A. Proof for Theorem 1

To prove Theorem 1, we first prove the following theorem:

Theorem 3. *If (1) $\bar{\mathbf{w}} = \Phi\theta^*$, (2) the prior on θ^* is $N(0, \lambda^2 I)$, and (3) the noises are i.i.d. sampled from $N(0, \sigma^2)$, then under* `CombLinTS` *algorithm with parameter (Φ, λ, σ), then we have*

$$R_{\text{Bayes}}(n) \leq 1 + K\lambda \min\left\{\sqrt{\ln\left(\frac{\lambda L n}{\sqrt{2\pi}}\right)}, \sqrt{d\ln\left(\frac{2dKn\lambda}{\sqrt{2\pi}}\right)}\right\}\sqrt{\frac{2dn\ln\left(1+\frac{nK\lambda^2}{d}\right)}{\ln\left(1+\frac{\lambda^2}{\sigma^2}\right)}}. \tag{10}$$

Notice that Theorem 1 follows immediately from Theorem 3. Specifically, if $\lambda \geq \sigma$, then we have

$$\begin{aligned} B_{\text{Bayes}}(n) &\leq 1 + K\lambda \min\left\{\sqrt{\ln\left(\frac{\lambda L n}{\sqrt{2\pi}}\right)}, \sqrt{d\ln\left(\frac{2dKn\lambda}{\sqrt{2\pi}}\right)}\right\}\sqrt{2dn\log_2\left(1+\frac{nK\lambda^2}{d}\right)} \\ &= \tilde{O}\left(K\lambda\sqrt{dn\min\left\{\ln(L), d\right\}}\right). \end{aligned} \tag{11}$$

We now outline the proof of Theorem 3, which is based on (Russo & Van Roy, 2013; Dani et al., 2008). Let $\mathcal{H}_t$ denote the "history" (i.e. all the available information) by the start of episode t. Note that from the Bayesian perspective, conditioning on $\mathcal{H}_t$, θ^* and θ_t are i.i.d. drawn from $N(\bar{\theta}_t, \Sigma_t)$ (see (Russo & Van Roy, 2013)). This is because that conditioning on $\mathcal{H}_t$, the posterior belief in θ^* is $N(\bar{\theta}_t, \Sigma_t)$ and based on Algorithm 2, θ_t is independently sampled from $N(\bar{\theta}_t, \Sigma_t)$. Since `ORACLE` is a fixed combinatorial optimization algorithm (even though it can be independently randomized), and $E, \mathcal{A}, \Phi$ are all fixed, then conditioning on $\mathcal{H}_t$, A^* and A^t are also i.i.d., furthermore, A^* is conditionally independent of θ_t, and A^t is conditionally independent of θ^*.

To simplify the exposition, $\forall \theta \in \mathbb{R}^d$ and $\forall A \subseteq E$, we define

$$g(A, \theta) = \sum_{e \in A} \langle \phi_e, \theta \rangle, \tag{12}$$

then we have $\mathbb{E}[f(A^*, \mathbf{w}_t)|\mathcal{H}_t, \theta^*, \theta_t, A^*, A^t] = g(A^*, \theta^*)$ and $\mathbb{E}[f(A^t, \mathbf{w}_t)|\mathcal{H}_t, \theta^*, \theta_t, A^*, A^t] = g(A^t, \theta^*)$, hence we have $\mathbb{E}[R_t|\mathcal{H}_t] = \mathbb{E}[g(A^*, \theta^*) - g(A^t, \theta^*)|\mathcal{H}_t]$. We also define the *upper confidence bound (UCB)* function $U_t : 2^E \to \mathbb{R}$ as

$$U_t(A) = \sum_{e \in A}\left[\left\langle \phi_e, \bar{\theta}_t \right\rangle + c\sqrt{\phi_e^T \Sigma_t \phi_e}\right], \tag{13}$$

where $c > 0$ is a constant to be specified. Notice that conditioning on $\mathcal{H}_t$, U_t is a deterministic function and A^*, A^t are i.i.d., then $\mathbb{E}[U_t(A^t) - U_t(A^*)|\mathcal{H}_t] = 0$ and

$$\mathbb{E}[R_t|\mathcal{H}_t] = \mathbb{E}[g(A^*, \theta^*) - U_t(A^*)|\mathcal{H}_t] + \mathbb{E}\left[U_t(A^t) - g(A^t, \theta^*)\middle|\mathcal{H}_t\right]. \tag{14}$$

One key observation is that

$$\begin{aligned} \mathbb{E}\left[U_t(A^t) - g(A^t, \theta^*)\middle|\mathcal{H}_t\right] &\overset{(a)}{=} \sum_{e \in E}\mathbb{E}\left[\mathbb{1}\left\{e \in A^t\right\}\left[\left\langle \phi_e, \bar{\theta}_t - \theta^* \right\rangle + c\sqrt{\phi_e^T\Sigma_t\phi_e}\right]\middle|\mathcal{H}_t\right] \\ &\overset{(b)}{=} \sum_{e \in E}\mathbb{E}\left[\mathbb{1}\left\{e \in A^t\right\}\middle|\mathcal{H}_t\right]\mathbb{E}\left[\left\langle \phi_e, \bar{\theta}_t - \theta^*\right\rangle\middle|\mathcal{H}_t\right] + c\mathbb{E}\left[\sum_{e \in A^t}\sqrt{\phi_e^T\Sigma_t\phi_e}\middle|\mathcal{H}_t\right] \\ &\overset{(c)}{=} c\mathbb{E}\left[\sum_{e \in A^t}\sqrt{\phi_e^T\Sigma_t\phi_e}\middle|\mathcal{H}_t\right], \end{aligned} \tag{15}$$

where (b) follows from the fact that A^t and θ^* are conditionally independent, and (c) follows from $\mathbb{E}[\theta^*|\mathcal{H}_t] = \bar{\theta}_t$. Hence $B_{\text{Bayes}}(n) = \sum_{t=1}^n \mathbb{E}[g(A^*, \theta^*) - U_t(A^*)] + c\sum_{t=1}^n \mathbb{E}\left[\sum_{e \in A^t}\sqrt{\phi_e^T\Sigma_t\phi_e}\right]$. We can show that (1)

$\sum_{t=1}^{n} \mathbb{E}[g(A^*, \theta^*) - U_t(A^*)] \leq 1$ if we choose

$$c \geq \min\left\{\sqrt{\ln\left(\frac{\lambda L n}{\sqrt{2\pi}}\right)}, \sqrt{d \ln\left(\frac{2dKn\lambda}{\sqrt{2\pi}}\right)}\right\}, \tag{16}$$

and (2) $\sum_{t=1}^{n} \mathbb{E}\left[\sum_{e \in A^t} \sqrt{\phi_e^T \Sigma_t \phi_e}\right] \leq K\lambda \sqrt{2dn \ln\left(1 + \frac{nK\lambda^2}{d}\right) / \ln\left(1 + \frac{\lambda^2}{\sigma^2}\right)}$. Thus, the bound in Theorem 3 holds. Please refer to the remainder of this section for the full proof.

A.1. Bound on $\sum_{t=1}^{n} \mathbb{E}[g(A^*, \theta^*) - U_t(A^*)]$

We first prove that if we choose

$$c \geq \min\left\{\sqrt{\ln\left(\frac{\lambda L n}{\sqrt{2\pi}}\right)}, \sqrt{d \ln\left(\frac{2dKn\lambda}{\sqrt{2\pi}}\right)}\right\}, \tag{17}$$

then $\sum_{t=1}^{n} \mathbb{E}[g(A^*, \theta^*) - U_t(A^*)] \leq 1$. To prove this result, we use the following inequality for truncated Gaussian distribution.

Lemma 1. *If $X \sim N(\mu, s^2)$, then we have*

$$\mathbb{E}[X \mathbb{1}\{X \geq 0\}] = \mu\left[1 - \Phi_G\left(\frac{-\mu}{s}\right)\right] + \frac{s}{\sqrt{2\pi}} \exp\left(-\frac{\mu^2}{2s^2}\right),$$

where Φ_G is the cumulative distribution function (CDF) of the standard Gaussian distribution $N(0,1)$. Furthermore, if $\mu \leq 0$, we have $\mathbb{E}[X \mathbb{1}\{X \geq 0\}] \leq \frac{s}{\sqrt{2\pi}} \exp\left(-\frac{\mu^2}{2s^2}\right)$.

Based on Lemma 1, we can prove the following lemmas:

Lemma 2. *If $c \geq \sqrt{\ln\left(\frac{\lambda L n}{\sqrt{2\pi}}\right)}$, then we have $\sum_{t=1}^{n} \mathbb{E}[g(A^*, \theta^*) - U_t(A^*)] \leq 1$.*

Proof. We have the following naive bound:

$$\begin{aligned} g(A^*, \theta^*) - U_t(A^*) &= \sum_{e \in A^*} \left[\left\langle \phi_e, \theta^* - \bar{\theta}_t \right\rangle - c\sqrt{\phi_e^T \Sigma_t \phi_e}\right] \\ &\leq \sum_{e \in A^*} \left[\left\langle \phi_e, \theta^* - \bar{\theta}_t \right\rangle - c\sqrt{\phi_e^T \Sigma_t \phi_e}\right] \mathbb{1}\left\{\left\langle \phi_e, \theta^* - \bar{\theta}_t \right\rangle - c\sqrt{\phi_e^T \Sigma_t \phi_e} \geq 0\right\} \\ &\leq \sum_{e \in E} \left[\left\langle \phi_e, \theta^* - \bar{\theta}_t \right\rangle - c\sqrt{\phi_e^T \Sigma_t \phi_e}\right] \mathbb{1}\left\{\left\langle \phi_e, \theta^* - \bar{\theta}_t \right\rangle - c\sqrt{\phi_e^T \Sigma_t \phi_e} \geq 0\right\}. \end{aligned}$$

Notice that conditioning on $\mathcal{H}_t$, $\left\langle \phi_e, \theta^* - \bar{\theta}_t \right\rangle - c\sqrt{\phi_e^T \Sigma_t \phi_e}$ is a Gaussian random variable with mean $-c\sqrt{\phi_e^T \Sigma_t \phi_e}$ and variance $\phi_e^T \Sigma_t \phi_e$. Thus, from Lemma 1, we have

$$\begin{aligned} &\mathbb{E}_{\theta^*, A^*}[g(A^*, \theta^*) - U_t(A^*) | \mathcal{H}_t] \\ &\overset{(a)}{\leq} \sum_{e \in E} \mathbb{E}_{\theta^*}\left[\left[\left\langle \phi_e, \theta^* - \bar{\theta}_t \right\rangle - c\sqrt{\phi_e^T \Sigma_t \phi_e}\right] \mathbb{1}\left\{\left\langle \phi_e, \theta^* - \bar{\theta}_t \right\rangle - c\sqrt{\phi_e^T \Sigma_t \phi_e} \geq 0\right\} \middle| \mathcal{H}_t\right] \\ &\overset{(b)}{\leq} \sum_{e \in E} \sqrt{\frac{\phi_e^T \Sigma_t \phi_e}{2\pi}} \exp\left(-\frac{c^2}{2}\right) \\ &\overset{(c)}{\leq} \exp\left(-\frac{c^2}{2}\right) \sum_{e \in E} \frac{\lambda \|\phi_e\|}{\sqrt{2\pi}} \leq \exp\left(-\frac{c^2}{2}\right) \frac{\lambda L}{\sqrt{2\pi}}, \end{aligned} \tag{18}$$

where the last two inequalities follow from the fact that $\phi_e^T \Sigma_t \phi_e \le \phi_e^T \Sigma_1 \phi_e \le \lambda^2 \|\phi_e\|^2 \le \lambda^2$, since $\|\phi_e\| \le 1$ by assumption[8]. Thus we have

$$\mathbb{E}\left[\sum_{t=1}^{n}\left[g(A^*,\theta^*) - U_t(A^*)\right]\right] \le \exp\left(-\frac{c^2}{2}\right)\frac{n\lambda L}{\sqrt{2\pi}}. \tag{19}$$

If we choose $c \ge \sqrt{2\ln\left(\frac{\lambda L n}{\sqrt{2\pi}}\right)}$, then we have $\mathbb{E}[\sum_{t=1}^{n}\left[g(A^*,\theta^*) - U_t(A^*)\right]] \le 1$. ◻

Lemma 3. *If $c \ge \sqrt{d\ln\left(\frac{2dKn\lambda}{\sqrt{2\pi}}\right)}$, then we also have $\sum_{t=1}^{n}\mathbb{E}[g(A^*,\theta^*) - U_t(A^*)] \le 1$.*

Proof. We use $v_1, \dots, v_d$ to denote a fixed set of d orthonormal eigenvectors of Σ_t, and $\Lambda_1^2, \dots, \Lambda_d^2$ to denote the associated eigenvalues. Notice that for $i \neq j$, we have $v_i^T \Sigma_t v_j = \Lambda_i^2 v_i^T v_j = 0$. $\forall i = 1, \dots, d$, we define $v_{i+d} = -v_i$ and $\Lambda_{i+d} = \Lambda_i$, which allows us to define the following conic decomposition:

$$\phi_e = \sum_{i=1}^{2d} \alpha_{ei} v_i, \quad \forall e \in E,$$

subject to the constraints that $\alpha_{ei} \ge 0$, $\forall (e,i)$. Notice that α_{ei}'s are uniquely determined. Furthermore, for i and j s.t. $|i-j| = d$, by definition of conic decomposition, we have $\alpha_{ei}\alpha_{ej} = 0$. In other words, α_e is a d-sparse vector.

Since we assume that $\|\phi_e\| \le 1$, we have that $\sum_{i=1}^{2d}\alpha_{ei}^2 \le 1$, $\forall e \in E$. Thus, for any e, we have that $\left\langle \phi_e, \theta^* - \bar{\theta}_t \right\rangle = \sum_{i=1}^{2d}\alpha_{ei}\left\langle v_i, \theta^* - \bar{\theta}_t\right\rangle$ and

$$\begin{aligned}\phi_e^T \Sigma_t \phi_e &= \left(\sum_{i=1}^{2d}\alpha_{ei} v_i^T\right)\Sigma_t\left(\sum_{j=1}^{2d}\alpha_{ei} v_j\right) \\ &= \sum_{i=1}^{2d}\sum_{j=1}^{2d}\alpha_{ei}\alpha_{e_j} v_i^T \Sigma_t v_j. \end{aligned} \tag{20}$$

Notice that for $i \neq j$, if $|i-j| \neq d$, then $v_i^T \Sigma_t v_j = 0$; on the other hand, if $|i-j| = d$, $\alpha_{ei}\alpha_{ej} = 0$. Thus, if $i \neq j$, we have $\alpha_{ei}\alpha_{e_j} v_i^T \Sigma_t v_j = 0$. Consequently,

$$\phi_e^T \Sigma_t \phi_e = \sum_{i=1}^{2d}\alpha_{ei}^2 v_i^T \Sigma_t v_i = \sum_{i=1}^{2d}\alpha_{ei}^2\Lambda_i^2.$$

Thus we have

$$\sqrt{\phi_e^T \Sigma_t \phi_e} = \sqrt{\sum_{i=1}^{2d}\alpha_{ei}^2\Lambda_i^2} \ge \frac{1}{\sqrt{d}}\sum_{i=1}^{2d}\alpha_{ei}\Lambda_i, \tag{21}$$

where the inequality follows from Cauchy-Schwartz inequality, specifically, define $s_i = 1$ if $\alpha_{ei}\Lambda_i \neq 0$, and $s_i = 0$ if $\alpha_{ei}\Lambda_i = 0$, then we have

$$\sum_{i=1}^{2d}\alpha_{ei}\Lambda_i = \sum_{i=1}^{2d}\alpha_{ei}\Lambda_i s_i \le \sqrt{\sum_{i=1}^{2d}s_i^2}\sqrt{\sum_{i=1}^{2d}\alpha_{ei}^2\Lambda_i^2} \le \sqrt{d}\sqrt{\sum_{i=1}^{2d}\alpha_{ei}^2\Lambda_i^2},$$

[8] Notice that in the derivation of Inequality (18), we implicitly assume that $\phi_e^T \Sigma_t \phi_e > 0$, $\forall e \in E$. It is worth pointing out that the case with $\phi_e^T \Sigma_t \phi_e = 0$ is a trivial case and this inequality still holds in this case.

where the last inequality follows from the fact that α_e is d-sparse. Thus, for any e, we have that

$$\left\langle \phi_e, \theta^* - \bar{\theta}_t \right\rangle - c\sqrt{\phi_e^T \Sigma_t \phi_e} \leq \sum_{i=1}^{2d} \alpha_{ei} \left\langle v_i, \theta^* - \bar{\theta}_t \right\rangle - \frac{c}{\sqrt{d}} \sum_{i=1}^{2d} \alpha_{ei} \Lambda_i. \tag{22}$$

Consequently, we have

$$\sum_{e \in A^*} \left[\left\langle \phi_e, \theta^* - \bar{\theta}_t \right\rangle - c\sqrt{\phi_e^T \Sigma_t \phi_e} \right] \leq \sum_{i=1}^{2d} \left(\left\langle v_i, \theta^* - \bar{\theta}_t \right\rangle - \frac{c\Lambda_i}{\sqrt{d}} \right) \left(\sum_{e \in A^*} \alpha_{ei} \right). \tag{23}$$

Define $X_i = \left\langle v_i, \theta^* - \bar{\theta}_t \right\rangle - \frac{c\Lambda_i}{\sqrt{d}}$, notice that conditioning on $\mathcal{H}_t$, we have $X_i | \mathcal{H}_t \sim N\left(-\frac{c\Lambda_i}{\sqrt{d}}, \Lambda_i^2 \right)$. Hence we have

$$\begin{aligned} \sum_{e \in A^*} \left\langle \phi_e, \theta^* - \bar{\theta}_t \right\rangle - c\sqrt{\phi_e^T \Sigma_t \phi_e} &\overset{(a)}{\leq} \sum_{i=1}^{2d} X_i \left[\sum_{e \in A^*} \alpha_{ei} \right] \\ &\overset{(b)}{\leq} \sum_{i=1}^{2d} X_i \mathbb{1}\{X_i \geq 0\} \left[\sum_{e \in A^*} \alpha_{ei} \right], \end{aligned}$$

where the inequality (b) follows from the fact that $X_i \leq X_i \mathbb{1}\{X_i \geq 0\}$ and $\left[\sum_{e \in A^*} \alpha_{ei} \right] \geq 0$. On the other hand, notice that $|A^*| \leq K$

$$\sum_{e \in A^*} \alpha_{ei} \leq \sqrt{|A^*|} \sqrt{\sum_{e \in A^*} \alpha_{ei}^2} \leq \sqrt{|A^*|} \sqrt{\sum_{e \in A^*} \sum_{j=1}^{d} \alpha_{ej}^2} \leq \sqrt{|A^*|} \sqrt{\sum_{e \in A^*} 1} = |A^*| \leq K.$$

Since $X_i \mathbb{1}\{X_i \geq 0\} \geq 0$, we have

$$\sum_{e \in A^*} \left\langle \phi_e, \theta^* - \bar{\theta}_t \right\rangle - c\sqrt{\phi_e^T \Sigma_t \phi_e} \leq K \sum_{i=1}^{2d} X_i \mathbb{1}\{X_i \geq 0\},$$

notice that the RHS does not include A^*. Hence we have

$$\begin{aligned} \mathbb{E}_{\theta^*}[g(A^*, \theta^*) - U_t(A^*) | \mathcal{H}_t] &= \mathbb{E}_{\theta^*}\left[\sum_{e \in A^*} \left\langle \phi_e, \theta^* - \bar{\theta}_t \right\rangle - c\sqrt{\phi_e^T \Sigma_t \phi_e} \middle| \mathcal{H}_t \right] \\ &\leq K \sum_{i=1}^{2d} \mathbb{E}_{\theta^*}[X_i \mathbb{1}\{X_i \geq 0\} | \mathcal{H}_t] \\ &\leq K \sum_{i=1}^{2d} \frac{\Lambda_i}{\sqrt{2\pi}} \exp\left(-\frac{c^2}{2d} \right) \leq \frac{2dK\lambda}{\sqrt{2\pi}} \exp\left(-\frac{c^2}{2d} \right), \end{aligned}$$

where the last inequality follows from the fact that $\Lambda_i \leq \lambda$. Hence we have

$$\sum_{t=1}^{n} \mathbb{E}[g(A^*, \theta^*) - U_t(A^*)] \leq \frac{2dKn\lambda}{\sqrt{2\pi}} \exp\left(-\frac{c^2}{2d} \right),$$

if we choose $c \geq \sqrt{2d \ln\left(\frac{2dKn\lambda}{\sqrt{2\pi}} \right)}$, then we have $\sum_{t=1}^{n} \mathbb{E}[f(A^*, \theta^*) - U_t(A^*)] \leq 1$. □

Combining the results from Lemma 2 and 3, we have proved that if

$$c \geq \min\left\{\sqrt{\ln\left(\frac{\lambda L n}{\sqrt{2\pi}}\right)}, \sqrt{d\ln\left(\frac{2dKn\lambda}{\sqrt{2\pi}}\right)}\right\},$$

then $\sum_{t=1}^{n}\mathbb{E}[g(A^*,\theta^*)-U_t(A^*)]\leq 1$.

A.2. Bound on $\sum_{t=1}^{n}\mathbb{E}\left[\sum_{e\in A^t}\sqrt{\phi_e^T\Sigma_t\phi_e}\right]$

In this subsection, we derive a bound on $\sum_{t=1}^{n}\mathbb{E}\left[\sum_{e\in A^t}\sqrt{\phi_e^T\Sigma_t\phi_e}\right]$. Our analysis is motivated by the analysis in (Dani et al., 2008). Specifically, we provide a worst-case bound on $\sum_{t=1}^{n}\sum_{e\in A^t}\sqrt{\phi_e^T\Sigma_t\phi_e}$, for any realization of random variable $\mathbf{w}_t$'s, θ_t's, A^t's, A^*, and θ^*.

Lemma 4. $\sum_{t=1}^{n}\sum_{e\in A^t}\sqrt{\phi_e^T\Sigma_t\phi_e}\leq K\lambda\sqrt{\frac{dn\log\left(1+\frac{nK\lambda^2}{d\sigma^2}\right)}{\log\left(1+\frac{\lambda^2}{\sigma^2}\right)}}$.

Proof. To simplify the exposition, we define

$$z_{t,k}=\sqrt{\phi_{a_k^t}^T\Sigma_t\phi_{a_k^t}}. \tag{24}$$

First, notice that Σ_t^{-1} is the Gramian matrix and satisfies

$$\Sigma_{t+1}^{-1}=\Sigma_t^{-1}+\frac{1}{\sigma^2}\sum_{k=1}^{|A^t|}\phi_{a_k^t}\phi_{a_k^t}^T. \tag{25}$$

Hence for any t, k, we have that

$$\begin{aligned}\det\left[\Sigma_{t+1}^{-1}\right] &\geq \det\left[\Sigma_t^{-1}+\frac{1}{\sigma^2}\phi_{a_k^t}\phi_{a_k^t}^T\right]=\det\left[\Sigma_t^{-\frac{1}{2}}\left(I+\frac{1}{\sigma^2}\Sigma_t^{\frac{1}{2}}\phi_{a_k^t}\phi_{a_k^t}^T\Sigma_t^{\frac{1}{2}}\right)\Sigma_t^{-\frac{1}{2}}\right]\\ &=\det\left[\Sigma_t^{-1}\right]\det\left[I+\frac{1}{\sigma^2}\Sigma_t^{\frac{1}{2}}\phi_{a_k^t}\phi_{a_k^t}^T\Sigma_t^{\frac{1}{2}}\right]=\det\left[\Sigma_t^{-1}\right]\left(1+\frac{1}{\sigma^2}\phi_{a_k^t}^T\Sigma_t\phi_{a_k^t}\right)\\ &=\det\left[\Sigma_t^{-1}\right]\left(1+\frac{z_{t,k}^2}{\sigma^2}\right).\end{aligned} \tag{26}$$

Hence we have that

$$\left(\det\left[\Sigma_{t+1}^{-1}\right]\right)^{|A^t|}\geq\left(\det\left[\Sigma_t^{-1}\right]\right)^{|A^t|}\prod_{k=1}^{|A^t|}\left(1+\frac{z_{t,k}^2}{\sigma^2}\right). \tag{27}$$

Remark 1. *This is where the extra $O(\sqrt{K})$ factor arises. Notice that this extra factor is purely due to linear generalization. Specifically, if $\Phi=I$, then Σ_t's and Σ_t^{-1}'s will be diagonal, and we have*

$$\det\left[\Sigma_{t+1}^{-1}\right]=\det\left[\Sigma_t^{-1}\right]\prod_{k=1}^{|A^t|}\left(1+\frac{z_{t,k}^2}{\sigma^2}\right). \tag{28}$$

Notice that Equation 27 further implies that

$$\left(\det\left[\Sigma_{t+1}^{-1}\right]\right)^{K}\geq\left(\det\left[\Sigma_t^{-1}\right]\right)^{K}\prod_{k=1}^{|A^t|}\left(1+\frac{z_{t,k}^2}{\sigma^2}\right), \tag{29}$$

since $\det\left[\Sigma_{t+1}^{-1}\right] \geq \det\left[\Sigma_t^{-1}\right]$ and $|A^t| \leq K$. Recall that $\det\left[\Sigma_1^{-1}\right] = \left(\frac{1}{\lambda^2}\right)^d$, we have that

$$\left(\det\left[\Sigma_{n+1}^{-1}\right]\right)^K \geq \left(\det\left[\Sigma_1^{-1}\right]\right)^K \prod_{t=1}^{n}\prod_{k=1}^{|A^t|}\left(1+\frac{z_{t,k}^2}{\sigma^2}\right) = \frac{1}{\lambda^{2dK}}\prod_{t=1}^{n}\prod_{k=1}^{|A^t|}\left(1+\frac{z_{t,k}^2}{\sigma^2}\right). \tag{30}$$

On the other hand, we have

$$\text{trace}\left[\Sigma_{n+1}^{-1}\right] = \text{trace}\left[\frac{1}{\lambda^2}I + \frac{1}{\sigma^2}\sum_{t=1}^{n}\sum_{k=1}^{|A^t|}\phi_{a_k^t}\phi_{a_k^t}^T\right] = \frac{d}{\lambda^2} + \frac{1}{\sigma^2}\sum_{t=1}^{n}\sum_{k=1}^{|A^t|}\|\phi_{a_k^t}\|^2 \leq \frac{d}{\lambda^2} + \frac{nK}{\sigma^2}, \tag{31}$$

where the last inequality follows from the assumption that $\|\phi_e\| \leq 1$, $\forall e \in E$ and $|A^t| \leq K$. From the trace-determinant inequality, we have

$$\frac{1}{d}\text{trace}\left[\Sigma_{n+1}^{-1}\right] \geq \left(\det\left[\Sigma_{n+1}^{-1}\right]\right)^{\frac{1}{d}},$$

which implies that

$$\left(\frac{1}{\lambda^2} + \frac{nK}{d\sigma^2}\right)^{dK} \geq \left(\frac{1}{d}\text{trace}\left[\Sigma_{n+1}^{-1}\right]\right)^{dK} \geq \left(\det\left[\Sigma_{n+1}^{-1}\right]\right)^K \geq \frac{1}{\lambda^{2dK}}\prod_{t=1}^{n}\prod_{k=1}^{|A^t|}\left(1+\frac{z_{t,k}^2}{\sigma^2}\right).$$

Taking the logarithm, we have

$$dK\log\left(1+\frac{nK\lambda^2}{d\sigma^2}\right) \geq \sum_{t=1}^{n}\sum_{k=1}^{|A^t|}\log\left(1+\frac{z_{t,k}^2}{\sigma^2}\right). \tag{32}$$

Notice that $z_{t,k}^2 = \phi_{a_k^t}^T\Sigma_t\phi_{a_k^t}$, hence we have that $0 \leq z_{t,k}^2 \leq \phi_{a_k^t}^T\Sigma_1\phi_{a_k^t} \leq \lambda^2\|\phi_{a_k^t}\|^2 \leq \lambda^2$. We have the following technical lemma:

Lemma 5. *For any real number $x \in [0, \lambda^2]$, we have $x \leq \frac{\lambda^2}{\log\left(1+\frac{\lambda^2}{\sigma^2}\right)}\log\left(1+\frac{x}{\sigma^2}\right)$.*

Proof. Define $h(x) = \frac{\lambda^2}{\log\left(1+\frac{\lambda^2}{\sigma^2}\right)}\log\left(1+\frac{x}{\sigma^2}\right) - x$, thus we only need to prove $h(x) \geq 0$ for $x \in [0, \lambda^2]$. Notice that $h(x)$ is a strictly concave function for $x \geq 0$, and $h(0) = 0$, $h(\lambda^2) = 0$. From Jensen's inequality, for any $x \in (0, \lambda^2)$, we have $h(x) > 0$. □

Hence we have that

$$\sum_{t=1}^{n}\sum_{k=1}^{|A^t|} z_{t,k}^2 \leq \frac{\lambda^2}{\log\left(1+\frac{\lambda^2}{\sigma^2}\right)}\sum_{t=1}^{n}\sum_{k=1}^{|A^t|}\log\left(1+\frac{z_{t,k}^2}{\sigma^2}\right) \leq \frac{dK\lambda^2\log\left(1+\frac{nK\lambda^2}{d\sigma^2}\right)}{\log\left(1+\frac{\lambda^2}{\sigma^2}\right)} \tag{33}$$

Finally, we have that

$$\sum_{t=1}^{n}\sum_{k=1}^{|A^t|} z_{t,k} \leq \sqrt{nK}\sqrt{\sum_{t=1}^{n}\sum_{k=1}^{|A^t|} z_{t,k}^2} \leq K\lambda\sqrt{\frac{dn\log\left(1+\frac{nK\lambda^2}{d\sigma^2}\right)}{\log\left(1+\frac{\lambda^2}{\sigma^2}\right)}}. \tag{34}$$

□

Recall that the above bound holds for any realization of random variables, thus, we have

$$\mathbb{E}\left[\sum_{t=1}^{n}\left[U_t(A^t)-g(A^t,\theta^*)\right]\right]=c\mathbb{E}\left[\sum_{t=1}^{n}\sum_{k=1}^{|A^t|}z_{t,k}\right]\leq cK\lambda\sqrt{\frac{dn\log\left(1+\frac{nK\lambda^2}{d}\right)}{\log\left(1+\frac{\lambda^2}{\sigma^2}\right)}}.$$

With

$$c=\min\left\{\sqrt{\ln\left(\frac{\lambda Ln}{\sqrt{2\pi}}\right)},\sqrt{d\ln\left(\frac{2dKn\lambda}{\sqrt{2\pi}}\right)}\right\}, \tag{35}$$

and combining the results in the previous subsection, we have proved Theorem 3.

B. Proof for Theorem 2

We start by writing an alternative formula for Σ_t and $\bar{\theta}_t$. Notice that based on Algorithm 1, we have:

$$\begin{aligned}\Sigma_t^{-1}&=\frac{1}{\lambda^2}I+\frac{1}{\sigma^2}\sum_{\tau=1}^{t-1}\sum_{k=1}^{|A^\tau|}\phi_{a_k^\tau}\phi_{a_k^\tau}^T\\ \Sigma_t^{-1}\bar{\theta}_t&=\frac{1}{\sigma^2}\sum_{\tau=1}^{t-1}\sum_{k=1}^{|A^\tau|}\phi_{a_k^\tau}\mathbf{w}_\tau\left(a_k^\tau\right)\end{aligned} \tag{36}$$

Interested readers might refer to Appendix C for the derivation of Equation (36). The proof proceeds as follows. We first construct a confidence set of θ^* based on the "self normalized bound" developed in (Abbasi-Yadkori et al., 2011). Then we derive a regret bound based on Lemma 4 derived above.

B.1. Confidence Set

Our construction of confidence set is motivated by the analysis in (Agrawal & Goyal, 2013). We start by defining some useful notation. Specifically, for any $t=1,2,\dots,n$, any $k=1,2,\dots,|A^t|$, we define

$$\eta_{t,k}=\mathbf{w}_t\left(a_k^t\right)-\bar{\mathbf{w}}\left(a_k^t\right).$$

One key observation is that $\eta_{t,k}$'s form a Martingale difference sequence (MDS)[9] since $\mathbf{w}(e)$'s are statistically independent under P. Moreover, since $\mathbf{w}_t\left(a_k^t\right)$ is bounded in interval $[0,1]$, $\eta_{t,k}$'s are sub-Gaussian with constant $R=1$. We further define

$$\begin{aligned}V_t&=\frac{\sigma^2}{\lambda^2}I+\sum_{\tau=1}^{t-1}\sum_{k=1}^{|A^\tau|}\phi_{a_k^\tau}\phi_{a_k^\tau}^T\\ \xi_t&=\sum_{\tau=1}^{t-1}\sum_{k=1}^{|A^\tau|}\phi_{a_k^\tau}\eta_{\tau,k}\end{aligned}$$

As we will see later, we define V_t and ξ_t to use the "self normalized bound" developed in (Abbasi-Yadkori et al., 2011) (see Theorem 1 of (Abbasi-Yadkori et al., 2011)). Notice that based on the above definition, we have $\Sigma_t^{-1}=\frac{1}{\sigma^2}V_t$, and

$$\bar{\theta}_t-\theta^*=\Sigma_t\left(\frac{1}{\sigma^2}\xi_t-\frac{1}{\lambda^2}\theta^*\right).$$

[9] Note that the notion of "time" is indexed by a pair (t,k), and follows the lexicographical order.

To see why the second equality holds, notice that

$$\Sigma_t^{-1}\bar{\theta}_t = \frac{1}{\sigma^2}\sum_{\tau=1}^{t-1}\sum_{k=1}^{|A^\tau|}\phi_{a_k^\tau}\left(\phi_{a_k^\tau}^T\theta^* + \eta_{\tau,k}\right) = \left(\Sigma_t^{-1} - \frac{1}{\lambda^2}I\right)\theta^* + \frac{1}{\sigma^2}\xi_t.$$

Hence, for any $e \in E$, we have

$$\begin{aligned} \left|\left\langle \phi_e, \bar{\theta}_t - \theta^* \right\rangle\right| &= \left|\phi_e^T\Sigma_t\left(\frac{1}{\sigma^2}\xi_t - \frac{1}{\lambda^2}\theta^*\right)\right| \\ &\le \|\phi_e\|_{\Sigma_t}\left\|\frac{1}{\sigma^2}\xi_t - \frac{1}{\lambda^2}\theta^*\right\|_{\Sigma_t} \\ &\le \|\phi_e\|_{\Sigma_t}\left[\frac{1}{\sigma^2}\|\xi_t\|_{\Sigma_t} + \frac{1}{\lambda^2}\|\theta^*\|_{\Sigma_t}\right], \end{aligned}$$

where the first inequality follows from the Cauchy-Schwarz inequality, and the second inequality follows from the triangular inequality. Notice that

$$\|\theta^*\|_{\Sigma_t} \le \|\theta^*\|_{\Sigma_1} = \lambda\|\theta^*\|_2,$$

hence we have

$$\left|\left\langle \phi_e, \bar{\theta}_t - \theta^* \right\rangle\right| \le \|\phi_e\|_{\Sigma_t}\left[\frac{1}{\sigma^2}\|\xi_t\|_{\Sigma_t} + \frac{1}{\lambda}\|\theta^*\|_2\right].$$

Moreover, we have

$$\frac{1}{\sigma^2}\|\xi_t\|_{\Sigma_t} = \frac{1}{\sigma^2}\|\xi_t\|_{\sigma^2 V_t^{-1}} = \frac{1}{\sigma}\|\xi_t\|_{V_t^{-1}}.$$

So we have

$$\left|\left\langle \phi_e, \bar{\theta}_t - \theta^* \right\rangle\right| \le \|\phi_e\|_{\Sigma_t}\left[\frac{1}{\sigma}\|\xi_t\|_{V_t^{-1}} + \frac{1}{\lambda}\|\theta^*\|_2\right]. \tag{37}$$

The above inequality always holds. We now provide a high probability bound on $\|\xi_t\|_{V_t^{-1}}$, based on the "self normalized bound" proposed in (Abbasi-Yadkori et al., 2011). From Theorem 1 of (Abbasi-Yadkori et al., 2011), we know for any $\delta \in (0,1)$, with probability at least $1-\delta$,

$$\|\xi_t\|_{V_t^{-1}} \le \sqrt{2\log\left(\frac{\det(V_t)^{1/2}\det(V_1)^{-1/2}}{\delta}\right)} \quad \forall t = 1, 2, \ldots.$$

Obviously, $\det(V_1) = \left[\frac{\sigma^2}{\lambda^2}\right]^d$, on the other hand, we have

$$[\det(V_t)]^{1/d} \le \frac{\text{trace}(V_t)}{d} = \frac{\sigma^2}{\lambda^2} + \frac{1}{d}\sum_{\tau=1}^{t-1}\sum_{k=1}^{|A^\tau|}\|\phi_{a_k^\tau}\|^2 \le \frac{\sigma^2}{\lambda^2} + \frac{(t-1)K}{d},$$

where the last inequality follows from the assumption that $\|\phi_e\| \le 1$. Hence, for $t \le n$, we have

$$[\det(V_t)]^{1/d} \le \frac{\sigma^2}{\lambda^2} + \frac{nK}{d}.$$

Thus, with probability at least $1-\delta$, we have

$$\|\xi_t\|_{V_t^{-1}} \le \sqrt{d\log\left(1 + \frac{nK\lambda^2}{d\sigma^2}\right) + 2\log\left(\frac{1}{\delta}\right)} \quad \forall t = 1, 2, \ldots, n.$$

Thus, we have the following lemma:

Lemma 6. *For any $\lambda, \sigma > 0$ and any $\delta \in (0,1)$, with probability at least $1-\delta$, we have*

$$\left|\left\langle \phi_e, \bar{\theta}_t - \theta^* \right\rangle\right| \leq \|\phi_e\|_{\Sigma_t} \left[\frac{1}{\sigma}\sqrt{d\log\left(1+\frac{nK\lambda^2}{d\sigma^2}\right)+2\log\left(\frac{1}{\delta}\right)} + \frac{\|\theta^*\|_2}{\lambda}\right], \tag{38}$$

for all $t = 1, 2, \ldots, n$, and for all $e \in E$.

Notice that $\|\phi_e\|_{\Sigma_t} = \sqrt{\phi_e^T \Sigma_t \phi_e}$, thus, the above lemma immediately implies the following lemma:

Lemma 7. *For any $\lambda, \sigma > 0$, any $\delta \in (0,1)$, and any*

$$c \geq \frac{1}{\sigma}\sqrt{d\log\left(1+\frac{nK\lambda^2}{d\sigma^2}\right)+2\log\left(\frac{1}{\delta}\right)} + \frac{\|\theta^*\|_2}{\lambda},$$

with probability at least $1-\delta$, we have

$$\left\langle \phi_e, \bar{\theta}_t \right\rangle - c\sqrt{\phi_e^T \Sigma_t \phi_e} \leq \langle \phi_e, \theta^* \rangle \leq \left\langle \phi_e, \bar{\theta}_t \right\rangle + c\sqrt{\phi_e^T \Sigma_t \phi_e},$$

for all $e \in E$ and $t = 1, 2, \ldots n$.

Notice that

$$\langle \phi_e, \theta^* \rangle \leq \left\langle \phi_e, \bar{\theta}_t \right\rangle + c\sqrt{\phi_e^T \Sigma_t \phi_e}$$

is exactly $\bar{\mathbf{w}}(e) \leq \hat{\mathbf{w}}_t(e)$.

B.2. Regret Analysis

We define event G as

$$G = \left\{ \left\langle \phi_e, \bar{\theta}_t \right\rangle - c\sqrt{\phi_e^T \Sigma_t \phi_e} \leq \langle \phi_e, \theta^* \rangle \leq \left\langle \phi_e, \bar{\theta}_t \right\rangle + c\sqrt{\phi_e^T \Sigma_t \phi_e} \ \forall e \in E, \, \forall t = 1, \ldots, n \right\}, \tag{39}$$

and use $\bar{G}$ to denote the complement of event G. Recall that Lemma 7 states that if

$$c \geq \frac{1}{\sigma}\sqrt{d\log\left(1+\frac{nK\lambda^2}{d\sigma^2}\right)+2\log\left(\frac{1}{\delta}\right)} + \frac{1}{\lambda}\|\theta^*\|_2, \tag{40}$$

then $\mathbb{P}(G) \geq 1-\delta$. Moreover, by definition, under event G, we have $\bar{\mathbf{w}}(e) \leq \hat{\mathbf{w}}_t(e)$, for all $t = 1, \ldots, n$ and any $e \in E$.

Notice that

$$\begin{aligned}
R(n) &= \sum_{t=1}^{n} \mathbb{E}\left[\sum_{e \in A^*} \mathbf{w}_t(e) - \sum_{e \in A^t} \mathbf{w}_t(e)\right] \\
&= \sum_{t=1}^{n} \mathbb{E}\left[\sum_{e \in A^*} \bar{\mathbf{w}}(e) - \sum_{e \in A^t} \bar{\mathbf{w}}(e)\right] \\
&= \mathbb{P}(G) \sum_{t=1}^{n} \mathbb{E}\left[\sum_{e \in A^*} \bar{\mathbf{w}}(e) - \sum_{e \in A^t} \bar{\mathbf{w}}(e) \middle| G\right] + \mathbb{P}(\bar{G}) \sum_{t=1}^{n} \mathbb{E}\left[\sum_{e \in A^*} \bar{\mathbf{w}}(e) - \sum_{e \in A^t} \bar{\mathbf{w}}(e) \middle| \bar{G}\right] \\
&\leq \sum_{t=1}^{n} \mathbb{E}\left[\sum_{e \in A^*} \bar{\mathbf{w}}(e) - \sum_{e \in A^t} \bar{\mathbf{w}}(e) \middle| G\right] + \mathbb{P}(\bar{G})\, nK,
\end{aligned}$$

where the last inequality follows from the naive bound on the realized regret. If c satisfies inequality (40), we have $\mathbb{P}\left(\bar{G}\right) \leq \delta$, hence we have

$$R(n) \leq \sum_{t=1}^{n} \mathbb{E}\left[\sum_{e \in A^*} \bar{\mathbf{w}}(e) - \sum_{e \in A^t} \bar{\mathbf{w}}(e) \middle| G\right] + nK\delta.$$

Finally, we bound $\sum_{t=1}^{n} \mathbb{E}\left[\sum_{e \in A^*} \bar{\mathbf{w}}(e) - \sum_{e \in A^t} \bar{\mathbf{w}}(e) \middle| G\right]$ using a worst-case bound conditioning on G (worst-case over all the possible random realizations), notice that conditioning on G, we have

$$\sum_{e \in A^*} \bar{\mathbf{w}}(e) \leq \sum_{e \in A^*} \hat{\mathbf{w}}_t(e) \leq \sum_{e \in A^t} \hat{\mathbf{w}}_t(e),$$

where the first inequality follows from the definition of event G, and the second inequality follows from that A^t is the *exact* solution of the combinatorial optimization problem $(E, \mathcal{A}, \hat{\mathbf{w}}_t)$. Thus we have

$$\begin{aligned}
\sum_{e \in A^*} \bar{\mathbf{w}}(e) - \sum_{e \in A^t} \bar{\mathbf{w}}(e) &\leq \sum_{e \in A^t} \hat{\mathbf{w}}_t(e) - \sum_{e \in A^t} \bar{\mathbf{w}}(e) \\
&= \sum_{e \in A^t} \left[\langle \phi_e, \bar{\theta}_t - \theta^* \rangle + c\sqrt{\phi_e^T \Sigma_t \phi_e} \right] \\
&\leq 2c \sum_{e \in A^t} \sqrt{\phi_e^T \Sigma_t \phi_e},
\end{aligned}$$

where the last inequality follows from the definition of G. Recall that from Lemma 4, we have

$$\sum_{t=1}^{n} \sum_{e \in A^t} \sqrt{\phi_e^T \Sigma_t \phi_e} \leq K\lambda \sqrt{\frac{dn \log\left(1 + \frac{nK\lambda^2}{d\sigma^2}\right)}{\log\left(1 + \frac{\lambda^2}{\sigma^2}\right)}}.$$

Thus we have

$$\sum_{t=1}^{n} \mathbb{E}\left[\sum_{e \in A^*} \bar{\mathbf{w}}(e) - \sum_{e \in A^t} \bar{\mathbf{w}}(e) \middle| G\right] \leq 2c\mathbb{E}\left[\sum_{t=1}^{n} \sum_{e \in A^t} \sqrt{\phi_e^T \Sigma_t \phi_e}\right] \leq 2cK\lambda \sqrt{\frac{dn \log\left(1 + \frac{nK\lambda^2}{d\sigma^2}\right)}{\log\left(1 + \frac{\lambda^2}{\sigma^2}\right)}},$$

which implies

$$R(n) \leq 2cK\lambda \sqrt{\frac{dn \log\left(1 + \frac{nK\lambda^2}{d\sigma^2}\right)}{\log\left(1 + \frac{\lambda^2}{\sigma^2}\right)}} + nK\delta.$$

C. Technical Lemma

In this section, we derive Equation (36). We first prove the following technical lemma:

Lemma 8. *For any $\phi, \bar{\theta} \in \mathbb{R}^d$, any positive definite $\Sigma \in \mathbb{R}^{d \times d}$, any $\sigma > 0$, and any $w \in \mathbb{R}$, if we define*

$$\begin{aligned}
\Sigma_{\text{new}} &= \Sigma - \frac{\Sigma \phi \phi^T \Sigma}{\phi^T \Sigma \phi + \sigma^2} \\
\bar{\theta}_{\text{new}} &= \left[I - \frac{\Sigma \phi \phi^T}{\phi^T \Sigma \phi + \sigma^2} \right] \bar{\theta} + \left[\frac{\Sigma \phi}{\phi^T \Sigma \phi + \sigma^2} \right] w,
\end{aligned}$$

then we have

$$\Sigma_{\text{new}}^{-1} = \Sigma^{-1} + \frac{1}{\sigma^2} \phi \phi^T \tag{41}$$

$$\Sigma_{\text{new}}^{-1} \bar{\theta}_{\text{new}} = \Sigma^{-1} \bar{\theta} + \frac{1}{\sigma^2} \phi w. \tag{42}$$

Proof. Notice that Equation (41) follows directly from the Woodbury matrix identity (matrix inversion lemma). We now prove Equation (42). Notice that we have

$$\begin{aligned}\bar{\theta}_{\text{new}} &= \left[I - \frac{\Sigma\phi\phi^T}{\phi^T\Sigma\phi + \sigma^2}\right]\bar{\theta} + \left[\frac{\Sigma\phi}{\phi^T\Sigma\phi + \sigma^2}\right] w \\ &= \left[\Sigma - \frac{\Sigma\phi\phi^T\Sigma}{\phi^T\Sigma\phi + \sigma^2}\right]\Sigma^{-1}\bar{\theta} + \left[\frac{\Sigma\phi}{\phi^T\Sigma\phi + \sigma^2}\right] w \\ &= \Sigma_{\text{new}}\Sigma^{-1}\bar{\theta} + \left[\frac{\Sigma\phi}{\phi^T\Sigma\phi + \sigma^2}\right] w,\end{aligned}$$

that is,

$$\Sigma_{\text{new}}^{-1}\bar{\theta}_{\text{new}} = \Sigma^{-1}\bar{\theta} + \left[\frac{\Sigma_{\text{new}}^{-1}\Sigma\phi}{\phi^T\Sigma\phi + \sigma^2}\right] w. \tag{43}$$

Notice that

$$\Sigma_{\text{new}}^{-1}\Sigma\phi = \left[\Sigma^{-1} + \frac{1}{\sigma^2}\phi\phi^T\right]\Sigma\phi = \phi + \frac{\phi^T\Sigma\phi}{\sigma^2}\phi = \frac{\sigma^2 + \phi^T\Sigma\phi}{\sigma^2}\phi. \tag{44}$$

Plug Equation (44) into Equation (43), we have Equation (42). □

Based on Lemma 8, by mathematical induction, we have

$$\begin{aligned}\Sigma_t^{-1} &= \Sigma_1^{-1} + \frac{1}{\sigma^2}\sum_{\tau=1}^{t-1}\sum_{k=1}^{|A^\tau|}\phi_{a_k^\tau}\phi_{a_k^\tau}^T \\ \Sigma_t^{-1}\bar{\theta}_t &= \Sigma_1^{-1}\bar{\theta}_1 \frac{1}{\sigma^2}\sum_{\tau=1}^{t-1}\sum_{k=1}^{|A^\tau|}\phi_{a_k^\tau}\mathbf{w}_\tau\left(a_k^\tau\right),\end{aligned}$$

further noting that $\Sigma_1 = \lambda^2 I$ and $\bar{\theta}_1 = 0$, we can derive Equation (36).

D. A Variant of Theorem 2 for Approximation Algorithms

By suitably redefining the realized regret, we can prove a variant of Theorem 2 in which `ORACLE` can be an approximation algorithm. Specifically, for a (possibly approximation) algorithm `ORACLE`, let $A^*(\mathbf{w})$ be the solution of `ORACLE` to the optimization problem $(E, \mathcal{A}, \mathbf{w})$, we say $\gamma \in [0, 1)$ is a *sub-optimality gap* of `ORACLE` if

$$f(A^*(\mathbf{w}), \mathbf{w}) \geq (1-\gamma)\max_{A\in\mathcal{A}} f(A, \mathbf{w}), \quad \forall \mathbf{w}. \tag{45}$$

Then we define the (scaled) realized regret R_t^γ as

$$R_t^\gamma = f\left(A^{\text{opt}}, \mathbf{w}_t\right) - \frac{f\left(A^t, \mathbf{w}_t\right)}{1-\gamma}, \tag{46}$$

where A^{opt} is the exact solution to the optimization problem $(E, \mathcal{A}, \bar{\mathbf{w}})$. The (scaled) cumulative regret $R^\gamma(n)$ is defined as

$$R^\gamma(n) = \sum_{t=1}^{n}\mathbb{E}\left[R_t^\gamma|\bar{\mathbf{w}}\right].$$

Under the assumptions that (1) the support of P is a subset of $[0, 1]^L$ (i.e. $\mathbf{w}_t(e) \in [0, 1]\ \forall t$ and $\forall e \in E$), (2) the item weight $\mathbf{w}(e)$'s are statistically independent under P, and (3) the oracle `ORACLE` has sub-optimality gap $\gamma \in [0, 1)$, we have the following variant of Theorem 2 when `CombLinUCB` is applied to coherent learning cases:

Theorem 4. *For any $\lambda, \sigma > 0$, any $\delta \in (0,1)$, and any c satisfying*

$$c \geq \frac{1}{\sigma}\sqrt{d \ln\left(1 + \frac{nK\lambda^2}{d\sigma^2}\right) + 2\ln\left(\frac{1}{\delta}\right)} + \frac{\|\theta^*\|_2}{\lambda}, \tag{47}$$

if $\bar{\mathbf{w}} = \Phi\theta^*$ and the above three assumptions hold, then under `CombLinUCB` *algorithm with parameter $(\Phi, \lambda, \sigma, c)$, we have*

$$R^\gamma(n) \leq \frac{2cK\lambda}{1-\gamma}\sqrt{\frac{dn\ln\left(1+\frac{nK\lambda^2}{d\sigma^2}\right)}{\ln\left(1+\frac{\lambda^2}{\sigma^2}\right)}} + nK\delta.$$

Proof. Notice that Lemma 7 in Section B.1 still holds. With G defined in Equation (39), we have

$$\begin{aligned}
R^\gamma(n) &= \sum_{t=1}^n \mathbb{E}\left[\sum_{e\in A^{\mathrm{opt}}} \mathbf{w}_t(e) - \frac{1}{1-\gamma}\sum_{e\in A^t}\mathbf{w}_t(e)\right] \\
&= \sum_{t=1}^n \mathbb{E}\left[\sum_{e\in A^{\mathrm{opt}}} \bar{\mathbf{w}}(e) - \frac{1}{1-\gamma}\sum_{e\in A^t}\bar{\mathbf{w}}(e)\right] \\
&= \mathbb{P}(G)\sum_{t=1}^n \mathbb{E}\left[\sum_{e\in A^{\mathrm{opt}}} \bar{\mathbf{w}}(e) - \frac{1}{1-\gamma}\sum_{e\in A^t}\bar{\mathbf{w}}(e)\middle| G\right] + \mathbb{P}\left(\bar{G}\right)\sum_{t=1}^n \mathbb{E}\left[\sum_{e\in A^{\mathrm{opt}}} \bar{\mathbf{w}}(e) - \frac{1}{1-\gamma}\sum_{e\in A^t}\bar{\mathbf{w}}(e)\middle| \bar{G}\right] \\
&\leq \sum_{t=1}^n \mathbb{E}\left[\sum_{e\in A^{\mathrm{opt}}} \bar{\mathbf{w}}(e) - \frac{1}{1-\gamma}\sum_{e\in A^t}\bar{\mathbf{w}}(e)\middle| G\right] + \mathbb{P}\left(\bar{G}\right) nK,
\end{aligned}$$

where the last inequality follows from the naive bound on R_t^γ. If c satisfies inequality (40), we have $\mathbb{P}\left(\bar{G}\right) \leq \delta$, hence we have

$$R^\gamma(n) \leq \sum_{t=1}^n \mathbb{E}\left[\sum_{e\in A^{\mathrm{opt}}} \bar{\mathbf{w}}(e) - \frac{1}{1-\gamma}\sum_{e\in A^t}\bar{\mathbf{w}}(e)\middle| G\right] + nK\delta.$$

Finally, we bound $\sum_{t=1}^n \mathbb{E}\left[\sum_{e\in A^{\mathrm{opt}}} \bar{\mathbf{w}}(e) - \frac{1}{1-\gamma}\sum_{e\in A^t}\bar{\mathbf{w}}(e)\middle| G\right]$ using a worst-case bound conditioning on G (worst-case over all the possible random realizations), notice that conditioning on G, we have

$$\sum_{e\in A^{\mathrm{opt}}} \bar{\mathbf{w}}(e) \leq \sum_{e\in A^{\mathrm{opt}}} \hat{\mathbf{w}}_t(e) \leq \max_{A\in\mathcal{A}}\sum_{e\in A}\hat{\mathbf{w}}_t(e) \leq \frac{1}{1-\gamma}\sum_{e\in A^t}\hat{\mathbf{w}}_t(e),$$

where

- The first inequality follows from the definition of event G. Specifically, under event G, $\bar{\mathbf{w}}(e) \leq \hat{\mathbf{w}}_t(e)$ for all $t = 1, \ldots, n$ and all $e \in E$.
- The second inequality follows from $A^{\mathrm{opt}} \in \mathcal{A}$.
- The last inequality follows from $A^t \leftarrow$ `ORACLE`$(E, \mathcal{A}, \hat{\mathbf{w}}_t)$ and `ORACLE` has sub-optimality gap γ (see Equation (45)).

Thus we have

$$\begin{aligned}
\sum_{e\in A^{\mathrm{opt}}} \bar{\mathbf{w}}(e) - \frac{1}{1-\gamma}\sum_{e\in A^t}\bar{\mathbf{w}}(e) &\leq \frac{1}{1-\gamma}\left[\sum_{e\in A^t}\hat{\mathbf{w}}_t(e) - \sum_{e\in A^t}\bar{\mathbf{w}}(e)\right] \\
&= \frac{1}{1-\gamma}\sum_{e\in A^t}\left[\left\langle \phi_e, \bar{\theta}_t - \theta^*\right\rangle + c\sqrt{\phi_e^T\Sigma_t\phi_e}\right] \\
&\leq \frac{2c}{1-\gamma}\sum_{e\in A^t}\sqrt{\phi_e^T\Sigma_t\phi_e},
\end{aligned}$$

where the last inequality follows from the definition of G. Recall that from Lemma 4, we also have

$$\sum_{t=1}^{n} \sum_{e \in A^t} \sqrt{\phi_e^T \Sigma_t \phi_e} \leq K\lambda \sqrt{\frac{dn \log\left(1 + \frac{nK\lambda^2}{d\sigma^2}\right)}{\log\left(1 + \frac{\lambda^2}{\sigma^2}\right)}}.$$

Putting the above inequalities together, we have proved the theorem. □