

Pure Exploration of Combinatorial Bandits

Shouyuan Chen

April 13, 2014

1 Preliminaries

1.1 Problems

Let n be the number of base arms. Let $\mathcal{M} \subseteq 2^{[n]}$ be the set of super arms.

In this note, we consider the following cases of \mathcal{M} .

Example 1 (Explore- m). $\mathcal{M}_{\text{TOP}_m}(n) = \{M \subseteq [n] \mid |M| = m\}$. This corresponds to finding the top m arms from $[n]$.

Example 2 (Explore- m -bandits). Suppose $n = mk$. Then $\mathcal{M}_{\text{BANDIT}_m}(n)$ contains all subsets $M \subseteq [n]$ with size m , such that

$$M \cap \{ik + 1, \dots, (i+1)k\} = 1, \quad \text{for all } i \in \{0, \dots, m-1\}.$$

This corresponds to finding the top arms from m bandits, where each bandit has k arms.

Example 3 (Perfect Matching). Let $G = (V, E)$ be a bipartite graph and $|E| = n$. For simplicity, let each edge $e \in E$ corresponds to a unique integer $i \in [n]$, and vice versa. Then $\mathcal{M}_{\text{MATCH}}(n, G)$ contains all subsets $M \subseteq [n]$ such that M corresponds to a perfect matching in G .

1.2 Diff-Sets

Definition 1 (Diff-set). An n -diff-set (or diff-set in short) is a pair of sets $c = (c_+, c_-)$, where $c_+ \subseteq [n]$, $c_- \subseteq [n]$ and $c_+ \cap c_- = \emptyset$.

Definition 2 (Difference of sets). Given any $M_1 \subseteq [n], M_2 \subseteq [n]$. We define $M_1 \ominus M_2 \triangleq C$, where $C = (C_+, C_-)$ is a diff-set and $C_+ = M_1 \setminus M_2$ and $C_- = M_2 \setminus M_1$.

Definition 3. Denote $\text{diff}[n]$ be the set of all possible n -diff-sets.

Definition 4 (Set operations of diff-sets). Let $C = (C_+, C_-), D = (D_+, D_-)$ be two diff-sets. We define $C \cap D \triangleq (C_+ \cap D_+, C_- \cap D_-)$ and $C \setminus D \triangleq (C_+ \setminus D_+, C_- \setminus D_-)$.

Further, for all $e \in [n]$, $e \in C \Leftrightarrow (e \in C_+) \vee (e \in C_-)$. And $|C| \triangleq |C_+| + |C_-|$.

Definition 5 (Valid diff-set). Given a set $M \subseteq [n]$ and a diff-set $C = (C_+, C_-)$, we call C a valid diff-set for M , iff $C_+ \cap M = \emptyset$ and $C_- \subseteq M$. In this case, we denote $C \prec M$.

Definition 6 (Negative diff-set). Given a diff-set $A = (A_+, A_-)$, we define $\neg A = (A_-, A_+)$.

1.2.1 diff-set operations

Definition 7 (Operators \oplus and \ominus). Given any $M \subseteq [n]$ and $C \in \text{diff}[n]$. If $C \prec M$, we define operator \oplus such that $M \oplus C \triangleq M \setminus C_- \cup C_+$. On the other hand if $\neg C \prec M$, we define operator \ominus such that $M \ominus C \triangleq M \oplus (\neg C) = M \setminus C_+ \cup C_-$.

Definition 8. Given two diff-sets $A = (A_+, A_-)$ and $B = (B_+, B_-)$. We denote $B \prec A$, if and only if $B_+ \cap A_+ = \emptyset$ and $A_+ \cap A_- = \emptyset$.

Definition 9. Given two diff-sets $A = (A_+, A_-)$ and $B = (B_+, B_-)$. If $B \prec A$, we define $A \oplus B = ((A_+ \cup B_+) \setminus (A_- \cup B_-), (A_- \cup B_-) \setminus (A_+ \cup B_+))$.

Lemma 1. Given two diff-sets $A = (A_+, A_-)$ and $B = (B_+, B_-)$. If $B \prec A$, then $A \oplus B$ is a diff-set.

Proof. Let $C = A \oplus B$. By definition, we have $C_+ = (A_+ \cup B_+) \setminus (A_- \cup B_-)$ and $C_- = (A_- \cup B_-) \setminus (A_+ \cup B_+)$.

We only need to show that $C_+ \cap C_- = \emptyset$.

$$\begin{aligned} C_+ \cap C_- &= ((A_+ \cup B_+) \setminus (A_- \cup B_-)) \cap ((A_- \cup B_-) \setminus (A_+ \cup B_+)) \\ &= (A_+ \cup B_+) \cap ((A_- \cup B_-) \setminus (A_+ \cup B_+)) \setminus (A_- \cup B_-) \\ &= \emptyset. \end{aligned}$$

□

Lemma 2. Given two diff-sets $A = (A_+, A_-)$ and $B = (B_+, B_-)$. If there exists $M \subseteq [n]$ such that $A \prec M$, and $B \prec (M \oplus A)$, then $B \prec A$ and $(M \oplus A \oplus B) \ominus M = A \oplus B$.

Proof. We first show that $B \prec A$. Since $B \prec (M \oplus A)$, we know that $B_+ \cap (M \setminus A_- \cup A_+) = \emptyset$. Therefore, we have

$$\begin{aligned} \emptyset &= B_+ \cap (M \setminus A_- \cup A_+) \\ &= (B_+ \cap (M \setminus A_-)) \cup (B_+ \cap A_+) \end{aligned}$$

We see that $B_+ \cap A_+ = \emptyset$.

On the other hand, we have $B_- \subseteq (M \setminus A_- \cup A_+)$, therefore

$$\begin{aligned} B_- \cap A_- &\subseteq (M \setminus A_- \cup A_+) \cap A_- \\ &= (M \setminus A_- \cap A_-) \cup (A_+ \cap A_-) \\ &= \emptyset. \end{aligned}$$

Hence we proved that $B \prec A$.

Define $D = (M \oplus A \oplus B) \ominus M$ and write $D = (D_+, D_-)$. Then,

$$\begin{aligned} D_+ &= (M \oplus A \oplus B) \setminus M \\ &= (M \setminus A_- \cup A_+ \setminus B_- \cup B_+) \setminus M \\ &= (A_+ \cup B_+) \setminus (A_- \cup B_-). \end{aligned}$$

Similarly, we have

$$\begin{aligned} D_- &= M \setminus (M \oplus A \oplus B) \\ &= M \setminus (M \setminus A_- \cup A_+ \setminus B_- \cup B_+) \\ &= (A_- \cup B_-) \setminus (A_+ \cup B_+). \end{aligned}$$

□

1.2.2 Diff-set class

Definition 10 (Decomposition of diff-set). Given $\mathcal{B} \subseteq \text{diff}[n]$ and $D \in \text{diff}[n]$, a decomposition of D on \mathcal{B} is a set $\{b_1, \dots, b_k\} \subseteq \mathcal{B}$ satisfying the following

1. For all $i \in [k]$ and $j \in [k]$, we write $b_i = (b_i^+, b_i^-)$ and $b_j = (b_j^+, b_j^-)$. Then, the following holds
 $b_i^+ \cap b_j^+ = \emptyset$, $b_i^+ \cap b_j^- = \emptyset$, $b_i^- \cap b_j^+ = \emptyset$ and $b_i^- \cap b_j^- = \emptyset$.
2. $D = b_1 \oplus b_2 \oplus \dots \oplus b_k$.

Lemma 3. Given $\mathcal{B} \subseteq \text{diff}[n]$ and $D \in \text{diff}[n]$. Let $\{b_1, \dots, b_k\} \subseteq \mathcal{B}$ be a decomposition of D on \mathcal{B} . Then,

1. Let $D = (D_+, D_-)$ and for all $i \in [k]$, we write $b_i = (b_i^+, b_i^-)$. Then $D_+ = b_1^+ \cup \dots \cup b_k^+$ and $D_- = b_1^- \cup \dots \cup b_k^-$.
2. For all $M \subseteq [n]$, if $D \prec M$, then, for all $i \in [k]$, we have $b_i \prec M$.

Proof. We prove (1) by induction. Let $D_i = b_1 \oplus \dots \oplus b_i$ and write $D_i = (D_i^+, D_i^-)$. We show that $D_i^+ = \bigcup_{j=1}^i b_j^+$ and $D_i^- = \bigcup_{j=1}^i b_j^-$ for all $i \in [k]$. For $i = 1$, this is trivially true. Then, assume that this is true for some $i > 1$. By definition $D_{i+1} = D_i \oplus b_{i+1}$, hence $D_{i+1}^+ = (D_i^+ \cup b_{i+1}^+) \setminus (D_i^- \cup b_{i+1}^-)$. Note that

$$\begin{aligned} (D_i^- \cup b_{i+1}^-) \cap (D_i^+ \cup b_{i+1}^+) &= (D_i^- \cap D_i^+) \cup (D_i^- \cap b_{i+1}^+) \cup (b_{i+1}^- \cap D_i^+) \cup (b_{i+1}^- \cap b_{i+1}^+) \\ &= (D_i^- \cap b_{i+1}^+) \cup (b_{i+1}^- \cap D_i^+) \\ &= \left(\left(\bigcup_{j=1}^i b_j^- \right) \cap b_{i+1}^+ \right) \cup \left(\left(\bigcup_{j=1}^i b_j^+ \right) \cap b_{i+1}^- \right) \\ &= \emptyset. \end{aligned}$$

Hence $D_{i+1}^+ = D_i^+ \cup b_{i+1}^+$. We can use the same method to show that $D_{i+1}^- = D_i^- \cup b_{i+1}^-$.

Next, we prove (2) using (1). To show that $b_i \prec M$, we only need to show that $b_i^+ \cap M = \emptyset$ and $b_i^- \subseteq M$. Since $D \prec M$, we know that $D_+ \cap M = \emptyset$ and $D_- \subseteq M$. By (1), we see that $b_i^+ \subseteq D_+$ and $b_i^- \subseteq D_-$. Therefore, we have $(b_i^+ \cap M) \subseteq (D_+ \cap M) = \emptyset$ and $b_i^- \subseteq D_- \subseteq M$. \square

Definition 11 (diff-set class). Given $\mathcal{M} \subseteq 2^{[n]}$. $\mathcal{B} \subseteq \text{diff}[n]$ is a diff-set class for \mathcal{M} , if the following hold.

1. $(\emptyset, \emptyset) \notin \mathcal{B}$.
2. For all $M \in \mathcal{M}$ and for all $b \in \mathcal{B}$, if $b \prec M$, then $M \oplus b \in \mathcal{M}$.
3. For all $M_1 \in \mathcal{M}$ and $M_2 \in \mathcal{M}$, where $M_1 \neq M_2$. Let $D = M_1 \ominus M_2$. Then, there exists a decomposition of D on \mathcal{B} .

Definition 12 (Rank of diff-set class). Let $\mathcal{B} \subseteq [n]$ be a diff-set class for some \mathcal{M} . We define

$$\text{rank}(\mathcal{B}) \triangleq \max_{b \in \mathcal{B}} |b|.$$

Example 4 (diff-set class for Explore-m). One diff-set class \mathcal{B} for $\mathcal{M}_{\text{TOP}_m}(n)$ is given by

$$\mathcal{B} = \{(\{b_1\}, \{b_2\}) \mid b_1 \neq b_2, b_1 \in [n], b_2 \in [n]\}.$$

Proof omitted. Further, we see that $\text{rank}(\mathcal{B}) = 2$.

Example 5 (diff-set class for Explore-m-badit). Let $n = mk$. One diff-set class \mathcal{B} for $\mathcal{M}_{\text{BANDIT}_m}(n)$ is given by

$$\mathcal{B} = \{(\{b_1\}, \{b_2\}) \mid b_1 \neq b_2, \exists i \in \{0, \dots, k-1\}, b_1 \in \{ik+1, \dots, (i+1)k\}, b_2 \in \{ik+1, \dots, (i+1)k\}\}.$$

Proof omitted. Further, we see that $\text{rank}(\mathcal{B}) = 2$.

Example 6 (diff-set class for Perfect Matching). *One diff-set class \mathcal{B} for $\mathcal{M}_{\text{MATCH}}(n, G)$ is the set of all augmenting cycles of G . More specifically,*

$$\mathcal{B} = \{(b_+, b_-) | b_+ \cup b_- \text{ is a cycle of } G\}.$$

Note $\text{rank}(\mathcal{B}) \leq n$.

1.3 Weights and confidence bounds

Definition 13 (Weight functions). *Define function $w : [n] \rightarrow \mathbb{R}^+$ which represents the weight of each base arm. Further, we slight abuse the notations, and extend the definition of w to diff-sets and sets as follows.*

1. For all $M \subseteq [n]$, we denote $w(M) = \sum_{e \in M} w(e)$.
2. For all $b = (b_+, b_-) \in \text{diff}[n]$, we denote $w(b) = \sum_{e \in b_+} w(e) - \sum_{e \in b_-} w(e)$.

Lemma 4. *Let $c \in \text{diff}[n], d \in \text{diff}[n]$. Let w be a weight function. Then,*

$$w(c \cup d) = w(c) + w(d) - w(c \cap d). \quad (1)$$

Proof. Let $c = (c_+, c_-)$ and $d = (d_+, d_-)$. We have

$$w(c \cup d) = w(c_+ \cup d_+) - w(c_- \cup d_-) \quad (2)$$

$$= w(c_+) + w(d_+) - w(c_+ \cap d_+) - w(c_-) - w(d_-) + w(c_- \cap d_-) \quad (3)$$

$$= w(c) + w(d) - (w(c_+ \cap d_+) - w(c_- \cap d_-)) \quad (4)$$

$$= w(c) + w(d) - w(c \cap d). \quad (5)$$

□

Definition 14 (Mean weight \bar{w}_t , sample size n_t). *Given $t > 0$. Define \bar{w}_t be a weight function such that, for all $e \in [n]$, $\bar{w}_t(e)$ equals to the empirical mean of e up to round t . Let $n_t : [n] \rightarrow \mathbb{N}$, such that $n_t(e)$ equals to number of plays of base arm e up to round t .*

Definition 15 (Confidence radius rad_t). *Given n and $t > 0$. Define $\text{rad}_t : [n] \rightarrow \mathbb{R}^+$ satisfying, for all $e \in [n]$,*

$$\text{rad}_t(e) = c_{\text{rad}} \log \left(\frac{c_\delta n t^2}{\delta} \right) \frac{1}{\sqrt{n_t(e)}},$$

where $c_{\text{rad}} > 0$ and $c_\delta > 0$ are some universal constants (specify later) and $\delta > 0$ is a parameter.

We extend the notation of rad_t to diff-sets and sets as follows.

1. For all $M \subseteq [n]$, $\text{rad}_t(M) \triangleq \sum_{e \in M} \text{rad}_t(e)$.
2. For all $b = (b_+, b_-) \in \text{diff}[n]$, $\text{rad}_t(b) \triangleq \text{rad}_t(b_+) + \text{rad}_t(b_-)$.

Definition 16 (UCB w_t^+). *Define $w_t^+ : [n] \rightarrow \mathbb{R}^+$, s.t., for all $e \in [n]$,*

$$w_t^+(e) = \bar{w}_t(e) + \text{rad}_t(e).$$

We extend the notation of w_t^+ to diff-sets and sets as follows.

1. For all $M \subseteq [n]$, $w_t^+(M) \triangleq \bar{w}_t(M) + \text{rad}_t(M)$.
2. For all $b = (b_+, b_-) \in \text{diff}[n]$, $w_t^+(b) \triangleq \bar{w}_t(b) + \text{rad}_t(b)$.

Lemma 5. Define random event

$$\xi = \{\forall e \in [n] \forall t > 0, |\bar{w}_t(e) - w(e)| \leq \text{rad}_t(e)\}.$$

Then, there exist constants c_{rad} and c_δ ,

$$\Pr[\xi] \geq 1 - \delta.$$

Proof. Hoeffding inequality and union bound. □

Corollary 1.

$$\xi \implies \forall t, \forall e \in [n] \ w_t^+(e) \geq w(e).$$

$$\xi \implies \forall t, \forall M \subseteq [n], \ w_t^+(M) \geq w(M).$$

$$\xi \implies \forall t, \forall b \in \text{diff}[n] \ w_t^+(b) \geq w(b).$$

1.4 Properties of rad_t

Lemma 6. Let $c \in \text{diff}[n], d \in \text{diff}[n]$. Then

$$\text{rad}_t(c \setminus d) = \text{rad}_t(c) - \text{rad}_t(c \cap d). \quad (6)$$

Proof. Let $c = (c_+, c_-)$ and $d = (d_+, d_-)$. We have

$$\begin{aligned} \text{rad}_t(c \setminus d) &= \text{rad}_t(c_+ \setminus d_+) + \text{rad}_t(c_- \setminus d_-) \\ &= \text{rad}_t(c_+) - \text{rad}_t(c_+ \cap d_+) + \text{rad}_t(c_-) - \text{rad}_t(c_- \cap d_-) \\ &= \text{rad}_t(c) - \text{rad}_t(c \cap d). \end{aligned}$$

□

Lemma 7. Let $C = (C_+, C_-)$ and $D = (D_+, D_-)$ be two diff-sets. If $D \prec C$, then

$$\text{rad}_t(C \oplus D) = \text{rad}_t(C) + \text{rad}_t(D) - 2\text{rad}_t(C_+ \cap D_-) - 2\text{rad}_t(C_- \cap D_+).$$

In addition, if $\neg D \prec C$, then

$$\text{rad}_t(C \ominus D) = \text{rad}_t(C) + \text{rad}_t(D) - 2\text{rad}_t(C_+ \cap D_+) - 2\text{rad}_t(C_- \cap D_-).$$

Proof. We prove the first part of the lemma. The second part follows from the first part and the definition of $\neg D$.

By definition, we have $C \oplus D = ((C_+ \cup D_+) \setminus (C_- \cup D_-), (C_- \cup D_-) \setminus (C_+ \cup D_+))$. Hence, we have

$$\text{rad}_t((C_+ \cup D_+) \setminus (C_- \cup D_-)) = \text{rad}_t(C_+ \cup D_+) - \text{rad}_t((C_+ \cup D_+) \cap (C_- \cup D_-)) \quad (7)$$

$$= \text{rad}_t(C_+) + \text{rad}_t(D_+) - \text{rad}_t((C_+ \cup D_+) \cap (C_- \cup D_-)), \quad (8)$$

where the second equality holds due to $C_+ \cap D_+ = \emptyset$ by the definition of $D \prec C$.

Similarly, we have

$$\text{rad}_t((C_- \cup D_-) \setminus (C_+ \cup D_+)) = \text{rad}_t(C_-) + \text{rad}_t(D_-) - \text{rad}_t((C_+ \cup D_+) \cap (C_- \cup D_-)).$$

Combine both equalities, we have

$$\text{rad}_t(C \oplus D) = \text{rad}_t((C_+ \cup D_+) \setminus (C_- \cup D_-)) + \text{rad}_t((C_- \cup D_-) \setminus (C_+ \cup D_+)) \quad (9)$$

$$= \text{rad}_t(C_+) + \text{rad}_t(D_+) + \text{rad}_t(C_-) + \text{rad}_t(D_-) - 2\text{rad}_t((C_+ \cup D_+) \cap (C_- \cup D_-)) \quad (10)$$

$$= \text{rad}_t(C) + \text{rad}_t(D) - 2\text{rad}_t((C_+ \cup D_+) \cap (C_- \cup D_-)). \quad (11)$$

□

2 Algorithm and Main Results

2.1 Algorithm

1. Input Parameter: $\delta \in (0, 1)$.
2. For $t = 1, \dots$,
3. Maintain \bar{w}_t and rad_t .
4. Let $M_t = \arg \max_{M \in \mathcal{M}} \bar{w}_t(M)$.
5. Let $D = \arg \max_{C \in \text{diff}[n], C \prec M_t} w_t^+(C)$.
6. If $w_t^+(D) \leq 0$. Then stop and return M_t .
7. Otherwise, find $p_t = \arg \min_{e \in D} \text{rad}_t(e)$.
8. Play p_t and observe outcome x_t .
9. Go back to step 2.

The step 5 of above procedure can be implemented by:

1. Let $M_t^+ = \arg \max_{M \in \mathcal{M}} \tilde{w}_t(M)$, where \tilde{w}_t is a weight function defined by:
 - (a) $\forall e \in M_t, \tilde{w}_t(e) = \bar{w}_t(e) - \text{rad}_t(e)$.
 - (b) $\forall e \notin M_t, \tilde{w}_t(e) = \bar{w}_t(e) + \text{rad}_t(e)$.
2. $D = M_t^+ \ominus M_t$

2.2 Main result

Definition 17 (Optimal diff-sets). *Given a diff-set class \mathcal{B} and the optimal set M_* . We define \mathcal{B}_{opt} as a subset of \mathcal{B} , and for all $b \in \mathcal{B}$, $b \in \mathcal{B}_{\text{opt}}$ if and only if, there exists $M \neq M_*$ and $M_* \ominus M$ can be decomposed as b, b_1, \dots, b_k on \mathcal{B} .*

Definition 18 (Hardness Δ_e of base arm e). *For each $e \in [n]$, we define its hardness Δ_e as follows*

$$\Delta_e = \min_{b \in \mathcal{B}_{\text{opt}}, e \in b} \frac{1}{\text{rank}(\mathcal{B})} w(b).$$

Definition 19 (Sufficient exploration). *For all $t > 0$, we define $E_t^3 \subseteq [n]$, such that, for all $e \in [n]$ $e \in E_t^3$ if and only if $\text{rad}_t(e) < \frac{1}{3}\Delta_e$.*

Corollary 2. *For all $t > 0$ and $e \in [n]$*

$$n_t(e) \geq O\left(\frac{1}{\Delta_e^2} \log(\Delta_e n / \delta)\right) \implies e \in E_t^3.$$

Theorem 1. *With probability at least $1 - \delta$, the algorithm returns M_* , and the number of samples used by the algorithm are at most*

$$\sum_{e \in [n]} \Delta_e^{-2} \log(\Delta_e n / \delta).$$

3 Proof of Main Results

Unless specified, we shall assume the random event ξ (defined in Lemma 5) holds in all the following proofs.

Lemma 8. *For any $t > 0$, if the algorithm terminates on round t , then $M_t = M_*$.*

Proof. Suppose $M_t \neq M_*$. Then $w(M_*) > w(M_t)$. Then, there exists $b \in \mathcal{B}$ such that $b \prec M_t$ and $w(b) > 0$. On the other hand, by Corollary 1, we have $w_t^+(b) > w(b)$. Hence $w_t^+(b) > 0$. This contradicts to the stopping condition of our algorithm. \square

Lemma 9. *For any $t > 0$. If $e \in E_t^3$, then $p_t \neq e$.*

Proof. Suppose that $p_t = e$. Let $D = M_t^+ \ominus M_t$. Let c, c_1, \dots, c_k be decomposition of D on \mathcal{B} . And since \mathcal{B} is a diff-set class, such decomposition exists. Assume, without loss of generality, that $e \in c$.

By Lemma Y, we know that

$$D_+ = c_+ \cup c_1^+ \cup \dots \cup c_k^+ \quad \text{and} \quad D_- = c_- \cup c_1^- \cup \dots \cup c_k^-. \quad (12)$$

We also denote $K = \text{rank}(\mathcal{B})$.

Case (1). Suppose that $c \in \mathcal{B}_{\text{opt}}$. Then $w(c) > 0$. Since $e \in E_t^3$, we have $\text{rad}_t(e) \leq \frac{1}{3}\Delta_e \leq \frac{1}{3K}w(c)$. In addition, $\forall g \in c_t, g \neq e$, $\text{rad}_t(g) \leq \text{rad}_t(e) \leq \frac{1}{3K}w(c)$. Hence, $\text{rad}_t(c) = \sum_{g \in c_t} \text{rad}_t(g) \leq \frac{|c_t|}{3K}w(c) \leq \frac{1}{3}w(c)$.

Hence, $\bar{w}_t(c) \geq w(c) - \text{rad}_t(c) \geq \frac{2}{3}w(c) > 0$. This means that $\bar{w}_t(M_t \oplus c) = \bar{w}_t(M_t) + \bar{w}_t(c) > \bar{w}_t(M_t)$. Therefore, $M_t \neq \max_{M \in \mathcal{M}} \bar{w}_t(M)$. This contradicts to the definition of M_t .

Case (2). Suppose that $c_t \notin \mathcal{B}_{\text{opt}}$. Then, one of the following mutually exclusive cases must hold.

Case (2.1). $(e \in M_* \wedge e \in c_+)$ or $(e \notin M_* \wedge e \in c_-)$.

Let the decomposition of $M_* \ominus (M_t \oplus D \ominus c)$ on \mathcal{B} be b, b_1, \dots, b_l , which exists due to \mathcal{B} is a diff-set class. Assume wlog that $e \in b$. We write $b = (b_+, b_-)$. It is easy to see that $b \in \mathcal{B}_{\text{opt}}$.

Define $\tilde{D} = (M_t \oplus D \ominus c) \ominus M_t$ and $D' = (M_t \oplus \tilde{D} \oplus b) \ominus M_t$. By Lemma 2, we know that $\tilde{D} = D \ominus c$ and $D' = \tilde{D} \oplus b$. We also write $\tilde{D} = (\tilde{D}_+, \tilde{D}_-)$ and $D' = (D'_+, D'_-)$. By definition, we have

$$\begin{aligned} \tilde{D}_+ &= (D_+ \cup c_-) \setminus (D_- \cup c_+) \\ &= (D_+ \cup c_- \setminus D_-) \cap (D_+ \cup c_- \setminus c_+) \\ &= D_+ \cap (D_+ \setminus c_-) \\ &= D_+ \setminus c_+. \end{aligned}$$

By the same method, we are able to show that $\tilde{D}_- = D_- \setminus c_-$. Therefore we have

$$\tilde{D}_+ \subseteq D_+ \quad \text{and} \quad \tilde{D}_- \subseteq D_-. \quad (13)$$

First, we show that $\text{rad}_t(c) \leq \frac{1}{3}w(b)$. Since $e \in E_t^3$, $e \in b$ and $b \in \mathcal{B}_{\text{opt}}$, we have $\text{rad}_t(e) \leq \frac{1}{3}\Delta_e \leq \frac{1}{3K}w(b)$. In addition, $\forall g \in c, g \neq e$, $\text{rad}_t(g) \leq \text{rad}_t(e) \leq \frac{1}{3K}w(b)$. Hence,

$$\begin{aligned} \text{rad}_t(c) &= \sum_{g \in c} \text{rad}_t(g) \\ &\leq \frac{|c|}{3K}w(b) \\ &\leq \frac{1}{3}w(b). \end{aligned} \quad (14)$$

Now, we show that $\text{rad}_t(\tilde{D}_+ \cap b_-) + \text{rad}_t(\tilde{D}_- \cap b_+) \leq \frac{1}{3}w(b)$. Since Eq. (13), we have $\forall g \in (\tilde{D}_+ \cap b_-) \cup$

$(\tilde{D}_- \cap b_+), g \neq e$, $\text{rad}_t(g) \leq \text{rad}_t(e) \leq \frac{1}{3K}w(b)$. Note that $|\tilde{D}_+ \cap b_-| + |\tilde{D}_- \cap b_+| \leq |b_+| + |b_-| \leq K$. Hence,

$$\begin{aligned} \text{rad}_t(\tilde{D}_+ \cap b_-) + \text{rad}_t(\tilde{D}_- \cap b_+) &= \sum_{g \in (\tilde{D}_+ \cap b_-) \cup (\tilde{D}_- \cap b_+)} \text{rad}_t(g) \\ &\leq \frac{K}{3K}w(b) \\ &\leq \frac{1}{3}w(b). \end{aligned} \quad (15)$$

Then, we have

$$\text{rad}_t(D') - \text{rad}_t(D) = \text{rad}_t(\tilde{D} \oplus b) - \text{rad}_t(D) \quad (16)$$

$$= \text{rad}_t(\tilde{D}) + \text{rad}_t(b) - 2\text{rad}_t(\tilde{D}_+ \cap b_-) - 2\text{rad}_t(\tilde{D}_- \cap b_+) - \text{rad}_t(D) \quad (17)$$

$$= \text{rad}_t(D \ominus c) + \text{rad}_t(b) - 2\text{rad}_t(\tilde{D}_+ \cap b_-) - 2\text{rad}_t(\tilde{D}_- \cap b_+) - \text{rad}_t(D) \quad (18)$$

$$\begin{aligned} &= \text{rad}_t(D) + \text{rad}_t(c) + \text{rad}_t(b) - 2\text{rad}_t(D_+ \cap c_+) - 2\text{rad}_t(D_- \cap c_-) \\ &\quad - 2\text{rad}_t(\tilde{D}_+ \cap b_-) - 2\text{rad}_t(\tilde{D}_- \cap b_+) - \text{rad}_t(D) \end{aligned} \quad (19)$$

$$\begin{aligned} &= \text{rad}_t(D) + \text{rad}_t(c) + \text{rad}_t(b) - 2\text{rad}_t(c_+) - 2\text{rad}_t(c_-) \\ &\quad - 2\text{rad}_t(\tilde{D}_+ \cap b_-) - 2\text{rad}_t(\tilde{D}_- \cap b_+) - \text{rad}_t(D) \end{aligned} \quad (20)$$

$$= \text{rad}_t(b) - \text{rad}_t(c) - 2\text{rad}_t(\tilde{D}_+ \cap b_-) - 2\text{rad}_t(\tilde{D}_- \cap b_+), \quad (21)$$

where Eq. (17) and Eq. (19) follow from Lemma 7, and Eq. (20) follows from Eq. (12).

By the definition of D , we have that $w_t^+(D) \geq w_t^+(D')$. This means that

$$\bar{w}_t(D) + \text{rad}_t(D) \geq \bar{w}_t(D') + \text{rad}_t(D') \quad (22)$$

$$= \bar{w}_t(D) - \bar{w}_t(c) + \bar{w}_t(b) + \text{rad}_t(D'). \quad (23)$$

By regrouping the above inequality, we have

$$\bar{w}_t(c) \geq \bar{w}_t(b) + \text{rad}_t(D') - \text{rad}_t(D) \quad (24)$$

$$= \bar{w}_t(b) + \text{rad}_t(b) - \text{rad}_t(c) - 2\text{rad}_t(\tilde{D}_+ \cap b_-) - 2\text{rad}_t(\tilde{D}_- \cap b_+) \quad (25)$$

$$\geq w(b) - \text{rad}_t(c) - 2\text{rad}_t(\tilde{D}_+ \cap b_-) - 2\text{rad}_t(\tilde{D}_- \cap b_+) \quad (26)$$

$$> w(b) - \frac{1}{3}w(b) - \frac{2}{3}w(b) \quad (27)$$

$$= 0, \quad (28)$$

where Eq. (27) follows from Eq. (14) and Eq. (15).

This contradicts to the definition of M_t .

Case (2.2). $(e \in M_* \wedge e \in c_-)$ or $(e \notin M_* \wedge e \in c_+)$.

Let the decomposition of $M_* \ominus (M_t \oplus D)$ on \mathcal{B} be b, b_1, \dots, b_l . Assume wlog that $e \in b$. We write that $b = (b_+, b_-)$. Note that $b \in \mathcal{B}_{\text{opt}}$ and hence $w(b) > 0$.

Define $D' = (M_t \oplus D \oplus b) \ominus M_t$. By Lemma 2, we know that $D' = D \oplus b$.

First, we show that $|D \setminus D'| \leq |b|$. Let $C = D \setminus D'$ and write $C = (C_+, C_-)$. We can bound $|C_+|$ as follows.

$$\begin{aligned} C_+ &= D_+ \setminus D'_+ \\ &= D_+ \setminus ((D_+ \cup b_+) \setminus (D_- \cup b_-)) \\ &= (D_+ \cap (D_- \cup b_-)) \cup (D_+ \setminus (D_+ \cup b_+)) \\ &= D_+ \cap b_-. \end{aligned}$$

Hence, we have $|C_+| \leq |b_-|$. Then, we move to bounding $|C_-|$

$$\begin{aligned}
C_- &= D_- \setminus D'_- \\
&= D_- \setminus ((D_- \cup b_-) \setminus (D_+ \cup b_+)) \\
&= (D_- \cap (D_+ \cup b_+)) \cup (D_- \setminus (D_- \cup b_-)) \\
&= D_- \cap b_+.
\end{aligned}$$

Thus $|C_-| \leq |b_+|$ and we proved that $|D \setminus D'| \leq |b|$.

Next, we show that $\text{rad}_t(D \setminus D') \leq \frac{1}{3}w(b)$. Since $e \in E_t^3$, $e \in b$ and $b \in \mathcal{B}_{\text{opt}}$, we have $\text{rad}_t(e) \leq \frac{1}{3}\Delta_e \leq \frac{1}{3K}w(b)$. In addition, $\forall g \in (D \setminus D'), g \neq e$, $\text{rad}_t(g) \leq \text{rad}_t(e) \leq \frac{1}{3K}w(b)$. Note that $|D \setminus D'| \leq |b| \leq K$. Hence, $\text{rad}_t(D \setminus D') = \sum_{g \in (D \setminus D')} \text{rad}_t(g) \leq \frac{K}{3K}w(b) \leq \frac{1}{3}w(b)$.

We also note that

$$w(D' \setminus D) - w(D \setminus D') = w(D' \setminus D) + w(D' \cap D) - w(D \cap D') - w(D \setminus D') \quad (29)$$

$$= w(D') - w(D) \quad (30)$$

$$= w(b), \quad (31)$$

where we have repeatedly applied Lemma 4.

Then, we show that $w_t^+(D') > w_t^+(D)$.

$$w_t^+(D') - w_t^+(D) = \bar{w}_t(D') - \bar{w}_t(D) + \text{rad}_t(D') - \text{rad}_t(D) \quad (32)$$

$$= \bar{w}_t(D' \setminus D) - \bar{w}_t(D \setminus D') + \text{rad}_t(D' \setminus D) - \text{rad}_t(D \setminus D') \quad (33)$$

$$\geq w(D' \setminus D) - w(D \setminus D') - 2\text{rad}_t(D \setminus D') \quad (34)$$

$$= w(b) - 2\text{rad}_t(D \setminus D') \quad (35)$$

$$> w(b) - \frac{2}{3}w(b) \quad (36)$$

$$= \frac{1}{3}w(b) > 0, \quad (37)$$

where Eq. (33) follows from Lemma 6 and Eq. (34) follows from the fact that $\bar{w}_t(D' \setminus D) + \text{rad}_t(D' \setminus D) \geq w(D' \setminus D)$ and that $\bar{w}_t(D \setminus D') + \text{rad}_t(D \setminus D') \geq w(D \setminus D')$, under the random event ξ .

This contradicts to the fact that D is chosen on round t . \square

4 Lower Bounds

Definition 20 (Hardness of arm). *Given \mathcal{M} , M_* and w . For any $e \in [n]$, we define its hardness Δ_e as follows*

$$\Delta_e = \begin{cases} \min_{M: e \in M, M \in \mathcal{M}} w(M_*) - w(M) & \text{if } e \notin M_*, \\ \min_{M: e \notin M, M \in \mathcal{M}} w(M_*) - w(M) & \text{if } e \in M_*. \end{cases}$$

Lemma 10.

$$\Delta_e = \min_{b: e \in b, b \in \mathcal{B}_{\text{opt}}} w(b).$$

Theorem 2. *Assume that, for each arm $i \in [n]$, its reward distribution is a Gaussian distribution with mean p_i and variance 1. Then, for any $\delta \in (0, e^{-16}/4)$ and any δ -correct algorithm \mathbb{A} . Let T denote the number of total samples used by algorithm \mathbb{A} . We have*

$$\mathbb{E}[T] \geq \sum_e \frac{1}{16\Delta_e^2} \log(1/\delta).$$

Proof. Fix $\delta > 0$, p_i for all $i \in [n]$ and a δ -correct policy \mathbb{A} . Assume that the reward distribution of an arm $i \in [n]$ is a Gaussian distribution with mean p_i and variance 1. Then, for any $e \in [n]$, let T_e denote the number of trials of arm e used by algorithm \mathbb{A} . In the rest of the proof, we will prove that for any $e \in [n]$, the number of trials of arm e is lower-bounded by

$$\mathbb{E}[T_e] \geq \frac{1}{c\Delta_e^2} \log(1/\delta), \quad (38)$$

where c is a constant whose value will be determined later. Notice that the theorem will follow immediately by summing up the above bounds for all $e \in [n]$.

Fix an arm $e \in [n]$. We now focus on proving Eq. (38). Consider two hypothesis H_0 and H_1 . Under each hypothesis, the reward distributions of every arm are still Gaussian distributions with unit variance, but the mean rewards of some arms might be altered. Under hypothesis H_0 , the mean reward of each arm is

$$H_0 : q_l = p_l, \quad \text{for all } l \in [n].$$

And under hypothesis H_1 , the mean reward of each arm is

$$H_1 : q_e = \begin{cases} p_e - 2\Delta_e & \text{if } e \in M_* \\ p_e + 2\Delta_e & \text{if } e \notin M_* \end{cases} \quad \text{and } q_l = p_l \quad \text{for all } l \neq e.$$

Define M_e be the “next-to-optimal” set as follows

$$M_e = \begin{cases} \arg \max_{M \in \mathcal{M}: e \in M} w(M) & \text{if } e \notin M_*, \\ \arg \max_{M \in \mathcal{M}: e \notin M} w(M) & \text{if } e \in M_*. \end{cases}$$

By definition, we know that $w(M_*) - w(M_e) = \Delta_e$.

Let w_0, w_1 be the weighting functions under H_0, H_1 respectively. Notice that $w_0(M_*) - w_0(M_e) = \Delta_e > 0$. On the other hand, $w_1(M_*) - w_1(M_e) = -\Delta < 0$. This means that under H_1 , M_* is not the optimal set.

For $l \in \{0, 1\}$, we use \mathbb{E}_l and \Pr_l to denote the expectation and probability, respectively, under the hypothesis H_l .

Define $\theta = 4\delta$. Define

$$t_e^* = \frac{1}{c\Delta_e^2} \log\left(\frac{1}{\theta}\right). \quad (39)$$

Recall that T_e denotes the total number of samples of arm e . Define the event $\mathcal{A} = \{T_e \leq 4t_e^*\}$.

First, we show that $\Pr_0[\mathcal{A}] \geq 3/4$. This can be proved by Markov inequality as follows.

$$\begin{aligned} \Pr_0[T_e > 4t_e^*] &\leq \frac{\mathbb{E}_0[T_e]}{4t_e^*} \\ &= \frac{t_e^*}{4t_e^*} = \frac{1}{4}. \end{aligned}$$

Let X_1, \dots, X_{T_e} denote the sequence of reward outcomes of arm e . We define $K_t(e)$ as the sum of outcomes of arm e up to round t , i.e. $K_t(e) = \sum_{i \in [t]} X_i$. Next, we define the event

$$\mathcal{C} = \left\{ \max_{1 \leq t \leq 4t_e^*} |K_t(e) - p_e t| < \sqrt{t_e^* \log(1/\theta)} \right\}.$$

We now show that $\Pr_0[\mathcal{C}] \geq 3/4$. First, notice that $K_t(e) - p_e t$ is a martingale under H_0 . Then, by

Kolmogorov's inequality, we have

$$\begin{aligned} \Pr_0 \left[\max_{1 \leq t \leq 4t_e^*} |K_t(e) - p_e t| \geq \sqrt{t_e^* \log(1/\theta)} \right] &\leq \frac{\mathbb{E}_0[(K_{4t_e^*}(e) - 4p_e t_e^*)^2]}{t_e^* \log(1/\theta)} \\ &= \frac{4t_e^*}{t_e^* \log(1/\theta)} \\ &< \frac{1}{4}, \end{aligned}$$

where the second inequality follows from the fact that $\mathbb{E}_0[(K_{4t_e^*}(e) - 4p_e t_e^*)^2] = 4t_e^*$; the last inequality follows since $\theta < e^{-16}$.

Then, we define the event \mathcal{B} as the event that the algorithm eventually returns M_* , i.e.

$$\mathcal{B} = \{O = M_*\}.$$

Since the probability of error of the algorithm is smaller than $\delta < 1/4$, we have $\Pr_0[\mathcal{B}] \geq 3/4$. Define \mathcal{S} be $\mathcal{S} = \mathcal{A} \cap \mathcal{B} \cap \mathcal{C}$. Then, by union bound, we have $\Pr_0[\mathcal{S}] \geq 1/4$.

Now, we show that if $\mathbb{E}_0[T_e] \leq t_e^*$, then $\Pr_1[\mathcal{B}] \geq \delta$. Let W be the history of the sampling process until the algorithm stops (including the sequence of arms chosen at each time and the sequence of observed outcomes). Define the likelihood function L_l as

$$L_l(w) = \Pr_l[W = w].$$

Let K be the shorthand of $K_e(T_e)$ and T stands for T_e .

Assume that the event \mathcal{S} occurred. We will bound the likelihood ratio $L_1(W)/L_0(W)$ under this assumption. To do this, we divide our analysis into two different cases.

Case (1): $e \notin M_*$. In this case, the reward distribution of arm e under H_1 is a Gaussian distribution with mean $p_e + 2\Delta_e$ and variance 1. Recall that the probability density function of a Gaussian distribution with mean μ and variance σ^2 is given by $f(x, \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$. Hence, we have

$$\begin{aligned} \frac{L_1(W)}{L_0(W)} &= \prod_{i=1}^{T_e} \exp\left(\frac{-(X_i - p_e - 2\Delta_e)^2 + (X_i - p_e)^2}{2}\right) \\ &= \prod_{i=1}^{T_e} \exp(\Delta_e(2X_i - 2p_e) - 2\Delta_e^2) \\ &= \exp(\Delta_e(2K - 2p_e T_e) - 2\Delta_e^2 T_e) \\ &= \exp(\Delta_e(2K - 2p_e T_e)) \exp(-2\Delta_e^2 T_e). \end{aligned} \tag{40}$$

Next, we bound each individual term on the right-hand side of Eq. (40). We begin with bounding the second term of Eq. (40).

$$\exp(-2\Delta_e^2 T_e) \geq \exp(-8\Delta_e^2 t_e^*) \tag{41}$$

$$= \exp\left(-\frac{8}{c} \log(1/\theta)\right) \tag{42}$$

$$= \theta^{8/c}, \tag{43}$$

where Eq. (41) follows from the assumption that event \mathcal{S} occurred, which implies that event \mathcal{A} occurred and therefore $T_e \leq 4t_e^*$; Eq. (42) follows from the definition of t_e^* .

Then, we bound the first term on the right-hand side of Eq. (40) as follows

$$\exp(\Delta_e(2K - 2p_e T_e)) \geq \exp\left(-2\Delta_e \sqrt{t_e^* \log(1/\theta)}\right) \tag{44}$$

$$= \exp\left(-\frac{2}{\sqrt{c}} \log(1/\theta)\right) \quad (45)$$

$$= \theta^{2/\sqrt{c}}, \quad (46)$$

where Eq. (44) follows from the assumption that event \mathcal{S} occurred, which implies that event \mathcal{C} and therefore $|2K - 2p_e T_e| \leq \sqrt{t_e^* \log(1/\theta)}$; Eq. (45) follows from the definition of t_e^* .

Combining Eq. (43) and Eq. (46), we can bound $L_1(W)/L_0(W)$ for this case as follows

$$\frac{L_1(W)}{L_0(W)} \geq \theta^{8/c+2/\sqrt{c}}. \quad (47)$$

(end of Case (1).)

Case (2): $e \in M_*$. In this case, we know that the mean reward of arm e under H_1 is $p_e - 2\Delta$. Therefore, the likelihood ratio $L_1(W)/L_0(W)$ is given by

$$\begin{aligned} \frac{L_1(W)}{L_0(W)} &= \prod_{i=1}^{T_e} \exp\left(\frac{-(X_i - p_e + 2\Delta_e)^2 + (X_i - p_e)^2}{2}\right) \\ &= \prod_{i=1}^{T_e} \exp(\Delta_e(2p_e - 2X_i) - 2\Delta_e^2) \\ &= \exp(\Delta_e(2p_e T_e - 2K)) \exp(-2\Delta_e^2 T_e). \end{aligned} \quad (48)$$

Notice that the right-hand side of Eq. (48) differs from Eq. (40) only in the first term. Now, we bound the first term as follows

$$\exp(\Delta_e(2K - 2p_e T_e)) \geq \exp(-2\Delta_e \sqrt{t_e^* \log(1/\theta)}) \quad (49)$$

$$= \exp\left(-\frac{2}{\sqrt{c}} \log(1/\theta)\right) \quad (50)$$

$$= \theta^{2/\sqrt{c}}, \quad (51)$$

where the inequalities hold due to reasons similar to Case (1): Eq. (49) follows from the assumption that event \mathcal{S} occurred, which implies that event \mathcal{C} and therefore $|2K - 2p_e T_e| \leq \sqrt{t_e^* \log(1/\theta)}$; Eq. (50) follows from the definition of t_e^* .

Combining Eq. (43) and Eq. (46), we can obtain the same bound of $L_1(W)/L_0(W)$ as in Eq. (47), i.e. $L_1(W)/L_0(W) \geq \theta^{8/c+2/\sqrt{c}}$.

(end of Case (2).)

At this point, we have proved that, if the event \mathcal{S} occurred, then the bound of likelihood ratio Eq. (47) holds, i.e. $\frac{L_1(W)}{L_0(W)} \geq \theta^{8/c+2/\sqrt{c}}$. We can choose the constant c to be sufficiently large such that $8/c+2/\sqrt{c} < 1$. Simple calculation shows that $c \geq 16$ suffices. Hence, we have

$$\begin{aligned} \frac{L_1(W)}{L_0(W)} &\geq \theta \\ &= 4\delta. \end{aligned} \quad (52)$$

Define 1_S as the indicator variable of event \mathcal{S} , i.e. $1_S = 1$ if and only if \mathcal{S} occurs and otherwise $1_S = 0$. Then, we have

$$\frac{L_1(W)}{L_0(W)} 1_S \geq 4\delta 1_S$$

holds regardless the occurrence of event \mathcal{S} . Therefore, we can obtain

$$\begin{aligned}
 \Pr_1[\mathcal{B}] &\geq \Pr_1[\mathcal{S}] = \mathbb{E}_1[1_S] \\
 &= \mathbb{E}_0 \left[\frac{L_1(W)}{L_0(W)} 1_S \right] \\
 &\geq 4\delta \mathbb{E}_0[1_S] \\
 &= 4\delta \Pr_0[\mathcal{S}] > \delta.
 \end{aligned}$$

Now we have proved that, if $\mathbb{E}_0[T_e] \leq t_e^*$, then $\Pr_1[\mathcal{B}] > \delta$. This means that, if $\mathbb{E}_0[T_e] \leq t_e^*$, algorithm \mathbb{A} will choose M_* as the output with probability at least δ , under hypothesis H_1 . However, under H_1 , we have shown that M_* is not the optimal set since $w_1(M_e) > w_1(M_*)$. Therefore, algorithm \mathbb{A} has a probability of error larger than δ under H_1 . This contradicts to the assumption that algorithm \mathbb{A} is a δ -correct algorithm. Hence, we must have $\mathbb{E}_0[T_e] > t_e^* = \frac{1}{16\Delta_e^2} \log(1/\delta)$. □

References