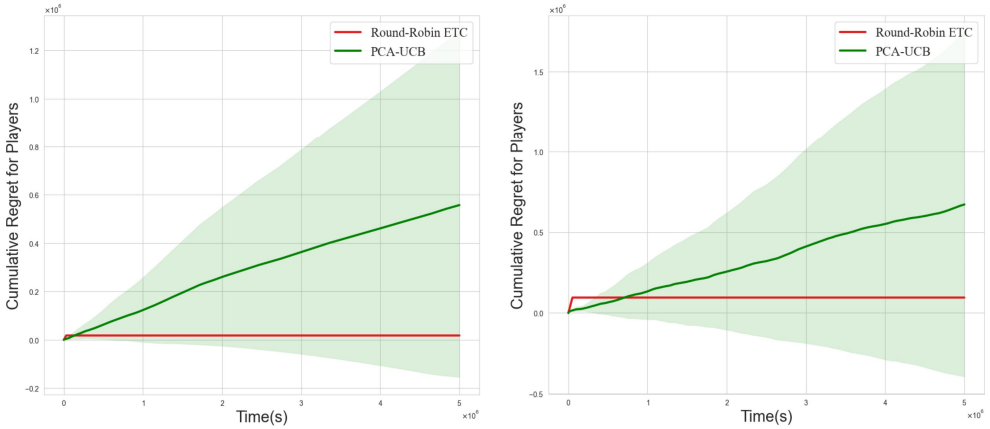
We estimate the average player-optimal stable regret and standard deviations of regret over 30 independent runs in each cases. The minimal gaps are all set to be 0.2.

Baseline Parameters. The authors in [Pokharel and Das, 2023] don't reveal their detailed strategies for the arm side. Thus, in the simulation, we assume that the arms choose the candidates with the highest UCB. As for the value of the random delay parameter λ, we set $\lambda = 0.9$ as in the simulations in [Pokharel and Das, 2023].



(a) Case 1 (b) Case 2

In case 1 (Figure 1(b)), there are 3 players and 3 arms. The preferences are as follows:

$$p_1 : a_1 \succ a_2 \succ a_3 \quad a_1 : p_1 \succ^a p_3 \succ^a p_2$$

$$p_2 : a_3 \succ a_2 \succ a_1 \quad a_2 : p_2 \succ^a p_3 \succ^a p_1$$

$$p_3 : a_2 \succ a_3 \succ a_1 \quad a_3 : p_3 \succ^a p_1 \succ^a p_2$$

In case 2 (Figure 1(a)), there are 5 players and 5 arms. The preferences are as follows:

$$p_1 : a_4 \succ a_1 \succ a_5 \succ a_2 \succ a_3 \quad a_1 : p_3 \succ^a p_1 \succ^a p_5 \succ^a p_2 \succ^a p_4$$

$$p_2 : a_5 \succ a_1 \succ a_2 \succ a_4 \succ a_3 \quad a_2 : p_5 \succ^a p_2 \succ^a p_1 \succ^a p_4 \succ^a p_3$$

$$p_3 : a_2 \succ a_5 \succ a_3 \succ a_1 \succ a_4 \quad a_3 : p_3 \succ^a p_1 \succ^a p_2 \succ^a p_5 \succ^a p_4$$

$$p_4 : a_5 \succ a_2 \succ a_1 \succ a_3 \succ a_4 \quad a_4 : p_1 \succ^a p_2 \succ^a p_5 \succ^a p_4 \succ^a p_3$$

$$p_5 : a_3 \succ a_5 \succ a_2 \succ a_4 \succ a_1 \quad a_5 : p_1 \succ^a p_4 \succ^a p_5 \succ^a p_3 \succ^a p_2$$