Appendix A. Additional Related Works

In this section, we provide additional related works concerning adaptivity in statistics via Lepskii method and adaptivity in the case of subgaussian bandits.

A.1. Adaptivity via Lepskii Method

In Bhatt et al. (2022), authors provide a novel technique to extend Catoni's M-estimator (Catoni, 2012) to the infinite variance setting. In principle, their procedure relies on the knowledge of both ϵ and the centered moment v, however, they propose a strategy based on the Lepskii method (Lepskii, 1992) to adapt to unknown v. While the Lepskii method is a popular choice in the adaptive statistics literature, we point out how it requires an upper bound on the quantity to estimate. Indeed, this method can be safely applied when adapting to unknown ϵ (since it can be at most 1), but when it comes to u (or the centered moment v), requiring an upper bound makes the approach *not* fully adaptive.

A.2. Adaptivity in Subgaussian Bandits

In the literature of subgaussian stochastic bandits, σ (the subgaussian proxy) is usually assumed to be known by the agent. However, many works consider settings in which this quantity is unknown. In this section, we discuss standard approaches to adapt to σ (or estimate it) in subgaussian bandits, and show the additional difficulties implied by the heavy-tailed setting.

The main difference between σ and u is that the former can be estimated from data while guaranteeing strong convergence properties. In Audibert et al. (2009), authors introduce `UCB-V`, a variation of the well-known `UCB-1` algorithm capable of using a data-driven estimation of the variance while keeping optimal performance. As customary in most of the literature, rewards are assumed to be bounded in a known range. However, in heavy-tailed bandits, it is not possible to make such an assumption, and the estimation of u cannot be carried on. Other works try to relax the assumption of bounded rewards by the means of other assumptions, *e.g.,* a known upper bound on kurtosis (Lattimore, 2017), or Gaussian rewards (Cowan et al., 2018).

Without additional assumptions, dealing with both the unknown range of the rewards and unknown σ comes at a cost. As shown in Hadiji and Stoltz (2023), when the range of the rewards is unknown and no additional knowledge on the distributions is available, it is impossible to be simultaneously optimal in both the instance-dependent sense and the worst-case one. The existence of such a trade-off shows how difficult is, even in subgaussian bandits, to attain optimal performances when no knowledge is given on the environment. As a consequence, also in fully adaptive heavy-tailed bandits, such an impossibility result holds. However, as we have discussed, thanks to a specific assumption not involving ϵ nor u we can provide optimal regret guarantees in both cases.

Appendix B. Proofs and Derivations

In this section, we prove the main theoretical results outlined in the paper.

B.1. Lower Bounds

Theorem 2 (Minimax lower bound – u-adaptive) *Fix $\epsilon \in (0,1]$. For every algorithm $\texttt{Alg}$, sufficiently large learning horizon $T \in \mathbb{N}$, and number of arms $K \in \mathbb{N}_{\geq 2}$, it holds that:*

$$\sup_{u \geq 0} \sup_{\boldsymbol{\nu} \in \mathcal{P}_{HT}(\epsilon,u)^K} \frac{R_T(\texttt{Alg}, \boldsymbol{\nu})}{u^{\frac{1}{1+\epsilon}}} = +\infty. \tag{5}$$

More precisely, for every $u' \geq u \geq 0$, under the same conditions above, there exist two instances $\boldsymbol{\nu} \in \mathcal{P}_{HT}(\epsilon, u)$ and $\boldsymbol{\nu}' \in \mathcal{P}_{HT}(\epsilon, u')$ such that:

$$\max \left\{ \frac{R_T(\texttt{Alg}, \boldsymbol{\nu})}{u^{\frac{1}{1+\epsilon}}}, \frac{R_T(\texttt{Alg}, \boldsymbol{\nu}')}{(u')^{\frac{1}{1+\epsilon}}} \right\} \geq c_1 \left(\frac{u'}{u} \right)^{\frac{\epsilon}{(1+\epsilon)^2}} T^{\frac{1}{1+\epsilon}}, \tag{6}$$

where $c_1 > 0$ is a constant independent of u, u', and T.

Proof We start by constructing two heavy-tailed bandit instances with a common maximum order of moment ϵ, but where $u' \geq u$. We use δ_x to denote the Dirac delta distribution centered on x

Base instance

$$\boldsymbol{\nu} = \begin{cases} \nu_1 = \delta_0, \\ \nu_2 = \left(1 - \Delta^{1+\frac{1}{\epsilon}} u^{-\frac{1}{\epsilon}}\right) \delta_0 + \Delta^{1+\frac{1}{\epsilon}} u^{-\frac{1}{\epsilon}} \delta_{u^{\frac{1}{\epsilon}} \Delta^{-\frac{1}{\epsilon}}}, \end{cases} \tag{20}$$

where $\Delta \in (0, u^{\frac{1}{1+\epsilon}})$. Thus, we have $\mu_1 = 0$ and $\mu_2 = \Delta$. Furthermore, $\mathbb{E}_{X \sim \nu_1}[|X|^{1+\epsilon}] = 0$ and $\mathbb{E}_{X \sim \nu_2}[|X|^{1+\epsilon}] = u$ Therefore, the optimal arm is arm 2 and $\boldsymbol{\nu} \in \mathcal{P}(\epsilon, u)^2$.

Alternative instance

$$\boldsymbol{\nu}' = \begin{cases} \nu_1' = \left(1 - (2\Delta)^{1+\frac{1}{\epsilon}} (u')^{-\frac{1}{\epsilon}}\right) \delta_0 + (2\Delta)^{1+\frac{1}{\epsilon}} (u')^{-\frac{1}{\epsilon}} \delta_{(u')^{\frac{1}{\epsilon}} (2\Delta)^{-\frac{1}{\epsilon}}}, \\ \nu_2' = \nu_2, \end{cases} \tag{21}$$

where $\Delta \in (0, \frac{1}{2}(u')^{\frac{1}{1+\epsilon}})$. Thus we have $\mu_1' = 2\Delta$ and $\mu_2' = \Delta$. Furthermore, $\mathbb{E}_{X \sim \nu_1'}[|X|^{1+\epsilon}] = u'$ and $\mathbb{E}_{X \sim \nu_2'}[|X|^{1+\epsilon}] = u$. Therefore, the optimal arm is arm 1 and $\boldsymbol{\nu} \in \mathcal{P}(\epsilon, u')^2$.

We seek to prove that for any algorithm $\texttt{Alg}$, it holds that:

$$\max \left\{ \frac{R_T(\texttt{Alg}, \boldsymbol{\nu})}{(uT)^{\frac{1}{1+\epsilon}}}, \frac{R_T(\texttt{Alg}, \boldsymbol{\nu}')}{(u'T)^{\frac{1}{1+\epsilon}}} \right\} \geq f(T, \epsilon, u, u'),$$

being f a function increasing in T. The proof merges the approach of (Bubeck et al., 2013b, Theorem 5) with that of (Lattimore and Szepesvári, 2020, Chapters 14.2, 14.3).

First, we observe that:

$$\max \left\{ \frac{R_T(\texttt{Alg}, \boldsymbol{\nu})}{(uT)^{\frac{1}{1+\epsilon}}}, \frac{R_T(\texttt{Alg}, \boldsymbol{\nu}')}{(u'T)^{\frac{1}{1+\epsilon}}} \right\} \geq \frac{R_T(\texttt{Alg}, \boldsymbol{\nu})}{(uT)^{\frac{1}{1+\epsilon}}} = \frac{\Delta \mathbb{E}_{\texttt{Alg}, \boldsymbol{\nu}}[N_1(T)]}{(uT)^{\frac{1}{1+\epsilon}}}, \tag{22}$$

where $\mathbb{E}_{\texttt{Alg},\boldsymbol{\nu}}[N_1(T)]$ is the expected number of times arm 1 is pulled over the horizon T. Second, recalling which are the optimal arms in the two instances and that $u' \geq u$, we have:

$$
\begin{aligned}
\max \left\{ \frac{R_T(\texttt{Alg}, \boldsymbol{\nu})}{(uT)^{\frac{1}{1+\epsilon}}}, \frac{R_T(\texttt{Alg}, \boldsymbol{\nu}')}{(u'T)^{\frac{1}{1+\epsilon}}} \right\} \geq & \\
& \geq (u'T)^{-\frac{1}{\epsilon+1}} \frac{\Delta T}{2} \max \left\{ \mathbb{P}_{\texttt{Alg},\boldsymbol{\nu}}\left(N_1(T) \geq T/2\right), \mathbb{P}_{\texttt{Alg},\boldsymbol{\nu}'}\left(N_1(T) < T/2\right) \right\} \\
& \geq \frac{\Delta}{4} (u')^{-\frac{1}{\epsilon+1}} T^{\frac{\epsilon}{\epsilon+1}} \left(\mathbb{P}_{\texttt{Alg},\boldsymbol{\nu}}\left(N_1(T) \geq T/2\right) + \mathbb{P}_{\texttt{Alg},\boldsymbol{\nu}'}\left(N_1(T) < T/2\right) \right) \\
& \geq \frac{\Delta}{8} (u')^{-\frac{1}{\epsilon+1}} T^{\frac{\epsilon}{\epsilon+1}} \exp\left(-\mathbb{E}_{\texttt{Alg},\boldsymbol{\nu}}[N_1(T)] D_{\text{KL}}(\nu_1 \| \nu_1') \right).
\end{aligned} \tag{23}
$$

where we used Bretagnolle-Huber inequality and divergence decomposition, together with $\max\{a, b\} \geq \frac{1}{2}(a+b)$ for $a, b \geq 0$. Let us now compute the KL-divergence, noting that $\nu_1 \ll \nu_1'$:

$$
\begin{aligned}
D_{\text{KL}}(\nu_1 \| \nu_1') &= \nu_1(0) \log \frac{\nu_1(0)}{\nu_1'(0)} \\
&= \log \frac{1}{1 - (2\Delta)^{1+\frac{1}{\epsilon}} (u')^{-\frac{1}{\epsilon}}} \leq c (2\Delta)^{1+\frac{1}{\epsilon}} (u')^{-\frac{1}{\epsilon}},
\end{aligned} \tag{24}
$$

for $\Delta \in (0, \left(\frac{1}{2}\right)^{\frac{2\epsilon+1}{1+\epsilon}} (u')^{\frac{1}{1+\epsilon}})$ and some constant $c \in (1, 2)$. Putting together Equations (22), (23) and (24), we have:

$$
\begin{aligned}
\max \left\{ \frac{R_T(\texttt{Alg}, \boldsymbol{\nu})}{(uT)^{\frac{1}{1+\epsilon}}}, \frac{R_T(\texttt{Alg}, \boldsymbol{\nu}')}{(u'T)^{\frac{1}{1+\epsilon}}} \right\} \geq & \\
& \geq \max \left\{ \frac{\Delta \mathbb{E}_{\texttt{Alg},\boldsymbol{\nu}}[N_1(T)]}{(uT)^{\frac{1}{1+\epsilon}}}, \frac{\Delta}{8} (u')^{-\frac{1}{\epsilon+1}} T^{\frac{\epsilon}{\epsilon+1}} \exp\left(-c \mathbb{E}_{\texttt{Alg},\boldsymbol{\nu}}[N_1(T)] (2\Delta)^{1+\frac{1}{\epsilon}} (u')^{-\frac{1}{\epsilon}} \right) \right\} \\
& \geq \frac{\Delta}{2} \left(\frac{\mathbb{E}_{\texttt{Alg},\boldsymbol{\nu}}[N_1(T)]}{(uT)^{\frac{1}{1+\epsilon}}} + \frac{1}{8} (u')^{-\frac{1}{\epsilon+1}} T^{\frac{\epsilon}{\epsilon+1}} \exp\left(-c \mathbb{E}_{\texttt{Alg},\boldsymbol{\nu}}[N_1(T)] (2\Delta)^{\frac{1+\epsilon}{\epsilon}} (u')^{-\frac{1}{\epsilon}} \right) \right) \\
& \geq \frac{\Delta}{2} \min_{x \in [0,T]} \left\{ \frac{x}{(uT)^{\frac{1}{1+\epsilon}}} + \frac{1}{8} (u')^{-\frac{1}{\epsilon+1}} T^{\frac{\epsilon}{\epsilon+1}} \exp\left(-cx (2\Delta)^{\frac{1+\epsilon}{\epsilon}} (u')^{-\frac{1}{\epsilon}} \right) \right\} =: g(x)
\end{aligned}
$$

The latter is a convex function of x and the minimization can be carried out in closed form, vanishing the derivative and finding:

$$
x^* = c^{-1} (2\Delta)^{-\frac{1+\epsilon}{\epsilon}} (u')^{\frac{1}{\epsilon}} \log \left(\frac{T u^{\frac{1}{\epsilon+1}}}{8 (u')^{\frac{1}{\epsilon} + \frac{1}{\epsilon+1}}} c (2\Delta)^{\frac{1+\epsilon}{\epsilon}} \right),
$$

which leads to:

$$
g(x^*) = \frac{\Delta}{2} (uT)^{-\frac{1}{\epsilon+1}} c^{-1} (2\Delta)^{-\frac{1+\epsilon}{\epsilon}} (u')^{\frac{1}{\epsilon}} \log \left(\frac{T u^{\frac{1}{\epsilon+1}}}{8 (u')^{\frac{1}{\epsilon} + \frac{1}{\epsilon+1}}} e c (2\Delta)^{\frac{1+\epsilon}{\epsilon}} \right).
$$

We choose Δ such that:

$$\frac{Tu^{\frac{1}{\epsilon+1}}}{8(u')^{\frac{1}{\epsilon}+\frac{1}{\epsilon+1}}}c(2\Delta)^{\frac{1+\epsilon}{\epsilon}} = e^{\epsilon},$$

resulting in $\Delta = 2^{\frac{2\epsilon-1}{1+\epsilon}} e^{\frac{\epsilon^2}{1+\epsilon}} (cT)^{-\frac{\epsilon}{\epsilon+1}} u^{-\frac{\epsilon}{(\epsilon+1)^2}} (u')^{\frac{1+2\epsilon}{(\epsilon+1)^2}}$. This implies, after some calculations, that:

$$g(x^*) = c^{-\frac{\epsilon}{\epsilon+1}} 2^{-\frac{2\epsilon+5}{\epsilon+1}} (1+\epsilon) e^{-\frac{\epsilon}{\epsilon+1}} u^{-\frac{\epsilon}{(\epsilon+1)^2}} (u')^{\frac{\epsilon}{(\epsilon+1)^2}} \geq c_1 \left(\frac{u'}{u}\right)^{\frac{\epsilon}{(\epsilon+1)^2}},$$

where $c_1 > 0$ is a value independent of T and both u and u'. Finally, we have that

$$\max\left\{\frac{R_T(\texttt{Alg},\boldsymbol{\nu})}{(uT)^{\frac{1}{1+\epsilon}}}, \frac{R_T(\texttt{Alg},\boldsymbol{\nu}')}{(u'T)^{\frac{1}{1+\epsilon}}}\right\} \geq c_1 \left(\frac{u'}{u}\right)^{\frac{\epsilon}{(\epsilon+1)^2}}.$$

We observe that $\Delta < \left(\frac{1}{2}\right)^{\frac{2\epsilon+1}{1+\epsilon}} (u')^{\frac{1}{1+\epsilon}}$ for sufficiently large T. This concludes the proof of the second statement. For the first statement, we observe that, since $u' \geq u$ can be taken arbitrarily large, the right-hand side of this inequality can be arbitrarily large. ■

Theorem 3 (Minimax lower bound – ϵ-adaptive) *Fix $u = 1$. For every algorithm $\texttt{Alg}$, sufficiently large learning horizon $T \in \mathbb{N}$, and number of arms $K \in \mathbb{N}_{\geq 0}$, it holds that:*

$$\sup_{\epsilon\in(0,1]} \sup_{\boldsymbol{\nu}\in\mathcal{P}_{HT}(\epsilon,u)^K} \frac{R_T(\texttt{Alg},\boldsymbol{\nu})}{T^{\frac{1}{1+\epsilon}}} \geq c_2 T^{\frac{1}{16}}. \tag{7}$$

More precisely, for every $\epsilon, \epsilon' \in (0,1]$ with $\epsilon' \leq \epsilon$, under the same conditions above, there exist two instances $\boldsymbol{\nu} \in \mathcal{P}_{HT}(\epsilon,u)$ and $\boldsymbol{\nu}' \in \mathcal{P}_{HT}(\epsilon',u)$ such that:

$$\max\left\{\frac{R_T(\texttt{Alg},\boldsymbol{\nu})}{T^{\frac{1}{1+\epsilon}}}, \frac{R_T(\texttt{Alg},\boldsymbol{\nu}')}{T^{\frac{1}{1+\epsilon'}}}\right\} \geq c_2 T^{\frac{\epsilon'(\epsilon-\epsilon')}{(1+\epsilon)(1+\epsilon')^2}}, \tag{8}$$

where $c_2 > 0$ is a constant independent of ϵ, ϵ', and T.

Proof We start by constructing two heavy-tailed bandit instances with different maximum orders of moment ϵ and ϵ', where $0 < \epsilon' < \epsilon < 1$. For the sake of simplicity, but without loss of generality, we will assume a common (and known to the algorithm) maximum moment of $u = 1$.

Base instance

$$\boldsymbol{\nu} = \begin{cases} \nu_1 = \delta_0, \\ \nu_2 = (1 + \Delta\gamma - \gamma^{1+\epsilon})\delta_0 + (\gamma^{1+\epsilon} - \Delta\gamma)\delta_{1/\gamma} \end{cases}, \tag{25}$$

where $\Delta \in [0, 1/2]$ and $\gamma = (2\Delta)^{\frac{1}{\epsilon}}$. Thus, we have $\mu_1 = 0$ and $\mu_2 = \Delta$. Furthermore, $\mathbb{E}_{X\sim\nu_1}[|X|^{\alpha}] = 0$ and $\mathbb{E}_{X\sim\nu_2}[|X|^{\alpha}] = 2^{\frac{1-\alpha}{\epsilon}} \Delta^{\frac{1+\epsilon-\alpha}{\epsilon}}$, which are guaranteed to be bounded by a constant smaller than 1 only if $\alpha \leq \epsilon + 1$. Thus, this instance admits moments finite only up to order $\epsilon + 1$, *i.e.*, $\boldsymbol{\nu} \in \mathcal{P}(\epsilon, 1)^2$. Moreover, the optimal arm is arm 2.

Alternative instance

$$\boldsymbol{\nu}' = \begin{cases} \nu_1' = (1-(\gamma')^{1+\epsilon'})\delta_0 + (\gamma')^{1+\epsilon'}\delta_{1/\gamma'}, \\ \nu_2' = \nu_2 \end{cases}, \tag{26}$$

where $\Delta \in [0, 1/2]$ and $\gamma' = (2\Delta)^{\frac{1}{\epsilon'}}$. Thus, we have $\mu_1' = 2\Delta$ and $\mu_2' = \Delta$. Furthermore, $\mathbb{E}_{X\sim\nu_1'}[|x|^\alpha] = (2\Delta)^{\frac{1+\epsilon'-\alpha}{\epsilon'}}$ and $\mathbb{E}_{X\sim\nu_2'}[|x|^\alpha] = 2^{\frac{1-\alpha}{\epsilon}}\Delta^{\frac{1+\epsilon-\alpha}{\epsilon}}$, which are guaranteed to be bounded by a constant smaller than 1 only if $\alpha \le \epsilon' + 1$. Thus, this instance admits moments finite only up to order $\epsilon'+1$, *i.e.,* $\boldsymbol{\nu} \in \mathcal{P}(\epsilon', 1)^2$. Moreover, the optimal arm is arm 1.

We will prove, that for any algorithm $\texttt{Alg}$ it holds that:

$$\max\left\{\frac{R_T(\texttt{Alg}, \boldsymbol{\nu})}{T^{\frac{1}{1+\epsilon}}}, \frac{R_T(\texttt{Alg}, \boldsymbol{\nu}')}{T^{\frac{1}{1+\epsilon'}}}\right\} \ge f(T, \epsilon, \epsilon'),$$

being f a function increasing in T. The proof emulates the analyses and steps performed to prove Theorem 2. First, we observe that:

$$\max\left\{\frac{R_T(\texttt{Alg}, \boldsymbol{\nu})}{T^{\frac{1}{1+\epsilon}}}, \frac{R_T(\texttt{Alg}, \boldsymbol{\nu}')}{T^{\frac{1}{1+\epsilon'}}}\right\} \ge \frac{R_T(\texttt{Alg}, \boldsymbol{\nu})}{T^{\frac{1}{1+\epsilon}}} = \frac{\Delta \mathbb{E}_{\texttt{Alg},\boldsymbol{\nu}}[N_1(T)]}{T^{\frac{1}{1+\epsilon}}}, \tag{27}$$

where $\mathbb{E}_{\texttt{Alg},\boldsymbol{\nu}}[N_1(T)]$ is the expected number of times arm 1 is pulled over the horizon T.

Second, recalling which are the optimal arms in the two instances and that $\epsilon' < \epsilon$, we have:

$$\begin{aligned} \max&\left\{\frac{R_T(\texttt{Alg}, \boldsymbol{\nu})}{T^{\frac{1}{1+\epsilon}}}, \frac{R_T(\texttt{Alg}, \boldsymbol{\nu}')}{T^{\frac{1}{1+\epsilon'}}}\right\} \ge \\ &\ge T^{-\frac{1}{\epsilon'+1}} \max\left\{\frac{\Delta T}{2}\mathbb{P}_{\texttt{Alg},\boldsymbol{\nu}}\left(N_1(T) \ge \frac{T}{2}\right), \frac{\Delta T}{2}\mathbb{P}_{\texttt{Alg},\boldsymbol{\nu}'}\left(N_1(T) < \frac{T}{2}\right)\right\} \\ &\ge \frac{\Delta}{4} T^{\frac{\epsilon'}{\epsilon'+1}}\left(\mathbb{P}_{\texttt{Alg},\boldsymbol{\nu}}\left(N_1(T) \ge \frac{T}{2}\right) + \mathbb{P}_{\texttt{Alg},\boldsymbol{\nu}'}\left(N_1(T) < \frac{T}{2}\right)\right) \\ &\ge \frac{\Delta}{8} T^{\frac{\epsilon'}{\epsilon'+1}} \exp\left(-\mathbb{E}_{\texttt{Alg},\boldsymbol{\nu}}[N_1(T)] D_{\text{KL}}(\nu_1 \| \nu_1')\right). \end{aligned} \tag{28}$$

where we used Bretagnolle-Huber inequality and divergence decomposition, together with $\max\{a, b\} \ge \frac{1}{2}(a+b)$ for $a, b \ge 0$. Let us now compute the KL-divergence, noting that $\nu_1 \ll \nu_1'$:

$$\begin{aligned} D_{\text{KL}}(\nu_1 \| \nu_1') &= \nu_1(0) \log \frac{\nu_1(0)}{\nu_1'(0)} \\ &= \log \frac{1}{1-(2\Delta)^{\frac{1+\epsilon'}{\epsilon'}}} \le c(2\Delta)^{\frac{1+\epsilon'}{\epsilon'}}, \end{aligned} \tag{29}$$

for $\Delta \in [0, 1/4]$ and some constant $c \in (1, 2)$. Putting together Equations (27), (28) and (29), we have:

$$\begin{aligned} \max&\left\{\frac{R_T(\texttt{Alg}, \boldsymbol{\nu})}{T^{\frac{1}{1+\epsilon}}}, \frac{R_T(\texttt{Alg}, \boldsymbol{\nu}')}{T^{\frac{1}{1+\epsilon'}}}\right\} \\ &\ge \max\left\{\frac{\Delta \mathbb{E}_{\texttt{Alg},\boldsymbol{\nu}}[N_1(T)]}{T^{\frac{1}{1+\epsilon}}}, \frac{\Delta}{8} T^{\frac{\epsilon'}{\epsilon'+1}} \exp\left(-c\mathbb{E}_{\texttt{Alg},\boldsymbol{\nu}}[N_1(T)](2\Delta)^{\frac{1+\epsilon'}{\epsilon'}}\right)\right\} \end{aligned}$$

$$\geq \frac{\Delta}{2}\left(\frac{\mathbb{E}[N_1(T)]}{T^{\frac{1}{1+\epsilon}}} + \frac{1}{8}T^{\frac{\epsilon'}{\epsilon'+1}} \exp\left(-c\mathbb{E}[N_1(T)](2\Delta)^{\frac{1+\epsilon'}{\epsilon'}}\right)\right)$$
$$\geq \frac{\Delta}{2}\min_{x\in[0,T]}\left\{\frac{x}{T^{\frac{1}{1+\epsilon}}} + \frac{1}{8}T^{\frac{\epsilon'}{\epsilon'+1}} \exp\left(-cx(2\Delta)^{\frac{1+\epsilon'}{\epsilon'}}\right)\right\} =: g(x).$$

The latter is a convex function of x and the minimization can be carried out in closed form vanishing the derivative and obtaining:

$$x^* = c^{-1}(2\Delta)^{-\frac{1+\epsilon'}{\epsilon'}} \log\left(\frac{T^{\frac{1}{\epsilon+1}+\frac{\epsilon'}{1+\epsilon'}}}{8}c(2\Delta)^{\frac{1+\epsilon'}{\epsilon'}}\right),$$

which leads to:

$$g(x^*) = \frac{\Delta}{2}T^{-\frac{1}{\epsilon+1}}c^{-1}(2\Delta)^{-\frac{1+\epsilon'}{\epsilon'}} \log\left(\frac{T^{\frac{1}{\epsilon+1}+\frac{\epsilon'}{1+\epsilon'}}}{8}ec(2\Delta)^{\frac{1+\epsilon'}{\epsilon'}}\right).$$

We take Δ such that:

$$\frac{T^{\frac{1}{\epsilon+1}+\frac{\epsilon'}{1+\epsilon'}}}{8}c(2\Delta)^{\frac{1+\epsilon'}{\epsilon'}} = 1,$$

resulting in $\Delta = 2^{\frac{2\epsilon'-1}{1+\epsilon'}}c^{-\frac{\epsilon'}{1+\epsilon'}}T^{-\frac{\epsilon'}{1+\epsilon'}\left(\frac{1}{\epsilon+1}+\frac{\epsilon'}{1+\epsilon'}\right)}$. This imply, after some calculations, that:

$$g(x^*) = 2^{\frac{-2\epsilon'-5}{1+\epsilon'}}c^{-\frac{\epsilon'}{1+\epsilon'}}T^{\frac{\epsilon'(\epsilon-\epsilon')}{(1+\epsilon')^2(1+\epsilon)}} \geq c_2 T^{\frac{\epsilon'(\epsilon-\epsilon')}{(1+\epsilon')^2(1+\epsilon)}}.$$

where $c_2 > 0$ is a value independent of T and can be always selected to be ϵ and ϵ'. Finally, we have that:

$$\max\left\{\frac{R_T(\mathtt{Alg},\boldsymbol{\nu})}{T^{\frac{1}{1+\epsilon}}}, \frac{R_T(\mathtt{Alg},\boldsymbol{\nu}')}{T^{\frac{1}{1+\epsilon'}}}\right\} \geq c_2 T^{\frac{\epsilon'(\epsilon-\epsilon')}{(1+\epsilon')^2(1+\epsilon)}}.$$

We observe that $\Delta < 1/4$ for sufficiently large T. We conclude by observing that the exponent of T is maximized by taking $\epsilon = 1$ and $\epsilon' = 1/3$. ■

Theorem 4 (Minimax lower bound under Assumption 1 - non-adaptive) *Fix $\epsilon \in (0,1]$ and $u \geq 0$. For every algorithm $\mathtt{Alg}$, sufficiently large learning horizon $T \in \mathbb{N}$, and number of arms $K \in \mathbb{N}_{\geq 2}$, it holds that:*

$$\sup_{\substack{\boldsymbol{\nu}\in\mathcal{P}_{HT}(\epsilon,u)^K \\ \boldsymbol{\nu} \text{ fulfills Assumption } 1}} R_T(\mathtt{Alg},\boldsymbol{\nu}) \geq c_3 K^{\frac{\epsilon}{1+\epsilon}}(uT)^{\frac{1}{1+\epsilon}}, \tag{9}$$

where $c_3 > 0$ is a constant independent of u, ϵ, K and T.

Proof We will construct instances using the following prototype of reward distribution, defined for $y \in (0, u^{\frac{1}{1+\epsilon}})$ and $\Delta \in (0, u^{\frac{1}{1+\epsilon}})$:

$$\rho_y = \left(1 - y^{1+\frac{1}{\epsilon}}u^{-\frac{1}{\epsilon}}\right)\delta_0 + \left(y^{1+\frac{1}{\epsilon}}u^{-\frac{1}{\epsilon}}\right)\delta_{-u^{\frac{1}{\epsilon}}\Delta^{-\frac{1}{\epsilon}}}. \tag{30}$$

The two instances are constructed by the means of Equation (30). Note that we have:

$$\mathop{\mathbb{E}}_{X\sim\rho_y}[X] = -y^{1+\frac{1}{\epsilon}}\Delta^{-\frac{1}{\epsilon}}, \tag{31}$$

$$\mathop{\mathbb{E}}_{X\sim\rho_y}[|X|^{1+\epsilon}] = y^{1+\frac{1}{\epsilon}}\Delta^{-1-\frac{1}{\epsilon}}u \le u, \tag{32}$$

for every $0 \le y \le \Delta$.

Base instance

$$\boldsymbol{\nu} = \begin{cases} \nu_1 = \rho_{\left(\frac{2}{3}\right)^{\frac{\epsilon}{1+\epsilon}}\Delta}, & \\ \nu_j = \rho_\Delta, & j \in [K]\setminus\{1\}. \end{cases}$$

Alternative instance

$$\boldsymbol{\nu}' = \begin{cases} \nu_1' = \rho_{\left(\frac{2}{3}\right)^{\frac{\epsilon}{1+\epsilon}}\Delta}, & \\ \nu_i' = \rho_{\left(\frac{1}{3}\right)^{\frac{\epsilon}{1+\epsilon}}\Delta}, & \\ \nu_j' = \rho_\Delta, & j \in [K]\setminus\{1, i\}, \end{cases}$$

where $i \in \operatorname{argmin}_{j\neq 1}\mathbb{E}_{\texttt{Alg},\boldsymbol{\nu}'}[N_j(T)]$. For the base instance, we have $\mu_1 = -2\Delta/3$ and $\mu_j = -\Delta$ for all $j \neq 1$; whereas for the alternative instance $\mu_j' = \mu_j$ for all $j \neq i$ and $\mu_i' = -\Delta/3$. Both instances satisfy Assumption 1, being the support a subset made of non-positive numbers. Moreover, for the base instance, the optimal arm is 1 and for the alternative instance, the optimal arm is i. Using the Bretagnolle-Huber inequality, we obtain:

$$\begin{aligned} R_T(\texttt{Alg},\boldsymbol{\nu}) + R_T(\texttt{Alg},\boldsymbol{\nu}') &\ge \frac{\Delta T}{6}\left(\mathbb{P}_{\texttt{Alg},\boldsymbol{\nu}}\left(N_1 \le \frac{T}{2}\right) + \mathbb{P}_{\texttt{Alg},\boldsymbol{\nu}'}\left(N_1 > \frac{T}{2}\right)\right) \\ &\ge \frac{\Delta T}{6}\exp\left(-\mathop{\mathbb{E}}_{\texttt{Alg},\boldsymbol{\nu}}[N_i(T)]D_{\mathrm{KL}}(\nu_i||\nu_i')\right) \end{aligned}$$

We recall that by the definition of i, we have that $\mathbb{E}_{\texttt{Alg},\boldsymbol{\nu}}[N_i(T)] \le \frac{T}{K-1}$. We now compute the Kullback-Leibler divergence between the two instances:

$$\begin{aligned} D_{\mathrm{KL}}(\nu_i||\nu_i') &= \Delta^{1+\frac{1}{\epsilon}}u^{-\frac{1}{\epsilon}}\log\left(\frac{\Delta^{1+\frac{1}{\epsilon}}u^{-\frac{1}{\epsilon}}}{\frac{1}{3}\Delta^{1+\frac{1}{\epsilon}}u^{-\frac{1}{\epsilon}}}\right) + \underbrace{(1-\Delta^{1+\frac{1}{\epsilon}}u^{-\frac{1}{\epsilon}})\log\left(\frac{1-\Delta^{1+\frac{1}{\epsilon}}u^{-\frac{1}{\epsilon}}}{1-\frac{1}{3}\Delta^{1+\frac{1}{\epsilon}}u^{-\frac{1}{\epsilon}}}\right)}_{\le 0} \\ &\le \Delta^{1+\frac{1}{\epsilon}}u^{-\frac{1}{\epsilon}}\log 3. \end{aligned}$$

Plugging this result, we finally get:

$$R_T(\texttt{Alg},\boldsymbol{\nu}) + R_T(\texttt{Alg},\boldsymbol{\nu}') \ge \frac{\Delta T}{6}\exp\left(-\frac{T}{K-1}\Delta^{1+\frac{1}{\epsilon}}u^{-\frac{1}{\epsilon}}\log 3\right).$$

We conclude the proof by noting that $\max\{x,y\} > \frac{1}{2}(x+y)$ and setting $\Delta = \frac{1}{2}\left(\frac{K-1}{T}u^{\frac{1}{\epsilon}}\frac{1}{\log 3}\right)^{\frac{\epsilon}{1+\epsilon}}$. Finally, we have:

$$\max\{R_T(\texttt{Alg},\boldsymbol{\nu}), R_T(\texttt{Alg},\boldsymbol{\nu}')\} \geq c_3 K^{\frac{\epsilon}{1+\epsilon}}(uT)^{\frac{1}{1+\epsilon}},$$

for some constant $c_3 > 0$ independent of T, u, ϵ and K. ■

B.2. Estimator

Lemma 2 (**(ϵ, u)-free Upper Confidence Bound**) *Let $\delta \in (0, 1/2)$ and $\mathbf{X} = \{X_1, \dots, X_s\}$ be a set of $s \in \mathbb{N}_{\geq 2}$ i.i.d. random variables satisfying $X_1 \sim \nu \in \mathcal{P}_{HT}(\epsilon, u)$, $\mu := \mathbb{E}[X_1]$, and $M > 0$ be a (possibly random) trimming threshold independent of $\mathbf{X}$. Then, under Assumption 1, it holds that:*

$$\mathbb{P}\Big(\mu - \widehat{\mu}_s(\mathbf{X}; M) \leq \sqrt{\frac{2V_s(\mathbf{X}; M)\log\delta^{-1}}{s}} + \frac{10M\log\delta^{-1}}{s}\Big) \geq 1 - 2\delta, \tag{11}$$

where $V_s(\mathbf{X}; M)$ is the sample variance of the trimmed random variables, defined as:

$$V_s(\mathbf{X}; M) := \frac{1}{s-1}\sum_{j\in[s]}(X_j \mathbb{1}_{\{|X_j|\leq M\}} - \widehat{\mu}_s(\mathbf{X}; M))^2. \tag{12}$$

Proof Since M is computed independently of $\mathbf{X}$, the trimmed samples $X_i\mathbb{1}_{\{|X_i|\leq M\}}$ remain independent. Thus, with probability at least $1-\delta$, we have:

$$\begin{aligned}
\mu - \widehat{\mu}_s(\mathbf{X}; M) &= \mathbb{E}[X_1] - \frac{1}{s}\sum_{t=1}^{s} X_t \mathbb{1}_{|X_t|\leq M} \\
&= \frac{1}{s}\sum_{t=1}^{s}\left(\mathbb{E}[X_1] - \mathbb{E}\left[X_t\mathbb{1}_{|X_t|\leq M}\right]\right) + \frac{1}{s}\sum_{t=1}^{s}\left(\mathbb{E}\left[X_t\mathbb{1}_{|X_t|\leq M}\right] - X_t\mathbb{1}_{|X_t|\leq M}\right) \\
&= \frac{1}{s}\sum_{t=1}^{s}\mathbb{E}[X_t\mathbb{1}_{|X_t|>M}] + \frac{1}{s}\sum_{t=1}^{s}\left(\mathbb{E}\left[X_t\mathbb{1}_{|X_t|\leq M}\right] - X_t\mathbb{1}_{|X_t|\leq M}\right) \\
&\overset{(*)}{\leq} \frac{1}{s}\sum_{t=1}^{s}\left(\mathbb{E}\left[X_t\mathbb{1}_{|X_t|\leq M}\right] - X_t\mathbb{1}_{|X_t|\leq M}\right) \\
&\overset{(**)}{\leq} \sqrt{\frac{2V_s(\mathbf{Y})\log 2\delta^{-1}}{s}} + \frac{14M\log 2\delta^{-1}}{3(s-1)} \\
&\leq \sqrt{\frac{2V_s(\mathbf{Y})\log 2\delta^{-1}}{s}} + \frac{10M\log 2\delta^{-1}}{s}
\end{aligned}$$

Note that in step $(*)$ we used Assumption 1 to make the first term vanish. In step $(**)$, instead, we used *empirical Bernstein inequality* (Maurer and Pontil, 2009) recalling that the trimmed random variables range in $[-M, M]$. We also use $\frac{1}{s-1} \leq \frac{2}{s}$ in the last step for $s \geq 2$. ■

Proposition 9 (Uniqueness of Solution of Equation (13), Wang et al. (2021)) *Let $\mathbf{X} = \{X_1, \ldots, X_s\}$ be a set of real numbers. If:*

$$0 < c \log \delta^{-1} < \sum_{j \in [s]} \mathbb{1}_{\{X_i \neq 0\}}, \tag{33}$$

then Equation (13) admits a unique positive solution.

Theorem 5 (Bounds on $\widehat{M}_s(\delta)$) *Let $\delta \in (0, 1/2)$ and $\mathbf{X}' = \{X_1', \ldots, X_s'\}$ be a set of $s \in \mathbb{N}_{\geq 1}$ i.i.d. random variables satisfying $X_1' \sim \nu \in \mathcal{P}_{HT}(\epsilon, u)$, and let $\widehat{M}_s(\delta)$ be the (random) positive root of Equation (13) with $c > 2$. Then, if $\widehat{M}_s(\delta)$ exists, with probability at least $1 - 2\delta$, it holds that:*

$$\widehat{M}_s(\delta) \leq \left(\frac{us}{(\sqrt{c} - \sqrt{2})^2 \log \delta^{-1}} \right)^{\frac{1}{1+\epsilon}} \quad \text{and} \quad \mathbb{P}\left(|X_1| > \widehat{M}_s(\delta) \right) \leq (\sqrt{c} + \sqrt{2})^2 \frac{\log \delta^{-1}}{s}. \tag{16}$$

Proof The proof makes use of the concentration inequality for self-bounding random variables (Maurer, 2006; Maurer and Pontil, 2009). Let $M > 0$, for every $i \in [\![s]\!]$, we define the random variable:

$$U_{i,M} \coloneqq \min \left\{ \left(\frac{X_i}{M} \right)^2, 1 \right\},$$

that ranges in $[0, 1]$. Furthermore, let: $Z_M(\mathbf{X}) \coloneqq \sum_{i=1}^{s} U_{i,M}$, ranging in $[0, s]$. Let us denote $\overline{U}_M(\mathbf{X}) \coloneqq Z_M(\mathbf{X})/s$, we observe that, given these definitions, the equation we want to solve for non-zero roots becomes:

$$\overline{U}_M(\mathbf{X}) - \frac{c \log \delta^{-1}}{s} = 0. \tag{34}$$

We start by showing that $Z_M(\mathbf{X})$ satisfies the assumptions of Theorem 13 of Maurer (2006), in particular, let $a \geq 1$, we have:

$$Z_M(\mathbf{X}) - \inf_{y \in \mathbb{R}} Z_M(\mathbf{X}_{y,k}) \leq 1, \quad \forall k \in [s], \tag{35}$$

$$\sum_{k=1}^{s} \left(Z_M(\mathbf{X}) - \inf_{y \in \mathbb{R}} Z_M(\mathbf{X}_{y,k}) \right)^2 \leq a Z_M(\mathbf{X}), \tag{36}$$

where $\mathbf{X}_{y,k}$ is obtained by replacing with y the k-th element X_k of the set $\mathbf{X}$. Indeed, Equation (35) follows as:

$$Z_M(\mathbf{X}) - \inf_{y \in \mathbb{R}} Z_M(\mathbf{X}_{y,k}) = U_{k,M} - \inf_{y \in \mathbb{R}} \min \left\{ \left(\frac{y}{M} \right)^2, 1 \right\} = U_{k,M} \leq 1, \quad \forall k \in [s].$$

Similarly, we set $a = 1$ and obtain Equation (36) as follows:

$$\begin{aligned}
\sum_{k=1}^{s} \left(Z_M(\mathbf{X}) - \inf_{y \in \mathbb{R}} Z_M(\mathbf{X}_{y,k}) \right)^2 &= \sum_{k=1}^{s} \left(U_{k,M} - \inf_{y \in \mathbb{R}} \min \left\{ \left(\frac{y}{M} \right)^2, 1 \right\} \right)^2 \\
&\leq \sum_{k=1}^{s} U_{k,M}^2
\end{aligned}$$

$$\leq \sum_{k=1}^{s} U_{k,M}$$
$$= Z_M(\mathbf{X}),$$

since $U_{k,M} \leq 1$. Using Theorem 13 from Maurer (2006) with $a = 1$, for the right tail of the distribution, we have for every $\epsilon > 0$:

$$\mathbb{P}\left(\mathbb{E}[Z_M(\mathbf{X})] - Z_M(\mathbf{X}) > s\epsilon\right) \leq \exp\left(\frac{-\epsilon^2 s^2}{2\mathbb{E}[Z_M(\mathbf{X})]}\right)$$

By letting $\epsilon = \sqrt{\frac{2\mathbb{E}[\overline{U}_M(\mathbf{X})] \log 2\delta^{-1}}{s}}$ and recalling the definition of $\overline{U}_M(\mathbf{X})$, we obtain:

$$\mathbb{P}\left(\mathbb{E}[\overline{U}_M(\mathbf{X})] - \overline{U}_M(\mathbf{X}) > \sqrt{\frac{2\mathbb{E}[\overline{U}_M(\mathbf{X})] \log \delta^{-1}}{s}}\right) \leq \delta,$$

which implies, after some algebraic manipulations (see Theorem 10 of (Maurer and Pontil, 2009)), the following:

$$\mathbb{P}\left(\sqrt{\mathbb{E}[\overline{U}_M(\mathbf{X})]} - \sqrt{\overline{U}_M(\mathbf{X})} > \sqrt{\frac{2\log \delta^{-1}}{s}}\right) \leq \delta.$$

A similar inequality holds for the left tail:

$$\mathbb{P}\left(Z_M(\mathbf{X}) - \mathbb{E}[Z_M(\mathbf{X})] > s\epsilon\right) \leq \exp\left(\frac{-\epsilon^2 s^2}{2\mathbb{E}[Z_M(\mathbf{X})] + \epsilon s}\right),$$

with similar steps, we obtain:

$$\mathbb{P}\left(\sqrt{\overline{U}_M(\mathbf{X})} - \sqrt{\mathbb{E}[\overline{U}_M(\mathbf{X})]} > \sqrt{\frac{2\log \delta^{-1}}{s}}\right) \leq \delta.$$

With a union bound over the two inequalities on the left and the right tail, we finally get:

$$\mathbb{P}\left(\left|\sqrt{\overline{U}_M(\mathbf{X})} - \sqrt{\mathbb{E}[\overline{U}_M(\mathbf{X})]}\right| > \sqrt{\frac{2\log \delta^{-1}}{s}}\right) \leq 2\delta. \tag{37}$$

Let us now define $\widehat{M}_s(\delta)$ random variable corresponding to the solution of the equation:

$$\overline{U}_{\widehat{M}_s(\delta)}(\mathbf{X}) = \frac{c \log \delta^{-1}}{s},$$

where $c > 0$. To control the bounds on $\widehat{M}$, we define the following auxiliary (non-random) quantities:

$$\sqrt{\overline{U}_M^{+}} := \sqrt{\mathbb{E}[\overline{U}_M(\mathbf{X})]} + \sqrt{\frac{2\log \delta^{-1}}{s}} \quad \text{and} \quad \sqrt{\overline{U}_M^{-}} := \sqrt{\mathbb{E}[\overline{U}_M(\mathbf{X})]} - \sqrt{\frac{2\log \delta^{-1}}{s}}. \tag{38}$$

Thanks to Equation 37, we have, for every $M \geq 0$, that $\mathbb{P}(\overline{U}_M^- \leq \overline{U}_M(\mathbf{X}) \leq \overline{U}_M^+) \geq 1 - 2\delta$. Furthermore, let $M^+(\delta), M^-(\delta) > 0$, the solutions of the following (non-random) equations:

$$U^+_{M^+(\delta)} = \frac{c \log \delta^{-1}}{s} \qquad \text{and} \qquad U^-_{M^-(\delta)} = \frac{c \log \delta^{-1}}{s}. \tag{39}$$

Since $\mathbb{P}(\overline{U}_M^- \leq \overline{U}_M(\mathbf{X}) \leq \overline{U}_M^+) \geq 1-2\delta$, it follows that $\mathbb{P}(M^-(\delta) \leq \widehat{M}_s(\delta) \leq M^+(\delta)) \geq 1-2\delta$. We now proceed at lower bounding $M^-(\delta)$ and upper bounding $M^+(\delta)$:

$$\begin{aligned}
\sqrt{\frac{c \log \delta^{-1}}{s}} &= \sqrt{U^-_{M^-(\delta)}} && (40)\\
&= \sqrt{\mathbb{E}[\overline{U}_{M^-(\delta)}(\mathbf{X})]} - \sqrt{\frac{2 \log \delta^{-1}}{s}} && (41)\\
&\geq \sqrt{\mathbb{P}\left(|X_1| \geq M^-(\delta)\right)} - \sqrt{\frac{2 \log \delta^{-1}}{s}} && (42)\\
&\geq \sqrt{\mathbb{P}\left(|X_1| \geq \widehat{M}_s(\delta)\right)} - \sqrt{\frac{2 \log \delta^{-1}}{s}}, && (43)
\end{aligned}$$

where the last but one inequality follows from:

$$\mathbb{E}[\overline{U}_M(\mathbf{X})] = \mathbb{E}\left[\min\left\{\left(\frac{X_1}{M}\right)^2, 1\right\}\right] \geq \mathbb{P}\left(\left(\frac{X_1}{M}\right)^2 \geq 1\right) = \mathbb{P}\left(|X_1| \geq M\right), \tag{44}$$

and the last inequality holds with probability $1 - \delta$ and follows from the fact that $\widehat{M}_s(\delta) \geq M^-(\delta)$. Similarly, we have:

$$\begin{aligned}
\sqrt{\frac{c \log \delta^{-1}}{s}} &= \sqrt{U^+_{M^+(\delta)}} && (45)\\
&= \sqrt{\mathbb{E}[\overline{U}_{M^-(\delta)}(\mathbf{X})]} + \sqrt{\frac{2 \log \delta^{-1}}{s}} && (46)\\
&\leq \sqrt{\frac{u}{(M^+(\delta))^{1+\epsilon}}} + \sqrt{\frac{2 \log \delta^{-1}}{s}} && (47)\\
&\leq \sqrt{\frac{u}{(\widehat{M}_s(\delta))^{1+\epsilon}}} + \sqrt{\frac{2 \log \delta^{-1}}{s}}, && (48)
\end{aligned}$$

where the last but one inequality follows from:

$$\mathbb{E}[\overline{U}_M(\mathbf{X})] = \mathbb{E}\left[\min\left\{\left(\frac{X_1}{M}\right)^2, 1\right\}\right] \leq M^{-1-\epsilon}\mathbb{E}\left[|X_1|^{1+\epsilon}\right] \leq M^{-1-\epsilon}u, \tag{49}$$

and the last inequality holds with probability $1 - \delta$ and follows from the fact that $\widehat{M}_s(\delta) \leq M^+(\delta)$. Thus, with probability $1 - 2\delta$, we have for $c > 2$:

$$\mathbb{P}\left(|X_1| > \widehat{M}_s(\delta)\right) \leq (\sqrt{c} + \sqrt{2})^2 \frac{\log \delta^{-1}}{s} \quad \text{and} \quad \widehat{M}_s(\delta) \leq \left(\frac{us}{(\sqrt{c} - \sqrt{2})^2 \log \delta^{-1}}\right)^{\frac{1}{1+\epsilon}}. \tag{50}$$

■

Theorem 6 **((ϵ, u)-dependent Concentration Bound)** *Let $\delta \in (0, 1/4)$, $\mathbf{X} = \{X_1, \dots, X_{s/2}\}$, and $\mathbf{X}' = \{X'_1, \dots, X'_{s/2}\}$ be two independent sets of $s/2 \in \mathbb{N}_{\geq 2}$ i.i.d. random variables satisfying $X_1 \sim \nu \in \mathcal{P}_{HT}(\epsilon, u)$, $\mu \coloneqq \mathbb{E}[X_1]$, and let $\widehat{M}_s(\delta)$ be the (random) positive root of Equation (13) with $c = (1+\sqrt{2})^2$. Then, if $\widehat{M}_s(\delta)$ exists, it holds that:*

$$\mathbb{P}\left(\left|\widehat{\mu}_s(\mathbf{X}; \widehat{M}_s(\delta)) - \mu\right| \leq 8u^{\frac{1}{1+\epsilon}} \left(\frac{\log \delta^{-1}}{s}\right)^{\frac{\epsilon}{1+\epsilon}} \right) \geq 1 - 4\delta. \tag{18}$$

Proof The result is obtained by combining an application of Bernstein's inequality and the bounds on the threshold $\widehat{M}_s(\delta)$ of Lemma 5. Furthermore since $\widehat{M}_s(\delta)$ is independent of $\mathbf{X}$, we can condition on the value of $\widehat{M}_s(\delta)$. With probability $1-\delta$, we have:

$$\begin{aligned}
&\widehat{\mu}_s(\mathbf{X}; \widehat{M}_s(\delta)) - \mu = \frac{1}{s}\sum_{i=1}^{s} X_i \mathbb{1}_{|X_i| \leq \widehat{M}_s(\delta)} - \mathbb{E}[X_1] \\
&\quad = \frac{1}{s}\sum_{i=1}^{s} \left(X_i \mathbb{1}_{|X_i| \leq \widehat{M}_s(\delta)} - \mathbb{E}\left[X_i \mathbb{1}_{|X_i| \leq \widehat{M}_s(\delta)}\right]\right) - \frac{1}{s}\sum_{i=1}^{s} \left(\mathbb{E}[X_1] - \mathbb{E}\left[X_t \mathbb{1}_{|X_i| \leq \widehat{M}_s(\delta)}\right]\right) \\
&\quad = \frac{1}{s}\sum_{i=1}^{s} \left(X_i \mathbb{1}_{|X_i| \leq \widehat{M}_s(\delta)} - \mathbb{E}\left[X_i \mathbb{1}_{|X_i| \leq \widehat{M}_s(\delta)}\right]\right) - \frac{1}{s}\sum_{i=1}^{s} \mathbb{E}[X_i \mathbb{1}_{|X_i| > \widehat{M}_s(\delta)}] \\
&\quad \leq \frac{1}{s}\sum_{i=1}^{s} \left(X_i \mathbb{1}_{|X_i| \leq \widehat{M}_s(\delta)} - \mathbb{E}\left[X_i \mathbb{1}_{|X_i| \leq \widehat{M}_s(\delta)}\right]\right) + \frac{1}{s}\sum_{i=1}^{s} \mathbb{E}[|X_i| \mathbb{1}_{|X_i| > \widehat{M}_s(\delta)}] \\
&\quad \overset{(*)}{\leq} \frac{1}{s}\sum_{i=1}^{s} \left(X_i \mathbb{1}_{|X_i| \leq \widehat{M}_s(\delta)} - \mathbb{E}\left[X_i \mathbb{1}_{|X_i| \leq \widehat{M}_s(\delta)}\right]\right) + \\
&\qquad + \frac{1}{s}\sum_{i=1}^{s} \left(\mathbb{E}\left[|X_i|^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}} \right) \left(\mathbb{E}\left[(\mathbb{1}_{|X_i| > \widehat{M}_s(\delta)})^{\frac{1+\epsilon}{\epsilon}} \right]^{\frac{\epsilon}{1+\epsilon}} \right) \\
&\quad \overset{(**)}{\leq} \sqrt{\frac{2\widehat{M}_s(\delta)^{1-\epsilon} u \log(\delta^{-1})}{s}} + \frac{\widehat{M}_s(\delta)\log(\delta^{-1})}{3s} + \frac{1}{s}\sum_{i=1}^{s} \left(u^{\frac{1}{1+\varepsilon}} \right) \left(\mathbb{E}\left[\mathbb{1}_{|X_i| > \widehat{M}_s(\delta)}\right]^{\frac{\epsilon}{1+\varepsilon}} \right) \\
&\quad \leq \sqrt{\frac{2\widehat{M}_s(\delta)^{1-\epsilon} u \log(\delta^{-1})}{s}} + \frac{\widehat{M}_s(\delta)\log(\delta^{-1})}{3s} + u^{\frac{1}{1+\varepsilon}} \mathbb{P}\left(|X_i| > \widehat{M}_s(\delta)\right)^{\frac{\epsilon}{1+\varepsilon}},
\end{aligned}$$

where step $(*)$ follows from Hölder inequality, while step $(**)$ is a consequence of Bernstein's inequality for bounded random variables. To proceed further, we use Lemma 5 in union bound with the previously applied inequality. Thus, with probability at least $1 - 3\delta$, we have:

$$\widehat{\mu}_s(\mathbf{X}; \widehat{M}_s(\delta)) - \mu \leq$$

$$
\begin{aligned}
&\le \sqrt{\frac{2\left(\frac{us}{(\sqrt{c}-\sqrt{2})^2 \log \delta^{-1}}\right)^{\frac{1-\epsilon}{1+\epsilon}} u \log\left(\delta^{-1}\right)}{s} + \frac{\left(\frac{us}{(\sqrt{c}-\sqrt{2})^2 \log \delta^{-1}}\right)^{\frac{1}{1+\epsilon}} \log\left(\delta^{-1}\right)}{3s}} \\
&\qquad\qquad + u^{\frac{1}{1+\varepsilon}} \left((\sqrt{c}+\sqrt{2})^2 \frac{\log \delta^{-1}}{s}\right)^{\frac{\epsilon}{1+\epsilon}} \\
&\le \left(\frac{\sqrt{2}}{(\sqrt{c}-\sqrt{2})^{\frac{1-\epsilon}{1+\epsilon}}} + \frac{1}{3(\sqrt{c}-\sqrt{2})^{\frac{2}{1+\epsilon}}} + (\sqrt{c}+\sqrt{2})^{\frac{2\epsilon}{1+\epsilon}}\right) u^{\frac{1}{1+\epsilon}} \left(\frac{\log \delta^{-1}}{n}\right)^{\frac{\epsilon}{1+\epsilon}} \\
&\le 5.6 u^{\frac{1}{1+\epsilon}} \left(\frac{\log \delta^{-1}}{s}\right)^{\frac{\epsilon}{1+\epsilon}},
\end{aligned}
$$

where in the last passage we set $c = (1+\sqrt{2})^2$ and bounded the resulting expression for $\epsilon \in (0, 1]$. A symmetric derivation leads to the second inequality. A union bound combined with renaming $s \leftarrow s/2$ and using $5.6\sqrt{2} \le 8$, concludes the proof. ■

Theorem 7 (Instance-Dependent Regret bound of `AdaR-UCB`) *Let $\boldsymbol{\nu} \in \mathcal{P}_{HT}(\epsilon, u)^K$ and $T \in \mathbb{N}_{\ge 2}$ be the learning horizon. Under Assumption 1,* `AdaR-UCB` *suffers a regret bounded as:*

$$
R_T(\texttt{AdaR-UCB}, \boldsymbol{\nu}) \le \sum_{i:\Delta_i > 0} \left[\left(120\left(\frac{u}{\Delta_i}\right)^{\frac{1}{\epsilon}} + \frac{24\Delta_i}{\mathbb{P}_{\nu_i}(X \ne 0)}\right) \log \frac{T}{2} + 20\Delta_i\right]. \tag{19}
$$

Proof For notational convenience, in this derivation, we will perform the substitution $T \leftarrow \lfloor T/2 \rfloor$ and $t \leftarrow \tau$. For every arm $i \in [K]$ and round $t \in [T]$, let us define the event:

$$
\mathcal{E}_{i,t} \coloneqq \left\{ \sum_{X \in \mathbf{X}'_i(t-1)} \mathbb{1}_{\{X \ne 0\}} - 4 \log t^3 > 0 \right\}. \tag{51}
$$

Under event $\mathcal{E}_{i,t}$ we do not incur in the forced exploration (FE) in line 4 ensuring that every arm has collected at least $4 \log t^3$ nonzero samples in $\mathbf{X}'_i$. Thus, we can decompose the expected number of pulls as follows:

$$
\begin{aligned}
\mathbb{E}[N_i^{\text{ALL}}(T)] &= \mathbb{E}\left[\sum_{t \in [T]} \mathbb{1}_{\{I_t = i \text{ and } \mathcal{E}_{i,t}\}}\right] + \mathbb{E}\left[\sum_{t \in [T]} \mathbb{1}_{\{I_t = i \text{ and } \mathcal{E}_{i,t}^{\complement}\}}\right] && (52) \\
&= \mathbb{E}[N_i(T)] + \mathbb{E}[N_i^{\text{FE}}(T)]. && (53)
\end{aligned}
$$

Part I: Bounding the expected number of pulls for forced exploration. We first bound the expected number of pulls $\mathbb{E}[N_i^{\text{FE}}(T)]$ due to the forced exploration. Considering only the samples collected due to forced exploration, thanks to independence among these samples, we can see the required number of pulls as a sum of geometric random variables. Thus, we can compute an upper bound on the expectation as:

$$\mathbb{E}_{\nu_i}[N_i^{\text{FE}}(T)] \le \frac{4\log T^3}{\mathbb{P}_{\nu_i}(|X|>0)}. \tag{54}$$

Part II: Bounding the expected number of pulls for optimistic exploration. We define for every arm $i \in [K]$ and every round $t \in [T]$, the upper confidence bound as:

$$B_i(t) = \widehat{\mu}_i(t) + \sqrt{\frac{2V_i(t)\log t^3}{N_i(t-1)}} + \frac{10\widehat{M}_i(t)\log t^3}{N_i(t-1)},$$

where $N_i(t-1)$ is the number of times arm i has been pulled up to time $t-1$, i.e., $N_i(t-1) = |\mathbf{X}_i(t-1)|$. We now show that if $I_t = i$, for an arm i such that $\Delta_i > 0$, then, one of the following four inequalities is true:

$$\text{either } B_1(t) \le \mu_1, \tag{55}$$

$$\text{or} \quad \widehat{\mu}_i(t) > \mu_i + 5.6u^{\frac{1}{1+\epsilon}}\left(\frac{\log t^3}{N_i(t-1)}\right)^{\frac{\epsilon}{1+\epsilon}}, \tag{56}$$

$$\text{or} \quad N_i(t-1) < 20\left(\frac{u}{\Delta_i^{1+\epsilon}}\right)^{\frac{1}{\epsilon}}\log t^3, \tag{57}$$

$$\text{or } \sqrt{V_i(t)} > \sqrt{\mathbb{E}[V_i(t)]} + 2\widehat{M}_i(t)\sqrt{\frac{\log t^3}{N_i(t-1)}}, \tag{58}$$

$$\text{or } \widehat{M}_i(t) \ge \left(\frac{uN_i(t-1)}{\log t^3}\right)^{\frac{1}{1+\epsilon}}. \tag{59}$$

Indeed, assume that all five inequalities are false. Then we have

$$\begin{aligned}
B_1(t) &\overset{(55)}{>} \mu_1 = \mu_i + \Delta_i \\
&\overset{(56)}{\ge} \widehat{\mu}_i(t) - 5.6u^{\frac{1}{1+\epsilon}}\left(\frac{\log t^3}{N_i(t-1)}\right)^{\frac{\epsilon}{1+\epsilon}} + \Delta_i \\
&\overset{(*)}{\ge} \widehat{\mu}_i(t) + \sqrt{\frac{2V_i(t)\log t^3}{N_i(t-1)}} + \frac{10\widehat{M}_i(t)\log t^3}{N_i(t-1)} \\
&= B_i(t).
\end{aligned}$$

The step marked with $(*)$ is a consequence of the fact that both (57), (58) and (59) are false. In particular, we need to show that

$$\Delta_i \ge 5.6u^{\frac{1}{1+\epsilon}}\left(\frac{\log t^3}{N_i(t-1)}\right)^{\frac{\epsilon}{1+\epsilon}} + \sqrt{\frac{2V_i(t)\log t^3}{N_i(t-1)}} + \frac{10\widehat{M}_i(t)\log t^3}{N_i(t-1)}. \tag{$*$}$$

To do so, we make use of the following inequality derived by exploiting the independence between $\mathbf{X}_i(t-1)$ and $\mathbf{X}'_i(t-1)$:

$$\mathbb{E}[V_i(t)] \le \mathbb{E}\left[X^2 \mathbb{1}_{|X| \le \widehat{M}_i(t)}\right] \le \mathbb{E}\left[|X|^{1+\epsilon}\right] \widehat{M}_i(t)^{1-\varepsilon} \le u\widehat{M}_i(t)^{1-\epsilon}. \tag{60}$$

Now, we make use of the fact that (57), (58), and (59) are false together with (60):

$$\begin{aligned}
\Delta_i &\overset{(57)}{\ge} 20u^{\frac{1}{1+\epsilon}} \left(\frac{\log t^3}{N_i(t-1)}\right)^{\frac{\epsilon}{1+\epsilon}} \\
&\ge (5.6 + \sqrt{2} + 10 + 2\sqrt{2}) u^{\frac{1}{1+\epsilon}} \left(\frac{\log t^3}{N_i(t-1)}\right)^{\frac{\epsilon}{1+\epsilon}} \\
&= 5.6u^{\frac{1}{1+\epsilon}} \left(\frac{\log t^3}{N_i(t-1)}\right)^{\frac{\epsilon}{1+\epsilon}} + \sqrt{\frac{2\log t^3 u \left(\frac{uN_i(t-1)}{\log t^3}\right)^{\frac{1-\epsilon}{1+\epsilon}}}{N_i(t-1)}} + \frac{(10+2\sqrt{2})\left(\frac{uN_i(t-1)}{\log t^3}\right)^{\frac{1}{1+\epsilon}} \log t^3}{N_i(t-1)} \\
&\overset{(59)}{\ge} 5.6u^{\frac{1}{1+\epsilon}} \left(\frac{\log t^3}{N_i(t-1)}\right)^{\frac{\epsilon}{1+\epsilon}} + \sqrt{\frac{2\log t^3 u\widehat{M}_i(t)^{1-\epsilon}}{N_i(t-1)}} + \frac{(10+2\sqrt{2})\widehat{M}_i(t)\log t^3}{N_i(t-1)} \\
&\overset{(60)}{\ge} 5.6u^{\frac{1}{1+\epsilon}} \left(\frac{\log t^3}{N_i(t-1)}\right)^{\frac{\epsilon}{1+\epsilon}} + \sqrt{\frac{2\mathbb{E}[V_i(t)]\log t^3}{N_i(t-1)}} + \frac{(10+2\sqrt{2})\widehat{M}_i(t)\log t^3}{N_i(t-1)} \\
&\overset{(58)}{\ge} 5.6u^{\frac{1}{1+\epsilon}} \left(\frac{\log t^3}{N_i(t-1)}\right)^{\frac{\epsilon}{1+\epsilon}} + \sqrt{\frac{2\log t^3}{N_i(t-1)}} \left[\sqrt{V_i(t)} - 2\widehat{M}_i(t)\sqrt{\frac{\log t^3}{N_i(t-1)}}\right] \\
&\qquad + \frac{(10+2\sqrt{2})\widehat{M}_i(t)\log t^3}{N_i(t-1)} \\
&\ge 5.6u^{\frac{1}{1+\epsilon}} \left[\frac{\log t^3}{N_i(t-1)}\right]^{\frac{\epsilon}{1+\epsilon}} + 2\sqrt{\frac{V_i(t)\log t^3}{N_i(t-1)}} + \frac{10\widehat{M}_i(t)\log t^3}{N_i(t-1)}.
\end{aligned} \tag{$*$}$$

Finally, as a consequence of $(*)$, we have $B_1(t) > B_i(t)$ but this is a contradiction since $T_t = i$. Thus, statements (55) to (59) cannot be false simultaneously. We now proceed with a union bound over all the possible values of $N_i(t-1)$ and of the previously introduced concentration inequalities to bound with $\frac{1}{t^3}$ the probabilities of events (55), (56), (58), and (59) to be true:

$$\begin{aligned}
&\mathbb{P}\left(\exists N_i(t-1) \in [t] \,:\, \{(55) \text{ is true}\} \text{ or } \{(56) \text{ is true}\} \text{ or } \{(58) \text{ is true}\} \text{ or } \{(59) \text{ is true}\}\right) \le \\
&\qquad \le 6\sum_{s=1}^{t} \frac{1}{t^3} = \frac{6}{t^2},
\end{aligned}$$

where for (58), we used the second inequality of Theorem 10 of (Maurer and Pontil, 2009) (bounding $1/(n-1) \le 2/n$) and for (59), we used Theorem 5. To proceed, we introduce the quantity:

$$v := \left\lceil 60\left(\frac{u}{\Delta_i^{1+\varepsilon}}\right)^{\frac{1}{\varepsilon}} \log T \right\rceil.$$

It's now time to bound the expected number of times each arm is pulled:

$$\begin{aligned}\mathbb{E}[N_i(T)] = \mathbb{E}\left[\sum_{t=1}^{T} \mathbb{1}_{\{I_t = i \text{and} \mathcal{E}_{i,t}\}}\right] &\leq v + \mathbb{E}\left[\sum_{t=v+1}^{T} \mathbb{1}_{\{I_t = i \text{ and } \{(57) \text{ is false }\}\}}\right] \\ &\leq v + \mathbb{E}\left[\sum_{t=v+1}^{T} \mathbb{1}_{\{I_t = i \text{ and } \{(55) \text{ or } (56) \text{ or } (58) \text{ or } (59) \text{ is true}\}\}}\right] \\ &\leq v + \sum_{t=v+1}^{T} \frac{6}{t^2} \\ &\leq v + 10.\end{aligned} \tag{61}$$

We now conclude the proof using the regret decomposition, considering the forced exploration through Equation (54) and that the effective number of pulls is doubled:

$$R_T(\texttt{AdaR-UCB}, \boldsymbol{\nu}) \leq \sum_{i:\Delta_i > 0} \left[\left(120\left(\frac{u}{\Delta_i}\right)^{\frac{1}{\epsilon}} + \frac{24\Delta_i}{\mathbb{P}_{\nu_i}(X \neq 0)}\right)\log\frac{T}{2} + 20\Delta_i\right].$$

■

Theorem 8 (Worst-Case Regret bound of AdaR-UCB) *Let $\boldsymbol{\nu} \in \mathcal{P}_{HT}(\epsilon, u)^K$ and $T \in \mathbb{N}_{\geq 2}$ be the learning horizon. Under Assumption 1,* `AdaR-UCB` *suffers a regret bounded as:*

$$R_T(\texttt{AdaR-UCB}, \boldsymbol{\nu}) \leq 46\left(K\log\frac{T}{2}\right)^{\frac{\epsilon}{1+\epsilon}} (uT)^{\frac{1}{1+\epsilon}} + \sum_{i:\Delta_i > 0}\left(\frac{24\Delta_i}{\mathbb{P}_{\nu_i}(X \neq 0)}\log\frac{T}{2} + 20\Delta_i\right).$$

Proof Let us fix $\Delta > 0$, to be chosen later. We have:

$$\begin{aligned}R_T(\texttt{AdaR-UCB}, \boldsymbol{\nu}) &= \sum_{i \in [K]} \Delta_i \left(2\mathbb{E}[N_i(T/2)] + \mathbb{E}_{\nu_i}[N_i^{\mathrm{FE}}(T/2)]\right) \\ &= \sum_{i:\Delta_i \leq \Delta} 2\Delta_i \mathbb{E}[N_i(T/2)] + \sum_{i:\Delta_i > \Delta} 2\Delta_i \mathbb{E}[N_i(T/2)] + \sum_{i:\Delta_i > 0} \frac{27\Delta_i}{\mathbb{P}_{\nu_i}(X \neq 0)} \log\frac{T}{2} \\ &\leq \Delta T + \sum_{i:\Delta_i > \Delta} 2\Delta_i \left(60\left(\frac{u}{\Delta_i^{1+\epsilon}}\right)^{\frac{1}{\epsilon}} \log\frac{T}{2} + 10\right) + \sum_{i:\Delta_i > 0} \frac{24\Delta_i}{\mathbb{P}_{\nu_i}(X \neq 0)} \log\frac{T}{2} \\ &\leq \Delta T + 2K\left(60\left(\frac{u}{\Delta}\right)^{\frac{1}{\epsilon}} \log\frac{T}{2}\right) + \sum_{i:\Delta_i > 0}\left(\frac{24\Delta_i}{\mathbb{P}_{\nu_i}(X \neq 0)}\log\frac{T}{2} + 20\Delta_i\right) \\ &\overset{(*)}{\leq} 120^{\frac{\epsilon}{1+\epsilon}}(1+\epsilon)\epsilon^{-\frac{\epsilon}{1+\epsilon}} \left(K\log\frac{T}{2}\right)^{\frac{\epsilon}{1+\epsilon}} (uT)^{\frac{1}{1+\epsilon}} + \sum_{i:\Delta_i > 0}\left(\frac{24\Delta_i}{\mathbb{P}_{\nu_i}(X \neq 0)}\log\frac{T}{2} + 20\Delta_i\right) \\ &\overset{(**)}{\leq} 46\left(K\log\frac{T}{2}\right)^{\frac{\epsilon}{1+\epsilon}} (uT)^{\frac{1}{1+\epsilon}} + \sum_{i:\Delta_i > 0}\left(\frac{24\Delta_i}{\mathbb{P}_{\nu_i}(X \neq 0)}\log\frac{T}{2} + 20\Delta_i\right),\end{aligned}$$

where the step marked with $(*)$ follows by a proper choice of Δ minimizing the bound:

$$T - 120Ku^{\frac{1}{\epsilon}}\epsilon^{-1}\Delta^{-\frac{1+\epsilon}{\epsilon}}\log\frac{T}{2} = 0 \implies \Delta = \left(\frac{120Ku^{\frac{1}{\epsilon}}\log\frac{T}{2}}{\epsilon T}\right)^{\frac{\epsilon}{1+\epsilon}},$$

and step marked with $(**)$ follows by bounding simple numerical bounds. ■

Appendix C. Efficient Numerical Resolution of Equation (13)

In this appendix, we present a computationally efficient strategy that can be implemented in Algorithm 1 to execute line 7, *i.e.*, the solution of the root-finding problem. In particular, to solve the equation:

$$f_s(\mathbf{X}'; M, \delta) \coloneqq \frac{1}{s} \sum_{j \in [s]} \frac{\min\{(X'_j)^2, M^2\}}{M^2} - \frac{c \log \delta^{-1}}{s} = 0. \tag{13}$$

Algorithm 2: Computationally Efficient Threshold Estimation.

---

1 Reward set $\mathbf{X}' = \{X'_1, \ldots, X'_s\}$, time counter τ, machine tolerance $\eta > 0$.
2 Initialize counter $h \leftarrow 0$, initial guess $x_0 \leftarrow \eta$, initial value $y_0 \leftarrow f_s(\mathbf{X}'; x_0, \tau^{-3})$.
3 **while** $y_h > 0$ **do**
4 $\quad x_{h+1} \leftarrow 2x_h$
5 $\quad y_{h+1} \leftarrow f_s(\mathbf{X}'; x_{h+1}, \tau^{-3})$
6 $\quad h \leftarrow h + 1$
7 **end**
8 Return x_h .

We propose Algorithm 2 to find an upper bound $\bar{M}_s(\tau^{-3})$ on the true solution $\widehat{M}_s(\tau^{-3})$ which is based on *bisection*. The strategy works as follows. We provide the minimum numerical tolerance of our machine $\eta > 0$, start from an initial guess $x_0 = \eta$, then, if this guess is an underestimation (*i.e.*, $f_s(\cdot, x_0)$ yields a positive value y_0) we proceed to iteratively double our guess until the real threshold has been passed (lines 4-6). In line 8, we return the final guess x_h. If the initial guess is already an overestimation of the threshold (*i.e.*, $f_s(\cdot, x_0)$ yields a negative value y_0), we simply have $x_0 = x_h = \eta$.

We point out that, by construction, the output of Algorithm 2 can be *at most* two times the true solution to Equation (13), *i.e.*, $\bar{M}_s(\tau^{-3}) \leq 2\widehat{M}_s(\tau^{-3})$. Thus, regret guarantees for Algorithm 1 remain the same (up to numerical constants) even when performing this approximation of the threshold. In particular, in the proof of Theorem 7, we can modify (59) as follows:

$$\bar{M}_i(t) \geq 2 \left(\frac{u N_i(t-1)}{\log t^3} \right)^{\frac{1}{1+\epsilon}},$$

and the final result remains the same up to multiplicative constants.

We now characterize the computational complexity of Algorithm 2, *i.e.*, the maximum number of steps to be performed before returning a solution.

Proposition 10 (Upper Bound on Algorithm 2 Number of Steps) *Let η be the minimum numerical tolerance, and assume $\eta \leq \widehat{M}_s(\tau^{-3})$. Then, in at most $\bar{h}_{\eta,\tau}(\epsilon, u)$ steps such that:*

$$\bar{h}_{\eta,\tau}(\epsilon, u) = \log_2\left(\frac{1}{\eta}\left(\frac{us}{\log(\tau^3)}\right)^{\frac{1}{1+\epsilon}}\right),$$

Algorithm 2, returns a solution $x_{\bar{h}_{\eta,\tau}(\epsilon,u)}$ s.t.

$$\mathbb{P}\left(\frac{x_{\bar{h}_{\eta,\tau}(\epsilon,u)}}{\widehat{M}_s(\tau^{-3})} \in [1,2]\right) \geq 1 - \frac{2}{\tau^3}.$$

Proposition 10 states an upper bound for the number of steps of Algorithm 2 as a function of both ϵ and u. However, we remark that these two are not required as input to the numerical solver. Moreover, it emerges a dependence on the inverse of the numerical tolerance of the machine on which the algorithm is run. Thanks to the logarithm, this dependence hardly becomes an issue. If we consider a very small tolerance of 10^{-16} (which is the standard tolerance of many programming languages) the number of steps becomes:

$$\bar{h}_{\eta,\tau}(\epsilon, u) = \log_2\left(\left(\frac{us}{\log(\tau^3)}\right)^{\frac{1}{1+\epsilon}}\right) + 16\log_2(10),$$

which is totally reasonable.