Contents

A Action Elimination with Gaussian Width

We first state an algorithm inspired by Lattimore and Szepesvári [2020] and prove a regret bound. This algorithm, while naive, incorporates the TIS inequality to obtain regret scaling with the Gaussian width. Furthermore, the analysis is simple and helps aid in the intuition of the proof of our main theorems.

For $\mathfrak{f} \in \{\mathrm{band}, \mathrm{semi}\}$, denote:

$$\gamma(A_{\mathfrak{f}}(\lambda), \mathcal{X}) := \mathbb{E}_{\eta \sim \mathcal{N}(0,I)} \left[\sup_{x \in \mathcal{X}} x^\top A_{\mathfrak{f}}(\lambda)^{-1/2} \eta \right]^2$$

Algorithm 2 Gaussian Width Action Elimination (GW-AE)

Input: Set of arms $\mathcal{X}$, confidence δ, largest gap $\Delta_{\max}$, rounding parameter $\zeta \in (0,1)$
$\hat{\mathcal{X}}_1 \leftarrow \mathcal{X}, \ell \leftarrow 1,$
while $|\hat{\mathcal{X}}_\ell| > 1$ **do**
  Let $\hat{\lambda}_\ell$ a minimizer of $\mathbb{E}_\eta[\max_{x \in \hat{\mathcal{X}}_\ell} x^\top A(\lambda)^{-1/2}\eta]^2 + \max_{x \in \hat{\mathcal{X}}_\ell} \|x\|^2_{A(\lambda)^{-1}}$
  $\epsilon_\ell = \Delta_{\max} 2^{-\ell}$, $\tau_\ell = 2(1+\zeta)\epsilon_\ell^{-2}(\gamma(A(\hat{\lambda}_\ell), \hat{\mathcal{X}}_\ell) + 2\sup_{x \in \hat{\mathcal{X}}_\ell} \|x\|^2_{A(\hat{\lambda}_\ell)^{-1}} \log(2\ell^2/\delta))$
  $\kappa_\ell \leftarrow \mathrm{ROUND}(\hat{\lambda}_\ell, \lceil \tau_\ell \rceil \vee q(\zeta), \zeta)$
  Pull arm x $\kappa_{\ell,x}$ times, compute $\hat{\theta}_\ell$ from this data
  $\hat{\mathcal{X}}_{\ell+1} \leftarrow \hat{\mathcal{X}}_\ell \backslash \{x \in \hat{\mathcal{X}}_\ell \ : \ \max_{x' \in \hat{\mathcal{X}}_\ell} (x' - x)^\top \hat{\theta}_\ell > 2\epsilon_\ell\}$
  $\ell \leftarrow \ell + 1$
end while

Here $\mathrm{ROUND}(\lambda, N, \zeta)$ is a rounding procedure which takes as input $\lambda \in \triangle_{\mathcal{X}}$, $N \in \mathbb{N}$, and $\zeta \in (0,1)$ and outputs an allocation $\kappa \in \mathbb{N}^{|\mathcal{X}|}$ such that:

$$\gamma(A(\kappa), \mathcal{X}) + \sup_{x \in \mathcal{X}} \|x\|^2_{A(\kappa)^{-1}} \leq (1+\zeta)\left(\gamma(A(\tau\lambda), \mathcal{X}) + \sup_{x \in \mathcal{X}} \|x\|^2_{A(\tau\lambda)^{-1}} \right)$$

and $\sum_{x \in \mathcal{X}} \kappa_x = N$, so long as $N \geq q(\zeta)$. From Katz-Samuels et al. [2020] and Allen-Zhu et al. [2020], we know such a rounding procedure exists and can be computed efficiently, and that it suffices to choose $q(\zeta) = O(d/\zeta^2)$.

Denote:

$$\bar{\gamma}_{\mathrm{ae}}(A_{\mathfrak{f}}) = \sup_{\epsilon > 0} \sup_{\mathcal{Y} \subseteq \mathcal{X}_\epsilon} \inf_{\lambda \in \triangle_{\mathcal{Y} \cup x_*}} \mathbb{E}_\eta \left[\sup_{x \in \mathcal{Y} \cup x_*} x^\top A_{\mathfrak{f}}(\lambda)^{-1/2} \eta \right]^2$$

where $\mathcal{X}_\epsilon := \{x \in \mathcal{X} \ : \ \Delta_x \leq \epsilon\}$.

Theorem 5. *For $\mathfrak{f} \in \{\mathrm{band}, \mathrm{semi}\}$, the absolute regret of GW-AE is bounded as:*

$$c_1 \Delta_{\max} \log(\Delta_{\max}/\Delta_{\min}) d + \frac{c_2(\bar{\gamma}_{\mathrm{ae}}(A_{\mathfrak{f}}) + d\log(\log(\Delta_{\max}/\Delta_{\min})/\delta))}{\Delta_{\min}}$$

with probability at least $1-\delta$ and minimax regret as:

$$c_1 \Delta_{\max} \log(\Delta_{\max}/\Delta_{\min}) d + c_2 \sqrt{(\bar{\gamma}_{\mathrm{ae}}(A_{\mathfrak{f}}) + d\log(\log(\Delta_{\max}/\Delta_{\min})/\delta))T}$$

with probability at least $1-\delta$. Here c_1, c_2 are absolute constants.

If desired, noting that $\tau_\ell \geq \epsilon_\ell^{-2}$ which implies that we will have at most $\mathcal{O}(\log(T))$ rounds, the $\log(\Delta_{\max}/\Delta_{\min})$ could be replaced with a term $\mathcal{O}(\log(T))$, as in Theorem 1.

While this result closely resembles Theorem 1, there are several major shortcomings. First, this algorithm does not plan as effectively as it only pulling arms with gap less than ϵ_ℓ, which could cause it to forego pulling informative yet suboptimal arms, something Algorithm 1 improves on. In particular, the regret bound stated for Algorithm 1 in Proposition 1 will not hold for this algorithm. In a sense, this algorithm can be thought of as being optimistic. Second, it is always the case that $\bar{\gamma}(A_{\mathfrak{f}}) \leq \bar{\gamma}_{\mathrm{ae}}(A_{\mathfrak{f}})$. The parameter $\bar{\gamma}_{\mathrm{ae}}(A_{\mathfrak{f}})$ could be tightened by altering the constant factors in Algorithm 2 so as to guarantee that, on the good event *all* arms with gap less than ϵ, for some ϵ, are in $\hat{\mathcal{X}}_\ell$. However, even with this tightening we will always have $\bar{\gamma}(A_{\mathfrak{f}}) \leq \bar{\gamma}_{\text{ae}}(A_{\mathfrak{f}})$. Finally, Algorithm 2 does not seem to admit a computationally feasible solution in the combinatorial bandit setting.

Proof. From Proposition 6, we will have that:

$$\begin{aligned} x^\top(\hat{\theta}_\ell - \theta_*) &\leq \mathbb{E}_{\eta\sim\mathcal{N}(0,I)}\left[\sup_{x\in\hat{\mathcal{X}}_\ell} x^\top A(\kappa_\ell)^{-1/2}\eta\right] + \sqrt{2\sup_{x\in\hat{\mathcal{X}}_\ell}\|x\|^2_{A(\kappa_\ell)^{-1}}\log(2\ell^2/\delta)} \\ &\leq \epsilon_\ell \end{aligned}$$

for all $x \in \hat{\mathcal{X}}_\ell$ simultaneously with probability $1-\delta/\ell^2$. The second inequality holds by our choice of τ_ℓ and Kiefer-Wolfowitz and Proposition 9. Let:

$$\mathcal{E}_{x,\ell}(\mathcal{V}) = \{|\langle x, \hat{\theta}_\ell - \theta^*\rangle| \leq \epsilon_\ell\}$$

where $\hat{\theta}_\ell$ is computed assuming $\mathcal{V}$ is the active set in the above algorithm. Then using the following calculation from Jamieson:

$$\begin{aligned} \mathbb{P}\left[\bigcup_{\ell=1}^\infty \bigcup_{x\in\hat{\mathcal{X}}_\ell} \mathcal{E}_{x,\ell}(\hat{\mathcal{X}}_\ell)^c\right] &\leq \sum_{\ell=1}^\infty \sum_{\mathcal{V}\subseteq\mathcal{X}} \mathbb{P}\left[\bigcup_{x\in\mathcal{V}} \mathcal{E}_{x,\ell}(\mathcal{V})^c\right]\mathbb{P}[\hat{\mathcal{X}}_\ell = \mathcal{V}] \\ &\leq \sum_{\ell=1}^\infty \sum_{\mathcal{V}\subseteq\mathcal{X}} \frac{\delta}{\ell^2}\mathbb{P}[\hat{\mathcal{X}}_\ell = \mathcal{V}] \\ &\leq \delta \end{aligned}$$

so the good event, that all the arm rewards are well estimated for all rounds, holds with high probability. Assume henceforth that the good event $\mathcal{E} = \cap_{\ell=1}^\infty \cap_{x\in\mathcal{X}} \mathcal{E}_{x,\ell}(\mathcal{V})$ holds. Following identically the argument from Jamieson, we will have that $x_* \in \hat{\mathcal{X}}_\ell$ and $\max_{x\in\hat{\mathcal{X}}_\ell}(x_*-x)^\top\theta_* \leq 8\epsilon_\ell$ for all ℓ. We assume the good event holds for the remainder of the proof.

We can now follow the same argument as Lemma 12 of Katz-Samuels et al. [2020]. Take $\mathcal{Y} \subseteq \mathcal{X}_\epsilon$ for some ϵ and let $\lambda_1 \in \triangle_{\mathcal{Y}}$ be the distribution that minimizes:

$$\max_{x\in\mathcal{Y}} \|x\|^2_{A(\lambda)^{-1}}$$

and $\lambda_2 \in \triangle_{\mathcal{Y}}$ the distribution that minimizes:

$$\mathbb{E}_\eta[\max_{x\in\mathcal{Y}} x^\top A(\lambda)^{-1/2}\eta]^2$$

Let $\lambda = \frac{1}{2}(\lambda_1+\lambda_2)$. Then we will have that:

$$2A(\lambda_i)^{-1} \succeq A(\lambda)^{-1}$$

From this it immediately follows that:

$$\max_{x\in\mathcal{Y}} \|x\|^2_{A(\lambda)^{-1}} \leq 2\max_{x\in\mathcal{Y}} \|x\|^2_{A(\lambda_1)^{-1}} \leq 2d$$

where the last inequality holds by Kiefer-Wolfowitz and Proposition 9. Also:

$$\mathbb{E}_\eta[\max_{x\in\mathcal{Y}} x^\top A(\lambda)^{-1/2}\eta]^2 \leq 2\mathbb{E}_\eta[\max_{x\in\mathcal{Y}} x^\top A(\lambda_2)^{-1/2}\eta]^2 \leq 2\bar{\gamma}_{\text{ae}}(A)$$

Since $\hat{\mathcal{X}}_\ell$ will always contain only arms with gap less than ϵ for some ϵ, we then have that:

$$\tau_\ell \leq c(1+\zeta)\epsilon_\ell^{-2}(2d\log(2\ell^2/\delta) + \bar{\gamma}_{\text{ae}}(A))$$

Using these bounds and noting that $\lceil \log_2(8\Delta_{\max}/(\Delta_{\min} \vee \nu)) \rceil$ upper bounds the number of rounds, we can upper bound the regret as:

$$\begin{aligned}
&\sum_{x \in \mathcal{X} \backslash \{x^*\}} \Delta_x T_x \\
&\leq T\nu + \sum_{\ell=1}^{\lceil \log_2(8\Delta_{\max}/(\Delta_{\min} \vee \nu)) \rceil} 8\epsilon_\ell(\tau_\ell + q(\zeta) + 1) \\
&\leq T\nu + 8\Delta_{\max}\lceil \log_2(8\Delta_{\max}/(\Delta_{\min} \vee \nu)) \rceil (q(\zeta)+1) + \sum_{\ell=1}^{\lceil \log_2(8\Delta_{\max}/(\Delta_{\min} \vee \nu)) \rceil} c(1+\zeta)\epsilon_\ell^{-1}(\bar{\gamma}_{\mathrm{ae}}(A), \hat{\mathcal{X}}_\ell) + 2d\log(2\ell^2/\delta)) \\
&\leq T\nu + 8\Delta_{\max}\lceil \log_2(8\Delta_{\max}/(\Delta_{\min} \vee \nu)) \rceil (q(\zeta)+1) + \sum_{\ell=1}^{\lceil \log_2(8\Delta_{\max}/(\Delta_{\min} \vee \nu)) \rceil} c(1+\zeta)\epsilon_\ell^{-1}(\bar{\gamma}_{\mathrm{ae}}(A) + 2d\log(2\ell^2/\delta)) \\
&\leq T\nu + 8\Delta_{\max}\lceil \log_2(8\Delta_{\max}/(\Delta_{\min} \vee \nu)) \rceil (q(\zeta)+1) + \sum_{\ell=1}^{\lceil \log_2(8\Delta_{\max}/(\Delta_{\min} \vee \nu)) \rceil} c(1+\zeta)\frac{2^\ell}{\Delta_{\max}}(\bar{\gamma}_{\mathrm{ae}}(A) + 2d\log(2\ell^2/\delta)) \\
&\leq T\nu + 8\Delta_{\max}\lceil \log_2(8\Delta_{\max}/(\Delta_{\min} \vee \nu)) \rceil (q(\zeta)+1) + \frac{c(1+\zeta)(\bar{\gamma}_{\mathrm{ae}}(A) + 2d\log(2\log_2^2(16\Delta_{\max}/(\Delta_{\min} \vee \nu))/\delta))}{\Delta_{\min} \vee \nu}
\end{aligned}$$

Optimizing this over ν gives the final regret of:

$$8\Delta_{\max}\lceil \log_2(8\Delta_{\max}/(\Delta_{\min})) \rceil (q(\zeta)+1) + \sqrt{c(1+\zeta)(\bar{\gamma}_{\mathrm{ae}}(A) + d\log(\log(\Delta_{\max}/(\Delta_{\min}))/\delta))T}$$

and choosing $\nu = 0$ gives the absolute regret bound. □

B Regret Bound Proofs

Proof of Theorem 2. Throughout we will let $\mathcal{R}_\ell$ denote the regret incurred in round ℓ, and $\mathcal{R}_{1:\ell}$ the regret incurred from rounds 1 through ℓ. We assume $A(\tau)$ corresponds to the type of feedback received. The first part of this proof closely mirrors the proof of Theorem 5 of Katz-Samuels et al. [2020]. We will prove this result for τ_ℓ being a (ν, ζ)-optimal solution to (3), where we calll a solution to (3) (ν, ζ)-optimal if $\widehat{\mathrm{OPT}} \leq \nu \mathrm{OPT} + \zeta$, where $\widehat{\mathrm{OPT}}$ is the value of the objective attained by the approximate solution, and OPT the value attained by the optimal solution.

Good event: We will define $\mathcal{S}_\ell$ as the following:

$$\mathcal{S}_\ell := \{x \in \mathcal{X} \ : \ \Delta_x \leq \epsilon_\ell\}$$

Let $\delta_k = \delta/(2k^3)$ and define the events:

$$\mathcal{E}_{k,j} = \left\{ \sup_{z,z' \in \mathcal{S}_j} |(z-z')^\top(\hat{\theta}_k - \theta_*)| \leq (1 + \sqrt{\pi \log(1/\delta_k)})\mathbb{E}_\eta \left[\sup_{z,z' \in \mathcal{S}_j} (z-z')^\top A(\tau_k)^{-1/2}\eta \right] \right\}$$

$$\mathcal{E} = \bigcap_{k=1}^\infty \bigcap_{j=0}^k \mathcal{E}_{k,j}$$

Proposition 6 gives that with probability at least $1 - \delta/k^3$:

$$\begin{aligned}
\sup_{z,z' \in \mathcal{S}_j} |(z-z')^\top(\hat{\theta}_k - \theta_*)| &\leq \mathbb{E}_\eta \left[\sup_{z,z' \in \mathcal{S}_j} (z-z')^\top A(\tau_k)^{-1/2}\eta \right] + \sqrt{2 \max_{z,z' \in \mathcal{S}_j} \|z-z'\|^2_{A(\tau_k)^{-1}} \log(1/\delta_k)} \\
&\overset{(a)}{\leq} (1 + \sqrt{\pi \log(1/\delta_k)})\mathbb{E}_\eta \left[\sup_{z,z' \in \mathcal{S}_j} (z-z')^\top A(\tau_k)^{-1/2}\eta \right]
\end{aligned}$$

where (a) follows by Lemma 11 of Katz-Samuels et al. [2020]. It follows then that $\mathbb{P}[\mathcal{E}_{k,j}^c] \leq \delta/k^3$, which implies that:

$$\mathbb{P}[\mathcal{E}^c] \leq \sum_{k=1}^{\infty}\sum_{j=0}^{k} \mathbb{P}[\mathcal{E}_{k,j}^c] \leq \sum_{k=1}^{\infty}\sum_{j=0}^{k} \frac{\delta}{k^3} \leq 3\delta$$

Estimation error: Henceforth we assume $\mathcal{E}$ holds. We proceed by induction to show that the gaps are always well-estimated. First we prove the base case. Let $k = 1$ and consider any $x \in \mathcal{X}$. Then:

$$\begin{aligned}
|(x_* - x)^\top(\hat{\theta}_1 - \theta_*)| &\leq \sup_{z,z' \in \mathcal{X}} |(z - z')^\top(\hat{\theta}_1 - \theta_*)| \\
&\leq (1 + \sqrt{\pi \log(1/\delta_1)})\mathbb{E}_\eta \left[\sup_{z,z' \in \mathcal{X}} (z - z')^\top A(\tau_1)^{-1/2}\eta \right] \\
&\overset{(a)}{=} 2(1 + \sqrt{\pi \log(1/\delta_1)})\mathbb{E}_\eta \left[\sup_{z \in \mathcal{X}} (x_1 - z')^\top A(\tau_1)^{-1/2}\eta \right] \\
&\overset{(b)}{\leq} \epsilon_1/8
\end{aligned}$$

where (a) follows by Proposition 7.5.2 of Vershynin [2018] and (b) follows since τ_1 is a feasible solution to (3).

For the inductive step, assume that, for all $x \in \mathcal{S}_k$:

$$|(x_* - x)^\top(\hat{\theta}_k - \theta_*)| \leq \epsilon_k/8$$

and for all $x \in \mathcal{S}_k^c$:

$$|(x_* - x)^\top(\hat{\theta}_k - \theta_*)| \leq \Delta_x/8$$

Consider round $k + 1$ and take $x \in \mathcal{S}_{k+1}^c$. There then exists some $k' \leq k$ such that $x \in \mathcal{S}_{k'} \backslash \mathcal{S}_{k'+1}$. Then:

$$\begin{aligned}
\frac{|(x_* - x)^\top(\hat{\theta}_{k+1} - \theta_*)|}{\Delta_x} &\leq \sup_{z,z' \in \mathcal{S}_{k'}} \frac{|(z - z')^\top(\hat{\theta}_{k+1} - \theta_*)|}{\Delta_x} \\
&\leq (1 + \sqrt{\pi \log(1/\delta_{k+1})})\mathbb{E}_\eta \left[\sup_{z,z' \in \mathcal{S}_{k'}} \frac{(z - z')^\top A(\tau_{k+1})^{-1/2}\eta}{\Delta_x} \right] \\
&\overset{(a)}{=} 2(1 + \sqrt{\pi \log(1/\delta_{k+1})})\mathbb{E}_\eta \left[\sup_{z \in \mathcal{S}_{k'}} \frac{(x_{k+1} - z)^\top A(\tau_{k+1})^{-1/2}\eta}{\Delta_x} \right] \\
&\overset{(b)}{\leq} 4(1 + \sqrt{\pi \log(1/\delta_{k+1})})\mathbb{E}_\eta \left[\sup_{z \in \mathcal{S}_{k'}} \frac{(x_{k+1} - z)^\top A(\tau_{k+1})^{-1/2}\eta}{\epsilon_{k+1} + \Delta_x} \right] \\
&\overset{(c)}{\leq} 8(1 + \sqrt{\pi \log(1/\delta_{k+1})})\mathbb{E}_\eta \left[\sup_{z \in \mathcal{S}_{k'}} \frac{(x_{k+1} - z)^\top A(\tau_{k+1})^{-1/2}\eta}{\epsilon_{k+1} + \Delta_z} \right] \\
&\leq 8(1 + \sqrt{\pi \log(1/\delta_{k+1})})\mathbb{E}_\eta \left[\sup_{z \in \mathcal{X}} \frac{(x_{k+1} - z)^\top A(\tau_{k+1})^{-1/2}\eta}{\epsilon_{k+1} + \Delta_z} \right] \\
&\overset{(d)}{\leq} 16(1 + \sqrt{\pi \log(1/\delta_{k+1})})\mathbb{E}_\eta \left[\sup_{z \in \mathcal{X}} \frac{(x_{k+1} - z)^\top A(\tau_{k+1})^{-1/2}\eta}{\epsilon_{k+1} + \hat{\Delta}_z} \right] \\
&\overset{(e)}{\leq} 1/8
\end{aligned}$$

where (a) follows by Proposition 7.5.2 of Vershynin [2018], (b) follows since $\Delta_x \geq \epsilon_{k+1}$ by virtue of the fact that $x \in \mathcal{S}_{k+1}^c$, so $\Delta_x \geq (\epsilon_{k+1} + \Delta_x)/2$, (c) follows since $\Delta_x \in [\epsilon_{k'+1}, \epsilon_{k'}]$ and for any $z \in \mathcal{S}_{k'}$, we will have theta $\Delta_z \leq \epsilon_{k'}$, so $\epsilon_{k+1} + \Delta_x \geq \epsilon_{k+1} + \epsilon_{k'+1} \geq \epsilon_{k+1} + \Delta_z/2$, (d) holds by the inductive hypothesis and Lemma 1 of Katz-Samuels et al. [2020] and taking $\hat{\Delta}_z$ to be the estimate of Δ_z at round $k + 1$, and (e) holds since τ_{k+1} is a feasible solution to (3). We can perform a similar calculation to get the same thing for $x \in \mathcal{S}_{k+1}$, allowing us to conclude that, for all $x \in \mathcal{S}_{k+1}$:

$$|(x_* - x)^\top(\hat{\theta}_{k+1} - \theta_*)| \leq \epsilon_{k+1}/8$$

and for all $x \in \mathcal{S}_{k+1}^c$:

$$|(x_* - x)^\top(\hat{\theta}_{k+1} - \theta_*)| \leq \Delta_x/8$$

From this and Lemma 1 of Katz-Samuels et al. [2020], it follows that for all ℓ and $x \in \mathcal{S}_\ell$:

$$\Delta_x \leq \hat{\Delta}_x + |\hat{\Delta}_x - \Delta_x| \leq \hat{\Delta}_x + \epsilon_\ell/2 \leq \hat{\Delta}_x + \epsilon_\ell$$

and for $x \in \mathcal{S}_\ell^c$:

$$\Delta_x \leq \hat{\Delta}_x + |\hat{\Delta}_x - \Delta_x| \leq 2\hat{\Delta}_x \leq 2\hat{\Delta}_x + 2\epsilon_\ell$$

So the objective of (3) upper bounds the real regret. Further, on the good event, using Lemma 1 from Katz-Samuels et al. [2020], for any ℓ and $x \in \mathcal{X}$, we have:

$$\frac{1}{2}(\epsilon_\ell + \Delta_x) \leq \epsilon_\ell + \hat{\Delta}_x \leq \frac{3}{2}(\epsilon_\ell + \Delta_x) \tag{6}$$

This implies that if we remove arm x from $\hat{\mathcal{X}}_\ell$:

$$\hat{\Delta}_x > 2\epsilon_\ell \implies \hat{\Delta}_x + \epsilon_\ell > 3\epsilon_\ell \implies \frac{3}{2}(\epsilon_\ell + \Delta_x) > 3\epsilon_\ell \implies \Delta_x > \epsilon_\ell$$

So, on the good event, if $\hat{\Delta}_x > 2\epsilon_\ell$, we will have identified the best arm correctly.

Bounding the Round Regret: From the previous section, we know that on the good event all our gaps will be well-estimated. From (6), it follows that the constraint in (3) is tighter than the following constraint:

$$\mathbb{E}_\eta \left[\max_{x \in \mathcal{X}} \frac{(x_\ell - x)^\top A(\tau)^{-1/2}\eta}{\epsilon_\ell + \Delta_x} \right] \leq \frac{1}{256(1 + \sqrt{\pi \log(2\ell^3/\delta)})} \tag{7}$$

so any τ satisfying this inequality is also a feasible solution to (3).

Consider drawing some η and let x_η be the point $x \in \mathcal{X}$ that achieves the maximum above (if the solution is not unique, break ties by choose x_η randomly from the $x \in \mathcal{X}$ for which the maximum is attained). If we assume that $x_\eta \in \mathcal{S}_\ell$, then it follows that:

$$\begin{aligned}
\max_{x \in \mathcal{X}} \frac{(x_\ell - x)^\top A(\tau)^{-1/2}\eta}{\epsilon_\ell + \Delta_x} &= \max_{x \in \mathcal{S}_\ell} \frac{(x_\ell - x)^\top A(\tau)^{-1/2}\eta}{\epsilon_\ell + \Delta_x} \\
&\leq \max_{x \in \mathcal{S}_\ell} \frac{(x_\ell - x)^\top A(\tau)^{-1/2}\eta}{\epsilon_\ell} \\
&\overset{(a)}{\leq} \sum_{j=1}^{\ell} \max_{x \in \mathcal{S}_j} \frac{(x_\ell - x)^\top A(\tau)^{-1/2}\eta}{\epsilon_j}
\end{aligned}$$

where (a) follows since we will always have:

$$\max_{x \in \mathcal{S}_j} \frac{(x_\ell - x)^\top A(\tau)^{-1/2}\eta}{\epsilon_j} \geq 0$$

since $x_\ell \in \mathcal{S}_j$ for $j \leq \ell$ by Lemma 1 of Katz-Samuels et al. [2020]. Assume that $x_\eta \in \mathcal{S}_k \backslash \mathcal{S}_{k+1}$. Then:

$$\begin{aligned}
\max_{x \in \mathcal{X}} \frac{(x_\ell - x)^\top A(\tau)^{-1/2}\eta}{\epsilon_\ell + \Delta_x} &= \max_{x \in \mathcal{S}_k \backslash \mathcal{S}_{k+1}} \frac{(x_\ell - x)^\top A(\tau)^{-1/2}\eta}{\epsilon_\ell + \Delta_x} \\
&\overset{(a)}{\leq} 2 \max_{x \in \mathcal{S}_k \backslash \mathcal{S}_{k+1}} \frac{(x_\ell - x)^\top A(\tau)^{-1/2}\eta}{\epsilon_k} \\
&\leq \max_{x \in \mathcal{S}_k} \frac{(x_\ell - x)^\top A(\tau)^{-1/2}\eta}{\epsilon_k} \\
&\leq \sum_{j=1}^{\ell} \max_{x \in \mathcal{S}_j} \frac{(x_\ell - x)^\top A(\tau)^{-1/2}\eta}{\epsilon_j}
\end{aligned}$$

where (a) uses the fact that for all $x \in \mathcal{S}_k \backslash \mathcal{S}_{k+1}$, $\Delta_x \in [\epsilon_{k+1}, \epsilon_k]$, and the last inequality follows as above. We therefore have that:

$$\mathbb{E}_\eta \left[\max_{x \in \mathcal{X}} \frac{(x_\ell - x)^\top A(\tau)^{-1/2}\eta}{\epsilon_\ell + \Delta_x} \right] \leq \sum_{j=1}^{\ell} \frac{1}{\epsilon_j} \mathbb{E}_\eta \left[\max_{x \in \mathcal{S}_j} (x_\ell - x)^\top A(\tau)^{-1/2}\eta \right]$$

Let λ_j^{gw} be the solution to:

$$\lambda_j^{\text{gw}} = \arg\min_{\lambda \in \triangle_{\mathcal{S}_j}} \mathbb{E}_\eta [\max_{x \in \mathcal{S}_j} (x_\ell - x)^\top A(\lambda)^{-1/2}\eta]$$

Let $\bar{\tau} = \ell^2 \sum_{j=1}^{\ell} \tau_j^{\text{gw}}$ and $\tau_j^{\text{gw}} = 65536 \bar{\gamma}(A) \epsilon_j^{-2} (1 + \sqrt{\pi \log(2\ell^3/\delta)})^2 \lambda_j^{\text{gw}}$. Then:

$$\begin{aligned}
\mathbb{E}_\eta \left[\max_{x \in \mathcal{S}_j} (x_\ell - x)^\top A(\tau_j^{\text{gw}})^{-1/2}\eta \right] &= \frac{\mathbb{E}_\eta \left[\max_{x \in \mathcal{S}_j} (x_\ell - x)^\top A(\lambda_j^{\text{gw}})^{-1/2}\eta \right]}{\sqrt{\tau_j^{\text{gw}}}} \\
&= \frac{\epsilon_j \mathbb{E}_\eta \left[\max_{x \in \mathcal{S}_j} (x_\ell - x)^\top A(\lambda_j^{\text{gw}})^{-1/2}\eta \right]}{\sqrt{\bar{\gamma}(A)} 256 (1 + \sqrt{\pi \log(2\ell^3/\delta)})} \\
&\leq \frac{\epsilon_j}{256(1 + \sqrt{\pi \log(2\ell^3/\delta)})}
\end{aligned}$$

Given this:

$$\begin{aligned}
\sum_{j=1}^{\ell} \frac{1}{\epsilon_j} \mathbb{E}_\eta \left[\max_{x \in \mathcal{S}_j} (x_\ell - x)^\top A(\bar{\tau})^{-1/2}\eta \right] &\overset{(a)}{\leq} \frac{1}{\ell} \sum_{j=1}^{\ell} \frac{1}{\epsilon_j} \mathbb{E}_\eta \left[\max_{x \in \mathcal{S}_j} (x_\ell - x)^\top A(\tau_j^{\text{gw}})^{-1/2}\eta \right] \\
&\leq \frac{1}{\ell} \sum_{j=1}^{\ell} \frac{1}{\epsilon_j} \frac{\epsilon_j}{256(1 + \sqrt{\pi \log(2\ell^3/\delta)})} \\
&= \frac{1}{256(1 + \sqrt{\pi \log(2\ell^3/\delta)})}
\end{aligned}$$

where (a) holds by the Sudakov-Fernique inequality (Theorem 7.2.11 of Vershynin [2018]). Thus, $\bar{\tau}$ satisfies (7) and so is a feasible solution to (3). Let τ_ℓ^* be the optimal solution to (3), then:

$$\sum_{x \in \mathcal{X}} 2(\epsilon_\ell + \hat{\Delta}_x) \tau_{\ell,x}^* \leq \sum_{x \in \mathcal{X}} 2(\epsilon_\ell + \hat{\Delta}_x) \bar{\tau}_x \leq \sum_{x \in \mathcal{X}} 3(\epsilon_\ell + \Delta_x) \bar{\tau}_x = \sum_{x \in \mathcal{X}} 3\Delta_x \bar{\tau}_x + 3\epsilon_\ell \bar{\tau}$$

The first term can be bounded by the regret bounded given in Lemma 2:

$$\sum_{x \in \mathcal{X}} 3\Delta_x \bar{\tau}_x \leq c_1 \Delta_{\max} \ell d + \frac{c_2 \ell^2 \log(\ell/\delta) \bar{\gamma}(A)}{\epsilon_\ell}$$

By construction we'll have that:

$$\bar{\tau} = c \sum_{k=1}^{\ell} \epsilon_k^{-2} \bar{\gamma}(A) (1 + \sqrt{\pi \log(2\ell^3/\delta)})^2 \leq c \bar{\gamma}(A) (1 + \sqrt{\pi \log(2\ell^3/\delta)})^2 \epsilon_\ell^{-2}$$

so:

$$3\epsilon_\ell \bar{\tau} \leq \frac{c \bar{\gamma}(A) \log(2\ell^3/\delta)}{\epsilon_\ell}$$

Recalling that τ_ℓ is a (ν, ζ)-optimal solution to (3), the above implies that:

$$\sum_{x \in \mathcal{X}} 2(\epsilon_\ell + \hat{\Delta}_x) \tau_{\ell,x} \leq (1+\nu) \sum_{x \in \mathcal{X}} 2(\epsilon_\ell + \hat{\Delta}_x) \tau_{\ell,x}^* + \zeta \leq (1+\nu) \left(c_1 \Delta_{\max} \ell d + \frac{c_2 \ell^2 \log(\ell/\delta) \bar{\gamma}(A)}{\epsilon_\ell} \right) + \zeta \quad (8)$$

We in fact play α_ℓ, as this will attain the same objective value and so the same regret bound. However, α_ℓ may not be integer, so we will pull every arm $\lceil \alpha_{\ell,x} \rceil$ times. Note that the rounded solution still meets the constraint from (3). Assume we are playing the rounded solution given by Lemma 1, then rounding the solution will incur additional regret of at most $\Delta_{\max} n_{\mathfrak{f}}$. Since $\sum_{x \in \mathcal{X}} 2(\epsilon_\ell + \hat{\Delta}_x)\tau_x$ upper bounds the real regret of playing τ_x, we'll have:

$$\mathcal{R}_\ell \leq (1+\nu)\left(c_1 \Delta_{\max} \ell d + \frac{c_2 \ell^2 \log(\ell/\delta)\bar{\gamma}(A)}{\epsilon_\ell}\right) + \Delta_{\max} n_{\mathfrak{f}} + \zeta$$

We can then bound the regret incurred after ℓ stages as:

$$\begin{aligned} \mathcal{R}_{1:\ell} &\leq \sum_{k=1}^{\ell} (1+\nu)\left(c_1 \Delta_{\max} k d + \frac{c_2 k^2 \log(k/\delta)\bar{\gamma}(A)}{\epsilon_k}\right) + \ell \Delta_{\max} n_{\mathfrak{f}} + \ell\zeta \\ &\leq c_1(1+\nu)\Delta_{\max}\ell^2 d + \ell\Delta_{\max} n_{\mathfrak{f}} + \ell\zeta + \sum_{k=1}^{\ell} \frac{c_2(1+\nu)k^2\log(\ell/\delta)\bar{\gamma}(A)}{\epsilon_k} \\ &\leq c_1(1+\nu)\Delta_{\max}\ell^2 d + \ell\Delta_{\max} n_{\mathfrak{f}} + \ell\zeta + c_2(1+\nu)\log(\ell/\delta)\bar{\gamma}(A)\sum_{k=1}^{\ell} k^2 2^k \\ &\leq c_1(1+\nu)\Delta_{\max}\ell^2 d + \ell\Delta_{\max} n_{\mathfrak{f}} + \ell\zeta + \frac{c_2(1+\nu)\ell^2\log(\ell/\delta)\bar{\gamma}(A)}{\epsilon_\ell} \end{aligned} \tag{9}$$

Minimax Regret: Denote the objective to (3) at round ℓ evaluated at τ_ℓ by:

$$f_\ell := \sum_{x \in \mathcal{X}} 2(\epsilon_\ell + \hat{\Delta}_x)\tau_{\ell,x}$$

By (8) we can upper bound:

$$\begin{aligned} f_\ell &\leq (1+\nu)\left(c_1 \Delta_{\max}\ell d + \frac{c_2 \ell^2 \log(\ell/\delta)\bar{\gamma}(A)}{\epsilon_\ell}\right) + \zeta \\ &\leq c_1(1+\nu)\Delta_{\max}\ell d + \zeta + \frac{c_2(1+\nu)\ell^2\log(\ell/\delta)\bar{\gamma}(A)}{\epsilon_\ell} \\ &=: C_1 + \frac{C_2}{\epsilon_\ell} \end{aligned} \tag{10}$$

Let $\bar{\ell}$ be the first round for which:

$$T\epsilon_\ell \leq C_1 + \frac{C_2}{\epsilon_\ell}$$

Note that, if ϵ_ℓ solves this with equality, then:

$$\epsilon_\ell = \frac{C_1}{2T} + \frac{1}{2}\sqrt{\frac{4C_2}{T} + \frac{C_1^2}{T^2}}$$

is the only non-negative solution. It follows then that:

$$\epsilon_{\bar{\ell}} \leq \frac{C_1}{2T} + \frac{1}{2}\sqrt{\frac{4C_2}{T} + \frac{C_1^2}{T^2}} \leq \frac{C_1}{T} + \sqrt{\frac{C_2}{T}}$$

Since $\epsilon_{\bar{\ell}}$ is the largest such solution, it follows that $2\epsilon_{\bar{\ell}}$ doesn't satisfy this inequality so:

$$2\epsilon_{\bar{\ell}} > \frac{C_1}{2T} + \frac{1}{2}\sqrt{\frac{4C_2}{T} + \frac{C_1^2}{T^2}} \geq \sqrt{\frac{C_2}{T}}$$

so in particular:

$$\frac{1}{\epsilon_{\bar{\ell}}} \leq \sqrt{\frac{4T}{C_2}}$$

Assume that $f_\ell \leq T\epsilon_\ell$ for all ℓ. Using the monotonicity of ϵ_ℓ, for $\ell \geq \bar{\ell}$, we'll have:

$$f_\ell \leq T\epsilon_\ell \leq T\epsilon_{\bar{\ell}} \leq C_1 + \sqrt{C_2 T}$$

Furthermore, by (9), we'll have that the total regret up to round $\bar{\ell}$ will be bounded as:

$$\begin{aligned}\mathcal{R}_{1:\bar{\ell}} &\leq c_1(1+\nu)\Delta_{\max}\bar{\ell}^2 d + \bar{\ell}\Delta_{\max}n_{\mathfrak{f}} + \bar{\ell}\zeta + \frac{c_2(1+\nu)\bar{\ell}^2\log(\bar{\ell}/\delta)\bar{\gamma}}{\epsilon_{\bar{\ell}}} \\ &\leq C_1\bar{\ell} + \bar{\ell}\Delta_{\max}n_{\mathfrak{f}} + \frac{C_2}{\epsilon_{\bar{\ell}}} \\ &\leq C_1\bar{\ell} + \bar{\ell}\Delta_{\max}n_{\mathfrak{f}} + \sqrt{4C_2T}\end{aligned}$$

So in this case, since by Lemma 3 there are at most $\ell_{\max}(T)$ rounds, and since $f_\ell + \Delta_{\max}n_{\mathfrak{f}}$ upper bounds the regret of round ℓ, we'll have that the total regret will be bounded as:

$$\mathcal{R}_T \leq \ell_{\max}(T)\left(C_1 + \Delta_{\max}n_{\mathfrak{f}} + 3\sqrt{C_2T}\right)$$

Now assume there is some round such that $f_\ell > T\epsilon_\ell$ and denote this round as ℓ_{mle}. By construction, it will be the case that the MLE at this point has gap at most $\epsilon_{\ell_{\text{mle}}}$, so the total regret incurred from playing the MLE for the remainder of time will be bounded as $T\epsilon_{\ell_{\text{mle}}}$. Further, note that by (10):

$$T\epsilon_{\ell_{\text{mle}}} < f_{\ell_{\text{mle}}} \leq C_1 + \frac{C_2}{\epsilon_{\ell_{\text{mle}}}}$$

By definition $\bar{\ell}$ is the first round where $T\epsilon_\ell \leq C_1 + \frac{C_2}{\epsilon_\ell}$, so it follows that $\ell_{\text{mle}} \geq \bar{\ell}$. We can then upper bound the total regret incurred as:

$$\mathcal{R}_T \leq \sum_{\ell=1}^{\bar{\ell}} f_\ell + \sum_{\ell=\bar{\ell}+1}^{\ell_{\text{mle}}-1} f_\ell + T\epsilon_{\ell_{\text{mle}}} + \ell_{\text{mle}}\Delta_{\max}n_{\mathfrak{f}}$$

From (9), as above, we can bound:

$$\sum_{\ell=1}^{\bar{\ell}} f_\ell \leq C_1\bar{\ell} + \frac{C_2}{\epsilon_{\bar{\ell}}} \leq C_1\bar{\ell} + \sqrt{4C_2T}$$

Since by definition we'll have that $f_\ell \leq T\epsilon_\ell$ for $\ell \in [\bar{\ell}+1, \ell_{\text{mle}}-1]$, the second term can be bounded as:

$$\sum_{\ell=\bar{\ell}+1}^{\ell_{\text{mle}}-1} f_\ell \leq T\sum_{\ell=\bar{\ell}+1}^{\ell_{\text{mle}}-1}\epsilon_\ell \leq (\ell_{\text{mle}} - \bar{\ell} - 2)T\epsilon_{\bar{\ell}} \leq (\ell_{\text{mle}} - \bar{\ell} - 2)(C_1 + \sqrt{C_2T})$$

Finally:

$$T\epsilon_{\ell_{\text{mle}}} \leq T\epsilon_{\bar{\ell}} \leq C_1 + \sqrt{C_2T}$$

Combining this, we have that:

$$\mathcal{R}_T \leq \ell_{\max}(T)(C_1 + \Delta_{\max}n_{\mathfrak{f}} + 4\sqrt{C_2T})$$

Absolute Regret: Assume:

$$T > \frac{C_1}{\Delta_{\min}} + \frac{C_2}{\Delta_{\min}^2}$$

then we'll have that $\epsilon_{\bar{\ell}} < \Delta_{\min}$, so the algorithm will exit before reaching round $\epsilon_{\bar{\ell}}$. In this case, since there are at most $\lceil\log(4\Delta_{\max}/\Delta_{\min})\rceil$ stages by Lemma 3 and since, as noted above, on the good event, once $|\hat{\mathcal{X}}_\ell| = 1$, we will have identified the best arm and so will incur 0 regret for the rest of time, (9) gives:

$$\begin{aligned}\mathcal{R}_T \leq c_1(1+\nu)\Delta_{\max}\log_2(\Delta_{\max}/\Delta_{\min})^2 d + \lceil\log(4\Delta_{\max}/\Delta_{\min})\rceil\Delta_{\max}n_{\mathfrak{f}} + \lceil\log(4\Delta_{\max}/\Delta_{\min})\rceil\zeta \\ + \frac{c_2(1+\nu)\bar{\gamma}(A)\log(\log(\Delta_{\max}/\Delta_{\min})/\delta)\log_2(\Delta_{\max}/\Delta_{\min})^2}{\Delta_{\min}}\end{aligned}$$

By definition, it will always be the case that $\epsilon_{\ell_{\text{mle}}} > \Delta_{\min}$, if it exists, as we would have otherwise exited the algorithm already. By (9), we'll then have:

$$\begin{aligned}
\mathcal{R}_T &\le \mathcal{R}_{1:\ell_{\text{mle}}} + T\epsilon_{\ell_{\text{mle}}} \\
&\le c_1(1+\nu)\Delta_{\max}\log_2(\Delta_{\max}/\Delta_{\min})^2 d + \lceil \log_2(4\Delta_{\max}/\Delta_{\min})\rceil \Delta_{\max} n_{\mathfrak{f}} + \lceil \log(4\Delta_{\max}/\Delta_{\min})\rceil \zeta \\
&\qquad + \frac{c_2(1+\nu)\bar{\gamma}(A)\log(\log(\Delta_{\max}/\Delta_{\min})/\delta)\log_2(\Delta_{\max}/\Delta_{\min})^2}{\epsilon_{\ell_{\text{mle}}}} + T\epsilon_{\bar{\ell}} \\
&\overset{(a)}{\le} c_1(1+\nu)\Delta_{\max}\log_2(\Delta_{\max}/\Delta_{\min})^2 d + \lceil \log_2(4\Delta_{\max}/\Delta_{\min})\rceil \Delta_{\max} n_{\mathfrak{f}} + \lceil \log(4\Delta_{\max}/\Delta_{\min})\rceil \zeta \\
&\qquad + \frac{c_2(1+\nu)\bar{\gamma}(A)\log(\log(\Delta_{\max}/\Delta_{\min})/\delta)\log_2(\Delta_{\max}/\Delta_{\min})^2}{\epsilon_{\ell_{\text{mle}}}} + C_1 + \frac{C_2}{\epsilon_{\bar{\ell}}} \\
&\le 2C_1\log_2(\Delta_{\max}/\Delta_{\min}) + \lceil \log_2(4\Delta_{\max}/\Delta_{\min})\rceil \Delta_{\max} n_{\mathfrak{f}} + \frac{2C_2}{\Delta_{\min}}
\end{aligned}$$

where (a) holds by the definition of $\bar{\ell}$. If round ℓ_{mle} is never reached, then the upper bound above still holds, as we can still bound $\mathcal{R}_T \le \mathcal{R}_{1:\ell_{\text{mle}}}$, the regret we would have incurred had we reached ℓ_{mle}.

Finally, by Theorem 4 we can choose $\nu = 4$, $\zeta = 2$, and we will be able to compute the solution efficiently. □

Proof of Theorem 1. The proof of this result is very similar to the proof of Theorem 2 but we include the points where it differs for the sake of completeness. Unless otherwise noted, all notation is defined as in the proof of Theorem 2.

Good event: Define the events:

$$\mathcal{E}_{k,j} = \left\{ \sup_{z,z' \in \mathcal{S}_j} |(z-z')^\top(\hat{\theta}_k - \theta_*)| \le \mathbb{E}_\eta\left[\sup_{z,z'\in\mathcal{S}_j}(z-z')^\top A(\tau_k)^{-1/2}\eta\right] + \sqrt{2\max_{z,z'\in\mathcal{S}_j}\|z-z'\|^2_{A(\tau_k)^{-1}}\log(1/\delta_k)} \right\}$$

$$\mathcal{E} = \bigcap_{k=1}^{\infty}\bigcap_{j=0}^{k}\mathcal{E}_{k,j}$$

Proposition 6 implies that $\mathbb{P}[\mathcal{E}_{k,j}^c] \le \delta/k^3$ so:

$$\mathbb{P}[\mathcal{E}^c] \le \sum_{k=1}^{\infty}\sum_{j=0}^{k}\mathbb{P}[\mathcal{E}_{k,j}^c] \le \sum_{k=1}^{\infty}\sum_{j=0}^{k}\frac{\delta}{k^3} \le 3\delta$$

Estimation error: Henceforth we assume $\mathcal{E}$ holds. We proceed by induction to show that the gaps are always well-estimated. First we prove the base case. Let $k=1$ and consider any $x \in \mathcal{X}$. Then:

$$\begin{aligned}
|(x_* - x)^\top(\hat{\theta}_1 - \theta_*)| &\le \sup_{z,z'\in\mathcal{X}}|(z-z')^\top(\hat{\theta}_1-\theta_*)| \\
&\le \mathbb{E}_\eta\left[\sup_{z,z'\in\mathcal{X}}(z-z')^\top A(\tau_1)^{-1/2}\eta\right] + \sqrt{2\max_{z,z'\in\mathcal{X}}\|z-z'\|^2_{A(\tau_1)^{-1}}\log(1/\delta_k)} \\
&\overset{(a)}{=} \mathbb{E}_\eta\left[\sup_{z\in\mathcal{X}}(x_1-z)^\top A(\tau_1)^{-1/2}\eta\right] + \sqrt{2\max_{z,z'\in\mathcal{X}}\|z-z'\|^2_{A(\tau_1)^{-1}}\log(1/\delta_k)} \\
&\overset{(b)}{\le} \epsilon_1/8
\end{aligned}$$

where (a) follows by Proposition 7.5.2 of Vershynin [2018] and (b) follows since τ_1 is a feasible solution to (2). For the inductive step, assume that, for all $x \in \mathcal{S}_k$:

$$|(x_* - x)^\top(\hat{\theta}_k - \theta_*)| \le \epsilon_k/8$$

and for all $x \in \mathcal{S}_k^c$:

$$|(x_* - x)^\top(\hat{\theta}_k - \theta_*)| \le \Delta_x/8$$

Consider round $k+1$ and take $x \in \mathcal{S}_{k+1}^c$. There then exists some $k' \leq k$ such that $x \in \mathcal{S}_{k'} \backslash \mathcal{S}_{k'+1}$. Then:

$$\begin{aligned}
\frac{|(x_* - x)^\top (\hat{\theta}_{k+1} - \theta_*)|}{\Delta_x} &\leq \sup_{z,z' \in \mathcal{S}_{k'}} \frac{|(z - z')^\top (\hat{\theta}_{k+1} - \theta_*)|}{\Delta_x} \\
&\leq \mathbb{E}_\eta \left[\sup_{z,z' \in \mathcal{S}_{k'}} (z' - z)^\top A(\tau_{k+1})^{-1/2} \eta \right] + \sqrt{2 \max_{z,z' \in \mathcal{S}_{k'}} \|z - z'\|^2_{A(\tau_{k+1})^{-1}} \log(1/\delta_{k+1})} \\
&\overset{(a)}{=} 2\mathbb{E}_\eta \left[\sup_{z \in \mathcal{S}_{k'}} \frac{(x_{k+1} - z)^\top A(\tau_{k+1})^{-1/2} \eta}{\Delta_x} \right] + \sqrt{8 \max_{z \in \mathcal{S}_{k'}} \frac{\|z\|^2_{A(\tau_{k+1})^{-1}}}{\Delta_x^2} \log(1/\delta_{k+1})} \\
&\overset{(b)}{\leq} 4\mathbb{E}_\eta \left[\sup_{z \in \mathcal{S}_{k'}} \frac{(x_{k+1} - z)^\top A(\tau_{k+1})^{-1/2} \eta}{\epsilon_{k+1} + \Delta_x} \right] + \sqrt{32 \max_{z \in \mathcal{S}_{k'}} \frac{\|z\|^2_{A(\tau_{k+1})^{-1}}}{(\epsilon_{k+1} + \Delta_x)^2} \log(1/\delta_{k+1})} \\
&\overset{(c)}{\leq} 8\mathbb{E}_\eta \left[\sup_{z \in \mathcal{S}_{k'}} \frac{(x_{k+1} - z)^\top A(\tau_{k+1})^{-1/2} \eta}{\epsilon_{k+1} + \Delta_z} \right] + \sqrt{128 \max_{z \in \mathcal{S}_{k'}} \frac{\|z\|^2_{A(\tau_{k+1})^{-1}}}{(\epsilon_{k+1} + \Delta_z)^2} \log(1/\delta_{k+1})} \\
&\leq 8\mathbb{E}_\eta \left[\sup_{z \in \mathcal{X}} \frac{(x_{k+1} - z)^\top A(\tau_{k+1})^{-1/2} \eta}{\epsilon_{k+1} + \Delta_z} \right] + \sqrt{128 \max_{z \in \mathcal{X}} \frac{\|z\|^2_{A(\tau_{k+1})^{-1}}}{(\epsilon_{k+1} + \Delta_z)^2} \log(1/\delta_{k+1})} \\
&\overset{(d)}{\leq} 16\mathbb{E}_\eta \left[\sup_{z \in \mathcal{X}} \frac{(x_{k+1} - z)^\top A(\tau_{k+1})^{-1/2} \eta}{\epsilon_{k+1} + \Delta_z} \right] + \sqrt{512 \max_{z \in \mathcal{X}} \frac{\|z\|^2_{A(\tau_{k+1})^{-1}}}{(\epsilon_{k+1} + \Delta_z)^2} \log(1/\delta_{k+1})} \\
&\overset{(e)}{\leq} 1/8
\end{aligned}$$

where (a) follows by Proposition 7.5.2 of Vershynin [2018], (b) follows since $\Delta_x \geq \epsilon_{k+1}$ by virtue of the fact that $x \in \mathcal{S}_{k+1}^c$, so $\Delta_x \geq (\epsilon_{k+1} + \Delta_x)/2$, (c) follows since $\Delta_x \in [\epsilon_{k'+1}, \epsilon_{k'}]$ and for any $z \in \mathcal{S}_{k'}$, we will have theta $\Delta_z \leq \epsilon_{k'}$, so $\epsilon_{k+1} + \Delta_x \geq \epsilon_{k+1} + \epsilon_{k'+1} \geq \epsilon_{k+1} + \Delta_z/2$, (d) holds by the inductive hypothesis and Lemma 1 of Katz-Samuels et al. [2020] and taking $\hat{\Delta}_z$ to be the estimate of Δ_z at round $k+1$, and (e) holds since τ_{k+1} is a feasible solution to (3). We can perform a similar calculation to get the same thing for $x \in \mathcal{S}_{k+1}$, allowing us to conclude that, for all $x \in \mathcal{S}_{k+1}$:

$$|(x_* - x)^\top (\hat{\theta}_{k+1} - \theta_*)| \leq \epsilon_{k+1}/8$$

and for all $x \in \mathcal{S}_{k+1}^c$:

$$|(x_* - x)^\top (\hat{\theta}_{k+1} - \theta_*)| \leq \Delta_x/8$$

From here the remaining calculations on the gap estimates performed in the proof of Theorem 2 hold almost identically.

Bounding the Round Regret: From (6), it follows that the constraint in (2) is tighter than the following constraint:

$$\mathbb{E}_\eta \left[\max_{x \in \mathcal{X}} \frac{(x_\ell - x)^\top A(\tau)^{-1/2} \eta}{\epsilon_\ell + \Delta_x} \right] + \sqrt{2 \max_{x \in \mathcal{X}} \frac{\|x\|^2_{A(\tau)^{-1}}}{(\epsilon_\ell + \Delta_x)^2} \log(2\ell^3/\delta)} \leq \frac{1}{256} \tag{11}$$

so any τ satisfying this inequality is also a feasible solution to (2).

From here we follow the same pattern as in the proof of Theorem 2. We handle each term in the constraint separately. For the second term, note that we can upper bound:

$$\begin{aligned}
\max_{x \in \mathcal{X}} \frac{\|x\|^2_{A(\tau)^{-1}}}{(\epsilon_\ell + \Delta_x)^2} &\leq \max \left\{ \max_{x \in \mathcal{S}_\ell} \frac{\|x\|^2_{A(\tau)^{-1}}}{(\epsilon_\ell + \Delta_x)^2}, \max_{j < \ell} \max_{x \in \mathcal{S}_j \backslash \mathcal{S}_{j+1}} \frac{\|x\|^2_{A(\tau)^{-1}}}{(\epsilon_\ell + \Delta_x)^2} \right\} \\
&\leq 2 \max \left\{ \epsilon_\ell^{-2} \max_{x \in \mathcal{S}_\ell} \|x\|^2_{A(\tau)^{-1}}, \max_{j < \ell} \epsilon_j^{-2} \max_{x \in \mathcal{S}_j \backslash \mathcal{S}_{j+1}} \|x\|^2_{A(\tau)^{-1}} \right\} \\
&\leq 2 \max \left\{ \epsilon_\ell^{-2} \max_{x \in \mathcal{S}_\ell} \|x\|^2_{A(\tau)^{-1}}, \max_{j < \ell} \epsilon_j^{-2} \max_{x \in \mathcal{S}_j} \|x\|^2_{A(\tau)^{-1}} \right\} \\
&\leq 2 \max_{j \leq \ell} \epsilon_j^{-2} \max_{x \in \mathcal{S}_j} \|x\|^2_{A(\tau)^{-1}}
\end{aligned}$$

We now choose $\bar{\tau} = \ell^2 \sum_{j=1}^{\ell} \tau_j(1) + 1048576 d \log(2\ell^3/\delta) \sum_{j=1}^{\ell} \epsilon_j^{-2} \lambda_j^{\mathrm{kf}}$, where λ_j^{kf} is the distribution minimizing $\max_{x \in \mathcal{S}_j} \|x\|_{A(\lambda)^{-1}}^2$. By the same argument as in the proof of Theorem 2, effectively ignoring the second term, we will have:

$$\mathbb{E}_\eta \left[\max_{x \in \mathcal{X}} \frac{(x_\ell - x)^\top A(\bar{\tau})^{-1/2} \eta}{\epsilon_\ell + \Delta_x} \right] \leq \frac{1}{512}$$

For the second term, by the Kiefer-Wolfowitz Theorem in the bandit case, and Proposition 9 in the semi-bandit case, we'll have:

$$\begin{aligned} \max_{j \leq \ell} \epsilon_j^{-2} \max_{x \in \mathcal{S}_j} \|x\|_{A(\bar{\tau})^{-1}}^2 & \leq \max_{j \leq \ell} \epsilon_j^{-2} \max_{x \in \mathcal{S}_j} \|x\|_{A(cd \log(2\ell^3/\delta) \epsilon_j^{-2} \lambda_j^{\mathrm{kf}})^{-1}}^2 \\ & \leq \frac{1}{cd \log(2\ell^3/\delta)} \max_{j \leq \ell} \max_{x \in \mathcal{S}_j} \|x\|_{A(\lambda_j^{\mathrm{kf}})^{-1}}^2 \\ & \leq \frac{1}{1048576 \log(2\ell^3/\delta)} \end{aligned}$$

So:

$$\sqrt{2 \max_{x \in \mathcal{X}} \frac{\|x\|_{A(\bar{\tau})^{-1}}^2}{(\epsilon_\ell + \Delta_x)^2} \log(2\ell^3/\delta)} \leq \frac{1}{512}$$

From this it follows $\bar{\tau}$ is a feasible solution to (2). Furthermore, by Lemma 2, the total regret incurred by playing $\ell^2 \sum_{j=1}^{\ell} \tau_j(1)$ is bounded by:

$$c_1 \Delta_{\max} \ell d + \frac{c_2 \ell^2 \bar{\gamma}(A)}{\epsilon_\ell}$$

and the total regret incurred playing $cd \log(2\ell^3/\delta) \sum_{j=1}^{\ell} \epsilon_j^{-2} \lambda_j^{\mathrm{kf}}$ is bounded as:

$$c_1 \Delta_{\max} \ell d + \frac{c_2 d \log(2\ell^3/\delta)}{\epsilon_\ell}$$

Following the same argument as in Theorem 2, it follows that:

$$\mathcal{R}_\ell \leq c_1 \Delta_{\max} \ell d + \frac{c_2(\ell^2 \bar{\gamma}(A) + d \log(2\ell^3/\delta))}{\epsilon_\ell}$$

From here the argument follows identically to the proof of Theorem 4, so we omit the remainder of the proof. □

Lemma 2. *Given an ℓ such that $\epsilon_\ell > \Delta_{\min}$, let λ_k be any distribution supported on $\mathcal{S}_k$ and for any ξ set:*

$$\tau_k = \xi \epsilon_k^{-2}$$

Play the distributions $\kappa_k \leftarrow \mathrm{ROUND}(\lambda_k, \lceil \tau_k \rceil \vee q(1/2), 1/2)$ for $k = 1, \ldots, \ell$, where ROUND is defined as in Section A. Then the total gap-dependent regret incurred by this procedure is bounded by:

$$c_1 \Delta_{\max} \ell d + \frac{c_2 \xi}{\epsilon_\ell}$$

Proof. We can think of this procedure as a deterministic variant of action elimination. We can bound the regret incurred as:

$$\begin{aligned} \sum_{x \in \mathcal{X} \backslash \{x^*\}} \Delta_x T_x & \leq \sum_{k=1}^{\ell} \epsilon_k (\tau_k + q(1/2) + 1) \\ & \leq \Delta_{\max} \ell (q(1/2) + 1) + \sum_{k=1}^{\ell} \epsilon_k \tau_k \\ & \leq \Delta_{\max} \ell (q(1/2) + 1) + \xi \sum_{k=1}^{\ell} \epsilon_k^{-1} \\ & \leq \Delta_{\max} \ell (q(1/2) + 1) + \xi \sum_{k=1}^{\ell} \frac{2^k}{\Delta_{\max}} \\ & \leq \Delta_{\max} \ell (q(1/2) + 1) + \frac{c \xi}{\epsilon_\ell} \end{aligned}$$

The results on the rounding procedure follow from Katz-Samuels et al. [2020], Allen-Zhu et al. [2020]. □

Lemma 3. *Given a T, Algorithm 1 will run for at most:*

$$\ell_{\max}(T) := \log_2\left(\frac{\max_{x\in\mathcal{X}}\|x\|_2}{\min_{x\in\mathcal{X}}\|x\|_2}\left(\mathrm{diam}(\mathcal{X})\|\theta\|_2\sqrt{T}+3\right)\right)+1$$

rounds. Furthermore, regardless of T, Algorithm 1 will run for at most:

$$\lceil \log_2(4\Delta_{\max}/\Delta_{\min})\rceil$$

rounds.

Proof. Note that τ_ℓ must satisfy:

$$\mathbb{E}_\eta\left[\max_{x\in\mathcal{X}}\frac{(x_\ell-x)^\top A(\tau_\ell)^{-1/2}\eta}{\epsilon_\ell+\hat{\Delta}_x}\right]\leq\frac{1}{128(1+\sqrt{\pi\log(2\ell^3/\delta)})}$$

However:

$$\begin{aligned}
\mathbb{E}_\eta\left[\max_{x\in\mathcal{X}}\frac{(x_\ell-x)^\top A(\tau_\ell)^{-1/2}\eta}{\epsilon_\ell+\hat{\Delta}_x}\right] &\overset{(a)}{\geq} \frac{1}{\sqrt{2\pi}}\max_{x,y\in\mathcal{X}}\left\|A(\tau_\ell)^{-1/2}\left(\frac{x}{\epsilon_\ell+\hat{\Delta}_x}-\frac{y}{\epsilon_\ell+\hat{\Delta}_y}\right)\right\|_2\\
&\geq \frac{1}{\sqrt{2\pi}}\left\|A(\tau_\ell)^{-1/2}\left(\frac{x^*}{\epsilon_\ell+\hat{\Delta}_{x^*}}-\frac{x_{\max}}{\epsilon_\ell+\hat{\Delta}_{x_{\max}}}\right)\right\|_2\\
&\overset{(b)}{\geq} \frac{1}{\sqrt{2\pi\tau_\ell}}\frac{1}{\max_{x\in\mathcal{X}}\|x\|_2}\left\|\frac{x^*}{\epsilon_\ell+\hat{\Delta}_{x^*}}-\frac{x_{\max}}{\epsilon_\ell+\hat{\Delta}_{x_{\max}}}\right\|_2\\
&\geq \frac{1}{\sqrt{2\pi\tau_\ell}}\frac{1}{\max_{x\in\mathcal{X}}\|x\|_2}\left(\frac{\|x^*\|_2}{\epsilon_\ell+\hat{\Delta}_{x^*}}-\frac{\|x_{\max}\|_2}{\epsilon_\ell+\hat{\Delta}_{x_{\max}}}\right)\\
&\overset{(c)}{\geq} \frac{1}{\sqrt{2\pi\tau_\ell}}\frac{1}{\max_{x\in\mathcal{X}}\|x\|_2}\left(\frac{2\|x^*\|_2}{3\epsilon_\ell}-\frac{2\|x_{\max}\|_2}{\Delta_{\max}}\right)\\
&\geq \frac{2}{3\sqrt{2\pi\tau_\ell}}\left(\frac{\min_{x\in\mathcal{X}}\|x\|_2}{\max_{x\in\mathcal{X}}\|x\|_2}\frac{1}{\epsilon_\ell}-\frac{3}{\Delta_{\max}}\right)
\end{aligned}$$

where (a) follows by Proposition 7.5.2 of Vershynin [2018], (b) follows since for any λ:

$$A(\lambda)\preceq(\max_{x\in\mathcal{X}}\|x\|_2^2)I$$

and (c) follows by (6). Thus:

$$\begin{aligned}
\tau_\ell &\geq \frac{4(128(1+\sqrt{\pi\log(2\ell^3/\delta)}))^2}{18\pi}\left(\frac{\min_{x\in\mathcal{X}}\|x\|_2}{\max_{x\in\mathcal{X}}\|x\|_2}\frac{1}{\epsilon_\ell}-\frac{3}{\Delta_{\max}}\right)^2\\
&\geq \left(\frac{\min_{x\in\mathcal{X}}\|x\|_2}{\max_{x\in\mathcal{X}}\|x\|_2}\frac{1}{\epsilon_\ell}-\frac{3}{\Delta_{\max}}\right)^2\\
&= \frac{1}{\Delta_{\max}^2}\left(\frac{\min_{x\in\mathcal{X}}\|x\|_2}{\max_{x\in\mathcal{X}}\|x\|_2}2^\ell-3\right)^2
\end{aligned}$$

where the final equality holds since $\epsilon_\ell=\Delta_{\max}2^{-\ell}$. If round ℓ is the last round the algorithm completes before terminating, we'll have that $T\geq\tau_\ell$, so:

$$T\geq\frac{1}{\Delta_{\max}^2}\left(\frac{\min_{x\in\mathcal{X}}\|x\|_2}{\max_{x\in\mathcal{X}}\|x\|_2}2^\ell-3\right)^2 \implies \log_2\left(\frac{\max_{x\in\mathcal{X}}\|x\|_2}{\min_{x\in\mathcal{X}}\|x\|_2}\left(\Delta_{\max}\sqrt{T}+3\right)\right)\geq\ell$$

The first conclusion follows by Lemma 4.

For the second conclusion, note that, as we showed above, on the good event we will have that for all $x \in \hat{\mathcal{X}}_\ell$, $\Delta_x \leq 2\epsilon_\ell$. Thus, once $\epsilon_\ell \leq \Delta_{\min}/4$, we can guarantee that for any $x \in \hat{\mathcal{X}}_\ell$, $\Delta_x \leq \Delta_{\min}/2$ which implies that x is the optimal arm so $|\hat{\mathcal{X}}_\ell| = 1$ and the algorithm will have terminated. It follows that:

$$\epsilon_\ell = \Delta_{\max} 2^{-\ell} \leq \Delta_{\min}/4 \implies \ell \leq \log_2(4\Delta_{\max}/\Delta_{\min})$$

□

Lemma 4.

$$\Delta_{\max} \leq \|\theta\|_2 \mathrm{diam}(\mathcal{X})$$

Proof.

$$\Delta_{\max} = \langle \theta, x^* - x_{\max} \rangle \leq \|\theta\|_2 \max_{x,y \in \mathcal{X}} \|x - y\|_2$$

□

C Pure Exploration Proofs

For the sake of clarity, we rewrite the pure exploration algorithm (see Algorithm 3).

Algorithm 3 Computationally Efficient Pure Exploration Algorithm Semi-Bandit Feedback

Input: Set of arms $\mathcal{X}$, largest gap $\Delta_{\max}$, confidence δ, total time T
$\hat{\mathcal{X}}_1 = \mathcal{X}, \hat{\theta}_0 = 0, \ell \leftarrow 1$
while $|\hat{\mathcal{X}}_\ell| > 1$ and total pulls less than T **do**
    $x_\ell \leftarrow \arg\max_{x \in \mathcal{X}} x^\top \hat{\theta}_{\ell-1}$, $\epsilon_\ell \leftarrow \Delta_{\max} 2^{-\ell}$
    Let τ_ℓ be a solution to:

$$\begin{aligned} &\arg\min_{\tau} \sum_{x \in \mathcal{X}} \tau_x \\ &\text{s.t. } \mathbb{E}_\eta \left[\max_{x \in \mathcal{X}} \frac{(x_\ell - x)^\top A(\tau)^{-1/2} \eta}{\epsilon_\ell + \hat{\Delta}_x} \right] \leq \frac{1}{128(1 + \sqrt{\pi \log(2\ell^3/\delta)})} \end{aligned} \tag{12}$$

    $\alpha_\ell \leftarrow \mathrm{SPARSE}(\tau_\ell, n_{\mathfrak{f}})$
    Pull arm x $\alpha_{\ell,x}$ times, compute $\hat{\theta}_\ell$
    if $\mathrm{MINGAP}(\widehat{\theta}_\ell, \mathcal{X}) \geq 3\epsilon_\ell/2$ **then**
        break
    end if
    Pull arm x $\lceil \tau_{\ell,x} \rceil$ times, compute $\hat{\theta}_\ell$ from this data, form gap estimates $\hat{\Delta}_x$ from $\hat{\theta}_\ell$
    $\ell \leftarrow \ell + 1$
end while
return $\arg\max_{x \in \mathcal{X}} x^\top \hat{\theta}_\ell$

Theorem (2) shows that we can solve (12) in polynomial-time, but note that it is easier to solve (12) approximately by calling stochastic Frank-Wolfe to solve

$$\inf_{\lambda \in \triangle} \mathbb{E}_\eta \left[\max_{x \in \mathcal{X}} \frac{(x_\ell - x)^\top A(\tau)^{-1/2} \eta}{\epsilon_\ell + \hat{\Delta}_x} \right]$$

and the convergence rate shown in Lemma 5 applies.

The MINGAP subroutine (Algorithm 4), originally provided in Chen et al. [2017], is a computationally scalable method to compute the empirical gap between the empirically best arm and the empirically second best arm. It uses at most d calls to the linear maximization oracle.

We note that the correctness and sample complexity proofs are quite similar to the proof of Theorem in Katz-Samuels et al. [2020], but we include it for the sake of completeness. The main contribution of our paper for

Algorithm 4 MINGAP

Input: $\mathcal{X}$, estimate $\tilde{\theta}$
$\tilde{x} \longleftarrow \arg\max_{x\in\mathcal{X}} \tilde{\theta}^\top x$
$\widehat{\Delta}_{min} \longleftarrow \infty$
for $i = 1, 2, \ldots, d$ s.t. $i \in \tilde{x}$ **do**

$$\tilde{\theta}^{(i)} = \begin{cases} \tilde{\theta}_j & j \neq i \\ -\infty & j = i \end{cases}$$

  $\tilde{x}^{(i)} \longleftarrow \arg\max_{x\in\mathcal{X}} x^\top \tilde{\theta}^{(i)}$
  if $\tilde{\theta}^\top(\tilde{x} - \tilde{x}^{(i)}) \leq \widehat{\Delta}_{min}$ **then**
    $\widehat{\Delta}_{min} \longleftarrow \tilde{\theta}^\top(\tilde{x} - \tilde{x}^{(i)})$
  end if
end for
return $\widehat{\Delta}_{min}$

the pure exploration problem is a computational method to solve (12) even when the number of variables $|\mathcal{X}|$ is exponential in the dimension.

Proof of Theorem 3. **Step 1: A good event and well-estimated gaps** Using the identical argument to the first two steps of the proof of Theorem 2, we have that with probability at least $1-\delta$ at every round k, for all $x \in \mathcal{S}_k$:

$$|(x_* - x)^\top(\hat{\theta}_{k+1} - \theta_*)| \leq \epsilon_k/8 \tag{13}$$

and for all $x \in \mathcal{S}_{k+1}^c$:

$$|(x_* - x)^\top(\hat{\theta}_{k+1} - \theta_*)| \leq \Delta_x/8. \tag{14}$$

For the remainder of the proof we suppose that this good event holds.

Step 2: Correctness. It is enough to show at round k, if $x_k \neq x_*$, then the $\textsc{Unique}(\mathcal{X}, \widehat{\theta}_k, \epsilon_k)$ returns false. Inspecting $\textsc{Unique}$, a sufficient condition is to show that $(x_k - x_*)^\top\widehat{\theta}_k - \epsilon_k \leq 0$. By (13) and (14), we have that

$$\begin{aligned}(x_k - x_*)^\top\widehat{\theta}_k - \epsilon_k &= (x_k - x_*)^\top(\widehat{\theta}_k - \theta) - \Delta_{x_k} - \epsilon_k \\ &\leq \max\left(\frac{\Delta_{x_k}}{8}, \frac{\epsilon_k}{8}\right) - \Delta_{x_k} - \epsilon_k \\ &\leq 0\end{aligned}$$

proving correctness.

Step 3: Bound the Sample Complexity. Letting $\tilde{x}_k = \arg\max_{x\neq x_k} \widehat{\theta}_k^\top x$, $\textsc{Unique}(\mathcal{Z}, \widehat{\theta}_k, \epsilon_k)$ at round k checks whether $\widehat{\theta}_k^\top(x_k - \tilde{x}_k)$ is at least ϵ_k, and terminates if it is. Thus, (13) and (14), the algorithm terminates and outputs x_* once $k \geq c\log(\Delta_{\max}/\Delta_{\min})$.

Thus, the sample complexity is upper bounded by

$$\sum_{k=1}^{c\log(\Delta_{\max}/\Delta_{\min})} \sum_{x\in\mathcal{X}} \lceil \alpha_{k,x} \rceil \leq c'[\log(\Delta_{\max}/\Delta_{\min})d + \sum_{k=1}^{c\log(\Delta_{\max}/\Delta_{\min})} \inf_{\lambda\in\triangle} \mathbb{E}_{\eta\sim N(0,I)}[\max_{x\in\mathcal{X}} \frac{(x_k - x)^\top A_{\text{semi}}(\lambda)^{-1/2}\eta}{2^{-k}\Gamma + \widehat{\theta}_k^\top(x_k - x)}]^2] \tag{15}$$

where we used the fact that the rounding procedure can use $O(d)$ points in the semi-bandit case. Thus, it suffices to upper bound the second term in the above expression. Fix $\lambda \in \triangle$. Then,

$$\mathbb{E}_{\eta \sim N(0,I)}[\max_{x \in \mathcal{X}} \frac{(x_k - x)^\top A_{\text{semi}}(\lambda)^{-1/2}\eta}{\epsilon_k + \widehat{\theta}_k^\top (x_k - x)}]^2 \leq c\mathbb{E}_{\eta \sim N(0,I)}[\max_{x \in \mathcal{X}} \frac{(x_k - x)^\top A_{\text{semi}}(\lambda)^{-1/2}\eta}{\epsilon_k + \Delta_x}]^2$$
$$\leq c'[\mathbb{E}_{\eta \sim N(0,I)}[\max_{x \in \mathcal{X}} \frac{(x_* - x)^\top A_{\text{semi}}(\lambda)^{-1/2}\eta}{\epsilon_k + \Delta_x}]^2$$
$$+ \mathbb{E}_{\eta \sim N(0,I)}[\max_{x \in \mathcal{X}} \frac{(z_* - x_k)^\top A_{\text{semi}}(\lambda)^{-1/2}\eta}{\epsilon_k + \Delta_x}]^2]$$

Fix $x_0 \in \mathcal{X} \setminus \{x_*\}$. The first term is bounded as follows.

$$\mathbb{E}_{\eta \sim N(0,I)}[\max_{x \in \mathcal{X}} \frac{(x_* - x)^\top A_{\text{semi}}(\lambda)^{-1/2}\eta}{\epsilon_k + \Delta_x}]^2$$
$$= \mathbb{E}_{\eta \sim N(0,I)}[\max_{x \in \mathcal{X} \setminus \{x_*\}} \max(\frac{(x_* - x)^\top A_{\text{semi}}(\lambda)^{-1/2}\eta}{\epsilon_k + \Delta_x}, 0)]^2$$
$$\leq 8\mathbb{E}_{\eta \sim N(0,I)}[\max_{x \in \mathcal{X} \setminus \{x_*\}} \frac{(x_* - x)^\top A_{\text{semi}}(\lambda)^{-1/2}\eta}{\epsilon_k + \Delta_x}]^2 + 8\frac{\|x_* - x_0\|^2_{A_{\text{semi}}(\lambda)^{-1}}}{\epsilon_k + \Delta_{x_0})^2} \tag{16}$$
$$\leq 8[\mathbb{E}_{\eta \sim N(0,I)}[\max_{x \in \mathcal{X} \setminus \{x_*\}} \frac{(x_* - x)^\top A_{\text{semi}}(\lambda)^{-1/2}\eta}{\Delta_x}]^2$$
$$+ \max_{x \neq x_*} \frac{\|x_* - x\|^2_{A_{\text{semi}}(\lambda)^{-1}}}{\Delta_x^2}] \tag{17}$$

where we obtained line (16) using exercise 7.6.9 in Vershynin [2018].

We also have that

$$\mathbb{E}_{\eta \sim N(0,I)}[\max_{x \in \mathcal{X}} \frac{(x_* - x_k)^\top A_{\text{semi}}(\lambda)^{-1/2}\eta}{\epsilon_k + \Delta_x}]^2 \leq \mathbb{E}_{\eta \sim N(0,I)}[\max(\frac{(x_* - x_k)^\top A_{\text{semi}}(\lambda)^{-1/2}\eta}{\epsilon_k}, 0)]^2$$
$$\leq c\frac{\|x_* - x_k\|^2_{A_{\text{semi}}(\lambda)^{-1}}}{\epsilon_k^2}$$
$$\leq c\frac{\|x_* - x_k\|^2_{A_{\text{semi}}(\lambda)^{-1}}}{\Delta_{x_k}^2} \tag{18}$$
$$\leq c \max_{x \in \mathcal{X} \setminus \{x_*\}} \frac{\|x_* - x\|^2_{A_{\text{semi}}(\lambda)^{-1}}}{\Delta_x^2} \tag{19}$$

where line (18) follows since (13), (14), and Lemma 1 in Katz-Samuels et al. [2020] imply that $x_k \in S_{k+2}$.

(15), (17), and (19) together imply that

$$\sum_{k=1}^{c\log(\Gamma/\Delta_{min})} \sum_{x \in \mathcal{X}} \lceil \alpha_{k,x} \rceil \leq c\log(\Delta_{\min}/\Delta_{\min})[d + \gamma^* + \rho^*],$$

completing the proof.

□

C.1 Lower Bound

In this section, we prove a lower bound for the combinatorial bandit setting with semi-bandit feedback. Fix a model θ and let $\nu_{\theta,i}$ denote the distribution of the observations when arm i is pulled. In this setting, at each round t, $Z^{(t)} \sim N(\theta, I)$ is drawn and

$$(\nu_{\theta,i})_j = \begin{cases} Z_j^{(t)} & j \in x_i \\ 0 & j \notin x_i \end{cases}.$$

Definition 1. *We say that an Algorithm is δ-PAC if for any instance $(\mathcal{X}, \theta_*)$, it returns $x \in \mathcal{X}$ with the largest mean with probability at least $1-\delta$.*

Theorem 6. *Fix an instance $(\theta_*, \mathcal{X})$ such that $\mathcal{X} \subset \{0,1\}^d$ and $x_* = \arg\max_{x\in\mathcal{X}} x^\top \theta$ is unique. Let $\mathcal{A}$ be a δ-PAC algorithm and let T be its total number of pulls on $(\theta_*, \mathcal{X})$. Then,*

$$\mathbb{E}_{\theta_*}[T] \geq \log(1/2.4\delta)\rho^* := \log(1/2.4\delta) \inf_{\lambda\in\triangle} \sup_{x\in\mathcal{X}\setminus\{x_*\}} \frac{\|x_* - x\|^2_{A_{\text{semi}}(\lambda)^{-1}}}{\theta^\top(x_*-x)^2}.$$

The proof is quite similar to the proof of Theorem 1 in Fiez et al. [2019].

Proof. For simplicity, label $\mathcal{X} = \{x_1, \dots, x_m\}$ and $x_* = x_1$. Define the set of alternative instances $\mathcal{O} = \{\theta : \arg\max_{x\in\mathcal{X}} x^\top\theta \neq x_1\}$. Let T_i denote the random number of times that x_i is pulled during the game. Then, noting that the standard transportation Lemma from Kaufmann et al. [2016] easily generalizes to semi-bandit feedback, we have that for any $\theta \in \mathcal{O}$,

$$\sum_{i=1}^n \mathbb{E}_{\theta_*}[T_i]\,\mathbf{KL}(\nu_{\theta_*,i}|\nu_{\theta,i}) \geq \ln(1/2.4\delta)$$

By a standard argument (see for example Theorem 1 Fiez et al. [2019]), this implies that

$$\mathbb{E}_{\theta_*}[T] \geq \ln(1/2.4\delta) \min_{\lambda\in\triangle}\max_{\theta\in\mathcal{O}} \frac{1}{\sum_{i=1}^m \lambda_i\,\mathbf{KL}(\nu_{\theta_*,i}|\nu_{\theta,i})}.$$

Let $\epsilon > 0$. For each $k \neq 1$, define

$$\theta^{(k)} = \theta_* - \frac{[(x_1-x_k)^\top\theta_* + \epsilon]A_{\text{semi}}(\lambda)^{-1}(x_1-x_k)}{(x_1-x_k)^\top A_{\text{semi}}(\lambda)^{-1}(x_1-x_k)}.$$

Note that

$$(x_k - x_1)^\top\theta^{(k)} = \epsilon$$

showing that $\theta^{(k)} \in \mathcal{O}$. Note that using the identity for the KL-divergence for a multivariate Gaussian, we have that

$$\begin{aligned}
\mathbf{KL}(\nu_{\theta_*,i}|\nu_{\theta^{(k)},i}) &= \frac{1}{2}\sum_{j\in x_i}(e_j^\top(\theta_* - \theta^{(k)})^2 \\
&= \frac{1}{2}(x_k^\top\theta_* + \epsilon)^2 \sum_{j\in x_i} \frac{(x_1-x_k)^\top A_{\text{semi}}(\lambda)^{-1} e_j e_j^\top A_{\text{semi}}(\lambda)^{-1}(x_1-x_k)}{[(x_1-x_k)^\top A_{\text{semi}}(\lambda)^{-1}(x_1-x_k)]^2}.
\end{aligned}$$

Then, we have that

$$\begin{aligned}
\mathbb{E}_{\theta_*}[T] &\geq \ln(1/2.4\delta)\min_{\lambda\in\triangle}\max_{\theta\in\mathcal{O}} \frac{1}{\sum_{i=1}^m \lambda_i\,\mathbf{KL}(\nu_{\theta_*,i}|\nu_{\theta,i})} \\
&\geq \ln(1/2.4\delta)\min_{\lambda\in\triangle}\max_{k\neq 1} \frac{1}{\sum_{i=1}^m \lambda_i\,\mathbf{KL}(\nu_{\theta_*,i}|\nu_{\theta^{(k)},i})} \\
&= 2\ln(1/2.4\delta)\min_{\lambda\in\triangle}\max_{k\neq 1} \frac{\|x_1-x_k\|^4_{A_{\text{semi}}(\lambda)^{-1}}}{(x_k^\top\theta_*+\epsilon)^2\sum_{i=1}^m\lambda_i\sum_{j\in x_i}(x_1-x_k)^\top A_{\text{semi}}(\lambda)^{-1}e_je_j^\top A_{\text{semi}}(\lambda)^{-1}(x_1-x_k)} \\
&= 2\ln(1/2.4\delta)\min_{\lambda\in\triangle}\max_{k\neq 1} \frac{\|x_1-x_k\|^4_{A_{\text{semi}}(\lambda)^{-1}}}{(x_k^\top\theta_*+\epsilon)^2(x_1-x_k)^\top A_{\text{semi}}(\lambda)^{-1}A_{\text{semi}}(\lambda)A_{\text{semi}}(\lambda)^{-1}(x_1-x_k)} \\
&= 2\ln(1/2.4\delta)\min_{\lambda\in\triangle}\max_{k\neq 1} \frac{\|x_1-x_k\|^2_{A_{\text{semi}}(\lambda)^{-1}}}{(x_k^\top\theta_*+\epsilon)^2}.
\end{aligned}$$

Since $\epsilon > 0$ was arbitrary, we may let $\epsilon \longrightarrow 0$, obtaining the result. ∎

Next, we state and prove a lower bound for the *non-interactive MLE*: it chooses an allocation $\{x_{I_1}, x_{I_2}, \dots, x_{I_T}\} \in \mathcal{X}$ prior to the game, then observes $y_{t,i} = \theta_{*,i} + \eta_{t,i}, \forall i \in x_{I_t}$ where $\eta_t \sim \mathcal{N}(0, I)$, and forms the MLE $\hat{\theta}_i = \frac{1}{T_i} \sum_{t=1, x_{I_t,i}=1}^{T} y_{t,i}$ and outputs $\widehat{x} = \arg\max_{x \in \mathcal{X}} z^\top \widehat{\theta}$. Since the non-interactive MLE may use knowledge of θ_* in choosing its allocation and the estimator and recommendation rules are very natural, we view the sample complexity of the non-interactive MLE as a good benchmark to measure the sample complexity of algorithms against. The following lower bound for the non-interactive MLE resembles Theorem 3 in Katz-Samuels et al. [2020].

Theorem 7. *Fix $\mathcal{X} \subset \{0,1\}^d$ and $\theta_* \in \mathbb{R}^d$. Let $\delta \in (0, 0.015]$. There exists a universal constant $c > 0$ such that if the non-interactive MLE uses less than $c(\gamma^* + \log(1/\delta)\rho^*)$ samples, it makes a mistake with probability at least δ.*

The proof is quite similar to the proof of Theorem 3 in Katz-Samuels et al. [2020], so we merely sketch it here.

Proof. Consider the combinatorial bandit protocol with $\mathcal{X} \subset \{0,1\}^d$ as the collection of sets: at each round $t \in \mathbb{N}$, the agent picks $J_t \in [d]$ and observes $\theta_{J_t} + N(0,1)$ (see Katz-Samuels et al. [2020] for a more precise definition). Let $T' \in \mathbb{N}$ and fix an allocation $I_1, \dots, I_{T'} \in [d]$. Define

$$\gamma^*_{\text{combi}}(I_1, \dots, I_{T'}) = \mathbb{E}_{\eta \sim N(0,I)}[\sup_{x \in \mathcal{X} \setminus \{x_*\}} \frac{(x_* - x)^\top (\sum_{s=1}^{T'} e_{I_s} e_{I_s}^\top)^{-1/2} \eta}{\Delta_x}]^2$$

$$\rho^*_{\text{combi}}(I_1, \dots, I_{T'}) = \sup_{x \in \mathcal{X} \setminus \{x_*\}} \frac{\|x_* - x\|^2_{(\sum_{s=1}^{T'} e_{I_s} e_{I_s}^\top)^{-1}}}{\Delta_x^2}.$$

Theorem 3 in Katz-Samuels et al. [2020] shows that there exists a universal constant $c > 0$ such that if $c \leq \gamma^*(I_1, \dots, I_{T'})$ or $c \leq \log(1/\delta)\rho^*(I_1, \dots, I_{T'})$, the with probability at least δ, the oracle MLE makes a mistake.

Now, consider the semi-bandit problem and wlog suppose that $\mathcal{X} = \{x_1, \dots, x_m\}$. Now, fix an allocation $x_{J_1}, \dots, x_{J_T} \in \mathcal{X}$ for the semi-bandit problem. Define $\lambda_i = \frac{1}{T} \sum_{s=1}^{T} \mathbb{1}\{J_s = i\}$. Suppose that $T \leq 1/2 \frac{1}{c} \log(1/\delta)\rho^* + \gamma^*] \leq \frac{1}{c} \max(\log(1/\delta)\rho^*, \gamma^*)$. Then,

$$cT \leq \gamma^* = \min_{\lambda \in \triangle} \gamma^*(\lambda) \leq \gamma^*(\lambda)$$

where

$$\gamma^*(\lambda) := \mathbb{E}_\eta \left[\sup_{x \in \mathcal{X} \setminus \{x_*\}} \frac{(x_* - x)^\top A_{\text{semi}}(\lambda)^{-1/2} \eta}{\theta_*^\top (x_* - x)} \right]^2.$$

Now, rearranging the above inequality,we have that

$$c \leq \gamma^*(T\lambda).$$

Note that the allocation $T\lambda$ for the semi-bandit problem specifies an allocation $I_1, \dots, I_{T'}$ for the combinatorial bandit problem and the stochastic process (and non-interactive MLE algorithm) is the same on both problems. Thus, $\gamma^*(T\lambda)$ can be interpreted as $\gamma^*_{\text{combi}}(I_1, \dots, I_{T'})$ in the combinatorial bandit protocol for some allocation $I_1, \dots, I_{T'}$, and we may apply the proof of Theorem 3 to obtain that with probability at least δ, the oracle MLE makes a mistake.

□

D Computational Complexity Results

D.1 Algorithmic Approach

In this section, we present the main computational algorithms and results in the paper, culiminating in the proof of Theorem 8, which immediately implies Theorem 4. For simplicity label $\mathcal{X} = \{x_1, \dots, x_m\}$. We can always find $\tilde{x}_1, \ldots, \tilde{x}_d \in \mathcal{X}$ such that $\cup_{i=1}^d \tilde{x}_i = [d]$ in d linear maximization oracle calls. For each $i \in [d]$, create a cost vector:

$$v_j^{(i)} = \begin{cases} \infty & j = i \\ 0 & j \neq i \end{cases}$$

and set $\tilde{x}_i = \arg\max_{x \in \mathcal{X}} x^\top v^{(i)}$. Thus, by reordering we may suppose that $\cup_{i=1}^d x_i = [d]$. Now, define

$$\tilde{\triangle} = \{\lambda \in \triangle : \lambda_i \geq \psi \ \forall i \in [d]\}$$

where $\psi \leq 1/d$. We optimize over $\tilde{\triangle}$ due to its computational benefits,e.g., controlling the second partial order derivatives of the Lagrangian of (5).

Algorithm 5 is the main algorithm (see Theorem 8 for its guarantee); it essentially does a grid search over the time horizon variable, $\tau \in [T]$. Note that for a fixed $\tau \in [T]$, we have that for all $\lambda \in \triangle$

$$\tau \sum_{x \in \mathcal{X}} [\beta + \bar{\theta}^\top (\bar{x} - x)]\lambda_x = \tau\beta + \tau \sum_{x \in \mathcal{X}} \bar{\theta}^\top (\bar{x} - x)\lambda_x$$

and thus we can ignore the term $\tau\beta$. Thus, Algorithm 5 calls Algorithm 6 to solve for a fixed $\tau \in [T]$ the following optimization problem.

$$\begin{aligned} \min_{\lambda \in \tilde{\triangle}} \tau \sum_{x \in \mathcal{X}} & \bar{\theta}^\top (\bar{x} - x)\lambda_x \\ \text{s.t. } & \mathbb{E}_\eta \left[\max_{x \in \mathcal{X}} \frac{(\bar{x} - x)^\top A_{\text{semi}}(\lambda)^{-1/2}\eta}{\beta + \bar{\theta}^\top (\bar{x} - x)} \right] \leq \sqrt{\tau} C \end{aligned} \tag{20}$$

To solve the above optimization problem, we convert it into a series of convex feasibility programs of the following form: $\exists ? \lambda \in \tilde{\triangle}$ such that

$$\begin{aligned} \tau \sum_{x \in \mathcal{X}} & \bar{\theta}^\top (\bar{x} - x)\lambda_x \leq \widehat{OPT} \\ \text{s.t. } & \mathbb{E}_\eta \left[\max_{x \in \mathcal{X}} \frac{(\bar{x} - x)^\top A_{\text{semi}}(\lambda)^{-1/2}\eta}{\beta + \bar{\theta}^\top (\bar{x} - x)} \right] \leq \sqrt{\tau} C \end{aligned}$$

and perform binary search over $\widehat{OPT}$. To solve each of these convex feasibility programs, we employ the Plotkin-Shmoys-Tardos reduction to online learning and apply Algorithm 7, a multiplicative weights update style algorithm. Lemmas 6 and 7 provide the guarantees for the multiplicative weights update algorithm and for the binary search procedure, respectively.

The Plotkin-Shmoys-Tardos reduction requires a method for solving for arbitrary $\kappa_1, \kappa_2 \in [0, 1]$:

$$\min_{\lambda \in \tilde{\triangle}} \mathcal{L}(\kappa_1, \kappa_2; \tau; \lambda) := \kappa_1 \tau \sum_{x \in \mathcal{X}} \bar{\theta}^\top (\bar{x} - x)\lambda_x + \kappa_2 (\mathbb{E}_\eta \left[\max_{x \in \mathcal{X}} \frac{(\bar{x} - x)^\top A_{\text{semi}}(\lambda)^{-1/2}\eta}{\beta + \bar{\theta}^\top (\bar{x} - x)} \right] - \sqrt{\tau} C).$$

To solve the above optimization problem, we use stochastic Frank-Wolfe (see Algorithm 8). Defining for a fixed $\eta \in \mathbb{R}^d$,

$$\mathcal{L}(\kappa_1, \kappa_2; \tau; \lambda; \eta) = \kappa_1 \tau \sum_{x \in \mathcal{X}} \bar{\theta}^\top (\bar{x} - x)\lambda_x + \kappa_2 (\max_{x \in \mathcal{X}} \frac{(\bar{x} - x)^\top A_{\text{semi}}(\lambda)^{-1/2}\eta}{\beta + \bar{\theta}^\top (\bar{x} - x)} - \sqrt{\tau} C).$$

we see that

$$\mathbb{E}_{\eta \sim N(0, I)} [\mathcal{L}(\kappa_1, \kappa_2; \tau; \lambda; \eta)] = \mathcal{L}(\kappa_1, \kappa_2; \tau; \lambda).$$

See Lemma 5 for our convergence result on stochastic Frank-Wolfe.

Finally, we note that each of our algorithms uses a global variable TOL, which for the theory we set to $\frac{(\sqrt{2}-1)C}{4}$. We note that C scales as $\frac{1}{\sqrt{\log(\frac{1}{\delta})}}$ and thus a polynomial dependence on $1/\text{TOL}$ results in a polynomial dependence on $\log(1/\delta)$.

Algorithm 5 Main

Input: Tolerance parameter $\text{TOL} \in (0,1)$, $\delta \in (0,1)$
$k \longleftarrow 1$, $\bar{\tau}_k \longleftarrow 2^k$
while $\bar{\tau}_k \leq T$ **do**
  $(\text{FEASIBLE}_k, \lambda_k) \longleftarrow \text{binSearch}(\bar{\tau}_k, \frac{\delta}{\log_2(T)})$
  $k \longleftarrow k+1$, $\bar{\tau}_k \longleftarrow 2^k$
end while
if FEASIBLE_k is False for all k **then**
  return "Program is not feasible"
end if
$\widehat{k}_* \longleftarrow \arg\min_k \{\bar{\tau}_k \sum_{x \in \mathcal{X}} \bar{\theta}^\top (\bar{x} - x)\lambda_{k,x} : \text{FEASIBLE}_k \text{ is True } \}$
return $(2\bar{\tau}_{\widehat{k}_*}, \lambda_{\widehat{k}_*})$

Algorithm 6 Binary Search (binSearch)

Input: $\bar{\tau} > 0$, $\delta \in (0,1)$, Tolerance parameter $\text{TOL} > 0$
$\text{LOW} \longleftarrow 0$, $\text{HIGH} \longleftarrow 2Td$
$(\text{FEASIBLE}, \bar{\lambda}) \longleftarrow \text{MW}(\bar{\tau}, \text{HIGH}, \frac{\delta}{\lceil \log_2(2Td/\text{TOL}) \rceil + 1})$ Check if program is feasible
if FEASIBLE is False **then**
  return $(\text{FEASIBLE}, \bar{\lambda})$
end if
while $\text{HIGH} - \text{LOW} \geq \text{TOL}$ **do** ▷ Initiate binary search
  $\widehat{OPT} \longleftarrow \frac{\text{LOW}+\text{HIGH}}{2}$
  $(\text{FEASIBLE}, \bar{\lambda}) \longleftarrow \text{MW}(\bar{\tau}, \widehat{OPT}, \frac{\delta}{\lceil \log_2(2Td/\text{TOL}) \rceil + 1})$
  if FEASIBLE **then**
    $\text{LOW} \longleftarrow \widehat{OPT}$
  else
    $\text{HIGH} \longleftarrow \widehat{OPT}$
  end if
end while
$(\text{FEASIBLE}, \bar{\lambda}) \longleftarrow \text{MW}(\bar{\tau}, \text{HIGH}, \frac{\delta}{\lceil \log_2(2Td/\text{TOL}) \rceil + 1})$
return $(\text{FEASIBLE}, \bar{\lambda})$

Algorithm 7 Multiplicative Weights Update Algorithm for Combinatorial Bandits with Semi-Bandit Feedback (MW)

Input: $\bar{\tau} > 0$, $\widehat{OPT} > 0$, Failure probability $\delta \in (0, 1)$, Tolerance parameter $\textsc{tol} > 0$
$\rho = \max(2dT, c\frac{d}{\beta\psi^{1/2}})$ for an appropriately chosen universal constant $c > 0$ (see the proof of Lemma 6)
$\eta = \min(\frac{\textsc{tol}}{4\rho}, 1/2)$, $R \longleftarrow \frac{16\rho^2 \ln(2)}{\textsc{tol}^2}$
$\textsc{Feasible} \longleftarrow \textsc{True}$ ▷ `Assume feasible program`
$w_i^{(1)} \longleftarrow 1$ for $i \in [2]$ ▷ `Initiate weights`
for $r = 1, 2, \ldots, R$ **do**
  $p_1^{(r)} \longleftarrow w_1^{(r)}/(w_1^{(r)} + w_2^{(r)})$ and $p_2^{(r)} \longleftarrow w_2^{(r)}/(w_1^{(r)} + w_2^{(r)})$
  $\lambda^{(r)} \longleftarrow \text{SFW}(p_1^{(r)}, p_2^{(r)}, \frac{\delta}{2R})$
  Define

$$h_1(\lambda) := \bar{\tau} \sum_{x \in \mathcal{X}} \bar{\theta}^\top (\bar{x} - x)\lambda_x - \widehat{OPT}$$

$$\widehat{h}_2(\lambda^{(r)}) = \text{estSup}(\lambda^{(r)}, \frac{\delta}{3R}) - \sqrt{\tau}C$$

$$\widehat{h}^{(r)}(\lambda^{(r)}) := p_1^{(r)} h_1(\lambda^{(r)}) + p_2^{(r)} \widehat{h}_2(\lambda^{(r)})$$

  if $\widehat{h}^{(r)}(\lambda^{(r)}) > 2\textsc{tol}$ **then**
    $\textsc{Feasible} \longleftarrow \textsc{False}$ ▷ `Declare infeasible program`
    Break
  end if
  $w_1^{(r+1)} \longleftarrow w_1^{(r)}(1 + \eta h_1(\lambda^{(r)}))$ ▷ `Update weights`
  $w_2^{(r+1)} \longleftarrow w_2^{(r)}(1 + \eta \widehat{h}_2(\lambda^{(r)}))$
end for
$\bar{\lambda}^{(r)} = \frac{1}{r}\sum_{s=1}^{r} \bar{\lambda}_s$
return $(\textsc{feasible}, \bar{\lambda}^{(r)})$

Algorithm 8 Stochastic Frank-Wolfe for Semi-Bandit Feedback (SFW)

Input: $\tau \geq 0$, $\kappa_1, \kappa_2 \in [0,1]$, $\delta \in (0,1)$.

$R_{\mathrm{SFW}} = \frac{8 \frac{1}{\beta \psi^{5/2}} d}{\mathrm{TOL}}$

$(q_r)_{r \in [R]} \in [0,1]^R$ such that $q_r = \frac{2}{r+1}$ and $(p_r)_{r \in [R]} \in \mathbb{N}^R$ such that $p_r = c \frac{1}{d \psi^2 q_r} \log(r^2/\delta)$ for an appropriately chosen universal constant $c > 0$ (see the proof of Lemma 5)

Initialize $\lambda_1 \in \tilde{\triangle}$ by setting $\lambda_{1,i} = 1/d$ if $i \in [d]$ and otherwise set $\lambda_{1,i} = 0$.

for $r = 1, 2, \ldots, R_{\mathrm{SFW}}$ **do**

  Draw $\eta^{(1)}, \ldots, \eta^{(p_r)} \sim N(0, I)$

  Compute

$$\tilde{\nabla}_r = \frac{1}{p_r} \sum_{j=1}^{p_r} \nabla \mathcal{L}(\kappa_1, \kappa_2; \tau; \lambda_r; \eta_j)$$

  Compute

$$i_r \longleftarrow \arg\max_{i \in [m]} -\tilde{\nabla}_{r,i} = -[\kappa_1 \tau \bar{\theta}^\top x_i + \kappa_2 \frac{1}{2} \frac{1}{p_r} \sum_{j=1}^{p_r} \frac{1}{[\beta + \bar{\theta}^\top (\bar{x} - \tilde{x}_j)]} \sum_{k \in (\bar{x} \Delta \tilde{x}_j) \cap x_i} \frac{\eta_k}{(\sum_{l : k \in x_l} \lambda_l)^{3/2}}]$$

  where

$$\tilde{x}_j = \arg\max_{x \in \mathcal{X}} \frac{\sum_{i \in \bar{x} \Delta x} \frac{\eta_i^{(j)}}{\sum_{x' : i \in x'} \lambda_{x'}}}{\beta + \bar{\theta}^\top (\bar{x} - x)}.$$

  is computed using Algorithm 9.

$$(v_r)_i = \begin{cases} \begin{cases} 0 & i \notin [d] \\ \psi & i \in [d] \setminus \{i_r\} \\ 1 - (d-1)\psi & i = i_r \end{cases}, & i_r \in [d] \\ \begin{cases} \psi & i \in [d] \\ 1 - d\psi & i = i_r \end{cases}, & i_r \notin [d] \end{cases}$$

$$\lambda_{r+1} \longleftarrow q_r v_r + (1 - q_r) \lambda_r$$

end for

return $\lambda_{R_{\mathrm{SFW}}}$

D.1.1 Subroutines

Algorithm 9, originally provided in Katz-Samuels et al. [2020], uses binary search and calls to the linear maximization oracle to compute

$$\frac{(\bar{x} - x)^\top A(\lambda)^{-1/2} \eta}{\beta + \bar{\theta}^\top (\bar{x} - x)}.$$

Algorithm 10 estimates

$$\mathbb{E}_\eta [\max_{x \in \mathcal{X}} \frac{(\bar{x} - x)^\top A_{\mathrm{semi}}(\lambda)^{-1/2} \eta}{\beta + \bar{\theta}^\top (\bar{x} - x)}].$$

Algorithm 9 computeMax

Define the following functions

$$g(\lambda;\eta;x) := \frac{(\bar{x}-x)^\top A(\lambda)^{-1/2}\eta}{\beta+\bar{\theta}^\top(\bar{x}-x)}$$
$$g(\lambda;\eta;r) := \max_{x\in\mathcal{X}} x^\top(A(\lambda)^{-1/2}\eta + r\bar{\theta}) - r(\beta+\bar{\theta}^\top x) - \bar{x}^\top A(\lambda)^{-1/2}\eta$$
$$g(\lambda;\eta;r;x) := x^\top(A(\lambda)^{-1/2}\eta + r\bar{\theta}) - r(\beta+\bar{\theta}^\top \bar{x}) - \bar{x}^\top A(\lambda)^{-1/2}\eta$$

Define

$$\textsc{low} = 0, \qquad \textsc{high} = 2$$

while $g(\lambda;\eta:\textsc{high}) \geq 0$ **do**
    $\textsc{high} \longleftarrow 2\cdot\textsc{high}$
end while
while $g(\lambda;\eta;\textsc{low}) \neq 0$ **do**
    if $g(\lambda;\eta;\frac{1}{2}(\textsc{high}+\textsc{low})) < 0$ **then**
        $\textsc{low} \longleftarrow \frac{1}{2}(\textsc{high}+\textsc{low})$
    else
        $\textsc{high} \longleftarrow \frac{1}{2}(\textsc{high}+\textsc{low})$
        $\textsc{low} \longleftarrow g(\lambda;\eta;x')$ for some $x' \in \arg\max_{x\in\mathcal{X}} g(\lambda;\eta;\textsc{low};x)$
    end if
end while
Return $\textsc{low}$

Algorithm 10 Estimate expected suprema (estimateSup)

Input: $\lambda \in \tilde{\triangle}$, failure probability $\delta > 0$, Tolerance parameter $\textsc{tol} > 0$, ,
$t = c\log(1/\delta)\frac{d}{\beta^2\psi\textsc{tol}^2}$
Draw $\eta_1,\ldots,\eta_t \sim N(0,I)$
Compute $g_s = \max_{x\in\mathcal{X}} \frac{(\bar{x}-x)^\top A_{\text{semi}}(\lambda)^{-1/2}\eta_s}{\beta+\bar{\theta}^\top(\bar{x}-x)}$ for $s = 1,\ldots,t$ using Algorithm 9.
return $\frac{1}{t}\sum_{s=1}^t g_s$

D.2 Main Optimization Proofs

For the sake of simplicity, we assume that T is a power of 2, and that the optimization problem is feasible. If the optimization problem is infeasible, we can determine this by applying stochastic Frank-Wolfe (see Lemma 5). For simplicity, we also assume that $\bar{\theta}^\top(\bar{x}-x) \leq \Delta_{\max} \leq 2d$ since typically it is assumed that $\|\theta\|_\infty \leq 1$ and whp $\left\|\widehat{\theta}_\ell\right\|_\infty = O(1)$ at every round ℓ. Further, note that whenever the algorithm is applied $C \leq 1$, and we assume this henceforth. We introduce the following functions to bound the number of linear maximization oracle calls:

$$\mathcal{A}(d,\beta,\psi,\textsc{tol},1/\delta,1/\xi) = O(\frac{d^2}{\textsc{tol}^3\beta^4\psi^8}[d+\log(\frac{d}{\beta\xi\psi\textsc{tol}})])$$
$$\mathcal{B}(d,\beta,\psi,\textsc{tol},1/\delta,1/\xi) = O(\log(1/\delta)\frac{d}{\beta^2\psi\textsc{tol}^2}[d+\log(\frac{d\Delta_{\max}}{\beta\xi})])$$
$$\mathcal{C}(d,\beta,\psi,\textsc{tol},1/\delta,1/\xi) = \frac{(dT)^2+\frac{d^2}{\beta^2\psi}}{\textsc{tol}^2}[\mathcal{A}(d,\beta,\psi,\textsc{tol},1/\delta,1/\xi)+\mathcal{B}(d,\beta,\psi,\textsc{tol},1/\delta,1/\xi)]$$

Note these are polynomial in $(d,\beta,\psi,1/\textsc{tol},\log(1/\delta),1/\xi)$. Our algorithms share a global parameter $\textsc{tol}$; it suffices to set $\textsc{tol} = \frac{(\sqrt{2}-1)C}{4}$. Define

$$M = \log_2(T)\log_2(\frac{2Td}{\textsc{tol}}).$$

We say a random variable X is sub-Gaussian with parameter σ^2 and write $X \in \mathcal{SG}(\sigma^2)$ if for all $\lambda \in \mathbb{R}$

$$\mathbb{E}[e^{\lambda(X-\mathbb{E}[X])}] \leq e^{\lambda^2\sigma^2/2}.$$

The following Lemma provides the convergence guarantee for stochastic Frank-Wolfe in the semi-bandit setting (see Algorithm 8).

Lemma 5. *Let $\delta \in (0,1)$, $\xi \in (0,1]$, $\kappa_1, \kappa_2 \in [0,1]$. With probability at least $1-\delta$ Algorithm 8 returns $\lambda_{R_{\text{SFW}}} \in \triangle$ such that*

$$\mathcal{L}(\kappa_1,\kappa_2;\tau;\lambda_{R_{\text{SFW}}}) \leq \min_{\lambda\in\tilde{\triangle}} \mathcal{L}(\kappa_1,\kappa_2;\tau;\lambda) + \text{TOL}$$

Furthermore, with probability at least $1-\frac{c\xi}{2^d}$, the number of oracle calls is bounded by

$$\mathcal{A}(d,\beta,\psi,1/\text{TOL},\log(1/\delta),1/\xi).$$

Proof. For simplicity, we focus on the case where $\kappa_1 = \kappa_2 = 1$ (the other cases are similar). We write $\mathcal{L}(\lambda)$ and $\mathcal{L}(\lambda;\eta)$ as abbreviations for $\mathcal{L}(\kappa_1,\kappa_2;\lambda)$ and $\mathcal{L}(\kappa_1,\kappa_2;\lambda;\eta)$.

Step 1: Bound the number of iterations of stochastic Frank-Wolfe. $\mathcal{L}(\lambda)$ is convex in λ by Proposition 7. Furthermore, $\max_{\lambda,\lambda'\in\tilde{\triangle}} \|\lambda-\lambda'\|_1 \leq 2$. Thus, by Proposition 8, it suffices to show

1. **Smoothness:** $\|\nabla\mathcal{L}(\lambda)-\nabla\mathcal{L}(\lambda')\|_\infty \leq L\|\lambda-\lambda'\|_1$ for an appropriate choice of L
2. **Small deviation with high probability:** p_r is chosen sufficiently large to ensure that with probability at least $1-\delta/r^2$

$$\left\|\tilde{\nabla}_r - \nabla\mathcal{L}(\lambda_{r-1})\right\|_\infty \leq \frac{Lq_r}{2}$$

Step 1.1: Smoothness. Let $\lambda,\lambda' \in \tilde{\triangle}$ and fix $i \in [m]$. It suffices to show that

$$|\frac{\partial\mathcal{L}(\lambda)}{\partial\lambda_i} - \frac{\partial\mathcal{L}(\lambda')}{\partial\lambda_i'}| \leq L\|\lambda-\lambda'\|_1.$$

For the sake of abbreviation, define $g(\lambda) := \frac{\partial\mathcal{L}(\lambda)}{\partial\lambda_i}$. By Lemma 13, we have that $\mathcal{L}(\lambda)$ is twice differentiable and that

$$\begin{aligned}\frac{\partial^2\mathcal{L}(\lambda)}{\partial\lambda_i\partial\lambda_j} &= \mathbb{E}[\frac{\partial^2\mathcal{L}(\lambda;\eta)}{\partial\lambda_i\partial\lambda_j}\mathbb{1}\{B\}] \\ &= \mathbb{E}[\frac{3}{4}\frac{1}{(\beta+\bar{\theta}^\top(\bar{x}-\tilde{x})}\sum_{k\in(\bar{x}\Delta\tilde{x})\cap x_i\cap x_j}\frac{\eta_k}{(\sum_{l:k\in x_l}\lambda_l)^{5/2}} : \tilde{x} = \arg\max_{x\in\mathcal{X}}\frac{\sum_{i\in\bar{x}\Delta x}\frac{1}{\sum_{x':i\in x'}\lambda_{x'}}}{\beta+\bar{\theta}^\top(\bar{x}-x)}]\end{aligned}$$

where

$$B = \{\eta : |\arg\max_{x\in\mathcal{X}}\frac{\sum_{i\in\bar{x}\Delta x}\frac{1}{\sum_{x':i\in x'}\lambda_{x'}}}{\beta+\bar{\theta}^\top(\bar{x}-x)}| = 1\}.$$

For any $\lambda \in \tilde{\triangle}$,

$$\begin{aligned}|\frac{\partial g(\lambda)}{\partial\lambda_j}| &= |\mathbb{E}[\frac{3}{4}\frac{1}{(\beta+\bar{\theta}^\top(\bar{x}-\tilde{x})}\sum_{k\in(\bar{x}\cap\tilde{x})\cap x_i\cap x_j}\frac{\eta_k}{(\sum_{l:k\in x_l}\lambda_l)^{5/2}}\mathbb{1}\{B\} : \tilde{x} = \arg\max_{x\in\mathcal{X}}\frac{\sum_{i\in\bar{x}\Delta x}\frac{1}{\sum_{x':i\in x'}\lambda_{x'}}}{\beta+\bar{\theta}^\top(\bar{x}-x)}.]| \\ &\leq \mathbb{E}[|\frac{3}{4}\frac{1}{(\beta+\bar{\theta}^\top(\bar{x}-\tilde{x})}\sum_{k\in(\bar{x}\cap\tilde{x})\cap x_i\cap x_j}\frac{\eta_k}{(\sum_{l:k\in x_l}\lambda_l)^{5/2}}|\mathbb{1}\{B\} : \tilde{x} = \arg\max_{x\in\mathcal{X}}\frac{\sum_{i\in\bar{x}\Delta x}\frac{1}{\sum_{x':i\in x'}\lambda_{x'}}}{\beta+\bar{\theta}^\top(\bar{x}-x)}.] \\ &\leq \mathbb{E}\frac{3}{4}\frac{1}{\beta}\sum_{k=1}^d|\eta_k| \\ &\leq c\frac{1}{\beta\psi^{5/2}}d.\end{aligned}$$

where we used Jensen's inequality and $c > 0$ is a universal constant.

Now, by the mean value theorem, there exists $s \in [0, 1]$ such that

$$\begin{aligned}|g(\lambda) - g(\lambda')| &\leq |\nabla g(s\lambda + (1-s)\lambda')^\top (\lambda - \lambda')| \\ &\leq \|\nabla g(s\lambda + (1-s)\lambda')\|_\infty \|\lambda - \lambda'\|_1 \\ &\leq c\frac{1}{\beta\psi^{5/2}} d \|\lambda - \lambda'\|_1\end{aligned}$$

where the second inequality follows by Holder's Inequality. Thus,

$$\|\nabla \mathcal{L}(\kappa; \lambda) - \nabla \mathcal{L}(\kappa; \lambda')\|_\infty \leq c\frac{1}{\beta\psi^{5/2}} d \|\lambda - \lambda'\|_1$$

For the sake of brevity, we write $L = \frac{1}{\beta\psi^{5/2}} d^{3/2}$ for the remainder of the proof.

Step 1.2: Small deviation with high probability. Now, we show that p_r is chosen sufficiently large to ensure that with probability at least $1 - \delta/r^2$

$$\left\|\widetilde{\nabla}_r - \nabla \mathcal{L}(\lambda_{r-1})\right\|_\infty \leq \frac{Lq_r}{2}. \tag{21}$$

Recall that

$$\begin{aligned}[\widetilde{\nabla}_r - \nabla \mathcal{L}(\lambda_{r-1})]_i &= [\widetilde{\nabla}_r - \nabla \mathcal{L}(\lambda_{r-1})]_i \\ &= \frac{1}{p_r}[\sum_{j=1}^{p_r} \frac{1}{2} \frac{1}{[\beta + \bar{\theta}^\top (\bar{x} - \tilde{x}_j)]} \sum_{k \in (\bar{x} \Delta \tilde{x}_j) \cap x_i} \frac{\eta_k}{(\sum_{l: k \in x_l} \lambda_l)^{3/2}} \\ &\quad - \mathbb{E}[(\frac{1}{2} \frac{1}{[\beta + \bar{\theta}^\top (\bar{x} - \tilde{x})]} \sum_{k \in (\bar{x} \Delta \tilde{x}) \cap x_i} \frac{\eta_k}{(\sum_{l: k \in x_l} \lambda_l)^{3/2}} : \tilde{x} = \arg\max_{x \in \mathcal{X}} \frac{\sum_{i \in \bar{x} \Delta x} \frac{1}{\sum_{x': i \in x'} \lambda_{x'}}}{\beta + \bar{\theta}^\top (\bar{x} - x)}.]\end{aligned}$$

where

$$\tilde{x}_j = \arg\max_{x \in \mathcal{X}} \frac{\sum_{i \in \bar{x} \Delta x} \frac{\eta_i^{(j)}}{\sum_{x': i \in x'} \lambda_{x'}}}{\beta + \bar{\theta}^\top (\bar{x} - x)}..$$

Note that

$$|\frac{1}{p_r} \sum_{j=1}^{p_r} \frac{1}{2} \frac{1}{[\beta + \bar{\theta}^\top (\bar{x} - \tilde{x}_j)]} \sum_{k \in \tilde{x}_j \cap x_i} \frac{\eta_k}{(\sum_{l: k \in x_l} \lambda_l)^{3/2}}| \leq \frac{1}{2} \frac{1}{p_r} \frac{1}{\beta\psi^{3/2}} \sum_{k=1}^{d} |\eta_k|.$$

Since

$$\frac{1}{2} \frac{1}{p_r} \frac{1}{\beta\psi^{3/2}} \sum_{k=1}^{d} |\eta_k| \in \mathcal{SG}(\frac{c\frac{1}{\beta^2\psi^3} d}{p_r})$$

we then have that by Lemma 2.6.8 in Vershynin [2018],

$$[\widetilde{\nabla}_r - \nabla \mathcal{L}(\lambda_{r-1})]_i \in \mathcal{SG}(\frac{c\frac{1}{\beta^2\psi^3} d}{p_r}).$$

Therefore, since $|\mathcal{X}| \leq 2^d$ and since $p_r = c\frac{\frac{1}{\beta^2\psi^3} d^2}{L^2 q_r^2}$ for an appropriately chosen universal constant, by a standard sub-Gaussian tail bound (21) follows.

Step 2: Bound the number of linear maximization oracle calls. Next, we bound the number of linear maximization oracle calls. At each round r, there is one linear maximization oracle call from finding the minimizing direction wrt the gradient over $\tilde{\triangle}$, but the dominant source of linear maximization oracles at each round is due to applying Algorithm 9 several times. Thus, it suffices to bound the number of linear maximization oracle calls due to Algorithm 9. Define the following event

$$\mathcal{E}_k = \{ \text{ the } k\text{th application of Algorithm 9 requires } O(d + \log(\frac{d}{\beta}) + \log(\Delta_{\max}\xi k^2)) \text{ oracle calls}\}$$
$$\mathcal{E} = \cap_k \mathcal{E}_k$$

Then, we have that

$$\Pr(\mathcal{E}) = \prod_{r=1}^{\infty} \Pr(\mathcal{E}_r | \cap_{s=1}^{r-1} \mathcal{E}_s) \geq \prod_{r=1}^{\infty} (1 - \frac{\xi}{2^d r^2}) = \frac{\sin(\pi \frac{\xi}{2^d})}{\pi + \frac{\xi}{2^d}} \geq 1 - \frac{\xi}{2^d}.$$

where we used the independence of each draw of a multivariate Gaussian in the algorithm and Lemma 9. The number of calls of Algorithm 9 at each iteration is upper bounded by $O(p_{R_{\mathrm{SFW}}})$ and, thus, the total number of oracle calls is upper bounded by

$$\begin{aligned}
&O(R_{\mathrm{SFW}} \cdot p_{R_{\mathrm{SFW}}}[d + \log(\frac{d}{\beta}) + \log(\Delta_{\max} R_{\mathrm{SFW}}/\xi)]) \\
&\quad \leq O(\frac{d^2}{\textsc{tol}^3 \beta^4 \psi^8}[d + \log(\frac{d}{\beta\xi}) + \log(\frac{d^2}{\textsc{tol}^3 \beta^4 \psi^8})] \\
&\quad = O(\frac{d^2}{\textsc{tol}^3 \beta^4 \psi^8}[d + \log(\frac{d}{\beta \xi \psi \textsc{tol}})]) \\
&\quad = \mathcal{A}(d, \beta, \psi, \textsc{tol}, 1/\delta, 1/\xi).
\end{aligned}$$

□

The following Lemma shows that the Multiplicative Weight Update algorithm (Algorithm 7) either finds an approximately feasible solution or if there is no approximately feasible solution, determines infeasibility.

Lemma 6. *Fix $\tau, \widehat{OPT} \geq 0$ and let $\delta \in (0, 1)$. Define*

$$P_\epsilon = \{\lambda \in \tilde{\triangle} : \mathbb{E}_\eta \left[\max_{x \in \mathcal{X}} \frac{(\bar{x} - x)^\top A_{\mathrm{semi}}(\lambda)^{-1/2} \eta}{\beta + \bar{\theta}^\top (\bar{x} - x)} \right] - \sqrt{\tau} C \leq \epsilon, \quad \tau \sum_{x \in \mathcal{X}} \bar{\theta}^\top (\bar{x} - x) \lambda_x - \widehat{OPT} \leq \epsilon\}$$

With probability at least $1 - \delta - \frac{1}{2^d M}$, if $MW(\tau, \widehat{OPT})$ does not declare infeasibility, then $MW(\tau, \widehat{OPT})$ returns $\bar{\lambda} \in P_{4\textsc{tol}}$ and if $MW(\tau, \widehat{OPT})$ declares infeasibility, then P_0 is infeasible. Furthermore, on the same event, $MW(\tau, \widehat{OPT})$ uses at most $\mathcal{C}(d, \beta, \psi, \textsc{tol}, 1/\delta, 1/\xi)$ linear maximization oracle calls.

Proof. The algorithm uses the Plotkin-Shmoys-Tardos reduction to online learning and essentially runs the multiplicative weights update algorithm (see Arora et al. [2012]) where there is an expert for each constraint. Define

$$\begin{aligned}
h_1(\lambda) &:= \bar{\tau} \sum_{x \in \mathcal{X}} \bar{\theta}^\top (\bar{x} - x) \lambda_x - \widehat{OPT} \\
h_2(\lambda) &:= \mathbb{E}_\eta \left[\max_{x \in \mathcal{X}} \frac{(\bar{x} - x)^\top A_{\mathrm{semi}}(\lambda)^{-1/2} \eta}{\beta + \bar{\theta}^\top (\bar{x} - x)} \right] - \sqrt{\bar{\tau}} C \\
h^{(r)}(\lambda) &:= p_1^{(r)} h_1(\lambda) + p_2^{(r)} h_2(\lambda).
\end{aligned}$$

At each round r, the algorithm chooses a distribution, $p_1^{(r)}$ and $p_2^{(r)}$, over the constraints and the adversary uses the stochastic Frank-Wolfe algorithm to find $\lambda^{(r)}$ such that

$$h^{(r)}(\lambda^{(r)}) \leq \min_{\lambda \in \tilde{\triangle}} h^{(r)}(\lambda) + \textsc{tol}.$$

The reward for expert/constraint 1 is $h_1(\lambda^{(r)})$ and the reward for expert/constraint 2 is $\widehat{h}_2(\lambda^{(r)})$.

Let $\mathcal{E}_r$ denote the event that $\lambda^{(r)} = \text{SFW}(p_1^{(r)}, p_2^{(r)}, \frac{\delta}{2R})$ satisfies

$$h^{(r)}(\lambda^{(r)}) \leq \min_{\lambda \in \tilde{\triangle}} h^{(r)}(\lambda) + \text{TOL}.$$

uses at most $\mathcal{A}(d, \beta, \psi, \text{TOL}, 1/\delta, 1/\xi)$ linear maximization oracle calls. Define $\mathcal{E} = \cap_r \mathcal{E}_r$ Further, define the following events

$$\mathcal{F}_r = \{|\text{estSup}(\lambda^{(r)},, \frac{\delta}{2R}) - \mathbb{E}_\eta \left[\max_{x \in \mathcal{X}} \frac{(\bar{x}-x)^\top A_{\text{semi}}(\lambda^{(r)})^{-1/2}\eta}{\beta + \bar{\theta}^\top(\bar{x}-x)}\right] | \leq \text{TOL}$$
$$\text{and estSup uses } \mathcal{B}(d, \beta, \psi, \text{TOL}, 1/\delta, 1/\xi) \text{ oracle calls}\}$$
$$\mathcal{F} = \cap_r \mathcal{F}_r$$

By Lemmas 5 and 8 applied with $\xi = \frac{1}{RM}$ and the law of total probability, we have that

$$\Pr(\mathcal{E}^c \cup \mathcal{F}^c) \leq \sum_{r=1}^R \Pr(\mathcal{E}_r^c \cup \mathcal{F}_r^c | \cap_{s=1}^{r-1} \mathcal{E}_s \cap \mathcal{F}_s) \leq \sum_{r=1}^R \frac{\delta}{R} + \frac{1}{2^d RM} = \delta + \frac{1}{2^d M}$$

Now, for the remainder of the proof we assume that $\mathcal{E} \cap \mathcal{F}$ occurs.

Suppose that at some round $r \in [R]$ Algorithm 8 returns $\lambda^{(r)}$ such that $\widehat{h}^{(r)}(\lambda^{(r)}) > 2\text{TOL}$. Then, since $\mathcal{F}$ implies that

$$|h^{(r)}(\lambda^{(r)}) - \widehat{h}^{(r)}(\lambda^{(r)})| \leq |\text{estSup}(\lambda^{(r)},, \frac{\delta}{2R}) - \mathbb{E}_\eta \left[\max_{x \in \mathcal{X}} \frac{(\bar{x}-x)^\top A_{\text{semi}}(\lambda^{(r)})^{-1/2}\eta}{\beta + \bar{\theta}^\top(\bar{x}-x)}\right] | \leq \text{TOL}$$

we have that on $\mathcal{E} \cap \mathcal{F}$

$$2\text{TOL} < \widehat{h}^{(r)}(\lambda^{(r)}) \leq \text{TOL} + h^{(r)}(\lambda^{(r)}) \leq \min_{\lambda \in \tilde{\triangle}} h^{(r)}(\lambda) + 2\text{TOL}.$$

Therefore, it follows that for every $\lambda \in \tilde{\triangle}$,

$$\max(\mathbb{E}_\eta \left[\max_{x \in \mathcal{X}} \frac{(\bar{x}-x)^\top A_{\text{semi}}(\lambda)^{-1/2}\eta}{\beta + \bar{\theta}^\top(\bar{x}-x)}\right] - \sqrt{\tau}C, \tau \sum_{x \in \mathcal{X}} \bar{\theta}^\top(\bar{x}-x)\lambda_x - \widehat{OPT}) > 0.$$

Thus, the algorithm correctly declares infeasibility of the convex feasibility program.

Next, suppose that the Algorithm 8 returns $\lambda^{(r)}$ such that $\widehat{h}^{(r)}(\lambda^{(r)}) \leq 2\text{TOL}$ at every round r. Then, we show that the algorithm returns $\bar{\lambda}^{(R)} \in P_{4\text{TOL}}$. To apply Theorem 9, a standard result for the multiplicative weights update algorithm, we must show that for any $\lambda^{(r)} \in \tilde{\triangle}$ returned during the execution of the Algorithm

$$\max(h_1(\lambda^{(r)}), \widehat{h}_2(\lambda^{(r)})) \leq \rho = \max(2dT, c\frac{d}{\beta\psi^{1/2}}) \tag{22}$$

where ρ is defined in Algorithm 7. We have that

$$h_1(\lambda^{(r)}) := \bar{\tau} \sum_{x \in \mathcal{X}} \bar{\theta}^\top(\bar{x}-x)\lambda_x^{(r)} - \widehat{OPT} \leq 2dT$$

since $\bar{\tau} \leq T$, $\widehat{OPT} \geq 0$, and we assume that $\bar{\theta}^\top(\bar{x}-x) \leq 2d$. Furthermore,

$$\begin{aligned}
\widehat{h}_2(\lambda^{(r)}) &= \text{estSup}(\lambda^{(r)}, \frac{\delta}{3R}) - \sqrt{\tau}C \\
&\leq \mathbb{E}_\eta \left[\max_{x \in \mathcal{X}} \frac{(\bar{x}-x)^\top A_{\text{semi}}(\lambda^{(r)})^{-1/2}\eta}{\beta + \bar{\theta}^\top(\bar{x}-x)}\right] | + \text{TOL} - \sqrt{\tau}C \\
&\leq \frac{1}{\beta\psi^{1/2}}\mathbb{E}[\sum_{i=1}^d |\eta_i|] \\
&\leq c\frac{d}{\beta\psi^{1/2}}
\end{aligned}$$

for a suitably chosen constant $c > 0$ where we used that fact that $\textsc{tol} = \frac{(\sqrt{2}-1)C}{4}$.

Thus, we have shown (22) and therefore may apply Theorem 9, which implies on $\mathcal{E} \cap \mathcal{F}$ that

$$\begin{aligned}
\frac{\sum_r \mathbb{E}_\eta\left[\max_{x\in\mathcal{X}} \frac{(\bar{x}-x)^\top A_{\text{semi}}(\lambda^{(r)})^{-1/2}\eta}{\beta+\bar{\theta}^\top(\bar{x}-x)}\right] - \sqrt{\tau}C}{R} &= \frac{\sum_r h_2(\lambda^{(r)})}{R} \\
&\leq \frac{\sum_r \widehat{h}_2(\lambda^{(r)}) + \textsc{tol}}{R} \\
&\leq 2\textsc{tol} + \frac{\sum_r \widehat{h}^{(r)}(\lambda^{(r)})}{R} \\
&\leq 4\textsc{tol}
\end{aligned}$$

Now, finally, applying Lemma 7, we have that

$$\mathbb{E}_\eta\left[\max_{x\in\mathcal{X}} \frac{(\bar{x}-x)^\top A_{\text{semi}}(\frac{1}{T}\sum_t \lambda^{(r)})^{-1/2}\eta}{\beta+\bar{\theta}^\top(\bar{x}-x)}\right] - \sqrt{\tau}C \leq \frac{\sum_r \mathbb{E}_\eta\left[\max_{x\in\mathcal{X}} \frac{(\bar{x}-x)^\top A_{\text{semi}}(\lambda^{(r)})^{-1/2}\eta}{\beta+\bar{\theta}^\top(\bar{x}-x)}\right] - \sqrt{\tau}C}{R} \leq 4\textsc{tol}$$

This shows that $\bar{\lambda}^{(R)}$ approximately satisfies one of the constraints; showing approximate satisfaction of the other constraint follows by a similar argument. Thus, we conclude that $\bar{\lambda}^{(R)} \in P_{4\textsc{tol}}$. □

The following Lemma shows that Algorithm 6 approximately solves the optimization problem (20).

Lemma 7. *Fix $\tau \in> 0$ and let $\delta \in (0,1)$. Let $\widetilde{\text{OPT}}_\tau$ be the value of*

$$\begin{aligned}
&\min_{\lambda\in\triangle} \tau \sum_{x\in\mathcal{X}} [\beta + \bar{\theta}^\top(\bar{x}-x)]\lambda_x \\
&\quad s.t.\ \mathbb{E}_\eta\left[\max_{x\in\mathcal{X}} \frac{(\bar{x}-x)^\top A_{\text{semi}}(\lambda)^{-1/2}\eta}{\beta+\bar{\theta}^\top(\bar{x}-x)}\right] \leq \sqrt{\tau}C.
\end{aligned}$$

If for all $\lambda \in \tilde{\triangle}$,

$$\mathbb{E}_\eta\left[\max_{x\in\mathcal{X}} \frac{(\bar{x}-x)^\top A_{\text{semi}}(\lambda)^{-1/2}\eta}{\beta+\bar{\theta}^\top(\bar{x}-x)}\right] > \sqrt{\tau}C + 4\textsc{tol}.$$

then with probability at least $1-\delta-\frac{1}{\log_2(T)2^d}$ Algorithm 6 declares the program infeasible. If

$$\mathbb{E}_\eta\left[\max_{x\in\mathcal{X}} \frac{(\bar{x}-x)^\top A_{\text{semi}}(\lambda)^{-1/2}\eta}{\beta+\bar{\theta}^\top(\bar{x}-x)}\right] \leq \sqrt{\tau}C$$

then with probability at least $1-\delta-\frac{1}{\log_2(T)2^d}$ Algorithm 6 returns $\bar{\lambda} \in \tilde{\triangle}$ such that

$$\begin{aligned}
\tau \sum_{x\in\mathcal{X}} \bar{\theta}^\top(\bar{x}-x)\bar{\lambda}_x &\leq \widetilde{\text{OPT}}_\tau + 4\textsc{tol} \\
\mathbb{E}_\eta\left[\max_{x\in\mathcal{X}} \frac{(\bar{x}-x)^\top A_{\text{semi}}(\lambda)^{-1/2}\eta}{\beta+\bar{\theta}^\top(\bar{x}-x)}\right] &\leq \sqrt{\tau}C + 4\textsc{tol}.
\end{aligned}$$

Furthermore, Algorithm 6 uses a number of oracle calls that is upper bounded by $\log_2(2Td/\textsc{tol}) \cdot \mathcal{C}(d,\beta,\psi,\textsc{tol},1/\delta,1/\xi)$.

Proof. Algorithm 6 applies Algorithm 7 at most $\log_2(2Td/\textsc{tol})$ times on a using a predetermined set of values for $\widehat{OPT} \in [0, 2Td]$, which we denote $\widehat{OPT}_1, \ldots, \widehat{OPT}_l$. Define the event

$$\begin{aligned}
\mathcal{E}_i = \{&\text{if MW}(\tau, \widehat{OPT}_i) \text{ does not declare infeasibility, then MW}(\tau, \widehat{OPT}_i) \text{ returns } \bar{\lambda} \in P_{4\textsc{tol}} \\
&\text{and if MW}(\tau, \widehat{OPT}_i) \text{ declares infeasibility,} P_0 \text{ is infeasible.}\} \\
&\cap \{\text{MW}(\tau, \widehat{OPT}_i) \text{ uses at most } \mathcal{C}(d,\beta,\psi,\textsc{tol},1/\delta,1/\xi) \text{ oracle calls }\} \\
\mathcal{E} = \cap_i &\mathcal{E}_i.
\end{aligned}$$

where P_ϵ is defined in Lemma 6. Then, by the union bound, we have that $\Pr(\mathcal{E}) \geq 1 - \delta - \frac{1}{2^d \log_2(T)}$. Suppose $\mathcal{E}$ occurs for the remainder of the proof.

First, consider the case that for all $\lambda \in \tilde{\triangle}$,

$$\mathbb{E}_\eta \left[\max_{x \in \mathcal{X}} \frac{(\bar{x} - x)^\top A_{\text{semi}}(\lambda)^{-1/2} \eta}{\beta + \bar{\theta}^\top (\bar{x} - x)} \right] > \sqrt{\tau} C + 4\text{TOL}.$$

Then, on the event $\mathcal{E}$, we have that the Algorithm 6 declares infeasibility of the program.

Now, suppose there exists $\lambda \in \tilde{\triangle}$ such that

$$\mathbb{E}_\eta \left[\max_{x \in \mathcal{X}} \frac{(\bar{x} - x)^\top A_{\text{semi}}(\lambda)^{-1/2} \eta}{\beta + \bar{\theta}^\top (\bar{x} - x)} \right] \leq \sqrt{\tau} C.$$

Note that for any $\lambda \in \triangle$, we have that

$$\tau \sum_{x \in \mathcal{X}} [\beta + \bar{\theta}^\top (\bar{x} - x)] \lambda_x = \beta \tau + \sum_{x \in \mathcal{X}} \bar{\theta}^\top (\bar{x} - x) \lambda_x$$

and thus the objective does not depend on β and β can be dropped from the objective. Using the event $\mathcal{E}$, if

$$\begin{aligned} Q(\widehat{OPT}) := \{\lambda \in \tilde{\triangle} : & \mathbb{E}_\eta \left[\max_{x \in \mathcal{X}} \frac{(\bar{x} - x)^\top A_{\text{semi}}(\lambda)^{-1/2} \eta}{\beta + \bar{\theta}^\top (\bar{x} - x)} \right] - \sqrt{\tau} C \leq 4\text{TOL}, \\ & \tau \sum_{x \in \mathcal{X}} \bar{\theta}^\top (\bar{x} - x) \lambda_x - \widehat{OPT} \leq 4\text{TOL}\} \end{aligned}$$

is empty, then Algorithm 7 declares the program infeasible; otherwise, Algorithm 7 finds $\bar{\lambda} \in Q(\widehat{OPT})$. Then, by a standard binary search argument, the result follows. □

The following Theorem establishes that Algorithm 5 approximately solves the main optimization problem (5). It directly implies Theorem 4.

Theorem 8. *Let $\delta \in (0, 1)$. Suppose* $\text{TOL} = \frac{(\sqrt{2}-1)C}{4}$, $\psi = \min(\frac{1}{4d\Delta_{\max} T}, \frac{1}{4d})$. *Let* OPT *be the value of*

$$\begin{aligned} \min_{\tau \in [T], \lambda \in \triangle} \tau & \sum_{x \in \mathcal{X}} [\epsilon + \bar{\theta}^\top (\bar{x} - x)] \lambda_x \\ \text{s.t. } & \mathbb{E}_\eta \left[\max_{x \in \mathcal{X}} \frac{(\bar{x} - x)^\top A_{\text{semi}}(\lambda)^{-1/2} \eta}{\beta + \bar{\theta}^\top (\bar{x} - x)} \right] \leq \sqrt{\tau} C. \end{aligned} \tag{23}$$

With probability at least $1 - \delta - \frac{1}{2^d}$, Algorithm 5 returns $(\bar{\tau}, \bar{\lambda})$ such that $\bar{\lambda} \in \triangle$, $\bar{\tau} \leq 2T$, and

$$\begin{aligned} \bar{\tau} \sum_{x \in \mathcal{X}} [\epsilon + \bar{\theta}^\top (\bar{x} - x)] \bar{\lambda}_x \leq 4\text{OPT} + 2 \\ \mathbb{E}_\eta \left[\max_{x \in \mathcal{X}} \frac{(\bar{x} - x)^\top A_{\text{semi}}(\bar{\lambda})^{-1/2} \eta}{\beta + \bar{\theta}^\top (\bar{x} - x)} \right] \leq \sqrt{\bar{\tau}} C. \end{aligned}$$

Furthermore, Algorithm 5 uses a number of oracle calls that is polynomial in $(d, \beta, \psi, \log(1/\delta))$

Proof. **Step 0.** Let $\widetilde{\text{OPT}}$ be the value of

$$\begin{aligned} \min_{\tau \in [T], \lambda \in \tilde{\triangle}} \tau & \sum_{x \in \mathcal{X}} [\beta + \bar{\theta}^\top (\bar{x} - x)] \lambda_x \\ \text{s.t. } & \mathbb{E}_\eta \left[\max_{x \in \mathcal{X}} \frac{(\bar{x} - x)^\top A_{\text{semi}}(\lambda)^{-1/2} \eta}{\beta + \bar{\theta}^\top (\bar{x} - x)} \right] \leq \sqrt{\tau} C. \end{aligned}$$

and let $\widetilde{\text{OPT}}_k$ be the value of

$$\min_{\lambda \in \triangle} \bar{\tau}_k \sum_{x \in \mathcal{X}} [\beta + \bar{\theta}^\top (\bar{x} - x)] \lambda_x$$
$$\text{s.t. } \mathbb{E}_\eta \left[\max_{x \in \mathcal{X}} \frac{(\bar{x} - x)^\top A_{\text{semi}}(\lambda)^{-1/2} \eta}{\beta + \bar{\theta}^\top (\bar{x} - x)} \right] \leq \sqrt{\bar{\tau}}_k C.$$

Let $\mathcal{E}_k$ denote the event that if for all $\lambda \in \tilde{\triangle}$,

$$\mathbb{E}_\eta \left[\max_{x \in \mathcal{X}} \frac{(\bar{x} - x)^\top A_{\text{semi}}(\lambda)^{-1/2} \eta}{\beta + \bar{\theta}^\top (\bar{x} - x)} \right] > \sqrt{\bar{\tau}_k} C + 4\text{TOL}.$$

then binSearch$(\bar{\tau}_k, \frac{\delta}{\log_2(T)})$ declares the program infeasible and if

$$\mathbb{E}_\eta \left[\max_{x \in \mathcal{X}} \frac{(\bar{x} - x)^\top A_{\text{semi}}(\lambda)^{-1/2} \eta}{\beta + \bar{\theta}^\top (\bar{x} - x)} \right] \leq \sqrt{\tau} C$$

then binSearch$(\bar{\tau}_k, \frac{\delta}{\log_2(T)})$ returns $\bar{\lambda}_k$ that satisfies

$$\tau \sum_{x \in \mathcal{X}} [\beta + \bar{\theta}^\top (\bar{x} - x) \bar{\lambda}_{k,x}] \leq \widetilde{\text{OPT}}_{\bar{\tau}_k} + 4\text{TOL}$$
$$\mathbb{E}_\eta \left[\max_{x \in \mathcal{X}} \frac{(\bar{x} - x)^\top A_{\text{semi}}(\bar{\lambda}_k)^{-1/2} \eta}{\beta + \bar{\theta}^\top (\bar{x} - x)} \right] \leq \sqrt{\tau} C + 4\text{TOL}.$$

Further, define $\mathcal{E} = \cap_k \mathcal{E}_k$. By Lemma 7 and a union bound, we have that $\Pr(\mathcal{E}) \geq 1 - \delta - \frac{1}{2^d}$. We suppose $\mathcal{E}$ holds for the rest of the proof.

Step 1. First, we show that Algorithm 5 returns $(\bar{\tau}, \bar{\lambda})$ such that

$$\bar{\tau} \sum_{x \in \mathcal{X}} [\beta + \bar{\theta}^\top (\bar{x} - x)] \bar{\lambda}_x \leq \widetilde{\text{OPT}} + 4\text{TOL}$$
$$\mathbb{E}_\eta \left[\max_{x \in \mathcal{X}} \frac{(\bar{x} - x)^\top A_{\text{semi}}(\bar{\lambda})^{-1/2} \eta}{\beta + \bar{\theta}^\top (\bar{x} - x)} \right] \leq \sqrt{\bar{\tau}} C.$$

By assumption the optimization problem in (5) is feasible and, hence, $\text{OPT} \neq \infty$ and thus by the event $\mathcal{E}$, the algorithm finds at least one nearly feasible solution, i.e., FEASIBLE_k is not False for all k. Let (τ_*, λ_*) attain the optimal value in the optimization problem (5). Let k_* such that $\bar{\tau}_{k_*} \in [\tau_*, 2\tau_*]$. By event $\mathcal{E}$ binSearch$(\bar{\tau}_{k_*}, \frac{\delta}{\log_2(T)})$ finds $\bar{\lambda}_{k_*}$ such that

$$\bar{\tau}_{k_*} \sum_{x \in \mathcal{X}} [\beta + \bar{\theta}^\top (\bar{x} - x)] \bar{\lambda}_{k_*,x} \leq \widetilde{\text{OPT}}_{k_*} + 4\text{TOL}$$
$$\mathbb{E}_\eta \left[\max_{x \in \mathcal{X}} \frac{(\bar{x} - x)^\top A_{\text{semi}}(\bar{\lambda}_{k_*})^{-1/2} \eta}{\beta + \bar{\theta}^\top (\bar{x} - x)} \right] \leq \sqrt{\bar{\tau}_{k_*}} C + 4\text{TOL}.$$

Algorithm 5 outputs $(\bar{\tau}, \lambda_{\widehat{k}_*})$, which satisfies by Lemma 7 and by construction,

$$\begin{aligned} \bar{\tau} \sum_{x \in \mathcal{X}} [\beta + \bar{\theta}^\top (\bar{x} - x) \bar{\lambda}_{\widehat{k}_*,x}] &= 2\bar{\tau}_{\widehat{k}_*} \sum_{x \in \mathcal{X}} [\beta + \bar{\theta}^\top (\bar{x} - x) \bar{\lambda}_{\widehat{k}_*,x}] \quad (24) \\ &\leq 2\bar{\tau}_{k_*} \sum_{x \in \mathcal{X}} [\beta + \bar{\theta}^\top (\bar{x} - x)] \bar{\lambda}_{k_*,x} \\ &\leq 2[\widetilde{\text{OPT}}_k + 4\text{TOL}] \\ &\leq 2\widetilde{\text{OPT}}_k + 1 \quad (25) \end{aligned}$$

where in the last line we used $\text{TOL} = \frac{(\sqrt{2}-1)C}{4} \leq 1/8$, which bounds the objective value of $(\bar{\tau}, \lambda_{\widehat{k}_*})$.

Next, we show feasiblity of $(\bar{\tau}, \lambda_{\widehat{k}_*})$. Observe that

$$\begin{aligned}\mathbb{E}_\eta\left[\max_{x\in\mathcal{X}}\frac{(\bar{x}-x)^\top A_{\text{semi}}(\bar{\lambda}_{k_*})^{-1/2}\eta}{\beta+\bar{\theta}^\top(\bar{x}-x)}\right] &\leq \sqrt{\bar{\tau}_{\widehat{k}_*}}C + 4\textsc{tol} \\ &\leq \sqrt{2\bar{\tau}_{\widehat{k}_*}}C \\ &= \sqrt{\bar{\tau}}C \end{aligned} \tag{26}$$

where we used the fact that $\bar{\tau} = 2\bar{\tau}_{\widehat{k}_*}$ and $\textsc{tol} = \frac{(\sqrt{2}-1)C}{4}$.

Step 2: Relate $\widetilde{\text{opt}}_k$ to $\widetilde{\text{opt}}$. Next, we show that

$$\widetilde{\textsc{opt}}_{k_*} \leq 2\widetilde{\textsc{opt}}.$$

Define the function

$$\begin{aligned} f(\lambda,\tau) =& \tau\sum_{x\in\mathcal{X}}[\beta+\bar{\theta}^\top(\bar{x}-x)]\lambda_x \\ &\text{s.t. } \mathbb{E}_\eta\left[\max_{x\in\mathcal{X}}\frac{(\bar{x}-x)^\top A_{\text{semi}}(\lambda)^{-1/2}\eta}{\beta+\bar{\theta}^\top(\bar{x}-x)}\right] \leq \sqrt{\tau}C. \end{aligned}$$

Recall that we let (τ_*, λ_*) attain the optimal value in the optimization problem (5). Let k_* such that $\bar{\tau}_{k_*} \in [\tau_*, 2\tau_*]$. Note that

$$\mathbb{E}_\eta\left[\max_{x\in\mathcal{X}}\frac{(\bar{x}-x)^\top A_{\text{semi}}(\lambda_*)^{-1/2}\eta}{\beta+\bar{\theta}^\top(\bar{x}-x)}\right] \leq \sqrt{\bar{\tau}_{k_*}}C$$

Thus,

$$\widetilde{\textsc{opt}}_{k_*} = f(\bar{\lambda}_k, \bar{\tau}_k) \leq f(\lambda_*, \bar{\tau}_k) \leq 2f(\lambda_*, \tau_*) = 2\widetilde{\textsc{opt}}, \tag{27}$$

where we used the fact that

$$\mathbb{E}_\eta\left[\max_{x\in\mathcal{X}}\frac{(\bar{x}-x)^\top A_{\text{semi}}(\lambda_*)^{-1/2}\eta}{\beta+\bar{\theta}^\top(\bar{x}-x)}\right] \leq \sqrt{\tau_*}C \leq \sqrt{\bar{\tau}_{k_*}}C.$$

This proves the claim.

Step 3: Relate $\widetilde{\text{opt}}$ to opt. Next, we show that

$$\widetilde{\textsc{opt}} \leq 2\textsc{opt} + T\psi d\Delta_{\max}.$$

Define

$$\check{\lambda}_i = \begin{cases} \frac{1}{d} : i \in [d] \\ 0 : i \not\in [d] \end{cases}.$$

and

$$\begin{aligned}\tilde{\lambda} &= \psi d\check{\lambda} + (1-\psi d)\lambda^* \\ \tilde{\tau} &= 2\tau_*.\end{aligned}$$

By the hypothesis, we have that $\psi d \leq \frac{1}{4}$ and, thus, $\tilde{\lambda}$ is a convex combination of $\check{\lambda}$ and λ^*.

Next, we show that $(\tilde{\lambda}, \tilde{\tau})$ are a feasible solution to (23) and show that it is approximately optimal. Note that

$$A_{\text{semi}}(\tilde{\lambda}) \geq (1-\psi d)A_{\text{semi}}(\lambda^*),$$

which implies that

$$\frac{1}{1-\psi d} A_{\mathrm{semi}}(\lambda^*)^{-1} \geq A_{\mathrm{semi}}(\tilde{\lambda})^{-1}.$$

Then, by Sudakov-Fernique, we have that

$$\begin{aligned}
\mathbb{E}_\eta\left[\max_{x\in\mathcal{X}} \frac{(\bar{x}-x)^\top A_{\mathrm{semi}}(\tilde{\lambda})^{-1/2}\eta}{\beta+\bar{\theta}^\top(\bar{x}-x)}\right] &\leq [1-\psi d]^{-1/2}\mathbb{E}_\eta\left[\max_{x\in\mathcal{X}} \frac{(\bar{x}-x)^\top A_{\mathrm{semi}}(\lambda^*)^{-1/2}\eta}{\beta+\bar{\theta}^\top(\bar{x}-x)}\right] \\
&\leq \sqrt{2}\mathbb{E}_\eta\left[\max_{x\in\mathcal{X}} \frac{(\bar{x}-x)^\top A_{\mathrm{semi}}(\lambda^*)^{-1/2}\eta}{\beta+\bar{\theta}^\top(\bar{x}-x)}\right] \\
&\leq \sqrt{2\tau^*}C \\
&= \sqrt{\tilde{\tau}}C
\end{aligned}$$

showing feasibility $(\tilde{\lambda}, \tilde{\tau})$. Furthermore, we have that

$$\begin{aligned}
\bar{\tau}\sum_{x\in\mathcal{X}} \tilde{\lambda}_x[\bar{\theta}^\top(\bar{x}-x)+\beta] &\leq 2\textsc{opt} + \tilde{\tau}\psi d \sum_{x\in\mathcal{X}} \tilde{\lambda}_x[\bar{\theta}^\top(\bar{x}-x)+\beta] \\
&\leq 2\textsc{opt} + T\psi d 2\Delta_{\max} \\
&\leq 2\textsc{opt} + 1 \qquad (28)
\end{aligned}$$

where in the last line we used $\psi = \min(\frac{1}{4d\Delta_{\max}T}, \frac{1}{4d})$.

Step 4: Putting it together. Putting together (26), (25), (27), and (28), we have that Algorithm 5 returns $(\bar{\tau}, \bar{\lambda})$ such that $\bar{\lambda} \in \triangle$, $\bar{\tau} \leq 2T$, and

$$\begin{aligned}
\bar{\tau}\sum_{x\in\mathcal{X}}[\beta+\bar{\theta}^\top(\bar{x}-x)]\bar{\lambda}_x &\leq 4\textsc{opt} + 2 \\
\mathbb{E}_\eta\left[\max_{x\in\mathcal{X}} \frac{(\bar{x}-x)^\top A_{\mathrm{semi}}(\bar{\lambda})^{-1/2}\eta}{\beta+\bar{\theta}^\top(\bar{x}-x)}\right] &\leq \sqrt{\bar{\tau}}C.
\end{aligned}$$

□

D.3 Miscellaneous Optimization Lemmas

Lemma 8. *Let $\lambda \in \tilde{\triangle}$. With probability at least $1-\delta-\frac{\xi}{2^d}$, Algorithm 10 returns $\widehat{\mu}$ such that*

$$|\widehat{\mu} - \mathbb{E}_\eta\left[\max_{x\in\mathcal{X}} \frac{(\bar{x}-x)^\top A_{\mathrm{semi}}(\lambda)^{-1/2}\eta}{\beta+\bar{\theta}^\top(\bar{x}-x)}\right]| \leq \textsc{tol}$$

and the number of linear maximization oracle calls is bounded above by

$$O(\log(1/\delta)\frac{d}{\beta^2\psi\textsc{tol}^2}[d+\log(\frac{d\Delta_{\max}}{\beta\xi})]).$$

Proof. We first show that

$$\max_{x\in\mathcal{X}} \frac{(\bar{x}-x)^\top A_{\mathrm{semi}}(\lambda)^{-1/2}\eta}{\beta+\bar{\theta}^\top(\bar{x}-x)} \in \mathcal{SG}(c\frac{d}{\beta^2\psi}).$$

Note that

$$|\max_{x\in\mathcal{X}} \frac{(\bar{x}-x)^\top A_{\mathrm{semi}}(\lambda)^{-1/2}\eta}{\beta+\bar{\theta}^\top(\bar{x}-x)}| \leq \frac{1}{\beta\psi^{1/2}}\sum_{i=1}^d |\eta_i|$$

and

$$\frac{1}{\beta\psi^{1/2}}\sum_{i=1}^{d}|\eta_i| \in \mathcal{SG}(c\frac{d}{\beta^2\psi}).$$

The estimation results by applying a standard subGaussian tail bound. The bound on the number of oracle calls follows since Algorithm 9 is applied $O(\log(1/\delta)\frac{d}{\beta^2\psi\text{TOL}^2})$ times and by Lemma 9 and a union bound. $\square$

The following Lemma shows that the binary search procedure in Algorithm 9 is efficient with very high probability and it follows immediately from the proof of Lemma 2 of Katz-Samuels et al. [2020].

Lemma 9. *Draw $\eta \sim N(0, I)$ and consider the optimization problem*

$$\tilde{x} = \arg\max_{x\in\mathcal{X}} \frac{(\bar{x}-x)^\top A_{\text{semi}}(\lambda)^{-1/2}\eta}{\beta + \bar{\theta}^\top(\bar{x}-x)}.$$

With probability at least $1-\frac{2\xi}{2^d}$, Algorithm 9 returns $\tilde{x}$ using at most $O(d+\log(\frac{d\Delta_{\max}}{\beta\xi}))$ oracle calls.

Next, we describe a result on the multiplicative weights update algorithm that follows immediately from Corollary 4 in Arora et al. [2012]. Consider the experts problem. The set of events is denoted by P. Suppose there are m experts. At each round t, the agent picks an expert $i \in [m]$ and the adversary picks an outcome $j^t \in P$ and the agent obtains reward $M(i, j^t)$. The multiplicative weights update algorithm mains a distribution D^t over the experts and chooses an expert randomly from D^t (see Arora et al. [2012] for details on how this distribution is chosen). The adversary may have knowledge of the D^t when choosing j^t. The following provides a lower bound on the expected reward obtained by the multiplicative weights update algorithm.

Theorem 9. *Let $\xi > 0$ denote an error parameter. Suppose there are m experts and $|M(i,j)| \leq \rho$. If the multiplicative weights algorithm sets the learning rate as $\epsilon = \min(\frac{\xi}{4\rho}, \frac{1}{2})$, after $T = \frac{16\rho^2\ln(m)}{\xi^2}$, then the multiplicative weights algorithm achieves the following bound on its average expected reward: for any expert i,*

$$\frac{\sum_t M(i,j^t)}{T} \leq \xi + \frac{\sum_t M(D^t, j^t)}{T}.$$

D.4 Convergence Lemmas

The objective in semi-feedback is convex (by a similar argument to the proof in Katz-Samuels et al. [2020]).

Proposition 7. *Fix $V \subset \mathbb{R}^d$.*

$$f(\lambda) = \mathbb{E}_{\eta\sim N(0,I)}[\max_{v\in V} v^\top A_{\text{semi}}(\lambda)^{-1/2}\eta]$$

is convex.

Proof. Fix $\lambda, \kappa \in \triangle^{|\mathcal{X}|}$ and $\alpha \in [0,1]$. By matrix convexity,

$$\text{diag}(\frac{1}{\sum_{x\in\mathcal{X}}\alpha\lambda_{x,i}+(1-\alpha)\kappa_{x,i}})^{1/2} \preceq \alpha\,\text{diag}(\frac{1}{\sum_{x\in\mathcal{X}}\lambda_{x,i}})^{1/2} + (1-\alpha)\,\text{diag}(\frac{1}{\sum_{x\in\mathcal{X}}\kappa_{x,i}})^{1/2}.$$

Furthermore, since the above matrices are diagonal,

$$\text{diag}(\frac{1}{\sum_{x\in\mathcal{X}}\alpha\lambda_{x,i}+(1-\alpha)\kappa_{x,i}}) \preceq [\alpha\,\text{diag}(\frac{1}{\sum_{x\in\mathcal{X}}\lambda_{x,i}})^{1/2} + (1-\alpha)\,\text{diag}(\frac{1}{\sum_{x\in\mathcal{X}}\kappa_{x,i}})^{1/2}]^2.$$

Then, by Sudakov-Fernique inequality (Theorem 7.2.11 in Vershynin [2018]),

$$\begin{aligned}
f(\alpha\lambda + (1-\alpha)\kappa) &= \mathbb{E}_{\eta \sim N(0, \mathrm{diag}(\frac{1}{\sum_{x \in \mathcal{X}} \alpha \lambda_{x,i} + (1-\alpha)\kappa_{x,i}}))} \sup_{v \in V} v^\top \eta \\
&\leq \mathbb{E}_{\eta \sim N(0, [\alpha \, \mathrm{diag}(\frac{1}{\sum_{x \in \mathcal{X}} \lambda_{x,i}})^{1/2} + (1-\alpha) \, \mathrm{diag}(\frac{1}{\sum_{x \in \mathcal{X}} \kappa_{x,i}})^{1/2}]^2)} \sup_{v \in V} z^\top \eta \\
&= \mathbb{E}_{\eta \sim N(0,I)} \sup_{v \in V} v^\top [\alpha \, \mathrm{diag}(\frac{1}{\sum_{x \in \mathcal{X}} \lambda_{x,i}})^{1/2} + (1-\alpha) \, \mathrm{diag}(\frac{1}{\sum_{x \in \mathcal{X}} \kappa_{x,i}})^{1/2}] \eta \\
&\leq \alpha \mathbb{E}_{\eta \sim N(0,I)} \sup_{v \in V} v^\top \, \mathrm{diag}(\frac{1}{\sum_{x \in \mathcal{X}} \lambda_{x,i}})^{1/2} \eta \\
&+ (1-\alpha) \mathbb{E}_{\eta \sim N(0,I)} \sup_{v \in V} v^\top \, \mathrm{diag}(\frac{1}{\sum_{x \in \mathcal{X}} \kappa_{x,i}})^{1/2} \eta \\
&= \alpha f(\lambda) + (1-\alpha) f(\kappa)
\end{aligned}$$

□

Next, we turn to analyzing stochastic Frank-Wolfe. Although a convergence result for stochastic frank wolfe is provided in Hazan and Luo [2016], our setup is slightly different, so we include a convergence analysis for our setting for the sake of completeness. The proof is quite similar to the proof in Hazan and Luo [2016].

Algorithm 11 Generic Stochastic Frank-Wolfe

1: **Input:** $f : \mathbb{R}^m \times \mathbb{R}^d \longrightarrow \mathbb{R}$, constraint set $\Omega \subset \mathbb{R}^m$, $(p_r)_r \in \mathbb{N}^\infty$, $(q_r)_r \in [0,1]^\infty$.
2: Initialize $w_1 \in \Omega$
3: **for** $r = 1, 2, \ldots$ **do**
4: Draw $\eta_1, \ldots, \eta_{p_r} \sim N(0, I)$
5: Compute

$$\tilde{\nabla}_r = \frac{1}{p_r} \sum_{j=1}^{p_r} \nabla f(w_r; \eta_j)$$

6: Compute

$$v_r = \arg\min_{v \in \Omega} \tilde{\nabla}_r^\top v$$

7:

$$w_{r+1} \longleftarrow q_r v_r + (1 - q_r) w_r$$

8: **end for**

Proposition 8. *Let $f : \mathbb{R}^m \times \mathbb{R}^d \longrightarrow \mathbb{R}$ and $\Omega \subset \mathbb{R}^m$. Define $f(x) = \mathbb{E}_{\eta \sim N(0,I)} f(x; \eta)$ and define*

$$w^* = \arg\min_{w \in \Omega} \mathbb{E}_\eta f(x; \eta).$$

Suppose that $\sup_{w, w' \in \Omega} \|w - w'\| \leq D$. Suppose that f is convex, $\|\nabla f(x) - \nabla f(y)\|_* \leq L \|x - y\|$, and p_r in Algorithm 11 is chosen such that with probability at least $1 - \delta / r^2$

$$\left\| \widetilde{\nabla}_r - \nabla f(w_{r-1}) \right\|_* \leq \frac{LDq_r}{2}$$

where $q_r = \frac{2}{k+1}$. Then, with probability at least $1 - c\delta$,

$$f(w_r) - f(w_*) \leq \frac{4LD^2}{r+2}.$$

Proof. The proof follows closely the analysis of SFW in Hazan and Luo [2016] but uses smoothness wrt $\|\cdot\|_*$. We have that

$$\begin{aligned}
f(w_r) &\leq f(w_{r-1}) + \nabla f(w_{r-1})^\top (w_r - w_{r-1}) + \frac{L}{2}\|w_r - w_{r-1}\|_1^2 && (29)\\
&= f(w_{r-1}) + q_r \nabla f(w_{r-1})^\top (v_r - w_{r-1}) + \frac{Lq_r^2}{2}\|v_r - w_{r-1}\|_1^2\\
&\leq f(w_{r-1}) + q_r \widetilde{\nabla}_r^\top (v_r - w_{r-1}) + q_r(\nabla f(w_{r-1}) - \widetilde{\nabla}_r)^\top (v_r - w_{r-1}) + \frac{LD^2q_r^2}{2}\\
&\leq f(w_{r-1}) + q_r \widetilde{\nabla}_r^\top (w_* - w_{r-1}) + q_r(\nabla f(w_{r-1}) - \widetilde{\nabla}_r)^\top (v_r - w_{r-1}) + \frac{LD^2q_r^2}{2} && (30)\\
&= f(w_{r-1}) + q_r \nabla f(w_{r-1})^\top (w_* - w_{r-1}) + q_r(\nabla f(w_{r-1}) - \widetilde{\nabla}_r)^\top (v_r - w_*) + \frac{LD^2q_r^2}{2}\\
&\leq f(w_{r-1}) + q_r \nabla f(w_{r-1})^\top (w_* - w_{r-1}) + q_r \left\|\nabla f(w_{r-1}) - \widetilde{\nabla}_r\right\|_* D + \frac{LD^2q_r^2}{2} - && (31)
\end{aligned}$$

where line (29) uses smoothness (Lemma 11), line (30) uses the optimality of v_r, and line (31) uses the definition of the dual norm. Now, define the event

$$\mathcal{E}_r = \{\left\|\widetilde{\nabla}_r - \nabla f(w_{r-1})\right\|_* \leq \frac{LDq_r}{2}\}.\mathcal{E} = \cap_r \mathcal{E}_r$$

By hypothesis, p_r is chosen such that with probability at least $1-\delta/r^2$, $\left\|\widetilde{\nabla}_r - \nabla f(w_{r-1})\right\|_* \leq \frac{LDq_r}{2}$. Therefore, we have that

$$\Pr(\mathcal{E}) = \prod_{r=1}^{\infty} \Pr(\mathcal{E}_r | \cap_{s=1}^{r-1} \mathcal{E}_s) \geq \prod_{r=1}^{\infty}(1 - \frac{\delta}{r^2}) = \frac{\sin(\pi\delta)}{\pi + \delta} \geq 1 - \delta.$$

Now, suppose $\mathcal{E}$ occurs. Then, we have that for all $r \in \mathbb{N}$,

$$f(w_r) - f(w_*) \leq (1 - q_r)[f(w_{r-1}) - f(w_*)] + LD^2q_r^2.$$

The proof is concluded by simple induction. □

The following Lemma shows that Algorithm 8 is an instantiation of stochastic Frank-Wolfe over $\tilde{\triangle}$.

Lemma 10. *Fix $v \in \mathbb{R}^m$. Let*

$$I := \arg\min_{i \in [m]} v_i.$$

Define

$$\bar{\lambda}_i = \begin{cases} \begin{cases} 0 & i \notin [d] \\ \psi & i \in [d] \setminus \{I\} \\ 1-(d-1)\psi & i = I \end{cases}, & I \in [d] \\ \begin{cases} \psi & i \in [d] \\ 1 - d\psi & i = I \end{cases}, & I \notin [d] \end{cases}$$

Then, $\bar{\lambda} \in \arg\min_{\lambda \in \tilde{\triangle}} v^\top \lambda$.

Proof. This follows by a straightforward case by case analysis. □

The following is standard smoothness Lemma from convex optimization.

Lemma 11. *Let $f : \mathbb{R}^m \longrightarrow \mathbb{R}$ satisfy $\|\nabla f(x) - \nabla f(y)\|_* \leq L\|x - y\|$. Then,*

$$f(x) - f(y) - \nabla f(y)^\top (-y) \leq \frac{L}{2}\|x - y\|^2.$$

Proof. This is standard (see Bubeck [2014]). □

D.5 Differentiability Lemmas

In this section, we show that $\mathcal{L}(\kappa_1, \kappa_2; \lambda)$ is twice-differentiable wrt λ. We set $\kappa_1, \kappa_2 = 1$ for simplicity and write $\mathcal{L}(\lambda)$ instead of $\mathcal{L}(\kappa_1, \kappa_2; \tau; \lambda)$ for the sake of brevity. The following Lemma shows that $\mathcal{L}(\kappa_1, \kappa_2; \tau; \lambda)$ is differentiable wrt λ.

Lemma 12. *Fix $i \in [m]$, and $\lambda \in \tilde{\triangle}$. Fix $\eta \in \mathbb{R}^d$ such there exists a neighborhood of η such that*

$$\tilde{x} = \arg\max_{x \in \mathcal{X}} \frac{\sum_{i \in \bar{x} \Delta x} \frac{\eta_i}{\sum_{x' : i \in x'} \lambda_{x'}}}{\beta + \bar{\theta}^\top (\bar{x} - x)}.$$

Then,

$$\frac{\partial \mathcal{L}(\lambda; \eta)}{\partial \lambda_i} = \tau \bar{\theta}^\top (\bar{x} - x_i) - \frac{1}{2} \frac{1}{[\beta + \bar{\theta}^\top (\bar{x} - \tilde{x})]} \sum_{k \in (\bar{x} \Delta \tilde{x}) \cap x_i} \frac{\eta_k}{(\sum_{j : k \in x_j} \lambda_j)^{3/2}}$$

Furthermore, $\mathcal{L}(\lambda)$ is differentiable at every $\lambda \in \tilde{\triangle}_\psi$ and

$$\frac{\partial \mathcal{L}(\lambda)}{\partial \lambda_i} = \mathbb{E}_{\eta \sim N(0,I)} [\frac{\partial \mathcal{L}(\lambda; \eta)}{\partial \lambda_i} \mathbb{1}\{B_\lambda\}]$$

where

$$B_\lambda = \{\eta : |\arg\max_{x \in \mathcal{X}} \frac{\sum_{i \in \bar{x} \Delta x} \frac{\eta_i}{\sum_{x' : i \in x'} \lambda_{x'}}}{\beta + \bar{\theta}^\top (\bar{x} - x)}| = 1\}.$$

Proof. The calculation of $\frac{\partial \mathcal{L}(\lambda; \eta)}{\partial \lambda_i}$ follows by the chain rule.

Fix $\lambda \in \tilde{\triangle}$. Since $\lambda \in \tilde{\triangle}$, we have that $A_{\mathrm{semi}}(\lambda)^{-1/2}$ is full rank.

Step 1: First, we show that $\mathcal{L}(\lambda; \eta)$ is Lipschitz with an absolutely integrable Lipschitz constant. Define

$$\mathcal{J}(\lambda; \eta; x) = \tau \sum_{x \in \mathcal{X}} \bar{\theta}^\top (\bar{x} - x) \lambda_x + (\frac{x^\top A_{\mathrm{semi}}(\lambda)^{-1/2} \eta}{\beta + \bar{\theta}^\top (\bar{x} - x)} - \sqrt{\tau} C).$$

and note that

$$\begin{aligned} |\frac{\partial \mathcal{J}(\lambda; \eta; x)}{\partial \lambda_i}| &= |\bar{\theta}^\top (\bar{x} - x_i) - \frac{1}{2} \frac{1}{[\beta + \bar{\theta}^\top (\bar{x} - x)]} \sum_{k \in (\bar{x} \Delta x) \cap x_i} \frac{\eta_k}{(\sum_{j : k \in x_j} \lambda_j)^{3/2}}| \\ &\leq |\bar{\theta}^\top (\bar{x} - x_i)| + \frac{1}{2} \frac{1}{[\beta + \bar{\theta}^\top (\bar{x} - x)]} \sum_{k \in (\bar{x} \Delta x) \cap x_i} \frac{|\eta_k|}{\psi^{3/2}} \\ &< |\bar{\theta}^\top (\bar{x} - x_i)| + \frac{1}{2} \frac{1}{\beta} \sum_{k \in (\bar{x} \Delta x) \cap x_i} \frac{|\eta_k|}{\psi^{3/2}} := C_\eta \end{aligned}$$

Let $\lambda, \lambda' \in \tilde{\triangle}$. Thus, by the mean value theorem, we have that for all $x \in \mathcal{X}$,

$$|\mathcal{J}(\lambda; \eta; x) - \mathcal{J}(\lambda; \eta; x)| \leq C_\eta \|\lambda - \lambda'\|_1$$

Since $\mathcal{L}(\lambda; \eta) := \max_{x \in \mathcal{X}} \mathcal{J}(\lambda; \eta; x)$ and the maximum of C_η-Lipschitz functions is C_η-Lipschitz, we have that

$$|\mathcal{L}(\lambda; \eta) - \mathcal{L}(\lambda; \eta)| \leq C_\eta \|\lambda - \lambda'\|_1$$

Step 2: Now, we show that the partial derivatives exist. Define the event

$$B_\lambda = \{\eta : |\arg\max_{x \in \mathcal{X}} \frac{(\bar{x} - x)^\top A_{\mathrm{semi}}(\lambda)^{-1/2} \eta}{\beta + \bar{\theta}^\top (\bar{x} - x)}| = 1\},$$

Since $A_{\text{semi}}(\lambda)^{-1/2}$ is full rank and each

$$\frac{x}{\beta + \bar{\theta}^\top(\bar{x} - x)}$$

is distinct, if $\eta \sim N(0, I)$, then with probability 1 B_λ holds and $\mathcal{L}(\lambda; \eta)$ is differentiable at λ.

Since $\mathcal{L}(\lambda; \eta)$ is C_η-Lipschitz (because $\lambda \in \tilde{\triangle}$, we have

$$|\frac{\mathcal{L}(\lambda + e_i h; \eta) - L(\lambda; \eta)}{h}| \leq C_\eta.$$

Since in addition $\mathbb{E}C_\eta < \infty$, by the dominated convergence theorem,

$$\begin{aligned}\lim_{h \longrightarrow 0} \mathbb{E}[\frac{\mathcal{L}(\lambda + e_i h; \eta) - L(\lambda; \eta)}{h}] &= \lim_{h \longrightarrow 0} \mathbb{E}[\frac{\mathcal{L}(\lambda + e_i h; \eta) - L(\lambda; \eta)}{h} \mathbb{1}\{B_\lambda\}] \\ &= \mathbb{E}[\lim_{h \longrightarrow 0} \frac{\mathcal{L}(\lambda + e_i h; \eta) - L(\lambda; \eta)}{h} \mathbb{1}\{B_\lambda\}] \\ &= \mathbb{E}[\nabla \mathcal{L}(\lambda; \eta)^\top e_i \mathbb{1}\{B_\lambda\}]\end{aligned}$$

where the last equality follows since on B_λ and $\lambda \in \tilde{\triangle}$, $\frac{\partial \mathcal{L}(\lambda;\eta)}{\partial \lambda_i}$ exists. Thus, the partial derivative $\frac{\partial \mathcal{L}(\lambda)}{\partial \lambda_i}$ exists at every point $\lambda \in \tilde{\triangle}$ and

$$\frac{\partial \mathcal{L}(\lambda)}{\partial \lambda_i} = \mathbb{E}[\nabla \mathcal{L}(\lambda; \eta)^\top e_i \mathbb{1}\{B_\lambda\}]$$

Step 3: We claim that the partial derivative is continuous in $\lambda \in \tilde{\triangle}$, which would show that that $\mathcal{L}(\lambda)$ is differentiable at every $\lambda \in \tilde{\triangle}$ Munkres [2018]. Let $\lambda^{(n)}$ be a sequence in $\tilde{\triangle}$ such that $\lambda^{(n)} \longrightarrow \lambda$. Note that since $\lambda^{(n)} \in \tilde{\triangle}$, we have that

$$\nabla \mathcal{L}(\lambda^{(n)}; \eta)^\top e_i \mathbb{1}\{B_{\lambda^{(n)}}\} = |\bar{\theta}^\top(\bar{x} - x_i)| + c\frac{1}{[\beta + \bar{\theta}^\top(\bar{x} - \tilde{x})]} \sum_{k \in (\bar{x} \Delta \tilde{x}) \cap x_i} \frac{|\eta_k|}{\psi^{3/2}}$$

for an appropriate universal constant $c > 0$, which has finite expectation. Further, since $\lambda^{(n)} \in \tilde{\triangle}$, the calculation showing that $\mathcal{L}(\lambda; \eta)$ is Lipschitz in λ implies that $A_{\text{semi}}(\lambda)^{-1/2}$ is Lipschitz in λ, so $A_{\text{semi}}(\lambda^{(n)})^{-1/2}$ can be made arbitrarily close to $A_{\text{semi}}(\lambda)^{-1/2}$. If $|\arg\max_{x \in \mathcal{X}} \frac{(\bar{x} - x)^\top A_{\text{semi}}(\lambda)^{-1/2} \eta}{\beta + \bar{\theta}^\top(\bar{x} - x)}| = 1$, this implies that:

$$\frac{(\bar{x} - x_\eta)^\top A_{\text{semi}}(\lambda)^{-1/2} \eta}{\beta + \bar{\theta}^\top(\bar{x} - x_\eta)} \geq \frac{(\bar{x} - x')^\top A_{\text{semi}}(\lambda)^{-1/2} \eta}{\beta + \bar{\theta}^\top(\bar{x} - x')} + \epsilon_\eta$$

for some $\epsilon_\eta > 0$, x_η the unique value the argmax is attained at, and $x' \neq x_\eta$. As we can make $A_{\text{semi}}(\lambda^{(n)})^{-1/2}$ arbitrarily close to $A_{\text{semi}}(\lambda)^{-1/2}$, it follows that for large enough n, we can guarantee:

$$\frac{(\bar{x} - x_\eta)^\top A_{\text{semi}}(\lambda^{(n)})^{-1/2} \eta}{\beta + \bar{\theta}^\top(\bar{x} - x_\eta)} \geq \frac{(\bar{x} - x')^\top A_{\text{semi}}(\lambda^{(n)})^{-1/2} \eta}{\beta + \bar{\theta}^\top(\bar{x} - x')} + \epsilon_\eta/2$$

so the maximizer will be unique. As this is true for all $\eta \in B_\lambda$, it follows that $\lim_{n\to\infty} B_{\lambda^{(n)}} \subseteq B_\lambda$. An identical argument implies $B_\lambda \subseteq \lim_{n\to\infty} B_{\lambda^{(n)}}$, so $\lim_{n\to\infty} B_{\lambda^{(n)}} = B_\lambda$. Then, by the dominated convergence theorem,

$$\begin{aligned}\lim_{n \longrightarrow \infty} \mathbb{E}[\nabla \mathcal{L}(\lambda^{(n)}; \eta)^\top e_i \mathbb{1}\{B_{\lambda^{(n)}}\}] &= \mathbb{E}[\lim_{n \longrightarrow \infty} \nabla \mathcal{L}(\lambda^{(n)}; \eta)^\top e_i \mathbb{1}\{B_{\lambda^{(n)}}\}] \\ &= \mathbb{E}[\nabla \mathcal{L}(\lambda; \eta)^\top e_i \mathbb{1}\{B_\lambda\}]\end{aligned}$$

where in the last line we used the continuity of $\nabla \mathcal{L}(\lambda; \eta)^\top e_i \mathbb{1}\{B_\lambda\}$ in λ on $\tilde{\triangle}$ for a fixed η. Thus, the partial derivatives are continuous, proving differentiability at every $\lambda \in \tilde{\triangle}$.

□

The following Lemma shows that $\mathcal{L}(\kappa_1, \kappa_2; \tau; \lambda)$ is twice-differentiable wrt λ.

Lemma 13. *$\mathcal{L}(\lambda)$ is twice-differentiable at every $\lambda \in \tilde{\triangle}$ and*

$$\frac{\partial^2 \mathcal{L}(\lambda)}{\partial \lambda_i \partial \lambda_j} = \mathbb{E}[\frac{\partial^2 \mathcal{L}(\lambda; \eta)}{\partial \lambda_i \partial \lambda_j} \mathbb{1}\{B_\lambda\}]$$

where

$$B_\lambda = \{\eta : |\arg\max_{x \in \mathcal{X}} \frac{(\bar{x} - x)^\top A_{\text{semi}}(\lambda)^{-1/2} \eta}{\beta + \bar{\theta}^\top (\bar{x} - x)}| = 1\}.$$

Proof. **Step 0: Setup.** From Lemma 12, $\mathcal{L}(\lambda)$ is differentiable at every $\lambda \in \tilde{\triangle}_\psi$. Therefore, it suffices to show that $\nabla \mathcal{L}(\lambda)$ is differentiable at every $\lambda \in \tilde{\triangle}_\psi$. It suffices to show that the 2nd order partial derivatives exist and are continuous. For the sake of abbreviation, define $g(\lambda) := \frac{\partial \mathcal{L}(\lambda)}{\partial \lambda_j}$ and $g(\lambda; \eta) := \frac{\partial \mathcal{L}(\lambda; \eta)}{\partial \lambda_j}$. Note that we have that

$$\frac{\partial g(\lambda; \eta)}{\partial \lambda_i} \mathbb{1}\{B_\lambda\} = \mathbb{1}\{B_\lambda\} \frac{3}{4} \frac{1}{(\beta + \bar{\theta}^\top (\bar{x} - \tilde{x}))} \sum_{k \in (\bar{x} \Delta \tilde{x}) \cap x_i \cap x_j} \frac{\eta_k}{(\sum_{l: k \in x_l} \lambda_l)^{5/2}} \tag{32}$$

$$\text{where } \tilde{x} = \arg\max_{x \in \mathcal{X}} \frac{\sum_{i \in \bar{x} \Delta x} \frac{\eta_i}{\sum_{x': i \in x'} \lambda_{x'}}}{\beta + \bar{\theta}^\top (\bar{x} - x)}.. \tag{33}$$

To begin, we show that the 2nd order partial derivatives exist using a truncation argument. Let $\varphi > 0$. Fix $\lambda \in \tilde{\triangle}$. Define

$$q(x; \eta) = \frac{\sum_{i \in \bar{x} \Delta x} \frac{\eta_i}{\sum_{x': i \in x'} \lambda_{x'}}}{\beta + \bar{\theta}^\top (\bar{x} - x)}$$

Define

$$B_\varphi = \{\eta : \tilde{x} = \arg\max_{x \in \mathcal{X}} q(x; \eta), \forall x' \neq \tilde{x} \quad \frac{q(x'; \eta)}{\|\eta\|_2} < \frac{q(\tilde{x}; \eta)}{\|\eta\|_2} - \varphi\}.$$

Note that

$$\lim_{\varphi \longrightarrow 0} B_\varphi = B_\lambda.$$

Step 1. First, we show that

$$\lim_{h \to 0} \mathbb{E}\left[(\frac{g(\lambda + he_i, \eta) - g(\lambda, \eta)}{h}) \mathbb{1}\{B_\varphi\}\right] = \mathbb{E}\left[\frac{\partial g(\lambda, \eta)}{\partial \lambda_i} \mathbb{1}\{B_\varphi\}\right] \tag{34}$$

Define

$$V_x = \frac{\bar{x} - x}{\beta + \bar{\theta}^\top (\bar{x} - x)}$$

for $x \in \mathcal{X}$. Note that since for any fixed $x \in \mathcal{X}$, $A_{\text{semi}}(\lambda)^{-1/2} V_x$ is Lipschitz in λ on $\tilde{\triangle}$, there exists L_ψ depending on $\psi, \beta, \bar{x}$ such that for all $x \in \mathcal{X}$

$$\left\| [A_{\text{semi}}(\lambda)^{-1/2} - A_{\text{semi}}(\lambda + he_i)^{-1/2}] V_x \right\|_2 \leq L_\psi h.$$

Let $h_{min} = \frac{\varphi}{4 L_\psi}$. Let $h \in [0, h_{min}]$. Let $\eta \in \mathbb{R}^d$ such that it satisfies B_φ and let $\tilde{x} = \arg\max_{x \in \mathcal{X}} q(x; \eta)$. Let $x \in \mathcal{X} \setminus \{\tilde{x}\}$. Then,

$$\begin{aligned}
\frac{\varphi}{4} + \frac{v_{\tilde{x}}^\top A(\lambda + he_i)^{-1/2} \eta}{\|\eta\|_2} &\geq \frac{v_{\tilde{x}}^\top A(\lambda)^{-1/2} \eta}{\|\eta\|_2} \\
&\geq \varphi + \frac{v_x^\top A(\lambda)^{-1/2} \eta}{\|\eta\|_2} \\
&\geq \frac{3\varphi}{4} + \frac{v_x^\top A(\lambda + he_i)^{-1/2} \eta}{\|\eta\|_2}
\end{aligned}$$

which implies that $\tilde{x} = \arg\max_{x \in \mathcal{X}} v_x^\top A(\lambda + he_i)^{-1/2}\eta$. Thus, on B_φ, for all $h \in [0, h_{min}]$, $\arg\max_{x \in \mathcal{X}} v_x^\top A(\lambda + he_i)^{-1/2}\eta$ is the same and hence $g(\lambda + he_i, \eta) = V_{\tilde{x}} A(\lambda + he_i)^{-1/2}\eta$ for all $h \in (0, h_{min})$ and is thus differentiable for all $h \in (0, h_{min})$. Thus, by the mean value theorem, we have that

$$(\frac{g(\lambda + he_i, \eta) - g(\lambda, \eta)}{h})\mathbb{1}\{B_\varphi\} = \frac{\partial g(\lambda + h' e_i; \eta)}{\partial \lambda_i}\mathbb{1}\{B_\varphi\}$$

for some $h' \in (0, h]$. Inspection of (33) shows that using $\lambda \in \tilde{\triangle}$

$$\mathbb{E}[|\frac{\partial g(\lambda; \eta)}{\partial \lambda_i}|] < \infty.$$

Thus, we may apply the dominating convergence theorem to obtain

$$\begin{aligned} \lim_{h \to 0} \mathbb{E}\left[(\frac{g(\lambda + he_i, \eta) - g(\lambda, \eta)}{h})\mathbb{1}\{B_\varphi\}\right] &= \mathbb{E}\left[\lim_{h \to 0}(\frac{g(\lambda + he_i, \eta) - g(\lambda, \eta)}{h})\mathbb{1}\{B_\varphi\}\right] \\ &= \mathbb{E}\left[\frac{\partial g(\lambda, \eta)}{\partial \lambda_i}\mathbb{1}\{B_\varphi\}\right] \end{aligned}$$

Step 2. Now, we show that

$$\lim_{\varphi \longrightarrow 0} \mathbb{E}\left[\frac{\partial g(\lambda, \eta)}{\partial \lambda_i}\mathbb{1}\{B_\varphi\}\right] = \mathbb{E}\left[\frac{\partial g(\lambda, \eta)}{\partial \lambda_i}\mathbb{1}\{B_\lambda\}\right]. \tag{35}$$

Define

$$Z(\eta) = \frac{3}{4}\frac{1}{(\beta + \bar{\theta}^\top(\bar{x} - \tilde{x})} \sum_{k \in (\bar{x}\Delta\tilde{x}) \cap x_i \cap x_j} \frac{|\eta_k|}{(\sum_{l: k \in x_l} \lambda_l)^{5/2}} \text{ where } \tilde{x} = \arg\max_{x \in \mathcal{X}} \frac{\sum_{i \in \bar{x}\Delta x} \frac{\eta_i}{\sum_{x': i \in x'} \lambda_{x'}}}{\beta + \bar{\theta}^\top(\bar{x} - x)}.$$

Note that for every $\varphi > 0$

$$|\frac{\partial g(\lambda, \eta)}{\partial \lambda_i}\mathbb{1}\{B_\varphi\}| \leq Z(\eta)$$

and $\mathbb{E}Z(\eta) < \infty$. Therefore, by the dominating convergence theorem,

$$\lim_{\varphi \longrightarrow 0} \mathbb{E}\left[\frac{\partial g(\lambda, \eta)}{\partial \lambda_i}\mathbb{1}\{B_\varphi\}\right] = \mathbb{E}\left[\lim_{\varphi \longrightarrow 0}\frac{\partial g(\lambda, \eta)}{\partial \lambda_i}\mathbb{1}\{B_\varphi\}\right] = \mathbb{E}\left[\frac{\partial g(\lambda, \eta)}{\partial \lambda_i}\mathbb{1}\{B_\lambda\}\right].$$

Step 3. Now, we show that

$$\lim_{\varphi \longrightarrow 0}\lim_{h \to 0} \mathbb{E}\left[(\frac{g(\lambda + he_i, \eta) - g(\lambda, \eta)}{h})\mathbb{1}\{B_\varphi\}\right] = \lim_{h \to 0} \mathbb{E}\left[(\frac{g(\lambda + he_i, \eta) - g(\lambda, \eta)}{h})\mathbb{1}\{B_\lambda\}\right] \tag{36}$$

By step 1, for every $\varphi > 0$,

$$\lim_{h \to 0} \mathbb{E}\left[(\frac{g(\lambda + he_i, \eta) - g(\lambda, \eta)}{h})\mathbb{1}\{B_\varphi\}\right] = \mathbb{E}\left[\frac{\partial g(\lambda, \eta)}{\partial \lambda_i}\mathbb{1}\{B_\varphi\}\right] \leq \mathbb{E}\left[|\frac{\partial g(\lambda, \eta)}{\partial \lambda_i}\mathbb{1}\{B_\lambda\}|\right] \leq C$$

for some constant $C > 0$. Therefore, by the bounded convergence theorem for limits, we have that

$$\lim_{\varphi \longrightarrow 0}\lim_{h \to 0} \mathbb{E}\left[(\frac{g(\lambda + he_i, \eta) - g(\lambda, \eta)}{h})\mathbb{1}\{B_\varphi\}\right] = \lim_{h \to 0}\lim_{\varphi \longrightarrow 0} \mathbb{E}\left[(\frac{g(\lambda + he_i, \eta) - g(\lambda, \eta)}{h})\mathbb{1}\{B_\varphi\}\right]$$

More formally, consider some sequence φ_m, h_n such that $\varphi_m \to 0$ as $m \to \infty$ and $h_n \to 0$ as $n \to \infty$. Let $a_{mn} = \mathbb{E}\left[(\frac{g(\lambda + h_n e_i, \eta) - g(\lambda, \eta)}{h_n})\mathbb{1}\{B_{\varphi_m}\}\right]$. If $\lim_{m \to \infty}\lim_{n \to \infty} a_{mn} = \lim_{n \to \infty}\lim_{m \to \infty} a_{mn}$ then the result is proven. Let $c_{mn} = a_{mn} - a_{m,n-1}$ and $c_{m0} = 0$. Note that for finite m, c_{mn} is uniformly bounded for all n. Then the Bounded Convergence Theorem applied to the counting measure gives that:

$$\lim_{m \to \infty} \sum_{n=0}^{\infty} c_{mn} = \sum_{n=0}^{\infty} \lim_{m \to \infty} c_{mn}$$

However, $\sum_{n=0}^{\infty} c_{mn} = \lim_{N\to\infty} \sum_{n=0}^{N} c_{mn}$, so the above implies:

$$\lim_{m\to\infty} \lim_{N\to\infty} \sum_{n=0}^{N} c_{mn} = \lim_{N\to\infty} \lim_{m\to\infty} \sum_{n=0}^{N} c_{mn}$$

By construction, we have $\sum_{n=0}^{N} c_{mn} = a_{mN}$, which proves the result.

Fix $h > 0$. Define

$$Y(h) = |\frac{g(\lambda + he_i, \eta) - g(\lambda, \eta)}{h})|\mathbb{1}\{B_\lambda\}.$$

Note that for every $\varphi > 0$

$$|(\frac{g(\lambda + he_i, \eta) - g(\lambda, \eta)}{h})\mathbb{1}\{B_\varphi\}| \leq Y(h)$$

and $EY(h) < \infty$. Thus, by the dominating convergence theorem,

$$\begin{aligned}\lim_{\varphi\longrightarrow 0} \mathbb{E}\left[(\frac{g(\lambda + he_i, \eta) - g(\lambda, \eta)}{h})\mathbb{1}\{B_\varphi\}\right] &= \mathbb{E}\left[\lim_{\varphi\longrightarrow 0}(\frac{g(\lambda + he_i, \eta) - g(\lambda, \eta)}{h})\mathbb{1}\{B_\varphi\}\right] \\ &= \mathbb{E}\left[(\frac{g(\lambda + he_i, \eta) - g(\lambda, \eta)}{h})\mathbb{1}\{B_\lambda\}\right].\end{aligned}$$

This completes the step.

Step 4. Putting together (34), (35), and (36), we have shown that

$$\begin{aligned}\lim_{h\longrightarrow 0} \mathbb{E}\left[[(\frac{g(\lambda + he_i, \eta) - g(\lambda, \eta)}{h})\mathbb{1}\{B_\lambda\}\right] &= \lim_{\varphi\longrightarrow 0} \lim_{h\longrightarrow 0} \mathbb{E}\left[(\frac{g(\lambda + he_i, \eta) - g(\lambda, \eta)}{h})\mathbb{1}\{B_\varphi\}\right] \\ &= \lim_{\varphi\longrightarrow 0} \mathbb{E}\left[\frac{\partial g(\lambda, \eta)}{\partial \lambda_i}\mathbb{1}\{B_\varphi\}\right] \\ &= \mathbb{E}\left[\frac{\partial g(\lambda, \eta)}{\partial \lambda_i}\mathbb{1}\{B_\lambda\}\right]\end{aligned}$$

Thus, we have that that the second order partial derivatives exist and derived an expression for them. Showing that the second order partial derivatives are continuous proceeds as in the proof of Lemma 12 (apply the dominating convergence theorem).

□

E Rounding

Theorem 10 (Caratheodory's Theorem)**.** *For any point y in the convex hull of a set $\mathcal{P} \subseteq \mathbb{R}^d$, y can be written as a convex combination of at most $d + 1$ points in $\mathcal{P}$.*

Proof. This is a standard result in convex geometry, see for instance Eggleston [1958]. □

Lemma 14. *Given any $\lambda \in \triangle_\mathcal{X}$, in the bandit setting, there exists a distribution $\lambda' \in \triangle_\mathcal{X}$ that is $(d^2 + d + 1)$-sparse and:*

$$A_{\mathrm{band}}(\lambda) = A_{\mathrm{band}}(\lambda'), \quad \sum_{x\in\mathcal{X}} \lambda_x x = \sum_{x\in\mathcal{X}} \lambda'_x x$$

In the semi-bandit setting, when $\mathcal{X} \subseteq \{0, 1\}^d$, there exists a distribution $\lambda' \in \triangle_\mathcal{X}$ that is $(d + 1)$-sparse and:

$$A_{\mathrm{semi}}(\lambda) = A_{\mathrm{semi}}(\lambda'), \quad \sum_{x\in\mathcal{X}} \lambda_x x = \sum_{x\in\mathcal{X}} \lambda'_x x$$

Proof. This is a direct corollary of Caratheodory's Theorem. Take $\lambda \in \triangle_{|\mathcal{X}|}$ and let $z_\lambda \in \mathbb{R}^{d+d^2}$, which we define as:

$$z_\lambda = \left[\sum_{x \in \mathcal{X}} \lambda_x x ; \mathrm{vec} \left(\sum_{x \in \mathcal{X}} \lambda_x x x^\top \right) \right]$$

Define the set:

$$\mathcal{V} := \left\{ \left[x ; \mathrm{vec}(x x^\top) \right] : x \in \mathcal{X} \right\} \subseteq \mathbb{R}^{d+d^2}$$

For any λ, we see that z_λ lies in the convex hull of $\mathcal{V}$. Caratheodory's Theorem then immediately implies the result in the bandit case, since $\mathrm{vec}\left(\sum_{x \in \mathcal{X}} \lambda_x x x^\top \right)$ uniquely determines $A_{\mathrm{band}}(\lambda)$.

In the semi-bandit case, we note that the diagonal of $A_{\mathrm{semi}}(\lambda)$ is equal to $\sum_{x \in \mathcal{X}} \lambda_x x$. Thus, we only need to consider a d-dimensional space, so Caratheodory implies we can find a $d+1$ sparse distribution. □

Proof of Lemma 1. Given some allocation τ, let λ the corresponding distribution, and $\bar{\tau} = \sum_{x \in \mathcal{X}} \tau_x$ (so $\tau = \bar{\tau} \lambda$).

Since we only care about the sparsity of λ, consider $\bar{\tau}$ fixed. Then, given a solution λ to (2) or (3), the value of the constraint and objective the solution achieves achieves are fully specified by $A_{\mathfrak{f}}(\lambda)$ and $\sum_{x \in \mathcal{X}} \lambda_x x$. To see the latter, note that $\sum_{x \in \mathcal{X}} (\epsilon + \Delta_x) \lambda_x = \epsilon + \sum_{x \in \mathcal{X}} \theta^\top (x_* - x) \lambda_x = \epsilon + \theta^\top x_* + \theta^\top \sum_{x \in \mathcal{X}} \lambda_x x$. Lemma 14 then implies that there exists a distribution λ that is $(d^2 + d + 1)$-sparse in the bandit case and $(d+1)$-sparse in the semi-bandit case that achieves the same value of the constraint and objective of (2) or (3).

To see the second part of the result, note that if we run the procedure of Theorem 4, we will run stochastic Frank Wolfe for a polynomial number of steps, each increasing the support of our distribution by at most 1, so we will obtain an approximate solution that has at most $n = \mathrm{poly}(d, \Delta_{\min}, T, 1/\delta)$ non-zero entries. By Theorem 6 in Maalouf et al. [2019], it then follows that we can compute the $(d+1)$-sparse distribution achieving the same value of the constraint and objective in time $\mathcal{O}(nd)$. □

F Gaussian Width Results

Proposition 9.

$$\inf_{\lambda \in \triangle_{\mathcal{X}}} \max_{x \in \mathcal{X}} \|x\|^2_{A_{\mathrm{semi}}(\lambda)^{-1}} = d \tag{37}$$

Proof. This proof closely mirrors the proof of Theorem 21.1 of Lattimore and Szepesvári [2020].

Let:

$$f(\lambda) = \log \det A_{\mathrm{semi}}(\lambda)$$

Noting that:

$$\frac{d}{dt} \det(A(t)) = \mathrm{Trace}\left(\mathrm{adj}(A(t)) \frac{d}{dt} A(t) \right)$$

and $A^{-1} = \mathrm{adj}(A)^\top / \det(A)$ Lattimore and Szepesvári [2020], we can compute the gradient of $f(\lambda)$ as:

$$\frac{d}{d\lambda_x} f(\lambda) = \frac{1}{\det A_{\mathrm{semi}}(\lambda)} \mathrm{Trace}\left(\mathrm{adj}(A_{\mathrm{semi}}(\lambda)) \, \mathrm{diag}(x x^\top) \right) = \mathrm{Trace}\left(A_{\mathrm{semi}}(\lambda)^{-1} \, \mathrm{diag}(x x^\top) \right)$$

Since $A_{\mathrm{semi}}(\lambda)$ is diagonal, we have:

$$\mathrm{Trace}\left(A_{\mathrm{semi}}(\lambda)^{-1} \, \mathrm{diag}(x x^\top) \right) = \sum_{i=1}^d \frac{x_i^2}{[A_{\mathrm{semi}}(\lambda)]_i} = x^\top A_{\mathrm{semi}}(\lambda)^{-1} x = \|x\|^2_{A_{\mathrm{semi}}(\lambda)^{-1}}$$

Note also that, by the identity above, for any λ:

$$\begin{aligned}
\sum_{x\in\mathcal{X}} \lambda_x \|x\|^2_{A_{\mathrm{semi}}(\lambda)^{-1}} &= \sum_{x\in\mathcal{X}} \lambda_x \, \mathrm{Trace}\left(A_{\mathrm{semi}}(\lambda)^{-1} \mathrm{diag}(xx^\top)\right) \\
&= \mathrm{Trace}\left(A_{\mathrm{semi}}(\lambda)^{-1}\left(\sum_{x\in\mathcal{X}} \lambda_x \, \mathrm{diag}(xx^\top)\right)\right) \\
&= \mathrm{Trace}\left(A_{\mathrm{semi}}(\lambda)^{-1} A_{\mathrm{semi}}(\lambda)\right) \\
&= \mathrm{Trace}(I) \\
&= d
\end{aligned}$$

Then, since $\log\det X$ is concave and $A_{\mathrm{semi}}(\lambda)$ is linear in λ, it follows that $f(\lambda)$ is concave. Applying standard first-order optimality conditions and denoting λ^* the solution to (37), we have, for any λ:

$$\begin{aligned}
0 &\geq \langle f(\lambda^*), \lambda - \lambda^* \rangle \\
&= \sum_{x\in\mathcal{X}} \lambda_x \|x\|^2_{A_{\mathrm{semi}}(\lambda^*)^{-1}} - \sum_{x\in\mathcal{X}} \lambda^*_x \|x\|^2_{A_{\mathrm{semi}}(\lambda^*)^{-1}} \\
&= \sum_{x\in\mathcal{X}} \lambda_x \|x\|^2_{A_{\mathrm{semi}}(\lambda^*)^{-1}} - d
\end{aligned}$$

Choosing λ to be the distribution putting all its mass on x, we have:

$$d \geq \|x\|^2_{A_{\mathrm{semi}}(\lambda^*)^{-1}}$$

To see the equality, note that the above implies:

$$d = \sum_{x\in\mathcal{X}} \lambda^*_x \|x\|^2_{A_{\mathrm{semi}}(\lambda^*)^{-1}} \leq \max_{x\in\mathcal{X}} \|x\|^2_{A_{\mathrm{semi}}(\lambda^*)^{-1}} \leq d$$

□

Proof of Proposition 2. Let $S = \{x \in \mathcal{X} : \Delta_x \leq \epsilon\}$ for some fixed $\epsilon > 0$. Therefore, $x^* \in S$. Define

$$\begin{aligned}
S_1 &= \{(x, x^*_{m+1:n+m}) : x \in \{0,1\}^m \text{ s.t. there exists } x' \in S \text{ s.t. } \Pi_{[m]} x' = x\} \\
S_2 &= \{(x^*_{1:m}, x) : x \in \{0,1\}^n \text{ s.t. there exists } x' \in S \text{ s.t. } \Pi_{[n+m]\setminus[m]} x' = x\}
\end{aligned}$$

where Π_A is the coordinate wise projection onto the coordinates $A \subset \mathbb{N}$. Then, using the fact that $\mathbb{E}[(x^*)^\top A(\lambda)^{-1/2}\eta] = 0$, we have that

$$\begin{aligned}
\min_{\lambda\in\triangle^{|S|}} \mathbb{E}[\sup_{x\in S} x^\top A(\lambda)^{-1/2}\eta]^2 &\leq \min_{\lambda\in\triangle^{|S|}} \mathbb{E}[\sup_{x_1\in S_1} \sum_{i=1}^{m} x_{1,i}[A(\lambda)^{-1/2}\eta]_i + \sup_{x_2\in S_2} \sum_{i=m+1}^{n+m} x_{2,i}[A(\lambda)^{-1/2}\eta]_i]^2 \\
&= \min_{\lambda\in\triangle^{|S|}} \mathbb{E}[\sup_{x_1\in S_1} \sum_{i=1}^{m} x_{1,i}[A(\lambda)^{-1/2}\eta]_i + \sup_{x_2\in S_2} \sum_{i=m+1}^{n+m} x_{2,i}[A(\lambda)^{-1/2}\eta]_i \\
&\quad + \sum_{i=1}^{n+m} x^*_i [A(\lambda)^{-1/2}\eta]_i]^2 \\
&= \min_{\lambda\in\triangle^{|S|}} \mathbb{E}[\sup_{x_1\in S_1} x_1^\top A(\lambda)^{-1/2}\eta + \sup_{x_2\in S_2} x_2^\top A(\lambda)^{-1/2}\eta]^2 \\
&\leq \min_{\lambda\in\triangle^{|S|}} c[\mathbb{E}[\sup_{x_1\in S_1} x_1^\top A(\lambda)\eta]^2 + \mathbb{E}[\sup_{x_2\in S_2} x_2^\top A(\lambda)\eta]^2] \\
&\leq \min_{\lambda\in\triangle^{|S|}} c'[k\log(m) \max_{x_1\in S_1} \|x_1\|^2_{A(\lambda)^{-1}} + \ell\log(n) \max_{x_2\in S_2} \|x_2\|^2_{A(\lambda)^{-1}}] \\
&\leq c''[k\log(m) \min_{\lambda\in\triangle^{|S|}} \max_{x_1\in S_1} \|x_1\|^2_{A(\lambda)^{-1}} + \ell\log(n) \min_{\lambda\in\triangle^{|S|}} \max_{x_2\in S_2} \|x_2\|^2_{A(\lambda)^{-1}}]
\end{aligned}$$

We begin by bounding the first term. Notice that $S_1 \subset S$ since $S = \{x \in \mathcal{X} : \Delta_x \leq \epsilon\}$ for some fixed $\epsilon > 0$ and thus if $x \in \{0,1\}^m$ s.t. there exists $x' \in S$ s.t. $\Pi_{[m]}x' = x$, then $(x, x^*_{m+1:n+m}) \in S$. Furthermore, the span of the vectors in S_1 has dimension at most $m+1$ since for any $x_1 \in S_1$, for all $i \geq m+1$, we have that

$$[x_1 - (\vec{0}_{1:m}, x^*_{m+1:n+m})]_i = 0.$$

Thus, by the Kiefer-Wolfowitz Theorem Lattimore and Szepesvári [2020]:

$$\min_{\lambda \in \triangle^{|S|}} \max_{x_1 \in S_1} \|x_1\|^2_{A(\lambda)^{-1}} \leq m+1.$$

and:

$$\min_{\lambda \in \triangle^{|S|}} \max_{x_2 \in S_2} \|x_2\|^2_{A(\lambda)^{-1}} \leq n.$$

Therefore,

$$\min_{\lambda \in \triangle^{|S|}} \mathbb{E}[\sup_{x \in S} x^\top A(\lambda)^{-1/2}\eta]^2 \leq c[k \log(m) m + \ell \log(n) n].$$

To lower bound $|\mathcal{X}|$, note that:

$$|\mathcal{X}| = \binom{m}{k}\binom{n}{\ell} \geq \left(\frac{m}{k}\right)^k \left(\frac{n}{\ell}\right)^\ell$$

For the second conclusion we set $\ell = \mathcal{O}(1)$, $k = \sqrt{m}$, and $n = m^{3/2}$ and apply our regret bound.

For the regret bound of competing algorithms, LinUCB will scale as $\widetilde{\mathcal{O}}(d\sqrt{T}) = \widetilde{\mathcal{O}}(m^{3/2}\sqrt{T})$. Given the above lower bound on $|\mathcal{X}|$, the regret of action elimination will scale as $\widetilde{\mathcal{O}}(m\sqrt{T})$. In the semi-bandit setting, Kveton et al. [2015] obtain a regret bound of $\widetilde{\mathcal{O}}(m\sqrt{T})$ and, ignoring logarithmic terms, Degenne and Perchet [2016] obtain the same bound. Other existing works [Combes et al., 2015, Perrault et al., 2020a] do not state minimax bounds but, using the standard analysis to obtain a minimax bound from a gap-dependent bound, their regret will also scale as $\widetilde{\mathcal{O}}(m\sqrt{T})$. Note that in this comparison we have ignored $\log(T)$ terms and have taken the dominate term to be the term with leading m dependence that hits the $\sqrt{T}$. □

Proof of Proposition 3. $\mathbb{E}_\eta[\max_{x \in \mathcal{X}} x^\top A(\lambda)^{-1/2}\eta]$ is the Gaussian width of the set $\{A(\lambda)^{-1/2}x \,:\, x \in \mathcal{X}\}$. By Proposition 7.5.2 of Vershynin [2018]:

$$\mathbb{E}_\eta[\max_{x \in \mathcal{X}} x^\top A(\lambda)^{-1/2}\eta] \leq c\sqrt{d}\,\mathrm{diam}(\{A(\lambda)^{-1/2}x \,:\, x \in \mathcal{X}\})$$

and:

$$\mathrm{diam}(\{A(\lambda)^{-1/2}x \,:\, x \in \mathcal{X}\}) = \max_{x_1, x_2 \in \mathcal{X}} \|A(\lambda)^{-1/2}(x_1 - x_2)\| \leq 2 \max_{x \in \mathcal{X}} \|x\|_{A(\lambda)^{-1}}$$

Taking the infimum over $\lambda \in \triangle_{\mathcal{X}}$, in the bandit feedback case Kiefer-Wolfowitz gives $\inf_{\lambda \in \triangle_{\mathcal{X}}} \max_{x \in \mathcal{X}} \|x\|_{A(\lambda)^{-1}} \leq \sqrt{d}$, and in the semi-bandit case, Proposition 9 gives the same result. Since $\mathcal{X}$ was chosen arbitrarily, it follows that $\bar{\gamma}(\mathcal{X}) \leq d^2$.

For the second bound, Exercise 7.5.10 of Vershynin [2018] gives that:

$$\mathbb{E}_\eta[\max_{x \in \mathcal{X}} x^\top A(\lambda)^{-1/2}\eta] \leq c\sqrt{\log|\mathcal{X}|}\,\mathrm{diam}(\{A(\lambda)^{-1/2}x \,:\, x \in \mathcal{X}\})$$

from which the result follows immediately. □

Proof of Proposition 4. If $\mathcal{X} \subseteq \{0,1\}^d$ and $k = \max_{x \in \mathcal{X}} \|x\|_1$, then $\mathcal{X}$ at most contains all subsets of size k and less so:

$$|\mathcal{X}| \leq \sum_{j=1}^k \binom{d}{j} \leq c \sum_{j=1}^k (d/j)^j \leq c \sum_{j=1}^k d^j = c\frac{d(d^k - 1)}{d-1} \leq cd^k$$

Thus, Proposition 3 gives:

$$\gamma^* \leq cdk \log d$$

□

Proof of Proposition 5. Consider the Top-k problem in the semi-bandit feedback regime, but augment the action set by adding the vector of all 1s to it. In this case, then, we can either query a subset of size k, or we can query every point at once. Assume that $\theta_i \geq 0$ for all i. Note that by our assumption on θ_i, $\mathbf{1}$ will always be in the action set regardless of how we are filtering on the gaps. If we put all our mass on $\mathbf{1}$, we will have that $A_{\text{semi}}(\lambda) = I$. Thus:

$$\begin{aligned}
\bar{\gamma}(A_{\text{semi}}) &= \sup_{\epsilon > 0} \inf_{\lambda \in \triangle_{\mathcal{X}_\epsilon}} \mathbb{E}_\eta[\sup_{x \in \mathcal{X}_\epsilon} x^\top A(\lambda)^{-1/2}\eta]^2 \\
&\leq \mathbb{E}_\eta[\sup_{x \in \mathcal{X}_\epsilon} x^\top \eta]^2 \\
&\leq \mathbb{E}_\eta[\max_{x \in \mathcal{X}} |x^\top \eta|]^2 \\
&\leq c\left(\mathbb{E}_\eta[\max_{x \in \mathcal{X}\setminus \mathbf{1}} |x^\top \eta|]^2 + \mathbb{E}_\eta[|\mathbf{1}^\top \eta|]^2\right) \\
&\leq c\left(\mathbb{E}_\eta[\max_{x \in \mathcal{X}\setminus \mathbf{1}} |x^\top \eta|]^2 + d\right) \\
&\leq c\left(k^2 \mathbb{E}_\eta[\max_{z : \|z\|_1 \leq 1} |z^\top \eta|]^2 + d\right) \\
&\leq c(k^2 \log d + d)
\end{aligned}$$

where the last inequality follows since the gaussian complexity is within a constant of the Gaussian width when the set contains 0, by Exercise 7.6.9 of Vershynin [2018]. The result then follows by choosing $k = \sqrt{d}$. □

Theorem 11 (Tsirelson-Ibragimov-Sudakov Inequality Tsirelson et al. [1976])**.** *Let $\mathcal{S} \subseteq \mathbb{R}^d$ be bounded. Let $(V_s)_{s \in \mathcal{S}}$ be a Gaussian process such that $\mathbb{E}[V_s] = 0$ for all $s \in \mathcal{S}$. Define $\sigma^2 = \sup_{s \in \mathcal{S}} \mathbb{E}[V_s^2]$. Then, for all $u > 0$:*

$$\mathbb{P}[|\sup_{s \in \mathcal{S}} V_s - \mathbb{E}\sup_{x \in \mathcal{S}}| \geq u] \leq 2\exp\left(\frac{-u^2}{2\sigma^2}\right)$$

Proof of Proposition 6. The proof in the bandit setting is identical to the proof given in Katz-Samuels et al. [2020] and we therefore omit it.

In the semibandit setting, we have that:

$$\hat{\theta}_i = \theta_i + \frac{1}{T_i}\sum_{t=1}^T x_{t,i}\eta_{t,i}$$

so $\mathbb{E}[\hat{\theta}_i] = \theta_i$ and:

$$\mathbb{E}[(\hat{\theta}_i - \theta_i)^2] = \frac{1}{T_i^2}\sum_{t=1}^T x_{t,i} = \frac{1}{T_i}$$

Furthermore, since the noise is uncorrelated between coordinates, we have $\mathbb{E}[(\hat{\theta}_i - \theta_i)(\hat{\theta}_j - \theta_j)] = 0$. Since $x_t \in \{0,1\}^d$, it follows then that:

$$\hat{\theta} \overset{\text{distribution}}{=} \theta_* + \widetilde{A}^{-1/2}\eta$$

for $\eta \sim \mathcal{N}(0, I)$. Now consider the Gaussian process $V_x := x^\top(\hat{\theta} - \theta_*) = x^\top \widetilde{A}^{-1/2}\eta$ for $x \in \mathcal{X}$. Noting that $\mathbb{E}[V_x^2] = x^\top \widetilde{A}^{-1}x \leq \max_{x \in \mathcal{X}} \|x\|^2_{\widetilde{A}^{-1}}$, we can then apply Theorem 11 to this process, which gives the result. □

G Lower Bound for Semi-Bandit Feedback and Optimistic Strategies

A policy π is *consistent* if for all θ and $p > 0$, $R_\theta^\pi(T) = o(T^p)$. Let T_x denote the number of times that $x \in \mathcal{X}$ is pulled and T_i the number of times that $i \in [d]$ is pulled.

Theorem 12. *Let π be a consistent policy such that $T_i \geq 1$ for all $i \in [d]$ with probability 1, $\theta \in \mathbb{R}^d$ such that there is a unique optimal arm in $\mathcal{X}$. Let $G_T = \mathbb{E}[\sum_{t=1}^T \operatorname{diag}(x_t x_t^\top)]$ where x_t is chosen at round $t \in [T]$. Then,*

$$\limsup_{T \longrightarrow \infty} \log(T) \|x\|^2_{G_T^{-1}} \leq \frac{\Delta_x^2}{2}$$

for all $x \in \mathcal{X}$. Furthermore,

$$\limsup_{T \longrightarrow \infty} \frac{R_\theta^\pi(T)}{\log(T)} \geq c(\mathcal{X}, \theta)$$

where

$$c(\mathcal{X}, \theta) := \min_{\tau \in [0,\infty)^{|\mathcal{X}|}} \sum_{x \in \mathcal{X}} \tau_x \Delta_x$$
$$s.t. \sum_{i \in x} \frac{1}{\sum_{x': i \in x'} \tau_{x'}} \leq \frac{\Delta_x^2}{2} \quad \forall x \in \mathcal{X} \setminus \{x_*\}.$$

Proof. We use a similar argument to the proof of Theorem 1 in Lattimore and Szepesvari [2017]. We construct an alternative instance θ' to obtain an asymptotic lower bound. Let P' denote the probability measure of the associated instance (which we will specify shortly). We note that the Divergence Lemma (Lemma 15.1 Lattimore and Szepesvári [2020]) is easily adapted to the semi-bandit feedback setting. Thus, by a standard argument that applies the Divergence Lemma and the Bretagnolle–Huber inequality (Theorem 14.2 in Lattimore and Szepesvári [2020]), we have that

$$\frac{1}{2} \|\theta - \theta'\|^2_{G_T} \geq \log\left(\frac{1}{2\mathbb{P}(E) + 2\mathbb{P}'(E^c)}\right) \tag{38}$$

for any event E. Define

$$\theta' = \theta + \frac{G_T^{-1}[x - x_*](\Delta_x + \epsilon)}{\|x - x_*\|^2_{G_T^{-1}}}.$$

Note that

$$(x - x_*)^\top \theta' = \epsilon > 0.$$

Let R'_T denote the regret of π on the alternative instance θ'. Choose $E = \{T_{x_*} \leq \frac{T}{2}\}$. We have that

$$R_T = \sum_x \mathbb{E}[T_x]\Delta_x \geq \Delta_{\min} \frac{T}{2} \mathbb{P}(T_{x_*} \leq T/2).$$

Furthermore,

$$R'_T = \sum_x \mathbb{E}[T_x]\Delta'_x \geq \frac{\epsilon T}{2} \mathbb{P}'(T_{x_*} \geq T/2).$$

Thus, assuming that $\epsilon \leq \Delta_{\min}$, we have that

$$\frac{R_T + R'_T}{\epsilon T} \geq \mathbb{P}(E) + \mathbb{P}'(E^c). \tag{39}$$

Then, inequalities (38) and (39) imply that

$$\frac{(\Delta_x + \epsilon)^2}{2\|x - x_*\|^2_{G_T^{-1}}} \geq \log\left(\frac{\epsilon T}{2[R_T + R'_T]}\right).$$

Dividing both sides by $\log(T)$, we have that

$$\frac{(\Delta_x + \epsilon)^2}{2 \|x - x_*\|^2_{G_T^{-1}}} \geq 1 - \frac{\log(1/2\epsilon)}{\log(T)} - \frac{\log(2R_T - R'_T)}{\log(T)}.$$

Consistency of the policy π implies that

$$\liminf_{T \longrightarrow \infty} \frac{(\Delta_x + \epsilon)^2}{2 \|x - x_*\|^2_{G_T^{-1}} \log(T)} \geq 1.$$

Rearranging, we have that

$$\frac{(\Delta_x + \epsilon)^2}{2} \geq \limsup_{T \longrightarrow \infty} \|x - x_*\|^2_{G_T^{-1}} \log(T).$$

This establishes the first claim in the lower bound. The second claim follows by a similar argument to the argument in Corollary 2 of Lattimore and Szepesvari [2017].

□

Proof of Proposition 1. **Proof of lower bound for optimism:** Define the following problem instance

$$\theta_i = \begin{cases} 1 & i = 1 \\ 1 - \epsilon & i \in \{2, \dots, m\} \\ -1 + \epsilon & i \in \{m+1, \dots, 2m-1\} \\ -1 & i \in \{2m, \dots, 2m + \sqrt{m}\} \end{cases}$$

with $\mathcal{X} = \{\{1\}, \dots, \{m\}, [2m + \sqrt{m}]\}$. Let $x^{(i)} = \{i\}$ for $i \leq m$ and $x^{(m+1)} = [2m + \sqrt{m}]$. Note that $\Delta_i = \epsilon$ if $i \leq m$ and $\Delta_{m+1} = \sqrt{m} + 1$. Then, the optimization problem in Theorem 12 becomes

$$\begin{aligned} \min_{\tau \in [0,\infty)^{|\mathcal{X}|}} & \sum_{i \leq m} \tau_i \epsilon + \tau_{m+1}(\sqrt{m} + 1) \\ \text{s.t. } & \frac{1}{\tau_i + \tau_{m+1}} \leq \epsilon^2/2 \quad \forall i \in \{2, \dots, m\} \\ & \sum_{i \in [m]} \frac{1}{\tau_i + \tau_{m+1}} + \frac{m + \sqrt{m}}{\tau_{m+1}} \leq \frac{(\sqrt{m}+1)^2}{2} \end{aligned}$$

Consider the solution is $\tau_{m+1} = \frac{4}{\epsilon^2}$ and $\tau_i = 0$ otherwise. This attains a value of

$$O(\frac{\sqrt{m}}{\epsilon^2}).$$

Now, consider the performance of the generic optimistic algorithm. Let T_i denote the number of times that arm i is chosen. Define the event

$$\mathcal{E} = \{|x^\top(\widehat{\theta}_t - \theta)| \leq \mathrm{CB}(x, \{x_s\}_{s \in [t-1]}) \forall x \in \mathcal{X},\ \forall t \in [T]\}.$$

Suppose $\mathcal{E}$ holds. Now, suppose that $T_{m+1} = 4\alpha \log(T)$. Then,

$$\begin{aligned} [x^{(m+1)}]^\top \widehat{\theta}_t + \mathrm{CB}(x^{(m+1)}, \{x_s\}_{s=1}^{t-1}) & \leq [x^{(m+1)}]^\top \theta + 2\,\mathrm{CB}(x^{(m+1)}, \{x_s\}_{s=1}^{t-1}) \\ & \leq -\sqrt{m} + 2\sqrt{\alpha \|x\|^2_{(\sum_{s=1}^{t-1} x_s x_s^\top)^{-1}} \log(T)} \\ & \leq 0. \end{aligned}$$

On the other hand, on $\mathcal{E}$, we have that $[x^{(1)}]^\top(\widehat{\theta}_t + \mathrm{CB}(x^{(1)}, \{x_s\}_{s=1}^{t-1})) \geq 1$ and hence $x^{(m)}$ is pulled at $4\alpha \log(T)$ times. Since $\mathbb{P}(\mathcal{E}^c) \leq \frac{1}{T}$, we have that

$$\mathbb{E}[T_{m+1}] \leq 4\alpha \log(T) + 1 \tag{40}$$

Recall that $G_T = \mathbb{E}[\sum_{t=1}^T \operatorname{diag}(x_t x_t^\top)]$. By Theorem 12, we have that

$$\limsup_{T\longrightarrow\infty} \log(T) \left\| x^{(1)} - x^{(i)} \right\|^2_{G_T^{-1}} \leq \epsilon^2/2$$

for all i, which together with (40) implies that

$$\mathbb{E}[T_i]/\log(T) = \Omega(1/\epsilon^2)$$

for all $i \in \{2, \dots, m\}$. Thus,

$$\limsup_{T\longrightarrow\infty} \frac{R_\theta^{optimistic}(T)}{\log(T)} = \Omega(m/\epsilon).$$

Proof of upper bound for Algorithm 1: From the proof of Theorem 2, we know that, for all ℓ simultaneously, with probability at least $1-\delta$:

$$\begin{aligned} \mathcal{R}_\ell \leq \min_\tau & \sum_{x\in\mathcal{X}} 2(\epsilon_\ell + \hat{\Delta}_x)\tau_x \\ \text{s.t. } & \mathbb{E}_\eta\left[\max_{x\in\mathcal{X}} \frac{(x_\ell - x)^\top A_{\text{semi}}(\tau)^{-1/2}\eta}{\epsilon_\ell + \hat{\Delta}_x}\right] \leq \frac{1}{128(1+\sqrt{\pi\log(2\ell^3/\delta)})} \end{aligned}$$

and a τ satisfying:

$$\mathbb{E}_\eta\left[\max_{x\in\mathcal{X}} \frac{(x_\ell - x)^\top A_{\text{semi}}(\tau)^{-1/2}\eta}{\epsilon_\ell + \Delta_x}\right] \leq \frac{1}{512(1+\sqrt{\pi\log(2\ell^3/\delta)})}$$

is also feasible for the problem above. Note that if we put all our mass on $\mathbf{1}$ we will have $A_{\text{semi}}(\tau) = \tau I$, so a feasible solution to the above problem requires that:

$$\left(512(1+\sqrt{\pi\log(2\ell^3/\delta)})\mathbb{E}_\eta\left[\max_{x\in\mathcal{X}} \frac{(x_\ell - x)^\top\eta}{\epsilon_\ell + \Delta_x}\right]\right)^2 \leq \tau$$

we can upper bound:

$$\begin{aligned} \mathbb{E}\left[\max_{x\in\mathcal{X}} \frac{(x_\ell - x)^\top\eta}{\epsilon_\ell + \Delta_x}\right] &= \mathbb{E}\left[\max\left\{\frac{x_\ell^\top\eta}{\epsilon_\ell+\epsilon} + \frac{\max_{i=1,\dots,m} -\eta_i}{\epsilon_\ell+\epsilon}, \frac{(x_\ell - \mathbf{1})^\top\eta}{\epsilon_\ell + \sqrt{m}}\right\}\right] \\ &\leq \frac{1}{\epsilon_\ell+\epsilon}\mathbb{E}[|x_\ell^\top\eta|] + \frac{1}{\epsilon_\ell+\epsilon}\mathbb{E}[\max_{i=1,\dots,m}|\eta_i|] + \frac{1}{\epsilon_\ell+\sqrt{m}}\mathbb{E}[|(x_\ell-\mathbf{1})^\top\eta|] \end{aligned}$$

Since x_ℓ is a candidate for the best arm at round ℓ, on the good event we must have that $\Delta_{x_\ell} \leq c\epsilon_\ell$. In particular, then, we will either have that $\|x_\ell\|_1 = 1$, or $\epsilon_\ell = O(\sqrt{m})$, so regardless of ℓ, $\frac{1}{\epsilon_\ell+\epsilon}\mathbb{E}[|x_\ell^\top\eta|] \leq c/\epsilon_\ell$. By Vershynin [2018], since each η_i has unit variance, we'll have $\mathbb{E}[\max_{i=1,\dots,m}|\eta_i|] \leq c\sqrt{\log(m)}$. Finally, noting that $x_\ell - \mathbf{1}$ has at most $c(m+\sqrt{m})$ non-zero entries, $(x_\ell - \mathbf{1})^\top\eta$ has variance bounded as $c(m+\sqrt{m})$, so $\mathbb{E}[|(x_\ell-\mathbf{1})^\top\eta|] \leq \mathcal{O}(\sqrt{m})$. We conclude that:

$$\mathbb{E}\left[\max_{x\in\mathcal{X}} \frac{(x_\ell - x)^\top\eta}{\epsilon_\ell + \Delta_x}\right] \leq \mathcal{O}\left(\frac{\sqrt{\log m}}{\epsilon_\ell}\right)$$

It follows that:

$$\tau \geq \mathcal{O}\left(\frac{\log(\ell^3/\delta)\log m}{\epsilon_\ell^2}\right)$$

is sufficient. Since this is a feasible solution, we'll then have that:

$$\mathcal{R}_\ell \leq \sum_{x\in\mathcal{X}} 2(\epsilon_\ell + \hat{\Delta}_x)\tau^*_{\ell,x} \leq \mathcal{O}\left((\epsilon_\ell + \sqrt{m})\frac{\log(\ell^3/\delta)\log m}{\epsilon_\ell^2}\right) \leq \mathcal{O}\left(\sqrt{m}\frac{\log(\ell^3/\delta)\log m}{\epsilon_\ell^2}\right)$$

where the last inequality holds since $\sqrt{m} = \Delta_{\max}$. Ignoring log factors that do not involve δ, and noting that there are at most $\log(\sqrt{m}/\epsilon)$ rounds, the total regret is bounded as:

$$\mathcal{O}\left(\sum_{\ell=1}^{\log(\sqrt{m}/\epsilon)} \frac{\sqrt{m}\log(1/\delta)}{\epsilon_\ell^2}\right) \leq \mathcal{O}\left(\frac{\sqrt{m}\log(1/\delta)}{m}4^{\log(\sqrt{m}/\epsilon)}\right) = \mathcal{O}\left(\frac{\sqrt{m}\log(1/\delta)}{\epsilon^2}\right)$$

Choosing $\delta = 1/T$ completes the proof.

$\square$

Failure of Thompson Sampling for semi-bandit feedback: We now provide a sketch as to why Thompson sampling fails on the instance in Proposition 1. Intuitively, Thompson Sampling is optimistic in a randomized fashion, so we would expect it to fail in the same way as optimistic algorithms. Slightly more formally, consider a typical version of Thompson sampling where at each round t, $\widetilde{\theta}_t \sim N(\widehat{\theta}_t, (\sum_{s=1}^{t-1} \mathrm{diag}(x_s x_s^\top))^{-1}))$ where x_s is the arm chosen at time s and $x_t = \arg\max_{x \in \mathcal{X}} x^\top \widetilde{\theta}_t$. Note that with high probability, we will have that:

$$|x^\top \widetilde{\theta}_t - x^\top \theta_*| \leq \sqrt{\alpha \|x\|^2_{(\sum_{s=1}^{t-1} \mathrm{diag}(x_s x_s^\top))^{-1}} \log(T)}$$

so we will essentially only pull an arm when $\sqrt{\alpha \|x\|^2_{(\sum_{s=1}^{t-1} \mathrm{diag}(x_s x_s^\top))^{-1}} \log(T)} > \Delta_x$. In the case of $\mathbf{1}$, we will have:

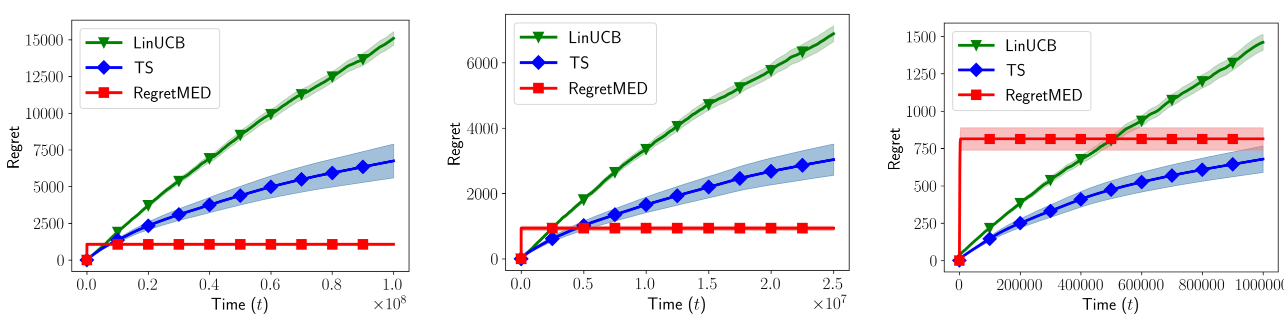
$$\|x\|^2_{(\sum_{s=1}^{t-1} \mathrm{diag}(x_s x_s^\top))^{-1}} \approx \frac{\sqrt{m}}{T_{m+1}}$$

where T_{m+1} are the total pulls of $\mathbf{1}$. Since $\Delta_{m+1} = \sqrt{m}$, the above inequality reduces to:

$$\sqrt{\frac{\alpha \sqrt{m} \log(T)}{T_{m+1}}} > \sqrt{m} \implies \frac{\log(T)}{\sqrt{m}} > T_{m+1}$$

so arm $\mathbf{1}$ will only be pulled a logarithmic number of times in T, which, as with optimism, is not sufficient to achieve optimal regret.

H Additional Experimental Results



(a) $\epsilon = 0.0005$ (b) $\epsilon = 0.001$ (c) $\epsilon = 0.005$

Figure 4: Regret against time plots for data points in Figure 3

We remark that, when running RegretMED, we do not use the exact constants specified in the algorithm. These constants are likely somewhat loose due to looseness in our analysis. In addition, we do not run the computationally efficient procedure derived formally but instead found that a much simpler heuristic—running stochastic Frank-Wolfe on the Lagrangian relaxation—works well in practice. We also do not use the precise value of $\Delta_{\max}$, and instead use an upper bound that can be computed using only knowledge of the arms.

The algorithms we compare against do not contain significant hyperparameters, and we choose reasonable values for the parameters they do require. In particular, for LinUCB, we use the regularization $\lambda = 1$.