# Nearly Optimal Catoni's M-estimator for Infinite Variance

**Sujay Bhatt, Guanhua Fang, Ping Li**
Cognitive Computing Lab
Baidu Research
10900 NE 8th St. Bellevue, WA 98004, USA
{sujaybhatt.hr, fanggh2018, pingli98}@gmail.com

**Gennady Samorodnitsky**
School of ORIE
Cornell University
220 Frank T Rhodes Hall, Ithaca, NY 14853, USA
gs18@cornell.edu

## Abstract

[1]In this paper, we extend the remarkable M-estimator of Catoni (2012) to situations where the variance is infinite. In particular, given a sequence of i.i.d random variables $\{X_i\}_{i=1}^n$ from distribution $\mathcal{D}$ over $\mathbb{R}$ with mean $\mu$, we only assume the existence of a known upper bound $\upsilon_\varepsilon > 0$ on the $(1+\varepsilon)^{th}$ central moment of the random variables, namely, for $\varepsilon \in (0, 1]$

$$\mathbb{E}_{X_1 \sim \mathcal{D}} \Big| X_1 - \mu \Big|^{1+\varepsilon} \le \upsilon_\varepsilon.$$

The extension is non-trivial owing to the difficulty in characterizing the roots of certain polynomials of degree smaller than 2. The proposed estimator has the same order of magnitude and the same asymptotic constant as in Catoni (2012), but for the case of bounded moments. We further propose a version of the estimator that does not require even the knowledge of $\upsilon_\varepsilon$, but adapts the moment bound in a data-driven manner. Finally, to illustrate the usefulness of the derived non-asymptotic confidence bounds, we consider an application in multi-armed bandits and propose best arm identification algorithms, in the fixed confidence setting, that outperform the state of the art.

## 1. Introduction

Mean estimation with an emphasis on achievable trade-off between accuracy and confidence plays a pivotal role in the design of algorithms for applications such as multi-armed bandits (Auer et al., 2002; Lattimore and Szepesvári, 2020) and reinforcement learning (Fiechter, 1994; Burnetas and Katehakis, 1997; Auer and Ortner, 2006; Agarwal et al., 2019). Here the fundamental problem is of designing an estimator $\widehat{\mu}_n = \widehat{\mu}_n(\{X_i\}_{i=1}^n)$ that for a given *confidence* $\delta \in (0, 1)$ has the smallest possible $\varrho = \varrho(n, \delta)$ such that $\mathbb{P}\Big\{ \Big| \widehat{\mu}_n - \mu \Big| > \varrho \Big\} \le \delta$.

### 1.1. Sub-Gaussian Estimators

Taking a cue from the asymptotic confidence bounds informed by the Central Limit Theorem for the empirical mean, Devroye et al. (2016) derive several $L-$sub-Gaussian estimators and also establish the tightest possible $L$ for different classes of distributions. More precisely, an estimator $\widehat{\mu}_n$ is $L-$sub-Gaussian if there is a constant $L > 0$ such that for random variables with variance $\sigma^2$ and (any) sample size $n$, it holds

$$\Big| \widehat{\mu}_n - \mu \Big| \le \frac{L\sigma\sqrt{\log(2/\delta)}}{\sqrt{n}}$$

with probability at least $1 - \delta$. Sub-Gaussian estimators are optimal up to constants, and estimators with $L \le \sqrt{2} + o(1)$ are nearly not improvable.

There are equivalent definitions of sub-Gaussian estimators; see for example, the monograph by Buldygin and Kozachenko (2000). Also see Li (2007, Chapter 4) for the use of sub-Gaussian distributions in the context of random projections and very sparse random projections.[2]

In the context of these $L-$sub-Gaussian estimators, the estimator proposed in Catoni (2012), which is henceforth referred to as *Catoni's estimator*, is significant owing to the fact that it is a (nearly) *optimal* sub-Gaussian estimator of the mean with $L = \sqrt{2} + o(1)$. Here optimal is to be understood in the sense that the Catoni's estimator comes close the best possible $L(= \sqrt{2})$. Given the sharpness of result, it is not surprising that Catoni's estimator has spawned a wide range of applications ranging from bandits (Bubeck et al., 2013) to empirical risk minimization (Brownlees et al., 2015).

*Proceedings of the $39^{th}$ International Conference on Machine Learning*, Baltimore, Maryland, USA, PMLR 162, 2022. 

[1]The work of Gennady Samorodnitsky was conducted as a consulting researcher at Baidu Research – Bellevue, WA 98004.

[2]https://hastie.su.domains/THESES/pingli_thesis.pdf.

### 1.2. Heavy-tailed Estimators

One crucial shortcoming of the Catoni estimator is that it requires the existence of finite second moment. This is a serious limitation in heavy-tail settings where the variance need not exist or be finite. Such heavy-tail distributions are well motivated and studied extensively in applications (Resnick, 2007), for example, *sparse recovery* with extremely heavy-tailed stable distributions (Li and Zhang, 2013; Li, 2016). Given the simplicity and near-optimality of the Catoni estimator under finite variance ($\varepsilon = 1$), it is natural to consider extending the underlying ideas to derive an estimator when the variance of $X \sim \mathcal{D}$ is infinite, i.e, $\mathbb{E}_{X_1 \sim \mathcal{D}} \Big| X_1 - \mu \Big|^{1+\varepsilon} \leq \upsilon_\varepsilon$ for $\varepsilon < 1$. Chen et al. (2021) take the first steps towards this and propose an estimator that approaches the same order as in Catoni's estimator as $\varepsilon \to 1$. However, Chen et al. (2021) assume a sub-optimal choice for the influence function inspired by Taylor-like expansions, resort to weak $C_r-$inequalities, and employ loose characterization of the roots of certain polynomials of degree smaller than 2 to construct the estimator. These lossy intermediate steps results in large asymptotic constants that are not the same as in Catoni (2012), ruling out near-optimality. The main contribution of this paper is to alleviate these shortcomings and obtain nearly optimal Catoni's estimator for $\varepsilon < 1$, using a different approach.

When $\varepsilon < 1$, Devroye et al. (2016) establish that the achievable $\varrho(n, \delta)$ is no more sub-Gaussian and is in fact given by the following result:

**Theorem** (Lower Bound)**.** There exists a distribution with mean $\mu$ and $(1+\varepsilon)^{th}$ central moment upper bound $\upsilon_\varepsilon$ such that for any mean estimator $\widehat{\mu}_n$ and $\delta \in (2e^{-n/4}, 1/2)$,

$$\mathbb{P}\Big\{ |\widehat{\mu}_n - \mu| > \Big( \frac{\upsilon_\varepsilon^{\frac{1}{\varepsilon}} \log(2/\delta)}{n} \Big)^{\frac{\varepsilon}{1+\varepsilon}} \Big\} \geq \delta.$$

This is established in Devroye et al. (2016, Theorem 3.1). See (Lugosi and Mendelson, 2019) for an excellent summary and related literature.

The derived extension of Catoni's estimator requires the knowledge of $\upsilon_\varepsilon$ to perform the M-estimation. This might be less desirable when such information is not available, whence it is impossible to provide any observable confidence intervals. We extend the results to define an adaptive estimator and to bound its deviations by bounds depending on the unknown moments. This extension borrows ideas from Lepskii's adaptation method (Lepskii, 1992) to provide a data-driven M-estimator that adapts to unknown $\upsilon_\varepsilon$.

To illustrate the usefulness of the non-asymptotic confidence bounds associated with Catoni's estimator for heavy-tails derived in this work, we propose best arm identification algorithms (BAI) in the fixed confidence setting (Even-Dar et al., 2006), which outperform the state of the art.

### 1.3. Main Contributions

We summarize our main contributions and key results:

(i) We provide a nearly-optimal M-estimator of the mean of random variables having bounded $(1+\varepsilon)^{th}$ moment. This implies that the asymptotic constant is the best possible and same as in Catoni (2012). Two key ideas contribute towards this: (a) we construct an influence function that favorably controls the $(1+\varepsilon)^{th}$ order variations from the mean; (b) we characterize tight upper bounds on the roots of certain polynomials of degree $1+\varepsilon$.

(ii) We provide an algorithm that extends the proposed M-estimator to adapt to an unknown moment bound $\upsilon_\varepsilon$ in a data-driven manner. The key ideas make use of Lepskii's classical adaptive estimation techniques (Lepskii, 1992). We further establish that the resulting M-estimator has close to near-optimal asymptotic bound.

(iii) Unlike the previous results on M-estimators under $(1+\varepsilon)^{th}$ moment assumption (see Chen et al. (2021) and the references therein), the minimum data required to obtain a tight high confidence bound for the proposed M-estimator is decoupled from the moment bound $\upsilon_\varepsilon$. This makes it possible to employ the M-estimators in online learning applications that need to adapt to unknown $\upsilon_\varepsilon$. We illustrate one such application in multi-armed bandits, where the objective is to identify the best arm with a high confidence. We propose a novel best arm identification algorithm that eliminates arms in phases, while sequentially adapting to the unknown $\upsilon_\varepsilon$. We further establish that the sample complexity is doubly-logarithmic in the problem parameters, while demonstrating excellent empirical performance.

### 1.4. Overview of Previous Work

Construction of tight confidence intervals for the unknown mean using finite samples from the distribution is a fundamental objective that has naturally generated an enormous interest in the statistics and machine learning community. A detailed investigation into this using M-estimation was initiated in the work of Catoni (2012), and there upon formalized in Devroye et al. (2016). Since this class of estimators only require a bounded second moment, a notion of 'robustness' is commonly associated (Lugosi and Mendelson, 2019) with them. While the focus of this work is primarily on M-estimators (Zhou et al., 2018), there are other estimators that obtain the optimal order such as Median-of-Means (Bubeck et al., 2013; Minsker, 2019; Lecué and Lerasle, 2020) and Trimmed Mean (Oliveira and Orenstein, 2019; Lugosi and Mendelson, 2021), each with their own merits and shortcomings. The key difference is that robustness in class of M-estimators is characterized by the extrema of an influence function which modulates the impact of outlier samples (Huber, 2004).

A powerful method for adapting to the unknown moments was introduced in Lepskii (1992). The versatility of the method stems from the fact that it can adapt to any unknown structure of the problem. Since then the idea has been used in M-estimation to adapt to unknown variance (Catoni, 2012; Minsker, 2018).

Best arm identification in the fixed confidence setting is a classical framework (Even-Dar et al., 2006) in multi-armed bandits that has a wealth of applications (Jamieson and Talwalkar, 2016; Kaufmann and Koolen, 2017). Two algorithms lil'UCB (Jamieson et al., 2014) and Exponential-Gap Elimination algorithm in (Karnin et al., 2013) obtain the optimal sample complexity for the best-arm problem in the fixed confidence setting under *sub-Gaussian* reward distributions, and are standard. In the asymptotic setting, Garivier and Kaufmann (2016) propose a best arm identification algorithm that matches the established tighter lower bound as $\delta$ reduces to zero. In the case of bounded $(1+\varepsilon)^{th}$ moment, Yu et al. (2018); Glynn and Juneja (2018) propose best arm identification algorithms based on trimmed mean estimator from Bubeck et al. (2013). In the asymptotic setting, Agrawal et al. (2020) propose a best arm identification algorithm that matches the established lower bound as $\delta$ reduces to zero.

## 2. Key Ideas from Finite Variance (Catoni, 2012)

In this section, we briefly review Catoni's estimator (Catoni, 2012) illustrating the key changes required in extending it to $\varepsilon \in (0,1)$. Let a sequence of i.i.d random variables $\{X_i\}_{i=1}^n$ be such that $\mathbb{E}X_1 = \mu$ and $\mathbb{E}|X_1-\mu|^2 \leq \upsilon$. Consider an "influence" function $\psi_c : \mathbb{R} \to \mathbb{R}$ that is non-decreasing and is such that for all $x \in \mathbb{R}$

$$-\log(1-x+x^2/2) \leq \psi_c(x) \leq \log(1+x+x^2/2) \quad (1)$$

(`Fact 1`) Note that the bounds that define the function $\psi_c(x)$ satisfy

$$(1-x+x^2/2)^{-1} \leq (1+x+x^2/2) \Leftrightarrow 1 \leq 1+x^4/4,\ \forall\, x.$$

An analogous fact will be required to state bounds appropriate for dealing with infinite variance. For $\alpha > 0$, Catoni's M-estimator $\widehat{\mu}_c$ is defined as a solution to the equation

$$\sum_{i=1}^n \psi_c\Big(\alpha(X_i - \widehat{\mu}_c)\Big) = 0.$$

The traditional empirical mean ($\psi_c(x) = x$) is unduly influenced by large values which might invalidate[3] the sub-Gaussian confidence bound when the distributional tails are not themselves sub-Gaussian. When instead $\psi_c(x)$ satisfies (1), it is similar to the linear function for small and moderate values of $x$, but its logarithmic rate of growth reduces the effect of large values and preserves the sub-Gaussian confidence bound.

Next, consider the function $r_{c,n}(\theta)$ defined as

$$r_{c,n}(\theta) = \sum_{i=1}^n \psi_c\Big(\alpha(X_i-\theta)\Big).$$

Note that $r_{c,n}(\theta)$ is a non-increasing function of $\theta$. Using the upper bound on $\psi_c(x)$ in (1), we have that

$$\mathbb{E}\Big[\exp(r_{c,n}(\theta))\Big] \leq \Big(1+\alpha(\mu-\theta)+\frac{\alpha^2(\upsilon^2+(\mu-\theta)^2)}{2}\Big)^n$$
$$\leq \exp\Big(n\alpha(\mu-\theta)+\frac{n\alpha^2(\upsilon^2+(\mu-\theta)^2)}{2}\Big). \quad (2)$$

From (2), a motivation for choosing the logarithm of a degree 2 polynomial in the bounds for the influence function is clear, as it helps to bound the expectation using the variance upper bound (`Fact 1`). This also informs the nature of influence functions required to obtain bounds using $(1+\varepsilon)^{th}$ moments; see Minsker (2018, Section 3.4).

Using the exponential Markov inequality, we now have that for any $\theta \in \mathbb{R}$ and $\delta \in (0,1)$

$$\mathbb{P}\Big\{r_{c,n}(\theta) \geq n\alpha(\mu-\theta)+\frac{n\alpha^2(\upsilon^2+(\mu-\theta)^2)}{2} + \log(2/\delta)\Big\} \leq \delta/2.$$

It is clear from the monotonicity of the $r_{c,n}$ that the roots of the quadratic polynomial in $\theta$,

$$n\alpha(\mu-\theta)+\frac{n\alpha^2(\upsilon^2+(\mu-\theta)^2)}{2}+\log(2/\delta),$$

play a pivotal role in characterizing the deviations of the estimator $\widehat{\mu}_c$ from $\mu$. For a suitable $\alpha$, the quadratic polynomial has at least one or two roots that can be obtained in closed form (`Fact 1`). If $\theta^+$ is the smallest root, then $r_{c,n}(\theta^+) < 0$ with probability at least $1-\delta/2$. A suitable choice $\alpha$ results in the following upper confidence bound (Catoni, 2012):

$$\widehat{\mu}_c \leq \theta^+ = \mu + \sqrt{\frac{2\upsilon\log(2/\delta)}{n-2\log(2/\delta)}}.$$

A lower bound can be derived using symmetric arguments. Note that an influence function designed using ideas outlined above (`Fact 1` and `Fact 2`) for the infinite variance setting will not have explicit expressions of the roots (`Fact 3`). In fact, as we shall see in the next section that this is precisely the difficulty that has to be overcome to extend Catoni's estimator to deal with infinite variance, while ensuring near-optimality.

[3] For examples of distributions that result in $O\Big(\sqrt{1/n\delta}\Big)$ confidence bound, see (Catoni, 2012; Lugosi and Mendelson, 2019).

## 3. Catoni's M-estimator for Heavy Tails

Let $\varepsilon \in (0,1]$ and let a sequence of i.i.d random variables $\{X_i\}_{i=1}^n$ be such that $\mathbb{E}(X_1) = \mu$ and $\mathbb{E}|X_1 - \mu|^{1+\varepsilon} \le \upsilon_\varepsilon$. For $C_\varepsilon > 0$, let $\psi : \mathbb{R} \to \mathbb{R}$ be a non-decreasing influence function such that for all $x \in \mathbb{R}$

$$-\log(1-x+C_\varepsilon|x|^{1+\varepsilon}) \le \psi(x) \le \log(1+x+C_\varepsilon|x|^{1+\varepsilon}). \tag{3}$$

Similarly, consider a Catoni's M-estimator $\widehat{\mu}_{c,\varepsilon}$ as a solution to the equation

$$\sum_{i=1}^{n} \psi\Big(\alpha(X_i - \widehat{\mu}_{c,\varepsilon})\Big) = 0 \tag{4}$$

using an influence function $\psi$ satisfying (3). If the solution is not unique, choose $\widehat{\mu}_{c,\varepsilon}$ to be the median solution.

**Lemma 3.1.** *A function $\psi$ satisfying (3) exists if and only if*

$$C_\varepsilon \ge \left(\frac{\varepsilon}{1+\varepsilon}\right)^{\frac{1+\varepsilon}{2}} \left(\frac{1-\varepsilon}{\varepsilon}\right)^{\frac{1-\varepsilon}{2}}.$$

Lemma 3.1 provides a necessary and sufficient condition for the selection of the coefficient in the influence function, similar to `Fact 1` in Section 2. We want to emphasize that many choices of $C_\varepsilon$ are possible that satisfy `Fact 1`, for example, Chen et al. (2021) choose $C_\varepsilon = \frac{1}{\varepsilon}$ inspired by Taylor-like expansions, and Minsker (2018) choose $C_\varepsilon = \frac{\varepsilon}{1+\varepsilon} \vee \sqrt{\frac{1-\varepsilon}{1+\varepsilon}}$. In our paper, we derive a **tightest** such choice that obtains a sharp bound as $|\widehat{\mu}_{c,\varepsilon} - \mu| < \texttt{rem.}\,(C_\varepsilon)^{1/(1+\varepsilon)}$. We further contribute a novel analysis that tightens the `rem.` term.

**Remark 1.** Unless otherwise stated, hereon we choose the coefficient $C_\varepsilon = \left(\frac{\varepsilon}{1+\varepsilon}\right)^{\frac{1+\varepsilon}{2}} \left(\frac{1-\varepsilon}{\varepsilon}\right)^{\frac{1-\varepsilon}{2}}$. When $\varepsilon = 1$, we recover the coefficient in Catoni (2012), namely $C_1 = 1/2$. (Here the standard convention $0^0 := 1$ applies.) Also, we assume throughout that the influence function $\psi$ satisfies (3).

**Theorem 3.2** (Main Result)**.** *Let $\{X_i\}_{i=1}^n$ be i.i.d random variables with mean $\mu$ and $\mathbb{E}|X_1 - \mu|^{1+\varepsilon} \le \upsilon_\varepsilon$. Let $\delta \in (0,1)$, $\tau > 0$ and $0 < h < 1$ be such that*

$$n \ge (1-h)^{-1} \frac{1+\varepsilon}{\varepsilon} \frac{(1+\tau)^{(1+\varepsilon)/\varepsilon}}{\tau^{1/\varepsilon}} C_\varepsilon^{1/\varepsilon} \log(2/\delta).$$

*Then for $\varepsilon \in (0,1]$, Catoni's M-estimator $\widehat{\mu}_{c,\varepsilon}$ with parameter*

$$\alpha = \frac{1}{\varepsilon^{\frac{1}{1+\varepsilon}}} \frac{1}{(h^{-\varepsilon} C_\varepsilon \upsilon_\varepsilon)^{1/(1+\varepsilon)}} \left(\frac{\log(2/\delta)}{n}\right)^{\frac{1}{1+\varepsilon}}$$

*satisfies, with probability at least $1-\delta$,*

$$|\widehat{\mu}_{c,\varepsilon} - \mu| < (1+\tau)(1+\varepsilon)^{1/2}(1-\varepsilon)^{\frac{1}{1+\varepsilon}-\frac{1}{2}} \times (h^{-\varepsilon}\upsilon_\varepsilon)^{1/(1+\varepsilon)} \left(\frac{\log(2/\delta)}{n}\right)^{\frac{\varepsilon}{1+\varepsilon}}. \tag{5}$$

Theorem 3.2 provides a non-asymptotic confidence bound for heavy tails. The estimator depends on scaling $\alpha$ and the confidence parameter $\delta$, and is $O\Big(\frac{\log(2/\delta)}{n}\Big)^{\frac{\varepsilon}{1+\varepsilon}}$. The key elements in the proof are as follows: (i) The 'right' influence function (3) coupled with a convexity upper bound,

$$(a+b)^{1+\varepsilon} \le \frac{a^{1+\varepsilon}}{h^\varepsilon} + \frac{b^{1+\varepsilon}}{(1-h)^\varepsilon}, \quad a, b \ge 0 \text{ and } h \in (0,1),$$

obtains a polynomial of degree $(1+\varepsilon)$ with smallest coefficients. (ii) Characterizing a tight upper bound on the roots of this polynomial yields the desired confidence bound.

We note that for $\varepsilon = 1$, we obtain the exact near-optimal estimator as in Catoni (2012); see Sec.3.3.1 for details.

### 3.1. Comparison with known bounds

Below we compare the non-asymptotic bounds in (5) with state-of-the-art mean estimators. Since trimmed mean has a larger bound than Median-of-Means (MoM) estimator (Lugosi and Mendelson, 2019), we only make a comparison with MoM estimator.

#### 3.1.1. Median of Means Estimator

We compare the constants obtained in (5) with the constants obtained for the Median-of-Means (MoM) estimator (Bubeck et al., 2013; Lugosi and Mendelson, 2019) that also features central moments. The MoM estimator $\widehat{\mu}_{MoM}$ with appropriate number of blocks has the bound

$$|\widehat{\mu}_{MoM} - \mu| \le 8(12)^{\frac{\varepsilon}{1+\varepsilon}} \upsilon_\varepsilon^{1/(1+\varepsilon)} \left(\frac{\log(2/\delta)}{n}\right)^{\frac{\varepsilon}{1+\varepsilon}},$$

with probability at least $1-\delta$; see Lugosi and Mendelson (2019, Theorem 3). For a similar length of the random sequence, for example even with a choice of $h = 0.5$, the constant in (5) is better than that using the MoM estimator for $\varepsilon \ge 0.05$.

#### 3.1.2. Chen et al.'s M-estimator

M-estimator derived by Chen et al. (2020) has the following bound, for $n \ge \left(\frac{2\upsilon_\varepsilon+1}{1+\varepsilon}\right)^{\frac{1+\varepsilon}{\varepsilon}} \frac{2(1+\varepsilon)\log(2/\delta)}{\upsilon_\varepsilon}$, with probability at least $1-\delta$

$$|\widehat{\mu}_M - \mu| \le \frac{2\Big(2(1+\varepsilon)\Big)^{\frac{\varepsilon}{1+\varepsilon}} \upsilon_\varepsilon^{1/1+\varepsilon}}{(1+\varepsilon) - \left(\frac{2(1+\varepsilon)\log(2/\delta)}{n\upsilon_\varepsilon}\right)^{\frac{\varepsilon}{1+\varepsilon}}} \left(\frac{\log(2/\delta)}{n}\right)^{\frac{\varepsilon}{1+\varepsilon}}. \tag{6}$$

The non-asymptotic bound in (5) is sharper than (6) for all $\varepsilon \in (0,1)$ even with $h = 0.5$. Note that the asymptotic constant in (6) is much larger than in Catoni (2012) for $\varepsilon = 1$, ruling out near-optimality (see Section 3.3.1). A minor generalization is later proposed in Chen et al. (2021) using a convexity upper bound, however, we compare with Chen et al. (2020) as the best parameter choices are not discussed in Chen et al. (2021). We note that the constant in (5) is still much sharper than that in Chen et al. (2021) owing to tight characterizations on the roots of degree $< 2$ polynomial.

### 3.2. $\delta-$independent estimator

It is clear that the estimator $\widehat{\mu}_{c,\varepsilon}$ is a function of $\delta$. This is not always desirable as a different point estimate results for each $\delta$. We can obtain an estimator *independent* of $\delta$, which however does not have sharp bounds. The dependence on $n$ is still optimal, but the dependence on $\delta$ is worse than (5).

**Corollary 3.3.** *Suppose that for* $0 < h < 1$,

$$n \geq (1 + \log(2/\delta))((1+\varepsilon)C_\varepsilon)^{1/\varepsilon} \frac{1+\varepsilon}{\varepsilon(1-h)}.$$

*Then with* $\alpha = \left(\frac{h^\varepsilon}{C_\varepsilon \upsilon_\varepsilon n}\right)^{1/1+\varepsilon}$, *Catoni's M-estimator* $\widehat{\mu}_{c,\varepsilon}$ *satisfies, with probability at least* $1-\delta$,

$$|\widehat{\mu}_{c,\varepsilon} - \mu| < \left(\frac{1+\varepsilon}{\varepsilon}\right)^{1/2} h^{\frac{-\varepsilon}{1+\varepsilon}} \left(\frac{1-\varepsilon}{\varepsilon}\right)^{1/(1+\varepsilon)-1/2} \times \left(\frac{\upsilon_\varepsilon^{1/(1+\varepsilon)}}{n^{\frac{\varepsilon}{1+\varepsilon}}}\right)(1+\log(2/\delta)).$$

### 3.3. $h-$independent estimator

The following is an immediate corollary of Theorem 3.2. It results from choosing the largest possible $h$ that satisfies the assumptions.

**Corollary 3.4.** *Suppose that for* $\tau > 0$,

$$n \geq \frac{1+\varepsilon}{\varepsilon} \frac{(1+\tau)^{(1+\varepsilon)/\varepsilon}}{\tau^{1/\varepsilon}} C_\varepsilon^{\frac{1}{\varepsilon}} \log(2/\delta).$$

*Then with*

$$\alpha = \frac{1}{\varepsilon^{\frac{1}{1+\varepsilon}}} \frac{1}{(C_\varepsilon \upsilon_\varepsilon)^{1/(1+\varepsilon)}} \left(\frac{\log(2/\delta)}{n}\right)^{\frac{1}{1+\varepsilon}} \times \left(1 - \frac{1+\varepsilon}{\varepsilon} \frac{(1+\tau)^{(1+\varepsilon)/\varepsilon}}{\tau^{1/\varepsilon}} C_\varepsilon^{\frac{1}{\varepsilon}} \frac{\log(2/\delta)}{n}\right)^{\varepsilon/(1+\varepsilon)},$$

*Catoni's M-estimator* $\widehat{\mu}_{c,\varepsilon}$ *satisfies, with probability at least* $1-\delta$,

$$|\widehat{\mu}_{c,\varepsilon} - \mu| < (1+\tau)(1+\varepsilon)^{1/2}(1-\varepsilon)^{\frac{1}{1+\varepsilon}-\frac{1}{2}} \upsilon_\varepsilon^{1/(1+\varepsilon)} \times \left(1 - \frac{1+\varepsilon}{\varepsilon} \frac{(1+\tau)^{(1+\varepsilon)/\varepsilon}}{\tau^{1/\varepsilon}} C_\varepsilon^{\frac{1}{\varepsilon}} \frac{\log(2/\delta)}{n}\right)^{\frac{-\varepsilon}{1+\varepsilon}} \times \left(\frac{\log(2/\delta)}{n}\right)^{\frac{\varepsilon}{1+\varepsilon}}. \tag{7}$$

**Remark 2.** We note that the dependence on the parameter $\tau > 0$ in (7) cannot be eliminated and is a free parameter. A value of $\tau \approx 0$ obtains a tight bound at the cost of an increase in the minimum number of samples as $n \propto \frac{1}{\tau^{1/\varepsilon}}$. A judicious choice hence depends on the budgetary constraints.

#### 3.3.1. Comparison with Catoni (2012) for $\varepsilon = 1$

It is interesting to compare (7) with Catoni (2012, Proposition 2.4) for $\varepsilon = 1$. For large $n$ we can take $\tau$ arbitrarily close to 0, hence in this case, the bound in (7) becomes

$$|\widehat{\mu}_{c,1} - \mu| < (2+\gamma)^{1/2} \upsilon_\varepsilon^{1/2} \left(\frac{\log(2/\delta)}{n}\right)^{1/2},$$

for any $\gamma > 0$ and $n$ large, which has the same order of magnitude and the same asymptotic constant as in Catoni (2012, Proposition 2.4).

## 4. Adapting to unknown $\upsilon_\varepsilon$

We provide an algorithm based on Lepski's method (Lepskii, 1992) to adapt to the unknown structure of the problem, specifically, we show how it adapts to the unknown moment bound in a data-driven manner. Lepski's adaptation method has been employed to select loss functions in regression problems (Sun et al., 2020), to adapt to unknown variance in mean estimation problems (Catoni, 2012), bandwidth selection in kernel density estimation (Kerkyacharian et al., 2001); to name a few. Note that this is different from estimating the moment bound, which requires further prior information. For ease of presentation, rewrite (7) as

$$|\widehat{\mu}_{c,\varepsilon} - \mu| \leq \Phi(\varepsilon,\tau,\delta) \upsilon_\varepsilon^{1/(1+\varepsilon)} \left(\frac{\log(2/\delta)}{n}\right)^{\frac{\varepsilon}{1+\varepsilon}}, \tag{8}$$

valid when $\upsilon_\varepsilon$ is a true moment bound. Here $\Phi(\varepsilon,\tau,\delta)$ absorbs the remaining terms.

Algorithm 1 adapts to the unknown moment bound as described. The main intuition is as follows: a crude moment

**Algorithm 1** Lepskii Moment Adaptation (LMA)

1: **LMA**$\Big(\varepsilon, \{X_i\}_{i\leq n}, \delta, \tau, A(>1), \sigma(>1)\Big)$
2: Compute $\widehat{\upsilon}_\varepsilon = \frac{1}{n}\sum_{i=1}^n |X_i|^{1+\varepsilon}$
3: Set $\vartheta_{\min} = \frac{1}{A}\widehat{\upsilon}_\varepsilon^{\frac{1}{1+\varepsilon}}$ and $\vartheta_{\max} = A\widehat{\upsilon}_\varepsilon^{\frac{1}{1+\varepsilon}}$
4: Set $\vartheta_j = \vartheta_{\min}\sigma^j$ and let

$$\mathcal{J} = \{j = 0,1,2,\cdots : \vartheta_{\min} \leq \vartheta_j \leq \sigma\vartheta_{\max}\}$$

5: Compute Catoni's M-estimators $\{\widetilde{\mu}_{\vartheta_j}\}_{j\in\mathcal{J}}$

$$|\widetilde{\mu}_{\vartheta_j} - \mu| \leq \Phi(\varepsilon,\tau,h)\vartheta_j \left(\frac{\log(2(1+\log_\sigma A^2)/\delta)}{n}\right)^{\frac{\varepsilon}{1+\varepsilon}}$$

6: Evaluate $\widehat{j}$ such that for all $k > j,\ k \in \mathcal{J}$

$$\widehat{j} = \min\Big\{j \in \mathcal{J} : |\widetilde{\mu}_{\vartheta_k} - \widetilde{\mu}_{\vartheta_j}| \leq 2\Phi(\varepsilon,\tau,h) \times \vartheta_j \left(\frac{\log(2(1+\log_\sigma A^2)/\delta)}{n}\right)^{\frac{\varepsilon}{1+\varepsilon}}\Big\}$$

7: **Output**: $\{\widetilde{\mu}_{\vartheta_{\widehat{j}}}, \vartheta_{\widehat{j}}\}$

estimate is obtained by the corresponding empirical mean. We estimate a non-centered moment, but it is easy to see that

$$\big(E|X-\mu|^{1+\varepsilon}\big)^{1/(1+\varepsilon)} \leq 2\big(E|X|^{1+\varepsilon}\big)^{1/(1+\varepsilon)}.$$

Next, we choose $A$ large enough so that, with a large probability the event

$$B_A = \left\{\widehat{\upsilon}_\varepsilon^{\frac{1}{1+\varepsilon}}/A \leq \upsilon_\varepsilon^{\frac{1}{1+\varepsilon}} \leq A\widehat{\upsilon}_\varepsilon^{\frac{1}{1+\varepsilon}}\right\}$$

occurs. There is no way to estimate the probability of the complement of $B_A$ without further assumptions on $X$, and the conclusions of the result below hold on the event $B_A$. However, the resulting confidence interval depends on $A$ only in a very modest manner, so one should choose $A$ large in order to make the probability of the event $B_A$ large. The algorithm proceeds by computing a series of M-estimators (4) with scaling $\alpha$ chosen as

$$\alpha(\vartheta_j) := \frac{1}{\vartheta_j}\frac{1}{(\varepsilon C_\varepsilon)^{1/(1+\varepsilon)}}\Big(\frac{\log(2(1+\log_\sigma A^2)/\delta)}{n}\Big)^{\frac{1}{1+\varepsilon}} \times \left(1-\frac{1+\varepsilon}{\varepsilon}\frac{(1+\tau)^{(1+\varepsilon)/\varepsilon}}{\tau^{1/\varepsilon}}C_\varepsilon^{\frac{1}{\varepsilon}}\frac{\log(2(1+\log_\sigma A^2)/\delta)}{n}\right)^{\frac{\varepsilon}{1+\varepsilon}}.$$

Intuitively, the definition of $\widehat{j}$ can be seen as choosing the best $\upsilon_\varepsilon$ from a point of view of the risk of the estimator (Birgé, 2001). The continuity property of the risk then ensures the desired adaptive performance of the final estimator $\widetilde{\mu}_{\vartheta_{\widehat{j}}}$ (Lepskii, 1992; Birgé, 2001).

**Theorem 4.1.** *Let $\delta \in (0,1)$ and $\widehat{\mu}_{c,\varepsilon}(\mathcal{D})$ denote the data-driven estimator output by Algorithm 1. Suppose that*

$$n \geq (1-h)^{-1}\frac{1+\varepsilon}{\varepsilon}\frac{(1+\tau)^{(1+\varepsilon)/\varepsilon}}{\tau^{1/\varepsilon}}C_\varepsilon^{\frac{1}{\varepsilon}}\log(2(1+\log_\sigma A^2)/\delta).$$

*Then, on the event $B_A$, outside of a part with probability at most $\delta$, the following holds:*

$$|\widehat{\mu}_{c,\varepsilon}(\mathcal{D})-\mu| \leq 3\sigma\Phi(\varepsilon,\tau,h)\upsilon_\varepsilon^{1/(1+\varepsilon)} \times \left(\frac{\log\left(2(1+2\log_\sigma A)\big/\delta\right)}{n}\right)^{\frac{\varepsilon}{1+\varepsilon}}.$$

Typical values for estimators that use Lepskii's method use $A \approx 100$ as a safe choice. Note that the confidence bound has a logarithmic dependence in $A$, so even larger values will only have minor effect.

# 5. Extension to non-i.i.d

In this section, we will show that the results derived in the previous sections can be extended to a special case of non-i.i.d random variables.

The extension of Catoni estimators to this more general setting is based on the application of standard martingale analysis (Freedman, 1975; Seldin et al., 2012) to establish the bounds for bounded functions of real-valued random variables. Since the influence function $\psi(\cdot)$ is bounded by logarithmic functions, a supermartingale can be constructed as a function of $\psi(\cdot)$ and the result then follows using Markov's inequality. Below we provide a crucial result that can be used to establish similar deviation bounds for the non-i.i.d case.

**Proposition 5.1.** *Suppose the process $\{X_t\}$ is a real-valued stochastic process adapted to the filtration $\{\mathcal{F}_t\}_{t\in\mathcal{N}}$ such that $\mathbb{E}[X_{t+1}|\mathcal{F}_t] = \mu$ for all $t \in \mathcal{N}$. Also, let $\mathbb{E}[|X_{t+1}-\mu|^{1+\varepsilon}|\mathcal{F}_t] \leq \upsilon_\varepsilon$. For any $n > 0$,*

$$\mathbb{P}\Big(r(\theta) \geq B_+(\theta)\Big) \leq \frac{\delta}{2},$$

*where $B_+(\theta) := (\mu-\theta)+h^{-\varepsilon}C_\varepsilon\alpha^\varepsilon\upsilon_\varepsilon+C_\varepsilon\alpha^\varepsilon(1-h)^{-\varepsilon}|\mu-\theta|^{1+\varepsilon}+\frac{\log(2/\delta)}{\alpha n}$.*

The result says that the random variables need not be i.i.d, and only need to have constant conditional expected value and a bounded conditional moment. Under these assumptions, the main results (Section 3 and Section 4) still hold using Proposition 5.1. Other types of non-i.i.d case (e.g. adversarial contamination, sequential detection) can be found in concurrent works (Bhatt et al., 2022b;a).

# 6. Application: Best Arm Identification with Fixed Confidence

Consider the best arm (arm with the largest mean) identification problem on a multi-armed bandit with $[K] := \{1,2,\cdots,K\}$ arms in a fixed confidence setting (Even-Dar et al., 2006). Here, given a confidence $\delta$, the objective is to pull the arms as few times as possible to identify the best arm with probability at least $1-\delta$. The distribution of rewards of arm $i \in [K]$ with mean $\mu(i)$ is such that $\mathbb{E}|X_1(i)-\mu(i)|^{1+\varepsilon} \leq \upsilon_\varepsilon$, for $\varepsilon \in (0,1]$. Let

$$G^\tau_{\varepsilon,h}(\delta) := (1+\tau)(1+\varepsilon)^{1/2}(1-\varepsilon)^{\frac{1}{1+\varepsilon}-\frac{1}{2}} \times (h^{-\varepsilon}\upsilon_\varepsilon)^{1/(1+\varepsilon)}\Big(\log(2/\delta)\Big)^{\frac{\varepsilon}{1+\varepsilon}}. \tag{9}$$

The confidence bound using the Catoni's estimator for heavy tails (4) can be rewritten as

$$|\widehat{\mu}_{c,\varepsilon}(t)-\mu| < G^\tau_{\varepsilon,h}(\delta)\Big(\frac{1}{t}\Big)^{\frac{\varepsilon}{1+\varepsilon}}. \tag{10}$$

Let $\widehat{\mu}_{c,\varepsilon}(j,t)$ denote the estimation of mean using (4) for arm $j \in [K]$ at time $t$. Let $* = \arg\max_{j\in[K]}\mu(j)$ and $\Delta_i := \mu(*)-\mu(i)$ denote the sub-optimality gap.

### 6.1. Simple Iterative Elimination with Catoni

In this section, we provide an intuitive algorithm for best arm identification with fixed confidence. With minor modifications, Algorithm 2 is essentially a successive elimination algorithm (Even-Dar et al., 2006) that eliminates sub-optimal arms according to (11). The key difference is the initial exploration required to obtain the desired bound using M-estimation. The elimination criterion is classical (Auer et al., 2002; Even-Dar et al., 2006), where an arm whose optimistic empirical estimate is still worse than the pessimistic estimate of the best arm (having the maximal reward) is eliminated. For $i \in S$, let the *elimination criterion* be specified as

$$\Big\{\widehat{\mu}_{c,\varepsilon}(i,t) + G^{\tau}_{\varepsilon,h}\big(\frac{\delta}{2Kt^2}\big)\Big(\frac{1}{t}\Big)^{\frac{\varepsilon}{1+\varepsilon}}\Big\} < \max_{j\in S}\Big\{\widehat{\mu}_{c,\varepsilon}(j,t) - G^{\tau}_{\varepsilon,h}\big(\frac{\delta}{2Kt^2}\big)\Big(\frac{1}{t}\Big)^{\frac{\varepsilon}{1+\varepsilon}}\Big\}. \qquad (11)$$

**Algorithm 2** Iterative Elimination with Catoni

**Input**: $\delta, K, h, \tau, \varepsilon, \upsilon_\varepsilon$
**Initialization**: Set $S := [K]$
Pull each arm in $S$, $t = \frac{1+\varepsilon}{\varepsilon(1-h)}\frac{(1+\tau)^{(1+\varepsilon)/\varepsilon}}{\tau^{1/\varepsilon}}C_\varepsilon^{1/\varepsilon}\log(4Kt^2/\delta)$ times
Compute $\widehat{\mu}_{c,\varepsilon}(i,t)$ using (4) with $\alpha = \frac{1}{\varepsilon^{\frac{1}{1+\varepsilon}}}\frac{1}{(h^{-\varepsilon}C_\varepsilon \upsilon_\varepsilon)^{1/(1+\varepsilon)}}\left(\frac{\log(4Kt^2/\delta)}{t}\right)^{\frac{1}{1+\varepsilon}}$
**while** $|S| > 1$ **do**
  Remove all arms $i$ from $S$ that satisfy (11)
  Update $S$ as the remaining arms and set $t = t+1$
  Pull each arm in $S$ once and update $\widehat{\mu}_{c,\varepsilon}(i,t)$
**end while**
**output**: $S$

*Parameter Choices*: There is an inverse relationship between the parameters $h, \tau$ and the number of samples $n$ necessary to ensure the bound in (5). So depending on the tolerance to the cost of initial exploration, the values of $\tau$ and $h$ can be chosen. Ideally, we could reduce the number of free parameters, namely $h$ as $\tau > 0$ can be arbitrary, by using the bound (7). While this has the added advantage of being sharper, characterizing theoretical sample complexity results is harder owing to the implicit nature of $\delta$ and $n$ dependence. So we use the bound in (5) and tune the parameters accordingly.

**Theorem 6.1** (Sample Complexity with $\upsilon_\varepsilon$ known)**.** *Let the confidence parameter $\delta \in (0,1)$ be one of the inputs to the algorithm. For Algorithm 2 to output $S = \{*\}$, with probability at least $1-\delta$, the number of pulls is bounded by $O\Big(\sum_{i\in[K]}\log\Big(\frac{K}{\delta\Delta_i^{\frac{1+\varepsilon}{\varepsilon}}}\Big)\Big(\frac{1}{\Delta_i}\Big)^{\frac{1+\varepsilon}{\varepsilon}}\Big)$.*

The sample complexity attains the order of successive elimination algorithm (Even-Dar et al., 2006) when $\varepsilon = 1$. Owing to the $O(1/n^{\frac{\varepsilon}{1+\varepsilon}})$ dependence in the confidence bound instead of $O(1/\sqrt{n})$, we need to explore further to estimate the parameters to a high confidence. For a small $\Delta_i$ with $\varepsilon < 1$, this translates to the additional samples required to distinguish the arms when the variance does not exist.

The moment bound $\upsilon_\varepsilon$ is an input to the algorithm. If such a bound is not known, adaptive estimators derived in Section 4 can be used instead on all arms. With a minor modification, Algorithm 2 can be used in conjunction with Algorithm 1 to identify the best arm (see Algorithm 4). Essentially the upper bound for $\upsilon_\varepsilon$ in (10) can be chosen as the maximum of the moment bound estimates from all the arms, and the event $B_A$ can be defined as the union of the corresponding events for all the arms. With this change, we have the next Theorem.

**Theorem 6.2** (Sample Complexity with $\upsilon_\varepsilon$ unknown)**.** *Using Algorithm 1 to adapt to unknown moment, for Algorithm 2 to output $S = \{*\}$, on the event $B_A$, outside of a part with probability at most $\delta$, the number of pulls is bounded by $O\Big(\sum_{i\in[K]}\log\Big(\frac{K(1+2\log_\sigma A)}{\delta\Delta_i^{\frac{1+\varepsilon}{\varepsilon}}}\Big)\Big(\frac{1}{\Delta_i}\Big)^{\frac{1+\varepsilon}{\varepsilon}}\Big)$.*

As expected, the number of samples required is larger than in Theorem 6.1. The additional samples are required to compensate for the increase in the width of the confidence bound owing to Algorithm 1.

### 6.2. Phase-based Elimination with Catoni

It is clear that in Algorithm 2, the Catoni's estimate is wastefully computed at every step and the initial exploration phase is larger than necessary due to the union bound. We next provide a phase-based algorithm and achieve better sample complexity in terms of $\Delta_i$.

Algorithm 3 eliminates arms conditioned on the phase, while using the standard elimination criterion (11). For $i \in S$, let the *elimination criterion* for the phase based algorithm be specified as

$$\Big\{\widehat{\mu}_{c,\varepsilon}(i,t_m) + G^{\tau}_{\varepsilon,h}\Big(\frac{\delta}{2K(m+1)^2}\Big)\Big(\frac{1}{t_m}\Big)^{\frac{\varepsilon}{1+\varepsilon}}\Big\} \qquad (12)$$
$$< \max_{j\in S}\Big\{\widehat{\mu}_{c,\varepsilon}(j,t_m) - G^{\tau}_{\varepsilon,h}\Big(\frac{\delta}{2K(m+1)^2}\Big)\Big(\frac{1}{t_m}\Big)^{\frac{\varepsilon}{1+\varepsilon}}\Big\},$$

where $m$ is the phase index.

The key reason for the gain in performance is that, it can be shown that once the phase index is such that

$$\Big(\gamma^m G^{\tau}_{\varepsilon,h}(\delta/2Km^2)\Big)^{\frac{1+\varepsilon}{\varepsilon}} \geq \Big\lceil\Big(\frac{4G^{\tau}_{\varepsilon,h}\big(\frac{\delta}{2Km^2}\big)}{\Delta_i}\Big)^{\frac{1+\varepsilon}{\varepsilon}}\Big\rceil,$$

then arm $i$ will be eliminated before phase $m$, whence we obtain $m = O(\log_\gamma(4/\Delta_i))$.

**Algorithm 3** Phase-based Iterative Elimination with Catoni

1: **Input**: $\delta, K, h, \tau, \varepsilon, \upsilon_\varepsilon, \gamma(>1)$
2: **Initialization**: Set $S := [K]$, phase index $m = 0$
3: Set

$$\mathrm{t}_{\text{init}} := \left\lceil \frac{1+\varepsilon}{\varepsilon(1-h)} \frac{(1+\tau)^{(1+\varepsilon)/\varepsilon}}{\tau^{1/\varepsilon}} C_\varepsilon^{1/\varepsilon} \log(4K/\delta) \right\rceil$$

4: Set $t_0 = 0$ and

$$t_1 = \max\left\{ \left(\gamma G^{\tau}_{\varepsilon,h}(\delta/2K)\right)^{\frac{1+\varepsilon}{\varepsilon}}, \mathrm{t}_{\text{init}} \right\}$$

5: **while** $|S| > 1$ **do**
6: Increase phase index $m$ by 1
7: Sample **every** arm in $S$ for $\max\{t_m - t_{m-1}, 0\}$ times
8: Compute $\widehat{\mu}_{c,\varepsilon}(i, t_m)$ using (4) for $i \in S$ with

$$\alpha = \frac{1}{(\varepsilon h^{-\varepsilon} C_\varepsilon \upsilon_\varepsilon)^{1/(1+\varepsilon)}} \left(\frac{\log(4K(m+1)^2/\delta)}{t_m}\right)^{\frac{1}{1+\varepsilon}}$$

9: Remove all arms $i$ from $S$ which satisfy (12)
10: Update $S$ as the remaining arms and set $m = m + 1$
11: Set $t_m = \left(\gamma^m G^{\tau}_{\varepsilon,h}(\delta/2Km^2)\right)^{\frac{1+\varepsilon}{\varepsilon}}$
12: **end while**
13: **Output**: $S$

Algorithm 3 considers a much smaller initial exploration phase (Step 3), which also results in considerable saving in sample complexity when $K$ is large, and this coupled with order optimal phase length yields the result below.

**Theorem 6.3** (Sample Complexity with $\upsilon_\varepsilon$ known)**.** *For Algorithm 3 to output $S = \{*\}$, with probability at least $1-\delta$, the number of pulls is bounded by*

$$\max\left\{ O\left( \sum_{i\in[K]} \log\left(\frac{K\log(1/\Delta_i)}{\delta}\right) \frac{1}{\Delta_i^{\frac{1+\epsilon}{\epsilon}}} \right), K t_{\text{init}} \right\}.$$

The sample complexity attains the lower bound for $\varepsilon = 1$ when using the state of the art algorithm lil'UCB (Jamieson et al., 2014) and Exponential-Gap Elimination algorithm in Karnin et al. (2013). When the moment bound $\upsilon_\varepsilon$ is unknown, Algorithm 1 can be used in conjunction with Algorithm 3 to identify the best arm. The complexity increases to compensate for the increased width of the confidence bound.

**Theorem 6.4** (Sample Complexity with $\upsilon_\varepsilon$ unknown)**.** *For Algorithm 4 to output $S = \{*\}$, with probability at least $1-\delta$, the number of pulls is bounded by*

$$\max\left\{ O\left( \sum_{i\in[K]} \log\left(\frac{\widehat{K}\log(1/\Delta_i)}{\delta}\right) \frac{1}{\Delta_i^{\frac{1+\epsilon}{\epsilon}}} \right), K t_{\text{init}} \right\},$$

*where $\widehat{K} := K(1 + 2\log_\sigma A)$.*

**Algorithm 4** Adaptive Elimination with Catoni

1: **Input**: $\delta, K, h, \tau, \varepsilon, \gamma(>1), A > 1, \sigma > 1$
2: **Initialization**: Set $S := [K]$, phase index $m = 0$ $\{$*Replace Step 8. in Algorithm 3 by the following two steps*$\}$
3: **Step 8a.** Compute $\{\widehat{\mu}_{c,\varepsilon}(i, t_m), \vartheta_{\widehat{j}}(i)\}$ using Algorithm 1
4: **Step 8a.** Set $\upsilon_\varepsilon = \max_{i\in S} \vartheta_{\widehat{j}}(i)$ and the number of (pseudo) arms $K := K(1 + 2\log_\sigma A)$
5: **Output**: $S$

The key idea in Algorithm 4 is that the minimum length of initial exploration ($t_{\text{init}}$) does not depend on $\upsilon_\varepsilon$, which is not the case in the current state of the art M-estimators like Chen et al. (2021). Note that in Algorithm 1, for $i \in S$, $\widetilde{\mu}_{c,\varepsilon}(i, t_m)$ is computed in Algorithm 4 over $\{X_l(i)\}_{l \le t_m}$ with

$$\alpha = \frac{1}{\vartheta_{\widehat{j}}(i)} \frac{1}{(h^{-\varepsilon} C_\varepsilon \varepsilon)^{1/(1+\varepsilon)}} \left(\frac{\log(4K(m+1)^2/\delta)}{t_m}\right)^{\frac{1}{1+\varepsilon}}.$$

**Remark 3.** In practice, the solution of (4) that corresponds to the estimate, is computed by root-finding using fixed-point iterative updates. While Catoni's estimator provides tight bounds and good results, the space complexity is linear, which perhaps is the only impediment in a sequential decision making application.

## 7. Experimental Results

We empirically validate the performance improvement in using our proposed algorithm based on Catoni's estimator for $\varepsilon \in (0, 1]$ by comparing it with the state of the art best arm identification algorithms proposed in Yu et al. (2018).

### 7.1. Theoretical Comparison

- *Successive Elimination (SE) vs Phase-based (PB)*: (i) SE requires Catoni estimation by root finding at every time step on all valid arms, while PB computes once at the end of each phase. In addition, in Algorithm 3, the length of the phases are carefully modulated to obtain the best possible order for sample complexity. (ii) The sample complexity using SE is at best $O(\log(K/\delta\Delta_i)/\Delta_i^{1+\varepsilon/\varepsilon})$ as opposed to $O(\log\left(\frac{K\log(1/\Delta_i)}{\delta}\right)/\Delta_i^{1+\varepsilon/\varepsilon})$ using PB, which cannot be improved, for example when $\varepsilon = 1$. Note that PB has similar order as Karnin et al. (2013) and Jamieson et al. (2014) for $\varepsilon = 1$ and also works for $\varepsilon < 1$.
- *SE-TEA (Yu et al. (2018)) vs Algorithm 2*: (i) SE-TEA is based on raw moments, so sensitive to the location of the mean, which is not desirable in bandit applications. (ii) The constants in the robust mean estimator

| | | K = 10 | K = 50 | K = 100 | K = 500 | K = 1000 |
|---|---|---|---|---|---|---|
| $\delta = 0.05$ | $\varepsilon = 0.85$ | 2.57 | 21.07 | 43.95 | 259.16 | 565.71 |
| | $\varepsilon = 1$ | 0.10 | 0.43 | 1.01 | 7.23 | 15.37 |
| $\delta = 0.1$ | $\varepsilon = 0.85$ | 2.19 | 19.33 | 40.39 | 248.25 | 550.76 |
| | $\varepsilon = 1$ | 0.06 | 0.34 | 0.99 | 5.97 | 13.42 |

*Table 1.* Average sample complexity ($\times 10^5$) over 50 iterations over a range of arms for different hyper-parameter values using Algorithm 3.

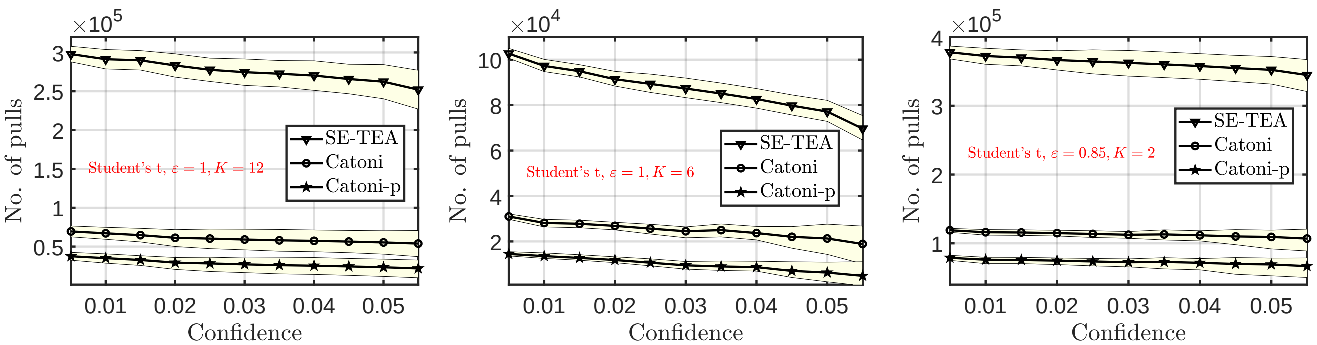


*Figure 1.* Average number of pulls for a fixed confidence over 50 iterations for different number of arms $K = 2, 6, 12$, and $\varepsilon = 1$ and $\varepsilon < 1$. Successive Elimination with Truncation Estimator(SE-TEA) algorithm is borrowed from Yu et al. (2018), Catoni corresponds to Algorithm 2, and Catoni-p to Algorithm 3. Significant performance improvement is observed using a sharper concentration bound.

are not even close to optimal (even disregarding the fact that central moment is much smaller than raw moment). These translate to poor empirical performance compared to Algorithm 3.

In terms of sample complexity bounds, Yu et al. (2018) obtain $O(\sum_i \log(K/\delta)/\Delta_i^{1+\varepsilon/\varepsilon})$. The reason for such a sharp, albeit **impossible** bound, is due to the fact that they use a *faulty* union bound argument (compare with Even-Dar et al. (2006)). For example when $\varepsilon = 1$, the lower bound is $\widetilde{O}(\sum_i \log\log(1/\Delta_i)/\Delta_i^2)$ as established in Jamieson et al. (2014). After correction, SE-TEA obtains the same order as Algorithm 2, however, with worse constants.

### 7.2. Numerical Comparison

The experimental setup is as follows: the rewards for all arms in all experiments are generated from the heavy tailed Student's-t distribution. The shifted mean of each arm is given using the following rule: $\mu(i) = 2 - (i - 1/K)^{0.6}$ for $i = 2, \cdots, K$ and $\mu(1) = 2$. The number of degrees of freedom for the first two figures from the left is $\nu = 3$ which corresponds to $\varepsilon = 1$ and $\upsilon_\varepsilon = 3$. The third figure uses $\nu = 2$ which has infinite variance and for $\varepsilon = 0.85$, an upper bound is $\upsilon_\varepsilon = 50$. Smaller values for $\tau$ and larger values for $h$ will have sharper bounds, and can be chosen depending on the tolerance to the cost of initial exploration. One choice parameters used is $\tau = 0.05$ and $h = 0.7$. Further, $\gamma(> 1)$ in Algorithm 3 is set to 1.1; this can be tuned further to improve performance. A significant improvement in performance using our algorithms is observed in Figure 1, more so in case of smaller $\varepsilon$, indicating the tightness for smaller $\varepsilon$. Table 1 shows the performance of Algorithm 3 for $K$ from 10 to 1000. Two facts contribute to algorithmic improvements over the state of the art. (i) Tighter bounds of $|\hat{\mu}_{c,\varepsilon} - \mu|$ obtained in Section 3 and Section 4. (ii) Phase-based elimination scheme (Algorithm 3).

We do not illustrate the performance of Algorithm 4, as Algorithm 1 can be incorporated into any of existing algorithms to adapt to the unknown parameters; and the algorithms using Catoni's estimator (Algorithm 2 & Algorithm 3) would continue to outperform by at least a similar margin owing to near-optimal confidence bounds.

## 8. Conclusion & Future Work

We provided a nearly-optimal M-estimator for computing the mean of a random variable having bounded $(1+\varepsilon)^{th}$ moment. This is based on the non-trivial extension of Catoni's estimator proposed in Catoni (2012) for the case of infinite variance. We also provided an algorithm to adapt to the situation of unknown moment bound using classical Lepskii's adaptive estimation method. We then provided adaptive best arm identification algorithms, based on this estimator, that have excellent empirical performance compared with trimmed mean baselines.

Useful and challenging future directions for the proposed M-estimator include extending the results to larger dimensions using ideas in Catoni and Giulini (2017) and Minsker (2018); and relaxing the requirement of $\varepsilon$ using ideas from Kagrecha et al. (2019); Ashutosh et al. (2021).

## References


Alekh Agarwal, Nan Jiang, Sham M Kakade, and Wen Sun. Reinforcement learning: Theory and algorithms. 2019.

Shubhada Agrawal, Sandeep Juneja, and Peter W. Glynn. Optimal $\delta$-correct best-arm selection for heavy-tailed distributions. In *Proceedings of the Conference on Algorithmic Learning Theory (ALT)*, pages 61–110, San Diego, CA, 2020.

Kumar Ashutosh, Jayakrishnan Nair, Anmol Kagrecha, and Krishna P. Jagannathan. Bandit algorithms: Letting go of logarithmic regret for statistical robustness. In *Proceedings of the 24th International Conference on Artificial Intelligence and Statistics (AISTATS)*, pages 622–630, Virtual Event, 2021.

Peter Auer and Ronald Ortner. Logarithmic online regret bounds for undiscounted reinforcement learning. In *Advances in Neural Information Processing Systems (NIPS)*, pages 49–56, Vancouver, Canada, 2006.

Peter Auer, Nicolò Cesa-Bianchi, and Paul Fischer. Finite-time analysis of the multiarmed bandit problem. *Mach. Learn.*, 47(2-3):235–256, 2002.

Sujay Bhatt, Guanhua Fang, and Ping Li. Offline change detection under contamination. In *Proceedings of the Thirty-Eighth Conference on Uncertainty in Artificial Intelligence (UAI)*, Eindhoven, The Netherlands, 2022a.

Sujay Bhatt, Guanhua Fang, Ping Li, and Gennady Samorodnitsky. Minimax M-estimation under adversarial corruption. In *Proceedings of the 39th International Conference on Machine Learning (ICML)*, Bartimore, MD, 2022b.

Lucien Birgé. An alternative point of view on lepski's method. *Lecture Notes-Monograph Series*, pages 113–133, 2001.

Christian Brownlees, Emilien Joly, and Gábor Lugosi. Empirical risk minimization for heavy-tailed losses. *The Annals of Statistics*, 43(6):2507–2536, 2015.

Sébastien Bubeck, Nicolò Cesa-Bianchi, and Gábor Lugosi. Bandits with heavy tail. *IEEE Trans. Inf. Theory*, 59(11): 7711–7717, 2013.

Valerii Vladimirovich Buldygin and YU V Kozachenko. *Metric characterization of random variables and random processes*, volume 188. American Mathematical Soc., 2000.

Apostolos N. Burnetas and Michael N. Katehakis. Optimal adaptive policies for markov decision processes. *Math. Oper. Res.*, 22(1):222–255, 1997.

Olivier Catoni. Challenging the empirical mean and empirical variance: a deviation study. In *Annales de l'IHP Probabilités et statistiques*, volume 48, pages 1148–1185, 2012.

Olivier Catoni and Ilaria Giulini. Dimension-free pac-bayesian bounds for matrices, vectors, and linear least squares regression. *arXiv preprint arXiv:1712.02747*, 2017.

Peng Chen, Xinghu Jin, Xiang Li, and Lihu Xu. A generalized catoni's m-estimator under finite $\alpha$-th moment assumption with $\alpha \in (1, 2)$. *arXiv preprint arXiv:2010.05008*, 2020.

Peng Chen, Xinghu Jin, Xiang Li, and Lihu Xu. A generalized catoni's m-estimator under finite $\alpha$-th moment assumption with $\alpha \in (1, 2)$. *Electronic Journal of Statistics*, 15(2):5523–5544, 2021.

Luc Devroye, Matthieu Lerasle, Gábor Lugosi, and Roberto I Oliveira. Sub-gaussian mean estimators. *The Annals of Statistics*, 44(6):2695–2725, 2016.

John C Duchi, Michael I Jordan, and Martin J Wainwright. Minimax optimal procedures for locally private estimation. *Journal of the American Statistical Association*, 113 (521):182–201, 2018.

Cynthia Dwork. Differential privacy: A survey of results. In Manindra Agrawal, Ding-Zhu Du, Zhenhua Duan, and Angsheng Li, editors, *Proceedings of the 5th International Conference on Theory and Applications of Models of Computation (TAMC)*, pages 1–19, Xi'an, China, 2008.

Cynthia Dwork and Aaron Roth. The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.*, 9(3-4):211–407, 2014.

Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor. Our data, ourselves: Privacy via distributed noise generation. In *Proceedings of the 25th Annual International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT)*, pages 486–503, St. Petersburg, Russia, 2006.

Eyal Even-Dar, Shie Mannor, and Yishay Mansour. Action elimination and stopping conditions for the multi-armed bandit and reinforcement learning problems. *J. Mach. Learn. Res.*, 7:1079–1105, 2006.

Claude-Nicolas Fiechter. Efficient reinforcement learning. In *Proceedings of the Seventh Annual ACM Conference on Computational Learning Theory (COLT)*, pages 88–97, New Brunswick, NJ, 1994.

David A Freedman. On tail probabilities for martingales. *The Annals of Probability*, pages 100–118, 1975.

Aurélien Garivier and Emilie Kaufmann. Optimal best arm identification with fixed confidence. In Vitaly Feldman, Alexander Rakhlin, and Ohad Shamir, editors, *Proceedings of the 29th Conference on Learning Theory (COLT)*, pages 998–1027, New York, 2016.

Peter Glynn and Sandeep Juneja. Selecting the best system and multi-armed bandits. *arXiv preprint arXiv:1507.04564*, 2018.

Peter J Huber. *Robust statistics*, volume 523. John Wiley & Sons, 2004.

Kevin Jamieson, Matthew Malloy, Robert Nowak, and Sébastien Bubeck. lil'ucb: An optimal exploration algorithm for multi-armed bandits. In *Conference on Learning Theory*, pages 423–439. PMLR, 2014.

Kevin G. Jamieson and Ameet Talwalkar. Non-stochastic best arm identification and hyperparameter optimization. In *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics (AISTATS)*, pages 240–248, Cadiz, Spain, 2016.

Anmol Kagrecha, Jayakrishnan Nair, and Krishna P. Jagannathan. Distribution oblivious, risk-aware algorithms for multi-armed bandits with unbounded rewards. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 11269–11278, Vancouver, Canada, 2019.

Zohar Shay Karnin, Tomer Koren, and Oren Somekh. Almost optimal exploration in multi-armed bandits. In *Proceedings of the 30th International Conference on Machine Learning (ICML)*, pages 1238–1246, Atlanta, GA, 2013.

Emilie Kaufmann and Wouter M. Koolen. Monte-carlo tree search by best arm identification. In *Advances in Neural Information Processing Systems (NIPS)*, pages 4897–4906, Long Beach, CA, 2017.

Gérard Kerkyacharian, Oleg Lepski, and Dominique Picard. Nonlinear estimation in anisotropic multi-index denoising. *Probability theory and related fields*, 121(2): 137–170, 2001.

Tor Lattimore and Csaba Szepesvári. *Bandit algorithms*. Cambridge University Press, 2020.

Guillaume Lecué and Matthieu Lerasle. Robust machine learning by median-of-means: theory and practice. *The Annals of Statistics*, 48(2):906–931, 2020.

OV Lepskii. Asymptotically minimax adaptive estimation. I: Upper bounds. optimally adaptive estimates. *Theory of Probability & Its Applications*, 36(4):682–697, 1992.

Ping Li. Stable random projections and conditional random sampling, two sampling techniques for modern massive datasets, PhD Thesis. 2007. URL https://hastie.su.domains/THESES/pingli_thesis.pdf.

Ping Li. One scan 1-bit compressed sensing. In *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics (AISTATS)*, pages 1515–1523, Cadiz, Spain, 2016.

Ping Li and Cun-Hui Zhang. Exact sparse recovery with L0 projections. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, pages 302–310, Chicago, IL, 2013.

Gábor Lugosi and Shahar Mendelson. Mean estimation and regression under heavy-tailed distributions: A survey. *Found. Comput. Math.*, 19(5):1145–1190, 2019.

Gábor Lugosi and Shahar Mendelson. Robust multivariate mean estimation: the optimality of trimmed mean. *The Annals of Statistics*, 49(1):393–410, 2021.

Stanislav Minsker. Sub-gaussian estimators of the mean of a random matrix with heavy-tailed entries. *The Annals of Statistics*, 46(6A):2871–2903, 2018.

Stanislav Minsker. Distributed statistical estimation and rates of convergence in normal approximation. *Electronic Journal of Statistics*, 13(2):5213–5252, 2019.

Roberto I Oliveira and Paulo Orenstein. The sub-gaussian property of trimmed means estimators. Technical report, Technical report, IMPA, 2019.

Sidney I Resnick. *Heavy-tail phenomena: probabilistic and statistical modeling*. Springer Science & Business Media, 2007.

Yevgeny Seldin, François Laviolette, Nicolò Cesa-Bianchi, John Shawe-Taylor, and Peter Auer. Pac-bayesian inequalities for martingales. *IEEE Trans. Inf. Theory*, 58 (12):7086–7093, 2012.

Qiang Sun, Wen-Xin Zhou, and Jianqing Fan. Adaptive huber regression. *Journal of the American Statistical Association*, 115(529):254–265, 2020.

Huazheng Wang, Qian Zhao, Qingyun Wu, Shubham Chopra, Abhinav Khaitan, and Hongning Wang. Global and local differential privacy for collaborative bandits. In *Proceedings of the Fourteenth ACM Conference on Recommender Systems (RecSys)*, pages 150–159, Virtual Event, Brazil, 2020.

Xiaotian Yu, Han Shao, Michael R. Lyu, and Irwin King. Pure exploration of multi-armed bandits with heavy-tailed payoffs. In *Proceedings of the Thirty-Fourth Conference* *on Uncertainty in Artificial Intelligence (UAI)*, pages 937–946, Monterey, CA, 2018.

Wen-Xin Zhou, Koushiki Bose, Jianqing Fan, and Han Liu. A new perspective on robust m-estimation: Finite sample theory and applications to dependence-adjusted multiple testing. *The Annals of Statistics*, 46(5):1904, 2018.

In this appendix, we include the following additional details to support the main claims and results. Specifically, we include detailed proofs of the main theorems to make the paper self-contained in Appendix A, and include an implication of the results in the upcoming important application of differential privacy in Appendix B.

## A. Proofs

**Proof of Lemma 3.1**

A necessary and sufficient condition for the existence of a function satisfying (3) is given by

$$(1 - x + C_\varepsilon x^{1+\varepsilon})(1 + x + C_\varepsilon x^{1+\varepsilon}) \geq 1, \; \forall\, x \geq 0.$$

Rearranging, this reduces to

$$2C_\varepsilon x^{1+\varepsilon} + C_\varepsilon^2 x^{2(1+\varepsilon)} \geq x^2, \; \forall\, x \geq 0,$$

which is equivalent to the condition

$$C_\varepsilon^2 x^{2\varepsilon} + 2C_\varepsilon x^{\varepsilon - 1} > 1, \; \forall\, x > 0. \tag{13}$$

The minimum of the expression in the left hand side over $x > 0$ is achieved at

$$x_* = \left(\frac{1-\varepsilon}{C_\varepsilon \varepsilon}\right)^{\frac{1}{1+\varepsilon}},$$

and substituting this value in (13) and solving for $C_\varepsilon$ produces the desired result. □

**Proof of Theorem 3.2**

As in Catoni (2012), define

$$r_n(\theta) = \sum_{i=1}^{n} \psi\Big(\alpha(X_i - \theta)\Big),$$

and note that $r_n(\theta)$ is non-increasing in $\theta \in \mathbb{R}$. Using the upper bound on the influence function in (3),

$$\begin{aligned}
\mathbb{E}\Big[\exp(r_n(\theta))\Big] &= \Big(\mathbb{E}\Big[\exp\Big(\psi\Big(\alpha(X_1 - \theta)\Big)\Big]\Big)^n \\
&\leq \Big(\mathbb{E}\Big[1 + \alpha(X_1 - \theta) + C_\varepsilon \alpha^{1+\varepsilon} |X_1 - \theta|^{1+\varepsilon}\Big]\Big)^n \\
&= \Big(1 + \alpha(\mu - \theta) + C_\varepsilon \alpha^{1+\varepsilon} \mathbb{E}|X_1 - \theta|^{1+\varepsilon}\Big)^n.
\end{aligned}$$

We will use a convexity upper bound as follows. For $a, b \geq 0$ and $0 < h < 1$,

$$\begin{aligned}
(a+b)^{1+\varepsilon} &= \Big(h\frac{a}{h} + (1-h)\frac{b}{1-h}\Big)^{1+\varepsilon}, \\
&\leq h\Big(\frac{a}{h}\Big)^{1+\varepsilon} + (1-h)\Big(\frac{b}{1-h}\Big)^{1+\varepsilon} = \frac{a^{1+\varepsilon}}{h^\varepsilon} + \frac{b^{1+\varepsilon}}{(1-h)^\varepsilon}.
\end{aligned} \tag{CB}$$

Therefore, for any $0 < h < 1$,

$$\mathbb{E}|X_1 - \theta|^{1+\varepsilon} \leq h^{-\varepsilon} \mathbb{E}|X_1 - \mu|^{1+\varepsilon} + (1-h)^{-\varepsilon} |\mu - \theta|^{1+\varepsilon}. \tag{14}$$

This leads to worse constants than in Catoni (2012), and is the price to pay for the generalization. Using the above bound, we obtain

$$\begin{aligned}
\mathbb{E}\Big[\exp(r_n(\theta))\Big] &\leq \Big(1 + \alpha(\mu - \theta) + h^{-\varepsilon} C_\varepsilon \alpha^{1+\varepsilon} v_\varepsilon + C_\varepsilon \alpha^{1+\varepsilon} (1-h)^{-\varepsilon} |\mu - \theta|^{1+\varepsilon}\Big)^n \\
&\leq \exp\Big(\alpha n(\mu - \theta) + n h^{-\varepsilon} C_\varepsilon \alpha^{1+\varepsilon} v_\varepsilon + n C_\varepsilon \alpha^{1+\varepsilon} (1-h)^{-\varepsilon} |\mu - \theta|^{1+\varepsilon}\Big).
\end{aligned}$$

Similarly, using the lower bound on the influence function in (3), we obtain by symmetric arguments

$$\mathbb{E}\Big[\exp(-r_n(\theta))\Big] \leq \exp\Big(-\alpha n(\mu - \theta) + n h^{-\varepsilon} C_\varepsilon \alpha^{1+\varepsilon} v_\varepsilon + n C_\varepsilon \alpha^{1+\varepsilon} (1-h)^{-\varepsilon} |\mu - \theta|^{1+\varepsilon}\Big).$$

Let $\delta \in (0,1)$. As in Catoni (2012), we define

$$B_+(\theta) = (\mu-\theta) + h^{-\varepsilon} C_\varepsilon \alpha^\varepsilon \upsilon_\varepsilon + C_\varepsilon \alpha^\varepsilon (1-h)^{-\varepsilon} |\mu-\theta|^{1+\varepsilon} + \frac{\log(2/\delta)}{\alpha n},$$
$$B_-(\theta) = (\mu-\theta) - h^{-\varepsilon} C_\varepsilon \alpha^\varepsilon \upsilon_\varepsilon - C_\varepsilon \alpha^\varepsilon (1-h)^{-\varepsilon} |\mu-\theta|^{1+\varepsilon} - \frac{\log(2/\delta)}{\alpha n}.$$

By the exponential Markov inequality, we have

$$\mathbb{P}\Big\{ r_n(\theta) \geq n\alpha B_+(\theta) \Big\} \leq \frac{\mathbb{E}\Big[\exp(r_n(\theta))\Big]}{\exp(\alpha n B_+(\theta))} \leq \delta/2,$$
$$\mathbb{P}\Big\{ r_n(\theta) \leq n\alpha B_-(\theta) \Big\} \leq \frac{\mathbb{E}\Big[\exp(-r_n(\theta))\Big]}{\exp(-\alpha n B_-(\theta))} \leq \delta/2. \tag{15}$$

Note that the function $B_+$ is a strictly convex function of $\theta$ and $B_+(\theta) \to \infty$ as $|\theta| \to \infty$. Therefore, $B_+$ has a unique minimum on $\mathbb{R}$, which is achieved at

$$\theta_* = \mu + \frac{1-h}{\alpha}\Big(\frac{1}{(1+\varepsilon)C_\varepsilon}\Big)^{\frac{1}{\varepsilon}},$$

so that

$$\min_{\theta\in\mathbb{R}} B_+(\theta) = B_+(\theta_*) = h^{-\varepsilon}\alpha^\varepsilon C_\varepsilon \upsilon_\varepsilon - \frac{\varepsilon}{1+\varepsilon}\frac{1-h}{\alpha}\Big(\frac{1}{(1+\varepsilon)C_\varepsilon}\Big)^{\frac{1}{\varepsilon}} + \frac{\log(2/\delta)}{\alpha n}.$$

Suppose that this minimum is non-positive, i.e.

$$h^{-\varepsilon}\alpha^{1+\varepsilon} C_\varepsilon \upsilon_\varepsilon + \frac{\log(2/\delta)}{n} \leq \frac{\varepsilon}{1+\varepsilon}(1-h)\Big(\frac{1}{(1+\varepsilon)C_\varepsilon}\Big)^{\frac{1}{\varepsilon}}. \tag{16}$$

Then the equation

$$B_+(\theta) = 0 \tag{17}$$

has a real root, and, if the inequality is strict, it has two real roots. Since $B_+(\mu) > 0$ and $\theta_* > \mu$, the roots are larger than $\mu$. Let $\theta_+(\alpha) \in [\mu, \theta_*]$ denote the smallest of these roots. Our next step is to bound the difference $\theta_+(\alpha) - \mu$.

Let us denote

$$K = (1-h)^{-\varepsilon}\alpha^\varepsilon C_\varepsilon,\ M = h^{-\varepsilon}\alpha^\varepsilon C_\varepsilon \upsilon_\varepsilon + \frac{\log(2/\delta)}{\alpha n},$$

and introduce a variable $z = \theta - \mu$, $\theta \geq \mu$. Now the equation (17) is rewritten as

$$K z^{1+\varepsilon} - z + M = 0,\ z \geq 0. \tag{18}$$

Denoting further

$$D = K^{\frac{1}{\varepsilon}} M,$$

and introducing the variable $y = K^{\frac{1}{\varepsilon}} z$ transforms (18) into the equation

$$y^{1+\varepsilon} - y + D = 0,\ y \geq 0. \tag{19}$$

Let the associated function $f$ be defined as

$$f(y) := y^{1+\varepsilon} - y + D.$$

If the condition

$$D \leq \frac{\tau^{1/\varepsilon}}{(1+\tau)^{\frac{1+\varepsilon}{\varepsilon}}} \tag{20}$$

holds for $\tau > 0$, then $f((1+\tau)D) \leq 0$, so the equation (19) has a positive solution $y(D)$ satisfying

$$y(D) \leq (1+\tau)D,$$

which is the same as

$$\theta_+(\alpha) - \mu \le (1+\tau)M.$$

Equivalently,

$$\theta_+(\alpha) - \mu \le (1+\tau)\Big(h^{-\varepsilon}\alpha^{\varepsilon} C_\varepsilon \upsilon_\varepsilon + \frac{\log(2/\delta)}{\alpha n}\Big),$$

assuming (20) holds. Note that the latter condition can be rewritten in the form

$$h^{-\varepsilon}\alpha^{1+\varepsilon} C_\varepsilon \upsilon_\varepsilon + \frac{\log(2/\delta)}{n} \le \frac{\tau^{1/\varepsilon}}{(1+\tau)^{(1+\varepsilon)/\varepsilon}}(1-h)C_\varepsilon^{-1/\varepsilon}. \tag{21}$$

This condition coincides with (16) for $\tau = 1/\varepsilon$ and is stronger than the latter for other values of $\tau$.

Similar analysis using $B_-(\theta)$ shows that the roots of the equation $B_-(\theta) = 0$ are smaller than $\mu$ and the largest of the roots, $\theta_-(\alpha)$, satisfies

$$\mu - \theta_-(\alpha) \le (1+\tau)\Big(h^{-\varepsilon}\alpha^{\varepsilon} C_\varepsilon \upsilon_\varepsilon + \frac{\log(2/\delta)}{\alpha n}\Big),$$

once again assuming that (21) holds.

Assuming (21) holds, from (15) with $\theta = \theta_+(\alpha)$, there is an event $\mathcal{E}_1$ with $\mathbb{P}\{\mathcal{E}_1\} \ge 1 - \delta/2$ such that, on $\mathcal{E}_1$,

$$r_n(\theta_+(\alpha)) < B_+(\theta_+(\alpha)) = 0.$$

Since $r_n$ is a non-increasing function of $\theta$, we see that, on $\mathcal{E}_1$, $\widehat{\mu}_{c,\varepsilon} < \theta_+(\alpha)$. It follows that on $\mathcal{E}_1$,

$$\widehat{\mu}_{c,\varepsilon} - \mu < (1+\tau)\Big(h^{-\varepsilon}\alpha^{\varepsilon} C_\varepsilon \upsilon_\varepsilon + \frac{\log(2/\delta)}{\alpha n}\Big). \tag{22}$$

Similarly, from (15) with $\theta = \theta_-(\alpha)$, there is an event $\mathcal{E}_2$ with $\mathbb{P}\{\mathcal{E}_2\} \ge 1 - \delta/2$ such that, on $\mathcal{E}_2$,

$$r_n(\theta_-(\alpha)) > B_-(\theta_-(\alpha)) = 0.$$

Again the monotonicity of $r$ implies that on $\mathcal{E}_2$, $\widehat{\mu}_{c,\varepsilon} > \theta_-(\alpha)$. It follows that on $\mathcal{E}_2$,

$$\mu - \widehat{\mu}_{c,\varepsilon} < (1+\tau)\Big(h^{-\varepsilon}\alpha^{\varepsilon} C_\varepsilon \upsilon_\varepsilon + \frac{\log(2/\delta)}{\alpha n}\Big). \tag{23}$$

It follows from (22) and (23) that, if (21) holds, then with probability at least $1-\delta$,

$$|\widehat{\mu}_{c,\varepsilon} - \mu| < (1+\tau)\Big(h^{-\varepsilon}\alpha^{\varepsilon} C_\varepsilon \upsilon_\varepsilon + \frac{\log(2/\delta)}{\alpha n}\Big). \tag{24}$$

*Choice of* $\alpha$:

We start by writing $\alpha$ in the form

$$\alpha = A\Big(\frac{\log(2/\delta)}{n}\Big)^{\frac{1}{1+\varepsilon}}, \tag{25}$$

for some $A > 0$ to be chosen. This form of $\alpha$ leads to the bound

$$|\widehat{\mu}_{c,\varepsilon} - \mu| < (1+\tau)\Big(\frac{\log(2/\delta)}{n}\Big)^{\frac{\varepsilon}{1+\varepsilon}}\Big(h^{-\varepsilon}A^{\varepsilon} C_\varepsilon \upsilon_\varepsilon + 1/A\Big). \tag{26}$$

Note that the condition (21) with $\alpha$ of the form (25) requires

$$h^{-\varepsilon} C_\varepsilon \upsilon_\varepsilon A^{1+\varepsilon}\frac{\log(2/\delta)}{n} + \frac{\log(2/\delta)}{n} \le (1-h)\frac{\tau^{1/\varepsilon}}{(1+\tau)^{(1+\varepsilon)/\varepsilon}} C_\varepsilon^{-1/\varepsilon}. \tag{27}$$

This implies that if $n$ satisfies

$$n \ge (1-h)^{-1}\frac{1+\varepsilon}{\varepsilon}\frac{(1+\tau)^{(1+\varepsilon)/\varepsilon}}{\tau^{1/\varepsilon}} C_\varepsilon^{1/\varepsilon}\log(2/\delta), \tag{28}$$

then the value of $A$ that minimizes the right hand side in (26),

$$A = \frac{1}{\varepsilon^{\frac{1}{1+\varepsilon}}} \frac{1}{(h^{-\varepsilon} C_\varepsilon \upsilon_\varepsilon)^{1/(1+\varepsilon)}}$$

satisfies (27). Using this choice of $A$, we conclude that for all $\delta \in (0,1)$ and for all $n$ satisfying (28), we have with probability at least $1-\delta$,

$$|\widehat{\mu}_{c,\varepsilon} - \mu| < (1+\tau)\frac{1+\varepsilon}{\varepsilon^{\varepsilon/(1+\varepsilon)}}(h^{-\varepsilon} C_\varepsilon \upsilon_\varepsilon)^{1/(1+\varepsilon)} \Big(\frac{\log(2/\delta)}{n}\Big)^{\frac{\varepsilon}{1+\varepsilon}}.$$

Finally, recalling that $C_\varepsilon = \left(\frac{\varepsilon}{1+\varepsilon}\right)^{\frac{1+\varepsilon}{2}} \left(\frac{1-\varepsilon}{\varepsilon}\right)^{\frac{1-\varepsilon}{2}}$, we obtain the required non-asymptotic confidence bound

$$|\widehat{\mu}_{c,\varepsilon} - \mu| < (1+\tau)(1+\varepsilon)^{1/2}(1-\varepsilon)^{\frac{1}{1+\varepsilon}-\frac{1}{2}}(h^{-\varepsilon} \upsilon_\varepsilon)^{1/(1+\varepsilon)} \Big(\frac{\log(2/\delta)}{n}\Big)^{\frac{\varepsilon}{1+\varepsilon}}$$

with probability at least $1-\delta$. □

**Proof of Theorem 6.1**

First, note that from Theorem 3.2,

$$\mathbb{P}\Big\{|\widehat{\mu}_{c,\varepsilon} - \mu| \le G^\tau_{\varepsilon,h}\Big(\frac{\delta}{2Kt^2}\Big)\Big(\frac{1}{t}\Big)^{\frac{\varepsilon}{1+\varepsilon}}\Big\} \ge 1 - \frac{\delta}{2Kt^2}.$$

From union bound, we therefore have that with probability at least $1-\delta$ for any time $t$ and any arm $i \in [K]$,

$$|\widehat{\mu}_{c,\varepsilon}(i,t) - \mu| \le G^\tau_{\varepsilon,h}\Big(\frac{\delta}{2Kt^2}\Big)\Big(\frac{1}{t}\Big)^{\frac{\varepsilon}{1+\varepsilon}}.$$

From Algorithm 2, a sub-optimal arm $i \in [K]$ is eliminated when the number of pulls $n$ is such that (cf. (11))

$$\Big\{\widehat{\mu}_{c,\varepsilon}(i,n) + G^\tau_{\varepsilon,h}\Big(\frac{\delta}{2Kn^2}\Big)\Big(\frac{1}{n}\Big)^{\frac{\varepsilon}{1+\varepsilon}}\Big\} \le \max_{j\in S}\Big\{\widehat{\mu}_{c,\varepsilon}(j,n) - G^\tau_{\varepsilon,h}\Big(\frac{\delta}{2Kn^2}\Big)\Big(\frac{1}{n}\Big)^{\frac{\varepsilon}{1+\varepsilon}}\Big\},$$
$$\max_{j\in S}\Big\{\widehat{\mu}_{c,\varepsilon}(j,n)\Big\} - \widehat{\mu}_{c,\varepsilon}(i,n) \ge 2G^\tau_{\varepsilon,h}\Big(\frac{\delta}{2Kn^2}\Big)\Big(\frac{1}{n}\Big)^{\frac{\varepsilon}{1+\varepsilon}}.$$

Note that on the relevant event the optimal arm is not eliminated in such as fashion, lest condition (11) leads to a contradiction. Furthermore, on this event the latter condition will be automatically satisfied if

$$\Delta_i \ge 4G^\tau_{\varepsilon,h}\Big(\frac{\delta}{2Kn^2}\Big)\Big(\frac{1}{n}\Big)^{\frac{\varepsilon}{1+\varepsilon}}.$$

With $q = 4(1+\tau)(1+\varepsilon)^{1/2}(1-\varepsilon)^{\frac{1}{1+\varepsilon}-\frac{1}{2}}(h^{-\varepsilon}\upsilon_\varepsilon)^{1/(1+\varepsilon)}$, this can be rewritten as

$$\Delta_i \ge q\Big(\log(4Kn^2/\delta)\Big)^{\frac{\varepsilon}{1+\varepsilon}}\Big(\frac{1}{n}\Big)^{\frac{\varepsilon}{1+\varepsilon}},$$
$$\Big(\frac{\Delta_i}{q}\Big)^{\frac{1+\varepsilon}{\varepsilon}} \ge \frac{\log(4Kn^2/\delta)}{n}.$$

Therefore, on the relevant event a suboptimal arm $i$ will eliminated after at most

$$O\Bigg(\log\Big(\frac{K}{\delta\Delta_i^{\frac{1+\varepsilon}{\varepsilon}}}\Big)\frac{1}{\Delta_i^{\frac{1+\varepsilon}{\varepsilon}}}\Bigg)$$

pulls.The result follows. □

**Proof of Theorem 6.3**

From Theorem 3.2, with probability at least $1-\frac{\delta}{Km^2}$, we have each of the following events

$$\widehat{\mu}_{c,\varepsilon}(i,t) \le \mu(i) + G^{\tau}_{\varepsilon,h}(\frac{\delta}{2Km^2})\Big(\frac{1}{t}\Big)^{\frac{\varepsilon}{1+\varepsilon}}, \tag{29}$$

and

$$\widehat{\mu}_{c,\varepsilon}(*,t) \ge \mu(*) - G^{\tau}_{\varepsilon,h}(\frac{\delta}{2Km^2})\Big(\frac{1}{t}\Big)^{\frac{\varepsilon}{1+\varepsilon}}. \tag{30}$$

We'll first establish that if the number of pulls of sub-optimal arm $i$ is larger than

$$n_{m_i} := \Big\lceil \Big(\frac{4G^{\tau}_{\varepsilon,h}(\frac{\delta}{2Km^2})}{\Delta_i}\Big)^{\frac{1+\varepsilon}{\varepsilon}} \Big\rceil,$$

at phase $m$ ($m$ will be determined later, i.e. (31)), then arm $i$ will be eliminated. Indeed, by (29) and (30), (and expressing $\Delta_i$ as a function $n_{m_i}$ and $G^{\tau}_{\varepsilon,h}$), we have

$$\begin{aligned}
\widehat{\mu}_{c,\varepsilon}(i,n_{m_i}) + G^{\tau}_{\varepsilon,h}(\frac{\delta}{2Km^2})\Big(\frac{1}{n_{m_i}}\Big)^{\frac{\varepsilon}{1+\varepsilon}} &\le \mu(i) + 2G^{\tau}_{\varepsilon,h}(\frac{\delta}{2Km^2})\Big(\frac{1}{n_{m_i}}\Big)^{\frac{\varepsilon}{1+\varepsilon}}\\
&< \mu(i) + \Delta_i - 2G^{\tau}_{\varepsilon,h}(\frac{\delta}{2Km^2})\Big(\frac{1}{n_{m_i}}\Big)^{\frac{\varepsilon}{1+\varepsilon}}\\
&= \mu(*) - 2G^{\tau}_{\varepsilon,h}(\frac{\delta}{2Km^2})\Big(\frac{1}{n_{m_i}}\Big)^{\frac{\varepsilon}{1+\varepsilon}}\\
&< \widehat{\mu}_{c,\varepsilon}(*,n_{m_i}) - G^{\tau}_{\varepsilon,h}(\frac{\delta}{2Km^2})\Big(\frac{1}{n_{m_i}}\Big)^{\frac{\varepsilon}{1+\varepsilon}}.
\end{aligned}$$

Note that the optimal arm will not be eliminated if (29) and (30) hold, as then the required relation

$$\widehat{\mu}^{\tau}_{c,\varepsilon}(i,n_{m_i}) - G^{\tau}_{\varepsilon,h}(\frac{\delta}{2Km^2})\Big(\frac{1}{n_{m_i}}\Big)^{\frac{\varepsilon}{1+\varepsilon}} > \widehat{\mu}^{\tau}_{c,\varepsilon}(*,n_{m_i}) + G^{\tau}_{\varepsilon,h}(\frac{\delta}{2Km^2})\Big(\frac{1}{n_{m_i}}\Big)^{\frac{\varepsilon}{1+\varepsilon}}$$

leads to a contradiction $\mu(i) \ge \mu(*)$ for a sub-optimal arm $i$. This implies that once the phase index $m_i^0$ is such that

$$m_i^0 := \min\Big\{m : \Big(\gamma^m G^{\tau}_{\varepsilon,h}(\delta/2Km^2)\Big)^{\frac{1+\varepsilon}{\varepsilon}} > n_{m_i}\Big\}, \tag{31}$$

then arm $i$ will be eliminated before phase $m_i^0$ with probability at least $1-\sum_{m=2}^{m_0}\frac{\delta}{Km^2} \ge 1-\frac{\delta}{K}$.

Solving (31), we have $\Big(\gamma^m G^{\tau}_{\varepsilon,h}(\delta/2Km^2)\Big)^{\frac{1+\varepsilon}{\varepsilon}} > \Big\lceil \Big(\frac{4G^{\tau}_{\varepsilon,h}(\frac{\delta}{2Km^2})}{\Delta_i}\Big)^{\frac{1+\varepsilon}{\varepsilon}}\Big\rceil$. It suffices to have $m > \log_\gamma(4/\Delta_i)$. We then know $m_i^0$ is bounded by $\lceil \log_\gamma(4/\Delta_i)\rceil$. Therefore, the number of times pulling arm $i$ is bounded by

$$\begin{aligned}
\Big(\gamma^{m_i^0} G^{\tau}_{\varepsilon,h}(\frac{\delta}{2Km^2})\Big)^{\frac{1+\epsilon}{\epsilon}} &= \Big(G^{\tau}_{\varepsilon,h}(\frac{\delta}{2K(m_i^0)^2})\Big)^{\frac{1+\epsilon}{\epsilon}} (\gamma^{m_i^0})^{\frac{1+\epsilon}{\epsilon}}\\
\le\quad & 2^{\frac{1+\epsilon}{\epsilon}}\Big(G^{\tau}_{\varepsilon,h}(\frac{\delta}{2K(m_i^0)^2})\Big)^{\frac{1+\epsilon}{\epsilon}}(\frac{4}{\Delta_i})^{\frac{1+\epsilon}{\epsilon}}\\
\le\quad & 2^{\frac{1+\epsilon}{\epsilon}}\big(2q\log(\frac{2K\log_\gamma(4/\Delta_i)}{\delta})\big)(\frac{4}{\Delta_i})^{\frac{1+\epsilon}{\epsilon}},
\end{aligned} \tag{32}$$

where $q=(1+\tau)(1+\varepsilon)^{1/2}(1-\varepsilon)^{\frac{1}{1+\varepsilon}-\frac{1}{2}}(h^{-\varepsilon}\upsilon_\varepsilon)^{1/(1+\varepsilon)}$. The result follows. $\square$

**Proof of Theorem 4.1**

Recall that on the event $B_A$, $\vartheta_{\min} > 0$ and $\vartheta_{\max} > 0$ sandwich the true bound $\upsilon_\varepsilon^{1/(1+\varepsilon)}$. Let $\vartheta_j = \vartheta_{\min}\sigma^j$. Define

$$\mathcal{J} := \mathcal{J}_\sigma = \{j = 0, 1, 2, \cdots, \ : \vartheta_{\min} \le \vartheta_j \le \sigma\vartheta_{\max}\}.$$

Note that the cardinality of $\mathcal{J}$ is upper bounded as

$$|\mathcal{J}| \le \Big(1 + \log_\sigma\Big(\frac{\vartheta_{\max}}{\vartheta_{\min}}\Big)\Big) = \big(1 + 2\log_\sigma A\big).$$

In the range $n$ specified in the theorem, with a slight abuse of notation, we have from (8) that for all $k \in \mathcal{J}$

$$|\widetilde{\mu}_{\vartheta_k} - \mu| \le \Phi(\varepsilon, \tau, h)\vartheta_k\Big(\frac{\log(2(1+\log_\sigma A^2)/\delta)}{n}\Big)^{\frac{\varepsilon}{1+\varepsilon}}$$

if $\upsilon_\varepsilon \le \vartheta_k$. Define

$$\widehat{j} = \min\Big\{j \in \mathcal{J} : |\widetilde{\mu}_{\vartheta_k} - \widetilde{\mu}_{\vartheta_j}| \le 2\Phi(\varepsilon, \tau, h)\vartheta_j\Big(\frac{\log(2(1+\log_\sigma A^2)/\delta)}{n}\Big)^{\frac{\varepsilon}{1+\varepsilon}}, \ \forall\, k > j,\ k \in \mathcal{J}\Big\}.$$

On the event $B_A$, we define $j^* = \inf\{j \in \mathcal{J} : \vartheta_j \ge \upsilon_\varepsilon^{\frac{1}{1+\varepsilon}}\}$, whence we have that

$$\upsilon_\varepsilon^{\frac{1}{1+\varepsilon}} \le \vartheta_{j^*} \le \sigma\upsilon_\varepsilon^{\frac{1}{1+\varepsilon}}.$$

Consider the event $B_A \cap \{\widehat{j} > j^*\}$, so that

$$\begin{aligned} B_A \cap \{\widehat{j} > j^*\} &\subseteq B_A \cap \bigcup_{j \in \mathcal{J}: j > j^*} \Big\{|\widetilde{\mu}_{\vartheta_j} - \widetilde{\mu}_{\vartheta_{j^*}}| > 2\Phi(\varepsilon, \tau, h)\vartheta_j\Big(\frac{\log(2(1+\log_\sigma A^2)/\delta)}{n}\Big)^{\frac{\varepsilon}{1+\varepsilon}}\Big\} \\ &\subseteq B_A \cap \bigcup_{j \in \mathcal{J}: j \ge j^*} \Big\{|\widetilde{\mu}_{\vartheta_j} - \mu| > \Phi(\varepsilon, \tau, h)\vartheta_j\Big(\frac{\log(2(1+\log_\sigma A^2)/\delta)}{n}\Big)^{\frac{\varepsilon}{1+\varepsilon}}\Big\}. \end{aligned}$$

Defining an event $\mathcal{E}$ as

$$\mathcal{E} = B_A \cap \bigcap_{j \in \mathcal{J}: j \ge j^*} \Big\{|\widetilde{\mu}_{\vartheta_j} - \widetilde{\mu}| \le \Phi(\varepsilon, \tau, \delta)\vartheta_j\Big(\frac{\log(2(1+\log_\sigma A^2)/\delta)}{n}\Big)^{\frac{\varepsilon}{1+\varepsilon}}\Big\},$$

we have that $\mathcal{E} \subseteq B_A \cap \{\widehat{j} \le j^*\}$. Using Theorem 3.2, the probability of $B_A \cap \mathcal{E}^c$ is upper bounded as

$$\mathbb{P}\{\mathcal{E}^c\} \le \sum_{j \in \mathcal{J}: \vartheta_j \ge \upsilon_\varepsilon^{\frac{1}{1+\varepsilon}}} \mathbb{P}\Big\{|\widetilde{\mu}_{\vartheta_j} - \widetilde{\mu}| > \Phi(\varepsilon, \tau, \delta)\vartheta_j\Big(\frac{\log(2(1+\log_\sigma A^2)/\delta)}{n}\Big)^{\frac{\varepsilon}{1+\varepsilon}}\Big\} \le \delta.$$

On the event $\mathcal{E}$, we have that $\widehat{j} \le j^*$, so that

$$\begin{aligned} |\widehat{\mu}_{\vartheta_{\widehat{j}}} - \mu| &\le |\widehat{\mu}_{\vartheta_{\widehat{j}}} - \widehat{\mu}_{\vartheta_{j^*}}| + |\widehat{\mu}_{\vartheta_{j^*}} - \mu| \\ &\le 3\Phi(\varepsilon, \tau, h)\vartheta_{j^*}\left(\frac{\log\Big(2\Big(1 + \log_a\Big(\frac{\vartheta_{\max}}{\vartheta_{\min}}\Big)\Big)/\delta\Big)}{n}\right)^{\frac{\varepsilon}{1+\varepsilon}}. \end{aligned}$$

The result follows by noting that $\vartheta_{j^*} \le \sigma\upsilon_\varepsilon^{\frac{1}{1+\varepsilon}}$. □

**Proof of Corollary 3.3**

The condition (16) is satisfied whenever $n$ and $\alpha$ are chosen as required. Under this condition, we obtain by (24) that

$$|\widehat{\mu}_{c,\varepsilon} - \mu| < \frac{1+\varepsilon}{\varepsilon}(h^{-\varepsilon}C_\varepsilon)^{1/(1+\varepsilon)}\Big(\frac{\upsilon_\varepsilon^{1/(1+\varepsilon)}}{n^{\frac{\varepsilon}{1+\varepsilon}}}\Big)(1 + \log(2/\delta))$$

with probability at least $1 - \delta$. Using the chosen value of $C_\varepsilon$, the result follows. □

**Proof of Proposition 5.1**

*Proof.* For any $i \leq n$, using the upper bound on the influence function, we have for $h \in (0,1)$

$$\begin{aligned}\mathbb{E}[\exp(\psi(\alpha(X_i-\theta)))] &\leq \mathbb{E}\Big[1+\alpha(X_i-\theta)+C_\varepsilon\alpha^{1+\varepsilon}|X_i-\theta|^{1+\varepsilon}\Big]\\ &\leq \Big(1+\alpha(\mu-\theta)+h^{-\varepsilon}C_\varepsilon\alpha^{1+\varepsilon}\mathbb{E}[|X_i-\mu|^{1+\varepsilon}]+C_\varepsilon\alpha^{1+\varepsilon}(1-h)^{-\varepsilon}|\mu-\theta|^{1+\varepsilon}\Big)\\ &\leq \exp\Big(\alpha(\mu-\theta)+h^{-\varepsilon}C_\varepsilon\alpha^{1+\varepsilon}\mathbb{E}[|X_i-\mu|^{1+\varepsilon}]+C_\varepsilon\alpha^{1+\varepsilon}(1-h)^{-\varepsilon}|\mu-\theta|^{1+\varepsilon}\Big). \end{aligned} \tag{33}$$

We shall use the standard methodology in dealing with martingales (Freedman, 1975; Seldin et al., 2012) to obtain a deviation probability for $r(\theta)$. Let $Z_0 = 1$, and for $i \geq 1$, let

$$Z_i = Z_{i-1}\exp(\psi(\alpha(X_i-\theta)))\exp\left(-\Big(\alpha(\mu-\theta)+h^{-\varepsilon}C_\varepsilon\alpha^{1+\varepsilon}\mathbb{E}[|X_i-\mu|^{1+\varepsilon}]+C_\varepsilon\alpha^{1+\varepsilon}(1-h)^{-\varepsilon}|\mu-\theta|^{1+\varepsilon}\Big)\right).$$

Clearly, using (33) and the definition of $Z_i$,

$$\mathbb{E}[Z_n] \leq \cdots \leq \mathbb{E}[Z_i] \leq \cdots \leq \mathbb{E}[Z_0] = 1, \tag{34}$$

where $Z_n$ is given as

$$Z_n = \exp\Big(\sum_{i=1}^{n}\psi(\alpha(X_i-\theta))\Big)\times\exp\left(-\Big(n\alpha(\mu-\theta)+nh^{-\varepsilon}C_\varepsilon\alpha^{1+\varepsilon}\upsilon_\varepsilon+nC_\varepsilon\alpha^{1+\varepsilon}(1-h)^{-\varepsilon}|\mu-\theta|^{1+\varepsilon}\Big)\right).$$

Now using (34) and Markov's inequality, we have

$$\mathbb{P}\Big(Z_n \geq \frac{2}{\delta}\Big) \leq \frac{\delta}{2}.$$

In other words,

$$\mathbb{P}\Big(r(\theta) \geq n\alpha(\mu-\theta)+nh^{-\varepsilon}C_\varepsilon\alpha^{1+\varepsilon}\upsilon_\varepsilon+nC_\varepsilon\alpha^{1+\varepsilon}(1-h)^{-\varepsilon}|\mu-\theta|^{1+\varepsilon}+\log\Big(\frac{2}{\delta}\Big)\Big) \leq \frac{\delta}{2}.$$

□

**Proof of Theorem 6.2 & Theorem 6.4**

Proofs follows from Theorem 6.1 and Theorem 6.3 with $\delta = \frac{\delta}{1+2\log_\sigma A}$. □

## B. Implication for Differential Privacy Problem

In this last section, we illustrate the usefulness of our result by considering mean estimation problem under a differential privacy setting.

Recall the classical definition of $\epsilon$**-differential privacy** (Dwork, 2008):

**Definition B.1.** The algorithm $\mathcal{A}$ is said to provide $\epsilon$-differential privacy if, for all datasets $\mathcal{D}$ and $\mathcal{D}'$ that differ only on a single element, and all events $S$,

$$\mathbb{P}(\mathcal{A}(\mathcal{D}) \in S) \leq \exp\{\epsilon\}\mathbb{P}(\mathcal{A}(\mathcal{D}') \in S), \tag{35}$$

where the probability is taken over the randomness of algorithm $\mathcal{A}$.

In the literature, a relaxation form is also considered. Dwork et al. (2006) proposed a framework of **approximate differential privacy**:

**Definition B.2.** The algorithm $\mathcal{A}$ is said to provide $(\epsilon, \delta)$-differential privacy if, for all datasets $\mathcal{D}$ and $\mathcal{D}'$ that differ only on a single element, and all events $S$,

$$\mathbb{P}(\mathcal{A}(\mathcal{D}) \in S) \leq \exp\{\epsilon\}\mathbb{P}(\mathcal{A}(\mathcal{D}') \in S) + \delta. \tag{36}$$

We restrict ourselves by considering that dataset $\mathcal{D}$ consists of independently and identically distributed (i.i.d) data samples. We hence introduce the **restricted differential privacy**:

**Definition B.3.** The algorithm $\mathcal{A}$ is said to satisfy restricted $(\epsilon, \delta)$-differential privacy if, for all i.i.d datasets $\mathcal{D}$ and $\mathcal{D}'$ that differ only on a single element, and all events $S$,

$$\mathbb{P}(\mathcal{A}(\mathcal{D}) \in S) \leq \exp\{\epsilon\}\mathbb{P}(\mathcal{A}(\mathcal{D}') \in S) + \delta, \tag{37}$$

where the probability is taken over the randomness of both $\mathcal{A}$ and $\mathcal{D}, \mathcal{D}'$.

Here, in the definition of restricted differential privacy, we require the samples in $\mathcal{D}$ to be i.i.d. The second dataset $\mathcal{D}'$ is obtained by replacing one sample in $\mathcal{D}$ by another independent copy. In this scenario, we can easily construct a series of algorithms with guarantees of restricted differential privacy as below.

**Algorithm 5** Restricted Differential Privacy for Heavy-tailed Data

1: **Input**: i.i.d dataset $\{X_i\}$, moment parameter $\varepsilon$, differential privacy parameter $\epsilon, \delta$, and an estimation algorithm $\mathcal{M}$.
2: Compute the estimator $\hat{\mu}$ of the mean via using algorithm $\mathcal{M}$.
3: Compute the $1 - \delta/2$ confidence error of $\hat{\mu}$ and denote it as $B_\delta$.
4: Randomly generate a noise term $W$ which follows Laplace$(2B_\delta/\epsilon)$.
5: **Output**: $\hat{\mu} + W$.

In the framework of Algorithm 5, the estimation algorithm $\mathcal{M}$ can be taken as any of Lugosi and Mendelson (2019), Chen et al. (2021) or our proposed method. The random variable $W$ follows a Laplace distribution (i.e., the probability density function of Laplace$(b) = \frac{1}{2b}\exp\{-|x|/b\}$). Larger value of $b$ will lead to larger magnitude of noise. In other words, the estimation quality is largely determined by $B_\delta$. According to the discussions in Section 3.1, our method provides a smaller $B_\delta$ compared with other two competing methods and hence returns a better estimator under the same algorithmic framework. Lastly, by using standard proof techniques (Dwork and Roth, 2014; Duchi et al., 2018), it can be easily shown that Algorithm 5 satisfies $(\epsilon, \delta)$ restricted differential privacy.

The proposed application framework raises a few questions that have important implications for the community. Firstly, it is interesting to see if the i.i.d restriction on the data $\mathcal{D}$ can be relaxed, to allow any dependence structure between samples, which are themselves *heavy-tailed*. Secondly, it is worthwhile investigating whether the framework we adopted here can be viewed as a global private method. That is, noise is added to the output (query) of the dataset (or noise is added only once, at the end of the process before sharing it with the third party/user) (Wang et al., 2020). Also, is there any reasonable local differential private framework to privatize each data sample $X_i$ before adopting any estimation algorithm $\mathcal{M}$?