A. Missing Proofs from Section 3

Before proving Lemma 4, we first need the following Lemma that bounds the probability of a single arm being misordered when deciding which arms to remove in a round:

Lemma 3. *Assume that the best arm was not eliminated prior to round m. Let $[x]_+ = \max(x, 0)$. Then for any arm $x_i \in S_m$,*

$$\mathbb{P}[\langle\hat{\theta}_m, x_i\rangle > \langle\hat{\theta}_m, x_1\rangle] \leq \exp\left(-\frac{\left[\Delta_i - (2 + 2\sqrt{2h_{|S_m|/4}})\gamma_{\max}\right]_+^2 T}{16\log_2(n)h_{|S_m|/4}\sigma^2}\right)$$

Proof. For simplicity, we will drop all subscripts m in this proof, so that we say $\hat{\theta}$ instead of $\hat{\theta}_m$ and $\mathcal{Z}$ instead of $\mathcal{Z}_m$ etc. By slight abuse of notation, let $y \in \mathbb{R}^N$ be the vector of true rewards for the sampled arms $\mathcal{Z}$, let $\gamma \in \mathbb{R}^N$ be the misspecification vector, and let $X \in \mathbb{R}^{N \times d}$ be the design matrix whose rose are the elements of $\mathcal{Z}$. That is, we have $y = X\theta + \gamma$. Let $s = \hat{y} - y$, all of whose components are independent 1-subgaussian random variables

Fix some index i. Notice that if i does not satisfy $\Delta_i > (2 + 2\sqrt{2h_{S_m}})|\gamma_{\max}|$, then the statement is trivially true. Therefore, we may safely assume $\Delta_i > (2 + 2\sqrt{2h_{S_m}})|\gamma_{\max}|$

Let $\theta' = \text{argmin}_{X\theta = X\theta'} \|\theta'\|_2$. Notice that $y = X\theta + \gamma = X\theta' + \gamma$ implies that using θ' does not introduce more misspecification. (In fact, $\theta = \theta'$ if X has rank d.) Further, by Lemma 10, we have that $x_i - x_j$ is in the span of $\mathcal{Z}$ for all x_i and x_j in S_m, so that $\langle\theta', x_i - x_j\rangle = \langle\theta, x_i - x_j\rangle$ for all i and j.

Then, after the player obtains the vector of rewards $\hat{y}$ by playing each arm in $\mathcal{Z}$, the estimate $\hat{\theta}$ is given by:

$$\begin{aligned}\hat{\theta} &= \underset{\theta \in span(\mathcal{Z})}{\text{argmin}} \sum_{x_i \in \hat{\mathcal{Z}}} (\langle\theta, x_i\rangle - \hat{y}_i)^2 \\ &= \text{argmin}\{\|\hat{\theta}\|_2 \text{ s.t. } X^T X\hat{\theta} = X^T\hat{y}\}\end{aligned}$$

Now, since for any v there exists u such that $X^T X u = X^T v$, we have:

$$\begin{aligned}\hat{\theta} &= (X^T X)^\ddagger X^T \hat{y} = (X^T X)^\ddagger X^T (X\theta + s + \gamma) \\ &= \theta' + (X^T X)^\ddagger X^T (s + \gamma)\end{aligned}$$

This implies:

$$\begin{aligned}\langle\hat{\theta}, x_i\rangle &> \langle\hat{\theta}, x_1\rangle \\ \langle\hat{\theta}, x_i - x_1\rangle - y_i + y_1 &\geq \Delta_i \\ \langle\hat{\theta} - \theta', x_i - x_1\rangle &> \Delta_i - \gamma_1 + \gamma_i \\ &\geq \Delta_i - 2\gamma_{\max}\end{aligned}$$

Substituting in our expression for $\hat{\theta}$, we have that the probability that the empirical average of the best arm is smaller than the empirical average of arm i is at most

$$\mathbb{P}[(x_i - x_1)^T (X^T X)^\ddagger X^T (s + \gamma) > \Delta_i - 2\gamma_{\max}]$$

Next, we will need some guarantee about the quality of the set $\mathcal{Z}$. By Theorem 5, `GetArms` returns $\mathcal{Z}$ that satisfies $|\mathcal{Z}| = N$ and

$$\sup_{x_i - x_j \in S_m} (x_j - x_i)^T \left(\sum_{z \in \mathcal{Z}} zz^T\right)^\ddagger (x_j - x_i) \leq 8\frac{h_{|S_m|/4}}{N} \tag{3}$$

Using this fact, we apply Cauchy-Schwarz to obtain:

$$\begin{aligned}
|(x_i - x_1)^T (X^T X)^\ddagger X^T \gamma| &\leq \|X(X^T X)^\ddagger (x_i - x_1)\|_2 \|\gamma\|_2 \\
&\leq \sqrt{N}\|\gamma\|_\infty \sqrt{\sup_{x_i, x_j \in S_m} (x_j - x_i)^T \left(X^T X\right)^\ddagger (x_j - x_i)} \\
&= \sqrt{N}\|\gamma\|_\infty \sqrt{\sup_{x_i, x_j \in S_m} (x_j - x_i)^T \left(\sum_{z \in \hat{\mathcal{Z}}} z z^T\right)^\ddagger (x_j - x_i)} \\
&\leq \sqrt{N}\sqrt{\frac{8h_{|S_m|/4}}{N}}\|\gamma\|_\infty \leq \sqrt{8h_{|S_m|/4}}\gamma_{\max}
\end{aligned}$$

Thus,

$$\mathbb{P}[(x_i - x_1)^T (X^T X)^\ddagger X^T (s + \gamma) > \Delta_i - 2\gamma_{\max}] \leq$$
$$\mathbb{P}\left[(x_i - x_1)^T (X^T X)^\ddagger X^T s > \Delta_i - \left(2 + \sqrt{8h_{\frac{|S_m|}{4}}}\right)\gamma_{\max}\right]$$

Then, since s is a mean zero, σ-subgaussian vector, for any $v \in \mathbb{R}^N$, $\langle v, s\rangle$ is a mean 0, $\|v\|_2\sigma$-subgaussian random variable. Applying the Hoeffding bound, for any $\epsilon > 0$,

$$\mathbb{P}[\langle v, s\rangle > \epsilon] \leq \exp\left(-\frac{\epsilon^2}{2\|v\|_2^2\sigma^2}\right)$$

We will substitute $X(X^T X)^\ddagger(x_i - x_1)$ for v in the above formulation of Hoeffding's bound. To start, we compute the norm:

$$\begin{aligned}
&\|X(X^T X)^\ddagger (x_i - x_1)\|_2^2 = (x_i - x_1)^T (X^T X)^\ddagger (x_i - x_1) \\
&\leq \sup_{x_i, x_j \in Z} (x_j - x_i)^T \left(\sum_{z \in \hat{\mathcal{Z}}} z z^T\right)^\ddagger (x_j - x_i) \leq 8\frac{h_{|S_m|/4}}{N}
\end{aligned}$$

Then substituting $N = \frac{T}{\log_2 n}$, and combining the above, we get that

$$\mathbb{P}[(x_i - x_1)^T (X^T X)^\ddagger X^T (s + \gamma) \geq \Delta_i - 2\gamma_{\max}] \leq$$
$$\exp\left(-\frac{(\Delta_i - (2 + 2\sqrt{2h_{|S_m|/4}})\gamma_{\max})^2 T}{16\log_2(n) h_{|S_m|/4}\sigma^2}\right)$$

□

Next, following (Karnin et al., 2013), we provide the missing proof of Lemma 4:

Lemma 4. *Assume that the best arm was not eliminated prior to round m, and let $[x]_+ = \max(0, x)$. Then the probability that the best arm is eliminated on round m is at most*

$$3\exp\left(-\frac{\left[\Delta_{\frac{1}{4}|S_m|} - (2 + 2\sqrt{2h_{\frac{1}{4}|S_m|}})\gamma_{\max}\right]_+^2 T}{16\log_2(n) h_{|S_m|/4}\sigma^2}\right)$$

Proof. If the best arm is thrown out at round m, there are at least $\frac{1}{2}|S_m|$ arms in S_m whose $\hat{y}$ estimates are higher than that of the best arm. Let $S'_m \subset S_m$ be the set of arms that excludes the $\frac{1}{4}|S_m|$ arms with the largest true means in S_m. If the best arm is thrown out, then at least $\frac{1}{3}$ of arms in S'_m must have higher $\hat{y}$ estimates than that of the best arm. Let N_m be the number of such arms. Define $D = \max\{(\Delta_{\frac{1}{4}|S_m|} - (2 + 2\sqrt{2h_{|S_m|/4}})\gamma_{\max})^2, 0\}$. Then using Lemma 3, the expected number of such arms is at most

$$\begin{aligned}\mathbb{E}[N_m] &= \sum_{x_i \in S'_m} \mathbb{P}[\langle \hat{\theta}, x_i \rangle \geq \langle \hat{\theta}, x_1 \rangle] \\ &\leq |S'_m| \exp\left(-\frac{DT}{16 \log_2(n) h_{|S_m|/4} \sigma^2}\right)\end{aligned}$$

Then, by Markov inequality, the probability of the best arm being thrown out at round m is at most

$$\mathbb{P}\left[N_m > \frac{1}{3}|S'_m|\right] \leq \frac{\mathbb{E}[N_m]}{\frac{1}{3}|S'_m|} \leq 3 \exp\left(-\frac{DT}{16 \log_2(n) h_{|S_m|/4} \sigma^2}\right)$$

□

B. Missing Proofs from Section 4

Before proving Theorem 5, we need the following Lemmas, which characterize the quality of the candidate set of arms returned by the method of (Allen-Zhu et al., 2017).

First, we have the following easy technical Lemma:

Lemma 10. *If for some sets S and $\mathcal{Z}$*

$$\sup_{x_i - x_j \in S} (x_j - x_i)^T (\sum_{z \in \mathcal{Z}} z z^T)^{\ddagger} (x_j - x_i) < \infty$$

then $S \subset span(\mathcal{Z})$.

Proof. Fix any $x = x_i - x_j \neq 0, x_i, x_j \in S$. Then, since

$$x^T (\sum_{z \in \mathcal{Z}} z z^T)^{\ddagger} x < \infty$$

there exists y s.t.

$$(\sum_{z \in \mathcal{Z}} z z^T) y = \sum_{z \in \mathcal{Z}} z (z^T y) = x$$

which implies that $x \in span(\mathcal{Z})$. □

Lemma 11. *Given the set $S \subset \mathcal{A} \subset \mathbb{R}^d$ the objective f_S above, and a number $N \geq d$ the output $\hat{\mathcal{Z}}$ of `OptDesign` satisfies $\hat{\mathcal{Z}} \subset \mathcal{A}, |\hat{\mathcal{Z}}| \leq N$ and $\sup_{x_i, x_j \in S} (x_j - x_i)^T \left(\sum_{z \in \hat{\mathcal{Z}}} z z^T\right)^{\ddagger} (x_j - x_i) \leq \frac{14}{N} \inf_v \inf_{\substack{\|\pi\|_1 \leq N \\ \pi \in [0,N]^{|\mathcal{A}|}}} \sup_{x_i \in S} (x_i - v)^T (\sum_{a \in \mathcal{A}} \pi(a) a a^T)^{\ddagger} (x_i - v) \leq 14 \frac{d}{N}$. That is, $\sup_{x_i, x_j \in S} (x_j - x_i)^T \left(\sum_{z \in \hat{\mathcal{Z}}} z z^T\right)^{\ddagger} (x_j - x_i) \leq \tilde{d}$ and also $\tilde{d} \leq d$.*

Proof. Let

$$\mathcal{Z} = \underset{\substack{\mathcal{Z}' \subset \mathcal{A} \\ |\mathcal{Z}'| \leq N}}{\text{argmin}} \{ \sup_{x_i, x_j \in S} (x_j - x_i)^T (\sum_{z \in \mathcal{Z}'} z z^T)^{\ddagger} (x_j - x_i) \} \tag{4}$$

Note that we consider $\mathcal{Z}$ to be a set *with multiplicity* - it is permitted for an element to appear multiple times in $\mathcal{Z}$. Computing $\mathcal{Z}$ exactly is NP-hard, but (Allen-Zhu et al., 2017) provides a way to approximate the solution. Note that equation 4 can be restated as follows:

$$c^* = \underset{\substack{\|c\|_1 \leq N \\ c \in \{0,1,2,\ldots,N\}^{|\mathcal{A}|}}}{\text{argmin}} \{ \sup_{x_i, x_j \in S} (x_j - x_i)^T (\sum_{a \in \mathcal{A}} c(a) a a^T)^{\ddagger} (x_j - x_i) \} \tag{5}$$

where c is treated as a vector in $|\mathcal{A}|$ indexed by $a \in \mathcal{A}$ and $\mathcal{Z}$ is related to c^* as

$$\mathcal{Z} = \{a \in \mathcal{A} \text{ repeated } c^*(a) \text{ times}\}$$

Then, define the continuous relaxation of the objective in 5 as:

$$\pi^* = \underset{\substack{\|\pi\|_1 \le N \\ \pi \in [0,N]^{|\mathcal{A}|}}}{\operatorname{argmin}} \{ \sup_{x_i, x_j \in S} (x_j - x_i)^T (\sum_{a \in \mathcal{A}} \pi(a) a a^T)^{\ddagger} (x_j - x_i) \} \tag{6}$$

Section 3 from (Allen-Zhu et al., 2017) guarantees that a polynomial-time continuous optimization procedure can find a fractional solution π such that

$$\sup_{x_i, x_j \in S} (x_j - x_i)^T (\sum_{a \in \mathcal{A}} \pi(a) a a^T)^{\ddagger} (x_j - x_i) \le 7/6 \sup_{x_i, x_j \in S} (x_j - x_i)^T (\sum_{a \in \mathcal{A}} \pi^*(a) a a^T)^{\ddagger} (x_j - x_i)$$

At the same time, Theorem 2.1 from (Allen-Zhu et al., 2017) (see Appendix C) provides a polynomial-time algorithm that rounds any fractional solution π to an integer solution c such that

$$\sup_{x_i, x_j \in S} (x_j - x_i)^T (\sum_{a \in \mathcal{A}} c(a) a a^T)^{\ddagger} (x_j - x_i) \le 3 \sup_{x_i, x_j \in S} (x_j - x_i)^T (\sum_{a \in \mathcal{A}} \pi(a) a a^T)^{\ddagger} (x_j - x_i)$$

Combining the two equations, we get that there exists a polynomial-time algorithm that finds an integer solution c (or, equivalently, set $\mathcal{Z}$) such that

$$\sup_{x_i, x_j \in S} (x_j - x_i)^T (\sum_{a \in \mathcal{A}} c(a) a a^T)^{\ddagger} (x_j - x_i) \le 7/2 \sup_{x_i, x_j \in S} (x_j - x_i)^T (\sum_{a \in \mathcal{A}} \pi^*(a) a a^T)^{\ddagger} (x_j - x_i)$$

Now, notice that for any solution π, we have:

$$\begin{aligned}
&\sup_{x_i, x_j \in S} (x_j - x_i)^T (\sum_{a \in \mathcal{A}} \pi(a) a a^T)^{\ddagger} (x_j - x_i) \\
&= \inf_v \sup_{x_i, x_j \in S} (x_j - v + v x_i)^T (\sum_{a \in \mathcal{A}} \pi(a) a a^T)^{\ddagger} (x_j - v + v x_i) \\
&\le \inf_v \sup_{x_i, x_j \in S} (x_i - v)^T (\sum_{a \in \mathcal{A}} \pi(a) a a^T)^{\ddagger} (x_i - v) + (x_j - v)^T (\sum_{a \in \mathcal{A}} \pi(a) a a^T)^{\ddagger} (x_j - v) + 2 (x_j - v)^T (\sum_{a \in \mathcal{A}} \pi^*(a) a a^T)^{\ddagger} (x_i - v)
\end{aligned}$$

Now, observe that $a^\top M b = \langle \sqrt{M} a, \sqrt{M} b \rangle \le \frac{a^\top M a}{2} + \frac{b^\top M b}{2}$ for any positive-semidefinite matrix M by Young inequality. Therefore:

$$\begin{aligned}
\sup_{x_i, x_j \in S} (x_j - x_i)^T (\sum_{a \in \mathcal{A}} \pi^*(a) a a^T)^{\ddagger} (x_j - x_i) &\le 4 \inf_v \inf_{\substack{\|\pi\|_1 \le N \\ \pi \in [0,N]^{|\mathcal{A}|}}} \sup_{x \in S} (x - v)^T (\sum_{a \in \mathcal{A}} \pi(a) a a^T)^{\ddagger} (x - v) \\
&\le \frac{4}{N} \inf_v \inf_{\substack{\|\pi\|_1 \le 1 \\ \pi \in [0,N]^{|\mathcal{A}|}}} \sup_{x \in S} (x - v)^T (\sum_{a \in \mathcal{A}} \pi(a) a a^T)^{\ddagger} (x - v) \\
&\le 4 \frac{d}{N}
\end{aligned}$$

where the last line follows from setting $v = 0$ and applying the Kiefer-Wolfowitz theorem (Kiefer & Wolfowitz, 1960).

Thus, combining everything, we get that

$$\begin{aligned}
&\sup_{x_i, x_j \in S} (x_j - x_i)^T (\sum_{a \in \mathcal{A}} c(a) a a^T)^{\ddagger} (x_j - x_i) \le \\
&7/2 \sup_{x_i, x_j \in S} (x_j - x_i)^T (\sum_{a \in \mathcal{A}} \pi^*(a) a a^T)^{\ddagger} (x_j - x_i) \le \\
&\frac{14}{N} \inf_v \inf_{\substack{\|\pi\|_1 \le 1 \\ \pi \in [0,N]^{|\mathcal{A}|}}} \sup_{x \in S} (x - v)^T (\sum_{a \in \mathcal{A}} \pi(a) a a^T)^{\ddagger} (x - v) \le \\
&14 \frac{d}{N}
\end{aligned}$$

□

Lemma 11 depends only on the budget N and the number of dimensions d. Notice that the bound does not depend on distrbution of the arms.

Lemma 12. *Given the objective f_S, the set $\mathcal{A}$, and a number $N \geq |S|$, the output $\mathcal{Z}$ of* `GetArms` *satisfies $|\mathcal{Z}| \leq N$ and $\sup_{x_i,x_j \in S}(x_j - x_i)^T \left(\sum_{z\in\mathcal{Z}} zz^T\right)^\ddagger (x_j - x_i) \leq 2\frac{|S|}{N}$.*

Proof. By Lemma 14, for any $x_i, x_j \in S$,

$$(x_j - x_i)^T(\sum_{x\in S} xx^T)^\ddagger(x_j - x_i) \leq 2$$

So, letting $\tilde{\mathcal{Z}} \leftarrow \{s \in S \text{ repeated } \frac{N}{S} \text{ times}\}$

$$\begin{aligned}
&\sup_{x_i,x_j\in\mathcal{S}}(x_j - x_i)^T(\sum_{z\in\tilde{\mathcal{Z}}} zz^T)^\ddagger(x_j - x_i) \\
&\leq \frac{|S|}{N}\sup_{x_i,x_j\in S}(x_j - x_i)^T(\sum_{x\in S} xx^T)^\ddagger(x_j - x_i) \\
&\leq 2\frac{|S|}{N}
\end{aligned}$$

□

While the above Lemma requires the number of pulls to be greater than the candidate set S, the next Lemma shows that it is possible to bound the performance of `OptDesign` even if $N < |S|$.

Lemma 13. *Given the objective f_S, the set $\mathcal{A}$, and a number N, the output $\mathcal{Z}$ of* `GetArms` *satisfies $|\mathcal{Z}| \leq N$ and $\sup_{x_i,x_j \in S}(x_j - x_i)^T \left(\sum_{z\in\mathcal{Z}} zz^T\right)^\ddagger (x_j - x_i) \leq 6\frac{|S|}{N}$.*

Proof. Given any fractional solution to f_S π, using Theorem 2.1 from (Allen-Zhu et al., 2017), `OptDesign` outputs an integer solution $\mathcal{Z}$ such that

$$\begin{aligned}
&\sup_{x_i,x_j\in s}(x_j - x_i)^T(\sum_{z\in\mathcal{Z}} zz^T)^\ddagger(x_j - x_i) \leq \\
&3\sup_{x_i,x_j\in\mathcal{Z}}(x_j - x_i)^T(\sum_{z\in\mathcal{Z}} \pi_z zz^T)^\ddagger(x_j - x_i) \leq \\
&3\sup_{z_i,z_j\in\mathcal{Z}}(x_j - x_i)^T(\sum_{z\in S} \frac{N}{|S|} zz^T)^\ddagger(x_j - x_i) \leq \\
&3\frac{|S|}{N}\sup_{x_i,x_j\in S}(x_j - x_i)^T(\sum_{z\in S} zz^T)^\ddagger(x_j - x_i) \leq 6\frac{|S|}{N}
\end{aligned}$$

where the last step uses Lemma 14 below. □

Now we prove Lemma 14:

Lemma 14. *Let $x_1, \ldots x_n$ be arbitrary vectors in $\mathbb{R}^d$. Let X be a matrix s.t. ith row of X corresponds to x_i for all i. Then for any j, k,*

$$(x_j - x_k)^T(X^TX)^\ddagger(x_j - x_k) \leq 2$$

Proof. For any j, $x_j = X^Te_j$ where e_j is the jth identity vector. Then for any j, k,

$$\begin{aligned}
&(x_j - x_k)^T(X^TX)^\ddagger(x_j - x_k) = \\
&(e_j - e_k)^TX(X^TX)^\ddagger X^T(e_j - e_k)
\end{aligned}$$

Let $e = e_j - e_k$, and let $u = (X^T X)^{\ddagger} X^T e$. Then,

$$X^T X u = X^T e \implies X^T(Xu - e) = 0 \implies$$
$$Xu - e \in ker(X^T) \implies Xu - e \perp Im(X)$$

Then there exists a w such that $e = Xu + Qw$, where Q is a matrix with columns orthogonal to columns of X. Then,

$$e^T X(X^T X)^{\ddagger} X^T e = u^T X^T X(X^T X)^{\ddagger} X^T X u$$

By definition of u, $u \perp ker(X^T X)$, and thus

$$(x_j - x_k)^T (X^T X)^{\ddagger} (x_j - x_k) =$$
$$u^T X^T X(X^T X)^{\ddagger} X^T X u =$$
$$u^T X^T X u = \|Xu\|_2^2 \leq \|e\|_2^2 = 2$$

because $\|e\|_2^2 = \|Xu\|_2^2 + \|Qw\|_2^2$ □

Given the above, we are ready to proof Theorem 5.

Proof of Theorem 5. Combining Lemmas 11, 12, and 13, we get that on inputs S, N and $\mathcal{A}$, `GetArms` produces a set $\mathcal{Z}$ such that if $\frac{T}{\lceil \log_2 n \rceil} = N \geq |S|$, `GetArms` returns a set $\mathcal{Z}$ that satisfies:

$$\sup_{x_i, x_j \in S} (x_j - x_i)^T (\sum_{z \in \mathcal{Z}} z z^T)^{\ddagger} (x_j - x_i) \leq \min\{\frac{8|S|/4}{N}, \frac{14\tilde{d}}{N}\} = 8\frac{h_{|S|/4}}{N}$$

Similarly, if $N < |S|$, `GetArms` returns a set $\mathcal{Z}$ that satisfies:

$$\sup_{x_i, x_j \in S} (x_j - x_i)^T (\sum_{z \in \mathcal{Z}} z z^T)^{\ddagger} (x_j - x_i) \leq \min\{14\frac{\tilde{d}}{N}, 24\frac{|S|/4}{N}\} = 8\frac{h_{|S|/4}}{N}$$

□

C. Properties of the Objective f_S

In this section we verify that the assumptions required to use the approach of (Allen-Zhu et al., 2017) in `GetArms` hold.

C.1. Equivalence of objectives

Lemma 15. *Assume we are given a set of n vectors $\mathcal{Z} \subset \mathbb{R}^d$ that span a k-dimensional subspace. Let $Z \in \mathbb{R}^{n \times d}$ be a matrix such that each row represents a vector in $\mathcal{Z}$. Let $Q \in \mathbb{R}^{k \times d}$ be a matrix with orthogonal rows such that $span(Q^T) = span(Z^T)$. Define $M = (QQ^T)^{-1/2} Q$. Then, for any symmetric, positive semi-definite matrix $A \in \mathbb{R}^{d \times d}$ such that $A^{\ddagger} z < \infty$ for all $z \in \mathcal{Z}$, it holds that MAM^T is invertible and for all $z \in \mathcal{Z}$*

$$(Mz)^T (MAM^T)^{-1} Mz = z^T A^{\ddagger} z$$

Proof.

$$MAM^T = (QQ^T)^{-1/2} QAQ^T (QQ^T)^{-1/2}$$

Since A is a symmetric PSD matrix, it has a symmetric PSD root $A^{1/2}$. Since $(QQ^T)^{-1/2}$ is invertible, $A^{1/2} Q^T (QQ^T)^{-1/2}$ has rank less than k iff there exists $v \neq 0$,

$$A^{1/2} Q^T v = A^{1/2} (\sum_{i=1}^{k} v_i q_i) = 0$$

where q_i correspond to rows of Q. By assumption, for each $z_i \in \mathcal{Z}$, there exists y_i s.t. $Ay_i = A^{1/2}A^{1/2}y_i = z_i$. Thus, $z_i \in Im(A^{1/2})$ for all i. Since $span(Z^T) = span(Q^T)$, it must be that $\sum_{i=1}^k v_i q_i \in Im(A^{1/2})$. Since the kernel of $A^{1/2}$ is orthogonal to the image of $A^{1/2}$, we get that $A^{1/2}(\sum_{i=1}^k v_i q_i) = 0$ iff $\sum_{i=1}^k v_i q_i = 0$, which is impossible, since q_i form an orthogonal set.

Thus, $B = A^{1/2}Q^T(QQ^T)^{-1/2}$ has rank k, and thus $MAM^T = B^T B \in \mathbb{R}^{k\times k}$ is invertible.

Moreover, note that if $\mathcal{Z}$ is not in the image of A, then MAM^T is not invertible.

For the second part of the Lemma, fix some $z \in \mathcal{Z}$. By assumption, there exists $y \in \mathbb{R}^d$ s.t. $A^{\ddagger}z = y$, which implies that $Ay = z$ and y is orthogonal to $ker(A)$. Since A is a symmetric PSD matrix, this means that $y \in Im(A)$. Let $Q_1 \in \mathbb{R}^{k_1\times d}$ be an orthogonal row matrix s.t. the rows of Q and Q_1 together span $Im(A)$ and rows of Q and Q_1 are orthogonal to each other. Notice that that implies that $A = [Q^T|Q_1^T]\Lambda[Q^T|Q_1^T]^T$ where Λ is diagonal. Then, $y = Q^T w + Q_1^T w_1$ for some w, w_1. Then substituting the definition of M and using the fact that $Ay = z$,

$$\begin{aligned}&(Mz)^T(MAM^T)^{-1}Mz = z^TQ^T(QAQ^T)^{-1}QAy =\\&z^TQ^T(QAQ^T)^{-1}QA(Q^Tw + Q_1^Tw_1) =\\&z^TQ^Tw + z^TQ^T(QAQ^T)^{-1}QAQ_1^Tw_1\end{aligned}$$

Then,

$$\begin{aligned}&QAQ_1^T = Q[Q^T|Q_1^T]\Lambda[Q^T|Q_1^T]^TQ_1^T =\\&[I_{k\times k}|\mathbf{0}_{k\times k_1}]\Lambda[\mathbf{0}_{k_1\times k}^T|I_{k_1\times k_1}]^T = \mathbf{0}\end{aligned}$$

Thus, the equation above becomes

$$(Mz)^T(MAM^T)^{-1}Mz = z^TQ^Tw$$

Since columns in Q_1 are orthogonal to columns of z,

$$\begin{aligned}&z^TA^{\ddagger}z = z^Ty = z^TQ^Tw =\\&(Mz)^T(MAM^T)^{-1}Mz\end{aligned}$$

□

Using the above, given a set of of arms $\mathcal{A} \subset \mathbb{R}^d$ and a set $\mathcal{Z} \subset \mathbb{R}^d$, define M as in Lemma 15. Then for any subset $\mathcal{A}' \subset \mathcal{A}$,

$$\begin{aligned}&\sup_{z\in\mathcal{Z}}(Mz)^T(\sum_{x\in\mathcal{A}'} Mx(Mx)^T)^{-1}Mz =\\&\sup_{z\in\mathcal{Z}} z^T(\sum_{x\in\mathcal{A}'} xx^T)^{\ddagger}z\end{aligned}$$

where we use

$$(Mz)^T(\sum_{x\in\mathcal{A}'} Mx(Mx)^T)^{-1}Mz = \infty$$

if $\sum_{x\in\mathcal{A}'} Mx(M_{\mathcal{Z}}x)^T$ is not invertible. Notice that if $\mathcal{Z} = \{z_i - z_j | z_i, z_j \in \mathcal{S}\}$ for some set of arms $\mathcal{S}$, and $\mathcal{S} \subset span(\mathcal{A})$, then $\mathcal{Z}$ is in the image of $\sum_{x\in\mathcal{A}} xx^T$.

Thus, if we are give a number N, sets $\mathcal{Z}, \mathcal{A}$, then

$$\begin{aligned}&\operatorname*{argmin}_{\substack{\mathcal{X}\subset\mathcal{A}\\|\mathcal{A}|\le N}} f_{\mathcal{Z}}(\mathcal{X}) =\\&\operatorname*{argmin}_{\substack{\mathcal{X}\subset\mathcal{A}\\|\mathcal{A}|\le N}} \sup_{z\in\mathcal{Z}} z^T(\sum_{x\in\mathcal{X}} xx^T)^{\ddagger}z =\\&\operatorname*{argmin}_{\substack{\mathcal{X}\subset\mathcal{A}\\|\mathcal{A}|\le N}} \sup_{z\in\mathcal{Z}}(Mz)^T(\sum_{x\in\mathcal{X}} Mx(M_{\mathcal{Z}}x)^T)^{-1}Mz\end{aligned}$$

The algorithm `GetArms` ensures that $\mathcal{Z} \subset \mathcal{A}$, and $N \geq \dim(span(\mathcal{Z}))$ whenever it attempts to find a solution to

$$\operatorname*{argmin}_{\substack{\mathcal{X} \subset \mathcal{A} \\ |\mathcal{A}| \leq N}} f_{\mathcal{Z}}(\mathcal{X})$$

Thus, every time the minimization algorithm for the above objective is called, there exists a subset $\mathcal{X}$ that makes $f_{\mathcal{Z}}$ finite. Moreover, $\sum_{x \in \mathcal{X}} Mx(Mx)^T$ is invertible.

Using this, we can now replace the minimization objective with

$$\operatorname*{argmin}_{\substack{\tilde{\mathcal{X}} \subset \tilde{\mathcal{A}} \\ |\tilde{\mathcal{A}}| \leq N}} \sup_{z \in \tilde{\mathcal{Z}}} z^T (\sum_{x \in \tilde{\mathcal{X}}} xx^T)^{-1} z$$

where $\tilde{\mathcal{A}} = M\mathcal{A}$ and $\tilde{\mathcal{Z}} = M\mathcal{Z}$. Notice that then, the solution to the objective $\tilde{X}$ forms an invertible matrix $\sum_{x \in \tilde{\mathcal{X}}} xx^T$. Thus, to find an approximate solution, we can use the procedure in (Allen-Zhu et al., 2017) that finds an approximate solution over the space of positive-definite matrices. Moreover, if $\mathcal{Z} \subset \mathbb{R}^d$ span $\mathbb{R}^d$, then $M = I$ and the two objective are exactly the same.

C.2. Solution over the set of positive-definite matrices

By the discussion in the previous section, we can assume that othe objective optimized in `GetArms` has the following form:

Given a set of arms $\mathcal{A} \subset \mathbb{R}^k$ that span $\mathbb{R}^k$, a subset $Z \in \mathcal{A}$ and number of pulls per round N,

$$f(M) = \sup_{\substack{z = z_i - z_j \\ z_i, z_j \in \mathcal{A}}} z^T M^{-1} z$$

Below, we will show that f satisfies the conditions required by the approximation algorithm (Allen-Zhu et al., 2017) hold.

f satisfies the following assumptions (see (Allen-Zhu et al., 2017)):

(A1) Monotonicity: for any $A, B \in \mathbb{S}_d^+$ with $A \preceq B$, $f(A) \geq f(B)$

(A2) Reciprocal sub-linearity: for any $A \in \mathbb{S}_d^+$ and $t \in (0, 1)$, $f(tA) \leq t^{-1} f(A)$

(A3) Polynomial-time approximability of continuous relaxation: for any fixed $\delta \in (0, 1)$, the continuous relaxation of 4 defined as

$$\min_{s \in C} F(s) = \min_{s \in C} f(\sum_{i=1}^{Nn} s_i x_i x_i^T) \text{ where}$$

$$C = \{s \in [0, N]^{Nn} : \sum_{i=1}^{Nn} s_i \leq N\}$$

can be solved with $(1 + \delta)$-relative error by a polynomial-time algorithm.

Assumptions (A1) and (A2) trivially hold. As for assumption (A3), notice that

Theorem 16. *f is a convex function in s over a convex set C.*

Proof. Given any $s_1, s_2 \in C$ and $t \in [0, 1]$,

$$\left(t \sum_{i=1}^{Nn} s_{1,i} x_i x_i^T + (1 - t) \sum_{i=1}^{Nn} s_{2,i} x_i x_i^T \right)^{-1} \preceq$$

$$t \left(\sum_{i=1}^{Nn} s_{1,i} x_i x_i^T \right)^{-1} + (1 - t) \left(\sum_{i=1}^{Nn} s_{2,i} x_i x_i^T \right)^{-1}$$

Let

$$z = \underset{\substack{z=z_i-z_j \\ z_i,z_j \in Z}}{\operatorname{argmax}}\, z^T \left(t \sum_{i=1}^{Nn} s_{1,i} x_i x_i^T + (1-t) \sum_{i=1}^{Nn} s_{2,i} x_i x_i^T \right) z$$

Then, using (A1),

$$\begin{aligned}
&F(ts_1 + (1-t)s_2) = \\
&z^T \left(t \sum_{i=1}^{n} s_{1,i} x_i x_i^T + (1-t) \sum_{i=1}^{n} s_{2,i} x_i x_i^T \right) z \le \\
&tz^T \left(\sum_{i=1}^{n} s_{1,i} x_i x_i^T)^{-1} z + (1-t) z^T (\sum_{i=1}^{n} s_{2,i} x_i x_i^T \right)^{-1} z \le \\
&tF(s_1) + (1-t)F(s_2)
\end{aligned}$$

□

Thus, since $F(s)$ is convex, there are a number of convex solvers that minimize F over C in polynomial time. For completeness, we also show that entropic mirror descent method from (Allen-Zhu et al., 2017) can be used to optimize F: Assumptions B1 and B3 are trivially satisfied, while for B2, F_λ is Lipschitz because it is convex.

Then, notice that equation (4) can be restated as follows: let $\{x_1, x_2, \ldots, x_{N\times n}\}$ be a set of arms such that each of the n arms in $\mathcal{A}$ is replicated exactly N times. Then, minimizing (4) is equivalent to

$$\underset{s\in S}{\operatorname{argmin}}\, F(s) = \underset{s \in S}{\operatorname{argmin}}\, f(\sum_{i=1}^{Nn} s_i x_i x_i^T) \text{ where}$$

$$S = \{s \in \{0, 1, \ldots, N\}^{Nn} : \sum_{i=1}^{Nn} s_i \le N\}$$

Thus, using Theorem 2.1 from (Allen-Zhu et al., 2017) (stated below, using our notation) and assuming $T \ge 45d \log_2 n$, we get that

Theorem 17. *Suppose $\varepsilon \in (0, 1/3]$, $Nn \ge N \ge 5d/\varepsilon^2$, $b \in \{1, 2, ..., N\}$, and $f : S_d^+ \to \mathbb{R}$ satisfies assumptions (A1) and (A2). Let $\pi \in C$ by any fractional solution so that $F(\pi) < \infty$. Then, in time complexity $\tilde{O}(Nnd^2)$ we can round π to an integral solution*

$$\hat{s} \in S \text{ satisfying } F(\hat{s}) \le (1 + 6\varepsilon) F(\pi)$$

where $C = \{c \in [0, N]^{Nn} : \sum_{i=1}^{Nn} c_i \le N\}$.

For completion, we also note that our algorithm is almost optimal, which follows from Theorem 1.4 from (Allen-Zhu et al., 2017):

Theorem 18. *Suppose $\varepsilon \in (0, 1/3]$, $Nn \ge N \ge 5d/\varepsilon^2$, $f : \mathbb{S}_d^+ \to R$ satisfies assumptions (A1)-(A3), and $\min_{s\in S} F(s) < +\infty$. Then, there exist a polynomial-time algorithm that outputs $\hat{s} \in S$ satisfying*

$$F(\hat{s}) \le (1 + 8\varepsilon) \min_{s\in S} F(s)$$

Thus, under assumption that $T > 45d \log_2 n$, using $N = \frac{T}{\log_2 n} \ge 45d$, we get that there exists an $\varepsilon \in (0, 1/3]$ s.t. $N\varepsilon^2 \ge 5d$ and so the hypotheses of Theorem 18 are satisfied. In particular, we have the following result:

Theorem 19. *If $T > 45d \log_2(n)$, then in `GetArms`, the output $\hat{\mathcal{X}}$ provided by the algorithm of (Allen-Zhu et al., 2017) satisfies:*

$$f_S(\hat{\mathcal{X}}) \le \frac{11}{3} \inf_{|\mathcal{X}|=N} f_S(\mathcal{X})$$

D. Proof of Theorem 6

Theorem 6. *Given a d-dimensional linear bandit pure-exploration algorithm, any $p \in (0, 1/2)$, and any $n \geq d$, there exists an problem instance on which the probability of identifying the best arm is at most $1 - \exp\left(\frac{-(15+o(1))T}{p(1-p)\tilde{H}_2}\right)$, where the $o(1)$ depends on n, T and p and goes to zero as $T \to \infty$.*

Proof. Suppose otherwise. Then given any d-armed multi-armed bandit problem with true rewards $y_1, \ldots, y_d$ contained in $[p, 1-p]$, we construct an $d+1$-dimensional linear bandit problem as follows: choose an arbitrary orthonormal set $z_1, \ldots, z_d \subset \mathbb{R}^{d+1}$. Set $\theta = \sum_{i=1}^{d} y_i z_i$. For $i > d$, set $x_i = z_i$ where now r_i is an arbitrary value in $[0, p]$, and z_i is an arbitary vector such that $\langle z - i, z_j \rangle = 0$ for all $i \leq d$. When our algorithm queries arm x_i for $i \leq d$, we provide an observed Bernoulli reward with mean $\langle x_i, \theta \rangle = y_i$, and for $i > d$, we provide $\langle x_i, \theta \rangle < p$. Then finding the best arm in this setting is equivelent to finding the best arm in the original multi-armed bandit problem. The value H_2 in the multi-armed bandit instance is $\max_{i \leq d} i\Delta_i^{-2} \geq \frac{1}{3}\tilde{H}_2$, and so the Theorem follows from (Audibert & Bubeck, 2010a) Theorem 4. ☐

E. Proof of Theorem 7

Theorem 7. *Given a linear bandit pure-exploration algorithm, there exists a problem instance on which the probability of identifying the best arm is at most $1 - \exp\left(-T \cdot (1/\sqrt{H_{LB}} + 2\sin\frac{\pi}{n})^2\right)$.*

We will use the following construction. Consider a problem instance with set A of n arms $x_1, x_2, \ldots, x_n$ that are equispaced on the unit circle $S^1 \subset \mathbb{R}^2$ (the construction can be extended to higher dimensions). The angle between every consecutive pair of arms is therefore $\frac{2\pi}{n}$. Let θ have unit norm and be aligned along x_1. Consider the set of parameter vectors M that are exactly aligned along one of the arms, i.e. $M := \{x_1, x_2, \cdots, x_n\}$. By rotational invariance, for any parameter vector $\lambda \in M$, the problem instances will have exactly the same value of H_{LB}. Now, we can use the same construction as in the proof of Theorem 1 from (Degenne et al., 2020), but with the added assumption that the strategy ψ has knowledge of the H_{LB} value for the two probvlem instances, to show the following:

Lemma 20. *For any linear-bandit fixed budget algorithm running for time T and achieving error probability δ, for all $\theta \in M$, we have*

$$\frac{T}{\log_2(1/\delta)} \geq \frac{1}{\max_{w \in D_d} \min_{\lambda \in M, \lambda \in \neg\theta} \|\theta - \lambda\|_{V_w}^2},$$

where $\neg\theta$ refers to the set $\{\lambda : \max_{a \in A} \lambda^T a > \max_{a \in A} \theta^T a\}$

Proof. Suppose there exists a linear bandit fixed-budget algorithm for which the above bound is not true. We will use such an algorithm. along with the value of H_{LB} from the construction to create a fixed-confidence strategy ψ to distinguish between two problem instances with parameter vector $\theta, \lambda \in M$, as described in the proof of Theorem 1 of (Degenne et al., 2020), from which the lemma follows. ☐

We can further show the following lemma, that relates $\min_{\lambda \in M, \lambda \in \neg\theta} \|\theta - \lambda\|^2$

Lemma 21. *For a fixed w and θ, we have*

$$\min_{\lambda \in M, \lambda \in \neg\theta} \|\theta - \lambda\|_{V_w} < \min_{\lambda \in S^1, \lambda \in \neg\theta} \|\theta - \lambda\|_{V_w} + 2\sin\frac{\pi}{n}$$

Proof. Let λ^* be the minimizer of $\min_{\lambda \in S^1, \lambda \in \neg\theta} \|\theta - \lambda\|_{V_w}$. By definition of M, there exists a $\lambda' \in M$ such that $\|\lambda' - \lambda^*\|_{V_w} \leq \|\lambda' - \lambda^*\|_2 \|V_w\|_2 = 2\sin\frac{\pi}{n}\|V_w\|_2 \leq 2\sin\frac{\pi}{n}$. The last inequality is true since $\|V_w\|_2 = \|\sum_{i=1}^{d} w_i x_i x_i^T\|_2 \leq \sum_{i=1}^{d} w_i \|x_i x_i^T\| = 1$, since the w_i lie on a simplex and x_i lie on the unit circle. ☐

Using the above two lemmas, we have

$$\begin{aligned}\frac{T}{\log_2(1/\delta)} &\geq \frac{1}{\max_{w\in D_d}\min_{\lambda\in M,\lambda\in\neg\theta}\|\theta-\lambda\|^2_{V_w}}\\ &= \frac{1}{\max_{w\in D_d}(\min_{\lambda\in M,\lambda\in\neg\theta}\|\theta-\lambda\|_{V_w})^2}\\ &\geq \frac{1}{\max_{w\in D_d}(\min_{\lambda\in S^1,\lambda\in\neg\theta}\|\theta-\lambda\|_{V_w}+2\sin\frac{\pi}{n})^2}\\ &\geq \frac{1}{(\frac{1}{\sqrt{H_{LB}}}+2\sin\frac{\pi}{n})^2}\end{aligned}$$

where the last inequality follows from Lemma 8 of (Degenne et al., 2020). The theorem follows.

F. Proofs for Section 6

Lemma 8. *The probability that $y_{1,m+1} < y_{1,m} - \epsilon$ is at most:*

$$3\exp\left(-\frac{\max\{(\epsilon-(2+\sqrt{2h_{|S_m|/4}})\gamma_{\max})^2,0\}T}{\log_2(n)h_{|S_m|/4}}\right)$$

Proof. Define $\Delta_{i,m} \leq \Delta_i$ to be the gap between the ith best arm and the best arm remaining in S_m. Then notice that the result of Lemma 3 still holds if we replace x_1 with the best arm remaining in S_m and Δ_i with $\Delta_{i,m}$

Let S^ϵ_m be the set of arms in S_m with y value less than $y_{1,m} - \epsilon$. Notice that in order for $y_{1,m+1}$ to be less than $y_{1,m} - \epsilon$, we must have $|S_m|/2$ elements of $|S^\epsilon_m|$ to have $\hat{y}$ values larger than those of the best arm left in S_m. Let $S^{\epsilon'}_m$ be the set of arms that excludes the $\frac{1}{4}|S_m|$ arms with highest true mean from S^ϵ_m. Then if $y_{1,m+1} < y_{1,m}$, we must have $\frac{1}{3}$ of the arms in $S^{\epsilon'}_m$ have higher $\hat{y}$ estimates than the best arm in S_m. Let N_m be the number of such arms. Define $D = \max\{(\epsilon-(2+\sqrt{h_{|S_m|/4}})\gamma_{\max})^2, 0\}$.

Then, since all arms in S^ϵ_m have $\Delta_{i,m} \geq \epsilon$, we have by Lemma 3 that

$$\mathbb{E}[N_m] \leq |S^{\epsilon'}_m|\exp\left(-\frac{DT}{\log_2(n)h_{|S_m|/4}}\right)$$

So using Markov inequality in exactly the same way as in the proof of Lemma 3, the conclusion follows. □