This is the online technical report of the manuscript "*Incentivized Online Learning Mechanism Design*". The remainder is organized as follows:

- Appendix A provides the proof of Theorem 1.
- Appendix B provides the proof of Theorem 2.
- Appendix C provides the proofs of Theorem 3 and Corollary 1.
- Appendix D provides two lemmas, which are used in Appendix B.
- Appendix E provides two lemmas, which are used in Appendix B.

Appendix A

Proof of Theorem 1. We first introduce several preliminary results in Step I, and then prove this theorem in Step II based on the results in Step I. For notation simplicity, we will suppress the time index t in this proof.

Step I: Recall that $\boldsymbol{c} = (\boldsymbol{c}_n : \forall n \in \mathcal{N})$ is the agents' real cost. We let $\hat{\boldsymbol{c}} = (\hat{\boldsymbol{c}}_{-\boldsymbol{n}}, \hat{\boldsymbol{c}}_n)$ denote the bidding cost of all agents, where $\hat{\boldsymbol{c}}_{-\boldsymbol{n}}$ is the bidding costs of all agents except agent n. When agent n truthfully reports his cost, i.e., $\hat{\boldsymbol{c}}_n = \boldsymbol{c}_n$, we let $\boldsymbol{x}^*$ denote the corresponding optimal assignment, i.e.,

$$\boldsymbol{x}^* \triangleq \max_{\boldsymbol{x} \in \mathcal{X}} \quad \mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}; \hat{\boldsymbol{r}}, \hat{\boldsymbol{c}}_{-\boldsymbol{n}}, \boldsymbol{c}_n). \tag{32}$$

If agent n does not truthfully report, i.e., $\hat{\boldsymbol{c}}_n = \bar{\boldsymbol{c}}_n$, we let $\bar{\boldsymbol{x}}^*$ denote the corresponding optimal assignment, i.e.,

$$\bar{\boldsymbol{x}}^* \triangleq \max_{\boldsymbol{x} \in \mathcal{X}} \quad \mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}; \hat{\boldsymbol{r}}, \hat{\boldsymbol{c}}_{-\boldsymbol{n}}, \bar{\boldsymbol{c}}_n). \tag{33}$$

Based on (32) and (33), we have

$$\mathcal{L}\left(\boldsymbol{x}^*, \boldsymbol{\lambda}; \hat{\boldsymbol{r}}, \hat{\boldsymbol{c}}_{-\boldsymbol{n}}, \boldsymbol{c}_n\right) \geq \mathcal{L}\left(\bar{\boldsymbol{x}}^*, \boldsymbol{\lambda}; \hat{\boldsymbol{r}}, \hat{\boldsymbol{c}}_{-n}, \boldsymbol{c}_n\right). \tag{34}$$

According to the definition of $\mathcal{L}^\star_{-n}(\cdot)$ in (23), we have

$$\mathcal{L}^\star_{-n}\left(\boldsymbol{\lambda}, \hat{\boldsymbol{r}}, \hat{\boldsymbol{c}}_{-\boldsymbol{n}}, \boldsymbol{c}_n\right) = \mathcal{L}^\star_{-n}\left(\boldsymbol{\lambda}, \hat{\boldsymbol{r}}, \hat{\boldsymbol{c}}_{-\boldsymbol{n}}, \bar{\boldsymbol{c}}_n\right). \tag{35}$$

Step II: We take agent n as a representative agent. When agent n truthfully bid, given the bidding outcome $(\hat{\boldsymbol{c}}_{-n}, \boldsymbol{c}_n)$, the payoff of agent n is given by

$$\begin{aligned}
U_n(\hat{\boldsymbol{c}}_{-\boldsymbol{n}}, \boldsymbol{c}_n) &= \left[\sum_{k=1}^K \hat{r}_k x^*_{n,k} - \lambda_n f_n(\boldsymbol{x}^*, \mathbf{1}_N) + \mathcal{L}_{-n}\left(\boldsymbol{x}^*; \hat{\boldsymbol{r}}, \hat{\boldsymbol{c}}_{-\boldsymbol{n}}, \boldsymbol{c}_n\right) - \mathcal{L}^\star_{-n}\left(\boldsymbol{\lambda}, \hat{\boldsymbol{r}}, \hat{\boldsymbol{c}}_{-\boldsymbol{n}}, \boldsymbol{c}_n\right)\right] - \sum_{k=1}^K c_{n,k} x^*_{n,k} \\
&= \left[\sum_{k=1}^K \hat{r}_k x^*_{n,k} - \lambda_n f_n(\boldsymbol{x}^*, \mathbf{1}_N) - \sum_{k=1}^K c_{n,k} x^*_{n,k} + \mathcal{L}_{-n}\left(\boldsymbol{x}^*; \hat{\boldsymbol{r}}, \hat{\boldsymbol{c}}_{-\boldsymbol{n}}, \boldsymbol{c}_n\right)\right] - \mathcal{L}^\star_{-n}\left(\boldsymbol{\lambda}, \hat{\boldsymbol{r}}, \hat{\boldsymbol{c}}_{-\boldsymbol{n}}, \boldsymbol{c}_n\right) \\
&= \mathcal{L}\left(\boldsymbol{x}^*, \boldsymbol{\lambda}; \hat{\boldsymbol{r}}, \hat{\boldsymbol{c}}_{-\boldsymbol{n}}, \boldsymbol{c}_n\right) - \mathcal{L}^\star_{-n}\left(\boldsymbol{\lambda}, \hat{\boldsymbol{r}}, \hat{\boldsymbol{c}}_{-\boldsymbol{n}}, \boldsymbol{c}_n\right),
\end{aligned} \tag{36}$$

where the last equality follows the definition of $\mathcal{L}(\cdot)$ and $\mathcal{L}_{-n}(\cdot)$.

When agent n does not truthfully bid, given the bidding outcome $(\hat{\boldsymbol{c}}_{-n}, \bar{\boldsymbol{c}}_n)$, the payoff of agent n is

$$\begin{aligned}
U_n(\hat{\boldsymbol{c}}_{-\boldsymbol{n}}, \bar{\boldsymbol{c}}_n) &= \left[\sum_{k=1}^K \hat{r}_k \bar{x}^*_{n,k} - \lambda_n f_n(\bar{\boldsymbol{x}}^*, \mathbf{1}_N) + \mathcal{L}_{-n}\left(\bar{\boldsymbol{x}}^*, \boldsymbol{\lambda}; \hat{\boldsymbol{r}}, \hat{\boldsymbol{c}}_{-\boldsymbol{n}}, \bar{\boldsymbol{c}}_n\right) - \mathcal{L}^\star_{-n}\left(\boldsymbol{\lambda}, \hat{\boldsymbol{r}}, \hat{\boldsymbol{c}}_{-\boldsymbol{n}}, \bar{\boldsymbol{c}}_n\right)\right] - \sum_{k=1}^K c_{n,k} \bar{x}^*_{n,k} \\
&= \left[\sum_{k=1}^K \hat{r}_k \bar{x}^*_{n,k} - \lambda_n f_n(\bar{\boldsymbol{x}}^*, \mathbf{1}_N) - \sum_{k=1}^K c_{n,k} \bar{x}^*_{n,k} + \mathcal{L}_{-n}\left(\bar{\boldsymbol{x}}^*, \boldsymbol{\lambda}; \hat{\boldsymbol{r}}, \hat{\boldsymbol{c}}_{-\boldsymbol{n}}, \bar{\boldsymbol{c}}_n\right)\right] - \mathcal{L}^\star_{-n}\left(\boldsymbol{\lambda}, \hat{\boldsymbol{r}}, \hat{\boldsymbol{c}}_{-\boldsymbol{n}}, \bar{\boldsymbol{c}}_n\right) \\
&\overset{(a)}{=} \mathcal{L}\left(\bar{\boldsymbol{x}}^*, \boldsymbol{\lambda}; \hat{\boldsymbol{r}}, \hat{\boldsymbol{c}}_{-n}, \boldsymbol{c}_n\right) - \mathcal{L}^\star_{-n}\left(\boldsymbol{\lambda}, \hat{\boldsymbol{r}}, \hat{\boldsymbol{c}}_{-\boldsymbol{n}}, \bar{\boldsymbol{c}}_n\right),
\end{aligned} \tag{37}$$

where the last equality follows the definition of $\mathcal{L}(\cdot)$ and $\mathcal{L}_{-n}(\cdot)$.

Comparing (36) and (37), we obtain

$$\begin{aligned}
& U_n(\hat{\boldsymbol{c}}_{-\boldsymbol{n}}, \boldsymbol{c}_n) - U_n(\hat{\boldsymbol{c}}_{-\boldsymbol{n}}, \bar{\boldsymbol{c}}_n) \\
=& \mathcal{L}\left(\boldsymbol{x}^*, \boldsymbol{\lambda}; \hat{\boldsymbol{r}}, \hat{\boldsymbol{c}}_{-\boldsymbol{n}}, \boldsymbol{c}_n\right) - \mathcal{L}^\star_{-n}\left(\boldsymbol{\lambda}, \hat{\boldsymbol{r}}, \hat{\boldsymbol{c}}_{-\boldsymbol{n}}, \boldsymbol{c}_n\right) - \mathcal{L}\left(\bar{\boldsymbol{x}}^*, \boldsymbol{\lambda}; \hat{\boldsymbol{r}}, \hat{\boldsymbol{c}}_{-n}, \boldsymbol{c}_n\right) + \mathcal{L}^\star_{-n}\left(\boldsymbol{\lambda}, \hat{\boldsymbol{r}}, \hat{\boldsymbol{c}}_{-\boldsymbol{n}}, \bar{\boldsymbol{c}}_n\right) \\
\overset{(a)}{=}& \mathcal{L}\left(\boldsymbol{x}^*, \boldsymbol{\lambda}; \hat{\boldsymbol{r}}, \hat{\boldsymbol{c}}_{-\boldsymbol{n}}, \boldsymbol{c}_n\right) - \mathcal{L}\left(\bar{\boldsymbol{x}}^*, \boldsymbol{\lambda}; \hat{\boldsymbol{r}}, \hat{\boldsymbol{c}}_{-n}, \boldsymbol{c}_n\right) \\
\overset{(b)}{\geq}& 0,
\end{aligned} \tag{38}$$

where the equality (a) follows (35) and the inequality (b) follows (34). This means that it is the dominant strategy for agent n to truthfully bid. Accordingly, the payoff of the agent n at the equilibrium is

$$U_n(\boldsymbol{c}) = \mathcal{L}\left(\boldsymbol{x}^*, \boldsymbol{\lambda}; \hat{\boldsymbol{r}}, \boldsymbol{c}\right) - \mathcal{L}^\star_{-n}\left(\boldsymbol{\lambda}, \hat{\boldsymbol{r}}, \boldsymbol{c}\right) \geq 0, \tag{39}$$

which indicates that $\hat{a}_n = 1$ for any agent $n \in \mathcal{N}$. This completes the proof of Theorem 1. $\square$

Appendix B

Proof of Theorem 2. We will prove this theorem in three steps.

Step I: We drive an upper bound for $\texttt{Vio}(T)$. Specifically, the updating policy (19) indicates that

$$\lambda_n^{t+1} = \big(\lambda_n^t + \eta\big[f_n(\hat{\boldsymbol{x}}^t, \mathbf{1}_N) - \phi_n\big]\big)^+ \geq \lambda_n^t + \eta\big[f_n(\hat{\boldsymbol{x}}^t, \mathbf{1}_N) - \phi_n\big], \tag{40}$$

which leads to

$$\lambda_n^{t+1} - \lambda_n^t \geq \eta\Big[f_n(\hat{\boldsymbol{x}}^t, \mathbf{1}_N) - \phi_n\Big]. \tag{41}$$

Summing the inequality over $t \in \{1, 2, ..., T\}$ yields

$$\sum_{t=1}^{T}\Big[f_n(\hat{\boldsymbol{x}}^t, \mathbf{1}_N) - \phi_n\Big] \leq \sum_{t=1}^{T}\frac{\lambda_n^{t+1} - \lambda_n^t}{\eta} = \frac{\lambda_n^{T+1}}{\eta}, \tag{42}$$

where the last equality follows the telescoping technique and $\boldsymbol{\lambda}^1 = \mathbf{0}_N$ in Algorithm 1. Therefore, the fairness violation has the following upper bound

$$\texttt{Vio}(T) \leq \frac{\mathbb{E}\|\boldsymbol{\lambda}^{T+1}\|_1}{\eta} \leq \Xi(\delta) + \frac{\sqrt{N}\Theta}{\delta\eta}, \tag{43}$$

where the last inequality follows from Lemma 1.

Substituting the step-size $\eta = \frac{4K+2\sqrt{6K\Phi T\ln T}}{T\Theta}$, we obtain

$$\texttt{Vio}(T) \leq \Xi(\delta) + \frac{1}{\delta}\cdot\frac{\sqrt{N}\Theta^2 T}{4K + 2\sqrt{6K\Phi T\ln T}} \leq \Xi(\delta) + \frac{\Theta^2\sqrt{NT}}{2\delta\sqrt{6K\Phi}} \leq \Xi(\delta) + \frac{\Theta^2}{4\delta}\sqrt{\frac{NT}{K\Phi}}, \tag{44}$$

which completes the proof of Step I for this theorem.

Step II: We derive an upper bound for the cumulative regret $\texttt{Reg}(T)$. The updating policy (19) leads to

$$\begin{aligned}
\big(\lambda_n^{t+1}\big)^2 &= \Big[\big(\lambda_n^t + \eta\big[f_n(\hat{\boldsymbol{x}}^t, \mathbf{1}_N) - \phi_n\big]\big)^+\Big]^2 \overset{(a)}{\leq} \Big[\lambda_n^t + \eta\big[f_n(\hat{\boldsymbol{x}}^t, \mathbf{1}_N) - \phi_n\big]\Big]^2 \\
&= \big(\lambda_n^t\big)^2 + \eta^2\big[f_n(\hat{\boldsymbol{x}}^t, \mathbf{1}_N) - \phi_n\big]^2 + 2\eta\lambda_n^t\big[f_n(\hat{\boldsymbol{x}}^t, \mathbf{1}_N) - \phi_n\big], \\
&\overset{(b)}{\leq} \big(\lambda_n^t\big)^2 + \eta^2\big|f_n(\hat{\boldsymbol{x}}^t, \mathbf{1}_N) - \phi_n\big| + 2\eta\lambda_n^t\big[f_n(\hat{\boldsymbol{x}}^t, \mathbf{1}_N) - \phi_n\big],
\end{aligned} \tag{45}$$

where the inequality (a) follows $(x)^+ \leq |x|$ and the inequality (b) follows $x^2 \leq x$ for any $x \in [0, 1]$. Now we derive an upper bound for $\sum_{n=1}^N |f_n(\hat{\boldsymbol{x}}^t, \mathbf{1}_N) - \phi_n|$ in the following

$$\begin{cases}
\displaystyle\sum_{n=1}^{N}|f_n(\hat{\boldsymbol{x}}^t, \mathbf{1}_N) - \phi_n| \overset{(a)}{\leq} \sum_{n=1}^{N}|f_n(\hat{\boldsymbol{x}}^t, \mathbf{1}_N)| + \sum_{n=1}^{N}|\phi_n| \overset{(b)}{\leq} \min(K, N) + \Phi, \\
\displaystyle\sum_{n=1}^{N}|f_n(\hat{\boldsymbol{x}}^t, \mathbf{1}_N) - \phi_n| \leq \sum_{n=1}^{N}1 = N,
\end{cases} \tag{46}$$

where the inequality (a) follows from the triangle inequality, the inequality (b) follows from $\sum_{n=1}^N |f_n(\boldsymbol{x}, \mathbf{1}_N)| \leq \min(K, N)$ for any $\boldsymbol{x} \in \mathcal{X}$. Hence we obtain $\sum_{n=1}^N |f_n(\hat{\boldsymbol{x}}^t, \mathbf{1}_N) - \phi_n| \leq \Theta$.

Based on (45) and $\sum_{n=1}^N |f_n(\hat{\boldsymbol{x}}^t, \mathbf{1}_N) - \phi_n| \leq \Theta$, we obtain

$$\begin{aligned}
\|\boldsymbol{\lambda}^{t+1}\|_2^2 - \|\boldsymbol{\lambda}^t\|_2^2 \leq& \eta^2\Theta + 2\eta\sum_{n=1}^{N}\lambda_n^t\left[f_n(\hat{\boldsymbol{x}}^t, \mathbf{1}_N) - \phi_n\right] \\
=& \eta^2\Theta + 2\eta S(\hat{\boldsymbol{x}}^t, \mathbf{1}_N; \hat{\boldsymbol{r}}^t, \hat{\boldsymbol{c}}^t) - 2\eta\left[S(\hat{\boldsymbol{x}}^t, \mathbf{1}_N; \hat{\boldsymbol{r}}^t, \hat{\boldsymbol{c}}^t) - \sum_{n=1}^{N}\lambda_n^t\left[f_n(\hat{\boldsymbol{x}}^t, \mathbf{1}_N) - \phi_n\right]\right] \\
\leq& \eta^2\Theta + 2\eta S(\hat{\boldsymbol{x}}^t, \mathbf{1}_N; \hat{\boldsymbol{r}}^t, \hat{\boldsymbol{c}}^t) - 2\eta\left[S(\tilde{\boldsymbol{x}}^t, \mathbf{1}_N; \hat{\boldsymbol{r}}^t, \hat{\boldsymbol{c}}^t) - \sum_{n=1}^{N}\lambda_n^t\left[f_n(\tilde{\boldsymbol{x}}^t, \mathbf{1}_N) - \phi_n\right]\right],
\end{aligned} \tag{47}$$

where the assignment $\tilde{\boldsymbol{x}}^t$ represents the optimal assignment policy in Baseline 1 and the last inequality follows the definition of (18). Taking the expectation leads to

$$\begin{aligned}
\mathbb{E}\left[\|\boldsymbol{\lambda}^{t+1}\|_2^2 - \|\boldsymbol{\lambda}^t\|_2^2\right] \leq& \Theta\eta^2 + 2\eta\mathbb{E}\Big[S(\hat{\boldsymbol{x}}^t, \mathbf{1}_N; \hat{\boldsymbol{r}}^t, \hat{\boldsymbol{c}}^t) - S(\tilde{\boldsymbol{x}}^t, \mathbf{1}_N; \hat{\boldsymbol{r}}^t, \hat{\boldsymbol{c}}^t)\Big] + 2\eta\sum_{n=1}^{N}\mathbb{E}\Big[\lambda_n^t\left[f_n(\tilde{\boldsymbol{x}}^t, \mathbf{1}_N) - \phi_n\right]\Big], \\
\leq& \Theta\eta^2 + 2\eta\mathbb{E}\Big[S(\hat{\boldsymbol{x}}^t, \mathbf{1}_N; \hat{\boldsymbol{r}}^t, \hat{\boldsymbol{c}}^t) - S(\tilde{\boldsymbol{x}}^t, \mathbf{1}_N; \hat{\boldsymbol{r}}^t, \hat{\boldsymbol{c}}^t)\Big],
\end{aligned} \tag{48}$$

where the last inequality follows that $\mathbb{E}\left[\lambda_n^t(f_n(\tilde{\boldsymbol{x}}^t, \mathbf{1}_N) - \phi_n)\right] = \mathbb{E}[\lambda_n^t] \cdot \mathbb{E}\left[f_n(\tilde{\boldsymbol{x}}^t, \mathbf{1}_N) - \phi_n\right] \leq 0$.

Summing the above inequality over $t \in \{1, 2, ..., T\}$ and using telescoping technique, we obtain

$$\left\|\boldsymbol{\lambda}^{T+1}\right\|_2^2 \leq T\Theta\eta^2 + 2\eta \sum_{t=1}^{T} \mathbb{E}\big[S(\hat{\boldsymbol{x}}^t, \mathbf{1}_N; \hat{\boldsymbol{r}}^t, \hat{\boldsymbol{c}}^t) - S(\tilde{\boldsymbol{x}}^t, \mathbf{1}_N; \hat{\boldsymbol{r}}^t, \hat{\boldsymbol{c}}^t)\big]. \tag{49}$$

On the RHS of (49), we have

$$\begin{aligned}
&\sum_{t=1}^{T} \mathbb{E}\Big[S(\hat{\boldsymbol{x}}^t, \mathbf{1}_N; \hat{\boldsymbol{r}}^t, \hat{\boldsymbol{c}}^t) - S(\tilde{\boldsymbol{x}}^t, \mathbf{1}_N; \hat{\boldsymbol{r}}^t, \hat{\boldsymbol{c}}^t)\Big] \\
=&\sum_{t=1}^{T} \mathbb{E}\left[\sum_{n=1}^{N}\sum_{k=1}^{K} \big(\hat{r}_k^t - \hat{c}_{n,k}^t\big)\hat{x}_{n,k}^t - \big(\hat{r}_k^t - \hat{c}_{n,k}^t\big)\tilde{x}_{n,k}^t\right] \\
\overset{(a)}{=}&\sum_{t=1}^{T} \mathbb{E}\left[\sum_{n=1}^{N}\sum_{k=1}^{K} \big(\hat{r}_k^t - c_{n,k}^t\big)\hat{x}_{n,k}^t - \big(\hat{r}_k^t - c_{n,k}^t\big)\tilde{x}_{n,k}^t\right] \\
=&\sum_{t=1}^{T} \mathbb{E}\left[\sum_{n=1}^{N}\sum_{k=1}^{K} \big(\bar{r}_k - c_{n,k}^t\big)\hat{x}_{n,k}^t - \big(\bar{r}_k - c_{n,k}^t\big)\tilde{x}_{n,k}^t\right] + \sum_{t=1}^{T} \mathbb{E}\left[\sum_{n=1}^{N}\sum_{k=1}^{K} \big(\hat{r}_k^t - \bar{r}_k\big)\hat{x}_{n,k}^t - \big(\hat{r}_k^t - \bar{r}_k\big)\tilde{x}_{n,k}^t\right] \\
\overset{(b)}{=}&- \mathtt{Reg}(T) + \sum_{t=1}^{T} \mathbb{E}\Big[W_1^t + W_2^t\Big]
\end{aligned} \tag{50}$$

where the equality (a) follows from the truthful bidding proved in Theorem 1, and the equality (b) follows the definition of $\mathtt{Reg}(T)$ and the definitions of W_1^t and W_2^t as follows:

$$\begin{aligned}
W_1^t &\triangleq \sum_{n=1}^{N}\sum_{k=1}^{K} \left(\hat{r}_k^t - \bar{r}_k\right) \hat{x}_{n,k}^t, \\
W_2^t &\triangleq \sum_{n=1}^{N}\sum_{k=1}^{K} \left(\bar{r}_k - \hat{r}_k^t\right) \tilde{x}_{n,k}^t.
\end{aligned} \tag{51}$$

Substituting (50) into (49) and using $\|\boldsymbol{\lambda}^{T+1}\|_2^2 \geq 0$, we obtain

$$\mathtt{Reg}(T) \leq \sum_{t=1}^{T} \mathbb{E}\Big[W_1^t + W_2^t\Big] + \frac{\Theta\eta T}{2} \leq 4K + 2\sqrt{6K[\mathtt{Vio}(T) + T\Phi]\ln T} + \frac{\Theta\eta T}{2}, \tag{52}$$

where the last inequality is based on Lemma 3 and Lemma 4 in Appendix E.

With the step-size $\eta \triangleq \frac{4K+2\sqrt{6K\Phi T\ln T}}{T\Theta}$, we obtain

$$\begin{aligned}
\mathtt{Reg}(T) \leq& 4K + 2\sqrt{6K\left[\mathtt{Vio}(T) + T\Phi\right]\ln T} + 2K + \sqrt{6KT\Phi\ln T} \\
\leq& 6K + 3\sqrt{6K\left[\mathtt{Vio}(T) + T\Phi\right]\ln T},
\end{aligned} \tag{53}$$

which completes the proof of Step II for this theorem.

Part III: Next we analyze the principal cumulative profit. The profit of the principal in slot t is given by

$$\begin{aligned}
&V(\hat{\boldsymbol{x}}^t, \mathbf{1}_N; \boldsymbol{r}^t) - \sum_{n=1}^{N} \hat{y}_n^t \\
=&\sum_{k=1}^{K}\sum_{n=1}^{N} r_k^t \hat{x}_{n,k}^t - \sum_{n=1}^{N}\left[\sum_{k=1}^{K} \hat{r}_k^t \hat{x}_{n,k}^t - \lambda_n^t f_n(\hat{\boldsymbol{x}}^t, \mathbf{1}_N) + \mathcal{L}_{-n}(\hat{\boldsymbol{x}}^t, \boldsymbol{\lambda}^t, \hat{\boldsymbol{r}}^t, \hat{\boldsymbol{c}}^t) - \mathcal{L}_{-n}^{\star}(\boldsymbol{\lambda}^t, \hat{\boldsymbol{r}}^t, \hat{\boldsymbol{c}}^t)\right] \\
=&- \sum_{n=1}^{N}\sum_{k=1}^{K} \left(\hat{r}_k^t - r_k^t\right) \hat{x}_{n,k}^t + \sum_{n=1}^{N} \lambda_n^t f_n(\hat{\boldsymbol{x}}^t, \mathbf{1}_N) + \sum_{n=1}^{N} \left[\mathcal{L}_{-n}^{\star}(\boldsymbol{\lambda}^t, \hat{\boldsymbol{r}}^t, \hat{\boldsymbol{c}}^t) - \mathcal{L}_{-n}(\hat{\boldsymbol{x}}^t, \boldsymbol{\lambda}^t, \hat{\boldsymbol{r}}^t, \hat{\boldsymbol{c}}^t)\right] \\
\geq&- \sum_{n=1}^{N}\sum_{k=1}^{K} \left(\hat{r}_k^t - r_k^t\right) \hat{x}_{n,k}^t,
\end{aligned} \tag{54}$$

where the last inequality follows that $\lambda_n^t f_n(\hat{\boldsymbol{x}}^t, \mathbf{1}_N) \geq 0$ and $\mathcal{L}_{-n}^{\star}(\boldsymbol{\lambda}^t, \hat{\boldsymbol{r}}^t, \hat{\boldsymbol{c}}^t) \geq \mathcal{L}_{-n}(\hat{\boldsymbol{x}}^t, \boldsymbol{\lambda}^t, \hat{\boldsymbol{r}}^t, \hat{\boldsymbol{c}}^t)$ for any $n \in \mathcal{N}$.

Taking the expectation and summing over $t \in \{1, 2, ..., T\}$, we obtain

$$0 \leq \mathtt{Pro}(T) + \sum_{t=1}^{T} \mathbb{E}\left[W_1^t\right] \leq \mathtt{Pro}(T) + \frac{5K}{2} + 2\sqrt{6KT\left[\Phi + \frac{\mathtt{Vio}(T)}{T}\right]\ln T}, \tag{55}$$

where the last inequality follows Lemma 3. This completes the proof of this theorem. □

APPENDIX C

Next we prove Theorem 3 and Corollary 1.

Proof of Theorem 3. We decompose the cumulative degradation as follows:

$$\begin{aligned}\texttt{Deg}(T) &= S^{\dagger}T - \sum_{t=1}^{T}\sum_{n=1}^{N}\sum_{k=1}^{K}\mathbb{E}\left[\left(\bar{r}_k - c_{n,k}^t\right)\hat{x}_{n,k}^t\right] \\ &= S^{\dagger}T - \sum_{t=1}^{T}\sum_{n=1}^{N}\sum_{k=1}^{K}\left(\mathbb{E}\left[\left(\bar{r}_k - c_{n,k}^t\right)\tilde{x}_{n,k}^t\right] - \mathbb{E}\left[\left(\bar{r}_k - c_{n,k}^t\right)\tilde{x}_{n,k}^t\right] + \mathbb{E}\left[\left(\bar{r}_k - c_{n,k}^t\right)\hat{x}_{n,k}^t\right]\right) \\ &= \sum_{t=1}^{T}\left[S^{\dagger} - \sum_{n=1}^{N}\sum_{k=1}^{K}\mathbb{E}\left[\left(\bar{r}_k - c_{n,k}^t\right)\tilde{x}_{n,k}^t\right]\right] + \texttt{Reg}(T),\end{aligned} \tag{56}$$

where $\tilde{\boldsymbol{x}}^t$ corresponds to the optimal assignment policy defined in Baseline 1. Moreover, based on Baseline 2, we have

$$\begin{aligned}S^{\dagger} - \sum_{n=1}^{N}\sum_{k=1}^{K}\mathbb{E}\left[\left(\bar{r}_k - c_{n,k}^t\right)\tilde{x}_{n,k}^t\right] &= \sum_{k=1}^{K}\left[\left(\bar{r}_k - c_{\min}\right)p_k^* - \sum_{n=1}^{N}\mathbb{E}\left[\left(\bar{r}_k - c_{n,k}^t\right)\tilde{x}_{n,k}^t\right]\right] \\ &\overset{(a)}{\leq} \sum_{k=1}^{K}\left[\left(\bar{r}_k - c_{\min}\right)p_k^* - \mathbb{E}\left[\left(\bar{r}_k - \min_{n\in\mathcal{N}} c_{n,k}^t\right)p_k^*\right]\right] \\ &= \sum_{k=1}^{K}\left(\mathbb{E}\left[\min_{n\in\mathcal{N}} c_{n,k}^t\right] - c_{\min}\right)p_k^*,\end{aligned} \tag{57}$$

where $\boldsymbol{p}^* = (p_k^* : \forall k \in \mathcal{K})$ is the optimal solution in (14), and the inequality (a) follows the definition of the optimal assignment policy in Baseline 1,

For notation simplicity, we define $Y_k \triangleq \min_{n\in\mathcal{N}} c_{n,k}^t$. Given the PDF $g_{n,k}(\cdot)$ and CDF $G_{n,k}(\cdot)$ of the agent n's cost $\{c_{n,k}^t : \forall t\}$, one can show that the CDF of Y_k is given by

$$P_k(c) = \Pr(Y_k \leq c) = 1 - \Pr\left(Y_k > c\right) = 1 - \Pr\left(c_{n,k}^t > c, \forall n \in \mathcal{N}\right) = \begin{cases} 0, & \textit{if } c < c_{\min}, \\ 1 - \left[1 - G_k(c)\right]^N, & \textit{if } c \geq c_{\min}, \end{cases} \tag{58}$$

Based on the above discussion, the expectation of Y_k is given by

$$\mathbb{E}\left[Y_k\right] = \int_0^1 z \mathrm{d}P_k(z) = c_{\min} + \int_{c_{\min}}^1 \left[1 - G_k(z)\right]^N \mathrm{d}z. \tag{59}$$

We substitute (59) into (57) and obtain

$$S^{\dagger} - \sum_{n=1}^{N}\sum_{k=1}^{K}\mathbb{E}\left[\left(\bar{r}_k - c_{n,k}^t\right)\tilde{x}_{n,k}^t\right] = \sum_{k=1}^{K} p_k^* \int_{c_{\min}}^1 \left[1 - G_k(z)\right]^N \mathrm{d}z = \sum_{k=1}^{K} p_k^* \int_0^1 \frac{\left[1 - G_k(z)\right]^N}{g_k(z)}\mathrm{d}G_k(z). \tag{60}$$

According to the mean value theorem for integral, there exists a constant ϵ such that

$$\sum_{k=1}^{K} p_k^* \int_0^1 \frac{\left[1 - G_k(z)\right]^N}{g_k(z)}\mathrm{d}G_k(z) \leq \sum_{k=1}^{K} p_k^* \epsilon \int_0^1 \left[1 - G_k(z)\right]^N \mathrm{d}G_k(z) = \frac{\alpha\epsilon}{N+1}, \tag{61}$$

where last equality follows from $\int_0^1 (1-z)^N \mathrm{d}z = \frac{1}{N+1}$ and $\sum_{k=1}^{K} p_k^* = \Phi = \alpha$. Therefore, we have

$$\texttt{Deg}(T) \leq \texttt{Reg}(T) + \frac{T\alpha\epsilon}{N+1}. \tag{62}$$

This completes the proof of this theorem. □

Proof of Corollary 1. We have $\phi_{\min} = \alpha/N$ and $\Theta = \min(K+\alpha, N)$. According to Theorem 2, we let $\delta = \frac{\phi_{\min}}{2}$ and $\Xi(\delta)$ is given by

$$\begin{aligned}\Xi(\delta) =& \frac{3\sqrt{N}\Theta^2}{\phi_{\min} - \delta}\ln\left(\frac{2\Theta}{\phi_{\min} - \delta}\right) + \frac{3\sqrt{N}\Theta}{2\delta} = \frac{6\sqrt{N}\Theta^2}{\phi_{\min}}\ln\left(\frac{4\Theta}{\phi_{\min}}\right) + \frac{3\sqrt{N}\Theta}{\phi_{\min}} \\ =& \frac{6\sqrt{N^3}\Theta^2}{\alpha}\ln\left(\frac{4N\Theta}{\alpha}\right) + \frac{3\sqrt{N^3}\Theta}{\alpha} \leq \frac{6T^{\frac{3\beta}{2}}\Theta^2}{\sqrt{\alpha}}\ln\left(\frac{4T^{\beta}\Theta}{\alpha^{\frac{2}{3}}}\right) + \frac{3T^{\frac{3\beta}{2}}\Theta}{\sqrt{\alpha}},\end{aligned} \tag{63}$$

where the last inequality follows that $N = \lfloor \alpha^{\frac{1}{3}} T^{\beta} \rfloor \le \alpha^{\frac{1}{3}} T^{\beta}$. Moreover, we have

$$\begin{aligned} \frac{\Theta^2}{4\delta}\sqrt{\frac{NT}{K\Phi}} &= \frac{(K+\alpha)^2}{2\phi_{\min}}\sqrt{\frac{NT}{K\alpha}} = \frac{(K+\alpha)^2}{2\alpha} N^{\frac{3}{2}}\sqrt{\frac{T}{K\alpha}} \\ &\le \frac{(K+\alpha)^2}{2\alpha}\sqrt{\alpha} T^{\frac{3\beta}{2}}\sqrt{\frac{T}{K\alpha}} = \frac{(K+\alpha)^2}{2\alpha K}\sqrt{KT^{1+3\beta}}, \end{aligned} \tag{64}$$

where the inequality follows from $N = \lfloor \alpha^{\frac{1}{3}} T^{\beta} \rfloor \le \alpha^{\frac{1}{3}} T^{\beta}$.

In this case, the fairness violation is given by

$$\mathtt{Vio}(T) \le \frac{6T^{\frac{3\beta}{2}}\Theta^2}{\sqrt{\alpha}} \ln\left(\frac{4T^{\beta}\Theta}{\alpha^{\frac{2}{3}}}\right) + \frac{3T^{\frac{3\beta}{2}}\Theta}{\sqrt{\alpha}} + \frac{(K+\alpha)^2}{2\alpha K}\sqrt{KT^{1+3\beta}} \le O\left(\frac{K}{\alpha}\sqrt{KT^{1+3\beta}}\right). \tag{65}$$

The cumulative degradation is

$$\mathtt{Deg}(T) \le \mathtt{Reg}(T) + \frac{T\alpha\epsilon}{N+1} = \mathtt{Reg}(T) + \frac{T\alpha\epsilon}{\lfloor \alpha^{\frac{1}{3}} T^{\beta} \rfloor + 1} \le \mathtt{Reg}(T) + \frac{T\alpha\epsilon}{\alpha^{\frac{1}{3}} T^{\beta}} = O(\alpha^{\frac{2}{3}} T^{1-\beta}), \tag{66}$$

Furthermore, the regret and principal's profit satisfy

$$\begin{aligned} &\mathtt{Reg}(T) \le O(\sqrt{\alpha K T \ln T}) \\ &\mathtt{Pro}(T) + O(\sqrt{\alpha K T \ln T}) \ge 0. \end{aligned} \tag{67}$$

This completes the proof of this corollary. □

APPENDIX D
LEMMA 1 AND LEMMA 2

Lemma 1. *In Algorithm 1, for each slot t, we have*

$$\frac{\mathbb{E}\|\boldsymbol{\lambda}^t\|_1}{\eta} \le \Xi(\delta) + \frac{\sqrt{N}\Theta}{\delta\eta}. \tag{68}$$

where $\Xi(\delta)$ is given by

$$\Xi(\delta) \triangleq \frac{3\sqrt{N}\Theta^2}{\phi_{\min}-\delta}\ln\left(\frac{2\Theta}{\phi_{\min}-\delta}\right) + \frac{3\sqrt{N}\Theta}{2\delta}. \tag{69}$$

Proof of Lemma 1. For notation simplicity, we define $Z^t \triangleq \|\boldsymbol{\lambda}^t/\eta\|_2$, where $\|\boldsymbol{\lambda}^t\|_2 = \sqrt{\sum_{n=1}^{N}(\lambda_n^t)^2}$. Based on the property of Z^t in Lemma 2, we are able to use Lemma 4 in [Neely] to show that the following inequality holds

$$\mathbb{E}\left[\mathrm{e}^{\gamma Z^t}\right] \le 1 + \frac{2\mathrm{e}^{\gamma(\Theta+\theta)}}{\gamma(\phi_{\min}-\delta)}, \tag{70}$$

where γ is given by

$$\gamma = \frac{\phi_{\min}-\delta}{\Theta^2 + \Theta(\phi_{\min}-\delta)/3}. \tag{71}$$

Based on the above inequality, we obtain

$$1 + \frac{2\mathrm{e}^{\gamma(\Theta+\theta)}}{\gamma(\phi_{\min}-\delta)} \ge \mathbb{E}\left[\mathrm{e}^{\gamma Z^t}\right] \overset{(a)}{=} \mathbb{E}\left[\mathrm{e}^{\frac{\gamma\|\boldsymbol{\lambda}^t\|_2}{\eta}}\right] \overset{(b)}{\ge} \mathbb{E}\left[\mathrm{e}^{\frac{\gamma\|\boldsymbol{\lambda}^t\|_1}{\eta\sqrt{N}}}\right] \overset{(c)}{\ge} \mathrm{e}^{\frac{\gamma\mathbb{E}\|\boldsymbol{\lambda}^t\|_1}{\eta\sqrt{N}}}, \tag{72}$$

where the equality (a) follows the definition $Z^t \triangleq \|\boldsymbol{\lambda}^t/\eta\|_2$, the inequality (b) follows $\|\boldsymbol{\lambda}^t\|_1 \le \sqrt{N}\|\boldsymbol{\lambda}^t\|_2$, and the inequality (c) follows from Jensen's inequality and the convexity of e^x. Rearranging the above inequality leads to

$$\begin{aligned} \frac{\mathbb{E}\|\boldsymbol{\lambda}^t\|_1}{\eta} &\le \frac{\sqrt{N}}{\gamma}\ln\left(1+\frac{2\mathrm{e}^{\gamma(\Theta+\theta)}}{\gamma(\phi_{\min}-\delta)}\right) \overset{(a)}{\le} \frac{\sqrt{N}}{\gamma}\ln\left(1+\frac{8\Theta^2\mathrm{e}^{\gamma(\Theta+\theta)}}{3(\phi_{\min}-\delta)^2}\right) \overset{(b)}{\le} \frac{\sqrt{N}}{\gamma}\ln\left(\frac{11\Theta^2\mathrm{e}^{\gamma(\Theta+\theta)}}{3(\phi_{\min}-\delta)^2}\right) \\ &\le \frac{\sqrt{N}}{\gamma}\left[2\ln\left(\frac{2\Theta}{\phi_{\min}-\delta}\right) + \gamma(\Theta+\theta)\right] \overset{(c)}{\le} \frac{3\sqrt{N}\Theta^2}{\phi_{\min}-\delta}\ln\left(\frac{2\Theta}{\phi_{\min}-\delta}\right) + \sqrt{N}\Theta + \sqrt{N}\theta \\ &= \frac{3\sqrt{N}\Theta^2}{\phi_{\min}-\delta}\ln\left(\frac{2\Theta}{\phi_{\min}-\delta}\right) + \sqrt{N}\Theta + \frac{\sqrt{N}\Theta}{2\delta} + \frac{\sqrt{N}\Theta}{\delta\eta}, \\ &\overset{(d)}{\le} \frac{3\sqrt{N}\Theta^2}{\phi_{\min}-\delta}\ln\left(\frac{2\Theta}{\phi_{\min}-\delta}\right) + \frac{3\sqrt{N}\Theta}{2\delta} + \frac{\sqrt{N}\Theta}{\delta\eta}, \end{aligned} \tag{73}$$

where the inequality (a) follows $\gamma \geq \frac{3}{4}\frac{\phi_{\min}-\delta}{\Theta^2}$ according to (71) and $\phi_{\min}-\delta \leq \Theta$. The inequality (b) follows that $1 \leq \frac{\Theta^2 e^{\gamma(\Theta+\theta)}}{(\phi_{\min}-\delta)^2}$, the inequality (c) follows $\gamma \geq \frac{3}{4}\frac{\phi_{\min}-\delta}{\Theta^2}$. The inequality (d) follows $1 \leq \frac{1}{\delta}$ for any $\delta \in (0, \phi_{\min})$. This completes the proof of this lemma. $\square$

Lemma 2. *For any slot t, we have*

$$\left|Z^{t+1} - Z^t\right| \leq \Theta. \tag{74}$$

For any $\delta \in (0, \phi_{\min})$, we let $\theta \triangleq \frac{\Theta}{2\delta} + \frac{\Theta}{\delta\eta}$. Then we have

$$\mathbb{E}\left[Z^{t+1} - Z^t \middle| Z^t\right] \leq \begin{cases} \Theta, & \text{if } Z^t < \theta, \\ -\left(\phi_{\min} - \delta\right), & \text{if } Z^t \geq \theta. \end{cases} \tag{75}$$

Proof of Lemma 2. We prove (74) and (75) in two steps, respectively.

Step I: According to the definition of Z^t, we have $|Z^{t+1} - Z^t| = \left|\frac{\|\boldsymbol{\lambda}^{t+1}\|_2 - \|\boldsymbol{\lambda}^t\|_2}{\eta}\right|$. Next we derive a upper bound for $\|\boldsymbol{\lambda}^{t+1}\|_2 - \|\boldsymbol{\lambda}^t\|_2$ as follows:

$$\|\boldsymbol{\lambda}^{t+1}\|_2 - \|\boldsymbol{\lambda}^t\|_2 \overset{(a)}{\leq} \|\boldsymbol{\lambda}^{t+1} - \boldsymbol{\lambda}^t\|_2 \overset{(b)}{\leq} \|\boldsymbol{\lambda}^{t+1} - \boldsymbol{\lambda}^t\|_1 = \sum_{n=1}^{N}\left|\lambda_n^{t+1} - \lambda_n^t\right| \overset{(c)}{\leq} \eta\sum_{n=1}^{N}\left|f_n(\hat{\boldsymbol{x}}^t, \mathbf{1}_N) - \phi_n\right| \overset{(d)}{\leq} \eta\Theta, \tag{76}$$

where the inequality (a) follows from the triangle inequality, the inequality (b) follows from $\|\boldsymbol{x}\|_2 \leq \|\boldsymbol{x}\|_1$, the inequality (c) follows from the updating policy (19), and the inequality (d) follows from (46). Therefore, we have $|Z^{t+1} - Z^t| \leq \Theta$.

Step II: To prove (75), we only need to show that $\mathbb{E}[Z^{t+1} - Z^t | Z^t] \leq -(\phi_{\min} - \delta)$ if $Z^t \geq \theta$. Based on the first inequality in (47), we have

$$\begin{aligned}
\|\boldsymbol{\lambda}^{t+1}\|_2^2 - \|\boldsymbol{\lambda}^t\|_2^2 \leq& \eta^2\Theta + 2\eta\sum_{n=1}^{N}\lambda_n^t\Big[f_n(\hat{\boldsymbol{x}}^t, \mathbf{1}_N) - \phi_n\Big] \\
=& \eta^2\Theta + 2\eta S(\hat{\boldsymbol{x}}^t, \mathbf{1}_N; \hat{\boldsymbol{r}}^t, \hat{\boldsymbol{c}}^t) - 2\eta\left(S(\hat{\boldsymbol{x}}^t, \mathbf{1}_N; \hat{\boldsymbol{r}}^t, \hat{\boldsymbol{c}}^t) - \sum_{n=1}^{N}\lambda_n^t\Big[f_n(\hat{\boldsymbol{x}}^t, \mathbf{1}_N) - \phi_n\Big]\right) \\
\overset{(a)}{\leq}& \eta^2\Theta + 2\eta\Theta - 2\eta\left(S(\hat{\boldsymbol{x}}^t, \mathbf{1}_N; \hat{\boldsymbol{r}}^t, \hat{\boldsymbol{c}}^t) - \sum_{n=1}^{N}\lambda_n^t\Big[f_n(\hat{\boldsymbol{x}}^t, \mathbf{1}_N) - \phi_n\Big]\right) \\
\overset{(b)}{\leq}& \eta^2\Theta + 2\eta\Theta - 2\eta\left(S(\mathbf{0}, \mathbf{1}_N; \hat{\boldsymbol{r}}^t, \hat{\boldsymbol{c}}^t) - \sum_{n=1}^{N}\lambda_n^t\Big[f_n(\mathbf{0}, \mathbf{1}_N) - \phi_n\Big]\right) \\
=& \eta^2\Theta + 2\eta\Theta - 2\eta\sum_{n=1}^{N}\lambda_n^t\phi_n \overset{(c)}{\leq} \eta^2\Theta + 2\eta\Theta - 2\eta\phi_{\min}\|\boldsymbol{\lambda}^t\|_1,
\end{aligned} \tag{77}$$

where the inequality (a) follows $S(\hat{\boldsymbol{x}}^t, \mathbf{1}_N, \hat{\boldsymbol{r}}^t, \hat{\boldsymbol{c}}^t) \leq \min(K, N) \leq \Theta$, the inequality (b) follows the definition of $\hat{\boldsymbol{x}}^t$ in (18), and the inequality (c) follows from the definition of $\phi_{\min}$.

Accordingly, we obtain

$$\begin{aligned}
\mathbb{E}\Big[Z^{t+1} - Z^t \Big| Z^t\Big] =& \frac{\mathbb{E}\Big[\|\boldsymbol{\lambda}^{t+1}\|_2 - \|\boldsymbol{\lambda}^t\|_2 \Big| \boldsymbol{\lambda}^t\Big]}{\eta} \overset{(a)}{\leq} \frac{\mathbb{E}\Big[\|\boldsymbol{\lambda}^{t+1}\|_2^2 - \|\boldsymbol{\lambda}^t\|_2^2 \Big| \boldsymbol{\lambda}^t\Big]}{2\|\boldsymbol{\lambda}^t\|_2\eta} \overset{(b)}{\leq} \frac{\eta\Theta + 2\Theta - 2\phi_{\min}\|\boldsymbol{\lambda}^t\|_1}{2\|\boldsymbol{\lambda}^t\|_2} \\
\overset{(c)}{\leq}& \frac{\eta\Theta + 2\Theta}{2\|\boldsymbol{\lambda}^t\|_2} - \phi_{\min} \overset{(d)}{\leq} -(\phi_{\min} - \delta),
\end{aligned} \tag{78}$$

where the inequality (a) follows $y - x \leq \frac{y^2 - x^2}{2x}$ for any x, y. The inequality (b) follows the last inequality in (77). The inequality (c) follows that $\|\boldsymbol{x}\|_2 \leq \|\boldsymbol{x}\|_1$. The inequality (d) holds if $Z^t = \|\boldsymbol{\lambda}^t/\eta\|_2 \geq \theta \triangleq \frac{\Theta}{2\delta} + \frac{\Theta}{\delta\eta}$. This completes the proof of this lemma. $\square$

APPENDIX E
LEMMA 3 & LEMMA 4

Lemma 3. *Define W_1^t as follows:*

$$W_1^t = \sum_{n=1}^{N}\sum_{k=1}^{K}\left(\hat{r}_k^t - \bar{r}_k\right)\hat{x}_{n,k}^t. \tag{79}$$

We have

$$\sum_{t=1}^{T}\mathbb{E}\{W_1^t\} \leq \frac{5K}{2} + 2\sqrt{6KT\left[\Phi + \frac{\texttt{Vio}(T)}{T}\right]\ln T}. \tag{80}$$

Proof of Lemma 3. To prove this lemma, we need to introduce some new notation.

- We define a set of events $E_k^t \triangleq \{\hat{r}_k^t < \bar{r}_k\}$. Moreover, we let $\bar{E}_k^t$ denote the complement of the event E_k^t.
- We define event F_k^t as follows:

$$F_k^t \triangleq \left\{ \tilde{r}_k^{t-1} - \bar{r}_k \leq \sqrt{\frac{3\ln(t)}{2H_k^{t-1}}} \right\} \tag{81}$$

- We suppose that arm k is pulled for the a-th time in slot t_k^a. Accordingly, we have $\sum_{n=1}^N \hat{x}_{n,k}^{t_k^a} = 1$. Moreover, we have $H_k^{t_k^a} = a$ and $H_k^{t_k^a - 1} = a - 1$.

Based on the definition of events E_k^t and $\bar{E}_k^t$, we have

$$\begin{aligned} \sum_{t=1}^T \mathbb{E}\{W_1^t\} =& \mathbb{E}\left\{ \sum_{t=1}^T \sum_{k=1}^K \left(\hat{r}_k^t - \bar{r}_k\right) \left(\sum_{n=1}^N \hat{x}_{n,k}^t \right) \right\} = \mathbb{E}\left\{ \sum_{t=1}^T \sum_{k=1}^K \left(\hat{r}_k^t - \bar{r}_k\right) \left(\sum_{n=1}^N \hat{x}_{n,k}^t \right) \cdot \left(\mathbb{I}_{E_k^t} + \mathbb{I}_{\bar{E}_k^t} \right) \right\} \\ \overset{(a)}{\leq}& \mathbb{E}\left\{ \sum_{t=1}^T \sum_{k=1}^K \left(\hat{r}_k^t - \bar{r}_k\right) \sum_{n=1}^N \hat{x}_{n,k}^t \cdot \mathbb{I}_{\bar{E}_k^t} \right\}, \end{aligned} \tag{82}$$

where the inequality (a) follows that event E_k^t corresponds to $\hat{r}_k^t - \bar{r}_k < 0$.

Given arm $k \in \mathcal{K}$, we calculate the sum of $(\hat{r}_k^t - \bar{r}_k)\sum_{n=1}^N \hat{x}_{n,k}^t \cdot \mathbb{I}_{\bar{E}_k^t}$ for all $t \in \mathcal{T}$, and obtain

$$\begin{aligned} \sum_{t=1}^T \left(\hat{r}_k^t - \bar{r}_k\right) \left(\sum_{n=1}^N \hat{x}_{n,k}^t \right) \cdot \mathbb{I}_{\bar{E}_k^t} &\overset{(a)}{=} \sum_{a=1}^{H_k^T} \left(\hat{r}_k^{t_k^a} - \bar{r}_k \right) \cdot \mathbb{I}_{\bar{E}_k^{t_k^a}} \overset{(b)}{\leq} 1 + \sum_{a=2}^{H_k^T} \left(\hat{r}_k^{t_k^a} - \bar{r}_k \right) \cdot \mathbb{I}_{\bar{E}_k^{t_k^a}} \\ &= 1 + \sum_{a=2}^{H_k^T} \left(\hat{r}_k^{t_k^a} - \bar{r}_k \right) \cdot \mathbb{I}_{\bar{E}_k^{t_k^a}} \cdot \left(\mathbb{I}_{F_k^{t_k^a}} + \mathbb{I}_{\bar{F}_k^{t_k^a}} \right) \\ &\overset{(c)}{\leq} 1 + \sum_{a=2}^{H_k^T} \left(\hat{r}_k^{t_k^a} - \bar{r}_k \right) \cdot \mathbb{I}_{\{\bar{E}_k^{t_k^a} \cap F_k^{t_k^a}\}} + \sum_{a=2}^{H_k^T} \mathbb{I}_{\{\bar{F}_k^{t_k^a}\}}, \end{aligned} \tag{83}$$

where the equality (a) follows from $\sum_{n=1}^N \hat{x}_{n,k}^{t_k^a} = 1$, the inequality (b) follows from $\hat{r}_k^{t_k^a} - \bar{r}_k \leq 1$, and the inequality (c) follows from $(\hat{r}_k^{t_k^a} - \bar{r}_k)\mathbb{I}_{\bar{E}_k^{t_k^a}} \leq 1$.

Next we derive the upper bound for the second term and third term in (83).

- A upper bound of $\sum_{a=2}^{H_k^T} \left(\hat{r}_k^{t_k^a} - \bar{r}_k \right) \cdot \mathbb{I}_{\{\bar{E}_k^{t_k^a} \cap F_k^{t_k^a}\}}$.

  When event $F_k^{t_k^a}$ happens, we have

$$\tilde{r}_k^{t_k^a - 1} - \bar{r}_k \leq \sqrt{\frac{3\ln(t_k^a)}{2H_k^{t_k^a - 1}}}. \tag{84}$$

  Moreover, the UCB updating rule (20) indicates that

$$\hat{r}_k^{t_k^a} \leq \tilde{r}_k^{t_k^a - 1} + \sqrt{\frac{3\ln(t_k^a)}{2H_k^{t_k^a - 1}}}. \tag{85}$$

  Combing (84) and (85), we have

$$\hat{r}_k^{t_k^a} - \bar{r}_k \leq 2\sqrt{\frac{3\ln(t_k^a)}{2H_k^{t_k^a - 1}}}, \tag{86}$$

  which leads to

$$\left(\hat{r}_k^{t_k^a} - \bar{r}_k \right) \cdot \mathbb{I}_{\{\bar{E}_k^{t_k^a} \cap F_k^{t_k^a}\}} \leq 2\sqrt{\frac{3\ln(t_k^a)}{2H_k^{t_k^a - 1}}}. \tag{87}$$

  Accordingly, we obtain

$$\begin{aligned} \sum_{a=2}^{H_k^T} \left(\hat{r}_k^{t_k^a} - \bar{r}_k \right) \cdot \mathbb{I}_{\{\bar{E}_k^{t_k^a} \cap F_k^{t_k^a}\}} &\leq \sum_{a=2}^{H_k^T} 2\sqrt{\frac{3\ln(t_k^a)}{2H_k^{t_k^a - 1}}} \overset{(a)}{\leq} \sum_{a=2}^{H_k^T} 2\sqrt{\frac{3\ln T}{2H_k^{t_k^a - 1}}} \overset{(b)}{=} \sqrt{6\ln T} \sum_{a=2}^{H_k^T} \sqrt{\frac{1}{a-1}} \\ &\overset{(c)}{\leq} \sqrt{6\ln T} \int_1^{H_k^T} \sqrt{\frac{1}{z-1}} \mathrm{d}z = \sqrt{6\ln T} \cdot 2\sqrt{H_k^T - 1} \leq 2\sqrt{6H_k^T \ln T}, \end{aligned} \tag{88}$$

where the inequality (a) follows from $t_k^a \le T$, the equality (b) is due to $H_k^{t_k^a-1} = a-1$, the inequality (c) follows the basic integral calculation.

- An upper bound of $\sum_{a=2}^{H_k^T} \mathbb{I}_{\{\bar{F}_k^{t_k^a}\}}$. Specifically, we have

$$\sum_{a=2}^{H_k^T} \mathbb{E}\{\mathbb{I}_{\{\bar{F}_k^{t_k^a}\}}\} \overset{(a)}{=} \sum_{a=2}^{H_k^T} \Pr\left(\tilde{r}_k^{t_k^a-1} - \bar{r}_k > \sqrt{\frac{3\ln(t_k^a)}{2H_k^{t_k^a-1}}}\right) \overset{(b)}{\le} \sum_{a=2}^{H_k^T} \frac{1}{(t_k^a)^3} \le \sum_{t=1}^{\infty} \frac{1}{t^3} \le 1 + \int_1^{\infty} \frac{1}{z^3} \mathrm{d}z = \frac{3}{2}, \tag{89}$$

where (a) follows the definition of event $\bar{F}_k^t$ and the inequality (b) follows Chernoff-Hoeffding inequality.

Based on the above two bounds, we express $\sum_{t=1}^{T} \mathbb{E}\{W_1^t\}$ as follows:

$$\sum_{t=1}^{T} \mathbb{E}\{W_1^t\} \le \mathbb{E}\left\{\sum_{k=1}^{K}\left(1 + 2\sqrt{6\ln T}\sqrt{H_k^T} + \frac{3}{2}\right)\right\} = \mathbb{E}\left\{\frac{5K}{2} + 2\sqrt{6\ln T}\sum_{k=1}^{K}\sqrt{H_k^T}\right\} \tag{90}$$

Based on Jensen inequality, we have

$$\mathbb{E}\left[\sum_{k=1}^{K}\sqrt{H_k^T}\right] \overset{(a)}{\le} \mathbb{E}\left[K\sqrt{\frac{\sum_{k=1}^{K} H_k^T}{K}}\right] = \sqrt{K}\mathbb{E}\sqrt{\sum_{k=1}^{K} H_k^T} \overset{(b)}{\le} \sqrt{K}\sqrt{\mathbb{E}\sum_{k=1}^{K} H_k^T} = \sqrt{K}\sqrt{\mathbb{E}\sum_{t=1}^{T}\sum_{n=1}^{N} f_n(\hat{\boldsymbol{x}}^t, \mathbf{1}_N)}, \tag{91}$$

where the inequalities (a) and (b) follow from Jensen inequality. Next we derive an upper bound for $\sum_{t=1}^{T}\sum_{n=1}^{N} \mathbb{E}[f_n(\hat{\boldsymbol{x}}^t, \mathbf{1}_N)]$. According to the definition in (8), we have

$$\mathtt{Vio}(T) = \sum_{n=1}^{N}\left(\sum_{t=1}^{T}\mathbb{E}\Big[f_n(\boldsymbol{x}^t, \boldsymbol{a}^t) - \phi_n\Big]\right)^+ \ge \sum_{n=1}^{N}\sum_{t=1}^{T}\mathbb{E}\Big[f_n(\boldsymbol{x}^t, \boldsymbol{a}^t) - \phi_n\Big] = \sum_{n=1}^{N}\sum_{t=1}^{T}\mathbb{E}\Big[f_n(\boldsymbol{x}^t, \boldsymbol{a}^t)\Big] - T\Phi, \tag{92}$$

where $\Phi = \sum_{n=1}^{N} \phi_n$. Accordingly, we obtain

$$\mathbb{E}\left[\sum_{k=1}^{K} H_k^T\right] = \sum_{t=1}^{T}\sum_{n=1}^{N}\mathbb{E}\Big[f_n(\hat{\boldsymbol{x}}^t, \mathbf{1}_N)\Big] \le \mathtt{Vio}(T) + T\Phi, \tag{93}$$

Therefore, we have

$$\sum_{t=1}^{T}\mathbb{E}\{W_1^t\} \le \frac{5K}{2} + 2\sqrt{6K\left[\mathtt{Vio}(T) + T\Phi\right]\ln T}, \tag{94}$$

which completes the proof of this lemma. □

Lemma 4. *Define W_2^t as follows:*

$$W_2^t = \sum_{n=1}^{N}\sum_{k=1}^{K}\left(\bar{r}_k - \hat{r}_k^t\right)\tilde{x}_{n,k}^t. \tag{95}$$

Moreover, we have

$$\sum_{t=1}^{T}\mathbb{E}\{W_2^t\} \le \frac{3K}{2}. \tag{96}$$

Proof of Lemma 4. We start with deriving an upper bound for W_2^t as follows

$$\begin{aligned}\mathbb{E}\left[W_2^t\right] =& \mathbb{E}\left[\sum_{k=1}^{K}\left(\bar{r}_k - \hat{r}_k^t\right)\sum_{n=1}^{N}\tilde{x}_{n,k}^t\right] \overset{(a)}{=} \mathbb{E}\left[\sum_{k=1}^{K}\left(\bar{r}_k - \hat{r}_k^t\right)\sum_{n=1}^{N}\tilde{x}_{n,k}^t\left(\mathbb{I}_{\{E_k^t\}} + \mathbb{I}_{\{\bar{E}_k^t\}}\right)\right] \\ \overset{(b)}{\le}& \mathbb{E}\left[\sum_{k=1}^{K}\left(\bar{r}_k - \hat{r}_k^t\right)\sum_{n=1}^{N}\tilde{x}_{n,k}^t \cdot \mathbb{I}_{\{E_k^t\}}\right],\end{aligned} \tag{97}$$

where the equality (a) follows from the definition of events E_k^t and $\bar{E}_k^t$, the inequality (b) follows that we have $\bar{r}_k - \hat{r}_k^t \le 0$ when event $\bar{E}_k^t$ happens. Next we consider two cases.

- Case of $t \le t_k^1$. Event E_k^t happens with probability zero. Hence we have

$$\mathbb{E}\left[\left(\bar{r}_k - \hat{r}_k^t\right)\sum_{n=1}^{N}\tilde{x}_{n,k}^t \cdot \mathbb{I}_{\{E_k^t\}}\right] = 0. \tag{98}$$

- Case of $t > t_k^1$. When event E_k^t happens, we have $\hat{r}_k^t < \bar{r}_k \leq 1$. Given the updating rule (20), we have

$$\hat{r}_k^t = \tilde{r}_k^{t-1} + \sqrt{\frac{3\ln(t)}{2H_k^{t-1}}}. \tag{99}$$

In this case, we have

$$\mathbb{E}\left[\left(\bar{r}_k - \hat{r}_k^t\right)\sum_{n=1}^{N}\tilde{x}_{n,k}^t \cdot \mathbb{I}_{\{E_k^t\}}\right] \overset{(a)}{\leq} \mathbb{E}\Big\{\mathbb{I}_{\{E_k^t\}}\Big\} \overset{(b)}{=} \Pr\left(\hat{r}_k^t \leq \bar{r}_k\right) \overset{(c)}{=} \Pr\left(\tilde{r}_k^{t-1} - \bar{r}_k \leq -\sqrt{\frac{3\ln(t)}{2H_k^{t-1}}}\right) \overset{(d)}{\leq} \frac{1}{t^3}, \tag{100}$$

where the inequality (a) is due to $(\bar{r}_k - \hat{r}_k^t)\sum_{n=1}^{N}\tilde{x}_{n,k}^t \leq \bar{r}_k - \hat{r}_k^t \leq 1$, the equality (b) follows the definition of event E_k^t, the equality (c) is due to (99), and the inequality (d) follows Chernoff-Hoeffding inequality.
Therefore, given arm $k \in \mathcal{K}$, we obtain

$$\sum_{t=1}^{T}\mathbb{E}\left[\left(\bar{r}_k - \hat{r}_k^t\right)\sum_{n=1}^{N}\tilde{x}_{n,k}^t\mathbb{I}_{\{E_k^t\}}\right] \leq \sum_{t=1}^{\infty}\frac{1}{t^3} \leq \frac{3}{2}. \tag{101}$$

The above two cases shows that

$$\sum_{t=1}^{T}\mathbb{E}\{W_2^t\} \leq \frac{3K}{2}, \tag{102}$$

which completes the proof of this lemma. $\square$