A Concentration Inequalities

We provide here required results of anytime bounds that hold in the i.i.d. setting for the interested reader.

Hoeffding on intervals

Lemma 8. *Let Z_t be a σ^2-subGaussian martingale difference sequence then for every $\delta > 0$ and every integers $T_1 \leq T_2 \in \mathbb{N}$*

$$\mathbb{P}\left\{\exists t \in [T_1; T_2], \overline{Z}_t \geq \sqrt{\frac{2\sigma^2}{t} \log\left(\frac{1}{\delta}\right) \phi\left(\frac{T_2}{T_1}\right)}\right\} \leq \delta$$

where the mapping $\phi(\cdot)$ is defined by

$$\phi(x) = \frac{1 + x + 2\sqrt{x}}{4\sqrt{x}}$$

and it holds that $1 - \frac{(x-1)^2}{16} \leq \frac{1}{\phi(x)} \leq 1$.

Proof. For any $\lambda > 0$, it holds that

$$\mathbb{E}\left[e^{\lambda t \overline{Z}_t} \middle| \mathcal{F}_{t-1}\right] = \mathbb{E}\left[e^{\lambda Z_t} \middle| \mathcal{F}_{t-1}\right] e^{\lambda(t-1)\overline{Z}_{t-1}} \leq e^{\frac{\lambda^2\sigma^2}{2}} e^{\lambda(t-1)\overline{Z}_{t-1}}$$

since Z_t is σ^2-subGaussian. Thus $X_t = e^{\lambda t \overline{Z}_t - \frac{\lambda^2\sigma^2 t}{2}}$ is a supermartingale and we derive from Markov's inequality that:

$$\mathbb{P}\left(\exists t \in [T_1, T_2], \lambda t \overline{Z}_t \geq \frac{\lambda^2\sigma^2}{2} t + \varepsilon\right) \leq e^{-\varepsilon} \tag{2}$$

As $t \mapsto \sqrt{t}$ is concave on $[T_1, T_2]$, we can find β such that $\beta\sqrt{t} \geq \frac{\lambda^2\sigma^2}{2} t + \varepsilon$ with equality in T_1 and T_2. From these two equalities, we choose λ and β as function of ε:

$$\lambda = \sqrt{\frac{2\varepsilon}{\sigma^2\sqrt{T_1 T_2}}} \qquad \beta = \varepsilon\left(\frac{\sqrt{T_1} + \sqrt{T_2}}{\sqrt{T_1 T_2}}\right)$$

Combining this with (2), we finally obtain:

$$\mathbb{P}\left(\exists t \in [T_1, T_2], \lambda t \overline{Z}_t \geq \beta\sqrt{t}\right) = \mathbb{P}\left(\exists t \in [T_1, T_2], \overline{Z}_t \geq \sqrt{\frac{\sigma^2\varepsilon}{t} \frac{(\sqrt{T_1} + \sqrt{T_2})^2}{4\sqrt{T_1 T_2}}}\right) \leq e^{-\varepsilon},$$

this yields the result. ∎

Anytime Hoeffding

Lemma 9. *Let Z_t be a σ^2-sub-Gaussian martingale difference sequence. Then, for any $\alpha > 0$ and $\delta > 0$ satisfying*

$$1 + \frac{1}{\log(\frac{\log(2)}{\delta})} < \alpha < \frac{\log(\frac{\log(2)}{\delta})}{8},$$

it holds that

$$\mathbb{P}\left\{\exists t \in \mathbb{N}, \overline{Z}_t \geq \sqrt{\frac{2\sigma^2}{t} \log(\frac{\log_2^\alpha(t)}{\delta})}\right\} \leq c_\alpha \delta \sqrt{\log\left(\frac{\delta}{\log(2)}\right)} + \delta(\frac{e^{\frac{\alpha}{2}}}{\log(2)} + 1)\,,$$

$$\textit{where} \;\; c_\alpha = \frac{e^{\frac{\alpha}{2}}}{\log(2)} \sqrt{\frac{1}{8\alpha}} \frac{1}{\alpha(1 - \frac{\alpha}{2\log(\frac{\log(2)}{\delta})}) - 1} \leq \frac{e^{\frac{\alpha}{2}}}{\log^\alpha(2)} \sqrt{\frac{1}{8\alpha}} \frac{16/15}{\alpha - 16/15},$$

so that $c_2 \approx 1.62$ and $\log_2(\cdot)$ is the natural logarithm in basis 2, with the extra assumption that $\log_2(1) = 1$.

If $\alpha, \delta > 0$ are such that

$$1 + \frac{1}{\log(\frac{1}{\delta})} \leq \alpha \leq \frac{1}{2}\log(\frac{1}{\delta})$$

then

$$\mathbb{P}\Big\{\exists t \in \mathbb{N}, \overline{Z}_t \geq \sqrt{\frac{2\sigma^2}{t}\log(\frac{\log_2^\alpha(t)}{\delta})}\Big\} \leq c'_\alpha \delta \log^{\frac{\alpha}{2}}(\frac{1}{\delta}) + \delta(1 + \sqrt{\frac{8\alpha}{\log(1/\delta)}}),$$

$$\textit{where } c'_\alpha = \Big(\frac{2e}{\alpha}\Big)^{\alpha/2}\zeta\big(\alpha - \frac{\alpha^2}{2\log(1/\delta)}\big) \leq \Big(\frac{2e}{\alpha}\Big)^{\alpha/2}\zeta(\frac{3\alpha}{4}),$$

so that $c'_2 \approx 7.11$.

Proof. We are going to use the fact that with probability at least $1-\delta$, for all $s \in [T_1, T_2]$,

$$s\Big(\overline{Z}_s - \mathbb{E}[Z]\Big) \leq \sqrt{2\sigma^2 s\phi(\frac{T_2}{T_1})\log(\frac{1}{\delta})}\,.$$

Define $\varepsilon_t = \sqrt{\frac{2\sigma^2}{t}\log(\frac{\log_2^\alpha(t)}{\delta})}$ so that with $\gamma = 1 + \eta > 1$,

$$\mathbb{P}\Big\{\exists t \in \mathbb{N}, \overline{Z}_t \geq \varepsilon_t\Big\} \leq \sum_{m=0}^{\infty}\mathbb{P}\Big\{\exists t \in [\gamma^m, \gamma^{m+1}], \overline{Z}_t \geq \varepsilon_t\Big\}$$

The case $m = 0$ (corresponding to $t = 1$) will be handled separatly at the cost of an extra δ term. Note that

$$\begin{aligned}
&\sum_{m=\lfloor\frac{1}{\log_2(\gamma)}\rfloor}\mathbb{P}\Big\{\exists t \in [\gamma^m, \gamma^{m+1}], \overline{Z}_t \geq \sqrt{\frac{2\sigma^2}{t}\log(\frac{\log_2^\alpha(\gamma^m)}{\delta})}\Big\}\\
&\leq \sum_{m=\lfloor\frac{1}{\log_2(\gamma)}\rfloor}\mathbb{P}\Big\{\exists t \in [\gamma^m, \gamma^{m+1}], \overline{Z}_t \geq \sqrt{\frac{2\sigma^2}{t}\log(\frac{\log_2^\alpha(\gamma^m)}{\delta})}\sqrt{\phi(\gamma)(1-\frac{\eta^2}{16})}\Big\}\\
&= \sum_{m=\lfloor\frac{1}{\log_2(\gamma)}\rfloor}\mathbb{P}\Big\{\exists t \in [\gamma^m, \gamma^{m+1}], \overline{Z}_t \geq \sqrt{\frac{2\sigma^2}{t}\phi(\frac{\gamma^{m+1}}{\gamma^m})\log\left(\Big(\frac{\log_2^\alpha(\gamma^m)}{\delta}\Big)^{1-\frac{\eta^2}{16}}\right)}\Big\}\,.
\end{aligned}$$

We can now apply Lemma 8. And this gives, assuming $\gamma \leq 2$ (i.e., $\eta < 1$) for the moment,

$$\begin{aligned}
\mathbb{P}\Big\{\exists t \in \mathbb{N}, \overline{Z}_t \geq \varepsilon_t\Big\} &\leq \sum_{m=\lfloor\frac{1}{\log_2(\gamma)}\rfloor}\Big(\frac{\delta}{\log_2^\alpha(\gamma^m)}\Big)^{1-\frac{\eta^2}{16}} + \delta\\
&= \Big(\frac{\delta}{\log_2^\alpha(\gamma)}\Big)^{1-\frac{\eta^2}{16}}\sum_{m=\lfloor\frac{1}{\log_2(\gamma)}\rfloor}\frac{1}{m^{\alpha(1-\frac{\eta^2}{16})}} + \delta\\
&\leq (\frac{\delta}{\log^\alpha(2)})^{1-\frac{\eta^2}{16}}\Big(1 + \frac{1}{\alpha(1-\frac{\eta^2}{16})-1}\frac{1}{\log_2(\gamma)}\Big(\Big) + \delta\Big)\\
&\leq (\frac{\delta}{\log^\alpha(2)})^{1-\frac{\eta^2}{16}}\Big(1 + \frac{1}{\alpha(1-\frac{\eta^2}{16})-1}\frac{1}{\eta}\Big) + \delta
\end{aligned}$$

The choice of $\eta^2 = 8\alpha/\log(\log(2)/\delta)$, which ensure that $\eta < 1$ as long as $\alpha < \frac{1}{8}\log(\frac{\log(2)}{\delta})$ gives

$$\begin{aligned}
\mathbb{P}\Big\{\exists t \in \mathbb{N}, \overline{Z}_t \geq \varepsilon_t\Big\} &\leq \frac{\delta}{\log^\alpha(2)}e^{\frac{\alpha}{2}}\Big(\sqrt{\frac{\log(\frac{\log(2)}{\delta})}{8\alpha}}\frac{1}{\alpha(1-\frac{\alpha}{2\log(\frac{\log(2)}{\delta})})-1} + 1\Big) + \delta\\
&\leq \delta\sqrt{\log\big(\frac{\log(2)}{\delta}\big)}\frac{1}{\log^\alpha(2)}e^{\frac{\alpha}{2}}\sqrt{\frac{1}{8\alpha}}\frac{16/15}{\alpha - 16/15} + \delta\Big(\frac{e^{\frac{\alpha}{2}}}{\log^\alpha(2)} + 1\Big)
\end{aligned}$$

We now consider the case where γ might be bigger than 2, but let us assume for now that $\gamma < 5$ (i.e., $\eta \leq 4$) and the exact same argument with the choice of $\eta^2 = 8\alpha/\log(1/\delta)$ gives

$$\begin{aligned}
\mathbb{P}\Big\{\exists t \in \mathbb{N}, \overline{Z}_t \geq \varepsilon_t\Big\} &\leq \sum_{m=1}^{\infty} \Big(\frac{\delta}{\log_2^{\alpha}(\gamma^m)}\Big)^{1-\frac{\eta^2}{16}} + \gamma\delta \\
&\leq \Big(\frac{\delta}{\log_2^{\alpha}(1+\eta)}\Big)^{1-\frac{\eta^2}{16}} \zeta(\alpha(1-\frac{\eta^2}{16})) + \gamma\delta \\
&\leq \delta^{1-\frac{\eta^2}{16}} \Big(\frac{4}{\eta}\Big)^{\alpha} \zeta(\alpha(1-\frac{\eta^2}{16})) + \gamma\delta \\
&= \delta e^{\frac{\alpha}{2}} \Big(\frac{2\log(1/\delta)}{\alpha}\Big)^{\alpha/2} \zeta\big(\alpha - \frac{\alpha^2}{2\log(1/\delta)}\big) + \Big(1 + \sqrt{\frac{8\alpha}{\log(1/\delta)}}\Big)\delta \\
&= \delta \log^{\frac{\alpha}{2}}\big(\frac{1}{\delta}\big) \Big(\frac{2e}{\alpha}\Big)^{\alpha/2} \zeta\big(\alpha - \frac{\alpha^2}{2\log(1/\delta)}\big) + \Big(1 + \sqrt{\frac{8\alpha}{\log(1/\delta)}}\Big)\delta
\end{aligned}$$

□

Corollary 10. *Let Z_t be a σ^2-sub-Gaussian martingale difference sequence. Then for $\delta > 0$ small enough*

$$\mathbb{P}\Big\{\exists t \in \mathbb{N}, \overline{Z}_t \geq \sqrt{\frac{2\sigma^2}{t}\log(\frac{\log^2(t)}{\delta})}\Big\} \leq \widetilde{\delta}\,,$$

where $\widetilde{\delta}$ is defined by either

$$\widetilde{\delta} = c_2 \frac{\delta}{\log^2(2)}\sqrt{\log(\frac{\log^3(2)}{\delta})} + \frac{\delta}{\log^2(2)}(\frac{e}{\log(2)} + 1)\,,$$

or, depending on the range of δ, by,

$$\widetilde{\delta} = c_2' \frac{\delta}{\log^2(2)} \log(\frac{\log^2(2)}{\delta}) + 5\frac{\delta}{\log^2(2)}.$$

In the first case, $\widetilde{\delta}$ is of the order of $\delta\sqrt{\log(\frac{1}{\delta})}$ and, in the second case, of the order of $\delta\log(\frac{1}{\delta})$.

The following concentration inequality will be useful for the non-iid case. In that framework, at some stage t, the total number of users in the population is random, denoted by n_t. We recall that r_n denotes the mean of the n-th user of the population and $\varepsilon_{n,s}$ is the random white noise for that user after he is in the population for s stages. In the remaining, we assume that the expectation of r_u is equal to r and that this random variable is σ_r^2-subGaussian, On the other hand, the expectation of ε is naturally 0 and this random variable is σ_ε^2-subGaussian. An algorithm is therefore a sampling policy $\mathcal{A}$ that indicates after seeing the first n values of r_u plus some empirical average noise $\overline{\varepsilon}_{u,t}$ at time $t \in \mathbb{N}$ whether to add a new user or not. We denote by $\mathcal{A}_t \in \{0,1\}$ the decision to include a new user or not at stage t. We denote by $\mathcal{T}_n \in \mathbb{N}$ the time where the n-th user is added, by $\tau_n = \mathcal{T}_{n+1} - \mathcal{T}_n$ the number of stages with exactly n users and by $\tau_{m:n} = \sum_{s=m}^{n-1} \tau_s$ the number of stages between the arrival of the m-th user and the n-th one. We also denote by $n+t$ the number of user at stage $t \in \mathbb{N}$,

Proposition 11. *For any algorithm and $\delta > 0$, it holds*

$$\mathbb{P}\Big\{\exists t \in \mathbb{N}, \frac{1}{n_t}\sum_{u=1}^{n_t} r_u + \overline{\varepsilon}_{u,t-\mathcal{T}_u+1} \leq r - \sqrt{\frac{2\Big(\sigma_r^2 + \frac{\sigma_\varepsilon^2 \log(en_t)}{n_t}\Big)}{n_t}\log\Big(\frac{(4n_t)^4}{6\delta}\max\{1, \frac{n_t\sigma_r^2}{\sigma_\varepsilon^2}\}\Big)}\Big\} \leq \widetilde{\delta}$$

Proof. We rewrite the statement of the proposition and notice that we just need to prove that, for any $n \in \mathbb{N}$,

$$\mathbb{P}\Big\{\exists 1 \leq s \leq \tau_n, \overline{r}_n + \frac{\overline{\varepsilon}_{1,s+\tau_{1:n}} + \ldots + \overline{\varepsilon}_{n,s}}{n} \leq r - \sqrt{\frac{2\Big(\sigma_r^2 + \frac{\sigma_\varepsilon^2\log(en)}{n}\Big)}{n}\log\Big(\frac{36n^4}{\delta}\max\{1, \frac{n\sigma_r^2}{\sigma_\varepsilon^2}\}\Big)}\Big\} \leq \frac{1}{3}\frac{\widetilde{\delta}}{n^{3/2}}\,.$$

The exponent $3/2$ will come from the fact that $\widetilde{\delta}$ is of the order of $\delta \log \frac{1}{\delta}$ (and not δ). We will even prove the following

$$\mathbb{P}\Big\{\exists 1 \leq s < \infty, \overline{r}_n + \frac{\overline{\varepsilon}_{1,s+\tau_{1:n}} + \ldots + \overline{\varepsilon}_{n,s}}{n} \leq r - \sqrt{\frac{2\Big(\sigma_r^2 + \frac{\sigma_\varepsilon^2 \log(en)}{n}\Big)}{n} \log\Big(\frac{36n^4}{\delta} \max\{1, \frac{n\sigma_r^2}{\sigma_\varepsilon^2}\}\Big)}\Big\} \leq \frac{1}{3}\frac{\widetilde{\delta}}{n^{3/2}}$$

We will decompose the considered event defined on $\{1 \leq s < \infty\}$ in two, depending whether $s \leq 6n^2 \max\{1, \frac{n\sigma_r^2}{\sigma_\varepsilon^2}\}$ or $s > 6n^2 \max\{1, \frac{n\sigma_r^2}{\sigma_\varepsilon^2}\}$.

In the first case, we aim at proving that for all $s \leq 6n^2 \max\{1, \frac{n\sigma_r^2}{\sigma_\varepsilon^2}\}$,

$$\mathbb{P}\Big\{\overline{r}_n + \frac{\overline{\varepsilon}_{1,s+\tau_{1:n}} + \ldots + \overline{\varepsilon}_{n,s}}{n} \leq r - \sqrt{\frac{2\Big(\sigma_r^2 + \frac{\sigma_\varepsilon^2 \log(en)}{n}\Big)}{n} \log\Big(\frac{36n^4}{\delta} \max\{1, \frac{n\sigma_r^2}{\sigma_\varepsilon^2}\}\Big)}\Big\} \leq \frac{1}{18}\frac{\widetilde{\delta}}{2n^{7/2}}\frac{1}{\max\{1, \frac{n\sigma_r^2}{\sigma_\varepsilon^2}\}}.$$

As usual, take $\eta > 0$ and let us try to upper bound

$$\mathbb{P}\Big\{r_1 + \ldots + r_n - nr + \overline{\varepsilon}_{1,s+\tau_{1:n}} + \ldots + \overline{\varepsilon}_{n,s} \geq \eta\Big\}$$

Introduce $\lambda > 0$ and Markov inequality yields

$$\mathbb{P}\Big\{r_1 + \ldots + r_n - nr + \overline{\varepsilon}_{1,s+\tau_{1:n}} + \ldots + \overline{\varepsilon}_{n,s} \geq \eta\Big\} \leq \mathbb{E}[e^{\lambda(r_1 + \ldots + r_n - nr + \overline{\varepsilon}_{1,s+\tau_{1:n}} + \ldots + \overline{\varepsilon}_{n,s})}]e^{-\lambda\eta}$$

Note that the realizations of $\overline{\varepsilon}_s^n$ and μ^n are independent of the other average once we have taken the decision to include that user, i.e., conditionally to $\mathcal{T}_n$

$$\begin{aligned}
\mathbb{E}[e^{\lambda(r_1 + \ldots + r_n - nr + \overline{\varepsilon}_{1,s+\tau_{1:n}} + \ldots + \overline{\varepsilon}_{n,s})}] &= \mathbb{E}\Big[\mathbb{E}\big[e^{\lambda(r_n - r + \overline{\varepsilon}_{n,s})}e^{\lambda(r_1 + \ldots + r_{n-1} - (n-1)r + \overline{\varepsilon}_{1,s+\tau_{1:n}} + \ldots + \overline{\varepsilon}_{n-1,s+\tau_{n-1:n}})}\big|\mathcal{T}_n\big]\Big] \\
&= \mathbb{E}\Big[\mathbb{E}\big[e^{\lambda(r_n - r + \overline{\varepsilon}_{n,s})}\big|\mathcal{T}_n\big]\mathbb{E}\big[e^{\lambda(r_1 + \ldots + r_{n-1} + \overline{\varepsilon}_{1,s+\tau_{1:n}} + \ldots + \overline{\varepsilon}_{n-1,s+\tau_{n-1:n}})}\big|\mathcal{T}_n\big]\Big] \\
&= \mathbb{E}\Big[\mathbb{E}\big[e^{(r_n - r + \overline{\varepsilon}_{n,s})}\big]\mathbb{E}\big[e^{\lambda(r_1 + \ldots + r_{n-1} - (n-1)r + \overline{\varepsilon}_{1,s+\tau_{1:n}} + \ldots + \overline{\varepsilon}_{n-1,s+\tau_{n-1:n}})}\big|\mathcal{T}_n\big]\Big] \\
&= \mathbb{E}\big[e^{(r_n - r + \overline{\varepsilon}_{n,s})}\big]\mathbb{E}\big[e^{\lambda(r_1 + \ldots + r_{n-1} - (n-1)r + \overline{\varepsilon}_{1,s+\tau_{1:n}} + \ldots + \overline{\varepsilon}_{n-1,s+\tau_{n-1:n}})}\big]
\end{aligned}$$

Let us rewrite, for clarity, the last expectation of the r.h.s. as

$$\mathbb{E}\big[e^{\lambda(r_1 + \ldots + r_{n-1} - (n-1)r + \overline{\varepsilon}_{1,s+\tau_{1:n}} + \ldots + \overline{\varepsilon}_{n-1,s+\tau_{n-1:n}})}\big] = \mathbb{E}\big[e^{\lambda(r_1 + \ldots + r_{n-1} - (n-1)r + \overline{\varepsilon}_{1,s+\tau_{n-1}+\tau_{1:n-1}} + \ldots + \overline{\varepsilon}_{n-1,s+\tau_{n-1}})}\big]$$

If we condition similarly to $\mathcal{T}_{n-1}$, we can focus on upper-bounding

$$\begin{aligned}
\mathbb{E}\Big[e^{\lambda(r_{n-1} - \mu + \overline{\varepsilon}_{n-1,s+\tau_{n-1}})}\Big] &= \sum_{j=s+1}^{\infty} \mathbb{E}\Big[e^{\lambda(r_{n-1} - \mu + \overline{\varepsilon}_{n-1,j})}\mathbb{1}\{s + \tau_{n-1} = j\}\Big] \\
&\leq \sum_{j=s+1}^{\infty} e^{\lambda^2(\frac{\sigma_r^2}{2} + \frac{\sigma_\varepsilon^2}{2j})}\mathbb{E}\Big[\mathbb{1}\{s + \tau_{n-1} = j\}\Big] \leq e^{\lambda^2(\frac{\sigma_r^2}{2} + \frac{\sigma_\varepsilon^2}{2(s+1)})}
\end{aligned}$$

Choosing $\lambda = \frac{\eta}{n\sigma_r^2 + \sigma_\varepsilon^2 \log(en)}$ gives that

$$\mathbb{P}\Big\{r_1 + \ldots + r_n - nr + \overline{\varepsilon}_{1,s+\tau_{1:n}} + \ldots + \overline{\varepsilon}_{n,s} \geq \eta\Big\} \leq e^{-\frac{\eta^2}{2\big(n\sigma_r^2 + \sigma_\varepsilon^2 \log(en)\big)}}$$

So, for all $n \in \mathbb{N}$ and all stage $s \leq 6n^2 \max\{1, \frac{n\sigma_r^2}{\sigma_\varepsilon^2}\}$ and with probability at least $1 - \frac{\delta}{2}$, it holds

$$\overline{r}_n + \frac{\overline{\varepsilon}_{1,s+\tau_{1:n}} + \ldots + \overline{\varepsilon}_{n,s}}{n} \geq \mu - \sqrt{\frac{2\Big(\sigma_r^2 + \frac{\sigma_\varepsilon^2 \log(en)}{n}\Big)}{n} \log\Big(\frac{36n^4}{\delta} \max\{1, \frac{n\sigma_r^2}{\sigma_\varepsilon^2}\}\Big)}$$

We now focus on the stages where $s > 6n^2 \max\{1, \frac{n\sigma_r^2}{\sigma_\varepsilon^2}\}$. But first, notice that Lemma 9 implies that:

$$\mathbb{P}\Big\{\exists n \in \mathbb{N}, \quad \overline{r}_n \leq r - \sqrt{2\sigma_r^2 \frac{\log \frac{12\log^2(n)}{\delta}}{n}}\Big\} \leq \frac{\widetilde{\delta}}{4},$$

and also, similarly,

$$\mathbb{P}\Big\{\exists i \in \mathbb{N}, s \in \mathbb{N}, \quad \overline{\varepsilon}_{i,s} \leq -\sqrt{2\sigma_\varepsilon^2 \frac{\log \frac{36i^2\log^2(s)}{\delta}}{s}}\Big\} \leq \frac{\widetilde{\delta}}{4}.$$

This implies that, with probability at least $1 - \frac{\widetilde{\delta}}{2}$, for every $n \in \mathbb{N}$ and $s \geq \underline{s} := 6n^2 \max\{1, \frac{n\sigma_r^2}{\sigma_\varepsilon^2}\}$

$$\overline{r}_n + \frac{\overline{\varepsilon}_{1,s+\tau_{1:n}} + \ldots + \overline{\varepsilon}_{n,s}}{n} \geq r - \left(\sqrt{2\sigma_r^2 \frac{\log \frac{12\log^2(n)}{\delta}}{n}} + \sqrt{\frac{2\sigma_\varepsilon^2}{\underline{s}} \log\left(\frac{36n^4}{\delta}\log^2(\underline{s})\right)}\right)$$

It only remains to notice that

$$\sqrt{2\sigma_r^2 \frac{\log \frac{18\log^2(n)}{\delta}}{n}} + \sqrt{\frac{2\sigma_\varepsilon^2}{\underline{s}} \log\left(\frac{36n^4}{\delta}\log^2(\underline{s})\right)} \leq \sqrt{\frac{2\left(\sigma_r^2 + \frac{\sigma_\varepsilon^2 \log(en)}{n}\right)}{n} \log\left(\frac{36n^4}{\delta} \max\{1, \frac{n\sigma_r^2}{\sigma_\varepsilon^2}\}\right)}$$

This inequality is a consequence of the fact that $\sqrt{a} + \sqrt{\lambda b} \leq \sqrt{a+b}$ as soon as $\lambda \leq \frac{1}{6\max\{1, a/b\}}$, and this gives the result. $\square$

B UCB$_\alpha$ and ETC: proofs.

Our exact theorems are the two following.

Theorem 12. *With probability greater than $1 - \tilde{\delta}$, UCB$_\alpha$ with $\alpha > 1$ returns the best arm at a stage t_d with*

$$\tau_d \leq (C_1 + C_2)\log\frac{2}{\delta} + 2C_1 \log\log(2C_1 \max\{\log\frac{2}{\delta}, 2\log(2C_1)\}) + 2C_2\log\log(2C_2\max\{\log\frac{2}{\delta}, 2\log(2C_2)\}),$$

with $C_1 = \frac{2\sigma^2}{\Delta^2}\min\{(\alpha+1)^2, \frac{16\alpha^2}{(\alpha-1)^2}\} + 1$ and $C_2 = \frac{(\alpha+1)^2}{(\alpha-1)^2}C_1 + 1$.

With the same probability, the regret $R(\tau_d)$ of UCB$_\alpha$ at the time of decision verifies

$$R(\tau_d) \leq \left(\frac{2\sigma^2}{\Delta}\min\{(\alpha+1)^2, \frac{16\alpha^2}{(\alpha-1)^2}\} + \Delta\right)\left(\log\frac{2}{\delta} + 2\log\log(2C_1\max\{\log\frac{2}{\delta}, 2\log(2C_1)\})\right)$$

Theorem 13. *With probability greater than $1 - \tilde{\delta}$, ETC returns the best arm at a stage τ_d with*

$$\tau_d \leq \frac{32\sigma^2}{\Delta^2}\left(\log\frac{1}{\delta} + 2\log\log(\frac{32\sigma^2}{\Delta^2}\max\{\log\frac{1}{\delta}, 2\log(\frac{32\sigma^2}{\Delta^2})\})\right).$$

With the same probability, the regret R_{ETC} of ETC at the time of decision verifies

$$R(\tau_d) \leq \frac{16\sigma^2}{\Delta}\left(\log\frac{1}{\delta} + 2\log\log(\frac{32\sigma^2}{\Delta^2}\max\{\log\frac{1}{\delta}, 2\log(\frac{32\sigma^2}{\Delta^2})\})\right).$$

We first prove a generic lemma, which will also be useful in the non-IID case.

Lemma 14. *Consider the two arms problem where $\mathcal{A}$ is the best arm.*

Let $\delta \in (0,1]$ and $(\varepsilon_n)_{n\in\mathbb{N}}$ be a sequence such that with probability $1-\delta$, for all $n_\mathcal{A}, n_\mathcal{B} \in \mathbb{N}^*$ we have the concentration inequalities $\hat{r}^\mathcal{A}(n_\mathcal{A}) + \varepsilon_{n_\mathcal{A}} \geq \mu_\mathcal{A}$ and $\hat{r}^\mathcal{B}(n_\mathcal{B}) - \varepsilon_{n_\mathcal{B}} \leq \mu_\mathcal{B}$. Suppose that for all $n \geq n_0 \in \mathbb{N}$, $\frac{1}{\varepsilon_{n+1}^2} - \frac{1}{\varepsilon_n^2} \leq C$.

Let an algorithm be such that it pulls each arm n_0 times, then pulls $\arg\max_{i\in(\mathcal{A},\mathcal{B})} \hat{r}^i(n_i) + \alpha\varepsilon_{n_i}$ for $\alpha > 1$, and takes a decision if for some $i, j \in (\mathcal{A}, \mathcal{B})$, $\hat{r}^i(n_i) - \varepsilon_{n_i} > \hat{r}^i(n_j) + \varepsilon_{n_j}$. Then with probability $1-\delta$, the algorithm correctly returns $\mathcal{A}$ and at all times after $2n_0$ and prior to the decision

$$\frac{1}{\varepsilon^2_{n_\mathcal{B}}} \le \frac{1}{\Delta^2}\min\{(\alpha+1)^2, \frac{16\alpha^2}{(\alpha-1)^2}\} + C\,,$$
$$\frac{1}{\varepsilon^2_{n_\mathcal{A}}} \le \frac{(\alpha+1)^2}{(\alpha-1)^2}\left(\frac{1}{\Delta^2}\min\{(\alpha+1)^2, \frac{16\alpha^2}{(\alpha-1)^2}\} + C\right) + C\,.$$

Proof. We prove that the number of pulls of the best arm $\mathcal{A}$ is bounded by a function of the number of pulls of $\mathcal{B}$, which is itself bounded since it is the worse arm.

Relation between $\mathcal{A}$ and $\mathcal{B}$ when $\mathcal{A}$ is pulled.

- No decision was taken yet: $\hat{r}^\mathcal{A}(n_\mathcal{A}) - \varepsilon_{n_\mathcal{A}} \le \hat{r}^\mathcal{B}(n_\mathcal{B}) + \varepsilon_{n_\mathcal{B}}$ (1),
- $\mathcal{A}$ is pulled: $\hat{r}^\mathcal{A}(n_\mathcal{A}) + \alpha\varepsilon_{n_\mathcal{A}} \ge \hat{r}^\mathcal{B}(n_\mathcal{B}) + \alpha\varepsilon_{n_\mathcal{B}}$ (2).

Subtract (1) from (2) to obtain $(\alpha+1)\varepsilon_{n_\mathcal{A}} \ge (\alpha-1)\varepsilon_{n_\mathcal{B}}$. Equivalently, $\frac{1}{\varepsilon^2_{n_\mathcal{A}}} \le \frac{(\alpha+1)^2}{(\alpha-1)^2}\frac{1}{\varepsilon^2_{n_\mathcal{B}}}$. Since the left-hand-side grows only when $\mathcal{A}$ is chosen and grows at most by C, we have that for all stages,

$$\frac{1}{\varepsilon^2_{n_\mathcal{A}}} \le \frac{(\alpha+1)^2}{(\alpha-1)^2}\frac{1}{\varepsilon^2_{n_\mathcal{B}}} + C\,.$$

Upper bound on $n_\mathcal{B}$. When $\mathcal{B}$ is pulled, $\hat{r}^\mathcal{A}(n_\mathcal{A}) + \alpha\varepsilon_{n_\mathcal{A}} \le \hat{r}^\mathcal{B}(n_\mathcal{B}) + \alpha\varepsilon_{n_\mathcal{B}}$. With probability $1-\tilde{\delta}$, for all $n_\mathcal{A}$ and $n_\mathcal{A}$ we also have the concentration inequalities $\hat{r}^\mathcal{A}(n_\mathcal{A}) + \varepsilon_{n_\mathcal{A}} \ge \mu_\mathcal{A}$ and $\hat{r}^\mathcal{B}(n_\mathcal{B}) - \varepsilon_{n_\mathcal{B}} \le \mu_\mathcal{B}$. Hence

$$\mu_\mathcal{A} + (\alpha-1)\varepsilon_{n_\mathcal{A}} \le \mu_\mathcal{B} + (\alpha+1)\varepsilon_{n_\mathcal{B}}\,.$$

From this inequality we can get that $\frac{1}{\varepsilon^2_{n_\mathcal{B}}} \le \frac{(\alpha+1)^2}{\Delta^2}$. Since the left-hand-side grows only when $\mathcal{B}$ is pulled and grows at most by 1, we have for all stages

$$\frac{1}{\varepsilon^2_{n_\mathcal{B}}} \le \frac{(\alpha+1)^2}{\Delta^2} + C\,.$$

In order to get another bound, relevant when α is big, we write that when no decision is taken and concentration holds, we have

$$\mu_\mathcal{A} - 2\varepsilon_{n_\mathcal{A}} \le \hat{r}^\mathcal{A}(n_\mathcal{A}) - \varepsilon_{n_\mathcal{A}} \le \hat{r}^\mathcal{B}(n_\mathcal{B}) + \varepsilon_{n_\mathcal{B}} \le \mu_\mathcal{B} + 2\varepsilon_{n_\mathcal{B}}\,.$$

This leads to $\varepsilon_{n_\mathcal{B}} + \varepsilon_{n_\mathcal{A}} \ge \frac{\Delta}{2}$. When $\mathcal{B}$ is pulled, we also have the inequality $(\alpha+1)\varepsilon_{n_\mathcal{B}} \ge (\alpha-1)\varepsilon_{n_\mathcal{A}}$, such that

$$\frac{\Delta}{2} \le \varepsilon_{n_\mathcal{B}}(1 + \frac{\alpha+1}{\alpha-1}) = \varepsilon_{n_\mathcal{B}}\frac{2\alpha}{\alpha-1}\,.$$

This gives a second inequality for n_B. For all stages,

$$\frac{1}{\varepsilon^2_{n_\mathcal{B}}} \le \frac{16\alpha^2}{\Delta^2(\alpha-1)^2} + C\,.$$

Bound on the decision time. We have the following inequalities for all stages prior to the decision,

$$\frac{1}{\varepsilon^2_{n_\mathcal{B}}} \le \frac{1}{\Delta^2}\min\{(\alpha+1)^2, \frac{16\alpha^2}{(\alpha-1)^2}\} + C\,,$$
$$\frac{1}{\varepsilon^2_{n_\mathcal{A}}} \le \frac{(\alpha+1)^2}{(\alpha-1)^2}\left(\frac{1}{\Delta^2}\min\{(\alpha+1)^2, \frac{16\alpha^2}{(\alpha-1)^2}\} + C\right) + C\,.$$

□

Proof of Theorem 12. We apply Lemma 14 with $\varepsilon_n = \sqrt{\frac{2\sigma^2}{n}\log(\frac{2\log^2 n}{\delta})}$, for which concentration holds with probability $1-\tilde{\delta}$ and for which we can take $n_0 = 3$ and $C = \frac{1}{2\sigma^2\log(2/\delta)} < \frac{1}{2\sigma^2}$. We obtain the following inequalities for all stages prior to the decision,

$$\frac{n_{\mathcal{B}}}{\log(\frac{2\log^2 n_{\mathcal{B}}}{\delta})} \le \frac{2\sigma^2}{\Delta^2}\min\{(\alpha+1)^2, \frac{16\alpha^2}{(\alpha-1)^2}\} + 1\,,$$

$$\frac{n_{\mathcal{A}}}{\log(\frac{2\log^2 n_{\mathcal{A}}}{\delta})} \le \frac{(\alpha+1)^2}{(\alpha-1)^2}\left(\frac{2\sigma^2}{\Delta^2}\min\{(\alpha+1)^2, \frac{16\alpha^2}{(\alpha-1)^2}\} + 1\right) + 1\,.$$

We introduce the notations $C_1 = \frac{2\sigma^2}{\Delta^2}\min\{(\alpha+1)^2, \frac{16\alpha^2}{(\alpha-1)^2}\} + 1$ and $C_2 = \frac{(\alpha+1)^2}{(\alpha-1)^2}C_1 + 1$. At the time of decision,

$$n_{\mathcal{B}} \le C_1\left(\log\frac{2}{\delta} + 2\log\log(2C_1\max\{\log\frac{2}{\delta},\ 2\log(2C_1)\})\right)$$

$$n_{\mathcal{A}} \le C_2\left(\log\frac{2}{\delta} + 2\log\log(2C_2\max\{\log\frac{2}{\delta},\ 2\log(2C_2)\})\right)$$

$$t = n_{\mathcal{A}} + n_{\mathcal{B}} \le (C_1 + C_2)\log\frac{1}{\delta} + L\,,$$

where L regroups the doubly logarithmic terms.

The regret is Δn_B. Thus with probability greater than $1-\tilde{\delta}$,

$$R_{t_d} \le \left(\frac{2\sigma^2}{\Delta}\min\{(\alpha+1)^2, \frac{16\alpha^2}{(\alpha-1)^2}\}+\Delta\right)\left(\log\frac{2}{\delta}+2\log\log(2C_1\max\{\log\frac{2}{\delta},\ 2\log(2C_1)\})\right)$$

□

Proof of Theorem 13. Let $n = t/2$. We consider only even stages t. Let $\varepsilon'_n = \sqrt{\frac{4\sigma^2}{n}\log(\frac{\log^2 n}{\delta})}$. As long as no decision is taken we have $\hat{r}^{\mathcal{A}}(n) - \hat{r}^{\mathcal{B}}(n) \le \varepsilon'_n$.

With probability $1-\tilde{\delta}$, for all $n \ge 1$, we have the concentration inequality $\hat{r}^{\mathcal{A}}(n) - \hat{r}^{\mathcal{B}}(n) \ge \Delta - \varepsilon'_n$.

Combining the two inequalities we obtain that as long as no decision is taken, $\varepsilon'_n \le \frac{\Delta}{2}$. That is,

$$\frac{n}{\log(\frac{\log^2 n}{\delta})} \le \frac{16\sigma^2}{\Delta^2} \quad \Rightarrow \quad n \le \frac{16\sigma^2}{\Delta^2}(\log\frac{1}{\delta} + 2\log\log(\frac{32\sigma^2}{\Delta^2}\max\{\log\frac{1}{\delta},\ 2\log(\frac{32\sigma^2}{\Delta^2})\}))\,.$$

This bound on n gives both a bound on the regret (Δn) and on the decision time ($2n$). □

Lemma 15. *Let $a, b > e$ and $a \ge b$. If $t \ge a + b\log\log(\max\{2a, 2b\log(2b)\})$, then $t \ge a + b\log\log(t)$.*

Proof. For $t \ge b$, the function $t \to t - b\log\log(t)$ is increasing. Let x be the solution of $x - b\log\log x = a$. We will show that the proposed t is bigger than x.

Case 1: $a \ge b\log\log(x)$. Then $x = a + b\log\log(x) \le 2a$ and $2a \ge a + b\log\log(2a)$, such that

$$a + b\log\log(2a) \ge a + b\log\log(a + b\log\log(2a))\,,$$

from which we conclude that $x \le a + b\log\log(2a)$. Since t is bigger than the latter, it is bigger than x.

Case 2: $a \le b\log\log(x)$. Then $x \le 2b\log\log x$. If $x > 2b\log 2b$ then $\frac{x}{\log\log x} > \frac{2b\log 2b}{\log\log(2b\log 2b)} \ge 2b$. We obtain that $x \le 2b\log 2b$. This implies that

$$a + b\log\log(2b\log(2b)) \ge a + b\log\log(a + b\log\log(2b\log(2b)))\,,$$

hence $x \le a + b\log\log(2b\log(2b))$.

In both cases, the proposed t is bigger than x, hence it verifies the wanted inequality. □

C UCB-MM$_\alpha$ and ETC-MM: proofs.

Lemma 16. *Assume that $\sigma_r \geq \frac{\Delta}{\sqrt{2}}$ and $\gamma > 4$, and let n be defined by the following equation*

$$\frac{2\Big(\sigma_r^2 + \frac{\sigma_\varepsilon^2 \log(en)}{n}\Big)}{n} \log\Big(\frac{36n^4}{\delta}\max\{1, \frac{n\sigma_r^2}{\sigma_\varepsilon^2}\}\Big) = \frac{\Delta^2}{\gamma} \tag{3}$$

then

$$n \leq \frac{2\gamma\sigma_r^2}{\Delta^2}\log(\frac{1}{\delta})(1+o_\delta(1)) + \frac{\sigma_\varepsilon^2}{\sigma_r^2}\log\log(\frac{1}{\delta})(1+o_\delta(1))$$

Proof. Consider first the following equation

$$\frac{2\Sigma^2}{n}\log(\frac{n^5}{\delta}) = C$$

and denote by n_0 its solution. It follows from straightforward computations that,

$$n_0 \leq \frac{2\Sigma^2}{C}\left(\log(\frac{1}{\delta}) + 5\log(\frac{2\Sigma^2}{C}) + 10\overline{\log}\Big(\log(\frac{1}{\delta}) + 5\log\frac{2\Sigma^2}{C}\Big)\right) = \frac{2\Sigma^2}{C}\log(\frac{1}{\delta})\left(1+o_\delta(1)\right),$$

where $\overline{\log}(X) = \max\{5, \log(X)\}$. We now go back to Equation (3), and assume for the moment that the solution n^* is such that $n^*\frac{\sigma_r^2}{\sigma_\varepsilon^2} \geq 1$. Moreover, it is clear that

$$n^* \geq \frac{2\gamma\sigma_r^2}{\Delta^2}\log\left(\frac{36}{\widetilde{\delta}}\frac{\sigma_r^2}{\sigma_\varepsilon^2}\left(\frac{2\gamma\sigma_r^2}{\Delta^2}\right)^5\right) =: \underline{n}$$

As a consequence, if we denote by $\underline{\Sigma}^2 = \sigma_r^2 + \frac{\sigma_\varepsilon^2\log(\underline{n})}{\underline{n}}$, then n^* is such that

$$\frac{2\underline{\Sigma}^2}{n^*}\log\left(\frac{(n^*)^5}{\underline{\delta}}\right) \leq \frac{\Delta^2}{\gamma}, \quad \text{where } \underline{\delta} = \frac{36\sigma_r^2}{\widetilde{\delta}\sigma_\varepsilon^2}.$$

So at the end, we have proved that

$$n^* = \frac{2\gamma\underline{\Sigma}^2}{\Delta^2}\log(\frac{1}{\delta})(1+o_\delta(1)) = \frac{2\gamma\sigma_r^2}{\Delta^2}\log(\frac{1}{\delta})(1+o_\delta(1)) + \frac{\sigma_\varepsilon^2}{\sigma_r^2}\log\log(\frac{1}{\delta})(1+o_\delta(1))$$

which gives the result. □

Corollary 17. *The decision time of UCB and ETC corresponds respectively to the solution of Equation (3) with $\gamma = 4$ for UCB and $\gamma = 8$ for ETC (and $\gamma = 16$ for ETC').*

Theorem 18. *Given $\delta > 0$ and $\alpha \geq 1$, the decision time of UCB-MM$_\alpha$ is such that, with probability at least $1 - \widetilde{\delta}$,*

$$t_d \leq \frac{2\sigma_r^2}{\Delta^2}\Big(2\frac{\alpha^2+1}{(\alpha-1)^2}\Big(\min\big\{(\alpha+1)^2, \frac{16\alpha^2}{(\alpha-1)^2}\big\} + \Delta^2\Big) + \Delta^2\Big)\log(\frac{1}{\delta})(1+o_\delta(1)) + 2\frac{\sigma_r^2}{\sigma_\varepsilon^2}\log\log(\frac{1}{\delta})(1+o_\delta(1))$$

Moreover, on the same event, the regret of UCB-MM$_\alpha$ at decision time satisfies

$$R_{t_d} \leq \Big(\frac{8\sigma_r}{\Delta^2}\min\big\{\frac{(\alpha+1)^2}{4}, \frac{4\alpha^2}{(\alpha-1)^2}\big\} + \Delta\Big)\log(\frac{1}{\delta})(1+o_\delta(1)) + \frac{\sigma_r^2}{\sigma_\varepsilon^2}\Delta\log\log(\frac{1}{\delta})(1+o_\delta(1))$$

Proof. The proof is almost identical to the iid case. The main difference is the change in error terms. Indeed, for $n \in \mathbb{N}^*$, we define

$$\varepsilon_n = \sqrt{\frac{2\Big(\sigma_r^2 + \frac{\sigma_\varepsilon^2\log(en)}{n}\Big)}{n}\log\Big(\frac{36n^4}{\widetilde{\delta}}\max\{1, \frac{n\sigma_r^2}{\sigma_\varepsilon^2}\}\Big)}.$$

We apply Lemma 14 with this ε_n, $n_0 = 0$ and $C = 1$. It yields that for all stages prior to the decision,

$$\frac{1}{\varepsilon^2_{n_\mathcal{B}}} \le \frac{1}{\Delta^2}\min\{(\alpha+1)^2, \frac{16\alpha^2}{(\alpha-1)^2}\} + 1\,,$$
$$\frac{1}{\varepsilon^2_{n_\mathcal{A}}} \le \frac{(\alpha+1)^2}{(\alpha-1)^2}\left(\frac{1}{\Delta^2}\min\{(\alpha+1)^2, \frac{16\alpha^2}{(\alpha-1)^2}\} + 1\right) + 1\,.$$

We now introduce the notations $\gamma_1 = \min\{(\alpha+1)^2, \frac{16\alpha^2}{(\alpha-1)^2}\} + \Delta^2$ and $\gamma_2 = \frac{(\alpha+1)^2}{(\alpha-1)^2}\gamma_1 + \Delta^2$. At the time of decision,

$$\begin{aligned}
n_\mathcal{B} &= \frac{2\gamma_1\sigma_r^2}{\Delta^2}\log(\frac{1}{\delta})(1+o_\delta(1)) + \frac{\sigma_r^2}{\sigma_\varepsilon^2}\log\log(\frac{1}{\delta})(1+o_\delta(1))\\
n_\mathcal{A} &= \frac{2\gamma_2\sigma_r^2}{\Delta^2}\log(\frac{1}{\delta})(1+o_\delta(1)) + \frac{\sigma_r^2}{\sigma_\varepsilon^2}\log\log(\frac{1}{\delta})(1+o_\delta(1))\\
t_d = n_\mathcal{A} + n_\mathcal{B} &= \frac{2(\gamma_2+\gamma_1)\sigma_r^2}{\Delta^2}\log(\frac{1}{\delta})(1+o_\delta(1)) + 2\frac{\sigma_r^2}{\sigma_\varepsilon^2}\log\log(\frac{1}{\delta})(1+o_\delta(1))\\
&= \frac{2\sigma_r^2}{\Delta^2}\Big(2\frac{\alpha^2+1}{(\alpha-1)^2}\big(\min\big\{(\alpha+1)^2, \frac{16\alpha^2}{(\alpha-1)^2}\big\} + \Delta^2\big) + \Delta^2\Big)\log(\frac{1}{\delta})(1+o_\delta(1))\\
&\qquad\qquad + 2\frac{\sigma_r^2}{\sigma_\varepsilon^2}\log\log(\frac{1}{\delta})(1+o_\delta(1))
\end{aligned}$$

As a consequence, on the same event,

$$R_{t_d} \le \Big(\frac{8\sigma_r}{\Delta^2}\min\big\{\frac{(\alpha+1)^2}{4}, \frac{4\alpha^2}{(\alpha-1)^2}\big\} + \Delta\Big)\log(\frac{1}{\delta})(1+o_\delta(1)) + \frac{\sigma_r^2}{\sigma_\varepsilon^2}\Delta\log\log(\frac{1}{\delta})(1+o_\delta(1))$$

□

D Static population

In the static setting, all users are allocated to populations $\mathcal{A}$ and $\mathcal{B}$ from the beginning of the test, and we only consider the case where the size of both populations are equal, even though the generalization to different population size is almost straightforwaed.

Finite fixed horizon: The first baseline is to wait until some horizon T and perform a statistical test based on a confidence bound on the uplift Δ.

Proposition 19. *In the static setting, the following holds with probability at least $1-\delta$,*

$$\Delta - (\hat{r}^\mathcal{A} - \hat{r}^\mathcal{B}_T) \le \sqrt{\frac{8\left(\sigma_r^2 + \frac{\sigma_\varepsilon^2}{T}\right)\log\frac{1}{\delta}}{n}}.$$

Therefore, the procedure waiting until the horizon T to select the $\mathcal{B}$ if $\hat{\Delta}_T$ is greater than the r.h.s. term has a linear regret of $R(T) = n\Delta/2$ and is guaranteed to be (δ, T)-PAC

$$n \ge \frac{32(\sigma_r^2 + \frac{\sigma_\varepsilon^2}{T})}{\Delta^2}\log\frac{1}{\delta}\,.$$

The proof is a direct consequence of standard concentration inequalities.

Adaptive decision time Instead of waiting for a fixed arbitrary horizon T, the decision can often be taken before, at the cost of using maximal concentration inequalities, that are valid at all stages.

Theorem 20. *In the static setting, it holds that, for all $t \in \mathbb{N}$ and with probability at least $1 - \widetilde{\delta}$,*

$$\Delta - (\hat{r}_t^{\mathcal{A}} - \hat{r}_t^{\mathcal{B}}) \leq \sqrt{\frac{8\sigma_r^2 \log \frac{2}{\delta}}{n}} + \sqrt{\frac{8\sigma_\varepsilon^2 \log \frac{3\log^2(t)}{\delta}}{tn}} \ .$$

As a consequence, ETC can take a correct decision with probability at least $1 - \delta$ if the number of users n is greater than $\frac{32\sigma_r^2}{\Delta^2} \log \frac{2}{\delta}$ and then, if we denote by $\eta := \Delta - \sqrt{\frac{32\sigma_r^2}{n} \log(\frac{1}{\delta})}$, the decision will be taken before the time step

$$\frac{32\sigma_\varepsilon^2}{\eta^2} \log\Big(\frac{1}{\delta}\Big)(1 + o_\delta(1)).$$