Pure Exploration of Combinatorial Bandits

Shouyuan Chen

March 26, 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). *$\mathcal{M}_{\mathsf{TOP}m}(n) = \{M \subseteq [n] \mid |M| = m\}$. This corresponds to finding the top m arms from $[n]$.*

Example 2 (Explore-m-bandits). *Suppose $n = mk$. Then $\mathcal{M}_{\mathsf{BANDIT}m}(n)$ 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 $\mathcal{M}_{\mathsf{MATCH}}(n, G)$ 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 $C_+ = M_1 \backslash M_2$ and $C_- = M_2 \backslash M_1$.*

Definition 3. *Denote $\mathsf{diff}[n]$ 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 $C \backslash D \triangleq (C_+ \backslash D_+, C_- \backslash D_-)$.*

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 $C \in \mathsf{diff}[n]$. If $C \prec M$, we define operator $\oplus$ such that $M \oplus C \triangleq M \backslash C_- \cup C_+$. On the other hand if $\neg C \prec M$, we define operator $\ominus$ such that $M \ominus C \triangleq M \oplus (\neg C) = M \backslash C_+ \cup C_-$.*

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 $A \oplus B = ((A_+ \cup B_+)\backslash(A_- \cup B_-), (A_- \cup B_-)\backslash(A_+ \cup B_+))$.*

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 $C_+ = (A_+ \cup B_+)\backslash(A_- \cup B_-)$ and $C_- = (A_- \cup B_-)\backslash(A_+ \cup B_+)$.

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

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

□

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 $B_+ \cap (M\backslash A_- \cup A_+) = \emptyset$. Therefore, we have

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

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

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

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

Hence we proved that $B \prec A$.

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

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

Similarly, we have

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

□

1.2.2 Diff-set class

Definition 10 (Decomposition of diff-set)**.** *Given $\mathcal{B} \subseteq \mathsf{diff}[n]$ and $D \in \mathsf{diff}[n]$, a decomposition of D on $\mathcal{B}$ is a set $\{b_1, \ldots, b_k\} \subseteq \mathcal{B}$ 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. $D = b_1 \oplus b_2 \oplus \ldots b_k$.

Lemma 3. *Given* $\mathcal{B} \subseteq \mathsf{diff}[n]$ *and* $D \in \mathsf{diff}[n]$. *Let* $\{b_1, \ldots, b_k\} \subseteq \mathcal{B}$ *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* $D_+ = b_1^+ \cup \ldots \cup b_k^+$ *and* $D_- = b_1^- \cup \ldots \cup b_k^-$.

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 $D_i = b_1 \oplus \ldots \oplus b_i$ and write $D_i = (D_i^+, D_i^-)$. We show that $D_i^+ = \bigcup_{j=1}^{i} b_i^+$ and $D_{i-} = \bigcup_{j=1}^{i} b_i^-$ 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 $D_{i+1}^+ = (D_i^+ \cup b_{i+1}^+) \backslash (D_i^- \cup b_{i+1}^-)$. 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 $b_i^- \subseteq D_- \subseteq M$. □

Definition 11 (diff-set class). *Given* $\mathcal{M} \subseteq 2^{[n]}$. $\mathcal{B} \subseteq \mathsf{diff}[n]$ *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*

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

Example 4 (diff-set class for Explore-m). *One diff-set class* $\mathcal{B}$ *for* $\mathcal{M}_{\mathsf{TOP}m}(n)$ *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 $\mathrm{rank}(\mathcal{B}) = 2$.

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

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

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

Example 6 (diff-set class for Perfect Matching). *One diff-set class $\mathcal{B}$ for $\mathcal{M}_{\mathsf{MATCH}}(n, G)$ 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 $\mathrm{rank}(\mathcal{B}) \le n$.

1.3 Weights and confidence bounds

Definition 13 (Weight functions). *Define function $w : [n] \to \mathbb{R}^+$ 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 $b = (b_+, b_-) \in \mathsf{diff}[n]$, we denote $w(b) = \sum_{e \in b_+} w(e) - \sum_{e \in b_-} w(e)$.*

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

$$w(c \cup d) = w(c) + w(d) - w(c \cap d). \tag{1}$$

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

$$\begin{aligned} w(c \cup d) &= w(c_+ \cup d_+) - w(c_- \cup d_-) &(2)\\ &= w(c_+) + w(d_+) - w(c_+ \cap d_+) - w(c_-) - w(d_-) + w(c_- \cap d_-) &(3)\\ &= w(c) + w(d) - (w(c_+ \cap d_+) - w(c_- \cap d_-)) &(4)\\ &= w(c) + w(d) - w(c \cap d). &(5) \end{aligned}$$

□

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 $n_t : [n] \to \mathbb{N}$, 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 $\mathrm{rad}_t : [n] \to \mathbb{R}^+$ satisfying, for all $e \in [n]$,*

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

where $c_{\mathrm{rad}} > 0$ 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 $M \subseteq [n]$, $\mathrm{rad}_t(M) \triangleq \sum_{e \in M} \mathrm{rad}_t(e)$.*
2. *For all $b = (b_+, b_-) \in \mathsf{diff}[n]$, $\mathrm{rad}_t(b) \triangleq \mathrm{rad}_t(b_+) + \mathrm{rad}_t(b_-)$.*

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

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

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

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

Lemma 5. *Define random event*

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

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

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

Proof. Hoeffding inequality and union bound. □

Corollary 1.

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

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

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

1.4 Properties of rad_t

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

$$\mathrm{rad}_t(c \backslash d) = \mathrm{rad}_t(c) - \mathrm{rad}_t(c \cap d). \tag{6}$$

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

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

□

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

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

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

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

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

$$\begin{aligned}\mathrm{rad}_t((C_+ \cup D_+) \backslash (C_- \cup D_-)) &= \mathrm{rad}_t(C_+ \cup D_+) - \mathrm{rad}_t((C_+ \cup D_+) \cap (C_- \cup D_-)) &(7)\\ &= \mathrm{rad}_t(C_+) + \mathrm{rad}_t(D_+) - \mathrm{rad}_t((C_+ \cup D_+) \cap (C_- \cup D_-)), &(8)\end{aligned}$$

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

Similarly, we have

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

Combine both equalities, we have

$$\begin{aligned}
\mathrm{rad}_t(C \oplus D) &= \mathrm{rad}_t((C_+ \cup D_+)\backslash(C_- \cup D_)) + \mathrm{rad}_t((C_- \cup D_-)\backslash(C_- \cap D_+)) && (9)\\
&= \mathrm{rad}_t(C_+) + \mathrm{rad}_t(D_+) + \mathrm{rad}_t(C_-) + \mathrm{rad}_t(D_-) - 2\,\mathrm{rad}_t((C_+ \cup D_+) \cap (C_- \cup D_-)) && (10)\\
&= \mathrm{rad}_t(C) + \mathrm{rad}_t(D) - 2\,\mathrm{rad}_t((C_+ \cup D_+) \cap (C_- \cup D_-)). && (11)
\end{aligned}$$

□

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 $D = \arg\max_{C \in \mathsf{diff}[n], C \prec M_t} w_t^+(C)$.
6. If $w_t^+(D) \leq 0$. Then stop and return M_t.
7. Otherwise, find $p_t = \arg\min_{e \in D} \mathrm{rad}_t(e)$.
8. Play p_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) $\forall e \in M_t$, $\tilde{w}_t(e) = \bar{w}_t(e) - \mathrm{rad}_t(e)$.
   (b) $\forall e \notin M_t$, $\tilde{w}_t(e) = \bar{w}_t(e) + \mathrm{rad}_t(e)$.
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 $\mathcal{B}_{\mathsf{opt}}$ as a subset of $\mathcal{B}$, and for all $b \in \mathcal{B}$, $b \in \mathcal{B}_{\mathsf{opt}}$ 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*

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

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 $\mathrm{rad}_t(e) < \frac{1}{3}\Delta_e$.*

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

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

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. ☐

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 $c, c_1, \ldots, c_k$ 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

$$D_+ = c_+ \cup c_1^+ \cup \ldots \cup c_k^+ \quad \text{and} \quad D_- = c_- \cup c_1^- \cup \ldots c_k^-. \tag{12}$$

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

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

Hence, $\bar{w}_t(c) \ge w(c) - \mathrm{rad}_t(c) \ge \frac{2}{3}w(c) > 0$. 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 $c_t \notin \mathcal{B}_{\mathsf{opt}}$. 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 $b, b_1, \ldots, b_l$, 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 $b\in\mathcal{B}_{\mathsf{opt}}$.

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

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

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

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

First, we show that $\mathrm{rad}_t(c) \le \frac{1}{3}w(b)$. Since $e\in E_t^3$, $e\in b$ and $b\in\mathcal{B}_{\mathsf{opt}}$, we have $\mathrm{rad}_t(e)\le\frac{1}{3}\Delta_e \le \frac{1}{3K}w(b)$.

In addition, $\forall g \in c, g \neq e$, $\mathrm{rad}_t(g) \leq \mathrm{rad}_t(e) \leq \frac{1}{3K} w(b)$. Hence,

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

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

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

Then, we have

$$\begin{aligned}
\mathrm{rad}_t(D') - \mathrm{rad}_t(D) &= \mathrm{rad}_t(\tilde{D} \oplus b) - \mathrm{rad}_t(D) && (16)\\
&= \mathrm{rad}_t(\tilde{D}) + \mathrm{rad}_t(b) - 2\,\mathrm{rad}_t(\tilde{D}_+ \cap b_-) - 2\,\mathrm{rad}_t(\tilde{D}_- \cap b_+) - \mathrm{rad}_t(D) && (17)\\
&= \mathrm{rad}_t(D \ominus c) + \mathrm{rad}_t(b) - 2\,\mathrm{rad}_t(\tilde{D}_+ \cap b_-) - 2\,\mathrm{rad}_t(\tilde{D}_- \cap b_+) - \mathrm{rad}_t(D) && (18)\\
&= \mathrm{rad}_t(D) + \mathrm{rad}_t(c) + \mathrm{rad}_t(b) - 2\,\mathrm{rad}_t(D_+ \cap c_+) - 2\,\mathrm{rad}_t(D_- \cap c_-) \\
&\quad - 2\,\mathrm{rad}_t(\tilde{D}_+ \cap b_-) - 2\,\mathrm{rad}_t(\tilde{D}_- \cap b_+) - \mathrm{rad}_t(D) && (19)\\
&= \mathrm{rad}_t(D) + \mathrm{rad}_t(c) + \mathrm{rad}_t(b) - 2\,\mathrm{rad}_t(c_+) - 2\,\mathrm{rad}_t(c_-) \\
&\quad - 2\,\mathrm{rad}_t(\tilde{D}_+ \cap b_-) - 2\,\mathrm{rad}_t(\tilde{D}_- \cap b_+) - \mathrm{rad}_t(D) && (20)\\
&= \mathrm{rad}_t(b) - \mathrm{rad}_t(c) - 2\,\mathrm{rad}_t(\tilde{D}_+ \cap b_-) - 2\,\mathrm{rad}_t(\tilde{D}_- \cap b_+), && (21)
\end{aligned}$$

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

$$\begin{aligned}
\bar{w}_t(D) + \mathrm{rad}_t(D) &\geq \bar{w}_t(D') + \mathrm{rad}_t(D') && (22)\\
&= \bar{w}_t(D) - \bar{w}_t(c) + \bar{w}_t(b) + \mathrm{rad}_t(D'). && (23)
\end{aligned}$$

By regrouping the above inequality, we have

$$\begin{aligned}
\bar{w}_t(c) &\geq \bar{w}_t(b) + \mathrm{rad}_t(D') - \mathrm{rad}_t(D) && (24)\\
&= \bar{w}_t(b) + \mathrm{rad}_t(b) - \mathrm{rad}_t(c) - 2\,\mathrm{rad}_t(\tilde{D}_+ \cap b_-) - 2\,\mathrm{rad}_t(\tilde{D}_- \cap b_+) && (25)\\
&\geq w(b) - \mathrm{rad}_t(c) - 2\,\mathrm{rad}_t(\tilde{D}_+ \cap b_-) - 2\,\mathrm{rad}_t(\tilde{D}_- \cap b_+) && (26)\\
&> w(b) - \frac{1}{3} w(b) - \frac{2}{3} w(b) && (27)\\
&= 0, && (28)
\end{aligned}$$

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 $b, b_1, \ldots, b_l$. Assume wlog that $e \in b$. We write that

$b = (b_+, b_-)$. Note that $b \in \mathcal{B}_{\mathsf{opt}}$ 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 $|D \backslash D'| \le |b|$. Let $C = D \backslash D'$ and write $C = (C_+, C_-)$. We can bound $|C_+|$ as follows.

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

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

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

Thus $|C_-| \le |b_-|$ and we proved that $|D \backslash D'| \le |b|$.

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

We also note that

$$\begin{aligned}
w(D' \backslash D) - w(D \backslash D') &= w(D' \backslash D) + w(D' \cap D) - w(D \cap D') - w(D \backslash D') && (29) \\
&= w(D') - w(D) && (30) \\
&= w(b), && (31)
\end{aligned}$$

where we have repeatedly applied Lemma 4.

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

$$\begin{aligned}
w_t^+(D') - w_t^+(D) &= \bar{w}_t(D') - \bar{w}_t(D) + \mathrm{rad}_t(D') - \mathrm{rad}_t(D) && (32) \\
&= \bar{w}_t(D' \backslash D) - \bar{w}_t(D \backslash D') + \mathrm{rad}_t(D' \backslash D) - \mathrm{rad}_t(D \backslash D') && (33) \\
&\ge w(D' \backslash D) - w(D \backslash D') - 2\,\mathrm{rad}_t(D \backslash D') && (34) \\
&= w(b) - 2\,\mathrm{rad}_t(D \backslash D') && (35) \\
&> w(b) - \frac{2}{3} w(b) && (36) \\
&= \frac{1}{3} w(b) > 0, && (37)
\end{aligned}$$

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

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