The best of both worlds: stochastic and adversarial bandits

Sebastien Bubeck and Aleksandrs Slivkins

$A \leftarrow \{1,\dots,K\}$, $p_i \leftarrow 1/K$, $\tau_i \leftarrow n$ ▷ initialization
for $t = 1,\dots,n$ **do** ▷ Main loop
    Play I_t at random from p ▷ Selection of the arm to play
    for $i = 1,\dots,K$ **do** ▷ Test of four properties for arm i
        if ▷ Test if arm i should be deactivated

$$i \in A, \text{ and } \max_{j\in A} \widetilde{H}_{j,t} - \widetilde{H}_{i,t} > 6c\sqrt{\frac{K\log(n)}{t}} \tag{1}$$

            then $A \leftarrow A \setminus \{i\}$, $\tau_i \leftarrow t$ and $q_i \leftarrow p_i$ ▷ Deactivation of arm i
        end if ▷ q_i denotes the probability of arm i at the moment when it was de-activated
        if one of the three following properties is satisfied
            then Stop and start Exp3 (also $\tau_0 \leftarrow t, \tau_i \leftarrow t$ for $i \in A$)

$$\left|\widetilde{H}_{i,t} - \widehat{H}_{i,t}\right| > c\sqrt{\frac{\log(n)}{T_i(t)}} + c\sqrt{\left(K\frac{\tau_i \wedge t}{t^2} + \frac{(t-\tau_i)\vee 0}{q_i\tau_i t}\right)\log(n)} \tag{2}$$

$$i \notin A, \text{ and } \max_{j\in A} \widetilde{H}_{j,t} - \widetilde{H}_{i,t} > 10c\sqrt{\frac{K\log(n)}{\tau_i}} \tag{3}$$

$$i \notin A, \text{ and } \max_{j\in A} \widetilde{H}_{j,t} - \widetilde{H}_{i,t} \le 2c\sqrt{\frac{K\log(n)}{\tau_i}} \tag{4}$$

        end if
    end for ▷ End of testing
    for $i = 1,\dots,K$ **do** ▷ Update of the probability of selecting arm i

$$p_i \leftarrow \frac{q_i\ \tau_i}{t+1}\mathbb{1}_{i\notin A} + \frac{1}{|A|}\left(1 - \sum_{j\notin A}\frac{q_j\ \tau_j}{t+1}\right)\mathbb{1}_{i\in A}. \tag{5}$$

    end for
end for

Claim: with probability at least $1-1/n$, for all $t \in [n], i \in [K]$,

$$\left|\widetilde{H}_{i,t} - H_{i,t}\right| \le^{\text{adv}} c\sqrt{\left(K\frac{\tau_i \wedge t}{t^2} + \frac{(t-\tau_i)\vee 0}{q_i\tau_i t}\right)\log(n)}, \qquad |\widehat{H}_{i,t} - \mu_i| \le^{\text{stocha}} c\sqrt{\frac{\log(n)}{T_i(t)}}$$

$$\left|\widetilde{H}_{i,t} - \mu_i\right| \le^{\text{stocha}} c\sqrt{\left(K\frac{\tau_i \wedge t}{t^2} + \frac{(t-\tau_i)\vee 0}{q_i\tau_i t}\right)\log(n)}, \qquad T_i(t) \le c q_i \tau_i \log(n)$$