APPENDIX A

We separately analyze the regret bound of the CD-MMAB algorithm in the exploration and exploitation phase and the calibration phase. The details are described as follows.

Case I) Regret in the exploration and exploitation phase $R_1(K)$*:*

Let $b_{s,n}^*$ denote AP n's optimal reservation decisions. A reservation decision $b_{s,n}^k$ can only be chosen if $\widehat{u}_{b_{s,n}^k}^{s,n}(k) \leq \widehat{u}_{b_{s,n}^*}^{s,n}(k)$. Applying Azuma-Hoeffding inequality, we can have the following inequalities:

$$\Pr\left[\left|\mathbb{E}\left[u_{b_{s,n}^k}^{s,n}(k)\right] - \overline{u}_{b_{s,n}^k}^{s,n}\right| \geq r_{b_{s,n}^k}^{s,n}(k)\right] \leq 2k^{-2\varpi}, \quad (1)$$

$$\overline{u}_{b_{s,n}^k}^{s,n} - r_{b_{s,n}^k}^{s,n}(k) = \widehat{u}_{b_{s,n}^k}^{s,n}(k) \leq \widehat{u}_{b_{s,n}^*}^{s,n}(k) = \mathbb{E}\left[u_{b_{s,n}^*}^{s,n}(k)\right], \quad (2)$$

$$\mathbb{E}\left[u_{b_{s,n}^k}^{s,n}(k)\right] \leq \overline{u}_{b_{s,n}^k}^{s,n} + r_{b_{s,n}^k}^{s,n}(k). \quad (3)$$

Thereupon, for AP n, the gap resulting from the suboptimal decision is bounded as

$$\mathbb{E}\left[u_{b_{s,n}^k}^{s,n}(k) - u_{b_{s,n}^*}^{s,n}(k)\right] \leq 2r_{b_{s,n}^k}^{s,n}(k). \quad (4)$$

Then, we have

$$\begin{aligned} R_1(K) &= \mathbb{E}\left[\sum\nolimits_{k=1}^{K}\left(U_{s,k} - U_s^*\right)\right] \leq 2\sum_{k=1}^{K}\sum_{n=1}^{N} r_{b_{s,n}^k}^{s,n}(k) \\ &= 2\sum_{k=1}^{K}\sum_{n=1}^{N}\sqrt{\varpi \ln(k)\frac{R^2}{f_{x,y}^{s,n}}}. \end{aligned} \quad (5)$$

By applying Cauchy-Schwartz inequality, we have

$$R_1(K) \leq 2\sqrt{\sum_{k=1}^{K}\sum_{n=1}^{N}\varpi\ln(k)} \times \sqrt{\sum_{k=1}^{K}\sum_{n=1}^{N}\frac{R^2}{f_{x,y}^{s,n}}}. \quad (6)$$

For the first term on the right-hand side, we have

$$\begin{aligned} \sum_{k=1}^{K}\sum_{n=1}^{N}\varpi\ln(k) &\leq N\int_1^K \varpi\ln(k)\,\mathrm{d}k + \mathcal{O}(1) \\ &\leq \mathcal{O}(NK\ln K) \end{aligned} \quad (7)$$

Let $X_{\max} = \max\{X_{s,n} | \forall s \in \mathcal{S}, n \in \mathcal{N}\}$, $Y_{\max} = \max\{Y_{s,n} | \forall s \in \mathcal{S}, n \in \mathcal{N}\}$, and $K_1 = (X_{\max}+1) \times (Y_{\max}+1)$, the second term on the right-hand side satisfies

$$\begin{aligned} \sum_{k=1}^{K}\sum_{n=1}^{N}\frac{R^2}{f_{x,y}^{s,n}} &= \sum_{n=1}^{N}\sum_{k=1}^{K}\frac{R^2}{f_{x,y}^{s,n}} \\ &= \sum_{n=1}^{N}\sum_{xy=0}^{K_1}\sum_{f_{x,y}^{s,n}=1}^{K}\frac{R^2}{f_{x,y}^{s,n}} \\ &\leq R^2\sum_{n=1}^{N}\sum_{xy=0}^{K_1}\ln\left(f_{x,y}^{s,n}\right). \end{aligned} \quad (8)$$

Since $\sum_{xy=0}^{K_1} f_{x,y}^{s,n} = K$and $\ln(x)$ is a convex function, applying Jensen's inequality on the upper equation has

$$\begin{aligned} \sum_{k=1}^{K}\sum_{n=1}^{N}\frac{R^2}{f_{x,y}^{s,n}} &\leq R^2\sum_{n=1}^{N}K_1\ln(K/K_1) \\ &= R^2NK_1\ln(K/K_1) \\ &= \mathcal{O}(NK_1\ln K). \end{aligned} \quad (9)$$

As a result, we have

$$R_1(K) \leq \mathcal{O}\left(N\sqrt{KK_1}\ln K\right). \quad (10)$$

Case II) Regret in the calibration phase $R_2(K)$*:*

Plugging Eq. (13) in the main text into $R_2(K)$, we have

$$\begin{aligned} R_2(K) &= \sum_{k=1}^{K}\left(U_s^* - U_{s,k}^*\right) \\ &= \ln\left\{\sum_{k=1}^{K}\sum_{n=1}^{N}\left[l_{n,a}^s(k) - l_{n,a^*}^s(k)\right]\right\}. \end{aligned} \quad (11)$$

For AP n, $\forall n \in \mathcal{N}$, its partial regret in the calibration phase can be written as

$$R_2^n(K) = \sum_{k=1}^{K}\left[l_{n,a}^s(k) - l_{n,a^*}^s(k)\right]. \quad (12)$$

Based on the Cramer-Chernoff method, we can rewrite $R_2^n(K)$ as

$$\begin{aligned} R_2^n(K) &= \sum_{k=1}^{K}\left[l_{n,a}^s(k) - l_{n,a^*}^s(k)\right] \\ &= \sum_{k=1}^{K}\left[l_{n,a'}^s(k) - \sum_{a\in\mathcal{A}_{s,n}} p_{n,a}^s(k)\,\widehat{l}_{n,a}^s(k)\right] \\ &+ \sum_{k=1}^{K}\left[\sum_{a\in\mathcal{A}_{s,n}} p_{n,a}^s(k)\,\widehat{l}_{n,a}^s(k) - \widehat{l}_{n,a^*}^s(k)\right] \\ &+ \sum_{k=1}^{K}\left[\widehat{l}_{n,a^*}^s(k) - l_{n,a^*}^s(k)\right] \\ &= \widetilde{R}_2^n(K) + \overline{R}_2^n(K) + \widehat{R}_2^n(K). \end{aligned} \quad (13)$$

By applying Markov's inequality, for any $\sigma > 0$, we have

$$\Pr\left\{\sum_{k\in\mathcal{K}}\sum_{a\in\mathcal{A}_{s,n}}\beta\left[\widehat{l}_{n,a}^s(k) - l_{n,a}^s(k)\right] > \sigma\right\} \leq e^{-\sigma} = \mu. \quad (14)$$

According to [20], with probability at least $1-\mu$, $\widehat{R}_2^n(K)$ satisfies

$$\begin{aligned} \widehat{R}_2^n(K) &= \sum_{k\in\mathcal{K}}\sum_{a\in\mathcal{A}_{s,n}}\left[\widehat{l}_{n,a}^s(k) - l_{n,a}^s(k)\right] \\ &\leq \ln\left(\left|\mathcal{A}_{s,n}\right|/\mu\right)/\beta\ . \end{aligned} \quad (15)$$

For $\widetilde{R}_n^2(K)$, we have

$$\begin{aligned} \widetilde{R}_n^2(K) &= \sum_{k\in\mathcal{K}}\left[l_{n,a'}^s(k) - \sum_{a\in\mathcal{A}_{s,n}} p_{n,a}^s(k)\,\widehat{l}_{n,a}^s(k)\right] \\ &= \sum_{k\in\mathcal{K}} l_{n,a'}^s(k) - \sum_{k\in\mathcal{K}}\sum_{a\in\mathcal{A}_{s,n}} l_{n,a}^s(k)\,1\{a=a'\} \\ &+ \sum_{k\in\mathcal{K}}\sum_{a\in\mathcal{A}_{s,n}} 1\{a=a'\}\frac{\beta l_{n,a}^s(k)}{p_{n,a}^s(k)+\beta} \end{aligned}$$

$$= \sum_{k\in\mathcal{K}} \sum_{a\in\mathcal{A}_{s,n}} \beta \widehat{l}_{n,a}^{s}(k). \tag{16}$$

According to [21], if $\psi \le 2\beta$, $\overline{R}_n^2(K)$ satisfies

$$\begin{aligned}\overline{R}_n^2(K) &= \sum_{k=1}^{K}\left[\sum_{a\in\mathcal{A}_{s,n}} p_{n,a}^{s}(k)\,\widehat{l}_{n,a}^{s}(k) - \widehat{l}_{n,a^*}^{s}(k)\right]\\ &\le \ln\left(|\mathcal{A}_{s,n}|\right)/\psi + \sum_{k\in\mathcal{K}}\sum_{a\in\mathcal{A}_{s,n}} \psi \widehat{l}_{n,a}^{s}(k)/2. \end{aligned}\tag{17}$$

Thereupon, when $\psi \le 2\beta$, $R_2^n(K)$ is upper-bounded by

$$\begin{aligned}R_2^n(K) \le{}& \ln\left(|\mathcal{A}_{s,n}|\right)/\psi \ + \ln\left(|\mathcal{A}_{s,n}|/\mu\right)/\beta\\ &+ \sum_{k\in\mathcal{K}}\sum_{a\in\mathcal{A}_{s,n}}\left(\beta + \frac{\psi}{2}\right)\widehat{l}_{n,a}^{s}(k).\end{aligned}\tag{18}$$

Similar to the proof of $\overline{R}_n^2(K)$, we have

$$\sum_{k\in\mathcal{K}}\sum_{a\in\mathcal{A}_{s,n}} \beta \widehat{l}_{n,a}^{s}(k) \le \sum_{k\in\mathcal{K}}\sum_{a\in\mathcal{A}_{s,n}} \beta l_{n,a}^{s}(k) + \ln\left(|\mathcal{A}_{s,n}|/\mu\right). \tag{19}$$

Thereupon,

$$\begin{aligned}R_2^n(K) \le{}& \ln\left(|\mathcal{A}_{s,n}|\right)/\psi \ + \left(\frac{1}{\beta}+1\right)\ln\left(|\mathcal{A}_{s,n}|/\mu\right)\\ &+ \sum_{k\in\mathcal{K}}\sum_{a\in\mathcal{A}_{s,n}}\left(\beta+\frac{\psi}{2}\right) l_{n,a}^{s}(k)\\ \le{}& \ln\left(|\mathcal{A}_{s,n}|\right)/\psi \ + \left(\frac{1}{\beta}+1\right)\ln\left(|\mathcal{A}_{s,n}|/\mu\right)\\ &+ \sum_{k\in\mathcal{K}}\left(\beta+\frac{\psi}{2}\right)|\mathcal{A}_{s,n}|.\end{aligned}\tag{20}$$

Let $A = \max\left\{|\mathcal{A}_{s,n}|\,|\forall s\in\mathcal{S}, n\in\mathcal{N}\right\}$. Since the playing of all agents is uncoupled,

$$R_2(K) \le \mathcal{O}\left(\ln(NKA)\right). \tag{21}$$

Obviously, $\ln(NKA) < N\sqrt{KK_1}\ln K$. Thus, based on the analysis above, we have

$$\begin{aligned}R(K) &= R_1(K) + R_2(K)\\ &\le \mathcal{O}\left(N\sqrt{KK_1}\ln K\right).\end{aligned}\tag{22}$$

The proof is complete.

Appendix B

The objection function in problem $\mathcal{P}_2$ can be rewritten as

$$\begin{aligned}&f\left(z_{s,m,n}^{k,t}\right)\\ &= \sum_{n\in\mathcal{N}}\left(\frac{1}{X_{s,n}^{k}R_n/\varsigma_s - \sum\limits_{m\in\mathcal{M}} c_{n,m} z_{s,n,m}^{k,t}\zeta_s\rho_m^{k,t}L}\right)\\ &+ \sum_{n\in\mathcal{N}}\left(\frac{1}{Y_{s,n}^{k}F/\xi_s - \sum\limits_{m\in\mathcal{M}} c_{n,m} z_{s,n,m}^{k,t}\zeta_s\rho_m^{k,t}L}\right).\end{aligned}\tag{23}$$

Its second-order derivative is given by

$$\begin{aligned}&\frac{\partial^2 f\left(z_{s,m,n}^{k,t}\right)}{\partial^2 z_{s,m,n}^{k,t}}\\ &= \sum_{n\in\mathcal{N}}\left[\frac{2\left(c_{n,m}\zeta_s\rho_m^{k,t}L\right)^2}{\left(X_{s,n}^{k}R_n/\varsigma_s - \sum\limits_{m\in\mathcal{M}} c_{n,m} z_{s,n,m}^{k,t}\zeta_s\rho_m^{k,t}L\right)^3}\right]\\ &+ \sum_{n\in\mathcal{N}}\left[\frac{2\left(c_{n,m}\zeta_s\rho_m^{k,t}L\right)^2}{\left(Y_{s,n}^{k}F/\xi_s - \sum\limits_{m\in\mathcal{M}} c_{n,m} z_{s,n,m}^{k,t}\zeta_s\rho_m^{k,t}L\right)^3}\right].\end{aligned}\tag{24}$$

When constraints C5 and C6 in problem $\mathcal{P}_0$ are satisfied, we have $\frac{\partial^2 f\left(z_{s,m,n}^{k,t}\right)}{\partial^2 z_{s,m,n}^{k,t}} \ge 0$. Since the other constraints of problem $\mathcal{P}_2$ are linear, Theorem 2 is proved.