1 f_1 + c_1 w_i f_1}{c_i w_1 f_1 + c_1 w_i f_i}\right) - (1-f_1)\log\left(1 - \frac{c_1 w_i (f_i - f_1)}{(c_i w_1 + c_1 w_i)(1-f_1)}\right).$
- $-(-\log f_1 + b\log\gamma) \leq \log\left(\frac{c_i w_1 f_1 + c_1 w_i f_1}{c_i w_1 f_1 + c_1 w_i f_i}\right) \leq -\log f_1 + b\log\gamma$
- $L_{1i} \leq -(1-f_1)\log\left(1 - \frac{c_1 w_i (f_i - f_1)}{(c_i w_1 + c_1 w_i)(1-f_1)}\right) \leq U_{1i}$, where $L_{1i} := \frac{c_1 w_i (f_i - f_1)}{(c_i w_1 + c_1 w_i)} - (f_i - f_1) f_1$ and $U_{1i} := \frac{c_1 w_i (f_i - f_1)}{(c_i w_1 + c_1 w_i)} + \frac{(f_i - f_1)^2}{2(1-f_1)}$.

This proves part (b)(i). Proof of part (b)(ii) is on similar lines.

B PROOF OUTLINE OF LEMMA 2

Let $w_i^*, f_i^*, i \in [K]$ solve the maxmin problem $V_c^*(\eta)$ and let $\tilde{w}_i^*, \tilde{f}_i^*, i \in [K]$ solve the maxmin problem $\tilde{V}_c^*(\eta)$. The statement then follows by observing that $w_1^*, \ldots, w_K^*$ are non-negative numbers adding up to 1, and using Lemma 1 on the inequality $|V_c^*(\eta) - \tilde{V}_c^*(\eta)| \leq \frac{w_1^*}{c_1}\left|d(f_1, f_i^*) - \tilde{d}_{1i}(w_1^*, w_i^*)\right| + \frac{w_i^*}{c_i}\left|d(f_i, f_i^*) - \tilde{d}_i(w_1^*, w_i^*)\right|$.

C PROOF OUTLINE OF LEMMA 3

We want to show that

$$f_i^* = \frac{c_i w_1^* f_1 + c_1 w_i^* f_i}{c_i w_1^* + c_1 w_i^*} \leq \frac{c_2 f_2}{c_j} + \Theta(f_1) \quad \forall i \neq 1 \tag{6}$$

For the above inequality to hold, we must have $w_j^* \to 0$ as $\gamma \to 0$ for all j such that $L_{2j} = 0$, because the inequality is otherwise violated.

Substituting the approximations of Lemma 1 in (3) gives us

$$c_1 w_2^* f_2 \log\left(\frac{c_2 w_1^* + c_1 w_2^*}{c_2 w_1^* \frac{f_1}{f_2} + c_1 w_2^*}\right) + \Theta(f_1) = c_1 w_j^* f_j \log\left(\frac{c_j w_1^* + c_1 w_j^*}{c_j w_1^* \frac{f_1}{f_j} + c_1 w_j^*}\right) + \Theta(f_1) \qquad \forall\ j \neq 1,2 \tag{7}$$

We will use the above inequality to obtain (6), by making a couple of transformations. Let us define $x_{1i} := \frac{c_1 w_i^*(f_i - f_1)}{c_i w_1^* f_1 + c_1 w_i^* f_i}$ and $x_i := \frac{c_i w_1^*(f_i - f_1)}{c_i w_1^* f_1 + c_1 w_i^* f_i}$. Simple algebraic manipulations will show that $\frac{f_1}{1 - x_{1i}} = f_i^* = \frac{f_i}{1 + x_i}$ and $\frac{x_{1i}}{x_i} = \frac{c_1 w_i^*}{c_i w_1^*}$.

Now, we divide both sides of (7) by w_1^* and use the transformation variables x_{1i} and x_i to rewrite (7) as

$$\frac{x_{12}}{x_2} c_2 f_2 \log(1 + x_2) + \Theta(f_1) = \frac{x_{1j}}{x_j} c_j f_j \log(1 + x_j) + \Theta(f_1)$$

$$\Rightarrow x_{12} c_2 f_2 + \Theta(f_1) \geq x_{1j} \frac{c_j f_j}{1 + x_j} + \Theta(f_1) = x_{1j} c_j f_j^* + \Theta(f_1)$$

where the last step uses $\log(1+y) \leq y$ and $\dfrac{\log(1+y)}{y} \geq \dfrac{1}{1+y}$ for $y > 0$. Next, we observe that $\frac{c_2 x_{12}}{c_j x_{1j}} = \dfrac{c_2\left(1 - \frac{f_1}{f_2^*}\right)}{c_j\left(1 - \frac{f_1}{f_j^*}\right)} = \frac{c_2}{c_j}\left(1 - \dfrac{\frac{f_1}{f_2^*} - \frac{f_1}{f_j^*}}{1 - \frac{f_1}{f_j^*}}\right) \leq \frac{c_2}{c_j}$. This helps us conclude $f_j^* \leq \frac{c_2}{c_j} f_2$, and (6) follows immediately. The exact same proof technique will work for $\tilde{w}_j^*$s.

D PROOF OUTLINE OF LEMMA 4

Substituting the approximations of Lemma 1 and using Lemma 3 in (4) and (5), observe that as $\gamma \to 0$, $w_i^*, \tilde{w}_i^*$ converge to positive numbers for all i such that $L_{2i} > 0$ for $\mathbf{U}_2$, and $L_{1i} > 0$ for $\mathbf{N}$. We also note that if w_i converges to a positive number as $\gamma \to 0$ for some $i \in [K]$, then it can be shown that w_i's are continuous in γ, hence there is a $\varepsilon_i > 0$ such that $w_i \geq \varepsilon_i$ whenever γ lies in some ζ_i neighborhood of 0. Let $\varepsilon = \min_i \varepsilon_i$, $\zeta = \min_i \zeta_i$.

We now observe that if $L_{12} = 0$, $\tilde{V}_c^*(\eta)$ equals $\frac{\tilde{w}_2^*}{c_2} f_2 \log\left(\frac{c_2 \tilde{w}_1^* + c_1 \tilde{w}_2^*}{c_2 \tilde{w}_1^* \frac{f_1}{f_2} + c_1 \tilde{w}_2^*}\right)$ and if $L_{12} > 0$, $\tilde{V}_c^*(\eta)$ equals $\left(c_2 \tilde{w}_1^* f_1 \log\left(\frac{c_2 \tilde{w}_1^* f_1 + c_1 \tilde{w}_2^* f_1}{c_2 \tilde{w}_1^* f_1 + c_1 \tilde{w}_2^* f_2}\right) + c_1 \tilde{w}_2^* f_2 \log\left(\frac{c_2 \tilde{w}_1^* f_2 + c_1 \tilde{w}_2^* f_2}{c_2 \tilde{w}_1^* f_1 + c_1 \tilde{w}_2^* f_2}\right)\right)$. We have already seen in our discussion in the previous paragraph that $0 < \varepsilon_i \leq \tilde{w}_i^* \leq 1$ whenever $\gamma < \min\{\zeta_1, \zeta_2\}$. It therefore follows from $L_{12} = 0$ or $L_{12} > 0$, as the case may be, that for γ small enough, $\tilde{V}_c^*(\eta)$ can be bounded above and below by positive multiples of f_2, letting us conclude that $\tilde{V}_c^*(\eta) = \Theta(f_2)$. It can be inferred from Lemma 2 that the same conclusion holds for $V_c^*(\eta)$. The conclusion on $\mathbb{E}_\eta[\tau_{c,\delta}]$ follows from (2).

E PROOF OUTLINE OF LEMMA 5

We first observe that for $x,y \in [0,1]$ such that $x+y<1$, $\log\left(\frac{x+y}{x\frac{f_1}{f_i}+y}\right) \geq \varepsilon_x\left(1-\frac{f_1}{f_i}\right)$, assuming that $x \geq \varepsilon_x$ for some $\varepsilon_x > 0$. Similarly, assuming the existence of an ε_y, we can show that $\log\left(\frac{f_1x+f_1y}{f_1x+f_iy}\right) \geq \frac{f_i-f_1}{f_1}\varepsilon_y$. The intermediate steps in obtaining these inequalities use $\log(1+x) \leq x$, $\forall x$. Let $g(x,y) = yf_i\log\left(\frac{x+y}{x\frac{f_1}{f_i}+y}\right)$ and $h(x,y) = f_1x\log\left(\frac{f_1x+f_1y}{f_1x+f_iy}\right) + f_iy\log\left(\frac{f_ix+f_iy}{f_1x+f_iy}\right)$. We will use subscripts of x and y with these functions to represent partial derivatives with respect to x and y.

It can easily be checked using the inequalities from our observations, that $|g_y(x,y)| \geq \varepsilon_x\left(1-\frac{f_1}{f_i}\right)f_i$, $|h_x(x,y)| \geq \frac{f_i-f_1}{f_1}\varepsilon_y$, $|h_y(x,y)| \geq \frac{f_i-f_1}{f_1}\varepsilon_x$, $g_x(x,y) \geq 0$ and $h_x(x,y) \geq 0$. We can now show by using the mean value theorem, that $\left|c_1\varepsilon\left(1-\frac{f_1}{f_i}\right)f_i|w_i^*-\tilde{w}_i^*|-c_1f_1\right| \leq |V_c^*(\eta)-\tilde{V}_c^*(\eta)| \leq C_Uf_1$, letting us conclude that $|w_i^*-\tilde{w}_i^*| \leq S_0\frac{f_1}{f_i}$, whenever $\alpha_1 > \alpha_i$. This proves part (a).

To prove part (b), we use lower bounds on the partial derivatives of h, along with the mean value theorem. The steps are similar to the proof of part(a).

F PROOF OUTLINE OF THEOREM 1

(a) The proofs in scenarios $\mathbf{U}_1$,$\mathbf{U}_2$ and $\mathbf{N}$ are similar, so we only show the details for scenario $\mathbf{U}_1$. Let $f_{ij} := \frac{f_i}{f_j}$. Substituting the approximations of Lemma 1 in (4) gives us that for all j such that $L_{2j}=0$,

$$\tilde{w}_2^*f_2\log\left(\frac{\tilde{w}_1^*+\tilde{w}_2^*}{\tilde{w}_1^*f_{12}+\tilde{w}_2^*}\right) = \tilde{w}_j^*f_j\log\left(\frac{\tilde{w}_1^*+\tilde{w}_j^*}{\tilde{w}_1^*f_{1j}+\tilde{w}_j^*}\right)$$
$$\Rightarrow \tilde{w}_j^* = \tilde{w}_2^*f_{2j}\frac{\log\left(\frac{\tilde{w}_1^*+\tilde{w}_2^*}{\tilde{w}_1^*f_{12}+\tilde{w}_2^*}\right)}{\log\left(\frac{\tilde{w}_1^*+\tilde{w}_j^*}{\tilde{w}_1^*f_{1j}+\tilde{w}_j^*}\right)}.$$

Now, it is easy to observe that

$$\log\left(\frac{\tilde{w}_1^*+\tilde{w}_i^*}{\tilde{w}_1^*f_{1i}+\tilde{w}_i^*}\right) \leq \min\left\{\log\left(\frac{1}{\tilde{w}_i^*}\right), \log\left(\frac{1}{\tilde{w}_1^*f_{1i}}\right)\right\},$$
$$\log\left(\frac{\tilde{w}_1^*+\tilde{w}_2^*}{\tilde{w}_1^*f_{12}+\tilde{w}_2^*}\right) \geq \varepsilon_1(1-f_{12}),$$

the latter inequality following from $\log(1+x) \leq x$. We have also seen in the proof of Lemma 3 that $\dfrac{\tilde{w}_1^*f_1+\tilde{w}_i^*f_i}{\tilde{w}_1^*+\tilde{w}_i^*} + \Theta(f_1) \leq f_2 + \Theta(f_1) \quad \forall i \neq 1$, from which $\tilde{w}_i^* \leq \dfrac{\tilde{w}_1^*f_{2i}}{1-f_{2i}} + O(f_{1i})$ follows. This gives us that

$$\log\left(\frac{\tilde{w}_1^*+\tilde{w}_i^*}{\tilde{w}_1^*f_{1i}+\tilde{w}_i^*}\right) \geq \log\left(\frac{\tilde{w}_1^*}{\tilde{w}_1^*f_{1i}+\frac{\tilde{w}_1^*f_{2i}}{1-f_{2i}}+O(f_{1i})}\right)$$
$$= \log\left(\frac{\tilde{w}_1^*(1-f_{2i})}{\tilde{w}_1^*f_{2i}+\Theta(f_{1i})}\right).$$

Combining our conclusions from the above equation, $\tilde{w}_j^* = \Theta\left(\dfrac{f_2}{f_j\log\left(\frac{1}{f_{2j}}\right)}\right) \;\forall\; j \neq 1,2$ follows directly. Lemma 5 gives us the statement of part (a) as conclusion.

(b) We provide the outline of proof for $\mathbf{U}_1$, and the proofs for $\mathbf{U}_2$ and $\mathbf{N}$ are similar. Substituting the approximations of Lemma 1 and using Lemma 3 in (5) gives us

$$\frac{c_2 \frac{c_1 \tilde{w}_2^*}{c_2 \tilde{w}_1^* + c_1 \tilde{w}_2^*}}{c_1 \log\left(\frac{c_2 \tilde{w}_1^* + c_1 \tilde{w}_2^*}{c_2 \tilde{w}_1^* f_{12} + c_1 \tilde{w}_2^*}\right) - \frac{c_1 c_2 \tilde{w}_1^*}{c_2 \tilde{w}_1^* + c_1 \tilde{w}_2^*}} + \sum_{j=3}^{K} \frac{c_j \frac{c_1 \tilde{w}_j^*}{c_j \tilde{w}_1^* + c_1 \tilde{w}_j^*}}{c_1 \log\left(\frac{c_j \tilde{w}_1^* + c_1 \tilde{w}_j^*}{c_j \tilde{w}_1^* f_{1j} + c_1 \tilde{w}_j^*}\right) - \frac{c_1 c_j \tilde{w}_1^*}{c_j \tilde{w}_1^* + c_1 \tilde{w}_j^*}} = 1$$

From our conclusion in part (a), we can observe that

$$\frac{c_2 \frac{c_1 \tilde{w}_2^*}{c_2 \tilde{w}_1^* + c_1 \tilde{w}_2^*}}{c_1 \log\left(\frac{c_2 \tilde{w}_1^* + c_1 \tilde{w}_2^*}{c_2 \tilde{w}_1^* f_{12} + c_1 \tilde{w}_2^*}\right) - \frac{c_1 c_2 \tilde{w}_1^*}{c_2 \tilde{w}_1^* + c_1 \tilde{w}_2^*}} + \Theta\left(\max_{j:L_{2j}=0} \frac{f_2}{f_j \log(\frac{1}{f_{2j}})}\right) = 1.$$

Algebraic manipulations in the equations from (b)(i) will yield

$$\frac{c_2 \frac{c_1 \hat{w}_2}{c_2 \hat{w}_1 + c_1 \hat{w}_2}}{c_1 \log\left(\frac{c_2 \hat{w}_1 + c_1 \hat{w}_2}{c_2 \hat{w}_1 f_{12} + c_1 \hat{w}_2}\right) - \frac{c_1 c_2 \hat{w}_1}{c_2 \hat{w}_1 + c_1 \hat{w}_2}} \geq \frac{c_2 \frac{c_1 \hat{w}_2}{c_2 \hat{w}_1 + c_1 \hat{w}_2}}{c_1 \log\left(\frac{c_2 \hat{w}_1 + c_1 \hat{w}_2}{c_1 \hat{w}_2}\right) - \frac{c_1 c_2 \hat{w}_1}{c_2 \hat{w}_1 + c_1 \hat{w}_2}} = 1.$$

Let $g(x,y) = \frac{c_2 \frac{c_1 y}{c_2 x + c_1 y}}{c_1 \log\left(\frac{c_2 x + c_1 y}{c_2 x \frac{p_1}{p_2} \gamma^{\alpha_1 - \alpha_2} + c_1 y}\right) - \frac{c_1 c_2 x}{c_2 x + c_1 y}}$. Combining the two above equations, we conclude that

$$|g(\hat{w}_1, \hat{w}_2) - g(\tilde{w}_1^*, \tilde{w}_2^*)| = \Theta\left(\max_{j:L_{2j}=0} \frac{f_2}{f_j \log(\frac{1}{f_{2j}})}\right).$$

We can now bound the patial derivatives of f, and use the mean value theorem to conclude that $|\hat{w}_1 - \tilde{w}_1^*|$, $|\hat{w}_2 - \tilde{w}_2^*| = \Theta\left(\max_{j:L_{2j}=0} \frac{f_2}{f_j \log(\frac{1}{f_{2j}})}\right)$. We combine this with Lemma 5 to conclude that $|w_j^* - \hat{w}_j| \to 0$ as $\gamma \to 0$, that is $w_j^* \to \hat{w}_j$ as $\gamma \to 0$, for all j such that $L_{2j} > 0$ or $L_{1j} > 0$.

For scenario $\mathbf{U}_2$, we bound the partial derivatives of $g_i(x,y) = y f_2 \log\left(\frac{c_2 x + c_1 y}{c_2 x f_{1i} c_1 y}\right)$ and $h_i(x,y) = \frac{\frac{c_1 c_i y}{c_2 x + c_1 y}}{c_1 \log\left(\frac{c_i x + c_1 \hat{w}_j}{c_i x f_{1i} + c_1 y}\right) - \frac{c_1 c_i x}{c_i x + c_1 y}}$, $i = 2, ..., K$, and again use the mean value theorem to conclude that $w_j^* \to \hat{w}_j$ as $\gamma \to 0$, for all j such that $L_{2j} > 0$ or $L_{1j} > 0$.

For $\mathbf{N}$, we bound partial derivatives of $g_i(x,y) = c_i x f_1 \log\left(\frac{c_i x f_1 + c_1 y f_1}{c_i x f_1 + c_1 y f_i}\right) + c_1 y f_i \log\left(\frac{c_i x f_i + c_1 y f_i}{c_i x f_1 + c_1 y f_i}\right)$ and $h_i(x,y) = \frac{f_1 \log\left(\frac{c_i x f_1 + c_1 y f_1}{c_i x f_1 + c_1 y f_i}\right) - \frac{c_1 y (f_i - f_1)}{c_i x + c_1 y}}{f_i \log\left(\frac{c_i x f_i + c_1 y f_i}{c_i x f_1 + c_1 y f_i}\right) + \frac{c_i x (f_i - f_1)}{c_i x + c_1 y}}$, $i = 2, ..., K$ and once again use the mean value theorem to conclude that $w_j^* \to \hat{w}_j$ as $\gamma \to 0$, for all j such that $L_{2j} > 0$ or $L_{1j} > 0$.

REFERENCES

Agrawal, S., S. Juneja, and P. Glynn. 2020. "Optimal δ-Correct Best-Arm Selection for Heavy-Tailed Distributions". In *Proceedings of the 31st International Conference on Algorithmic Learning Theory*, edited by A. Kontorovich and G. Neu, 61–110. Columbia University, New York, New York, USA: Proceeedings of Machine Learning Research.

Agrawal, S., S. Juneja, K. Shanmugam, and A. Suggala. 2023. "Optimal Best-Arm Identification in Bandits with Access to Offline Data". *arXiv preprint arXiv:2306.09048*.

Agrawal, S., W. M. Koolen, and S. Juneja. 2021. "Optimal Best-Arm Identification Methods for Tail-Risk Measures". In *Advances in Neural Information Processing Systems*, edited by M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang, and J. W. Vaughan, Volume 34, 25578–25590. Red Hook, NY 12571 USA: Curran Associates, Inc.

Ahn, D. and T. Kim. 2023. "Risk-Sensitive Ordinal Optimization". In *2023 Winter Simulation Conference (WSC)*, 3364–3375 https://doi.org/10.1109/WSC60868.2023.10407133.

Batur, D. and F. Choobineh. 2010. "A Quantile-based Approach to System Selection". *European Journal of Operational Research* 202(3):764–772.

Bekki, J. M., J. W. Fowler, G. T. Mackulak, and B. L. Nelson. 2007. "Using Quantiles in Ranking and Selection Procedures". In *2007 Winter Simulation Conference*, 1722–1728 https://doi.org/10.1109/WSC.2007.4419795.

Bhattacharjee, A., S. Vijayan, and S. Juneja. 2023. "Best Arm Identification in Rare Events". In *Proceedings of the Thirty-Ninth Conference on Uncertainty in Artificial Intelligence*, edited by R. J. Evans and I. Shpitser, 163–172. Columbia University, New York, New York, USA: Proceeedings of Machine Learning Research.

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

Chernoff, H. 1959. "Sequential Design of Experiments". *The Annals of Mathematical Statistics* 30(3):755–770.

Even-Dar, E., S. Mannor, Y. Mansour, and S. Mahadevan. 2006. "Action Elimination and Stopping Conditions for the Multi-Armed Bandit and Reinforcement Learning Problems". *Journal of Machine Learning Research* 7(39):1079–1105.

Garivier, A. and E. Kaufmann. 2016. "Optimal Best Arm Identification with Fixed Confidence". In *29th Annual Conference on Learning Theory*, edited by V. Feldman, A. Rakhlin, and O. Shamir, 998–1027. Columbia University, New York, New York, USA: Proceeedings of Machine Learning Research.

Glynn, P. and S. Juneja. 2004. "A Large Deviations Perspective on Ordinal Optimization.". In *2004 Winter Simulation Conference (WSC)*, 577–585 https://doi.org/10.1109/WSC.2004.1371364.

Kalyanakrishnan, S., A. Tewari, P. Auer, and P. Stone. 2012. "PAC Subset Selection in Stochastic Multi-armed Bandits". In *Proceedings of the 29th International Coference on International Conference on Machine Learning*, ICML'12, 227–234. Madison, WI, USA: Omnipress.

Kaufmann, E. 2020. *Contributions to the Optimal Solution of Several Bandit Problems*. Ph.D. thesis, Université de Lille, Lille, Hauts-de-France. https://theses.hal.science/tel-03825097/file/HDR_EmilieKaufmann.pdf, accessed on September 23rd, 2024.

Kaufmann, E., O. Cappé, and A. Garivier. 2016. "On the Complexity of Best-arm Identification in Multi-armed Bandit Models". *Journal of Machine Learning Research* 17(1):1–42.

Kaufmann, E. and W. M. Koolen. 2021. "Mixture Martingales Revisited With Applications to Sequential Tests and Confidence Intervals". *Journal of Machine Learning Research* 22(246):1–44.

Kim, S.-H. and B. L. Nelson. 2006. "Selecting the Best System". In *Simulation*, edited by S. G. Henderson and B. L. Nelson, Volume 13 of *Handbooks in Operations Research and Management Science*, 501–534. Radarweg 29, 1043 NX Amsterdam, The Netherlands: Elsevier https://doi.org/https://doi.org/10.1016/S0927-0507(06)13017-0.

Paulson, E. 1964. "A Sequential Procedure for Selecting the Population with the Largest Mean from k Normal Populations". *The Annals of Mathematical Statistics* 35(1):174–180.

AUTHOR BIOGRAPHIES

ANIRBAN BHATTACHARJEE is a research associate and teaching fellow at the Centre for Data, Learning and Decision Sciences at Ashoka University. He is pursuing his Ph.D. from the Tata Institute of Fundamental Research, Mumbai, under Prof. Sandeep Juneja. His research interests are in applied probability, sequential learning and simulation methods. His email address is anirban.bhattacharjee1@ashoka.edu.in.

SANDEEP JUNEJA is a professor of computer science and the director for the Centre for Data, Learning and Decision Sciences at Ashoka University. His research interests lie in applied probability including in sequential learning, mathematical finance, Monte Carlo methods, game theoretic analysis of queues and epidemiological modelling. He is currently the area editor for Operations Research in simulation. Earlier, he has been on editorial boards of Stochastic Systems, Mathematics of Operations Research, Management Science, and ACM TOMACS. His email address is sandeep.juneja@ashoka.edu.in, and his web page is https://www.tcs.tifr.res.in/~sandeepj/.