A USEFUL LEMMAS

We introduce the additional notation of

\[ \xi_i = \sum_{s=1}^{t} e_{\pi_s}(x_s - \mu_{\pi_s}), \quad (24) \]
\[ \sigma_i^t = \sqrt{(V_t^{-1})_{ii}}, \quad (25) \]
\[ N_t = \text{diag}(n_t), \quad (26) \]

to be used in the proofs of our results. The following lemmas are proved in Section 2.

**Lemma A.1.** With probability at least 1 − 1/\(n\), for any \(i \in [N]\) and \(t \geq 1\),

\[ |\hat{\mu}_i^t - \mu_i| \leq \sigma_i^t \left( \frac{R}{\gamma} \sqrt{\log \left( \frac{|V_t|}{\sigma^2|L_\lambda|} \right)} + \|\mu\|_{L_\lambda} \right). \quad (27) \]

**Lemma A.2.** For all \(i \in [N]\) and \(t \geq 0\),

\[ \sigma_i^t \leq \sqrt{\frac{(\sigma_0^t)^2}{1 + (\sigma_0^t)^2 n_t^i / \gamma}}. \quad (28) \]

**Lemma A.3.** Let \(d_T\) be the effective dimension. Then

\[ \log \frac{|V_T|}{|L_\lambda|} \leq 2d_T \log \left( 1 + \frac{T}{\gamma \lambda} \right). \quad (29) \]

B PROOF OF PROPOSITION 2.2

For Algorithm 1 to succeed, it must be that \(\hat{\mu}_i \geq \tau\) for each \(i\) such that \(\mu_i \geq \tau + \varepsilon\) and \(\hat{\mu}_i < \tau\) for each \(i\) such that \(\mu_i < \tau - \varepsilon\) (we can make this inequality strict or non-strict without changing probabilistic statements since \(\hat{\mu}\) is a continuous random variable). For a given \(i\), this is satisfied if \(|\hat{\mu}_i - \mu_i| \leq |\mu_i - \tau|\). We show this for the case that \(\mu_i \geq \tau + \varepsilon\). If \(\hat{\mu}_i \geq \mu_i\) in this case, then the necessary condition is satisfied. If \(\hat{\mu}_i < \mu_i\), then

\[ (30) \quad \mu_i - \tau = |\mu_i - \tau| \geq |\hat{\mu}_i - \mu_i| = \mu_i - \hat{\mu}_i \]
\[ \Rightarrow \tau \leq \hat{\mu}_i. \quad (31) \]

The case where \(\mu_i \leq \tau - \varepsilon\) is analogous. Thus, a sufficient condition for the success of Algorithm 1 is that \(|\hat{\mu}_i - \mu_i| \leq |\mu_i - \tau|\) for all \(i\) such that \(|\mu_i - \tau| \geq \varepsilon\). If we use Lemmas A.1, A.2, and A.3, we know that with probability at least 1 − \(\delta\),

\[ |\hat{\mu}_i^t - \mu_i| \leq \sigma_i^t \left( \frac{R}{\gamma} \sqrt{\log \left( \frac{|V_t|}{\sigma^2|L_\lambda|} \right)} + \|\mu\|_{L_\lambda} \right) \]
\[ \leq \sqrt{\frac{(\sigma_0^t)^2}{1 + (\sigma_0^t)^2 n_t^i / \gamma}} \left( \frac{R}{\gamma} \sqrt{2d_T \log \left( 1 + \frac{T}{\gamma \lambda} \right)} - 2 \log \delta + \|\mu\|_{L_\lambda} \right) \]
\[ \leq \sqrt{\frac{\gamma}{n_t^i}} \left( \frac{R}{\gamma} \sqrt{2d_T \log \left( 1 + \frac{T}{\gamma \lambda} \right)} - 2 \log \delta + \|\mu\|_{L_\lambda} \right). \quad (32) \]

Thus Algorithm 1 succeeds with probability at least 1 − \(\delta\) if, for all \(i\) such that \(|\mu_i - \tau| \geq \varepsilon\),

\[ \sqrt{\frac{\gamma}{n_t^i}} \left( \frac{R}{\gamma} \sqrt{2d_T \log \left( 1 + \frac{T}{\gamma \lambda} \right)} - 2 \log \delta + \|\mu\|_{L_\lambda} \right) \leq |\mu_i - \tau|. \quad (35) \]
Because Algorithm 1 has an equal sampling allocation for each arm, for $T = kN$ we have that $n_i^T = k = T/N$. Then since for each $i$ the left-hand side of (35) is the same, we can write the complete sufficient condition as

$$\sqrt{\frac{\gamma N}{T}} \left( \frac{R}{\gamma} \sqrt{2d_T \log \left( 1 + \frac{T}{\gamma \lambda} \right)} - 2 \log \delta + \|\mu\|_{L_\lambda} \right) \leq \min \{|\mu_i - \tau| : |\mu_i - \tau| \geq \varepsilon\}. \quad (36)$$

The smallest $\delta$ for which this inequality holds is

$$\delta = \exp \left\{ - \frac{\gamma^2}{2R^2} \left( \frac{T}{\gamma H} - \|\mu\|_{L_\lambda} \right)^2 + d_T \log \left( 1 + \frac{T}{\gamma \lambda} \right) \right\}, \quad (37)$$

provided $\|\mu\|_{L_\lambda} \leq \sqrt{\frac{T}{\gamma H}}$, where $\tilde{H} \triangleq N/ \min \{|\mu_i - \tau|^2 : |\mu_i - \tau| \geq \varepsilon\}$.

C PROOF OF THEOREM 3.1

The proof follows the same general strategy as that of Theorem 2 of Locatelli et al. (2016).

C.1 A Favorable Event

Let

$$\delta = \exp \left\{ - \frac{\gamma^2}{2R^2} \left( \frac{1}{3M + 1} \sqrt{\frac{T}{\gamma H}} - \|\mu\|_{L_\lambda} \right)^2 + d_T \log \left( 1 + \frac{T}{\gamma \lambda} \right) \right\}, \quad (38)$$

and consider for the rest of the proof an event of probability at least $1 - \delta$ that gives us the result of Lemma A.1. On this event then, for all $i \in [N]$,

$$|\hat{\mu}_i^t - \mu_i| \leq \sigma_i^t \left( \frac{R}{\gamma} \sqrt{2d_T \log \left( 1 + \frac{T}{\gamma \lambda} \right)} + \|\mu\|_{L_\lambda} \right) \leq \sigma_i^t \left( \frac{R}{\gamma} \sqrt{2d_T \log \left( 1 + \frac{T}{\gamma \lambda} \right)} - 2 \log \delta + \|\mu\|_{L_\lambda} \right) \leq \frac{\sigma_i^t}{3M + 1} \sqrt{\frac{T}{\gamma H}}, \quad (39)$$

where the second inequality comes from Lemma A.3 and the third inequality comes from plugging in $\delta$ using the fact that $\|\mu\|_{L_\lambda} \leq \frac{1}{3M+1} \sqrt{\frac{T}{\gamma H}}$.

C.2 A Helpful Arm

At time $T$, there must exist an arm $k$ such that $n_k^T \geq \frac{T}{H \Delta_k}$. If this were not true, then

$$T = \sum_{i=1}^N n_i^T < \sum_{i=1}^N \frac{T}{H \Delta_i^2} = T, \quad (40)$$

which is a contradiction. Let $t \leq T$ be the last time this arm was pulled, and consider this time for the rest of the proof. Note that $n_k^t = n_k^T \geq \frac{T}{H \Delta_k}$. 
C.3 Bounding the Other Arms using the Helpful Arm

When \( n_i^t \geq 1 \), using Lemma A.2,

\[
\sigma_i^t \sqrt{n_i^t + \alpha} \leq \sqrt{\frac{\sigma_i^0(n_i^t + \alpha)}{1 + \sigma_i^0 \gamma n_i^t / \gamma}} \\
\leq \sqrt{\frac{\gamma(n_i^t + \alpha)}{n_i^t}} \\
\leq \sqrt{\gamma(1 + \alpha)}.
\]  

(41)

So, including the case of \( n_i^t = 0 \),

\[
\sigma_i^t \sqrt{n_i^t + \alpha} \leq \max \left\{ \sigma_i^0 \sqrt{\alpha}, \sqrt{\gamma(1 + \alpha)} \right\} \leq \sqrt{\gamma} M,
\]  

(42)

where the last inequality comes from the fact that \( \sigma_i^0 \leq 1/\sqrt{\lambda} \).

We know that

\[
|\hat{\mu}_i^t - \mu_i| \geq ||\hat{\mu}_i^t - \tau| - |\mu_i - \tau|| = |\hat{\Delta}_i^t - \Delta_i|,
\]  

(43)

so we can find a lower bound:

\[
z_k^t = \hat{\Delta}_k^t \sqrt{n_k^t + \alpha} \\
\geq \left( \Delta_k - \frac{\sigma_k^t}{3M+1} \sqrt{\frac{T}{\gamma H}} \right) \sqrt{n_k^t} \\
\geq \sqrt{\frac{T}{H}} \frac{3M}{3M+1}. 
\]  

(44)

where the last inequality comes from our bound on \( n_k^t \) and from (41) with \( \alpha = 0 \). For the upper bound,

\[
z_i^t = \hat{\Delta}_i^t \sqrt{n_i^t + \alpha} \\
\leq \left( \Delta_i + \frac{\sigma_i^t}{3M+1} \sqrt{\frac{T}{\gamma H}} \right) \sqrt{n_i^t + \alpha} \\
\leq \Delta_i \sqrt{n_i^t + \alpha} + \frac{M}{3M+1} \sqrt{\frac{T}{H}}.
\]  

(45)

Since we pulled arm \( k \) on round \( t \), \( z_k^t \leq z_i^t \), so

\[
\sqrt{\frac{T}{H}} \frac{3M}{3M+1} \leq \Delta_i \sqrt{n_i^t + \alpha} + \frac{M}{3M+1} \sqrt{\frac{T}{H}},
\]  

(46)

\[
\Rightarrow \frac{1}{3M+1} \sqrt{\frac{T}{H}} \leq \frac{\Delta_i \sqrt{n_i^t + \alpha}}{2M}. 
\]  

(47)

C.4 Wrapping Up

Finally, we have that

\[
|\hat{\mu}_i^T - \mu_i| \leq \frac{\sigma_i^T}{3M+1} \sqrt{\frac{T}{\gamma H}} \leq \frac{\Delta_i \sigma_i^t \sqrt{n_i^t + \alpha}}{2 \sqrt{\gamma} M} \leq \frac{\Delta_i}{2},
\]  

(48)

where the second inequality comes from the fact that \( \sigma_i^0 \) is decreasing in \( t \) and from (47). Now for \( i \) such that \( \mu_i \geq \tau + \varepsilon \), we have

\[
\hat{\mu}_i^T \geq \mu_i - \frac{\Delta_i}{2} = \mu_i - \frac{\mu_i - \tau + \varepsilon}{2} = \frac{\tau + \mu_i - \varepsilon}{2} \geq \tau.
\]  

(49)
For $i$ such that $\mu_i \leq \tau - \varepsilon$, we have
\[
\hat{\mu}_i \leq \mu_i + \frac{\Delta_i}{2} = \mu_i + \frac{\tau - \mu_i + \varepsilon}{2} = \frac{\tau + \mu_i + \varepsilon}{2} \leq \tau.
\] (50)

D PROOF OF PROPOSITION 3.6

The proof of this proposition is the same as the proof of proposition 2.2 until the choice of the sampling allocation $n_t^i = \beta_i t$. Continuing from (35), we must choose $\beta$ such that, for all $i$ such that $|\mu_i - \tau| \geq \varepsilon$,
\[
\sqrt{\frac{\gamma}{T}} \left( \frac{R}{\gamma} \sqrt{2dT \log \left( \frac{1}{\delta} \gamma \lambda \right)} - 2 \log \delta + \|\mu\|_{L_\lambda} \right) \leq \sqrt{\beta_i} |\mu_i - \tau|.
\] (51)

To optimize this inequality such that it holds for the smallest possible $\delta$, we must make the right-hand side as large as possible. That is, we must choose $\beta$ that maximizes
\[
\min_{i : |\mu_i - \tau| \geq \varepsilon} \sqrt{\beta_i} |\mu_i - \tau|.
\] (52)

To maximize this minimum, we must choose $\beta$ that makes all of the terms the same. With the constraint that $\sum_i \beta_i = 1$, this means that we must choose
\[
\beta_i = \begin{cases} 
(H_* |\mu_i - \tau|^2)^{-1} & \text{if } |\mu_i - \tau| \geq \varepsilon \\
0 & \text{otherwise}
\end{cases}
\] (53)

where
\[
H_* = \sum_{j : |\mu_j - \tau| \geq \varepsilon} |\mu_j - \tau|^{-2}.
\] (54)

With this choice of $\beta$, the smallest $\delta$ for which the inequality holds is
\[
\delta = \exp \left\{ - \frac{\gamma^2}{2R^2} \left( \sqrt{\frac{T}{\gamma H_*}} - \|\mu\|_{L_\lambda} \right)^2 + dT \log \left( \frac{T}{\gamma \lambda} \right) \right\},
\] (55)

provided $\|\mu\|_{L_\lambda} \leq \sqrt{\frac{T}{\gamma H_*}}$.

E PROOF OF LEMMAS

E.1 Proof of Lemma A.1

To prove Lemma A.1, we first need the following lemma, which is a direct consequence of Theorem 1 of Abbasi-Yadkori et al. (2011):

Lemma E.1. For any $\delta > 0$, with probability at least $1 - \delta$, for all $t \geq 0$,
\[
\|\xi_t\|_{V_t^{-1}}^2 \leq R^2 \log \left( \frac{|V_t\|_{\delta^2 [L_\lambda]} \|V_t\|_{\delta^2 [L_\lambda]} \right) .
\] (56)

Using Lemma E.1, the proof of Lemma A.1 follows that of Lemma 3 of Valko et al. (2014). Let $N_t = \text{diag}(n_t)$, and note that $x_t = (N_t \mu + \xi_t)/\gamma$. Then
\[
\|\hat{\mu}_t^i - \mu_i\| = \|e_i, V_t^{-1}(N_t \mu + \xi_i)/\gamma - \mu_i\| \\
= \|e_i, V_t^{-1} \xi_i/\gamma - V_t^{-1} (V_t - N_t/\gamma) \mu_i\| \\
\leq \|e_i, \xi_i/\gamma\|_{V_t^{-1}} + \|e_i, L_\lambda \mu\|_{V_t^{-1}} \\
\leq \sigma_t \left( \|\xi_i/\gamma\|_{V_t^{-1}} + \|L_\lambda \mu\|_{V_t^{-1}} \right),
\] (57)

Daniel LeJeune, Gautam Dasarathy, Richard G. Baraniuk
where the last inequality comes from Cauchy-Schwarz and the fact that $\sigma_t^2 = \|e_i\|_{V_t^{-1}}^2$. The first term is bounded by Lemma E.1 and the second term is bounded as follows:

$$\|L_\lambda \mu\|_{V_t^{-1}}^2 = \mu^T L_\lambda V_t^{-1} L_\lambda \mu$$

$$= \mu^T \left( L_\lambda - N_t^{1/2} \left( \gamma I + N_t^{1/2} L_\lambda N_t^{1/2} \right)^{-1} N_t^{1/2} \right) \mu$$

$$\leq \mu^T L_\lambda \mu = \|\mu\|_{L_\lambda}^2,$$

where the second equality comes from the Woodbury matrix identity, and the first inequality is from the subtrahend being positive semidefinite.

### E.2 Proof of Lemma A.2

From the Sherman–Morrison formula, for $t \geq 1$,

$$(\sigma_t^i)^2 = e_i^T (V_{t-1} + e_{\pi_t}^T e_{\pi_t}^\top / \gamma)^{-1} e_i$$

$$= e_i^T \left( V_{t-1} - \frac{V_{t-1} e_{\pi_t} e_{\pi_t}^T V_{t-1}}{\gamma + e_{\pi_t} V_{t-1} e_{\pi_t}} \right) e_i$$

$$= (\sigma_{t-1}^i)^2 - \frac{(e_i^T V_{t-1} e_{\pi_t})^2}{\gamma + (\sigma_{t-1}^i)^2},$$

so $\sigma_t^i$ is decreasing in $t$. When $\pi_t = i$, the update depends only on the previous value $\sigma_{t-1}^i$. Consider the setting where $\pi_t = i \forall t \geq 1$. Then $(\sigma_t^i)^2 = \gamma (\sigma_0^i)^2 / (\gamma + t(\sigma_0^i)^2)$, which can be shown by induction. It clearly holds for $t = 0$. For $t \geq 1$,

$$(\sigma_t^i)^2 = (\sigma_{t-1}^i)^2 \left( 1 - \frac{(\sigma_{t-1}^i)^2}{\gamma + (\sigma_{t-1}^i)^2} \right)$$

$$= \gamma (\sigma_{t-1}^i)^2 \frac{\gamma + (\sigma_{t-1}^i)^2}{\gamma + (\sigma_{t-1}^i)^2}$$

$$= \frac{(\sigma_0^i)^2}{(\gamma + (t-1)(\sigma_0^i)^2)} \left( \gamma + \frac{\gamma (\sigma_0^i)^2}{\gamma + (t-1)(\sigma_0^i)^2} \right)$$

$$= \frac{\gamma (\sigma_0^i)^2}{\gamma + t(\sigma_0^i)^2}.$$  \hspace{1cm} (60)

In the setting where we do not have $\pi_t = i$ for all $t \geq 1$, since $\sigma_t^i$ is decreasing even when $\pi_t \neq i$, we can upper bound $\sigma_t^i$ with what its value would be if at each time $t$ such that $\pi_t \neq i$ we do not update $\sigma_t^i$. This would mean that by time $t$, $\sigma_t^i$ has been updated $n_t^i$ times, yielding the stated bound.

### E.3 Proof of Lemma A.3

This lemma is derived from Lemma 6 of Valko et al. (2014). If $QAQ^\top$ is the eigendecomposition of $L_\lambda$, then let $V_T$ and $A$ in the notation of Valko et al. (2014) be equal to $\gamma Q^\top V_T Q$ and $\gamma A$, respectively, in our notation. The result follows by the invariance of determinants under unitary transformations.