Appendix

1 Useful concentration inequalities

We first state the following versions of the standard Hoeffding and Chernoff bounds, which will be used in many proofs later on.

Proposition 1.1. *Let $X_i (1 \le i \le n)$ be independent random variables with values in $[0,1]$. Let $X = \frac{1}{n}\sum_{i=1}^{n} X_i$. The following statements hold:*

1. *For every $t > 0$, we have that*
$$\Pr\left[\, |X - \mathbf{E}[X]| > t \right] < 2\exp(-2t^2 n).$$

2. *Suppose $\mathbf{E}[X_i] < a$ for some real $0 \le a \le 1$. For every $0 < t < 1$, we have that*
$$\Pr\left[X > a + t\right] < \left(\left(\frac{a}{a+t}\right)^{a+t}\left(\frac{1-a}{1-a-t}\right)^{1-a-t}\right)^{n}.$$

3. *For every $0 < \epsilon < 1$, we have that*
$$\Pr\left[X < (1-\epsilon)\,\mathbf{E}[X]\right] < \exp\left(-\frac{\epsilon^2 n\,\mathbf{E}[X]}{2}\right),$$
$$\Pr\left[X > (1+\epsilon)\,\mathbf{E}[X]\right] < \exp\left(-\frac{\epsilon^2 n\,\mathbf{E}[X]}{3}\right).$$

Besides the above standard Hoeffding and Chernoff bounds, we also need the following Chernoff-type concentration inequality.

Proposition 1.2. *Let $X_i (1 \le i \le K)$ be independent random variables. Each X_i takes value a_i ($a_i \ge 0$) with probability at most $\exp(-a_i^2 t)$ for some $t \ge 0$, and 0 otherwise. Let $X = \frac{1}{K}\sum_{i=1}^{K} X_i$. For every $\epsilon > 0$, when $t \ge \frac{2}{\epsilon^2}$, we have that*
$$\Pr\left[X > \epsilon\right] < \exp\left(-\epsilon^2 t K/2\right).$$

Proof of Proposition 1.2. The proof is similar to that for the standard Chernoff bound. First, we observe that

$$\begin{aligned}
\Pr[X > \epsilon] &= \Pr\left[\sum_{i=1}^{K} X_i > \epsilon K\right] = \Pr\left[\sum_{i=1}^{K} \epsilon t X_i > \epsilon^2 t K\right] = \Pr\left[\exp\left(\sum_{i=1}^{K} \epsilon t X_i\right) > \exp(\epsilon^2 t K)\right] \\
&\le \mathbf{E}\left[\frac{\exp\left(\sum_{i=1}^{K} \epsilon t X_i\right)}{\exp(\epsilon^2 t K)}\right] = \frac{\prod_{i=1}^{K} \mathbf{E}[\exp(\epsilon t X_i)]}{\exp(\epsilon^2 t K)}, \qquad (1)
\end{aligned}$$

where the first inequality follows from Markov inequality and the last equality holds due to independence. Now, we claim that

$$\mathbf{E}[\exp(\epsilon t X_i)] \le \exp\left(\epsilon^2 t K/2\right).$$

By the definition of X_i, combined with the fact that $a(\epsilon - a) \le \epsilon^2/4$ for any real value a, it holds that

$$\mathbf{E}[e^{\epsilon t X_i}] \le \exp(\epsilon a t - a^2 t) + 1 = \exp(a(\epsilon - a)t) + 1 \le \exp\left(\epsilon^2 t/4\right) + 1.$$

When $\epsilon^2 t \ge 2$, we have $\exp\left(\epsilon^2 t/2\right) - \exp\left(\epsilon^2 t/4\right) > 1.06 > 1$ and hence $\mathbf{E}[e^{\epsilon t X_i}] \le \exp\left(\epsilon^2 t/2\right)$. Plugging this bound into (1), we get that

$$\Pr[X > \epsilon] \le \frac{\prod_{i=1}^{K} \exp\left(\epsilon^2 t/2\right)}{\exp(\epsilon^2 t K)} = \exp\left(-\frac{\epsilon^2 t K}{2}\right).$$

The proof is completed. □

2 Proof of the Correctness of the QE Algorithm (Lemma 4.1)

Lemma 4.1 (restated). *Assume that $K \le |S|/4$ and let V be the output of $\mathrm{QE}(S, Q)$ (Algorithm 2). For every $0 < \delta < 1$, with probability $1 - \delta$, we have that $\mathsf{val}_K(V) \ge \mathsf{val}_K(S) - \epsilon$, where $\epsilon = \sqrt{\frac{|S|}{Q}\left(10 + \frac{4\ln(2/\delta)}{K}\right)}$.*

Let $p = \theta_{\mathsf{ind}_{|S|/2}(S)}$ be the median of the means of the arms in S. Let $\tau = \min_{i \in V}(\widehat{\theta}_i)$ be the minimum empirical mean for the selected arms in V. For each arm i among the top K arms in S, we define the random variable $X_i = \mathbf{1}\{\widehat{\theta}_{\mathsf{ind}_i(S)} < p + \frac{\epsilon}{2}\}$ and $X = \frac{1}{K}\sum_{i=1}^{K}(\theta_{\mathsf{ind}_i(S)} - p)X_i$, where $\mathbf{1}\{\cdot\}$ is the indicator function. We further define two events $\mathcal{E}_1 = \{X \le \epsilon\}$ and $\mathcal{E}_2 = \{\tau < p + \frac{\epsilon}{2}\}$. Our first claim is that $\mathcal{E}_1$ and $\mathcal{E}_2$ together imply our conclusion $\mathsf{val}_K(V) \ge \mathsf{val}_K(S) - \epsilon$.

Lemma 2.1. *$\mathcal{E}_1$ and $\mathcal{E}_2$ imply that $\mathsf{val}_K(V) \ge \mathsf{val}_K(S) - \epsilon$.*

Proof. Suppose both $\mathcal{E}_1$ and $\mathcal{E}_2$ hold. We first claim that

$$\frac{1}{K}\sum_{i=1}^{K}\theta_{\mathsf{ind}_i(V)} \ge \frac{1}{K}\sum_{i=1}^{K}\left((1 - X_i)\theta_{\mathsf{ind}_i(S)} + X_i p\right). \tag{2}$$

To see this claim, consider arm $\mathsf{ind}_i(S)$ for some $i \in [K]$. If $X_i = 0$ (i.e., $\widehat{\theta}_{\mathsf{ind}_i(S)} \ge p + \frac{\epsilon}{2}$), together with $\mathcal{E}_2$, we have that

$$\widehat{\theta}_{\mathsf{ind}_i(S)} \ge p + \frac{\epsilon}{2} > \tau = \min_{i \in V}(\widehat{\theta}_i).$$

Hence, the arm should be included in the output set V. Moreover, since it is one of the best K arms in S, it is also one of the best K arms in V. Hence, for each term on the right hand side of (2) with $X_i = 0$, there is exactly one term with the same value on the left hand side.

Since there are $|S|/2 \ge K + |S|/4$ arms with means greater or equal to p, after removing $|S|/4$ of them, there are still at least K such arms. Therefore we know that the best K arms of V all have means greater than or equal to p. In other words, each term on the left hand side of (2) is greater than or equal to p. This proves (2). Now, we can see that

$$\begin{aligned}
\mathsf{val}_K(V) &= \frac{1}{K}\sum_{i=1}^{K}\theta_{\mathsf{ind}_i(V)} \ge \frac{1}{K}\sum_{i=1}^{K}\left((1 - X_i)\theta_{\mathsf{ind}_i(S)} + X_i p\right) \\
&= \frac{1}{K}\sum_{i=1}^{K}\theta_{\mathsf{ind}_i(S)} - \frac{1}{K}\sum_{i=1}^{K}(\theta_{\mathsf{ind}_i(S)} - p)X_i \ge \mathsf{val}_K(S) - \epsilon,
\end{aligned}$$

where the last inequality is due to $\mathcal{E}_1$. $\square$

In light of Lemma 2.1, it suffices to show that the probability that both $\mathcal{E}_1$ and $\mathcal{E}_2$ happen is at least $1-\delta$. First, we bound $\Pr[\mathcal{E}_1]$ in the following lemma.

Lemma 2.2. $\Pr[\mathcal{E}_1] \geq 1 - \dfrac{\delta}{2}$.

Proof of Lemma 2.2. Recall that $Q_0 = \frac{Q}{|S|}$ is the number of samples taken from each arm in S. By the definition of ϵ in Lemma 2.2, we trivially have that $\epsilon \geq \max\left\{\sqrt{\frac{10}{Q_0}}, \sqrt{\frac{4\ln(2/\delta)}{Q_0 K}}\right\}$. For each $i \in [K]$, let $\eta_i = \max\{\theta_{\mathsf{ind}_i(S)} - p - \frac{\epsilon}{2}, 0\}$ and let $Y_i = \eta_i X_i$. By Proposition 1.1(1), we have that

$$\begin{aligned}\Pr[Y_i = \eta_i] = \Pr[X_i = 1] = \Pr\left[\widehat{\theta}_{\mathsf{ind}_i(S)} < p + \frac{\epsilon}{2}\right] = \Pr\left[\widehat{\theta}_{\mathsf{ind}_i(S)} < \theta_{\mathsf{ind}_i(S)} - (\theta_{\mathsf{ind}_i(S)} - p - \frac{\epsilon}{2})\right] \\ \leq \exp\left(-2\left(\theta_{\mathsf{ind}_i(S)} - p - \frac{\epsilon}{2}\right)^2 \cdot Q_0\right) \leq \exp\left(-\eta_i^2 \cdot 2Q_0\right).\end{aligned}$$

Applying Proposition 1.2 on Y_i's, we can get that

$$\Pr\left[\frac{1}{K}\sum_{i=1}^{K} Y_i > \frac{\epsilon}{2}\right] \leq \exp\left(-\frac{\epsilon^2 Q_0 K}{4}\right) \leq \frac{\delta}{2},$$

where the last inequality holds because $\epsilon \geq \sqrt{\frac{4\ln(2/\delta)}{Q_0 K}}$. Observe that $Y_i \geq (\theta_{\mathsf{ind}_S(i)} - p)X_i - \frac{\epsilon}{2}$ for all $i \in [K]$. Therefore, with probability at least $1 - \frac{\delta}{2}$, we have that

$$X = \frac{1}{K}\sum_{i=1}^{K}(\theta_{\mathsf{ind}_i(S)} - p)X_i \leq \frac{1}{K}\sum_{i=1}^{K} Y_i + \frac{\epsilon}{2} \leq \epsilon.$$

This completes the proof of Lemma 2.2. $\square$

Next, we bound the probability that $\mathcal{E}_2$ happens in the following lemma.

Lemma 2.3. $\Pr[\mathcal{E}_2] \geq 1 - \dfrac{\delta}{2}$.

Proof. First, we can see that $\mathcal{E}_2$ holds if and only if there are no more than $3|S|/4$ arms with empirical mean larger than $p + \frac{\epsilon}{2}$. Define the indictor random variable $Z_i = \mathbf{1}\{\widehat{\theta}_{\mathsf{ind}_i(S)} \geq p + \frac{\epsilon}{2}\}$. Hence, it suffice to show that $\Pr\left[\sum_{i=|S|/2}^{|S|} Z_i < \frac{|S|}{4}\right] \geq 1 - \frac{\delta}{2}$.

Let us only consider the arms with indices $i \in [|S|/2, |S|]$ (i.e., $\theta_i \leq p$). By Proposition 1.1(1), we have

$$\Pr\left[\widehat{\theta}_{\mathsf{ind}_i(S)} \geq p + \frac{\epsilon}{2}\right] \leq \Pr\left[\widehat{\theta}_{\mathsf{ind}_i(S)} \geq \theta_{\mathsf{ind}_i(S)} + \frac{\epsilon}{2}\right] \leq \exp\left(-\frac{\epsilon^2}{2} \cdot Q_0\right). \tag{3}$$

From (3), we can see that $\mathbf{E}[Z_i] < \exp\left(-\frac{\epsilon^2}{2} \cdot Q_0\right)$. Let $\mu = \max_{i \in [|S|/2, |S|]} \mathbf{E}[Z_i]$ and we have $\mu < \exp\left(-\frac{\epsilon^2}{2} \cdot Q_0\right)$. Then, by Proposition 1.1(2), we have

$$\begin{aligned}\Pr\left[\sum_{i=|S|/2}^{|S|} Z_i > \frac{|S|}{4}\right] &\leq \left(\left(\frac{\mu}{1/2}\right)^{1/2}\left(\frac{1-\mu}{1/2}\right)^{1/2}\right)^{|S|/2} \leq \left(\sqrt{2\mu} \cdot \sqrt{2}\right)^{|S|/2} \\ &\leq \exp\left(\frac{|S|}{2}\left(\ln(2) - \frac{\epsilon^2}{2} \cdot Q_0\right)\right) \leq \exp\left(-\frac{|S|}{2} \cdot \frac{\epsilon^2}{4} \cdot Q_0\right) \leq \exp\left(-\frac{\epsilon^2 K}{2} \cdot Q_0\right) \leq \frac{\delta}{2}.\end{aligned}$$

where the third to last inequality follows because of $\epsilon > \sqrt{\frac{10}{Q_0}}$, the second to last inequality uses the assumption that $|S|/4 \geq K$ and the last inequality holds because we assume that $\epsilon \geq \sqrt{\frac{4\ln(2/\delta)}{Q_0 K}}$. $\square$

Proof of Lemma 4.1. By Lemma 2.2, Lemma 2.3, and a union bound, we have $\Pr[\mathcal{E}_1 \text{ and } \mathcal{E}_2] \geq 1-\delta$. By Lemma 2.1, we have $\Pr[\mathsf{val}_K(T) \geq \mathsf{val}_K(S) - \epsilon] \geq \Pr[\mathcal{E}_1 \text{ and } \mathcal{E}_2]$ and then the lemma follows. $\square$

3 Proof of the Correctness of the AR Algorithm (Lemma 4.2 in the Main Text)

Lemma 4.2 (restated). *Let (S', T') be the output of the algorithm $\mathrm{AR}(S, T, Q, K)$ (Algorithm 3). For every $0 < \delta < 1$, with probability $1-\delta$, we have that*

$$\mathsf{tot}_{K-|T'|}(S') + \mathsf{tot}_{|T'|}(T') \geq \mathsf{tot}_{K-|T|}(S) + \mathsf{tot}_{|T|}(T) - \epsilon K,$$

where $\epsilon = \sqrt{\frac{|S|}{Q}\left(4 + \frac{\ln(2/\delta)}{K}\right)}$.

Proof. Recall that $Q_0 = \frac{Q}{|S|}$ is the number of samples taken from each arm in S. Also recall that $K' = K - |T|$.

We need to define a few notations. Let $U_1 = T' \setminus T$ denote the set of arms we added to T' in this round. Let $U_2 = S \setminus (S' \cup U_1)$ be the set of arms we discarded in this round. Let U_1^* be the set of $|U_1|$ arms in S with largest θ_i's; let U_2^* be the set of $|U_2|$ arms in S with smallest θ_i's. Ideally, if $U_1 = U_1^*$ and $U_2 = U_2^*$, we do not lose anything in this round (i.e., $\mathsf{tot}_{K-|T'|}(S') + \mathsf{tot}_{|T'|}(T') = \mathsf{tot}_{K-|T|}(S) + \mathsf{tot}_{|T|}(T)$). When $U_1 \neq U_1^*$ and/or $U_2 \neq U_2^*$, we can bound the difference between $\mathsf{tot}_{K-|T'|}(S') + \mathsf{tot}_{|T'|}(T')$ and $\mathsf{tot}_{K-|T|}(S) + \mathsf{tot}_{|T|}(T)$ by the sum of the difference between U_1^* and U_1, and the difference between U_2^* and U_2. More concretely, we claim that

$$\left(\mathsf{tot}_{K-|T'|}(S') + \mathsf{tot}_{|T'|}(T')\right) - \left(\mathsf{tot}_{K-|T|}(S) + \mathsf{tot}_{|T|}(T)\right) \geq \left(\sum_{i \in U_2^*} \theta_i - \sum_{i \in U_2} \theta_i\right) - \left(\sum_{i \in U_1} \theta_i - \sum_{i \in U_1^*} \theta_i\right). \quad (4)$$

The proof of (4) is not difficult, but somewhat tedious, and we present it at the end of this section. From now on, we assume (4) is true.

For every $t \leq K$, for every set $U \subseteq S$ of t arms (i.e. $|U| = t$), by Proposition 1.1, we have

$$\Pr\left[\left|\sum_{i \in U} \widehat{\theta}_i - \sum_{i \in U} \theta_i\right| > \frac{\epsilon K}{4}\right] \leq 2\exp\left(-\frac{\epsilon^2}{8} \cdot Q_0 \frac{K^2}{t}\right) \leq 2\exp\left(-\frac{\epsilon^2}{8} \cdot Q_0 K\right).$$

By a union bound over all subset of size at most K, we have that

$$\begin{aligned} \Pr\left[\forall U \subseteq S, |U| \leq K : \left|\sum_{i \in U} \widehat{\theta}_i - \sum_{i \in U} \theta_i\right| \leq \frac{\epsilon K}{4}\right] \quad &\geq \quad 1 - 2 \cdot 2^{|S|} \exp\left(-\frac{\epsilon^2}{8} \cdot Q_0 K\right) \\ &\geq \quad 1 - 2\exp\left(|S| - \frac{\epsilon^2}{8} \cdot Q_0 K\right) \geq 1 - \delta, \end{aligned}$$

where we used the facts that $|S| < 4K$ and $\epsilon \geq \sqrt{\frac{1}{Q_0}\left(4 + \frac{\ln(2/\delta)}{K}\right)}$.

Thus, with probability at least $1-\delta$, all of the following four inequalities hold:

$$\left|\sum_{i\in U_1}\widehat{\theta}_i - \sum_{i\in U_1}\theta_i\right| \le \frac{\epsilon K}{4}, \qquad \left|\sum_{i\in U_1^*}\widehat{\theta}_i - \sum_{i\in U_1^*}\theta_i\right| \le \frac{\epsilon K}{4},$$
$$\left|\sum_{i\in U_2}\widehat{\theta}_i - \sum_{i\in U_2}\theta_i\right| \le \frac{\epsilon K}{4}, \qquad \left|\sum_{i\in U_2^*}\widehat{\theta}_i - \sum_{i\in U_2^*}\theta_i\right| \le \frac{\epsilon K}{4}.$$

Therefore we have

$$\sum_{i\in U_1}\theta_i \ge \sum_{i\in U_1}\widehat{\theta}_i - \frac{\epsilon K}{4} \ge \sum_{i\in U_1^*}\widehat{\theta}_i - \frac{\epsilon K}{4} \ge \sum_{i\in U_1^*}\theta_i - \frac{\epsilon K}{2}, \quad \text{and} \tag{5}$$

$$\sum_{i\in U_2}\theta_i \le \sum_{i\in U_2}\widehat{\theta}_i - \frac{\epsilon K}{4} \le \sum_{i\in U_2^*}\widehat{\theta}_i - \frac{\epsilon K}{4} \le \sum_{i\in U_2^*}\theta_i - \frac{\epsilon K}{2}. \tag{6}$$

Combining (4), (5) and (6), we get (7) $\ge -\epsilon K$, which concludes the proof. □

Proof of (4). For ease of notation, for any subset S of arms, we let $\theta(S) = \sum_{i\in S}\theta_i$. One can easily see that

$$\begin{aligned}
&\big(\mathsf{tot}_{K-|T'|}(S') + \mathsf{tot}_{|T'|}(T')\big) - \big(\mathsf{tot}_{K-|T|}(S) + \mathsf{tot}_{|T|}(T)\big)\\
&\quad = \mathsf{tot}_{K-|T'|}(S') - \mathsf{tot}_{K-|T|}(S) + \theta(U_1)\\
&\quad = \mathsf{tot}_{K-|T'|}(S') - \mathsf{tot}_{K-|T'|}(S\setminus U_1^*) + (\theta(U_1) - \theta(U_1^*))\\
&\quad \ge \mathsf{tot}_{K-|T'|}(S') - \mathsf{tot}_{K-|T'|}(S\setminus U_1) + (\theta(U_1) - \theta(U_1^*))\,.
\end{aligned} \tag{7}$$

Let $\widetilde{U}_2$ be the $|U_2|$ arms with the smallest means in $S\setminus U_1$. By definition we have 1) $|\widetilde{U}_2| = |U_2^*|$; 2) $\widetilde{U}_2 \cap U_1 = U_2\cap U_1 = \emptyset$; 3) $\theta(\widetilde{U}_2) \ge \theta(U_2^*)$.

Since $|U_1| + |U_2| + (K - |T'|) \le |S|$, the $(K-|T'|)$ arms with largest means in $S\setminus U_1$ do not intersect with the $|U_2|$ arms with smallest means in $S\setminus U_1$ (namely $\widetilde{U}_2$). Therefore, we have that

$$\mathsf{tot}_{K-|T'|}(S\setminus U_1) = \mathsf{tot}_{K-|T'|}((S\setminus U_1)\setminus \widetilde{U}_2). \tag{8}$$

On the other hand, for every set W of arms, define $\mathsf{tot}_t^{\min}(W)$ to be the sum of the t smallest means among the arms in W. Let $t = |S| - |U_1| - |U_2| - (K-|T'|)$. Since $\widetilde{U}_2$ consists of the arms with the smallest means in $S\setminus U_1$, we have

$$\mathsf{tot}_t^{\min}((S\setminus U_1)\setminus\widetilde{U}_2) \ge \mathsf{tot}_t^{\min}((S\setminus U_1)\setminus U_2).$$

Together with the facts that

$$\mathsf{tot}_t^{\min}((S\setminus U_1)\setminus\widetilde{U}_2) = \theta((S\setminus U_1)\setminus\widetilde{U}_2) - \mathsf{tot}_{K-|T'|}((S\setminus U_1)\setminus\widetilde{U}_2), \text{ and}$$
$$\mathsf{tot}_t^{\min}((S\setminus U_1)\setminus U_2) = \theta(S\setminus U_1)\setminus U_2) - \mathsf{tot}_{K-|T'|}((S\setminus U_1)\setminus U_2),$$

we can see that

$$\theta(S\setminus U_1)\setminus\widetilde{U}_2) - \mathsf{tot}_{K-|T'|}((S\setminus U_1)\setminus\widetilde{U}_2) \ge \theta(S\setminus U_1)\setminus U_2) - \mathsf{tot}_{K-|T'|}((S\setminus U_1)\setminus U_2).$$

Equivalently, we have that

$$\mathsf{tot}_{K-|T'|}((S\setminus U_1)\setminus U_2) - \mathsf{tot}_{K-|T'|}((S\setminus U_1)\setminus\widetilde{U}_2) \ge \theta(\widetilde{U}_2) - \theta(U_2). \tag{9}$$

By combining (7), (8) and (9), and the observations that $S' = (S\setminus U_1)\setminus U_2$ and $\theta(\widetilde{U}_2) \ge \theta(U_2^*)$, we have proved (4). □

4 Proof of the Main Complexity Result (Theorem 4.3 in the Main Text)

Theorem 4.3 (restated). *For every $0 < \delta < 1$ and sample budget $Q > 0$, with probability at least $1-\delta$, the output of OptMAI algorithm T is an ϵ-optimal solution (i.e., $\mathsf{val}_K(T) \geq \mathsf{val}_K([n]) - \epsilon$) with $\epsilon = O\left(\sqrt{\frac{n}{Q}\left(1+\frac{\ln(1/\delta)}{K}\right)}\right)$. Moreover, each arm is sampled by at most $O(Q/n^{0.3})$ times.*

Proof. Recall r is the counter of the number of iterations in Algorithm 1. Let r_0 be the first r such that we have $|S_r| < 4K$. Let R be the final value of r. For any positive integer r, let

$$\delta_r = e^{-.1r}(1-e^{-.1})\delta \quad \text{and} \quad \epsilon_r = O\left(\sqrt{\frac{(\frac{3}{4})^r n}{(1-\beta)\beta^r Q}\left(1+\frac{\ln(1/\delta_r)}{K}\right)}\right).$$

For $r < r_0$, by Lemma 4.1, with probability $1-\delta_r$, we have that $\mathsf{val}_K(S_{r+1}) \geq \mathsf{val}_K(S_r) - \epsilon_r$. By union bound, with probability $1-\sum_{r=0}^{r_0-1}\delta_r$, we have that

$$\mathsf{val}_K(S_{r_0}) \geq \mathsf{val}_K([n]) - \sum_{r=0}^{r_0-1}\epsilon_r. \tag{10}$$

For $r : r_0 \leq r < R$, by Lemma 4.2, with probability $1-\delta_r$, we have that

$$\left(\mathsf{tot}_{K-|T_{r+1}|}(S_{r+1}) + \mathsf{tot}_{|T_{r+1}|}(T_{r+1})\right) - \left(\mathsf{tot}_{K-|T_r|}(S_r) + \mathsf{tot}_{|T_r|}(T_r)\right) \geq K\cdot\epsilon_r$$

Since T_R has exactly K elements and $T_{r_0} = \emptyset$, by union bound, with probability $1-\sum_{r=r_0}^{R-1}\delta_r$, we can see that

$$\mathsf{val}_K(T_R) \geq \mathsf{val}_K(S_{r_0}) - \sum_{r=r_0}^{R-1}\epsilon_r. \tag{11}$$

Now, by a union bound over both (10) and (11), we have that, with probability $1-\sum_{r=0}^{R-1}\delta_r \geq 1-\delta$,

$$\begin{aligned}
\mathsf{val}_K([n]) - \mathsf{val}_K(T_R) &\leq \sum_{r=0}^{R-1}\epsilon_r = \sum_{r=0}^{R-1} O\left(\sqrt{\left(\frac{3/4}{\beta(1-\beta)}\right)^r\left(\frac{n}{Q}\right)\left(1+\frac{\ln(1/\delta_r)}{K}\right)}\right)\\
&\leq \sum_{r=0}^{R-1} O\left(\sqrt{\left(\frac{3/4}{\beta(1-\beta)}\right)^r\left(\frac{n}{Q}\right)\left(1+\frac{\ln(1/\delta)+0.1r+\ln(1-e^{-.1})}{K}\right)}\right)\\
&= O\left(\sqrt{\frac{n}{Q}\left(1+\frac{\ln(1/\delta)}{K}\right)}\right).
\end{aligned}$$

Finally, let us upper bound the number of samples made to an arbitrary arm. At the r-th round, every arm is sampled by at most $\beta^r(1-\beta)Q/|S_r|$ times. Since $|S_r| \leq (3/4)^r n$, we know that every arm is sampled by at most $(4\beta/3)^r(1-\beta)Q/n$ times at the r-th round. Recall that R is the total number of rounds performed by the algorithm and we have $R \leq \log_{4/3} n$. We upper bound the number of samples made to an arm by

$$\begin{aligned}
\sum_{r=0}^{R}\frac{(4\beta/3)^r(1-\beta)Q}{n} &\leq \sum_{r=0}^{\log_{4/3} n}\frac{(4\beta/3)^r(1-\beta)Q}{n}\\
&= \sum_{r=0}^{\log_{4/3} n}\frac{e^{.2r}Q}{n} = O(e^{.2\ln n/(\ln(4/3))}Q/n) = O(n^{.7}Q/n) = O(Q/n^{.3}).
\end{aligned}$$

This completes the proof of the theorem. ∎

5 Proof of the Complexity Result for $K \geq n/2$ (Theorem 4.5 in the Main Text)

Theorem 4.5 (restated). *For any $0 < \delta < 1$ and $K \geq n/2$, with probability at least $1-\delta$, there is an algorithm that can find an ϵ-optimal solution T and the number of samples used is at most*

$$O\left(\left(\frac{n-K}{K} \cdot \frac{n}{\epsilon^2}\right)\left(\frac{n-K}{K} + \frac{\ln(1/\delta)}{K}\right)\right). \tag{12}$$

Proof. Instead of directly finding the best K arms, we attempt to find the worst $n-K$ arms for $n/2 \leq K < n$. In fact, for any $\epsilon', \delta > 0$, we can find a set T' of $n-K$ arms such that

$$\sum_{i \in T'} \theta_i - \sum_{i=K+1}^{n} \theta_i \leq (n-K)\epsilon', \tag{13}$$

(we call such a set T' an ϵ'-*worst solution*) with probability $1-\delta$, using at most $O\left(\frac{n}{\epsilon'^2}\left(1 + \frac{\ln(1/\delta)}{n-K}\right)\right)$ samples. This can be done by running OptMAI in the following way: whenever we obtain a sample of value x, we use $1-x$ as the sample value and identify the top $n-K$ arms. Using $O\left(\frac{n}{\epsilon'^2}\left(1 + \frac{\ln(1/\delta)}{n-K}\right)\right)$ samples, with probability $1-\delta$, we have an ϵ'-optimal solution T': $\sum_{i=K+1}^{n}(1-\theta_i) - \sum_{i \in T'}(1-\theta_i) \leq (n-K)\epsilon'$. This is equivalent to an ϵ'-worst solution in (13). By setting $\epsilon' = \frac{K}{n-K} \cdot \epsilon$ and $T = [n] \setminus T'$, we can see that (13) implies that $\sum_{i=1}^{K} \theta_i - \sum_{i \in T} \theta_i \leq K\epsilon$. with the sample complexity in (12).

When $K = n$, the sample complexity will be zero and (12) is still correct. Therefore, we obtain the result in Theorem 4.5. □

6 An Alternative to the AR Procedure (Remark Remark 4.1 in the Main Text)

We can replace the AR procedure by the following uniform sampling procedure $B(S_r, K, \epsilon', \delta')$, when the number of remaining arms $|S_r|$ is at most $4K$. Using this alternative procedure, we can achieve the same asymptotic sampling complexity, and its analysis is slightly simpler. However, its performance in practice is worse than the AR procedure. Note that the condition $|S_r| \leq 4K$ is crucial for the uniform sampling procedure to achieve the desired sample complexity (otherwise, we need to pay an extra $\log(n)$ factor. See Section 7 for more information).

More specifically, the algorithm takes as input the remaining subset of arms $S_r \subseteq [n]$, an integer K, and two real numbers $\epsilon', \delta' > 0$ as input, and outputs a set $T \subseteq S_r$ such that $|T| = K$. We set $\epsilon' = \epsilon/2, \delta' = \delta/2$. Note that we only run $B(S_r, K, \epsilon', \delta')$ once and its output T is our final output of the entire algorithm. The algorithm proceeds as follows.

- Sample each arm $i \in S_r$ for

  $$Q_0 = Q_B(K, \epsilon', \delta') = \frac{2(K \ln(e|S_r|/K) + \ln(2/\delta'))}{\epsilon'^2 K} = O\left(\frac{1}{\epsilon'^2}\left(1 + \frac{\ln(1/\delta')}{K}\right)\right).$$

  times and let $\widehat{\theta_i}$ be the empirical mean of arm i.
- Output the set $T \subseteq S_r$ which is the set of K arms with the largest empirical means.

It is easy to see the number of samples is bounded by $|S_r|Q_0 \leq O\left(\frac{K}{\epsilon'^2}\left(1 + \frac{\ln(1/\delta')}{K}\right)\right)$. The above algorithm can achieve the following performance guarantee.

Lemma 6.1. *Let T be the output of the algorithm $B(S_r, K, \epsilon', \delta')$. With probability $1-\delta'$, we have that* $\mathsf{val}_K(T) \geq \mathsf{val}_K(S_r) - \epsilon'$.

Proof. For every set $U \subseteq S_r$ of K arms (i.e. $|U| = K$), by Proposition 1.1(1), we have

$$\Pr\left[\left|\frac{1}{|U|}\sum_{i\in U}\widehat{\theta}_i - \frac{1}{|U|}\sum_{i\in U}\theta_i\right| > \frac{\epsilon'}{2}\right] \leq 2\exp\left(-\frac{\epsilon'^2}{2}\cdot Q_0 K\right).$$

By union bound over all subsets of size K, we have that

$$\Pr\left[\forall U \subseteq S_r, |U| = K : \left|\frac{1}{|U|}\sum_{i\in U}\widehat{\theta}_i - \frac{1}{|U|}\sum_{i\in U}\theta_i\right| \leq \frac{\epsilon'}{2}\right] \geq 1 - 2\binom{|S_r|}{k}\exp\left(-\frac{\epsilon'^2}{2}\cdot Q_0 K\right)$$

$$\geq 1 - 2\left(\frac{e|S_r|}{K}\right)^K \exp\left(-\frac{\epsilon'^2}{2}\cdot Q_0 K\right) = 1 - 2\exp\left(K\ln(e|S_r|/K) - \frac{\epsilon'^2}{2}\cdot Q_0 K\right) \geq 1-\delta'.$$

Let T^* be the set of K arms in S_r with largest θ_i's. With probability at least $1-\delta'$, we have

$$\left|\frac{1}{|T|}\sum_{i\in T}\widehat{\theta}_i - \frac{1}{|T|}\sum_{i\in T}\theta_i\right| \leq \frac{\epsilon'}{2}, \qquad \left|\frac{1}{|T^*|}\sum_{i\in T^*}\widehat{\theta}_i - \frac{1}{|T^*|}\sum_{i\in T^*}\theta_i\right| \leq \frac{\epsilon'}{2}.$$

Therefore, we can get that

$$\mathsf{val}_K(T) = \frac{1}{|T|}\sum_{i\in T}\theta_i \geq \frac{1}{|T|}\sum_{i\in T}\widehat{\theta}_i - \frac{\epsilon'}{2} \geq \frac{1}{|T^*|}\sum_{i\in T^*}\widehat{\theta}_i - \frac{\epsilon'}{2} \geq \frac{1}{|T^*|}\sum_{i\in T^*}\theta_i - \epsilon' = \mathsf{val}_K(S_r) - \epsilon'.$$

□

If we set $Q = O\left(\frac{n}{\epsilon^2}\left(1 + \frac{\ln(1/\delta)}{K}\right)\right)$ in OptMAI, the proof of Theorem 4.3 show that, after the QE stage, the set S_r of remaining arms satisfies that $\mathsf{val}_K([n]) - \mathsf{val}_K(S_r) \leq \epsilon/2$ with probability at least $1-\delta/2$. Combined with the conclusion of Lemma 6.1 and $\epsilon' = \epsilon/2, \delta' = \delta/2$, we get that $\mathsf{val}_K([n]) - \mathsf{val}_K(T) \leq \epsilon$ with probability at least $1-\delta$. The number of samples used in both QE and the uniform sampling stages is at most $Q + |S_r|Q_0 = O\left(\frac{n}{\epsilon^2}\left(1 + \frac{\ln(1/\delta)}{K}\right)\right)$, which is the same as the sample complexity stated in Corollary 4.4.

7 Naive Uniform Sampling (Remark Remark 4.2 in the Main Text)

As we have seen in Section 6, we can use a uniform sampling procedure to replace AR. We show in this section that, simply using following naive uniform sampling as the entire algorithm is not sufficient to achieve the linear sample complexity.

Naive Uniform Sampling:

- Sample each arm $i \in S_r$ for Q_0 times and let $\widehat{\theta}_i$ be the empirical mean of arm i.
- Output the set $T \subseteq S_r$ which is the set of K arms with the largest empirical means.

In fact, when $K = 1$, the work [1] shows that $Q_0 = O(\frac{1}{\epsilon^2}\ln\left(\frac{n}{\delta}\right))$ (which is $\log(n)$ factor worse than the optimal bound) is enough to identify an ϵ-optimal arm with probability at least $1-\delta$. For general K, by following the same proof of Lemma 6.1, we can show that

$$Q_0 = O\left(\frac{1}{\epsilon^2}\left(\ln\left(\frac{n}{K}\right) + \frac{\ln(1/\delta)}{K}\right)\right) \tag{14}$$

suffices for identifying an ϵ-optimal solution with probability at least $1-\delta$. Moreover, we can also show the bound (14) is essential tight for naive uniform sampling, as in the following theorem.

Theorem 7.1. *Given any $0 < \epsilon, \delta < 0.01$, and $K \le n/2$, suppose that the naive uniform sampling algorithm with parameter Q_0 can find an ϵ-optimal solution with probability at least $1-\delta$. Then, it must hold that* $Q_0 = \max\left\{\Omega\left(\frac{1}{\epsilon^2}\ln\left(\frac{n}{K}\right)\right), \Omega\left(\frac{1}{\epsilon^2}\frac{\ln(1/\delta)}{K}\right)\right\}$.

Proof. In fact, the second lower bound $\Omega\left(\frac{1}{\epsilon^2}\frac{\ln(1/\delta)}{K}\right)$ holds for any algorithm (including the uniform sampling algorithm), which is proved in Lemma 5.4. So, we only focus on the first lower bound in this proof. Let C be a sufficiently large constant ($C > 6000$ suffices). First we consider the case where $1 \le K \le n/C$. Given the instance which consists of K Bernoulli arms with mean $1/2 + 4\epsilon$ (denoted as set A) and $n - K$ Bernoulli arms with mean $1/2$ (denoted as set B), it is easy to see that any ϵ-optimal solution must contain at least $\frac{3}{4}K$ arms from A. Let $Q_0 = \frac{1}{400\epsilon^2}\ln\left(\frac{n}{K}\right)$. Now, we show that with probability at least 0.05, there are at least $K/4$ arms from B whose empirical mean is at least $1/2 + 8\epsilon$ and at least $K/4$ arms from A whose empirical mean is smaller than $1/2 + 8\epsilon$. We denote the former event by $\mathcal{E}_1$ and the later by $\mathcal{E}_2$. Note that if the event that both $\mathcal{E}_1$ and $\mathcal{E}_2$ happen implies that we fail to find an ϵ-optimal solution.

First, let us consider $\mathcal{E}_1$. Let $Y_i = \mathbf{1}\{\widehat{\theta}_i \ge 1/2 + 8\epsilon\}$. For any arm i in B, we have that

$$\Pr[Y_i = 1] = \Pr[\widehat{\theta}_i \ge 1/2 + 8\epsilon] = \left(\frac{1}{2}\right)^{Q_0} \sum_{i=(1/2+8\epsilon)Q_0}^{Q_0} \binom{Q_0}{i} \ge \left(\frac{1}{2}\right)^{200Q_0\epsilon^2} \ge \left(\frac{n}{K}\right)^{-1/2},$$

where the second to last inequality follows from the fact that $\sum_{k\le\alpha m}\binom{m}{k} \ge 2^{mH(\alpha)-\ln m}$ ($H(\alpha)$ is the binary entropy function) [2] and the Taylor expansion of $H(\alpha)$ around $1/2$: $H(1/2-\epsilon) \simeq 1 - 2\epsilon^2/\ln 2 + o(\epsilon^2)$. Therefore, in expectation, there are at least $(n-K)\left(\frac{n}{K}\right)^{-1/2}$ arms in B whose empirical mean is at least $1/2 + 8\epsilon$, i.e., $\mathbf{E}[\sum_{i\in B} Y_i] \ge (n-K)\left(\frac{n}{K}\right)^{-1/2}$. Using Proposition 1.1(3), we can see that (for $n \ge CK \ge C$)

$$\Pr\left[\sum_{i\in B} Y_i < \frac{K}{4}\right] \le \exp\left(-\left(\frac{1}{2}\right)^2 (n-K)\left(\frac{n}{K}\right)^{-1/2}/2\right) < 0.05,$$

where we use the fact that $(n-K)\left(\frac{n}{K}\right)^{-1/2} \ge K/2$ in the first inequality, and that $(n-K)\left(\frac{n}{K}\right)^{-1/2} \ge \frac{n-1}{\sqrt{n}}$ for $1 \le K \le n/2$ in the second. Hence, with probability at least 0.95, there are at least $K/4$ arms in B whose empirical mean is at least $1/2 + 8\epsilon$.

For any arm in A, using Proposition 1.1(1), we can see that

$$\Pr[\widehat{\theta}_i \ge 1/2 + 8\epsilon] \le \exp\left(-32\epsilon^2 Q_0\right) < 0.5,$$

in which the last inequality holds because $Q_0 \ge \frac{1}{400\epsilon^2}\ln C$. Let $Z_i = \mathbf{1}\{\widehat{\theta}_i \ge 1/2 + 8\epsilon\}$ and $\mu = \exp\left(-32\epsilon^2 Q_0\right) < 0.5$. Then, by Proposition 1.1(2), we have

$$\Pr[\neg\mathcal{E}_2] = \Pr\left[\sum_{i\in A} Z_i > \frac{3K}{4}\right] \le \left(\left(\frac{\mu}{3/4}\right)^{3/4}\left(\frac{1-\mu}{1/4}\right)^{1/4}\right)^K < 0.877.$$

The last inequality holds since $(\frac{\mu}{3/4})^{3/4}(\frac{1-\mu}{1/4})^{1/4}$ is an increasing function on $[0, 0.5]$, thus is maximized at $u = 0.5$. Hence, $\Pr[\mathcal{E}_2] \ge 0.1$. So, we have $\Pr[\mathcal{E}_1 \text{ and } \mathcal{E}_2] \ge 0.05$ and the proof is complete for the case $K \le n/C$. When, $n/C \le K \le n/2$, the desired bound becomes $\Omega(1/\epsilon^2)$, which follows from Theorem 5.1. $\square$

8 Lower bounds

8.1 Proof of the First Lower Bound (Lemma 5.2 in the Main Text)

Lemma 5.2 (restated). *Let $\mathcal{A}$ be an algorithm in Theorem 5.1. There is an algorithm $\mathcal{B}$, which correctly outputs whether a Bernoulli arm X has the mean $\frac{1}{2}+4\epsilon$ or the mean $\frac{1}{2}$ with probability at least 0.51, and $\mathcal{B}$ uses at most $\frac{200Q}{n}$ samples.*

Proof. We note that in step 3(1), only when $\mathcal{A}$ attempts to sample the j-th (artificial) arm, we *actually* take a sample from X. Since $\mathcal{B}$ stops and output the mean $\frac{1}{2}+4\epsilon$ if the number of trials on X reaches $\frac{200Q}{n}$, $\mathcal{B}$ takes at most $\frac{200Q}{n}$ samples form X. Now, we $\mathcal{B}$ can correctly output the mean of X with probability at least 0.51.

We first show that when the Bernoulli arm X has the mean $\frac{1}{2}$, $\mathcal{B}$ decides correctly with probability at least 0.51. Assuming that X has the mean $\frac{1}{2}$, among the n arms in the algorithm $\mathcal{A}$, the arms in $S\setminus\{j\}$ have the mean $\frac{1}{2}+4\epsilon$, while others have the mean $\frac{1}{2}$. For each $i\in[n]$, let the random variable q_i be the number of samples taken from the i-th arm by $\mathcal{A}$. We have

$$\sum_{i\in[n]} \mathbf{E}[q_i] \le Q.$$

Let the random variable q_X be the number of samples taken from arm X and $S'=S\setminus\{j\}$. Observe that when conditioned on S', for $\mathcal{A}$, j is the same as any other arms in $[n]\setminus S'$, hence, j is uniformly distributed among $[n]\setminus S'$. We have

$$\begin{aligned}
\mathbf{E}[q_X] &= \mathbf{E}_{S'}\left[\mathbf{E}[q_X \mid S']\right] = \mathbf{E}_{S'}\left[\frac{1}{n-K+1}\sum_{i\in[n]\setminus S'}\mathbf{E}[q_i\mid S']\right]\\
&\le \frac{1}{n-K+1}\mathbf{E}_{S'}\left[\sum_{i\in[n]}\mathbf{E}[q_i\mid S']\right] \le \frac{2}{n}\sum_{i\in[n]}\mathbf{E}[q_i] \le \frac{2Q}{n},
\end{aligned}$$

where in the second equality we use the fact that j is uniformly distributed among $[n]\setminus S'$ conditioned on S', and in the second inequality we used the assumption that $K\le n/2$. Therefore, by Markov's inequality,

$$\Pr\left[q_X \ge \frac{200Q}{n}\right] < 0.01.$$

Let T be the output of the algorithm $\mathcal{A}$. It is easy to see that $\mathsf{val}_K([n]) = \frac{1}{2}+4\epsilon\cdot(1-\frac{1}{K})$, and $\mathsf{val}_K(T) = \frac{1}{2}+4\epsilon\cdot\frac{|S'\cap T|}{K}$. When $\mathcal{A}$ finds an ϵ-optimal solution (i.e., $\mathsf{val}_K(T)\ge\mathsf{val}_K([n])-\epsilon$), we have

$$\frac{1}{2}+4\epsilon\cdot\frac{|S'\cap T|}{K} \ge \frac{1}{2}+4\epsilon\cdot\left(1-\frac{1}{K}\right)-\epsilon,$$

from which we can get that $|S'\cap T|\ge\frac{3}{4}K-1$. Since $\mathcal{A}$ can find an ϵ-optimal solution with probability at least 0.8, for any fixed subset S', we have that

$$\Pr\left[|S'\cap T|\ge\frac{3}{4}K-1\right] \ge \Pr[\mathsf{val}_K(T)\ge\mathsf{val}_K([n])-\epsilon] \ge 0.8.$$

Conditioned on S', j is uniformly distributed among $[n]\setminus S'$ and is independent from T. Therefore, we have

$$\begin{aligned}
\Pr[j\in T] &= \mathbf{E}_{S'}\left[\Pr[j\in T\mid S']\right] = \mathbf{E}_{S',T}\left[\frac{|([n]\setminus S')\cap T|}{|[n]\setminus S'|}\right]\\
&\le 0.2\times 1 + 0.8\cdot\frac{\frac{1}{4}K+1}{n-K+1} \le 0.2+0.8\times(0.25+0.1) = 0.48,
\end{aligned}$$

where in the last inequality, we used the assumption that $10 \le K \le n/2$. Therefore, when X has the mean $\frac{1}{2}$, we have that

$$\Pr\left[\mathcal{B} \text{ decides that } X \text{ has the mean } \frac{1}{2}\right] = \Pr\left[j \notin T \text{ and } q_X \le \frac{200Q}{n}\right] \ge 0.52 - 0.01 = 0.51.$$

Now we assume that X has the mean $\frac{1}{2} + 4\epsilon$. The proof is similar as before. Among the n arms, the arms in S have the mean $\frac{1}{2} + 4\epsilon$, while others have the mean $\frac{1}{2}$. Again, let T be the output of the algorithm $\mathcal{A}$. Since with probability at least 0.8, we have that $\mathsf{val}_K(T) \ge \mathsf{val}_K([n]) - \epsilon$, we have

$$\Pr\left[|S \cap T| \ge \frac{3}{4}K\right] \ge 0.8.$$

Since j is a uniformly distributed in S, and conditioned on S, j is independent from T, we have that

$$\Pr[j \in T] = \mathbf{E}_S\left[\Pr[j \in T \mid S]\right] = \mathbf{E}_{S,T}\left[\frac{|S \cap T|}{|S|}\right] \ge 0.8 \cdot \frac{3}{4} \ge 0.6.$$

In sum, when X has the mean $\frac{1}{2} + 4\epsilon$, we have that

$$\Pr\left[\mathcal{B} \text{ decides that } X \text{ has mean } \frac{1}{2} + 4\epsilon\right] \ge \Pr\left[j \in T\right] \ge 0.6 > 0.51.$$

In either case, $\mathcal{B}$ makes the right decision with probability at least 0.51. $\square$

8.2 Proof of the Second Lower Bound (Lemma 5.4 and Theorem 5.5 in the Main Text)

Lemma 5.4 (restated). *Fix real numbers δ, ϵ such that $0 < \delta, \epsilon \le 0.01$, and integers K, n such that $K \le n/2$. Let $\mathcal{A}$ be a deterministic algorithm (i.e., the only randomness comes from the sampling the arms), so that for any set of n Bernoulli arms with means $\theta_1, \theta_2, \ldots, \theta_n$,*

- *$\mathcal{A}$ makes at most Q samples in expectation;*
- *with probability at least $1 - \delta$, $\mathcal{A}$ outputs a set T of size K with $\mathsf{val}_K(T) \ge \mathsf{val}_K([n]) - \epsilon$.*

Then, we have that $Q \ge \frac{n \ln(1/\delta)}{20000\epsilon^2 K}$.

Proof of Lemma 5.4. Let $t = \lfloor \frac{n}{K} \rfloor \ge 2$ and we divide the first tK arms into t groups. The j-th group consists of the arms with the index in $[(j-1)K + 1, jK]$ for $j \in [t]$. We first construct t hypotheses $H_1, H_2, \ldots, H_t$ as follows. In H_1, we let $\theta_i = 1/2 + 4\epsilon$ for arms in the first group and let $\theta_i = 1/2$ for the remaining arms. In H_j, where $2 \le j \le t$, we let $\theta_i = 1/2 + 4\epsilon$ when i is in the first group, $\theta_i = 1/2 + 8\epsilon$ when i is in the j-th group, and $\theta_i = 1/2$ otherwise. For each $j \in [t]$, let $\Pr_{H_j}[\cdot]$ denote the probability of the event in $[\cdot]$ under the hypothesis H_j and $\mathbf{E}_{H_j}[\cdot]$ the expected value of the random variable in $[\cdot]$ under the hypothesis H_j.

For each arm $i \in [n]$, let the random variable q_i be the number of times that $\mathcal{A}$ samples the i-th arm before termination. For each $j \in [t]$, let the random variable $\widetilde{q}_j = \sum_{i=(j-1)K+1}^{jK} q_i$ be the total number of trials of the arms in the j-th group. Since $\mathcal{A}$ makes at most Q samples in expectation, we know that $\mathbf{E}_{H_1}\left[\sum_{j=2}^{t} \widetilde{q}_j\right] \le \mathbf{E}_{H_1}\left[\sum_{i\in[n]} q_i\right] \le Q$. By an averaging argument, there exists j_0 with $2 \le j_0 \le t$ such that

$$\mathbf{E}_{H_1}[\widetilde{q}_{j_0}] \le \frac{Q}{t-1} \le \frac{2Q}{t}.$$

Using Markov's inequality and letting $Q_0 = \frac{8Q}{t}$, we have $\Pr_{H_1}[\widetilde{q}_{j_0} \ge Q_0] \le \frac{\mathbf{E}_{H_1}[\widetilde{q}_{j_0}]}{Q_0} \le \frac{1}{4}$, and hence,

$$\Pr_{H_1}[\widetilde{q}_{j_0} \le Q_0] \ge \frac{3}{4}. \tag{15}$$

Now we only focus on the hypotheses H_1 and H_{j_0}. Let $\mathsf{val}_K^{H_j}(T)$ ($\mathsf{val}_K^{H_j}([n])$ resp.) be $\mathsf{val}_K(T)$ ($\mathsf{val}_K([n])$ resp.) under the hypothesis H_j. In other words, $\mathsf{val}_K^{H_j}(T)$ is the average mean value of the best K arms in T, if the the means of the arms are dictated by hypothesis H_j.

Now, we assume for contradiction that $Q < \frac{n\ln(1/\delta)}{20000\epsilon^2 K}$ (i.e., $Q_0 < \frac{\ln(1/\delta)}{1250\epsilon^2}$). Let T denote the output of $\mathcal{A}$. First, using the assumption that

$$\Pr_{H_1}[\mathsf{val}_K^{H_1}(T) \ge \mathsf{val}_K^{H_1}([n]) - \epsilon] \ge 1-\delta \tag{16}$$

(i.e., if the underlying hypothesis is H_1, T is an ϵ-optimal solution), we can prove that:

$$\Pr_{H_{j_0}}[\mathsf{val}_K^{H_1}(T) \ge \mathsf{val}_K^{H_1}([n]) - \epsilon] \ge \frac{\sqrt{\delta}}{4}. \tag{17}$$

We further observe that when $\mathsf{val}_K^{H_1}(T) \ge \mathsf{val}_K^{H_1}([n]) - \epsilon$, T must consist of more than $\frac{3}{4}K$ arms from the first group; while when $\mathsf{val}_K^{H_{j_0}}(T) \ge \mathsf{val}_K^{H_{j_0}}([n]) - \epsilon$, T must consist of more than $\frac{3}{4}K$ arms from the j_0-th group. Therefore, the two events are mutually exclusive and we have:

$$\begin{aligned}\Pr_{H_{j_0}}\left[\mathsf{val}_K^{H_{j_0}}(T) \ge \mathsf{val}_K^{H_{j_0}}([n]) - \epsilon\right] &\le 1 - \Pr_{H_{j_0}}\left[\mathsf{val}_K^{H_1}(T) \ge \mathsf{val}_K^{H_1}([n]) - \epsilon\right] \\ &\le 1 - \frac{\sqrt{\delta}}{4} \le 1 - 2\delta,\end{aligned}$$

where the last inequality holds because $\delta < 0.01$. This essentially says that if the underlying hypothesis is H_{j_0}, the probability that $\mathcal{A}$ finds an ϵ-optimal solution is not large enough, which contradicts the performance guarantees of the Algorithm $\mathcal{A}$, and thus we conclude our proof.

Therefore, the remaining task is to prove (17). We first define a sequence of random variables $Z_0, Z_1, Z_2, \ldots, Z_{Q_0}$ where $Z_0 = 0$. For each $i \in [Q_0]$, if the i-th trial of the j_0-th group by $\mathcal{A}$ results in 1, let $Z_i = Z_{i-1}+1$; if the result is 0, let $Z_i = Z_{i-1} - 1$; if $\mathcal{A}$ terminates before the i-th trial of the j_0-th group, let $Z_i = Z_{i-1}$. Under hypothesis H_0, the sequence $\{Z_0, Z_1, Z_2, \ldots, Z_{Q_0}\}$ forms a martingale since arms in the j_0-th group are independent zero-mean random variables. Therefore, by Azuma-Hoeffding's inequality, we have

$$\Pr_{H_1}\left[|Z_{Q_0}| \le \sqrt{5Q_0}\right] > 1 - 2\exp\left(-\frac{(\sqrt{5Q_0})^2}{2Q_0}\right) > \frac{3}{4}. \tag{18}$$

By a union bound over (15), (16) and (18), we have

$$\Pr_{H_1}\left[\mathsf{val}_K^{H_1}(T) \ge \mathsf{val}_K^{H_1}([n]) - \epsilon \text{ and } \widetilde{q}_{j_0} \le Q_0 \text{ and } |Z_{Q_0}| \le \sqrt{5Q_0}\right] \ge 1 - \delta - \frac{1}{4} - \frac{1}{4} \ge \frac{1}{4}. \tag{19}$$

For ease of notation, we use $\mathcal{E}$ to denote the event that all of the following three events happen: (1) $\mathsf{val}_K^{H_1}(T) \ge \mathsf{val}_K^{H_1}([n]) - \epsilon$, (2) $\widetilde{q}_{j_0} \le Q_0$ and (3) $|Z_{Q_0}| \le \sqrt{5Q_0}$.

Suppose that $\mathcal{A}$ uses exactly Q' trials. We call a string $y = ((i_1, b_1), (i_2, b_2), \ldots, (i_{Q'}, b_{Q'}))$ a transcript for a particular execution of $\mathcal{A}$ if the r-th trial ($1 \le r \le Q'$) performed by $\mathcal{A}$ is the i_r-th arm and the result is $b_r \in \{0,1\}$. Let $\mathcal{Y}$ be the set of transcripts for $\mathcal{A}$. For each $y \in \mathcal{Y}$, we define the following quantities:

- Let $u_0^i(y)$ be the number of $(i,0)$ pairs in y and $u_1^i(y)$ be the number of $(i,1)$ pairs in y;

- Let $q_i(y) = u_0^i(y) + u_1^i(y)$ be the number of times $\mathcal{A}$ takes sample from the i-th arm in y;
- For all $j \in [t]$, let $\widetilde{q}_j(y) = \sum_{i=(j-1)K+1}^{jK} q_i(y)$ be the number of times to sample from the j-th group in y.
- Let $\widetilde{u}_0^j(y) = \sum_{i=(j-1)K+1}^{jK} u_0^i(y)$ be the number of times that sampling from the j-th group results 0; $\widetilde{u}_1^j(y) = \sum_{i=(j-1)K+1}^{jK} u_1^i(y)$ be the number of times that sampling from the j-th group results 1;
- let $T(y)$ be the output of $\mathcal{A}$ when the transcript generated by $\mathcal{A}$ is y (note that the output of $\mathcal{A}$ is completed determined by y since $\mathcal{A}$ is deterministic).

Let the random variable Y be the transcript generated by $\mathcal{A}$. We use $\mathcal{E}_y$ to denote the event that $\mathsf{val}_K^{H_1}(T(y)) \geq \mathsf{val}^{H_1}(K) - \epsilon$ and $\widetilde{q}_{j_0} \leq Q_0$ and $|\widetilde{u}_0^{j_0}(y) - \widetilde{u}_1^{j_0}(y)| \leq \sqrt{5Q_0}$. It is not hard to see that an equivalent way to write (19) is as follows:

$$\sum_{y \in \mathcal{Y}} \mathbf{1}\{\mathcal{E}_y\} \cdot \Pr_{H_1}[Y = y] \geq \frac{1}{4}. \tag{20}$$

Now, we claim that for any $y \in \mathcal{Y}$, we have that

$$\frac{\Pr_{H_{j_0}}[Y = y]}{\Pr_{H_1}[Y = y]} \geq \sqrt{\delta}. \tag{21}$$

Therefore, we have

$$\begin{aligned}\Pr_{H_{j_0}}\left[\mathsf{val}_K^{H_1}(T) \geq \mathsf{val}_K^{H_1}([n]) - \epsilon\right] &\geq \Pr_{H_{j_0}}[\mathcal{E}] = \sum_{y \in \mathcal{Y}} \mathbf{1}\{\mathcal{E}_y\} \cdot \Pr_{H_{j_0}}[Y = y] \\ &\geq \sum_{y \in \mathcal{Y}} \mathbf{1}\{\mathcal{E}_y\} \cdot \Pr_{H_1}[Y = y] \cdot \sqrt{\delta} \geq \frac{\sqrt{\delta}}{4},\end{aligned}$$

where the penultimate inequality is because of (21); and the last inequality is because of (20). Therefore, we finish the proof of the Eq. (17), which concludes the proof the lemma.

What remains is to prove the claim (21). Fix a $y \in \mathcal{Y}$. We first express $\Pr_{H_1}[Y = y]$ and $\Pr_{H_{j_0}}[Y = y]$ in terms of $\widetilde{u}$ and $\widetilde{q}$:

1. $\Pr_{H_1}[Y = y] = \left(\frac{1}{2} + 4\epsilon\right)^{\widetilde{u}_1^1(y)} \left(\frac{1}{2} - 4\epsilon\right)^{\widetilde{u}_0^1(y)} \prod_{j=2}^t \left(\frac{1}{2}\right)^{\widetilde{q}_j(y)}$;
2. $\Pr_{H_{j_0}}[Y = y] = \left(\frac{1}{2} + 4\epsilon\right)^{\widetilde{u}_1^1(y)} \left(\frac{1}{2} - 4\epsilon\right)^{\widetilde{u}_0^1(y)} (1 - 16\epsilon)^{\widetilde{u}_0^{j_0}(y)} (1 + 16\epsilon)^{\widetilde{u}_1^{j_0}(y)} \prod_{j=2}^t \left(\frac{1}{2}\right)^{\widetilde{q}_j(y)}$.

Taking the ratio, we obtain that

$$\begin{aligned}\frac{\Pr_{H_{j_0}}[Y = y]}{\Pr_{H_1}[Y = y]} &= (1 - 16\epsilon)^{\widetilde{u}_0^{j_0}(y)} (1 + 16\epsilon)^{\widetilde{u}_1^{j_0}(y)} = (1 - 16\epsilon)^{\frac{\widetilde{q}_{j_0}(y)}{2} + \frac{\widetilde{u}_0^{j_0}(y) - \widetilde{u}_1^{j_0}(y)}{2}} (1 + 16\epsilon)^{\frac{\widetilde{q}_{j_0}(y)}{2} - \frac{\widetilde{u}_0^{j_0}(y) - \widetilde{u}_1^{j_0}(y)}{2}} \\ &= (1 - 256\epsilon^2)^{\frac{\widetilde{q}_{j_0}(y)}{2}} \left(\frac{1 - 16\epsilon}{1 + 16\epsilon}\right)^{\frac{\widetilde{u}_0^{j_0}(y) - \widetilde{u}_1^{j_0}(y)}{2}} \geq (1 - 256\epsilon^2)^{\frac{\widetilde{q}_{j_0}(y)}{2}} (1 - 32\epsilon)^{\left|\frac{\widetilde{u}_0^{j_0}(y) - \widetilde{u}_1^{j_0}(y)}{2}\right|}. \end{aligned} \tag{22}$$

When both $\widetilde{q}_{j_0} \leq Q_0$ and $|u_0^{j_0}(y) - u_1^{j_0}(y)| \leq \sqrt{5Q_0}$ hold, we have (recall that $Q_0 \leq \frac{\ln(1/\delta)}{1250\epsilon^2}$)

$$\begin{aligned}(1 - 256\epsilon^2)^{\frac{\widetilde{q}_{j_0}(y)}{2}} (1 - 32\epsilon)^{\left|\frac{\widetilde{u}_0^{j_0}(y) - \widetilde{u}_1^{j_0}(y)}{2}\right|} &\geq (1 - 256\epsilon^2)^{Q_0/2} (1 - 32\epsilon)^{\sqrt{5Q_0}/2} \\ &\geq (1 - 256\epsilon^2)^{\frac{\ln(1/\delta)}{2500\epsilon^2}} (1 - 32\epsilon)^{\frac{\sqrt{\ln(1/\delta)}}{30\epsilon}} \geq \delta^{1/4} \cdot \delta^{1/4} = \sqrt{\delta},\end{aligned}$$

where in the penultimate inequality we used the assumption that $0 < \epsilon, \delta \leq 0.01$. This proves (21). □

Theorem 5.5 (restated). *Fix real numbers δ, ϵ such that $0 < \delta, \epsilon \le 0.01$, and integers K, n, such that $K \le n/2$. Let $\mathcal{A}$ be a (possibly randomized) algorithm so that for any set of n Bernoulli arms with the mean $\theta_1, \theta_2, \dots, \theta_n$,*

- *$\mathcal{A}$ makes at most Q samples in expectation;*
- *With probability at least $1-\delta$, $\mathcal{A}$ outputs a set T of size K with $\mathsf{val}_K(T) \ge \mathsf{val}_K([n]) - \epsilon$.*

We have that $Q = \Omega\big(\frac{n \ln(1/\delta)}{\epsilon^2 K}\big)$.

Proof of Theorem 5.5. We show that essentially the same lower bound also holds for any randomized algorithm. The following argument is standard and we include it for completeness. Fix $0 < \epsilon, \delta < 1/2$. We assume, for contradiction, that there is a randomized algorithm $\mathcal{A}$ which can achieve the same performance guarantee stated as in the theorem, but the expected number Q of samples is no more than $\frac{n \ln(1/\delta)}{100000\epsilon^2 K}$. We can view the randomized algorithm $\mathcal{A}$ as a deterministic algorithm with a sequence S of random bits. We use R to denote the randomness from the arms. Note that if we fix S and R, the execution and the output of the algorithm are fixed. We use $\mathcal{A}(S,R) = 1$ to denote the event that the output of $\mathcal{A}$ is an ϵ-optimal solution. Let us use $Q(S,R)$ to denote the number of samples taken by $\mathcal{A}$. The performance guarantee of $\mathcal{A}$ is that

$$\Pr_{S,R}[\mathcal{A}(S,R)=1] = \mathbf{E}_{S,R}[\mathcal{A}(S,R)] = \mathbf{E}_S\, \mathbf{E}_R[\mathcal{A}(S,R) \mid S] \ge 1-\delta.$$

This is equivalent to say that $\mathbf{E}_S\, \mathbf{E}_R[1-\mathcal{A}(S,R) \mid S] \le \delta$. By Markov inequality, we have that $\Pr_S\Big[\mathbf{E}_R[1-\mathcal{A}(S,R) \mid S] \ge 2\delta\Big] \le 1/2$. Equivalently, we have that

$$\Pr_S\Big[\mathbf{E}_R[\mathcal{A}(S,R) \mid S] \ge 1-2\delta\Big] \ge 1/2. \tag{23}$$

By our assumption, we have $\mathbf{E}_{S,R}\, Q(S,R) \le \frac{n\ln(1/\delta)}{100000\epsilon^2 K}$. So, by Markov inequality,

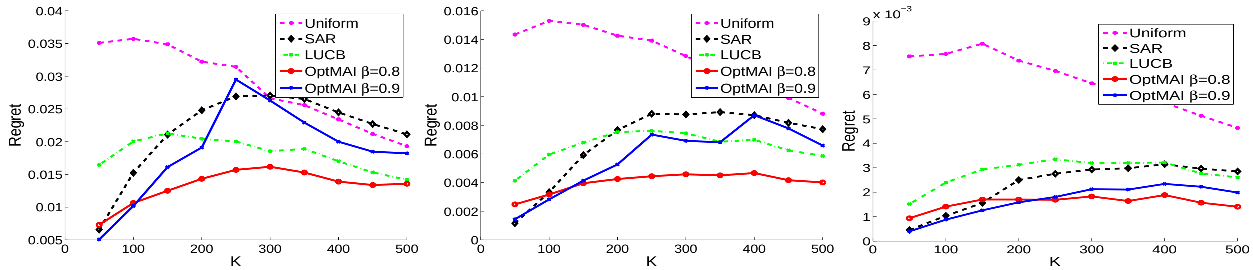
$$\Pr_S\Big[\mathbf{E}_R[Q(S,R) \mid S] \le \frac{n\ln(1/\delta)}{40000\epsilon^2 K}\Big] \ge \frac{3}{5}. \tag{24}$$

Combining (23) and (24), we know there is a particular random sequence S such that both $\mathbf{E}_R[\mathcal{A}(S,R) \mid S] \ge 1-2\delta$ and $\mathbf{E}_R[Q(S,R) \mid S] \le \frac{n\ln(1/\delta)}{40000\epsilon^2 K}$ hold. Since the algorithm $\mathcal{A}$ with a particular sequence S is simply a deterministic algorithm, this contradicts the lower bound we proved for any deterministic algorithm in Lemma 5.4. $\square$

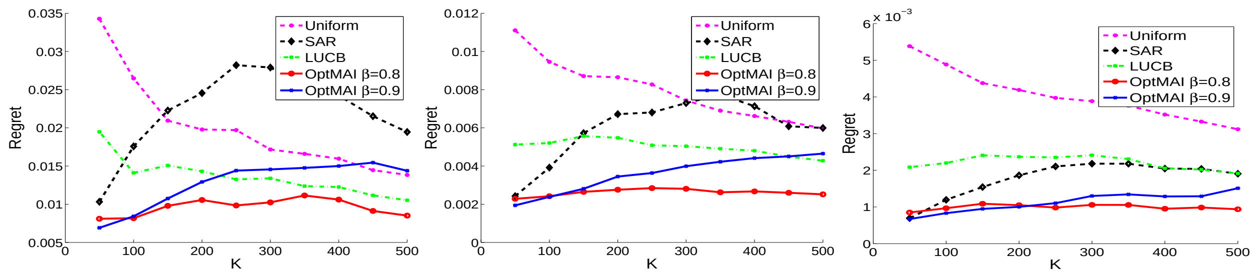
Theorem 5.6 (restated). *Fix real numbers δ, ϵ such that $0 < \delta, \epsilon \le 0.01$, and integers K, n such that $K \ge n/2$. Let $\mathcal{A}$ be a (possibly randomized) algorithm such that for any set of n Bernoulli arms, $\mathcal{A}$ can output an ϵ-optimal set T of size K, with probability at least $1-\delta$, using at most Q samples in expectation. We have that*

$$Q = \Omega\left(\Big(\frac{n-K}{K}\cdot\frac{n}{\epsilon^2}\Big)\Big(\frac{n-K}{K} + \frac{\ln(1/\delta)}{K}\Big)\right).$$

Proof. In fact, in the proof of Theorem 4.5, we have established the equivalence between identifying an ϵ-optimal solution of size K and an ϵ'-optimal solution of size $n-K$, where $\epsilon' = \frac{K}{n-K}\cdot\epsilon$. Since $n-K \le n/2$, we can use the lower bounds developed in Theorem 5.1 and Theorem 5.5, which show that Q should be at least $Q = \Omega\left(\frac{n}{\epsilon'^2}\left(1+\frac{\ln(1/\delta)}{n-K}\right)\right)$. Plugging in $\epsilon' = \frac{K}{n-K}\cdot\epsilon$, we obtain the desired lower bound. $\square$



(a) $\theta \sim \text{TN}(0.5, 0.2)$, $Q = 20 \cdot n$ (b) $\theta \sim \text{TN}(0.5, 0.2)$, $Q = 50 \cdot n$ (c) $\theta \sim \text{TN}(0.5, 0.2)$, $Q = 100 \cdot n$



(d) $\theta \sim \text{Beta}(4, 1)$, $Q = 20 \cdot n$ (e) $\theta \sim \text{Beta}(4, 1)$, $Q = 50 \cdot n$ (f) $\theta \sim \text{Beta}(4, 1)$, $Q = 100 \cdot n$

Figure 1: Performance comparison on simulated data.

9 Additional Experiments

In this section, we provide additional simulated experimental results when using the following two different ways to generate $\{\theta_i\}_{i=1}^n$:

1. $\theta \sim \text{TN}(0.5, 0.2)$: each θ_i is generated from a truncated normal distribution with mean 0.5, the standard deviation 0.2 and the support $[0, 1]$ (Figure 1(a) to Figure 1(c)).
2. $\theta \sim \text{Beta}(4, 1)$: each θ_i is generated from a Beta distribution with the parameters $(4, 1)$. The $\{\theta_i\}$ from $\text{Beta}(4, 1)$ are close to the workers' accuracy in real crowdsourcing applications, where most workers perform reasonably well and the averaged accuracy is around 80% (Figure 1(d) to Figure 1(f)).

We note that the number of total arms is set to $n = 1000$. We vary the total budget $Q = 20n, 50n, 100n$ and $K = 10, 20, \ldots, 500$. We use different ways to generate $\{\theta_i\}_{i=1}^n$ and report the comparison among different algorithms. It can be seen from Figure 1 that our method outperforms the SAR and LUCB in most of the scenarios. In addition, we also observe that when K is large, the setting of $\beta = 0.8$ outperforms that of $\beta = 0.9$; while for small K, $\beta = 0.9$ is a better choice.

References

[1] Eyal Even-Dar, Shie Mannor, and Yishay Mansour. Action elimination and stopping conditions for the multi-armed bandit and reinforcement learning problems. *Journal of machine learning research*, 7:1079–1105, 2006. 8

[2] L Graham Ronald, E Knuth Donald, and Patashnik Oren. *Concrete Mathematics; A Foundation for Computer Science*. Addison-Wesley, 1989. 9