A Proof of Theorem 1

In this section, we establish a proof of Theorem 1, and all the required lemmas stated in Section 4. We fix a finite $p \geq 1$, a $\sigma \in \mathbb{R}_+^G$, and $T \geq 2G$.

A.1 Properties of $R_{p,T}$

The goal of this section is to establish the necessary analytical properties of $R_{p,T}$. Let $\Delta(G)$ be the set of distributions on $[G]$, and $\mathcal{K}$ be the projection of $\Delta(G)$ in $\mathbb{R}_+^{G-1}$:

$$\mathcal{K} := \left\{ \lambda \in [0,1]^{G-1} | \sum_g \lambda_g \leq 1 \right\}. \tag{12}$$

$\mathcal{K}$ is also the $G-1$-dimensional unit simplexe. For each vector $n \in \mathbb{R}_+^G$ with $\sum_{t=1}^T n_g = T$, we set for each $g \in [G-1]$ $\lambda_g := \frac{n_g}{T}$. We have $\lambda \in \mathcal{K}$. Moreover, we have from (1):

$$R_{p,T}(n,\sigma) = \left\| \left\{ \frac{\sigma_g^2}{n_g} \right\}_{g=1}^G \right\|_p = \frac{1}{T} \left\| \left\{ \frac{\sigma_1^2}{\lambda_1}, \ldots, \frac{\sigma_{G-1}^2}{\lambda_{G-1}}, \frac{\sigma_G^2}{1-\lambda_1-\ldots-\lambda_{G-1}} \right\} \right\|_p. \tag{13}$$

which motivates us to introduce the re-scaled function r:

$$r_p : \lambda \in \mathcal{K} \to r(\lambda) := \left\| \left\{ \frac{\sigma_1^2}{\lambda_1}, \ldots, \frac{\sigma_{G-1}^2}{\lambda_{G-1}}, \frac{\sigma_G^2}{1-\lambda_1-\ldots-\lambda_{G-1}} \right\} \right\|_p. \tag{14}$$

so that (13) can be written as

$$R_{p,T}(n,\sigma) = \frac{1}{T} r_p(\lambda). \tag{15}$$

From the definition of $\mathcal{K}$ in (12), the interior of $\mathcal{K}$, denoted $\mathcal{K}^\circ$, is

$$\mathcal{K}^\circ = \left\{ \lambda \in \mathcal{K} | \forall g \in [G-1], \lambda_g > 0, \quad \sum_{g \in [G-1]} \lambda_g < 1 \right\}.$$

Moreover, the function r is $\mathcal{C}^3$ in $\mathcal{K}^\circ$. From Taylor's theorem, we have for $\lambda, \lambda' \in \mathcal{K}^\circ$:

$$\left| r_p(\lambda) - r_p(\lambda') - (\lambda-\lambda')\nabla r_p(\lambda') - \frac{1}{2}\langle \mathcal{H}(\lambda')(\lambda-\lambda'), \lambda-\lambda' \rangle \right| \leq \|\lambda-\lambda'\|_\infty^3 \sup_{\substack{u \in [\lambda,\lambda'] \\ x,y,z \in \mathbb{N} \\ x+y+z=3 \\ \{g,h,i\} \subset [G-1]}} \left| \frac{1}{x!y!z!} \frac{\partial^3 r_p(u)}{\partial \lambda_g \partial \lambda_h \partial \lambda_i} \right| \tag{16}$$

where $\nabla, \mathcal{H}$ are respectively the gradient and hessian operators. In what follows, we will derive expressions for the relevant derivatives of r in (16).

Lemma 7. *For $g, h, i \in [G-1]$, we introduce the following functions in $\mathcal{K}^\circ$:*

$$H_g : \lambda \in \mathcal{K}^\circ \to \frac{\sigma_G^{2p}}{(1-\lambda_1-\ldots-\lambda_{G-1})^{p+1}} - \frac{\sigma_g^{2p}}{\lambda_g^{p+1}}$$
$$G_{h,g} := \frac{1}{p+1} \frac{\partial}{\partial \lambda_g} H_h$$
$$I_{g,h,i} := \frac{1}{p+2} \frac{\partial}{\partial \lambda_i} G_{g,h}$$

The following holds:

1. $\nabla r_p = r_p^{1-p}(H_1, \ldots, H_{G-1})$
2. $\mathcal{H}_{g,h} = (1-p) r_p^{1-2p} H_g H_h + (p+1) r_p^{1-p} G_{g,h}$

3.

$$\frac{\partial^3}{\partial\lambda_g\partial\lambda_h\partial_i} r_p = (1-p)(1-2p)H_g H_h H_i r_p^{1-3p} \\ + (1-2p)(1+p)\left(H_g G_{h,i} + H_h G_{i,g} + H_g G_{h,i}\right) \\ + r_p^{1-2p} + (p+1)(p+2) I_{g,h,i} r_p^{1-p}$$

Proof. Fix a $\lambda \in \mathcal{K}^\circ$ and $g, h, i \in [G-1]$.

Expression of the gradient: On the one hand, we have from the definition of r_p in (14)

$$r_p^p(\lambda) = \left\| \left\{ \frac{\sigma_1^2}{\lambda_1}, \dots, \frac{\sigma_{G-1}^2}{\lambda_{G-1}}, \frac{\sigma_G^2}{1-\lambda_1-\ldots-\lambda_{G-1}} \right\} \right\|_p^p = \sum_{h \leq G-1} \frac{\sigma_h^{2p}}{\lambda_h^p} + \frac{\sigma_G^{2p}}{(1-\lambda_1-\ldots-\lambda_{G-1})^p},$$

so that

$$\frac{\partial}{\partial\lambda_g}(r_p^p)(\lambda) = p\frac{\sigma_G^{2p}}{(1-\lambda_1-\ldots-\lambda_{G-1})^{p+1}} = pH_g(\lambda) \tag{17}$$

On the other hand, from the formula $(f^p)' = of' f^{p-1}$, we have

$$\frac{\partial}{\partial\lambda_g}(r_p^p)(\lambda) = p r_p^{p-1}(\lambda) \frac{\partial}{\partial\lambda_g} r_p(\lambda) \tag{18}$$

By combining Equations (17) and (18), we obtain

$$\frac{\partial}{\partial\lambda_g} r_p^p = pH_g = p r_p^{p-1} \frac{\partial}{\partial\lambda_g} r_p$$

so that

$$\frac{\partial}{\partial\lambda_g} r_p = r_p^{1-p} H_g \tag{19}$$

Therefore $\nabla r_p = \left(\frac{\partial}{\partial\lambda_1} r_p, \dots, \frac{\partial}{\partial\lambda_{G-1}} r_p\right) = r_p^{1-p}(H_1, \dots, H_{G-1})$, which proves the first equation of Lemma 7.

Expression of the Hessian: We have

$$\begin{aligned} \mathcal{H}_{g,h} &= \frac{\partial^2}{\partial\lambda_g\partial\lambda_h} r_p \\ &= \frac{\partial}{\partial\lambda_g}\left(r_p^{1-p} H_h\right) \\ &= (1-p) r_p^{-p} \left(\frac{\partial}{\partial\lambda_g} r_p\right) H_h + r_p^{1-p} \frac{\partial}{\partial\lambda_g} H_h \\ &= (1-p) r_p^{1-2p} H_g H_h + (p+1) r_p^{1-p} G_{g,h}. \end{aligned}$$

where the first equality is due to the definition of the Hessian, the second equality is due to the expression of the gradient from Equation (19), the third equality applies the product rule to the derivative, and the fourth equality applies the definition of $G_{g,h}$ in Lemma 7. This proves the second Equality of Lemma 7.

Third derivatives. Finally, we have,

$$\begin{aligned} \frac{\partial^3}{\partial\lambda_g\partial\lambda_h\partial_i} r_p &= \frac{\partial}{\partial\lambda_i}\mathcal{H}_{g,h} \\ &= \frac{\partial}{\partial\lambda_i}\left\{(1-p) r_p^{1-2p} H_g H_h + (p+1) r_p^{1-p} G_{g,h}\right\} \\ &= (1-p)\frac{\partial}{\partial\lambda_i}\left\{r_p^{1-2p} H_g H_h\right\} + (p+1)\frac{\partial}{\partial\lambda_i}\left\{r_p^{1-p} G_{g,h}\right\} \end{aligned}$$

where the first equality is due to the definition of the Hessian, the second inequality is due to the Hessian expression established previously, and the third equality is due to the linearity of derivation. In a similar fashion to the previous case, we derivate each of the products $r_p^{1-2p} H_g H_h$ and $r_p^{1-p} G_{g,h}$ separately. On the one hand,

$$\begin{aligned}\frac{\partial}{\partial \lambda_i}\{r_p^{1-2p} H_g H_h\} &= H_g H_h \frac{\partial}{\partial \lambda_i}\{r_p^{1-2p}\} + r_p^{1-2p}\left(H_g \frac{\partial}{\partial \lambda_i} H_h + H_h \frac{\partial}{\partial \lambda_i} H_g\right)\\ &= H_g H_h (1-2p) r_p^{-2p} r_p^{1-p} H_i + r_p^{1-2p}\left(H_g (p+1) G_{h,i} + H_h (p+1) G_{g,i}\right)\\ &= (1-2p) r_p^{1-3p} H_g H_h H_i + (p+1) r_p^{1-2p}\left(H_g G_{h,i} + H_h G_{g,i}\right),\end{aligned}$$

where the first equality is due to the derivation product rule, the second equality is due to the definition of $G_{g,h}$ introduced in Lemma 7, and the third equality is due to a reordering of the terms. On the other hand, we have by folowing the exact same steps

$$\begin{aligned}\frac{\partial}{\partial \lambda_i}\left\{r_p^{1-p} G_{g,h}\right\} &= (1-p) r_p^{-p} \frac{\partial}{\partial \lambda_i} r_p + r_p^{1-p} \frac{\partial}{\partial \lambda_i} G_{g,h}\\ &= (1-p) r_p^{-p} r_p^{1-p} H_i G_{g,h} + (p+2) r_p^{1-p} I_{g,h,i}\\ &= (1-p) r_p^{1-2p} H_i G_{g,h} + (p+2) r_p^{1-p} I_{g,h,i}.\end{aligned}$$

Then, we replace the previous two expressions in the formula for $\frac{\partial^3}{\partial \lambda_g \partial \lambda_h \partial_i} r_p$ to obtain:

$$\begin{aligned}\frac{\partial^3}{\partial \lambda_g \partial \lambda_h \partial_i} r_p &= (1-p)(1-2p) H_g H_h H_i r_p^{1-3p}\\ &\quad + (1-2p)(1+p)\left(H_g G_{h,i} + H_h G_{i,g} + H_g G_{h,i}\right)\\ &\quad + r_p^{1-2p} + (p+1)(p+2) I_{g,h,i} r_p^{1-p}\end{aligned}$$

where we use the symmetry of $G_{g,h} = G_{h,g}$. This derives the third equation of Lemma 7 and concludes the proof. □

A.2 Proof of Lemma 1

The proof of Lemma 1 consists of using optimality conditions on r_p:

Lemma 1. *[Benchmark analysis] For each $t \in \mathbb{N}$ and $p \in [1, +\infty]$, let $n_{g,t}^* := \frac{\sigma_g^{\frac{2p}{p+1}} t}{\sum_{h \in [G]} \sigma_h^{\frac{2p}{p+1}}}$. Then,*

$$R_{p,t}^*(\sigma) = R_{p,t}(n_t^*, \sigma) = \frac{1}{t} R_{p,t}(n_1^*, \sigma). \tag{9}$$

Proof. From the definition of r_p in (14), we have

$$R_{p,T}^*(\sigma) = \inf_{\substack{n \in \mathbb{R}_+^G\\ \sum_{t=1}^T n_g = T}} R_{p,T}(n, \sigma) = \frac{1}{T} \inf_{\lambda \in \mathcal{K}} r_p(\lambda).$$

For any $\lambda \in \mathcal{K} - \mathcal{K}^\circ$, at least one element of the set $\{\lambda_1, \ldots, \lambda_{G-1}, 1 - \lambda_1 - \ldots - \lambda_{G-1}\}$ should be 0, therefore for such a λ, we must have $r_p(\lambda) = +\infty$. Therefore $\operatorname{argmin}_{\mathcal{K}} r_p = \operatorname{argmin}_{\mathcal{K}^\circ} r_p$. Moreover, since r_p is differentiable in $\mathcal{K}^\circ$, we must also have $\operatorname{argmin}_{\mathcal{K}^\circ} r_p \subset \nabla r_p^{-1}(\{0\})$. Therefore,

$$\begin{aligned}\lambda \in \nabla r_p^{-1}(\{0\}) &\iff r_p^{1-p}(\lambda)(H_1(\lambda), \ldots, H_{G-1}(\lambda)) = 0_{G-1}\\ &\iff H_1(\lambda) = \ldots = H_{G-1}(\lambda)) = 0\\ &\iff \frac{\sigma_1^{2p}}{\lambda_1^{p+1}} = \ldots = \frac{\sigma_{G-1}^{2p}}{\lambda_{G-1}^{p+1}} = \frac{\sigma_G^{2p}}{(1 - \lambda_1 - \ldots \lambda_{G-1})^{p+1}}\\ &\iff \forall g \in [G-1], \quad \lambda_g = \frac{\sigma_g^{\frac{2p}{p+1}}}{\sum_h \sigma_h^{\frac{2p}{p+1}}}.\end{aligned}$$

where the first equivalence is due to the expression of the gradient from Lemma 7, the second equivalence is due to $r_p > 0$, the third equivalence is due to the definition of H introduced in Lemma 7, and where we solve the system of equations in the last equivalence. Therefore there is a unique solution to $\nabla r_p(\lambda) = 0$. By continuity of r_p in the compact set $\mathcal{K}$, $\text{argmin}_{\mathcal{K}} r_p \neq \emptyset$. Therefore the minimum of r_p is attained exactly once at the point $\left\{ \frac{\sigma_g^{\frac{2p}{p+1}}}{\sum_h \sigma_h^{\frac{2p}{p+1}}} \right\}_{g\in[G-1]}$. Therefore by setting

$$\lambda_g^* := \frac{\sigma_g^{\frac{2p}{p+1}}}{\sum_h \sigma_h^{\frac{2p}{p+1}}}, \quad \forall g \in [G-1]$$

$$\lambda_G^* = 1 - \lambda_1^* - \ldots - \lambda_{G-1}^* = \frac{\sigma_G^{\frac{2p}{p+1}}}{\sum_h \sigma_h^{\frac{2p}{p+1}}}$$

we have $\lambda^* \in (\lambda - \lambda^*)_G$ and $r_p(\{\lambda_g^*\}_{g\in[G-1]}) = \inf_{\lambda\in\mathcal{K}} r_p(\lambda) = TR^*_{p,T}(\sigma)$. Lemma 1 follows by setting $n^* := T\lambda^*$. $\square$

A.3 Proof of Lemma 2

In this section, we establish the properties of $\text{UCB}_t(\sigma_g)$, which essentially from Assumption 1. Following the notations of Section 3, we introduce the following event:

$$\mathcal{A}_T := \bigcap_{g\in[G],2\leq t\leq T} \left\{ |\hat{\sigma}_{g,t} - \sigma_g| \leq \frac{C_T}{\sqrt{t}} \right\}$$

Based on Corollary 1 of [13], we have $\mathbb{P}_\pi(\mathcal{A}_T) \geq 1 - 2GT^{-2.5}$, and conditionally on $\mathcal{A}_T$,

$$\forall g \in [G], t \geq 2G, \quad |\hat{\sigma}_{g,t} - \sigma_g| \leq \frac{C_T}{\sqrt{n_{g,t}}}.$$

We are now ready to prove Lemma 2.

Lemma 2. *With probability at least* $1 - \tilde{O}(T^{-2})$,

$$0 \leq \mathit{UCB}_t(\sigma_g)^{\frac{2p}{p+1}} - \sigma_g^{\frac{2p}{p+1}} \leq \frac{4C_T}{\sqrt{n_{g,t}}} \frac{p}{p+1} \left(\sigma_g + \frac{2C_T}{\sqrt{n_{g,t}}} \right)^{\frac{p-1}{p+1}}.$$

Proof. Conditionally on $\mathcal{A}_T$, we have:

$$\begin{aligned}
\text{UCB}_t(\sigma_g) &= \hat{\sigma}_{g,t} + \frac{C_T}{\sqrt{n_{g,t}}} \\
&= \sigma_g + \frac{C_T}{\sqrt{n_{g,t}}} + (\hat{\sigma}_{g,t} - \sigma_g) \\
&\in \sigma_g + \frac{C_T}{\sqrt{n_{g,t}}} + \left[-\frac{C_T}{\sqrt{n_{g,t}}}, \frac{C_T}{\sqrt{n_{g,t}}} \right] \\
&= \sigma_g + \left[0, \frac{2C_T}{\sqrt{n_{g,t}}} \right],
\end{aligned}$$

where the first equation is due to the definition of UCB_t introduced in (7), and the bounding is due to the definition of $\mathcal{A}_T$. Next, notice that the function $x \to (\sigma_g + x)^{\frac{2p}{p+1}}$ is increasing, which implies that

$$\sigma_g^{\frac{2p}{p+1}} \leq \text{UCB}_t(\sigma_g)^{\frac{2p}{p+1}} \leq \left(\sigma_g + \frac{2C_T}{\sqrt{n_{g,t}}} \right)^{\frac{2p}{p+1}}, \tag{20}$$

which proves the leftmost inequality in Lemma 2. To prove the rightmost inequality, notice that $x \to (\sigma_g + x)^{\frac{2p}{p+1}}$ is also convex, therefore by Jensen's inequality

$$\left(\sigma_g + \frac{2C_T}{\sqrt{n_{g,t}}} \right)^{\frac{2p}{p+1}} - \sigma_g^{\frac{2p}{p+1}} \leq \frac{2C_T}{\sqrt{n_{g,t}}} \frac{2p}{p+1} \left(\sigma_g + \frac{2C_T}{\sqrt{n_{g,t}}} \right)^{\frac{2p}{p+1}-1} = \frac{4C_T}{\sqrt{n_{g,t}}} \frac{p}{p+1} \left(\sigma_g + \frac{2C_T}{\sqrt{n_{g,t}}} \right)^{\frac{p-1}{p+1}},$$

which concludes the proof. $\square$

A.4 Proof of Lemma 3

The goal of this section is to prove Lemma 3, which consists of bounding with high probability $n - n^*$. To do so, we will design an alternative sequence $\tilde{n}$, that is simultaneously easy to analyze, and upper bounds n with high probability. The motivation for the choice of $\tilde{n}$ comes from the choice of Variance-UCB.

We assume through the whole section what the event $\mathcal{A}_T$ is achieved. For convenience, we view the right hand side of Lemma 2 as a quantity of its own, and introduce the width function

$$w_g : x > 0 \to w_g(x) := \frac{4C_T}{\sqrt{x}} \frac{p}{p+1} \left(\sigma_g + \frac{2C_T}{\Sigma_p \sqrt{x}} \right)^{\frac{p-1}{p+1}},$$

so that the inequality in Lemma 2 can be rewritten as

$$0 \le \mathrm{UCB}_t(\sigma_g)^{\frac{2p}{p+1}} - \sigma_g^{\frac{2p}{p+1}} \le \Sigma_p w_g(n_{g,t}). \tag{21}$$

We start by proving the following lemma, which follows from the decisions Variance-UCB makes:

Lemma 8. *For $t \ge 2G$, we have:*

$$n_{X_{t+1},t} - t w_{X_{t+1}}(n_{X_{t+1},t}) \le n^*_{X_{t+1},t}$$

Proof. The proof exploits the greedy property of the algorithm. At $t = 2G$, each group is sampled exactly twice, and we have $\mathrm{UCB}_{2G}(\sigma_g)^{\frac{2p}{p+1}} < +\infty$ for every group $g \in [G]$. By choice of X_{t+1}, we must have:

$$\forall g \in [G], \quad \frac{\mathrm{UCB}_t(\sigma_g)^{\frac{2p}{p+1}}}{n_{g,t}} \le \frac{\mathrm{UCB}_t(\sigma_{X_{t+1}})^{\frac{2p}{p+1}}}{n_{X_{t+1},t}} \tag{22}$$

On the one hand, from the leftmost inequality in (21), we have

$$\forall g \in [G], \quad \frac{\sigma_g^{\frac{2p}{p+1}}}{n_{g,t}} \le \frac{\mathrm{UCB}_t(\sigma_g)^{\frac{2p}{p+1}}}{n_{g,t}}. \tag{23}$$

On the other hand, from the rightmost inequality in (21), we have

$$\frac{\mathrm{UCB}_t(\sigma_{X_{t+1}})^{\frac{2p}{p+1}}}{n_{X_{t+1},t}} \le \frac{\sigma_{X_{t+1}}^{\frac{2p}{p+1}} + \Sigma_p w_{X_{t+1}}(n_{X_{t+1},t})}{n_{X_{t+1},t}}. \tag{24}$$

Therefore by combining both Inequalities (23) and (24) in Inequality (22), we obtain

$$\forall g \in [G], \quad \frac{\sigma_g^{\frac{2p}{p+1}}}{n_{g,t}} \le \frac{\sigma_{X_{t+1}}^{\frac{2p}{p+1}} + \Sigma_p w_{X_{t+1}}(n_{X_{t+1},t})}{n_{X_{t+1},t}} \tag{25}$$

After multiplying both sides by $n_{g,t} n_{X_{t+1},t}$ and summing over $g \in [G]$, we get:

$$n_{X_{t+1},t} \underbrace{\sum_{g \in [G]} \sigma_g^{\frac{2p}{p+1}}}_{=\Sigma_p} \le \left(\sigma_{X_{t+1}}^{\frac{2p}{p+1}} + \Sigma_p w_{X_{t+1}}(n_{X_{t+1},t}) \right) \underbrace{\sum_{g \in [G]} n_{g,t}}_{=t}$$

we then divide both sides by $\Sigma_p > 0$, and use the formula $n^*_{g,t} = \frac{\sigma_g^{\frac{2p}{p+1}}}{\Sigma_p} t$ (see Lemma 1) to obtain

$$n_{X_{t+1},t} \le n^*_{X_{t+1},t} + t w_{X_{t+1}}(n_{X_{t+1},t})$$

Lemma 8 follows by substracting $w_{X_{t+1}}(n_{X_{t+1},t})$ from both sides. ☐

Lemma 8 states that the possible excess between the number of samples output by the algorithm and the optimal number of samples is not too big, and can be controlled by the width w. Since the width decreases in the number of samples, the function

$$x \to x - t w_g(x)$$

must be increasing and has therefore an inverse function that is also increasing, which we denote $W_g^t(x)$. We introduce the following sequence, which mimics the behavior stated in Lemma 8:

$$\begin{aligned} \tilde{n}_{g,t} &= n_{g,t} && \text{For } t = 1, \dots, 2G \\ \tilde{n}_{g,t+1} &= \tilde{n}_{g,t} + \mathbb{1}\left(\tilde{n}_{g,t} \leq W_g^t\left(n_{g,t}^*\right)\right) && \text{For } t \geq 2G \end{aligned}$$

The sequence is easier to analyze and upper bounds the true number of samples n:

Lemma 9. *We have $\tilde{n} \geq n$.*

Proof. By construction, the result holds for $t = 1, \dots, 2G$. Assume for the sake of contradiction that the result does not hold for a $g \in [G]$ and $t + 1 > 2G$, and take such a t minimal. For such a pair (g, t) We have:

$$\begin{aligned} \mathbb{1}(X_{t+1} = g) &= n_{g,t+1} - n_{g,t} \\ &> \tilde{n}_{g,t+1} - \tilde{n}_{g,t} \\ &= \mathbb{1}(\tilde{n}_{g,t} \leq W_g^t(n_{g,t}^*)) \\ &= \mathbb{1}(n_{g,t} \leq W_g^t(n_{g,t}^*)), \end{aligned}$$

where the first step follows from the definition of n, the second step follows from the minimality of t, the third step follows from the definition of $\tilde{n}$, and the last step follows from the minimality of t.

Therefore $1 \geq \mathbb{1}(X_{t+1} = g) > \mathbb{1}(n_{g,t} \leq W_g^t(n_{g,t}^*)) \geq 0$, which implies $\mathbb{1}(X_{t+1} = g) = 1$ and $\mathbb{1}(n_{g,t} \leq W_g^t(n_{g,t}^*)) = 0$, so that

$$n_{X_{t+1},t} > W_g^t(n_{X_{t+1},t}^*).$$

By taking the inverse of the increasing function W_g^t on both sides in the previous equality, we get

$$n_{X_{t+1},t} - t w_{X_{t+1}}(n_{X_{t+1},t}) > n_{X_{t+1},t}^*,$$

contradicting Lemma 8. Therefore the assumption is wrong and $\tilde{n} \geq n$, which completes the proof. ☐

Lemma 10. *For a fixed g,the sequence $\{W_g^t(n_{g,t}^*)\}_{t \geq 1}$ is increasing. Consequently,*

$$\tilde{n}_{g,t} \leq W_g^t(n_{g,t}^*)^+ + 2$$

Proof. For a fixed $x > 0$ and $t \geq 1$, we have

$$(x - (t+1)w_g(x)) - (x - t w_g(x)) = -w_g(x) < 0,$$

therefore the sequence of functions $\{x \to x - t w_g(x)\}_{t \geq 1}$ is simply decreasing in t, therefore the sequence of its inverse functions $\{x \to W_g^t(x)\}_{t \geq 1}$ is simply increasing in t. Consequently,

$$W_g^{t+1}(n_{g,t+1}^*) \geq W_g^t(n_{g,t+1}^*). \tag{26}$$

Moreover, the function W_g^t is increasing in $\mathbb{R}_+$, thus

$$W_g^t(n_{g,t+1}^*) \geq W_g^t(n_{g,t}^*). \tag{27}$$

By combining Equations (26) and (27), we obtain

$$W_g^{t+1}(n_{g,t+1}^*) \geq W_g^t(n_{g,t}^*),$$

which proves that the sequence $\{W_g^t(n_{g,t}^*)\}_{t \geq 1}$ is increasing, thus completing the proof for the first part of Lemma 10.

We derive the second part by induction on $t \geq 1$. For $t \leq 2G$, the result holds immediately as $\tilde{n}_{g,t} = n_{g,t}$ and $n_{g,t} \leq 2$. We assume the result holds for a $t \geq 2G$. We distinguish two cases:

- $\tilde{n}_{g,t} \leq W_g^t\left(n_{g,t}^*\right)$: This implies that:

$$\begin{aligned}\tilde{n}_{g,t+1} &= \tilde{n}_{g,t} + 1\\ &\leq W_g^t\left(n_{g,t}^*\right) + 1\\ &\leq W_p^{t+1,g}\left(n_{g,t+1}^*\right) + 1\\ &\leq W_p^{t+1,g}\left(n_{g,t+1}^*\right)^+ + 2,\end{aligned}$$

  where the first step stems from the definition of $\tilde{n}$, the second step stems from the induction hypothesis, and the third step stems from the first part of the proof.

- $\tilde{n}_{g,t} > W_g^t\left(n_{g,t}^*\right)$: This implies that:

$$\begin{aligned}\tilde{n}_{g,t+1} &= \tilde{n}_{g,t}\\ &\leq W_g^t\left(n_{g,t}^*\right)^+ + 2\\ &\leq W_p^{t+1,g}\left(n_{g,t+1}^*\right)^+ + 2,\end{aligned}$$

  where the first step stems from the definition of $\tilde{n}$, the second step stems from the induction hypothesis, and the third step stems from the first part of the proof.

Studying both cases concludes our induction and proves the inequality for all $t \geq 1$. This concludes the proof of Lemma 10. □

Next, we derive an upper bound on W_g^t:

Lemma 11. *For $x > 0$, we have $W_g^t(x) \leq \frac{x}{1-t\frac{w_g(x)}{x}}$.*

Proof. Let $x, y > 0$ with $0 \leq y - tw_g(y) = x$, so that $y = W_g^t(x)$ by definition of W_g^t. we have:

$$\begin{aligned}\frac{x}{1-t\frac{w_g(x)}{x}} &= \frac{y - tw_g(y)}{1 - t\frac{w_g(y-tw_g(y))}{x-tw_g(x)}}\\ &= y\frac{1-t\frac{w_g(y)}{y}}{1-t\frac{w_g(y-tw_g(y))}{y-tw_g(y)}}\\ &\geq y = W_g^t(x),\end{aligned}$$

which concludes the proof. □

We are now ready to prove Lemma 3.

Lemma 3. *Variance-UCB collects a vector of samples n that, with probability at least $1 - \tilde{O}(T^{-2})$, satisfies,*

$$n_{g,T} - n_{g,T}^* \leq 3 + \frac{4C_T p}{\Sigma_p(p+1)}\left(\sigma_g + \frac{2C_T}{\sqrt{n_{g,T}^*}}\right)^{\frac{p-1}{p+1}}\sqrt{n_{g,T}^*} = \Theta(\sqrt{T}).$$

Proof. Conditionally on $\mathcal{A}_T$, all the previous lemmas stated in this section hold. Consequently,

$$\begin{aligned}n_{g,T} &\leq \tilde{n}_{g,T}\\ &\leq 2 + W_g^t(n_{g,T}^*)^+\\ &\leq 2 + \frac{n_{g,T}^*}{1 - T\frac{w_g(n_{g,T}^*)}{n_{g,T}^*}}\\ &\leq 2 + n_{g,T}^* + 1 + Tw_g(n_{g,T}^*)\\ &= 3 + n_{g,T}^* + \frac{4C_T p}{\Sigma_p(p+1)}\left(\sigma_g + \frac{2C_T}{\sqrt{n_{g,T}^*}}\right)^{\frac{p-1}{p+1}}\sqrt{n_{g,T}^*},\end{aligned}$$

where the first step stems from Lemma 9, the second step stems from Lemma 10, the third step stems from first order approximations and the last step stems from Lemma 11. □

A.5 Proof of Lemma 4

We follow the notations from Appendix A.1. The goal of this section if to prove Lemma 4. We will do so by combining the optimality of λ^* and by using the properties of $R_{p,T}$ established in A.1.

Lemma 4. *Let $\sigma \in \mathbb{R}_+^G$ and $n' \in \mathbb{R}_+^G$ such that $\sum_{g\in[G]} n'_g = T$. Then,*

$$\frac{R_{p,T}(n',\sigma) - R_{p,T}(n^*,\sigma)}{R_{p,T}(n^*,\sigma)} \leq \frac{p+1}{2}\sum_{g\in[G]}\frac{(n'_g - n^*_{g,T})^2}{Tn^*_{g,T}} + \frac{7(p+2)^2\Sigma_p^2}{\sigma_{\min}^2}\max_g\left(\frac{n^*_{g,T}}{n'_g}\right)^{3p+3}\frac{\|n'-n^*\|_\infty^3}{T^3}$$
$$= O\left(\frac{\|n-n^*\|_\infty^2}{T^2}R_p(\sigma)\right).$$

Proof. For $\lambda \in \mathcal{K}$ be fixed, from Taylor's inequality, we have

$$\left|r_p(\lambda) - r_p(\lambda^*) - \langle \lambda-\lambda^*, \nabla r_p(\lambda^*)\rangle - \frac{1}{2}\langle \mathcal{H}(\lambda^*)(\lambda-\lambda^*), \lambda-\lambda^*\rangle\right| \leq \|\lambda-\lambda^*\|_\infty^3 \sup_{\substack{u\in[\lambda,\lambda^*]\\ x,y,z\in\mathbb{N}\\ x+y+z=3\\ \{g,h,i\}\subset[G-1]}}\left|\frac{1}{x!y!z!}\frac{\partial^3 r_p(u)}{\partial\lambda_g\partial\lambda_h\partial\lambda_i}\right|$$

We use the derivative formulas obtained in 7 to obtain simple forms for $\nabla r_p(\lambda^*)$, $\mathcal{H}(\lambda^*)$, and $\frac{\partial^3 r_p(u)}{\partial\lambda_g\partial\lambda_h\partial\lambda_i}$. We start by upperbounding $\frac{\partial^3 r_p(u)}{\partial\lambda_g\partial\lambda_h\partial\lambda_i}$. From Lemma 1, we have for each $g \in [G]$:

$$\sigma_g^{2p} = \left(\sum_{h\in[G]}\sigma_h^{\frac{2p}{p+1}}\right)^{p+1}(\lambda_g^*)^{p+1} = \Sigma_p^{p+1}(\lambda_g^*)^{p+1}.$$

For a fixed $u \in [\lambda,\lambda^*]$, $\sum_g u_g = \sum_g \lambda_g = \sum_g \lambda_g^* = 1$, thus the coordinate g_0 achieving the maximal $\frac{\lambda_g^*}{u_g}$ must have $u_{g_0} \leq \lambda_{g_0}^*$. Since $u \in [\lambda,\lambda^*]$ this implies that $\lambda_{g_0} \leq u_{g_0} \leq \lambda_{g_0}^*$ and consequently $\max_g \frac{\lambda_g^*}{u_g} \leq \max_g \frac{\lambda_g^*}{\lambda_g}$. Hence we get:

$$\begin{aligned}
|H_g(u)| &= \Sigma_p^{p+1}\left|\left(\frac{\lambda_G^*}{u_G}\right)^{p+1} - \left(\frac{\lambda_g^*}{u_g}\right)^{p+1}\right| && \leq \Sigma_p^{p+1}\left(\max_g\frac{\lambda_g^*}{\lambda_g}\right)^{p+1},\\
|G_{g,h}(u)| &= \Sigma_p^{p+1}\left|\frac{(\lambda_G^*)^{p+1}}{(u_G)^{p+2}} + \frac{(\lambda_g^*)^{p+1}}{(u_g)^{p+2}}\mathbb{1}(g=h)\right| && \leq \frac{2\Sigma_p^{p+1}}{\min_g\lambda_g^*}\left(\max_g\frac{\lambda_g^*}{\lambda_g}\right)^{p+2,}\\
|I_{g,h,i}(u)| &= \Sigma_p^{p+1}\left|\frac{(\lambda_G^*)^{p+1}}{(u_G)^{p+3}} - \frac{(\lambda_g^*)^{p+1}}{(u_g)^{p+3}}\mathbb{1}(g=h=i)\right| && \leq \frac{\Sigma_p^{p+1}}{(\min_g\lambda_g^*)^2}\left(\max_g\frac{\lambda_g^*}{\lambda_g}\right)^{p+3}.
\end{aligned}$$

Moreover, since $p \geq 1$, all $1-p, 1-2p, 1-3p$ are non-positive, and for $j \in 1-p, 1-2p, 1-3p$ we have from the minimality of λ^*:

$$r_p(u)^j \leq r_p(\lambda^*)^j = (\Sigma_p)^{\frac{1}{p}+1},$$

therefore:

$$\begin{aligned}
|H_gH_hH_ir_p^{1-3p}(u)| &\leq \left(\max_g \frac{\lambda_g^*}{\lambda_g}\right)^{3p+3} \Sigma_p^{3p+3} r_p(\lambda^*)^{1-3p} \\
&= \left(\max_g \frac{\lambda_g^*}{\lambda_g}\right)^{3p+3} \Sigma_p^{3p+3+(1-3p)(1+1/p)} \\
&= \left(\max_g \frac{\lambda_g^*}{\lambda_g}\right)^{3p+3} \Sigma_p^{1+1/p} \\
&= \left(\max_g \frac{\lambda_g^*}{\lambda_g}\right)^{3p+3} r_p(\lambda^*) \\
|\left(H_gG_{h,i} + H_hG_{i,g} + H_gG_{h,i}\right) r_p^{1-2p}(u)| &\leq 3\frac{2\Sigma_p^{p+1}}{\min_g \lambda_g^*} \left(\max_g \frac{\lambda_g^*}{\lambda_g}\right)^{p+2} \left(\Sigma_p^{p+1} \left(\max_g \frac{\lambda_g^*}{\lambda_g}\right)^{p+1}\right) \\
&= \frac{6}{\min_g \lambda_g^*} \left(\max_g \frac{\lambda_g^*}{\lambda_g}\right)^{2p+3} \Sigma_p^{2p+2} r_p^{1-2p}(\lambda^*) \\
&= \frac{6}{\min_g \lambda_g^*} \left(\max_g \frac{\lambda_g^*}{\lambda_g}\right)^{2p+3} r_p(\lambda^*, \sigma) \\
|I_{g,h,i} r_p^{1-p}(u)| &\leq \frac{\Sigma_p^{p+1}}{(\min_g \lambda_g^*)^2} \left(\max_g \frac{\lambda_g^*}{\lambda_g}\right)^{p+3} r_p^{1-p}(\lambda^*) \\
&= \frac{1}{(\min_g \lambda_g^*)^2} \left(\max_g \frac{\lambda_g^*}{\lambda_g}\right)^{p+3} r_p(\lambda^*, \sigma)
\end{aligned}$$

Hence by using the expression of the third derivatives, we get:

$$\begin{aligned}
\left|\frac{\partial^3 r_p(u)}{\partial\lambda_g\partial\lambda_h\partial\lambda_i}\right| &\leq \left|(1-p)(1-2p)H_gH_hH_ir_p^{1-3p}(u)\right| \\
&+ \left|(1-2p)(1+p)\left(H_gG_{h,i} + H_hG_{i,g} + H_gG_{h,i}\right) r_p^{1-2p(u)}\right| \\
&+ \left|(p+1)(p+2)I_{g,h,i}r_p^{1-p}(u)\right| \\
&\leq 2p^2 \left(\max_g \frac{\lambda_g^*}{\lambda_g}\right)^{3p+3} r_p(\lambda^*) + \frac{12(p+1)^2}{\min_g \lambda_g^*} \left(\max_g \frac{\lambda_g^*}{\lambda_g}\right)^{2p+3} r_p(\lambda^*, \sigma) \\
&+ \frac{(p+2)^2}{(\min_g \lambda_g^*)^2} \left(\max_g \frac{\lambda_g^*}{\lambda_g}\right)^{p+3} r_p(\lambda^*, \sigma) \\
&\leq \frac{7(p+2)^2}{(\min_g \lambda_g^*)^2} \left(\max_g \frac{\lambda_g^*}{\lambda_g}\right)^{3p+3} r_p(\lambda^*).
\end{aligned}$$

As a consequence, by taking the sup over u, g, h, i, x, y, z, we get

$$\sup_{\substack{u\in[\lambda,\lambda^*] \\ x,y,z\in\mathbb{N} \\ x+y+z=3 \\ \{g,h,i\}\subset[G-1]}} \left|\frac{1}{x!y!z!}\frac{\partial^3 r_p(u)}{\partial\lambda_g\partial\lambda_h\partial\lambda_i}\right| \leq \frac{7(p+2)^2}{(\min_g \lambda_g^*)^2} \max_g \left(\frac{\lambda_g^*}{\lambda_g}\right)^{3p+3} r_p(\lambda^*) \tag{28}$$

We now simplify both $\nabla r_p(\lambda^*)$ and $\mathcal{H}(\lambda^*)$. First, since λ^* is optimal and an interior point of $\mathcal{K}$, we have

$$\nabla r_p(\lambda^*) = 0_G \tag{29}$$

Setting the value λ^* in Equation (30) implies that for $g, h \in [G]$:

$$\mathcal{H}_{g,h}(\lambda^*) = (p+1)r_p^{1-p}(\lambda^*)G_{g,h}(\lambda^*)$$

From the definition of $G_{g,h}$ we have:

$$G_{g,h}(\lambda^*) = \Sigma_p^{p+1}\left(\frac{1}{\lambda_G^*} + \frac{1}{\lambda_g^*}\mathbb{1}(g=h)\right)$$

Hence:

$$\mathcal{H}(\lambda^*) = (p+1) r_p^{1-p}(\lambda^*) \Sigma_p^{p+1} \left(\frac{1}{\lambda_G^*} + \frac{1}{\lambda_g^*} \mathbb{1}(g=h) \right) = (p+1) r_p(\lambda^*) \left(\frac{1}{\lambda_G^*} + \frac{1}{\lambda_g^*} \mathbb{1}(g=h) \right)$$

As a consequence,

$$\begin{aligned}
\langle \mathcal{H}(\lambda^*)(\lambda - \lambda^*), \lambda - \lambda^* \rangle &= \sum_{g,h \in [G-1]} \mathcal{H}_{g,h} (\lambda - \lambda^*)_g (\lambda - \lambda^*)_h \\
&= (p+1) r_p(\lambda^*) \sum_{g,h \in [G-1]} \left(\frac{1}{\lambda_G^*} + \frac{1}{\lambda_g^*} \mathbb{1}(g=h) \right) (\lambda - \lambda^*)_g (\lambda - \lambda^*)_h \\
&= (p+1) r_p(\lambda^*) \left\{ \frac{\sum_{g,h \in [G-1]} (\lambda - \lambda^*)_g (\lambda - \lambda^*)_h}{\lambda_G^*} + \sum_{g \in [G-1]} \frac{(\lambda - \lambda^*)_g^2}{\lambda_g^*} \right\} \\
&= (p+1) r_p(\lambda^*) \left\{ \frac{\left(\sum_{g \in [G-1]} (\lambda - \lambda^*)_g \right)^2}{\lambda_G^*} + \sum_{g \in [G-1]} \frac{(\lambda - \lambda^*)_g^2}{\lambda_g^*} \right\} \\
&= (p+1) r_p(\lambda^*) \left\{ \frac{(\lambda - \lambda^*)_G^2}{\lambda_G^*} + \sum_{g \in [G-1]} \frac{(\lambda - \lambda^*)_g^2}{\lambda_g^*} \right\} \\
&= (p+1) r_p(\lambda^*) \sum_{g \in [G]} \frac{(\lambda_g - \lambda_g^*)^2}{\lambda_g^*},
\end{aligned}$$

where the first step steps from the definition of the scalar product, and the second step stems from the expression of the Hessian we derived in Lemma 7. In the third step, we distribute the sum over the terms $\frac{1}{\lambda_G^*}$ and $\frac{1}{\lambda_g^*}\mathbb{1}(g=h)$. In the fourth step, we factorize the first sum, and in the fifth step, we use that $\sum_{g \in [G-1]} (\lambda - \lambda^*)_g = (1 - \lambda_G) - (1 - \lambda_G^*) = \lambda_G^* - \lambda_G$. Therefore:

$$\frac{1}{2} \langle \mathcal{H}(\lambda^*)(\lambda - \lambda^*), \lambda - \lambda^* \rangle = \frac{p+1}{2} r_p(\lambda^*) \sum_{g \in [G]} \frac{(\lambda_g - \lambda_g^*)^2}{\lambda_g^*}. \tag{30}$$

Therefore, by combining Equations (29), (30) and Inequality (28), in Taylor Inequality 16, we get:

$$\left| \frac{r_p(\lambda, \sigma)}{r_p^*} - 1 - \frac{p+1}{2} \sum_{g \in [G]} \frac{(\lambda_g - \lambda_g^*)^2}{\lambda_g^*} \right| \leq \frac{7(p+2)^2}{(\min_g \lambda_g^*)^2} \max_g \left(\frac{\lambda_g^*}{\lambda_g} \right)^{3p+3} \|\lambda - \lambda^*\|_\infty^3.$$

The proof of Lemma 4 follows by setting $n' = T\lambda$, $n^* = T\lambda^*$. In this case $R_{p,T}(n', \sigma) = \frac{1}{T} r_p(\lambda)$ and $R_{p,T}^* = \frac{1}{T} r_p(\lambda^*)$. □

A.6 Putting everything together

We are now ready to complete the proof of Theorem 1.

Theorem 1. *For any $\mathcal{D}$ that satisfies Assumptions 1 and 2 and for any finite p, the regret of Variance-UCB is at most $\tilde{O}(T^{-2})$, i.e.,*

$$\textit{Regret}_{p,T}(\textit{Variance-UCB}, \mathcal{D}) = \tilde{O}(T^{-2}).$$

Proof. First, notice that

$$\begin{aligned}
\text{Regret}_{p,T}(\text{Variance-UCB}) &= \mathbb{E}_\pi[R_{p,T}(n, \sigma) - R_{p,T}^*(\sigma)] \\
&= \mathbb{E}[R_{p,T}(n, \sigma) - R_{p,T}^*(\sigma) | \mathcal{A}_T] \mathbb{P}_\pi(\mathcal{A}_T) + \mathbb{E}[R_{p,T}(n, \sigma) - R_{p,T}^*(\sigma) | \mathcal{A}_T^c] \mathbb{P}_\pi(\mathcal{A}_T^c) \\
&\leq \mathbb{E}[R_{p,T}(n, \sigma) - R_{p,T}^*(\sigma) | \mathcal{A}_T] + \|\sigma^2\|_p \mathbb{P}_\pi(\mathcal{A}_T^c),
\end{aligned}$$

where the first step stems from the definition of regret introduced in 5, the second step stems from the law of total expectation, and the third step stems from both $\mathbb{P}(\mathcal{A}_T) \leq 1$ and $R_{p,T}(n,\sigma) - R^*_{p,T}(\sigma) \leq \|\sigma^2\|_p$. It remains to show that each term in the rightmost side is in $\tilde{O}(T^{-2})$. First, we know that $\mathbb{P}_\pi(\mathcal{A}_T) \leq 2GT^{-2.5} = \tilde{O}(T^{-2})$. Moreover, we have conditionally on $\mathcal{A}_T$:

$$\begin{aligned}\frac{R_{p,T}(n,\sigma) - R_{p,T}(n^*,\sigma)}{R_{p,T}(n^*,\sigma)} &\leq \frac{p+1}{2}\sum_{g\in[G]}\frac{(n_g - n^*_{g,T})^2}{Tn^*_{g,T}} + \frac{7(p+2)^2\Sigma_p^2}{\sigma^2_{\min}}\max_g\left(\frac{n^*_{g,T}}{n_g}\right)^{3p+3}\frac{\|n-n^*\|^3_\infty}{T^3}\\ &\leq \frac{p+1}{2}\frac{G\|n-n^*\|^2_\infty}{T\min_g n^*_{g,T}} + \frac{7(p+2)^2\Sigma_p^2}{\sigma^2_{\min}}\max_g\left(\frac{n^*_{g,T}}{n_g}\right)^{3p+3}\frac{\|n-n^*\|^3_\infty}{T^3},\end{aligned}$$

where the first inequality stems from Lemma 4, and the second inequality stems from $(n_{g,T} - n^*_{g,T})^2 \leq \|n-n^*\|^2_\infty$. Since $\sum_g n_{g,T} = T$, from Lemma 3 we have:

$$\begin{aligned} n_{g,T} - n^*_{g,T} &= -\sum_{h\neq g} n_{h,T} - n^*_{h,T}\\ &\geq -3(G-1) - \sum_{h\neq g}\frac{4C_Tp}{\Sigma_p(p+1)}\left(\sigma_g + \frac{2C_T}{\sqrt{n^*_{h,T}}}\right)^{\frac{p-1}{p+1}}\sqrt{n^*_{h,T}}\\ &\geq -3(G-1) - G\max_h\frac{4C_Tp}{\Sigma_p(p+1)}\left(\sigma_g + \frac{2C_T}{\sqrt{n^*_{h,T}}}\right)^{\frac{p-1}{p+1}}\sqrt{n^*_{h,T}}\\ &\geq -3G - \frac{4GC_Tp}{\Sigma_p(p+1)}\left(\min\sigma_g + \frac{2C_T}{\sqrt{\min_h n^*_{h,T}}}\right)^{\frac{p-1}{p+1}}\sqrt{\min_h n^*_{h,T}},\end{aligned}$$

where the first step stems from $\sum_h n_{h,T} - n^*_{h,T} = \sum_h n_{h,T} - \sum_h n^*_{h,T} = T - T = 0$, the second step stems from Lemma 3, and the last steps stem from taking the max over the sum. The last inequality implies

$$\|n-n^*\|_\infty \leq 3G + \frac{4GC_Tp}{\Sigma_p(p+1)}\left(\min\sigma_g + \frac{2C_T}{\sqrt{\min_h n^*_{h,T}}}\right)^{\frac{p-1}{p+1}}\sqrt{\min_h n^*_{h,T}}. \tag{31}$$

In particular, $\|n-n^*\|_\infty = \tilde{O}(\sqrt{\min_h n^*_{h,t}}) = \tilde{O}(\sqrt{T})$ and

$$\max_g \frac{n^*_{g,T}}{n_{g,T}} \leq \frac{1}{1 - \frac{\|n-n^*\|_\infty}{\min_h n^*}} = \frac{1}{1-\tilde{O}(T^{-0.5})} = \tilde{O}(1).$$

Therefore,

$$\begin{aligned}\frac{p+1}{2}\frac{G\|n-n^*\|^2_\infty}{T\min_g n^*_{g,T}} + \frac{7(p+2)^2\Sigma_p^2}{\sigma^2_{\min}}\max_g\left(\frac{n^*_{g,T}}{n_g}\right)^{3p+3}\frac{\|n-n^*\|^3_\infty}{T^3} &= \frac{p+1}{2}\frac{G\tilde{O}(T)}{T\Theta(T)} + \frac{7(p+2)^2\Sigma_p^2}{\sigma^2_{\min}}\tilde{O}(1)\frac{\tilde{O}(T^{1.5})}{T^3}\\ &= \tilde{O}(T^{-1}).\end{aligned}$$

Recall from Lemma 1 that $R_{p,T}(n^*,\sigma) = \Theta(T^{-1})$. Thus by taking the conditional expectation on $\mathcal{A}_T$, we have

$$\mathbb{E}[R_{p,T}(n,\sigma) - R^*_{p,T}(\sigma)|\mathcal{A}_T] \leq \tilde{O}(T^{-1}R^*_{p,T}(\sigma)) = \tilde{O}(T^{-2}),$$

which concludes the proof of Theorem 1. □

B Proof of Theorem 2

Through this section, we fix a policy π.

In this section, we give a proof to 2. First, we establish an initial lower bound that captures the trade-off between how hard it is to distinguish two instances and how hard it is to optimize both under the same action (B.1). Next, we provide a specific counter example which regret is at least $\Theta(T^{-2})$ (B.2).

B.1 Proof of Lemma 6

Lemma 6. *Let π be a fixed policy and $\mathcal{D}^a$, $\mathcal{D}^b$ be two instances with standard deviation vectors σ^a, σ^b, respectively. Then,*

$$\max\{\textit{Regret}_{p,T}(\pi, \mathcal{D}^a), \textit{Regret}_{p,T}(\pi, \mathcal{D}^b)\} \geq d(\sigma^a, \sigma^b) \exp\left(-\sum_{g\in[G]} \mathbb{E}_{\pi,\mathcal{D}^a}[n_{g,t}] KL(\mathcal{D}^a_g || \mathcal{D}^b_g)\right). \tag{11}$$

Proof. Let $\mathcal{D}^a$ and $\mathcal{D}^b$ be two instances with standard deviations σ^a and σ^b, and let $\delta \geq d(\sigma^a, \sigma^b)$. Let X be a random variable over $\{a, b\}$, we have

$$\begin{aligned}
\max(\text{Regret}_{p,T}(\pi, \mathcal{D}^a), \text{Regret}_{p,T}(\pi, \mathcal{D}^b)) &\geq \mathbb{E}_X[\text{Regret}_{p,T}(\pi, \mathcal{D}^X)] \\
&\geq \mathbb{E}[R_{p,T}(n, \sigma^X) - R^*_{p,T}(\sigma^X) | R_{p,T}(n, \sigma^X) - R^*_{p,T}(\sigma^X) > \delta] \\
&\quad \times \mathbb{P}_{\pi,X}\left(R_{p,T}(n, \sigma^X) - R^*_{p,T}(\sigma^X) > \delta\right) \\
&\geq \delta \mathbb{P}_{X,\pi}\left(R_{p,T}(n, \sigma^X) - R^*_{p,T}(\sigma^X) > \delta\right)
\end{aligned}$$

Let $\hat{x}$ be the following (random) classifier:

$$\hat{x} := \begin{cases} a & \text{If } R_{p,T}(n, \sigma^a) - R^*_{p,T}(\sigma^a) \leq \delta \\ b & \text{If } R_{p,T}(n, \sigma^b) - R^*_{p,T}(\sigma^b) \leq \delta \\ \text{Indifferent} & \text{Otherwise} \end{cases}$$

Since $\delta \geq d(\sigma^a, \sigma^b)$, $\hat{x}$ is well defined. Moreover, $\mathbb{P}_{\pi.X}\left(R_{p,T}(n, \sigma^X) - R^*_{p,T}(\sigma^X) > \delta\right) \geq \mathbb{P}_X(\hat{x} \neq X) \geq \inf_{\hat{x}} \mathbb{P}_X(\hat{x} \neq X)$, where the infinimum is taken over all the classifiers of $\{a, b\}$. Moreover, by Pinsker's inequality, $\inf_{\hat{x}} \mathbb{P}_X(\hat{x} \neq X) \geq \exp\left(-KL(\mathcal{D}^a||\mathcal{D}^b)\right)$. Since we have a bandits feedback, we have (see [29]) $KL(\mathcal{D}^a||\mathcal{D}^b) = \sum_{g\in[G]} \mathbb{E}_{\pi,\mathcal{D}^a}[n_{g,t}] KL(\mathcal{D}^a_g||\mathcal{D}^b_g)$, which concludes the proof of Lemma 6. □

B.2 The counter-examples

We recall the two instances we consider

$$\mathcal{D}^a : \begin{cases} \mathcal{D}^a_1 \sim \mathcal{N}\left(0, 1 + \frac{1}{\sqrt{T}}\right) \\ \mathcal{D}^a_g \sim \mathcal{N}(0, 1), \quad \forall g \neq 1 \end{cases} \qquad \mathcal{D}^b : \begin{cases} \mathcal{D}^b_2 \sim \mathcal{N}\left(0, 1 + \frac{1}{\sqrt{T}}\right) \\ \mathcal{D}^a_g \sim \mathcal{N}(0, 1), \quad \forall g \neq 2 \end{cases}$$

We start by upper bounding the KL-divergence between the two instances:

Lemma 12. *We have:* $\mathbb{E}_{\pi,\mathcal{D}^a}[n_{1,T}] KL(\mathcal{D}^a_1||\mathcal{D}^b_1) + \mathbb{E}_{\pi,\mathcal{D}^a}[n_{2,T}] KL(\mathcal{D}^a_2||\mathcal{D}^b_2) \leq \frac{1}{2}$

Proof. For convenience, we set $\nu := \frac{1}{\sqrt{T}}$. We use the formula for the $KL-$divergence of two univariate normal distributions of zero mean:

$$KL(\mathcal{D}^a_1||\mathcal{D}^b_1) = \frac{1}{2}\left(\log\left(\frac{1+\nu}{1}\right) + \frac{1-(1+\nu)}{1+\nu}\right).$$

The taylor expansion of the expression above can be derived by combining the expansions of both the functions $x \to \log(1+x)$ and $x \to \frac{1}{1+x}$:

$$\begin{aligned}
\frac{1}{2}\left(\log\left(\frac{1+\nu}{1}\right) + \frac{1-(1+\nu)}{1+\nu}\right) &= \frac{1}{2}\left(-\sum_{k\geq 1}\frac{(-1)^k}{k}\nu^k - \nu\sum_{k\geq 0}(-1)^k\nu^k\right) \\
&= \frac{1}{2}\sum_{k\geq 1}(-1)^k\nu^k\left(1 - \frac{1}{k}\right) \\
&= \frac{\nu^2}{2}\sum_{k\geq 0}(-1)^k\nu^k\left(1 - \frac{1}{k+2}\right)
\end{aligned}$$

Similarly,

$$\begin{aligned}
KL(\mathcal{D}_2^a||\mathcal{D}_2^b) &= \frac{1}{2}\left(\log\left(\frac{1}{1+\nu}\right) + \frac{1+\nu-1}{1}\right) \\
&= \frac{1}{2}\left(\nu - \log(1+\nu)\right) \\
&= \nu - \sum_{k\geq 1}\frac{(-1)^k}{k}\nu^k \\
&= \frac{\nu^2}{2}\sum_{k\geq 0}\frac{(-1)^k}{k+2}\nu^k
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathbb{E}_{\pi,\mathcal{D}^a}[n_{1,T}]KL(\mathcal{D}_1^a||\mathcal{D}_1^b) + \mathbb{E}_{\pi,\mathcal{D}^a}[n_{2,T}]KL(\mathcal{D}_2^a||\mathcal{D}_2^b) &\leq T\left(KL(\mathcal{D}_1^a||\mathcal{D}_1^b) + KL(\mathcal{D}_2^a||\mathcal{D}_2^b)\right) \\
&= \frac{T\nu^2}{2}\sum_{k\geq 0}\left(\frac{(-1)^k}{k+2} + 1 - \frac{(-1)^k}{k+2}\right)\nu^k \\
&= \frac{T\nu^2}{2}\sum_{k\geq 0}(-\nu)^k \\
&= \frac{T\nu^2}{2(1+\nu)} \\
&\leq \frac{T\nu^2}{2} \leq \frac{1}{2}
\end{aligned}$$

where we use in the first inequality that $n_{1,T}, n_{2,T} \leq T$. This concludes the proof. □

Next, we derive a simpler form for $d_p(\sigma^a, \sigma^b)$. We do so by exploiting the symmetries in σ^a, σ^b,

Lemma 13. *Let $\boldsymbol{u}_2$ denote the unit vector $(1,1)^T$. We have:*

$$d_p(\sigma^a, \sigma^b) = r_p\left(\frac{1}{2}\boldsymbol{u}, \sigma^a\right) - r_p^*(\sigma^a)$$

Proof. For $x \in \{a, b\}$, let $\mathcal{S}_\epsilon^a := \{\epsilon > 0 | r_p(\lambda, \sigma^x) - r_p^*(\sigma^x) \leq \epsilon\}$. Recall that

$$d_p(\sigma^a, \sigma^b) = \inf\{\delta \geq 0 | \mathcal{S}_\epsilon^a \cap \mathcal{S}_\epsilon^b \neq \emptyset\}$$

By symmetry of the problem, we have:

$$r_p^*(\sigma^a) = r_p^*(\sigma^b)$$

and for each $(\lambda_1, \lambda_2, \lambda') \in \mathcal{K}$:

$$r_p((\lambda_1, \lambda_2, \lambda'), \sigma^a) = r_p((\lambda_2, \lambda_1, \lambda'), \sigma^b)$$

On the one hand, if $r_p\left(\frac{1}{2}\mathbf{u}, \sigma^a\right) - r^*(\sigma^a) \leq \epsilon$, we must also have $r_p\left(\frac{1}{2}\mathbf{u}, \sigma^b\right) - r^*(\sigma^b) \leq \epsilon$, thus for $\epsilon \geq r_p\left(\frac{1}{2}\mathbf{u}, \sigma^a\right) - r^*(\sigma^a)$, we have $\mathcal{S}_\epsilon^a \cap \mathcal{S}_\epsilon^b \neq \emptyset$ and therefore $d_p(\sigma^a, \sigma^b) \leq r_p\left(\frac{1}{2}\mathbf{u}, \sigma^a\right) - r^*(\sigma^a)$. On the other hand, let $\epsilon \geq 0$ such that $\mathcal{S}_\epsilon^a \cap \mathcal{S}_\epsilon^b \neq \emptyset$. Each of $\mathcal{S}_\epsilon^a$ and $\mathcal{S}_\epsilon^b$ is convex therefore $\mathcal{S}_\epsilon^a \cap \mathcal{S}_\epsilon^b$ is also convex. Let $\lambda \in \mathcal{S}_\epsilon^a \cap \mathcal{S}_\epsilon^b$. By symmetry properties we must also have $1 - \lambda \in \mathcal{S}_\epsilon^a \cap \mathcal{S}_\epsilon^b$, and by convexity of the intersection, we must also have $\frac{1}{2}\mathbf{u} = \frac{1}{2}(\lambda + (1-\lambda)) \in \mathcal{S}_\epsilon^a \cap \mathcal{S}_\epsilon^b$, therefore $\epsilon \geq r_p\left(\frac{1}{2}\mathbf{u}, \sigma^a\right) - r^*(\sigma^a)$. By taking the inf we get $d_p(\sigma^a, \sigma^b) \geq r_p\left(\frac{1}{2}\mathbf{u}, \sigma^a\right) - r^*(\sigma^a)$, which concludes the proof. □

We are now ready to complete the proof of 2. It remains to show that $d_p(\sigma^a, \sigma^b) = r_p\left(\frac{1}{2}\mathbf{u}, \sigma^a\right) - r_p^*(\sigma^a) = \Theta(T^{-2})$.

Finally, we are ready to prove Theorem 2.

Theorem 2. *Let p be finite. For any online policy π, there exists an instance $\mathcal{D}_\pi$ such that for any $T \geq 1$,*

$$\textit{Regret}_{p,T}(\pi, \mathcal{D}_\pi) \geq (p+1)\left(\kappa_0 - \frac{\kappa_p}{\sqrt{T}}\right) T^{-2},$$

where κ_0 is a universal constant and κ_p is a constant that only depends on p.

Proof. For simplicity, we show the proof for $G = 2$. By using the proof of Lemma 4, we derive the following lower bound by setting $n' = \frac{1}{2T}\mathbf{u}$.

$$\frac{r_p\left(\frac{1}{2}\mathbf{u}, \sigma^a\right)}{r^*(\sigma^a)} - 1 \geq \frac{p+1}{2} \sum_{g \in \{1,2\}} \frac{\left(\frac{1}{2} - \lambda_g^*(\sigma^a)\right)^2}{\lambda_g^*(\sigma^a)} - \frac{7(p+2)^2}{(\min_g \lambda_g^*)^2} \max_g \left(\frac{\lambda_g^*}{1/2}\right)^{3p+3} \|\mathbf{u}/2 - \lambda^*\|_\infty^3. \tag{32}$$

it is convenient to write $\nu = \frac{1}{2\sqrt{T}}$ and $\lambda_1^*(\sigma^a) = \frac{1}{2} + f(\nu)$. In this case $\lambda_2^*(\sigma^a) = 1 - \lambda_1^*(\sigma^a) = \frac{1}{2} - f(\nu)$. Thus:

$$\sum_{g \in \{1,2\}} \frac{\left(\frac{1}{2} - \lambda_g^*(\sigma^a)\right)^2}{\lambda_g^*(\sigma^a)} = f(\nu)^2 \sum_{g \in \{1,2\}} \frac{\frac{1}{2} + f(\nu) + \frac{1}{2} - f(\nu)}{\left(\frac{1}{2} + f(\nu)\right)\left(\frac{1}{2} - f(\nu)\right)} = \frac{f(\nu)^2}{\frac{1}{4} - f(\nu)^2}$$

Since $\nu > 0$, we must have $\lambda_1^*(\sigma^a) > \lambda_2^*(\sigma^a)$, therefore $f(\nu) > 0$ and $\min_g \lambda_g^* = \frac{1}{2} - f(\nu)$, $\max_g \left(\frac{\lambda_g^*}{1/2}\right)^{3p+3} = \left(1 + \frac{1}{2} f(\nu)\right)^{3p+3}$. Finally $\|\mathbf{u}/2 - \lambda^*\|_\infty = \max(f(\nu), -f(\nu)) = |f(\nu)| = f(\nu)$. Therefore Equation 32 can be simplified into:

$$\frac{r_p\left(\frac{1}{2}\mathbf{u}, \sigma^a\right)}{r^*(\sigma^a)} - 1 \geq \frac{p+1}{2} \frac{f(\nu)^2}{\frac{1}{4} - f(\nu)^2} - \frac{7(p+2)^2}{\left(\frac{1}{2} - f(\nu)\right)^2} \left(1 + \frac{1}{2} f(\nu)\right)^{3p+3} f(\nu)^3 \tag{33}$$

It remains to derive simpler bounds for $f(\nu)$. On the one hand, we have from Lemma 1:

$$\begin{aligned}
f(\nu) = \frac{(1+\nu)^{\frac{p}{p+1}}}{1 + (1+\nu)^{\frac{p}{p+1}}} - \frac{1}{2} &\geq \frac{\sqrt{1+\nu}}{1 + \sqrt{1+\nu}} - \frac{1}{2} \\
&= \frac{1}{1 + (1+\nu)^{-1/2}} - \frac{1}{2} \\
&= \frac{1}{1 + 1 - \frac{1}{2}\nu + o(\nu)} - \frac{1}{2} \\
&= \frac{1}{2 - \frac{\nu + o(\nu)}{2}} - \frac{1}{2} \\
&= \frac{1}{2}\left(1 + \frac{\nu + o(\nu)}{2} + o\left(\frac{\nu + o(\nu)}{2}\right) - 1\right) \\
&= \frac{\nu}{4} + o(\nu).
\end{aligned}$$

where the first step is due to $p \to (1+\nu)^{\frac{p}{p+1}}$ is increasing in $p \geq 1$, and the next steps consist of deriving the first order approximation of $\frac{\sqrt{1+\nu}}{1+\sqrt{1+\nu}} - \frac{1}{2}$. Therefore, there exists a universal constant $\kappa < +\infty$ such that $\forall \nu \leq \frac{1}{2}$:

$$f(\nu) \geq \kappa \nu.$$

On the other hand, we have for $\nu \in [0, 1/2]$:

$$\begin{aligned}
\frac{(1+\nu)^{\frac{p}{p+1}}}{1 + (1+\nu)^{\frac{p}{p+1}}} - \frac{1}{2} &\leq \frac{1 + \frac{p}{p+1}\nu}{2 + \frac{p}{p+1}\nu} - \frac{1}{2} \\
&\leq \frac{1 + \frac{p}{p+1}\nu}{2} - \frac{1}{2} \\
&\leq \frac{1}{2}\nu \leq \frac{1}{4},
\end{aligned}$$

where the first inequality stems from $(1+x)^{\frac{p}{p+1}} \leq 1 + \frac{p}{p+1}x$ for $x \geq 0$, the second inequality stems from $\frac{p}{p+1}\nu \geq 0$, and the last inequality stems from $\frac{p}{p+1}\nu \leq \nu \leq 1/2$. We replace the previous inequality into (33) to get:

$$\begin{aligned}\frac{r_p\left(\frac{1}{2}\mathbf{u},\sigma^a\right)}{r^*(\sigma^a)} - 1 &\geq (p+1)\nu^2\left(\frac{\kappa^2/2}{1/4-0} - \frac{7\frac{(p+2)^2}{p+1}}{(1/2-1/2\times 1/4)^2}\left(1+\frac{1/2}{4}\right)^{3p+3}\frac{\nu}{2}\right)\\ &\geq (p+1)\nu^2\left(2\kappa^2 - \frac{7(p+1)}{18/64}\left(\frac{9}{8}\right)^{3p+3}\nu\right)\\ &= 2\kappa^2(p+1)\nu^2 - \frac{7(p+1)^2}{18/64}\left(\frac{9}{8}\right)^{3p+3}\nu^3,\end{aligned}$$

where the second step stems from $(p+2)^2 \geq p+1$. Thus, for $\nu \leq \frac{1}{2}$, we have:

$$\frac{r_p\left(\frac{1}{2}\mathbf{u},\sigma^a\right)}{r_p^*(\sigma^a)} - 1 \geq 2\kappa^2(p+1)\nu^2 - \frac{7(p+1)^2}{18/64}\left(\frac{9}{8}\right)^{3p+3}\nu^3 \tag{34}$$

We then use Inequality (34), Lemma 12 in the inequality stated in Lemma 6 to derive $\forall\nu \leq \frac{1}{2}$:

$$\max\left(\text{Regret}_{p,T}(\pi,\mathcal{D}^a),\text{Regret}_T(\pi,\mathcal{D}^b)\right) \geq \frac{1}{T}\left(2\kappa^2(p+1)\nu^2 - \frac{7(p+1)^2}{18/64}\left(\frac{9}{8}\right)^{3p+3}\nu^3\right)^+\exp\left(-\frac{T\nu^2}{2}\right)$$

By recalling that $\nu = \frac{1}{2\sqrt{T}} \leq \frac{1}{2}$, we have:

$$\max\left(\text{Regret}_{p,T}(\pi,\mathcal{D}^a),\text{Regret}_T(\pi,\mathcal{D}^b)\right) \geq \frac{(p+1)}{T^2}\left(2\kappa^2 - \frac{24(p+1)}{\sqrt{T}}\left(\frac{9}{8}\right)^{3p+3}\right)^+\exp\left(-\frac{1}{8}\right)$$

which concludes the proof by setting $\kappa_p := 24\exp\left(-\frac{1}{8}\right)(p+1)\left(\frac{9}{8}\right)^{3p+3}$ and $\kappa_0 := 2\exp\left(-\frac{1}{8}\right)\kappa^2$. □

C Upper and lower bounds when $p = \infty$

The proof for Theorem 3 (upper bound when $p = \infty$) follows the same high-level steps as the proof of Theorem 1 (upper bound when $p \in \mathbb{R}$). However, some adjustments of the proofs are necessary. Table 2 summarizes the changes that are required.

Result	**Does it hold for $p = +\infty$?**	**Comments**
Lemma 1	Yes	Proof needs modification
Lemmas 2, 3	Yes	Same proof
Lemma 4	No	Replaced by Lemma 14
Lemmas 6, 12, 13	Yes	Same proof
$d(\sigma^a,\sigma^b) = \Theta(T^{-2})$	No	Replaced by $d(\sigma^a,\sigma^b) = \Theta(T^{-1.5})$

Table 2: Summary of the possible extensions to $p = +\infty$

In Appendix C.1, we prove Lemma 1 for the case where $p = +\infty$. In Appendix C.2, we introduce and prove Lemma 14, the replacement for Lemma 4. In Appendix C.4, we prove Theorem 4.

C.1 Extending Lemma 1

Lemma 1. *[Benchmark analysis] For each $t \in \mathbb{N}$ and $p \in [1,+\infty]$, let $n^*_{g,t} := \frac{\sigma_g^{\frac{2p}{p+1}}t}{\sum_{h\in[G]}\sigma_h^{\frac{2p}{p+1}}}$. Then,*

$$R^*_{p,t}(\sigma) = R_{p,t}(n^*_t,\sigma) = \frac{1}{t}R_{p,t}(n^*_1,\sigma). \tag{9}$$

Proof. For the case where $p = +\infty$, r_∞ is still continuous on $\mathcal{S}_G$ and infinite whenever a coordinate is zero. Therefore:

$$\emptyset \neq \text{argmin}_{\mathcal{S}_G} r_\infty = \text{argmin}_{\mathcal{S}_G \cap (\mathbb{R}_+^*)^G} r_\infty$$

Let $\lambda \in \text{argmin}_{\mathcal{S}_G} r_\infty$ with $\lambda > 0$. Set $A := \{g \in [G] | \frac{\sigma_g^2}{\lambda_g} \text{ is maximal}\}$ and $B := G - A$. For the sake of contradiction, we assume that $B \neq \emptyset$ and let us consider an element $g \in B$. We will construct a $\lambda' \in \mathcal{S}_G$ satisfying $r_\infty(\lambda') < r_\infty(\lambda)$ and contradicting the minimality of λ First, set:

$$\kappa := \frac{\frac{\lambda_g}{\sigma_g^2} - \frac{\sum_{a \in A} \lambda_a}{\sum_{a \in A} \sigma_g^2}}{\frac{1}{\sum_{a \in A} \sigma_a^2} + \frac{1}{\sigma_g^2}}$$

and for each $h \in [G]$, set:

$$\lambda'_h = \begin{cases} \lambda_h & \text{for } h \in B - \{g\} \\ \lambda_h + \frac{\sigma_h^2}{\sum_{a \in A} \sigma_a^2} \kappa & \text{for } h \in A \\ \lambda_h - \kappa & \text{for } h = g \end{cases}$$

Note that:

$$\sum_{h \in [G]} \lambda'_h = \sum_{h \in [G]} \lambda_h + \sum_{h \in A} \frac{\sigma_h^2}{\sum_{a \in A} \sigma_a^2} \kappa - \kappa = \sum_{h \in [G]} \lambda_h = 1$$

Let us now prove that $\lambda' > 0$. For $h \in B - \{g\}$, the result is true since $\lambda'_h = \lambda_h \geq 0$. Moreover, by construction of A and g, we have for each $h \in A$:

$$\frac{\lambda_g}{\sigma_g^2} > \frac{\lambda_h}{\sigma_h^2} = \frac{\sum_{a \in A} \lambda_a}{\sum_{a \in A} \sigma_a^2}$$

In particular, $\kappa > 0$ and $\lambda'_h > 0$ for each $h \in A$. Lastly, we have

$$\begin{aligned} \lambda'_g = \lambda_g - \kappa &= \frac{1}{\frac{1}{\sum_{a \in A} \sigma_a^2} + \frac{1}{\sigma_g^2}} \left(\frac{\lambda_g}{\sigma_g^2} + \frac{\lambda_g}{\sum_{a \in A} \sigma_a^2} - \left(\frac{\lambda_g}{\sigma_g^2} - \frac{\sum_{a \in A} \lambda_a}{\sum_{a \in A} \sigma_g^2} \right) \right) \\ &= \frac{1}{\frac{1}{\sum_{a \in A} \sigma_a^2} + \frac{1}{\sigma_g^2}} \left(\frac{\lambda_g + \sum_{a \in A} \lambda_a}{\sum_{a \in A} \sigma_a^2} \right) > 0 \end{aligned}$$

We next prove that $r_\infty(\lambda', \sigma) < r_\infty(\lambda, \sigma)$. For $h \in B - \{g\}$, we have:

$$\frac{\sigma_h^2}{\lambda'_h} = \frac{\sigma_h^2}{\lambda_h} < r_\infty(\lambda, \sigma)$$

Next, we have for any $a \in A$, $\frac{\sigma_h^2}{\lambda_a} = r_\infty(\lambda, \sigma)$, thus:

$$\frac{\sigma_h^2}{\lambda'_h} = \frac{\sigma_h^2}{\lambda_h + \frac{\sigma_h^2}{\sum_{a \in A} \sigma_a^2} \kappa} = \frac{1}{\underbrace{\frac{\lambda_h}{\sigma_h^2}}_{= \frac{1}{r_\infty(\lambda, \sigma)}} + \underbrace{\frac{1}{\sum_{a \in A} \sigma_a^2} \kappa}_{>0}} < r_\infty(\lambda, \sigma)$$

Finally, by choice of κ, we have:

$$\frac{\sigma_g^2}{\lambda'_g} = \frac{\sum_{a \in} \sigma_a^2}{\sum_{a \in A} \lambda'_a} < r_\infty(\lambda, \sigma)$$

Therefore $r_\infty(\lambda', \sigma) = \max_{h \in [G]} \frac{\sigma_g^2}{\lambda_g} < r_\infty(\lambda, \sigma)$, contradicting the minimality of λ. Hence our assumption is wrong and $B = \emptyset$, or equivalently, $A = [G]$, which implies that for any possible minimizer λ, we must have:

$$\forall g, h \in [G], \quad \frac{\sigma_g^2}{\lambda_g} = \frac{\sigma_h^2}{\lambda_h}$$

which implies that $\text{argmin} r_\infty = \{\lambda^*(\infty)\}$

$\square$

C.2 Extending Lemma 4 to R_∞

The upper bound Lemma 4 goes to $+\infty$ as $p = +\infty$ and is no longer insightful. In this section, we provide a suitable bound:

Lemma 14. *Let $\sigma \in \mathbb{R}_+^G$ and $n' \in \mathbb{R}_+^G$ such that $\sum_{g\in[G]} n'_g = T$. Then, We have:*

$$\frac{R_\infty(n',\sigma)}{R^*_{\infty,T}(\sigma)} \le 1 + \max_g \left|1 - \frac{n'_{g,T}}{n^*_{g,T}}\right| + \max_g \left(\frac{n^*_{g,T}}{n'_{g,T}}\right)^3 \max_g \left|1 - \frac{n^*_{g,T}}{n'_{g,T}}\right|^2$$

Proof. □

From Lemma 1, we have:

$$\frac{\sigma_1^2}{\lambda_1^*} = \ldots = \frac{\sigma_G^2}{\lambda_G^*} = r^*_\infty$$

Thus for $g \in [G]$, we have:

$$\begin{aligned}\frac{\sigma_g^2}{\lambda_g} &\le r^*_\infty + \left|\frac{\sigma_g^2}{\lambda_g} - \frac{\sigma_g^2}{\lambda_g^*}\right| \\ &\le r^*_\infty + \frac{\sigma_g^2}{(\lambda_g^*)^2}|\lambda_g - \lambda_g^*| + \sigma_g^2 \max\left(\frac{1}{\lambda_g}, \frac{1}{\lambda_g^*}\right)^3 |\lambda_g - \lambda_g^*|^2 \\ &\le r^*_\infty \left(1 + \max_g \left|1 - \frac{\lambda_g}{\lambda_g^*}\right| + \max_g \left(\frac{\lambda_g^*}{\lambda_g}\right)^3 \max_g \left|1 - \frac{\lambda_g^*}{\lambda_g}\right|\right)\end{aligned}$$

The proof follows from taking the maximum over $g \in [G]$ and setting $n^* = T\lambda^*$, and $n' = T\lambda$.

C.3 Proof of Theorem 3

Theorem 3. *For any $\mathcal{D}$ that satisfies Assumptions 1 and 2,*

$$\textit{Regret}_{\infty,T}(\textit{Variance-UCB}, \mathcal{D}) = \tilde{O}(T^{-3/2}).$$

Proof. Similarly to Appendix A.6, we decompose the regret into two terms, and Inequality (**??**) still holds:

$$\text{Regret}_{\infty,T}(\text{Variance-UCB}) \le \mathbb{E}[R_{\infty,T}(n,\sigma) - R^*_{\infty,T}(\sigma)|\mathcal{A}_T] + \|\sigma^2\|_\infty \mathbb{P}_\pi(\mathcal{A}_T^c). \tag{35}$$

We now show that each term in the right hand side of 35 is in $\tilde{O}(T^{-1.5})$. We have $\mathbb{P}_\pi(\mathcal{A}_T^c) = \tilde{O}(T^{-1.5})$. It remains to show that $\mathbb{E}[R_{\infty,T}(n,\sigma) - R^*_{\infty,T}(\sigma)|\mathcal{A}_T] = \tilde{O}(T^{-1.5})$. For this, we apply Lemma 14 for n_T. Conditionally on $\mathcal{A}_T$, we have:

$$\begin{aligned}\frac{R_\infty(n_T,\sigma)}{R^*_\infty(\sigma)} &\le 1 + \max_g \left|1 - \frac{n_{g,T}}{n^*_{g,T}}\right| + \max_g \left(\frac{n^*_{g,T}}{n_{g,T}}\right)^3 \max_g \left|1 - \frac{n^*_{g,T}}{n_{g,T}}\right|^2 &(36)\\ &\le 1 + \frac{\|n_T - n_T^*\|_\infty}{\min_g n^*_{g,T}} + \max_g \left(\frac{n^*_{g,T}}{n_{g,T}}\right)^3 \frac{\|n_T - n_T^*\|^2_\infty}{(\min_g n^*_{g,T})^2} &(37)\end{aligned}$$

From Lemma 1, $\min_g n^*_{g,T} = \Theta(T)$ and that $R^*_{\infty,T}(\sigma) = \Theta(T^{-1})$. From A.6, we know that $\|n_T - n_T^*\|_\infty = \tilde{O}(\sqrt{T})$ and that $\frac{n^*_{g,T}}{n_{g,T}} = \tilde{O}(1)$. Therefore, conditionally on $\mathcal{A})_T$,

$$\begin{aligned}R_\infty(n_T,\sigma) - R^*_{\infty,T}(\sigma) &\le \left(\frac{\|n_T - n_T^*\|_\infty}{\min_g n^*_{g,T}} + \max_g \left(\frac{n^*_{g,T}}{n_{g,T}}\right)^3 \frac{\|n_T - n_T^*\|^2_\infty}{(\min_g n^*_{g,T})^2}\right) R^*_{\infty,T}(\sigma) = \tilde{O}\left(\frac{R^*_{\infty,T}(\sigma)}{\sqrt{T}}\right) \\ &= \tilde{O}\left(\frac{T^{-1}}{\sqrt{T}}\right) = \tilde{O}(T^{-1.5}).\end{aligned}$$

Where the terms in $\tilde{O}$ do not depend on the randomness of Variance-UCB, but only on the instance of the problem. Therefore by taking the conditional expectation on $\mathcal{A}_T$ we get $\mathbb{E}[R_{\infty,T}(n,\sigma) - R^*_{\infty,T}(\sigma)|\mathcal{A}_T] = \tilde{O}(T^{-1.5})$, which completes the proof. □

C.4 Proof of Theorem 4

Theorem 4. *For any online policy π, there exists an instance $\mathcal{D}_\pi$ such that for any $T \geq 1$,*

$$\textit{Regret}_{\infty,T}(\pi, \mathcal{D}_\pi) \geq \frac{2}{5} T^{-3/2}.$$

Proof. Notice that Lemmas 6, 12 and 13 still hold for $p = +\infty$. Similarly to the proof for Theorem 2, it remains to derive a lower bound on $r_\infty\left(\frac{1}{2}\mathbf{u}, \sigma^a\right) - r^*_\infty(\sigma^a)$. We have for $\frac{1}{2\sqrt{T}} \leq \frac{1}{2}$,

$$\begin{aligned}
r_\infty\left(\frac{1}{2}\mathbf{u}, \sigma^a\right) &= \max\left(\frac{(\sigma_1^a)^2}{1/2}, \frac{(\sigma_2^a)^2}{1/2}\right) \\
&= \max\left(\frac{1 + \frac{1}{2\sqrt{T}}}{1/2}, \frac{1}{1/2}\right) \\
&= 2\left(1 + \frac{1}{2\sqrt{T}}\right) \\
r_\infty(\sigma^a) &= (\sigma_1^a)^2 + (\sigma_2^a)^2 \\
&= 2 + \frac{1}{2\sqrt{T}}
\end{aligned}$$

therefore:

$$r_\infty\left(\frac{1}{2}\mathbf{u}, \sigma^a\right) - r^*_\infty(\sigma^a) = 2\left(1 + \frac{1}{2\sqrt{T}}\right) - \left(2 + \frac{1}{2\sqrt{T}}\right) = \frac{1}{2\sqrt{T}} \tag{38}$$

We replace 38 and the inequality from Lemma 12 in the inequality from Lemma 6 to derive:

$$\max\left(\text{Regret}_{\infty,T}(\pi, \mathcal{D}^a), \text{Regret}_{\infty,T}(\pi, \mathcal{D}^b)\right) \geq \frac{\frac{1}{2\sqrt{T}}}{T} \exp\left(\frac{-T \frac{1}{2\sqrt{T}}^2}{2}\right)$$

thus we obtain:

$$\max\left(\text{Regret}_{\infty,T}(\pi, \mathcal{D}^a), \text{Regret}_{\infty,T}(\pi, \mathcal{D}^b)\right) \geq \frac{\exp(-1/8)}{2T\sqrt{T}} \geq \frac{2/5}{T\sqrt{T}}$$

$\square$