Feasible Arm Identification: Supplementary Material

Julian Katz-Samuels Clayton Scott

June 5, 2018

A Outline

In Section C, we prove our lower bound for the feasible identification problem (Theorem 1). In Section D, we prove the upper bound for MD-UCBE (Theorem 2). In Section E, we prove the upper bound for MD-SAR (Theorem 3). In Section F, we prove Proposition 1 and the upper bound for MD-APT (Theorem 4). In Section G, we prove the key lemmas that unify our analyses of the three algorithms, namely, Lemmas 1, 2, and 3. In Section H, we prove some useful technical lemmas. In Section I, we extend our results to the feasible arm identification problem where P is convex. Finally, in Section J, we present additional experimental results.

Regarding the lower and upper bound proofs (Theorems 1, 2, 3, and 4), we note that we may assume that the realizations for each arm are drawn before the game has begun. Therefore, the empirical mean of an arm after t pulls is well-defined even if that arm has not been pulled t times.

B Notation

Let T_i denote the number of pulls of arm i after T rounds. Let $X_{i,j,t}$ denote the tth realization of the jth coordinate of ν_i. For the sake of brevity, we write Δ_i instead of $\Delta_i^{(\epsilon)}$.

C Lower Bound Proof

We note that the proof of Theorem 1 has some similarities to the proof of Theorem 1 of Locatelli et al. [2016]. The most important technical differences are *(i)* our novel lower bound construction with multidimensional distributions and *(ii)* our simple "chaining" argument that iteratively applies the well-known change-of-measure equation (6) in Audibert and Bubeck [2010] to relate $\mathcal{B}^0$ and $\mathcal{B}^i$.

Proof of Theorem 1. **Step 1: All of the problems have the same complexity.** The difference between problem $\mathcal{B}^0$ and $\mathcal{B}^i$ is the ith arm, i.e., the distributions ν_i and ν_i'. Since $\boldsymbol{\mu}_i \in P$ and $\boldsymbol{\mu}_i' \notin P$, by definition of H, it suffices to show that $\text{dist}(\boldsymbol{\mu}_i, \partial P) = \text{dist}(\boldsymbol{\mu}_i', P)$. By Lemma H.1, there is $m \in [M]$ such that $\boldsymbol{\tau}_m \in \{\boldsymbol{x} : \boldsymbol{a}_m^t \boldsymbol{x} = b_m\}$ and $\boldsymbol{\tau}_m$ is the projection of $\boldsymbol{\mu}_i$ onto $\{\boldsymbol{x} : \boldsymbol{a}_m^t \boldsymbol{x} = b_m\}$. Let $\boldsymbol{\tau}_i'$ denote the projection of $\boldsymbol{\mu}_i'$ onto $\{\boldsymbol{x} : \boldsymbol{a}_m^t \boldsymbol{x} = b_m\}$. We claim that $\boldsymbol{\tau}_i = \boldsymbol{\tau}_i'$. Using the closed form solution of the projection of a vector onto a hyperplane [Boyd and Vandenberghe, 2004],

$$
\begin{aligned}
\boldsymbol{\tau}_i &= \boldsymbol{\mu}_i + (\boldsymbol{a}_m^t \boldsymbol{\mu}_i - b_m)\boldsymbol{a}_m, \\
\boldsymbol{\tau}_i' &= \boldsymbol{\mu}_i' + (b_m - \boldsymbol{a}_m^t \boldsymbol{\mu}_i')\boldsymbol{a}_m \\
&= \boldsymbol{\mu}_i + 2(\boldsymbol{\tau}_i - \boldsymbol{\mu}_i) + (b_m - \boldsymbol{a}_m^t(\boldsymbol{\mu}_i + 2(\boldsymbol{\tau}_i - \boldsymbol{\mu}_i))\boldsymbol{a}_m \\
&= 2\boldsymbol{\tau}_i - \boldsymbol{\mu}_i - (b_m - \boldsymbol{a}_m^t \boldsymbol{\mu}_i)\boldsymbol{a}_m \\
&= \boldsymbol{\tau}_i,
\end{aligned}
$$

establishing the claim.

Then,

$$\text{dist}(\boldsymbol{\mu}_i', P) \geqslant \text{dist}(\boldsymbol{\mu}_i', \{\boldsymbol{x} : \boldsymbol{a}_m^t \boldsymbol{x} \leqslant b_m\}) = \left\|\boldsymbol{\mu}_i' - \boldsymbol{\tau}_i\right\|_2 \geqslant \text{dist}(\boldsymbol{\mu}_i', P)$$

where the last inequality follows since $\boldsymbol{\tau}_i \in P$. Thus,

$$\text{dist}(\boldsymbol{\mu}_i', P) = \left\|\boldsymbol{\mu}_i' - \boldsymbol{\tau}_i\right\|_2 = \left\|\boldsymbol{\mu}_i - \boldsymbol{\tau}_i\right\|_2 = \text{dist}(\boldsymbol{\mu}_i, \partial P).$$

Thus, $\mathcal{B}^0, \ldots, \mathcal{B}^K$ have the same problem complexity.

Step 2: Change of Measure For all $i \in [K]$, since $\text{dist}(\boldsymbol{\mu}_i, \partial P) > \epsilon$, there exists $d_i > 0$ such that $\text{dist}(\boldsymbol{\mu}_i, \partial P) = d_i + \epsilon$. We note that

$$\left\|\boldsymbol{\mu}_i' - \boldsymbol{\mu}_i\right\|_2 = 2\left\|\boldsymbol{\tau}_i - \boldsymbol{\mu}_i\right\|_2 = 2(d_i + \epsilon) = \Delta_i + d_i \leqslant 2\Delta_i. \tag{1}$$

Note that we can write ν_i as a product distribution $\nu_{i,1} \otimes \nu_{i,2} \otimes \ldots \otimes \nu_{i,D}$ where $\nu_{i,j} \sim N(\mu_{i,j}, 1)$ and $\nu_i' \coloneqq \nu_{i,1}' \otimes \nu_{i,2}' \otimes \ldots \otimes \nu_{i,D}'$ where $\nu_{i,j}' \sim N(\mu_{i,j}', 1)$. Let $l \leqslant D$ and define

$$\nu_i^{\prime,(l)} = \nu_{i,1}' \otimes \nu_{i,2}' \otimes \ldots \otimes \nu_{i,l-1}' \otimes \nu_{i,l}' \otimes \nu_{i,l+1} \otimes \ldots \otimes \nu_{i,D}.$$

Let $\mathcal{B}^{i,(l)}$ denote the product distribution

$$\nu_1 \otimes \ldots \otimes \nu_{i-1} \otimes \nu_i^{\prime,(l)} \otimes \nu_{i+1} \otimes \ldots \otimes \nu_K.$$

Define

$$\text{KL}_{k,l} \coloneqq \text{KL}(\nu_{k,l}', \nu_{k,l}) = \frac{1}{2}(\mu_{k,l}' - \mu_{k,l})^2$$

where we used the KL-divergence between two multivariate Gaussian random variables. Next, define for $1 \leqslant k \leqslant K$, $1 \leqslant l \leqslant D$, and $1 \leqslant t \leqslant T$,

$$\widehat{\text{KL}}_{k,l,t} \coloneqq \frac{1}{t}\sum_{s=1}^{t} \log(\frac{d\nu_{k,l}'(X_{k,l,s})}{d\nu_{k,l}(X_{k,l,s})}) = \frac{1}{t}\sum_{s=1}^{t}[\frac{1}{2}(\mu_{k,l}^2 - (\mu_{k,l}')^2) + (\mu_{k,l}' - \mu_{k,l})X_{k,l,s}]$$

where we used the definition of the pdf of Gaussian random variables. Note that $\mathbb{E}_{\nu_{k,l}}\widehat{\text{KL}}_{k,l,t} = \text{KL}_{k,l}$ and that

$$\text{Var}_{\nu_{k,l}}[\frac{1}{2}(\mu_{k,l}^2 - (\mu_{k,l}')^2) + (\mu_{k,l}' - \mu_{k,l})X_{k,l,s}] = (\mu_{k,l}' - \mu_{k,l})^2\text{Var}_{\nu_{k,l}}(X_{k,l,s}) = (\mu_{k,l}' - \mu_{k,l})^2.$$

Define the event

$$\Theta = \{\forall k \leqslant K, \forall t \leqslant T, \forall l \leqslant D : \widehat{\mathrm{KL}}_{k,l,t} - \mathrm{KL}_{k,l} \leqslant 2|\mu_{k,l} - \mu'_{k,l}|\sqrt{\frac{\log(4(\log(T)+1)KD)}{t}}\}.$$

Claim: $\mathbb{P}_{\mathcal{B}^0}(\Theta) \geqslant \frac{3}{4}$. Fix $k \leqslant K$ and $l \leqslant D$. $\widehat{\mathrm{KL}}_{k,l,t} - \mathrm{KL}_{k,l}$ is a sum of centered Gaussian random variables with variance $(\mu'_{k,l} - \mu_{k,l})^2$. Therefore, the sub-Gaussian norm of each term in the sum is $|\mu'_{k,l} - \mu_{k,l}|$. Let $u \in \{0, \dots, \lceil \log(T) \rceil\}$. By Lemma H.2,

$$\mathbb{P}_{\mathcal{B}^0}(\exists t \in [2^u, 2^{u+1}] : \widehat{\mathrm{KL}}_{k,l,t} - \mathrm{KL}_{k,l} \geqslant 2|\mu'_{k,l} - \mu_{k,l}|\sqrt{\frac{\log(4(\log(T)+1)KD)}{t}}) \leqslant \frac{1}{4(\log(T)+1)KD}.$$

Then a union bound over $k \leqslant K$, $u \in \{0, \dots, \lceil \log(T) \rceil\}, l \leqslant D$ yields that

$$\mathbb{P}_{\mathcal{B}^0}(\Theta^c) \leqslant \frac{1}{4} \tag{2}$$

establishing the claim.

Next, let $i \in [K]$ and define the event $\mathcal{A}_i = \{i \in \widehat{S}\}$. We lower bound $\mathbb{P}_{\mathcal{B}^i}(\mathcal{A}_i)$. Recall that T_i denotes the number of pulls of arm i after T rounds and let

$$t_i = \mathbb{E}_{\mathcal{B}^0} T_i.$$

and define the event

$$\Theta_i = \Theta \cap \mathcal{A}_i \cap \{T_i \leqslant 6t_i\}.$$

We use equation (6) from Audibert and Bubeck [2010], whose argument we briefly restate in the interest of making our paper more self-contained. Let E denote an event. Then,

$$\begin{aligned}
&\mathbb{E}_{\mathcal{B}^{i,(D-1)}}[\mathbf{1}\{E\}\exp(-T_i\widehat{\mathrm{KL}}_{i,D,T_i})] = \mathbb{E}_{\mathcal{B}^{i,(D-1)}}[\mathbf{1}\{E\}\prod_{s=1}^{T_i}\frac{d\nu_{i,D}}{d\nu'_{i,D}}(X_{i,D,s})] \\
&= \int \dots \int \mathbf{1}\{E\}\prod_{s=1}^{T_i}\frac{d\nu_{i,D}}{d\nu'_{i,D}}(X_{i,D,s})\Big[\prod_{k\neq i}\prod_{s=1}^{T_k} d\nu_{k,l}(X_{k,l,s})\Big]\Big[\prod_{s=1}^{T_i}\prod_{l\neq D} d\nu_{i,l}(X_{i,l,s})\Big]\prod_{s=1}^{T_i} d\nu'_{i,D}(X_{i,D,s}) \\
&= \mathbb{E}_{\mathcal{B}^i}[\mathbf{1}\{E\}].
\end{aligned} \tag{3}$$

We have the following series of inequalities:

$$\begin{aligned}
\mathbb{P}_{\mathcal{B}^i}(\mathcal{A}_i) &\geqslant \mathbb{P}_{\mathcal{B}^i}(\Theta_i) \\
&= \mathbb{E}_{\mathcal{B}^{i,(D-1)}}[\mathbf{1}\{\Theta_i\}\exp(-T_i\widehat{\mathrm{KL}}_{i,D,T_i})] && (4) \\
&\geqslant \mathbb{E}_{\mathcal{B}^{i,(D-1)}}[\mathbf{1}\{\Theta_i\}\exp(-T_i\frac{1}{2}(\mu_{i,D} - \mu'_{i,D})^2 - 2|\mu_{i,D} - \mu'_{i,D}|\sqrt{T_i\log(4(\log(T)+1)KD)})] && (5) \\
&\geqslant \mathrm{Pr}_{\mathcal{B}^{i,(D-1)}}(\Theta_i)\exp(-3t_i(\mu_{i,D} - \mu'_{i,D})^2 - 2|\mu_{i,D} - \mu'_{i,D}|\sqrt{6t_i\log(4(\log(T)+1)KD)})) && (6)
\end{aligned}$$

where equality (4) follows by equation (3), inequality (5) follows by Θ, and inequality (6) follows by $\{T_i \leqslant 6t_i\}$. Observe that we can repeat lines (4), (5), and (6) for $\Pr_{\mathcal{B}^{i,(D-1)}}(\Theta_i)$. Continuing in this manner for $l = 1, \ldots, D-1$ yields:

$$
\begin{aligned}
\mathbb{P}_{\mathcal{B}^i}(\mathcal{A}_i) &\geqslant \Pr_{\mathcal{B}^{i,(D-1)}}(\Theta_i)\exp(-3t_i(\mu_{i,D}-\mu'_{i,D})^2 - 2|\mu_{i,D}-\mu'_{i,D}|\sqrt{6t_i\log(4(\log(T)+1)KD)})) \\
&\geqslant \Pr_{\mathcal{B}^0}(\Theta_i)\exp(-3t_i\sum_{l=1}^{D}(\mu_{i,l}-\mu'_{i,l})^2 - |\mu_{i,l}-\mu'_{i,l}|\sqrt{24t_i\log(4(\log(T)+1)KD)})) && (7)\\
&\geqslant \Pr_{\mathcal{B}^0}(\Theta_i)\exp(-12t_i\Delta_i^2 - \left\|\boldsymbol{\mu}_i-\boldsymbol{\mu}'_i\right\|_1\sqrt{24t_i\log(4(\log(T)+1)KD)})) && (8)\\
&\geqslant \Pr_{\mathcal{B}^0}(\Theta_i)\exp(-12t_i\Delta_i^2 - \left\|\boldsymbol{\mu}_i-\boldsymbol{\mu}'_i\right\|_2\sqrt{D}\sqrt{24t_i\log(4(\log(T)+1)KD)})) && (9)\\
&\geqslant \Pr_{\mathcal{B}^0}(\Theta_i)\exp(-12t_i\Delta_i^2 - 2\Delta_i\sqrt{24t_iD\log(4(\log(T)+1)KD)})) && (10)\\
&\geqslant \Pr_{\mathcal{B}^0}(\Theta_i)\exp(-13t_i\Delta_i^2 - 24D\log(4(\log(T)+1)KD))) && (11)
\end{aligned}
$$

Line (8) follows by (1), line (9) follows by applying the inequality $\left\|\boldsymbol{x}\right\|_1 \leqslant \left\|\boldsymbol{x}\right\|_2\sqrt{D}$, line (10) follows by (1), and line (11) follows by the inequality $2ab \leqslant a^2+b^2$ with $a = \Delta_i\sqrt{t_i}$.

Step 3: Lower bounding $\Pr_{\mathcal{B}^0}(\Theta_i)$. Suppose that for some i it holds that $\Pr_{\mathcal{B}^0}(\mathcal{A}_i) < \frac{1}{2}$. Then,

$$\Pr_{\mathcal{B}^0}(\cap_{k\in[K]}\mathcal{A}_k) \leqslant \Pr_{\mathcal{B}^0}(\mathcal{A}_i) < \frac{1}{2}.$$

Observe that under $\mathcal{B}^0$, the event $(\cap_{k\in[K]}\mathcal{A}_k)^c$ implies that $\mathcal{L}_{T,\epsilon}(\widehat{S}) = 1$ since for all $k \in [K]$, $\boldsymbol{\mu}_k \in P$ and $\mathrm{dist}(\boldsymbol{\mu}_k, \partial P) \geqslant \epsilon$. Thus, the theorem follows since

$$\max_{i\in\{0,\ldots,K\}} \mathbb{E}_{\mathcal{B}^i}(\mathcal{L}_{T,\epsilon}(\widehat{S})) \geqslant \mathbb{E}_{\mathcal{B}^0}(\mathcal{L}_{T,\epsilon}(\widehat{S})) \geqslant \Pr_{\mathcal{B}^0}((\cap_{k\in[K]}\mathcal{A}_k)^c) > \frac{1}{2}.$$

Therefore, we may suppose for the remainder of the proof that $\Pr_{\mathcal{B}^0}(\mathcal{A}_j) \geqslant \frac{1}{2}$ for all $j \in [K]$. Fix $i \in [K]$. By Markov's inequality,

$$\Pr_{\mathcal{B}^0}(T_i > 6t_i) \leqslant \frac{\mathbb{E}_{\mathcal{B}^0}[T_i]}{6t_i} = \frac{1}{6}.$$

Then, using the above two inequalities and inequality (2), by a union bound,

$$\Pr_{\mathcal{B}^0}(\Theta_i^c) \leqslant \frac{1}{4} + \frac{1}{2} + \frac{1}{6} = \frac{11}{12},$$

concluding this step of the proof.

Step 4: Putting it together.

$$
\begin{aligned}
\max_{i\in\{1,\ldots,K\}} \Pr_{\mathcal{B}^i}(\mathcal{A}_i) &\geqslant \frac{1}{K}\sum_{i=1}^{K}\Pr_{\mathcal{B}^i}(\mathcal{A}_i) \\
&\geqslant \frac{1}{K}\sum_{i=1}^{K}\Pr_{\mathcal{B}^0}(\Theta_i)\exp(-13t_i\Delta_i^2 - 24D\log(4(\log(T)+1)KD))) && (12)\\
&\geqslant \frac{1}{12}\frac{1}{K}\sum_{i=1}^{K}\exp(-13t_i\Delta_i^2 - 24D\log(4(\log(T)+1)KD))) && (13)
\end{aligned}
$$

where in inequality (12) we used (11), in inequality (13) we used $\Pr_{\mathcal{B}^0}(\Theta_i) \geqslant \frac{1}{12}$. We claim that since $\sum_{i=1}^K t_i = T$, there exists some j such that $t_j \leqslant \frac{T}{H\Delta_j^2}$. Towards a contradiction, suppose that for all $i \in [K]$ $t_i > \frac{T}{H\Delta_i^2}$. Then,

$$T = \sum_{i \in [K]} t_i > \sum_{i \in [K]} \frac{T}{H\Delta_i^2} = T,$$

which is a contradiction. Then,

$$\begin{aligned} &\frac{1}{K}\sum_{i=1}^{K} \exp(-13 t_i \Delta_i^2 - 24D\log(4(\log(T)+1)KD))) \\ &\qquad \geqslant \exp(-13 t_j \Delta_j^2 - 24D\log(4(\log(T)+1)KD)) - \log(K)) \\ &\qquad \geqslant \exp(-13\frac{T}{H} - 24D\log(4(\log(T)+1)KD)) - \log(K)). \end{aligned}$$

Observe that under $\mathcal{B}^i$, the event $\mathcal{A}_i$ implies that $\mathcal{L}_{T,\epsilon}(\widehat{S}) = 1$ since $\text{dist}(\boldsymbol{\mu}_i', P) > \epsilon$. Thus,

$$\begin{aligned} \max_{i \in \{0,\ldots,K\}} \mathbb{E}_{\mathcal{B}^i}(\mathcal{L}_{T,\epsilon}(\widehat{S})) &\geqslant \max_{i \in \{1,\ldots,K\}} \mathbb{E}_{\mathcal{B}^i}(\mathcal{L}_{T,\epsilon}(\widehat{S})) \\ &\geqslant \max_{i \in \{1,\ldots,K\}} \Pr_{\mathcal{B}^i}(\mathcal{A}_i) \\ &\geqslant \frac{1}{12}\exp(-13\frac{T}{H} - 24D\log(4(\log(T)+1)KD)) - \log(K)) \\ &\geqslant \exp(-13\frac{T}{H} - 25D\log(48(\log(T)+1)KD))). \end{aligned}$$

□

D MD-UCBE Upper Bound Proof

Proof of Theorem 2. **Step 1: Defining an appropriate event.**

Let $\mathcal{N}$ be a minimal $\frac{1}{2}$-net on $\mathcal{S}^{D-1}$. Let $\delta > 0$ (we choose it later). Define the event

$$\Xi = \{\forall i \in [K], \forall \boldsymbol{y} \in \mathcal{N}, \forall r \in \{1,\ldots,T\} : |\boldsymbol{y}^t(\widehat{\boldsymbol{\mu}}_{i,r} - \boldsymbol{\mu}_i)| \leqslant \sqrt{\frac{a\delta^2}{4r}}\}.$$

By Lemma 2, on Ξ, for all $i \in [K]$ and for all $r \in [T]$,

$$\|\widehat{\boldsymbol{\mu}}_{i,r} - \boldsymbol{\mu}_i\|_2 \leqslant \sqrt{\frac{a\delta^2}{r}} \tag{14}$$

and

$$\Pr(\Xi) \geqslant 1 - 2(\log(T)+1)K5^D \exp(-a\frac{\delta^2}{16R^2}).$$

For the remainder of the proof, we suppose that Ξ holds.

Step 2: Lower bound the number of pulls for some arm.

Fix T. Recall that T_i denotes the number of pulls of arm i after T rounds. We claim that there exists an arm k such that it has been pulled after initialization and such that $T_k - 1 \geqslant \frac{T-K}{H\Delta_k^2}$ (for the remainder of the proof, let k denote one of these arms). If not, then we obtain the following contradiction.

$$T - K = \sum_{i=1}^{K}(T_i - 1) < \sum_{i=1}^{K}\frac{T-K}{H\Delta_i^2} = T - K.$$

For the remainder of the proof, let t denote the last time at which arm k was pulled. Then,

$$T_k(t) = T_k - 1 \geqslant \frac{T-K}{H\Delta_k^2}. \tag{15}$$

Step 3: Lower bound the number of pulls for each arm.

Lemma 1 and event Ξ imply that

$$|\widehat{\Delta}_{i,T_i(t)} - \Delta_i| \leqslant 2\sqrt{\frac{a\delta^2}{T_i(t)}} \tag{16}$$

for all $i \in [K]$. We choose $\delta = \frac{1}{10}$.

Arm k was pulled at time t, so that we have for all $i \in [K]$,

$$\widehat{\Delta}_{k,T_k(t)} - \sqrt{\frac{a}{T_k(t)}} \leqslant \widehat{\Delta}_{i,T_i(t)} - \sqrt{\frac{a}{T_i(t)}}. \tag{17}$$

Now,

$$\Delta_k + \sqrt{\frac{a}{T_i(t)}} \leqslant \widehat{\Delta}_{k,T_k(t)} + \sqrt{\frac{a}{T_i(t)}} + \frac{1}{5}\sqrt{\frac{a}{T_k(t)}} \tag{18}$$

$$\leqslant \frac{6}{5}\sqrt{\frac{a}{T_k(t)}} + \widehat{\Delta}_{i,T_i(t)} \tag{19}$$

where in inequality (18) we apply (16) and in inequality (19) we apply (17).

Rearranging (19), we obtain

$$\sqrt{\frac{a}{T_i(t)}} \leqslant \frac{6}{5}\sqrt{\frac{a}{T_k(t)}} - \Delta_k + \widehat{\Delta}_{i,T_i(t)}$$

$$\leqslant \Delta_k - \Delta_k + \widehat{\Delta}_{i,T_i(t)} \tag{20}$$

$$\leqslant \Delta_i + \frac{1}{5}\sqrt{\frac{a}{T_i(t)}}. \tag{21}$$

where inequality (20) follows by (15) and $0 \leqslant a \leqslant \frac{25}{36}\frac{T-K}{H}$ and inequality (21) follows by (16).

Rearranging (21) implies that

$$\frac{4}{5}\sqrt{\frac{a}{T_i(t)}} \leqslant \Delta_i.$$

Thus,

$$\frac{1}{5}\sqrt{\frac{a}{T_i}} \leqslant \frac{1}{5}\sqrt{\frac{a}{T_i(t)}} \leqslant \frac{\Delta_i}{4} < \frac{\Delta_i}{2} \tag{22}$$

Step 4: Putting it together
Combining (22), event Ξ, and (14) yields for all $i \in [K]$,

$$\left\|\widehat{\boldsymbol{\mu}}_{i,T_i(T+1)} - \boldsymbol{\mu}_i\right\|_2 < \frac{\Delta_i}{2}.$$

Then, by Lemma 3, it follows that $\mathcal{L}_{T,\epsilon}(\widehat{S}) = 0$. □

E MD-SAR Upper Bound Proof

As in Algorithm 2, define $\overline{\log}(x) = \frac{1}{2} + \sum_{i=2}^{x} \frac{1}{i}$.

Proof of Theorem 3. **Step 1: Defining an event and bounding probability** Let $\mathcal{N}$ be a minimal $\frac{1}{2}$-net on $\mathcal{S}^{D-1}$. Let $\delta > 0$ (we choose it later). Define the events

$$\Xi_1 = \{\forall i \in [K], \forall \boldsymbol{y} \in \mathcal{N}, \forall r \in \{1,\ldots,T\} : |\boldsymbol{y}^t(\widehat{\boldsymbol{\mu}}_{i,r} - \boldsymbol{\mu}_i)| \leqslant \sqrt{\frac{(T-K)\delta^2}{\overline{\log}(K)H_2 r}}\},$$

$$\Xi_2 = \{\forall k \in [K-1], \forall l \in \{(K),\ldots,(K+1-k)\} \text{ and } j \in [K] \text{ s.t. } 2\Delta_j < \Delta_l : \widehat{\Delta}_{l,n_k} - \widehat{\Delta}_{j,n_k} > 0\}.$$

Then, by Lemma 2,

$$\begin{aligned}\Pr(\Xi_1) &\geqslant 1 - 2(\log(T)+1)K5^D \exp(-\frac{T-K}{16\overline{\log}(K)H_2}\frac{\delta^2}{R^2}) \\ &\geqslant 1 - 2(\log(T)+1)K5^D \exp(-\frac{T-K}{16\log(2K)H_2}\frac{\delta^2}{R^2}) \end{aligned} \tag{23}$$

where line (23) follows by $\log(K+1) - \frac{1}{2} \leqslant \overline{\log}(K) \leqslant \log(K) + \frac{1}{2} \leqslant \log(2K)$ [Audibert and Bubeck, 2010].

Next, we bound $\Pr(\Xi_2)$. By a union bound,

$$\Pr(\Xi_2^c) \leqslant \sum_{k\in[K-1]} \sum_{l\in\{(K),\ldots,(K+1-k)\},j:2\Delta_j<\Delta_l} \Pr(\widehat{\Delta}_{j,n_k} - \widehat{\Delta}_{l,n_k} \geqslant 0).$$

Fix a round $k \in \{1,\ldots,K-1\}$, let $l \in \{(K),\ldots,(K+1-k)\}$, and let $j \in [K]$ such that $2\Delta_j < \Delta_l$. Then,

$$\begin{aligned}\Pr(\widehat{\Delta}_{j,n_k} - \widehat{\Delta}_{l,n_k} \geqslant 0) &= \Pr((\widehat{\Delta}_{j,n_k} - \Delta_j) - (\widehat{\Delta}_{l,n_k} - \Delta_l) \geqslant \Delta_l - \Delta_j) \\ &\leqslant \Pr((\widehat{\Delta}_{j,n_k} - \Delta_j) - (\widehat{\Delta}_{l,n_k} - \Delta_l) > \frac{1}{2}\Delta_l) \\ &\leqslant \Pr(|(\widehat{\Delta}_{j,n_k} - \Delta_j) - (\widehat{\Delta}_{l,n_k} - \Delta_l)| > \frac{1}{2}\Delta_l) \\ &\leqslant \Pr(|\widehat{\Delta}_{j,n_k} - \Delta_j| + |\widehat{\Delta}_{l,n_k} - \Delta_l| > \frac{1}{2}\Delta_l) \\ &\leqslant \Pr(|\widehat{\Delta}_{j,n_k} - \Delta_j| > \frac{1}{4}\Delta_l) + \Pr(|\widehat{\Delta}_{l,n_k} - \Delta_l| > \frac{1}{4}\Delta_l)\end{aligned}$$

Define the event

$$\Sigma_i = \{\forall \boldsymbol{y} \in \mathcal{N}, : |\boldsymbol{y}^t(\widehat{\boldsymbol{\mu}}_{i,n_k} - \boldsymbol{\mu}_i)| \leqslant \frac{1}{16}\Delta_l\}.$$

Under Σ_i, Lemma H.3 implies that

$$\|\widehat{\boldsymbol{\mu}}_{i,n_k} - \boldsymbol{\mu}_i\|_2 \leqslant 2 \sup_{\boldsymbol{y}\in\mathcal{N}} \boldsymbol{y}^t(\widehat{\boldsymbol{\mu}}_{i,n_k} - \boldsymbol{\mu}_i) \leqslant \frac{1}{8}\Delta_l.$$

Thus, by Lemma 1, Σ_i implies that $|\widehat{\Delta}_i(n_k) - \Delta_i| \leqslant \frac{1}{4}\Delta_l$. Using the contrapositive of this implication,

$$\begin{aligned}
\Pr(|\widehat{\Delta}_{j,n_k} - \Delta_j| > \frac{1}{4}\Delta_l) + \Pr(|\widehat{\Delta}_{l,n_k} - \Delta_l| > \frac{1}{4}\Delta_l) &\leqslant \Pr(\Sigma_j^c) + \Pr(\Sigma_l^c) \\
&\leqslant \sum_{\boldsymbol{y}\in\mathcal{N}} [\Pr(|\boldsymbol{y}^t(\widehat{\boldsymbol{\mu}}_{l,n_k} - \boldsymbol{\mu}_l)| > \frac{1}{16}\Delta_l) \\
&\quad + \Pr(|\boldsymbol{y}^t(\widehat{\boldsymbol{\mu}}_{j,n_k} - \boldsymbol{\mu}_j)| > \frac{1}{16}\Delta_l)] \\
&\leqslant 4 \cdot 5^D \exp(-\frac{n_k \Delta_l^2}{512R^2}) \qquad (24) \\
&\leqslant 4 \cdot 5^D \exp(-\frac{n_k \Delta_{(K+1-k)}^2}{512R^2}).
\end{aligned}$$

where line (24) follows by Lemma H.5 and since $\mathcal{N}$ is a $\frac{1}{2}$-net by construction, we have $|\mathcal{N}| \leqslant 5^D$ by Lemma H.4. Then,

$$\begin{aligned}
\Pr(\Xi_2^c) &\leqslant \sum_{k\in[K-1]} \sum_{l\in\{(K),\dots,(K+1-k)\}, j:2\Delta_j<\Delta_l} 4 \cdot 5^D \exp(-\frac{n_k \Delta_{(K+1-k)}^2}{512R^2}) \\
&\leqslant \sum_{k\in[K-1]} kK^2 4 \cdot 5^D \exp(-\frac{n_k \Delta_{(K+1-k)}^2}{512R^2}) \\
&\leqslant 4K^3 5^D \exp(-\frac{T-K}{512R^2 H_2})
\end{aligned}$$

where we used the fact that

$$n_k \Delta_{(K+1-k)}^2 \geqslant \frac{T-K}{\overline{\log}(K)(K+1-k)\Delta_{(K+1-k)}^{-2}} \geqslant \frac{T-K}{\overline{\log}(K) H_2}.$$

For the remainder of the proof, we suppose $\Xi_1 \cap \Xi_2$ holds.

Step 2: Lower bounding the number of pulls This step of the proof is similar to the proof of Theorem 3 in [Audibert and Bubeck, 2010]; we repeat it for the sake of completeness. Consider phase k. At least one of the arms $l \in \{(K), \dots, (K+1-k)\}$ has not been eliminated. Then, by Ξ_2, we have that $\widehat{\Delta}_{l,n_k} > \widehat{\Delta}_{j,n_k}$ for any arm j satisfying $2\Delta_j < \Delta_{(K+1-k)}$. Thus, at the end of phase k, MD-SAR does not eliminate any arm j such that $2\Delta_j < \Delta_{(K+1-k)}$.

Now, fix an arm j. Recall that T_j denotes the number of pulls of arm j after T rounds. We consider two distinct cases: *(i)* there exists $m \in [K]$ such that $\Delta_{(m-1)} \leqslant 2\Delta_j < \Delta_{(m)}$ and *(ii)* there exists no such

$m \in [K]$. Suppose *(i)* holds. Since $2\Delta_j < \Delta_{(m)}$, the arm j is eliminated some time after the $K+2-m$ phase so that

$$\Delta_j^2 T_j \geqslant \Delta_j^2 n_{K+2-m} = \frac{\Delta_j^2}{\Delta_{(m-1)}^2} \frac{T-K}{\overline{\log}(K)(m-1)\Delta_{(m-1)}^{-2}} \geqslant \frac{\Delta_j^2}{\Delta_{(m-1)}^2} \frac{T-K}{\overline{\log}(K)H_2} \geqslant \frac{T-K}{4\overline{\log}(K)H_2}.$$

Next, suppose *(ii)* holds. Then, $2\Delta_j \geqslant \Delta_{(K)}$, so that

$$\Delta_j^2 T_j \geqslant \frac{1}{4}\Delta_{(K)}^2 n_1 = \frac{T-K}{4\overline{\log}(K)K\Delta_{(K)}^{-2}} \geqslant \frac{T-K}{4\overline{\log}(K)H_2}.$$

Thus, we have that for all $j \in [K]$,

$$T_j \geqslant \frac{T-K}{4\overline{\log}(K)H_2\Delta_j^2}. \tag{25}$$

Step 3: Putting it together. Using Lemma 2, Ξ_1, and (25), we have that for all $i \in [K]$,

$$\left\|\widehat{\boldsymbol{\mu}}_{i,T_i(T+1)} - \boldsymbol{\mu}_i\right\|_2 \leqslant 2\sqrt{\frac{(T-K)\delta^2}{\overline{\log}(K)H_2T_i}} \leqslant 4\delta\Delta_i.$$

We choose $\delta = \frac{1}{9}$. Then, by Lemma 3, it follows that $\mathcal{L}_{T,\epsilon}(\widehat{S}) = 0$.

□

F MD-APT Upper Bound Proof

Proof of Proposition 1. Let $\mathcal{N}$ be a minimal $\frac{1}{2}$-net of $\mathcal{S}^{D-1}$. By Lemma H.4, $|\mathcal{N}| \leqslant 5^D$. Then,

$$\begin{aligned}
\mathbb{E}[\mathcal{L}_{T,\epsilon}(\widehat{S})] &\leqslant \Pr(\exists i : \widehat{\boldsymbol{\mu}}_{i,T_i(T+1)} \in P \text{ and } \boldsymbol{\mu}_i \notin P \\
&\qquad \text{or } \widehat{\boldsymbol{\mu}}_{i,T_i(T+1)} \notin P \text{ and } \boldsymbol{\mu}_i \in P) \\
&\leqslant \Pr(\exists i : \left\|\widehat{\boldsymbol{\mu}}_{i,T_i(T+1)} - \boldsymbol{\mu}_i\right\|_2 > \Delta_i) \\
&\leqslant \sum_{i\in[K]} \Pr(\left\|\widehat{\boldsymbol{\mu}}_{i,T_i(T+1)} - \boldsymbol{\mu}_i\right\|_2 > \Delta_i) \\
&\leqslant \sum_{i\in[K]}\sum_{\boldsymbol{y}\in\mathcal{N}} \Pr(|\boldsymbol{y}^t(\widehat{\boldsymbol{\mu}}_{i,T_i(T+1)} - \boldsymbol{\mu}_i)| > \frac{\Delta_i}{2}) && (26)\\
&\leqslant \sum_{i\in[K]}\sum_{y\in\mathcal{N}} 2\exp(-\frac{1}{8}\frac{\Delta_i^2 T_i}{R^2}) && (27)\\
&= 2K5^D\exp(-\frac{1}{8}\frac{T}{HR^2}).
\end{aligned}$$

Line (26) follows since by Lemma H.3, if $\left\|\widehat{\boldsymbol{\mu}}_{i,T_i(T+1)} - \boldsymbol{\mu}_i\right\|_2 > \Delta_i$, then there exists $\boldsymbol{y} \in \mathcal{N}$ such that $|\boldsymbol{y}^t(\widehat{\boldsymbol{\mu}}_{i,T_i(T+1)} - \boldsymbol{\mu}_i)| > \frac{\Delta_i}{2}$. Line (27) follows by Lemma H.5. □

Proof of Theorem 4. **Step 1: Defining an appropriate event.**

Let $\mathcal{N}$ be a minimal $\frac{1}{2}$-net on $\mathcal{S}^{D-1}$. Let $\delta > 0$ (we choose it later). Define the event

$$\Xi = \{\forall i \in [K], \forall \boldsymbol{y} \in \mathcal{N}, \forall r \in \{1, \dots, T\} : |\boldsymbol{y}^t(\widehat{\boldsymbol{\mu}}_{i.r} - \boldsymbol{\mu}_i)| \leqslant \sqrt{\frac{T\delta^2}{4Hr}}\}.$$

By Lemma 2, on Ξ, for all $i \in [K]$ and for all $r \in [T]$

$$\|\widehat{\boldsymbol{\mu}}_{i.r} - \boldsymbol{\mu}_i\|_2 \leqslant \sqrt{\frac{T\delta^2}{Hr}}. \tag{28}$$

and

$$\mathbb{P}(\Xi) \geqslant 1 - 2(\log(T) + 1)K5^D \exp(-T\frac{\delta^2}{16R^2H})$$

For the remainder of the proof, we suppose that Ξ holds.

Step 2: Lower bound the number of pulls for some arm.

Fix T. Recall that T_i denotes the number of pulls of arm i after T rounds. We claim that there exists an arm k that has been pulled after initialization and such that $T_k - 1 \geqslant \frac{T-K}{H\Delta_k^2}$ (for the remainder of the proof, let k denote one of these arms). If not, then we obtain the following contradiction:

$$T - K = \sum_{i=1}^{K}(T_i - 1) < \sum_{i=1}^{K}\frac{T-K}{H\Delta_i^2} = T - K.$$

Since $T \geqslant 2K$, $T_k - 1 \geqslant \frac{T}{2H\Delta_k^2}$.

For the remainder of the proof, let $t \leqslant T$ denote the last round that arm k was pulled. Then,

$$T_k(t) = T_k - 1 \geqslant \frac{T}{2H\Delta_k^2}. \tag{29}$$

Step 3: Lower bound the number of pulls for each arm.

Lemma 1 and event Ξ imply that

$$|\widehat{\Delta}_{i,T_i(t)} - \Delta_i| \leqslant 2\sqrt{\frac{T\delta^2}{HT_i(t)}} \tag{30}$$

for all $i \in [K]$.

At time t, we pulled arm k, so that for all $i \in [K]$, $\widehat{\Delta}_{k,T_k(t)}\sqrt{T_k(t)} \leqslant \widehat{\Delta}_{i,T_i(t)}\sqrt{T_i(t)}$. Then, using (29) and (30),

$$\begin{aligned}\widehat{\Delta}_{k,T_k(t)}\sqrt{T_k(t)} &\geqslant (\Delta_k - 2\sqrt{\frac{T\delta^2}{HT_k(t)}})\sqrt{T_k(t)} \\ &\geqslant (\Delta_k - 2\sqrt{2\delta^2\Delta_k^2})\sqrt{T_k(t)}.\end{aligned} \tag{31}$$

We require that $\delta < \frac{1}{2\sqrt{2}}$ so that (31) is positive. Thus, we can apply (31) and (29) to obtain that

$$\widehat{\Delta}_{k,T_k(t)}\sqrt{T_k(t)} \geqslant (\frac{1}{\sqrt{2}} - 2\delta)\sqrt{\frac{T}{H}}. \tag{32}$$

Next, applying (30), we obtain

$$\begin{aligned}\widehat{\Delta}_{i,T_i(t)}\sqrt{T_i(t)} &= \widehat{\Delta}_{i,T_i(t)}\sqrt{T_i(t)} \\ &\leqslant (\Delta_i + 2\sqrt{\frac{T\delta^2}{HT_i(t)}})\sqrt{T_i(t)} \\ &\leqslant \Delta_i\sqrt{T_i(t)} + 2\delta\sqrt{\frac{T}{H}}. \end{aligned} \tag{33}$$

Combining inequalities $T_i(t) \leqslant T_i$, (33), $\widehat{\Delta}_{k,T_k(t)}\sqrt{T_k(t)} \leqslant \widehat{\Delta}_{i,T_i(t)}\sqrt{T_i(t)}$, and (32) yields that

$$\begin{aligned}\Delta_i\sqrt{T_i} + 2\delta\sqrt{\frac{T}{H}} &\geqslant \Delta_i\sqrt{T_i(t)} + 2\delta\sqrt{\frac{T}{H}} \\ &\geqslant \widehat{\Delta}_{i,T_i(t)}\sqrt{T_i(t)} \\ &\geqslant \widehat{\Delta}_{k,T_k(t)}\sqrt{T_k(t)} \\ &\geqslant (\frac{1}{\sqrt{2}} - 2\delta)\sqrt{\frac{T}{H}}. \end{aligned} \tag{34}$$

Rearranging (34) yields for all $i \in [K]$

$$(1 - 4\sqrt{2}\delta)^2\frac{T}{2H\Delta_i^2} \leqslant T_i. \tag{35}$$

Step 4: Putting it together.
Combining (35) with (28) and Ξ respectively, we obtain

$$\left\|\widehat{\boldsymbol{\mu}}_{i,T_i(T+1)} - \boldsymbol{\mu}_i\right\|_2 \leqslant \Delta_i\delta(1 - 4\sqrt{2}\delta).$$

We choose $\delta = \frac{\sqrt{2}}{9}$. Thus, by Lemma 3, $\mathcal{L}_{T,\epsilon}(\widehat{S}) = 0$. □

G Key Lemmas

In this section, we prove the Lemmas of Section 6, namely, Lemmas 1, 2, and 3.

Proof of Lemma 1. For the sake of brevity, we write $\widehat{\boldsymbol{\mu}}_i$ instead of $\widehat{\boldsymbol{\mu}}_{i,t}$ and $\widehat{\Delta}_i$ instead of $\widehat{\Delta}_{i,t}$. We separate the analysis into 4 cases.

Case 1: $A\boldsymbol{\mu}_i \leqslant \boldsymbol{b}$ and $A\widehat{\boldsymbol{\mu}}_i \leqslant \boldsymbol{b}$.

Let j be such that $\Delta_i = |b_j - \boldsymbol{a}_j^t\boldsymbol{\mu}_i| + \epsilon$ and let $\widehat{j}$ such that $\widehat{\Delta}_i = |b_{\widehat{j}} - \boldsymbol{a}_{\widehat{j}}^t\widehat{\boldsymbol{\mu}}_i| + \epsilon$. Then, by definition of j and $\widehat{j}$,

$$0 \leqslant b_j - \boldsymbol{a}_j^t\boldsymbol{\mu}_i \leqslant b_{\widehat{j}} - \boldsymbol{a}_{\widehat{j}}^t\boldsymbol{\mu}_i \tag{36}$$

$$0 \leqslant b_{\widehat{j}} - \boldsymbol{a}_{\widehat{j}}^t\widehat{\boldsymbol{\mu}}_i \leqslant b_j - \boldsymbol{a}_j^t\widehat{\boldsymbol{\mu}}_i. \tag{37}$$

Note that it suffices to bound

$$|\Delta_i - \widehat{\Delta}_i| = ||b_j - \boldsymbol{a}_j^t\boldsymbol{\mu}_i| - |b_{\widehat{j}} - \boldsymbol{a}_{\widehat{j}}^t\widehat{\boldsymbol{\mu}}_i||.$$

Then,

$$\begin{aligned}
(b_j - \boldsymbol{a}_j^t \boldsymbol{\mu}_i) - (b_{\hat{j}} - \boldsymbol{a}_{\hat{j}}^t \widehat{\boldsymbol{\mu}}_i) &\leqslant (b_{\hat{j}} - \boldsymbol{a}_{\hat{j}}^t \boldsymbol{\mu}_i) - (b_{\hat{j}} - \boldsymbol{a}_{\hat{j}}^t \widehat{\boldsymbol{\mu}}_i) && (38)\\
&= \boldsymbol{a}_{\hat{j}}^t (\widehat{\boldsymbol{\mu}}_i - \boldsymbol{\mu}_i) && (39)\\
&\leqslant \gamma. && (40)
\end{aligned}$$

where line (38) used line (36) and line (40) follows by Cauchy-Schwarz inequality, $\|\boldsymbol{a}_j\|_2 = 1$, and the hypothesis. Next,

$$\begin{aligned}
(b_{\hat{j}} - \boldsymbol{a}_{\hat{j}}^t \widehat{\boldsymbol{\mu}}_i) - (b_j - \boldsymbol{a}_j^t \boldsymbol{\mu}_i) &\leqslant (b_j - \boldsymbol{a}_j^t \widehat{\boldsymbol{\mu}}_i) - (b_j - \boldsymbol{a}_j^t \boldsymbol{\mu}_i) && (41)\\
&= \boldsymbol{a}_j^t (\boldsymbol{\mu}_i - \widehat{\boldsymbol{\mu}}_i)\\
&\leqslant \gamma. && (42)
\end{aligned}$$

where line (41) used line (37) and line (42) follows by Cauchy-Schwarz inequality, $\|\boldsymbol{a}_j\|_2 = 1$, and the hypothesis.

Case 2: $A\boldsymbol{\mu}_i \nleqslant \boldsymbol{b}$ and $A\widehat{\boldsymbol{\mu}}_i \nleqslant \boldsymbol{b}$. Then,

$$\begin{aligned}
|\Delta_i - \widehat{\Delta}_i| &= |\, \|\boldsymbol{\mu}_i - \mathrm{Proj}_P(\boldsymbol{\mu}_i)\|_2 - \|\widehat{\boldsymbol{\mu}}_i - \mathrm{Proj}_P(\widehat{\boldsymbol{\mu}}_i)\|_2 \,|\\
&\leqslant \|(\boldsymbol{\mu}_i - \widehat{\boldsymbol{\mu}}_i) - (\mathrm{Proj}_P(\boldsymbol{\mu}_i) - \mathrm{Proj}_P(\widehat{\boldsymbol{\mu}}_i))\|_2 && (43)\\
&\leqslant \|\boldsymbol{\mu}_i - \widehat{\boldsymbol{\mu}}_i\|_2 + \|\mathrm{Proj}_P(\boldsymbol{\mu}_i) - \mathrm{Proj}_P(\widehat{\boldsymbol{\mu}}_i)\|_2 && (44)\\
&\leqslant 2\,\|\boldsymbol{\mu}_i - \widehat{\boldsymbol{\mu}}_i\|_2 && (45)\\
&\leqslant 2\gamma
\end{aligned}$$

where line (43) used the reverse triangle inequality, (44) used the triangle inequality, and (45) used the fact that projection onto a convex set is contractive (Proposition 2.2.1 [Bertsekas, 2009]).

Case 3: $A\boldsymbol{\mu}_i \nleqslant \boldsymbol{b}$ and $A\widehat{\boldsymbol{\mu}}_i \leqslant \boldsymbol{b}$.

We claim that $\mathrm{dist}(\widehat{\boldsymbol{\mu}}_i, \partial P) \leqslant \|\widehat{\boldsymbol{\mu}}_i - \boldsymbol{\mu}_i\|_2$. Suppose not. Then, since $\widehat{\boldsymbol{\mu}}_i \in P$ and $\boldsymbol{\mu}_i \notin P$, there exists $\theta \in [0,1]$ such that $\boldsymbol{z} = \theta\boldsymbol{\mu}_i + (1-\theta)\widehat{\boldsymbol{\mu}}_i \in \partial P$. Then,

$$\|\widehat{\boldsymbol{\mu}}_i - \boldsymbol{z}\|_2 \leqslant \|\widehat{\boldsymbol{\mu}}_i - \boldsymbol{\mu}_i\|_2 < \mathrm{dist}(\widehat{\boldsymbol{\mu}}_i, \partial P),$$

which is a contradiction. Thus, the claim follows. Then,

$$\mathrm{dist}(\widehat{\boldsymbol{\mu}}_i, \partial P) \leqslant \|\widehat{\boldsymbol{\mu}}_i - \boldsymbol{\mu}_i\|_2 \leqslant \gamma.$$

Next, since $\widehat{\boldsymbol{\mu}}_i \in P$,

$$\mathrm{dist}(\boldsymbol{\mu}_i, P) \leqslant \|\boldsymbol{\mu}_i - \widehat{\boldsymbol{\mu}}_i\|_2 \leqslant \gamma.$$

Thus,

$$|\Delta_i - \widehat{\Delta}_i| = |\,\mathrm{dist}(\widehat{\boldsymbol{\mu}}_i, \partial P) - \mathrm{dist}(\boldsymbol{\mu}_i, P)| \leqslant \max(\mathrm{dist}(\widehat{\boldsymbol{\mu}}_i, \partial P), \mathrm{dist}(\boldsymbol{\mu}_i, P)) \leqslant \gamma.$$

Case 4: $A\boldsymbol{\mu}_i \leqslant \boldsymbol{b}$ and $A\widehat{\boldsymbol{\mu}}_i \nleqslant \boldsymbol{b}$. This case is similar to case 3. Since $\boldsymbol{\mu}_i \in P$ and $\widehat{\boldsymbol{\mu}}_i \notin P$,

$$\mathrm{dist}(\widehat{\boldsymbol{\mu}}_i, P) \leqslant \|\widehat{\boldsymbol{\mu}}_i - \boldsymbol{\mu}_i\|_2 \leqslant \gamma.$$

Next, since $\boldsymbol{\mu}_i \in P$ and $\widehat{\boldsymbol{\mu}}_i \notin P$, by a similar argument used in case 3,

$$\begin{aligned}\operatorname{dist}(\boldsymbol{\mu}_i, \partial P) &\leqslant \|\boldsymbol{\mu}_i - \widehat{\boldsymbol{\mu}}_i\|_2 \\ &\leqslant \gamma.\end{aligned}$$

Thus,

$$|\Delta_i - \widehat{\Delta}_i| = |\operatorname{dist}(\widehat{\boldsymbol{\mu}}_i, \partial P) - \operatorname{dist}(\boldsymbol{\mu}_i, P)| \leqslant \max(\operatorname{dist}(\widehat{\boldsymbol{\mu}}_i, \partial P), \operatorname{dist}(\boldsymbol{\mu}_i, P)) \leqslant \gamma.$$

□

Proof of Lemma 2. First, we bound the norm of $\widehat{\boldsymbol{\mu}}_{i,r} - \boldsymbol{\mu}_i$ on Ξ for all $i \in [K]$ and for all $r \in [T]$. Fix $i \in [K]$ and $r \in [T]$. Recall that $\mathcal{N}$ is a minimal $\frac{1}{2}$-net. Using the event Ξ and Lemma H.3,

$$\|\widehat{\boldsymbol{\mu}}_{i,r} - \boldsymbol{\mu}_i\|_2 \leqslant 2 \sup_{\boldsymbol{y} \in \mathcal{N}} \boldsymbol{y}^t(\widehat{\boldsymbol{\mu}}_{i,r} - \boldsymbol{\mu}_i) \leqslant \sqrt{\frac{\omega^2}{r}}.$$

Next, we give the probability bound. Since ν_i is R-sub-Gaussian, by definition, we have that if $\boldsymbol{X} \sim \nu_i$, then

$$\sup_{\boldsymbol{y} \in \mathcal{N}} \left\|\boldsymbol{X}^t \boldsymbol{y}\right\|_{\psi_2} \leqslant \sup_{\boldsymbol{y} \in \mathcal{S}^{D-1}} \left\|\boldsymbol{X}^t \boldsymbol{y}\right\|_{\psi_2} = \|\nu_i\|_{\psi_2} \leqslant R.$$

Thus, by Lemma H.2 and a union bound, for each $i \in [K]$, $\boldsymbol{y} \in \mathcal{N}$, and $u \in \{0, \ldots, \lfloor \log(T) \rfloor\}$:

$$\Pr(\exists v \in [2^u, 2^{u+1}], |\boldsymbol{y}^t(\widehat{\boldsymbol{\mu}}_{i,v} - \boldsymbol{\mu}_i)| \geqslant \sqrt{\frac{\omega^2}{4v}}) \leqslant 2\exp(-\frac{\omega^2}{16R^2}).$$

Taking a union bound over all $i \in [K]$, $\boldsymbol{y} \in \mathcal{N}$, and $u \in \{0, \ldots, \lfloor \log(T) \rfloor\}$ yields

$$\begin{aligned}\Pr(\Xi) &\geqslant 1 - 2(\log(T)+1)K|\mathcal{N}|\exp(-\frac{\omega^2}{16R^2}) \\ &\geqslant 1 - 2(\log(T)+1)K5^D\exp(-\frac{\omega^2}{16R^2})\end{aligned}$$

where in the last line we used $|\mathcal{N}| \leqslant 5^D$ by Lemma H.4. □

Proof of Lemma 3. Fix $i \in [K]$. For the sake of brevity, we write $\widehat{\boldsymbol{\mu}}_i$ instead of $\widehat{\boldsymbol{\mu}}_{i,t}$. First, suppose $A\boldsymbol{\mu}_i \leqslant \boldsymbol{b} - \epsilon\boldsymbol{1}$. Fix any $j \in [M]$. Then,

$$\begin{aligned}\boldsymbol{a}_j^t\widehat{\boldsymbol{\mu}}_i - b_j &= \boldsymbol{a}_j^t(\widehat{\boldsymbol{\mu}}_i - \boldsymbol{\mu}_i) + \boldsymbol{a}_j^t\boldsymbol{\mu}_i - b_j \\ &< \frac{\Delta_i}{2} + \boldsymbol{a}_j^t\boldsymbol{\mu}_i - b_j \qquad (46) \\ &\leqslant \frac{1}{2}(b_j - \boldsymbol{a}_j^t\boldsymbol{\mu}_i + \epsilon) + \boldsymbol{a}_j^t\boldsymbol{\mu}_i - b_j \\ &\leqslant \frac{1}{2}(\boldsymbol{a}_j^t\boldsymbol{\mu}_i - b_j + \epsilon) \\ &\leqslant \frac{1}{2}(-\epsilon + \epsilon) \\ &= 0\end{aligned}$$

where line (46) follows by the Cauchy-Schwarz inequality, $\|\boldsymbol{a}_j\|_2 = 1$, and the hypothesis.

Next, suppose $\text{dist}(\boldsymbol{\mu}_i, P) \geqslant \epsilon$. Then,

$$\|\widehat{\boldsymbol{\mu}}_i - \boldsymbol{\mu}_i\|_2 < \frac{\Delta_i}{2} = \frac{1}{2}(\text{dist}(\boldsymbol{\mu}_i, P) + \epsilon) \leqslant \text{dist}(\boldsymbol{\mu}_i, P).$$

Thus, $\widehat{\boldsymbol{\mu}}_i \notin P$ since otherwise we have a contradiction.

□

H Technical Lemmas

Lemma H.1. *Let* $P = \{\boldsymbol{x} \in \mathbb{R}^D : A\boldsymbol{x} \leqslant \boldsymbol{b}\}$ *with* $A \in \mathbb{R}^{M \times D}$. *Let* $\boldsymbol{\mu} \in P$. *Then,*

$$\text{dist}(\boldsymbol{\mu}, \partial P) = \min_{i=1,\ldots,M} \text{dist}(\boldsymbol{\mu}, \{\boldsymbol{x} : \boldsymbol{a}_i^t \boldsymbol{x} = b_i\}).$$

Proof. It is not hard to establish that $\partial P = P \cap (\cup_{i=1}^M \{\boldsymbol{x} : \boldsymbol{a}_i^t \boldsymbol{x} = b_i\})$. We claim that

$$\text{dist}(\boldsymbol{\mu}, \cup_{i=1}^M \{\boldsymbol{x} : \boldsymbol{a}_i^t \boldsymbol{x} = b_i\}) = \text{dist}(\boldsymbol{\mu}, P \cap (\cup_{i=1}^M \{\boldsymbol{x} : \boldsymbol{a}_i^t \boldsymbol{x} = b_i\})).$$

Since $\cup_{i=1}^M \{\boldsymbol{x} : \boldsymbol{a}_i^t bx = b_i\}$ is closed, there exists $\boldsymbol{y} \in \cup_{i=1}^M \{\boldsymbol{x} : \boldsymbol{a}_i^t bx = b_i\}$ such that

$$\|\boldsymbol{\mu} - \boldsymbol{y}\|_2 = \text{dist}(\boldsymbol{\mu}, \cup_{i=1}^M \{\boldsymbol{x} : \boldsymbol{a}_i^t bx = b_i\}).$$

We claim that $\boldsymbol{y} \in P$. Suppose not (towards a contradiction). Then, there exists $\theta \in (0, 1)$ such that $\boldsymbol{z} = (1 - \theta)\boldsymbol{\mu} + \theta \boldsymbol{y} \in \partial P$. Then,

$$\text{dist}(\boldsymbol{\mu}, (\cup_{i=1}^M \{\boldsymbol{x} : \boldsymbol{a}_i^t \boldsymbol{x} = b_i\})) \leqslant \|\boldsymbol{z} - \boldsymbol{\mu}\|_2 < \|\boldsymbol{y} - \boldsymbol{\mu}\|_2 = \text{dist}(\boldsymbol{\mu}, \cup_{i=1}^M \{\boldsymbol{x} : \boldsymbol{a}_i^t bx = b_i\}),$$

which is a contradiction, establishing the claim. Then,

$$\begin{aligned}\min_{i=1,\ldots,M} \text{dist}(\boldsymbol{\mu}, \{\boldsymbol{x} : \boldsymbol{a}_i^t \boldsymbol{x} = b_i\}) &= \text{dist}(\boldsymbol{\mu}, \cup_{i=1}^M \{\boldsymbol{x} : \boldsymbol{a}_i^t \boldsymbol{x} = b_i\}) \\ &= \text{dist}(\boldsymbol{\mu}, P \cap (\cup_{i=1}^M \{\boldsymbol{x} : \boldsymbol{a}_i^t \boldsymbol{x} = b_i\})) \\ &= \text{dist}(\boldsymbol{\mu}, \partial P).\end{aligned}$$

□

Lemma H.2. *Suppose that* $X_1, \ldots, X_T$ *are centered scalar* R*-sub-Gaussian random variables. Then,* $\forall u \in \{0, \ldots, \lceil \log(T) \rceil\}$,

$$\Pr(\exists v \in [2^u, 2^{u+1}] : \frac{1}{v} \sum_{i=1}^{v} X_i \geqslant \sqrt{\frac{x}{v}}) \leqslant \exp(-\frac{x}{4R^2}).$$

Proof. Define $S_v = \sum_{i=1}^v X_i$. Fix $u \in \{0, \ldots, \lceil \log(T) \rceil\}$. Let $m = 2^{u+1}$. Hoeffding's maximal inequality yields (see Step 2 of Lemma 1 of [Jamieson et al., 2014])

$$\begin{aligned}\Pr(\exists v \in [m] : \frac{1}{v} S_v \geqslant \frac{\sqrt{x}}{v}) &= \Pr(\exists v \in [m] : S_v \geqslant \sqrt{x}) \\ &\leqslant \exp(-\frac{x}{2R^2 m}).\end{aligned}$$

Then,

$$\Pr(\exists v \in [2^u, 2^{u+1}] : \frac{1}{v} S_v \geqslant \frac{\sqrt{x}}{v}) \leqslant \Pr(\exists v \in [m] : \frac{1}{v} S_v \geqslant \frac{\sqrt{x}}{v})$$
$$\leqslant \exp(-\frac{x}{2R^2 m}).$$

Finally,

$$\begin{aligned}\Pr(\exists v \in [2^u, 2^{u+1}] : \frac{1}{v} S_v \geqslant \sqrt{\frac{x}{v}}) &\leqslant \Pr(\exists v \in [2^u, 2^{u+1}] : \frac{1}{v} S_v \geqslant \frac{\sqrt{x 2^u}}{v}) \\ &\leqslant \exp(-\frac{x 2^u}{2R^2 m}) \\ &= \exp(-\frac{x}{4R^2}).\end{aligned}$$

□

Lemma H.3. *Let $\epsilon > 0$ and $\mathcal{N}_\epsilon$ be an ϵ-net of $\mathcal{S}^{D-1}$. For any $\boldsymbol{y} \in \mathbb{R}^D$,*

$$\|\boldsymbol{y}\|_2 \leqslant \frac{1}{1-\epsilon} \sup_{\boldsymbol{z} \in \mathcal{N}_\epsilon} \boldsymbol{y}^t \boldsymbol{z}.$$

Proof. Let $\boldsymbol{z}_0 \in \mathcal{N}_\epsilon$ such that $\left\| \frac{\boldsymbol{y}}{\|\boldsymbol{y}\|_2} - \boldsymbol{z}_0 \right\|_2 \leqslant \epsilon$. Then, by Cauchy-Schwarz,

$$\|\boldsymbol{y}\|_2 = \frac{\boldsymbol{y}^t \boldsymbol{y}}{\|\boldsymbol{y}\|_2} = \boldsymbol{y}^t (\frac{\boldsymbol{y}}{\|\boldsymbol{y}\|_2} - \boldsymbol{z}_0) + \boldsymbol{y}^t \boldsymbol{z}_0 \leqslant \|\boldsymbol{y}\|_2 \left\| \frac{\boldsymbol{y}}{\|\boldsymbol{y}\|_2} - \boldsymbol{z}_0 \right\|_2 + \boldsymbol{y}^t \boldsymbol{z}_0 \leqslant \epsilon \|\boldsymbol{y}\|_2 + \boldsymbol{y}^t \boldsymbol{z}_0.$$

Rearranging the inequality, we obtain

$$\|\boldsymbol{y}\|_2 \leqslant \frac{1}{1-\epsilon} \boldsymbol{y}^t \boldsymbol{z}_0 \leqslant \frac{1}{1-\epsilon} \sup_{\boldsymbol{z} \in \mathcal{N}_\epsilon} \boldsymbol{y}^t \boldsymbol{z}.$$

□

The following Lemma appears in [Vershynin et al., 2017] (see Corollary 4.2.13).

Lemma H.4. *Let $\epsilon > 0$ and $\mathcal{N}_\epsilon$ be a minimal ϵ-net of $\mathcal{S}^{D-1}$. Then, $|\mathcal{N}_\epsilon| \leqslant (\frac{2}{\epsilon} + 1)^D$.*

We state without proof general Hoeffding's inequality for sub-Gaussian random variables (see Theorem 2.6.2 in Vershynin et al. [2017]).

Lemma H.5. *Suppose that $X_1, \ldots, X_n$ are i.i.d. scalar R-sub-Gaussian random variables with mean $\boldsymbol{\mu} \in \mathbb{R}$. Then, for all $t > 0$,*

$$\Pr(|\frac{1}{n} \sum_{i=1}^n X_i - \mu| > t) \leqslant 2 \exp(-\frac{t^2 n}{2R^2}).$$

I Feasible Arm Identification with a Convex Region: Statistical Results

To begin, we introduce some notation. Let $\delta > 0$ and $\boldsymbol{x} \in \mathbb{R}^D$. Define $B_\delta(\boldsymbol{x}) = \{\boldsymbol{x} \in \mathbb{R}^D : \|\boldsymbol{x}\|_2 \leqslant \delta\}$.

Proposition I.1. *Let P be a compact convex set with positive volume. There exists a sequence of polyhedra $\{P_n\}$ such that*

$$\lim_{n \longrightarrow \infty} \sup_{\boldsymbol{y} \in \mathbb{R}^D} |\operatorname{dist}(\boldsymbol{y}, \partial P_n) - \operatorname{dist}(\boldsymbol{y}, \partial P)| = 0.$$

Further, if there exists some $\tau > 0$ such that $\operatorname{dist}(\boldsymbol{\mu}_i, \partial P) \geqslant \tau \ \forall i \in [K]$, then for all $\delta > 0$, there exists a large enough n such that

$$(1-\delta)H_P \leqslant H_{P_n} \leqslant (1+\delta)H_P.$$

Proof. **Step 1: Defining a sequence of approximations.** Define a *dyadic cube of side length* 2^{-n} as a set of the form

$$[\frac{i_1}{2^{-n}}, \frac{i_1+1}{2^{-n}}] \times \ldots \times [\frac{i_D}{2^{-n}}, \frac{i_D+1}{2^{-n}}]$$

where $i_1, \ldots i_D$ are integers and $n \in \mathbb{N}$. Let E_n denote the set of dyadic cubes with sidelength 2^{-n}. Define

$$P_n = \operatorname{conv}(\bigcup_{E \in E_n, E \subset P} E).$$

Note that for any n, P_n is a polyhedron with a finite number of constraints.

Step 2: For large n, ∂P_n is a good approximation of ∂P. Next, we claim that $\forall \delta > 0$, there exists N such that $n \geqslant N$ implies that $\sup_{\boldsymbol{x} \in \partial P_n} \operatorname{dist}(\boldsymbol{x}, \partial P) \leqslant \delta$. Suppose not. Then, there exists $\delta > 0$ such that $\forall n \in \mathbb{N}$ there exists $\boldsymbol{x}_n$ such that $\operatorname{dist}(\boldsymbol{x}_n, \partial P) > \delta$. Since P is compact and $\{\boldsymbol{x}_n\}_{n \in \mathbb{N}} \subset P$, there exists a convergent subsequence $\{\boldsymbol{x}_{n_j}\}$ with limit $\boldsymbol{x} \in P$. Then, $\operatorname{dist}(\boldsymbol{x}, \partial P) \geqslant \delta$, which implies that $\boldsymbol{x} \in P^\circ$ and $B_\delta(\boldsymbol{x}) \subset P$. By definition of P_n, there exists N such that $n \geqslant N$ implies that $B_{\frac{\delta}{2}}(\boldsymbol{x}) \subset P_N$. Thus, $n \geqslant N$ implies that $\operatorname{dist}(\boldsymbol{x}, \partial P_n) \geqslant \frac{\delta}{2}$. Thus, $\boldsymbol{x}_{n_j}$ cannot converge to $\boldsymbol{x}$, which is a contradiction. So, the claim is true.

Next, we claim that $\forall \delta > 0$, there exists N such that $n \geqslant N$ implies that $\sup_{\boldsymbol{x} \in \partial P} \operatorname{dist}(\boldsymbol{x}, \partial P_n) \leqslant \delta$. Suppose not. Then, $\exists \delta > 0$ such that $\forall n \in \mathbb{N}$ there exists $\boldsymbol{x}_n \in \partial P$ such that $\operatorname{dist}(\boldsymbol{x}_n, \partial P_n) > \delta$. ∂P is bounded and closed so that ∂P is compact. Thus, $\{\boldsymbol{x}_n\}$ has a convergent subsequence $\{\boldsymbol{x}_{n_j}\}$ with limit point $\boldsymbol{x} \in \partial P$. $\boldsymbol{x}$ has the property that $\operatorname{dist}(\boldsymbol{x}, \partial P_n) \geqslant \delta$ for all $n \in \mathbb{N}$. Let $\boldsymbol{y} \in P^\circ$ (such a point exists since P has positive volume). Then, since P is convex, by the line segment principle (Proposition 1.4.1 [Bertsekas, 2009]), every point of the form $\boldsymbol{z}_\theta = (1-\theta)\boldsymbol{x} + \theta \boldsymbol{y}$ for $\theta \in (0, 1]$ is such that $\boldsymbol{z}_\theta \in P^\circ$. So there exists $\boldsymbol{w} \in B_\delta(\boldsymbol{x}) \cap P^\circ$. For large enough n, $\boldsymbol{w} \in P_n$. Since $\|\boldsymbol{w} - \boldsymbol{x}\|_2 < \delta$, we have a contradiction and thus the claim follows.

Step 3: Distance to ∂P_n approaches uniformly distance to ∂P. Formally, we show that

$$\lim_{n \longrightarrow \infty} \sup_{\boldsymbol{y} \in \mathbb{R}^D} |\operatorname{dist}(\boldsymbol{y}, \partial P_n) - \operatorname{dist}(\boldsymbol{y}, \partial P)| = 0. \tag{47}$$

Let $\delta > 0$. Let n large enough so that $\sup_{\boldsymbol{x} \in \partial P_n} \operatorname{dist}(\boldsymbol{x}, \partial P) \leqslant \delta$ and $\sup_{\boldsymbol{x} \in \partial P} \operatorname{dist}(\boldsymbol{x}, \partial P_n) \leqslant \delta$. Fix $\boldsymbol{y} \in \mathbb{R}^D$. Let $\boldsymbol{x}_p \in \partial P$ such that $\|\boldsymbol{y} - \boldsymbol{x}_p\|_2 = \operatorname{dist}(\boldsymbol{y}, \partial P)$ and $\boldsymbol{x}_{p_n} \in \partial P_n$ such that $\|\boldsymbol{y} - \boldsymbol{x}_{p_n}\|_2 = \operatorname{dist}(\boldsymbol{y}, \partial P_n)$.

Let $\boldsymbol{z} \in \partial P$ such that $\|\boldsymbol{x}_{p_n} - \boldsymbol{z}\|_2 \leqslant \delta$. Then, by the reverse triangle inequality,

$$| \, \|\boldsymbol{z} - \boldsymbol{y}\|_2 - \|\boldsymbol{x}_{p_n} - \boldsymbol{y}\|_2 \, | \leqslant \delta.$$

Then,

$$\operatorname{dist}(\boldsymbol{y}, \partial P) - \|\boldsymbol{y} - \boldsymbol{x}_{p_n}\|_2 \leqslant \operatorname{dist}(\boldsymbol{y}, \partial P) - \|\boldsymbol{z} - \boldsymbol{y}\|_2 + \delta \leqslant \delta.$$

Let $\boldsymbol{w} \in \partial P_n$ such that $\|\boldsymbol{x}_p - \boldsymbol{w}\|_2 \leqslant \delta$. By the reverse triangle inequality,

$$| \, \|\boldsymbol{y} - \boldsymbol{x}_p\|_2 - \|\boldsymbol{y} - \boldsymbol{w}\|_2 \, | \leqslant \delta.$$

Then,

$$\operatorname{dist}(\boldsymbol{y}, \partial P_n) - \|\boldsymbol{y} - \boldsymbol{x}_p\|_2 \leqslant \operatorname{dist}(\boldsymbol{y}, \partial P_n) - \|\boldsymbol{y} - \boldsymbol{w}\|_2 + \delta \leqslant \delta.$$

This establishes (47).

Step 4: Approximation of Problem Complexity. Suppose there exists some $\tau > 0$ such that $\operatorname{dist}(\boldsymbol{\mu}_i, \partial P) \geqslant \tau \ \forall i \in [K]$. Let $\gamma_1 = 1 - \frac{1}{\sqrt{1+\delta}}$, $\gamma_2 = \frac{1}{\sqrt{1-\delta}} - 1$ and $\gamma = \min(\gamma_1, \gamma_2)$. Let n large enough such that $\forall \boldsymbol{y} \in \mathbb{R}^D$,

$$|\operatorname{dist}(\boldsymbol{y}, \partial P_n) - \operatorname{dist}(\boldsymbol{y}, \partial P)| \leqslant \gamma\tau \leqslant \gamma \min_i \operatorname{dist}(\boldsymbol{\mu}_i, \partial P).$$

Then,

$$\begin{aligned} H_P &= \sum_{i \in [K]} [\operatorname{dist}(\boldsymbol{\mu}_i, \partial P) + \epsilon]^{-2} \\ &\leqslant \sum_{i \in [K]} [\operatorname{dist}(\boldsymbol{\mu}_i, \partial P_N)(1-\gamma) + \epsilon]^{-2} \\ &\leqslant \sum_{i \in [K]} [\operatorname{dist}(\boldsymbol{\mu}_i, \partial P_N) + \epsilon]^{-2}(1-\gamma)^{-2} \\ &\leqslant \sum_{i \in [K]} [\operatorname{dist}(\boldsymbol{\mu}_i, \partial P_N) + \epsilon]^{-2}(1+\delta) \\ &= (1+\delta) H_{P_n} \end{aligned}$$

Similarly, $H_P \geqslant (1-\delta) H_{P_n}$.

□

Theorem I.1. *Let P be a convex set with positive volume and $\epsilon \geqslant 0$ such that $P_\epsilon^\circ \coloneqq \{\boldsymbol{x} \in P : \operatorname{dist}(\boldsymbol{x}, \partial P) > \epsilon)\}$ is nonempty. Let $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K \in P_\epsilon^\circ$. Then, for any $\delta > 0$, there exists a collection of $K+1$ problems $\mathcal{B}^0, \ldots, \mathcal{B}^K$ such that for any algorithm,*

$$\max_{i \in \{0, \ldots, K\}} \mathbb{E}_{\mathcal{B}^i}(\mathcal{L}_{T,P,\epsilon}(\widehat{S})) \geqslant \exp(-13 \frac{T}{(1-\delta) H_P} - 25 D \log(48(\log(T)+1)KD))).$$

where

$$H_P = \sum_{i=1}^K [\operatorname{dist}(\boldsymbol{\mu}_i, \partial P) + \epsilon]^{-2}.$$

Proof. **Step 1: Reduce convex set to compact convex set.** Let $r_i := 2\ \mathrm{dist}(\boldsymbol{\mu}_i, \partial P)$ for all $i \in [K]$. Clearly, there exists large enough $B > 0$ such that $P_0 = P \cap \{\boldsymbol{x} \in \mathbb{R}^D : \|\boldsymbol{x}\|_2 \leqslant B\}$ has the property that for all $i \in [K]$, if we replace $\boldsymbol{\mu}_i$ with any $\tilde{\boldsymbol{\mu}}_i \in B_r(\boldsymbol{\mu}_i)$, then $\mathrm{dist}(\tilde{\boldsymbol{\mu}}_i, \partial P) = \mathrm{dist}(\tilde{\boldsymbol{\mu}}_i, \partial P_0)$ and

$$\mathcal{L}_{T,P,\epsilon}(S) = \mathcal{L}_{T,P_0,\epsilon}(S) \tag{48}$$

for all $S \subset [K]$. Further, for the feasible identification problem with $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K$ as the means, $H_P = H_{P_0}$.

Step 2: Define approximation. Since $\mathrm{dist}(\boldsymbol{\mu}_i, \partial P_0) > \epsilon$ for all $i \in [K]$, there exists $\gamma \in (0, \epsilon)$ such that for all $i \in [K]$,

$$\mathrm{dist}(\boldsymbol{\mu}_i, \partial P_0) > \epsilon + \gamma. \tag{49}$$

Fix $\delta > 0$. By Proposition I.1, there exists a polyhedron P_{approx} such that

$$(1 - \delta)H_{P_0} \leqslant H_{P_{approx}} \leqslant (1 + \delta)H_{P_0}$$

and

$$\sup_{\boldsymbol{y} \in \mathbb{R}^D} |\,\mathrm{dist}(\boldsymbol{y}, \partial P_{approx}) - \mathrm{dist}(\boldsymbol{y}, \partial P_0)| < \frac{\gamma}{2}. \tag{50}$$

By (49) and (50), $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K \in P_{approx}$ and for all $i \in [K]$, $\mathrm{dist}(\boldsymbol{\mu}_i, \partial P_{approx}) > \epsilon + \frac{\gamma}{2}$ so that $\mathcal{L}_{T,P_0,\epsilon}(S) = \mathcal{L}_{T,P_{approx},\epsilon}(S)$ for all $S \subset [K]$.

Step 3: Apply lower bound for polyhedra. Apply the lower bound construction from Theorem 1 to P_{approx} to define $K + 1$ collections of distributions $\mathcal{B}^i$ for $i \in \{0, 1, \ldots, K\}$ (see Theorem 1 for their definitions). We claim that for every problem $\mathcal{B}^i$, $\mathcal{L}_{T,P_0,\epsilon}(S) = \mathcal{L}_{T,P_{approx},\epsilon}(S)$ for all $S \subset [K]$. We briefly sketch the proof. First,

$$\mathrm{dist}(\boldsymbol{\mu}'_i, \partial P_0) + \frac{\gamma}{2} > \mathrm{dist}(\boldsymbol{\mu}'_i, \partial P_{approx}) = \mathrm{dist}(\boldsymbol{\mu}_i, \partial P_{approx}) > \mathrm{dist}(\boldsymbol{\mu}_i, \partial P_0) - \frac{\gamma}{2} \geqslant \epsilon + \frac{\gamma}{2}.$$

Thus, $\mathrm{dist}(\boldsymbol{\mu}'_i, \partial P_0) > \epsilon$. Further, $\boldsymbol{\mu}'_i \notin P_{approx}$, $\mathrm{dist}(\boldsymbol{\mu}'_i, P_{approx}) \geqslant \epsilon + \frac{\gamma}{2}$, and (50) imply that $\boldsymbol{\mu}'_i \notin P$. The claim follows from this observation.

Thus, by Theorem 1, for all $\widehat{S} \subset [K]$,

$$\begin{aligned}
\max_{i \in \{0,\ldots,K\}} \mathbb{E}_{\mathcal{B}^i} \mathcal{L}_{T,P,\epsilon}(\widehat{S}) &= \max_{i \in \{0,\ldots,K\}} \mathbb{E}_{\mathcal{B}^i} \mathcal{L}_{T,P_0,\epsilon}(\widehat{S}) \qquad (51)\\
&= \max_{i \in \{0,\ldots,K\}} \mathbb{E}_{\mathcal{B}^i} \mathcal{L}_{T,P_{approx},\epsilon}(\widehat{S})\\
&\geqslant \exp(-13\frac{T}{H_{P_{approx}}} - 25D\log(48(\log(T) + 1)KD)))\\
&\geqslant \exp(-13\frac{T}{H_P(1 - \delta)} - 25D\log(48(\log(T) + 1)KD)))
\end{aligned}$$

where line (51) follows by (48).

□

Theorem I.2. *Let P be a convex set with positive volume and $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K \in \mathbb{R}^D$. Suppose that there are some $B > 0$ and $\gamma > 0$ such that it is known that*

1. $\max(\,\mathrm{dist}(\boldsymbol{\mu}_i, \partial P), \|\boldsymbol{\mu}_i\|_2) \leqslant B$ *for all* $i \in [K]$,

2. $|\operatorname{dist}(\boldsymbol{\mu}_i, \partial P) - \epsilon| \geqslant \gamma$ *for all* $i \in [K]$, *and*

3. $\operatorname{dist}(\boldsymbol{\mu}_i, \partial P) \geqslant \gamma$.

Then, there exists an algorithm that given any $\delta > 0$, *achieves*

$$\mathbb{E}[\mathcal{L}_{T,P,\epsilon}(\widehat{S})] \leqslant 2(\log(T) + 1)K5^D \exp(-\frac{T}{1296R^2 H_P(1+\delta)}).$$

Proof. **Step 1: Define the algorithm.** The algorithm is as follows:

1. Set $P_0 = P \cap \overline{B_{2B}(\mathbf{0})}$.

2. Use the construction from Proposition I.1 to approximate P_0 with P_{approx} such that

$$\sup_{\boldsymbol{y} \in \mathbb{R}^D} |\operatorname{dist}(\boldsymbol{y}, \partial P_0) - \operatorname{dist}(\boldsymbol{y}, \partial P_{approx})| \leqslant \frac{\gamma}{2} \tag{52}$$
$$(1-\delta)H_{P_0} \leqslant H_{approx} \leqslant H_{P_0}(1+\delta).$$

3. Run MD-APT with the K given arms, P_{approx}, and ϵ and return its answer $\widehat{S}$.

We note that because it is known that $\operatorname{dist}(\boldsymbol{\mu}_i, \partial P) \geqslant \gamma$, step 2 of the algorithm is valid.
Step 2: Distance of $\boldsymbol{\mu}_i$ to ∂P is equal to the distance of $\boldsymbol{\mu}_i$ to ∂P_0. First, we claim that

$$\operatorname{dist}(\boldsymbol{\mu}_i, \partial P) = \operatorname{dist}(\boldsymbol{\mu}_i, \partial P_0). \tag{53}$$

Let $\boldsymbol{x} \in \partial P$ such that $\|\boldsymbol{x} - \boldsymbol{\mu}_i\|_2 = \operatorname{dist}(\boldsymbol{\mu}_i, \partial P)$. Then,

$$\|\boldsymbol{x}\|_2 \leqslant \|\boldsymbol{\mu}_i\|_2 + \|\boldsymbol{x} - \boldsymbol{\mu}_i\|_2 \leqslant 2B.$$

Thus, $\boldsymbol{x} \in P_0$. Since $P_0 \subset P$, $\boldsymbol{x} \in \partial P_0$. Therefore, $\operatorname{dist}(\boldsymbol{\mu}_i, \partial P_0) \leqslant \operatorname{dist}(\boldsymbol{\mu}_i, \partial P)$. Towards a contradiction, suppose that $\operatorname{dist}(\boldsymbol{\mu}_i, \partial P_0) < \operatorname{dist}(\boldsymbol{\mu}_i, \partial P)$. Let $\boldsymbol{y} \in \partial P_0$ such that $\|\boldsymbol{\mu}_i - \boldsymbol{y}\| = \operatorname{dist}(\boldsymbol{\mu}_i, \partial P_0)$. Then,

$$\|\boldsymbol{y}\|_2 \leqslant \|\boldsymbol{\mu}_i\|_2 + \|\boldsymbol{y} - \boldsymbol{\mu}_i\|_2 < 2B. \tag{54}$$

Recall the fact

$$\partial(A \cap B) \subset \partial A \cup \partial B.$$

Therefore, by this fact, $\boldsymbol{y} \in \partial P_0$ and (54) imply that $\boldsymbol{y} \in \partial P$. Thus, we have a contradiction. This establishes (53).

Step 3: $\mathcal{L}_{T,P,\epsilon}(S) = \mathcal{L}_{T,P_0\epsilon}(S)$ for every $S \subset [K]$. Next, we show that $\mathcal{L}_{T,P,\epsilon}(S) = \mathcal{L}_{T,P_0,\epsilon}(S)$ for all $S \subset [K]$. Suppose $\boldsymbol{\mu}_i \in S^{\text{int}}_{P,\epsilon}$. Then, by (53), $\operatorname{dist}(\boldsymbol{\mu}_i, \partial P_0) = \operatorname{dist}(\boldsymbol{\mu}_i, \partial P) \geqslant \epsilon$. Further, by hypothesis, $\|\boldsymbol{\mu}_i\|_2 \leqslant B$ so that $\boldsymbol{\mu}_i \in P_0$. Thus, $\boldsymbol{\mu}_i \in S^{\text{int}}_{P_0,\epsilon}$.

Next, suppose that $\boldsymbol{\mu}_i \in S^{\text{out}}_{P,\epsilon}$. Then, $P_0 \subset P$ implies that $\boldsymbol{\mu}_i \notin P_0$ and $\operatorname{dist}(\boldsymbol{\mu}_i, P_0) = \operatorname{dist}(\boldsymbol{\mu}_i, P) > \epsilon$ by (53). Thus, $\boldsymbol{\mu}_i \in S^{\text{out}}_{P_0,\epsilon}$.

Next, suppose that $\boldsymbol{\mu}_i \notin S^{\text{out}}_{P,\epsilon}$ and $\boldsymbol{\mu}_i \notin S^{\text{int}}_{P,\epsilon}$. Then, either *(i)* $\boldsymbol{\mu}_i \in P$ and $\operatorname{dist}(\boldsymbol{\mu}_i, \partial P) < \epsilon$ or *(ii)* $\boldsymbol{\mu}_i \notin P$ and $\operatorname{dist}(\boldsymbol{\mu}_i, \partial P) \leqslant \epsilon$. Suppose *(i)*. Then, by (53), it follows that $\boldsymbol{\mu}_i \notin S^{\text{out}}_{P_0,\epsilon}$ and $\boldsymbol{\mu}_i \notin S^{\text{int}}_{P_0,\epsilon}$. Suppose *(ii)*. Then, $P_0 \subset P$ and (53) imply that $\boldsymbol{\mu}_i \notin S^{\text{out}}_{P_0,\epsilon}$ and $\boldsymbol{\mu}_i \notin S^{\text{int}}_{P_0,\epsilon}$. This establishes the claim.

Step 4: Putting it together. (52) and the hypotheses imply that $\mathcal{L}_{T,P_0,\epsilon}(S) = \mathcal{L}_{T,\epsilon,P_{approx}}(S)$ for all $S \subset [K]$.

Thus, let $\widehat{S}$ denote the output of MD-APT with the K given arms, P_{approx}, and ϵ. By Theorem 4,

$$\begin{aligned}
\mathbb{E}[\mathcal{L}_{T,P,\epsilon}(S)] &= \mathbb{E}[\mathcal{L}_{T,P_0,\epsilon}(S)] \\
&= \mathbb{E}[\mathcal{L}_{T,P_{approx},\epsilon}(S)] \\
&\leqslant 2(\log(T)+1)K5^D \exp(-\frac{T}{1296R^2 H_{P_approx}}) \\
&\leqslant 2(\log(T)+1)K5^D \exp(-\frac{T}{1296R^2 H_P(1+\delta)}).
\end{aligned}$$

□

J Additional Experiments

In this section, we present a couple more experiments. First, we present another variant of experiment 3, linear progression of arms on a cube, where there are no irrelevant arms. We set $\epsilon = 0$. We use $\boldsymbol{\mu}_{0:3} = (.75)^{\otimes 5} + (0:3) \times .05$, $\boldsymbol{\mu}_4 = (.95)^{\otimes 5}$, $\boldsymbol{\mu}_5 = (1.05)^{\otimes 5}$, $\boldsymbol{\mu}_{6:9} = (1.25)^{\otimes 5} - (0:3) \times .05$, $\boldsymbol{\mu}_{10:19} = (1.15)^{\otimes 5}$. In comparison to experiment 3, we make it slightly easier to determine whether the arms $\boldsymbol{\mu}_4$ and $\boldsymbol{\mu}_5$ belong to the polyhedron because otherwise the difficulty of the problem prevents any algorithm from achieving substantial progress after 2000 time steps. Figure 1 presents the results. MD-SAR performs substantially better than MD-APT. MD-APT pulls arm 4, which minimizes Δ_i, too much. MD-APT pulls arm 4 time on average 1006.27 times, whereas MD-SAR pulls arm 4 on average 319.59 times.

We also repeat the crowdsourcing experiment with a slightly different setup. Now, we draw samples from a Gaussian distribution for each worker with mean calculated from the dataset in Snow et al. [2008] and variance over all the ratings over all the workers. The results are very similar to the results in the main text.

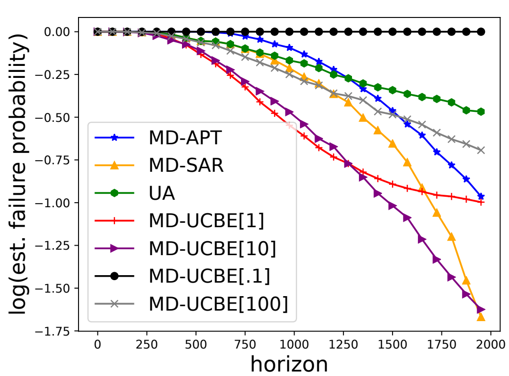

Figure 1: Linear Progression on Cube, no Irrelevant Arms

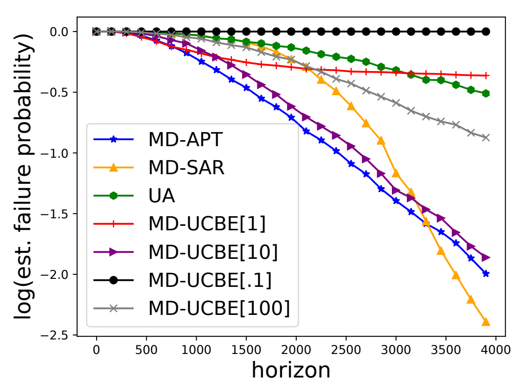

Figure 2: Crowdsourcing Experiment with Simulated Data

References

J.-Y. Audibert and S. Bubeck. Best arm identification in multi-armed bandits. *Conference on Learning Theory*, 2010.

D. Bertsekas. *Convex optimization theory*. Belmont: Athena Scientific, 2009.

S. Boyd and L. Vandenberghe. Convex optimization. *Cambridge University press*, 2004.

K. Jamieson, M. Malloy, R. Nowak, and S. Bubeck. lil'ucb: An optimal exploration algorithm for multi-armed bandits. *Conference on Learning Theory*, pages 424–439, 2014.

A. Locatelli, M. Gutzeit, and A. Carpentier. An optimal algorithm for the thresholding bandit problem. *Proceedings of The 33rd International Conference on Machine Learning*, pages 1690–1698, 2016.

R. Snow, B. O'Connor, D. Jurafsky, and A. Ng. Cheap and fast—but is it good?: evaluating non-expert annotations for natural language tasks. *Proceedings of the conference on empirical methods in natural language processing*, pages 254–263, 2008.

R. Vershynin, P. Hsu, C. Ma, J. Nelson, E. Schnoor, D. Stoger, T. Sullivan, and T. Tao. High-dimensional probability: An introduction with applications in data science. 2017.