HAVER: Instance-Dependent Error Bounds for Maximum Mean Estimation and Applications to Q-Learning and Monte Carlo Tree Search

Tuan Ngo Nguyen
University of Arizona
tnguyen9210@arizona.edu

Jay Barrett
Cornell University
jab864@cornell.edu

Kwang-Sung Jun
University of Arizona
kjun@cs.arizona.edu

Abstract

We study the problem of estimating the *value* of the largest mean among K distributions via samples from them (rather than estimating *which* distribution has the largest mean), which arises from various machine learning tasks including Q-learning and Monte Carlo Tree Search (MCTS). While there have been a few proposed algorithms, their performance analyses have been limited to their biases rather than a precise error metric. In this paper, we propose a novel algorithm called HAVER (Head AVERaging) and analyze its mean squared error. Our analysis reveals that HAVER has a compelling performance in two respects. First, HAVER estimates the maximum mean as well as the oracle who knows the identity of the best distribution and reports its sample mean. Second, perhaps surprisingly, HAVER exhibits even better rates than this oracle when there are many distributions near the best one. Both of these improvements are the first of their kind in the literature, and we also prove that the naive algorithm that reports the largest empirical mean does not achieve these bounds. Finally, we confirm our theoretical findings via numerical experiments where we implement HAVER in bandit, Q-learning, and MCTS algorithms. In these experiments, HAVER consistently outperforms the baseline methods, demonstrating its effectiveness across different applications.

Proceedings of the 28th International Conference on Artificial Intelligence and Statistics (AISTATS) 2025, Mai Khao, Thailand. PMLR: Volume 258.

1 Introduction

We consider the problem of estimating the maximum mean value among K distributions based on samples from them, which we call the *maximum mean estimation* (MME) problem. Specifically, for each $i \in [K] := \{1, \ldots, K\}$, we are passively given samples $X_{i,1}, \ldots, X_{i,N_i}$ from the i-th distribution ν_i. The learner must estimate the largest mean $\max_{i\in[K]} \mathbb{E}_{X\sim\nu_i}[X]$ as accurately as possible. The MME problem arises from various machine learning tasks. In Q-learning at each time step, an agent updates its state-action value estimates $\hat{Q}(s, a)$ based on the observed reward and the estimated value of the next state s'. The latter requires an accurate estimate of the largest state-action value $\max_a Q^*(s', a)$ where Q^* is the true state-action value. An inaccurate estimator for the expectation can adversely impact the learning process. Similarly, in Monte Carlo tree search (Coulom, 2006; Kocsis and Szepesvári, 2006), one faces to estimate the value of each node at a non-leaf node, which is the maximum value of its children node. Accurate estimation of the largest value is paramount to quickly filter out unpromising nodes, which allows budget-efficient identification of the best action.

Arguably, the simplest estimator is to take the largest empirical mean (LEM). However, LEM has a positive bias, which can be detrimental to its accuracy when the number of samples is small or the number of distributions K is large. To overcome such an issue, there have been a number of studies on the problem of estimating the maximum mean value. The earliest work we are aware of is van Hasselt (2010, 2013), where the authors propose a double estimator (DE) based on sample splitting that is guaranteed to be negatively biased. In another work of Lan et al. (2020), the authors propose the maxmin estimator that also uses a sample splitting but chooses the largest of the smallest estimator from multiple sampling buckets. Another idea is to consider a weighted estimator (WE) that computes a weighted average over the empirical means (Tesauro et al., 2010; D'Eramo et al., 2016). While these studies report some success in downstream applications, their theoretical justification is limited to characterizing the variance or the direction of the bias rather than precise error metrics such as mean squared error (MSE). Furthermore, to our knowledge, no studies have ever shown that their proposed methods enjoy orderwise better MSE compared to LEM. We discuss more on related work in Section 6.

In this paper, we focus on the MSE as the error metric of interest. What are the good rates of MSE in the MME problem? We identify two desiderata centered around a reference estimator that we call *the oracle* who knows the identity of the true best arm and thus reports its sample mean:

D1: Can we achieve a worst-case MSE bound that is as good as the oracle?

D2: Can we achieve an instance-dependent MSE bound that can be strictly better than the oracle?

As a sanity check, we first show that LEM fails to satisfy either of these criteria. This means that the two criteria above are not trivial to achieve.

As our main contribution, we propose a novel estimator called HAVER (Head AVERaging). Our analysis shows that HAVER achieves not only **D1** but also, perhaps surprisingly, **D2**. In particular, we derive a generic instance-dependent upper bound on the MSE of HAVER. This generic bound can further be upper-bound by an instance-independent quantity to satisfy **D1**. For **D2**, we evaluate the generic upper bound of HAVER under several practical instances including instances whose suboptimality gaps or sample counts vary with specific rates. Finally, we conduct empirical studies across various settings, including bandits and Q-learning environments, to show that HAVER consistently outperforms prior methods.

Our theoretical analysis provides a convenient framework for analyzing MSE of maximum mean estimators by leveraging the equality of expectation with the tail bound of nonnegative random variables; i.e.,

$$\mathbb{E}[(X-\mathbb{E}[X])^2] = \int_{\varepsilon=0}^{\infty} \mathbb{P}((X-\mathbb{E}[X])^2 > \varepsilon)\, \mathrm{d}\varepsilon \quad (1)$$

which allows us to leverage tight concentration inequalities along with careful event decomposition. We believe our analysis framework can be of independent interest to the machine learning research community.

2 Problem Definition and Preliminaries

Notations. Throughout the paper, c_1 and c_2 are a universal and positive constant, which may have different values for different expressions. Both $\tilde{\Theta}(\cdot)$, $\tilde{\mathcal{O}}(\cdot)$, and "$\lesssim$" omits logarithmic factors on K and N, but not $1/\delta$. We denote $A_i, A_{i+1}, \ldots, A_j$ by $A_{i:j}$.

The maximum mean estimation (MME) problem. In the MME problem, we are given K distributions, which we call *arms*, hereafter, following the vocabulary of multi-armed bandits (Thompson, 1933; Lattimore and Szepesvári, 2020). We denote by ν_i the distribution of arm i and define $\mu_i = \mathbb{E}_{X\sim\nu_i}[X]$. The learner is given N_i observations denoted by $X_{i,1}, \ldots, X_{i,N_i}$ that follow ν_i for every arm $i \in [K]$. We define $D := \{X_{1,1:N_1}, \ldots, X_{K,1:N_K}\}$ as the aggregate set of all samples from all the arms. Note that this problem is a passive sampling setting in that N_i is arbitrarily given from the environment rather than being chosen by the learner. The goal of the learner is to estimate the largest mean among the K arms, i.e., $\max_{i\in[K]} \mu_i$, as accurately as possible. Without loss of generality, we assume that the arms are ordered in decreasing order of their mean, i.e., $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_K$. Thus, the maximum mean is μ_1. This assumption is for notational convenience only, and the learner is not aware of this ordering. We define $\Delta_i = \mu_1 - \mu_i$ to be the suboptimality gap of arm $i \in [K]$.

Let $\hat{\mu}$ be the estimator computed by a maximum mean estimation algorithm. As a performance measure, we choose the mean squared error (MSE) metric:

$$\mathrm{MSE}(\hat{\mu}) = \mathbb{E}\left[(\hat{\mu}-\mu_1)^2\right] ,$$

which is a standard measure of error in statistics.

Definition 1 (Sub-Gaussian distribution)**.** If X is σ-sub-Gaussian, then for any $\varepsilon \geq 0$,

$$\mathbb{P}(X \geq \varepsilon) \leq \exp\left(-\frac{\varepsilon^2}{2\sigma^2}\right)$$

We assume that the distributions are sub-Gaussian as follows.

Assumption 2 (Sub-Gaussian distributions)**.** For each arm $i \in [K]$, the distribution ν_i is 1-sub-Gaussian, and this sub-Gaussian parameter is known to the learner.

Note that the choice of 1 for the sub-Gaussian parameter is for brevity only, and all our results can be extended to the generic σ^2-sub-Gaussian assumption.

We also make the following standard assumption in the maximum mean estimation literature (van Hasselt, 2010, 2013; D'Eramo et al., 2016; Lan et al., 2020).

Assumption 3 (i.i.d samples)**.** For each arm $i \in [K]$, each sample $X_{i,j}$ is drawn i.i.d from the distribution ν_i, $\forall j \in [N_i]$.

Largest Empirical Mean (LEM) estimator. The most immediate and intuitive estimator for the MME problem is the largest empirical mean (LEM) estimator that returns

$$\hat{\mu}^{\text{LEM}} := \max_{i\in[K]} \hat{\mu}_i$$

where $\hat{\mu}_i := \frac{1}{N_i}\sum_{j=1}^{N_i} X_{i,j}$ is the empirical mean.

We analyze the MSE of LEM in the following theorem.

Theorem 4. *LEM achieves*

$$\text{MSE}(\hat{\mu}^{LEM}) = \mathcal{O}\left(\frac{\log(2K)}{\min_i N_i}\right) .$$

Our bound is improved upon the prior work of van Hasselt (2013) in the following sense. van Hasselt (2013) provides separate upper bounds for bias and variance as follows:

$$\text{Bias}^2(\hat{\mu}^{\text{LEM}}) \le \frac{K-1}{K}\sum_{i=1}^{K}\frac{1}{N_i}, \;\; \text{Var}(\hat{\mu}^{\text{LEM}}) \le \sum_{i=1}^{K}\frac{1}{N_i}.$$

One can derive an MSE bound from this since MSE is the sum of the squared bias and the variance, which becomes

$$\text{MSE}(\hat{\mu}^{\text{LEM}}) = \mathcal{O}\left(\sum_{i=1}^{K}\frac{1}{N_i}\right) .$$

When all sample sizes are equal, i.e., $N_i = N, \forall i \in [K]$, our bound is a factor of $K/\log(2K)$ tighter.

It is natural to ask what properties an ideal estimator can possess and how LEM measures up to these standards. First, consider an oracle that has prior knowledge of the optimal arm, which has the maximum mean μ_1. The oracle would base its estimation only on samples drawn from this arm. Consequently, the MSE of the oracle would be $\Theta\left(\frac{1}{N_1}\right)$ which we refer to as the **oracle rate**. Compared to this oracle rate, LEM does not perform as reliably. We show that, in a two-arms instance where both arms follow Gaussian distributions with the same mean, the MSE of LEM has a lower bound proportional to $\frac{1}{\min_{i\in[2]} N_i}$, which is worse than the oracle rate. This implies that LEM cannot always obtain the oracle rate in general. In certain applications, the number of samples drawn from suboptimal arms can be significantly smaller compared to those from the optimal arm. As a result, using LEM in such scenarios can lead to arbitrarily large errors. For instance, in Q-learning, the agent gradually learns to favor the optimal action, leading to disproportionally few samples from the suboptimal actions. This slows down the internal maximum mean estimation and, consequently, the learning process.

Second, consider an extreme problem instance where all arms share the same mean, i.e., $\mu_1 = \cdots = \mu_K$, and the number of samples across all arms are identical $\forall i$, $N_i = N$. In this instance, LEM achieves an MSE bound that is proportional to $\frac{1}{N_1}$, which matches the oracle rate. However, can we do even better? Average estimator (AE) is another estimator that is used in the standard Monte Carlo tree search (MCTS) algorithm called UCT (Kocsis and Szepesvári, 2006):

$$\hat{\mu}^{\text{AE}} := \frac{1}{K}\sum_{i=1}^{K}\sum_{j=1}^{N_i} X_{i,j} .$$

In this extreme case, the AE achieves an MSE bound that is proportional to $\frac{1}{KN}$, which is a factor of K improvement over the oracle rate.

In summary, a good estimator should (i) perform as well as the oracle rate, and (ii) achieve acceleration over the oracle rate in special instances. This raises the following question: does there exist an estimator that satisfies both properties? In the next section, we answer this question in the affirmative by proposing a novel estimator called HAVER.

3 Warmup: The Oracle Rate via Maximum Lower Confidence Bound

We begin by investigating whether the maximum lower bound confidence bound estimator (MLCB) can achieve the desired properties, which is known as the *pessimistic principle* and has been useful for estimating *which arm* has the highest mean in various settings including off-policy reinforcement learning (Jin et al., 2021) which is different from estimating the value of the best arm.

For every arm i, we define its empirical mean as $\hat{\mu}_i := \frac{1}{N_i}\sum_{j=1}^{N_i} X_{i,j}$ and its confidence width as $\gamma_i := \sqrt{\frac{16}{N_i}\log\left(\left(\frac{KT}{N_i}\right)^2\right)}$, where $T = \sum_{j\in[K]} N_j$. We propose the MLCB algorithm, which chooses an arm with the highest lower confidence bound, defined as

$$\hat{r} := \arg\max_{i\in[K]} \hat{\mu}_i - \gamma_i$$

and outputs

$$\hat{\mu}^{\text{MLCB}} := \hat{\mu}_{\hat{r}} .$$

In the following theorem, we show that MLCB achieves the oracle rate up to logarithmic factors.

Theorem 5. *MLCB achieves*

$$\text{MSE}(\hat{\mu}^{MLCB}) = \tilde{\mathcal{O}}\left(\frac{1}{N_1}\right)$$

Algorithm 1 HAVER

Input: A set of K arms with samples $\left\{X_{i,1:N_i}\right\}_{i=1}^{K}$.

For each arm $i \in [K]$, compute its empirical mean: $\hat{\mu}_i = \frac{1}{N_i}\sum_{j=1}^{N_i} X_{i,j}$.

Find the pivot arm $\hat{r}$:
$\hat{r} = \arg\max_{i\in[K]} \hat{\mu}_i - \gamma_i$
where $\gamma_i = \sqrt{\frac{18}{N_i}\log\left(\left(\frac{KS}{N_i}\right)^4\right)}$,
and $S = N_{\max}\sum_{j\in[K]} N_j$

Form a candidate set $\mathcal{B}$:
$\mathcal{B} = \left\{i \in [K] : \hat{\mu}_i \geq \hat{\mu}_{\hat{r}} - \gamma_{\hat{r}},\, \gamma_i \leq \frac{3}{2}\gamma_{\hat{r}}\right\}$

Compute the weighted average of the empirical means of candidate arms in set $\mathcal{B}$:
$\hat{\mu}^{\mathcal{B}} = \frac{1}{Z}\sum_{i\in\mathcal{B}} N_i\hat{\mu}_i$ where $Z = \sum_{i\in\mathcal{B}} N_i$

Output: $\hat{\mu}^{\mathcal{B}}$

We remark that, though the algorithmic principle was inspired by the well-known pessimism, adapting it to our problem of estimating the *value* of the best arm with respect to the MSE metric is not straightforward since the existing work using pessimism cares about a high probability guarantee with a given specified confidence level δ, so setting γ_i as $\sqrt{\frac{2}{N_i}\log(K/\delta)}$ is enough. This is a matter of considering the worst-case behavior of the estimator under high probability events, and invoking concentration bounds with the specified level is enough. In contrast, analyzing MSE requires studying the tail of the estimator at all the confidence levels (see (1)), which requires carefully setting the right confidence width γ_i to account for the event where the confidence bound fails to hold.

4 Better than the Oracle: HAVER (Head AVERaging)

In this section, we propose a novel estimator called HAVER (Head AVERaging). The key idea is the following observation: for many problem instances of interest, there exists a set of good arms whose means are *very close* to the maximum mean, so it might be a good idea to average out a few empirical top arms, which may reduce the variance at the price of introducing some bias.

At the same time, we still want to be as good as the oracle, so we should be leveraging some form of pessimism that MLCB employ.

With this intuition in mind, we now describe HAVER in detail whose full pseudocode can be found in Algorithm 1.

For each arm $i \in [K]$, we define $\gamma_i = \sqrt{\frac{18}{N_i}\log\left(\left(\frac{KS}{N_i}\right)^4\right)}$ as its confidence width where $S = N_{\max}\sum_{j\in[K]} N_j$ and $N_{\max} = \max_{j\in[K]} N_j$. HAVER consists of three key steps. First, we select a pivot arm with maximum lower confidence bound $\hat{r} := \arg\max_{i\in[K]} \hat{\mu}_i - \gamma_i$. Next, we form a candidate set of arms, $\mathcal{B} := \left\{i \in [K] : \hat{\mu}_i \geq \hat{\mu}_{\hat{r}} - \gamma_{\hat{r}},\, \gamma_i \leq \frac{3}{2}\gamma_{\hat{r}}\right\}$, whose empirical means exceed the maximum lower confidence bound threshold and whose sample sizes are not significantly smaller than that of the pivot arm. Note that the intuition of $\mathcal{B}$ is to include arms that are statistically not distinguishable from the pivot arm. However, it is essential to include the constraint $\gamma_i \leq \frac{3}{2}\gamma_{\hat{r}}$, which forces us to exclude arms with much lower sample size than the pivot arm to make sure we do not inadvertently include bad arms that are being overrepresented due to low sample sizes (i.e., large variances). Finally, HAVER computes the average of the empirical means of the arms in the candidate set, weighted by the normalized sample size for each arm.

The following definitions are required for our main theorem presented next. We define $s := \arg\max_{i\in[K]} \mu_i - \gamma_i$ as the ground truth pivot arm. We define

$$\mathcal{R} := \left\{r \in [K] \colon \mu_s - \frac{4}{3}\gamma_s + \frac{2}{3}\gamma_r \leq \mu_r \leq \mu_s - \frac{1}{6}\gamma_s + \frac{1}{6}\gamma_r\right\}$$

as the statistically-plausible candidate set for the pivot arm $\hat{r}$. For any arm $r \in \mathcal{R}$, we define

$$\mathcal{B}^*(r) := \left\{i \in [K] : \mu_i \geq \mu_s - \frac{1}{6}\gamma_s,\, \gamma_i \leq \frac{3}{2}\gamma_r\right\}$$

as the set of *good* arms whose means are *very close* to the maximum mean and whose sample sizes are *within a constant factor* from the ground truth pivot arm's sample size. Also, for any arm $r \in \mathcal{R}$, we define $\mathcal{B}^+(r) := \left\{i : \mu_i \geq \mu_s - \frac{4}{3}\gamma_s - \frac{4}{3}\gamma_i,\, \gamma_i \leq \frac{3}{2}\gamma_r\right\}$ as the set of *nearly good* arms whose the condition is relaxed from $\mathcal{B}^*(r)$. On the event of r being the pivot arm, $\mathcal{B}^*(r)$ represents the set of arms that should belong to the set $\mathcal{B}$ with overwhelming probability, and $[K]\backslash\mathcal{B}^+(r)$ represents the set of arms that should not belong to the set $\mathcal{B}$ with overwhelming probability.

Theorem 6. *HAVER achieves*

$$\begin{aligned}
&\mathrm{MSE}(\hat{\mu}^{HAVER})\\
&= \tilde{\mathcal{O}}\left(\left(\max_{r\in\mathcal{R}} \frac{1}{\sum_{j\in\mathcal{B}^*(r)} N_j}\sum_{i\in\mathcal{B}^+(r)} N_i\Delta_i\right)^2 \wedge \frac{1}{N_1}\right)\\
&+ \tilde{\mathcal{O}}\left(\max_{r\in\mathcal{R}}\max_{k=0}^{d(r)} \max_{\substack{\mathcal{S}:\mathcal{B}^*(r)\subseteq\mathcal{S}\subseteq\mathcal{B}^+(r)\\ |\mathcal{S}|=n_*(r)+k}} \frac{k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}{\left(\sum_{j\in\mathcal{S}} N_j\right)^2} \wedge \frac{1}{N_1}\right)\\
&+ \tilde{\mathcal{O}}\left(\frac{1}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j} \wedge \frac{1}{N_1}\right) + \tilde{\mathcal{O}}\left(\frac{1}{KN_1}\right).
\end{aligned}$$

where $d(r) = |\mathcal{B}^+(r)| - |\mathcal{B}^*(r)|$ *and* $n^*(r) = |\mathcal{B}^*(r)|$.

The theorem above provides a fine-grained, instance-dependent upper bound on the MSE of HAVER, though with a rather complicated form. While we leave detailed explanation on the bound above to the end of this section, one can still observe that the bound above is of order between $\frac{1}{N_1}$ and $\frac{1}{KN_1}$. This demonstrates that HAVER meets the first criterion D1: it performs at least as well as the oracle rate. The primary reason HAVER obtains the oracle rate, $\frac{1}{N_1}$, is due to its use of the pessimism principle when choosing the pivot arm based on the maximum lower confidence bound in a similar way to MLCB estimator. However, due to a careful averaging over the set $\mathcal{B}$, HAVER is able to achieve instance-dependent accelerated rates in various problems as we show below.

The case of equal number of samples. In the next three corollaries, we focus on instances where the sample sizes are equal. This setup allows us to better understand the influence of the arm means to accelerated rates.

Assumption 7 (Equal number of samples)**.** The number of samples across all arms is the same, $\forall i \in [K]$, $N_i = N$ for some $N > 0$.

Corollary 8. Under Assumption 7 (equal sample sizes), let $\gamma := \sqrt{\frac{18}{N}\log\left((K^2N)^4\right)}$, $\mathcal{B}^* := \left\{i \in [K] : \Delta_i \le \frac{1}{6}\gamma\right\}$, and $\mathcal{B}^+ := \left\{i \in [K] : \Delta_i \le \frac{8}{3}\gamma\right\}$. Then,

$$\begin{aligned}\mathrm{MSE}(\hat{\mu}^{\mathrm{HAVER}}) = \tilde{\mathcal{O}}&\left(\left(\frac{1}{|\mathcal{B}^*|}\sum_{i=1}^{|\mathcal{B}^+|}\Delta_i\right)^2 \wedge \frac{1}{N}\right)\\ &+ \tilde{\mathcal{O}}\left(\frac{1}{N}\left(\log\left(\frac{|\mathcal{B}^+|}{|\mathcal{B}^*|}\right)\right)^2 \wedge \frac{1}{N}\right)\\ &+ \tilde{\mathcal{O}}\left(\frac{1}{|\mathcal{B}^*|\,N}\right) + \tilde{\mathcal{O}}\left(\frac{1}{KN}\right).\end{aligned}$$

The bound above is further simplified from Theorem 6, and exhibits an interesting behavior as a function of $|\mathcal{B}^*|$ and $|\mathcal{B}^+|$. Assuming that cardinality of $\mathcal{B}^*$ fixed, HAVER is particularly beneficial for problem instances where the gap between $\mathcal{B}^*$ and $\mathcal{B}^+$ is small, as it leads to a simultaneous decrease in the first two terms. For many cases, the log ratio in the second term is less than 1 and contribute to the acceleration. Indeed, we next present a few such instances where HAVER achieves strictly accelerated rates compared to the oracle.

Corollary 9. Under Assumption 7 (equal number of samples), consider the K^*-best instance where $\forall i \in [K^*]$, $\mu_i = \mu_1$. If $N > \frac{256}{\Delta^2_{K^*+1}}\log\left(\frac{256K^2}{\Delta^2_{K^*+1}e}\right)$, HAVER achieves

$$\mathrm{MSE}(\hat{\mu}^{\mathrm{HAVER}}) = \tilde{\mathcal{O}}\left(\frac{1}{K^*N}\right).$$

The proof is in Appendix B.1. When the number of samples grows and exceeds the inverse gap of the K^*+1 arm, HAVER becomes more effective at rejecting the bad arms. This enables HAVER to achieve an acceleration by a factor of K^*.

Corollary 10. Under Assumption 7 (equal number of samples), consider the Poly(α) instance where $\forall i \ge 2, \Delta_i = \left(\frac{i}{K}\right)^\alpha$ where $\alpha \ge 1$. If $N \le \frac{1}{2}\log(K^2N)\left(\frac{K}{\alpha}\log(2)\right)^{2\alpha}$, HAVER achieves

$$\mathrm{MSE}(\hat{\mu}^{\mathrm{HAVER}}) = \tilde{\mathcal{O}}\left(\frac{1}{(\alpha \wedge K)N}\right).$$

The proof is in Appendix B.2. In this problem instance, HAVER achieves a factor α acceleration. This result is intuitive: as α increases, there are more good arms, allowing HAVER to capture many more good arms and thus achieve greater acceleration. However, when N becomes too large, acceleration diminishes. This is because the empirical means of all arms concentrate more around their means, causing HAVER to only include the best arm in $\mathcal{B}$, thereby loosing the potential for acceleration. In the next corollary, we assume all arm means are identical and shift our focus to the impact of varying sample sizes.

The case of all arms having equal means. We now turn to a case where all arms have equal means to focus on the effect of the sample counts on the accelerated rate.

Corollary 11. Consider the all-best instance where $\forall i \in [K]$, $\mu_i = \mu_1$ and the number of samples are characterized by $N_i = (K-i+1)^\beta$. In the regime of $\beta \in (0,1)$, HAVER achieves

$$\mathrm{MSE}(\hat{\mu}^{\mathrm{HAVER}}) = \tilde{\mathcal{O}}\left(\frac{1}{KN_1}\right).$$

The proof is in Appendix B.3. In this instance, the number of samples are characterized by $N_i = (K-i+1)^\beta$. For $\beta \in (0,1)$, the number of samples forms a smoothly decreasing polynomial curve. In this regime, HAVER achieves a factor of K-fold acceleration. This corollary provides a different perspective on the problem, showing that HAVER can achieve acceleration in both changing arm means and their sample sizes.

Detailed interpretation of the bound presented in Theorem 6. Through these corollaries, we have demonstrated that HAVER can indeed achieve acceleration over the oracle in certain instances. Therefore, we can conclude that HAVER satisfies both of the desired criteria.

Next, we provide a more detailed into HAVER's theorem and offer some intuition on where this acceleration comes from. HAVER can achieve acceleration by forming a candidate set of good arms, $\mathcal{B}$, and taking weighted average of their empirical means. To delve deeper, we need to understand the intuition of $\mathcal{R}$, $\mathcal{B}^*(r)$, and $\mathcal{B}^+(r)$. Since the algorithm revolves around the pivot arm, it is natural to define a set $\mathcal{R}$ that includes all the arms that can be pivot with non-trivial probability. For any pivot arm r, we define the set $\mathcal{B}^*(r)$, referred to as the set of *good* arms, based on conditions that ensure it closely aligns with the candidate $\mathcal{B}$, by matching both the mean deviation and sample size constraints. However, noticing that $\mathcal{B}$ is highly likely to include nearly-good arms that are marginally close to the good arms, we introduce the set of *nearly-good* arms $\mathcal{B}^+(\hat{r})$, to further refine our analysis.

Since the set $\mathcal{B}$ is a random set, its composition varies depending on the observed samples. To analyze, we first identify a set of statistically-plausible set of pivot arms, which is $\mathcal{R}$ we define above. Assuming $\hat{r} \in \mathcal{R}$, which is a likely event, we identify four events based on the set $\mathcal{B}$. The first event, $G_0 = \{\mathcal{B} = \mathcal{B}^*(\hat{r})\}$, represents the ideal scenario where the candidate set $\mathcal{B}$ exactly aligns with $\mathcal{B}^*(\hat{r})$. This event contributes to the third term in the theorem. In this ideal event, HAVER's MSE is in order of $\frac{1}{\sum_{j \in \mathcal{B}^*(r)} N_j}$. Depending on the instances, HAVER will achieve acceleration as shown in Corollaries 9, 10, and 11. The second event, $G_1 = \{\mathcal{B}^*(\hat{r}) \subset \mathcal{B} \subseteq \mathcal{B}^+(\hat{r})\}$, is also highly likely and captures scenarios where $\mathcal{B}$ contains the good arms but may include additional nearly-good arms. This event contributes to the first and second terms in the theorem. The next two events, $G_2 = \{\exists i \in \mathcal{B}^*(\hat{r}),\, i \notin \mathcal{B}\}$ and $G_3 = \{\mathcal{B}^*(\hat{r}) \subseteq \mathcal{B},\, \exists i \notin \mathcal{B}^+(\hat{r}),\, i \in \mathcal{B}\}$, are highly unlikely to occur. The event G_2 captures the scenario where $\mathcal{B}$ misses an arm from $\mathcal{B}^*(\hat{r})$, while the event G_3 corresponds to $\mathcal{B}$ including supoptimal arm that is not part of $\mathcal{B}^+(\hat{r})$. These unlikely events contribute to the potential K-fold acceleration.

5 Experiment

In this section, we present a comprehensive empirical analysis to evaluate and compare the performance of HAVER, LEM, DE, and WE across three distinct scenarios: (1) the multi-armed bandit setting, (2) Q-learning applied to the Grid World problem, and (3) Monte Carlo Tree Search (MCTS) applied to the FrozenLake environment in OpenAI Gym. Our first experiment setting, in the multi-armed bandits (MAB) domain, is a controlled environment designed to examine the scaling effect as we increase the number of samples and number of distributions. Within this setting, we consider three problem instances K^*-best and Poly(α) and uniform distributed means. The first two instances align with our Corollary 8 and 9, and the last instance can be considered an almost general case where we do not assume any structure based on the distributions' means. Our second and third experimental settings, focused on Q-learning and Monte Carlo Tree Search (MCTS), are designed as a fully general environment. In these settings, we do not impose any implicit structural assumptions on the problem. In particular, when the agent is at a particular state, the reward distributions for the available actions are not required to follow any specific structure.

We notice several potential challenges that arise in practice and suggest solutions to address them. First, the variance of an arm's distribution is often unknown to the user. To mitigate this issue, we use an unbiased estimate of the variance, i.e., $\hat{\sigma}_i^2 = \frac{\sum_{j=1}^{N_i}(X_{i,j} - \hat{\mu}_i)^2}{N_i - 1}$. Second, since we are working with estimates of variances, we should incorporate these estimates into the average weights in the averaging step. For each arm $i \in [K]$, its average weight becomes $\frac{N_i}{\hat{\sigma}_i^2}$. However, $\hat{\sigma}_i^2$ can be arbitrarily small, particularly in early episodes in Q-learning, which causes the denominator to explode. To prevent this, we add a small tunable hyperparameter ε to $\hat{\sigma}_i^2$. This parameter is set to 0.01 in our experiments. The detailed pseudocode is available in Appendix A.1.

5.1 Multi-Armed Bandits

The multi-armed bandit environment can serve as a controlled environment for our experiments. It is important to note that the standard multi-armed bandit problem, particularly best-arm identification with a fixed budget, differs from the maximum mean estimation problem. In the former, a learner adaptively selects arms to sample and, after exhausting the budget, identifies the optimal arm. In contrast, in the maximum mean estimation problem, the samples are fixed and provided by the environment. Our focus is solely on estimating the value of the largest mean among K distributions.

We follow a similar experiment setup as formulated in van Hasselt (2013). Consider K ads, each with a click rate μ_i. For simplicity, we assume the return for each ad follows the Bernoulli distribution (i.e., each click has a reward of one), so the ad's mean return is equal to its click rate μ_i. Each ad is equally shown to customers N times, and the return is recorded each time. We compare the results of HAVER, LEM, DE, and WE in two different settings for three problem instances. In the default configuration, we set the number of ads $K = 50$ and the number of samples per ad is $N = 500$, and mean click rates are from the interval $[0.002, 0.005]$. In the first setting, we vary the number of samples per ad $N = \{100, 200, \ldots, 1000\}$, which is equivalent to the total number of samples ranging from 5000 to 50000. In the second setting, we vary the number of ads $K = \{30, 40, \ldots, 100\}$.

We consider two problem instances (K^*-best and Poly(α)) mentioned in our previous corollaries, as well as a generic uniformly-sampled instance. For the K^*-best instance, the arms are divided into two halves: the arms in the first half have mean click rates of 0.005, while the remaining arms have mean click rates of 0.002. For the Poly(α), we set $\alpha = 2$ and normalize the mean click rates to the interval $[0.002, 0.005]$. Lastly, for the generic instance, the mean click rates are uniformly sampled from the interval $[0.002, 0.005]$. For each problem instance, we repeat the two experiment settings accordingly. We present the results in terms of $\text{MSE} = \text{Bias}^2 + \text{Var}$ for all these experiments, averaging the results over 1000 trials.

We observe similar patterns over the three problem instances. We present the results for the generic instance in Figure 1 while the results for the other two instances are included in the appendix due to space constraints. The results of the first and second experimental settings are shown at the top and bottom, respectively. Overall, HAVER's MSE effectively decays with both the number of samples and the number of ads, which aligns with our theorem and corollaries. In the first setting, unsurprisingly, the MSE decreases for all estimators as the number of samples per ad increases, with HAVER consistently achieving the lowest MSE in all cases. Additionally, HAVER has the lowest squared bias among the estimators, suggesting its effectiveness in capturing the good arms. In the second setting, the MSE of both HAVER and DE decreases as the number of ads increases, whereas LEM's performance deteriorates with a larger number of ads. Interestingly, the MSE of WE does not decay with an increasing number of ads.

5.2 Q-Learning Applied to the Simple Grid World

The maximum mean estimation problem is often discussed in the context of Q-learning (Watkins, 1989). The update of Q-learning is

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha_t \left(r_t + \eta \max_a Q(s_{t+1}, a) - Q(s_t, a_t)\right).$$

where $Q(s, a)$ represents the value of action a in state s at time t, the reward r_t is drawn from a fixed distribution as a function of s_t and a_t, and η is a discount factor. Here, given the history of visits and action selections in the state s_{t+1}, we want to estimate the maximum value of $Q(s_{t+1}, a)$ among the actions. Errors propagation problems resulting from inaccurate estimation can have a negative effect on the learning process.

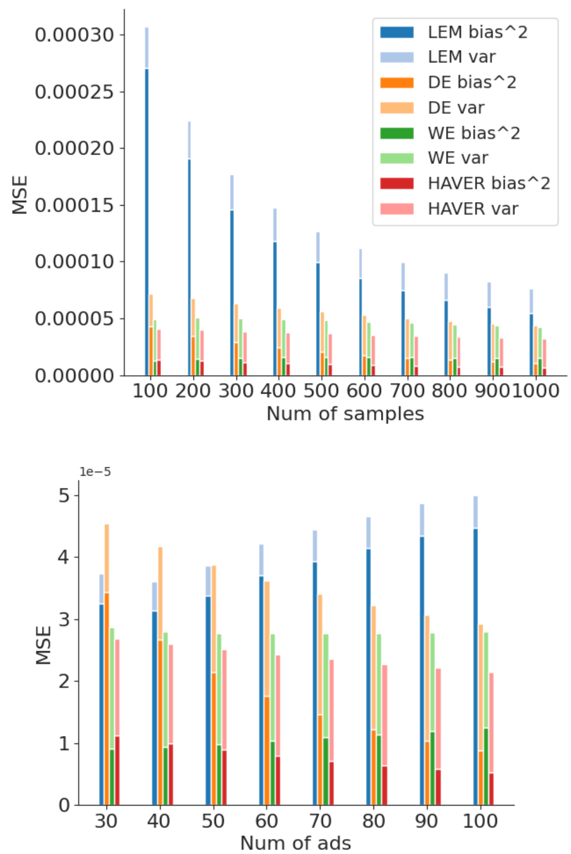

Figure 1: Uniform sampling instance. The results are averaged over 1000 trials.

In this experiment, we use a 3x3 grid world environment setup similar to that in van Hasselt (2010). The agent starts in the lower-left cell, and the terminal cell is located in upper-right corner. The agent's goal to maximize its cumulative reward before reaching the terminal cell. There are two kinds of rewards. At each step, the agent performs an action and gains a random reward drawn from the standard Gaussian distribution $\mathcal{N}(\mu = 0, \sigma^2 = 1)$. Upon reaching the terminal cell, the agent gains a fixed reward of 5. The discount factor is set to $\eta = 0.95$. We run Q-learning for 10000 steps, and each time the agent reaches the terminal state, it is reset to the initial position.

We focus on the estimation at the initial state, as it serves as a natural indicator of the agent's performance. If the agent performs well at the initial state, it is likely to perform well in subsequent states. For the initial state, the maximum action value is $5\eta^4 \approx 4.073$.

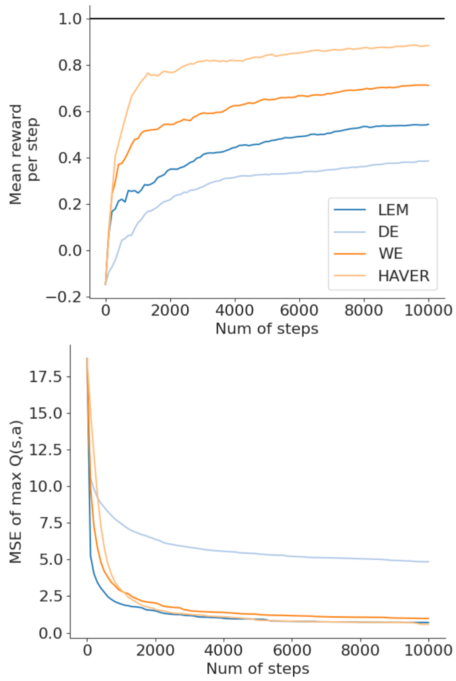

Figure 2: Q-learning in the regular grid world environment. The results are averaged over 1000 trials. The optimal mean reward per step is the black line.

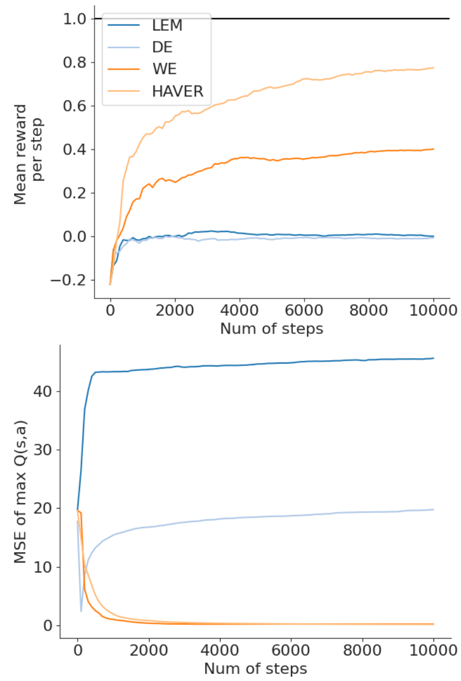

Figure 3: Q-learning in the inflated grid world with the number of actions at each state is duplicated to 4. The results are averaged over 1000 trials. The optimal mean reward per step is the black line.

In addition, we consider the optimal average reward per step, which is 1, as a direct measure of the agent's performance. To investigate the effect of increasing the number of actions, we design a unique grid world setting where each action-up, down, left, right- is duplicated M times, resulting in a total of $4 \cdot M$ actions. We refer to this setting as the inflated grid world.

We compare the performance of HAVER, LEM, DE, and WE estimators within Q-learning framework in two settings: the regular grid world and the inflated grid world where $M = 4$. The results for both settings are shown in Figure 2 and 3 respectively. In both settings, HAVER consistently outperforms the other estimators in achieving higher mean rewards and in lower MSE in $Q(s, a)$ estimation. HAVER's performance is more pronounced in the inflated grid world setting, where it converges to the mean reward much faster than the other estimators. We observe that the HAVER's MSE in $Q(s, a)$ estimation tends to lag behind in earlier time steps, likely due to the high bias introduced during the averaging computation.

5.3 Monte Carlo Tree Search (MCTS) Applied to the FrozenLake Environment

Maximum mean estimation is an integral part of the Monte Carlo Tree Search (MCTS) algorithm. During the backup step in MCTS, the agent aims to estimate the maximum value of a node based on the highest values of its child nodes. In this study, we investigate the effectiveness of maximum mean estimators in MCTS within the FrozenLake environment as implemented in OpenAI Gym (Brockman et al., 2016). In this environment, the agent's objective is to navigate an icy grid starting from the top-left cell to the bottom-right goal position. A key challenge is the presence of holes scattered across the grid; if the agent falls into a hole, the episode ends immediately with a reward of 0. Successfully reaching the goal position, however, grants a reward of 10, the maximum possible reward. If the agent fails to reach the goal position within a time limit of 40 steps, it receives a reward of 0.

We evaluate the performance of HAVER, LEM, and AE (the average estimator utilized in the original MCTS implementation by Kocsis and Szepesvári (2006)) in grid environments: a smaller 4x4 grid and a larger 8x8 grid. Both environments are deterministic, meaning that when the agent selects a direction, it moves exactly in that direction without any stochasticity. Our experiments vary the number of MCTS simulations to assess the impact of simulation count on the agent's performance. For the smaller 4x4 grid, we test a range of 32, 64, 128, and 256 MCTS simulations. For the larger 8x8 grid, which presents a larger state space and higher complexity, we extend the range of MCTS simulations to 600, 800, and 1000 to provide the agent with sufficient exploration capabilities. Figure 4 shows that HAVER outperforms the other estimators, achieving higher rewards across all tested simulation counts.

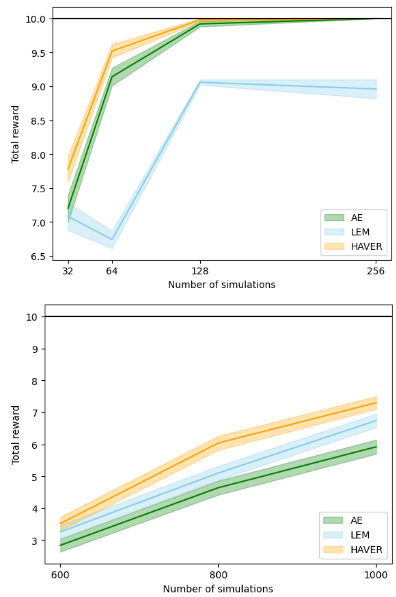

Figure 4: MCTS applied to the FrozenLake environments (top: 4x4 environment, bottom: 8x8 environment). The results are averaged over 500 trials. The optimal total reward is the black line.

6 Related Work

Maximum mean estimation (MME). MME can be taken as a rather naive task since the naive estimator of LEM seems to be very reasonable. However, as reported by Smith and Winkler (2006), LEM suffers from an overestimation bias, which can be large when the sample size is small or the number of distributions K is large. In the machine learning community, van Hasselt (2010) is the first one to formally study the MME problem, to our knowledge. In particular, the author reports that the overestimation problem of LEM is harmful to the performance of Q-learning (Watkins, 1989), which is a popular algorithm for reinforcement learning, and instead proposes a negatively biased estimator called double estimator (DE) that works similarly to the 2-fold cross-validation. There are numerous follow-up studies that propose various MME algorithms including weighted estimator (WE) (D'Eramo et al., 2016; D'Eramo et al., 2021) and maxmin estimator (Lan et al., 2020). The maximum mean estimation problem also arises from Monte Carlo tree search (MCTS) (Coulom, 2006; Kocsis and Szepesvári, 2006), which is an online planning methodology that forms a search tree from the current state s and aims to identify the best action by repeatedly running randomized rollouts in a simulator. Interestingly, the de facto standard algorithm called UCT (Kocsis and Szepesvári, 2006) uses the average estimator (AE) $\frac{1}{\sum_{i'=1}^{K} N_{i'}} \sum_{i=1}^{K} \sum_{j=1}^{N_i} X_{i,j}$ as an estimator for the value of each state. While this is not a consistent estimator in general, it is consistent in the MCTS setting since UCT is designed to take the best action most of the time asymptotically, i.e., $\frac{N_1}{\sum_i N_i} \to 1$. To address the issue of the slow convergence of AE in the nonasymptotic regime, Tesauro et al. (2010) proposed a Bayesian estimator that is essentially WE and proved its effectiveness in MCTS. In Dam et al. (2024), The estimator WE was further generalized to the weighted power mean estimator and applied to MCTS. However, to our knowledge, all of these studies lack precise error analysis and instead provide bounds on bias and variance separately.

Best-arm identification. A closely-related problem to MME is the best-arm identification problem (Chernoff, 1959; Even-Dar et al., 2006; Bubeck et al., 2009), especially the fixed budget setting. In this setting, at time step $t \in \{1, \ldots, T\}$, the learner chooses an arm I_t from K arms and observes a reward r_t from the chosen arm. After T time steps, the learner must output the estimated best arm, i.e., the arm with the largest mean reward, and the learner's performance is measured by the accuracy of the estimation. BAI has been studied extensively (Kalyanakrishnan et al., 2012; Karnin et al., 2013; Jamieson et al., 2014; Zhao et al., 2023), resulting in applications in adaptive crowdsourcing (Tanczos et al., 2017), hyperparameter optimization (Li et al., 2018), and accelerating the k-medoids algorithm (Bagaria et al., 2021). However, BAI is different from MME in two respects: MME (i) requires estimation of the largest mean *value* rather than the *identity* of the arm with the largest mean and (ii) does not perform active sampling; i.e., the samples are chosen from an external entity rather than being chosen by the learner.

7 Conclusion and Future Work

We have identified two key desiderata of an ideal estimator: (i) it should achieve an MSE rate as good as the oracle rate and (ii) it can strictly outperform the oracle rate in an instance-dependent manner. Then, we have proposed a novel estimator HAVER and proved that it satisfies the two desiderata. Our result is first of its kind and opens up numerous interesting future research directions. Finally, it would be interesting and relevant to practice to remove the i.i.d. assumption and instead consider nonstationary distributions, which we believe are better models of sample generation in Q-learning.

Acknowledgements

Tuan Ngo Nguyen and Kwang-Sung Jun were supported in part by the National Science Foundation under grant CCF-2327013.

References

Abramowitz, M. and Stegun, I. A. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55. US Government printing office, 1968.

Bagaria, V., Baharav, T. Z., Kamath, G. M., and David, N. T. Bandit-based monte carlo optimization for nearest neighbors. *IEEE Journal on Selected Areas in Information Theory*, 2(2):599–610, 2021.

Brockman, G., Cheung, V., Pettersson, L., Schneider, J., Schulman, J., Tang, J., and Zaremba, W. Openai gym. *arXiv preprint arXiv:1606.01540*, 2016.

Bubeck, S., Munos, R., and Stoltz, G. Pure Exploration in Multi-armed Bandits Problems. In *Proceedings of the International Conference on Algorithmic Learning Theory (ALT)*, pages 23–37, 2009.

Chernoff, H. Sequential design of experiments. *The Annals of Mathematical Statistics*, 30(3):755–770, 1959.

Coulom, R. Efficient selectivity and backup operators in monte-carlo tree search. In *5th International Conference on Computers and Games*, pages 72–83, 2006.

Dam, T., D'Eramo, C., Peters, J., and Pajarinen, J. A unified perspective on value backup and exploration in monte-carlo tree search. *Journal of Artificial Intelligence Research*, 81:511–577, 2024.

D'Eramo, C., Cini, A., Nuara, A., Pirotta, M., Alippi, C., Peters, J., and Restelli, M. Gaussian approximation for bias reduction in q-learning. *The Journal of Machine Learning Research*, 22(1):12690–12740, 2021.

D'Eramo, C., Restelli, M., and Nuara, A. Estimating maximum expected value through gaussian approximation. In *Proceedings of the International Conference on Machine Learning (ICML)*, pages 1032–1040. PMLR, 2016.

Even-Dar, E., Mannor, S., and Mansour, Y. Action Elimination and Stopping Conditions for the Multi-Armed Bandit and Reinforcement Learning Problems. *Journal of Machine Learning Research*, 7: 1079–1105, 2006.

Jamieson, K., Malloy, M., Nowak, R., and Bubeck, S. lil' UCB : An Optimal Exploration Algorithm for Multi-Armed Bandits. In *Proceedings of the Conference on Learning Theory (COLT)*, pages 423–439, 2014.

Jin, Y., Yang, Z., and Wang, Z. Is pessimism provably efficient for offline rl? In *Proceedings of the International Conference on Machine Learning (ICML)*, pages 5084–5096, 2021.

Kalyanakrishnan, S., Tewari, A., Auer, P., and Stone, P. PAC Subset Selection in Stochastic Multi-armed Bandits. In *Proceedings of the International Conference on Machine Learning (ICML)*, pages 655–662, 2012.

Karnin, Z., Koren, T., and Somekh, O. Almost Optimal Exploration in Multi-Armed Bandits. In *Proceedings of the International Conference on Machine Learning (ICML)*, pages 1238–1246, 2013.

Kocsis, L. and Szepesvári, C. Bandit based monte-carlo planning. In *Proceedings of the European Conference on Machine Learning (ECML)*, pages 282–293. Springer, 2006.

Lan, Q., Pan, Y., Fyshe, A., and White, M. Maxmin q-learning: Controlling the estimation bias of q-learning. In *Proceedings of the International Conference on Learning Representations (ICLR)*, 2020.

Lattimore, T. and Szepesvári, C. *Bandit Algorithms*. Cambridge University Press, 2020.

Li, L., Jamieson, K., DeSalvo, G., Rostamizadeh, A., and Talwalkar, A. Hyperband: A Novel Bandit-Based Approach to Hyperparameter Optimization. *Journal of Machine Learning Research*, 18(185):1–52, 2018.

Smith, J. E. and Winkler, R. L. The optimizer's curse: Skepticism and postdecision surprise in decision analysis. *Management Science*, 52(3):311–322, 2006.

Tanczos, E., Nowak, R., and Mankoff, B. A KL-LUCB algorithm for Large-Scale Crowdsourcing. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 5894–5903, 2017.

Tesauro, G., Rajan, V., and Segal, R. Bayesian inference in monte-carlo tree search. In *Proceedings of the Twenty-Sixth Conference on Uncertainty in Artificial Intelligence*, pages 580–588, 2010.

Thompson, W. R. On the Likelihood that One Unknown Probability Exceeds Another in View of the Evidence of Two Samples. *Biometrika*, 25(3/4):285, 1933.

van Hasselt, H. Double q-learning. *Advances in Neural Information Processing Systems (NeurIPS)*, 2010.

van Hasselt, H. Estimating the maximum expected value: an analysis of (nested) cross validation and the maximum sample average. *arXiv preprint arXiv:1302.7175*, 2013.

Watkins, C. Learning from delayed rewards. *PhD thesis, Cambridge University, Cambridge, England*, 1989.

Zhao, Y., Stephens, C., Szepesvári, C., and Jun, K.-S. Revisiting simple regret: Fast rates for returning a good arm. *Proceedings of the International Conference on Machine Learning (ICML)*, 2023.

Checklist

1. For all models and algorithms presented, check if you include:
   (a) A clear description of the mathematical setting, assumptions, algorithm, and/or model. [Yes]
   (b) An analysis of the properties and complexity (time, space, sample size) of any algorithm. [Yes]
   (c) (Optional) Anonymized source code, with specification of all dependencies, including external libraries. [No]
2. For any theoretical claim, check if you include:
   (a) Statements of the full set of assumptions of all theoretical results. [Yes]
   (b) Complete proofs of all theoretical results. [Yes] (most proofs are in the appendix)
   (c) Clear explanations of any assumptions. [Yes]
3. For all figures and tables that present empirical results, check if you include:
   (a) The code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL). [No]
   (b) All the training details (e.g., data splits, hyperparameters, how they were chosen). [Not Applicable]
   (c) A clear definition of the specific measure or statistics and error bars (e.g., with respect to the random seed after running experiments multiple times). [No Applicable]
   (d) A description of the computing infrastructure used. (e.g., type of GPUs, internal cluster, or cloud provider). [No Applicable]
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets, check if you include:
   (a) Citations of the creator If your work uses existing assets. [Not Applicable]
   (b) The license information of the assets, if applicable. [Not Applicable]
   (c) New assets either in the supplemental material or as a URL, if applicable. [Not Applicable]
   (d) Information about consent from data providers/curators. [Not Applicable]
   (e) Discussion of sensible content if applicable, e.g., personally identifiable information or offensive content. [Not Applicable]
5. If you used crowdsourcing or conducted research with human subjects, check if you include:
   (a) The full text of instructions given to participants and screenshots. [Not Applicable]
   (b) Descriptions of potential participant risks, with links to Institutional Review Board (IRB) approvals if applicable. [Not Applicable]
   (c) The estimated hourly wage paid to participants and the total amount spent on participant compensation. [Not Applicable]

Supplementary Materials

Appendix

Table of Contents

A More on Experiments

A.1 HAVER Estimator in Practice

The following Algorithm 2 is a practical version of HAVER. There are two key parts. First, we compute the unbiased estimate of the variance, i.e., $\hat{\sigma}_i^2 = \frac{\sum_{j=1}^{N_i} \left(X_{i,j} - \hat{\mu}_i\right)^2}{N_i - 1}$. Second, we incorporate these variance estimates into the the average weights in the averaging step. For each arm $i \in [K]$, its average weight becomes $\frac{N_i}{\hat{\sigma}_i^2}$. However, $\hat{\sigma}_i^2$ can be arbitrarily small, so it can causes the denominator to explode. To prevent this, we add a small tunable hyperparameter ε to $\hat{\sigma}_i^2$. This hyperparameter is set to 0.01 in our experiments.

Algorithm 2 HAVER Estimator in Practice

Input: A set of K arms with samples $\left\{X_{i,1:N_i}\right\}_{i=1}^K$, hyperparameter ε.

For each arm i, compute its empirical mean $\hat{\mu}_i$ and variance $\hat{\sigma}_i^2$

$$\hat{\mu}_i = \frac{1}{N_i} \sum_{j=1}^{N_i} X_{i,j}, \qquad \hat{\sigma}_i^2 = \frac{\sum_{j=1}^{N_i} \left(X_{i,j} - \hat{\mu}_i\right)^2}{N_i - 1}.$$

Find pivot arm $\hat{r}$

$$\hat{r} = \arg\max_{i \in [K]} \hat{\mu}_i - \gamma_i$$

where $\sqrt{\frac{18}{N_i} \log\left(\left(\frac{KS}{N_i}\right)^4\right)}$,

$S = N_{\max} \sum_{j \in [K]} N_j$, and $N_{\max} = \max_{i \in [K]} N_i$.

Form a set $\mathcal{B}$ of desirable arms

$$\mathcal{B} = \left\{i \in [K] : \hat{\mu}_i \geq \hat{\mu}_{\hat{r}} - \gamma_{\hat{r}},\, \gamma_i \leq \tfrac{3}{2}\gamma_{\hat{r}}\right\}.$$

Compute the weighted average of empirical means of desirable arms in set $\mathcal{B}$

$$\hat{\mu}^{\mathcal{B}} = \frac{1}{Z} \sum_{i \in \mathcal{B}} \frac{N_i}{\hat{\sigma}_i^2 + \varepsilon} \hat{\mu}_i,$$

where $Z = \sum_{i \in \mathcal{B}} \frac{N_i}{\hat{\sigma}_i^2 + \varepsilon}$.

return $\hat{\mu}^{\mathcal{B}}$.

A.2 Multi-Armed Bandits for Internet Ads

The following Figures 5 and 6 respectively shows the results for the K^*-best and Polynomial(α) instances mentioned in the main paper.

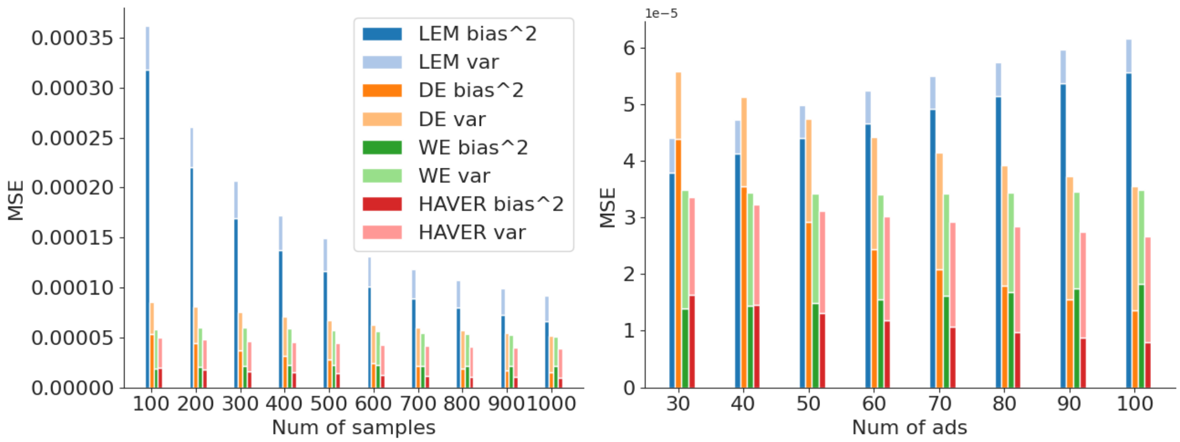

Figure 5: K^*-best instance. The results are averaged over 1000 trials.

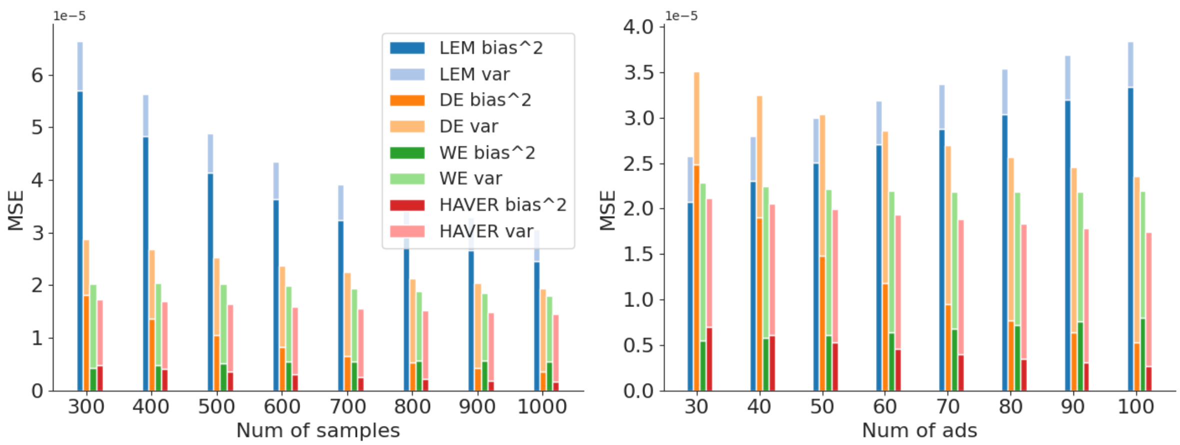

Figure 6: Polynomial(α) instance with $\alpha = 2$. The results are averaged over 1000 trials.

B HAVER's Corollaries

Corollary 8. *Under Assumption 7 (equal number of samples), let* $\gamma := \sqrt{\frac{18}{N}\log\left((K^2N)^4\right)}$, $\mathcal{B}^* := \left\{i \in [K] : \Delta_i \le \frac{1}{6}\gamma\right\}$, *and* $\mathcal{B}^+ := \left\{i \in [K] : \Delta_i \le \frac{8}{3}\gamma\right\}$. *HAVER achieves*

$$\begin{aligned}
&\mathrm{MSE}(\hat{\mu}^{HAVER})\\
&= \tilde{\mathcal{O}}\left(\left(\frac{1}{|\mathcal{B}^*|}\sum_{i=1}^{|\mathcal{B}^+|}\Delta_i\right)^2 \wedge \frac{1}{N}\right)\\
&\quad + \tilde{\mathcal{O}}\left(\left(\frac{1}{N}\left(\log\left(\frac{|\mathcal{B}^+|}{|\mathcal{B}^*|}\right)\right)^2\right) \wedge \frac{1}{N}\right)\\
&\quad + \tilde{\mathcal{O}}\left(\frac{1}{|\mathcal{B}^*|\,N}\right) + \tilde{\mathcal{O}}\left(\frac{1}{KN}\right).
\end{aligned}$$

Proof. Under Assumption 7 (equal number of samples), the definitions used in Theorem 6 are reduced to the following definitions.

First, $\forall i \in [K]$, the confidence width γ_i becomes γ

$$\gamma_i = \sqrt{\frac{18}{N_i}\log\left(\left(\frac{KN_{\max}\sum_{j\in[K]}N_j}{N_i}\right)^4\right)} = \sqrt{\frac{18}{N}\log\left((K^2N)^4\right)} =: \gamma.$$

Second, the ground truth of the pivot arm s becomes 1: $s = \arg\max_i \hat{\mu}_i - \frac{1}{6}\gamma_i = \arg\max_i \hat{\mu}_i - \frac{1}{6}\gamma = 1$ Third, $\mathcal{B}^*(r)$ and $\mathcal{B}^+(r)$ become $\mathcal{B}^*$ and $\mathcal{B}^+$ respectively as follow:

$$\begin{aligned}
\mathcal{B}^*(r) &= \left\{i \in [K] : \mu_i \ge \mu_s - \frac{1}{6}\gamma_s,\ \gamma_i \le \frac{3}{2}\gamma_r\right\}\\
&= \left\{i \in [K] : \mu_i \ge \mu_1 - \frac{1}{6}\gamma\right\}\\
&= \left\{i \in [K] : \Delta_i \le \frac{1}{6}\gamma\right\} =: \mathcal{B}^*
\end{aligned}$$

$$\begin{aligned}
\mathcal{B}^+(r) &= \left\{i \in [K] : \mu_i \ge \mu_s - \frac{4}{3}\gamma_s - \frac{4}{3}\gamma_i,\ \gamma_i \le \frac{3}{2}\gamma_r\right\}\\
&= \left\{i \in [K] : \mu_i \ge \mu_1 - \frac{4}{3}\gamma - \frac{4}{3}\gamma\right\}\\
&= \left\{i \in [K] : \Delta_i \le \frac{8}{3}\gamma\right\} =: \mathcal{B}^+.
\end{aligned}$$

Additionally, $d(r)$ and $n_*(r)$ becomes d and n_*: $d(r) = |\mathcal{B}^+(r)| - |\mathcal{B}^*(r)| = |\mathcal{B}^+| - |\mathcal{B}^*| =: d$ and $n_*(r) = |\mathcal{B}^*(r)| = |\mathcal{B}^*| =: n_*$. We adopt a slight abuse of notation for $\mathcal{B}^*$, $\mathcal{B}^+$, d and n_*. From Theorem 6, we have

$$\begin{aligned}
&\mathrm{MSE}(\hat{\mu}^{\mathrm{HAVER}})\\
&= \tilde{\mathcal{O}}\left(\left(\max_{r\in\mathcal{R}}\frac{1}{\sum_{j\in\mathcal{B}^*(r)}N_j}\sum_{i\in\mathcal{B}^+(r)}N_i\Delta_i\right)^2 \wedge \frac{1}{N_1}\right)\\
&\quad + \tilde{\mathcal{O}}\left(\left(\max_{r\in\mathcal{R}}\max_{k=0}^{d(r)}\max_{\substack{\mathcal{S}:\mathcal{B}^*(r)\subseteq\mathcal{S}\subseteq\mathcal{B}^+(r)\\|\mathcal{S}|=n_*(r)+k}}\frac{k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)}N_j}{\left(\sum_{j\in\mathcal{S}}N_j\right)^2}\right) \wedge \frac{1}{N_1}\right)
\end{aligned}$$

$$+ \tilde{\mathcal{O}}\left(\left(\frac{1}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}\right) \wedge \frac{1}{N_1}\right)$$
$$+ \tilde{\mathcal{O}}\left(\frac{1}{KN_1}\right).$$

The first component becomes

$$\max_{r\in\mathcal{R}} \frac{1}{\sum_{j\in\mathcal{B}^*(r)} N_j} \sum_{i\in\mathcal{B}^+(r)} N_i\Delta_i^2 = \left(\frac{1}{|\mathcal{B}^*|}\sum_{i=1}^{|\mathcal{B}^+|}\Delta_i\right)^2.$$

The second component becomes

$$\begin{aligned}
&\max_{r\in\mathcal{R}} \max_{k=0}^{d(r)} \max_{\substack{\mathcal{S}:\mathcal{B}^*(r)\subseteq\mathcal{S}\subseteq\mathcal{B}^+(r)\\ |\mathcal{S}|=n_*(r)+k}} \frac{k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}{\left(\sum_{j\in\mathcal{S}} N_j\right)^2}\\
&= \max_{k=0}^{d} \frac{k^2 N}{\left((n_*(r)+k)N\right)^2}\\
&= \max_{k=0}^{d} \frac{k^2}{\left(n_*(r)+k\right)^2 N}\\
&\leq \frac{1}{N}\left(\log\left(\frac{n_*+d}{n_*}\right)\right)^2 && \text{(use Lemma 34)}\\
&= \frac{1}{N}\left(\log\left(\frac{|\mathcal{B}^+|}{|\mathcal{B}^*|}\right)\right)^2.
\end{aligned}$$

The third component becomes

$$\frac{1}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j} = \frac{1}{|\mathcal{B}^*|\,N}.$$

Thus, we have

$$\begin{aligned}
&\mathrm{MSE}(\hat{\mu}^{\mathrm{HAVER}})\\
&= \tilde{\mathcal{O}}\left(\left(\frac{1}{|\mathcal{B}^*|}\sum_{i=1}^{|\mathcal{B}^+|}\Delta_i\right)^2 \wedge \frac{1}{N}\right)\\
&\quad + \tilde{\mathcal{O}}\left(\left(\frac{1}{N}\left(\log\left(\frac{|\mathcal{B}^+|}{|\mathcal{B}^*|}\right)\right)^2\right) \wedge \frac{1}{N}\right)\\
&\quad + \tilde{\mathcal{O}}\left(\left(\frac{1}{|\mathcal{B}^*|\,N}\right) \wedge \frac{1}{N}\right)\\
&\quad + \tilde{\mathcal{O}}\left(\frac{1}{KN}\right)\\
&= \tilde{\mathcal{O}}\left(\left(\frac{1}{|\mathcal{B}^*|}\sum_{i=1}^{|\mathcal{B}^+|}\Delta_i\right)^2 \wedge \frac{1}{N}\right)\\
&\quad + \tilde{\mathcal{O}}\left(\left(\frac{1}{N}\left(\log\left(\frac{|\mathcal{B}^+|}{|\mathcal{B}^*|}\right)\right)^2\right) \wedge \frac{1}{N}\right)
\end{aligned}$$

$$+\tilde{\mathcal{O}}\left(\frac{1}{|\mathcal{B}^*|\,N}\right)+\tilde{\mathcal{O}}\left(\frac{1}{KN}\right).$$

□

Corollary 12. Under Assumption 7 (equal number of samples). Consider the all-best instance where $\forall i \in [K], \mu_i = \mu_1$, HAVER achieves

$$\mathrm{MSE}(\hat{\mu}^{\mathrm{HAVER}}) = \tilde{\mathcal{O}}\left(\frac{1}{KN}\right).$$

Proof. From Corollary 8, we have $\mathcal{B}^* = \left\{i \in [K] : \Delta_i \le \frac{1}{6}\gamma\right\}$ and $\mathcal{B}^+ = \left\{i \in [K] : \Delta_i \le \frac{8}{3}\gamma\right\}$, and

$$\begin{aligned}&\mathrm{MSE}(\hat{\mu}^{\mathrm{HAVER}})\\ &=\tilde{\mathcal{O}}\left(\left(\frac{1}{|\mathcal{B}^*|}\sum_{i=1}^{|\mathcal{B}^+|}\Delta_i\right)^2 \wedge \frac{1}{N}\right)\\ &\quad+\tilde{\mathcal{O}}\left(\left(\frac{1}{N}\left(\log\left(\frac{|\mathcal{B}^+|}{|\mathcal{B}^*|}\right)\right)^2\right)\wedge\frac{1}{N}\right)\\ &\quad+\tilde{\mathcal{O}}\left(\frac{1}{|\mathcal{B}^*|\,N}\right)+\tilde{\mathcal{O}}\left(\frac{1}{KN}\right).\end{aligned}$$

For the all-best instance, we can see that $\mathcal{B}^* = [K]$, $\mathcal{B}^+ = [K]$. Therefore, from Corollary 8, the first term $\left(\frac{1}{|\mathcal{B}^*|}\sum_{i=1}^{|\mathcal{B}^+|}\Delta_i\right)^2$ becomes 0 since $\forall i \in \mathcal{B}^+ = [K]$, $\Delta_i = 0$. Also, the second term $\frac{1}{N}\left(\log\left(\frac{|\mathcal{B}^+|}{|\mathcal{B}^*|}\right)\right)^2 = 0$ since $\log\left(\frac{|\mathcal{B}^+|}{|\mathcal{B}^*|}\right) = 0$. Also, we have $\frac{1}{|\mathcal{B}^*|N} = \frac{1}{KN}$.

Therefore, from from Corollary 8, HAVER achieves

$$\mathrm{MSE}(\hat{\mu}^{\mathrm{HAVER}}) = \tilde{\mathcal{O}}\left(\frac{1}{KN}\right).$$

□

B.1 Proof of Corollary 9

Corollary 9. *Under Assumption 7 (equal number of samples), consider the K^*-best instance where $K^* \le K$ $\forall i \in [K^*]$, $\mu_i = \mu_1$. If $N > \frac{256}{\Delta_{K^*+1}^2}\log\left(\frac{256K^2}{\Delta_{K^*+1}^2 e}\right)$, HAVER achieves*

$$\mathrm{MSE}(\hat{\mu}^{HAVER}) = \tilde{\mathcal{O}}\left(\frac{1}{K^*N}\right).$$

Proof. From Corollary 8, we have $\mathcal{B}^* = \left\{i \in [K] : \Delta_i \le \frac{1}{6}\gamma\right\}$, $\mathcal{B}^+ = \left\{i \in [K] : \Delta_i \le \frac{8}{3}\gamma\right\}$, and

$$\begin{aligned}&\mathrm{MSE}(\hat{\mu}^{\mathrm{HAVER}})\\ &=\tilde{\mathcal{O}}\left(\left(\frac{1}{|\mathcal{B}^*|}\sum_{i=1}^{|\mathcal{B}^+|}\Delta_i\right)^2 \wedge \frac{1}{N}\right)\\ &\quad+\tilde{\mathcal{O}}\left(\left(\frac{1}{N}\left(\log\left(\frac{|\mathcal{B}^+|}{|\mathcal{B}^*|}\right)\right)^2\right)\wedge\frac{1}{N}\right)\end{aligned}$$

$$+\tilde{\mathcal{O}}\left(\frac{1}{|\mathcal{B}^*|\,N}\right)+\tilde{\mathcal{O}}\left(\frac{1}{KN}\right).$$

Lemma 39 states that $N > \frac{256}{\Delta_{K^*+1}^2}\log\left(\frac{256K^2}{\Delta_{K^*+1}^2 e}\right)$ implies $\Delta_{K^*+1} > \frac{8}{3}\sqrt{\frac{18}{N}\log\left(K^2N\right)}$, Thus, we have $\Delta_{K^*+1} > \frac{8}{3}\gamma$. Therefore, we have $\mathcal{B}^* = [K^*]$, $\mathcal{B}^+ = [K^*]$.

Therefore, from Corollary 8, the first term $\frac{1}{|\mathcal{B}^*|}\sum_{i=1}^{|\mathcal{B}^+|}\Delta_i$ becomes 0 since $\forall i \in \mathcal{B}^+ = [K^*]$, $\Delta_i = 0$. Also, the second term $\frac{1}{N}\left(\log\left(\frac{|\mathcal{B}^+|}{|\mathcal{B}^*|}\right)\right)^2 = 0$ since $\log\left(\frac{|\mathcal{B}^+|}{|\mathcal{B}^*|}\right) = 0$. Also, we have $\frac{1}{|\mathcal{B}^*|N} = \frac{1}{K^*N}$.

Therefore, from from Corollary 8, HAVER achieves

$$\mathrm{MSE}(\hat{\mu}^{\mathrm{HAVER}}) = \tilde{\mathcal{O}}\left(\frac{1}{K^*N}\right)+\tilde{\mathcal{O}}\left(\frac{1}{KN}\right)=\tilde{\mathcal{O}}\left(\frac{1}{K^*N}\right).$$

□

B.2 Proof of Corollary 10

Corollary 10. *Under Assumption 7 (equal number of samples), consider the Poly(α) instance where* $\forall i \geq 2$, $\Delta_i = \left(\frac{i}{K}\right)^{\alpha}$ *where* $\alpha \geq 0$. *If* $N \leq \frac{1}{2}\log(K^2N)\left(\frac{K}{\alpha}\log(2)\right)^{2\alpha}$, *HAVER achieves*

$$\mathrm{MSE}(\hat{\mu}^{HAVER}) = \tilde{\mathcal{O}}\left(\frac{1}{(\alpha\wedge K)N}\right).$$

Proof. From Corollary 8, we have $\mathcal{B}^* = \left\{i \in [K] : \Delta_i \leq \frac{1}{6}\gamma\right\}$ and $\mathcal{B}^+ = \left\{i \in [K] : \Delta_i \leq \frac{8}{3}\gamma\right\}$, and

$$\begin{aligned}&\mathrm{MSE}(\hat{\mu}^{\mathrm{HAVER}})\\ &=\tilde{\mathcal{O}}\left(\left(\frac{1}{|\mathcal{B}^*|}\sum_{i=1}^{|\mathcal{B}^+|}\Delta_i\right)^2\wedge\frac{1}{N}\right)\\ &\quad+\tilde{\mathcal{O}}\left(\left(\frac{1}{N}\left(\log\left(\frac{|\mathcal{B}^+|}{|\mathcal{B}^*|}\right)\right)^2\right)\wedge\frac{1}{N}\right)\\ &\quad+\tilde{\mathcal{O}}\left(\frac{1}{|\mathcal{B}^*|\,N}\right)+\tilde{\mathcal{O}}\left(\frac{1}{KN}\right).\end{aligned}$$

We observe the condition in set $\mathcal{B}^*$, $\forall i \in [K]$

$$\Delta_i \leq \frac{1}{6}\gamma = \frac{1}{6}\sqrt{\frac{18}{N}\log\left(\left(K^2N\right)^4\right)} = \sqrt{2\log\left(K^2N\right)}\sqrt{\frac{1}{N}}$$

We denote $a_1 = \sqrt{2\log\left(K^2N\right)}$. We rewrite $\mathcal{B}^* = \left\{\forall i \in [K] : a_1\sqrt{\frac{1}{N}}\right\}$ We observe the condition in set $\mathcal{B}^+$, $\forall i \in [K]$

$$\Delta_i \leq \frac{8}{3}\gamma = \frac{8}{3}\sqrt{\frac{18}{N}\log\left(\left(K^2N\right)^4\right)} = 16\sqrt{2\log\left(K^2N\right)}\sqrt{\frac{1}{N}}.$$

We denote $a_2 = 16\sqrt{2\log\left(K^2N\right)}$. We rewrite $\mathcal{B}^+ = \left\{\forall i \in [K] : a_2\sqrt{\frac{1}{N}}\right\}$. We have the ratio $\frac{2a_2}{a_1} = 32$. The assumption $N \leq \frac{1}{2}\log(K^2N)\left(\frac{K}{\alpha}\log(2)\right)^{2\alpha}$ satisfies $N \leq \left(\frac{a_1^2}{4}\wedge a_2\right)\left(\frac{K}{\alpha}\log(2)\right)^{2\alpha}$. Within this context, we can use Lemma 38. For the first term, we use Lemma 38 to bound

$$\begin{aligned}&\left(\frac{1}{|\mathcal{B}^*|}\sum_{i=1}^{|\mathcal{B}^+|}\Delta_i\right)^2\\ \leq&\left(\frac{2a_2\left(\frac{4a_2}{a_1}\right)^{\frac{1}{\alpha}}}{(\alpha+1)\sqrt{N}}\right)^2\\ \leq&\left(\frac{32\sqrt{2\log\left(K^2N\right)}\,(64)^{\frac{1}{\alpha}}}{(\alpha+1)\sqrt{N}}\right)^2\\ =&\frac{2048\log\left(K^2N\right)(64)^{\frac{2}{\alpha}}}{(\alpha+1)^2 N}.\end{aligned}$$

For the second term, we use Lemma 38 to bound

$$\frac{|\mathcal{B}^+|}{|\mathcal{B}^*|}\leq\left(\frac{2a_2}{a_1}\right)^{\frac{1}{\alpha}}=32^{\frac{1}{\alpha}}.$$

Hence,

$$\frac{1}{N}\left(\log\left(\frac{|\mathcal{B}^+|}{|\mathcal{B}^*|}\right)\right)^2\leq\frac{1}{N}\left(\log\left(32^{\frac{1}{\alpha}}\right)\right)^2=\frac{1}{\alpha^2N}\log^2(32)$$

For the third term, we use Lemma 38 to bound

$$\frac{1}{|\mathcal{B}^*|}\leq\frac{1}{K\left(\frac{\sqrt{2\log(K^2N)}}{2}\sqrt{\frac{1}{N}}\right)^{\frac{1}{\alpha}}}$$

From the assumption $N\leq\frac{1}{2}\log(K^2N)\left(\frac{K}{\alpha}\log(2)\right)^{2\alpha}$, we have

$$\begin{aligned}&N\leq\frac{1}{2}\log(K^2N)\left(\frac{K}{\alpha}\log(2)\right)^{2\alpha}\\ \Leftrightarrow&\sqrt{\frac{1}{N}}\geq\frac{1}{\sqrt{\frac{1}{2}\log(K^2N)\left(\frac{K}{\alpha}\log(2)\right)^{2\alpha}}}\\ \Leftrightarrow&\frac{\sqrt{2\log\left(K^2N\right)}}{2}\sqrt{\frac{1}{N}}\geq\frac{\frac{\sqrt{2\log(K^2N)}}{2}}{\sqrt{\frac{1}{2}\log(K^2N)\left(\frac{K}{\alpha}\log(2)\right)^{2\alpha}}}\\ \Leftrightarrow&\frac{\sqrt{2\log\left(K^2N\right)}}{2}\sqrt{\frac{1}{N}}\geq\frac{1}{\left(\frac{K}{\alpha}\log(2)\right)^{\alpha}}\\ \Leftrightarrow&\left(\frac{\sqrt{2\log\left(K^2N\right)}}{2}\sqrt{\frac{1}{N}}\right)^{\frac{1}{\alpha}}\geq\frac{1}{\frac{K}{\alpha}\log(2)}\\ \Leftrightarrow&K\left(\frac{\sqrt{2\log\left(K^2N\right)}}{2}\sqrt{\frac{1}{N}}\right)^{\frac{1}{\alpha}}\geq\frac{\alpha}{\log(2)}.\end{aligned}$$

Therefore,

$$\frac{1}{|\mathcal{B}^*|\,N}\leq\frac{1}{K\left(\frac{\sqrt{2\log(K^2N)}}{2}\sqrt{\frac{1}{N}}\right)^{\frac{1}{\alpha}}N}\leq\frac{\log(2)}{\alpha N}.$$

Combining all the terms, we have

$$\begin{aligned}
&\mathrm{MSE}(\hat{\mu}^{\mathrm{HAVER}})\\
=&\tilde{\mathcal{O}}\left(\left(\frac{1}{|\mathcal{B}^*|}\sum_{i=1}^{|\mathcal{B}^+|}\Delta_i\right)^2 \wedge \frac{1}{N}\right)\\
&+\tilde{\mathcal{O}}\left(\left(\frac{1}{N}\left(\log\left(\frac{|\mathcal{B}^+|}{|\mathcal{B}^*|}\right)\right)^2\right)\wedge\frac{1}{N}\right)\\
&+\tilde{\mathcal{O}}\left(\frac{1}{|\mathcal{B}^*|\,N}\right)+\tilde{\mathcal{O}}\left(\frac{1}{KN}\right)\\
=&\tilde{\mathcal{O}}\left(\frac{2048\log\left(K^2N\right)(64)^{\frac{2}{\alpha}}}{(\alpha+1)^2N}\wedge\left(\frac{1}{N}\right)\right)\\
&+\tilde{\mathcal{O}}\left(\frac{1}{\alpha^2N}\log^2(32)\wedge\left(\frac{1}{N}\right)\right)\\
&+\tilde{\mathcal{O}}\left(\frac{\log(2)}{\alpha N}\right)+\tilde{\mathcal{O}}\left(\frac{1}{KN}\right)\\
=&\tilde{\mathcal{O}}\left(\frac{1}{(\alpha\wedge K)N}\right).
\end{aligned}$$

□

B.3 Proof of Corollary 11

Corollary 11. *Consider the all-best instance where $\forall i\in[K]$, $\mu_i=\mu_1$ and the number of samples are characterized by $N_i=(K-i+1)^\beta$ where $\beta\in(0,1)$, HAVER achieves*

$$\mathrm{MSE}(\hat{\mu}^{HAVER})=\tilde{\mathcal{O}}\left(\frac{1}{KN_1}\right).$$

Proof. In this instance, we have $s=\arg\max_{i\in[K]}\mu_i-\frac{1}{6}\gamma_i=1$. The reasons are (1) $\forall i\in[K]$, $\mu_i=\mu_1$ in the all-best instance and (2) the number of samples $N_i=(K-i+1)^\beta$ decreases and γ_i increases from arm 1 to K, thus γ_1 is the smallest. From Theorem 6, we have

$$\mathcal{B}^*(r)=\left\{i\in[K]:\mu_i\geq\mu_1-\frac{1}{6}\gamma_1,\ \gamma_i\leq\frac{3}{2}\gamma_r\right\}=\left\{i\in[K]:\Delta_i\leq\frac{1}{6}\gamma_1,\ \gamma_i\leq\frac{3}{2}\gamma_r\right\}$$

and

$$\mathcal{B}^+(r):=\left\{i\in[K]:\mu_i\geq\mu_1-\frac{4}{3}\gamma_1-\frac{4}{3}\gamma_i,\ \gamma_i\leq\frac{3}{2}\gamma_r\right\}=\left\{i\in[K]:\Delta_i\leq\frac{4}{3}\gamma_i+\frac{4}{3}\gamma_1,\ \gamma_i\leq\frac{3}{2}\gamma_r\right\}.$$

In the all-best instance, we have $\forall i\in[K]$, $\Delta_i=0$. Therefore, in this instance, for every $r\in\mathcal{R}$, the first condition of $\mathcal{B}^*(r)$, $\forall i\in[K]$, $\Delta_i\leq\frac{1}{6}\gamma_1$ is always satisfied and the first condition of $\mathcal{B}^+(r)$, $\forall i\in[K]$, $\Delta_i\leq\frac{4}{3}\gamma_i+\frac{4}{3}\gamma_1$, is always satisfied. Also, the second conditions of both $\mathcal{B}^*(r)$ and $\mathcal{B}^+(r)$ are the same. Thus, we have $\forall r\in\mathcal{R}$, $\mathcal{B}^+(r)=\mathcal{B}^*(r)$.

From Theorem 6, we have

$$\begin{aligned}
&\mathrm{MSE}(\hat{\mu}^{\mathrm{HAVER}})\\
=&\tilde{\mathcal{O}}\left(\left(\max_{r\in\mathcal{R}}\frac{1}{\sum_{j\in\mathcal{B}^*(r)}N_j}\sum_{i\in\mathcal{B}^+(r)}N_i\Delta_i\right)^2\wedge\frac{1}{N_1}\right)
\end{aligned}$$

$$
\begin{aligned}
&+ \tilde{\mathcal{O}}\left(\left(\max_{r\in\mathcal{R}} \max_{k=0}^{d} \max_{\substack{S:\mathcal{B}^*(r)\subseteq\mathcal{S}\subseteq\mathcal{B}^+(r)\\ |\mathcal{S}|=n_*(r)+k}} \frac{k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}{\left(\sum_{j\in\mathcal{S}} N_j\right)^2}\right) \wedge \frac{1}{N_1}\right)\\
&+ \tilde{\mathcal{O}}\left(\left(\frac{1}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}\right) \wedge \frac{1}{N_1}\right)\\
&+ \tilde{\mathcal{O}}\left(\frac{1}{KN_1}\right).
\end{aligned}
$$

For the first term, $\left(\max_{r\in\mathcal{R}} \frac{1}{\sum_{j\in\mathcal{B}^*(r)} N_j} \sum_{i\in\mathcal{B}^+(r)} N_i\Delta_i\right)^2$ becomes 0 since $\forall i\in[K]$, $\Delta_i = 0$.

For the second term, $\left(\max_{r\in\mathcal{R}} \max_{k=0}^{d(r)} \max_{\substack{S:\mathcal{B}^*(r)\subseteq\mathcal{S}\subseteq\mathcal{B}^+(r)\\ |\mathcal{S}|=n_*(r)+k}} \frac{k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}{\left(\sum_{j\in\mathcal{S}} N_j\right)^2}\right)$ becomes 0 since $\forall r\in\mathcal{R}$, $d(r) = \left|\mathcal{B}^+(r)\right| - \left|\mathcal{B}^*(r)\right| = 0$ as $\mathcal{B}^+(r) = \mathcal{B}^*(r)$.

For the third term, we have

$$
\frac{1}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j} = \frac{1}{\sum_{j\in\mathcal{B}^*(1)} N_j}.
$$

Because the first condition of set $\mathcal{B}^*$ is always satisfied, set $\mathcal{B}^*$ becomes just a matter of arm counts condition. Therefore, $r = 1$ results in minimal number of arms in $\mathcal{B}^*(r)$ because (1) the optimal arm 1 is always in $\mathcal{R}$ and (2) among arm $r\in\mathcal{R}$, arm $r = 1$ has the highest number of samples making the condition $\gamma_i \le \frac{3}{2}\gamma_r$ most restrictive. An arm i is included in $\mathcal{B}^*(1)$ if it satisfies

$$
\begin{aligned}
&\gamma_i \le \frac{3}{2}\gamma_1\\
\Leftrightarrow &\frac{18}{N_i}\log\left(\left(\frac{KS}{N_i}\right)^{2\lambda}\right) \le \frac{9}{4}\frac{18}{N_1}\log\left(\left(\frac{KS}{N_1}\right)^{2\lambda}\right)\\
\Leftrightarrow &\frac{1}{N_i}\log\left(\frac{KS}{N_i}\right) \le \frac{9}{4}\frac{1}{N_1}\log\left(\frac{KS}{N_1}\right)\\
\Leftrightarrow &N_i \ge \frac{4}{9}N_1\frac{\log\left(\frac{KS}{N_i}\right)}{\log\left(\frac{KS}{N_1}\right)}\\
\Rightarrow &N_i \ge \frac{4}{9}N_1 && \text{(use } N_i \le N_1\text{)}\\
\Leftrightarrow &(K-i+1)^\beta \ge \frac{4}{9}K^\beta\\
\Leftrightarrow &K-i+1 \ge \left(\frac{4}{9}\right)^{\frac{1}{\beta}}K\\
\Leftrightarrow &i \le K\left(1-\left(\frac{4}{9}\right)^{\frac{1}{\beta}}\right).
\end{aligned}
$$

Thus, we have

$$
\sum_{j\in\mathcal{B}^*(1)} N_j \ge \sum_{j=1}^{K\left(1-\left(\frac{4}{9}\right)^{\frac{1}{\beta}}\right)} N_j.
$$

Let $\omega = 1 - \left(\frac{4}{9}\right)^{\frac{1}{\beta}}$. We have

$$\sum_{j=1}^{K\left(1-\left(\frac{4}{9}\right)^{\frac{1}{\beta}}\right)} N_j = \sum_{j=1}^{\omega K} N_j = \sum_{j=1}^{\omega K} (K-j+1)^\beta \geq \int_{j=1}^{\omega K} (K-j+1)^\beta \,\mathrm{d}j.$$

We do change of variables of the integral. Let $u = K - j + 1$, then $\mathrm{d}u = -\,\mathrm{d}j$. When $j = 1$, $u = K$. When $j = \omega K$, $u = K - \omega K$. The integral becomes

$$\begin{aligned}
&\int_{j=1}^{\omega K} (K-j+1)^\beta \,\mathrm{d}j \\
=& \int_{j=K}^{K-\omega K} -u^\beta \,\mathrm{d}u \\
=& \int_{j=K-\omega K}^{K} u^\beta \,\mathrm{d}u \\
=& \left(\frac{u^{\beta+1}}{\beta+1}\right)\Big|_{K-\omega K}^{K} \\
=& \frac{K^{\beta+1} - (K-\omega K)^{\beta+1}}{\beta+1} \\
=& \frac{K^{\beta+1} - \big((1-\omega)K\big)^{\beta+1}}{\beta+1}.
\end{aligned}$$

Consider function $f(x) = x^{\beta+1}$ for any $x \in \mathbb{R}$. We have $f'(x) = (\beta+1)x^\beta$ It is trivial to see that $f(x)$ is convex. Thus, we have $f(x) - f(y) \geq f'(y)(x-y)$. We plug in $x = K$ and $y = (1-\omega)K$. We have

$$K^{\beta+1} - \big((1-\omega)K\big)^{\beta+1} \geq (\beta+1)\left((1-\omega)K\right)^\beta \omega K.$$

Therefore,

$$\begin{aligned}
&\int_{j=1}^{\omega K} (K-j+1)^\beta \,\mathrm{d}j \\
=& \frac{K^{\beta+1} - \big((1-\omega)K\big)^{\beta+1}}{\beta+1} \\
\geq& \frac{(\beta+1)\left((1-\omega)K\right)^\beta \omega K}{\beta+1} \\
=& \left((1-\omega)K\right)^\beta \omega K.
\end{aligned}$$

Plugging $\omega = 1 - \left(\frac{4}{9}\right)^{\frac{1}{\beta}}$, we have

$$\begin{aligned}
&\left((1-\omega)K\right)^\beta \omega K \\
=& \left(\left(1 - 1 + \left(\frac{4}{9}\right)^{\frac{1}{\beta}}\right)K\right)^\beta \left(1 - \left(\frac{4}{9}\right)^{\frac{1}{\beta}}\right)K \\
=& \frac{4}{9}\left(1 - \left(\frac{4}{9}\right)^{\frac{1}{\beta}}\right)K^{\beta+1} \\
\geq& \frac{4}{9}\frac{5}{9}K^{\beta+1} && \text{(use } (1-\left(\tfrac{4}{9}\right)^{\frac{1}{\beta}}) > \tfrac{5}{9} \text{ with } \beta \in (0,1)\text{)} \\
=& \frac{20}{81}K^{\beta+1}.
\end{aligned}$$

Therefore,

$$\sum_{j \in \mathcal{B}^*(r)} N_j \geq \sum_{j=1}^{K\left(1-\left(\frac{4}{9}\right)^{\frac{1}{\beta}}\right)} N_j \geq \int_{j=1}^{\omega K} (K-j+1)^\beta \, \mathrm{d}j \geq \frac{4}{27} K^{\beta+1}.$$

Therefore, for the third term, we have

$$\begin{aligned}
&\frac{1}{\min_{r \in \mathcal{R}} \sum_{j \in \mathcal{B}^*(r)} N_j} \\
&= \frac{1}{\sum_{j \in \mathcal{B}^*(1)} N_j} \\
&\leq \frac{1}{\frac{20}{81} K^{\beta+1}} \\
&= \frac{81}{20} \left(\frac{1}{K N_1} \right). && \text{(use } N_1 = K^\beta\text{)}
\end{aligned}$$

Combining all the terms, we have

$$\begin{aligned}
&\mathrm{MSE}(\hat{\mu}^{\mathrm{HAVER}}) \\
&= \tilde{\mathcal{O}}\left(\left(\max_{r \in \mathcal{R}} \frac{1}{\sum_{j \in \mathcal{B}^*(r)} N_j} \sum_{i \in \mathcal{B}^+(r)} N_i \Delta_i \right)^2 \wedge \frac{1}{N_1} \right) \\
&\quad + \tilde{\mathcal{O}}\left(\left(\max_{r \in \mathcal{R}} \max_{k=0}^{d} \max_{\substack{S: \mathcal{B}^*(r) \subseteq \mathcal{S} \subseteq \mathcal{B}^+(r) \\ |\mathcal{S}| = n_*(r)+k}} \frac{k \sum_{j \in \mathcal{S} \setminus \mathcal{B}^*(r)} N_j}{\left(\sum_{j \in \mathcal{S}} N_j \right)^2} \right) \wedge \frac{1}{N_1} \right) \\
&\quad + \tilde{\mathcal{O}}\left(\left(\frac{1}{\min_{r \in \mathcal{R}} \sum_{j \in \mathcal{B}^*(r)} N_j} \right) \wedge \frac{1}{N_1} \right) \\
&\quad + \tilde{\mathcal{O}}\left(\frac{1}{K N_1} \right) \\
&= \tilde{\mathcal{O}}\left(\left(\frac{1}{K N_1} \right) \wedge \frac{1}{N_1} \right) \\
&\quad + \tilde{\mathcal{O}}\left(\frac{1}{K N_1} \right) \\
&= \tilde{\mathcal{O}}\left(\frac{1}{K N_1} \right).
\end{aligned}$$

□

C HAVER's Theorem 6

Recall the following definitions that would be used in HAVER.

For each arm $i \in [K]$, we define

$$\gamma_i := \sqrt{\frac{18}{N_i} \log\left(\left(\frac{KS}{N_i}\right)^{2\lambda}\right)},$$

as its confidence width where $S = N_{\max} \sum_{j \in [K]} N_j$ and $\lambda = 2$. We define

$$\hat{r} := \arg\max_{i \in [K]} \hat{\mu}_i - \gamma_i$$

as the pivot arm with maximum lower confidence bound, and define

$$s := \arg\max_{i \in [K]} \mu_i - \frac{1}{6}\gamma_i$$

as the ground truth version of the pivot arm. We define

$$\mathcal{B} := \left\{ i \in [K] : \hat{\mu}_i \geq \hat{\mu}_{\hat{r}} - \gamma_{\hat{r}},\, \gamma_i \leq \frac{3}{2}\gamma_{\hat{r}} \right\}$$

as the candidate set of arms. We define

$$\mathcal{R} := \left\{ r \in [K] : \mu_s - \frac{4}{3}\gamma_s + \frac{2}{3}\gamma_r \leq \mu_r \leq \mu_s - \frac{1}{6}\gamma_s + \frac{1}{6}\gamma_r \right\}$$

as the statistically-plausible set of pivot arms $\hat{r}$. For any $r \in \mathcal{R}$, we define

$$\mathcal{B}^*(r) := \left\{ i \in [K] : \mu_i \geq \mu_s - \frac{1}{6}\gamma_s,\, \gamma_i \leq \frac{3}{2}\gamma_r \right\},$$

$$\mathcal{B}^+(r) := \left\{ i \in [K] : \mu_i \geq \mu_s - \frac{4}{3}\gamma_s - \frac{4}{3}\gamma_i,\, \gamma_i \leq \frac{3}{2}\gamma_r \right\},$$

$d(r) = \left|\mathcal{B}^+(r)\right| - \left|\mathcal{B}^*(r)\right|$, and $n_*(r) = \left|\mathcal{B}^*(r)\right|$.

In addition, throughout the proofs, we define $(x)_+ = \max(0, x)$.

Theorem 6. *HAVER achieves*

$$\begin{aligned}
&\mathrm{MSE}(\hat{\mu}^{HAVER}) \\
&= \tilde{\mathcal{O}}\left(\left(\max_{r \in \mathcal{R}} \frac{1}{\sum_{j \in \mathcal{B}^*(r)} N_j} \sum_{i \in \mathcal{B}^+(r)} N_i \Delta_i \right)^2 \wedge \frac{1}{N_1} \right) \\
&+ \tilde{\mathcal{O}}\left(\left(\max_{r \in \mathcal{R}} \max_{k=0}^{d(r)} \max_{\substack{S: \mathcal{B}^*(r) \subseteq \mathcal{S} \subseteq \mathcal{B}^+(r) \\ |\mathcal{S}| = n_*(r) + k}} \frac{k \sum_{j \in \mathcal{S} \setminus \mathcal{B}^*(r)} N_j}{\left(\sum_{j \in \mathcal{S}} N_j\right)^2} \right) \wedge \frac{1}{N_1} \right) \\
&+ \tilde{\mathcal{O}}\left(\left(\frac{1}{\min_{r \in \mathcal{R}} \sum_{j \in \mathcal{B}^*(r)} N_j} \right) \wedge \frac{1}{N_1} \right) \\
&+ \tilde{\mathcal{O}}\left(\frac{1}{KN_1} \right),
\end{aligned}$$

where $d(r) = \left|\mathcal{B}^+(r)\right| - \left|\mathcal{B}^*(r)\right|$ *and* $n_*(r) = \left|\mathcal{B}^*(r)\right|$.

Proof. We have

$$\begin{aligned}
&\mathrm{MSE}(\hat{\mu}^{\mathrm{HAVER}}) \\
&= \mathbb{E}\left[\left(\hat{\mu}^{\mathcal{B}} - \mu_1 \right)^2 \right]
\end{aligned}$$

$$
\begin{aligned}
&= \int_0^\infty \mathbb{P}\left(\left(\hat{\mu}^{\mathcal{B}} - \mu_1\right)^2 > \varepsilon\right) \mathrm{d}\varepsilon \\
&= \int_0^\infty \mathbb{P}\left(\left|\hat{\mu}^{\mathcal{B}} - \mu_1\right| > \varepsilon\right) 2\varepsilon \,\mathrm{d}\varepsilon && \text{(change of variable)} \\
&= \int_0^\infty \mathbb{P}\left(\left|\hat{\mu}^{\mathcal{B}} - \mu_1\right| > \varepsilon,\, \exists i, |\hat{\mu}_i - \mu_i| \ge \frac{1}{3}\gamma_i\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&\quad + \int_0^\infty \mathbb{P}\left(\left|\hat{\mu}^{\mathcal{B}} - \mu_1\right| > \varepsilon,\, \forall i, |\hat{\mu}_i - \mu_i| < \frac{1}{3}\gamma_i\right) 2\varepsilon \,\mathrm{d}\varepsilon.
\end{aligned}
$$

For the first term, we use Lemma 33 to obtain

$$
\int_0^\infty \mathbb{P}\left(\left|\hat{\mu}^{\mathcal{B}} - \mu_1\right| > \varepsilon,\, \exists i, |\hat{\mu}_i - \mu_i| \ge \frac{1}{3}\gamma_i\right) 2\varepsilon \,\mathrm{d}\varepsilon = \tilde{\mathcal{O}}\left(\frac{1}{KN_1}\right).
$$

For the second term, we claim that the condition

$$
\forall i, |\hat{\mu}_i - \mu_i| < \frac{1}{3}\gamma_i \tag{2}
$$

implies that $\hat{r} \in \mathcal{R}$. To see this, by the definition of $\hat{r}$, we have

$$
\begin{aligned}
&\hat{\mu}_{\hat{r}} - \hat{\mu}_s \ge \gamma_{\hat{r}} - \gamma_s \\
&\Rightarrow \mu_{\hat{r}} + \frac{1}{3}\gamma_{\hat{r}} - \mu_s + \frac{1}{3}\gamma_s \ge \gamma_{\hat{r}} - \gamma_s && \text{(use the condition (2))} \\
&\Leftrightarrow \mu_{\hat{r}} \ge \mu_s + \frac{2}{3}\gamma_{\hat{r}} - \frac{4}{3}\gamma_s.
\end{aligned}
$$

In addition, by definition of s, we have

$$
\mu_{\hat{r}} \le \mu_s - \frac{1}{6}\gamma_s + \frac{1}{6}\gamma_{\hat{r}}.
$$

These conditions imply that $\hat{r} \in \mathcal{R} = \left\{r \in [K] : \mu_s - \frac{4}{3}\gamma_s + \frac{2}{3}\gamma_r \le \mu_r \le \mu_s - \frac{1}{6}\gamma_s + \frac{1}{6}\gamma_r\right\}$.

With $\hat{r} \in \mathcal{R}$, we decompose the second term

$$
\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\left|\hat{\mu}^{\mathcal{B}} - \mu_1\right| < \varepsilon, \hat{r} \in \mathcal{R}\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&= \int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, \hat{r} \in \mathcal{R}\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&\quad + \int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon,\, \hat{r} \in \mathcal{R}\right) 2\varepsilon \,\mathrm{d}\varepsilon.
\end{aligned}
$$

For the first subterm, we use Lemma 13 to obtain

$$
\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, \hat{r} \in \mathcal{R}\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&= \tilde{\mathcal{O}}\left(\left(\max_{r \in \mathcal{R}} \frac{1}{\sum_{j \in \mathcal{B}^*(r)} N_j} \sum_{i \in \mathcal{B}^+(r)} N_i \Delta_i\right)^2 \wedge \frac{1}{N_1}\right) \\
&\quad + \tilde{\mathcal{O}}\left(\left(\max_{r \in \mathcal{R}} \max_{k=0}^{d(r)} \max_{\substack{S : \mathcal{B}^*(r) \subseteq \mathcal{S} \subseteq \mathcal{B}^+(r) \\ |\mathcal{S}| = n_*(r) + k}} \frac{k \sum_{j \in \mathcal{S} \setminus \mathcal{B}^*(r)} N_j}{\left(\sum_{j \in \mathcal{S}} N_j\right)^2}\right) \wedge \frac{1}{N_1}\right) \\
&\quad + \tilde{\mathcal{O}}\left(\left(\frac{1}{\min_{r \in \mathcal{R}} \sum_{j \in \mathcal{B}^*(r)} N_j}\right) \wedge \frac{1}{N_1}\right) \\
&\quad + \tilde{\mathcal{O}}\left(\frac{1}{KN_1}\right).
\end{aligned}
$$

For the second subterm, we use Lemma 19 to obtain

$$
\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon,\, \hat{r} \in \mathcal{R}\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
=&\tilde{\mathcal{O}}\left(\left(\max_{r\in\mathcal{R}} \max_{k=0}^{d(r)} \max_{\substack{\mathcal{S}:\mathcal{B}^*(r)\subseteq\mathcal{S}\subseteq\mathcal{B}^+(r)\\ |\mathcal{S}|=n_*(r)+k}} \frac{k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}{\left(\sum_{j\in\mathcal{S}} N_j\right)^2}\right) \wedge \frac{1}{N_1}\right) \\
&+\tilde{\mathcal{O}}\left(\left(\frac{1}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}\right) \wedge \frac{1}{N_1}\right) \\
&+\tilde{\mathcal{O}}\left(\frac{1}{KN_1}\right).
\end{aligned}
$$

Combining all the results completes our proof. $\square$

Lemma 13. HAVER achieves

$$
\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, \hat{r} \in \mathcal{R}\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
=&\tilde{\mathcal{O}}\left(\left(\max_{r\in\mathcal{R}} \frac{1}{\sum_{j\in\mathcal{B}^*(r)} N_j} \sum_{i\in\mathcal{B}^+(r)} N_i\Delta_i\right)^2 \wedge \frac{1}{N_1}\right) \\
&+\tilde{\mathcal{O}}\left(\left(\max_{r\in\mathcal{R}} \max_{k=0}^{d(r)} \max_{\substack{\mathcal{S}:\mathcal{B}^*(r)\subseteq\mathcal{S}\subseteq\mathcal{B}^+(r)\\ |\mathcal{S}|=n_*(r)+k}} \frac{k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}{\left(\sum_{j\in\mathcal{S}} N_j\right)^2}\right) \wedge \frac{1}{N_1}\right) \\
&+\tilde{\mathcal{O}}\left(\left(\frac{1}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}\right) \wedge \frac{1}{N_1}\right) \\
&+\tilde{\mathcal{O}}\left(\frac{1}{KN_1}\right),
\end{aligned}
$$

where $d(r) = \left|\mathcal{B}^+(r)\right| - \left|\mathcal{B}^*(r)\right|$ and $n_*(r) = \left|\mathcal{B}^*(r)\right|$.

Proof. We consider 4 complementary events:

- $G_0 = \left\{\mathcal{B} = \mathcal{B}^*(\hat{r})\right\}$
- $G_1 = \left\{\mathcal{B}^*(\hat{r}) \subset \mathcal{B} \subseteq \mathcal{B}^+(\hat{r})\right\}$
- $G_2 = \left\{\exists i \in \mathcal{B}^*(\hat{r}) \,\text{s.t.}\, i \notin \mathcal{B}\right\}$
- $G_3 = \left\{\mathcal{B}^*(\hat{r}) \subseteq \mathcal{B},\, \exists i \notin \mathcal{B}^+(\hat{r}) \,\text{s.t.}\, i \in \mathcal{B}\right\}$

We have

$$
\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, \hat{r} \in \mathcal{R}\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
=&\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, \hat{r} \in \mathcal{R},\, G_0\right) 2\varepsilon \,\mathrm{d}\varepsilon + \int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, \hat{r} \in \mathcal{R},\, G_1\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&+\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, \hat{r} \in \mathcal{R},\, G_2\right) 2\varepsilon \,\mathrm{d}\varepsilon + \int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, \hat{r} \in \mathcal{R},\, G_3\right) 2\varepsilon \,\mathrm{d}\varepsilon.
\end{aligned}
$$

Using Lemma 14 and 15 for event G_0 and G_1 respectively, we have

$$
\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, \hat{r} \in \mathcal{R},\, G_0\right) 2\varepsilon \,\mathrm{d}\varepsilon
$$

$$= \mathcal{O}\left(\left(\max_{r\in\mathcal{R}} \frac{1}{\sum_{j\in\mathcal{B}^*(r)} N_j} \sum_{i\in\mathcal{B}^*(r)} N_i\Delta_i\right)^2 + \frac{\log(K)}{\min_{r\in\mathcal{R}} \sum_{j\in\mathcal{B}^*(r)} N_j}\right)$$

and

$$\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, \hat{r}\in\mathcal{R},\, G_1\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&= \mathcal{O}\left(\left(\max_{r\in\mathcal{R}} \max_{S:\mathcal{B}^*(r)\subseteq\mathcal{S}\subseteq\mathcal{B}^+(r)} \frac{1}{\sum_{j\in\mathcal{S}} N_j} \sum_{i\in\mathcal{S}} N_i\Delta_i\right)^2\right) \\
&\quad + \mathcal{O}\left(\max_{r\in\mathcal{R}} \max_{k=0}^{d(r)} \max_{\substack{S:\mathcal{B}^*(r)\subseteq\mathcal{S}\subseteq\mathcal{B}^+(r)\\ |\mathcal{S}|=n_*(r)+k}} \frac{k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}{\left(\sum_{j\in\mathcal{S}} N_j\right)^2} \log\left(\frac{ed(r)}{k}\right)\log(K)\right) \\
&\quad + \mathcal{O}\left(\frac{\log(K)}{\min_{r\in\mathcal{R}} \sum_{j\in\mathcal{B}^*(r)} N_j}\right).
\end{aligned}$$

Combining the first two events, we have

$$\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, \hat{r}\in\mathcal{R},\, G_0\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&+ \int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, \hat{r}\in\mathcal{R},\, G_1\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&= \mathcal{O}\left(\left(\max_{r\in\mathcal{R}} \max_{S:\mathcal{B}^*(r)\subseteq\mathcal{S}\subseteq\mathcal{B}^+(r)} \frac{1}{\sum_{j\in\mathcal{S}} N_j} \sum_{i\in\mathcal{S}} N_i\Delta_i\right)^2\right) \\
&\quad + \mathcal{O}\left(\max_{r\in\mathcal{R}} \max_{k=0}^{d(r)} \max_{\substack{S:\mathcal{B}^*(r)\subseteq\mathcal{S}\subseteq\mathcal{B}^+(r)\\ |\mathcal{S}|=n_*(r)+k}} \frac{k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}{\left(\sum_{j\in\mathcal{S}} N_j\right)^2} \log\left(\frac{ed(r)}{k}\right)\log(K)\right) \\
&\quad + \mathcal{O}\left(\frac{\log(K)}{\min_{r\in\mathcal{R}} \sum_{j\in\mathcal{B}^*(r)} N_j}\right) \\
&= \mathcal{O}\left(\left(\max_{r\in\mathcal{R}} \frac{1}{\sum_{j\in\mathcal{B}^*(r)} N_j} \sum_{i\in\mathcal{B}^+(r)} N_i\Delta_i\right)^2\right) \\
&\quad + \mathcal{O}\left(\max_{r\in\mathcal{R}} \max_{k=0}^{d(r)} \max_{\substack{S:\mathcal{B}^*(r)\subseteq\mathcal{S}\subseteq\mathcal{B}^+(r)\\ |\mathcal{S}|=n_*(r)+k}} \frac{k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}{\left(\sum_{j\in\mathcal{S}} N_j\right)^2} \log(3K)\log(K)\right) \qquad \left(\text{use } \tfrac{ed(r)}{k} \le 3K\right) \\
&\quad + \mathcal{O}\left(\frac{\log(K)}{\min_{r\in\mathcal{R}} \sum_{j\in\mathcal{B}^*(r)} N_j}\right) \\
&= \tilde{\mathcal{O}}\left(\left(\max_{r\in\mathcal{R}} \frac{1}{\sum_{j\in\mathcal{B}^*(r)} N_j} \sum_{i\in\mathcal{B}^+(r)} N_i\Delta_i\right)^2\right) \\
&\quad + \tilde{\mathcal{O}}\left(\max_{r\in\mathcal{R}} \max_{k=0}^{d(r)} \max_{\substack{S:\mathcal{B}^*(r)\subseteq\mathcal{S}\subseteq\mathcal{B}^+(r)\\ |\mathcal{S}|=n_*(r)+k}} \frac{k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}{\left(\sum_{j\in\mathcal{S}} N_j\right)^2}\right)
\end{aligned}$$

$$+ \tilde{\mathcal{O}}\left(\frac{1}{\min_{r\in\mathcal{R}} \sum_{j\in\mathcal{B}^*(r)} N_j}\right).$$

We use Lemma 28 and 29 to show that $\mathbb{P}(G_2) \le \frac{2K^2}{K^{2\lambda}}$ and $\mathbb{P}(G_3) \le \frac{2K^2}{K^{2\lambda}}$ respectively.

Consequently, we use Lemma 18 for event G_2 to obtain the integral

$$\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, \hat{r}\in\mathcal{R},\, G_2\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&= \mathcal{O}\left(\min_m \left(\frac{1}{K^2 N_m}\log\left(\frac{KS}{N_m}\right) + \frac{1}{K^2}\Delta_m^2\right)\right) \\
&\le \mathcal{O}\left(\left(\frac{1}{K^2 N_1}\log\left(\frac{KS}{N_1}\right)\right)\right) && \text{(upper bounded by } m=1) \\
&\le \tilde{\mathcal{O}}\left(\frac{1}{KN_1}\right).
\end{aligned}$$

We use Lemma 18 for event G_3 and get the same result. Furthermore, Lemma 17 states that

$$\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&= \mathcal{O}\left(\min_m \left(\frac{1}{N_m}\log\left(\frac{KS}{N_m}\right) + \Delta_m^2\right)\right) \\
&= \mathcal{O}\left(\frac{1}{N_1}\log\left(\frac{KS}{N_1}\right)\right) && \text{(upper bounded by } m=1) \\
&= \tilde{\mathcal{O}}\left(\frac{1}{N_1}\right).
\end{aligned}$$

Combining all the results completes our proof.

□

Lemma 14. In the event of $G_0 = \{\mathcal{B} = \mathcal{B}^*(\hat{r})\}$, HAVER achieves

$$\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, \hat{r}\in\mathcal{R},\, G_0\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&\le 4\left(\max_{r\in\mathcal{R}} \frac{1}{\sum_{j\in\mathcal{B}^*(r)} N_j} \sum_{i\in\mathcal{B}^*(r)} N_i\Delta_i\right)^2 + 48\left(\frac{\log(K)}{\min_{r\in\mathcal{R}} \sum_{j\in\mathcal{B}^*(r)} N_j}\right).
\end{aligned}$$

Proof. We have

$$\begin{aligned}
&\mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, \hat{r}\in\mathcal{R},\, G_0\right) \\
&= \mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j}\sum_{i\in\mathcal{B}} N_i\hat{\mu}_i - \mu_1 < -\varepsilon,\, \hat{r}\in\mathcal{R},\, G_0\right) \\
&= \mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j}\sum_{i\in\mathcal{B}} N_i\left(\hat{\mu}_i - \mu_1\right) < -\varepsilon,\, \hat{r}\in\mathcal{R},\, G_0\right) \\
&= \mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j}\sum_{i\in\mathcal{B}} N_i\left(\hat{\mu}_i - \mu_i\right) < -\varepsilon + \frac{1}{\sum_{j\in\mathcal{B}} N_j}\sum_{i\in\mathcal{B}} N_i\Delta_i,\, \hat{r}\in\mathcal{R},\, G_0\right)
\end{aligned}$$

$$= \mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}^*(\hat{r})} N_j} \sum_{i\in\mathcal{B}^*(\hat{r})} N_i\left(\hat{\mu}_i - \mu_i\right) < -\varepsilon + \frac{1}{\sum_{j\in\mathcal{B}^*(\hat{r})} N_j} \sum_{i\in\mathcal{B}^*(\hat{r})} N_i\Delta_i,\, \hat{r}\in\mathcal{R},\, G_0\right)$$

$$\text{(use } G_0 = \left\{\mathcal{B} = \mathcal{B}^*(\hat{r})\right\}\text{)}$$

$$\le \sum_{r\in\mathcal{R}} \mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}^*(r)} N_j} \sum_{i\in\mathcal{B}^*(r)} N_i\left(\hat{\mu}_i - \mu_i\right) < -\varepsilon + \frac{1}{\sum_{j\in\mathcal{B}^*(r)} N_j} \sum_{i\in\mathcal{B}^*(r)} N_i\Delta_i\right)$$

$$\le \sum_{r\in\mathcal{R}} \exp\left(-\frac{1}{2}\left(\sum_{j\in\mathcal{B}^*(r)} N_j\right)\left(\varepsilon - \frac{1}{\sum_{j\in\mathcal{B}^*(r)} N_j}\sum_{i\in\mathcal{B}^*(r)} N_i\Delta_i\right)_+^2\right)$$

$$\le K\exp\left(-\frac{1}{2}\left(\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j\right)\left(\varepsilon - \max_{r\in\mathcal{R}}\frac{1}{\sum_{j\in\mathcal{B}^*(r)} N_j}\sum_{i\in\mathcal{B}^*(r)} N_i\Delta_i\right)_+^2\right).$$

Let $q_1 = q_2 = 1$, $z_1 = \frac{1}{16}\left(\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j\right)$, and

$$\varepsilon_0 = \left(2\max_{r\in\mathcal{R}}\frac{1}{\sum_{j\in\mathcal{B}^*(r)} N_j}\sum_{i\in\mathcal{B}^*(r)} N_i\Delta_i\right) \vee \left(\sqrt{\frac{16\log(K)}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}}\right).$$

We claim that, in the regime of $\varepsilon \ge \varepsilon_0$, we have

$$\mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, \hat{r}\in\mathcal{R},\, G_0\right) \le q_1\exp\left(-z_1\varepsilon^2\right)$$

and in the regime of $\varepsilon < \varepsilon_0$, we have $\mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, \hat{r}\in\mathcal{R},\, G_0\right) \le q_2$.

The second claim is trivial. We prove the first claim as follows. In the regime of $\varepsilon \ge \varepsilon_0$, we have

$$\mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, \hat{r}\in\mathcal{R},\, G_0\right)$$

$$\le K\exp\left(-\frac{1}{2}\left(\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j\right)\left(\varepsilon - \max_{r\in\mathcal{R}}\frac{1}{\sum_{j\in\mathcal{B}^*(r)} N_j}\sum_{i\in\mathcal{B}^*(r)} N_i\Delta_i\right)_+^2\right)$$

$$\le K\exp\left(-\frac{1}{2}\left(\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j\right)\left(\frac{\varepsilon}{2}\right)^2\right) \qquad \text{(use } \varepsilon \ge \varepsilon_0\text{)}$$

$$= K\exp\left(-\frac{1}{8}\left(\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j\right)\varepsilon^2\right)$$

$$= \exp\left(-\frac{1}{8}\left(\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j\right)\varepsilon^2 + \log(K)\right)$$

$$\le \exp\left(-\frac{1}{16}\left(\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j\right)\varepsilon^2\right). \qquad \text{(use } \varepsilon \ge \varepsilon_0\text{)}$$

With the above claim, we use Lemma 31 to bound the integral

$$\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, \hat{r}\in\mathcal{R},\, G_0\right) 2\varepsilon\,\mathrm{d}\varepsilon$$

$$\le q_1\varepsilon_0^2 + q_2\frac{1}{z_1}$$

$$
\begin{aligned}
&= \left(\left(2 \max_{r\in\mathcal{R}} \frac{1}{\sum_{j\in\mathcal{B}^*(r)} N_j} \sum_{i\in\mathcal{B}^*(r)} N_i\Delta_i \right) \vee \left(\sqrt{\frac{16\log(K)}{\min_{r\in\mathcal{R}} \sum_{j\in\mathcal{B}^*(r)} N_j}} \right) \right)^2 + 16 \left(\frac{1}{\min_{r\in\mathcal{R}} \sum_{j\in\mathcal{B}^*(r)} N_j} \right) \\
&\le 4 \left(\max_{r\in\mathcal{R}} \frac{1}{\sum_{j\in\mathcal{B}^*(r)} N_j} \sum_{i\in\mathcal{B}^*(r)} N_i\Delta_i \right)^2 + 16 \left(\frac{\log(K)}{\min_{r\in\mathcal{R}} \sum_{j\in\mathcal{B}^*(r)} N_j} \right) + 16 \left(\frac{1}{\min_{r\in\mathcal{R}} \sum_{j\in\mathcal{B}^*(r)} N_j} \right) \\
&= 4 \left(\max_{r\in\mathcal{R}} \frac{1}{\sum_{j\in\mathcal{B}^*(r)} N_j} \sum_{i\in\mathcal{B}^*(r)} N_i\Delta_i \right)^2 + 8 \left(\frac{\log(K^2)}{\min_{r\in\mathcal{R}} \sum_{j\in\mathcal{B}^*(r)} N_j} \right) + 16 \left(\frac{1}{\min_{r\in\mathcal{R}} \sum_{j\in\mathcal{B}^*(r)} N_j} \right) \\
&\le 4 \left(\max_{r\in\mathcal{R}} \frac{1}{\sum_{j\in\mathcal{B}^*(r)} N_j} \sum_{i\in\mathcal{B}^*(r)} N_i\Delta_i \right)^2 + 24 \left(\frac{\log(K^2)}{\min_{r\in\mathcal{R}} \sum_{j\in\mathcal{B}^*(r)} N_j} \right) \qquad \text{(use } 1 \le \log(K^2) \text{ with } K \ge 2\text{)} \\
&\le 4 \left(\max_{r\in\mathcal{R}} \frac{1}{\sum_{j\in\mathcal{B}^*(r)} N_j} \sum_{i\in\mathcal{B}^*(r)} N_i\Delta_i \right)^2 + 48 \left(\frac{\log(K)}{\min_{r\in\mathcal{R}} \sum_{j\in\mathcal{B}^*(r)} N_j} \right).
\end{aligned}
$$

□

Lemma 15. In the event of $G_1 = \left\{\mathcal{B}^*(\hat{r}) \subset \mathcal{B} \subseteq \mathcal{B}^+(\hat{r})\right\}$, HAVER achieves

$$
\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, \hat{r} \in \mathcal{R},\, G_1\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&\le 64 \left(\max_{r\in\mathcal{R}} \max_{S:\mathcal{B}^*(r)\subseteq\mathcal{S}\subseteq\mathcal{B}^+(r)} \frac{1}{\sum_{j\in\mathcal{S}} N_j} \sum_{i\in\mathcal{S}} N_i\Delta_i \right)^2 \\
&\quad + 512 \left(\max_{r\in\mathcal{R}} \max_{k=0}^{d(r)} \max_{\substack{S:\mathcal{B}^*(r)\subseteq\mathcal{S}\subseteq\mathcal{B}^+(r) \\ |\mathcal{S}|=n_*(r)+k}} \frac{k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}{\left(\sum_{j\in\mathcal{S}} N_j\right)^2} \log\left(\frac{ed(r)}{k}\right) \log(K) \right) \\
&\quad + 192 \left(\frac{\log(K)}{\min_{r\in\mathcal{R}} \sum_{j\in\mathcal{B}^*(r)} N_j} \right),
\end{aligned}
$$

where $d(r) = \left|\mathcal{B}^+(r)\right| - \left|\mathcal{B}^*(r)\right|$ and $n_*(r) = \left|\mathcal{B}^*(r)\right|$.

Proof. We decompose the integral as follows

$$
\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, \hat{r} \in \mathcal{R},\, G_1\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&= \int_0^\infty \mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j} \sum_{i\in\mathcal{B}} N_i \left(\hat{\mu}_i - \mu_1\right) < -\varepsilon,\, \hat{r} \in \mathcal{R},\, G_1 \right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&\le \int_0^\infty \mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j} \sum_{i\in\mathcal{B}^*(\hat{r})} N_i \left(\hat{\mu}_i - \mu_1\right) < -\frac{\varepsilon}{2},\, \hat{r} \in \mathcal{R},\, G_1 \right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&\quad + \int_0^\infty \mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j} \sum_{i\in\mathcal{B}\setminus\mathcal{B}^*(\hat{r})} N_i \left(\hat{\mu}_i - \mu_1\right) < -\frac{\varepsilon}{2},\, \hat{r} \in \mathcal{R},\, G_1 \right) 2\varepsilon \,\mathrm{d}\varepsilon.
\end{aligned}
$$

We bound the probability and integral of these terms respectively. For the first term, we have

$$
\mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j} \sum_{i\in\mathcal{B}^*(\hat{r})} N_i \left(\hat{\mu}_i - \mu_1\right) < -\frac{\varepsilon}{2},\, \hat{r} \in \mathcal{R},\, G_1 \right)
$$

$$
\begin{aligned}
&= \mathbb{P}\left(\sum_{i\in\mathcal{B}^*(\hat r)} N_i\left(\hat\mu_i-\mu_1\right) < -\frac{\varepsilon}{2}\sum_{j\in\mathcal{B}} N_j,\, \hat r\in\mathcal{R},\, G_1\right)\\
&\le \mathbb{P}\left(\sum_{i\in\mathcal{B}^*(\hat r)} N_i\left(\hat\mu_i-\mu_1\right) < -\frac{\varepsilon}{2}\sum_{j\in\mathcal{B}^*(\hat r)} N_j,\, \hat r\in\mathcal{R},\, G_1\right) \qquad \left(\text{since } \mathcal{B}^*(\hat r)\subset\mathcal{B},\, \textstyle\sum_{j\in\mathcal{B}^*(\hat r)} N_j < \sum_{j\in\mathcal{B}} N_j\right)\\
&= \mathbb{P}\left(\sum_{i\in\mathcal{B}^*(\hat r)} N_i\left(\hat\mu_i-\mu_i\right) < -\frac{\varepsilon}{2}\sum_{j\in\mathcal{B}^*(\hat r)} N_j + \sum_{i\in\mathcal{B}^*(\hat r)} N_i\Delta_i,\, \hat r\in\mathcal{R},\, G_1\right)\\
&\le \sum_{r\in\mathcal{R}}\mathbb{P}\left(\sum_{i\in\mathcal{B}^*(r)} N_i\left(\hat\mu_i-\mu_i\right) < -\frac{\varepsilon}{2}\sum_{j\in\mathcal{B}^*(r)} N_j + \sum_{i\in\mathcal{B}^*(r)} N_i\Delta_i\right)\\
&\le \sum_{r\in\mathcal{R}}\exp\left(-\frac{1}{2}\frac{1}{\sum_{j\in\mathcal{B}^*(r)} N_j}\left(\frac{\varepsilon}{2}\sum_{j\in\mathcal{B}^*(r)} N_j - \sum_{i\in\mathcal{B}^*(r)} N_i\Delta_i\right)_+^2\right)\\
&= \sum_{r\in\mathcal{R}}\exp\left(-\frac{1}{8}\left(\sum_{j\in\mathcal{B}^*(r)} N_j\right)\left(\varepsilon - \frac{2}{\sum_{j\in\mathcal{B}^*(r)} N_j}\sum_{i\in\mathcal{B}^*(r)} N_i\Delta_i\right)_+^2\right)\\
&\le K\exp\left(-\frac{1}{8}\left(\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j\right)\left(\varepsilon - \max_{r\in\mathcal{R}}\frac{2}{\sum_{j\in\mathcal{B}^*(r)} N_j}\sum_{i\in\mathcal{B}^*(r)} N_i\Delta_i\right)_+^2\right).
\end{aligned}
$$

For this term, let $q_1 = 1$, $q_2 = 1$, $z_1 = \frac{1}{32}\left(\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j\right)$, and

$$
\varepsilon_0 = \left(4\max_{r\in\mathcal{R}}\frac{1}{\sum_{j\in\mathcal{B}^*(r)} N_j}\sum_{i\in\mathcal{B}^*(r)} N_i\Delta_i\right)\vee\left(\sqrt{\frac{64\log(K)}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}}\right).
$$

We claim that, in the regime of $\varepsilon\ge\varepsilon_0$, we have

$$
\mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j}\sum_{i\in\mathcal{B}^*(r)} N_i\left(\hat\mu_i-\mu_1\right) < -\frac{\varepsilon}{2},\, \hat r\in\mathcal{R},\, G_1\right)\le q_1\exp\left(-z_1\varepsilon^2\right)
$$

and in the regime of $\varepsilon\ge\varepsilon_0$, we have

$$
\mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j}\sum_{i\in\mathcal{B}^*(\hat r)} N_i\left(\hat\mu_i-\mu_1\right) < -\frac{\varepsilon}{2},\, \hat r\in\mathcal{R},\, G_1\right)\le q_2.
$$

The second claim is trivial. We prove the first claim as follows. In the regime of $\varepsilon\ge\varepsilon_0$, we have

$$
\begin{aligned}
&\mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j}\sum_{i\in\mathcal{B}^*} N_i\left(\hat\mu_i-\mu_1\right) < -\frac{\varepsilon}{2},\, \hat r\in\mathcal{R},\, G_1\right)\\
&\le K\exp\left(-\frac{1}{8}\left(\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j\right)\left(\varepsilon - \max_{r\in\mathcal{R}}\frac{2}{\sum_{j\in\mathcal{B}^*(r)} N_j}\sum_{i\in\mathcal{B}^*} N_i\Delta_i\right)_+^2\right)\\
&\le K\exp\left(-\frac{1}{8}\left(\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j\right)\left(\frac{\varepsilon}{2}\right)^2\right) \qquad (\text{use } \varepsilon\ge\varepsilon_0)\\
&= K\exp\left(-\frac{1}{32}\left(\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j\right)\varepsilon^2\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \exp\left(-\frac{1}{32}\left(\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j\right)\varepsilon^2 + \log(K)\right) \\
&\le \exp\left(-\frac{1}{64}\left(\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j\right)\varepsilon^2\right). \qquad \text{(use } \varepsilon \ge \varepsilon_0\text{)}
\end{aligned}
$$

With the above claim, we use Lemma 31 to bound the integral

$$
\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j}\sum_{i\in\mathcal{B}^*} N_i\left(\hat{\mu}_i - \mu_1\right) < -\frac{\varepsilon}{2},\, \hat{r}\in\mathcal{R},\, G_1\right) 2\varepsilon\,\mathrm{d}\varepsilon \\
&\le q_1\varepsilon_0^2 + q_2\frac{1}{z_1} \\
&= \left(\left(4\max_{r\in\mathcal{R}}\frac{1}{\sum_{j\in\mathcal{B}^*(r)} N_j}\sum_{i\in\mathcal{B}^*(r)} N_i\Delta_i\right) \vee \left(\sqrt{\frac{64\log(K)}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}}\right)\right)^2 + 64\left(\frac{1}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}\right) \\
&\le 16\left(\max_{r\in\mathcal{R}}\frac{1}{\sum_{j\in\mathcal{B}^*(r)} N_j}\sum_{i\in\mathcal{B}^*(r)} N_i\Delta_i\right)^2 + 64\left(\frac{\log(K)}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}\right) + 64\left(\frac{1}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}\right) \\
&= 16\left(\max_{r\in\mathcal{R}}\frac{1}{\sum_{j\in\mathcal{B}^*(r)} N_j}\sum_{i\in\mathcal{B}^*(r)} N_i\Delta_i\right)^2 + 32\left(\frac{\log(K^2)}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}\right) + 64\left(\frac{1}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}\right) \\
&\le 16\left(\max_{r\in\mathcal{R}}\frac{1}{\sum_{j\in\mathcal{B}^*(r)} N_j}\sum_{i\in\mathcal{B}^*(r)} N_i\Delta_i\right)^2 + 96\left(\frac{\log(K^2)}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}\right) \qquad \text{(use } 1 \le \log(K^2) \text{ with } K\ge 2\text{)} \\
&= 16\left(\max_{r\in\mathcal{R}}\frac{1}{\sum_{j\in\mathcal{B}^*(r)} N_j}\sum_{i\in\mathcal{B}^*(r)} N_i\Delta_i\right)^2 + 192\left(\frac{\log(K)}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}\right).
\end{aligned}
$$

For the second term, we have

$$
\begin{aligned}
&\mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j}\sum_{i\in\mathcal{B}\setminus\mathcal{B}^*(\hat{r})} N_i\left(\hat{\mu}_i - \mu_1\right) < -\frac{\varepsilon}{2},\, \hat{r}\in\mathcal{R},\, G_1\right) \\
&= \mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j}\sum_{i\in\mathcal{B}\setminus\mathcal{B}^*(\hat{r})} N_i\left(\hat{\mu}_i - \mu_i\right) < -\frac{\varepsilon}{2} + \frac{1}{\sum_{j\in\mathcal{B}} N_j}\sum_{i\in\mathcal{B}\setminus\mathcal{B}^*(\hat{r})} N_i\Delta_i,\, \hat{r}\in\mathcal{R},\, G_1\right) \\
&\le \sum_{r\in\mathcal{R}}\sum_{k=1}^{d(r)}\sum_{\substack{\mathcal{S}:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\ |\mathcal{S}|=n_*(r)+k}} \mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{S}} N_j}\sum_{i\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_i\left(\hat{\mu}_i - \mu_i\right) < -\frac{\varepsilon}{2} + \frac{1}{\sum_{j\in\mathcal{S}} N_j}\sum_{i\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_i\Delta_i\right) \\
&\qquad \text{(use } d(r) = \left|\mathcal{B}^+(r)\right| - \left|\mathcal{B}^*(r)\right| \text{ and } n_*(r) = \left|\mathcal{B}^*(r)\right|\text{)} \\
&= \sum_{r\in\mathcal{R}}\sum_{k=1}^{d(r)}\sum_{\substack{\mathcal{S}:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\ |\mathcal{S}|=n_*(r)+k}} \exp\left(-\frac{1}{2}\left(\frac{\left(\sum_{j\in\mathcal{S}} N_j\right)^2}{\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}\right)\left(\frac{\varepsilon}{2} - \frac{1}{\sum_{j\in\mathcal{S}} N_j}\sum_{i\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_i\Delta_i\right)_+^2\right) \\
&= \sum_{r\in\mathcal{R}}\sum_{k=1}^{d(r)}\sum_{\substack{\mathcal{S}:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\ |\mathcal{S}|=n_*(r)+k}} \exp\left(-\frac{1}{8}\left(\frac{\left(\sum_{j\in\mathcal{S}} N_j\right)^2}{\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}\right)\left(\varepsilon - \frac{2}{\sum_{j\in\mathcal{S}} N_j}\sum_{i\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_i\Delta_i\right)_+^2\right).
\end{aligned}
$$

For brevity, we denote

$$M = \left(\min_{r\in\mathcal{R}} \min_{k=1}^{d(r)} \min_{\substack{S:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\ |\mathcal{S}|=n_*(r)+k}} \frac{\left(\sum_{j\in\mathcal{S}} N_j\right)^2}{\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j} \right)$$

and

$$H = \max_{r\in\mathcal{R}} \max_{S:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)} \frac{1}{\sum_{j\in\mathcal{S}} N_j} \sum_{i\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_i\Delta_i.$$

For $\forall r \in \mathcal{R}$,$\forall k \in \{1,\dots,d(r)\}$, and $\forall S$, $\mathcal{B}^*(r) \subset \mathcal{S} \subseteq \mathcal{B}^+(r), |\mathcal{S}| = n_*(r)+k$, we have

$$\left(\frac{\left(\sum_{j\in\mathcal{S}} N_j\right)^2}{\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j} \right) \geq M$$

and

$$\frac{1}{\sum_{j\in\mathcal{S}} N_j} \sum_{i\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_i\Delta_i \leq H.$$

Thus, we have

$$\begin{aligned}
&\mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j} \sum_{i\in\mathcal{B}\setminus\mathcal{B}^*(\hat{r})} N_i\left(\hat{\mu}_i - \mu_1\right) < -\frac{\varepsilon}{2},\, \hat{r}\in\mathcal{R},\, G_1 \right)\\
&\leq \sum_{r\in\mathcal{R}} \sum_{k=1}^{d(r)} \sum_{\substack{S:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\ |\mathcal{S}|=n_*(r)+k}} \exp\left(-\frac{1}{8} \left(\frac{\left(\sum_{j\in\mathcal{S}} N_j\right)^2}{\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j} \right) \left(\varepsilon - \frac{2}{\sum_{j\in\mathcal{S}} N_j} \sum_{i\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_i\Delta_i \right)_+^2 \right)\\
&\leq \sum_{r\in\mathcal{R}} \sum_{k=1}^{d(r)} \sum_{\substack{S:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\ |\mathcal{S}|=n_*(r)+k}} \exp\left(-\frac{1}{8} M\left(\varepsilon - 2H\right)_+^2 \right)\\
&= \sum_{r\in\mathcal{R}} \sum_{k=1}^{d(r)} \binom{d(r)}{k} \exp\left(-\frac{1}{8} M\left(\varepsilon - 2H\right)_+^2 \right).
\end{aligned}$$

For this term, let $q_1 = q_2 = 1$, $z_1 = \frac{M}{128}$, and

$$\varepsilon_0 = (4H) \vee \left(\sqrt{ \max_{r\in\mathcal{R}} \max_{k=1}^{d(r)} \max_{\substack{S:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\ |\mathcal{S}|=n_*(r)+k}} \frac{128k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}{\left(\sum_{j\in\mathcal{S}} N_j\right)^2} \log\left(\frac{ed(r)}{k}\right) \log(K^2) } \right).$$

We claim that, in the regime of $\varepsilon \geq \varepsilon_0$, we have

$$\mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j} \sum_{i\in\mathcal{B}\setminus\mathcal{B}^*(r)} N_i\left(\hat{\mu}_i - \mu_1\right) < -\frac{\varepsilon}{2},\, \hat{r}\in\mathcal{R},\, G_1 \right) \leq q_1 \exp\left(-z_1\varepsilon^2\right)$$

and in the regime of $\varepsilon < \varepsilon_0$, we have

$$\mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j} \sum_{i\in\mathcal{B}\setminus\mathcal{B}^*(r)} N_i\left(\hat{\mu}_i - \mu_1\right) < -\frac{\varepsilon}{2},\, \hat{r}\in\mathcal{R},\, G_1 \right) \leq q_2.$$

The second claim is trivial. We prove the first claim as follows. In the regime of $\varepsilon \geq \varepsilon_0$, we have

$$\mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j} \sum_{i\in\mathcal{B}\setminus\mathcal{B}^*(r)} N_i\left(\hat{\mu}_i - \mu_1\right) < -\frac{\varepsilon}{2},\, \hat{r}\in\mathcal{R},\, G_1 \right)$$

$$
\begin{aligned}
&\leq \sum_{r\in\mathcal{R}} \sum_{k=1}^{d(r)} \binom{d(r)}{k} \exp\left(-\frac{1}{8}M\left(\varepsilon-2H\right)_+^2\right) \\
&\leq \sum_{r\in\mathcal{R}} \sum_{k=1}^{d(r)} \binom{d(r)}{k} \exp\left(-\frac{1}{8}M\left(\frac{\varepsilon}{2}\right)^2\right) && \text{(use } \varepsilon \geq \varepsilon_0\text{)} \\
&= \sum_{r\in\mathcal{R}} \sum_{k=1}^{d(r)} \binom{d(r)}{k} \exp\left(-\frac{1}{32}M\varepsilon^2\right) \\
&\leq \sum_{r\in\mathcal{R}} \sum_{k=1}^{d(r)} \left(\frac{ed(r)}{k}\right)^k \exp\left(-\frac{1}{32}M\varepsilon^2\right) && \text{(use Stirling's formula, Lemma 46)} \\
&= \sum_{r\in\mathcal{R}} \sum_{k=1}^{d(r)} \exp\left(-\frac{1}{32}M\varepsilon^2 + k\log\left(\frac{ed(r)}{k}\right)\right) \\
&\leq \sum_{r\in\mathcal{R}} \sum_{k=1}^{d(r)} \exp\left(-\frac{1}{64}M\varepsilon^2\right) && \text{(use } \varepsilon \geq \varepsilon_0\text{)} \\
&\leq Kd\exp\left(-\frac{1}{64}M\varepsilon^2\right) \\
&= \exp\left(-\frac{1}{64}M\varepsilon^2 + \ln(Kd)\right) \\
&\leq \exp\left(-\frac{1}{128}M\varepsilon^2\right). && \text{(use } \varepsilon \geq \varepsilon_0\text{)}
\end{aligned}
$$

With the above claim, we use Lemma 31 to bound the integral

$$
\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j} \sum_{i\in\mathcal{B}\setminus\mathcal{B}^*(r)} N_i\left(\hat{\mu}_i-\mu_1\right) < -\frac{\varepsilon}{2},\, \hat{r}\in\mathcal{R},\, G_1\right) 2\varepsilon\,\mathrm{d}\varepsilon \\
&\leq q_1\varepsilon_0^2 + q_2\frac{1}{z_1} \\
&\leq \left((4H) \vee \left(\sqrt{\max_{r\in\mathcal{R}} \max_{k=1}^{d(r)} \max_{\substack{S:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\ |\mathcal{S}|=n_*(r)+k}} \frac{128k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}{\left(\sum_{j\in\mathcal{S}} N_j\right)^2} \log\left(\frac{ed(r)}{k}\right)\log(K^2)}\right)\right)^2 + \frac{128}{M} \\
&\leq 16H^2 + 128\left(\max_{r\in\mathcal{R}} \max_{k=1}^{d(r)} \max_{\substack{S:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\ |\mathcal{S}|=n_*(r)+k}} \frac{k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}{\left(\sum_{j\in\mathcal{S}} N_j\right)^2} \log\left(\frac{ed(r)}{k}\right)\log(K^2)\right) + \frac{128}{M} \\
& && \text{(use } (A\vee B)^2 \leq A^2+B^2\text{)} \\
&= 16H^2 + 128\left(\max_{r\in\mathcal{R}} \max_{k=1}^{d(r)} \max_{\substack{S:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\ |\mathcal{S}|=n_*(r)+k}} \frac{k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}{\left(\sum_{j\in\mathcal{S}} N_j\right)^2} \log\left(\frac{ed(r)}{k}\right)\log(K^2)\right) \\
&\quad + 128\left(\max_{r\in\mathcal{R}} \max_{k=1}^{d(r)} \max_{\substack{S:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\ |\mathcal{S}|=n_*(r)+k}} \frac{128\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}{\left(\sum_{j\in\mathcal{S}} N_j\right)^2}\right) && \text{(plug in } M\text{)}
\end{aligned}
$$

$$= 16H^2 + 256\left(\max_{r\in\mathcal{R}}\max_{k=1}^{d(r)}\max_{\substack{S:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\|\mathcal{S}|=n_*(r)+k}}\frac{k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)}N_j}{\left(\sum_{j\in\mathcal{S}}N_j\right)^2}\log\left(\frac{ed(r)}{k}\right)\log(K^2)\right)$$

(use $1\le\log\left(\frac{ed(r)}{k}\right)\,\forall k\in\{1,\dots,d\}$ and $1\le\log(K^2)\,\forall K\ge 2$)

$$= 16H^2 + 512\left(\max_{r\in\mathcal{R}}\max_{k=1}^{d(r)}\max_{\substack{S:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\|\mathcal{S}|=n_*(r)+k}}\frac{k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)}N_j}{\left(\sum_{j\in\mathcal{S}}N_j\right)^2}\log\left(\frac{ed(r)}{k}\right)\log(K)\right)$$

$$= 16\left(\max_{r\in\mathcal{R}}\max_{S:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)}\frac{1}{\sum_{j\in\mathcal{S}}N_j}\sum_{i\in\mathcal{S}}N_i\Delta_i\right)^2 \qquad \text{(plug in } H\text{)}$$

$$+ 512\left(\max_{r\in\mathcal{R}}\max_{k=1}^{d(r)}\max_{\substack{S:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\|\mathcal{S}|=n_*(r)+k}}\frac{k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)}N_j}{\left(\sum_{j\in\mathcal{S}}N_j\right)^2}\log\left(\frac{ed(r)}{k}\right)\log(K)\right).$$

Combining the two terms, we have

$$\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}}-\mu_1<-\varepsilon,\,\hat{r}\in\mathcal{R},\,G_1\right)2\varepsilon\,\mathrm{d}\varepsilon$$

$$\le \int_0^\infty \mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}}N_j}\sum_{i\in\mathcal{B}^*(r)}N_i\left(\hat{\mu}_i-\mu_1\right)<-\frac{\varepsilon}{2},\,\hat{r}\in\mathcal{R},\,G_1\right)2\varepsilon\,\mathrm{d}\varepsilon$$

$$+\int_0^\infty \mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}}N_j}\sum_{i\in\mathcal{B}\setminus\mathcal{B}^*(r)}N_i\left(\hat{\mu}_i-\mu_1\right)<-\frac{\varepsilon}{2},\,\hat{r}\in\mathcal{R},\,G_1\right)2\varepsilon\,\mathrm{d}\varepsilon$$

$$\le 16\left(\max_{r\in\mathcal{R}}\frac{1}{\sum_{j\in\mathcal{B}^*(r)}N_j}\sum_{i\in\mathcal{B}^*(r)}N_i\Delta_i\right)^2+192\left(\frac{\log(K)}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)}N_j}\right)$$

$$+16\left(\max_{r\in\mathcal{R}}\max_{S:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)}\frac{1}{\sum_{j\in\mathcal{S}}N_j}\sum_{i\in\mathcal{S}}N_i\Delta_i\right)^2$$

$$+512\left(\max_{r\in\mathcal{R}}\max_{k=1}^{d(r)}\max_{\substack{S:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\|\mathcal{S}|=n_*(r)+k}}\frac{k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)}N_j}{\left(\sum_{j\in\mathcal{S}}N_j\right)^2}\log\left(\frac{ed(r)}{k}\right)\log(K)\right)$$

$$\le 16\left(\max_{r\in\mathcal{R}}\max_{S:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)}\frac{1}{\sum_{j\in\mathcal{S}}N_j}\sum_{i\in\mathcal{S}}N_i\Delta_i+\max_{r\in\mathcal{R}}\frac{1}{\sum_{j\in\mathcal{B}^*(r)}N_j}\sum_{i\in\mathcal{B}^*(r)}N_i\Delta_i\right)^2$$

(use $\forall a,b\ge 0,\ a^2+b^2\le(a+b)^2$)

$$+512\left(\max_{r\in\mathcal{R}}\max_{k=1}^{d(r)}\max_{\substack{S:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\|\mathcal{S}|=n_*(r)+k}}\frac{k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)}N_j}{\left(\sum_{j\in\mathcal{S}}N_j\right)^2}\log\left(\frac{ed(r)}{k}\right)\log(K)\right)$$

$$+192\left(\frac{\log(K)}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)(r)}N_j}\right)$$

$$\le 64\left(\max_{r\in\mathcal{R}}\max_{S:\mathcal{B}^*(r)\subseteq\mathcal{S}\subseteq\mathcal{B}^+(r)}\frac{1}{\sum_{j\in\mathcal{S}}N_j}\sum_{i\in\mathcal{S}}N_i\Delta_i\right)^2$$

$$
+ 512 \left(\max_{r \in \mathcal{R}} \max_{k=0}^{d(r)} \max_{\substack{\mathcal{S}: \mathcal{B}^*(r) \subseteq \mathcal{S} \subseteq \mathcal{B}^+(r) \\ |\mathcal{S}| = n_*(r) + k}} \frac{k \sum_{j \in \mathcal{S} \setminus \mathcal{B}^*(r)} N_j}{\left(\sum_{j \in \mathcal{S}} N_j \right)^2} \log \left(\frac{ed(r)}{k} \right) \log(K) \right)
$$

$$
+ 192 \left(\frac{\log(K)}{\min_{r \in \mathcal{R}} \sum_{j \in \mathcal{B}^*(r)} N_j} \right).
$$

□

Lemma 16. Let $m \in [K]$ be any arm. In HAVER, we have

$$
\mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon \right) \le 2 \exp\left(-\frac{1}{2} N_m \left(\varepsilon - \gamma_m - \Delta_m \right)_+^2 \right).
$$

Proof. We separate the probability into two terms

$$
\begin{aligned}
&\mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon \right) \\
&= \mathbb{P}\left(\hat{r} = m,\, \hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon \right) + \mathbb{P}\left(\hat{r} \in [K] \setminus \{m\},\, \hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon \right).
\end{aligned}
$$

For the first term, we have

$$
\begin{aligned}
&\mathbb{P}\left(\hat{r} = m,\, \hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon \right) \\
&= \mathbb{P}\left(\hat{r} = m,\, \frac{N_m}{\sum_{j \in \mathcal{B}} N_j} \hat{\mu}_m + \sum_{i \in \mathcal{B} \setminus \{m\}} \frac{N_i}{\sum_{j \in \mathcal{B}} N_j} \hat{\mu}_i - \mu_1 < -\varepsilon \right) \\
&\le \mathbb{P}\left(\hat{r} = m,\, \frac{N_m}{\sum_{j \in \mathcal{B}} N_j} \hat{\mu}_m + \sum_{i \in \mathcal{B} \setminus \{m\}} \frac{N_i}{\sum_{j \in \mathcal{B}} N_j} \left(\hat{\mu}_m - \gamma_m \right) - \mu_1 < -\varepsilon \right) && \text{(use } \forall i \in \mathcal{B},\, \hat{\mu}_m - \gamma_m \le \hat{\mu}_i) \\
&= \mathbb{P}\left(\hat{\mu}_m - \mu_1 < -\varepsilon + \sum_{i \in \mathcal{B} \setminus \{m\}} \frac{N_i}{\sum_{j \in \mathcal{B}} N_j} \gamma_m \right) \\
&\le \mathbb{P}\left(\hat{\mu}_m - \mu_1 < -\varepsilon + \gamma_m \right) && \text{(use } \textstyle\sum_{i \in \mathcal{B} \setminus \{m\}} \frac{N_i}{\sum_{j \in \mathcal{B}} N_j} \le 1) \\
&= \mathbb{P}\left(\hat{\mu}_m - \mu_m < -\varepsilon + \gamma_m + \mu_1 - \mu_m \right) \\
&= \mathbb{P}\left(\hat{\mu}_m - \mu_m < -\varepsilon + \gamma_m + \Delta_m \right) \\
&\le \exp\left(-\frac{N_m \left(\varepsilon - \gamma_m - \Delta_m \right)_+^2}{2} \right).
\end{aligned}
$$

For the second term, we have

$$
\begin{aligned}
&\mathbb{P}\left(\hat{r} \in [K] \setminus \{m\},\, \hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon \right) \\
&= \mathbb{P}\left(\hat{r} \in [K] \setminus \{m\},\, \frac{N_{\hat{r}}}{\sum_{j \in \mathcal{B}} N_j} \hat{\mu}_{\hat{r}} + \sum_{i \in \mathcal{B} \setminus \{\hat{r}\}} \frac{N_i}{\sum_{j \in \mathcal{B}} N_j} \hat{\mu}_i - \mu_1 < -\varepsilon \right) \\
&\le \mathbb{P}\left(\hat{r} \in [K] \setminus \{m\},\, \frac{N_{\hat{r}}}{\sum_{j \in \mathcal{B}} N_j} \hat{\mu}_{\hat{r}} + \sum_{i \in \mathcal{B} \setminus \{\hat{r}\}} \frac{N_i}{\sum_{j \in \mathcal{B}} N_j} \left(\hat{\mu}_{\hat{r}} - \gamma_{\hat{r}} \right) - \mu_1 < -\varepsilon \right) && \text{(use } \forall i \in \mathcal{B},\, \hat{\mu}_i \ge \hat{\mu}_{\hat{r}} - \gamma_{\hat{r}}) \\
&= \mathbb{P}\left(\hat{r} \in [K] \setminus \{m\},\, \hat{\mu}_{\hat{r}} - \mu_1 < -\varepsilon + \sum_{i \in \mathcal{B} \setminus \{\hat{r}\}} \frac{N_i}{\sum_{j \in \mathcal{B}} N_j} \gamma_{\hat{r}} \right) \\
&\le \mathbb{P}\left(\hat{r} \in [K] \setminus \{m\},\, \hat{\mu}_m - \gamma_m + \gamma_{\hat{r}} - \mu_1 < -\varepsilon + \sum_{i \in \mathcal{B} \setminus \{\hat{r}\}} \frac{N_i}{\sum_{j \in \mathcal{B}} N_j} \gamma_{\hat{r}} \right) && \text{(by def of } \hat{r},\, \hat{\mu}_{\hat{r}} > \hat{\mu}_m - \gamma_m + \gamma_{\hat{r}})
\end{aligned}
$$

$$
\begin{aligned}
&= \mathbb{P}\left(\hat{r} \in [K] \backslash \{m\}, \hat{\mu}_m - \mu_1 < -\varepsilon + \gamma_m - \gamma_{\hat{r}} + \sum_{i \in \mathcal{B} \backslash \{\hat{r}\}} \frac{N_i}{\sum_{j \in \mathcal{B}} N_j} \gamma_{\hat{r}}\right) \\
&\leq \mathbb{P}\left(\hat{r} \in [K] \backslash \{m\}, \hat{\mu}_m - \mu_1 < -\varepsilon + \gamma_m - \gamma_{\hat{r}} + \gamma_{\hat{r}}\right) && \left(\text{use } \textstyle\sum_{i \in \mathcal{B} \backslash \{m\}} \frac{N_i}{\sum_{j \in \mathcal{B}} N_j} \leq 1\right) \\
&= \mathbb{P}\left(\hat{r} \in [K] \backslash \{m\}, \hat{\mu}_m - \mu_1 < -\varepsilon + \gamma_m\right) \\
&\leq \mathbb{P}\left(\hat{\mu}_m - \mu_1 < -\varepsilon + \gamma_m\right) \\
&= \mathbb{P}\left(\hat{\mu}_m - \mu_m < -\varepsilon + \gamma_m + \mu_1 - \mu_m\right) \\
&= \mathbb{P}\left(\hat{\mu}_m - \mu_m < -\varepsilon + \gamma_m + \Delta_m\right) \\
&\leq \exp\left(-\frac{N_m\left(\varepsilon - \gamma_m - \Delta_m\right)_+^2}{2}\right).
\end{aligned}
$$

Combining the two terms, we have

$$
\begin{aligned}
&\mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon\right) \\
&= \mathbb{P}\left(\hat{r} = m, \hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon\right) + \mathbb{P}\left(\hat{r} \in [K] \backslash \{m\}, \hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon\right) \\
&\leq 2\exp\left(-\frac{N_m\left(\varepsilon - \gamma_m - \Delta_m\right)_+^2}{2}\right).
\end{aligned}
$$

□

Lemma 17. In HAVER, we have

$$
\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon\right) 2\varepsilon \, \mathrm{d}\varepsilon \leq \min_m \left(\frac{352}{N_m} \log\left(\frac{KS}{N_m}\right) + 4\Delta_m^2\right).
$$

Proof. Lemma 16 states that, for any arm $m \in [K]$,

$$
\mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon\right) \leq 2\exp\left(-\frac{1}{2} N_m \left(\varepsilon - \gamma_m - \Delta_m\right)_+^2\right)
$$

Let $q_1 = q_2 = 1$, $\varepsilon_0 = 2\left(\gamma_m + \Delta_m\right)$, and $z_1 = \frac{1}{16} N_m$. We claim that, in the regime of $\varepsilon \geq \varepsilon_0$, we have

$$
\mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon\right) \leq q_1 \exp\left(-z_1 \varepsilon^2\right)
$$

and in the regime of $\varepsilon < \varepsilon_0$, we have $\mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon\right) \leq q_2$.

The second claim is trivial. We prove the first claim as follows. In the regime of $\varepsilon \geq \varepsilon_0$, we have

$$
\begin{aligned}
&\mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon\right) \\
&\leq 2\exp\left(-\frac{1}{2} N_m \left(\varepsilon - \gamma_m - \Delta_m\right)_+^2\right) \\
&\leq 2\exp\left(-\frac{1}{2} N_m \left(\frac{\varepsilon}{2}\right)^2\right) && \left(\text{use } \varepsilon \geq \varepsilon_0 = 2\left(\gamma_m + \Delta_m\right)\right) \\
&= 2\exp\left(-\frac{1}{8} N_m \varepsilon^2\right) \\
&= \exp\left(-\frac{1}{8} N_m \varepsilon^2 + \log(2)\right) \\
&\leq \exp\left(-\frac{1}{16} N_m \varepsilon^2\right). && \left(\text{use } \varepsilon \geq \varepsilon_0 \geq \gamma_m\right)
\end{aligned}
$$

With the above claim, we use Lemma 31 to bound the integral

$$
\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon\right) 2\varepsilon \, \mathrm{d}\varepsilon
$$

$$
\begin{aligned}
&\leq q_2\varepsilon_0^2 + q_1\frac{1}{z_1} \\
&= \left(2\left(\gamma_m + \Delta_m\right)\right)^2 + \frac{16}{N_m} \\
&\leq 4\gamma_m^2 + 4\Delta_m^2 + \frac{16}{N_m} \\
&\leq 4\frac{18}{N_m}\log\left(\left(\frac{KS}{N_m}\right)^{2\lambda}\right) + 4\Delta_m^2 + \frac{16}{N_m} && \left(\text{use } \gamma_m = \sqrt{\tfrac{18}{N_m}\log\left(\left(\tfrac{KS}{N_m}\right)^{2\lambda}\right)}\right) \\
&= \frac{72}{N_m}\log\left(\left(\frac{KS}{N_m}\right)^{4}\right) + 4\Delta_m^2 + \frac{16}{N_m} && (\text{recall } \lambda = 2) \\
&\leq \frac{88}{N_m}\log\left(\left(\frac{KS}{N_m}\right)^{4}\right) + 4\Delta_m^2 && \left(\text{use } 1 \leq \log\left(\left(\tfrac{KS}{N_m}\right)^{4}\right) \text{ with } K \geq 2\right) \\
&\leq \frac{352}{N_m}\log\left(\frac{KS}{N_m}\right) + 4\Delta_m^2 \\
&\leq \min_m \left(\frac{352}{N_m}\log\left(\frac{KS}{N_m}\right) + 4\Delta_m^2\right).
\end{aligned}
$$

□

Lemma 18. If an event G sastisfies $\mathbb{P}(G) \leq \frac{2K^2}{K^{2\lambda}}$, then HAVER achieves

$$
\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, G\right) \mathrm{d}\varepsilon \leq \min_{m\in[K]}\left(\frac{704}{K^2 N_m}\log\left(\frac{KS}{N_m}\right) + \frac{8}{K^2}\Delta_m^2\right).
$$

Proof. Using Lemma 16, for any arm m, we have

$$
\begin{aligned}
&\mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, G\right) \\
&\leq \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon\right) \\
&\leq 2\exp\left(-\frac{1}{2}N_m\left(\varepsilon - \gamma_m - \Delta_m\right)_+^2\right).
\end{aligned}
$$

We also have

$$
\mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, G\right) \leq \mathbb{P}(G) \leq \frac{2K^2}{K^{2\lambda}}.
$$

Thus,

$$
\mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, G\right) \leq 2\exp\left(-\frac{1}{2}N_m\left(\varepsilon - \gamma_m - \Delta_m\right)_+^2\right) \wedge \frac{2K^2}{K^{2\lambda}}.
$$

Let $q_1 = \frac{2K}{K^\lambda}$, $q_2 = \frac{2K^2}{K^{2\lambda}}$, $\varepsilon_0 = 2(\gamma_m + \Delta_m)$ and $z_1 = \frac{1}{16}N_m$. We claim that, in the regime of $\varepsilon \geq \varepsilon_0$, we have

$$
\mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, G\right) \leq q_1\exp\left(-z_1\varepsilon^2\right)
$$

and in the regime of $\varepsilon < \varepsilon_0$, we have $\mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, G\right) \leq q_2$.

We prove the claim as follows. In the regime of $\varepsilon \geq \varepsilon_0$, we have

$$
\begin{aligned}
&\mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, G\right) \\
&\leq 2\exp\left(-\frac{1}{2}N_m\left(\varepsilon - \gamma_m - \Delta_m\right)_+^2\right) \wedge \frac{2K^2}{K^{2\lambda}} \\
&\leq \sqrt{2\exp\left(-\frac{1}{2}N_m\left(\varepsilon - \gamma_m - \Delta_m\right)_+^2\right)\cdot\frac{2K^2}{K^{2\lambda}}}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{2K}{K^{\lambda}} \exp\left(-\frac{1}{4} N_m \left(\varepsilon - \gamma_m - \Delta_m\right)_+^2\right) \\
&\leq \frac{2K}{K^{\lambda}} \exp\left(-\frac{1}{4} N_m \left(\frac{\varepsilon}{2}\right)^2\right) && \text{(use } \varepsilon \geq \varepsilon_0 = 2\left(\gamma_m + \Delta_m\right)\text{)} \\
&= \frac{2K}{K^{\lambda}} \exp\left(-\frac{1}{16} N_m \varepsilon^2\right).
\end{aligned}
$$

In the regime of $\varepsilon < \varepsilon_0$, we have

$$
\begin{aligned}
&\mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, G\right) \\
&\leq 2 \exp\left(-\frac{1}{2} N_m \left(\varepsilon - \gamma_m + \Delta_m\right)_+^2\right) \wedge \frac{2K^2}{K^{2\lambda}} \\
&\leq \frac{2K^2}{K^{2\lambda}}.
\end{aligned}
$$

With the above claim, we use Lemma 31 to bound the integral (todo fix the constant)

$$
\begin{aligned}
&\int_0^{\infty} \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, G\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&\leq q_2 \varepsilon_0^2 + q_1 \frac{1}{z_1} \\
&= \frac{2K^2}{K^{2\lambda}} \left(2\left(\gamma_m + \Delta_m\right)\right)^2 + \frac{2K}{K^{\lambda}} \frac{16}{N_m} \\
&\leq \frac{8K^2}{K^{2\lambda}} \gamma_m^2 + \frac{8K}{K^{2\lambda}} \Delta_m^2 + \frac{32}{K^{\lambda} N_m} \\
&\leq \frac{8K^2}{K^{2\lambda}} \frac{18}{N_m} \log\left(\left(\frac{KS}{N_m}\right)^{2\lambda}\right) + \frac{8K}{K^{2\lambda}} \Delta_m^2 + \frac{32}{K^{\lambda} N_m} && \text{(use } \gamma_m = \sqrt{\tfrac{18}{N_m} \log\left(\left(\tfrac{KS}{N_m}\right)^{2\lambda}\right)}\text{)} \\
&= \frac{144K^2}{K^4 N_m} \log\left(\left(\frac{KS}{N_m}\right)^4\right) + \frac{8K}{K^4} \Delta_m^2 + \frac{32}{K^2 N_m} && \text{(recall } \lambda = 2\text{)} \\
&\leq \frac{176}{K^2 N_m} \log\left(\left(\frac{KS}{N_m}\right)^4\right) + \frac{8}{K^3} \Delta_m^2 && \text{(use } 1 \leq \log\left(\left(\tfrac{KS}{N_m}\right)^4\right) \text{ with } K \geq 2\text{)} \\
&\leq \frac{704}{K^2 N_m} \log\left(\frac{KS}{N_m}\right) + \frac{8}{K^3} \Delta_m^2.
\end{aligned}
$$

We conclude the proof by using the fact that the result above applies to any choice of $m \in [K]$.

$\square$

Lemma 19. HAVER achieves

$$
\begin{aligned}
&\int_0^{\infty} \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon,\, \hat{r} \in \mathcal{R}\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&= \tilde{\mathcal{O}}\left(\left(\max_{r \in \mathcal{R}} \max_{k=0}^{d(r)} \max_{\substack{\mathcal{S}: \mathcal{B}^*(r) \subseteq \mathcal{S} \subseteq \mathcal{B}^+(r) \\ |\mathcal{S}| = n_*(r) + k}} \frac{k \sum_{j \in \mathcal{S} \setminus \mathcal{B}^*(r)} N_j}{\left(\sum_{j \in \mathcal{S}} N_j\right)^2}\right) \wedge \frac{1}{N_1}\right) \\
&\quad + \tilde{\mathcal{O}}\left(\left(\frac{1}{\min_{r \in \mathcal{R}} \sum_{j \in \mathcal{B}^*(r)} N_j}\right) \wedge \frac{1}{N_1}\right) \\
&\quad + \tilde{\mathcal{O}}\left(\frac{1}{K N_1}\right),
\end{aligned}
$$

where $d(r) = \left|\mathcal{B}^+(r)\right| - \left|\mathcal{B}^*(r)\right|$ and $n_*(r) = \left|\mathcal{B}^*(r)\right|$.

Proof. We consider 4 complementary events:

- $G_0 = \left\{\mathcal{B} = \mathcal{B}^*(\hat{r})\right\}$
- $G_1 = \left\{\mathcal{B}^*(\hat{r}) \subset \mathcal{B} \subseteq \mathcal{B}^+(\hat{r})\right\}$
- $G_2 = \left\{\exists i \in \mathcal{B}^*(\hat{r})\, \text{s.t.}\, i \notin \mathcal{B}\right\}$
- $G_3 = \left\{\mathcal{B}^*(\hat{r}) \subseteq \mathcal{B},\, \exists i \notin \mathcal{B}^+(\hat{r})\, \text{s.t.}\, i \in \mathcal{B}\right\}$

We have

$$\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon,\, \hat{r} \in \mathcal{R}\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
= &\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon,\, \hat{r} \in \mathcal{R},\, G_0\right) 2\varepsilon \,\mathrm{d}\varepsilon + \int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon,\, \hat{r} \in \mathcal{R},\, G_1\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&+ \int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon,\, \hat{r} \in \mathcal{R},\, G_2\right) 2\varepsilon \,\mathrm{d}\varepsilon + \int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon,\, \hat{r} \in \mathcal{R},\, G_3\right) 2\varepsilon \,\mathrm{d}\varepsilon.
\end{aligned}$$

Use Lemma 20 and 21 for event G_0 and G_1 respectively, we have

$$\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon,\, \hat{r} \in \mathcal{R},\, G_0\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
= &\mathcal{O}\left(\frac{\log(K)}{\min_{r\in\mathcal{R}} \sum_{j\in\mathcal{B}^*(r)} N_j}\right)
\end{aligned}$$

and

$$\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon,\, \hat{r} \in \mathcal{R},\, G_1\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
= &\mathcal{O}\left(\max_{r\in\mathcal{R}} \max_{k=0}^{d(r)} \max_{\substack{S:\mathcal{B}^*(r)\subseteq\mathcal{S}\subseteq\mathcal{B}^+(r)\\ |\mathcal{S}|=n_*(r)+k}} \frac{k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}{\left(\sum_{j\in\mathcal{S}} N_j\right)^2} \log\left(\frac{ed(r)}{k}\right) \log(K)\right) \\
&+ \mathcal{O}\left(\frac{\log(K)}{\min_{r\in\mathcal{R}} \sum_{j\in\mathcal{B}^*(r)} N_j}\right).
\end{aligned}$$

Combining the first two events, we have

$$\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon,\, \hat{r} \in \mathcal{R},\, G_0\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&+ \int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon,\, \hat{r} \in \mathcal{R},\, G_1\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
= &\mathcal{O}\left(\max_{r\in\mathcal{R}} \max_{k=0}^{d(r)} \max_{\substack{S:\mathcal{B}^*(r)\subseteq\mathcal{S}\subseteq\mathcal{B}^+(r)\\ |\mathcal{S}|=n_*(r)+k}} \frac{k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}{\left(\sum_{j\in\mathcal{S}} N_j\right)^2} \log\left(\frac{ed(r)}{k}\right) \log(K)\right) \\
&+ \mathcal{O}\left(\frac{\log(K)}{\min_{r\in\mathcal{R}} \sum_{j\in\mathcal{B}^*(r)} N_j}\right) \\
\leq &\mathcal{O}\left(\max_{r\in\mathcal{R}} \max_{k=0}^{d(r)} \max_{\substack{S:\mathcal{B}^*(r)\subseteq\mathcal{S}\subseteq\mathcal{B}^+(r)\\ |\mathcal{S}|=n_*(r)+k}} \frac{k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}{\left(\sum_{j\in\mathcal{S}} N_j\right)^2} \log\left(3K\right) \log(K)\right) \qquad \left(\text{use } \tfrac{ed(r)}{k} \leq 3K\right) \\
&+ \mathcal{O}\left(\frac{\log(K)}{\min_{r\in\mathcal{R}} \sum_{j\in\mathcal{B}^*(r)} N_j}\right) \\
= &\tilde{\mathcal{O}}\left(\max_{r\in\mathcal{R}} \max_{k=0}^{d(r)} \max_{\substack{S:\mathcal{B}^*(r)\subseteq\mathcal{S}\subseteq\mathcal{B}^+(r)\\ |\mathcal{S}|=n_*(r)+k}} \frac{k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}{\left(\sum_{j\in\mathcal{S}} N_j\right)^2}\right.
\end{aligned}$$

$$+\tilde{\mathcal{O}}\left(\frac{1}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)}N_j}\right).$$

We use Lemma 28 and 29 to show that $\mathbb{P}(G_2)\le\frac{2K^2}{K^{2\lambda}}$ and $\mathbb{P}(G_3)\le\frac{2K^2}{K^{2\lambda}}$ respectively.

Consequently, we use Lemma 22 for event G_2 to obtain the integral

$$\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}}-\mu_1>\varepsilon,\,\hat{r}\in\mathcal{R},\,G_2\right)2\varepsilon\,\mathrm{d}\varepsilon=\tilde{\mathcal{O}}\left(\frac{1}{KN_1}\right).$$

We use use Lemma 22 for event G_3 and get the same result. Furthermore, Lemma 27 states that

$$\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}}-\mu_1>\varepsilon\right)2\varepsilon\,\mathrm{d}\varepsilon=\tilde{\mathcal{O}}\left(\frac{1}{N_1}\right).$$

Combining all the results completes our proof. □

Lemma 20. In the event of $G_0=\{\mathcal{B}=\mathcal{B}^*(\hat{r})\}$, HAVER achieves

$$\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}}-\mu_1>\varepsilon,\,\hat{r}\in\mathcal{R},\,G_0\right)2\varepsilon\,\mathrm{d}\varepsilon\le 10\left(\frac{\log(K)}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)}N_j}\right).$$

Proof. We have

$$\begin{aligned}
&\mathbb{P}\left(\hat{\mu}^{\mathcal{B}}-\mu_1>\varepsilon,\,\hat{r}\in\mathcal{R},\,G_0\right)\\
=&\,\mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}}N_j}\sum_{i\in\mathcal{B}}N_i\hat{\mu}_i-\mu_1>\varepsilon,\,\hat{r}\in\mathcal{R},\,G_0\right)\\
=&\,\mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}}N_j}\sum_{i\in\mathcal{B}}N_i\left(\hat{\mu}_i-\mu_1\right)>\varepsilon,\,\hat{r}\in\mathcal{R},\,G_0\right)\\
=&\,\mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}}N_j}\sum_{i\in\mathcal{B}}N_i\left(\hat{\mu}_i-\mu_i\right)>\varepsilon+\frac{1}{\sum_{j\in\mathcal{B}}N_j}\sum_{i\in\mathcal{B}}N_i\Delta_i,\,\hat{r}\in\mathcal{R},\,G_0\right)\\
=&\,\mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}^*(\hat{r})}N_j}\sum_{i\in\mathcal{B}^*(\hat{r})}N_i\left(\hat{\mu}_i-\mu_i\right)>\varepsilon+\frac{1}{\sum_{j\in\mathcal{B}^*(\hat{r})}N_j}\sum_{i\in\mathcal{B}^*(\hat{r})}N_i\Delta_i,\,\hat{r}\in\mathcal{R},\,G_0\right)\quad(\text{use } G_0=\{\mathcal{B}=\mathcal{B}^*(\hat{r})\})\\
\le&\sum_{r\in\mathcal{R}}\mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}^*(r)}N_j}\sum_{i\in\mathcal{B}^*(r)}N_i\left(\hat{\mu}_i-\mu_i\right)>\varepsilon+\frac{1}{\sum_{j\in\mathcal{B}^*(r)}N_j}\sum_{i\in\mathcal{B}^*(r)}N_i\Delta_i\right)\\
\le&\sum_{r\in\mathcal{R}}\exp\left(-\frac{1}{2}\left(\sum_{j\in\mathcal{B}^*(\hat{r})}N_j\right)\left(\varepsilon+\frac{1}{\sum_{j\in\mathcal{B}^*(r)}N_j}\sum_{i\in\mathcal{B}^+(r)}N_i\Delta_i\right)^2\right)\\
\le&\sum_{r\in\mathcal{R}}\exp\left(-\frac{1}{2}\left(\sum_{j\in\mathcal{B}^*(r)}N_j\right)\varepsilon^2\right)\\
\le&\,K\exp\left(-\frac{1}{2}\left(\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)}N_j\right)\varepsilon^2\right).
\end{aligned}$$

Let $q_1=q_2=1$, $z_1=\frac{1}{4}\left(\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)}N_j\right)$, and $\varepsilon_0=\sqrt{\frac{4\log(K)}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)}N_j}}$. We claim that, in the regime of $\varepsilon\ge\varepsilon_0$, we have

$$\mathbb{P}\left(\hat{\mu}^{\mathcal{B}}-\mu_1>\varepsilon,\,\hat{r}\in\mathcal{R},\,G_0\right)\le q_1\exp\left(-z_1\varepsilon^2\right)$$

and in the regime of $\varepsilon < \varepsilon_0$, we have

$$\mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon,\, \hat{r} \in \mathcal{R},\, G_0\right) \leq q_2.$$

The second claim is trivial. We prove the first claim as follows. In the regime of $\varepsilon \geq \varepsilon_0$, we have

$$\begin{aligned}
&\mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon,\, \hat{r} \in \mathcal{R},\, G_0\right) \\
&\leq K \exp\left(-\frac{1}{2}\left(\min_{r\in\mathcal{R}} \sum_{j\in\mathcal{B}^*(r)} N_j\right)\varepsilon^2\right) \\
&\leq \exp\left(-\frac{1}{2}\left(\min_{r\in\mathcal{R}} \sum_{j\in\mathcal{B}^*(r)} N_j\right)\varepsilon^2 + \log(K)\right) \\
&\leq \exp\left(-\frac{1}{4}\left(\min_{r\in\mathcal{R}} \sum_{j\in\mathcal{B}^*(r)} N_j\right)\varepsilon^2\right). && \text{(use } \varepsilon \geq \varepsilon_0\text{)}
\end{aligned}$$

With the above claim, we use Lemma 31 to bound the integral

$$\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon,\, G_0\right) 2\varepsilon\,\mathrm{d}\varepsilon \\
&\leq q_1\varepsilon_0^2 + q_2\frac{1}{z_1} \\
&= 4\left(\frac{\log(K)}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}\right) + 4\left(\frac{4}{\sum_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}\right) \\
&= 2\left(\frac{\log(K^2)}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}\right) + 4\left(\frac{4}{\sum_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}\right) \\
&\leq 5\left(\frac{\log(K^2)}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}\right) && \text{(use } K \geq 2\text{)} \\
&\leq 10\left(\frac{\log(K)}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}\right).
\end{aligned}$$

□

Lemma 21. In the event of $G_1 = \left\{\mathcal{B}^*(\hat{r}) \subset \mathcal{B} \subseteq \mathcal{B}^+(\hat{r})\right\}$, HAVER achieves

$$\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon,\, \hat{r} \in \mathcal{R},\, G_1\right) 2\varepsilon\,\mathrm{d}\varepsilon \\
&= 128\left(\max_{r\in\mathcal{R}}\max_{k=0}^{d(r)} \max_{\substack{S:\mathcal{B}^*(r)\subseteq\mathcal{S}\subseteq\mathcal{B}^+(r)\\ |\mathcal{S}|=n_*(r)+k}} \frac{k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}{\left(\sum_{j\in\mathcal{S}} N_j\right)^2}\log\left(\frac{ed(r)}{k}\right)\log(K)\right) + 48\left(\frac{\log(K)}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}\right),
\end{aligned}$$

where $d(r) = \left|\mathcal{B}^+(r)\right| - \left|\mathcal{B}^*(r)\right|$ and $n_*(r) = \left|\mathcal{B}^*(r)\right|$.

Proof. We decompose the integral as follow:

$$\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon,\, \hat{r} \in \mathcal{R},\, G_1\right) 2\varepsilon\,\mathrm{d}\varepsilon \\
&= \int_0^\infty \mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j}\sum_{i\in\mathcal{B}} N_i\left(\hat{\mu}_i - \mu_1\right) > \varepsilon,\, \hat{r} \in \mathcal{R},\, G_1\right) 2\varepsilon\,\mathrm{d}\varepsilon
\end{aligned}$$

$$\le \int_0^\infty \mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j}\sum_{i\in\mathcal{B}^*(\hat{r})} N_i\left(\hat{\mu}_i-\mu_1\right) > \frac{\varepsilon}{2},\, \hat{r}\in\mathcal{R},\, G_1\right) 2\varepsilon\,\mathrm{d}\varepsilon$$
$$+ \int_0^\infty \mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j}\sum_{i\in\mathcal{B}\setminus\mathcal{B}^*(\hat{r})} N_i\left(\hat{\mu}_i-\mu_1\right) > \frac{\varepsilon}{2},\, \hat{r}\in\mathcal{R},\, G_1\right) 2\varepsilon\,\mathrm{d}\varepsilon.$$

We bound the probability and integral of these terms respectively. For the first term, we have

$$\begin{aligned}
&\mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j}\sum_{i\in\mathcal{B}^*(\hat{r})} N_i\left(\hat{\mu}_i-\mu_1\right) > \frac{\varepsilon}{2},\, \hat{r}\in\mathcal{R},\, G_1\right)\\
&= \mathbb{P}\left(\sum_{i\in\mathcal{B}^*(\hat{r})} N_i\left(\hat{\mu}_i-\mu_1\right) > \frac{\varepsilon}{2}\sum_{j\in\mathcal{B}} N_j,\, \hat{r}\in\mathcal{R},\, G_1\right)\\
&\le \mathbb{P}\left(\sum_{i\in\mathcal{B}^*(\hat{r})} N_i\left(\hat{\mu}_i-\mu_1\right) > \frac{\varepsilon}{2}\sum_{j\in\mathcal{B}^*(\hat{r})} N_j,\, \hat{r}\in\mathcal{R},\, G_1\right) \qquad \left(\text{since } \mathcal{B}^*(\hat{r})\subset\mathcal{B},\, \textstyle\sum_{j\in\mathcal{B}^*(\hat{r})} N_j < \sum_{j\in\mathcal{B}} N_j\right)\\
&= \mathbb{P}\left(\sum_{i\in\mathcal{B}^*(\hat{r})} N_i\left(\hat{\mu}_i-\mu_i\right) > \frac{\varepsilon}{2}\sum_{j\in\mathcal{B}^*(\hat{r})} N_j + \sum_{i\in\mathcal{B}^*(\hat{r})} N_i\Delta_i,\, \hat{r}\in\mathcal{R},\, G_1\right)\\
&\le \sum_{r\in\mathcal{R}} \mathbb{P}\left(\sum_{i\in\mathcal{B}^*(r)} N_i\left(\hat{\mu}_i-\mu_i\right) \ge \frac{\varepsilon}{2}\sum_{j\in\mathcal{B}^*(r)} N_j + \sum_{i\in\mathcal{B}^*(r)} N_i\Delta_i\right)\\
&\le \sum_{r\in\mathcal{R}} \exp\left(-\frac{1}{2}\frac{1}{\sum_{j\in\mathcal{B}^*(r)} N_j}\left(\frac{\varepsilon}{2}\sum_{j\in\mathcal{B}^*(r)} N_j + \sum_{i\in\mathcal{B}^*(r)} N_i\Delta_i\right)^2\right)\\
&\le \sum_{r\in\mathcal{R}} \exp\left(-\frac{1}{2}\frac{1}{\sum_{j\in\mathcal{B}^*(r)} N_j}\left(\frac{\varepsilon}{2}\sum_{j\in\mathcal{B}^*(r)} N_j\right)^2\right)\\
&= \sum_{r\in\mathcal{R}} \exp\left(-\frac{1}{8}\left(\sum_{j\in\mathcal{B}^*(r)} N_j\right)\varepsilon^2\right)\\
&\le K\exp\left(-\frac{1}{8}\left(\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j\right)\varepsilon^2\right).
\end{aligned}$$

For this term, let $q_1 = q_2 = 1$, $z_1 = \frac{1}{16}\left(\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j\right)$, and $\varepsilon_0 = \sqrt{\frac{16\log(K)}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}}$. We claim that, in the regime of $\varepsilon \ge \varepsilon_0$, we have

$$\mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j}\sum_{i\in\mathcal{B}^*(\hat{r})} N_i\left(\hat{\mu}_i-\mu_1\right) > \frac{\varepsilon}{2},\, \hat{r}\in\mathcal{R},\, G_1\right) \le q_1\exp\left(-z_1\varepsilon^2\right)$$

and in the regime of $\varepsilon < \varepsilon_0$, we have

$$\mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j}\sum_{i\in\mathcal{B}^*(\hat{r})} N_i\left(\hat{\mu}_i-\mu_1\right) > \frac{\varepsilon}{2},\, \hat{r}\in\mathcal{R},\, G_1\right) \le q_2.$$

The second claim is trivial. We prove the first claim as follows. In the regime of $\varepsilon \ge \varepsilon_0$, we have

$$\mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j}\sum_{i\in\mathcal{B}^*(\hat{r})} N_i\left(\hat{\mu}_i-\mu_1\right) > \frac{\varepsilon}{2},\, \hat{r}\in\mathcal{R},\, G_1\right)$$

$$
\begin{aligned}
&\leq K \exp\left(-\frac{1}{8}\left(\min_{r\in\mathcal{R}} \sum_{j\in\mathcal{B}^*(r)} N_j\right)\varepsilon^2\right) \\
&= \exp\left(-\frac{1}{8}\left(\min_{r\in\mathcal{R}} \sum_{j\in\mathcal{B}^*(r)} N_j\right)\varepsilon^2 + \log(K)\right) \\
&\leq \exp\left(-\frac{1}{16}\left(\min_{r\in\mathcal{R}} \sum_{j\in\mathcal{B}^*(r)} N_j\right)\varepsilon^2\right) \qquad \text{(use } \varepsilon \geq \varepsilon_0\text{)}
\end{aligned}
$$

With the above claim, we use Lemma 31 to bound the integral

$$
\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j} \sum_{i\in\mathcal{B}^*(\hat{r})} N_i\left(\hat{\mu}_i - \mu_1\right) > \frac{\varepsilon}{2},\, \hat{r}\in\mathcal{R},\, G_1\right) 2\varepsilon\,\mathrm{d}\varepsilon \\
&\leq q_1\varepsilon_0^2 + q_2\frac{1}{z_1} \\
&= 16\left(\frac{\log(K)}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}\right) + 16\left(\frac{1}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}\right) \\
&= 8\left(\frac{\log(K^2)}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}\right) + 16\left(\frac{1}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}\right) \\
&\leq 24\left(\frac{\log(K^2)}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}\right) \qquad \text{(use } K \geq 2\text{)} \\
&= 48\left(\frac{\log(K)}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}\right).
\end{aligned}
$$

For the second term, we have

$$
\begin{aligned}
&\mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j} \sum_{i\in\mathcal{B}\setminus\mathcal{B}^*(\hat{r})} N_i\left(\hat{\mu}_i - \mu_1\right) > \frac{\varepsilon}{2},\, \hat{r}\in\mathcal{R},\, G_1\right) \\
&= \mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j} \sum_{i\in\mathcal{B}\setminus\mathcal{B}^*(\hat{r})} N_i\left(\hat{\mu}_i - \mu_i\right) > \frac{\varepsilon}{2} + \frac{1}{\sum_{j\in\mathcal{B}} N_j} \sum_{i\in\mathcal{B}\setminus\mathcal{B}^*(\hat{r})} N_i\Delta_i,\, \hat{r}\in\mathcal{R},\, G_1\right) \\
&\leq \mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j} \sum_{i\in\mathcal{B}\setminus\mathcal{B}^*(\hat{r})} N_i\left(\hat{\mu}_i - \mu_i\right) > \frac{\varepsilon}{2},\, \hat{r}\in\mathcal{R},\, G_1\right) \\
&\leq \sum_{r\in\mathcal{R}} \sum_{k=1}^{d(r)} \sum_{\substack{\mathcal{S}:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\ |\mathcal{S}|=n_*(r)+k}} \mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{S}} N_j} \sum_{i\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_i\left(\hat{\mu}_i - \mu_i\right) > \frac{\varepsilon}{2},\, G_1\right) \\
&\qquad\qquad \text{(use } d(r) = \left|\mathcal{B}^+(r)\right| - \left|\mathcal{B}^*(r)\right| \text{ and } n_*(r) = \left|\mathcal{B}^*(r)\right|\text{)} \\
&\leq \sum_{r\in\mathcal{R}} \sum_{k=1}^{d(r)} \sum_{\substack{\mathcal{S}:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\ |\mathcal{S}|=n_*(r)+k}} \exp\left(-\frac{1}{2}\left(\frac{\left(\sum_{j\in\mathcal{S}} N_j\right)^2}{\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(\hat{r})} N_j}\right)\left(\frac{\varepsilon}{2}\right)^2\right) \\
&= \sum_{r\in\mathcal{R}} \sum_{k=1}^{d(r)} \sum_{\substack{\mathcal{S}:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\ |\mathcal{S}|=n_*(r)+k}} \exp\left(-\frac{1}{8}\left(\frac{\left(\sum_{j\in\mathcal{S}} N_j\right)^2}{\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}\right)\varepsilon^2\right)
\end{aligned}
$$

$$\le \sum_{r\in\mathcal{R}}\sum_{k=1}^{d(r)}\sum_{\substack{S:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\|\mathcal{S}|=n_*(r)+k}}\exp\left(-\frac{1}{8}\left(\min_{r\in\mathcal{R}}\min_{k=1}^{d(r)}\min_{\substack{S:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\|\mathcal{S}|=n_*(r)+k}}\frac{\left(\sum_{j\in\mathcal{S}}N_j\right)^2}{\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)}N_j}\right)\varepsilon^2\right)$$

$$=\sum_{r\in\mathcal{R}}\sum_{k=1}^{d(r)}\binom{d(r)}{k}\exp\left(-\frac{1}{8}\left(\min_{r\in\mathcal{R}}\min_{k=1}^{d(r)}\min_{\substack{S:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\|\mathcal{S}|=n_*(r)+k}}\frac{\left(\sum_{j\in\mathcal{S}}N_j\right)^2}{\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)}N_j}\right)\varepsilon^2\right).$$

For brevity, we denote

$$M=\min_{r\in\mathcal{R}}\min_{k=1}^{d(r)}\min_{\substack{S:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\|\mathcal{S}|=n_*(r)+k}}\frac{\left(\sum_{j\in\mathcal{S}}N_j\right)^2}{\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)}N_j}$$

For this term, let $q_1=q_2=1$, $z_1=\frac{M}{32}$, and

$$\varepsilon_0=\sqrt{\max_{r\in\mathcal{R}}\max_{k=1}^{d(r)}\max_{\substack{S:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\|\mathcal{S}|=n_*(r)+k}}\frac{32k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)}N_j}{\left(\sum_{j\in\mathcal{S}}N_j\right)^2}\log\left(\frac{ed(r)}{k}\right)\log\left(K^2\right)}.$$

We claim that, in the regime of $\varepsilon\ge\varepsilon_0$, we have

$$\mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}}N_j}\sum_{i\in\mathcal{B}\setminus\mathcal{B}^*(r)}N_i\left(\hat{\mu}_i-\mu_1\right)>\frac{\varepsilon}{2},\,\hat{r}\in\mathcal{R},\,G_1\right)\le q_1\exp\left(-z_1\varepsilon^2\right)$$

and in the regime of $\varepsilon<\varepsilon_0$, we have

$$\mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}}N_j}\sum_{i\in\mathcal{B}\setminus\mathcal{B}^*(r)}N_i\left(\hat{\mu}_i-\mu_1\right)>\frac{\varepsilon}{2},\,\hat{r}\in\mathcal{R},\,G_1\right)\le q_2.$$

The second claim is trivial. We prove the first claim as follows. In the regime of $\varepsilon\ge\varepsilon_0$, we have

$$\begin{aligned}
&\mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}}N_j}\sum_{i\in\mathcal{B}\setminus\mathcal{B}^*(r)}N_i\left(\hat{\mu}_i-\mu_1\right)>\frac{\varepsilon}{2},\,\hat{r}\in\mathcal{R},\,G_1\right)\\
&\le\sum_{r\in\mathcal{R}}\sum_{k=1}^{d(r)}\binom{d(r)}{k}\exp\left(-\frac{1}{8}M\varepsilon^2\right)\\
&\le\sum_{r\in\mathcal{R}}\sum_{k=1}^{d(r)}\left(\frac{ed(r)}{k}\right)^k\exp\left(-\frac{1}{8}M\varepsilon^2\right) && \text{(use Stirling's formula, Lemma 46)}\\
&=\sum_{r\in\mathcal{R}}\sum_{k=1}^{d(r)}\exp\left(-\frac{1}{8}M\varepsilon^2+k\log\left(\frac{ed(r)}{k}\right)\right)\\
&\le\sum_{r\in\mathcal{R}}\sum_{k=1}^{d(r)}\exp\left(-\frac{1}{16}M\varepsilon^2\right) && \text{(use }\varepsilon\ge\varepsilon_0\text{)}\\
&\le\sum_{r\in\mathcal{R}}\sum_{k=1}^{d(r)}\exp\left(-\frac{1}{16}M\varepsilon^2\right)\\
&=Kd\exp\left(-\frac{1}{16}M\varepsilon^2\right)\\
&=\exp\left(-\frac{1}{16}M\varepsilon^2+\ln(Kd)\right)
\end{aligned}$$

$$\leq \exp\left(-\frac{1}{32}M\varepsilon^2\right). \qquad \text{(use } \varepsilon \geq \varepsilon_0\text{)}$$

With the above claim, we use Lemma 31 to bound the integral

$$\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j}\sum_{i\in\mathcal{B}\setminus\mathcal{B}^*(r)} N_i\left(\hat{\mu}_i-\mu_1\right) > \frac{\varepsilon}{2},\, \hat{r}\in\mathcal{R},\, G_1\right) 2\varepsilon\,\mathrm{d}\varepsilon \\
&\leq q_1\varepsilon_0^2 + q_2\frac{1}{z_1} \\
&= \left(\sqrt{\max_{r\in\mathcal{R}}\max_{k=1}^{d(r)}\max_{\substack{S:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\|\mathcal{S}|=n_*(r)+k}} \frac{32k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}{\left(\sum_{j\in\mathcal{S}} N_j\right)^2}\log\left(\frac{ed(r)}{k}\right)\log\left(K^2\right)}\right)^2 + \frac{32}{M} \\
&= 32\left(\max_{r\in\mathcal{R}}\max_{k=1}^{d(r)}\max_{\substack{S:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\|\mathcal{S}|=n_*(r)+k}} \frac{32k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}{\left(\sum_{j\in\mathcal{S}} N_j\right)^2}\log\left(\frac{ed(r)}{k}\right)\log\left(K^2\right)\right) \\
&\quad + 32\left(\max_{r\in\mathcal{R}}\max_{k=1}^{d(r)}\max_{\substack{S:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\|\mathcal{S}|=n_*(r)+k}} \frac{\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}{\left(\sum_{j\in\mathcal{S}} N_j\right)^2}\right) \qquad \text{(plug in } M\text{)} \\
&\leq 64\left(\max_{r\in\mathcal{R}}\max_{k=1}^{d(r)}\max_{\substack{S:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\|\mathcal{S}|=n_*(r)+k}} \frac{k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}{\left(\sum_{j\in\mathcal{S}} N_j\right)^2}\log\left(\frac{ed(r)}{k}\right)\log\left(K^2\right)\right) \\
&\qquad \text{(use } 1\leq\log\left(\tfrac{ed(r)}{k}\right)\ \forall k\in[1,d] \text{ and } 1\leq\log(K^2)\ \forall K\geq 2\text{)} \\
&= 128\left(\max_{r\in\mathcal{R}}\max_{k=1}^{d(r)}\max_{\substack{S:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\|\mathcal{S}|=n_*(r)+k}} \frac{k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}{\left(\sum_{j\in\mathcal{S}} N_j\right)^2}\log\left(\frac{ed(r)}{k}\right)\log(K)\right).
\end{aligned}$$

Combining all the terms, we have

$$\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}}-\mu_1 > \varepsilon,\, \hat{r}\in\mathcal{R},\, G_1\right) 2\varepsilon\,\mathrm{d}\varepsilon \\
&\leq \int_0^\infty \mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j}\sum_{i\in\mathcal{B}^*(r)} N_i\left(\hat{\mu}_i-\mu_1\right) > \frac{\varepsilon}{2},\, \hat{r}\in\mathcal{R},\, G_1\right) 2\varepsilon\,\mathrm{d}\varepsilon \\
&\quad + \int_0^\infty \mathbb{P}\left(\frac{1}{\sum_{j\in\mathcal{B}} N_j}\sum_{i\in\mathcal{B}\setminus\mathcal{B}^*(r)} N_i\left(\hat{\mu}_i-\mu_1\right) > \frac{\varepsilon}{2},\, \hat{r}\in\mathcal{R},\, G_1\right) 2\varepsilon\,\mathrm{d}\varepsilon \\
&\leq 48\left(\frac{\log(K)}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}\right) + 128\left(\max_{r\in\mathcal{R}}\max_{k=1}^{d(r)}\max_{\substack{S:\mathcal{B}^*(r)\subset\mathcal{S}\subseteq\mathcal{B}^+(r)\\|\mathcal{S}|=n_*(r)+k}} \frac{k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}{\left(\sum_{j\in\mathcal{S}} N_j\right)^2}\log\left(\frac{ed(r)}{k}\right)\log(K)\right) \\
&\leq 48\left(\frac{\log(K)}{\min_{r\in\mathcal{R}}\sum_{j\in\mathcal{B}^*(r)} N_j}\right) + 128\left(\max_{r\in\mathcal{R}}\max_{k=0}^{d(r)}\max_{\substack{S:\mathcal{B}^*(r)\subseteq\mathcal{S}\subseteq\mathcal{B}^+(r)\\|\mathcal{S}|=n_*(r)+k}} \frac{k\sum_{j\in\mathcal{S}\setminus\mathcal{B}^*(r)} N_j}{\left(\sum_{j\in\mathcal{S}} N_j\right)^2}\log\left(\frac{ed(r)}{k}\right)\log(K)\right)
\end{aligned}$$

□

Lemma 22. If event G sastisfies $\mathbb{P}(G) \le \frac{2K^2}{K^{2\lambda}}$, then HAVER achieves

$$\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon,\, G\right) 2\varepsilon \,\mathrm{d}\varepsilon \le \frac{3088}{KN_1} \log(KS).$$

Proof. We have

$$\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon, G\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
\le &\int_0^\infty \mathbb{P}\left(\hat{r} = 1,\, \hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon, G\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&+ \int_0^\infty \sum_{\substack{i: i \neq 1, \\ \gamma_i \le 2\gamma_1}} \mathbb{P}\left(\hat{r} = i,\, \hat{\mu}^{\mathcal{B}} - \mu_i > \varepsilon, G\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&+ \int_0^\infty \sum_{\substack{i: i \neq 1, \\ \gamma_i > 2\gamma_1}} \mathbb{P}\left(\hat{r} = i,\, \hat{\mu}^{\mathcal{B}} - \mu_i > \varepsilon, G\right) 2\varepsilon \,\mathrm{d}\varepsilon.
\end{aligned}$$

For the first term, we have

$$\begin{aligned}
&\mathbb{P}\left(\hat{r} = 1,\, \hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon, G\right) \\
\le\, &\mathbb{P}\left(\hat{r} = 1,\, \hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon\right) \\
\le\, &\exp\left(-\frac{1}{2} N_1 \left(\varepsilon - \frac{1}{2}\gamma_1\right)_+^2\right). && \text{(use Lemma 24)}
\end{aligned}$$

We also have $\mathbb{P}(G) \le \frac{2K^2}{K^{2\lambda}}$, thus

$$\mathbb{P}\left(\hat{r} = 1,\, \hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon, G\right) \le \exp\left(-\frac{1}{2} N_1 (\varepsilon - \gamma_1)_+^2\right) \wedge \frac{2K^2}{K^{2\lambda}}.$$

Let $q_1 = \frac{2K}{K^\lambda}$, $q_2 = \frac{2K^2}{K^{2\lambda}}$, $z_1 = \frac{N_1}{16}$, and $\varepsilon_0 = \gamma_1$. We claim that in the regime of $\varepsilon \ge \varepsilon_0$, we have

$$\mathbb{P}\left(\hat{r} = 1,\, \hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon, G\right) \le q_1 \exp\left(-z_1 \varepsilon^2\right)$$

and in the regime of $\varepsilon < \varepsilon_0$, we have $\mathbb{P}\left(\hat{r} = 1,\, \hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon, G\right) \le q_2$.

We prove the claim as follows. In the regime of $\varepsilon \ge \varepsilon_0$, we have

$$\begin{aligned}
&\mathbb{P}\left(\hat{r} = 1,\, \hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon, G\right) \\
\le\, &\exp\left(-\frac{1}{2} N_1 \left(\varepsilon - \frac{1}{2}\gamma_1\right)_+^2\right) \wedge \frac{2K^2}{K^{2\lambda}} \\
\le\, &\sqrt{\exp\left(-\frac{1}{2} N_1 \left(\varepsilon - \frac{1}{2}\gamma_1\right)_+^2\right) \cdot \frac{2K^2}{K^{2\lambda}}} \\
=\, &\frac{2K}{K^\lambda} \exp\left(-\frac{1}{4} N_1 \left(\varepsilon - \frac{1}{2}\gamma_1\right)_+^2\right) \\
\le\, &\frac{2K}{K^\lambda} \exp\left(-\frac{1}{4} N_1 \left(\frac{\varepsilon}{2}\right)^2\right) && \text{(use } \varepsilon \ge \gamma_1\text{)} \\
=\, &\frac{2K}{K^\lambda} \exp\left(-\frac{1}{16} N_1 \varepsilon^2\right).
\end{aligned}$$

In the regime of $\varepsilon < \varepsilon_0$, we have

$$\mathbb{P}\left(\hat{r} = 1,\, \hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon, G\right)$$

$$\leq \exp\left(-\frac{1}{2}N_1\left(\varepsilon-\frac{1}{2}\gamma_1\right)_+^2\right)\wedge\frac{2K^2}{K^{2\lambda}}$$
$$\leq \frac{2K^2}{K^{2\lambda}}.$$

With the above claim, we use Lemma 31 to bound the integral

$$\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{r}=1,\,\hat{\mu}^{\mathcal{B}}-\mu_1>\varepsilon,G\right)2\varepsilon\,\mathrm{d}\varepsilon\\
&\leq q_2\varepsilon_0^2+q_1\frac{1}{z_1}\\
&=\frac{2K^2}{K^{2\lambda}}\gamma_1^2+\frac{2K}{K^\lambda}\frac{16}{N_1}\\
&=\frac{2K^2}{K^{2\lambda}}\frac{18}{N_1}\log\left(\left(\frac{KS}{N_1}\right)^{2\lambda}\right)+\frac{2K}{K^\lambda}\frac{16}{N_1} && \left(\text{use } \gamma_i=\sqrt{\tfrac{18}{N_i}\log\left(\left(\tfrac{KS}{N_i}\right)^{2\lambda}\right)}\right)\\
&=\frac{36K^2}{K^4N_1}\log\left(\left(\frac{KS}{N_1}\right)^{4}\right)+\frac{32K}{K^2N_1} && (\text{recall } \lambda=2)\\
&=\frac{36}{K^2N_1}\log\left(\left(\frac{KS}{N_1}\right)^{4}\right)+\frac{32}{KN_1}\\
&=\frac{68}{KN_1}\log\left(\left(\frac{KS}{N_1}\right)^{4}\right) && \left(\text{use } 1\leq\log\left(\left(\tfrac{KS}{N_1}\right)^4\right)\text{ with } K\geq 2\right)\\
&=\frac{272}{KN_1}\log\left(\frac{KS}{N_1}\right).
\end{aligned}$$

For the second term, we denote

$$h(\varepsilon)=\sum_{\substack{i:i\neq 1,\\ \gamma_i\leq 2\gamma_1}}\mathbb{P}\left(\hat{r}=i,\,\hat{\mu}^{\mathcal{B}}-\mu_i>\varepsilon,G\right)$$

by the probability of interest. We have

$$\begin{aligned}
&h(\varepsilon)\\
&=\sum_{\substack{i:i\neq 1,\\ \gamma_i\leq 2\gamma_1}}\mathbb{P}\left(\hat{r}=i,\,\hat{\mu}^{\mathcal{B}}-\mu_i>\varepsilon,G\right)\\
&\leq\sum_{\substack{i:i\neq 1,\\ \gamma_i\leq 2\gamma_1}}\mathbb{P}\left(\hat{r}=i,\,\hat{\mu}^{\mathcal{B}}-\mu_i>\varepsilon\right)\\
&\leq\sum_{\substack{i:i\neq 1,\\ \gamma_i\leq 2\gamma_1}}\exp\left(-\frac{1}{2}N_i\left(\varepsilon-\gamma_1+\Delta_i\right)_+^2\right) && (\text{use Lemma 25})\\
&\leq\sum_{\substack{i:i\neq 1,\\ \gamma_i\leq 2\gamma_1}}\exp\left(-\frac{1}{2}N_i\left(\varepsilon-\gamma_1\right)_+^2\right)\\
&\leq K\exp\left(-\frac{1}{2}\min_{\substack{i:i\neq 1,\\ \gamma_i\leq 2\gamma_1}}\left(\varepsilon-\gamma_1\right)_+^2\right)\\
&=K\exp\left(-\frac{1}{2}N_u\left(\varepsilon-\gamma_1\right)_+^2\right).
\end{aligned}$$

We also have $\mathbb{P}(G)\leq\frac{2K^2}{K^{2\lambda}}$ thus

$$h(\varepsilon)$$

$$
\begin{aligned}
&= \sum_{\substack{i: i \neq 1, \\ \gamma_i \leq 2\gamma_1}} \mathbb{P}\left(\hat{r} = i,\, \hat{\mu}^{\mathcal{B}} - \mu_i > \varepsilon, G\right) \\
&\leq \sum_{\substack{i: i \neq 1, \\ \gamma_i \leq 2\gamma_1}} \mathbb{P}(G) \\
&\leq \sum_{\substack{i: i \neq 1, \\ \gamma_i \leq 2\gamma_1}} \frac{2K^2}{K^{2\lambda}} && \left(\text{use } \mathbb{P}(G) \leq \tfrac{2K^2}{K^{2\lambda}}\right) \\
&\leq \frac{2K^3}{K^{2\lambda}}.
\end{aligned}
$$

Let $q_1 = \frac{2K}{K^\lambda}$, $q_2 = \frac{2K^3}{K^{2\lambda}}$, $z_1 = \frac{N_u}{32}$, and $\varepsilon_0 = 2\gamma_1 \vee \sqrt{\frac{32\log(K)}{N_u}}$. We claim that in the regime of $\varepsilon \geq \varepsilon_0$, we have

$$
h(\varepsilon) \leq q_1 \exp\left(-z_1 \varepsilon^2\right)
$$

and in the regime of $\varepsilon < \varepsilon_0$, we have $h(\varepsilon) \leq q_2$.

We prove the claim as follows. In the regime of $\varepsilon \geq \varepsilon_0$, we have

$$
\begin{aligned}
& h(\varepsilon) \\
&\leq K \exp\left(-\frac{1}{2} N_u (\varepsilon - \gamma_1)_+^2\right) \wedge \frac{2K^3}{K^{2\lambda}} \\
&\leq \sqrt{K \exp\left(-\frac{1}{2} N_u (\varepsilon - \gamma_1)_+^2\right) \cdot \frac{2K^3}{K^{2\lambda}}} \\
&= \exp\left(-\frac{1}{4} N_u (\varepsilon - \gamma_1)_+^2\right) \\
&\leq \frac{2K^2}{K^\lambda} \exp\left(-\frac{1}{4} N_u \left(\frac{\varepsilon}{2}\right)^2\right) && (\text{use } \varepsilon \geq \varepsilon_0 = 2\gamma_1) \\
&= \frac{2K^2}{K^\lambda} \exp\left(-\frac{1}{16} N_u \varepsilon^2\right) \\
&= \frac{2K}{K^\lambda} \exp\left(-\frac{1}{16} N_u \varepsilon^2 + \log(K)\right) \\
&\leq \frac{2K}{K^\lambda} \exp\left(-\frac{1}{32} N_u \varepsilon^2\right). && \left(\text{use } \varepsilon \geq \varepsilon_0 \geq \sqrt{\tfrac{32\log(K)}{N_u}}\right)
\end{aligned}
$$

In the regime of $\varepsilon < \varepsilon_0$, we have

$$
\begin{aligned}
& h(\varepsilon) \\
&\leq \exp\left(-\frac{1}{2} N_u (\varepsilon - \gamma_1)_+^2\right) \wedge \frac{2K^3}{K^{2\lambda}} \\
&\leq \frac{2K^3}{K^{2\lambda}}.
\end{aligned}
$$

With the above claim, we use Lemma 31 to bound the integral

$$
\begin{aligned}
& \int_0^\infty h(\varepsilon) 2\varepsilon \,\mathrm{d}\varepsilon \\
&\leq q_2 \varepsilon_0^2 + q_1 \frac{1}{z_1} \\
&= \frac{2K^3}{K^{2\lambda}} \left(2\gamma_1 \vee \sqrt{\frac{32\log(K)}{N_u}}\right)^2 + \frac{2K}{K^\lambda} \frac{32}{N_u} \\
&\leq \gamma_1^2 + \frac{8K^3}{K^{2\lambda}} \frac{32\log(K)}{N_u} + \frac{2K}{K^\lambda} \frac{32}{N_u}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{8K^3}{K^{2\lambda}} \frac{18}{N_1} \log\left(\left(\frac{KS}{N_1}\right)^{2\lambda}\right) + \frac{8K^3}{K^{2\lambda}} \frac{32\log(K)}{N_u} + \frac{2K}{K^\lambda} \frac{32}{N_u} && \left(\text{use } \gamma_i = \sqrt{\tfrac{18}{N_i}\log\left(\left(\tfrac{KS}{N_i}\right)^{2\lambda}\right)}\right) \\
&= \frac{144K^3}{K^4 N_1} \log\left(\left(\frac{KS}{N_1}\right)^{4}\right) + \frac{256K^3\log(K)}{K^4 N_u} + \frac{64K}{K^2 N_u} && (\text{recall } \lambda = 2) \\
&= \frac{144}{KN_1} \log\left(\left(\frac{KS}{N_1}\right)^{4}\right) + \frac{128\log(K^2)}{KN_u} + \frac{64}{KN_u} \\
&\le \frac{144}{KN_1} \log\left(\left(\frac{KS}{N_1}\right)^{4}\right) + \frac{192\log(K^2)}{KN_u} && (\text{use } 1 \le \log(K^2) \text{ with } K \ge 2) \\
&\le \frac{144}{KN_1} \log\left(\left(\frac{KS}{N_1}\right)^{4}\right) + \frac{384\log(K)}{KN_u}.
\end{aligned}
$$

Lemma 42 states that for any $i \neq 1$ the condition $\gamma_i < 2\gamma_1$ implies that $N_i > \frac{N_1}{4} \frac{\log(K)}{\log\left(\frac{KS}{N_1}\right)}$. Thus, $N_u \ge \frac{N_1}{4} \frac{\log(K)}{\log\left(\frac{KS}{N_1}\right)}$.

$$
\begin{aligned}
&\frac{144}{KN_1} \log\left(\left(\frac{KS}{N_1}\right)^{4}\right) + \frac{384\log(K)}{KN_u} \\
&\le \frac{144}{KN_1} \log\left(\left(\frac{KS}{N_1}\right)^{4}\right) + \frac{384\log(K)}{KN_1} \frac{4\log\left(\frac{KS}{N_1}\right)}{\log(K)} \\
&= \frac{144}{KN_1} \log\left(\left(\frac{KS}{N_1}\right)^{4}\right) + \frac{1436}{KN_1} \log\left(\frac{KS}{N_1}\right) \\
&= \frac{2112}{KN_1} \log\left(\frac{KS}{N_1}\right).
\end{aligned}
$$

For the third term, we focus on arm i such that $i \neq 1$, $\gamma_i > 2\gamma_1$. We have

$$
\begin{aligned}
&\mathbb{P}\left(\hat{r} = i, \hat{\mu}^{\mathcal{B}} - \mu_i > \varepsilon, G\right) \\
&\le \mathbb{P}\left(\hat{r} = i, \hat{\mu}^{\mathcal{B}} - \mu_i > \varepsilon\right) \\
&\le \mathbb{P}\left(\hat{r} = i, \hat{\mu}_i - \mu_i > \varepsilon - \frac{1}{2}\gamma_i + \Delta_i\right) && (\text{use Lemma 23}) \\
&\le \exp\left(-\frac{1}{2} N_i \left(\varepsilon - \gamma_i + \Delta_i\right)_+^2\right).
\end{aligned}
$$

Lemma 30 states that

$$
\mathbb{P}\left(\hat{r} = i,\, \hat{\mu}_{\hat{r}} - \gamma_{\hat{r}} > \hat{\mu}_1 - \gamma_1\right) \le \left(\frac{N_i}{KS}\right)^{2\lambda}.
$$

We also have

$$
\begin{aligned}
&\mathbb{P}\left(\hat{r} = i, \hat{\mu}^{\mathcal{B}} - \mu_i > \varepsilon, G\right) \\
&\le \mathbb{P}(G) \\
&\le \frac{2K^2}{K^{2\lambda}}.
\end{aligned}
$$

Thus, we have

$$
\begin{aligned}
&\mathbb{P}\left(\hat{r} = i, \hat{\mu}^{\mathcal{B}} - \mu_i > \varepsilon, G\right) \\
&\le \exp\left(-\frac{1}{2} N_i \left(\varepsilon - \gamma_i + \Delta_i\right)_+^2\right) \wedge \left(\frac{N_i}{KS}\right)^{2\lambda} \wedge \frac{2K^2}{K^{2\lambda}}.
\end{aligned}
$$

Let $q_1 = \frac{2K}{K^\lambda}\left(\frac{N_i}{KS}\right)^\lambda$, $q_2 = \frac{2K}{K^\lambda}\left(\frac{N_i}{KS}\right)^\lambda$, $z_1 = \frac{N_i}{16}$, and $\varepsilon_0 = 2\gamma_i$. We claim that in the regime of $\varepsilon \geq \varepsilon_0$, we have

$$\mathbb{P}\left(\hat{r} = i, \hat{\mu}^{\mathcal{B}} - \mu_i > \varepsilon, G\right) \leq q_1 \exp\left(-z_1 \varepsilon^2\right)$$

and in the regime of $\varepsilon < \varepsilon_0$, we have $\mathbb{P}\left(\hat{r} = i, \hat{\mu}^{\mathcal{B}} - \mu_i > \varepsilon, G\right) \leq q_2$.

We prove the claim as follows. In the regime of $\varepsilon \geq \varepsilon_0$, we have

$$\begin{aligned}
&\mathbb{P}\left(\hat{r} = i, \hat{\mu}^{\mathcal{B}} - \mu_i > \varepsilon, G\right) \\
&\leq \exp\left(-\frac{1}{2}N_i\left(\varepsilon - \gamma_i\right)_+^2\right) \wedge \left(\frac{N_i}{KS}\right)^{2\lambda} \wedge \frac{2K^2}{K^{2\lambda}} \\
&\leq \sqrt{\exp\left(-\frac{1}{2}N_i\left(\varepsilon - \gamma_i\right)_+^2\right) \cdot \left(\frac{N_i}{KS}\right)^{2\lambda} \cdot \frac{2K^2}{K^{2\lambda}}} \\
&\leq \left(\frac{N_i}{KS}\right)^{\lambda} \frac{2K}{K^\lambda} \exp\left(-\frac{1}{4}N_i\left(\varepsilon - \gamma_i\right)_+^2\right) \\
&\leq \frac{2K}{K^\lambda}\left(\frac{N_i}{KS}\right)^{\lambda} \exp\left(-\frac{1}{16}N_i\varepsilon^2\right). && \text{(use } \varepsilon \geq eps_0 = 2\gamma_i\text{)}
\end{aligned}$$

In the regime of $\varepsilon < \varepsilon_0$, we have

$$\begin{aligned}
&\mathbb{P}\left(\hat{r} = i, \hat{\mu}^{\mathcal{B}} - \mu_i > \varepsilon, G\right) \\
&\leq \left(\frac{N_i}{KS}\right)^{2\lambda} \wedge \frac{2K^2}{K^{2\lambda}} \\
&\leq \sqrt{\left(\frac{N_i}{KS}\right)^{2\lambda} \cdot \frac{2K^2}{K^{2\lambda}}} \\
&\leq \frac{2K}{K^\lambda}\left(\frac{N_i}{KS}\right)^{\lambda}.
\end{aligned}$$

With the claim above, we use Lemma 31 to bound the integral

$$\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{r} = i,\, \hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon, G\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&\leq q_2\varepsilon_0^2 + q_1\frac{1}{z_1} \\
&\leq \frac{2K}{K^\lambda}\left(\frac{N_i}{KS}\right)^{\lambda} 4\gamma_i^2 + \frac{2K}{K^\lambda}\left(\frac{N_i}{KS}\right)^{\lambda}\frac{16}{N_i} \\
&\leq \frac{8K}{K^\lambda}\left(\frac{N_i}{KS}\right)^{\lambda}\frac{18}{N_i}\log\left(\left(\frac{KS}{N_i}\right)^{2\lambda}\right) + \frac{2K}{K^\lambda}\left(\frac{N_i}{KS}\right)^{\lambda}\frac{16}{N_i} \\
&= \frac{8K}{K^\lambda}\left(\frac{N_i}{KN_{\max}\sum_{j\in[K]}N_j}\right)^{\lambda}\frac{18}{N_i}\log\left(\left(\frac{KS}{N_i}\right)^{2\lambda}\right) + \frac{2K}{K^\lambda}\left(\frac{N_i}{KN_{\max}\sum_{j\in[K]}N_j}\right)^{\lambda}\frac{16}{N_i} \\
& && \text{(use } S = N_{\max}\textstyle\sum_{j\in[K]}N_j\text{)} \\
&\leq \frac{144K}{K^\lambda N_i}\left(\frac{1}{K\sum_{j\in[K]}N_j}\right)^{\lambda}\log\left(\left(\frac{KS}{N_i}\right)^{2\lambda}\right) + \frac{32K}{K^\lambda N_i}\left(\frac{1}{K\sum_{j\in[K]}N_j}\right)^{\lambda} \\
&= \frac{144K}{K^2 N_i}\left(\frac{1}{K\sum_{j\in[K]}N_j}\right)^{2}\log\left(\left(\frac{KS}{N_i}\right)^{4}\right) + \frac{32K}{K^2 N_i}\left(\frac{1}{K\sum_{j\in[K]}N_j}\right)^{2} && \text{(recall } \lambda = 2\text{)} \\
&= \frac{144}{K^3\left(\sum_{j\in[K]}N_j\right)^2}\log\left((KS)^4\right) + \frac{32}{K^3\left(\sum_{j\in[K]}N_j\right)^2} && \text{(use } N_i \geq 1\text{)}
\end{aligned}$$

$$\leq \frac{176}{K^3 \left(\sum_{j\in[K]} N_j\right)^2} \log\left((KS)^4\right) \qquad \text{(use } 1 \leq \log((KS)^4) \text{ with } K \geq 2\text{)}$$
$$= \frac{704}{K^3 \left(\sum_{j\in[K]} N_j\right)^2} \log\left((KS)\right).$$

Thus, for the third term, we obtain the integral

$$\begin{aligned}
&\sum_{\substack{i:i\neq 1,\\ \gamma_i > 2\gamma_1}} \int_0^\infty \mathbb{P}\left(\hat{r} = i,\, \hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon, G\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&\leq \sum_{\substack{i:i\neq 1,\\ \gamma_i > 2\gamma_1}} \frac{704}{K^3 \left(\sum_{j\in[K]} N_j\right)^2} \log\left(KS\right) \\
&\leq \frac{704}{K^2 \left(\sum_{j\in[K]} N_j\right)^2} \log\left(KS\right).
\end{aligned}$$

Combining all the terms, we have

$$\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon, G\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&\leq \int_0^\infty \mathbb{P}\left(\hat{r} = 1,\, \hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon, G\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&\quad + \int_0^\infty \sum_{\substack{i:i\neq 1,\\ \gamma_i \leq 2\gamma_1}} \mathbb{P}\left(\hat{r} = i,\, \hat{\mu}^{\mathcal{B}} - \mu_i > \varepsilon, G\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&\quad + \int_0^\infty \sum_{\substack{i:i\neq 1,\\ \gamma_i > 2\gamma_1}} \mathbb{P}\left(\hat{r} = i,\, \hat{\mu}^{\mathcal{B}} - \mu_i > \varepsilon, G\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&\leq \frac{272}{KN_1} \log\left(\frac{KS}{N_1}\right) \\
&\quad + \frac{2112}{KN_1} \log\left(\frac{KS}{N_1}\right) \\
&\quad + \frac{704}{K^2 \left(\sum_{j\in[K]} N_j\right)^2} \log\left(KS\right) \\
&\leq \frac{3088}{KN_1} \log\left(KS\right).
\end{aligned}$$

□

Lemma 23. In HAVER, we have

$$\hat{\mu}_{\hat{r}} - \gamma_{\hat{r}} \leq \hat{\mu}^{\mathcal{B}} \leq \hat{\mu}_{\hat{r}} + \frac{1}{2}\gamma_{\hat{r}}.$$

Proof. For the first inequality, by the definition of $\mathcal{B}$, $\forall i \in \mathcal{B}$, $\hat{\mu}_i \geq \hat{\mu}_{\hat{r}} - \gamma_{\hat{r}}$, therefore

$$\begin{aligned}
&\sum_{i\in\mathcal{B}} w_i \hat{\mu}_i \geq \hat{\mu}_{\hat{r}} - \gamma_{\hat{r}} \\
\Leftrightarrow\, &\hat{\mu}^{\mathcal{B}} \geq \hat{\mu}_{\hat{r}} - \gamma_{\hat{r}}.
\end{aligned}$$

For the second inequality, by the definition of $\hat{r}$, $\forall i \in \mathcal{B}$, $\hat{\mu}_{\hat{r}} - \gamma_{\hat{r}} \geq \hat{\mu}_i - \gamma_i$, therefore $\forall i \in \mathcal{B}$,

$$\begin{aligned}
&\hat{\mu}_i \leq \hat{\mu}_{\hat{r}} - \gamma_{\hat{r}} + \gamma_i \\
\Rightarrow\, &\hat{\mu}_i \leq \hat{\mu}_{\hat{r}} - \gamma_{\hat{r}} + \frac{3}{2}\gamma_{\hat{r}} \qquad \text{(use } \forall i \in \mathcal{B},\ \gamma_i \leq \tfrac{3}{2}\gamma_{\hat{r}}\text{)}
\end{aligned}$$

$$
\begin{aligned}
&\Leftrightarrow \hat{\mu}_i \le \hat{\mu}_{\hat{r}} + \frac{1}{2}\gamma_{\hat{r}} \\
&\Leftrightarrow \sum_{i \in \mathcal{B}} w_i \hat{\mu}_i \le \hat{\mu}_{\hat{r}} + \frac{1}{2}\gamma_{\hat{r}} \\
&\Leftrightarrow \hat{\mu}^{\mathcal{B}} \le \hat{\mu}_{\hat{r}} + \frac{1}{2}\gamma_{\hat{r}}.
\end{aligned}
$$

□

Lemma 24. In HAVER, we have

$$
\mathbb{P}\left(\hat{r} = 1,\, \hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon\right) \le \exp\left(-\frac{1}{2} N_1 \left(\varepsilon - \frac{1}{2}\gamma_1\right)_+^2\right).
$$

Proof. We have

$$
\begin{aligned}
&\mathbb{P}\left(\hat{r} = 1,\, \hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon\right) \\
&\le \mathbb{P}\left(\hat{r} = 1,\, \hat{\mu}_{\hat{r}} + \frac{1}{2}\gamma_{\hat{r}} - \mu_1 > \varepsilon\right) && \text{(use Lemma 23, } \hat{\mu}^{\mathcal{B}} \le \hat{\mu}_{\hat{r}} + \tfrac{1}{2}\gamma_{\hat{r}}\text{)} \\
&= \mathbb{P}\left(\hat{r} = 1,\, \hat{\mu}_1 - \mu_1 > \varepsilon - \frac{1}{2}\gamma_1\right) \\
&\le \exp\left(-\frac{1}{2} N_1 \left(\varepsilon - \frac{1}{2}\gamma_1\right)_+^2\right).
\end{aligned}
$$

□

Lemma 25. Let $i \in [K]$ be any arm such that $i \ne 1$, $\gamma_i \le 2\gamma_1$. In HAVER, we have

$$
\mathbb{P}\left(\hat{r} = i,\, \hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon\right) \le \exp\left(-\frac{1}{2} N_i \left(\varepsilon - \gamma_1 + \Delta_i\right)_+^2\right).
$$

Proof. We have

$$
\begin{aligned}
&\mathbb{P}\left(\hat{r} = i,\, \hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon\right) \\
&\le \mathbb{P}\left(\hat{r} = i,\, \hat{\mu}_{\hat{r}} + \frac{1}{2}\gamma_{\hat{r}} - \mu_1 > \varepsilon\right) && \text{(use Lemma 23, } \hat{\mu}^{\mathcal{B}} \le \hat{\mu}_{\hat{r}} + \tfrac{1}{2}\gamma_{\hat{r}}\text{)} \\
&= \mathbb{P}\left(\hat{r} = i,\, \hat{\mu}_i + \frac{1}{2}\gamma_i - \mu_1 > \varepsilon\right) \\
&= \mathbb{P}\left(\hat{r} = i,\, \hat{\mu}_i - \mu_1 > \varepsilon - \frac{1}{2}\gamma_i\right) \\
&\le \mathbb{P}\left(\hat{r} = i,\, \hat{\mu}_i - \mu_i > \varepsilon - \frac{1}{2}\gamma_i + \Delta_i\right) \\
&\le \mathbb{P}\left(\hat{r} = i,\, \hat{\mu}_i - \mu_i > \varepsilon - \gamma_1 + \Delta_i\right) && \text{(use } \gamma_i \le 2\gamma_1\text{)} \\
&\le \exp\left(-\frac{1}{2} N_i \left(\varepsilon - \gamma_1 + \Delta_i\right)_+^2\right).
\end{aligned}
$$

□

Lemma 26. Let $i \in [K]$ be any arm such that $i \ne 1$, $\gamma_i > 2\gamma_1$. In HAVER, we have

$$
\mathbb{P}\left(\hat{r} = i,\, \hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon\right) \le \exp\left(-\frac{1}{2} N_i \left(\varepsilon - \gamma_i + \Delta_i\right)_+^2\right).
$$

Proof. We have

$$
\mathbb{P}\left(\hat{r} = i, \hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon\right)
$$

$$
\begin{aligned}
&\le \mathbb{P}\left(\hat{r} = i,\, \hat{\mu}_{\hat{r}} + \frac{1}{2}\gamma_{\hat{r}} - \mu_1 > \varepsilon\right) && \text{(use Lemma 23, } \hat{\mu}^{\mathcal{B}} \le \hat{\mu}_{\hat{r}} + \tfrac{1}{2}\gamma_{\hat{r}}\text{)}\\
&= \mathbb{P}\left(\hat{r} = i,\, \hat{\mu}_i + \frac{1}{2}\gamma_i - \mu_1 > \varepsilon\right)\\
&= \mathbb{P}\left(\hat{r} = i,\, \hat{\mu}_i - \mu_1 > \varepsilon - \frac{1}{2}\gamma_i\right)\\
&= \mathbb{P}\left(\hat{r} = i,\, \hat{\mu}_i - \mu_i > \varepsilon - \frac{1}{2}\gamma_i + \Delta_i\right)\\
&\le \exp\left(-\frac{1}{2}N_i\left(\varepsilon - \frac{1}{2}\gamma_i + \Delta_i\right)_+^2\right)\\
&\le \exp\left(-\frac{1}{2}N_i\left(\varepsilon - \gamma_i + \Delta_i\right)_+^2\right).
\end{aligned}
$$

□

Lemma 27. In HAVER, we have

$$
\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon\right) 2\varepsilon \, \mathrm{d}\varepsilon \le \frac{720}{N_1}\log\left(KS\right).
$$

Proof. We have

$$
\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon\right) 2\varepsilon \, \mathrm{d}\varepsilon\\
\le &\int_0^\infty \mathbb{P}\left(\hat{r} = 1,\, \hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon\right) 2\varepsilon \, \mathrm{d}\varepsilon\\
&+ \int_0^\infty \sum_{\substack{i: i \ne 1,\\ \gamma_i \le 2\gamma_1}} \mathbb{P}\left(\hat{r} = i,\, \hat{\mu}^{\mathcal{B}} - \mu_i > \varepsilon, G\right) 2\varepsilon \, \mathrm{d}\varepsilon\\
&+ \int_0^\infty \sum_{\substack{i: i \ne 1,\\ \gamma_i > 2\gamma_1}} \mathbb{P}\left(\hat{r} = i,\, \hat{\mu}^{\mathcal{B}} - \mu_i > \varepsilon, G\right) 2\varepsilon \, \mathrm{d}\varepsilon.
\end{aligned}
$$

For the first term, from Lemma 24, we have

$$
\begin{aligned}
&\mathbb{P}\left(\hat{r} = 1,\, \hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon\right)\\
\le &\exp\left(-\frac{1}{2}N_1\left(\varepsilon - \frac{1}{2}\gamma_1\right)_+^2\right)
\end{aligned}
$$

Let $q_1 = q_2 = 1$, $z_1 = \frac{N_1}{8}$, and $\varepsilon_0 = \gamma_1$. We claim that in the regime of $\varepsilon \ge \varepsilon_0$, we have

$$
\mathbb{P}\left(\hat{r} = 1,\, \hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon\right) \le q_1 \exp\left(-z_1\varepsilon^2\right)
$$

and in the regime of $\varepsilon < \varepsilon_0$, we have $\mathbb{P}\left(\hat{r} = 1,\, \hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon, G\right) \le q_2$.

The second claim is trivial. We prove the first claim as follows. In the regime of $\varepsilon \ge \varepsilon_0$, we have

$$
\begin{aligned}
&\mathbb{P}\left(\hat{r} = 1,\, \hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon\right)\\
\le &\exp\left(-\frac{1}{2}N_1\left(\varepsilon - \frac{1}{2}\gamma_1\right)_+^2\right)\\
\le &\exp\left(-\frac{1}{2}N_1\left(\frac{\varepsilon}{2}\right)^2\right) && \text{(use } \varepsilon \ge \varepsilon_0 = \gamma_1\text{)}\\
= &\exp\left(-\frac{1}{8}N_1\varepsilon^2\right).
\end{aligned}
$$

With the above claim, we use Lemma 31 to bound the integral

$$\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{r} = 1,\, \hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&\le q_2 \varepsilon_0^2 + q_1 \frac{1}{z_1} \\
&= \gamma_1^2 + \frac{16}{N_1} \\
&= \frac{18}{N_1} \log\left(\left(\frac{KS}{N_1}\right)^{2\lambda}\right) + \frac{8}{N_1} && \left(\text{use } \gamma_i = \sqrt{\tfrac{18}{N_i} \log\left(\left(\tfrac{KS}{N_i}\right)^{2\lambda}\right)}\right) \\
&= \frac{18}{N_1} \log\left(\left(\frac{KS}{N_1}\right)^{4}\right) + \frac{8}{N_1} && (\text{recall } \lambda = 2) \\
&= \frac{26}{N_1} \log\left(\left(\frac{KS}{N_1}\right)^{4}\right) && \left(\text{use } 1 \le \log\left(\left(\tfrac{KS}{N_1}\right)^4\right) \text{ with } K \ge 2\right) \\
&= \frac{104}{N_1} \log\left(\frac{KS}{N_1}\right).
\end{aligned}$$

For the second term, we denote

$$h(\varepsilon) = \sum_{\substack{i: i \neq 1, \\ \gamma_i \le 2\gamma_1}} \mathbb{P}\left(\hat{r} = i,\, \hat{\mu}^{\mathcal{B}} - \mu_i > \varepsilon\right)$$

by the probability of interest. We have

$$\begin{aligned}
&h(\varepsilon) \\
&= \sum_{\substack{i: i \neq 1, \\ \gamma_i \le 2\gamma_1}} \mathbb{P}\left(\hat{r} = i,\, \hat{\mu}^{\mathcal{B}} - \mu_i > \varepsilon\right) \\
&\le \sum_{\substack{i: i \neq 1, \\ \gamma_i \le 2\gamma_1}} \exp\left(-\frac{1}{2} N_i \left(\varepsilon - \gamma_1 + \Delta_i\right)_+^2\right) && (\text{use Lemma 25}) \\
&\le \sum_{\substack{i: i \neq 1, \\ \gamma_i \le 2\gamma_1}} \exp\left(-\frac{1}{2} N_i \left(\varepsilon - \gamma_1\right)_+^2\right) \\
&\le K \exp\left(-\frac{1}{2} \min_{\substack{i: i \neq 1, \\ \gamma_i \le 2\gamma_1}} \left(\varepsilon - \gamma_1\right)_+^2\right) \\
&= K \exp\left(-\frac{1}{2} N_u \left(\varepsilon - \gamma_1\right)_+^2\right).
\end{aligned}$$

Let $q_1 = q_2 = 1$, $z_1 = \frac{N_u}{16}$, and $\varepsilon_0 = 2\gamma_1 \vee \sqrt{\frac{16 \log(K)}{N_u}}$. We claim that in the regime of $\varepsilon \ge \varepsilon_0$, we have

$$h(\varepsilon) \le q_1 \exp\left(-z_1 \varepsilon^2\right)$$

and in the regime of $\varepsilon < \varepsilon_0$, we have $h(\varepsilon) \le q_2$.

The second claim is trivial. We prove the first claim as follows. In the regime of $\varepsilon \ge \varepsilon_0$, we have

$$\begin{aligned}
&h(\varepsilon) \\
&= \exp\left(-\frac{1}{2} N_u \left(\varepsilon - \gamma_1\right)_+^2\right) \\
&\le \exp\left(-\frac{1}{2} N_u \left(\frac{\varepsilon}{2}\right)^2\right) && (\text{use } \varepsilon \ge \varepsilon_0 = 2\gamma_1)
\end{aligned}$$

$$
\begin{aligned}
&= \exp\left(-\frac{1}{8}N_u\varepsilon^2\right) \\
&= \exp\left(-\frac{1}{8}N_u\varepsilon^2 + \log(K)\right) \\
&\le \exp\left(-\frac{1}{16}N_u\varepsilon^2\right). && \left(\text{use } \varepsilon \ge \varepsilon_0 \ge \sqrt{\tfrac{16\log(K)}{N_u}}\right)
\end{aligned}
$$

With the above claim, we use Lemma 31 to bound the integral

$$
\begin{aligned}
&\int_0^\infty h(\varepsilon) 2\varepsilon \,\mathrm{d}\varepsilon \\
&\le q_2\varepsilon_0^2 + q_1\frac{1}{z_1} \\
&= \left(2\gamma_1 \vee \sqrt{\frac{16\log(K)}{N_u}}\right)^2 + \frac{16}{N_u} \\
&\le 4\gamma_1^2 + \frac{16\log(K)}{N_u} + \frac{16}{N_u} \\
&= 4\frac{18}{N_1}\log\left(\left(\frac{KS}{N_1}\right)^{2\lambda}\right) + \frac{16\log(K)}{N_u} + \frac{16}{N_u} && \left(\text{use } \gamma_i = \sqrt{\tfrac{18}{N_i}\log\left(\left(\tfrac{KS}{N_i}\right)^{2\lambda}\right)}\right) \\
&= \frac{72}{N_1}\log\left(\left(\frac{KS}{N_1}\right)^4\right) + \frac{16\log(K)}{N_u} + \frac{16}{N_u} && (\text{recall } \lambda = 2) \\
&= \frac{72}{N_1}\log\left(\left(\frac{KS}{N_1}\right)^4\right) + \frac{8\log(K^2)}{N_u} + \frac{16}{N_u} \\
&\le \frac{72}{N_1}\log\left(\left(\frac{KS}{N_1}\right)^4\right) + \frac{24\log(K^2)}{N_u} && (\text{use } 1 \le \log(K^2) \text{ with } K \ge 2) \\
&= \frac{72}{N_1}\log\left(\left(\frac{KS}{N_1}\right)^4\right) + \frac{48\log(K)}{N_u}.
\end{aligned}
$$

Lemma 42 states that for any $i \ne 1$ the condition $\gamma_i < 2\gamma_1$ implies that $N_i > \frac{N_1}{4}\frac{\log(K)}{\log\left(\frac{KS}{N_1}\right)}$. Thus, $N_u \ge \frac{N_1}{4}\frac{\log(K)}{\log\left(\frac{KS}{N_1}\right)}$.

$$
\begin{aligned}
&\frac{72}{N_1}\log\left(\left(\frac{KS}{N_1}\right)^4\right) + \frac{48\log(K)}{N_u} \\
&\le \frac{72}{N_1}\log\left(\left(\frac{KS}{N_1}\right)^4\right) + \frac{48\log(K)}{N_1}\frac{4\log\left(\frac{KS}{N_1}\right)}{\log(K)} \\
&= \frac{72}{N_1}\log\left(\left(\frac{KS}{N_1}\right)^4\right) + \frac{192}{N_1}\log\left(\frac{KS}{N_1}\right) \\
&= \frac{480}{N_1}\log\left(\frac{KS}{N_1}\right).
\end{aligned}
$$

For the third term, we focus on arm $i \in [K]$ such that $i \ne 1$, $\gamma_i > 2\gamma_1$. From Lemma 26, we have

$$
\begin{aligned}
&\mathbb{P}\left(\hat{r} = i, \hat{\mu}^{\mathcal{B}} - \mu_i > \varepsilon\right) \\
&\le \exp\left(-\frac{1}{2}N_i\left(\varepsilon - \gamma_i + \Delta_i\right)_+^2\right).
\end{aligned}
$$

Lemma 30 states that

$$
\mathbb{P}\left(\hat{r} = i,\, \hat{\mu}_{\hat{r}} - \gamma_{\hat{r}} > \hat{\mu}_1 - \gamma_1\right) \le \left(\frac{N_i}{KS}\right)^{2\lambda}.
$$

Thus, we have

$$\begin{aligned}&\mathbb{P}\left(\hat{r}=i,\hat{\mu}^{\mathcal{B}}-\mu_i>\varepsilon\right)\\ &\le \exp\left(-\frac{1}{2}N_i\left(\varepsilon-\gamma_i+\Delta_i\right)_+^2\right)\wedge\left(\frac{N_i}{KS}\right)^{2\lambda}.\end{aligned}$$

Let $q_1=\left(\frac{N_i}{KS}\right)^{\lambda}$, $q_2=\left(\frac{N_i}{KS}\right)^{2\lambda}$, $z_1=\frac{N_i}{16}$, and $\varepsilon_0=2\gamma_i$ We claim that in the regime of $\varepsilon\ge\varepsilon_0$, we have

$$\mathbb{P}\left(\hat{r}=i,\hat{\mu}^{\mathcal{B}}-\mu_i>\varepsilon\right)\le q_1\exp\left(-z_1\varepsilon^2\right)$$

and in the regime of $\varepsilon<\varepsilon_0$, we have $\mathbb{P}\left(\hat{r}=i,\hat{\mu}^{\mathcal{B}}-\mu_i>\varepsilon\right)\le q_2$.

We prove the claim as follows. In the regime of $\varepsilon\ge\varepsilon_0$, we have

$$\begin{aligned}&\mathbb{P}\left(\hat{r}=i,\hat{\mu}^{\mathcal{B}}-\mu_i>\varepsilon\right)\\ &\le \exp\left(-\frac{1}{2}N_i\left(\varepsilon-\gamma_i\right)_+^2\right)\wedge\left(\frac{N_i}{KS}\right)^{2\lambda}\\ &\le\sqrt{\exp\left(-\frac{1}{2}N_i\left(\varepsilon-\gamma_i\right)_+^2\right)\cdot\left(\frac{N_i}{KS}\right)^{2\lambda}}\\ &\le\left(\frac{N_i}{KS}\right)^{\lambda}\exp\left(-\frac{1}{4}N_i\left(\varepsilon-\gamma_i\right)_+^2\right)\\ &\le\left(\frac{N_i}{KS}\right)^{\lambda}\exp\left(-\frac{1}{8}N_i\varepsilon^2\right). && \text{(use } \varepsilon\ge\varepsilon_0=2\gamma_i\text{)}\end{aligned}$$

In the regime of $\varepsilon<\varepsilon_0$, we have $\mathbb{P}\left(\hat{r}=i,\hat{\mu}^{\mathcal{B}}-\mu_i>\varepsilon\right)\le\left(\frac{N_i}{KS}\right)^{2\lambda}$.

With the above claim, we use Lemma 31 to bound the integral

$$\begin{aligned}&\int_0^\infty\mathbb{P}\left(\hat{r}=i,\,\hat{\mu}^{\mathcal{B}}-\mu_1>\varepsilon,G\right)2\varepsilon\,\mathrm{d}\varepsilon\\ \le\,&q_2\varepsilon_0^2+q_1\frac{1}{z_1}\\ \le\,&\left(\frac{N_i}{KS}\right)^{\lambda}4\gamma_i^2+\left(\frac{N_i}{KS}\right)^{2\lambda}\frac{16}{N_i}\\ \le\,&\left(\frac{N_i}{KS}\right)^{\lambda}\frac{18}{N_i}\log\left(\left(\frac{KS}{N_i}\right)^{2\lambda}\right)+\left(\frac{N_i}{KS}\right)^{2\lambda}\frac{16}{N_i}\\ =\,&\left(\frac{N_i}{KN_{\max}\sum_{j\in[K]}N_j}\right)^{\lambda}\frac{18}{N_i}\log\left(\left(\frac{KS}{N_i}\right)^{2\lambda}\right)+\left(\frac{N_i}{KN_{\max}\sum_{j\in[K]}N_j}\right)^{\lambda}\frac{16}{N_i} && \text{(use } S=N_{\max}\textstyle\sum_{j\in[K]}N_j\text{)}\\ \le\,&\frac{18}{N_i}\left(\frac{1}{K\sum_{j\in[K]}N_j}\right)^{\lambda}\log\left(\left(\frac{KS}{N_i}\right)^{2\lambda}\right)+\frac{16}{N_i}\left(\frac{1}{K\sum_{j\in[K]}N_j}\right)^{\lambda} && \text{(use } N_i\le N_{\max}\text{)}\\ =\,&\frac{18}{N_i}\left(\frac{1}{K\sum_{j\in[K]}N_j}\right)^{2}\log\left(\left(\frac{KS}{N_i}\right)^{4}\right)+\frac{16}{N_i}\left(\frac{1}{K\sum_{j\in[K]}N_j}\right)^{2} && \text{(recall } \lambda=2\text{)}\\ =\,&\frac{18}{K\left(\sum_{j\in[K]}N_j\right)^2}\log\left((KS)^4\right)+\frac{16}{K\left(\sum_{j\in[K]}N_j\right)^2} && \text{(use } N_i\ge 1\text{)}\\ \le\,&\frac{34}{K\left(\sum_{j\in[K]}N_j\right)^2}\log\left((KS)^4\right) && \text{(use } 1\le\log((KS)^4)\text{ with } K\ge 2\text{)}\end{aligned}$$

$$= \frac{136}{K\left(\sum_{j\in[K]} N_j\right)^2} \log\left((KS)\right).$$

Thus, for the third term, we obtain the integral

$$\begin{aligned}
&\sum_{\substack{i:i\neq 1,\\ \gamma_i>2\gamma_1}} \int_0^\infty \mathbb{P}\left(\hat{r}=i,\, \hat{\mu}^{\mathcal{B}}-\mu_1>\varepsilon, G\right) 2\varepsilon\,\mathrm{d}\varepsilon \\
\leq& \sum_{\substack{i:i\neq 1,\\ \gamma_i>2\gamma_1}} \frac{136}{K\left(\sum_{j\in[K]} N_j\right)^2} \log\left((KS)\right) \\
\leq& \frac{136}{\left(\sum_{j\in[K]} N_j\right)^2} \log\left((KS)\right).
\end{aligned}$$

Combining all the terms, we have

$$\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}}-\mu_1>\varepsilon, G\right) 2\varepsilon\,\mathrm{d}\varepsilon \\
\leq& \int_0^\infty \mathbb{P}\left(\hat{r}=1,\, \hat{\mu}^{\mathcal{B}}-\mu_1>\varepsilon, G\right) 2\varepsilon\,\mathrm{d}\varepsilon \\
&+ \int_0^\infty \sum_{\substack{i:i\neq 1,\\ \gamma_i\leq 2\gamma_1}} \mathbb{P}\left(\hat{r}=i,\, \hat{\mu}^{\mathcal{B}}-\mu_i>\varepsilon, G\right) 2\varepsilon\,\mathrm{d}\varepsilon \\
&+ \int_0^\infty \sum_{\substack{i:i\neq 1,\\ \gamma_i> 2\gamma_1}} \mathbb{P}\left(\hat{r}=i,\, \hat{\mu}^{\mathcal{B}}-\mu_i>\varepsilon, G\right) 2\varepsilon\,\mathrm{d}\varepsilon \\
\leq& \frac{104}{N_1}\log\left(\frac{KS}{N_1}\right) \\
&+ \frac{480}{N_1}\log\left(\frac{KS}{N_1}\right) \\
&+ \frac{136}{\left(\sum_{j\in[K]} N_j\right)^2} \log\left((KS)\right) \\
\leq& \frac{720}{N_1}\log\left(KS\right).
\end{aligned}$$

□

Lemma 28. In the event of $G_2 = \left\{\exists i \in \mathcal{B}^*(\hat{r})\,\text{s.t.}\, i \notin \mathcal{B}\right\}$, HAVER achieves

$$\mathbb{P}\left(G_2\right) \leq \frac{2K^2}{K^{2\lambda}}.$$

Proof. Let i be an arm that satisfies the event G_2, i.e., $i \in \mathcal{B}^*(\hat{r})$, $i \notin \mathcal{B}$.

By the definition of $\mathcal{B}$, $i \notin \mathcal{B}$ means either (1) $\left(\hat{\mu}_i < \hat{\mu}_{\hat{r}} - \gamma_{\hat{r}},\, \gamma_i \leq \frac{3}{2}\gamma_{\hat{r}}\right)$ or (2) $(\gamma_i > f\gamma_{\hat{r}})$.

The condition (1) implies that

$$\begin{aligned}
&\hat{\mu}_i < \hat{\mu}_{\hat{r}} - \gamma_{\hat{r}} \\
\Leftrightarrow& \hat{\mu}_i - \mu_i - \hat{\mu}_{\hat{r}} + \mu_{\hat{r}} \leq -\gamma_{\hat{r}} - \mu_i + \mu_{\hat{r}} \\
\Rightarrow& \hat{\mu}_i - \mu_i - \hat{\mu}_{\hat{r}} + \mu_{\hat{r}} \leq -\gamma_{\hat{r}} + \mu_{\hat{r}} - \mu_s + \frac{1}{6}\gamma_s && \left(\text{use } i \in \mathcal{B}^*(\hat{r}),\, \mu_i \geq \mu_s - \tfrac{1}{6}\gamma_s\right) \\
\Rightarrow& \hat{\mu}_i - \mu_i - \hat{\mu}_{\hat{r}} + \mu_{\hat{r}} \leq -\gamma_{\hat{r}} + \frac{1}{6}\gamma_s - \frac{1}{6}\gamma_s + \frac{1}{6}\gamma_{\hat{r}} && \left(\text{by def of } s,\, \mu_s - \tfrac{1}{6}\gamma_s \geq \mu_{\hat{r}} - \tfrac{1}{6}\gamma_{\hat{r}}\right) \\
\Leftrightarrow& \hat{\mu}_i - \mu_i - \hat{\mu}_{\hat{r}} + \mu_{\hat{r}} \leq -\frac{5}{6}\gamma_{\hat{r}}.
\end{aligned}$$

Thus, we have

$$
\begin{aligned}
&\mathbb{P}\left(\exists i \in \mathcal{B}^*(\hat{r})\,\text{s.t.}\, i \notin \mathcal{B},\, \hat{\mu}_i - \mu_i + \mu_{\hat{r}} - \hat{\mu}_{\hat{r}} < -\frac{5}{6}\gamma_{\hat{r}}\right) \\
\leq& \sum_{r\in[K]} \sum_{i\in\mathcal{B}^*(r)} \mathbb{P}\left(\hat{\mu}_i - \mu_i + \mu_r - \hat{\mu}_r < -\frac{5}{6}\gamma_r\right) \\
\leq& \sum_{r\in[K]} \sum_{i\in\mathcal{B}^*(r)} \left(\mathbb{P}\left(\mu_r - \hat{\mu}_r < -\frac{1}{3}\gamma_r\right) + \mathbb{P}\left(\hat{\mu}_i - \mu_i < -\frac{1}{2}\gamma_r\right)\right) \\
\leq& \sum_{r\in[K]} \sum_{i\in\mathcal{B}^*(r)} \left(\exp\left(\frac{-N_r\gamma_r^2}{18}\right) + \exp\left(\frac{-N_i\gamma_r^2}{8}\right)\right) \\
\leq& \sum_{r\in[K]} \sum_{i\in\mathcal{B}^*(r)} \left(\exp\left(\frac{-N_r\gamma_r^2}{18}\right) + \exp\left(\frac{-N_i\left(\frac{2}{3}\right)^2\gamma_i^2}{8}\right)\right) && \left(\text{use } i \in \mathcal{B}^*(r),\, \gamma_i \leq \tfrac{3}{2}\gamma_r\right) \\
=& \sum_{r\in[K]} \sum_{i\in\mathcal{B}^*(r)} \left(\exp\left(\frac{-N_r\gamma_r^2}{18}\right) + \exp\left(\frac{-N_i\gamma_i^2}{18}\right)\right) \\
\leq& \sum_{r\in[K]} \sum_{i\in\mathcal{B}^*(r)} \left(\left(\frac{N_r}{KS}\right)^{2\lambda} + \left(\frac{N_s}{KS}\right)^{2\lambda}\right) && \left(\text{by def } \forall i \in [K],\, \gamma_i = \sqrt{\tfrac{18}{N_i}\log\left(\left(\tfrac{KS}{N_i}\right)^{2\lambda}\right)}\right) \\
\leq& \sum_{r\in[K]} \sum_{i\in\mathcal{B}^*(r)} \frac{2}{K^{2\lambda}} && (\text{use } \forall i,\, N_i \leq S) \\
\leq& \frac{2K^2}{K^{2\lambda}}.
\end{aligned}
$$

The condition (2) contradicts with $i \in \mathcal{B}^*(\hat{r})$ (i.e, $\gamma_i \leq \frac{3}{2}\gamma_{\hat{r}}$), thus,

$$\mathbb{P}\left(\exists i \in \mathcal{B}^*(\hat{r})\,\text{s.t.}\, i \notin \mathcal{B},\, \gamma_i > f\gamma_{\hat{r}}\right) = 0.$$

Therefore,

$$
\begin{aligned}
&\mathbb{P}(G_2) \\
\leq& \mathbb{P}\left(\exists i \in \mathcal{B}^*(\hat{r})\,\text{s.t.}\, i \notin \mathcal{B},\, \hat{\mu}_i < \hat{\mu}_{\hat{r}} - \gamma_{\hat{r}},\, \gamma_i \leq \frac{3}{2}\gamma_{\hat{r}}\right) + \mathbb{P}\left(\exists i \in \mathcal{B}^*(\hat{r})\,\text{s.t.}\, i \notin \mathcal{B},\, \gamma_i > \frac{3}{2}\gamma_{\hat{r}}\right) \\
\leq& \frac{2K^2}{K^{2\lambda}}.
\end{aligned}
$$

□

Lemma 29. In the event of $G_3 = \left\{\mathcal{B}^*(\hat{r}) \subseteq \mathcal{B},\, \exists i \notin \mathcal{B}^+(\hat{r})\,\text{s.t.}\, i \in \mathcal{B}\right\}$, HAVER achieves

$$\mathbb{P}(G_3) \leq \frac{2K^2}{K^{2\lambda}}.$$

Proof. Let i be an arm that satisfies the event G_3, i.e., $i \notin \mathcal{B}^+(r)$, $i \in \mathcal{B}$.

By the definition of $\mathcal{B}^+(r)$, $i \notin \mathcal{B}^+(r)$ means either (1) $(\mu_i < \mu_s - \frac{4}{3}\gamma_s - \frac{4}{3}\gamma_i,\, \gamma_i \leq \frac{3}{2}\gamma_{\hat{r}})$ or (2) $(\gamma_i > \frac{3}{2}\gamma_{\hat{r}})$.

By the definition of $\mathcal{B}$, since $i \in \mathcal{B}$, we have $\hat{\mu}_i \geq \hat{\mu}_{\hat{r}} - \gamma_{\hat{r}}$. With the condition (1), we have

$$
\begin{aligned}
&\hat{\mu}_i \geq \hat{\mu}_{\hat{r}} - \gamma_{\hat{r}} \\
\Leftrightarrow& \hat{\mu}_i - \hat{\mu}_{\hat{r}} \geq -\gamma_{\hat{r}} \\
\Rightarrow& \hat{\mu}_i - \hat{\mu}_s + \gamma_s - \gamma_{\hat{r}} \geq -\gamma_{\hat{r}} && (\text{by def of } \hat{r},\, \hat{\mu}_{\hat{r}} \geq \hat{\mu}_s - \gamma_s + \gamma_{\hat{r}}) \\
\Leftrightarrow& \hat{\mu}_i - \hat{\mu}_s \geq -\gamma_s \\
\Leftrightarrow& \hat{\mu}_i - \mu_i + \mu_s - \hat{\mu}_s \geq -\gamma_s + \mu_s - \mu_i
\end{aligned}
$$

$$\Rightarrow \hat{\mu}_i - \mu_i + \mu_s - \hat{\mu}_s \geq -\gamma_s + \frac{4}{3}\gamma_s + \frac{4}{3}\gamma_i \qquad \text{(use condition (1))}$$
$$\Leftrightarrow \hat{\mu}_i - \mu_i + \mu_s - \hat{\mu}_s \geq \frac{1}{3}\gamma_s + \frac{4}{3}\gamma_i.$$

Thus, we have

$$\begin{aligned}
&\mathbb{P}(G_3) \\
&= \mathbb{P}\left(\exists i \notin \mathcal{B}^+,\, i \in \mathcal{B},\, \hat{\mu}_i - \hat{\mu}_{\hat{r}} \geq -\gamma_{\hat{r}}\right) \\
&\leq \mathbb{P}\left(\exists i \notin \mathcal{B}^+,\, i \in \mathcal{B},\, \hat{\mu}_i - \mu_i + \mu_s - \hat{\mu}_s \geq \frac{1}{3}\gamma_s + \frac{4}{3}\gamma_i\right) \\
&\leq \sum_{i \notin \mathcal{B}^+} \mathbb{P}\left(\hat{\mu}_i - \mu_i + \mu_s - \hat{\mu}_s \geq \frac{1}{3}\gamma_s + \frac{4}{3}\gamma_i\right) \\
&\leq \sum_{i \notin \mathcal{B}^+} \left(\mathbb{P}\left(\mu_s - \hat{\mu}_s \geq \frac{1}{3}\gamma_s\right) + \mathbb{P}\left(\hat{\mu}_i - \mu_i \geq \frac{4}{3}\gamma_i\right)\right) \\
&\leq \sum_{i \notin \mathcal{B}^+} \left(\exp\left(-\frac{N_s\gamma_s^2}{18}\right) + \exp\left(-\frac{16N_i\gamma_i^2}{18}\right)\right) \\
&\leq \sum_{i \notin \mathcal{B}^+} \left(\exp\left(-\frac{N_s\gamma_s^2}{18}\right) + \exp\left(-\frac{N_i\gamma_i^2}{18}\right)\right) \\
&= \sum_{i \notin \mathcal{B}^+} \left(\left(\frac{N_s}{KS}\right)^{2\lambda} + \left(\frac{N_i}{KS}\right)^{2\lambda}\right) && \text{(by def } \forall i \in [K],\, \gamma_i = \sqrt{\tfrac{18}{N_i}\log\left(\left(\tfrac{KS}{N_i}\right)^{2\lambda}\right)}\text{)} \\
&\leq \sum_{i \notin \mathcal{B}^+} \frac{2}{K^{2\lambda}} && \text{(use } \forall i,\, N_i \leq S\text{)} \\
&\leq \frac{2K^2}{K^{2\lambda}}.
\end{aligned}$$

The condition (2) contradicts with $i \in \mathcal{B}$ (i.e., $\gamma_i \leq \frac{3}{2}\gamma_{\hat{r}}$), thus,

$$\mathbb{P}\left(\exists i \notin \mathcal{B}^+,\, i \in \mathcal{B},\, \gamma_i > \frac{3}{2}\gamma_{\hat{r}}\right) = 0.$$

Therefore,

$$\begin{aligned}
&\mathbb{P}(G_3) \\
&\leq \mathbb{P}\left(\exists i \notin \mathcal{B}^+,\, i \in \mathcal{B},\, \mu_i < \mu_s - \frac{4}{3}\gamma_s - \frac{4}{3}\gamma_i,\, \gamma_i \leq \frac{3}{2}\gamma_{\hat{r}}\right) + \mathbb{P}\left(\exists i \notin \mathcal{B}^+,\, i \in \mathcal{B},\, \gamma_i > \frac{3}{2}\gamma_{\hat{r}}\right) \\
&\leq \frac{2K^2}{K^{2\lambda}}.
\end{aligned}$$

□

Lemma 30. Let any arm i such that $i \neq 1$, $\gamma_i > 2\gamma_1$, HAVER achieves

$$\mathbb{P}\left(\hat{r} = i,\, \hat{\mu}_{\hat{r}} - \gamma_{\hat{r}} > \hat{\mu}_1 - \gamma_1\right) \leq \left(\frac{N_i}{KS}\right)^{2\lambda}.$$

Proof. By the definition of $\hat{r}$, we have $\hat{\mu}_{\hat{r}} - \gamma_{\hat{r}} \geq \hat{\mu}_1 - \gamma_1$. Therefore, with $\hat{r} = i$, we have

$$\begin{aligned}
&\mathbb{P}\left(\hat{r} = i,\, \hat{\mu}_{\hat{r}} - \gamma_{\hat{r}} \geq \hat{\mu}_1 - \gamma_1\right) \\
&\leq \mathbb{P}\left(\hat{r} = i,\, \hat{\mu}_i - \mu_i - \hat{\mu}_1 + \mu_1 \geq \gamma_i - \gamma_1 + \Delta_i\right) \\
&\leq \mathbb{P}\left(\hat{r} = i,\, \hat{\mu}_i - \mu_i - \hat{\mu}_1 + \mu_1 \geq \frac{1}{2}\gamma_i + \Delta_i\right) && \text{(use } \gamma_i > 2\gamma_1\text{)}
\end{aligned}$$

$$
\begin{aligned}
&\le \exp\left(-\frac{1}{2}\frac{1}{\left(\frac{1}{N_i}+\frac{1}{N_1}\right)}\left(\frac{1}{2}\gamma_i+\Delta_i\right)^2\right) \\
&\le \exp\left(-\frac{1}{8}\frac{1}{\left(\frac{1}{N_i}+\frac{1}{N_1}\right)}\left(\gamma_i+\Delta_i\right)^2\right).
\end{aligned}
$$

Lemma 43 states that for any $i \neq 1$ the condition $\gamma_i > 2\gamma_1$ implies that $N_i < N_1$. Therefore,

$$
\begin{aligned}
&\exp\left(-\frac{1}{8}\frac{1}{\left(\frac{1}{N_i}+\frac{1}{N_1}\right)}\left(\gamma_i+\Delta_i\right)^2\right) \\
&\le \exp\left(-\frac{1}{8}\frac{1}{\left(\frac{1}{N_i}+\frac{1}{N_i}\right)}\left(\gamma_i+\Delta_i\right)^2\right) \\
&= \exp\left(-\frac{1}{16}N_i\left(\gamma_i+\Delta_i\right)^2\right) \\
&\le \exp\left(-\frac{1}{16}N_i\gamma_i^2\right) && \text{(use } \Delta_i \ge 1\text{)} \\
&\le \exp\left(-\frac{1}{18}N_i\gamma_i^2\right) \\
&= \left(\frac{N_i}{KS}\right)^{2\lambda}. && \left(\text{use } \gamma_i = \sqrt{\tfrac{18}{N_i}\log\left(\left(\tfrac{KS}{N_i}\right)^{2\lambda}\right)}\right)
\end{aligned}
$$

□

Lemma 31. Let q_1, q_2, z_1, and ε_0 be some positive real values. Let f be a function of ε such that in the regime of $\varepsilon \ge \varepsilon_0$, we have

$$
f(\varepsilon) \le q_1 \exp\left(-z_1\varepsilon^2\right).
$$

and in the regime of $\varepsilon < \varepsilon_0$, we have $f(\varepsilon) \le q_2$. Then, we can bound the integral

$$
\int_0^\infty f(\varepsilon)2\varepsilon \,\mathrm{d}\varepsilon \le q_2\varepsilon_0^2 + q_1\frac{1}{z_1}.
$$

Proof. We bound the integral as follows

$$
\begin{aligned}
&\int_0^\infty f(\varepsilon)2\varepsilon \,\mathrm{d}\varepsilon \\
&= \int_0^{\varepsilon_0} f(\varepsilon)2\varepsilon \,\mathrm{d}\varepsilon + \int_{\varepsilon_0}^\infty f(\varepsilon)2\varepsilon \,\mathrm{d}\varepsilon \\
&\le \int_0^{\varepsilon_0} f(\varepsilon)2\varepsilon \,\mathrm{d}\varepsilon + \int_{\varepsilon_0}^\infty q_1\exp\left(-z_1\varepsilon^2\right)2\varepsilon \,\mathrm{d}\varepsilon && \text{(in the regime } \varepsilon \ge \varepsilon_0\text{, use } f(\varepsilon) \le q_1\exp\left(-z_1\varepsilon^2\right)\text{)} \\
&\le \int_0^{\varepsilon_0} q_2 2\varepsilon \,\mathrm{d}\varepsilon + \int_{\varepsilon_0}^\infty q_1\exp\left(-z_1\varepsilon^2\right)2\varepsilon \,\mathrm{d}\varepsilon && \text{(in the regime } \varepsilon \le \varepsilon_0\text{, use } f(\varepsilon) \le q_2\text{)} \\
&= q_2\varepsilon^2\Big|_0^{\varepsilon_0} - \frac{q_1}{z_1}\exp\left(-z_1\varepsilon^2\right)\Big|_{\varepsilon_0}^\infty \\
&= q_2\varepsilon_0^2 + q_1\frac{1}{z_1}\exp\left(-z_1\varepsilon_0^2\right) \\
&\le q_2\varepsilon_0^2 + q_1\frac{1}{z_1}. && \text{(use } \exp\left(-z_1\varepsilon_0^2\right) \le 1\text{)}
\end{aligned}
$$

□

Lemma 32. Let q_1 and z_1 be some positive real values. Let f be a function of ε. If we have

$$f(\varepsilon) \leq q_1 \exp\left(-z_1 \varepsilon^2\right).$$

Then, we can bound the integral

$$\int_0^\infty f(\varepsilon) 2\varepsilon \, \mathrm{d}\varepsilon \leq q_1 \frac{1}{z_1}.$$

Proof. We bound the integral as follows

$$\begin{aligned}
&\int_0^\infty f(\varepsilon) 2\varepsilon \, \mathrm{d}\varepsilon \\
\leq &\int_0^\infty q_1 \exp\left(-z_1 \varepsilon^2\right) 2\varepsilon \, \mathrm{d}\varepsilon \\
= &- q_1 \frac{1}{z_1} \exp\left(-z_1 \varepsilon^2\right) \Big|_0^\infty \\
= &\, q_1 \frac{1}{z_1}.
\end{aligned}$$

□

Lemma 33. With $q^2 \geq \frac{1}{9}$, in HAVER, we have

$$\int_0^\infty \mathbb{P}\left(\left|\hat{\mu}^{\mathcal{B}} - \mu_1\right| > \varepsilon, \, \exists i \in [K], |\hat{\mu}_i - \mu_i| \geq q\gamma_i\right) 2\varepsilon \, \mathrm{d}\varepsilon \leq \tilde{\mathcal{O}}\left(\frac{1}{KN_1}\right).$$

Proof. We denote $G = \left\{\exists i \in [K], |\hat{\mu}_i - \mu_i| \geq q\gamma_i\right\}$. We have

$$\begin{aligned}
&\mathbb{P}\left(\exists i \in [K], |\hat{\mu}_i - \mu_i| \geq q\gamma_i\right) \\
\leq &\sum_{i \in [K]} \mathbb{P}\left(|\hat{\mu}_i - \mu_i| \geq q\gamma_i\right) \\
\leq &\sum_{i \in [K]} 2 \exp\left(-\frac{1}{2} N_i q^2 \gamma_i^2\right) \\
= &\sum_{i \in [K]} 2 \exp\left(-\frac{1}{2} N_i q^2 \frac{18}{N_i} \ln\left(\left(\frac{KS}{N_i}\right)^{2\lambda}\right)\right) && \left(\text{use } \gamma_i = \sqrt{\tfrac{18}{N_i} \ln\left(\left(\tfrac{KS}{N_i}\right)^{2\lambda}\right)}\right) \\
\leq &\sum_{i \in [K]} 2 \left(\frac{N_i}{KS}\right)^{2\lambda} && \left(\text{use } q^2 \geq \tfrac{1}{9}\right) \\
\leq &\sum_{i \in [K]} 2 \left(\frac{1}{K}\right)^{2\lambda} && \left(\text{use } \forall i \in [K], \, N_i \leq S\right) \\
\leq &\frac{2K^2}{K^{2\lambda}}. && (\text{reminder we set } \lambda = 2)
\end{aligned}$$

Since G satisfies $\mathbb{P}(G) \leq \frac{2K^2}{K^{2\lambda}}$, we use Lemma 18 to obtain

$$\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon, \, G\right) 2\varepsilon \, \mathrm{d}\varepsilon \\
\leq &\min_m \left(\frac{704}{K^2 N_m} \log\left(\frac{KS}{N_m}\right) + \frac{8}{K^2} \Delta_m^2\right) \\
\leq &\frac{704}{K^2 N_1} \log\left(\frac{KS}{N_1}\right). && (\text{upper bounded by } m = 1)
\end{aligned}$$

Also, we use Lemma 22 to obtain

$$\int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon,\, G\right) 2\varepsilon \,\mathrm{d}\varepsilon \le \frac{3088}{KN_1} \log\left(KS\right).$$

Therefore,

$$\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\left|\hat{\mu}^{\mathcal{B}} - \mu_1\right| < \varepsilon,\, G\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&= \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 < -\varepsilon,\, G\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&\quad + \int_0^\infty \mathbb{P}\left(\hat{\mu}^{\mathcal{B}} - \mu_1 > \varepsilon,\, G\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&\le \frac{704}{K^2 N_1} \log\left(\frac{KS}{N_1}\right) + \frac{3088}{KN_1} \log\left(KS\right) \\
&\le \frac{3792}{KN_1} \log\left(KS\right).
\end{aligned}$$

□

Lemma 34. Let $n, d \ge 1$ be two integers. Then,

$$\max_{k=0}^{d} \frac{k^2}{\left(n+k\right)^2} \le \left(\log\left(\frac{n+d}{n}\right)\right)^2.$$

Proof. We have

$$\begin{aligned}
&\max_{k=0}^{d} \frac{k^2}{\left(n+k\right)^2} \\
&\le \left(\max_{k=0}^{d} \frac{k}{\left(n+k\right)}\right)^2 \\
&\le \left(\sum_{k=0}^{d} \frac{1}{\left(n+k\right)}\right)^2
\end{aligned}$$

Then, we bound the summation term with integral

$$\begin{aligned}
&\sum_{k=0}^{d} \frac{1}{\left(n+k\right)} \\
&\le \int_{k=0}^{d} \frac{1}{\left(n+k\right)} \\
&= \log\left(n+k\right)\Big|_0^d \\
&= \log\left(\frac{n+d}{n}\right).
\end{aligned}$$

Therefore,

$$\max_{k=0}^{d} \frac{k^2}{\left(n+k\right)^2} \le \left(\log\left(\frac{n+d}{n}\right)\right)^2.$$

□

Lemma 35. A sufficient condition for

$$K\left(c_1\sqrt{\frac{1}{N}}\right)^{\frac{1}{\alpha}} - 1 \ge K\left(\frac{c_1}{2}\sqrt{\frac{1}{N}}\right)^{\frac{1}{\alpha}}$$

is

$$N \le \frac{c_1^2}{4}\left(\frac{K}{\alpha}\log(2)\right)^{2\alpha}.$$

Proof. It suffices to prove the contraposition. Thus, we assume $K\left(c_1\sqrt{\frac{1}{N}}\right)^{\frac{1}{\alpha}} - 1 < K\left(\frac{c_1}{2}\sqrt{\frac{1}{N}}\right)^{\frac{1}{\alpha}}$ is true. Then,

$$\begin{aligned}
&K\left(c_1\sqrt{\frac{1}{N}}\right)^{\frac{1}{\alpha}} - 1 < K\left(\frac{c_1}{2}\sqrt{\frac{1}{N}}\right)^{\frac{1}{\alpha}} \\
\Leftrightarrow\ & 1 > K\left(c_1\sqrt{\frac{1}{N}}\right)^{\frac{1}{\alpha}}\left(1-\left(\frac{1}{2}\right)^{\frac{1}{\alpha}}\right) \\
\Leftrightarrow\ & 1 > \left(c_1\sqrt{\frac{1}{N}}\right)\left(K\left(1-\left(\frac{1}{2}\right)^{\frac{1}{\alpha}}\right)\right)^{\alpha} \\
\Leftrightarrow\ & N > c_1^2\left(K\left(1-\left(\frac{1}{2}\right)^{\frac{1}{\alpha}}\right)\right)^{2\alpha} \\
\Leftrightarrow\ & N > c_1^2\left(K\left(\frac{2^{\frac{1}{\alpha}}-1}{2^{\frac{1}{\alpha}}}\right)\right)^{2\alpha} \\
\Leftrightarrow\ & N > \frac{c_1^2}{4}\left(K\left(2^{\frac{1}{\alpha}}-1\right)\right)^{2\alpha} \\
\Rightarrow\ & N > \frac{1}{4c_1^2}\left(K\left(\frac{\log(2)}{\alpha}\right)\right)^{2\alpha} && \left(\text{use } \forall x,\ \frac{\log(x)}{\alpha} \le x^{\frac{1}{\alpha}}-1 \text{ choose } x=2\right) \\
\Leftrightarrow\ & N > \frac{c_1^2}{4}\left(\frac{K}{\alpha}\log(2)\right)^{2\alpha}.
\end{aligned}$$

Therefore, the sufficient condition for

$$K\left(\frac{1}{c_1}\sqrt{\frac{1}{N}}\right)^{\frac{1}{\alpha}} - 1 \ge K\left(\frac{1}{2c_1}\sqrt{\frac{1}{N}}\right)^{\frac{1}{\alpha}}$$

is

$$N \le \frac{c_1^2}{4}\left(\frac{K}{\alpha}\log(2)\right)^{2\alpha}.$$

□

Lemma 36. A sufficient condition for

$$K\left(c_2\sqrt{\frac{1}{N}}\right)^{\frac{1}{\alpha}} + 1 \le K\left(2c_2\sqrt{\frac{1}{N}}\right)^{\frac{1}{\alpha}}$$

is

$$N \le c_2^2\left(\frac{K}{\alpha}\log(2)\right)^{2\alpha}.$$

Proof. It suffices to prove the contraposition. Thus, we assume $K\left(c_2\sqrt{\frac{1}{N}}\right)^{\frac{1}{\alpha}} + 1 > K\left(2c_2\sqrt{\frac{1}{N}}\right)^{\frac{1}{\alpha}}$ is true.

Then,

$$
\begin{aligned}
& K\left(c_2\sqrt{\frac{1}{N}}\right)^{\frac{1}{\alpha}} + 1 > K\left(2c_2\sqrt{\frac{1}{N}}\right)^{\frac{1}{\alpha}} \\
\Leftrightarrow & 1 > K\left(c_2\sqrt{\frac{1}{N}}\right)^{\frac{1}{\alpha}}\left(2^{\frac{1}{\alpha}} - 1\right) \\
\Leftrightarrow & 1 > \left(c_2\sqrt{\frac{1}{N}}\right)\left(K\left(2^{\frac{1}{\alpha}} - 1\right)\right)^{\alpha} \\
\Leftrightarrow & N > c_2^2\left(K\left(2^{\frac{1}{\alpha}} - 1\right)\right)^{2\alpha} \\
\Rightarrow & N > c_2^2\left(K\left(\frac{\log(2)}{\alpha}\right)\right)^{2\alpha} && \left(\text{use } \forall x,\ \tfrac{\log(x)}{\alpha} \le x^{\frac{1}{\alpha}} - 1 \text{ choose } x = 2\right) \\
\Leftrightarrow & N > c_2^2\left(\frac{K}{\alpha}\log(2)\right)^{2\alpha}.
\end{aligned}
$$

Therefore, the sufficient condition for

$$K\left(c_2\sqrt{\frac{1}{N}}\right)^{\frac{1}{\alpha}} + 1 \le \left(2c_2\sqrt{\frac{1}{N}}\right)^{\frac{1}{\alpha}}$$

is

$$N \le c_2^2\left(\frac{K}{\alpha}\log(2)\right)^{2\alpha}.$$

□

Lemma 37. Consider the Poly(α) instance where $\forall i \ge 2$, $\Delta_i = \left(\frac{i}{K}\right)^{\alpha}$. Suppose $\mathcal{S} = \left\{i : \Delta_i \le c\sqrt{\frac{1}{N}}\right\}$. Assuming $K\left(\frac{c}{\sqrt{N}}\right)^{\frac{1}{\alpha}} \ge 2$, we have

$$|\mathcal{S}| = \left\lfloor K\left(\frac{c}{\sqrt{N}}\right)^{\frac{1}{\alpha}}\right\rfloor.$$

Proof. We have

$$
\begin{aligned}
& \Delta_i \le c\sqrt{\frac{1}{N}} \\
\Leftrightarrow & \left(\frac{i}{K}\right)^{\alpha} \le c\sqrt{\frac{1}{N}} \\
\Leftrightarrow & \frac{i}{K} \le \frac{c^{\frac{1}{\alpha}}}{N^{\frac{1}{2\alpha}}} \\
\Leftrightarrow & i \le \frac{Kc^{\frac{1}{\alpha}}}{N^{\frac{1}{2\alpha}}} \\
\Leftrightarrow & i \le K\left(\frac{c}{\sqrt{N}}\right)^{\frac{1}{\alpha}}.
\end{aligned}
$$

Therefore, $|\mathcal{S}| = \max\left\{i : \Delta_i \le c\sqrt{\frac{1}{N}}\right\} = \left\lfloor K\left(\frac{c}{\sqrt{N}}\right)^{\frac{1}{\alpha}}\right\rfloor$.

□

Lemma 38. Consider the Poly(α) instance where $\forall i \ge 2$, $\Delta_i = \left(\frac{i}{K}\right)^{\alpha}$. Suppose $\mathcal{B}^* = \left\{i : \Delta_i \le a_1\sqrt{\frac{1}{N}}\right\}$ and

$\mathcal{B}^+ = \left\{ i : \Delta_i \leq a_2\sqrt{\frac{1}{N}} \right\}$. Assuming $N \leq \left(\frac{a_1^2}{4} \wedge a_2^2\right)\left(\frac{K}{\alpha}\log(2)\right)^{2\alpha}$, we have

$$\frac{1}{|\mathcal{B}^*|} \leq \frac{1}{K\left(\frac{a_1}{2}\frac{1}{\sqrt{N}}\right)^{\frac{1}{\alpha}}},$$

$$\frac{\left|\mathcal{B}^+\right|}{|\mathcal{B}^*|} \leq \left(\frac{2a_2}{a_1}\right)^{\frac{1}{\alpha}},$$

and

$$\frac{1}{|\mathcal{B}^*|}\sum_{i=1}^{|\mathcal{B}^+|}\Delta_i \leq \frac{2a_2\left(\frac{4a_2}{a_1}\right)^{\frac{1}{\alpha}}}{(\alpha+1)\sqrt{N}}.$$

Proof. We prove the first inequality as follows. From Lemma 37, we have $|\mathcal{B}^*| = \left\lfloor K\left(\frac{a_1}{\sqrt{N}}\right)^{\frac{1}{\alpha}} \right\rfloor \geq K\left(\frac{c_1}{\sqrt{N}}\right)^{\frac{1}{\alpha}} - 1$.

We have the assumption $N \leq \frac{a_1^2}{4}\left(\frac{K}{\alpha}\log(2)\right)^{2\alpha}$. Lemma 35 states that condition $N \leq \frac{a_1^2}{4}\left(\frac{K}{\alpha}\log(2)\right)^{2\alpha}$ implies

$$K\left(\frac{a_1}{\sqrt{N}}\right)^{\frac{1}{\alpha}} - 1 \geq K\left(\frac{a_1}{2}\frac{1}{\sqrt{N}}\right)^{\frac{1}{\alpha}}.$$

Thus,

$$|\mathcal{B}^*| \geq K\left(\frac{a_1}{\sqrt{N}}\right)^{\frac{1}{\alpha}} - 1 \geq K\left(\frac{a_1}{2}\frac{1}{\sqrt{N}}\right)^{\frac{1}{\alpha}}. \tag{3}$$

Therefore,

$$\frac{1}{|\mathcal{B}^*|} \leq \frac{1}{K\left(\frac{a_1}{2}\frac{1}{\sqrt{N}}\right)^{\frac{1}{\alpha}}},$$

which concludes the first inequality.

We prove the second inequality as follows. From Lemma 37, we have $\left|\mathcal{B}^+\right| = \left\lfloor K\left(\frac{a_2}{\sqrt{N}}\right)^{\frac{1}{\alpha}} \right\rfloor \leq K\left(\frac{a_2}{\sqrt{N}}\right)^{\frac{1}{\alpha}}$. Thus,

$$\frac{\left|\mathcal{B}^+\right|}{|\mathcal{B}^*|} \leq \frac{K\left(\frac{a_2}{\sqrt{N}}\right)^{\frac{1}{\alpha}}}{K\left(\frac{a_1}{2}\frac{1}{\sqrt{N}}\right)^{\frac{1}{\alpha}}} = \left(\frac{2a_2}{a_1}\right)^{\frac{1}{\alpha}},$$

which concludes the second inequality.

We prove the third inequality as follows. We have the assumption $N \leq a_2^2\left(\frac{K}{\alpha}\log(2)\right)^{2\alpha}$. Lemma 36 states that condition $N \leq a_2^2\left(\frac{K}{\alpha}\log(2)\right)^{2\alpha}$ implies

$$K\left(\frac{a_2}{\sqrt{N}}\right)^{\frac{1}{\alpha}} + 1 \leq K\left(\frac{2a_2}{\sqrt{N}}\right)^{\frac{1}{\alpha}}.$$

Thus, we have

$$\left|\mathcal{B}^+\right| + 1 \leq K\left(\frac{a_2}{\sqrt{N}}\right)^{\frac{1}{\alpha}} + 1 \leq K\left(\frac{2a_2}{\sqrt{N}}\right)^{\frac{1}{\alpha}}. \tag{4}$$

$$\frac{\left|\mathcal{B}^+\right|}{|\mathcal{B}^*|} \leq \frac{K\left(a_2\sqrt{\frac{1}{N}}\right)^{\frac{1}{\alpha}}}{K\left(\frac{1}{a_1}\sqrt{\frac{1}{N}}\right)^{\frac{1}{\alpha}} - 1}$$

$$\begin{aligned}&\le \frac{K\left(a_2\sqrt{\frac{1}{N}}\right)^{\frac{1}{\alpha}}}{K\left(\frac{1}{2a_1}\sqrt{\frac{1}{N}}\right)^{\frac{1}{\alpha}}}\\ &= (2a_1a_2)^{\frac{1}{\alpha}}.\end{aligned}$$

We have

$$\begin{aligned}\sum_{i=1}^{|\mathcal{B}^+|}\Delta_i &= \sum_{i=1}^{|\mathcal{B}^+|}\left(\frac{i}{K}\right)^{\alpha}\\ &= \frac{1}{K^\alpha}\sum_{i=1}^{|\mathcal{B}^+|} i^\alpha\\ &\le \frac{1}{K^\alpha}\int_{i=1}^{|\mathcal{B}^+|+1} i^\alpha\\ &= \frac{1}{K^\alpha}\frac{\left(|\mathcal{B}^+|+1\right)^{\alpha+1}-1}{\alpha+1}\\ &\le \frac{1}{K^\alpha}\frac{\left(|\mathcal{B}^+|+1\right)^{\alpha+1}}{\alpha+1}\\ &\le \frac{1}{K^\alpha}\frac{\left(K\left(\frac{2a_2}{\sqrt{N}}\right)^{\frac{1}{\alpha}}\right)^{\alpha+1}}{\alpha+1}. && \text{(from 4)}\end{aligned}$$

Therefore,

$$\begin{aligned}\frac{1}{|\mathcal{B}^*|}\sum_{i=1}^{|\mathcal{B}^+|}\Delta_i &\le \frac{1}{|\mathcal{B}^*|}\frac{1}{K^\alpha}\frac{\left(K\left(\frac{2a_2}{\sqrt{N}}\right)^{\frac{1}{\alpha}}\right)^{\alpha+1}}{\alpha+1}\\ &\le \frac{1}{K\left(\frac{a_1}{2}\frac{1}{\sqrt{N}}\right)^{\frac{1}{\alpha}}}\frac{1}{K^\alpha}\frac{\left(K\left(\frac{2a_2}{\sqrt{N}}\right)^{\frac{1}{\alpha}}\right)^{\alpha+1}}{\alpha+1} && \text{(from 3)}\\ &= \frac{(2a_2)^{1+\frac{1}{\alpha}}}{\left(\frac{a_1}{2}\right)^{\frac{1}{\alpha}}}\cdot\frac{1}{(\alpha+1)\sqrt{N}}\\ &= \frac{2a_2\left(\frac{4a_2}{a_1}\right)^{\frac{1}{\alpha}}}{(\alpha+1)\sqrt{N}},\end{aligned}$$

which concludes the third inequality.

□

Lemma 39. A sufficient condition for $\Delta > \sqrt{\frac{c}{N}\log(K^2N)}$ is

$$N > \frac{2c}{\Delta^2}\log\left(\frac{2cK^2}{\Delta^2 e}\right).$$

Proof. It suffices to prove the contraposition. Thus, we assume $\Delta \le \sqrt{\frac{c}{N}\log(K^2N)}$ is true. Then,

$$\Delta \le \sqrt{\frac{c}{N}\log(K^2N)}$$

$$
\begin{aligned}
&\Leftrightarrow N \leq \frac{c}{\Delta^2} \log\left(K^2 N\right) \\
&\Leftrightarrow N = \frac{c}{\Delta^2} \log(N) + \frac{c}{\Delta^2} \log\left(K^2\right).
\end{aligned}
$$

We set $A = \frac{c}{\Delta^2}$ and $B = \frac{c}{\Delta^2} \log\left(K^2\right)$. Thus, we have

$$
\begin{aligned}
&\Delta \leq \sqrt{\frac{c}{N} \log(K^2 N)} \\
&\Leftrightarrow N = \frac{c}{\Delta^2} \log(N) + \frac{c}{\Delta^2} \log\left(K^2\right) \\
&\Leftrightarrow N = A \log(N) + B \\
&\Leftrightarrow N = A \log\left(\frac{N}{2A} \cdot 2A\right) + B \\
&\Leftrightarrow N = A \log\left(\frac{N}{2A}\right) + A \log(2A) + B \\
&\Rightarrow N \leq A\left(\frac{N}{2A} - 1\right) + A \log(2A) + B \\
&\Leftrightarrow N = \frac{N}{2} - A + A \log(2A) + B \\
&\Leftrightarrow N = \frac{N}{2} + A \log\left(\frac{2A}{e}\right) + B \\
&\Leftrightarrow N = 2A \log\left(\frac{2A}{e}\right) + 2B \\
&\Leftrightarrow N = \frac{2c}{\Delta^2} \log\left(\frac{2c}{\Delta^2 e}\right) + \frac{2c}{\Delta^2} \log\left(K^2\right) \\
&\Leftrightarrow N = \frac{2c}{\Delta^2} \log\left(\frac{2cK^2}{\Delta^2 e}\right).
\end{aligned}
$$

Therefore, a sufficient condition is

$$
N > \frac{2c}{\Delta^2} \log\left(\frac{2cK^2}{\Delta^2 e}\right).
$$

□

D MLCB's Theorem 5

Recall the following definitions that would be used in MLCB.

For each arm $i \in [K]$, we define

$$
\gamma_i := \sqrt{\frac{16}{N_i} \log\left(\left(\frac{KT}{N_i}\right)^2\right)},
$$

as its confidence width where $T = \sum_{j \in [K]} N_j$. We define

$$
\hat{r} := \arg\max_{i \in [K]} \hat{\mu}_i - \gamma_i
$$

as the output arm that has the highest lower confidence bound.

Theorem 5. *MLCB achieves*

$$
\mathrm{MSE}(\hat{\mu}^{MLCB}) = \tilde{\mathcal{O}}\left(\frac{1}{N_1}\right).
$$

Proof. We have

$$
\mathrm{MSE}(\hat{\mu}^{\mathrm{MLCB}})
$$

$$
\begin{aligned}
&= \mathbb{E}\left[(\hat{\mu}_{\hat{r}} - \mu_1)^2\right] \\
&= \int_0^\infty \mathbb{P}\left((\hat{\mu}_{\hat{r}} - \mu_1)^2 > \varepsilon\right) \mathrm{d}\varepsilon \\
&= \int_0^\infty \mathbb{P}\left(|\hat{\mu}_{\hat{r}} - \mu_1| > \varepsilon\right) 2\varepsilon \,\mathrm{d}\varepsilon && \text{(change of variable)} \\
&= \int_0^\infty \mathbb{P}\left(\hat{\mu}_{\hat{r}} - \mu_1 < -\varepsilon\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&\quad + \int_0^\infty \mathbb{P}\left(\hat{\mu}_{\hat{r}} - \mu_1 > \varepsilon\right) 2\varepsilon \,\mathrm{d}\varepsilon.
\end{aligned}
$$

We use claim Lemma 40 for the first term

$$\int_0^\infty \mathbb{P}\left(\hat{\mu}_{\hat{r}} - \mu_1 < -\varepsilon\right) 2\varepsilon \,\mathrm{d}\varepsilon \le \frac{84}{N_1} \log\left(\frac{KT}{N_1}\right).$$

and use claim Lemma 41 for the second term

$$\int_0^\infty \mathbb{P}\left(\hat{\mu}_{\hat{r}} - \mu_1 > \varepsilon\right) 2\varepsilon \,\mathrm{d}\varepsilon \le \frac{60}{N_1} \log\left(\frac{KT}{N_1}\right).$$

Combining the two terms concludes our proof. □

Lemma 40. MLCB achieves

$$\int_0^\infty \mathbb{P}\left(\hat{\mu}_{\hat{r}} - \mu_1 < -\varepsilon\right) 2\varepsilon \,\mathrm{d}\varepsilon \le \frac{84}{N_1} \log\left(\frac{KT}{N_1}\right),$$

where $\hat{\mu}_{\hat{r}} = \max_{i \in [K]} \hat{\mu}_i - \gamma_i$.

Proof. We have

$$
\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{\mu}_{\hat{r}} - \mu_1 < -\varepsilon\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
\le &\int_0^\infty \mathbb{P}\left(\hat{r} = 1,\, \hat{\mu}_{\hat{r}} - \mu_1 < -\varepsilon\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&+ \int_0^\infty \mathbb{P}\left(\hat{r} \ne 1,\, \hat{\mu}_{\hat{r}} - \mu_1 < -\varepsilon\right) 2\varepsilon \,\mathrm{d}\varepsilon.
\end{aligned}
$$

For the first term, we have

$$
\begin{aligned}
&\mathbb{P}\left(\hat{r} = 1,\, \hat{\mu}_{\hat{r}} - \mu_1 < -\varepsilon\right) \\
&= \mathbb{P}\left(\hat{r} = 1,\, \hat{\mu}_1 - \mu_1 < -\varepsilon\right) \\
&\le \exp\left(-\frac{1}{2} N_1 \varepsilon^2\right).
\end{aligned}
$$

We use Lemma 32 (with $q_1 = 1$ and $z_1 = \frac{1}{2} N_1$) to bound the integral

$$
\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{r} = 1,\, \hat{\mu}_{\hat{r}} - \mu_1 < -\varepsilon\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&\le q_1 \frac{1}{z_1} \\
&= \frac{2}{N_1}.
\end{aligned}
$$

For the second term, we have

$$
\begin{aligned}
&\mathbb{P}\left(\hat{r} \ne 1,\, \hat{\mu}_{\hat{r}} - \mu_1 < -\varepsilon\right) \\
&\le \mathbb{P}\left(\hat{r} \ne 1,\, \hat{\mu}_{\hat{r}} - \mu_1 < -\varepsilon\right) \\
&\le \mathbb{P}\left(\hat{r} \ne 1,\, \hat{\mu}_1 - \gamma_1 + \gamma_{\hat{r}} - \mu_1 < -\varepsilon\right) && \text{(use def of } \hat{r},\ \hat{\mu}_{\hat{r}} \ge \hat{\mu}_1 - \gamma_1 + \gamma_{\hat{r}}) \\
&= \mathbb{P}\left(\hat{r} \ne 1,\, \hat{\mu}_1 - \mu_1 < -\varepsilon + \gamma_1 - \gamma_{\hat{r}}\right) \\
&\le \mathbb{P}\left(\hat{r} \ne 1,\, \hat{\mu}_1 - \mu_1 < -\varepsilon + \gamma_1\right)
\end{aligned}
$$

$$
\begin{aligned}
&\le \mathbb{P}\left(\hat{\mu}_1 - \mu_1 < -\varepsilon + \gamma_1\right) \\
&\le \exp\left(-\frac{1}{2} N_1 \left(\varepsilon - \gamma_1\right)_+^2\right).
\end{aligned}
$$

Let $q_1 = q_2 = 1$, $z_1 = \frac{1}{8} N_1$, and $\varepsilon_0 = 2\gamma_1$. We claim that, in the regime of $\varepsilon \ge \varepsilon_0$, we have

$$
\mathbb{P}\left(\hat{r} \neq 1,\, \hat{\mu}_{\hat{r}} - \mu_1 < -\varepsilon\right) \le q_1 \exp\left(-z_1 \varepsilon^2\right)
$$

and in the regime of $\varepsilon < \varepsilon_0$, we have

$$
\mathbb{P}\left(\hat{r} \neq 1,\, \hat{\mu}_{\hat{r}} - \mu_1 < -\varepsilon\right) \le q_2.
$$

The second claim is trivial. We prove the first claim as follows. In the regime of $\varepsilon \ge \varepsilon_0$, we have

$$
\begin{aligned}
&\mathbb{P}\left(\hat{r} \neq 1,\, \hat{\mu}_{\hat{r}} - \mu_1 < -\varepsilon\right) \\
&\le \exp\left(-\frac{1}{2} N_1 \left(\varepsilon - \gamma_1\right)_+^2\right) \\
&\le \exp\left(-\frac{1}{2} N_1 \left(\frac{\varepsilon}{2}\right)^2\right) && \text{(use } \varepsilon \ge \varepsilon_0\text{)} \\
&= \exp\left(-\frac{1}{8} N_1 \varepsilon^2\right).
\end{aligned}
$$

With the above claim, we use Lemma 31 to bound the integral

$$
\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{r} \neq 1,\, \hat{\mu}_{\hat{r}} - \mu_1 < -\varepsilon\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
\le\, & q_2 \varepsilon_0^2 + q_1 \frac{1}{z_1} \\
=\, & 4\gamma_1^2 + \frac{8}{N_1} \\
=\, & 4 \frac{16}{N_1} \log\left(\frac{KT}{N_1}\right) + \frac{8}{N_1} \\
=\, & \frac{64}{N_1} \log\left(\frac{KT}{N_1}\right) + \frac{8}{N_1}.
\end{aligned}
$$

Combining the two terms, we have

$$
\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{\mu}_{\hat{r}} - \mu_1 < -\varepsilon\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
\le\, & \int_0^\infty \mathbb{P}\left(\hat{r} = 1,\, \hat{\mu}_{\hat{r}} - \mu_1 < -\varepsilon\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
& + \int_0^\infty \mathbb{P}\left(\hat{r} \neq 1,\, \hat{\mu}_{\hat{r}} - \mu_1 < -\varepsilon\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
\le\, & \frac{2}{N_1} + \frac{64}{N_1} \log\left(\frac{KT}{N_1}\right) + \frac{8}{N_1} \\
=\, & \frac{2}{N_1} + \frac{32}{N_1} \log\left(\left(\frac{KT}{N_1}\right)^2\right) + \frac{8}{N_1} \\
\le\, & \frac{42}{N_1} \log\left(\left(\frac{KT}{N_1}\right)^2\right) && \text{(use } 1 \le \log(K^2) \text{ with } K \ge 2\text{)} \\
=\, & \frac{84}{N_1} \log\left(\frac{KT}{N_1}\right).
\end{aligned}
$$

□

Lemma 41. MLCB achieves

$$\int_0^\infty \mathbb{P}\left(\hat{\mu}_{\hat{r}} - \mu_1 > \varepsilon\right) 2\varepsilon \, \mathrm{d}\varepsilon \le \frac{60}{N_1} \log\left(\frac{KT}{N_1}\right),$$

where $\hat{\mu}_{\hat{r}} = \max_{i \in [K]} \hat{\mu}_i - \gamma_i$.

Proof. We have

$$\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{\mu}_{\hat{r}} - \mu_1 > \varepsilon\right) 2\varepsilon \, \mathrm{d}\varepsilon \\
\le &\int_0^\infty \mathbb{P}\left(\hat{r} = 1,\, \hat{\mu}_1 - \mu_1 > \varepsilon\right) 2\varepsilon \, \mathrm{d}\varepsilon \\
&+ \sum_{\substack{i: i \neq 1, \\ \gamma_i < 2\gamma_1}} \int_0^\infty \mathbb{P}\left(\hat{r} = i,\, \hat{\mu}_i - \mu_1 > \varepsilon\right) 2\varepsilon \, \mathrm{d}\varepsilon + \sum_{\substack{i: i \neq 1, \\ \gamma_i > 2\gamma_1}} \int_0^\infty \mathbb{P}\left(\hat{r} = i,\, \hat{\mu}_i - \mu_1 > \varepsilon\right) 2\varepsilon \, \mathrm{d}\varepsilon.
\end{aligned}$$

We bound the probability and integral of each term separately as follows. For the first term, we have

$$\begin{aligned}
&\mathbb{P}\left(\hat{r} = 1,\, \hat{\mu}_{\hat{r}} - \mu_1 > \varepsilon\right) \\
=\,&\mathbb{P}\left(\hat{r} = 1,\, \hat{\mu}_1 - \mu_1 > \varepsilon\right) \\
\le\,&\exp\left(-\frac{1}{2} N_1 \varepsilon^2\right).
\end{aligned}$$

We use Lemma 32 (with $q_1 = 1$ and $z_1 = \frac{1}{2} N_1$) to bound the integral

$$\int_0^\infty \mathbb{P}\left(\hat{r} = 1,\, \hat{\mu}_{\hat{r}} - \mu_1 > \varepsilon\right) 2\varepsilon \, \mathrm{d}\varepsilon \le \frac{2}{N_1}.$$

For the second term, for any $i \neq 1$, $\gamma_i < 2\gamma_1$ we have

$$\begin{aligned}
&\sum_{\substack{i: i \neq 1, \\ \gamma_i < 2\gamma_1}} \mathbb{P}\left(\hat{r} = i,\, \hat{\mu}_i - \mu_1 > \varepsilon\right) \\
=&\sum_{\substack{i: i \neq 1, \\ \gamma_i < 2\gamma_1}} \mathbb{P}\left(\hat{r} = i,\, \hat{\mu}_i - \mu_i > \varepsilon + \Delta_i\right) \\
\le&\sum_{\substack{i: i \neq 1, \\ \gamma_i < 2\gamma_1}} \exp\left(-\frac{1}{2} N_i \left(\varepsilon + \Delta_i\right)^2\right) \\
\le&\sum_{\substack{i: i \neq 1, \\ \gamma_i < 2\gamma_1}} \exp\left(-\frac{1}{2} N_i \varepsilon^2\right) \\
\le&\sum_{\substack{i: i \neq 1, \\ \gamma_i < 2\gamma_1}} \exp\left(-\frac{1}{2} \min_{\substack{i: i \neq 1, \\ \gamma_i < 2\gamma_1}} N_i \varepsilon^2\right) \\
\le\,&K \exp\left(-\frac{1}{2} \min_{\substack{i: i \neq 1, \\ \gamma_i < 2\gamma_1}} N_i \varepsilon^2\right).
\end{aligned}$$

For this term, let $q_1 = q_2 = 1$, $z_1 = \frac{1}{4} \min_{\substack{i: i \neq 1, \\ \gamma_i < 2\gamma_1}} N_i$, and $\varepsilon_0 = \sqrt{\max_{\substack{i: i \neq 1, \\ \gamma_i < 2\gamma_1}} \frac{4\log(K)}{N_i}}$. We claim that, in the regime of $\varepsilon \ge \varepsilon_0$, we have

$$\sum_{\substack{i: i \neq 1, \\ \gamma_i < 2\gamma_1}} \mathbb{P}\left(\hat{r} = i,\, \hat{\mu}_i - \mu_1 > \varepsilon\right) \le \exp\left(-z_1 \varepsilon^2\right)$$

and, in the regime of $\varepsilon < \varepsilon_0$, we have

$$\sum_{\substack{i:i\neq 1,\\ \gamma_i<2\gamma_1}} \mathbb{P}\left(\hat{r}=i,\, \hat{\mu}_i-\mu_1>\varepsilon\right) \le q_2.$$

The second claim is trivial. We prove the first claim as follows. In the regime of $\varepsilon \ge \varepsilon_0$, we have

$$\begin{aligned}
&\sum_{\substack{i:i\neq 1,\\ \gamma_i<2\gamma_1}} \mathbb{P}\left(\hat{r}=i,\, \hat{\mu}_i-\mu_1>\varepsilon\right)\\
&\le K\exp\left(-\frac{1}{2}\min_{\substack{i:i\neq 1,\\ \gamma_i<2\gamma_1}} N_i\varepsilon^2\right)\\
&\le \exp\left(-\frac{1}{2}\min_{\substack{i:i\neq 1,\\ \gamma_i<2\gamma_1}} N_i\varepsilon^2+\log(K)\right)\\
&\le \exp\left(-\frac{1}{4}\min_{\substack{i:i\neq 1,\\ \gamma_i<2\gamma_1}} N_i\varepsilon^2\right). && \text{(use } \varepsilon\ge\varepsilon_0)
\end{aligned}$$

With the above claim, we use Lemma 31 to bound the integral

$$\begin{aligned}
&\int_0^\infty \sum_{\substack{i:i\neq 1,\\ \gamma_i<2\gamma_1}} \mathbb{P}\left(\hat{r}=i,\, \hat{\mu}_i-\mu_1>\varepsilon\right)2\varepsilon\,\mathrm{d}\varepsilon\\
&\le q_2\varepsilon_0^2+q_1\frac{1}{z_1}\\
&=\max_{\substack{i:i\neq 1,\\ \gamma_i<2\gamma_1}}\frac{4\log(K)}{N_i}+\max_{\substack{i:i\neq 1,\\ \gamma_i<2\gamma_1}}\frac{4}{N_i}\\
&=\max_{\substack{i:i\neq 1,\\ \gamma_i<2\gamma_1}}\frac{2\log(K^2)}{N_i}+\max_{\substack{i:i\neq 1,\\ \gamma_i<2\gamma_1}}\frac{4}{N_i}\\
&\le\max_{\substack{i:i\neq 1,\\ \gamma_i<2\gamma_1}}\frac{6\log(K^2)}{N_i} && \text{(use } 1\le\log(K^2) \text{ with } K\ge 2)\\
&=\max_{\substack{i:i\neq 1,\\ \gamma_i<2\gamma_1}}\frac{12\log(K)}{N_i}.
\end{aligned}$$

Lemma 42 states that for any $i\neq 1$ the condition $\gamma_i<2\gamma_1$ implies that $N_i>\frac{N_1}{4}\frac{\log(K)}{\log\left(\frac{KT}{N_1}\right)}$. Therefore,

$$\begin{aligned}
&\max_{\substack{i:i\neq 1,\\ \gamma_i<2\gamma_1}}\frac{12\log(K)}{N_i}\\
&\le\max_{\substack{i:i\neq 1,\\ \gamma_i<2\gamma_1}}\frac{12\log(K)}{N_1}\frac{4\log\left(\frac{KT}{N_1}\right)}{\log(K)}\\
&=\frac{48}{N_1}\log\left(\frac{KT}{N_1}\right).
\end{aligned}$$

From this point, let i be an arm such that $i\neq 1$, $\gamma_i>2\gamma_1$. For the third term, we have

$$\begin{aligned}
&\mathbb{P}\left(\hat{r}=i,\, \hat{\mu}_i-\mu_1>\varepsilon\right)\\
&=\mathbb{P}\left(\hat{r}=i,\, \hat{\mu}_i-\mu_i>\varepsilon+\Delta_i\right)\\
&\le\exp\left(-\frac{1}{2}N_i\left(\varepsilon+\Delta_i\right)^2\right)
\end{aligned}$$

$$\le \exp\left(-\frac{1}{2}N_i\varepsilon^2\right).$$

In addition, we have

$$\begin{aligned}
&\mathbb{P}\left(\hat{r}=i\right)\\
&\le \mathbb{P}\left(\hat{r}=i,\, \hat{\mu}_i-\gamma_i>\hat{\mu}_1-\gamma_1\right)\\
&= \mathbb{P}\left(\hat{r}=i,\, \hat{\mu}_i-\mu_i-\hat{\mu}_1+\mu_1>\gamma_i-\gamma_1+\Delta_i\right)\\
&\le \mathbb{P}\left(\hat{r}=i,\, \hat{\mu}_i-\mu_i-\hat{\mu}_1+\mu_1>\frac{1}{2}\gamma_i+\Delta_i\right) && \text{(use } \gamma_i>2\gamma_1\text{)}\\
&\le \exp\left(-\frac{1}{2}\frac{1}{\left(\frac{1}{N_i}+\frac{1}{N_1}\right)}\left(\frac{1}{2}\gamma_i+\Delta_i\right)^2\right)\\
&\le \exp\left(-\frac{1}{8}\frac{1}{\left(\frac{1}{N_i}+\frac{1}{N_1}\right)}\gamma_1^2\right).
\end{aligned}$$

Lemma 43 states that for any $i\neq 1$ the condition $\gamma_i>2\gamma_1$ implies that $N_i<N_1$. Therefore,

$$\begin{aligned}
&\exp\left(-\frac{1}{8}\frac{1}{\left(\frac{1}{N_i}+\frac{1}{N_1}\right)}\gamma_i^2\right)\\
&\le \exp\left(-\frac{1}{8}\frac{1}{\left(\frac{1}{N_i}+\frac{1}{N_i}\right)}\gamma_i^2\right) && \text{(use } N_i<N_1\text{)}\\
&= \exp\left(-\frac{1}{16}N_i\gamma_i^2\right)\\
&= \exp\left(-\frac{1}{16}N_i\frac{16}{N_i}\log\left(\left(\frac{KT}{N_i}\right)^2\right)\right)\\
&= \left(\frac{N_i}{KT}\right)^2.
\end{aligned}$$

Combining the two inequalities, we have

$$\begin{aligned}
&\mathbb{P}\left(\hat{r}=i,\, \hat{\mu}_i-\mu_1>\varepsilon\right)\\
&\le \left(\frac{N_i}{KT}\right)^2\wedge\exp\left(-\frac{1}{2}N_i\varepsilon^2\right)\\
&\le \sqrt{\left(\frac{N_i}{KT}\right)^2\exp\left(-\frac{1}{2}N_i\varepsilon^2\right)}\\
&= \frac{N_i}{KT}\exp\left(-\frac{1}{4}N_i\varepsilon^2\right).
\end{aligned}$$

We use Lemma 32 (with $q_1=\frac{N_i}{KT}$ and $z_1=\frac{1}{4}N_i$) to bound the integral

$$\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{r}=i,\, \hat{\mu}_i-\mu_1>\varepsilon\right)2\varepsilon\,\mathrm{d}\varepsilon\\
&\le q\frac{1}{z_1}\\
&\le \frac{N_i}{KT}\frac{4}{N_i}
\end{aligned}$$

$$\leq \frac{4}{KT}.$$

Therefore, for the third term, we obtain the integral

$$\begin{aligned}
&\sum_{\substack{i:i\neq 1,\\ \gamma_i>2\gamma_1}} \int_0^\infty \mathbb{P}\left(\hat{r}=i,\,\hat{\mu}_i-\mu_1>\varepsilon\right) 2\varepsilon\,\mathrm{d}\varepsilon \\
&\leq \sum_{\substack{i:i\neq 1,\\ \gamma_i>2\gamma_1}} \frac{4}{K\sum_{j\in[K]} N_j} \\
&\leq K\frac{4}{KT} \\
&= \frac{4}{T}.
\end{aligned}$$

Combining all the terms, we have

$$\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{\mu}_{\hat{r}}-\mu_1>\varepsilon\right) 2\varepsilon\,\mathrm{d}\varepsilon \\
&\leq \int_0^\infty \mathbb{P}\left(\hat{r}=1,\,\hat{\mu}_1-\mu_1>\varepsilon\right) 2\varepsilon\,\mathrm{d}\varepsilon \\
&\quad + \sum_{\substack{i:i\neq 1,\\ \gamma_i<2\gamma_1}} \int_0^\infty \mathbb{P}\left(\hat{r}=i,\,\hat{\mu}_i-\mu_1>\varepsilon\right) 2\varepsilon\,\mathrm{d}\varepsilon + \sum_{\substack{i:i\neq 1,\\ \gamma_i>2\gamma_1}} \int_0^\infty \mathbb{P}\left(\hat{r}\neq i,\,\hat{\mu}_i-\mu_1>\varepsilon\right) 2\varepsilon\,\mathrm{d}\varepsilon \\
&\leq \frac{2}{N_1}+\frac{48}{N_1}\log\left(\frac{KT}{N_1}\right)+\frac{4}{S} \\
&= \frac{2}{N_1}+\frac{48}{N_1}\log\left(\frac{KT}{N_1}\right)+\frac{4}{\sum_{j\in[K]} N_j} && \left(\text{use } T=\textstyle\sum_{j\in[K]} N_j\right) \\
&= \frac{2}{N_1}+\frac{24}{N_1}\log\left(\left(\frac{KT}{N_1}\right)^2\right)+\frac{4}{\sum_{j\in[K]} N_j} \\
&\leq \frac{30}{N_1}\log\left(\left(\frac{KT}{N_1}\right)^2\right) && \left(\text{use } 1\leq\log(K^2) \text{ with } K\geq 2\right) \\
&= \frac{60}{N_1}\log\left(\frac{KT}{N_1}\right).
\end{aligned}$$

□

Lemma 42. Let $\gamma_i = \sqrt{\frac{a}{N_i}\log\left(\left(\frac{KT}{N_i}\right)^s\right)}$ where $T \geq N_i$ and $s \geq 1$. For any $i \neq 1$, the condition $\gamma_i < b\gamma_1$ implies that $N_i > \frac{N_1}{b^2}\frac{\log(K)}{\log\left(\frac{KT}{N_1}\right)}$.

Proof. We have

$$\begin{aligned}
&\gamma_i < b\gamma_1 \\
&\Leftrightarrow \frac{a}{N_i}\log\left(\left(\frac{KT}{N_i}\right)^s\right) < b^2\frac{a}{N_1}\log\left(\left(\frac{KT}{N_1}\right)^s\right) \\
&\Leftrightarrow \frac{1}{N_i}\log\left(\frac{KT}{N_i}\right) < \frac{b^2}{N_1}\log\left(\frac{KT}{N_1}\right) \\
&\Leftrightarrow N_i > \frac{N_1}{b^2}\frac{\log\left(\frac{KT}{N_i}\right)}{\log\left(\frac{KT}{N_1}\right)}
\end{aligned}$$

$$\Rightarrow N_i > \frac{N_1}{b^2} \frac{\log(K)}{\log\left(\frac{KT}{N_1}\right)}. \qquad \text{(use } N_i \leq T\text{)}$$

$\square$

Lemma 43. Let $\gamma_i = \sqrt{\frac{a}{N_i} \log\left(\left(\frac{KT}{N_i}\right)^s\right)}$ where $T \geq N_i$ and $s \geq 1$. For $b \geq 2$, for any $i \neq 1$, the condition $\gamma_i > b\gamma_1$ implies that $N_i < N_1$.

Proof. We have

$$\begin{aligned}
&\gamma_i > b\gamma_1 \\
\Leftrightarrow\ & \frac{a}{N_i} \log\left(\left(\frac{KT}{N_i}\right)^s\right) > b^2 \frac{a}{N_1} \log\left(\left(\frac{KT}{N_1}\right)^s\right) \\
\Leftrightarrow\ & \frac{1}{N_i} \log\left(\frac{KT}{N_i}\right) > \frac{b^2}{N_1} \log\left(\frac{KT}{N_1}\right) \\
\Leftrightarrow\ & N_i < \frac{N_1}{b^2} \frac{\log\left(\frac{KT}{N_i}\right)}{\log\left(\frac{KT}{N_1}\right)}.
\end{aligned}$$

For any $c \geq 0$, there are two cases: (1) $N_i \geq c$ and (2) $N_i < c$.

The case $N_i \geq c$ implies that

$$N_i < \frac{N_1}{b^2} \frac{\log\left(\frac{KT}{N_i}\right)}{\log\left(\frac{KT}{N_1}\right)} < \frac{N_1}{b^2} \frac{\log\left(\frac{KT}{c}\right)}{\log\left(\frac{KT}{N_1}\right)}.$$

Therefore, for any $c \geq 0$, we have

$$N_i < \frac{N_1}{b^2} \frac{\log\left(\frac{KT}{c}\right)}{\log\left(\frac{KT}{N_1}\right)} \vee c.$$

Choosing $c = N_1$, we have

$$N_i < \frac{N_1}{b^2} \frac{\log\left(\frac{KT}{N_1}\right)}{\log\left(\frac{KT}{N_1}\right)} \vee N_1 < \frac{N_1}{b^2} \vee N_1 \leq N_1.$$

$\square$

E LEM's Lemma 4

Lemma 44. Assuming the arms follow Gaussian distributions, in two-arm instances with equal means (i.e., $\mu_1 = \mu_2$), there exists an absolute constant c_1 such that

$$\mathrm{MSE}(\hat{\mu}^{\mathrm{LEM}}) \geq c_1 \left(\frac{1}{\min_{i\in[2]} N_i}\right).$$

Proof. The proof naturally follows Lemma 45 with $\Delta_2 = 0$.

$\square$

Lemma 45. Assuming the arms follow Gaussian distributions, in two-arm instances, there exists an absolute constant c_1 such that

$$\mathrm{MSE}(\hat{\mu}^{\mathrm{LEM}}) \geq c_1 \left(\frac{1}{N_1} + \frac{1}{N_2} \exp\left(-14 N_2 \Delta_2^2\right)\right).$$

Proof. We denote by $\hat{a} = \arg\max_{i\in[2]} \hat{\mu}_i$ the arm with the larger empirical mean between the two arms.

We decompose the MSE as follows

$$
\begin{aligned}
&\mathrm{MSE}(\hat{\mu}^{\mathrm{LEM}}) \\
&= \mathbb{E}\left[(\hat{\mu}_{\hat{a}} - \mu_1)^2\right] \\
&= \int_0^\infty \mathbb{P}\left((\hat{\mu}_{\hat{a}} - \mu_1)^2 > \varepsilon^2\right) \mathrm{d}\varepsilon \\
&= \int_0^\infty \mathbb{P}\left(|\hat{\mu}_{\hat{a}} - \mu_1| > \varepsilon\right) 2\varepsilon \,\mathrm{d}\varepsilon && \text{(change of variable)} \\
&= \int_0^\infty \left(\mathbb{P}\left(\hat{\mu}_{\hat{a}} - \mu_1 > \varepsilon\right) + \mathbb{P}\left(\hat{\mu}_{\hat{a}} - \mu_1 < -\varepsilon\right)\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&\ge \int_0^\infty \mathbb{P}\left(\hat{\mu}_{\hat{a}} - \mu_1 > \varepsilon\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&= \int_0^\infty \mathbb{P}\left(\hat{a} = 1,\, \hat{\mu}_1 - \mu_1 > \varepsilon\right) 2\varepsilon \,\mathrm{d}\varepsilon + \int_0^\infty \mathbb{P}\left(\hat{a} = 2,\, \hat{\mu}_2 - \mu_1 > \varepsilon\right) 2\varepsilon \,\mathrm{d}\varepsilon.
\end{aligned}
$$

We bound the first probability as follows. We have

$$
\begin{aligned}
&\mathbb{P}\left(\hat{a} = 1,\, \hat{\mu}_1 - \mu_1 > \varepsilon\right) \\
&= \mathbb{P}\left(\hat{\mu}_1 - \hat{\mu}_2 \ge 0,\, \hat{\mu}_1 - \mu_1 > \varepsilon\right) && (\text{use } \hat{\mu}_1 = \hat{\mu}_{\hat{a}} \ge \hat{\mu}_2) \\
&= \mathbb{P}\left(\hat{\mu}_1 - \mu_1 - \hat{\mu}_2 + \mu_2 \ge -\Delta_2,\, \hat{\mu}_1 - \mu_1 > \varepsilon\right).
\end{aligned}
$$

Let $X_1 = \hat{\mu}_1 - \mu_1$ and $X_2 = \hat{\mu}_2 - \mu_2$, we further have

$$
\begin{aligned}
&\mathbb{P}\left(X_1 - X_2 \ge -\Delta_2,\, X_1 > \varepsilon\right) \\
&= \int_{x_1,x_2} \mathbb{1}\left\{x_1 - x_2 \ge -\Delta_2,\, x_1 > \varepsilon\right\} \mathrm{d}\mathbb{P}(x_2)\,\mathrm{d}\mathbb{P}(x_1) \\
&= \int_{x_1=\varepsilon}^\infty \int_{x_2=-\infty}^{x_1+\Delta_2} 1 \,\mathrm{d}\mathbb{P}(x_1)\,\mathrm{d}\mathbb{P}(x_2) \\
&\ge \int_{x_1=\varepsilon}^\infty \int_{x_2=-\infty}^{0} 1 \,\mathrm{d}\mathbb{P}(x_1)\,\mathrm{d}\mathbb{P}(x_2) && (\text{since } x_1 + \Delta_2 \ge \varepsilon + \Delta_2 \ge 0) \\
&= \int_{x_1=\varepsilon}^\infty \frac{1}{2} \,\mathrm{d}\mathbb{P}(x_2) && (x_1 \text{ follows Gaussian distribution with zero mean}) \\
&= \frac{1}{2}\,\mathbb{P}\left(X_1 \ge \varepsilon\right) \\
&= \frac{1}{2}\,\mathbb{P}\left(\hat{\mu}_1 - \mu_1 \ge \varepsilon\right) \\
&\ge \frac{1}{2}\frac{1}{4\sqrt{\pi}} \exp\left(-\frac{7}{2}N_1\varepsilon^2\right) && (\text{use Lemma 47 with variance } \sigma^2 = \tfrac{1}{N_1}) \\
&= \frac{1}{8\sqrt{\pi}} \exp\left(-\frac{7}{2}N_1\varepsilon^2\right).
\end{aligned}
$$

Thus, we have

$$
\begin{aligned}
&\int_0^\infty \mathbb{P}\left(\hat{a} = 1,\, \hat{\mu}_1 - \mu_1 > \varepsilon\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&\ge \int_0^\infty \frac{1}{8\sqrt{\pi}} \exp\left(-\frac{7}{2}N_1\varepsilon^2\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&= \frac{1}{8\sqrt{\pi}} \left(-\frac{2}{7N_1} \exp\left(-\frac{7}{2}N_1\varepsilon^2\right)\right) \Big|_0^\infty \\
&= \frac{2}{56\sqrt{\pi}}\frac{1}{N_1}.
\end{aligned}
$$

Next, we bound the second probability term,

$$\begin{aligned}
&\mathbb{P}\left(\hat{a}=2,\, \hat{\mu}_2-\mu_1>\varepsilon\right) \\
&= \mathbb{P}\left(\hat{\mu}_2-\hat{\mu}_1\geq 0,\, \hat{\mu}_2-\mu_1>\varepsilon\right) && (\text{use } \hat{\mu}_2=\hat{\mu}_{\hat{a}}\geq\hat{\mu}_1) \\
&= \mathbb{P}\left(\hat{\mu}_2-\mu_2-\hat{\mu}_1+\mu_1\geq\Delta_2,\, \hat{\mu}_2-\mu_2>\varepsilon+\Delta_2\right).
\end{aligned}$$

Let $X_2=\hat{\mu}_2-\mu_2$ and $X_1=\hat{\mu}_1-\mu_1$, we further have

$$\begin{aligned}
&\mathbb{P}\left(X_2-X_1\geq\Delta_2,\, X_2>\varepsilon+\Delta_2\right) \\
&= \int_{x_1,x_2} \mathbb{1}\left\{x_2-x_1\geq\Delta_2,\, x_2>\varepsilon+\Delta_2\right\} \mathrm{d}\mathbb{P}(x_1)\,\mathrm{d}\mathbb{P}(x_2) \\
&= \int_{x_2=\varepsilon+\Delta_2}^{\infty}\int_{x_1=-\infty}^{x_2-\Delta_2} 1\,\mathrm{d}\mathbb{P}(x_1)\,\mathrm{d}\mathbb{P}(x_2) \\
&\geq \int_{x_2=\varepsilon+\Delta_2}^{\infty}\int_{x_1=-\infty}^{0} 1\,\mathrm{d}\mathbb{P}(x_1)\,\mathrm{d}\mathbb{P}(x_2) && (\text{since } x_2-\Delta_2\geq\varepsilon+\Delta_2-\Delta_2\geq 0) \\
&= \int_{x_2=\varepsilon+\Delta_2}^{\infty} \frac{1}{2}\,\mathrm{d}\mathbb{P}(x_2) && (x_2 \text{ follows Gaussian distribution with zero mean}) \\
&= \frac{1}{2}\,\mathbb{P}\left(X_2>\varepsilon+\Delta_2\right) \\
&= \frac{1}{2}\,\mathbb{P}\left(\hat{\mu}_2-\mu_2>\varepsilon+\Delta_2\right) \\
&\geq \frac{1}{2}\frac{1}{4\sqrt{\pi}}\exp\left(-\frac{7}{2}N_2\left(\varepsilon+\Delta_2\right)^2\right) && (\text{use Lemma 47 with variance } \sigma^2=\tfrac{1}{N_2}) \\
&= \frac{1}{8\sqrt{\pi}}\exp\left(-\frac{7}{2}N_2\left(\varepsilon+\Delta_2\right)^2\right).
\end{aligned}$$

In the regime of $\varepsilon\geq\Delta_2$, we have

$$\begin{aligned}
&\mathbb{P}\left(\hat{a}=2,\, \hat{\mu}_2-\mu_1>\varepsilon\right) \\
&\geq \frac{1}{8\sqrt{\pi}}\exp\left(-\frac{7}{2}N_2\left(\varepsilon+\Delta_2\right)^2\right) \\
&\geq \frac{1}{8\sqrt{\pi}}\exp\left(-\frac{7}{2}N_2\left(2\varepsilon\right)^2\right) && (\text{use } \varepsilon\geq\Delta_2) \\
&= \frac{1}{8\sqrt{\pi}}\exp\left(-14N_2\varepsilon^2\right).
\end{aligned}$$

Thus, we have

$$\begin{aligned}
&\int_0^{\infty}\mathbb{P}\left(\hat{a}=2,\, \hat{\mu}_2-\mu_1>\varepsilon\right)2\varepsilon\,\mathrm{d}\varepsilon \\
&= \int_0^{\Delta_2}\mathbb{P}\left(\hat{a}=2,\, \hat{\mu}_2-\mu_1>\varepsilon\right)2\varepsilon\,\mathrm{d}\varepsilon + \int_{\Delta_2}^{\infty}\mathbb{P}\left(\hat{a}=2,\, \hat{\mu}_2-\mu_1>\varepsilon\right)2\varepsilon\,\mathrm{d}\varepsilon \\
&\geq \int_{\Delta_2}^{\infty}\mathbb{P}\left(\hat{a}=2,\, \hat{\mu}_2-\mu_1>\varepsilon\right)2\varepsilon\,\mathrm{d}\varepsilon \\
&\geq \int_{\Delta_2}^{\infty}\frac{1}{8\sqrt{\pi}}\exp\left(-14N_2\varepsilon^2\right)2\varepsilon\,\mathrm{d}\varepsilon \\
&= \frac{1}{8\sqrt{\pi}}\left(-\frac{1}{14N_2}\exp\left(-14N_2\varepsilon^2\right)\right)\Big|_{\Delta_2}^{\infty} \\
&= \frac{2}{112\sqrt{\pi}}\frac{1}{N_2}\exp\left(-14N_2\Delta_2^2\right).
\end{aligned}$$

Combining the results, we have

$$\int_0^{\infty}\mathbb{P}\left(\hat{\mu}_{\hat{a}}-\mu_1>\varepsilon\right)2\varepsilon\,\mathrm{d}\varepsilon$$

$$
\begin{aligned}
&= \int_0^\infty \mathbb{P}\left(\hat{a}=1,\, \hat{\mu}_1-\mu_1>\varepsilon\right) 2\varepsilon \,\mathrm{d}\varepsilon + \int_0^\infty \mathbb{P}\left(\hat{a}=2,\, \hat{\mu}_2-\mu_1>\varepsilon\right) 2\varepsilon \,\mathrm{d}\varepsilon \\
&\geq \frac{2}{56\sqrt{\pi}} \frac{1}{N_1} + \frac{2}{112\sqrt{\pi}} \frac{1}{N_2} \exp\left(-14 N_2 \Delta_2^2\right).
\end{aligned}
$$

which concludes our proof. □

Lemma 46 (Stirling's formula)**.** Let k and n be two positive integers such that $1 \leq k \leq n$. Then,

$$
\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{en}{k}\right)^k.
$$

Lemma 47 (Anti-concentration inequality (Abramowitz and Stegun, 1968))**.** For a Gaussian random variable $X \sim \mathcal{N}(\mu, \sigma^2)$ and any $\varepsilon > 0$, we have

$$
\mathbb{P}\left(|X-\mu|>\varepsilon\right) > \frac{1}{4\sqrt{\pi}} \exp\left(-\frac{7\varepsilon^2}{2\sigma^2}\right).
$$