Organization of the Appendix

In the appendix, we present the missing proofs in this paper. In Appendix A, we first discuss a specific instance mentioned in Section 1, showing that our upper bound strictly improves previous algorithms. In Appendix B, we prove Fact 2.2 in Section 2. In Appendix C, we prove the Instance Embedding lemma (Lemma 3.1) and the relatively technical Lemma 3.3, which relates a general instance of Best-1-Arm to a symmetric instance. In Appendix D, we discuss the implementation of the building blocks of our algorithm, prove a few useful and observations, and finally complete the missing proofs of other lemmas and theorems.

A Specific Best-k-Arm Instance

We show that our upper bound results (Theorem 1.2 and Theorem 1.3) strictly improve the state-of-the-art algorithm for Best-k-Arm obtained in [CGL16] by calculating the sample complexity of both algorithms on a specific Best-k-Arm instance.

We consider a family of instances parametrized by integer n and $\varepsilon \in (0, 1/4)$. Each instance consists of n arms with mean 0, n arms with mean 1/2, along with two arms with means $1/4+\varepsilon$ and $1/4-\varepsilon$ respectively. We are required to identify the top $n+1$ arms. By definition, the gap of every arm with mean 0 or 1/2 is $1/4+\varepsilon$, while the gaps of the remaining two arms are 2ε. As ε tends to zero, the arms with gap $1/4+\varepsilon$ become relatively simple: an algorithm can decide whether to include them in the answer or not with few samples. The hardness of the instance is then concentrated on the two arms with close means.

For simplicity, we assume that the confidence level, δ, is set to a constant. Then the $O(H \ln \delta^{-1})$ term in the upper bounds are dominated by the $O(\widetilde{H})$ term. By a direct calculation, we have

$$\widetilde{H} = \Theta(n + \varepsilon^{-2} \ln\ln \varepsilon^{-1}).$$

Let m be the integer that satisfies $2\varepsilon \in (\varepsilon_{m+1}, \varepsilon_m]$. Then we have

$$|G_1^{\mathsf{large}}| = |G_1^{\mathsf{small}}| = n, \text{ and}$$
$$|G_m^{\mathsf{large}}| = |G_m^{\mathsf{small}}| = 1.$$

It follows from the definition of $\widetilde{H}^{\mathsf{large}}$ and $\widetilde{H}^{\mathsf{small}}$ that

$$\widetilde{H}^{\mathsf{large}} = \widetilde{H}^{\mathsf{small}} = O(n \ln n + \varepsilon^{-2}).$$

By Theorem 1.2, for constant δ, our algorithm takes

$$O(\widetilde{H} + \widetilde{H}^{\mathsf{large}} + \widetilde{H}^{\mathsf{small}}) = O(n \ln n + \varepsilon^{-2} \ln\ln \varepsilon^{-1})$$

samples on this instance.

On the other hand, the upper bound achieved by PAC-SamplePrune algorithm is

$$O(\widetilde{H} + H \ln n) = O(n \ln n + \varepsilon^{-2} \ln\ln \varepsilon^{-1} + \varepsilon^{-2} \ln n).$$

Note that if $\varepsilon = 1/n$, our algorithm takes $O(n^2 \ln\ln n)$ samples, while PAC-SamplePrune takes $O(n^2 \ln n)$ samples. This indicates that there is a logarithmic gap between the state-of-the-art upper bound and the instance-wise lower bound, while we narrow down the gap to a doubly-logarithmic factor.

B Missing Proof in Section 2

Fact 2.2 (restated) *For $0 \le y \le y_0 \le x_0 \le x \le 1$, $d(x, y) \ge d(x_0, y_0)$.*

Proof of Fact 2.2. Taking the partial derivative yields

$$\frac{\partial d(x,y)}{\partial x} = \ln \frac{x(1-y)}{y(1-x)},$$
$$\frac{\partial d(x,y)}{\partial y} = \frac{y-x}{y(1-y)}.$$

Therefore when $x \ge y$, $d(x, y)$ is increasing in x and decreasing in y, which proves the fact. □

C Missing Proofs in Section 3

C.1 Proof of Lemma 3.1

Lemma 3.1 (restated) *Let $\mathcal{I}$ be a Best-k-Arm instance. Let A be an arm among the top k arms, and $\mathcal{I}^{\mathsf{emb}}$ be a Best-1-Arm instance consisting of A and a subset of arms in $\mathcal{I}$ outside the top k arms. If some algorithm $\mathbb{A}$ solves $\mathcal{I}$ with probability $1-\delta$ while taking less than N samples on A in expectation, there exists another algorithm $\mathbb{A}^{\mathsf{emb}}$ that solves $\mathcal{I}^{\mathsf{emb}}$ with probability $1-\delta$ while taking less than N samples on A in expectation.*

Proof of Lemma 3.1. We construct the following algorithm $\mathbb{A}^{\mathsf{emb}}$ for $\mathcal{I}^{\mathsf{emb}}$. Given the instance $\mathcal{I}^{\mathsf{emb}}$, $\mathbb{A}^{\mathsf{emb}}$ first augments the instance into $\mathcal{I}$ by adding a fictitious arm for each arm in $\mathcal{I} \setminus \mathcal{I}^{\mathsf{emb}}$. Then $\mathbb{A}^{\mathsf{emb}}$ simulates $\mathbb{A}$ on the Best-k-Arm instance $\mathcal{I}$. When $\mathbb{A}$ requires a sample from an arm in $\mathcal{I}^{\mathsf{emb}}$, $\mathbb{A}^{\mathsf{emb}}$ draws a sample and sends it to $\mathbb{A}$. If $\mathbb{A}$ requires a sample from an arm outside $\mathcal{I}^{\mathsf{emb}}$, $\mathbb{A}^{\mathsf{emb}}$ generates a fictitious sample on its own and then sends it to $\mathbb{A}$. When $\mathbb{A}$ terminates and returns a subset S of k arms, $\mathbb{A}^{\mathsf{emb}}$ terminates and returns an arbitrary arm in $S \cap \mathcal{I}^{\mathsf{emb}}$.

Note that when $\mathbb{A}^{\mathsf{emb}}$ runs on instance $\mathcal{I}^{\mathsf{emb}}$, the algorithm $\mathbb{A}$ simulated by $\mathbb{A}^{\mathsf{emb}}$ effectively runs on the instance $\mathcal{I}$. It follows that with probability $1-\delta$, $\mathbb{A}$ returns the correct answer of the Best-k-Arm instance $\mathcal{I}$, and thus A is the only arm in both $\mathcal{I}^{\mathsf{emb}}$ and the set S returned by $\mathbb{A}$. Therefore, $\mathbb{A}^{\mathsf{emb}}$ correctly solves the Best-1-Arm instance $\mathcal{I}^{\mathsf{emb}}$ with probability at least $1-\delta$. Moreover, the expected number of samples drawn from arm A is less than N by our assumptions. □

C.2 Proof of Lemma 3.3

Lemma 3.3 (restated) *Let $\mathcal{I}$ be an instance of Best-1-Arm consisting of one arm with mean μ and n arms with means on $[\mu-\Delta,\mu)$. There exist universal constants δ and c (independent of n and Δ) such that for all algorithm $\mathbb{A}$ that correctly solves $\mathcal{I}$ with probability $1-\delta$, the expected number of samples drawn from the optimal arm is at least $c\Delta^{-2}\ln n$.*

Proof of Lemma 3.3. Let δ_0 and c_0 be the constants in Lemma 3.4. We claim that Lemma 3.3 holds for constants $\delta=\delta_0/3$ and $c=c_0\delta_0/30$.

Suppose for a contradiction that when algorithm $\mathbb{A}$ runs on Best-1-Arm instance $\mathcal{I}$, it outputs the correct answer with probability $1-\delta$ and the optimal arm A_0 is sampled less than $c\Delta^{-2}\ln n$ times in expectation.

Overview. Our proof follows the following five steps.

Step 1. We apply Instance Embedding to obtain a smaller yet denser (in the sense that all suboptimal arms have almost identical means) instance $\mathcal{I}^{\mathsf{dense}}$, together with a new algorithm $\mathbb{A}^{\mathsf{new}}$ that solves $\mathcal{I}^{\mathsf{dense}}$ by taking few samples on the optimal arm with high probability.

Step 2. We obtain a symmetric instance $\mathcal{I}^{\mathsf{sym}}$ from $\mathcal{I}^{\mathsf{dense}}$ by making the suboptimal arms identical to each other. We also define an algorithm $\mathbb{A}^{\mathsf{sym}}$ for instance $\mathcal{I}^{\mathsf{sym}}$.

Step 3. To analyze algorithm $\mathbb{A}^{\mathsf{sym}}$ on instance $\mathcal{I}^{\mathsf{sym}}$, we define the notion of "mixed arms", which return a fixed number of samples from one distribution, and then switch to another distribution permanently. We transform $\mathcal{I}^{\mathsf{dense}}$ into an intance $\mathcal{I}^{\mathsf{mix}}$ with mixed arms.

Step 4. We show by Change of Distribution that when $\mathbb{A}^{\mathsf{new}}$ runs on $\mathcal{I}^{\mathsf{mix}}$, it also returns the correct answer with few samples on the optimal arm.

Step 5. We show that the execution of $\mathbb{A}^{\mathsf{sym}}$ on $\mathcal{I}^{\mathsf{sym}}$ is, in a sense, equivalent to the execution of $\mathbb{A}^{\mathsf{new}}$ on $\mathcal{I}^{\mathsf{mix}}$. This finally leads to a contradiction to Lemma 3.4.

The reductions involved in the proof is illustrated in Figure 1.

Step 1: Construct $\mathcal{I}^{\mathsf{dense}}$ and $\mathbb{A}^{\mathsf{new}}$. We first construct a new Best-1-Arm instance $\mathcal{I}^{\mathsf{dense}}$ in which the sub-optimal arms have almost identical means. Let μ_0 denote the mean of the optimal arm A_0. We divide the interval $[\mu_0-\Delta,\mu_0]$ into $n^{0.9}$ segments, each with length $\Delta/n^{0.9}$. Set $m=n^{0.1}$. By the pigeonhole principle, we can assume that $A_1,A_2,\ldots,A_m$ are m arms with means in the same interval. Let μ_i denote the mean of arm A_i. By construction, $\mu_0-\mu_i\le\Delta$ for all $1\le i\le m$ and $|\mu_i-\mu_j|\le\Delta/n^{0.9}$ for all $1\le i,j\le m$.

We simply let $\mathcal{I}^{\mathsf{dense}}=\{A_0,A_1,A_2,\ldots,A_m\}$. By Instance Embedding (Lemma 3.1), there exists an algorithm $\mathbb{A}^{\mathsf{new}}$ that solves $\mathcal{I}^{\mathsf{dense}}$ with probability $1-\delta$ while taking less than $c\Delta^{-2}\ln n$ samples on A_0 in expectation. We will focus on instance $\mathcal{I}^{\mathsf{dense}}$ in the rest of our proof.

Recall that $\Pr_{\mathbb{A},\mathcal{I}}$ and $\mathrm{E}_{\mathbb{A},\mathcal{I}}$ denote the probability and expectation when algorithm $\mathbb{A}$ runs on instance $\mathcal{I}$ respectively. Let τ_i denote the number of samples taken on A_i. Then we have

$$\mathrm{E}_{\mathbb{A}^{\mathsf{new}},\mathcal{I}^{\mathsf{dense}}}[\tau_0]\le c\Delta^{-2}\ln n.$$

Let $N=c\delta^{-1}\Delta^{-2}\ln n$. By Markov's inequality,

$$\Pr_{\mathbb{A}^{\mathsf{new}},\mathcal{I}^{\mathsf{dense}}}[\tau_0\ge N]\le\frac{c\Delta^{-2}\ln n}{N}=\delta.$$

Let $\mathcal{E}$ denote the event that the algorithm returns the correct answer while taking at most N samples on arm A_0. The union bound implies that

$$\Pr_{\mathbb{A}^{\mathsf{new}},\mathcal{I}^{\mathsf{dense}}}[\mathcal{E}]\ge 1-2\delta.$$

Step 2: Construct $\mathcal{I}^{\mathsf{sym}}$ and $\mathbb{A}^{\mathsf{sym}}$. Let $\mathcal{I}^{\mathsf{sym}}$ be the Best-1-Arm instance consisting of arm A_0 and $m=n^{0.1}$ copies of arm A_1. We define algorithm $\mathbb{A}^{\mathsf{sym}}$ as follows. Given instance $\mathcal{I}^{\mathsf{sym}}$, $\mathbb{A}^{\mathsf{sym}}$ simulates algorithm $\mathbb{A}^{\mathsf{new}}$ as if $\mathbb{A}^{\mathsf{new}}$ is running on instance $\mathcal{I}^{\mathsf{dense}}$. When $\mathbb{A}^{\mathsf{new}}$ requires a sample from an arm A that has not been pulled N times (recall that $N=c\delta^{-1}\Delta^{-2}\ln n$), $\mathbb{A}^{\mathsf{sym}}$ draws a sample from A and sends it to $\mathbb{A}^{\mathsf{new}}$. When the number of pulls on A exceeds N for the first time, $\mathbb{A}^{\mathsf{sym}}$ assigns a random number $\pi(A)$ from $\{1,2,\ldots,m\}$ to arm A, such that $\pi(A)$ is different from every number that has already been assigned to another arm. If this step cannot be performed because all numbers in $\{1,2,\ldots,m\}$ have been used up, $\mathbb{A}^{\mathsf{sym}}$ simply terminates without returning an answer.[4] After that, upon each pull of A, $\mathbb{A}^{\mathsf{sym}}$ sends a sample drawn from $\mathcal{N}(\mu_{\pi(A)},1)$ to $\mathbb{A}^{\mathsf{new}}$. (Recall that μ_i denotes the mean of arm A_i in $\mathcal{I}^{\mathsf{dense}}$.) Finally, $\mathbb{A}^{\mathsf{sym}}$ outputs what $\mathbb{A}^{\mathsf{new}}$ outputs.

Step 3: Construct mixed arms and $\mathcal{I}^{\mathsf{mix}}$. In order to analyze the execution of $\mathbb{A}^{\mathsf{sym}}$ on instance $\mathcal{I}^{\mathsf{sym}}$, it is

[4] As shown in the analysis in Step 5, we only care the behavior of $\mathbb{A}^{\mathsf{sym}}$ when the labels are not used up.

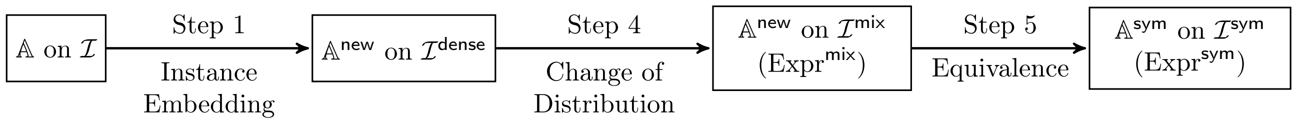

Figure 1: Each rectangle denotes the execution of an algorithm on an instance. The arrows specify the step in which each reduction is performed and the major technique involved in the reduction.

helpful to define m "mixed arms". For $1 \le i \le m$, the i-th mixed arm, denoted by M_i, returns a sample drawn from $\mathcal{N}(\mu_1, 1)$ (i.e., the reward distribution of arm A_1) when it is pulled for the first N times. After N pulls, M_i returns samples from $\mathcal{N}(\mu_i, 1)$ as A_i does. For ease of notation, we also let M_0 denote A_0. Let $\mathcal{I}^{\mathsf{mix}}$ denote the Best-1-Arm instance $\{M_0, M_1, M_2, \dots, M_m\}$.

Step 4: Run $\mathbb{A}^{\mathsf{new}}$ on $\mathcal{I}^{\mathsf{mix}}$. Now suppose we run $\mathbb{A}^{\mathsf{new}}$ on instance $\mathcal{I}^{\mathsf{mix}}$. In fact, we may view each arm (either A_i or M_i) as two separate "semi-arms". When $\mathbb{A}^{\mathsf{new}}$ samples arm A_i in the first N times, it pulls the first semi-arm of A_i. After A_i has been pulled N times, $\mathbb{A}^{\mathsf{new}}$ pulls the second semi-arm. From this perspective, $\mathcal{I}^{\mathsf{mix}}$ is simply obtained from $\mathcal{I}^{\mathsf{dense}}$ by changing the first semi-arm of each arm A_i ($1 \le i \le m$) from $\mathcal{N}(\mu_i, 1)$ to $\mathcal{N}(\mu_1, 1)$. Since the first semi-arm is sampled at most N times by $\mathbb{A}^{\mathsf{new}}$, it follows from Change of Distribution (Lemma 2.1) that

$$\begin{aligned} & d\left(\Pr_{\mathbb{A}^{\mathsf{new}},\mathcal{I}^{\mathsf{dense}}}[\mathcal{E}], \Pr_{\mathbb{A}^{\mathsf{new}},\mathcal{I}^{\mathsf{mix}}}[\mathcal{E}]\right) \\ \le & \sum_{i=1}^{m} N \cdot \mathrm{KL}\left(\mathcal{N}(\mu_i, 1), \mathcal{N}(\mu_1, 1)\right) \\ = & \frac{N}{2}\sum_{i=1}^{m}(\mu_i - \mu_1)^2 \\ \le & \frac{c\delta^{-1}\Delta^{-2}\ln n}{2} \cdot n^{0.1} \cdot (\Delta/n^{0.9})^2 \le \frac{c}{2\delta} n^{-1.7}\ln n. \end{aligned}$$

Here the second step follows from

$$\mathrm{KL}(\mathcal{N}(\mu_1, 1), \mathcal{N}(\mu_2, 1)) = (\mu_1 - \mu_2)^2/2.$$

The third step is due to $N = c\delta^{-1}\Delta^{-2}\ln n$, $m = n^{0.1}$, and $|\mu_1 - \mu_i| \le \Delta/n^{0.9}$.

For sufficiently large n, we have

$$\frac{c}{2\delta} n^{-1.7}\ln n < d(1 - 2\delta, 1 - 3\delta).$$

Recall that $\Pr_{\mathbb{A}^{\mathsf{new}},\mathcal{I}^{\mathsf{dense}}}[\mathcal{E}] \ge 1 - 2\delta$. It follows from the monotonicity of $d(\cdot,\cdot)$ (Fact 2.2) that

$$\Pr_{\mathbb{A}^{\mathsf{new}},\mathcal{I}^{\mathsf{mix}}}[\mathcal{E}] \ge 1 - 3\delta.$$

Step 5: Analyze $\mathbb{A}^{\mathsf{sym}}$ and derive a contradiction to Lemma 3.4. For clarity, let $\mathrm{Expr}^{\mathsf{mix}}$ denote the experiment that $\mathbb{A}^{\mathsf{new}}$ runs on $\mathcal{I}^{\mathsf{mix}}$, and $\mathrm{Expr}^{\mathsf{sym}}$ denote the experiment that $\mathbb{A}^{\mathsf{sym}}$ runs on $\mathcal{I}^{\mathsf{sym}}$. Step 4 implies that event $\mathcal{E}$ happens with probability at least $1 - 3\delta$ in experiment $\mathrm{Expr}^{\mathsf{mix}}$.

In the following, we derive the likelihood of an arbitrary execution of $\mathrm{Expr}^{\mathsf{mix}}$ in which event $\mathcal{E}$ happens, and prove that this execution has the same likelihood in experiment $\mathrm{Expr}^{\mathsf{sym}}$. As a result, $\mathbb{A}^{\mathsf{sym}}$ also returns the correct answer with probability at least $1 - 3\delta$. Moreover, according to our construction, $\mathbb{A}^{\mathsf{sym}}$ always takes at most N samples on arm A_0. On the other hand, since $\mu_0 - \mu_1 \le \Delta$, Lemma 3.4 implies that no algorithm can solve $\mathbb{A}^{\mathsf{sym}}$ correctly with probability $1 - \delta_0 = 1 - 3\delta$ while taking less than

$$c_0\Delta^{-2}\ln m = 30c\delta_0^{-1} \cdot \Delta^{-2} \cdot (0.1\ln n) = N$$

samples on A_0 in expectation. This leads to a contradiction and finishes the proof.

Technicalities: equivalence between $\mathrm{Expr}^{\mathsf{mix}}$ and $\mathrm{Expr}^{\mathsf{sym}}$. For ease of notation, we assume in the following that algorithm $\mathbb{A}^{\mathsf{new}}$ is deterministic.[5] Then the only randomness in experiment $\mathrm{Expr}^{\mathsf{mix}}$ stems from the random permutation of arms at the beginning, and the samples drawn from the arms.

We consider an arbitrary run of experiment $\mathrm{Expr}^{\mathsf{mix}}$ in which event $\mathcal{E}$ happens (i.e., $\mathbb{A}^{\mathsf{new}}$ returns the optimal arm before taking more than N samples from it). For $0 \le i \le m$, let $\sigma(i)$ denote the index of the i-th arm received by algorithm $\mathbb{A}^{\mathsf{new}}$. (i.e., the i-th arm received by $\mathbb{A}^{\mathsf{new}}$ is $M_{\sigma(i)}$.) By definition, σ is a uniformly random permutation of $\{0, 1, \dots, m\}$. Let obs_i denote the sequence of samples that $\mathbb{A}^{\mathsf{new}}$ observes from the i-th arm. Then the likelihood of this execution is given by

$$\frac{1}{(m+1)!}\sum_{\sigma}\prod_{i=0}^{m} f_{M_{\sigma(i)}}(\mathrm{obs}_i). \tag{8}$$

Here we sum over all random permutations σ on $\{0, 1, 2, \dots, m\}$, and $f_{M_{\sigma(i)}}(\mathrm{obs}_i)$ denote the probability density of observing obs_i on arm $M_{\sigma(i)}$.

Now we compute the likelihood that in experiment $\mathrm{Expr}^{\mathsf{sym}}$, the algorithm $\mathbb{A}^{\mathsf{new}}$ simulated by $\mathbb{A}^{\mathsf{sym}}$ ob-

[5] In fact this assumption is without loss of generality: the argument still holds conditioning on the randomness of $\mathbb{A}^{\mathsf{new}}$.

serves the same sequence of samples. Let λ denote the random permutation of arms given to $\mathbb{A}^{\mathsf{sym}}$. We define

$$p^* = \lambda^{-1}(0),$$

$$\mathsf{Long} = \{i \in \{0, 1, 2, \ldots, m\} : |\mathrm{obs}_i| > N\}, \text{ and}$$
$$\mathsf{Short} = \{0, 1, \ldots, m\} \setminus (\mathsf{Long} \cup \{p^*\}).$$

In other words, p^* is the position of the optimal arm A_0 in $\mathcal{I}^{\mathsf{sym}}$. Long denote the positions of suboptimal arms that have been sampled more than N times, while Short denote the remaining suboptimal arms. Note that since less than N samples are taken on the optimal arm, p^* is excluded from both sets.

Another source of randomness in $\mathrm{Expr}^{\mathsf{sym}}$ is the random numbers $\pi(\cdot)$ that $\mathbb{A}^{\mathsf{sym}}$ assigns to different arms. In this specific execution, function $\pi(\cdot)$ chosen by $\mathbb{A}^{\mathsf{sym}}$ is a random injection from Long to $\{1, 2, \ldots, m\}$. By our construction of $\mathbb{A}^{\mathsf{sym}}$, for each $i \in \mathsf{Long}$, the algorithm $\mathbb{A}^{\mathsf{new}}$ simulated by $\mathbb{A}^{\mathsf{sym}}$ first observes N samples drawn from $\mathcal{N}(\mu_1, 1)$ (i.e., the reward distribution of arm A_1) on the i-th arm. After that, $\mathbb{A}^{\mathsf{new}}$ starts to observe samples drawn from $\mathcal{N}(\mu_{\pi(i)}, 1)$. Recall that the mixed arm $M_{\pi(i)}$ also returns samples in this pattern. Therefore, the likelihood of observations on the i-th arm is exactly

$$f_{M_{\pi(i)}}(\mathrm{obs}_i). \tag{9}$$

In fact, we may express the likelihood for all arms as in (9) by extending π into a permutation on $\{0, 1, 2, \ldots, m\}$. First, we set $\pi(p^*) = 0$. Recall that the optimal arm is sampled less than N times, all the samples observed from it are drawn from $\mathcal{N}(\mu_0, 1)$, which is exactly the reward distribution of $M_0 = M_{\pi(p^*)}$. Therefore the likelihood of observations obs_{p^*} is given by

$$f_{M_{\pi(p^*)}}(\mathrm{obs}_{p^*}).$$

Second, we let $R = \{1, 2, \ldots, m\} \setminus \pi(\mathsf{Long})$ denote the available labels among $\{1, 2, \ldots, m\}$. We define π on Short by matching Short with R uniformly at random. Note that since all arms in Short are sampled at most N times, $\mathbb{A}^{\mathsf{new}}$ simulated by $\mathbb{A}^{\mathsf{sym}}$ always observes samples drawn from $\mathcal{N}(\mu_1, 1)$, which agrees with the first N samples from every mixed arm M_i ($i \neq 0$). Therefore, the likelihood of observations on the i-th arm where $i \in \mathsf{Short}$ is also given by

$$f_{M_{\pi(i)}}(\mathrm{obs}_i).$$

According to our analysis above, the samples from the i-th arm observed by the simulated $\mathbb{A}^{\mathsf{new}}$ in experiment $\mathrm{Expr}^{\mathsf{sym}}$ follows the same distribution as samples drawn from $M_{\pi(i)}$. Moreover, π is a uniformly random permutation with the only condition that $\pi(p^*) = 0$, which is equivalent to $\pi^{-1}(0) = p^* = \lambda^{-1}(0)$. Therefore, the likelihood is given by

$$\frac{1}{m! \cdot (m+1)!} \sum_{\pi^{-1}(0) = \lambda^{-1}(0)} \prod_{i=0}^{m} f_{M_{\pi(i)}}(\mathrm{obs}_i). \tag{10}$$

Note that conditioning on $\lambda^{-1}(0) = \pi^{-1}(0)$, π is still a uniformly random permutation on $\{0, 1, 2, \ldots, m\}$. Therefore the two likelihoods in (8) and (10) are equal. This finishes the proof of the equivalence. $\square$

D Missing Proofs in Section 4

D.1 Building Blocks

D.1.1 PAC algorithm for Best-k-Arm

On an instance of Best-k-Arm with n arms, the PAC-SamplePrune algorithm in [CGL16] is guaranteed to return a ε-optimal answer of Best-k-Arm with probability $1 - \delta$, using

$$O(n\varepsilon^{-2}(\ln \delta^{-1} + \ln k))$$

samples. Here a subset of k arms $T \subseteq \mathcal{I}$ is called ε-optimal, if after adding ε to the mean of each arm in T, T becomes the best k arms in $\mathcal{I}$.

We implement our PAC-Best-k$(S, k, \varepsilon, \delta)$ subroutine as follows. Recall that PAC-Best-k is expected to return a partition $(S^{\mathsf{large}}, S^{\mathsf{small}})$ of the arm set S. If $k \le |S|/2$, we directly run PAC-SamplePrune on the Best-k-Arm instance S and return its output as S^{large}. We let $S^{\mathsf{small}} = S \setminus S^{\mathsf{large}}$. Otherwise, we negate the mean of all arms in S and run PAC-SamplePrune to find the top $|S| - k$ arms in the negated instance.[6] Finally, we return the output of PAC-SamplePrune as S^{small} and let $S^{\mathsf{large}} = S \setminus S^{\mathsf{small}}$. In the following we prove Lemma 4.1.

Proof of Lemma 4.1. By construction, the algorithm PAC-Best-k$(S, k, \varepsilon, \delta)$ takes

$$O(|S|\varepsilon^{-2}[\ln \delta^{-1} + \ln \min(k, |S| - k)])$$

samples. In the following we prove that if $k \le |S|/2$, the set T returned by PAC-SamplePrune is ε-optimal with probability $1 - \delta$. The case $k > |S|/2$ can be proved by an analogous argument.

Let S' denote the instance in which the mean of every arm in T is increased by ε. By definition of ε-optimality,

[6] More precisely, when the algorithm requires a sample from an arm, we draw a sample and return the opposite.

T contains the best k arms in S'. Note that the k-th largest mean is S' is at least $\mu_{[k]}$. Thus for each arm $A \in T$, μ_A must be at least $\mu_{[k]} - \varepsilon$, since otherwise even after μ_A increases by ε, A is still not among the best k arms.

It also holds that every arm in $S \setminus T$ must have a mean smaller than or equal to $\mu_{[k+1]} + \varepsilon$. Suppose for a contradiction that $A \in S \setminus T$ has a mean $\mu_A > \mu_{[k+1]} + \varepsilon$. Then every arm with mean less than or equal to $\mu_{[k+1]}$ in S still have a mean smaller than μ_A in S'. This implies that A is among the best k arms in S', which contradicts our assumption that $A \notin T$. □

D.1.2 PAC algorithms for Best-1-Arm

By symmetry, it suffices to implement the subroutine EstMean-Large and prove its property. In order to estimate the mean of the largest arm in S, we first call PAC-Best-k$(S, 1, \varepsilon/2, \delta/2)$ to find an approximately largest arm. Then we sample the arm $2\varepsilon^{-2}\ln(4/\delta)$ times, and finally return its empirical mean. We prove Lemma 4.2 as follows.

Proof of Lemma 4.2. Let A^* denote the largest arm in S, and let A_0 denote the arm returned by PAC-Best-k$(S, 1, \varepsilon/2, \delta/2)$. According to Lemma 4.1, with probability $1 - \delta/2$, $\mu_{A_0} \in [\mu_{A^*} - \varepsilon/2, \mu_{A^*}]$. It follows that, with probability $1 - \delta/2$,

$$\left|\mu_{A_0} - \max_{A \in S} \mu_A\right| \le \varepsilon/2.$$

Let $\hat{\mu}$ denote the empirical mean of arm A_0. By a Chernoff bound, with probability $1 - \delta/2$,

$$|\hat{\mu} - \mu_{A_0}| \le \varepsilon/2.$$

It follows from a union bound that with probability $1 - \delta$,

$$\left|\hat{\mu} - \max_{A \in S} \mu_A\right| \le \varepsilon.$$

Finally, we note that PAC-Best-k consumes $O(|S|\varepsilon^{-2}\ln\delta^{-1})$ samples as $k = 1$, while sampling A_0 takes $O(\varepsilon^{-2}\ln\delta^{-1})$ samples. This finishes the proof. □

D.1.3 Elimination procedures

We use the Elimination procedure defined in [CL15] as our subroutine Elim-Small$(S, \theta^{\mathsf{small}}, \theta^{\mathsf{large}}, \delta)$. The other building block Elim-Large$(S, \theta^{\mathsf{small}}, \theta^{\mathsf{large}}, \delta)$ can be implemented either using a procedure symmetric to Elimination, or simply by running Elim-Small$(S', -\theta^{\mathsf{large}}, -\theta^{\mathsf{small}}, \delta)$, where S' is obtained from S by negating the arms. In the following, we prove Lemma 4.3.

Proof of Lemma 4.3. Let T denote the set of arms returned by Elim-Small$(S, \theta^{\mathsf{small}}, \theta^{\mathsf{large}}, \delta)$. Lemma B.4 in [CL15] guarantees that with probability $1 - \delta$, the following three properties are satisfied: (1) Elim-Small takes $O(|S|\varepsilon^{-2}\ln\delta^{-1})$ samples, where $\varepsilon = \theta^{\mathsf{large}} - \theta^{\mathsf{small}}$; (2)

$$\left|\{A \in T : \mu_A \le \theta^{\mathsf{small}}\}\right| \le |T|/10;$$

(3) Let A^* be the largest arm in S. If $\mu_{A^*} \ge \theta^{\mathsf{large}}$, then $A^* \in T$.

In fact, the proof of Lemma B.4 does not rely on the fact that A^* is the largest arm in S. Thus property (3) holds for any fixed arm in S. This proves the properties of Elim-Small. The properties of Elim-Large hold due to the symmetry. □

D.2 Observations

D.2.1 Proof of Observation 4.2

Proof of Observation 4.2. Let A denote the arm with the largest mean in S_r^{small}. Recall that μ_r^{small} denote the mean of the $(k_r^{\mathsf{large}} + 1)$-th largest arm in S_r. The correctness of PAC-Best-k and Lemma 4.1 guarantee that $\mu_A \le \mu_r^{\mathsf{small}} + \varepsilon_r/8$. Note that μ_r^{small} is the k_r^{small}-th smallest mean in S_r, while μ_A is the largest mean among the k_r^{small} arms in $S_r^{\mathsf{small}} \subseteq S_r$. So it also holds that $\mu_A \ge \mu_r^{\mathsf{small}}$. Thus we have

$$\mu_A \in [\mu_r^{\mathsf{small}}, \mu_r^{\mathsf{small}} + \varepsilon_r/8].$$

Moreover, as EstMean-Large returns correctly conditioning on $\mathcal{E}_r^{\mathsf{good}}$, by Lemma 4.2 we have

$$\theta_r^{\mathsf{large}} \in [\mu_r^{\mathsf{small}} - \varepsilon_r/8, \mu_r^{\mathsf{small}} + \varepsilon_r/4].$$

The second property follows from a symmetric argument. □

D.2.2 Proof of Observation 4.3

Proof of Observation 4.3. Recall that $\mathcal{E}^{\mathsf{valid}}$ denotes the event that the execution of Bilateral-Elimination is valid. We condition on $\mathcal{E}^{\mathsf{valid}}$ in the following proof. In particular, conditioning on $\mathcal{E}^{\mathsf{valid}}$, $\mathcal{E}_{r-1}^{\mathsf{good}}$ happens and T_{r-1} along with the best k_{r-1}^{large} arms in S_{r-1} constitute the correct answer of the original instance.

Let $\mu_{r-1}^{\mathsf{large}}$ and $\mu_{r-1}^{\mathsf{small}}$ be the k_{r-1}^{large}-th and the $(k_{r-1}^{\mathsf{large}} + 1)$-th largest mean in S_{r-1}. As the arm with mean $\mu_{r-1}^{\mathsf{large}}$ is among the correct answer, we have $\mu_{r-1}^{\mathsf{large}} \ge \mu_{[k]}$, where $\mu_{[k]}$ is the k-th largest mean in the original instance. We also have $\mu_{r-1}^{\mathsf{small}} \le \mu_{[k+1]}$ for the same reason.

Since $\mathcal{E}_{r-1}^{\mathsf{good}}$ happens, by Observation 4.2 we have

$$\theta_{r-1}^{\mathsf{large}} \le \mu_{r-1}^{\mathsf{small}} + \varepsilon_{r-1}/4 \le \mu_{[k+1]} + \varepsilon_{r-1}/4.$$

Then the larger threshold used in Elim-Large is upper bounded by

$$\theta^{\mathsf{large}}_{r-1} + \varepsilon_{r-1}/4 \le \mu_{[k+1]} + \varepsilon_{r-1}/2 = \mu_{[k+1]} + \varepsilon_r.$$

Let T denote the set of arms returned Elim-Large in round $r-1$. We partition T into the following three parts:

$$T^{(1)} = \left\{A \in T : \mu_A > \mu_{[k+1]} + \varepsilon_r\right\},$$

$$T^{(2)} = \left\{A \in T : \mu_{[k]} \le \mu_A \le \mu_{[k+1]} + \varepsilon_r\right\},$$

$$T^{(3)} = \left\{A \in T : \mu_A \le \mu_{[k+1]}\right\}.$$

By Lemma 4.3 and the correctness of Elim conditioning on $\mathcal{E}^{\mathsf{good}}_{r-1}$, we have

$$|T^{(1)}| \le |T|/10.$$

It follows that

$$|T^{(2)}| + |T^{(3)}| \ge 9|T|/10 \ge |T|/2.$$

By definition of arm groups, every arm in $T^{(2)}$ is in $G^{\mathsf{large}}_{\ge r}$. In order to bound $T^{(3)}$, we say that an arm is misclassified into S^{large}_{r-1}, if the arm is not among the best k^{large}_{r-1} arms in S_{r-1}, but is included in S^{large}_{r-1}. We may define misclassification into S^{small}_{r-1} similarly. As $|S^{\mathsf{large}}_{r-1}| = k^{\mathsf{large}}_{r-1}$, the numbers of arms misclassified into both sides are the same.

Since the arms in $T^{(3)}$ are misclassified into S^{large}_{r-1}, there are at least $|T^{(3)}|$ other arms misclassified into S^{small}_{r-1}. Lemma 4.1 (along with the correctness of PAC-Best-k) guarantees that all arms misclassified into S^{small}_{r-1} have means smaller than or equal to $\mu_{[k+1]} + \varepsilon_{r-1}/8$. Thus by definition of arm groups, all these $|T^{(3)}|$ arms are also in $G^{\mathsf{large}}_{\ge r}$. Therefore, we have

$$|G^{\mathsf{large}}_{\ge r}| \ge |T^{(2)}| + |T^{(3)}| \ge |T|/2.$$

Note that $|T| = k^{\mathsf{large}}_r$. Therefore we conclude that $k^{\mathsf{large}}_r \le 2|G^{\mathsf{large}}_{\ge r}|$. The bound on k^{small}_r can be proved using a symmetric argument. □

D.3 Proof of Lemma 4.4

Lemma 4.4 (restated) $\Pr\left[\mathcal{E}^{valid}\right] \ge 1 - \delta$.

Proof of Lemma 4.4. We prove the lemma by upper bounding the probability of $\overline{\mathcal{E}^{\mathsf{valid}}}$, the complement of $\mathcal{E}^{\mathsf{valid}}$.

Split $\overline{\mathcal{E}^{\mathsf{valid}}}$. Let $\mathcal{E}^{\mathsf{bad}}_r$ denote the event that Bilateral-Elimination is valid at round r, yet it becomes invalid at round $r+1$. Then we have

$$\Pr\left[\overline{\mathcal{E}^{\mathsf{valid}}}\right] = \sum_{r=1}^{\infty} \Pr\left[\mathcal{E}^{\mathsf{bad}}_r\right].$$

By definition of validity, event $\mathcal{E}^{\mathsf{bad}}_r$ happens in one of the following two cases:

- Case 1: $\mathcal{E}^{\mathsf{good}}_r$ does not happen.
- Case 2: $\mathcal{E}^{\mathsf{good}}_r$ happens, yet T_{r+1} together with the best k^{large}_{r+1} arms in S_{r+1} is no longer the correct answer.

The probability of Case 1 is upper bounded by $5\delta_r$ according to Observation 4.1. We focus on bounding the probability of Case 2 in the following.

Misclassified arms. Recall that μ^{large}_r and μ^{small}_r denote the means of the k^{large}_r-th and the $(k^{\mathsf{large}}_r + 1)$-th largest arms in S_r respectively. Conditioning on the validity of the execution at round r, the arm with mean μ^{large}_r is among the best k arms in the original instance, while the arm with mean μ^{small}_r is not. Thus we have

$$\mu^{\mathsf{large}}_r \ge \mu_{[k]} > \mu_{[k+1]} \ge \mu^{\mathsf{small}}_r.$$

Define

$$U^{\mathsf{large}}_r = \{A \in S^{\mathsf{large}}_r : \mu_A \le \mu^{\mathsf{small}}_r\}$$

and

$$U^{\mathsf{small}}_r = \{A \in S^{\mathsf{small}}_r : \mu_A \ge \mu^{\mathsf{large}}_r\}.$$

In other words, U^{large}_r and U^{small}_r denote the set of arms "misclassified" by the PAC-Best-k subroutine into S^{large}_r and S^{small}_r in round r.

Bound the number of misclassified arms. Note that since $|U^{\mathsf{large}}_r| \le |S^{\mathsf{large}}_r| = k^{\mathsf{large}}_r$, and in addition, less than k^{small}_r arms in S_r have means smaller than or equal to μ^{small}_r,

$$|U^{\mathsf{large}}_r| \le \min(k^{\mathsf{large}}_r, k^{\mathsf{small}}_r).$$

For the same reason, it holds that

$$|U^{\mathsf{small}}_r| \le \min(k^{\mathsf{large}}_r, k^{\mathsf{small}}_r).$$

With high probability, no misclassified arms are removed. By Observation 4.2, conditioning on $\mathcal{E}^{\mathsf{good}}_r$, we have

$$\theta^{\mathsf{large}}_r \ge \mu^{\mathsf{small}}_r - \varepsilon_r/8.$$

Therefore, when Elim-Large in Line 11 is called at round r, the smaller threshold is at least

$$\theta^{\mathsf{large}}_r + \varepsilon_r/8 \ge \mu^{\mathsf{small}}_r,$$

which is larger than the mean of every arm in U_r^{large}. By Lemma 4.3 and a union bound, with probability

$$1 - |U_r^{\mathsf{large}}|\delta_r' \geq 1 - \min(k_r^{\mathsf{large}}, k_r^{\mathsf{small}})\delta_r' = 1 - \delta_r,$$

no arms in U^{large} are removed by Elim-Large. For the same reason, with probability $1-\delta_r$, no arms in U^{small} are removed by Elim-Small.

Bound the probability of Case 2. Thus, with probability at least $1-2\delta_r$ conditioning on $\mathcal{E}_r^{\mathsf{good}}$, Elim-Large only removes arms with means larger than or equal to μ_r^{large}, and Elim-Small only removes arms with means smaller than or equal to μ_r^{small}. Consequently, every arm in S_r with mean greater than or equal to μ_r^{large} either moves to T_{r+1} or stays in S_{r+1}, which implies that Case 2 does not happen.

Therefore, the Case 2 happens with probability at most $2\delta_r$, and it follows that

$$\Pr\left[\mathcal{E}_r^{\mathsf{bad}}\right] \leq 5\delta_r + 2\delta_r = 7\delta_r.$$

Finally, we have

$$\Pr\left[\overline{\mathcal{E}^{\mathsf{valid}}}\right] \leq \sum_{r=1}^{\infty} 7\delta_r \leq \sum_{r=1}^{\infty} \frac{7\delta}{20r^2} \geq \delta.$$

□

D.4 Missing Calculation in the Proof of Lemma 4.6

Lemma 4.6 (restated) *Conditioning on event $\mathcal{E}^{\mathsf{valid}}$, Bilateral-Elimination takes $O(H\ln\delta^{-1} + \widetilde{H}^{\mathsf{large}} + \widetilde{H}^{\mathsf{small}} + \widetilde{H})$ samples.*

Proof (continued). Recall that

$$H_r^{(1)} = \left(|G_{\geq r}^{\mathsf{large}}| + |G_{\geq r}^{\mathsf{small}}|\right)\varepsilon_r^{-2}(\ln\delta^{-1} + \ln r),$$

$$H_r^{(2,\mathsf{large})} = \varepsilon_r^{-2}|G_{\geq r}^{\mathsf{large}}|\ln|G_{\geq r}^{\mathsf{small}}|,$$

$$H_r^{(2,\mathsf{small})} = \varepsilon_r^{-2}|G_{\geq r}^{\mathsf{small}}|\ln|G_{\geq r}^{\mathsf{large}}|.$$

Our goal is to show that

$$\sum_{r=1}^{\infty} H_r^{(1)} = O\left(H\ln\delta^{-1} + \widetilde{H}\right),$$

$$\sum_{r=1}^{\infty} H_r^{(2,\mathsf{large})} = O\left(\widetilde{H}^{\mathsf{large}}\right), \text{ and}$$

$$\sum_{r=1}^{\infty} H_r^{(2,\mathsf{small})} = O\left(\widetilde{H}^{\mathsf{small}}\right).$$

Upper bound the $H^{(1)}$ term: It follows from a directly calculation that

$$\begin{aligned}\sum_{r=1}^{\infty} H_r^{(1)} &= \sum_{r=1}^{\infty}\sum_{i=r}^{\infty}\left(|G_i^{\mathsf{large}}| + |G_i^{\mathsf{small}}|\right)\varepsilon_r^{-2}(\ln\delta^{-1} + \ln r)\\ &= \sum_{i=1}^{\infty}\left(|G_i^{\mathsf{large}}| + |G_i^{\mathsf{small}}|\right)\sum_{r=1}^{i}\varepsilon_r^{-2}(\ln\delta^{-1} + \ln r)\\ &= O\left(\sum_{i=1}^{\infty}\left(|G_i^{\mathsf{large}}| + |G_i^{\mathsf{small}}|\right)\varepsilon_i^{-2}(\ln\delta^{-1} + \ln i)\right)\\ &= O\left(\sum_{i=1}^{n}\Delta_{[i]}^{-2}\left(\ln\delta^{-1} + \ln\ln\Delta_{[i]}^{-1}\right)\right).\end{aligned}$$

Here the second step interchanges the order of summation. The third step holds since the inner summation is always dominated by the last term. Finally, the last step is due to the fact that $\Delta_A = \Theta(\varepsilon_i)$ for every arm $A \in G_i^{\mathsf{large}} \cup G_i^{\mathsf{small}}$. Therefore we have

$$\sum_{r=1}^{\infty} H_r^{(1)} = O(H\ln\delta^{-1} + \widetilde{H}).$$

Upper bound $H^{(2,\mathsf{large})}$ and $H^{(2,\mathsf{small})}$: By definition of $H_r^{(2,\mathsf{large})}$, we have

$$\begin{aligned}\sum_{r=1}^{\infty} H_r^{(2,\mathsf{large})} &= \sum_{r=1}^{\infty}\sum_{i=r}^{\infty}\varepsilon_r^{-2}|G_i^{\mathsf{large}}|\ln|G_{\geq r}^{\mathsf{small}}|\\ &= \sum_{i=1}^{\infty}|G_i^{\mathsf{large}}|\sum_{r=1}^{i}\varepsilon_r^{-2}\ln|G_{\geq r}^{\mathsf{small}}|.\end{aligned}$$

Therefore we conclude that

$$\sum_{r=1}^{\infty} H_r^{(2,\mathsf{large})} = O(\widetilde{H}^{\mathsf{large}}).$$

The bound on the sum of $H_r^{(2,\mathsf{small})}$ follows from an analogous calculation. □

D.5 Proof of Theorem 1.3

Theorem 1.3 (restated) *For every Best-k-Arm instance, the following statements hold:*

1. $\widetilde{H}^{\mathsf{large}} + \widetilde{H}^{\mathsf{small}} = O\left(\left(H^{\mathsf{large}} + H^{\mathsf{small}}\right)\ln\ln n\right)$.
2. $\widetilde{H}^{\mathsf{large}} + \widetilde{H}^{\mathsf{small}} = O\left(H\ln k\right)$.

Proof of Theorem 1.3. **First Upper Bound.** Recall that

$$H^{\mathsf{large}} = \sum_{i=1}^{\infty}\left|G_i^{\mathsf{large}}\right| \cdot \max_{j\leq i}\varepsilon_j^{-2}\ln\left|G_{\geq j}^{\mathsf{small}}\right|, \text{ and}$$

$$\widetilde{H}^{\mathsf{large}} = \sum_{i=1}^{\infty} \left|G_i^{\mathsf{large}}\right| \sum_{j=1}^{i} \varepsilon_j^{-2} \ln \left|G_{\geq j}^{\mathsf{small}}\right|.$$

For brevity, let N_r denote $\varepsilon_r^{-2} \ln |G_{\geq r}^{\mathsf{small}}| = 4^r \ln |G_{\geq r}^{\mathsf{small}}|$. We fix the value i. Then the i-th term in $\widetilde{H}^{\mathsf{large}}$ reduces to $\left|G_i^{\mathsf{large}}\right| \sum_{r=1}^{i} N_r$. Let $r^* = \mathrm{argmax}_{1 \leq r \leq i} N_r$. Thus the i-th term in H^{large} is simply $\left|G_i^{\mathsf{large}}\right| N_{r^*}$, which is in general smaller than $\left|G_i^{\mathsf{large}}\right| \sum_{r=1}^{i} N_r$. However, we will show that the ratio between the two terms is bounded by $O(\ln \ln n)$.

By definition of r^*, we have $N_{r^*} \geq N_i$. Substituting N_{r^*} and N_i yields

$$4^{r^*} \ln \left|G_{\geq r^*}^{\mathsf{small}}\right| \geq 4^i \ln \left|G_{\geq i}^{\mathsf{small}}\right|.$$

It follows that

$$4^{i - r^*} \ln \left|G_{\geq i}^{\mathsf{small}}\right| \leq \ln \left|G_{\geq r^*}^{\mathsf{small}}\right| \leq \ln n,$$

and thus $i - r^* = O(\ln \ln n)$.

Let $1 \leq r_1 \leq r^*$ be the smallest integer such that $N_{r_1} \geq 2^{r_1 - r^*} N_{r^*}$. By substituting N_{r_1} and N_{r^*}, we obtain

$$4^{r_1} \ln \left|G_{\geq r_1}^{\mathsf{small}}\right| \geq 2^{r_1 - r^*} \cdot 4^{r^*} \ln \left|G_{\geq r^*}^{\mathsf{small}}\right|,$$

which further implies that

$$2^{r^* - r_1} \ln |G_{\geq r^*}^{\mathsf{small}}| \leq \ln |G_{\geq r_1}^{\mathsf{small}}| \leq \ln n$$

and thus $r^* - r_1 = O(\ln \ln n)$.

Therefore we have $i - r_1 = O(\ln \ln n)$, and we can bound the sum of N_r as follows:

$$\begin{aligned}
\sum_{r=1}^{i} N_r &= \sum_{r=1}^{r_1 - 1} N_r + \sum_{r=r_1}^{i} N_r \\
&\leq N_{r^*} \sum_{r=1}^{r_1 - 1} 2^{r - r^*} + (i - r_1 + 1) N_{r^*} \\
&\leq (i - r_1 + 2) N_{r^*} = O(N_{r^*} \ln \ln n).
\end{aligned}$$

Here the second step follows from $N_r < 2^{r - r^*} N_{r^*}$ for $r < r_1$ (by definition of r_1) and $N_r \leq N_{r^*}$ for $r \geq r_1$ (by definition of r^*).

It then follows from a direct summation over all i that

$$\widetilde{H}^{\mathsf{large}} = O(H^{\mathsf{large}} \ln \ln n).$$

The bound on $\widetilde{H}^{\mathsf{small}}$ can be proved similarly.

Second Upper Bound. Note that

$$\begin{aligned}
\widetilde{H}^{\mathsf{large}} &= \sum_{i=1}^{\infty} \left|G_i^{\mathsf{large}}\right| \sum_{j=1}^{i} \varepsilon_j^{-2} \ln \left|G_{\geq j}^{\mathsf{small}}\right| \\
&= \sum_{j=1}^{\infty} \varepsilon_j^{-2} \ln \left|G_{\geq j}^{\mathsf{small}}\right| \sum_{i=j}^{\infty} \left|G_i^{\mathsf{large}}\right| \qquad (11) \\
&= \sum_{i=1}^{\infty} \varepsilon_i^{-2} \left|G_{\geq i}^{\mathsf{large}}\right| \ln \left|G_{\geq i}^{\mathsf{small}}\right|.
\end{aligned}$$

Here the second step interchanges the order of summation. By symmetry we also have

$$\widetilde{H}^{\mathsf{small}} = \sum_{i=1}^{\infty} \varepsilon_i^{-2} \left|G_{\geq i}^{\mathsf{small}}\right| \ln \left|G_{\geq i}^{\mathsf{large}}\right|. \qquad (12)$$

It can be easily verified that for $1 \leq x \leq y$, we have

$$x \ln y + y \ln x \leq (x + y)(2 \ln x + 1). \qquad (13)$$

Note that $\min\left(|G_{\geq i}^{\mathsf{large}}|, |G_{\geq i}^{\mathsf{small}}|\right) \leq k$ for all i. Therefore we can bound $\widetilde{H}^{\mathsf{large}} + \widetilde{H}^{\mathsf{small}}$ as follows:

$$\begin{aligned}
&\widetilde{H}^{\mathsf{large}} + \widetilde{H}^{\mathsf{small}} \\
=& \sum_{i=1}^{\infty} \varepsilon_i^{-2} \left(|G_{\geq i}^{\mathsf{large}}| \ln |G_{\geq i}^{\mathsf{small}}| + |G_{\geq i}^{\mathsf{small}}| \ln |G_{\geq i}^{\mathsf{large}}|\right) \\
=& O\left(\sum_{i=1}^{\infty} \varepsilon_i^{-2} \left(|G_{\geq i}^{\mathsf{large}}| + |G_{\geq i}^{\mathsf{small}}|\right) \ln \min\left(|G_{\geq i}^{\mathsf{large}}|, |G_{\geq i}^{\mathsf{small}}|\right)\right) \\
=& O\left(\sum_{i=1}^{\infty} \varepsilon_i^{-2} \left(|G_{\geq i}^{\mathsf{large}}| + |G_{\geq i}^{\mathsf{small}}|\right) \ln k\right) = O(H \ln k).
\end{aligned}$$

The first step follows from (11) and (12). The second step is due to (13). The third step is due to the observation that $\min\left(|G_{\geq i}^{\mathsf{large}}|, |G_{\geq i}^{\mathsf{small}}|\right) \leq k$. Finally, the last step follows from a simple rearrangement of the summation:

$$\begin{aligned}
&\sum_{i=1}^{\infty} \varepsilon_i^{-2} \left(|G_{\geq i}^{\mathsf{large}}| + |G_{\geq i}^{\mathsf{small}}|\right) \\
=& \sum_{i=1}^{\infty} \varepsilon_i^{-2} \sum_{j=i}^{\infty} \left(|G_j^{\mathsf{large}}| + |G_j^{\mathsf{small}}|\right) \\
=& \sum_{j=1}^{\infty} \left(|G_j^{\mathsf{large}}| + |G_j^{\mathsf{small}}|\right) \sum_{i=1}^{j} \varepsilon_i^{-2} \\
=& O\left(\sum_{j=1}^{\infty} \varepsilon_j^{-2} \left(|G_j^{\mathsf{large}}| + |G_j^{\mathsf{small}}|\right)\right) = O(H).
\end{aligned}$$

□