A Proof of Lemma 1

Proof. Let $\bar{t}$ be the first time when the condition in equation (2) is met, then at time $\bar{t}-1$ we have that

$$\widehat{\Delta}_{V,U}(\bar{t}-1) - \sum_{i \in U \oplus V} \beta_i(\bar{t}-1) \leq 0,$$

which implies that

$$\Delta_{V,U} - 2 \sum_{i \in U \oplus V} \beta_i(\bar{t}-1) \leq 0,$$

and thus

$$\sum_{i \in U \oplus V} \sqrt{\frac{\log \frac{4K'(\bar{t}-1)^2}{\delta}}{2T_i(\bar{t}-1)}} \geq \frac{\Delta_{V,U}}{2}.$$

Since the algorithm is selecting arms in $U \oplus V$ using a round-robin strategy, at any time $T_i(t) = T_j(t) \pm 1$ for any pair of arms i, j. Thus, let j be the least pulled arm at round $\bar{t}-1$ (i.e., $j \in \arg\min T_i(\bar{t}-1)$), then the previous inequality can be written as

$$\overline{d}_{U,V} \sqrt{\frac{\log \frac{4K'(\bar{t}-1)^2}{\delta}}{2T_j(\bar{t}-1)}} \geq \frac{\Delta_{V,U}}{2},$$

which leads to the statement. □

B The Complexity in Equation (4) is Well-defined

In order to prove that our complexity measure in equation (4) is well-defined, we have to show that the $\max$ operator actually returns a value, that is, that there exists at least one element in the argument of the $\max$ operator. This is done in the following proposition.

Proposition 1. *Let Q_i be the set of the decision sets in the argument of the* $\max$ *operator in equation* (4)*, i.e.,*

$$Q_i = \{U \in \mathcal{C} : i \in U \oplus \mathsf{C}_U\}. \tag{10}$$

Then $Q_i \neq \varnothing$ for all i in $\mathcal{K}$.

Proof. We distinguish the following two cases:

Case 1) $i \notin U^*$

According to the assumption we made in Section 2, there exists a decision set $V \in \mathcal{C}$ such that $i \in V$. Note that i does not necessarily belong to $V \oplus \mathsf{C}_V$. We construct a sequence of decision sets $\mathcal{X} = \{X_1, \ldots, X_p\}$ such that $X_1 = V$, for all $j \in \{1, \ldots, p-1\}$, $i \in X_j$ and $X_{j+1} = \mathsf{C}_{X_j}$, and $i \notin X_p$.[8] As a result, setting $U = X_{p-1}$ and $\mathsf{C}_U = X_p$, we have that $i \in U \oplus \mathsf{C}_U$, and thus, $Q_i \neq \varnothing$.

Case 2) $i \in U^*$

Let $V = \underset{U \in \mathcal{C}\,:\, i \notin U}{\arg\max}\, \mu_U$. Then C_V exists, because $i \in U^*$. Moreover $i \in \mathsf{C}_V$, as otherwise we obtain the contradiction $\mu_{\mathsf{C}_V} > \mu_V = \max_{U \in \mathcal{C}\,:\, i \notin U} \mu_U \geq \mu_{\mathsf{C}_V}$. Therefore, $i \in \mathsf{C}_V \backslash V$, so that $V \in Q_i$ and $Q_i \neq \varnothing$. □

[8] Note that $\mathcal{X}$ is not only finite but also contains a decision set X_p, such that $i \notin X_p$. The first claim comes from the definition of complement that gives us $\mu_{X_{j+1}} > \mu_{X_j},\ \forall j \in \{1, \ldots, p-1\}$, and the fact that $\mathcal{C}$ is finite. For the second claim note that as $i \notin U^*$, $\mathcal{X}$ has at least two elements.

C Useful Properties of the Complexity $H_{U,V}$

In this appendix, we prove several useful properties of the complexity $H_{U,V}$ later, particularly in the proof that our complexity measure is not higher than the measure used by Chen et al. [2014].

The first proposition shows that the distance $\overline{d}_{U,V}$ between two decision sets U and V follows a triangle inequality.

Proposition 2. *For any three distinct decision sets $U, V, W \in \mathcal{C}$, we have, $U \oplus V \subseteq (U \oplus W) \cup (W \oplus V)$ and $\overline{d}_{U,V} \leq \overline{d}_{U,W} + \overline{d}_{W,V}$. Moreover, if $(U \oplus W) \cap (W \oplus V) = \varnothing$ then $U \oplus V = (U \oplus W) \cup (W \oplus V)$ and $\overline{d}_{U,V} = \overline{d}_{U,W} + \overline{d}_{W,V}$.*

Proof. To prove the statements for $U \oplus V$, we first prove that

$$\begin{aligned} U\backslash V &= \big((U\backslash V)\backslash W\big) \cup \big((U\backslash V) \cap W\big) \\ &= \Big((U\backslash W)\backslash\big((U\backslash W) \cap (V\backslash W)\big)\Big) \cup \Big((W\backslash V)\backslash\big((W\backslash V) \cap (W\backslash U)\big)\Big) \quad (11) \\ &\subseteq (U\backslash W) \cup (W\backslash V). \quad (12) \end{aligned}$$

Similar to (12), we may prove

$$V\backslash U \subseteq (V\backslash W) \cup (W\backslash U). \quad (13)$$

Taking the union of (12) and (13) gives

$$(U\backslash V) \cup (V\backslash U) \subseteq (U\backslash W) \cup (W\backslash V) \cup (V\backslash W) \cup (W\backslash U),$$

and the first claim of the proposition

$$U \oplus V \subseteq (U \oplus W) \cup (W \oplus V). \quad (14)$$

follows by definition of $\oplus$. The second claim is straightforward from (14) and the definition of symmetric distance $\overline{d}$, i.e.,

$$\overline{d}_{U,V} \leq \overline{d}_{U,W} + \overline{d}_{W,V}.$$

To prove the second part of the proposition, we start from (11), that is,

$$U\backslash V = \Big((U\backslash W)\backslash \underbrace{\big(\overbrace{(U\backslash W)}^{X} \cap \overbrace{(V\backslash W)}^{Y}\big)}_{A}\Big) \cup \Big((W\backslash V)\backslash \underbrace{\big(\overbrace{(W\backslash V)}^{S} \cap \overbrace{(W\backslash U)}^{Z}\big)}_{B}\Big),$$

from which by our assumption $(U \oplus W) \cap (W \oplus V) = \varnothing$ we obtain

$$A \cup B = (X \cup S) \cap \underbrace{\overbrace{(X \cup Z)}^{U\oplus W} \cap \overbrace{(Y \cup S)}^{W\oplus V}}_{\varnothing} \cap (Y \cup Z) = \varnothing \implies A = \varnothing \text{ and } B = \varnothing. \quad (15)$$

From (15), (C) may be written as

$$U\backslash V = (U\backslash W) \cup (W\backslash V). \quad (16)$$

Similarly, one can show that

$$V\backslash U = (V\backslash W) \cup (W\backslash U). \quad (17)$$

Taking union from both sides of (16) and (17), we obtain

$$U \oplus V = (U \oplus W) \cup (W \oplus V),$$

and as a result $\overline{d}_{U,V} = \overline{d}_{U,W} + \overline{d}_{W,V}$, which completes the proof of the second part of the proposition. □

The next proposition proves useful properties for the complexity of two decision sets.

Proposition 3. *For any three decision sets $U, V, W \in \mathcal{C}$ with $\mu_U < \mu_V < \mu_W$, we have*

$$H_{U,W} \leq \max\left(H_{U,V}, H_{V,W}\right). \quad (18)$$

Furthermore, if $(U \oplus V) \cap (V \oplus W) = \varnothing$, then

$$H_{U,W} \geq \min\left(H_{U,V}, H_{V,W}\right), \quad (19)$$

and finally, the above two inequalities are strict if $H_{U,V} \neq H_{V,W}$.

Proof. We write

$$\Delta_{W,U} = \mu_W - \mu_U = \mu_W - \mu_V + \mu_V - \mu_U = \Delta_{W,V} + \Delta_{V,U}. \tag{20}$$

By assumption, we have $\mu_W > \mu_V$ and $\mu_V > \mu_U$, and thus, $\Delta_{W,V} > 0$ and $\Delta_{V,U} > 0$. As a result, we may write

$$\frac{\overline{d}_{U,W}}{\Delta_{W,U}} \overset{(a)}{\leq} \frac{\overline{d}_{U,V} + \overline{d}_{V,W}}{\Delta_{W,V} + \Delta_{V,U}} \overset{(b)}{\leq} \max\left(\frac{\overline{d}_{U,V}}{\Delta_{V,U}}, \frac{\overline{d}_{V,W}}{\Delta_{W,V}}\right), \tag{21}$$

where **(a)** follows from Proposition 2 and (20), and **(b)** follows from the fact that for any positive $a, b, c, d \geq 0$, it holds that $\frac{a+b}{c+d} \leq \max\left(\frac{a}{c}, \frac{b}{d}\right)$.[9] From (21), we get

$$\frac{\overline{d}^2_{U,W}}{\Delta^2_{W,U}} \leq \max\left(\frac{\overline{d}^2_{U,V}}{\Delta^2_{V,U}}, \frac{\overline{d}^2_{V,W}}{\Delta^2_{W,V}}\right),$$

which gives us (18).

The second statement is similarly proved as

$$\frac{\overline{d}_{U,W}}{\Delta_{W,U}} \overset{(a)}{=} \frac{\overline{d}_{U,V} + \overline{d}_{V,W}}{\Delta_{W,V} + \Delta_{V,U}} \overset{(b)}{\geq} \min\left(\frac{\overline{d}_{U,V}}{\Delta_{V,U}}, \frac{\overline{d}_{V,W}}{\Delta_{W,V}}\right), \tag{22}$$

where **(a)** is true under the assumption that $(V \oplus U) \cap (V \oplus W) = \varnothing$ from Proposition 2, and **(b)** follows from the fact that for any positive $a, b, c, d \geq 0$, it holds that $\min\left(\frac{a}{c}, \frac{b}{d}\right) \leq \frac{a+b}{c+d}$.[10] From (22) it follows that

$$\frac{\overline{d}^2_{U,W}}{\Delta^2_{W,U}} \geq \min\left(\frac{\overline{d}^2_{U,V}}{\Delta^2_{V,U}}, \frac{\overline{d}^2_{V,W}}{\Delta^2_{W,V}}\right),$$

which gives us (19).

The proof of the very last statement, the strict inequalities, comes directly from the fact that when $\frac{a}{c} \neq \frac{b}{d}$, the two inequalities at step **(b)** of (21) and (22) become strict. $\square$

In the last proposition of this section, we show that the complexity of discriminating between a decision set $U \neq U^*$ and its complement $V = C_U$ is less than the complexity of discriminating between $V = C_U$ and its complement $W = C_V = C_{C_U}$, provided that the complement of U is not the best decision set U^*, i.e., $V = C_U \neq U^*$.

Proposition 4. *For any decision set $U \neq U^*$ with $V = \mathsf{C}_U \neq U^*$ and $W = \mathsf{C}_V = C_{C_U}$, it holds that $H_{U,V} \leq H_{V,W}$.*

Proof. We prove the statement by contradiction. Let us assume that $H_{U,V} > H_{V,W}$. Since $V \neq U^*$ and by definition of complement, we have $\mu_U < \mu_V < \mu_W$. As a result, $H_{U,W} \leq \max\left(H_{U,V}, H_{V,W}\right)$ from Proposition 3. Note that this inequality is strict, whenever $H_{U,V} \neq H_{V,W}$, again according to Proposition 3, and since we assumed that $H_{U,V} > H_{V,W}$, we have $H_{U,W} < H_{U,V}$. This gives us

$$H_{U,V} = H_{U,\mathsf{C}_U} = \min_{Z \in \mathcal{C}: \mu_Z > \mu_U} H_{U,Z} \leq H_{U,W} < H_{U,V},$$

which leads to the contradiction that $H_{U,V} < H_{U,V}$. $\square$

D Equivalence of the Different Notions of Arm Complexity

In this section, we give two alternative notions of complexity of an arm that are equivalent to the original definition H_i of equation 4. In the analysis of the algorithms (see Appendices G and H) we will use the definition of the complexity that is the most handy. The equivalence proof requires the results of Appendix C, especially Proposition 4. We start with the definition of the alternative complexity notions and two intermediate results that will be needed for the equivalence proof given at the end of this section.

[9] Here is the proof: Assume without loss of generality that $\frac{a}{c} \leq \frac{b}{d}$. Then $\frac{a+b}{c+d} \leq \frac{bc/d+b}{c+d} = \frac{b(c/d+1)}{d(c/d+1)} = \frac{b}{d} = \max\left(\frac{a}{c}, \frac{b}{d}\right)$.

[10] The proof is analogous to the previous footnote: Assume without loss of generality that $\frac{a}{c} \leq \frac{b}{d}$. Then $\frac{a+b}{c+d} \geq \frac{a+da/c}{c+d} = \frac{a(d/c+1)}{c(d/c+1)} = \frac{a}{c} = \min\left(\frac{a}{c}, \frac{b}{d}\right)$.

Definition 7. *Our two notions of complexity for an arm $i \in \mathcal{K}$, $\mathcal{H}^1$ and $\mathcal{H}^2$, are defined as*

$$\mathcal{H}_i^1 = \begin{cases} \max\limits_{U \in \mathcal{C}: i \in U \setminus \mathsf{C}_U} H_{U,\mathsf{C}_U} & \textit{if } i \notin U^*, \\ \max\limits_{U \in \mathcal{C}: i \in \mathsf{C}_U \setminus U} H_{U,\mathsf{C}_U} & \textit{if } i \in U^*. \end{cases} \qquad \mathcal{H}_i^2 = \begin{cases} \max\limits_{U \in \mathcal{C}: i \in U} H_{U,\mathsf{C}_U} & \textit{if } i \notin U^*, \\ \max\limits_{U \in \mathcal{C}: i \notin U} H_{U,\mathsf{C}_U} & \textit{if } i \in U^*. \end{cases}$$

The following proposition plays an important role in proving the equivalence $\mathcal{H}_i^1 = H_i, \forall i \in \mathcal{K}$. It shows that if $i \in (U \oplus \mathsf{C}_U)$, then $\mathcal{H}_i^1 \geq H_{U,\mathsf{C}_U}$.

Proposition 5. *For any decision set $U \in \mathcal{C}$ such that $U \neq U^*$, and any arm $i \in (U \oplus \mathsf{C}_U)$, we have $\mathcal{H}_i^1 \geq H_{U,\mathsf{C}_U}$.*

Proof. We consider the following two cases for a fixed arm $i \in (U \oplus \mathsf{C}_U)$:

Case 1) $i \in U^*$

If $i \in \mathsf{C}_U \setminus U$, the result follows directly from the definition of $\mathcal{H}_i^1$. If $i \in U \setminus \mathsf{C}_U$, similar to the proof of Proposition 1 in Appendix B, we construct a sequence of decision sets $\{X_1, \dots, X_p\}$ such that $X_1 = U$, for all $j \in \{2, \dots, p-1\}$, $i \notin X_j$ and $X_{j+1} = \mathsf{C}_{X_j}$, and $i \in X_p$. Note that $\{X_2, \dots, X_p\}$ is a sequence of decision sets and their complements that do not contain arm i until a set X_p is generated that contains i.[11] From Proposition 4, we have that $H_{X_j,X_{j+1}} \geq H_{X_{j-1},X_j}, \forall j \in \{2, \dots, p-1\}$. Now starting from the definition of $\mathcal{H}_i^1$, we may write

$$\mathcal{H}_i^1 = \max_{Z: i \in \mathsf{C}_Z \setminus Z} H_{Z,\mathsf{C}_Z} \overset{\text{(a)}}{\geq} H_{X_{p-1},X_p} \geq H_{X_1,X_2} = H_{U,\mathsf{C}_U},$$

which proves the claim of the proposition. Note that **(a)** comes from the fact that $i \in \mathsf{C}_{X_{p-1}} \setminus X_{p-1}$ by definition of the sequence.

Case 2) $i \notin U^*$

If $i \in U \setminus \mathsf{C}_U$, the result follows directly from the definition of $\mathcal{H}_i^1$. When $i \in \mathsf{C}_U \setminus U$, we construct a sequence of decision sets $\{X_1, \dots, X_p\}$ such that $X_1 = U$, for all $j \in \{2, \dots, p-1\}$, $i \in X_j$ and $X_{j+1} = \mathsf{C}_{X_j}$, and $i \notin X_p$. This is a sequence of decision sets and their complements that contain arm i until a set X_p is generated that does not contain i. As a result $i \in X_{p-1} \setminus \mathsf{C}_{X_{p-1}}$. From Proposition 4, we have that $H_{X_j,X_{j+1}} \geq H_{X_{j-1},X_j}, \forall j \in \{2, \dots, p-1\}$. Now starting from the definition of $\mathcal{H}_i^1$, we may write

$$\mathcal{H}_i^1 = \max_{Z: i \in Z \setminus \mathsf{C}_Z} H_{Z,\mathsf{C}_Z} \geq H_{X_{p-1},X_p} \geq H_{X_1,X_2} = H_{U,\mathsf{C}_U}, \tag{23}$$

which proves the claim of the proposition. □

Proposition 6. *For any decision set $U \in \mathcal{C}$ such that $U \neq U^*$, and any arm $i \in (U \oplus U^*)$, we have $\mathcal{H}_i^1 \geq H_{U,\mathsf{C}_U}$.*

Proof. Let us construct a sequence of decision sets $\{X_1, \dots, X_p\}$ such that $X_1 = U$, for all $j \in \{2, \dots, p-1\}$, $X_{j+1} = \mathsf{C}_{X_j}$, and $X_p = U^*$. This sequence is well-defined and has at least two elements, since $U \neq U^*$ and U^* is unique. If we prove that for any $j \in \{2, \dots, p\}$, we have that for all $i \in (X_1 \oplus X_j)$, $\mathcal{H}_i^1 \geq H_{X_1,X_2}$, then $j = p$ will gives us the proof of the proposition. Now let us prove this statement. The proof is by induction on j.

Base Step: $j = 2$. In this case, the claim follows directly from Proposition 5.

Inductive Step: Here we assume that for $j = j'$, we have that $\mathcal{H}_i^1 \geq H_{X_1,X_2}, \forall i \in (X_1 \oplus X_{j'})$, and we want to show that $\mathcal{H}_i^1 \geq H_{X_1,X_2}, \forall i \in (X_1 \oplus X_{j'+1})$. From Proposition 5 and the construction of the sequence, we have $\mathcal{H}_i^1 \geq H_{X_{j'},X_{j'+1}}, \forall i \in (X_{j'}, X_{j'+1})$. By repeated application of Proposition 4, we can show that $\forall i \in (X_1 \oplus X_{j'}) \cup (X_{j'} \oplus X_{j'+1})$, we have $\mathcal{H}_i^1 \geq \min(H_{X_1,X_{j'}}, H_{X_{j'},X_{j'+1}}) \geq H_{X_1,X_2}$. Moreover, from Proposition 2, we know that $X_1 \oplus X_{j'+1} \subseteq (X_1 \oplus X_{j'}) \cup (X_{j'} \oplus X_{j'+1})$, and thus, we obtain $\forall i \in (X_1 \oplus X_{j'+1})$ that $\mathcal{H}_i^1 \geq H_{X_1,X_2}$, which proves the inductive step. □

We are now ready to prove the main result of this section, the equivalence of the different notions of arm complexity.

Lemma 3. *For any arm $i \in \mathcal{K}$, we have $H_i = \mathcal{H}_i^1 = \mathcal{H}_i^2$.*

[11] Note that such a sequence is finite, because by the definition of complement of a decision set, we have $\mu_{X_{j+1}} > \mu_{X_j}$, the number of decision sets is finite, and $i \in U^*$.

Proof. **Step 1:** We first prove that $H_i = \mathcal{H}_i^1, \forall i \in \mathcal{K}$.

From the definition of $\mathcal{H}_i^1$, it is immediate to see that $\mathcal{H}_i^1 \leq \max_{U \in \mathcal{C}: i \in U \oplus \mathsf{C}_U} H_{U,\mathsf{C}_U} = H_i$, and from Proposition 5, we may write $H_i = \max_{U \in \mathcal{C}: i \in U \oplus \mathsf{C}_U} H_{U,\mathsf{C}_U} \leq \mathcal{H}_i^1$. These together prove Step 1.

Step 2: We now want to prove $\mathcal{H}_i^1 = \mathcal{H}_i^2, \forall i \in \mathcal{K}$.

From the definitions of $\mathcal{H}_i^1$ and $\mathcal{H}_i^2$, it is immediate to write $\mathcal{H}_i^1 \leq \mathcal{H}_i^2$. To prove the reverse, we consider the following two cases:

Case 1) $i \notin U^*$

In this case, we may write

$$\mathcal{H}_i^2 = \max_{U \in \mathcal{C}: i \in U} H_{U,\mathsf{C}_U} = \max_{U \in \mathcal{C}: i \in (U \oplus U^*)} H_{U,\mathsf{C}_U} \overset{\text{(a)}}{\leq} \mathcal{H}_i^1,$$

where **(a)** is from Proposition 6.

Case 2) $i \in U^*$

In this case, we may write

$$\mathcal{H}_i^2 = \max_{U \in \mathcal{C}: i \notin U} H_{U,\mathsf{C}_U} = \max_{U \in \mathcal{C}: i \in (U \oplus U^*)} H_{U,\mathsf{C}_U} \overset{\text{(a)}}{\leq} \mathcal{H}_i^1,$$

where **(a)** is from Proposition 6.

The two cases together prove Step 2. □

E Proof of Lemma 2

Let $U \in \mathcal{C}$ be a decision set with complement C_U. Let b be an exchange set that satisfies constraints **(b)–(e)** of Definition 2 for the decision set pair (U, C_U). Let $V = U \pm b$ be the decision set resulted from applying the transformation b to U. We now define the exchange set $d = (C_U \backslash V, V \backslash C_U)$ as the exchange set that completes the transformation of U to C_U after applying b to U. It is easy to show that $d = \big((C_U \backslash U) \backslash b_+, (U \backslash C_U) \backslash b_-\big)$. We now prove the following two propositions that are used in the proof of Lemma 2.

Proposition 7. *For any decision set $U \in \mathcal{C}$, any exchange set b that satisfies constraints **(b)–(e)** of Definition 2 for the decision set pair (U, C_U), and any exchange set d that completes the transformation of U to C_U after applying b to U, i.e., $d = \big((C_U \backslash U) \backslash b_+, (U \backslash C_U) \backslash b_-\big)$, we have*

$$\Delta_{C_U,U} = \Delta_{b_+,b_-} + \Delta_{d_+,d_-} > 0, \tag{24}$$

$$\overline{d}_{U,C_U} = \overline{d}_{b_+,b_-} + \overline{d}_{d_+,d_-}, \tag{25}$$

so that $H_{U,C_U} = \frac{(\overline{d}_{b_+,b_-} + \overline{d}_{d_+,d_-})^2}{(\Delta_{b_+,b_-} + \Delta_{d_+,d_-})^2}$.

Proof. We begin with the proof of (24). By definition of C_U, $\mu_{C_U} > \mu_U$, so that Δ_{b_+,b_-} and Δ_{d_+,d_-} cannot be both negative. Now to prove the equality, first note that $\mu_{d_+} = \mu_{C_U \backslash U} - \mu_{b_+}$ and $\mu_{d_-} = \mu_{U \backslash C_U} - \mu_{b_-}$ from the definition of d and the fact that $b_+ \subseteq C_U \backslash U$. Further we have $b_- \subseteq U \backslash C_U$ from constraints (b) and (c) of Definition 2. Therefore,

$$\Delta_{C_U,U} = \mu_{C_U} - \mu_U = \mu_{C_U \backslash U} - \mu_{U \backslash C_U} = \underbrace{\big(\mu_{C_U \backslash U} - \mu_{b_+}\big)}_{\mu_{d_+}} + \mu_{b_+} - \underbrace{\big(\mu_{U \backslash C_U} - \mu_{b_-}\big)}_{\mu_{d_-}} - \mu_{b_-} = \Delta_{b_+,b_-} + \Delta_{d_+,d_-},$$

which proves (24).

Now let us turn to showing (25):

$$\begin{aligned}
\overline{d}_{b_+,b_-} + \overline{d}_{d_+,d_-} = |b_+ \oplus b_-| + |d_+ \oplus d_-| &\overset{\text{(a)}}{=} |b_+| + |b_-| + |d_+| + |d_-| \\
&\overset{\text{(b)}}{=} |b_+| + |b_-| + |C_U \backslash U \backslash b_+| + |U \backslash C_U \backslash b_-| \\
&\overset{\text{(c)}}{=} |b_+| + |b_-| + |C_U \backslash U| - |b_+| + |U \backslash C_U| - |b_-| = |U \oplus C_U| = \overline{d}_{U,C_U},
\end{aligned}$$

where **(a)** comes from the fact that $b_+ \cap b_- = d_+ \cap d_- = \varnothing$, **(b)** is from the definition of d, and **(c)** follows from constraints (b) and (c) of Definition 2. $\square$

We are now ready to prove Lemma 2.

Proof of Lemma 2. The proof is by contradiction. Suppose $U \not\perp \mathsf{C}_U$. Since independence is symmetric, we also have $\mathsf{C}_U \not\perp U$. This means that there exists a non-empty exchange set $b = (b_+, b_-)$, different than the independent exchange set $(C_U \backslash U, U \backslash C_U)$ of (U, C_U), that satisfies constraints **(b)–(e)** of Definition 2. From the exchange set b, we define the exchange set d that completes the transformation of U to C_U after applying b to U as $d = \big((C_U \backslash U)\backslash b_+, (U \backslash C_U)\backslash b_-\big)$.

Since b satisfies constraints (b) and (c), we may write

$$\mu_{C_U \mp b} = \mu_{C_U} - \mu_{b_+} + \mu_{b_-} = \mu_{C_U} - \Delta_{b_+,b_-},$$

which gives

$$\Delta_{C_U \mp b, U} = \mu_{C_U \mp b} - \mu_U = \Delta_{C_U,U} - \Delta_{b_+,b_-}. \tag{26}$$

Since b is not empty, $C_U \mp b$ is closer to U than C_U, and hence, $\overline{d}_{U,C_U \mp b} < \overline{d}_{U,C_U}$. Now consider the following three cases (note that as shown in Proposition 7, Δ_{b_+,b_-} and Δ_{d_+,d_-} cannot be both negative):

Case 1) $\Delta_{b_+,b_-} \leq 0$

In this case, by (26) we may write

$$H_{U,C_U \mp b} = \frac{\overline{d}^2_{U,C_U \mp b}}{\Delta^2_{U,C_U \mp b}} = \frac{\overline{d}^2_{U,C_U \mp b}}{(\Delta_{U,C_U} - \Delta_{b_+,b_-})^2} \leq \frac{\overline{d}^2_{U,C_U \mp b}}{\Delta^2_{U,C_U}} \overset{\text{(a)}}{<} \frac{\overline{d}^2_{U,C_U}}{\Delta^2_{U,C_U}} \overset{\text{(b)}}{=} \min_{V \in \mathcal{C}: \mu_V > \mu_U} H_{U,V}, \tag{27}$$

where **(a)** comes from the fact that $\overline{d}_{U,C_U \mp b} < \overline{d}_{U,C_U}$ and **(b)** is from the definition of the complement C_U. Moreover, in this case, from (26), we have $\Delta_{C_U,U} \leq \Delta_{C_U \mp b,U}$, which gives us $\mu_{C_U} \leq \mu_{C_U \mp b}$. Since $\mu_U < \mu_{C_U}$ by definition, we have that $(C_U \mp b) \in \{V \in \mathcal{C} : \mu_V > \mu_U\}$ and hence $H_{U,C_U \mp b} \geq H_{U,C_U}$ by definition of C_U, which contradicts equation (27).

Case 2) $\Delta_{b_+,b_-} > 0$ and $\Delta_{d_+,d_-} \leq 0$

Here we first show that

$$\mu_{U \pm b} = \mu_{U \backslash b_- \cup b_+} \overset{\text{(a)}}{=} \mu_U - \mu_{b_-} + \mu_{b_+} = \mu_U + \Delta_{b_+,b_-},$$

where **(a)** comes from constraints (b) and (c) of Definition 2, which gives us

$$\Delta_{U \pm b, U} = \mu_{U \pm b} - \mu_U = \Delta_{b_+,b_-}. \tag{28}$$

It is also straightforward to see that

$$\overline{d}_{U,U \pm b} = |U \oplus U \pm b| = \overline{d}_{b_+,b_-}. \tag{29}$$

Now similar to (27), we may write

$$H_{U,U \pm b} = \frac{\overline{d}^2_{U,U \pm b}}{\Delta^2_{U,U \pm b}} \overset{\text{(a)}}{=} \frac{\overline{d}^2_{b_+,b_-}}{\Delta^2_{b_+,b_-}} \overset{\text{(b)}}{<} \frac{(\overline{d}_{b_+,b_-} + \overline{d}_{d_+,d_-})^2}{(\Delta_{b_+,b_-} + \Delta_{d_+,d_-})^2} \overset{\text{(c)}}{=} H_{U,C_U} = \min_{V \in \mathcal{C}: \mu_V > \mu_U} H_{U,V}, \tag{30}$$

where **(a)** comes from (28) and (29), **(b)** is from the fact that $\overline{d}_{d_+,d_-} > 0$ and $\Delta_{d_+,d_-} \leq 0$, and finally **(c)** is from Proposition 7. Moreover, since $\Delta_{b_+,b_-} > 0$, from (28) we have $\mu_U < \mu_{U \pm b}$, which means that $(U \pm b) \in \{V \in \mathcal{C} : \mu_V > \mu_U\}$, and thus, $H_{U,U \pm b}$ should be bigger than or equal to H_{U,C_U}, which contradicts equation (30).

Case 3) $\Delta_{b_+,b_-} > 0$ and $\Delta_{d_+,d_-} > 0$

From Proposition 7, we have

$$H_{U,C_U} = \frac{(\overline{d}_{b_+,b_-} + \overline{d}_{d_+,d_-})^2}{(\Delta_{b_+,b_-} + \Delta_{d_+,d_-})^2} \overset{\text{(a)}}{\geq} \min\left(\frac{\overline{d}^2_{b_+,b_-}}{\Delta^2_{b_+,b_-}}, \frac{\overline{d}^2_{d_+,d_-}}{\Delta^2_{d_+,d_-}}\right) = \min(H_{b_+,b_-}, H_{d_+,d_-}), \tag{31}$$

where **(a)** comes from footnote 10 Appendix C. The inequality in (31) is strict whenever $H_{b_+,b_-} \neq H_{d_+,d_-}$. We now consider the following three cases that all end up contradicting that $\mathsf{C}_U = \arg\min_{V \in \mathcal{C}: \mu_V > \mu_u} H_{U,V}$.

Case 3.1) $H_{b_+,b_-} < H_{d_+,d_-}$

From equation 31, we have $H_{U,C_U} > H_{b_+,b_-} \overset{\text{(a)}}{=} H_{U\pm b,U}$, where **(a)** is from (28) and (29). At the same time we have $H_{U,C_U} = \arg\min_{V:\mu_V > \mu_U} H_{U,V} \overset{\text{(b)}}{\leq} H_{U\pm b,U}$, which leads to a contradiction. Note that **(b)** holds because $\Delta_{b_+,b_-} > 0$, and thus, $\mu_U < \mu_{U\pm b}$ from (28).

Case 3.2) $H_{b_+,b_-} > H_{d_+,d_-}$

From equation 31, we have $H_{U,C_U} > H_{d_+,d_-}$. Since $d_- \subseteq U$ and $d_+ \cap U = \varnothing$ from the definition of d, we may write

$$\Delta_{U\pm d,U} = \mu_{U\pm d} - \mu_U = \mu_U - \mu_{d_-} + \mu_{d_+} - \mu_U = \Delta_{d_+,d_-},$$

and

$$\overline{d}_{U\pm d,U} = \overline{d}_{d_+,d_-},$$

which gives us $H_{U,C_U} > H_{d_+,d_-} = H_{U\pm d,U}$. At the same time $H_{U,C_U} = \arg\min_{V:\mu_V > \mu_U} H_{U,V} \leq H_{U\pm d,U}$, which leads to a contradiction.

Case 3.3) $H_{b_+,b_-} = H_{d_+,d_-}$

From equation 31, we have $H_{U,C_U} \geq H_{b_+,b_-}$. If the inequality is strict, then we obtain the contradiction as in Case 3.1. Thus let us assume that $H_{U,C_U} = H_{b_+,b_-}$. From (28) and (29), we have $H_{U,C_U} = H_{b_+,b_-} = H_{U,U\pm b}$, and from (28) and the fact that $\Delta_{b_+,b_-} > 0$, we have $\mu_{U\pm b} > \mu_U$, and from (25) and (29), we have $\overline{d}_{U\pm b,U} < \overline{d}_{U,C_U}$. This creates a contradiction because in case $H_{U,C_U} = H_{U,U\pm b}$, according to Definition 4, the tie should be broken in favor of the set with the smaller symmetric distance, and thus, C_U should be $U \pm b$. □

F Proof of Theorem 3

We begin this section with the definition of the $*$-complement of a decision set. For this, let $Q(U)$ be the set of decision sets V such that $U \perp V$ and the exchange set $b = (V\backslash U, U\backslash V)$ satisfies constraints **(b)–(e)** of Definition 2 for the pair of decision sets (U, U^*).[12]

Definition 8. *The $*$-complement of a decision set $U \in \mathcal{C}$ with $U \neq U^*$, denoted by C_U^*, is defined as*

$$\mathsf{C}_U^* = \arg\min_{V \in Q(U)} H_{U,V}.$$

We have to show that the argument of the *argmin* in Definition 8 is not empty, i.e., $Q(U) \neq \varnothing$. For this purpose, we build a sequence of decision sets $\{V_1, \ldots, V_p\}$ such that $V_1, \ldots, V_{p-1}$ are all not independent of U and $U \perp V_p$, that is, we stop the sequence as soon as we reach a decision set V_p independent from U. To construct such a sequence, we start with $V_1 = U^*$ and for $k \in \{1, \ldots, p-1\}$, we generate $V_{k+1} = U \pm b_{k+1}$, where $b_1 = (b_{1,+}, b_{1,-}) = (U^*\backslash U, U\backslash U^*)$ and $b_{k+1} = (b_{k+1,+}, b_{k+1,-}) \subset b_k = (b_{k,+}, b_{k,-})$ is an exchange set that satisfies constraints (b)-(e) of Definition 2 for the pair of decision sets (U, V_k). Note that b_{k+1} exists by definition as V_k is not independent of U and this is why we can build iteratively the sequence $\{V_1, \ldots, V_p\}$ until we find V_p with $U \perp V_p$. Since $V_{k+1} = U \pm b_{k+1}$, we have $|V_{k+1} \oplus U| = |(b_{k+} \oplus b_{k-})| < |V_k \oplus U|$, which means that the size of the exchange sets b_k is decreasing, and thus, the sequence eventually has to end. From the construction of the b_k's, it is clear that they are all subsets of $b_1 = (U^*\backslash U, U\backslash U^*)$, and thus, $(V_p\backslash U, U\backslash V_p)$ satisfies constraints **(b)–(e)** of Definition 2 for the pair of decision sets (U, U^*). This proves that $Q \neq \varnothing$ and the argument of the *argmin* in Definition 8 is not empty. Also, note that $\mu_{\mathsf{C}_U^*} > \mu_U$ as intuitively C_U^* is a decision set made by replacing parts of U by parts of U^*.

We are now ready to give the proof of Theorem 3.

Proof of Theorem 3. We only consider the case where $i \notin U^*$ in detail, the case $i \in U^*$ is symmetric. Let $H*_i$ and $\mathcal{H}*_i^2$ be defined as H_i and $\mathcal{H}_i^2$, respectively, but using the $*$-complement C_U^* of Definition 8 instead of C_U. Then similar to Lemma 3 one can show the equivalence of the $*$-complement complexities, i.e., $\mathcal{H}*_i^2 = H*_i$.

[12]Note that since $U \perp V$, the exchange set $b = (V\backslash U, U\backslash V)$ is the only non-empty exchange set that satisfies constraints **(b)–(e)** of Definition 2 for the pair of decision sets (U, C_U).

Therefore, we have the following series of inequalities

$$H_i \overset{(a)}{=} \mathcal{H}_i^2 \overset{(b)}{=} \max_{U\in\mathcal{C}:i\in U} H_{U,\mathsf{C}_U} \overset{(c)}{\leq} \max_{U\in\mathcal{C}:i\in U} H_{U,\mathsf{C}^*_U} \overset{(d)}{=} \max_{U\in\mathcal{C}:i\in U\setminus\mathsf{C}^*_U} H_{U,\mathsf{C}^*_U} \overset{(e)}{=} H_{U_i,V_i},$$

where **(a)** holds by Lemma 3, **(b)** uses the definition of $\mathcal{H}_i^2$, **(c)** uses $H_{U,C_U} = \arg\min_{V:\mu_V>\mu_U} H_{U,V} \leq H_{U,\mathsf{C}^*_U}$ as $\mu_{\mathsf{C}^*_U} > \mu_U$, **(d)** uses the equivalence of complexity based on the $*$-complement, **(e)** introduces U_i to denote the decision set attaining the maximum in the above equation and $V_i = \mathsf{C}^*_{U_i}$. By the definition of the $*$-complement, $b' = (V_i\setminus U_i, U_i\setminus V_i)$ satisfies constraints **(b)**-**(e)** of Definition 2 for the pair of decision sets (U, U^*). As a result, $U^* \mp b' \in \mathcal{C}$ and $i \in U^* \mp b'$ (see constraint (e) in Definition 2 and $i \in b'_-$). By the definition of b', we have $\mu_{U_i} = \mu_{V_i} + \mu_{b'_+} - \mu_{b'_-}$, and thus, we may write

$$\Delta_{V_i,U_i} = \mu_{b'_+} - \mu_{b'_-} = \mu^* - (\mu^* - \mu_{b'_+} + \mu_{b'_-}) = \Delta_{U^*,U^*\mp b'} \geq \min_{U:i\in U} \Delta_{U^*,U},$$

where the last inequality follows from the fact that $i \in U^* \mp b'$. We note that for any independent pair of sets such as V_i, U_i, any well defined exchange class $\mathcal{B}$ should include the (unique) exchange set $b' = (V_i\setminus U_i, U_i\setminus V_i)$ that allows to move from one set to another. As a result for any exchange class $\mathcal{B}$, width$(\mathcal{B}) = \max_{(b_+,b_-)\in\mathcal{B}} |b_+|+|b_-| \geq |b'_+|+|b'_-| = |U_i \oplus V_i| = \overline{d}_{U_i,V_i}$. Therefore, width$(\mathcal{C}) = \min_{\mathcal{B}\in\text{Exchange}(\mathcal{C})}$ width$(\mathcal{B}) \geq \overline{d}_{U_i,V_i}$, which together with $\Delta_{V_i,U_i} \geq \min_{U:i\in U} \Delta_{U^*,U}$ leads to the desired outcome

$$H_i \leq \frac{\overline{d}^2_{U_i,V_i}}{\Delta^2_{V_i,U_i}} \leq \frac{\text{width}(\mathcal{C})^2}{\min_{U:i\in U} \Delta^2_{U^*,U}} = H_i^{\odot}.$$

$\square$

G Fixed Budget Results: Proof of Theorem 1

In this section, we provide a proof of Theorem 1. In the following, we will mainly work with the complexity $\mathcal{H}_i^2$ (and the corresponding simplicity G) as defined in equation (7). Recall that this formulation is equivalent to H_i. In the following, we use $[N]$ to denote the set $\{1, 2, \ldots, N\}$. We also introduce two numerical constants $0 < c_1 < 1$ and $0 < c_2 < 1/2$ such that $c_2 \geq \frac{c_1}{1-2c_2} \geq c_2$, whose exact values will be chosen later. Finally, we consider a permutation π of the arms that orders the arms with respect to the values G_i, that is, $G_{\pi(1)} \geq G_{\pi(2)} \geq \ldots G_{\pi(K)}$. To simplify notation, in the following, we will simply write $G_{(i)}$ instead of $G_{\pi(i)}$.

We now introduce a high-probability event which serves as a basis for the proof of the correctness of the algorithm. This event states that at the end of each phase k the estimated values of the arms will differ from their real values by at most $G(k)$.

Lemma 4. *Let $G_{(1)} \geq G_{(2)} \geq \ldots \geq G_{(K)}$ be an ordering of arms by decreasing complexity* [13]*. The event ξ defined as*

$$\xi = \left\{\forall i \in \mathcal{K}, k \in [K], \;\; |\widehat{\mu}_i(k) - \mu_i| \leq c_1 G_{(k)}\right\} \tag{32}$$

holds with probability

$$\mathbb{P}(\xi) \geq 1 - 2K^2 \exp\left(-\frac{2c_1^2(n-K)}{\overline{\log}(K)\overline{H}}\right).$$

Proof. By Hoeffding's inequality and a union bound, the probability of the complementary event $\overline{\xi}$ of ξ can be bounded as

[13] Notice that the i-th simplest arm is the $(K+1-i)$-th most complex arm.

Parameters: number of rounds n, set of arms $\mathcal{K}$, decision set $\mathcal{C}$, and cumulative pulls scheme $n_0, n_1, \dots, n_K$.
Let $\mathcal{K}_1 = \mathcal{K}$, $k = 1$, and $J_n = \varnothing$
while $|\mathcal{K}_k| \geq 1$ **do**
  Pull each arm $i \in \mathcal{K}_k$ for $n_k - n_{k-1}$ rounds.
  Compute $\widehat{U}^*(k) = \arg\max_{U \in \mathcal{C}} \widehat{\mu}_U(k)$.
  Find $j_k = \max_{i \in \mathcal{K}_k} \widehat{G}_i(k)$.
  if $j_k \in \widehat{U}^*(k)$ **then**
    The arm j_k is accepted and $J_n = J_n \cup \{j_k\}$.
  end if
  Deactivate arms j_k, i.e., set $\mathcal{K}_{k+1} = \mathcal{K}_k \backslash j_k$.
  $k \longleftarrow k+1$
end while
Return J_n

Figure 5: The modified fixed budget algorithm.

follows, provided we use the proposed pulls scheme $n_k = \left\lceil \frac{n-K}{\overline{\log}(K)(K+1-k)} \right\rceil,\ k \in \mathcal{K}$:

$$\begin{aligned}
\mathbb{P}(\bar{\xi}) &= \sum_{i=1}^{K} \sum_{k=1}^{K} \mathbb{P}\left(|\widehat{\mu}_i(k) - \mu_i| > c_1 G_{(k)}\right) \\
&\leq \sum_{i=1}^{K} \sum_{k=1}^{K} 2\exp\left(-2n_k c_1^2 G_{(k)}^2\right) \\
&\leq \sum_{i=1}^{K} \sum_{k=1}^{K} 2\exp\left(-\frac{2c_1^2(n-K)}{\overline{\log}(K)(K+1-k)H_{(K+1-k)}}\right) \\
&\leq 2K^2 \exp\left(-\frac{2c_1^2(n-K)}{\overline{\log}(K)\overline{H}}\right).
\end{aligned}$$

□

For the proof of Theorem 1 we analyze a slightly modified algorithm described in Figure 5, where for each arm that is deactivated it is immediately (and not only after all arms have been deactivated) decided, whether it shall be contained in the set returned by the algorithm at the end. On the event ξ, the correctness of both algorithms is the same, which can be deduced from statement (ii) of the induction hypothesis in Definition 9 below.

We will now prove Theorem 1 by showing that on event ξ, the optimal set U^* is identified at the end of the phases. That is, the algorithm neither accepts an arm not in U^*, nor is any arm in U^* rejected.

G.1 The Induction Hypothesis

The proof proceeds by induction over the phases of the algorithm. We first introduce the induction hypothesis.

Definition 9. *The induction hypothesis is defined by the two following properties. At the beginning of phase k we have:*

(i) *All accepted arms belong to the optimal set, i.e., $J_n(k-1) \subseteq U^*$, and all rejected arms (i.e., arms which have been deactivated but never accepted) do not belong to the optimal set, i.e., $(\mathcal{K}_k \backslash J_n(k-1)) \cap U^* = \varnothing$.*

(ii) *If arm $i \notin \mathcal{K}_k$ and it has been deactivated during phase $l \in [k-1]$, then $G_i \geq (1-2c_2)G_{(l)}$.*

Statement (i) is the classical desired property, while statement (ii) is specific to our approach and implies that by having been pulled n_l times the arm i has been sampled sufficiently often w.r.t. its complexity. Indeed, recall that a set is not necessarily compared to the optimal set U^* whose arms most probably belongs to $\mathcal{K}_k$. Therefore we need to show that an arm i contained in a set V that is likely to be used as a complement set by some "active" set U has been sampled often enough (i.e., proportionally to its complexity H_i), especially if it has been removed in a previous phase $l < k$.

We continue with a few properties implied by the induction hypothesis together with the high-probability event ξ of Lemma 4. We start with concentration inequalities on event ξ for $\widehat{\mu}$, $\widehat{\Delta}$, and $\widehat{G}$.

Proposition 8. *Assume that the induction hypothesis (Definition 9) at the beginning of phase k as well as event ξ hold. Then for any arm $i \in \mathcal{K}$,*

$$|\widehat{\mu}_i(k) - \mu_i| \le c_1 \max\left\{\frac{G_i}{1-2c_2}, G_{(k)}\right\}. \tag{33}$$

Furthermore for any pair $(U, V) \in \mathcal{C}^2$ such that $V = \mathsf{C}_U$ we have

$$\begin{aligned} \left|\widehat{\Delta}_{V,U}(k) - \Delta_{V,U}\right| &\le c_2 \overline{d}_{U,V} \max\{G_{V,U}, G_{(k)}\} \\ \text{and} \quad \left|\widehat{G}_{V,U}(k) - G_{V,U}\right| &\le c_2 \max\left\{G_{V,U}, G_{(k)}\right\}. \end{aligned}$$

where $\widehat{\Delta}_{V,U}(k)$ and $\widehat{G}_{V,U}(k)$ are the gaps and the simplicity computed at the end of the phase k. Finally, for the special case of pairs (U^*, U),

$$\begin{aligned} \left|\widehat{\Delta}_{U^*,U}(k) - \Delta_{U^*,U}\right| &\le c_2 \overline{d}_{U,U^*} \max\{G_{V,U}, G_{(k)}\} \\ \text{and} \quad \left|\widehat{G}_{U^*,U}(k) - G_{U^*,U}\right| &\le c_2 \max\{G_{V,U}, G_{(k)}\}, \end{aligned}$$

with $V = \mathsf{C}_U$.

Proof. First note that if $i \in \mathcal{K}_k$, then on event ξ we have $|\widehat{\mu}_i(k) - \mu_i| \le c_1 G_{(k)}$. Thus, let us assume that $i \notin \mathcal{K}_k$. Let l be the phase at which arm i has been deactivated, with $l \in \{1, \dots, k-1\}$. We have

$$|\widehat{\mu}_i(k) - \mu_i| \overset{(a)}{\le} c_1 G_{(l)} \overset{(b)}{\le} \frac{c_1}{1-2c_2} G_i,$$

where **(a)** is implied by the event ξ and the fact that $\widehat{\mu}_i(k) = \widehat{\mu}_i(l)$, **(b)** uses property (ii) of the induction hypothesis. Summarizing, independent of whether $i \in \mathcal{K}_k$ or not, we have for any i

$$|\widehat{\mu}_i(k) - \mu_i| \le c_1 \max\left\{\frac{G_i}{1-2c_2}, G_{(k)}\right\},$$

which shows the first claim. Let us now focus on a pair of sets U, V such that $i \in U \oplus V$ and $V = \mathsf{C}_U$ or $V = U^*$. Then by the previous inequality,

$$\begin{aligned} |\widehat{\mu}_i(k) - \mu_i| &\le c_1 \max\left\{\frac{G_i}{1-2c_2}, G_{(k)}\right\} \\ &\overset{(a)}{\le} \max\left\{c_2 G_i, c_1 G_{(k)}\right\} \\ &\overset{(b)}{\le} c_2 \max\left\{G_{V,U}, G_{(k)}\right\}, \end{aligned}$$

where **(a)** follows from the choice of the constants such that $c_2 \ge \frac{c_1}{1-2c_2}$. For **(b)** we use Proposition 5 for $V = \mathsf{C}_U$, the definition of $G_{V,U}$, the fact that $\Delta_{V,U} > 0$ as well as $c_1 \le c_2$. As a result we obtain

$$\begin{aligned} \left|\widehat{\Delta}_{V,U}(k) - \Delta_{V,U}\right| &\le \sum_{i \in U \oplus V} |\widehat{\mu}_i(k) - \mu_i| \\ &\le c_2 \overline{d}_{V,U} \max\{G_{V,U}, G_{(k)}\}, \end{aligned}$$

which proves the first part of the second statement. The second part then simply follows from

$$\begin{aligned} \widehat{G}_{V,U}(k) &= \frac{\widehat{\Delta}_{V,U}(k)}{\overline{d}_{U,V}} \\ &\overset{(c)}{\ge} \frac{\Delta_{V,U} - c_2 \overline{d}_{V,U} \max\{G_{V,U}, G_{(k)}\}}{\overline{d}_{V,U}} \\ &= G_{V,U} - c_2 \max\{G_{V,U}, G_{(k)}\}, \end{aligned}$$

where **(c)** follows from the first statement. The missing inequality to conclude the second part of the second statement can be obtained analogously. Finally, the last statement follows along the same lines, only replacing Proposition 5 in step **(b)** above by Proposition 6. □

The following Proposition shows that at the beginning of each phase k there is an arm a_k which has a larger simplicity than the k-th largest simplicity. In the remaining proof this arm will serve as a reference arm, as this is the arm that should be deactivated at the end of phase k.

Proposition 9. *Let*

$$a_k = \operatorname*{arg\,min}_{i\in\mathcal{K}_k} H_i = \operatorname*{arg\,max}_{i\in\mathcal{K}_k} G_i$$

be the simplest arm among those left at the beginning of phase k. Then $G_{a_k} \geq G_{(k)}$.

Proof. At the beginning of phase k, only $K - |\mathcal{K}_k| = k-1$ arms have been deactivated, i.e., $|\mathcal{K}_k| = K+1-k$. Hence the simplest arm left in $\mathcal{K}_k$ (i.e., a_k) cannot be more difficult than the arm that would be left if all the $k-1$ simpler arms were deactivated (i.e., $G_{(k)}$), which gives the claimed $G_{(k)} \leq G_{a_k}$. □

The following Proposition shows that at the end of each phase k, the reference arm a_k belongs to $\widehat{U}^*$ if and only if it actually belongs to U^*. This allows us to show that the simplicity of a_k can be well estimated at the end of phase k.

Proposition 10. *Assume that the induction hypothesis (Definition 9) at the beginning of phase k as well as event ξ hold. Then $a_k \in U^*$ if and only if $a_k \in \widehat{U}^*(k)$, where $\widehat{U}^*(k)$ is the estimated optimal set at the end of phase k.*

Proof. We prove in detail that $a_k \notin U^*$ implies $a_k \notin \widehat{U}^*(k)$. The reverse can be shown along a similar line of arguments. The proof is by contradiction. Thus, let us assume that $a_k \notin U^*$ and $a_k \in \widehat{U}^*(k)$. Note that this implies that $\widehat{U}^*(k) \neq U^*$. Let $W = \mathsf{C}_{\widehat{U}^*(k)}$ be the complement of the estimated optimal set (note that W exists, since $\widehat{U}^*(k)$ is not optimal). We have

$$\begin{aligned}\widehat{\Delta}_{W,\widehat{U}^*(k)}(k) &\overset{(a)}{\geq} \Delta_{W,\widehat{U}^*(k)} - c_2 \overline{d}_{W,\widehat{U}^*(k)} \max\left\{G_{W,\widehat{U}^*(k)}, G_{(k)}\right\} \\ &= \overline{d}_{W,\widehat{U}^*(k)}\left(G_{W,\widehat{U}^*(k)} - c_2 \max\left\{G_{W,\widehat{U}^*(k)}, G_{(k)}\right\}\right),\end{aligned}$$

where **(a)** uses Proposition 8. We consider two cases.

Case 1) $G_{W,\widehat{U}^*(k)} \geq G_{(k)}$

In this case,

$$\begin{aligned}\widehat{\Delta}_{W,\widehat{U}^*(k)}(k) &\geq \overline{d}_{W,\widehat{U}^*(k)}\left(G_{W,\widehat{U}^*(k)} - c_2 G_{W,\widehat{U}^*(k)}\right) \\ &= \overline{d}_{W,\widehat{U}^*(k)} G_{W,\widehat{U}^*(k)}(1-c_2) > 0,\end{aligned}$$

where the last inequality follows from the fact that $0 < c_2 < 1$ and $G_{W,\widehat{U}^*(k)} > 0$ by definition of W. It follows that $\widehat{\mu}_{\widehat{U}^*(k)}(k) < \widehat{\mu}_W(k)$, which contradicts that $\widehat{U}^*(k)$ is the empirical best set.

Case 2) $G_{W,\widehat{U}^*(k)} < G_{(k)}$)

We write

$$\begin{aligned}\widehat{\Delta}_{W,\widehat{U}^*(k)}(k) &\geq \overline{d}_{W,\widehat{U}^*(k)}\left(G_{W,\widehat{U}^*(k)} - c_2 G_{(k)}\right) \\ &\overset{(b)}{\geq} \overline{d}_{W,\widehat{U}^*(k)}\left(G_{a_k} - c_2 G_{(k)}\right) \\ &\overset{(c)}{\geq} \overline{d}_{W,\widehat{U}^*(k)} G_{(k)}(1-c_2) > 0,\end{aligned}$$

where **(b)** holds since $a_k \in \widehat{U}^*(k)$, so that $G_{a_k} = \min_{U:a_k\in U} \max_{V:\mu_V>\mu_U} G_{V,U} \leq \max_{V:\mu_V>\mu_{\widehat{U}^*(k)}} G_{V,\widehat{U}^*(k)} = G_{W,\widehat{U}^*(k)}$. Concerning **(c)**, this holds by Proposition 9, as $G_{a_k} \geq G_{(k)}$. Similar as before we obtain the contradiction $\widehat{\mu}_{\widehat{U}^*(k)}(k) < \widehat{\mu}_W(k)$. □

The following Proposition gives a lower bound on the estimated simplicity of the reference arm a_k at the end of phase k depending on $G(k)$. This will be used later to show that the algorithm does not remove other arms than a_k in each phase k.

Proposition 11. *Assume that the induction hypothesis (Definition 9) at the beginning of phase k as well as event ξ hold. Then $\widehat{G}_{a_k}(k) \geq (1-c_2)G_{(k)}$.*

Proof. We give a detailed proof for the case where $a_k \notin \widehat{U}^*(k)$. The other case follows from symmetric arguments. Let U_k be a set that defines the estimated simplicity of a_k, that is

$$U_k \in \arg\min_{U: a_k \in U} \widehat{G}_{\widehat{\mathsf{C}}_U(k),U}(k),$$

and let $W_k = \mathsf{C}_{U_k}$. Note that W_k is well-defined because $U_k \neq U^*$: Indeed, from Proposition 10, since $a_k \notin \widehat{U}^*(k)$, also $a_k \notin U^*$, so that because $a_k \in U_k$, we have $U_k \neq U^*$.

Then we have

$$\begin{aligned}
\widehat{G}_{a_k}(k) &= \widehat{G}_{\widehat{\mathsf{C}}_U(k),U_k}(k) \\
&= \max_{V: \widehat{\mu}_V(k) > \widehat{\mu}_{U_k}(k)} \widehat{G}_{V,U_k}(k) \\
&\geq \widehat{G}_{W_k,U_k}(k) \\
&\overset{(a)}{\geq} G_{W_k,U_k} - c_2 \max\left\{G_{W_k,U_k}, G_{(k)}\right\},
\end{aligned}$$

where **(a)** follows from Proposition 8. We have, as $a_k \notin U^*$, that

$$G_{a_k} = \min_{U: a_k \in U} G_{\mathsf{C}_U,U} \leq G_{W_k,U_k}.$$

Furthermore, from Proposition 9 we have that $G_{(k)} \leq G_{a_k}$ and thus $G_{(k)} \leq G_{W_k,U_k}$. Thus the previous expression simplifies to

$$\widehat{G}_{a_k}(k) \geq G_{W_k,U_k} - c_2 G_{W_k,U_k} \geq (1-c_2)G_{(k)}.$$

□

The following lemma is rather technical. It shows that if the estimated simplicity of a sub-optimal decision U_k is defined with respect to V_k, then this estimated simplicity will be larger than the true simplicity of the decision set V_k. The proof shows that if the estimated simplicity of U_k were smaller than the true simplicity of V_k, then it would surely be also smaller than the estimated simplicity of V_k defined with respect to W_k, leading to a contradiction. For notational convenience, in the following we will drop the dependency of the estimated quantities on the phase k (e.g., write $\widehat{G}_{V_k,U_k}$ instead of $\widehat{G}_{V_k,U_k}(k)$).

Proposition 12. *Assume that the induction hypothesis (Definition 9) at the beginning of phase k as well as event ξ hold. Further assume that $U_k, V_k, W_k \in \mathcal{C}$ such that $U_k \neq \widehat{U}^*(k)$, $V_k = \widehat{\mathsf{C}}_{U_k}(k) \neq U^*$, $W_k = \mathsf{C}_{V_k}$, and $G_{W_k,V_k} \geq G_{(k)}$. Then $\widehat{G}_{V_k,U_k}(k) \geq (1-c_2)G_{W_k,V_k}$.*

Proof. We start by showing that $\widehat{\mu}_{U_k} < \widehat{\mu}_{V_k} < \widehat{\mu}_{W_k}$. First, $\widehat{\mu}_{U_k} < \widehat{\mu}_{V_k}$ comes from the definition of V_k as the (estimated) complement of U_k. Furthermore,

$$\begin{aligned}
\widehat{\Delta}_{W_k,V_k} &\overset{(a)}{\geq} \Delta_{W_k,V_k} - c_2 \overline{d}_{W_k,V_k} \max\{G_{W_k,V_k}, G_{(k)}\} \\
&\overset{(b)}{\geq} \Delta_{W_k,V_k} - c_2 \overline{d}_{W_k,V_k} G_{W_k,V_k} \\
&\overset{(c)}{=} \Delta_{W_k,V_k} - c_2 \Delta_{W_k,V_k} > 0,
\end{aligned}$$

where **(a)** follows from Proposition 8 and the fact that $\Delta_{W_k,V_k} > 0$ (since W_k is the (exact) complement of V_k), **(b)** follows from the assumption that $G_{W_k,V_k} \geq G_{(k)}$, and **(c)** is obtained from the definition of simplicity G_{W_k,V_k} and the fact that $0 < c_2 < 1$. This completes the proof of the claim that $\widehat{\mu}_{U_k} < \widehat{\mu}_{V_k} < \widehat{\mu}_{W_k}$.

Next, we show that $\widehat{G}_{W_k,V_k} \leq \widehat{G}_{V_k,U_k}$. First note that by Proposition 3[14] we obtain from $\widehat{\mu}_{U_k} < \widehat{\mu}_{V_k} < \widehat{\mu}_{W_k}$ that $\widehat{G}_{W_k,U_k} \geq \min\left\{\widehat{G}_{W_k,V_k}, \widehat{G}_{V_k,U_k}\right\}$, where the inequality is strict whenever $\widehat{G}_{W_k,V_k} \neq \widehat{G}_{V_k,U_k}$. Now if we assume that $\widehat{G}_{W_k,V_k} > \widehat{G}_{V_k,U_k}$, the previous inequality becomes strict as $\widehat{G}_{W_k,U_k} > \widehat{G}_{V_k,U_k}$. Then we would obtain the contradiction

$$\widehat{G}_{V_k,U_k} = \max_{V:\widehat{\mu}_V > \widehat{\mu}_U} \widehat{G}_{V,U_k} \geq \widehat{G}_{W_k,U_k} > \widehat{G}_{V_k,U_k},$$

thus implying the claimed $\widehat{G}_{W_k,V_k} \leq \widehat{G}_{V_k,U_k}$.

We now can conclude with

$$\begin{aligned}\widehat{G}_{V_k,U_k} &\geq \widehat{G}_{W_k,V_k}\\ &\overset{(a)}{\geq} G_{W_k,V_k} - c_2 \max\{G_{W_k,V_k}, G_{(k)}\}\\ &\overset{(b)}{\geq} G_{W_k,V_k} - c_2 G_{W_k,V_k},\end{aligned}$$

where **(a)** follows from Proposition 8 and the fact that $\Delta_{W_k,V_k} > 0$ and **(b)** holds due to the assumption that $G_{W_k,V_k} \geq G_{(k)}$. □

G.2 The Induction Step

We now move on to prove the induction step. We do this in two separate lemmas, one for each property in the induction hypothesis.

Lemma 5. *Assume that the induction hypothesis at the beginning of phase k as well as event ξ hold. Then property (i) of Definition 9 holds at phase $k+1$ as well.*

Proof. Property (i) states that no error is made until the beginning of phase k. To prove that this is still true at the beginning of phase $k+1$ we need to prove that during phase k no error is made. An error occurs when either the algorithm deactivates and rejects an arm $j_k \in U^*$, or the algorithm deactivates and accepts an arm $j_k \notin U^*$. We show by contradiction that both cases cannot happen.

We give a detailed proof for the case where $j_k \in U^*$ is rejected. The argument for the second source of error, when $j \notin U^*$ is accepted, is similar.

The strategy of the proof will be to compare the estimated simplicity of the rejected arm to the simplicity of the arm a_k, that is, the arm with the highest simplicity at the end of phase k, and which should be the targeted arm to be deactivated. Using Proposition 11, we know that the simplicity of a_k is of order of $G_{(k)}$. Therefore it remains to prove that G_{j_k} is smaller than $G_{(k)}$.

As j_k is rejected, $j_k \notin \widehat{U}^*(k)$ and hence, $U^* \neq \widehat{U}^*$. Furthermore, $a_k \neq j_k$, as otherwise Proposition 10 would imply that $j_k \in \widehat{U}^*$, a contradiction to the assumption that arm j_k is rejected.
As j_k has been deactivated during phase k, we have $\widehat{G}_{j_k} \geq \widehat{G}_{a_k}$, since the algorithm deactivates the arm with the largest simplicity, that is, $j_k = \arg\max_{i\in\mathcal{K}_k} \widehat{G}_i$. Let $V_k = \widehat{\mathsf{C}}_{U^*}$ be the estimated complement of the optimal set (which exists, as $U^* \neq \widehat{U}^*$) and $W_k = \mathsf{C}_{V_k}$ be the (exact) complement of V_k. Then

$$\begin{aligned}\widehat{G}_{j_k} &= \min_{U:j\in U}\max_{V:\widehat{\mu}_V > \widehat{\mu}_U} \widehat{G}_{V,U}\\ &\overset{(a)}{\leq} \widehat{G}_{V_k,U^*}\\ &\overset{(b)}{\leq} G_{V_k,U^*} + c_2 \max\left\{G_{W_k,V_k}, G_{(k)}\right\}\\ &\overset{(c)}{\leq} c_2 \max\left\{G_{W_k,V_k}, G_{(k)}\right\},\end{aligned}$$

where **(a)** is because $j_k \in U^*$ and V_k is the (estimated) complement of U^*, **(b)** holds by the last statement of Proposition 8, and **(c)** follows from $\mu_{V_k} < \mu^*$, whence $\Delta_{V_k,U^*} < 0$ and $G_{V_k,U^*} < 0$.

[14] More precisely, we rely on an equivalent version based on estimated values and estimated simplicity.

We now show by contradiction that $G_{W_k,V_k} \leq G_{(k)}$. Thus, assume that $G_{W_k,V_k} > G_{(k)}$. Then by Proposition 12, $\widehat{G}_{V_k,U^*} \geq (1-c_2)G_{W_k,V_k}$. Together with the previous inequality this gives

$$(1-c_2)G_{W_k,V_k} \leq \widehat{G}_{V_k,U^*} \leq c_2 \max\{G_{W_k,V_k}, G_{(k)}\} \leq c_2 G_{W_k,V_k},$$

which is a contradiction, since $c_2 < 1/2$, and we conclude that $G_{W_k,V_k} \leq G_{(k)}$.

Now using inequality (c) from before, we get $\widehat{G}_{j_k} \leq c_2 G_{(k)}$. Using Proposition 11, for a_k we have that $\widehat{G}_{a_k} \geq (1-c_2)G_{(k)}$. Since $c_2 < 1/2$, it holds that $\widehat{G}_{j_k} \leq c_2 G_{(k)} < (1-c_2)G_{(k)} \leq \widehat{G}_{a_k}$. As a result $\widehat{G}_{j_k} < \widehat{G}_{a_k}$, which contradicts the fact that j_k would be the deactivated arm during phase k, as by definition of $j_k = \arg\max_{i \in \mathcal{K}_k} \widehat{G}_i$. Thus we can conclude that the algorithm does not reject any arm from U^*.

$\square$

Lemma 6. *Assume that the induction hypothesis at the beginning of phase k as well as event ξ hold. Then property (ii) of Definition 9 holds at phase $k+1$ as well.*

Proof. Let j_k be the arm which is deactivated at the end of phase k, i.e., $j_k = \arg\max_{i \in \mathcal{K}_k} \widehat{G}_i(k)$. Again, we only consider the case where $j_k \notin U^*$, as the proof for the case $j_k \in U^*$ is symmetrical. We have to show that j_k satisfies $G_{j_k} \geq (1-2c_2)G_{(k)}$, which will be done in five steps.

Step 1. We first notice that Lemma 5 implies that no error is made during phase k, since property (i) still holds at the beginning of phase $k+1$. As a result, we have that $j_k \in U^*$ if and only if $j_k \in \widehat{U}^*$ which means in our current case that $j_k \notin \widehat{U}^*$. Let U_k and V_k be the sets which define the (exact) simplicity of j_k, that is,

$$U_k = \arg\min_{U: j_k \in U} G_{\mathsf{C}_U, U}$$

and $V_k = \mathsf{C}_{U_k}$, so that $G_{j_k} = G_{V_k,U_k}$. Further, let $W_k = \widehat{\mathsf{C}}_{U_k}$, noting that W_k is well defined since $U_k \neq \widehat{U}^*$. Indeed, $j_k \notin U^*$ and $j_k \in U_k$, whence $U_k \neq U^*$.

We claim that $G_{W_k,U_k} \leq G_{j_k}$. Indeed, if $\mu_{W_k} \leq \mu_{U_k}$ then we trivially have $G_{W_k,U_k} \leq 0 \leq G_{j_k}$. Furthermore, if $\mu_{W_k} > \mu_{U_k}$ we have by definition of U_k and V_k,

$$G_{W_k,U_k} \leq \max_{W: \mu_W > \mu_{U_k}} G_{W,U_k} = G_{V_k,U_k} = G_{j_k},$$

which proves our claim $G_{W_k,U_k} \leq G_{j_k}$.

Step 2. Next, we note that

$$\widehat{G}_{j_k} = \min_{U: j_k \in U} \widehat{G}_{\widehat{\mathsf{C}}_U, U} \leq \max_{\widehat{\mu}_V > \widehat{\mu}_{U_k}} \widehat{G}_{V,U_k} = \widehat{G}_{W_k,U_k}. \tag{34}$$

Step 3. Here we show that $G_i \leq \max\{G_{j_k}, G_{Z_k,W_k}\}$ for $i \in W_k \oplus U_k$. We distinguish the following cases:

Case 1) $i \in W_k \backslash U_k$
Case 1.1) $i \in U^*$: In this case $i \in U_k \oplus U^*$ and thus we can apply Proposition 6 to U_k and $\mathsf{C}_{U_k} = V_k$ and obtain $G_i \leq G_{V_k,U_k} = G_{j_k}$.
Case 1.2) $i \notin U^*$: Let $Z_k = \mathsf{C}_{W_k}$, noting that since i is in W_k but not in U^*, we have $W_k \neq U^*$. Then

$$G_i = \min_{U: i \in U} G_{\mathsf{C}_U, U} \leq G_{\mathsf{C}_{W_k}, W_k} = G_{Z_k,W_k},$$

where the first equality follows from the fact that $i \notin U^*$, while the inequality is due to the fact that $i \in W_k$.

Case 2) $i \in U_k \backslash W_k$
Case 2.1) $i \in U^*$: Let $Z_k = \mathsf{C}_{W_k}$, noting that since i is in U^* but not in W_k, it holds that $W_k \neq U^*$. Then

$$G_i = \min_{U: i \notin U} G_{\mathsf{C}_U, U} \leq G_{\mathsf{C}_{W_k}, W_k} = G_{Z_k,W_k},$$

where the first equality follows from the fact that $i \in U^*$, while the inequality is due to the fact that $i \notin W_k$.
Case 2.2) $i \notin U^*$: In this case we have by definition that

$$G_i = \min_{U: i \in U} G_{\mathsf{C}_U, U} \leq G_{\mathsf{C}_U, U_k} = G_{V_k, U_k} = G_{j_k},$$

which completes the proof that $G_i \leq \max\{G_{j_k}, G_{Z_k, W_k}\}$ for $i \in W_k \oplus U_k$.

Step 4. Next, we are going to show that $\widehat{G}_{j_k} \leq G_{j_k} + c_2 \max\{G_{j_k}, G_{(k)}\}$, a version of Proposition 8 for arm j_k. We distinguish two cases:

Case 1) $W_k = U^*$

In this case, we have

$$\widehat{G}_{j_k} \overset{(a)}{\leq} \widehat{G}_{W_k, U_k} \overset{(b)}{\leq} G_{W_k, U_k} + c_2 \max\{G_{V_k, U_k}, G_{(k)}\} \overset{(c)}{\leq} G_{j_k} + c_2 \max\left\{G_{j_k}, G_{(k)}\right\},$$

where **(a)** is obtained from equation (34), **(b)** is a result of Proposition 8, and **(c)** follows from $G_{W_k, U_k} \leq G_{j_k}$ of Step 1.

Case 2) $W_k \neq U^*$

Let $Z_k = \mathsf{C}_{W_k}$. We prove by contradiction that $G_{Z_k, W_k} \leq \max\left\{G_{V_k, U_k}/(1-2c_2), G_{(k)}\right\}$. Thus, assume that $G_{Z_k, W_k} > \max\left\{G_{V_k, U_k}/(1-2c_2), G_{(k)}\right\}$. Then

$$\begin{aligned}
\widehat{G}_{W_k, U_k} &\overset{(a)}{\leq} G_{W_k, U_k} + c_1 \frac{1}{\bar{d}_{U_k \oplus W_k}} \sum_{i \in U_k \oplus W_k} \max\left\{\frac{G_i}{1-2c_2}, G_{(k)}\right\} \\
&\overset{(b)}{\leq} G_{W_k, U_k} + c_1 \max\left\{\frac{G_{j_k}}{1-2c_2}, \frac{G_{Z_k, W_k}}{1-2c_2}, G_{(k)}\right\} \\
&\overset{(c)}{\leq} G_{V_k, U_k} + c_1 \frac{G_{Z_k, W_k}}{1-2c_2} \\
&\overset{(d)}{\leq} G_{V_k, U_k} + c_2 G_{Z_k, W_k},
\end{aligned} \tag{35}$$

where **(a)** is using for all $i \in U_k \oplus W_k$ the Equation 33 in Proposition 8, **(b)** follows from Step 3, and **(c)** is obtained by $G_{W_k, U_k} \leq G_{j_k} = G_{V_k, U_k}$ (Step 1) and the assumption on G_{Z_k, W_k}, finally **(d)** is obtained by $\frac{c_1}{1-2c_2} \leq c_2$. By Proposition 12, we have $\widehat{G}_{W_k, U_k} \geq (1-c_2) G_{Z_k, W_k}$, which together with (35) gives

$$(1-c_2) G_{Z_k, W_k} \leq \widehat{G}_{W_k, U_k} \leq G_{V_k, U_k} + c_2 G_{Z_k, W_k} \leq c_2 G_{Z_k, W_k},$$

which is a contradiction due to $0 < c_2 < 1/2$. This finishes the proof of $G_{Z_k, W_k} \leq \max\{G_{V_k, U_k}/(1-2c_2), G_{(k)}\}$.

By (35) and the results of Steps 2 and 1, we finally get

$$\widehat{G}_{j_k} \leq \widehat{G}_{W_k, U_k} \leq G_{W_k, U_k} + c_1 \max\left\{\frac{G_{j_k}}{1-2c_2}, \frac{G_{Z_k, W_k}}{1-2c_2}, G_{(k)}\right\} \leq G_{j_k} + c_2 \max\left\{G_{j_k}, G_{(k)}\right\}.$$

Step 5. From Proposition 11 and the fact that j_k is the deactivated arm (i.e., $j_k = \arg\max_{i \in \mathcal{K}_k} \widehat{G}_i$) we have

$$(1-c_2) G_{(k)} \leq \widehat{G}_{a_k} \leq \widehat{G}_{j_k} \leq G_{j_k} + c_2 \max\left\{G_{j_k}, G_{(k)}\right\}.$$

We conclude by considering the two possible cases for the $\max$ term.
Case 1) $G_{j_k} > G_{(k)}$
We have $(1-c_2) G_{(k)} \leq G_{j_k} + c_2 G_{j_k}$. Since $\frac{1-c_2}{1+c_2} \geq 1 - 2c_2$, we get

$$G_{j_k} \geq \frac{1-c_2}{1+c_2} G_{(k)} \geq (1-2c_2) G_{(k)}.$$

Case 2) $G_{j_k} \leq G_{(k)}$
Here we have $(1-c_2) G_{(k)} \leq G_{j_k} + c_2 G_{(k)}$, whence

$$(1-2c_2) G_{(k)} \leq G_{j_k},$$

which concludes the proof. □

G.3 Proof of Theorem 1

With the results of the previous sections, the proof of Theorem 1 is immediate. First, assume that event ξ holds. Note that properties (i) and (ii) of the induction assumption hold for phase $k = 1$. Lemmas 5 and 6 prove the induction step, showing that properties (i) and (ii) hold for all phases k. It remains to consider the error probability for event ξ. This is handled by Lemma 4, where we finally choose $c_1 = 1/8$, $c_2 = 1/4$ so that $0 < c_1 < 1$, $0 < c_2 < 1/2$ and $c_2 \geq \frac{c_1}{1-2c_2}$. $\square$

H Fixed Confidence Results: Proof of Theorem 2

We first introduce a high-probability event corresponding to the confidence bounds used by our algorithm.

Lemma 7. *The event ξ defined as*

$$\xi = \{\forall i \in \mathcal{K}, \forall t > 0, |\widehat{\mu}_i(t) - \mu_i| \leq \beta_i(t-1)\} \tag{36}$$

holds with probability $1 - \delta$, where

$$\widehat{\mu}_i(t) = \frac{1}{T_i(t)} \sum_{t=1}^{T_i(t)} X_{i,t} \quad \textit{and} \quad \beta_i(t-1) = \sqrt{\frac{\log \frac{4Kt^2}{\delta}}{2T_i(t)}}.$$

Proof. By Chernoff-Hoeffding's inequality, the definition of the confidence intervals $\beta_i(t-1)$, and a union bound over all $T_i(t) \in \{0, \ldots, t\},\ t = 1, \ldots, \infty$. $\square$

Recall that

$$\mathcal{U}'_t = \left\{U : \forall V \in \mathcal{C}, \widehat{\Delta}^+_{U,V}(t) > -\overline{d}_{U,V} \max_{W \in \mathcal{C}} \widehat{G}^+_{W,U}(t)/2\right\}$$

and

$$\widehat{G}^+(t) = \max_{U \in \mathcal{U}'_t, V \in \mathcal{C}} \widehat{G}^+_{V,U}(t) \text{ and } (U_t, V_t) = \operatorname*{arg\,max}_{U \in \mathcal{U}'_t, V \in \mathcal{C}, U \neq V} \widehat{G}^+_{V,U}(t).$$

The following lemma gives a lower bound on $\widehat{G}^+(t)$.

Lemma 8. *On the event ξ, for all time steps t, $\widehat{G}^+(t) \geq \frac{1}{2} G_{I(t)}$.*

Proof. First note that on event ξ, for any t and any pair of decisions $U, V \in \mathcal{C}$, we have $\widehat{\Delta}^+_{U,V}(t) \geq \Delta_{U,V}$ and consequently $\widehat{G}^+_{U,V} \geq G_{U,V}$. The proof proceeds by distinguishing two main cases. We show the details for the case when $I(t) \notin U^*$. The case $I(t) \in U^*$ can be dealt with using similar arguments.

Case 1) $I(t) \in U_t$
We introduce $W_t = \mathsf{C}_{U_t}$, which exists since $U_t \neq U^*$, as $I(t) \notin U^*$ and $I(t) \in U_t$. We have

$$\begin{aligned} \widehat{G}^+(t) &= \max_{U \in \mathcal{U}'_t, V \in \mathcal{C}} \widehat{G}^+_{V,U}(t) \geq \widehat{G}^+_{W_t,U_t}(t) = \frac{\widehat{\Delta}^+_{W_t,U_t}(t)}{\overline{d}_{U_t,W_t}} \\ &\geq \frac{\Delta_{W_t,U_t}}{\overline{d}_{U_t,W_t}} \overset{(a)}{\geq} \min_{U: I(t) \in U} \frac{\Delta_{\mathsf{C}_U,U}}{\overline{d}_{U,\mathsf{C}_U}} \overset{(b)}{=} G_{I(t)}, \end{aligned}$$

where **(a)** follows from the fact that $W_t = \mathsf{C}_{U_t}$ and $I(t) \in U_t$, and **(b)** is due to $I(t) \notin U^*$ so that its complexity is defined as the minimum over decisions U to which it belongs.

Case 2) $I(t) \in V_t$
Let $W_t = \mathsf{C}_{V_t}$, noting that W_t is well-defined since $V_t \neq U^*$, as $I(t) \notin U^*$ and $I(t) \in V_t$.
Case 2.1) $V_t \in \mathcal{U}'_t$: Similar to Case 1, we have

$$\begin{aligned} \widehat{G}^+(t) &= \max_{U \in \mathcal{U}'_t, V \in \mathcal{C}} \widehat{G}^+_{V,U}(t) \overset{(a)}{\geq} \widehat{G}^+_{W_t,V_t}(t) = \frac{\widehat{\Delta}^+_{W_t,V_t}(t)}{\overline{d}_{V_t,W_t}} \\ &\geq \frac{\Delta_{W_t,V_t}}{\overline{d}_{V_t,W_t}} \overset{(b)}{\geq} \min_{U: I(t) \in U} \frac{\Delta_{\mathsf{C}_U,U}}{\overline{d}_{U,\mathsf{C}_U}} \overset{(c)}{=} G_{I(t)}, \end{aligned}$$

where **(a)** holds by $V_t \in \mathcal{U}'_t$ and the definition of $\widehat{G}^+(t)$, **(b)** follows from $W_t = \mathsf{C}_{V_t}$ and our assumption $I(t) \in V_t$, and **(c)** is due to $I(t) \notin U^*$ so that its complexity is defined as the minimum over decisions U to which it belongs.

Case 2.2) $V_t \notin \mathcal{U}'_t$: In this case, by definition of $\mathcal{U}'(t)$ there exists a decision set Z_t such that $\widehat{G}^+_{V_t,Z_t}(t) \leq -\frac{1}{2}\max_{W\in\mathcal{C}} \widehat{G}^+_{W,V_t}(t)$. Therefore, we have

$$\widehat{G}^+_{V_t,Z_t}(t) \leq -\frac{1}{2}\max_{W\in\mathcal{C}} \widehat{G}^+_{W,V_t}(t) \leq -\frac{1}{2}\widehat{G}^+_{W_t,V_t}(t) \tag{37}$$

$$\leq -\frac{1}{2}G_{W_t,V_t} \leq -\frac{1}{2}G_{I(t)}, \tag{38}$$

where the last equality follows from the fact that $I(t) \notin U^*$ and the definition of $G_{I(t)}$ in that case. We now focus on the three decision sets V_t, U_t, Z_t and define the value $\dot{\mu}_i$ associated to arms $i \in V_t \cup U_t \cup Z_t$ as

$$\dot{\mu}_i = \begin{cases} \widehat{\mu}_i^-(t) & \text{if } i \in U_t, \\ \widehat{\mu}_i^+(t) & \text{if } i \in (V_t \cup Z_t)\backslash U_t. \end{cases}$$

We also define $\dot{G}_{Z_t,U_t}$, $\dot{G}_{V_t,U_t}$, $\dot{G}_{Z_t,V_t}$ as well as $\dot{\Delta}_{Z_t,U_t}$, $\dot{\Delta}_{V_t,U_t}$, $\dot{\Delta}_{Z_t,V_t}$ obtained by using $\dot{\mu}_i$ instead of μ_i in their computation. Then we get

$$\begin{aligned}
\dot{\Delta}_{Z_t,V_t}(t) &= \sum_{i\in Z_t\backslash V_t} \dot{\mu}_i - \sum_{i\in V_t\backslash Z_t} \dot{\mu}_i = \sum_{i\in (Z_t\backslash V_t)\backslash U_t} \dot{\mu}_i + \sum_{i\in (Z_t\backslash V_t)\cap U_t} \dot{\mu}_i - \sum_{i\in (V_t\backslash Z_t)\backslash U_t} \dot{\mu}_i - \sum_{i\in (V_t\backslash Z_t)\cap U_t} \dot{\mu}_i \\
&= \sum_{i\in (Z_t\backslash V_t)\backslash U_t} \widehat{\mu}_i^+(t) + \sum_{i\in (Z_t\backslash V_t)\cap U_t} \widehat{\mu}_i^-(t) - \sum_{i\in (V_t\backslash Z_t)\backslash U_t} \widehat{\mu}_i^+(t) - \sum_{i\in (V_t\backslash Z_t)\cap U_t} \widehat{\mu}_i^-(t) \\
&\geq \sum_{i\in (Z_t\backslash V_t)\backslash U_t} \widehat{\mu}_i^-(t) + \sum_{i\in (Z_t\backslash V_t)\cap U_t} \widehat{\mu}_i^-(t) - \sum_{i\in (V_t\backslash Z_t)\backslash U_t} \widehat{\mu}_i^+(t) - \sum_{i\in (V_t\backslash Z_t)\cap U_t} \widehat{\mu}_i^+(t) \\
&= -\left(\sum_{i\in (V_t\backslash Z_t)\backslash U_t} \widehat{\mu}_i^+(t) + \sum_{i\in (V_t\backslash Z_t)\cap U_t} \widehat{\mu}_i^+(t) - \sum_{i\in (Z_t\backslash V_t)\backslash U_t} \widehat{\mu}_i^-(t) - \sum_{i\in (Z_t\backslash V_t)\cap U_t} \widehat{\mu}_i^-(t) \right) \\
&= -\Delta^+_{V_t,Z_t}(t).
\end{aligned} \tag{39}$$

Furthermore,

$$\widehat{G}^+_{Z_t,U_t}(t) \overset{(a)}{=} \dot{G}_{Z_t,U_t} \overset{(b)}{\geq} \min\left(\dot{G}_{Z_t,V_t}, \dot{G}_{V_t,U_t}\right) \tag{40}$$

$$\overset{(c)}{\geq} \min\left(\dot{G}_{Z_t,V_t}, \widehat{G}^+_{V_t,U_t}(t)\right) \tag{41}$$

$$\overset{(d)}{\geq} \min\left(-\widehat{G}^+_{V_t,Z_t}(t), \widehat{G}^+_{V_t,U_t}(t)\right) \tag{42}$$

$$\overset{(e)}{\geq} \min\left(\frac{1}{2}G_{I(t)}, \widehat{G}^+_{V_t,U_t}(t)\right),$$

where **(a)** and **(c)** are obtained by the definition $\dot{G}$, **(b)** follows from $\dot{\mu}_{U_t} < \dot{\mu}_{V_t} < \dot{\mu}_{Z_t}$ and an analogue of Proposition 3 with strict inequality in case of $\dot{G}_{V_t,U_t} \neq \dot{G}_{Z_t,V_t}$, **(d)** uses equation (39), and **(e)** is by equation (37).

Now let us assume $\widehat{G}^+(t) = \widehat{G}^+_{V_t,U_t}(t) < \frac{1}{2}G_{I(t)}$, from which we will derive a contradiction. From equation (37) and by definition of $\dot{G}$, we have $\widehat{G}^+(t) = \widehat{G}^+_{V_t,U_t}(t) < \frac{1}{2}G_{I(t)} \leq \dot{G}_{Z_t,V_t}$, so we have strict inequality in **(b)** of equation (40). Consequently, we can derive from (40) the contradiction

$$\widehat{G}^+_{Z_t,U_t}(t) > \min\left(\widehat{G}^+_{V_t,U_t}(t), \frac{1}{2}G_{I(t)}\right) \geq \widehat{G}^+_{V_t,U_t}(t) = \max_{V\in\mathcal{C}} \widehat{G}^+_{V,U_t}(t),$$

which completes the proof of $\widehat{G}^+(t) \geq \frac{1}{2}G_{I(t)}$. $\square$

The following Lemma shows that any set U in $\mathcal{U}_t$ also is in $\mathcal{U}'_t$.

Lemma 9. *On the event ξ, for all time steps t, $\mathcal{U}_t \subset \mathcal{U}'_t$.*

Proof. It is obviously sufficient to show that the threshold $-\mathcal{T}_{U,V}(t) \le 0$. On the event ξ, we have $\max_{W\in\mathcal{C}} \widehat{G}^+_{W,U}(t) \ge \widehat{G}^+_{U^*,U}(t) \ge G_{U^*,U} \ge 0$. By definition of $\mathcal{T}_{U,V}(t)$, this implies that $-\mathcal{T}_{U,V}(t) \le 0$. □

The following Lemma gives an upper bound on $\widehat{G}^+(t)$.

Lemma 10. *Assume that event ξ holds. Then $8\beta_{I(t)}(t-1) \ge \widehat{G}^+(t)$ for all t.*

Proof. First note that

$$\widehat{G}^+_{U_t,V_t}(t) \overset{(a)}{>} -\frac{1}{2}\max_{W\in\mathcal{C}} \widehat{G}^+_{W,U_t}(t) \overset{(b)}{=} -\frac{1}{2}\widehat{G}^+_{V_t,U_t}(t), \tag{43}$$

where **(a)** follows from $U_t \in \mathcal{U}'_t$ and the definition of $\mathcal{U}'_t$ and **(b)** from the definition of U_t and V_t. Moreover, we have

$$\begin{aligned}\widehat{G}^+_{V_t,U_t}(t) &\overset{(a)}{=} \frac{2}{\bar{d}_{U_t,V_t}} \sum_{i\in U_t\oplus V_t} \beta_i(t-1) - \widehat{G}^+_{U_t,V_t}(t)\\ &\overset{(b)}{\le} \frac{2}{\bar{d}_{U_t,V_t}} \sum_{i\in U_t\oplus V_t} \beta_i(t-1) + \frac{1}{2}\widehat{G}^+_{V_t,U_t}(t),\end{aligned}$$

where **(a)** is because $\widehat{\Delta}^+_{V_t,U_t}(t) + \Delta^+_{U_t,V_t}(t) = 2\sum_{i\in U_t\oplus V_t} \beta_i(t-1)$, and **(b)** is because of equation (43). Hence, we obtain

$$\widehat{G}^+(t) = \widehat{G}^+_{V_t,U_t}(t) \le \frac{4}{\bar{d}_{U_t,V_t}} \sum_{i\in U_t\oplus V_t} \beta_i(t-1). \tag{44}$$

Moreover, as we will demonstrate in the following, we have

$$\frac{1}{\bar{d}_{U_t,V_t}} \sum_{i\in U_t\oplus V_t} \beta_i(t-1) \le 2\beta_{I(t)}(t-1). \tag{45}$$

We show this in detail for the case when $I(t) \in U_t$, the case $I(t) \in V_t$ is similar. For $I(t) \in U_t$, we have $\sum_{i\in U_t\setminus V_t} \beta_i(t-1) \ge \sum_{i\in V_t\setminus U_t} \beta_i(t-1)$ and consequently,

$$\sum_{i\in U_t\oplus V_t} \beta_i(t-1) = \sum_{i\in U_t\setminus V_t} \beta_i(t-1) + \sum_{i\in V_t\setminus U_t} \beta_i(t-1) \le 2\sum_{i\in U_t\setminus V_t} \beta_i(t-1).$$

Since for $I(t) \in U_t$, we have for all $i \in U_t\setminus V_t$ that $\beta_i(t-1) \le \beta_{I(t)}(t-1)$, we therefore obtain

$$\sum_{i\in U_t\oplus V_t} \beta_i(t-1) \le 2\sum_{i\in U_t\setminus V_t} \beta_i(t-1) \le 2d_{U_t,V_t}\beta_{I(t)}(t-1)$$

and consequently,

$$\frac{1}{\bar{d}_{U_t,V_t}} \sum_{i\in U_t\oplus V_t} \beta_i(t-1) \le 2\,\frac{d_{U_t,V_t}}{\bar{d}_{U_t,V_t}}\beta_{I(t)}(t-1) \le 2\beta_{I(t)}(t-1),$$

which proves equation (45). Finally, combining (44) and (45) gives the claim of the lemma. □

Finally, we are ready to give the proof of Theorem 2.

Proof of Theorem 2. First note that by Lemma 9, on event ξ we have $\mathcal{U}_t \subseteq \mathcal{U}'_t$, so the algorithm is well-defined, since as long as $|\mathcal{U}_t| > 1$ also $|\mathcal{U}'_t| > 1$ and there is always an arm to be pulled.

Next, we show that with probability of at least $1-\delta$, the algorithm returns the optimal set U^*. Indeed, assume that ξ holds and that U^* is rejected from $\mathcal{U}_t$ at some step t. Then there exists a set V such that $0 \geq \widehat{\Delta}^+_{U^*,V}(t) = \mu^+_{U^*}(t) - \mu^-_V(t) \geq \mu_{U^*} - \mu_V > 0$, which is a contradiction. Hence, the claim follows from Lemma 7.

Finally, let us consider the sample complexity of our algorithm. By Lemma 10 and Lemma 8 we have for any pulled arm i that $8\beta_i(t) \geq \widehat{G}^+(t) \geq \frac{1}{2}G_i$. Summing over all arms $i \in \mathcal{K}$, this gives for each t that $T_i(t) \leq 128H_i \log(4Kt^2/\delta)$. Therefore,

$$\sum_{i \in \mathcal{K}} T_i(t) = t \leq \sum_{i \in \mathcal{K}} 128H_i \log(4Kt^2/\delta) \leq 128H \log(4Kt^2/\delta).$$

Thus, as soon as t reaches $t \geq 128H \log(4Kt^2/\delta)$, the algorithm stops. Denoting this step by $\tilde{n}$ and using Lemma 8 of Antos et al. [2010] in order to solve this equation gives $\tilde{n} \leq O\left(H \log(HK/\delta)\right)$. $\square$