Algorithm 1 ϵ-greedy Algorithm

Initialize $\widehat{\theta}(j) \leftarrow 0, \text{count}(j) \leftarrow 0, j \in \{1,2,3\}$

1: **for** $t = 1, 2 \ldots, N$ **do**

2:

$$I(t) \leftarrow \begin{cases} \underset{j \in \{1,2,3\}}{\arg\max}\ \widehat{\theta}(j) & w.p.\ 1-\epsilon \\ \text{randomly chosen from} \{1,2,3\} & w.p.\ \epsilon \end{cases}$$

3: $\quad \text{count}(I(t)) \leftarrow \text{count}(I(t)) + 1$

4: $\quad \widehat{\theta}(I(t)) \leftarrow \widehat{\theta}(I(t)) + \frac{1}{\text{count}(I(t))} \left[r_{I(t)} - \widehat{\theta}(I(t)) \right]$

5: **end for**