Appendix: Proof of Theorem 3

Proof

Only a sketch of this proof showing the differences with the corresponding steps in a similar derivation for UCB3 are given. The probability that the arm $j$ is chosen at time $t$ is given by:

$$P[I_n = j] = c_n^j + (1 - \sum_{j=1}^{K} c_n^j)P[\tilde{X}_{j,T_j(n-1)}^h \geq \tilde{X}_{s,T_s(n-1)}^h]$$

Moreover,

$$P[\tilde{X}_{j,T_j(n)} \geq \tilde{X}_{j,T_j(n)}^*] \leq P[\tilde{X}_{j,T_j(n)}^h \geq \mu_j + \frac{\Delta_j}{2}] + P[\tilde{X}_{j,T_j(n)}^h \leq \mu_* - \frac{\Delta_j}{2}]. \quad (1)$$

Denoting $\frac{1}{2} \sum_{t=1}^{n} c_t^j$ by $x_0^j$, it can be shown that the first term above is upper bounded by:

$$P[\tilde{X}_{j,T_j(n)}^h \geq \mu_j + \frac{\Delta_j}{2}] \leq \left( x_0^j P[T_j^R(n) \leq x_0^j] + \frac{2}{\Delta_j^2} e^{-\Delta_j^2 |x_0^j|/2} \right) e^{-H_j \Delta_j^2/2}, \quad (2)$$

where, we get the extra factor $e^{-H_j \Delta_j^2/2}$ from an application of Hoeffding’s inequality incorporating the historic data and $T_j^R(n)$ is the number of times arm $j$ is selected at random in the first $n$ draws. Since $d \leq \Delta_j$ for all $j$ we can replace $e^{-H_j \Delta_j^2/2}$ with $e^{-H_j d^2/2}$.

It can further be shown that:

$$P[T_j^R(n) \leq x_0^j] \leq e^{-x_0^j/5}, \quad (3)$$

using Bernstein’s inequality.

Finally, we can lower bound, $x_0^j$ as follows:

$$x_0^j = \frac{1}{2} \sum_{t=1}^{n} c_t^j$$

$$= \frac{1}{2} \sum_{t=1}^{n} \frac{1}{K} + \frac{K}{2} \sum_{t=\frac{d}{2}+1}^{n} \frac{c}{d^2 (e^{H_j d^2/2} - 1) + t}$$

$$\geq \frac{c}{2d^2} \log \left( \frac{cK (e^{H_j d^2/2} - 1) + ne}{e^{H_j d^2/2} / e} \right). \quad (4)$$

Using (1), (2), (3) and (4), it can be shown that:

$$P[I_n = j] \leq \frac{c}{d^2 (e^{H_j d^2/2} - 1) + n})$$

$$+ \left( \frac{c}{2d^2} P_{j}^{\text{mean}} \log \left( \frac{1}{P_j} \right) + \frac{2}{d^2} P_j^2 \right) e^{-H_j d^2/2}$$

$$+ \left( \frac{c}{2d^2} P_*^{\text{mean}} \log \left( \frac{1}{P_*} \right) + \frac{2}{d^2} P_*^2 \right) e^{-H_* d^2/2} \quad (5)$$
where

\[ P_j := \frac{cK e^{H_d d^2 / c}}{\frac{cK}{d^2} (e^{H_d d^2 / c} - 1) + n - 1}. \]

Thus, for \( c \geq 10 \), the last four terms in (5) are \( o(\frac{1}{n}) \) since \( d < 1 \).