Correction

**Mean-variance of Observations:** Equation (9) in Lemma 2 should be replaced with

\[
R_{\pi}^{(1)}(T) = \sum_{i=1}^{K} E[\tau_i(T)] \Delta_i - \mathbb{E}[\sum_{i=1}^{K} \tau_i(\pi_i - \mu_i)^2] + \frac{1}{T^2} \mathbb{E}[\sum_{i=1}^{K} \tau_i(T)(\sum_{j=1}^{K} \tau_{ij}^2)].
\]  

(9)

Consequently, equation (13) in Theorem 2 should be replaced with

\[
R_{MV-UCB}^{(1)}(T) \leq \sum_{i \neq *} \left( \frac{4b^2 \log T}{\Delta_i^2} + 5 \right) (\Delta_i + 2\Gamma_{\text{max}}^2 + \frac{1}{4} \Delta_i^2 + 8). \quad (13)
\]

**Mean-variance of Total Reward:** Equation (16) in Lemma 3 should be replaced with

\[
R_{\pi}^{(2)}(T) = \sum_{i=1}^{K} E[\tau_i(T)] \Delta_i + \mathbb{E}[\left( \sum_{i \neq *} (\tau_i - E[\tau_i]) \Gamma_{i,*} \right)^2]
\]

\[
+ 2 \mathbb{E}[\left( \sum_{i=1}^{K} \sum_{s=1}^{X_i(s)} (X_i(s) - \mu_i) \right)(\sum_{i \neq *} (\tau_i - E[\tau_i]) \Gamma_{i,*})]. \quad (16)
\]

Due to this, the lower bound results stated in Subsection 4.1. may no longer hold. Theorem 4 still holds with the MV-UCB algorithm replaced with the following modified version.

**Initialization:** \(\Delta_0 = 1, S_0 = \{1, 2, \ldots, K\}\).

Repeat until the end of time horizon:

- At step \(m = 0, 1, 2, \ldots\)

1) Play each arm in \(S_m\) until the total number of times it has been played is \(n_m = \frac{6b^2 \log T}{\Delta_m^2}\).

2) Remove from \(S_m\) all arms satisfying the following condition to form \(S_{m+1}\).

\[
\bar{\xi}_i(n_m) - b \sqrt{\frac{\log T}{n_m}} > \min_{j \in S_m} \{\bar{\xi}_j(n_m) + b \sqrt{\frac{\log T}{n_m}}\}.
\]

3) Set \(\Delta_{m+1} = \frac{\Delta_m}{2}\).