# Beyond the Best: Estimating Distribution Functionals in Infinite-Armed Bandits

**Yifei Wang**
Department of Electrical Engineering
Stanford University
Stanford, CA 94305
wangyf18@stanford.edu

**Tavor Z. Baharav**
Department of Electrical Engineering
Stanford University
Stanford, CA 94305
tavorb@stanford.edu

**Yanjun Han**
Institute for Data, Systems, and Society
Massachusetts Institute of Technology
Cambridge, MA 02142
yjhan@mit.edu

**Jiantao Jiao**
Department of Electrical Engineering and Computer Sciences and Department of Statistics
University of California, Berkeley
Berkeley, CA 94720
jiantao@eecs.berkeley.edu

**David Tse**
Department of Electrical Engineering
Stanford University
Stanford, CA 94305
dntse@stanford.edu

## Abstract

In the infinite-armed bandit problem, each arm's average reward is sampled from an unknown distribution, and each arm can be sampled further to obtain noisy estimates of the average reward of that arm. Prior work focuses on identifying the best arm, i.e., estimating the maximum of the average reward distribution. We consider a general class of distribution functionals beyond the maximum, and propose unified meta algorithms for both the offline and online settings, achieving optimal sample complexities. We show that online estimation, where the learner can sequentially choose whether to sample a new or existing arm, offers no advantage over the offline setting for estimating the mean functional, but significantly reduces the sample complexity for other functionals such as the median, maximum, and trimmed mean. The matching lower bounds utilize several different Wasserstein distances. For the special case of median estimation, we identify a curious thresholding phenomenon on the indistinguishability between Gaussian convolutions with respect to the noise level, which may be of independent interest.

## 1 Introduction

In the infinite-armed bandit problem (Berry et al., 1997), at each time instance the learner can either sample an arm that has been previously observed, or sample from a new arm, whose average reward

36th Conference on Neural Information Processing Systems (NeurIPS 2022).

is drawn from an unknown distribution $F$. The learner's goal is to identify arms with large average reward, with the objective being achieving either small cumulative regret (Berry et al., 1997; Wang et al., 2008; Bonald and Proutiere, 2013), or small simple regret (Carpentier and Valko, 2015). This setting differs from the classical multi-armed bandit formulation as the number of observed arms is not fixed a priori and needs to be carefully chosen by the algorithm.

We consider the problem of estimating some functional $g(F)$ of an underlying distribution $F$, as is illusrated in Figure 1. From this point of view, the classical infinite-armed bandit problem can be viewed as an *online* sampling algorithm to estimate the *maximum* of the distribution $F$. [1] Once we cast the infinite-armed bandit problem in this manner, it immediately suggests several additional questions. For example, what about *offline* sampling algorithms? Indeed, online sampling requires continual interactions with the environment which may be infeasible in certain applications, and recent work in online and offline reinforcement learning have demonstrated the significant value of both formulations (Rashidinejad et al., 2021; Zhang et al., 2021; Schrittwieser et al., 2021). Additionally, it is worth estimating functionals beyond the maximum: in many practical scenarios, including mean estimation in single-cell RNA-sequencing (Zhang et al., 2020) and Benjamini Hochberg (BH) threshold estimation in multiple hypothesis testing (Zhang et al., 2019), we are interested in the mean, median (quantile), or trimmed mean of the underlying distribution $F$. The estimation of quantiles is similar to estimation of the BH threshold, as both depend on the order statistics of the underlying distribution. Estimating the median or trimmed mean has further applications in robust statistic for instance, maintaining the fidelity of an estimator in the presence of adversarial corruption or outliers. Another natural setting where such problems arise is in large-scale distributed learning (Son and Simon, 2012). Here, a server / platform wants to estimate how much test-users like their newly released product. Users return a noisy realization of their affinity for the product, and the platform can decide to pay the user further to spend more time with the product, to test it further. For many natural objectives which are robust to a small fraction of adversarial users, e.g. trimmed mean, median, or quantile estimation, we see that our algorithm will enable estimation of the desired quantity to high accuracy while minimizing the total cost (number of samples taken). Since sampling is expensive, it is critical to identify the optimal method to collect samples, and identify the improvements afforded by adaptivity. For example, do online methods offer significant gains over offline methods? Are the fundamental limits of estimating the median and trimmed mean different from that of the maximum?

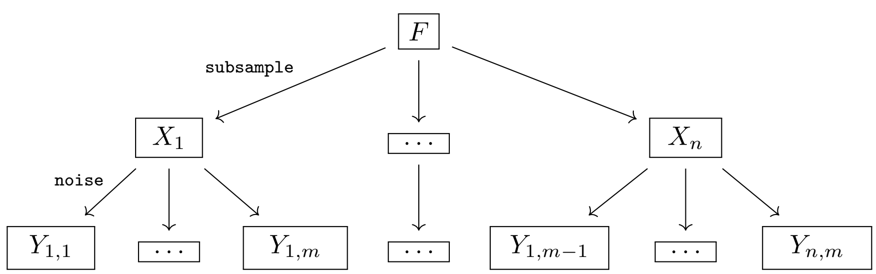


Figure 1: Problem setting. Level 0: underlying distribution $F(x)$. Level 1: unobserved samples $X_1, \ldots, X_n \sim F(x)$. Level 2: noisy observations $Y_{i,j} \sim \mathcal{N}(X_i, 1)$.

In this paper we initiate the study of distribution functional estimation in both online and offline settings and obtain both information theoretic limits and efficient algorithms for estimating the mean, median, trimmed mean, and maximum. We propose unified meta algorithms for both offline and online settings, and provide matching upper and lower boundsfor the sample complexity of estimating the aforementioned functionals in Table 1.

We also reveal new insights on the fundamental differences between the offline and online algorithms, as well as the fundamental differences between different functionals. To determine these sharp statistical limits, we use the Wasserstein-2 distance to upper bound the KL divergence in the offline setting, while the Wasserstein-$\infty$ distance is applied in the online setting instead. This approach leads to valid sample complexity lower bounds for general functionals $g$, which turn out to be tight for estimating the mean and maximum. However, a curious thresholding phenomenon, which is not

[1] To be precise, the objectives in infinite-armed bandit works (Berry et al., 1997; Wang et al., 2008; Bonald and Proutiere, 2013; Carpentier and Valko, 2015) are slightly different, minimizing simple or cumulative regret.

captured by the previous approach and does not occur for the *mean* and *maximum*, appears in the *median* and *trimmed mean* analyses: the KL divergence does not change smoothly with the noise level and enjoys a phase transition after the noise level exceeds some threshold. This phenomenon calls for different treatments under different estimation targets and could be of independent interest.

| Functional | Offline complexity | Online complexity | Comments |
|---|---|---|---|
| Mean | $\Theta(\varepsilon^{-2})$ | $\Theta(\varepsilon^{-2})$ | No gain from online sampling |
| Median | $\Theta(\varepsilon^{-3})$ | $\tilde{\Theta}(\varepsilon^{-2.5})$ | Holds for any quantile not on the boundary |
| Maximum | $\Theta(\varepsilon^{-(2+\beta)})$ | $\tilde{\Theta}(\varepsilon^{-\max(\beta,2)})$ | Depends on the tail regularity $\beta$ |
| Trimmed mean | $\tilde{\Theta}(\varepsilon^{-3})$ | $\tilde{\Theta}(\varepsilon^{-2.5})$ | $g(F) = \mathbb{E}\{X \mid X \in [F^{-1}(\alpha), F^{-1}(1-\alpha)]\}$ |

Table 1: Sample complexity of estimating different functionals $g(F)$, where $F$ is the cumulative distribution function (CDF) of the distribution to estimate. The trimmed mean result holds for a fixed $\alpha \in (0, 1/2)$. Here $\varepsilon$ is the target accuracy and we use $\Theta$ to denote the matching upper and lower bounds up to constants not depending on $\varepsilon$. Additionally, we use $\tilde{\Theta}, \gtrsim$, and $\lesssim$ to suppress constants and logarithmic factors in $\varepsilon$, and $\varepsilon^c$ for any fixed $c$ arbitrarily close to zero. If $h(\varepsilon) \lesssim f(\varepsilon)$ and $f(\varepsilon) \lesssim h(\varepsilon)$ then we denote this as $f(\varepsilon) \asymp h(\varepsilon)$. For maximum estimation, we assume that the distribution satisfies $\mathbb{P}(X \geq F^{-1}(1) - \varepsilon) \asymp \varepsilon^{\beta}$. Other assumptions on $F$ are detailed in Section 3.

The rest of this paper is structured as follows. In Section 1.1 we discuss the relevant literature. We then formulate our distribution functional estimation problem in Section 2. Our unified meta algorithms for the offline and online settings are presented in Section 3, where we show the sample complexity upper bounds. We present information theoretic lower bounds proofs via Wasserstein distance for the online and offline settings in Section 4, and discuss a special thresholding phenomenon arising in median estimation in Section 5. Section 6 concludes this work.

### 1.1 Related works

The field of multi-armed bandits has seen broad interest and utility since its formalization in 1985 (Lai et al., 1985). Across clinical trials, multi-agent learning, online recommendation systems, and beyond (Lattimore and Szepesvári, 2020), multi-armed bandits have proven to be an excellent framework for modeling and solving complex tasks regarding exploration in an unknown environment. In the classical multi-armed bandit setting we have a set of $n$ distributions, where the player sequentially pulls one arm per round and observes a sample drawn from the associated reward distribution. In the infinite-armed bandit setting (Berry et al., 1997), the average arm reward for each arm is sampled i.i.d. from an unknown distribution, i.e., we have infinitely many available arms. There are many possible objectives that can be formulated in this online learning problem, from cumulative/simple regret minimization (Wang et al., 2008; Bonald and Proutiere, 2013; Carpentier and Valko, 2015; Li and Xia, 2017) to identification tasks (for example identifying an arm whose average reward is $\varepsilon$ close to the largest average reward) (Aziz et al., 2018; Chaudhuri and Kalyanakrishnan, 2017, 2019). Many works have studied best-arm identification, and we now have essentially matching instance-dependent upper and lower bounds (Jamieson and Nowak, 2014; Kaufmann et al., 2016). One could also use the average reward estimate of the identified best arm to estimate the maximum of the average reward distribution in the infinite-armed bandit setting (Carpentier and Valko, 2015; Aziz et al., 2018; Chaudhuri and Kalyanakrishnan, 2017, 2019).

From a statistical perspective, the sample complexity in the offline setting is closely related to deconvolution distribution estimation (Cordy and Thomas, 1997; Wasserman, 2004; Hall and Lahiri, 2008; Delaigle et al., 2008; Dattner et al., 2011). Nevertheless, these previous works mainly focus on the expected L2 difference between the underlying distribution function and its estimation. This simplified setting does not allow for consideration of the trade-off inherent in our setting between the number of points and the (variable) number of observations per point. Additionally, these past works did not calculate the specific sample complexity for more general functionals like quantile and trimmed mean. Since the noise is treated as fixed and uniform, there has been no study of the online setting where adaptive resampling can enable dramatic sample complexity improvements. In particular, the challenge is that we have noisy observations, which makes deriving lower bounds even in offline cases a significant challenge that has not been dealt with in the past, let alone analyzing the online case. The dramatic performance gains afforded by adaptive resampling for functional estimation, combined with its lack of formal study, motivates the focus of this work.

## 2 Problem formulation

We are interested in estimating the distribution functional $g(F) \in \mathbb{R}$ of an underlying distribution with cumulative distribution function (CDF) $F$. We study a class of indicator-based functionals $g$ defined as follows.

**Definition 1** (Indicator-based functionals)**.** *The functional $g$ can be represented as*

$$g(F) = \mathbb{E}\left[X | X \in S(F)\right] \tag{1}$$

*for some set $S(F)$, where $X \sim F$. The set $S(F)$ is defined as follows:*

$$S(F) = [F^{-1}(\alpha_1), F^{-1}(\alpha_2)], 0 \leq \alpha_1 \leq \alpha_2 \leq 1. \tag{2}$$

We denote $S(F)$ by $S$ throughout this work when $F$ is clear from context. This class encompasses many natural functionals of interest, which we formulate in Table 2. In Appendix A, we discuss extending our results to more general functionals, and show that our approach can extend to smooth reweighting functions $h(X)$ and more complex sets $S$.

| Functional | $g(F)$ | $\alpha_1$ | $\alpha_2$ | Comment |
|---|---|---|---|---|
| Mean | $\mathbb{E}[X]$ | 0 | 1 | |
| Quantile | $F^{-1}(\alpha)$ | $\alpha$ | $\alpha$ | $\alpha \in (0, 1)$, e.g. $\alpha = 1/2$ for median |
| Maximum | $F^{-1}(1)$ | 1 | 1 | $\alpha_1 = \alpha_2 = 0$ for minimum |
| Trimmed mean | $\mathbb{E}[X \| F(X) \in [\alpha_1, \alpha_2]]$ | $\alpha$ | $1 - \alpha$ | $\alpha \in (0, 1/2)$ |

Table 2: Indicator-based functionals.

As in the infinite-armed bandit setting, we only have access to noisy observations of samples drawn from the distribution with CDF $F$. We can either choose to sample from a point $X$ which we already have some noisy observations of, or draw a new point $X$ from $F$. We then observe $Y = X + Z$, where $Z \sim \mathcal{N}(0, 1)$ is independent of everything observed so far.

In this paper we characterize the online and offline sample complexities of these problems, and in Section 3 propose online and offline algorithms achieving them. For $\varepsilon > 0$ and $\delta \in (0, 1)$, we call an estimator $\hat{G}$ an $(\varepsilon, \delta)$-PAC approximation of $g(F)$ if $\mathbb{P}(|\hat{G} - g(F)| > \varepsilon) \leq \delta$.

## 3 Offline and online algorithms

### 3.1 Offline estimation algorithms

We study a special class of offline algorithms, which uniformly obtain observations of the points following the underlying distribution. To be precise, based on prior information regarding the distribution in question, $F$, it will choose an appropriate number of points $n$ and number of samples per point $m$ to obtain an $(\varepsilon, \delta)$-PAC approximation of $g(F)$. Specifically, the latent variables $X_1, \ldots, X_n$ are drawn from $F$, and our observations $\{Y_{i,j}\}_{j=1}^m$ are drawn i.i.d. from $\mathcal{N}(X_i, 1)$, independently for each $i$. For $i \in [n]$, denote $\hat{X}_i = m^{-1} \sum_{j=1}^m Y_{i,j}$ as the empirical mean of the observations for arm $i$. Then, we can write $\hat{X}_i = X_i + \hat{Z}_i$ where $\hat{Z}_i \sim \mathcal{N}(0, 1/m)$, independent across $i$. Define $S_n \triangleq \{i : X_{(\lfloor \alpha_1 n \rfloor)} \leq X_i \leq X_{(\lfloor \alpha_2 n \rfloor)}\}$ as the set of arms relevant for estimating the functional $g$, and define our $n$ sample estimate of $g$ as $g_n(\hat{X}_1, \ldots, \hat{X}_n) \triangleq |S_n|^{-1} \sum_{i \in S_n} \hat{X}_i$. Here $X_{(i)}$ denotes the $i$-th order statistic, that is the $i$-th smallest entry in $X_1, \ldots, X_n$. Then, $G_{n,m} = g_n(\hat{X}_1, \ldots, \hat{X}_n)$, where each $X_i$ has been sampled $m$ times, serves as a natural estimator for $g(F)$ from the noisy observations. With this, we can state the following theorem:

**Theorem 1** (Offline PAC sample complexity)**.** *An $(\varepsilon, \delta)$-PAC offline uniform-sampling-based algorithm for estimating $g(F)$ requires $\Theta(nm)$ samples where $n, m$ depend on $\varepsilon, \delta$, the functional $g$, and information about $F$, with orderwise dependence on $\varepsilon$ detailed in Table 3.*

For the rest of this section, we discuss in greater detail our assumptions on the underlying distribution. We defer the proofs and calculations for $n$ and $m$ to Appendix C, as well as discussion regarding the trimmed mean to Appendix B.

| Functional | $m$ | $n$ |
|---|---|---|
| Mean | $\Theta(1)$ | $\Theta(\varepsilon^{-2})$ |
| Median | $\Theta(\varepsilon^{-1})$ | $\Theta(\varepsilon^{-2})$ |
| Maximum | $\Theta(\varepsilon^{-2})$ | $\Theta(\varepsilon^{-\beta})$ |
| Trimmed mean | $\Theta\left(\varepsilon^{-1}\log\left(\varepsilon^{-1}\right)\right)$ | $\Theta(\varepsilon^{-2})$ |

Table 3: Choice of $(m, n)$ for estimating different functionals to accuracy $\varepsilon$.

#### 3.1.1 Mean

To guarantee that the empirical mean is a good estimator for the true mean, we impose assumptions on the tail of the distribution $F$:

**Assumption 1.** *The distribution $F$ satisfies $\mathrm{Var}_{X\sim F}[X] \leq c$.*

Assumption 1 ensures that estimation of the mean of the distribution can be accomplished with finite samples. The following proposition gives the sample complexity of the offline algorithm.

**Proposition 1.** *Suppose that Assumption 1 is satisfied. By choosing $m = 1$ and $n \geq \delta^{-1}(1+c)\varepsilon^{-2}$, the estimator $G_{n,m}$ is an $(\varepsilon, \delta)$-PAC approximation of $g(F)$. Thus, the offline algorithm requires $O(\varepsilon^{-2})$ samples.*

#### 3.1.2 Median

For median estimation we require different assumptions than the mean, as listed below.

**Assumption 2.** *There exist constants $c_1, c_2 > 0$ such that*

- $F'(x) \geq c_1$ *for* $|x - \mathrm{median}(F)| \lesssim \varepsilon$.
- $|F''(x)| \leq c_2$ *for* $|x - \mathrm{median}(F)| \lesssim \sqrt{\varepsilon}$.

The first assumption ensures that the median of $F$ is unique. The second assumption precludes the distribution from being dumbbell-shaped (very little mass near the median), in which case estimating the true median is meaningless and can be arbitrarily difficult. The following proposition gives a suitable choice of $(n, m)$ for providing an $(\varepsilon, \delta)$-PAC approximation of $g(F)$.

**Proposition 2.** *Suppose that Assumption 2 holds. Then, by choosing $m \geq 4(c_2+1)/(c_1\varepsilon)$ and $n \geq 28\log(1/\delta)/(c_1\varepsilon)^2$, the estimator $G_{n,m}$ is an $(\varepsilon, \delta)$-PAC approximation of $g(F)$. Thus, the offline algorithm requires $O(\varepsilon^{-3})$ samples.*

#### 3.1.3 Maximum

For maximum estimation, we require an assumption on the tail of $F$ as is common in the infinite-armed bandit literature.

**Assumption 3.** *There exist constants $0 < c_1 < c_2$ and $\beta > 0$ such that*

- $1 - F(F^{-1}(1) - t) \in [c_1 t^\beta, c_2 t^\beta]$, *for all* $0 \leq t \lesssim \varepsilon$.

This assumption is also known as the $\beta$-regularity of $F$ around $F^{-1}(1)$, see (Wang et al., 2008). We present a suitable choice of $(n, m)$ in the following proposition.

**Proposition 3.** *Suppose that Assumption 3 holds. By choosing $n \geq c_1^{-1}2^\beta\varepsilon^{-\beta}\log(2/\delta)$ and $m \geq 4\varepsilon^{-2}\log(2n/\delta)$, the estimator $G_{n,m}$ is an $(\varepsilon, \delta)$-PAC approximation of $g(F)$. Therefore, the offline algorithm requires $O(\varepsilon^{-\beta-2})$ samples.*

### 3.2 Online estimation algorithm

We now present our general algorithm (Algorithm 1), an elimination-based $(\varepsilon, \delta)$-PAC algorithm that efficiently estimates $g(F)$, where $g$ is a known input functional and $F$ is an unknown distribution from which we are able to sample $X_i$ independently, and observe noisy observations $Y_{i,j}$ of $X_i$.

**Algorithm 1** `Meta Algorithm`

**Input:** target accuracy $\varepsilon$, error probability $\delta$, functional $g$ parameterized by $(\alpha_1, \alpha_2)$
Compute $(n, m)$ for $(\varepsilon/2, \delta/2)$-PAC estimation of $g(F)$ based on Theorem 1
Construct active set $A_1 = [n]$, and define $b_0 = 1$ and $t_0 = 0$
**for** $r = 1, 2, \dots$ **do**
  Define $b_r = 2^{-r}$ and $t_r = \min(m, \lceil 8b_r^{-2} \log(16n \log(m)/\delta) \rceil)$
  Pull each arm in $A_r$ for $t_r - t_{r-1}$ times, construct $\hat{\mu}(r)$
  Compute $A_{r+1} = \left\{ i : |\hat{\mu}_i(r) - \hat{\mu}_{(\lfloor \alpha_1 n \rfloor)}(r)| \le b_r \text{ or } |\hat{\mu}_i(r) - \hat{\mu}_{(\lfloor \alpha_2 n \rfloor)}(r)| \le b_r \right\}$
  **if** $t_r \equiv m$ **then**
    **Break**, exit For loop
  **end if**
**end for**
Construct $\hat{S}_n = \{ i : \hat{\mu}_{(\lfloor \alpha_1 n \rfloor)}(r) \le \hat{\mu}_i(r) \le \hat{\mu}_{(\lfloor \alpha_2 n \rfloor)}(r) \}$
**if** $\alpha_1 \equiv \alpha_2$ **then**
  **return** $\frac{1}{|\hat{S}_n|} \sum_{i \in \hat{S}_n} \hat{\mu}_i(r)$
**else**
  Draw one observation from each $i \in \hat{S}_n$, construct $\tilde{\mu}_i$
  **return** $\frac{1}{|\hat{S}_n|} \sum_{i \in \hat{S}_n} \tilde{\mu}_i$
**end if**

In order to exploit the Bayesian nature of the problem, we analyze the algorithm in two parts. First, we use the fact that our arms are drawn from a common distribution to find some $n, m$ as in Theorem 1 such that the plug-in estimator $G_{n,m}$ will be an $(\varepsilon/2, \delta/2)$-PAC approximation of $g(F)$. Second, we show that our adaptive algorithm is an $(\varepsilon/2, \delta/2)$-PAC approximation of $G_{n,m}$, but is able to accomplish this using significantly fewer than $n \times m$ samples.

Notationally, we denote by $\hat{\mu}(r)$ the estimated mean vector of all arms at round $r$, and denote the $i$-th entry of this vector by $\hat{\mu}_i(r)$. We have that with high probability each arm's mean estimate stays within its width $b_r = 2^{-r}$ confidence interval for each round $r$. To analyze our algorithm, we denote $\mu_1^{\text{uni}}, \dots, \mu_n^{\text{uni}}$ as the estimates of $X_1, \dots, X_n$ generated by the offline algorithm after sampling each arm $m$ times. Then, we see that for the offline algorithm the arms relevant for the estimation task and the corresponding $n$ sample estimator are

$$S_n = \left\{ i : \mu_{(\lfloor \alpha_1 n \rfloor)}^{\text{uni}} \le \mu_i^{\text{uni}} \le \mu_{(\lfloor \alpha_2 n \rfloor)}^{\text{uni}} \right\}, \quad G_{n,m} = \frac{1}{|S_n|} \sum_{i \in S_n} \mu_i^{\text{uni}}. \tag{3}$$

Here $S_n$ indicates the arms that the offline algorithm believes are in $S$. We show that our online algorithm is able to efficiently estimate the set $S_n$ as $\hat{S}_n$, determining whether or not arms are in $S_n$, sampling these arms in $S_n$ sufficiently, and returning a plug-in estimator. By construction each arm is only pulled by the adaptive algorithm at most $m$ times, as we know from the analysis of the offline algorithm that for the utilized $n, m$, if each arm is pulled $m$ times then the output is an $(\varepsilon/2, \delta/2)$-PAC estimate of $g(F)$. Thus, the online algorithm's objective is essentially emulating the output of the offline algorithm, for which it only needs to sample any arm at most $m$ times.

Note that when $\alpha_1 \neq \alpha_2$, we have many samples $X_i$ that are within $S$, with $|S_n| \ge \lfloor n(\alpha_2 - \alpha_1) \rfloor$. In order to avoid issues of dependence, we discard all previous samples (as arms in $\hat{S}_n$ will have been sampled different numbers of times), and see that since we have $\Theta(n)$ arms in $\hat{S}_n$ we can construct a sufficiently accurate estimate by sampling each arm in $\hat{S}_n$ once. Algorithmically, we denote this as obtaining one fresh observation and constructing $\tilde{\mu}_i$.

To upper bound the sample complexity of our algorithm, we see that each arm only needs to be sampled to determine whether it is in $S_n$ or not. As we show in Appendix D, the number of samples $N(i)$ needed for point $X_i$ satisfies

$$N(i) \le \min \left( m, \frac{256 \log \left( \frac{16n \log m}{\delta} \right)}{\left[ \text{dist}(X_i, \partial \, \text{Conv}(\{\mu_i^{\text{uni}} : i \in S_n\})) \right]^2} \right), \tag{4}$$

with probability at least $1 - \delta/4$ for all arms simultaneously, where $\partial A$ denotes the boundary of a set $A$, $\text{Conv}(A)$ denotes the convex hull of a set $A$, and $\text{dist}(X, A) = \min_{a \in A} |X - a|$. In the limit as $\varepsilon \to 0$ we show that $\mu^{\text{uni}}_{(\lfloor \alpha_1 n \rfloor)} \to F^{-1}(\alpha_1)$ (similarly with $\mu^{\text{uni}}_{(\lfloor \alpha_2 n \rfloor)}$). This allows us to state the following theorem regarding the expected sample complexity of Algorithm 1 with respect to the distribution's relevant set of values $S$ rather than the estimated indices $S_n$.

**Theorem 2** (Meta algorithm)**.** *For a functional $g$ satisfying Definition 1, Algorithm 1 provides an $(\varepsilon, \delta)$-PAC estimate of $g(F)$ with $M$ samples when given the requisite inputs. Here $m$ and $n$ are calculated as in Theorem 1, and the number of samples $M$ required satisfies*

$$\mathbb{E}[M] = O\left(n\mathbb{E}\left[\min\left(m, \frac{\log(n/\delta)}{[\text{dist}(X, \partial S)]^2}\right)\right]\right). \tag{5}$$

The proof of this Theorem is deferred to Appendix D.

Evaluating this expression for different functionals under their corresponding assumptions yields the stated sample complexity upper bounds, as we show in Appendix D.3.

## 4 Lower bounds via Wasserstein distance

In this section we derive general lower bounds on the sample complexity of functional estimation for both offline and online algorithms, where two different Wasserstein distances play important roles. These Wasserstein-based lower bounds yield tight results for mean and maximum estimation.

### 4.1 General lower bounds based on Wasserstein distance

A classical technique for proving minimax lower bounds is Le Cam's two-point method (Le Cam et al., 2000): let $F_1$ and $F_2$ be two distributions with $|g(F_1) - g(F_2)| \geq 2\varepsilon$, and let $p_{\pi,F_1}$ and $p_{\pi,F_2}$ be the probability distributions of all observations queried by policy $\pi$ under the true population distributions $F_1$ and $F_2$, respectively. One version of Le Cam's two-point lower bound (Tsybakov, 2009, Theorem 2.2) gives

$$\inf_{\widehat{g}} \sup_{F \in \{F_1, F_2\}} \mathbb{P}_F(|\widehat{g} - g(F)| \geq \varepsilon) \geq \frac{1}{4} \exp\left(-D_{\text{KL}}(p_{\pi,F_1} \| p_{\pi,F_2})\right).$$

Consequently, to construct a lower bound on the PAC sample complexity of estimating $g(F)$, it suffices to find the largest $\varepsilon$ such that there exist $F_1, F_2$ with $|g(F_1) - g(F_2)| \geq 2\varepsilon$ while $D_{\text{KL}}(p_{\pi,F_1} \| p_{\pi,F_2}) = O(1)$.

A key step in the above analysis is to upper bound the KL divergence $D_{\text{KL}}(p_{\pi,F_1} \| p_{\pi,F_2})$, which differs significantly between offline and online algorithms. For offline algorithms, the learner samples $n$ arms i.i.d. from $F$ with average rewards $X_1, \cdots, X_n \sim F$, and each arm is pulled $m$ times with Gaussian observations. Consequently, $p_{\pi,F} == (F * \mathcal{N}(0, 1/m))^{\otimes n}$, where $p^{\otimes n}$ denotes the $n$-fold product distribution and $*$ denotes the convolution operation. The following lemma presents an upper bound on the KL divergence for offline algorithms.

**Lemma 1.** *For any offline algorithm $\pi$ defined in Section 3.1, it holds that*

$$D_{\text{KL}}(p_{\pi,F_1} \| p_{\pi,F_2}) \leq \frac{mn}{2} \mathcal{W}_2^2(F_1, F_2),$$

*where $\mathcal{W}_2(P, Q)$ is the Wasserstein-2 distance defined as $\mathcal{W}_2^2(P, Q) = \inf_{\gamma \in \Gamma} \mathbb{E}_{(X,Y) \sim \Gamma}[(X - Y)^2]$, with $\Gamma$ being the class of all couplings between $P$ and $Q$.*

For online algorithms the distribution $p_{\pi,F}$ is no longer a product distribution as actions can depend on past observations. As a result, the KL divergence becomes larger, but still enjoys an upper bound based on another Wasserstein distance.

**Lemma 2.** *For any online algorithm $\pi$ which queries $T$ samples, it holds that*

$$D_{\text{KL}}(p_{\pi,F_1} \| p_{\pi,F_2}) \leq \frac{T}{2} \mathcal{W}_\infty^2(F_1, F_2),$$

*where $\mathcal{W}_\infty(P, Q)$ is the Wasserstein-$\infty$ distance: $\mathcal{W}_\infty(P, Q) = \inf_{\gamma \in \Gamma} \text{esssup}_{(X,Y) \sim \Gamma} |X - Y|$, with $\Gamma$ being the class of all couplings between $P$ and $Q$.*

As $\mathcal{W}_2(P,Q) \le \mathcal{W}_\infty(P,Q)$, the upper bound of Lemma 2 is no smaller than that of Lemma 1, showing the stronger power of online algorithms. The following corollary is then immediate from Lemmas 1 and 2.

**Corollary 2.1.** *The sample complexity of* $(\varepsilon, .1)$*-PAC estimation of* $g(F)$ *is*

$$\Omega(1/\min\{\mathcal{W}_2^2(F_1,F_2) : F_1, F_2 \in \mathcal{F}, |g(F_1) - g(F_2)| \ge 2\varepsilon\})$$

*for offline algorithms, and is*

$$\Omega(1/\min\{\mathcal{W}_\infty^2(F_1,F_2) : F_1, F_2 \in \mathcal{F}, |g(F_1) - g(F_2)| \ge 2\varepsilon\})$$

*for online algorithms.*

In the remainder of this section, we show that Corollary 2.1 leads to tight lower bounds for mean and maximum estimations for both offline and online settings.

### 4.2 Lower bounds for mean estimation

Consider two distributions $F_1$ and $F_2$ which are Dirac masses supported on $1/2 - \varepsilon$ and $1/2 + \varepsilon$, respectively. Clearly $\mathcal{W}_2(F_1,F_2) = \mathcal{W}_\infty(F_1,F_2) = 2\varepsilon$, which is the best possible as $\mathcal{W}_2(F_1,F_2) \ge |\mathrm{mean}(F_1) - \mathrm{mean}(F_2)| \ge 2\varepsilon$. Corollary 2.1 gives the following lower bounds.

**Corollary 2.2.** *The* $(\varepsilon, .1)$*-PAC sample complexity for mean estimation is* $\Omega(\varepsilon^{-2})$ *for both offline and online algorithms.*

### 4.3 Lower bounds for maximum estimation

For maximum estimation, the Wasserstein distances $\mathcal{W}_2$ and $\mathcal{W}_\infty$ behave differently, as summarized in the following lemma. Let $\mathcal{F}_\beta$ be the class of densities satisfying Assumption 3.

**Lemma 3.** *For* $\varepsilon \in (0, 1/2)$*, it holds that*

$$\begin{aligned}
\min\{\mathcal{W}_2(F_1,F_2) : F_1,F_2 \in \mathcal{F}_\beta, |\max(F_1) - \max(F_2)| \ge 2\varepsilon\} &= O(\varepsilon^{\beta/2+1}); \\
\min\{\mathcal{W}_\infty(F_1,F_2) : F_1,F_2 \in \mathcal{F}_\beta, |\max(F_1) - \max(F_2)| \ge 2\varepsilon\} &= O(\varepsilon); \\
\min\{D_{\mathrm{KL}}(F_1\|F_2) : F_1,F_2 \in \mathcal{F}_\beta, |\max(F_1) - \max(F_2)| \ge 2\varepsilon\} &= O(\varepsilon^{\beta}).
\end{aligned}$$

Note that we have included another term $D_{\mathrm{KL}}(F_1\|F_2)$ in Lemma 3 as it can provide a better lower bound than using $\mathcal{W}_\infty$ if $\beta \ge 2$, as $D_{\mathrm{KL}}(p_{\pi,F_1}\|p_{\pi,F_2}) \le T \cdot D_{\mathrm{KL}}(F_1\|F_2)$ always holds due to the data-processing inequality (i.e. assuming that all arm rewards are independent). Consequently, we have the following corollary on the sample complexity of maximum estimation.

**Corollary 2.3.** *The* $(\varepsilon, .1)$*-PAC sample complexity for maximum estimation over* $\mathcal{F}_\beta$ *is* $\Omega(\varepsilon^{-(\beta+2)})$ *for offline algorithms, and* $\Omega(\varepsilon^{-\max\{\beta,2\}})$ *for online algorithms.*

## 5 Lower bounds via thresholding phenomenon

Although the Wasserstein distance-based approach in Section 4 provides general lower bounds for both offline and online algorithms, and these lower bounds are tight for mean and maximum estimation, sometimes this approach can be loose. For example, Lemma 3 shows that using the $\mathcal{W}_\infty$ distance might be looser than using the original KL divergence for maximum estimation. This section provides tight lower bounds for median estimation, revealing a curious thresholding phenomenon.

### 5.1 Thresholding phenomenon for offline algorithms

Let $\mathcal{F}$ denote the set of distributions satisfying Assumption 2. To use Le Cam's two-point method to prove lower bounds for offline algorithms for median estimation, the key quantity is:

$$\mathrm{KL}_\sigma(\varepsilon) \triangleq \min\{D_{\mathrm{KL}}(F_1 * \mathcal{N}(0,\sigma^2) \| F_2 * \mathcal{N}(0,\sigma^2)) : F_1,F_2 \in \mathcal{F}, |F_1^{-1}(1/2) - F_2^{-1}(1/2)| \ge 2\varepsilon\}.$$

Its inverse $\mathrm{KL}_\sigma^{-1}(\varepsilon)$ is referred to as the modulus of smoothness of the median with respect to the KL divergence under Gaussian convolution. The Wasserstein-based approach to upper bound $\mathrm{KL}_\sigma(\varepsilon)$

in Lemma 1 is the following: let $\mathcal{W}_{2,\sigma}(\varepsilon)$ be the counterpart of the above quantity with the KL divergence replaced by the Wasserstein-2 distance, Lemma 1 shows that

$$\mathrm{KL}_\sigma(\varepsilon) \le \frac{\mathcal{W}_{2,\sigma}(\varepsilon)^2}{2\sigma^2} = \Theta\left(\frac{\varepsilon^{2.5}}{\sigma^2}\right), \tag{6}$$

an upper bound decreasing continuously with $\sigma$, where the proof of the last identity is presented in the Appendix. However, this upper bound is not tight, as shown in the following lemma.

**Lemma 4.** *For $\varepsilon \in (0, 1/4)$, $\mathrm{KL}_\sigma(\varepsilon)$ can be characterized as follows:*

$$\mathrm{KL}_\sigma(\varepsilon) \begin{cases} \in [C_1\varepsilon^2, C_2\varepsilon^2] & \text{if } \sigma \le c\varepsilon^{1/2}, \\ \le C(\theta,\kappa)\varepsilon^\kappa & \text{if } \sigma \ge \varepsilon^{1/2-\theta}, \end{cases}$$

*where $\theta \in (0, 1/4), \kappa \in \mathbb{N}$ are arbitrary fixed parameters, and $c, C_1, C_2, C(\theta,\kappa)$ are absolute constants with the last one depending only on $(\theta,\kappa)$.*

Lemma 4 shows a thresholding phenomenon as follows: when $\sigma$ increases from 0 to 1, the quantity $\mathrm{KL}_\sigma(\varepsilon)$ stabilizes at $\Theta(\varepsilon^2)$ whenever $\sigma \lesssim \varepsilon^{1/2}$; however, when $\sigma$ exceeds this threshold slightly (i.e. $\sigma \gtrsim \varepsilon^{1/2-\theta}$ for any constant $\theta > 0$), this quantity immediately drops to $o(\varepsilon^\kappa)$ for every possible $\kappa$. The main intuition behind this thresholding phenomenon is that, if $\sigma = O(\varepsilon^{1/2})$, the "bandwidth" of $F_1 - F_2$ exceeds that of $\mathcal{N}(0, \sigma^2)$, and the convolution is effectively using $\mathcal{N}(0, \sigma^2)$ as a Gaussian kernel (which preserves polynomials up to order 2) for smoothing $F_1 - F_2$ (which is second-order differentiable). In contrast, when $\sigma \gg \varepsilon^{1/2}$, the "bandwidth" of $F_1 - F_2$ could be smaller than $\mathcal{N}(0, \sigma^2)$, and the convolution is effectively using $F_1 - F_2$ as a kernel (which could preserve polynomials up to any desired order) for smoothing $\mathcal{N}(0, 1)$ (which is infinitely differentiable). Approximation theory tells us that the latter approximation error could be much smaller than the former, leading to the thresholding phenomenon. We remark that this phenomenon is not captured by using the $\mathcal{W}_2$ distance.

This thresholding behavior has an important consequence for median estimation. By Lemma 4 with $\sigma = 1/\sqrt{m}$, PAC learning requires that $m = \Omega(\varepsilon^{2\theta-1})$ for any offline algorithm, as otherwise the KL divergence could be made arbitrarily small. When $m$ is large enough, the first line of Lemma 4 then requires $n = \Omega(\varepsilon^{-2})$ to result in a large KL divergence for PAC learning, which comes from the idendity that

$$\mathrm{D_{KL}}(P^{\otimes n} \| Q^{\otimes n}) = n\mathrm{D_{KL}}(P\|Q).$$

Consequently, we have the following corollary for median estimation using offline algorithms.

**Corollary 2.4.** *Fix any $\theta > 0$. The $(\varepsilon, .1)$-PAC sample complexity for median estimation is $\Omega(\varepsilon^{-3+\theta})$ for any offline algorithm.*

### 5.2 Thresholding phenomenon for online algorithms

To prove the PAC lower bound for online algorithms, one first wonders if the same observation in Lemma 4 could still work. However, a close inspection of the proof reveals an issue: the optimizers $(F_1, F_2)$ in the definition of $\mathrm{KL}_\sigma(\varepsilon)$ are different under the regimes $\sigma = O(\varepsilon^{1/2})$ and $\sigma = \Omega(\varepsilon^{1/2-\theta})$. An online learning algorithm could first identify the right scenario and then choose a proper sample size to tackle the problem, and thus the above lower bound arguments break down.

To resolve this issue, we aim to choose a proper pair of distributions $(F_1, F_2)$ with $|\mathrm{median}(F_1) - \mathrm{median}(F_2)| \ge 2\varepsilon$, and investigate the behavior of $D_{\mathrm{KL}}(F_1 * \mathcal{N}(0, \sigma^2) \| F_2 * \mathcal{N}(0, \sigma^2))$ as a function of $\sigma$ *with* $(F_1, F_2)$ *fixed along the line*. The following lemma shows that, even for some fixed pair $(F_1, F_2)$, a similar thresholding phenomenon still holds for the KL divergence.

**Lemma 5.** *Fix any $\varepsilon, \theta \in (0, 1/4)$, and $\kappa \in \mathbb{N}$. There exist two distributions $F_1, F_2 \in \mathcal{F}$ with $|\mathrm{median}(F_1) - \mathrm{median}(F_2)| \ge 2\varepsilon$, and*

$$D_{\mathrm{KL}}(F_1 * \mathcal{N}(0, \sigma^2) \| F_2 * \mathcal{N}(0, \sigma^2)) \begin{cases} \in [C_1\varepsilon^{3/2}, C_2\varepsilon^{3/2}] & \text{if } \sigma \le c\varepsilon^{1/2}, \\ \le C(\theta,\kappa)\varepsilon^\kappa & \text{if } \sigma \ge \varepsilon^{1/2-\theta}, \end{cases}$$

*where $c, C_1, C_2, C(\theta,\kappa)$ are absolute constants with the last one depending only on $(\theta,\kappa)$.*

Compared with Lemma 4, Lemma 5 still shows a similar thresholding phenomenon for the KL divergence when $\sigma \gg \varepsilon^{1/2}$, but the KL divergence becomes larger for small $\sigma$ due to the additional constraint that $(F_1, F_2)$ is held fixed. Under the choice of $(F_1, F_2)$ in Lemma 5, each arm should be pulled at least $\Omega(\varepsilon^{2\theta-1})$ times, while $\Omega(\varepsilon^{\theta-3/2})$ arms need to be pulled in view of the first line. The following theorem makes the above intuition formal.

**Theorem 3.** *The $(\varepsilon, .1)$-PAC sample complexity for median estimation is $\Omega(\varepsilon^{-5/2+\theta})$ for any fixed $\theta > 0$ and any online algorithm.*

The formal proof of Theorem 3 is more complicated and requires an explicit computation of the KL divergence $D_{\mathrm{KL}}(p_{\pi,F_1} \| p_{\pi,F_2})$. We relegate the full proof to Appendix F. This thresholding phenomenon of the noise level also applies to the case of trimmed mean, which is discussed further and an analogous result is proved in Appendix B.

## 6 Conclusion

In this work we formulated and studied offline and online algorithms for estimating functionals of distributions. We developed unified algorithms for estimating the mean, median, maximum, and trimmed mean, providing sample complexity upper bounds. We additionally proved information theoretic lower bounds in these settings, which show that our algorithms are optimal up to $\varepsilon^c$ where $c$ is a fixed constant arbitrarily close to zero. We used different Wasserstein distances to construct information theoretic lower bounds for mean and maximum estimation, and showed how fundamentally different techniques are required for median and trimmed mean estimation. The lower bounds for median and trimmed mean estimation elucidate an interesting thresholding phenomenon of the noise level to distinguish two distributions after Gaussian convolution, which may be of independent interest. Interesting directions of future work include extending our analysis to non-indicator-based functionals, such as the BH threshold and analyzing the limiting behavior as $\delta \to \infty$.

## Acknowledgements

Yifei Wang and David Tse were partially supported by NSF Grants CCF-1909499. Tavor Z. Baharav was supported in part by the NSF GRFP and the Alcatel-Lucent Stanford Graduate Fellowship. Yanjun Han is supported by a Simons-Berkeley research fellowship and Norbert Wiener postdoctoral fellowship. Jiantao Jiao was partially supported by NSF Grants IIS-1901252, and CCF-1909499.

## References

Aziz, M., Anderton, J., Kaufmann, E., and Aslam, J. (2018). Pure exploration in infinitely-armed bandit models with fixed-confidence. In *Algorithmic Learning Theory*, pages 3–24. PMLR.

Berry, D. A., Chen, R. W., Zame, A., Heath, D. C., and Shepp, L. A. (1997). Bandit problems with infinitely many arms. *The Annals of Statistics*, 25(5):2103–2116.

Bonald, T. and Proutiere, A. (2013). Two-target algorithms for infinite-armed bandits with bernoulli rewards. *Advances in Neural Information Processing Systems*, 26.

Carpentier, A. and Valko, M. (2015). Simple regret for infinitely many armed bandits. In *International Conference on Machine Learning*, pages 1133–1141. PMLR.

Chaudhuri, A. R. and Kalyanakrishnan, S. (2017). Pac identification of a bandit arm relative to a reward quantile. In *Thirty-First AAAI Conference on Artificial Intelligence*.

Chaudhuri, A. R. and Kalyanakrishnan, S. (2019). Pac identification of many good arms in stochastic multi-armed bandits. In *International Conference on Machine Learning*, pages 991–1000. PMLR.

Cordy, C. B. and Thomas, D. R. (1997). Deconvolution of a distribution function. *Journal of the American Statistical Association*, 92(440):1459–1465.

Dattner, I., Goldenshluger, A., and Juditsky, A. (2011). On deconvolution of distribution functions. *The Annals of Statistics*, pages 2477–2501.

Delaigle, A., Hall, P., and Meister, A. (2008). On deconvolution with repeated measurements. *The Annals of Statistics*, 36(2):665–685.

Hall, P. and Lahiri, S. N. (2008). Estimation of distributions, moments and quantiles in deconvolution problems. *The Annals of Statistics*, 36(5):2110–2134.

Jamieson, K. and Nowak, R. (2014). Best-arm identification algorithms for multi-armed bandits in the fixed confidence setting. In *Information Sciences and Systems (CISS), 2014 48th Annual Conference on*, pages 1–6. IEEE.

Kaufmann, E., Cappé, O., and Garivier, A. (2016). On the complexity of best-arm identification in multi-armed bandit models. *The Journal of Machine Learning Research*, 17(1):1–42.

Lai, T. L., Robbins, H., et al. (1985). Asymptotically efficient adaptive allocation rules. *Advances in applied mathematics*, 6(1):4–22.

Lattimore, T. and Szepesvári, C. (2020). *Bandit algorithms*. Cambridge University Press.

Le Cam, L., LeCam, L. M., and Yang, G. L. (2000). *Asymptotics in statistics: some basic concepts*. Springer Science & Business Media.

Li, H. and Xia, Y. (2017). Infinitely many-armed bandits with budget constraints. In *Thirty-First AAAI Conference on Artificial Intelligence*.

Rashidinejad, P., Zhu, B., Ma, C., Jiao, J., and Russell, S. (2021). Bridging offline reinforcement learning and imitation learning: A tale of pessimism. *Advances in Neural Information Processing Systems*, 34.

Schrittwieser, J., Hubert, T., Mandhane, A., Barekatain, M., Antonoglou, I., and Silver, D. (2021). Online and offline reinforcement learning by planning with a learned model. *Advances in Neural Information Processing Systems*, 34.

Son, L. K. and Simon, D. A. (2012). Distributed learning: Data, metacognition, and educational implications. *Educational Psychology Review*, 24(3):379–399.

Tsybakov, A. (2009). *Introduction to Nonparametric Estimation*. Springer-Verlag.

Wang, Y., Audibert, J.-Y., and Munos, R. (2008). Algorithms for infinitely many-armed bandits. *Advances in Neural Information Processing Systems*, 21.

Wasserman, L. (2004). *All of statistics: a concise course in statistical inference*, volume 26. Springer.

Zhang, M., Zou, J., and Tse, D. (2019). Adaptive monte carlo multiple testing via multi-armed bandits. In *International Conference on Machine Learning*, pages 7512–7522. PMLR.

Zhang, M. J., Ntranos, V., and Tse, D. (2020). Determining sequencing depth in a single-cell rna-seq experiment. *Nature communications*, 11(1):1–11.

Zhang, Z., Ji, X., and Du, S. (2021). Is reinforcement learning more difficult than bandits? a near-optimal algorithm escaping the curse of horizon. In *Conference on Learning Theory*, pages 4528–4531. PMLR.

## Checklist

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes] Results are stated and properly qualified.
   (b) Did you describe the limitations of your work? [Yes]
   (c) Did you discuss any potential negative societal impacts of your work? [N/A]
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [Yes] All theorem's have clearly stated assumptions
   (b) Did you include complete proofs of all theoretical results? [Yes] All proofs are detailed in the Appendix
3. If you ran experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [N/A]
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [N/A]
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [N/A]
   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [N/A]
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
   (a) If your work uses existing assets, did you cite the creators? [N/A]
   (b) Did you mention the license of the assets? [N/A]
   (c) Did you include any new assets either in the supplemental material or as a URL? [N/A]
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [N/A]
   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A]
5. If you used crowdsourcing or conducted research with human subjects...
   (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]
   (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]
   (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]

# A Extensions of the formulation

The formulation of the functional $g$ can be extended in several different ways. First, we note that we can extend the set $S(F)$ to a finite union of disjoint closed intervals, i.e., $S(F) = \cup_{i=1}^k S_i(F)$, where $S_i(F)$ is a closed interval for $i \in [k]$. This is because $\mathbb{E}[X|S(F)]$ can be estimated based on estimations of $\mathbb{E}[X|S_i(F)]$ via

$$\mathbb{E}[X|S(F)] = \frac{\mathbb{E}[X\mathbb{I}(X \in S(F))]}{\mathbb{P}(X \in S(F))} = \frac{\sum_{i=1}^k \mathbb{P}(X \in S_i(F))\mathbb{E}[X|S_i(F)]}{\sum_{i=1}^k \mathbb{P}(X \in S_i(F))}. \quad (7)$$

Observe that this definition naturally extends to cases where the distribution is continuous, where the density of $F$ at $X$ can be substituted for $\mathbb{P}(X \in S(F))$ for singleton sets $S(F)$. We can also consider a more general class of functionals

$$g(F) = \mathbb{E}\left[h(X)|X \in S(F)\right], \quad (8)$$

where $h$ is a differentiable function. However, when we take the limit $\varepsilon \to 0$, we see that for any fixed distribution $F$ and fixed function $h$ the reweighting induced by $h$ does not matter. Assuming that we knew whether $X_i \in S(F)$ for each $i$, we would simply want to sample $X_i \propto h'(X_i)\varepsilon^{-r}$ for some $r$. Since $h$ is differentiable, this is simply reweighting by a constant factor, which does not show up in our $\varepsilon$ dependence. Thus, we can safely only consider the weighting functional $h(x) = x$, which retains the central elimination aspect of this setting (determining whether a point is relevant or not). Loosely speaking, for any differentiable function $h$ and smooth and compactly supported $F$, we have that in the limit as $\varepsilon \to 0$ it degenerates to one of these settings.

# B Results for trimmed mean

In this section, we present our upper and lower bound analysis for trimmed mean via both online and offline sampling algorithms.

## B.1 Upper bound for offline algorithms

For trimmed mean, the following statements are assumed to hold:

**Assumption 4.** *There exist constants $c_0, c_1, \ldots, c_5$ such that*

- $\int x^2 dF(x) \le c_0$.
- $F'(x) \ge c_1$ *for* $|x - F^{-1}(\alpha)| \lesssim \varepsilon$ *and* $|x - F^{-1}(1-\alpha)| \lesssim \varepsilon$.
- $|F^{(2)}(x)| \le c_2$ *for* $|x - F^{-1}(\alpha)| \lesssim \sqrt{\varepsilon}$ *and* $|x - F^{-1}(1-\alpha)| \lesssim \sqrt{\varepsilon}$.
- $F'(x) \le c_3$ *for* $|x - F^{-1}(\alpha)| \lesssim \varepsilon$ *and* $|x - F^{-1}(1-\alpha)| \lesssim \varepsilon$.
- $\max\{|F^{-1}(\alpha)|, |F^{-1}(1-\alpha)|\} \le c_4$, $\min\{|F^{-1}(\alpha)|, |F^{-1}(1-\alpha)|\} \ge c_5$.

The first assumption is to ensure that the mean and variance of $F$ is upper bounded, which is slightly stronger than the assumption for mean. The second assumption is to ensure that the $\alpha$ and $1-\alpha$ quantiles of $F$ is well-defined. The third assumption ensures that the distribution has Lipschitz-continuous density around the quantiles. The forth assumption precludes the distributions which have lots of mass around the $\alpha$ and $1-\alpha$ quantiles. The fifth assumption ensures that the $\alpha$ and $1-\alpha$ quantiles are upper-bounded and bounded away from 0. The following proposition gives the choice of $(n, m)$ to obtain the $(\varepsilon, \delta)$-PAC approximation of the trimmed mean.

**Proposition 4.** *Suppose that Assumption 4 holds. Then, by choosing $m \ge C_1\varepsilon^{-1}\log\varepsilon^{-1}$ and $n \ge C_2\varepsilon^{-2}\delta^{-1}$, the estimator $G_{n,m}$ is an $(\varepsilon, \delta)$-PAC approximation of $g(F)$. Here $C_1, C_2$ are constants which can be expressed by $c_0, \ldots, c_5$. Thus, the offline sampling algorithm takes overall $O(\varepsilon^{-3}\log(1/\varepsilon))$ samples.*

## B.2 Lower bounds for offline algorithms

Similar to the analysis for estimating median, we consider the following quantity

$$\mathrm{KL}_\sigma(\varepsilon) \triangleq \min\{D_{\mathrm{KL}}(F_1 * \mathcal{N}(0,\sigma^2) \| F_2 * \mathcal{N}(0,\sigma^2)) : F_1, F_2 \in \mathcal{F}, |g(F_1) - g(F_2)| \ge 2\varepsilon\}.$$

Analogously, we have the following bounds on the above quantity with respect to the magnitude of noise $\sigma$.

**Lemma 6.** *For $\varepsilon \in (0, 1/4)$, the following characterization of $\mathrm{KL}_\sigma(\varepsilon)$ holds as a function of $\sigma$:*

$$\mathrm{KL}_\sigma(\varepsilon) \begin{cases} \in [C_1\varepsilon^2, C_2\varepsilon^2] & \text{if } \sigma \leq c\varepsilon^{1/2}, \\ \leq C(\theta,\kappa)\varepsilon^\kappa & \text{if } \sigma \geq \varepsilon^{1/2-\theta}, \end{cases}$$

*where $\theta \in (0, 1/4), \kappa \in \mathbb{N}$ are arbitrary parameters, and $c, C_1, C_2, C(\theta, \kappa)$ are absolute constants with the last one depending only on $(\theta, \kappa)$.*

In the same manner, we have the following corollary.

**Corollary 3.1.** *Fix any $\theta > 0$. The $(\varepsilon, .1)$-PAC sample complexity for trimmed mean estimation is $\Omega(\varepsilon^{-3+\theta})$ for any offline algorithm.*

### B.3 Lower bounds for online algorithms

Analogous to the results for median, we start with the following lemma to give bounds of KL divergence between two distributions after the convolution.

**Lemma 7.** *Fix any $\varepsilon, \theta \in (0, 1/4)$, and $\kappa \in \mathbb{N}$. There exists two distributions $F_1, F_2 \in \mathcal{F}$ with $|g(F_1) - g(F_2)| \geq 2\varepsilon$, and*

$$D_{\mathrm{KL}}(F_1 * \mathcal{N}(0,\sigma^2) \| F_2 * \mathcal{N}(0,\sigma^2)) \begin{cases} \in [C_1\varepsilon^{3/2}, C_2\varepsilon^{3/2}] & \text{if } \sigma \leq c\varepsilon^{1/2}, \\ \leq C(\theta,\kappa)\varepsilon^\kappa & \text{if } \sigma \geq \varepsilon^{1/2-\theta}, \end{cases}$$

*where $c, C_1, C_2, C(\theta, \kappa)$ are absolute constants with the last one depending only on $(\theta, \kappa)$.*

We then show the lower bound for trimmed mean via online sampling algorithms.

**Theorem 4.** *Suppose that $\varepsilon > 0$. Denote $\mathcal{F}$ as the set of distributions satisfying Assumption 4. Consider an online algorithm $\pi$ with a fixed budget $t$ which outputs $\hat{G}$. Then, for any $\theta \in (0, 1/4)$, there exists at least one distribution $F \in \mathcal{F}$, such that*

$$\mathbb{P}(|\hat{G} - g(F)| > \varepsilon) \geq \frac{1}{4}\exp\left(-ct\varepsilon^{2.5-2\theta}\right), \tag{9}$$

*where $c > 0$ is a constant.*

## C Proofs of upper bounds for offline algorithms

### C.1 Mean

Here we present the proof of Proposition 1.

*Proof.* Let $X \sim F$ and $Z \sim \mathcal{N}(0, 1/m)$ are independent random variables. Then, we have

$$\mathbb{E}[X + Z] = \mathbb{E}[X] + \mathbb{E}[Z] = \mathbb{E}[X].$$

This implies that $g(F_m) = g(F)$ for any $m \geq 1$. Therefore, we can simply take $m = 1$. Then, we note that

$$\mathrm{Var}[X + Z] = \mathrm{Var}[X] + \mathrm{Var}[Z] \leq c + 1.$$

This implies that $\mathrm{Var}_{\hat{X}\sim F_1}[\hat{X}] \leq c + 1$. According to the Chebyshev inequality, we have

$$\mathbb{P}(|G_{n,m} - g(F_m)| \geq \varepsilon) \leq \frac{\mathrm{Var}_{\hat{X}\sim F_1}[\hat{X}]}{n\varepsilon^2} \leq \frac{c+1}{n\varepsilon^2}.$$

Therefore, by taking $n \geq \delta^{-1}(c + 1)\varepsilon^{-2}$, we have

$$\mathbb{P}(|G_{n,m} - g(F_m)| \leq \varepsilon) \geq 1 - \delta.$$

Hence, it takes $mn = O(\varepsilon^{-2})$ samples to provide an $(\varepsilon, \delta)$-PAC approximation of $g(F)$. $\square$

### C.2 Median

Consider the following conditions

(A1) For $x \in \mathbb{R}$, there exists $c_1, t_1 > 0$ such that for all $t$ satisfying $0 \le |t - x| \le t_1$, $F'(t) \ge c_1$.

(A2) For $x \in \mathbb{R}$, there exists $c_2, t_2 > 0$ such that

$$|F'(x_1) - F'(x_2)|| \le c_2.$$

for $x_1, x_2 \in [x - t_2, x + t_2]$.

We can view Assumption 2 as follows. Let $\eta = g(F) = F^{-1}(0.5)$. $F$ satisfies (A1) with $(\eta, c_1, t_1)$ and satisfies (A2) with $(\eta, c_2, t_2)$ while $t_1 \gtrsim \varepsilon$ and $t_2 \gtrsim \sqrt{\varepsilon}$. Denote $\rho(x) = F'(x)$. Let $\rho_m = \rho * \varphi_{1/m}$ as the pdf of the distribution of $\hat{X}_i$. We start with Lemma 8 to show that under suitable choice of $m$, $F_m$ also satisfies (A1).

**Lemma 8.** *Let $\eta = g(F)$. Assume that $F$ satisfies (A1) with $(\eta, c_1, t_1)$. Suppose that $m^{-1/2} \le t_1/2$. Then, $F_m$ satisfies (A1) with $(\eta, c_1/4, t_1)$.*

*Proof.* It is sufficient to show that for $x \in [\eta - t_1, \eta + t_1]$, $\rho_m(x) > c_1/4$. As $m^{-1/2} < t_1/2$,

$$\int_0^{t_1} \varphi_{1/m}(x)dx \ge \int_0^{2m^{-1/2}} \varphi_{1/m}(x)dx = \int_0^2 \varphi_1(x)dx \ge 1/4.$$

Therefore, for $x \in [\eta, \eta + t_1]$, we have

$$\rho_m(x) = \int_z \varphi_{1/m}(z)\rho(x - z)dz \ge c_1 \int_0^{t_1} \varphi_{1/m}(z)dz \ge c_1/4.$$

Similarly, for $x \in [\eta - t_1, \eta]$, we have

$$\rho_m(x) = \int_z \varphi_{1/m}(z)\rho(x - z)dz \ge c_1 \int_{-t_1}^0 \varphi_{1/m}(z)dz = c_1 \int_0^{t_1} \varphi_{1/m}(z)dz \ge c_1/4.$$

This completes the proof. □

To prove Proposition 2, we introduce the following proposition to give a point-wise bound on the difference between $F(x)$ and $F_m(x)$.

**Proposition 5.** *Suppose that $F$ satisfies (A2) with $(x, c_2, t_2)$ and $\sqrt{4\log(2m^{-1/2})}m^{-1/2} \le t_2$. Then, we have*

$$|F_m(x) - F(x)| \le \frac{c_2 + 1}{2} m^{-1}.$$

*Proof.* With $k = \sqrt{4\log(2m^{-1/2})}$, we have

$$\int_{-\infty}^{-km^{-1/2}} \varphi_{1/m}(y)dy = \int_{km^{-1/2}}^{\infty} \varphi_{1/m}(y)dy \le e^{-k^2/2} \le \frac{1}{4} m^{-1}.$$

For $|y| \le t_2$, as $|F^{(2)}(x - y)| \le c_2$, it follows that

$$|F(x - y) - F(x) - y\rho(x)| \le \frac{c_2 y^2}{2}.$$

Note that $km^{-1/2} = \sqrt{4\log(2m^{-1/2})}m^{-1/2} \le t_2$, we have

$$
\begin{aligned}
&|F_m(x) - F(x)| \\
=&\left|\int (F(x-y) - F(x))\varphi_{1/m}(y)dy\right| \\
\le&\int_{-\infty}^{-km^{-1/2}} \varphi_{1/m}(y)|F(x-y) - F(x)|dy + \int_{km^{-1/2}}^{\infty} \varphi_{1/m}(y)|F(x-y) - F(x)|dy \\
&+\left|\int_{-km^{-1/2}}^{km^{-1/2}} (F(x-y) - F(x) - y\rho(t))\varphi_{1/m}(y)dy\right| + \left|\int_{-km^{-1/2}}^{km^{-1/2}} y\rho(t)\varphi_{1/m}(y)dy\right| \\
\le&\frac{m^{-1}}{2} + \int_{-km^{-1/2}}^{km^{-1/2}} |F(x-y) - F(x) - y\rho(x)|\varphi_{1/m}(y)dy \\
\le&\frac{m^{-1}}{2} + \frac{c_2}{2}\int_{-km^{-1/2}}^{km^{-1/2}} y^2\varphi_{1/m}(y)dy \le \frac{c_2+1}{2}m^{-1}.
\end{aligned}
$$

This completes the proof. □

We restate Proposition 2 as follows and present the proof.

**Proposition 6.** *Suppose that $\delta \in (0,1)$. Assume that (A1) holds at $\eta$ with $(c_1, t_1)$ and (A2) holds at $\eta$ with $(c_2, t_2)$. Suppose that $t_1 \ge \varepsilon/2$ and $t_2 \gtrsim \sqrt{\varepsilon}$. Then, with $m \ge \frac{4(c_2+1)\varepsilon^{-1}}{c_1}$ and $n \ge \frac{28\varepsilon^{-2}\log\delta^{-1}}{c_1^2}$, $G_{n,m}$ is an $(\varepsilon, \delta)$-PAC approximation of $g(F)$.*

*Proof.* Suppose that we use $n$ points and $m$ samples per point. From our choice of $m$, we have

$$
\sqrt{4\log(2m^{1/2})}m^{-1/2} \le t_2, \frac{c_2+1}{2}m^{-1} \le \frac{c_1\varepsilon}{8}.
$$

Let $\eta_m = g(F_m)$ and $\eta = g(F)$. From Proposition 5, we have

$$
|F_m(\eta_m) - F_m(\eta)| = |F(\eta) - F_m(\eta)| \le c_1\varepsilon/8.
$$

From Lemma 8, we note that $F_m$ satisfies (A1) with $(\eta, c_1/4, t_1)$. If $|\eta_\sigma - \eta| \ge t_1$, then, we have

$$
|F_m(\eta_\sigma) - F_m(\eta)| \ge \min\{|F_m(\eta + t_1) - F_m(\eta)|, |F_m(\eta - t_1) - F_m(\eta)|\} \ge \frac{c_1 t_1}{4},
$$

which leads to a contradiction. Therefore, we have

$$
c_1\varepsilon/8 \ge |F_m(\eta_\sigma) - F_m(\eta)| \ge c_1|\eta_\sigma - \eta|/4.
$$

This implies that $|\eta - \eta_m| \le \varepsilon/2$. As $\varepsilon \le t_1/2$, we note that $F_m$ satisfies (A1) with $(\eta_m, c_1/4, t_1/2)$. From the choice of $n$, according to Lemma 13, we have

$$
\mathbb{P}(|G_{n,m} - \eta_m| \le \varepsilon/2) \ge 1 - \delta.
$$

Under the event $|G_{n,m} - \eta_m| \le \varepsilon/2$, we have

$$
|G_{n,m} - \eta| \le |G_{n,m} - \eta_m| + |\eta_m - \eta| \le \varepsilon.
$$

This completes the proof. □

### C.3 Maximum

In this case, the estimator for the noiseless samples writes $G_n = \max_{i\in[n]} X_n$. We first show that for sufficiently large $n$, $F(G_n)$ can be close to 1.

**Proposition 7.** *Suppose that $\varepsilon > 0$ and $\delta \in (0,1)$. Then, for $n \ge \varepsilon^{-1}\log(2/\delta)$, we have $\mathbb{P}(F(G_n) \ge 1 - \varepsilon) \ge 1 - \delta/2$.*

*Proof.* Consider a fixed number of points $n$. Note that $G_n = \max_{i\in[n]} X_i$. Therefore, we have

$$
\begin{aligned}
&\mathbb{P}(F(G_n) \le 1 - \varepsilon) = \mathbb{P}(F(X_i) \le 1 - \varepsilon, \forall i \in [n]) \\
=&(1-\varepsilon)^n \le \exp(\varepsilon^{-1}\log(2/\delta)\log(1-\varepsilon)) \le \delta/2.
\end{aligned}
$$

Here we utilize that $\log(1-\varepsilon) \le -\varepsilon$. This completes the proof. □

Then, based on the $\beta$-regularity of $F$, we show that $G_n$ can be close to $g(F)$ when $n$ is large.

**Proposition 8.** *Let $\varepsilon > 0$ and $\delta \in (0,1)$. Denote $\eta = g(F)$. Suppose that Assumption 3 holds. Then, with $n = c_1^{-1}\varepsilon^{-\beta}\log(2/\delta)$ points, we have $\mathbb{P}(|G_n - \eta| \le \varepsilon) \ge 1 - \delta/2$.*

*Proof.* From Proposition 7, we note that

$$\mathbb{P}(F(G_n) \ge 1 - c_1\varepsilon^\beta) \ge 1 - \delta/2.$$

According to Assumption 3, $F(G_n) \ge 1 - c_1\varepsilon^\beta$ implies that $G_n \ge \eta - \varepsilon$. As $G_n = \max_{i\in[n]} X_i \le \eta$, this completes the proof. □

We first choose $n \ge c_1^{-1}(\varepsilon/2)^{-\beta}\log(2/\delta)$. From Proposition 8, this guarantees that $\mathbb{P}(|G_n - \eta| \le \varepsilon/2) \ge 1 - \delta/2$. Then, by choosing $m \ge 4\varepsilon^{-2}\log(2n\delta^{-1})$, we have

$$\mathbb{P}(|X_i - \hat{X}_i| \le \varepsilon/2) \ge 1 - e^{m^{-1}\varepsilon^{-2}} \ge 1 - \delta/(2n).$$

Here we utilize the tail bound of Gaussian distributions and the fact that $X_i - \hat{X}_i \sim \mathcal{N}(0, 1/m)$. As $G_n = \max_{i\in[n]} X_i$ and $G_{n,m} = \max_{i\in[n]} \hat{X}_i$, conditioned on $\{|X_i - \hat{X}_i| \le \varepsilon/2, \forall i \in [n]\}$, we have $|G_n - G_{n,m}| \le \varepsilon/2$. Therefore, it follows that

$$\mathbb{P}(|G_n - G_{n,m}| \le \varepsilon/2) \ge P\left(|X_i - \hat{X}_i| \le \varepsilon/2, \forall i \in [n]\right) \ge 1 - n\delta/(2n) = 1 - \delta/2.$$

In summary, we have $\mathbb{P}(|\eta - G_{n,m}| \le \varepsilon) \ge \mathbb{P}(|G_n - G_{n,m}| \le \varepsilon/2) + \mathbb{P}(|G_n - \eta| \le \varepsilon/2) - 1 \ge 1 - \delta$.

### C.4 Trimmed mean

Consider the following conditions

- (B1) There exists constant $c_0 > 0$ such that $\int_{-\infty}^{\infty} x^2 dF(x) \le c_0$.
- (B2) For $x \in \mathbb{R}$, there exists $c_1, t_1 > 0$ such that for all $t$ satisfying $0 \le |t - x| \le t_1$, $F'(t) \ge c_1$.
- (B3) For $x \in \mathbb{R}$, there exists $c_2, t_2 > 0$ such that for $x_1, x_2 \in [x - t_2, x + t_2]$,
$$|F'(x_1) - F'(x_2)| \le c_2.$$
- (B4) For $x \in \mathbb{R}$, there exists $c_3, t_3 > 0$ such that for all $t$ satisfying $0 \le |t - x| \le t_3$, $F'(t) \le c_3$.
- (B5) There exists constants $c_4, c_5 > 0$ such that $\max\{|F^{-1}(\alpha)|, |F^{-1}(1-\alpha)|\} \le c_4$, $\min\{|F^{-1}(\alpha)|, |F^{-1}(1-\alpha)|\} \ge c_5$.

We can view Assumption 4 as follows. $F$ satisfies (B1) with $c_0$ and (B5) with $c_4, c_5$. At $F^{-1}(\alpha)$ and $F^{-1}(1-\alpha)$, $F$ satisfies (B2) with $(c_1, t_1)$, satisfies (B3) with $(c_2, t_2)$ and satisfies (B4) with $(c_3, t_3)$. Here $t_1, t_3 \gtrsim \varepsilon$ and $t_2 \gtrsim \sqrt{\varepsilon}$.

We first show that for $n \ge O(\varepsilon^{-2})$, the empirical estimator of the trimmed mean from noiseless samples will be close to the trimmed mean.

**Proposition 9.** *Suppose that Assumption 4 holds. Let $\delta \in (0,1)$. Suppose that $\varepsilon > 0$ is sufficiently small. For*

$$n \ge (2\alpha - 1)^{-2}(4c_3c_4 + 1)^{-2}\varepsilon^{-2}\max\{16c_0\delta^{-1}, 4\log(4/\delta)c_1^{-2}\},$$

*with probability at least $1 - \delta$, we have*

$$\left|\frac{1}{n}\sum_{i=\lfloor \alpha n\rfloor}^{\lfloor (1-\alpha)n\rfloor} X_{(i)} - \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} x dF(x)\right| \le \varepsilon.$$

We defer the proof to Appendix C.4.1. Then, we prove for the noisy case.

**Lemma 9.** *Suppose that Assumption 4 holds. Then, we have*

$$\begin{aligned}&\left|\int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} x dF(x) - \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} x dF_m(x)\right| \\ \le& m^{-1}\left(4\sqrt{c_0} + k^2(2c_2c_4 + 2c_3) + 4 + 2\sqrt{c_0}c_2k^2m^{-1/2}\right) = O(m^{-1}\log(m)),\end{aligned}$$

*where $k = \sqrt{2\log(m/2)}$.*

We leave the proof in Appendix C.4.2. From the median proof, analogously, we also have

$$|F^{-1}(\alpha) - F_m^{-1}(\alpha)| \leq O(m^{-1}\log(m)) \quad , \quad |F^{-1}(1-\alpha) - F_m^{-1}(1-\alpha)| \leq O(m^{-1}\log(m)).$$

Then, we have the bound

$$\begin{aligned}
&\left|\int_{F_m^{-1}(\alpha)}^{F_m^{-1}(1-\alpha)} x dF_m(x) - \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} x dF(x)\right| \\
\leq &\left|\int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} x dF_m(x) - \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} x dF(x)\right| \\
&+ \left|\int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} x dF_m(x) - \int_{F_m^{-1}(\alpha)}^{F_m^{-1}(1-\alpha)} x dF_m(x)\right| \\
\leq & O(m^{-1}\log(m)).
\end{aligned}$$

Here we utilize that $|x|F_m'(x)$ is upper bounded. Therefore, by choosing $m = O(\varepsilon^{-1}\log(1/\varepsilon))$, we have

$$\left|\int_{F_m^{-1}(\alpha)}^{F_m^{-1}(1-\alpha)} x dF_m(x) - \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} x dF(x)\right| \leq \varepsilon/2.$$

By choosing $\varepsilon$ sufficiently small, $F_m$ also satisfies Assumption 4 with constants $(2c_0, c_1/2, 2c_2, 2c_3, 2c_4, c_5/2)$. Therefore, with $n \geq O(\varepsilon^{-2}\delta^{-1})$, we have

$$\mathbb{P}\left(\left|\int_{F_m^{-1}(\alpha)}^{F_m^{-1}(1-\alpha)} x dF_m(x) - G_{m,n}\right| \leq \varepsilon/2\right) \geq 1-\delta.$$

This completes the proof.

### C.4.1 Proof of Proposition 9

*Proof.* For $\xi > 0$, denote the event

$$E(\xi) = \{|X_{(\lfloor \alpha n \rfloor)} - F^{-1}(\alpha)| \leq \xi, |X_{(\lfloor (1-\alpha)n \rfloor)} - F^{-1}(1-\alpha)| \leq \xi\}.$$

Choose $\xi < c_5$. From Lemma 14, with $n \geq 4\log(4/\delta)c_1^{-2}\xi^{-2}$, we have $\mathbb{P}(E(\xi)) \geq 1-\delta/2$. Conditioned on $E(\xi)$, we note that

$$\frac{1}{n}\sum_{i=\lfloor \alpha n \rfloor}^{\lfloor (1-\alpha)n \rfloor} X_{(i)} \geq \frac{1}{n}\sum_{i\in[n]} X_i \mathbb{I}(X_i \in [F^{-1}(\alpha) + \mathtt{sign}(F^{-1}(\alpha))\xi, F^{-1}(1-\alpha) - \mathtt{sign}(F^{-1}(1-\alpha))\xi]),$$

and

$$\frac{1}{n}\sum_{i=\lfloor \alpha n \rfloor}^{\lfloor (1-\alpha)n \rfloor} X_{(i)} \leq \frac{1}{n}\sum_{i\in[n]} X_i \mathbb{I}(X_i \in [F^{-1}(\alpha) - \mathtt{sign}(F^{-1}(\alpha))\xi, F^{-1}(1-\alpha) + \mathtt{sign}(F^{-1}(1-\alpha))\xi]).$$

We introduce the following lemma to show the convergence of trimmed mean.

**Lemma 10.** *Let $a < b$. Then, with $n \geq c_0\varepsilon^{-2}\delta^{-1}$, we have*

$$\mathbb{P}\left(\left|\int_a^b x dF(x) - \frac{1}{n}\sum_{i=1}^n X_i \mathbb{I}(X_i \in [a,b])\right| \leq \varepsilon\right) \geq 1-\delta.$$

*Proof.* Denote $Y_i = X_i \mathbb{I}(X_i \in [a,b])$. Then, $\mathbb{E}[Y_i] = \int_a^b x dF(x)$. Note that $\mathrm{Var}[Y_i] \leq \int_a^b x^2 dF(x) \leq \int_{-\infty}^{\infty} x^2 dF(x) \leq c_0$. Therefore, from the Chebyshev inequality, we have

$$P\left(\left|\int_a^b x dF(x) - \frac{1}{n}\sum_{i=1}^n X_i \mathbb{I}(X_i \in [a,b])\right| \geq \varepsilon\right) \leq \frac{c_0}{n\varepsilon^2} \leq \delta.$$

This completes the proof. □

From Lemma 10, for $n \geq 16c_0\xi^{-2}\delta^{-1}$, with probability at least $1-\delta/4$, we have

$$
\begin{aligned}
&\frac{1}{n}\sum_{i\in[n]} X_i \mathbb{I}(X_i \in [F^{-1}(\alpha) + \texttt{sign}(F^{-1}(\alpha))\xi, F^{-1}(1-\alpha) - \texttt{sign}(F^{-1}(1-\alpha))\xi]) \\
\geq &\int_{F^{-1}(\alpha)+\texttt{sign}(F^{-1}(\alpha)}^{F^{-1}(1-\alpha)-\texttt{sign}(F^{-1}(\alpha)} x dF(x) - \xi \geq \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} x dF(x) - (4c_3c_4+1)\xi,
\end{aligned}
$$

and

$$
\begin{aligned}
&\frac{1}{n}\sum_{i\in[n]} X_i \mathbb{I}(X_i \in [F^{-1}(\alpha) - \texttt{sign}(F^{-1}(\alpha))\xi, F^{-1}(1-\alpha) + \texttt{sign}(F^{-1}(1-\alpha))\xi]) \\
\leq &\int_{F^{-1}(\alpha)-\texttt{sign}(F^{-1}(\alpha)}^{F^{-1}(1-\alpha)+\texttt{sign}(F^{-1}(\alpha)} x dF(x) + \xi \leq \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} x dF(x) + (4c_3c_4+1)\xi.
\end{aligned}
$$

Here we utilize that $\xi \leq c_5 \leq c_4$ and $|xF'(x)| \leq 2c_3c_4$ around $F^{-1}(\alpha)$ and $F^{-1}(1-\alpha)$. Combining the above bound with (C.4.1) and (C.4.1), with probability at least $1-3\delta/4$, we have

$$
\left|\frac{1}{n}\sum_{i=\lfloor \alpha n\rfloor}^{\lfloor (1-\alpha)n\rfloor} X_{(i)} - \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} x dF(x)\right| \leq (4c_3c_5+1)\xi.
$$

Therefore, by letting $\xi = \frac{1}{(2\alpha-1)(4c_3c_5+1)}\varepsilon$, we complete the proof. □

### C.4.2 Proof of Lemma 9

*Proof.* Denote $\sigma = 1/\sqrt{m}$. We also denote $F_\sigma =: F_m$. According to the Cauchy-Schwartz inequality, we have

$$
\left(\int |x| dF(x)\right)^2 \leq \left(\int x^2 dF(x)\right)\left(\int 1 dF(x)\right) \leq c_0,
$$

which implies that $\int |x| dF(x) \leq \sqrt{c_0}$. Let $k > 0$ be a constant. Note that

$$
\int_{k\sigma}^{\infty} \varphi_{\sigma^2}(x)dx = \int_{k}^{\infty} \varphi_1(x)dx, \int_{k\sigma}^{\infty} x\varphi_{\sigma^2}(x)dx = \sigma\int_{k}^{\infty} x\varphi_1(x)dx.
$$

It follows that

$$
\int_{k}^{\infty} x\varphi_1(x)dx = \frac{1}{\sqrt{2\pi}}\int_{k}^{\infty} e^{-\frac{x^2}{2}} d\frac{x^2}{2} = \frac{1}{\sqrt{2\pi}} e^{-k^2/2}.
$$

We note that $\int_k^\infty \varphi_1(x)dx \leq e^{-k^2/2}$. By taking $k = \sqrt{-2\log(\sigma^2/2)}$, then, we have

$$
\int_{k\sigma}^{\infty} \varphi_{\sigma^2}(x)dx \leq e^{-k^2/2} \leq \frac{1}{2}\sigma^2, \int_{k\sigma}^{\infty} x\varphi_{\sigma^2}(x)dx \leq \sigma e^{-k^2/2} \leq \frac{1}{2}\sigma^2.
$$

We can compute that

$$
\begin{aligned}
&\left|\int_a^b x dF_\sigma(x) - \int_a^b yF'(y)dy\right| \\
=&\left|\int_{a\leq y+z\leq b} (y+z)F'(y)\varphi_{\sigma^2}(z)dydz - \int_a^b yF'(y)dy\right| \\
\leq&\left|\int_{a\leq y+z\leq b} yF'(y)\varphi_{\sigma^2}(z)dydz - \int_a^b yF'(y)dy\right| + \left|\int_{a\leq y+z\leq b} zF'(y)\varphi_{\sigma^2}(z)dydz\right|.
\end{aligned}
$$

In the following two lemmas, we show that both terms in the last line are upper bounded by $O(\sigma^2)$.

**Lemma 11.** *We have the bound*

$$\left|\int_{a\le y+z\le b} yF'(y)\varphi_{\sigma^2}(z)dydz - \int_a^b yF'(y)dy\right| \le 4\sqrt{c_0}\sigma^2 + k^2\sigma^2((|b|+|a|)c_2 + 2c_4).$$

**Lemma 12.** *We have the bound*

$$\left|\int_{a\le y+z\le b} zF'(y)\varphi_{\sigma^2}(z)dydz\right| \le 4\sigma^2 + 2\sqrt{c_0}c_2k^2\sigma^3.$$

The proofs are left in Appendix C.4.3 and C.4.4. In summary, we have the bound

$$\begin{aligned}
&\left|\int_a^b xdF(x) - \int_a^b xdF_\sigma(x)\right| \\
\le& 4\sqrt{c_0}\sigma^2 + k^2\sigma^2((|b|+|a|)c_2+2c_4) + 4\sigma^2 + 2\sqrt{c_0}c_2k^2\sigma^3 \\
=& \sigma^2\left(4\sqrt{c_0} + k^2((|b|+|a|)c_2+2c_4) + 4 + 2\sqrt{c_0}c_2k^2\sigma\right)
\end{aligned}$$

This completes the proof. □

### C.4.3 Proof of Lemma 11

We first upper-bound the LHS in (11) by the following parts:

$$\begin{aligned}
&\left|\int_{a\le y+z\le b} yF'(y)\varphi_{\sigma^2}(z)dydz - \int_a^b yF'(y)dy\right| \\
\le&\left|\int_{a+k\sigma}^{b-k\sigma}\left(\int_{a-y}^{b-y}\varphi_{\sigma^2}(z)dz\right)yF'(y)dzdy - \int_{a+k\sigma}^{b-k\sigma} yF'(y)dy\right| && (10)\\
&+\left|\left(\int_{-\infty}^{a-k\sigma}+\int_{b+k\sigma}^{\infty}\right)\left(\int_{a-y}^{b-y}\varphi_{\sigma^2}(z)dz\right)yF'(y)dy\right| && (11)\\
&+\left|\int_{a-k\sigma}^{a+k\sigma}\left(\int_{a-y}^{b-y}\varphi_{\sigma^2}(z)dz\right)yF'(y)dy - \int_a^{a+k\sigma} yF'(y)dy\right| && (12)\\
&+\left|\int_{b-k\sigma}^{b+k\sigma}\left(\int_{a-y}^{b-y}\varphi_{\sigma^2}(z)dz\right)yF'(y)dy - \int_{b-k\sigma}^{b} yF'(y)dy\right|. && (13)
\end{aligned}$$

For the term (10), as $y\in[a+k\sigma, b-k\sigma]$, we have $[-k\sigma,k\sigma]\subseteq[a-y,b-y]$, which implies that

$$\left|\int_{a-y}^{b-y}\varphi_{\sigma^2}(z)dz - 1\right| \le \left(\int_{-\infty}^{-k\sigma}+\int_{k\sigma}^{\infty}\right)\varphi_{\sigma^2}(z)dz \le \sigma^2.$$

Hence, we have

$$\begin{aligned}
&\left|\int_{a+k\sigma}^{b-k\sigma}\left(\int_{a-y}^{b-y}\varphi_{\sigma^2}(z)dz\right)yF'(y)dzdy - \int_{a+k\sigma}^{b-k\sigma} yF'(y)dy\right| \\
\le&\sigma^2\int_{a+k\sigma}^{b-k\sigma}|y|F'(y)dy \le \sigma^2\int_{-\infty}^{\infty}|y|F'(y)dy \le \sqrt{c_0}\sigma^2.
\end{aligned}$$

For the term (11), we note that

$$\begin{aligned}
&\left|\left(\int_{-\infty}^{a-k\sigma}+\int_{b+k\sigma}^{\infty}\right)\left(\int_{y-a}^{y-b}\varphi_{\sigma^2}(z)dz\right)yF'(y)dy\right| \\
\le&\frac{1}{2}\sigma^2\left(\int_{-\infty}^{a-k\sigma}+\int_{b+k\sigma}^{\infty}\right)|y|F'(y)dy \le \frac{\sqrt{c_0}}{2}\sigma^2.
\end{aligned}$$

Here we utilize that for $y \geq b + k\sigma$ or $y \leq a - k\sigma$, we have

$$\int_{a-y}^{b-y} \varphi_{\sigma^2}(z)dz \leq \int_{k\sigma}^{\infty} \varphi_{\sigma^2}(z)dz \leq \frac{1}{2}\sigma^2.$$

For the term (12), we note that

$$\begin{aligned}
&\left|\int_{a-k\sigma}^{a+k\sigma} yF'(y)\left(\int_{a-y}^{b-y} \varphi_{\sigma^2}(z)\right) dzdy - \int_{a}^{a+k\sigma} yF'(y)dy\right| \\
\leq&\left|\int_{0}^{k\sigma} (a+t)F'(a+t)\left(\int_{-(b-a)-t}^{b-a-t} \varphi_{\sigma^2}(z)dz\right) dt - \int_{a}^{a+k\sigma} yF'(y)dy\right| \\
&+\left|\int_{0}^{k\sigma} ((a-t)F'(a-t) - (a+t)F'(a+t))\left(\int_{t}^{b-a+t} \varphi_{\sigma^2}(z)\right) dzdt\right| \\
\leq&\left|\int_{0}^{k\sigma} (a+t)F'(a+t)\left(\int_{-(b-a)-t}^{b-a-t} \varphi_{\sigma^2}(z)dz\right) dt - \int_{0}^{k\sigma} (a+t)F'(a+t)dt\right| \\
&+\int_{0}^{k\sigma} |a|F'(a-t) - F'(a+t)| + t|F'(a-t) + F'(a+t)|dt \\
\leq&\sigma^2\int_{0}^{k\sigma} |a+t|F'(a+t)dt + \int_{0}^{k\sigma} (2|a|c_2 t + 2c_4 t)\, dt \\
\leq&\sqrt{c_0}\sigma^2 + k^2\sigma^2(|a|c_2 + c_4).
\end{aligned}$$

Similarly, for the term (13), we have the bound

$$\left|\int_{b-k\sigma}^{b+k\sigma} yF'(y)\left(\int_{a-y}^{b-y} \varphi_{\sigma^2}(z)\right) dzdy - \int_{b}^{b+k\sigma} yF'(y)dy\right| \leq \sqrt{c_0}\sigma^2 + k^2\sigma^2(|b|c_2 + c_4).$$

#### C.4.4 Proof of Lemma 12

For the LHS in (12), we can decompose it into

$$\begin{aligned}
&\left|\int_{a\leq y+z\leq b} zF'(y)\varphi_{\sigma^2}(z)dydz\right| \\
\leq&\left|\int_{a+k\sigma}^{b-k\sigma} F'(y)\left(\int_{a-y}^{b-y} z\varphi_{\sigma^2}(z)\right) dzdy\right| &(14)\\
&+\left|\left(\int_{-\infty}^{a-k\sigma} + \int_{b+k\sigma}^{\infty}\right)\left(\int_{a-y}^{b-y} z\varphi_{\sigma^2}(z)dz\right) F'(y)dy\right| &(15)\\
&+\left|\int_{a-k\sigma}^{a+k\sigma} F'(y)\left(\int_{a-y}^{b-y} z\varphi_{\sigma^2}(z)\right) dzdy\right| &(16)\\
&+\left|\int_{b-k\sigma}^{b+k\sigma} F'(y)\left(\int_{a-y}^{b-y} z\varphi_{\sigma^2}(z)\right) dzdy\right| &(17)
\end{aligned}$$

For the term (14), we note that

$$\left|\int_{a+k\sigma}^{b-k\sigma} F'(y)\left(\int_{a-y}^{b-y} z\varphi_{\sigma^2}(z)\right) dzdy\right| \leq \quad \sigma^2\int_{a+k\sigma}^{b-k\sigma} F'(y)dy \leq \sigma^2.$$

Here we utilize that for $y \in [a + k\sigma, b - k\sigma]$,

$$\begin{aligned}\left|\int_{a-y}^{b-y} z\varphi_{\sigma^2}(z)dz\right| &= \left|\left(\int_{-\infty}^{a-y} + \int_{b-y}^{\infty}\right) z\varphi_{\sigma^2}(z)dz\right| \\ &\leq \left(\int_{-\infty}^{a-y} + \int_{b-y}^{\infty}\right) |z|\varphi_{\sigma^2}(z)dz \\ &\leq \left(\int_{-\infty}^{-k\sigma} + \int_{k\sigma}^{\infty}\right) |z|\varphi_{\sigma^2}(z)dz \leq \sigma^2.\end{aligned}$$

We can bound the term (15) by

$$\begin{aligned}&\left|\left(\int_{-\infty}^{a-k\sigma} + \int_{b+k\sigma}^{\infty}\right)\left(\int_{a-y}^{b-y} z\varphi_{\sigma^2}(z)dz\right) F'(y)dy\right| \\ \leq &\frac{1}{2}\sigma^2 \left(\int_{-\infty}^{a-k\sigma} + \int_{b+k\sigma}^{\infty}\right) F'(y)dy \leq \frac{1}{2}\sigma^2.\end{aligned}$$

Here we utilize that for $y \leq a - k\sigma$ or $y \geq b + k\sigma$,

$$\left|\int_{a-y}^{b-y} z\varphi_{\sigma^2}(z)dz\right| \leq \int_{a-y}^{b-y} |z|\varphi_{\sigma^2}(z)dz \leq \int_{k\sigma}^{\infty} |z|\varphi_{\sigma^2}(z)dz \leq \frac{1}{2}\sigma^2.$$

For the term (16), we note that

$$\begin{aligned}&\left|\int_{a-k\sigma}^{a+k\sigma} F'(y)\left(\int_{a-y}^{b-y} z\varphi_{\sigma^2}(z)\right) dzdy\right| \\ =&\left|\int_{0}^{k\sigma} F'(a+t)\left(\int_{a-b-t}^{b-a-t} z\varphi_{\sigma^2}(z)\right) dzdt\right| \\ &+\left|\int_{0}^{k\sigma} (F'(a-t) - F'(a+t))\left(\int_{t}^{b-a+t} z\varphi_{\sigma^2}(z)\right) dzdt\right| \\ \leq &\sigma^2 \int_{0}^{k\sigma} F'(a+t)dt + \int_{0}^{k\sigma} 2c_2 t\sigma\sqrt{c_0}dt \\ \leq &\sigma^2 + \sqrt{c_0}c_2k^2\sigma^3.\end{aligned}$$

Here we utilize that for sufficiently small $\sigma$ such that $2k\sigma \leq b - a$,

$$\left|\int_{a-b-t}^{b-a-t} z\varphi_{\sigma^2}(z)\right| \leq \int_{k\sigma}^{\infty} |z|\varphi_{\sigma^2}(z) + \int_{-\infty}^{-k\sigma} |z|\varphi_{\sigma^2}(z) \leq \sigma^2.$$

Similarly, we can bound the term (17) by

$$\left|\int_{b-k\sigma}^{b+k\sigma} F'(y)\left(\int_{a-y}^{b-y} z\varphi_{\sigma^2}(z)\right) dzdy\right| \leq \sigma^2 + \sqrt{c_0}c_2k^2\sigma^3.$$

This completes the proof.

### C.5 Auxiliary results

We introduce the following auxilary lemmas to extend the median results to quantile.

**Lemma 13.** *Let* $\alpha \in (0,1)$. *Suppose that* $0 < \varepsilon < \min\{\alpha, 1-\alpha\}/5$ *and* $\delta \in (0,1)$. *For* $n \geq 4\varepsilon^{-2}\log(2/\delta)$ *points, with probability at least* $1-\delta$, $\hat{\eta} = X_{(\lfloor \alpha n \rfloor)}$ *satisfies*

$$|F(\hat{\eta}) - \alpha| \leq \varepsilon.$$

*Proof.* Assume that $\varepsilon < \min\{\alpha, 1-\alpha\}/5$. Consider the random variable $Z_i = 1$ if $F(X_i) \le \alpha - \varepsilon$ and 0 otherwise. Let $Z = \sum_{i=1}^n Z_i$. By the Chernoff bound,

$$\mathbb{P}(F(\hat{\eta}) \le \alpha - \varepsilon) \le \mathbb{P}(Z \ge \alpha n) \le \mathbb{P}(Z \ge (1 + \varepsilon/\alpha)\mathbb{E}[Z_1]) \le \exp\left(-\frac{4\varepsilon^2 n}{15\alpha}\right).$$

On the other hand, consider the random variable $Z_i' = 1$ if $F(X_i) \ge \alpha + \varepsilon$ and 0 otherwise. Let $Z' = \sum_{i=1}^n Z_i'$. According to the Chernoff bound,

$$\mathbb{P}(F(\hat{\eta}) \ge \alpha + \varepsilon) \le \mathbb{P}(Z' \ge (1-\alpha)n) \le \mathbb{P}(Z' \ge (1+\varepsilon/(1-\alpha))\mathbb{E}[Z_i']) \le \exp\left(-\frac{4\varepsilon^2 n}{15(1-\alpha)}\right).$$

In summary, we have

$$\mathbb{P}(|F(\hat{\eta}) - \alpha| \le \varepsilon) \le 1 - 2\exp\left(-\frac{4\varepsilon^2 n}{15}\right) \le 1 - 2\exp\left(-\frac{\varepsilon^2 n}{4}\right).$$

Therefore, by taking $n = 4\varepsilon^{-2}\log(2/\delta)$, we have $\mathbb{P}(|F(\hat{\eta}) - \alpha| \le \varepsilon) \le \delta$. This completes the proof. □

**Lemma 14.** *Assume that $F$ satisfies (B2) At $F^{-1}(\alpha)$ with $(c_1, t_1)$. Suppose that $\varepsilon \le \min\{t_1, \min\{\alpha, 1-\alpha\}/(5c_1)\}$ and $\delta \in (0,1)$. With $n \ge 4\log(2/\delta)c_1^{-2}\varepsilon^{-2}$ points, we have $\mathbb{P}(|X_{(\lfloor \alpha n \rfloor)} - F^{-1}(\alpha)| \le \varepsilon) \ge 1 - \delta$.*

*Proof.* Note that $c_1\varepsilon \le c_1 t_1 \le \frac{1}{5}\min\{\alpha, 1-\alpha\}$. From Lemma 13, with $n \ge 4\log(2/\delta)c_1^{-2}\varepsilon^{-2}$, we have

$$\mathbb{P}(|F(X_{(\lfloor \alpha n \rfloor)}) - F(F^{-1}(\alpha))| \le c_1\varepsilon) \ge 1 - \delta.$$

Let $\eta = F^{-1}(\alpha)$ and $\hat{\eta} = X_{(\lfloor \alpha n \rfloor)}$. If $|\hat{\eta} - \eta| > c_1 t_1$, as $F$ satisfies (B2) at $\eta$ with $(c_1, c_3, t_1)$, we have

$$|F(X_{(\lfloor \alpha n \rfloor)}) - F(F^{-1}(\alpha))| \ge \min\{|F(\eta + t_1) - F(\eta)|, |F(\eta - t_1) - F(\eta)|\} \ge c_1 t_1 > c_1\varepsilon,$$

which leads to a contradiction. If $|\hat{\eta} - \eta| \le c_1 t_1$, then, in the same manner,

$$c_1\varepsilon|F(X_{(\lfloor \alpha n \rfloor)}) - F(F^{-1}(\alpha))| \ge c_1|\hat{\eta}|.$$

This implies that $|X_{(\lfloor \alpha n \rfloor)} - F^{-1}(\alpha)| \le \varepsilon$. □

# D Proofs of upper bounds for online algorithms

In this Appendix we provide the proof of Theorem 2. As discussed, in order to exploit the Bayesian nature of the problem, we analyze the algorithm in two parts. First, we use the fact that our arms are drawn from a common distribution to find some $n, m$ as in Theorem 1 such that the plug-in estimator $G_{n,m}$ will be an $(\varepsilon/2, \delta/2)$-PAC approximation of $g(F)$. Second, we show that our adaptive algorithm is an $(\varepsilon/2, \delta/2)$-PAC approximation of $G_{n,m}$, but is able to accomplish this using significantly fewer samples. We begin by proving the correctness of our algorithm, afterwards analyzing its sample complexity.

## D.1 Correctness

To show correctness, we need to condition on the $n \times m$ matrix of observed samples $Y$, where $A_{i,j} = X_i + Z_{i,j}$, where $Z_{i,j}$ are i.i.d. $\mathcal{N}(0,1)$. We couple the randomness in the analysis of the offline and online algorithms, considering our random arm pulls for both to be jointly generated, and the same matrix $Y$ fed into each algorithm. Analyzing the online algorithm, we show that it recovers the result of the offline sampling algorithm within error $\varepsilon/2$ with probability at least $1 - \delta/4$.

Notationally, let $\mu_1^{\text{uni}}, \ldots, \mu_n^{\text{uni}}$ be the estimates of samples of offline sampling algorithm and online sampling algorithms with given $(m, n)$. Let $N(i)$ be the number of samples for point $X_i$ from the online algorithm.

Defining $g_n$ as the $n$-sample version of the functional $g$, we proceed by showing that the output of our algorithm is close to the output of the $n, m$ offline sampling algorithm, which is close to $g(F)$. Concretely, for our algorithm output $\hat{G}$, we have that

$$\mathbb{P}(|g(F) - \hat{G}| \geq \varepsilon) \leq \mathbb{P}(|g(F) - G_{n,m}| \geq \varepsilon/2) + \mathbb{P}(|G_{n,m} - \hat{G}| \geq \varepsilon/2).$$

We see from the previous arguments regarding offline sampling that for $n, m$ as selected, we have that

$$\mathbb{P}(|g(F) - G_{n,m}| \geq \varepsilon/2) \leq \delta/2.$$

Now all that remains is to show that the second term is small. We show that when $\alpha_1 = \alpha_2$, the online algorithm exactly recovers the output of the offline sampling algorithm on the event that the confidence intervals hold. When $\alpha_1 \neq \alpha_2$ (the case of the trimmed mean), we show that our estimate $\hat{G}$ is within $\varepsilon/2$ of $G_{n,m}$ with probability at least $1 - \delta/4$ on the event that the confidence intervals hold.

We begin by defining $\xi_1$ as the good event where our arms stay within their confidence intervals:

$$\xi_1 = \bigcap_{r \in \mathbb{N}, i \in [n]} \{|\hat{\mu}_i(r) - \mu_i^{\text{uni}}| \leq b_r\}.$$

We give a lower bound on the probability of the good event $\xi$ in the following lemma.

**Lemma 15** (Confidence intervals)**.** *The event $\xi_1$ defined in* (D.1)*, where the confidence intervals of $\hat{\mu}_i^r$ about $\mu_i^{uni}$ hold, satisfies* $\mathbb{P}(\xi_1) \geq 1 - \delta/4$*.*

*Proof.* Recall that

$$t_r \geq 8 b_r^{-2} \log\left(\frac{16 n \log m}{\delta}\right).$$

With this choice of $t_r$, we have

$$\begin{aligned}
\mathbb{P}(\xi_1^c) &\leq \sum_{r \in \mathbb{N}, i \in [n]} \mathbb{P}(|\hat{\mu}_i(r) - \mu_i^{\text{uni}}| > b_r) \\
&\leq n \sum_{r \in \mathbb{N}} \mathbb{P}(|\hat{\mu}_1(r) - \mu_1| \geq b_r/2) + \mathbb{P}(|\mu_1 - \mu_1^{\text{uni}}| \geq b_r/2) \\
&\leq 4n \sum_{r \leq \lceil \log(m) \rceil} \mathbb{P}(|\hat{\mu}_1(r) - \mu_1| \geq b_r/2) \\
&\leq 4n \sum_{r \leq \lceil \log(m) \rceil} 2 \cdot \exp\left(-2 t_r (b_r/2)^2\right) \leq \delta/4.
\end{aligned}$$

This completes the proof. ☐

On this good event $\xi_1$, we show that our online algorithm exactly recovers the partitioning of arms performed by the offline sampling algorithm. That is, for a given matrix of observations $Y$ overloading notation we can see that $g_n(Y)$, the output of the offline sampling algorithm, satisfies

$$g_n(Y) = \frac{1}{|S_n|} \sum_{i \in S_n} \mu_i^{\text{uni}},$$

where $S_n$ is the set of relevant arms (i.e. those close to the boundary of $S$). Note that $S_n$ is a function of $g$. In the following lemma, we show that the online algorithm correctly identifies the arms in this set.

**Lemma 16.** *On the event $\xi_1$ we have that $\hat{S}_n$ is identical to $S_n$.*

The proof of this Lemma conditions on the good event where all confidence intervals hold, and shows that in this case the boundaries of $\hat{S}_n$ stay accurate throughout the course of the algorithm, and no arms are spuriously eliminated.

*Proof.* In this proof we focus on showing that $\hat{S}_n$ correctly partitions those elements smaller than $\mu^{\text{uni}}_{(\lfloor \alpha_1 n \rfloor)}$ from those greater than this threshold. Identical arguments hold for the analysis of $\mu^{\text{uni}}_{(\lfloor \alpha_2 n \rfloor)}$, which together imply the correctness of $\hat{S}_n$.

In round $r$, we use $(r)$ to represent the corresponding values before the sampling, for example, $\hat{\mu}(r)$. Let $t_i$ be the round that the $i$-th arm is eliminated from the active set. Suppose that the algorithm ends in $T$ rounds. We denote

$$\begin{aligned} G(r) =& \{i : \hat{\mu}_i(r) > \hat{\mu}_{(\lfloor \alpha_1 n \rfloor)}(r) + b_{\min\{r,t_i\}}\}, \\ L(r) =& \{i : \hat{\mu}_i(r) < \hat{\mu}_{(\lfloor \alpha_1 n \rfloor)}(r) - b_{\min\{r,t_i\}}\}, \\ U(r) =& \{i : |\hat{\mu}_i(r) - \hat{\mu}_{(\lfloor \alpha_1 n \rfloor)}(r)| \le b_{\min\{r,t_i\}}\}. \end{aligned}$$

From the definition of $G(r)$ and $L(r)$, it is easy to observe that $|L(T+1)| \le \lfloor \alpha_1 n \rfloor$ and $|G(T+1)| \le n - \lfloor \alpha_1 n \rfloor$. We note that $|U(T+1)| = 0$ and this implies that $|G(T+1)| = n - \lfloor \alpha_1 n \rfloor$ and $|L(T+1)| = \lfloor \alpha_1 n \rfloor$. From the definition of $L(T+1)$, it consists of $\lfloor \alpha_1 n \rfloor$ points with minimal $\hat{\mu}_i(T+1)$, i.e.,

$$\hat{\mu}_{(\lfloor \alpha_1 n \rfloor)}(T+1) = \max_{i \in L(T+1)} \hat{\mu}_i(T+1).$$

Conditioned on the good event $\xi_1$, we have

$$\hat{\mu}_i(T+1) - b_{t_i} \le \mu^{\text{uni}}_i \le \hat{\mu}_i(T+1) + b_{t_i}, i \in [n].$$

Note that $b_{T+1} = 0$. Therefore, for arbitrary $i \in L(T+1)$ and $j \in G(T+1)$, we have

$$\mu^{\text{uni}}_i \le \hat{\mu}_i(T+1) + b_{t_i} \le \hat{\mu}_{(\lfloor \alpha_1 n \rfloor)}(t) \le \hat{\mu}_j(T+1) - b_{t_j} \le \mu^{\text{uni}}_j.$$

Hence, the maximal element in $\{\mu^{\text{uni}}_i\}_{i \in L(T+1)}$ is $\mu^{\text{uni}}_{(\lfloor \alpha_1 n \rfloor)}$, which is the $\alpha_1$-th quantile of $\{\mu^{\text{uni}}_i\}_{i=1}^n$. This also implies that

$$\{i : \hat{\mu}_i(T+1) \ge \hat{\mu}_{(\lfloor \alpha_1 n \rfloor)}(T+1)\} = \{i : \mu^{\text{uni}}_i \ge \mu^{\text{uni}}_{(\lfloor \alpha_1 n \rfloor)}\}.$$

On the other hand, analogously, we note that

$$\{i : \hat{\mu}_i(T+1) \le \hat{\mu}_{(\lfloor \alpha_2 n \rfloor)}(T+1)\} = \{i : \mu^{\text{uni}}_i \le \mu^{\text{uni}}_{(\lfloor \alpha_2 n \rfloor)}\}.$$

By combining the above two equations together, we completes the proof. □

We now split our analysis into cases. When $\alpha_1 = \alpha_2$, we see that all arms in $\hat{S}_n$ will be pulled exactly $m$ times, and so for $i \in \hat{S}_n$ we have that $\mu_i(r) = \mu^{\text{uni}}_i$ for the final round $r$. This implies that $\hat{G} = G_{n,m}$ for this given $Y$.

When $\alpha_1 \neq \alpha_2$, we have that some arms in the set $\hat{S}_n$ have not been pulled $m$ times; they were determined to be in $S_n$ using fewer samples, and removed from the active set as they did not require further sampling. Thus, we will not have that $\hat{G} = G_{n,m}$. Instead, we show that because there are so many points in $\hat{S}_n$, by sampling each of them only once and averaging the results, we obtain $\hat{G}$ which is within $\varepsilon/2$ of $G_{n,m}$ with probability at least $1 - \delta/4$.

**Lemma 17.** *On the event $\xi_1$, Algorithm 1 satisfies $\mathbb{P}(|G_{n,m} - \hat{G}| \ge \varepsilon/2) \le \delta/4$.*

*Proof.* If $\alpha_1 = \alpha_2$ then $G_{n,m} = \hat{G}$, and so the result holds trivially.

If $\alpha_1 \neq \alpha_2$, then on the good event $\xi_1$ where our confidence intervals hold, we have that our online algorithm correctly identifies $S_n(Y)$. Then,

$$\begin{aligned} \mathbb{P}(|G_{n,m} - \hat{G}| \ge \varepsilon/2 \mid \xi_1) &\le 2\mathbb{P}\left(\left|\frac{1}{|S_n|}\sum_{i \in S_n}\left(\tilde{\mu}_i - \mu^{\text{uni}}_i\right)\right| \ge \varepsilon/2\right) \\ &= 2\mathbb{P}\left(\left|\mathcal{N}\left(0, \frac{1+1/m}{|S_n|}\right)\right| \ge \varepsilon/2\right) \le \delta/4, \end{aligned}$$

as this sum is normally distributed with variance decaying with $S_n$, and so for sufficiently large $n$ we have the desired result (as when $\alpha_1 \neq \alpha_2$, $|S_n| \ge \lfloor n(\alpha_2 - \alpha_1) \rfloor$). □

Thus, we see that our algorithm's output $\hat{G}$ will be close to $g(F)$ with high probability.

### D.2 Sample complexity analysis

We now turn to bounding the sample complexity of our online algorithm. For simplicity, we overload $S_n$ in our analysis as $S_n = \text{Conv}(\{\mu_i^{\text{uni}} : i \in S_n\})$. Useful in this analysis will be the distance from $X$ to the boundary of $S_n$ (essentially the gap of $X$), which we define as

$$\text{dist}(X, \partial S_n) = \min\left(|X - \mu^{\text{uni}}_{(\lfloor \alpha_1 n \rfloor)}|, |X - \mu^{\text{uni}}_{(\lfloor \alpha_2 n \rfloor)}|\right)$$

To this end, we provide the following Lemma:

**Lemma 18.** *On the good event $\xi_1$, we have that a given arm $X_i$ will be pulled $N(i)$ times where*

$$N(i) \le \min\left(m, \frac{256 \log\left(\frac{16 n \log m}{\delta}\right)}{[\text{dist}(X_i, \partial S_n)]^2}\right).$$

*Proof.* To begin, no arm can be pulled more than $m$ times by our adaptive algorithm, due to the structure of $t_r$. We now additionally see that by the construction of our $b_r$ confidence intervals, we have that On the good event where our confidence intervals hold, we see that an arm $X_i$ must be eliminated when $2b_r \le \text{dist}(X_i, \partial S_n)$. Due to the iterative halving of $b_r$, this means that arm $i$ must be eliminated by round $r$ where $4b_r \ge \text{dist}(X_i, \partial S_n)$, and so $b_r^{-2} \le 16\,[\text{dist}(X_i, \partial S_n)]^{-2}$. ☐

This is to say, it cannot be pulled more than $m$ times, and if it is far from the boundary of $S_n$ then it can be determined whether it is in the set or not using many fewer samples, only scaling with $[\text{dist}(X, \partial S_n)]^{-2}$.

Thus, the total sample complexity of our online algorithm (for a given matrix of observed samples $Y$ with corresponding arm mean vector $\mu^{\text{uni}}$) is upper bounded by

$$\sum_{i=1}^{n} N(i) \le \sum_{i=1}^{n} \min\left(m, \frac{256 \log\left(\frac{n \log m}{\delta}\right)}{[\text{dist}(X, \partial S_n)]^2}\right).$$

We know by the Glivenko-Cantelli theorem that in the limit $\mu^{\text{uni}}_{(\lfloor \alpha_1 n \rfloor)} \to F^{-1}(\alpha_1)$, but we require finite sample rates to give a useful bound.

**Lemma 19.** *For $n = \Theta(\varepsilon^{-2})$, with probability at least $1 - \delta/8$ in the randomness in $Y$, we have for $i \in \{1, 2\}$ that simultaneously*

$$|\mu^{\text{uni}}_{(\lfloor \alpha_i n \rfloor)} - F^{-1}(\alpha_i)| = O\left(\sqrt{\frac{\log(1/\delta)}{m}}\right).$$

*Denote this event as $\xi_2$.*

*Proof.* This lemma is simply a statement about the correctness of the offline sampling algorithm for estimating the $\alpha_1$/$\alpha_2$-th quantiles, which we have already proven. ☐

Note that we are not conditioning on $\xi_2$ occurring in order for our algorithm to provide the correct output; we are simply utilizing this event to bound our algorithm's sample complexity.

On the good event in Lemma 19 and the good event where our algorithm correctly outputs an $\varepsilon$ accurate estimate and has sample complexity as in (18), we have that

$$\begin{aligned} N(i) &\le \min\left(m, \frac{256 \log\left(\frac{16 n \log m}{\delta}\right)}{\left[\min\left(|X_i - \mu^{\text{uni}}_{(\lfloor \alpha_1 n \rfloor)}|, |X_i - \mu^{\text{uni}}_{(\lfloor \alpha_2 n \rfloor)}|\right)\right]^2}\right) \\ &\le \begin{cases} m & \text{when } \text{dist}(X, \partial S_n) \le C\sqrt{\frac{\log(1/\delta)}{m}}, \\ \frac{1024 \log\left(\frac{16 n \log m}{\delta}\right)}{[\text{dist}(X_i, \partial S)]^2} & \text{when } \text{dist}(X, \partial S_n) > C\sqrt{\frac{\log(1/\delta)}{m}}, \end{cases} \\ &\le \min\left(m, \frac{\log\left(\frac{n \log m}{\delta}\right)}{[\text{dist}(X_i, \partial S)]^2}\right), \end{aligned}$$

where $C$ is an absolute constant and $\text{dist}(X_i, \partial S)$ is defined analogously as $\text{dist}(X_i, \partial S) = \min\left(|X_i - F^{-1}(\alpha_1)|, |X_i - F^{-1}(\alpha_2)|\right)$.

We then have that our sample complexity $M$ is bounded as, conditioned on $\xi_1$ we have that

$$\begin{aligned}
\mathbb{E}[M] &= O\left(\mathbb{E}\left[\sum_{i=1}^{n} \min\left(m, \log\left(\frac{n\log m}{\delta}\right)[\text{dist}(X_i, \partial S_n)]^{-2}\right)\right]\right) \\
&\leq O\left(n\mathbb{E}\left[\min\left(m, \log(n/\delta)\,[\text{dist}(X_i, \partial S_n)]^{-2}\right)|\xi_2\right] + nm\mathbb{P}(\xi_2^c)\right) \\
&\overset{(a)}{\leq} O\left(n\log(n/\delta)\mathbb{E}\left[\min\left(m, \text{dist}(X, \partial S)^{-2}\right)\right]\right).
\end{aligned}$$

where we defined the good event $\xi_2$ as in Lemma 19 where our $\mu^{\text{uni}}_{(\lfloor \alpha_i n\rfloor)}$ are within $\sqrt{\log(1/\delta)/m}$ of their distributional values, i.e. $F^{-1}(\alpha_i)$. We utilize the fact that $n \geq m$ to simplify the sample complexity. (a) comes from that $\mathbb{P}(\xi_2^c) \leq \delta/8$ from Lemma 19 and that for events $E$ with probability greater than $1/2$ and positive random variables $X$, we have that $\mathbb{E}[X|E] \leq 2\mathbb{E}[X]$. This gives us the desired result.

**Theorem 5** (Restating Theorem 2)**.** *Algorithm 1 succeeds in estimating $g(F)$ to within accuracy $\varepsilon$ with probability at least $1-\delta$, and requires at most*

$$O\left(n\log(n/\delta)\mathbb{E}\left[\min\left(m, \text{dist}(X, \partial S)^{-2}\right)\right]\right) \tag{18}$$

*observations in expectation.*

### D.3 Functional-specific upper bounds

From Theorem 2, we are able to derive the upper bound sampling complexity of online algorithms in Table 1 for mean, median, maximum and trimmed mean estimation by analyzing (18) under the functional specific assumptions.

#### D.3.1 Mean estimation

*Proof.* For mean estimation, from Theorem 1 we have that $n = \Theta(\varepsilon^{-2})$ and $m = \Theta(1)$ is sufficient. Therefore, we have an expected sample complexity of

$$\begin{aligned}
\mathbb{E}[M] =& O\left(n\log(n/\delta)\mathbb{E}\left[\min\left(m, \text{dist}(X, \partial S)^{-2}\right)\right]\right) \\
=& O(n\log(n/\delta)) = O(\varepsilon^{-2}\log(1/\varepsilon)).
\end{aligned} \tag{19}$$

This completes the proof. ☐

#### D.3.2 Median estimation

*Proof.* For median estimation, from Theorem 1 we have that $n = \Theta(\varepsilon^{-2})$ and $m = \Theta(\varepsilon^{-1})$ is sufficient. We can compute that

$$\begin{aligned}
&\mathbb{E}\left[\min\left(m, \text{dist}(X, \partial S)^{-2}\right)\right] \\
=& O(1)\int \min\left(\varepsilon^{-1}, (x - F^{-1}(0.5))^{-2}\right) dF(x) \\
=& O(1)\left(\int_{-\infty}^{F^{-1}(0.5)-\sqrt{\varepsilon}} (x - F^{-1}(0.5))^{-2} dF(x) + \int_{F^{-1}(0.5)+\sqrt{\varepsilon}}^{\infty} (x - F^{-1}(0.5))^{-2} dF(x)\right) \\
&+ O(1)\left(\int_{F^{-1}(0.5)-\sqrt{\varepsilon}}^{F^{-1}(0.5)+\sqrt{\varepsilon}} \varepsilon^{-1} dx\right).
\end{aligned}$$

The first term can be bounded using integration by parts, where we note that

$$\begin{aligned}
&\int_{-\infty}^{F^{-1}(0.5)-\sqrt{\varepsilon}} (x-F^{-1}(0.5))^{-2}dF(x) \\
=& -\int_{-\infty}^{F^{-1}(0.5)-\sqrt{\varepsilon}} F'(x)d(x-F^{-1}(0.5))^{-1} \\
=& -F'(x)(x-F^{-1}(0.5))^{-1}\big|_{-\infty}^{F^{-1}(0.5)-\sqrt{\varepsilon}} + \int_{-\infty}^{F^{-1}(0.5)-\sqrt{\varepsilon}} F^{(2)}(x)(x-F^{-1}(0.5))^{-1}dx \\
\leq& \sqrt{\varepsilon^{-1}}F'(F^{-1}(0.5)-\sqrt{\varepsilon}) + \sqrt{\varepsilon^{-1}}\int_{-\infty}^{F^{-1}(0.5)-\sqrt{\varepsilon}} F^{(2)}(x) \\
\leq& 2\sqrt{\varepsilon^{-1}}F'(F^{-1}(0.5)-\sqrt{\varepsilon}) = O(\sqrt{\varepsilon^{-1}}).
\end{aligned}$$

Here we utilize that $F'(x)$ is upper bounded at $F^{-1}(0.5)-\sqrt{\varepsilon}$. Similarly, we have

$$\int_{F^{-1}(0.5)+\sqrt{\varepsilon}}^{\infty} (x-F^{-1}(0.5))^{-2}dF(x) \leq O(\sqrt{\varepsilon}).$$

In summary, we have

$$\mathbb{E}\left[\min\left(m, \text{dist}(X,\partial S)^{-2}\right)\right] \leq O(\sqrt{\varepsilon^{-1}}),$$

and this implies that

$$\begin{aligned}
\mathbb{E}[M] =& O\left(n\log(n/\delta)\mathbb{E}\left[\min\left(m, \text{dist}(X,\partial S)^{-2}\right)\right]\right) \leq O(n\log(n/\delta)\varepsilon^{-0.5}) \\
=& O(\varepsilon^{-2.5}\log(1/\varepsilon)).
\end{aligned}$$

This completes the proof. □

#### D.3.3 Maximum estimation

*Proof.* For maximum estimation, from Theorem 1 we have that $n = \Theta(\varepsilon^{-\beta})$ and $m = \Theta(\varepsilon^{-2})$. We note that

$$\begin{aligned}
&\mathbb{E}\left[\min\left(m, \text{dist}(X,\partial S)^{-2}\right)\right] \\
=& O(1)\int \min\left(\varepsilon^{-2}, (x-F^{-1}(1))^{-2}\right)dF(x) \\
=& O(1)\left(\int_{-\infty}^{F^{-1}(1)-\varepsilon} (x-F^{-1}(1))^{-2}dF(x) + \int_{F^{-1}(1)-\varepsilon}^{\infty} \varepsilon^{-2}dF(x)\right).
\end{aligned}$$

For $\beta < 2$, we can compute that

$$\begin{aligned}
\int_{-\infty}^{F^{-1}(1)-\varepsilon} (x-F^{-1}(1))^{-2}dF(x) =& \int_{-\infty}^{F^{-1}(1)-\varepsilon} F'(x)(x-F^{-1}(1))^{-2}dx \\
=& \int_{\varepsilon}^{\infty} F'(F^{-1}(1)-x)x^{-2}dx \\
\leq& \int_{\varepsilon}^{\infty} c_2\beta x^{\beta-1}x^{-2}dx = \frac{c_2\beta}{2-\beta}\varepsilon^{\beta-2},
\end{aligned}$$

and

$$\int_{F^{-1}(1)-\varepsilon}^{\infty} \varepsilon^{-2}dF(x) = \varepsilon^{-2}(1-F(F^{-1}(1)-\varepsilon)) \leq c_2\varepsilon^{\beta-2}.$$

This implies that

$$\mathbb{E}\left[\min\left(m, \text{dist}(X,\partial S)^{-2}\right)\right] = O(\varepsilon^{\beta-2}).$$

As a result, we have

$$\mathbb{E}[M] = O\left(n\log(n/\delta)\mathbb{E}\left[\min\left(m, \text{dist}(X,\partial S)^{-2}\right)\right]\right) = O(\varepsilon^{-2}\log(1/\varepsilon)).$$

For $\beta = 2$, we can compute that

$$\int_{-\infty}^{F^{-1}(1)-\varepsilon} (x - F^{-1}(1))^{-2} dF(x) = \int_{-\infty}^{F^{-1}(1)-\varepsilon} F'(x)(x - F^{-1}(1))^{-2} dx$$
$$= \int_{\varepsilon}^{F^{-1}(1)-F^{-1}(0)} F'(F^{-1}(1) - x) x^{-2} dx$$
$$\leq \int_{\varepsilon}^{F^{-1}(1)-F^{-1}(0)} 2c_2 x x^{-2} dx = O(\log \varepsilon^{-1}).$$

Hence, we have

$$\mathbb{E}[M] = O\left(n \log(n/\delta) \mathbb{E}\left[\min\left(m, \text{dist}(X, \partial S)^{-2}\right)\right]\right) \leq O(\varepsilon^{-2} \log(1/\varepsilon)).$$

For $\beta > 2$, we note that

$$\int_{-\infty}^{F^{-1}(1)-\varepsilon} (x - F^{-1}(1))^{-2} dF(x) = \int_{-\infty}^{F^{-1}(1)-\varepsilon} F'(x)(x - F^{-1}(1))^{-2} dx$$
$$= \int_{\varepsilon}^{\infty} F'(F^{-1}(1) - x) x^{-2} dx$$
$$\leq \int_{\varepsilon}^{F^{-1}(1)-F^{-1}(0)} c_2 \beta x^{\beta-1} x^{-2} dx = O(1).$$

Hence, we have

$$\mathbb{E}[M] = O\left(n \log(n/\delta) \mathbb{E}\left[\min\left(m, \text{dist}(X, \partial S)^{-2}\right)\right]\right) \leq O(\varepsilon^{-\beta} \log(1/\varepsilon)).$$

In summary, we have

$$\mathbb{E}[M] = O\left(n \log(n/\delta) \mathbb{E}\left[\min\left(m, \text{dist}(X, \partial S)^{-2}\right)\right]\right) \leq O(\varepsilon^{-\max\{\beta,2\}} \log(1/\varepsilon)).$$

This completes the proof. □

#### D.3.4 Trimmed mean estimation

*Proof.* For trimmed mean, we note that the analysis is similar to the case of median. This gives that $O\left(n \log(n/\delta) \mathbb{E}\left[\min\left(m, \text{dist}(X, \partial S)^{-2}\right)\right]\right) \leq O(n \log(n) \varepsilon^{-0.5} \log(1/\varepsilon)) = O(\varepsilon^{-2.5} \log^2(1/\varepsilon))$. □

## E Proofs in Section 4

### E.1 Proof of Lemma 1

*Proof.* Denote $p_{F_1}$ and $p_{F_2}$ as the pdf of $F_1$ and $F_2$ respectively. Let $\sigma = 1/\sqrt{m}$ and let $\varphi_{\sigma^2}$ be the pdf of $\mathcal{N}(0, \sigma^2)$. As $p_{\pi,F_1} = (p_{F_1} * \varphi_{\sigma^2})^{\otimes n}$ and $p_{\pi,F_2} = (p_{F_2} * \varphi_{\sigma^2})^{\otimes n}$, we have

$$\mathrm{D}_{\mathrm{KL}}(p_{\pi,F_1} \| p_{\pi,F_2}) = n \mathrm{D}_{\mathrm{KL}}(p_{F_1} * \varphi_{\sigma^2} \| p_{F_2} * \varphi_{\sigma^2}).$$

On the other hand, we note that

$$p_{F_1} * \varphi_{\sigma^2}(y) = \int_{-\infty}^{\infty} p_{F_1}(x) \varphi_{\sigma^2}(y - x) dx = \mathbb{E}_{X \sim F_1}[\varphi_{\sigma^2}(y - X)].$$

Similarly, we have $p_{F_1} * \varphi_{\sigma^2}(y) = \mathbb{E}_{X' \sim F_2}[\varphi_{\sigma^2}(y - X')]$.

Let $\gamma \in \Gamma$ be a coupling of $F_1$ and $F_2$. Namely, it is a joint distribution of $(X, X')$ and its marginal distribution on $X$ $(X')$ are $F_1$ $(F_2)$. Then, utilizing the convexity of KL divergence, we have

$$\mathrm{D}_{\mathrm{KL}}(p_{F_1} * \varphi_{\sigma^2} \| p_{F_2} * \varphi_{\sigma^2}) = \mathrm{D}_{\mathrm{KL}}(\mathbb{E}_{(X,X') \sim \gamma}[\varphi_{\sigma^2}(y - X)] \| \mathbb{E}_{(X,X') \sim \gamma}[\varphi_{\sigma^2}(y - X)])$$
$$\leq \mathbb{E}_{(X,X') \sim \gamma} \mathrm{D}_{\mathrm{KL}}(\varphi_{\sigma^2}(y - X) \| \varphi_{\sigma^2}(y - X')) = \mathbb{E}_{(X,X') \sim \gamma} \frac{\|X - X'\|_2^2}{2\sigma^2}.$$

By taking the infimum w.r.t. all possible coupling $\gamma$, we note that

$$\mathrm{D}_{\mathrm{KL}}(p_{\pi,F_1} \| p_{\pi,F_2}) = n \mathrm{D}_{\mathrm{KL}}(F_1 * \mathcal{N}(0, \sigma^2) \| F_2 * \mathcal{N}(0, \sigma^2))$$
$$\leq n \inf_{\gamma \in \Gamma} \mathbb{E}_{(X,X') \sim \gamma} \frac{\|X - X'\|_2^2}{2\sigma^2} = \frac{mn}{2} \mathcal{W}(F_1, F_2)^2.$$

This completes the proof. □

### E.2 Proof of Lemma 2

*Proof.* For a given underlying distribution $F_1$ and a given algorithm $\pi$, the joint distribution of $\{(X_i, A_i, Y_i)\}_{i=1}^t$ has the following probability density function

$$p_{\pi,F_1}(\{(x_i,a_i,y_i)\}_{i=1}^t) = \prod_{i=1}^t p_{F_1}(x_i)p_\pi(a_i|(a_j,y_j)_{j=1}^{i-1})p(y_i|a_i,\{x_j\}_{j=1}^i).$$

Thus, we can also write

$$p_{\pi,F_1}(\{(x_i,a_i,y_i)\}_{i=1}^t) = p_\pi(\{(a_i,y_i)\}_{i=1}^t|\{x_i\}_{i=1}^t)\prod_{i=1}^t p_{F_1}(x_i),$$

where

$$p_\pi(\{(a_i,y_i)\}_{i=1}^t|\{x_i\}_{i=1}^t) = \prod_{i=1}^t p_\pi(a_i|(a_j,y_j)_{j=1}^{i-1})p(y_i|x_{a_i}).$$

Thus, the marginal distribution on $\{(a_i,y_i)\}_{i=1}^t$ follows

$$\begin{aligned} p_{\pi,F_1}(\{(a_i,y_i)\}_{i=1}^t) &= \int p_{\pi,F_1}(\{(x_i,a_i,y_i)\}_{i=1}^t)dz_1\ldots dz_t \\ &= \mathbb{E}_{(X_i)_{i=1}^t\sim F_1}[p_\pi(\{(a_i,y_i)\}_{i=1}^t|\{X_i\}_{i=1}^t)]. \end{aligned}$$

Let $F_2$ be a distribution different from $F_1$. We want to bound the KL divergence from $p_{\pi,F_1}(\{(a_i,y_i)\}_{i=1}^t)$ to $p_{\pi,F_2}(\{(a_i,y_i)\}_{i=1}^t)$. Let $\gamma\in\Gamma$ be a joint distribution with marginals $F_1$ and $F_2$. For simplicity, we write $\mathbb{E}_\gamma = \mathbb{E}_{(X_i,X_i')_{i=1}^t\sim\gamma}$. Utilizing the convexity of KL divergence, we note that

$$\begin{aligned} \mathrm{D_{KL}}(p_{\pi,F_1}\|p_{\pi,F_2}) &= \mathrm{D_{KL}}(\mathbb{E}_\gamma[p_\pi(\cdot|\{X_i\}_{i=1}^t)]\|\mathbb{E}_\gamma[p_\pi(\cdot|\{X_i'\}_{i=1}^t)]) \\ &\le \mathbb{E}_\gamma \mathrm{D_{KL}}(p_\pi(\cdot|\{X_i\}_{i=1}^t)\|p_\pi(\cdot|\{X_i'\}_{i=1}^t)). \end{aligned} \tag{20}$$

Given the pair of underlying states $(\{X_i\}_{i=1}^t,\{X_i'\}_{i=1}^t)$, we can compute that

$$\begin{aligned} &\mathrm{D_{KL}}(p_\pi(\cdot|\{X_i\}_{i=1}^t))\|p_\pi(\cdot|\{X_i'\}_{i=1}^t))) \\ =&\mathbb{E}_{\{(A_i,Y_i)\}_{i=1}^t\sim p_\pi(\cdot|\{X_i\}_{i=1}^t)}\left[\sum_{i=1}^t \log\frac{p(Y_i|X_{A_i})}{p(Y_i|X'_{A_i})}\right] \\ =&\mathbb{E}_{\{A_i\}_{i=1}^t\sim p_\pi(\cdot|\{X_i\}_{i=1}^t)}\left[\sum_{i=1}^t\sum_{j=1}^t \mathbb{I}(A_j=i)\frac{1}{2}|X_i-X_i'|^2\right] \\ =&\sum_{i=1}^t \frac{C_i(\{X_j\}_{j=1}^t)}{2}|X_i-X_i'|^2. \end{aligned}$$

Here we let $C_i(\{X_j\}_{j=1}^t) = \mathbb{E}_{\{A_i\}_{i=1}^t\sim p_\pi(\cdot|\{X_i\}_{i=1}^t)}\left[\sum_{j=1}^t \mathbb{I}(A_j=i)\right]$. This implies that

$$\mathbb{E}_\gamma \mathrm{D_{KL}}(p_{\pi,F_1}(\cdot|\{X_j\}_{j=1}^t)\|p_{\pi,F_2}(\cdot|\{X_i'\}_{i=1}^t)) = \quad \mathbb{E}_\gamma\left[\sum_{i=1}^t \frac{C_i(\{X_j\}_{j=1}^t)}{2}|X_i-X_i'|^2\right].$$

We note that $\sum_{i=1}^t C_i(\{X_j\}_{j=1}^t) = t$ and

$$\mathbb{E}_\gamma|X_i-X_i'|^2 \le \mathrm{esssup}_{(X,X')\sim\gamma}|X-X'|^2, i=1,\ldots,t$$

This implies that

$$\mathbb{E}_\gamma \mathrm{D_{KL}}(p_{\pi,F_1}(\cdot|\{X_j\}_{j=1}^t)\|p_{\pi,F_2}(\cdot|\{X_i'\}_{i=1}^t)) \le \frac{t}{2}\cdot \mathrm{esssup}_{(X,X')\sim\gamma}|X-X'|^2.$$

By taking the infinimum w.r.t. $\gamma$, we have

$$\begin{aligned} \mathrm{D_{KL}}(p_{\pi,F_1}\|p_{\pi,F_2}) &\le \inf_{\gamma\in\Gamma}\mathbb{E}_\gamma \mathrm{D_{KL}}(p_{\pi,F_1}(\cdot|\{X_j\}_{j=1}^t)\|p_{\pi,F_2}(\cdot|\{X_i'\}_{i=1}^t)) \\ &\le t\inf_{\gamma\in\Gamma}\mathrm{esssup}_{(X,X')\sim\gamma}\frac{|X-X'|^2}{2} = \frac{t}{2}\mathcal{W}_\infty(F_1,F_2)^2. \end{aligned}$$

This completes the proof. □

### E.3 Proof of Lemma 3

*Proof.* We first give the example for the Wasserstein-2 distance. Let $G(s)$ be defined as

$$G^{-1}(s) = \begin{cases} g_1(s), & t \in [0, 2\varepsilon], \\ s, & s \in [2\varepsilon, 1], \end{cases}$$

where $g(s)$ is a monotonic cubic interpolation satisfying that

$$g_1(0) = \varepsilon, g_1'(0) = 1, g_1(2\varepsilon) = 2\varepsilon, g_1'(2\varepsilon) = 1,$$

We note that the image of $G^{-1}$ is $[\varepsilon, 1]$. Therefore, the domain of $G$ is $[\varepsilon, 1]$. We also note that for $s \in [0, \varepsilon]$, we have

$$|G(s) - s| \leq \varepsilon.$$

Consider the following two distributions. We consider a distribution with

$$F_1(x) = 1 - (1 - x)^\beta$$

in its support $[0, 1]$ and another distribution with CDF

$$F_2(x) = 1 - (G(1 - x))^\beta$$

in its support $[0, 1 - \varepsilon]$. We can verify that $F_1$ and $F_2$ satisfy Assumption 3 and $|\max(F_1) - \max(F_2)| = \varepsilon$. We note that

$$F_1^{-1}(s) = 1 - (1 - s)^{1/\beta}, F_2^{-1}(s) = 1 - G^{-1}((1 - s)^{1/\beta}),$$

We can verify that $F_1$ and $F_2$ satisfy Assumption 3 and $|\max(F_1) - \max(F_2) |= \varepsilon$. The Wasserstein-2 distance between $F_1$ and $F_2$ can be computed as

$$\begin{aligned} \mathcal{W}_2(F_1, F_2)^2 &= \int_0^1 \left(F_1^{-1}(s) - F_2^{-1}(s)\right)^2 ds \\ &= \int_0^{\varepsilon^\beta} \left(s^{1/\beta} - G(s^{1/\beta})\right)^2 ds \\ &\leq \int_0^{\varepsilon^\beta} \varepsilon^2 ds = \varepsilon^{\beta+2}. \end{aligned}$$

We then give the example for the Wasserstein-$\infty$ distance. Consider a distribution with CDF

$$F_1(x) = 1 - (1 - x)^\beta$$

in its support $[0, 1]$ and another distribution with CDF

$$F_2(x) = 1 - (1 - x + \varepsilon)^\beta$$

in its support $[\varepsilon, 1 + \varepsilon]$. Let $\gamma$ be the joint distribution of $(X, X + \varepsilon)$, where $X$ follows $F_1$. Then, $\gamma \in \Gamma$ is the coupling of $F_1$ and $F_2$. We can compute that

$$\text{esssup}_{(X,X') \sim \gamma} |X - X'| = \varepsilon.$$

This implies that $\mathcal{W}_\infty(F_1, F_2) \leq \varepsilon$.

Finally, we give the example for the KL divergence. Consider two distributions with following CDFs:

$$F_1(x) = \begin{cases} \dfrac{1 - (1 - x)^\beta}{1 - \varepsilon^\beta}, & x \in [0, 1 - \varepsilon], \\ 0, & x \in (-\infty, 0), \\ 1, & x \in (1 - \varepsilon, \infty). \end{cases}$$

$$F_2(x) = \begin{cases} 1 - (1 - x)^\beta, & x \in [0, 1], \\ 0, & x \in (-\infty, 0), \\ 1, & x \in (1 - \varepsilon, \infty). \end{cases}$$

We can verify that $F_1$ and $F_2$ satisfy Assumption 3 and $|\max(F_1) - \max(F_2) |= \varepsilon$. We note that

$$\sup_{x\in[0,1-\varepsilon]} \frac{p_{F_1}(x)}{p_{F_2}(x)} = \frac{1}{1-\varepsilon^\beta} =: \zeta.$$

Thus, according to the reverse Pinsker inequality, we have

$$\mathrm{D}_{\mathrm{KL}}(F_1\|F_2) \le \frac{\log\zeta}{1-\zeta^{-1}} \mathrm{D}_{\mathrm{TV}}(F_1\|F_2).$$

We note that $\lim_{\varepsilon\to 0} \frac{\log\zeta}{1-\zeta^{-1}} = \lim_{\varepsilon\to 0} \frac{-\log(1-\varepsilon^\beta)}{\varepsilon^\beta} = 1$. For $x \in [0, 1-\varepsilon]$.

$$F_1(x) - F_2(x) = \frac{\varepsilon^\beta}{1-\varepsilon^\beta} F_2(x).$$

Therefore, we have

$$\mathrm{D}_{\mathrm{TV}}(F_1, F_2) = \max_{x\in[0,1]} (F_1(x) - F_2(x)) \le \frac{\varepsilon^\beta}{1-\varepsilon^\beta} = O(\varepsilon^\beta).$$

This implies that

$$\mathrm{D}_{\mathrm{KL}}(F_1\|F_2) \le O(\varepsilon^\beta).$$

This completes the proof. □

# F Proof of lower bounds for median estimation

We start with an auxiliary lemma to pointwisely bound the log-likelihood difference of two distributions.

**Lemma 20.** *Consider two densities supported on $[-1, 1]$ with pdf $p(x)$ and $q(x)$ such that $p(x), q(x) \ge 1/4$ and $|p(x) - q(x)| \le \varepsilon \cdot \mathbb{1}(|x| \le \zeta)$ for all $x \in [-1, 1]$, and $\int_{-\zeta}^{\zeta} x^\ell (p(x) - q(x))dx = 0$ for all $\ell = 0, \cdots, k$. Then for $\sigma \le 1/2$,*

$$\left|\log\frac{p * \varphi_{\sigma^2}(x)}{q * \varphi_{\sigma^2}(x)}\right| \le C\varepsilon \left(\frac{\zeta}{\sigma}\right)^{k+2}, \quad \forall x \in \mathbb{R},$$

*where $\varphi_{\sigma^2}$ is the density function of $\mathcal{N}(0, \sigma^2)$, and $C > 0$ is an absolute constant.*

*Proof.* Write $h = p - q$, then

$$\begin{aligned}
|h * \varphi_{\sigma^2}(x)| &= \left|\int_{-\zeta}^{\zeta} h(y)\varphi_{\sigma^2}(x-y)dy\right| \\
&\overset{\text{(a)}}{=} \left|\int_{-\zeta}^{\zeta} h(y)\varphi_{\sigma^2}(x) \sum_{\ell=0}^{\infty} H_\ell\left(\frac{x}{\sigma}\right)\frac{y^\ell}{\ell!\sigma^\ell}dy\right| \\
&= \left|\sum_{\ell=0}^{\infty} \varphi_{\sigma^2}(x) H_\ell\left(\frac{x}{\sigma}\right)\int_{-\zeta}^{\zeta} h(y)\frac{y^\ell}{\ell!\sigma^\ell}dy\right| \\
&\overset{\text{(b)}}{\le} 2\varepsilon\zeta\left(\frac{\zeta}{\sigma}\right)^{k+1} \cdot \sum_{\ell=k+1}^{\infty} \frac{\varphi_{\sigma^2}(x)}{\ell!}\left|H_\ell\left(\frac{x}{\sigma}\right)\right| \\
&\overset{\text{(c)}}{\le} 2\varepsilon\zeta\left(\frac{\zeta}{\sigma}\right)^{k+1} \cdot \sum_{\ell=k+1}^{\infty} \varphi_{\sigma^2}(x) \sum_{m=0}^{\lfloor \ell/2\rfloor} \frac{(|x|/\sigma)^{\ell-2m}}{(\ell-2m)!m!2^m} \\
&\le 2\varepsilon\zeta\left(\frac{\zeta}{\sigma}\right)^{k+1} \cdot \sum_{m=0}^{\infty} \frac{\varphi_{\sigma^2}(x)}{m!2^m} \sum_{\ell=2m}^{\infty} \frac{(|x|/\sigma)^{\ell-2m}}{(\ell-2m)!} \\
&\le 2\varepsilon\zeta\left(\frac{\zeta}{\sigma}\right)^{k+1} \cdot 2\varphi_{\sigma^2}(x)e^{|x|/\sigma} \\
&= 4e^{1/2}\varepsilon\zeta\left(\frac{\zeta}{\sigma}\right)^{k+1} \cdot \varphi_{\sigma^2}(|x|-\sigma),
\end{aligned}$$

where $H_\ell(x) = \ell! \cdot \sum_{m=0}^{\lfloor \ell/2 \rfloor} \frac{(-1)^m x^{\ell-2m}}{m!(\ell-2m)!2^m}$ is the Hermite polynomial, (c) uses its analytical form, and (a) uses its exponential generating function:

$$\sum_{\ell=0}^{\infty} H_\ell(x) \frac{t^\ell}{\ell!} = \exp\left(xt - \frac{x^2}{2}\right).$$

As for the step (b), we use the assumed property of $h$ to conclude that

$$\left| \int_{-\zeta}^{\zeta} y^\ell h(y) dy \right| \le 2\varepsilon \zeta^{\ell+1} \cdot \mathbb{1}(\ell > k).$$

On the other hand, to lower bound the denominator $q * \phi_{\sigma^2}(x)$, we have the following observations: as $\sigma \le 1/2$,

$$\begin{aligned}
\int_{-1}^{1} \mathbb{1}(\varphi_{\sigma^2}(x-y) \ge \varphi_{\sigma^2}(|x|-\sigma)/e^2) dy &\ge \int_{-1}^{1} \mathbb{1}(\varphi_{\sigma^2}(x-y) \ge \varphi_{\sigma^2}(2\sigma)) dy \\
&\ge \int_{-1}^{1} \mathbb{1}(0 \le y \cdot \mathrm{sign}(x) \le \sigma) dy \\
&\ge \sigma, \quad \text{if } |x| \le 2\sigma; \\
\int_{-1}^{1} \mathbb{1}(\varphi_{\sigma^2}(x-y) \ge \varphi_{\sigma^2}(|x|-\sigma)) dy &\ge \int_{-1}^{1} \mathbb{1}(\sigma \le y \cdot \mathrm{sign}(x) \le 2\sigma) dy \\
&\ge \sigma, \quad \text{if } |x| > 2\sigma.
\end{aligned}$$

Consequently, by Markov's inequality,

$$\begin{aligned}
q * \varphi_{\sigma^2}(x) &\ge \frac{1}{4} \int_{-1}^{1} \varphi_{\sigma^2}(x-y) dy \\
&\ge \frac{\varphi_{\sigma^2}(|x|-\sigma)}{4e^2} \cdot \int_{-1}^{1} \mathbb{1}(\varphi_{\sigma^2}(x-y) \ge \varphi_{\sigma^2}(|x|-\sigma)/e^2) dy \\
&\ge \frac{\varphi_{\sigma^2}(|x|-\sigma)}{4e^2} \cdot \sigma.
\end{aligned}$$

A combination of the above inequalities leads to

$$\left| \frac{p * \varphi_{\sigma^2}(x)}{q * \varphi_{\sigma^2}(x)} - 1 \right| = \frac{|h * \varphi_{\sigma^2}(x)|}{q * \varphi_{\sigma^2}(x)} \le 16 e^{5/2} \varepsilon \left(\frac{\zeta}{\sigma}\right)^{k+2},$$

and therefore the claimed result. □

Then, we introduce a lemma for constructing two distributions with matched moments.

**Lemma 21.** *Let $\varepsilon > 0$. For any $k = 1, 2, \dots$, there exists a constant $b > 0$ and a function $h(x)$ supported in $[-\sqrt{\varepsilon}, \sqrt{\varepsilon}]$ such that*

$$\int_0^\infty h(x) dx = \varepsilon, \quad \int h(x) x^i dx = 0, i = 0, 1, \dots, 2k,$$

*For $i > k$,*

$$\left| \int h(x) x^{2i-1} dx \right| \le 2b\varepsilon^{i+1/2}, \quad \int h(x) x^{2i} dx = 0.$$

*We further have $|h(x)| \le b\sqrt{\varepsilon}$ and $h(x)$ is $b$-Lipschitz continuous. Here the constant $b$ only depends on $k$ and not on $\varepsilon$.*

*Proof.* Note that we only need to prove the lemma for $\varepsilon = 1$, as $h(x) = \sqrt{\varepsilon} h_0(x/\sqrt{\varepsilon})$ only properly scales the moments and preserves Lipschitzness. For $h_0$, consider the following form

$$h_0(x) = \begin{cases} -h_1(-x), & x \in [-1, 0] \\ h_1(x), & x \in [0, 1] \\ 0, & \text{otherwise} \end{cases},$$

where $h_1(x)$ is a polynomial taking the form

$$h_1(x) = \sum_{i=1}^{k+2} a_i x^i.$$

Let $(a_1, \ldots, a_{k+2})$ be the unique solution to the following linear system:

$$\sum_{i=1}^{k+2} a_i = 0, \quad \sum_{i=1}^{k+2} \frac{a_i}{i+1} = 1, \quad \sum_{i=1}^{k+2} \frac{a_i}{2j+i} = 0, \ j = 1, \ldots, k.$$

Let $b \triangleq \sum_{i=1}^{k+2} i|a_i|$. Then clearly $h_1(0) = h_1(1) = 0$, $|h_1(x)| \le \sum_{i=1}^{k+2} |a_i| \le b$, and $|h_1'(x)| \le \sum_{i=1}^{k+2} i|a_i| \le b$. It remains to check the odd moments of $h_0$ (all even moments are zero by symmetry). Specifically,

$$\int_0^\infty h_0(x)dx = \int_0^\infty h_1(x)dx = \sum_{i=1}^{k+2} \frac{a_i}{i+1} = 1;$$

$$\int_{-\infty}^\infty h_0(x)x^{2j-1}dx = 2\int_0^\infty h_1(x)x^{2j-1}dx = 2\sum_{i=1}^{k+2} \frac{a_i}{2j+i} = 0, \quad 1 \le j \le k;$$

$$\left|\int_{-\infty}^\infty h_0(x)x^{2j-1}dx\right| = 2\left|\int_0^1 h_1(x)x^{2j-1}dx\right| \le 2b, \quad j > k.$$

This completes the proof. □

### F.1 Proof of Lemma 4

For $\sigma \le c\varepsilon^{1/2}$, we consider two Gaussian distribution $F_1$ as the CDF of $N(0,1)$ and $F_2$ as the CDF of $N(3\varepsilon, 1)$. Then, $g(F_2) - g(F_1) = 3\varepsilon$ and $\mathrm{D_{KL}}(F_1 \| F_2) = O(\varepsilon^2)$. From the data-processing inequality, we have

$$\mathrm{D_{KL}}(F_1 * \mathcal{N}(0,\sigma^2) \| F_2 * \mathcal{N}(0,\sigma^2)) \le \mathrm{D_{KL}}(F_1 \| F_2) = O(\varepsilon^2).$$

To show the lower bound $\Omega(\varepsilon^2)$, without loss of generality we assume that $F_1^{-1}(0.5) = 0$ and $F_2^{-1}(0.5) \ge \varepsilon$. Proposition 5 and the density lower bound in Assumption 2 show that

$$\begin{aligned} F_1 * \varphi_{\sigma^2}(0) &\ge F_1(0) - \frac{c_2+1}{2}\sigma^2 = \frac{1}{2} - \frac{c_2+1}{2}\sigma^2, \\ F_2 * \varphi_{\sigma^2}(0) &\le F_2(0) + \frac{c_2+1}{2}\sigma^2 \le F_2(\varepsilon) - c_1\varepsilon + \frac{c_2+1}{2}\sigma^2 \le \frac{1}{2} - c_1\varepsilon + \frac{c_2+1}{2}\sigma^2. \end{aligned}$$

Consequently, for $\sigma \le c\varepsilon^{1/2}$ with a small constant $c > 0$, it holds that

$$F_1 * \varphi_{\sigma^2}(0) - F_2 * \varphi_{\sigma^2}(0) = \Omega(\varepsilon).$$

Therefore, Pinsker's inequality gives

$$\begin{aligned} \mathrm{D_{KL}}(F_1 * \mathcal{N}(0,\sigma^2) \| F_2 * \mathcal{N}(0,\sigma^2)) &\ge \mathrm{D_{TV}}(F_1 * \mathcal{N}(0,\sigma^2) \| F_2 * \mathcal{N}(0,\sigma^2))^2 \\ &\ge |F_2 * \varphi_{\sigma^2}(0) - F_1 * \varphi_{\sigma^2}(0)|^2 = \Omega(\varepsilon^2). \end{aligned}$$

For $\sigma \ge \varepsilon^{1/2-\theta}$, let $F_1$ be uniform on $[-1,1]$. Then, $g(F_1) = 0$. Clearly $F_1$ satisfies Assumption 2. To construct $F_2$, we take the construction of $h(x)$ in Lemma 21 with $k \ge 2\kappa/\theta$ and support on $[-\varepsilon_1, \varepsilon_1]$, with $\varepsilon_1 = b\varepsilon/c_2$. Here $b$ is the Lipschitz constant in Lemma 21, and $c_2$ is the smoothness constant in Assumption 2. The density of $F_2$ is then taken to be

$$p_{F_2}(x) = p_{F_1}(x) + \frac{c_2}{b}h(x).$$

As long as $\varepsilon$ is sufficiently small, we have $|p_{F_2}(x) - 1/2| \le c_2/b \cdot b\sqrt{\varepsilon_1} \le 1/4$ everywhere on $x \in [-1,1]$. In other words, $p_{F_2}(x) \in [1/4, 3/4]$ on its support. Moreover, $p_{F_2}'(x) \le c_2/b \cdot b = c_2$. This shows that $F_2$ satisfies Assumption 2 as well.

We first show that the median difference between $F_1$ and $F_2$ is at least $\varepsilon$. In fact, by the density upper bound $p_{F_2}(x) \le 3/4$, we have

$$g(F_2) \ge \frac{4}{3}\left(\frac{1}{2} - F_2(0)\right) = -\frac{4}{3}\cdot\frac{c_2}{b}\int_{-\varepsilon_1}^{0} h(x)dx = \frac{4\varepsilon}{3} > \varepsilon.$$

Next we upper bound the KL divergence between Gaussian convolutions. By choosing $\varepsilon$ sufficiently small, we have $\varepsilon_1 \le \varepsilon^{1-\theta}$. From Lemma 20 and the property of $h(x)$, we immediately have

$$\begin{aligned}
&\mathrm{D}_{\mathrm{KL}}(F_1 * \mathcal{N}(0,\sigma^2)\|F_2 * \mathcal{N}(0,\sigma^2)) \\
\le&\max_{x\in\mathbb{R}} \frac{\log(p_{F_1} * \varphi_{\sigma^2}(x))}{\log(p_{F_2} * \varphi_{\sigma^2}(x))} \\
=&O\left(\left(\frac{\sqrt{\varepsilon_1}}{\sigma}\right)^{k+2}\right) = O(\varepsilon^{k\theta/2}) = O(\varepsilon^{\kappa}).
\end{aligned}$$

### F.2 Proof of Lemma 5

We construct the same pair of distributions $(F_1, F_2)$ as in Lemma 4, and it suffices to prove that when $\sigma \le c\varepsilon^{1/2}$, we have

$$\mathrm{D}_{\mathrm{KL}}(F_2 * \mathcal{N}(0,\sigma^2)\|F_1 * \mathcal{N}(0,\sigma^2)) = \Theta(\varepsilon^{1.5}).$$

For the upper bound, we simply use the data-processing inequality:

$$\begin{aligned}
\mathrm{D}_{\mathrm{KL}}(F_2 * \mathcal{N}(0,\sigma^2)\|F_1 * \mathcal{N}(0,\sigma^2)) &\le \mathrm{D}_{\mathrm{KL}}(F_2\|F_1) \le \chi^2(F_2\|F_1) \\
&= \int_{-1}^{1} \frac{(p_{F_1}(x) - p_{F_2}(x))^2}{p_{F_1}(x)}dx \\
&\le \left(\frac{c_2}{b}\right)^2 \int_{-1}^{1} \frac{h(x)^2}{1/4}dx \\
&\le 4\left(\frac{c_2}{b}\right)^2 \cdot \int_{-\sqrt{\varepsilon_1}}^{\sqrt{\varepsilon_1}} (b\sqrt{\varepsilon_1})^2 dx = O(\varepsilon^{1.5}).
\end{aligned}$$

For the lower bound, the same proof of Lemma 4 shows that $\mathrm{D}_{\mathrm{TV}}(F_1 * \mathcal{N}(0,\sigma^2), F_2 * \mathcal{N}(0,\sigma^2)) = \Omega(\varepsilon)$. A naïve application of Pinsker's inequality only leads to an $\Omega(\varepsilon^2)$ lower bound on the KL divergence. A better lower bound is obtained by noticing that the signed measure $(F_1 - F_2) * \mathcal{N}(0,\sigma^2)$ is effectively supported on $[-\Theta(\sqrt{\varepsilon_1}), \Theta(\sqrt{\varepsilon_1})]$.

To this end, recall from the proof of Lemma 4 that

$$F_1 * \mathcal{N}(0,\sigma^2)(0) - F_2 * \mathcal{N}(0,\sigma^2)(0) = \Omega(\varepsilon).$$

On the other hand, Proposition 5 tells that

$$\begin{aligned}
|F_1 * \mathcal{N}(0,\sigma^2)(-\sqrt{\varepsilon_1}) - F_2 * \mathcal{N}(0,\sigma^2)(-\sqrt{\varepsilon_1})| &\le |F_1(-\sqrt{\varepsilon_1}) - F_2(-\sqrt{\varepsilon_1})| + (c_2+1)\sigma^2 \\
&= (c_2+1)\sigma^2.
\end{aligned}$$

Therefore, for $\sigma \le c\varepsilon^{1/2}$ with a small enough $c > 0$, we have

$$\begin{aligned}
&\int_{-\sqrt{\varepsilon_1}}^{0} |p_{F_1} * \mathcal{N}(0,\sigma^2)(x) - p_{F_2} * \mathcal{N}(0,\sigma^2)(x)|dx \\
\ge&|F_1 * \mathcal{N}(0,\sigma^2)(0) - F_1 * \mathcal{N}(0,\sigma^2)(-\sqrt{\varepsilon_1}) - (F_2 * \mathcal{N}(0,\sigma^2)(0) - F_2 * \mathcal{N}(0,\sigma^2)(-\sqrt{\varepsilon_1}))| \\
\ge&F_1 * \mathcal{N}(0,\sigma^2)(0) - F_2 * \mathcal{N}(0,\sigma^2)(0) - |F_1 * \mathcal{N}(0,\sigma^2)(-\sqrt{\varepsilon_1}) - F_2 * \mathcal{N}(0,\sigma^2)(-\sqrt{\varepsilon_1})| \\
=&\Omega(\varepsilon).
\end{aligned}$$

Let $p(x)$ and $q(x)$ be the shorthands of $p_{F_2} * \mathcal{N}(0, \sigma^2)(x)$ and $p_{F_1} * \mathcal{N}(0, \sigma^2)(x)$, respectively. The KL-divergence can be lower bounded as follows:

$$\begin{aligned}
\mathrm{D}_{\mathrm{KL}}(F_2 * \mathcal{N}(0,\sigma^2) \| F_1 * \mathcal{N}(0,\sigma^2)) &= \int_{-\infty}^{\infty} p(x) \log \frac{p(x)}{q(x)} dx \\
&= \int_{-\infty}^{\infty} \left( p(x) \log \frac{p(x)}{q(x)} - p(x) + q(x) \right) dx \\
&\overset{\text{(a)}}{\geq} \int_{-\sqrt{\varepsilon_1}}^{0} \left( p(x) \log \frac{p(x)}{q(x)} - p(x) + q(x) \right) dx \\
&\overset{\text{(b)}}{\geq} \Omega(1) \cdot \int_{-\sqrt{\varepsilon_1}}^{0} (p(x) - q(x))^2 dx \\
&\overset{\text{(c)}}{\geq} \Omega(1) \cdot \frac{1}{\sqrt{\varepsilon_1}} \cdot \left( \int_{-\sqrt{\varepsilon_1}}^{0} |p(x) - q(x)| dx \right)^2 \\
&= \Omega(\varepsilon^{1.5}),
\end{aligned}$$

where

- (a) is due to the non-negativity of $a \log(a/b) - a + b \geq 0$;
- (b) follows from $a \log(a/b) - a + b \asymp (a-b)^2$ whenever $a, b = \Theta(1)$. The latter follows from $|p(x) - p_{F_1}(x)| = O(\sigma) = O(1)$ from Proposition 5, and similarly for $q(x)$;
- (c) makes use of the Cauchy-Schwarz inequality.

This completes the proof.

### F.3 Proof of Theorem 3

From Le Cam's two-point lower bound, it is sufficient to show that the following proposition holds.

**Proposition 10.** *Suppose that $\varepsilon > 0$. Let $\mathcal{F}$ denote the set of distributions satisfying Assumption 2. Consider an online algorithm $\pi$ with a fixed budget $t$ which outputs $\hat{G}$. Given the distribution $F \in \mathcal{F}$ of the underlying arms and the algorithm $\pi$, let $p_{\pi,F}(\{(a_i, y_i)\}_{i=1}^{t})$ denote the distribution of the action-observation pairs up to the $t$-th iteration. Then, for any $\theta \in (0, 1/4)$, there exist $F_1, F_2 \in \mathcal{F}$ with median $g(F_1)$ and $g(F_2)$ such that $|g(F_1) - g(F_2)| \geq \varepsilon$ and*

$$\mathrm{D}_{\mathrm{KL}}(p_{\pi,F_1} \| p_{\pi,F_2}) \leq O(\varepsilon^{2.5-\theta} t).$$

We start with a general log-sum inequality.

**Lemma 22.** *Suppose that $p, q, K$ are probability density functions. Then, we have the following inequality:*

$$(p * K) \log \frac{p * K}{q * K} \leq \left( p \log \frac{p}{q} \right) * K.$$

Then, we observe that the pair of distributions $(F_1, F_2)$ constructed in the proof of Lemma 5 satisfies the following property.

**Proposition 11.** *Suppose that $\theta > 0$ is a given constant. Let $\mathcal{F}$ denote the set of distributions satisfying Assumption 2. Then, there exists two distribution $F_1, F_2 \in \mathcal{F}$ with median $g(F_1)$ and $g(F_2)$ such that $g(F_2) = g(F_1) + \varepsilon$ and they satisfy that*

$$\mathrm{D}_{\mathrm{KL}}(F_1 \| F_2) = O(\varepsilon^{1.5}),$$

*Denote $\varphi_{\sigma^2}$ as the pdf of $\mathcal{N}(0, \sigma^2)$. For sufficiently small $\varepsilon$ and for $\sigma$ satisfying $\sigma^2 \geq \varepsilon^{1-\theta}$, we further have*

$$\left| \log \frac{p_{F_1} * \varphi_{\sigma^2}(x)}{p_{F_2} * \varphi_{\sigma^2}(x)} \right| \leq O(\varepsilon^3).$$

We then continue with the proof of Proposition 10.

*Proof.* Consider two densities defined in Proposition 11 with the parameter $\theta$. Suppose that $x_1, \dots, x_t$ are i.i.d. samples from either $F_1$ or $F_2$. Then, we note that $a_i \in \mathbb{N}$ for $i \in [t]$ and $y_i \sim \mathcal{N}(x_{a_i}, 1)$ for $i \in [t]$. For simplicity, we write $a^t = (a_1, \dots, a_t)$ and $y^t = (y_1, \dots, y_t)$. Hence, we can write the probability distribution of $(a^t, y^t)$ as follows

$$
\begin{aligned}
p_{\pi,F_1}(a^t, y^t) &= \mathbb{E}_{\{x_i\}_{i=1}^t \sim F_1}\left[\prod_{i=1}^t \left(p_\pi(a_i|a^{i-1}, y^{i-1})\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}(y_i - x_{a_i})^2\right)\right)\right] \\
&= \prod_{i=1}^t p_\pi(a_i|a^{i-1}, y^{i-1}) \prod_{j\in\mathbb{N}} \mathbb{E}_{x_j\sim F_1}\left[\prod_{i\le t, a_i=j}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}(y_i - x_j)^2\right)\right].
\end{aligned}
$$

Let we write $n_j = \sum_{i\le t} \mathbb{1}(a_i = j)$ and $\bar{y}_j = \frac{1}{n_j}\sum_{i\le t, a_i=j} y_i$. Then we can write

$$
\mathbb{E}_{x_j\sim F_1}\left[\prod_{i\le t, a_i=j}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}(y_i - x_j)^2\right)\right] = (p_{F_1} * K_j(\cdot, \bar{y}_j)) \cdot f_j(\{y_i\}_{a_i=j}),
$$

Here we denote

$$
K_j(x, y) = (2\pi)^{-n_j/2}\exp(-n_j(x-y)^2/2)),
$$

and

$$
f_j(\{y_i\}_{a_i=j}) = \exp\left(n_j\bar{y}_j^2/2 - \sum_{i:a_i=j} y_j^2\right).
$$

Therefore, we can write the log-likelihood ratio as

$$
\begin{aligned}
\log\frac{p_{\pi,F_1}(a^t, y^t)}{p_{\pi,F_2}(a^t, y^t)} &= \sum_{j\in\mathbb{N}}\log\frac{\mathbb{E}_{x_j\sim F_1}\left[\prod_{i\le t, a_i=j}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}(y_i - x_j)^2\right)\right]}{\mathbb{E}_{x_j\sim F_2}\left[\prod_{i\le t, a_i=j}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}(y_i - x_j)^2\right)\right]} \\
&= \sum_{j\in\mathbb{N}}\log\frac{\mathbb{E}_{x_j\sim F_1}\left[\exp\left(-\frac{n_j}{2}(\bar{y}_j - x_j)\right)\right]}{\mathbb{E}_{x_j\sim F_2}\left[\exp\left(-\frac{n_j}{2}(\bar{y}_j - x_j)\right)\right]}.
\end{aligned}
$$

Then, we can compute that

$$
\begin{aligned}
\mathrm{D_{KL}}(p_{\pi,F_1}\|p_{\pi,F_2}) &= \int\sum_{a^T} p_{\pi,F_1}(a^T, y^T)\log\frac{p_{\pi,F_1}(a^t, y^t)}{p_{\pi,F_2}(a^t, y^t)}dy^t \\
&= \sum_{j\in\mathbb{N}}\int\sum_{a^T}\prod_{i=1}^t p_\pi(a_i|a^{i-1}, y^{i-1})\prod_{k\in\mathbb{N}}\mathbb{E}_{x_k\sim F_1}\left[\prod_{i:a_i=k}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}(y_i - x_k)^2\right)\right] \\
&\quad\times\log\frac{\mathbb{E}_{x_j\sim F_1}\left[\exp\left(-\frac{n_j}{2}(\bar{y}_j - x_j)\right)\right]}{\mathbb{E}_{x_j\sim F_2}\left[\exp\left(-\frac{n_j}{2}(\bar{y}_j - x_j)\right)\right]}dy^t.
\end{aligned}
$$

For $n_j \le \varepsilon^{\theta-1}$, from Proposition 11, we have

$$
\left|\log\frac{F_1 * \mathcal{N}(0, 1/n_i)}{F_2 * \mathcal{N}(0, 1/n_i)}(\bar{y}_j)\right| \le O(\varepsilon^3).
$$

On the other hand, for $n_j \geq \varepsilon^{\theta-1}$, by utilizing Lemma 22, we note that

$$
\begin{aligned}
&\int \sum_{a^T} \prod_{i=1}^{t} p_\pi(a_i|a^{i-1}, y^{i-1}) \prod_{k\in\mathbb{N}} \mathbb{E}_{x_k\sim F_1}\left[\prod_{i:a_i=k} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(y_i - x_k)^2\right)\right] \\
&\times \log \frac{\mathbb{E}_{x_j\sim F_1}\left[\exp\left(-\frac{n_j}{2}(\bar{y}_j - x_j)\right)\right]}{\mathbb{E}_{x_j\sim F_2}\left[\exp\left(-\frac{n_j}{2}(\bar{y}_j - x_j)\right)\right]} dy^t \\
=& \iint \sum_{a^T} \prod_{i=1}^{t} p_\pi(a_i|a^{i-1}, y^{i-1}) \prod_{k\neq j} \mathbb{E}_{x_k\sim F_1}\left[\prod_{i:a_i=k} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(y_i - x_k)^2\right)\right] \\
&\times f(\{y_i\}_{a_i=j})[p_{F_1} * K_j(\cdot, \bar{y}_j)](x_j) \log \frac{[p_{F_1} * K_j(\cdot, \bar{y}_j)](x_j)}{[p_{F_2} * K_j(\cdot, \bar{y}_j)](x_j)} dx_j dy^t \\
\leq& \iint \sum_{a^T} \prod_{i=1}^{t} p_\pi(a_i|a^{i-1}, y^{i-1}) \prod_{k\neq j} \mathbb{E}_{x_k\sim F_1}\left[\prod_{i:a_i=k} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(y_i - x_k)^2\right)\right] \\
&\times f_j(\{y_i\}_{a_i=j}) p_{F_1}(x_j) \log \frac{p_{F_1}(x_j)}{p_{F_2}(x_j)} K_j(x_j, \bar{y}_j) dx_j dy^t \\
=& \iint p_{F_1}(x_j) \log \frac{p_{F_1}(x_j)}{p_{F_2}(x_j)} \sum_{a^T} \prod_{i=1}^{t} p_\pi(a_i|a^{i-1}, y^{i-1}) \\
&\times \int \prod_{k\neq j} \mathbb{E}_{x_k\sim F_1}\left[\prod_{i:a_i=k} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(y_i - x_k)^2\right)\right] \\
&\times \left[\prod_{i:a_i=j} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(y_i - x_j)^2\right)\right] dx_j dy^t \\
=& \int p_{F_1}(x_j) \log \frac{p_{F_1}(x_j)}{p_{F_2}(x_j)} dx_j = O(\varepsilon^{1.5}).
\end{aligned}
$$

Therefore, we have

$$
\mathrm{D}_{\mathrm{KL}}(p_{\pi,F_1} \| p_{\pi,F_2}) \leq \sum_{i\in\mathbb{N}} \left(O(\varepsilon^{1.5})\mathbb{I}(n_i \geq \varepsilon^{\theta-1}) + O(\varepsilon^3)\mathbb{I}(n_i \leq \varepsilon^{\theta-1})\right).
$$

Note that

$$
t = \sum_{i\in\mathbb{N}} n_i \geq \varepsilon^{\theta-1} \sum_{i\in\mathbb{N}} \mathbb{I}(n_i \geq \varepsilon^{\theta-1}),
$$

and

$$
t \geq \sum_{i\in\mathbb{N}} \mathbb{I}(n_i \leq \varepsilon^{\theta-1}).
$$

The above inequalities imply that

$$
\mathrm{D}_{\mathrm{KL}}(p_{\pi,F_1} \| p_{\pi,F_2}) \leq O(\varepsilon^{2.5-\theta} t).
$$

This completes the proof. □

### F.4 Proof of Lemma 22

*Proof.* We note that the function $f(x) = x \log x$ is strictly convex. Suppose that $x \in \mathbb{R}$. We note that

$$
\int \frac{q(y)K(x-y)}{(K*q)(x)} dy = 1.
$$

By the Jensen's inequality, we have

$$
\int f\left(\frac{p(y)}{q(y)}\right) \frac{q(y)K(x-y)}{(K*q)(x)} dy \geq f\left(\int \frac{p(y)}{q(y)} \frac{q(y)K(x-y)}{(K*q)(x)} dy\right).
$$

This implies that

$$\frac{\left(K * \left(p \log \frac{p}{q}\right)\right)(x)}{(K * q)(x)} \geq \frac{(p * K)(x)}{(q * K)(x)} \log \frac{(p * K)(x)}{(q * K)(x)}.$$

This completes the proof.

□

# G Proof of Lower Bounds for trimmed mean estimation

## G.1 Proof of Lemma 6

Firstly, for $\sigma \leq C\varepsilon^{1/2}$, we consider two Gaussian distribution $F_1$ as the CDF of $N(0,1)$ and $F_2$ as the CDF of $N(3\varepsilon, 1)$. Then, $g(F_2) - g(F_1) = 3\varepsilon$ and $\mathrm{D}_{\mathrm{KL}}(F_1 \| F_2) = \frac{9\varepsilon^2}{2}$. From the data-processing inequality, we have

$$\mathrm{D}_{\mathrm{KL}}(F_1 * \mathcal{N}(0,\sigma^2) \| F_2 * \mathcal{N}(0,\sigma^2)) \leq \mathrm{D}_{\mathrm{KL}}(F_1 \| F_2) = O(\varepsilon^2).$$

To show the lower bound $\Omega(\varepsilon^2)$, without loss of generality we assume that $\int_{F_1^{-1}(\alpha)}^{F_1^{-1}(1-\alpha)} x dF_1(x) = 0$ and $\int_{F_2^{-1}(\alpha)}^{F_2^{-1}(1-\alpha)} x dF_2(x) \leq -\varepsilon$. Lemma 9 and the density lower bound in Assumption 4 show that

$$\begin{aligned}
\int_{(F_1*\varphi_{\sigma^2})^{-1}(\alpha)}^{(F_1*\varphi_{\sigma^2})^{-1}(1-\alpha)} x dF_1 * \varphi_{\sigma^2}(x) &\geq \int_{F_1^{-1}(\alpha)}^{F_1^{-1}(1-\alpha)} x dF_1(x) - \tilde{C}\sigma^2 = -\tilde{C}\sigma^2, \\
\int_{(F_2*\varphi_{\sigma^2})^{-1}(\alpha)}^{(F_2*\varphi_{\sigma^2})^{-1}(1-\alpha)} x dF_2 * \varphi_{\sigma^2}(x) &\leq \int_{F_2^{-1}(\alpha)}^{F_1^{-1}(1-\alpha)} x dF_2(x) + \tilde{C}\sigma^2 \leq -\varepsilon - \tilde{C}\sigma^2,
\end{aligned}$$

Here $\tilde{C} > 0$ is a constant. Consequently, for $\sigma \leq c\varepsilon^{1/2}$ with a small constant $c > 0$, it holds that

$$\int_{(F_1*\varphi_{\sigma^2})^{-1}(\alpha)}^{(F_1*\varphi_{\sigma^2})^{-1}(1-\alpha)} x dF_1 * \varphi_{\sigma^2}(x) - \int_{(F_2*\varphi_{\sigma^2})^{-1}(\alpha)}^{(F_2*\varphi_{\sigma^2})^{-1}(1-\alpha)} x dF_2 * \varphi_{\sigma^2}(x) = \Omega(\varepsilon).$$

From the algorithm for trimmed mean estimation, we can distinguish $F_1 * \varphi_{\sigma^2}$ and $F_2 * \varphi_{\sigma^2}$ using $O(\varepsilon^{-2})$ samples. This implies that $\mathrm{D}_{\mathrm{KL}}(F_1 * \mathcal{N}(0,\sigma^2) \| F_2 * \mathcal{N}(0,\sigma^2)) \geq \Omega(\varepsilon^2)$.

For $\sigma \geq \varepsilon^{1-\theta}$, let $F_1$ be uniform on $[1,2]$. Without the loss of generality, we may assume that $\sigma^2 \geq 4$, $c_1 \leq 0.5$, $c_3 \geq 2$, $c_4 \geq 2$, $c_5 \leq 1$. Then, $F_1$ satisfies Assumption 4. By taking $k \geq \frac{2\kappa}{\theta}$, we can construct $h(x)$ satisfying the conditions in Lemma 21 with $\varepsilon_1 = 4b'\varepsilon$, where $b' = \max\{\frac{b}{c_2}, 4\}$. Consider the distribution $F_2$ with pdf $p_{F_2}(x) = p_{F_1}(x) + (b')^{-1}h(x-1-\alpha)$. By choosing $\varepsilon$ sufficiently small such that $\sqrt{\varepsilon_1} \leq \min\{\alpha, 1-2\alpha\}$, the density function $p_{F_2}$ is supported in $[1,2]$ and $F_2^{-1}(1-\alpha) = 2-\alpha$.

As $|h(x)| \leq b\sqrt{\varepsilon_1}$, for $\varepsilon_1 \leq \frac{1}{4c_2^2}$, we have $(b')^{-1}h(x) \leq \frac{c_2}{b}h(x) \leq c_2\sqrt{\varepsilon} \leq \frac{1}{2}$. This implies that $p_{F_2}(x) \in [1/2, 3/2]$ for $x \in [1,2]$. Therefore, $p_{F_2}$ is a density function. Because $p'_{F_1}(x) = 0$ for $x \in [1,2]$, $p_{F_2}(x)$ is $c_2$-Lipschitz continuous in $[1,2]$.

Note that $F_2(1+\alpha) = F_1(1+\alpha) + (b')^{-1}\int_{-\infty}^{0} h(x)dx = \alpha - 4\varepsilon$. Hence, it follows that

$$4\varepsilon = F_2(F_2^{-1}(\alpha)) - F_2(1+\alpha) \leq 3/2(F_2^{-1}(\alpha) - 1 - \alpha).$$

This implies that $F_2^{-1}(\alpha) \geq 1 + \alpha + \frac{2}{3} \cdot 4\varepsilon \geq 1 + \alpha + 2\varepsilon$. Therefore,

$$\begin{aligned}
\int_{F_2^{-1}(\alpha)}^{F_2^{-1}(1-\alpha)} x dF_2(x) &\leq \int_{1+\alpha+\varepsilon}^{2-\alpha} x dF_2(x) = 2 - 2\alpha - 2\varepsilon - \int_{\varepsilon}^{\sqrt{\varepsilon_1}} (b')^{-1}h(x)dx \\
&\leq 2 - 2\alpha - 2\varepsilon + 2\varepsilon\sqrt{(b')^{-1}} \leq 2 - 2\alpha - \varepsilon.
\end{aligned}$$

Here we utilize that $b' \geq 4$. Note that $\int_{F_1^{-1}(\alpha)}^{F^{-1}(1-\alpha)} x dF_1(x) = 2 - 2\alpha$. This implies that

$$g(F_2) - g(F_1) \leq -\varepsilon.$$

By choosing $\varepsilon$ sufficiently small, we have $\varepsilon_1 \le \varepsilon^{1-\theta/2}$. From Lemma 20 and the property of $h(x)$, we immediately have

$$\mathrm{D_{KL}}(F_1 * \mathcal{N}(0,\sigma^2) \| F_2 * \mathcal{N}(0,\sigma^2)) \le \max_{x\in\mathbb{R}} \frac{\log(p_{F_1} * \varphi_{\sigma^2}(x))}{\log(p_{F_2} * \varphi_{\sigma^2}(x))}$$
$$\le O(1)\cdot \frac{b\sqrt{\varepsilon_1}\sqrt{\varepsilon_1}^{2k+2}}{\sigma^{2k+2}} = O(\varepsilon^{(2k+1)\theta/2}) = O(\varepsilon^{\kappa}).$$

This completes the proof.

### G.2 Proof of Lemma 7

We construct the same pair of $F_1$ and $F_2$ as in Lemma 6. It is sufficient to prove that when $\sigma \le c\varepsilon^{1/2}$,

$$\mathrm{D_{KL}}(F_2 * \mathcal{N}(0,\sigma^2) \| F_1 * \mathcal{N}(0,\sigma^2)) = \Theta(\varepsilon^{1.5}).$$

For the upper bound, we simply use the data-processing inequality:

$$\begin{aligned}\mathrm{D_{KL}}(F_2 * \mathcal{N}(0,\sigma^2) \| F_1 * \mathcal{N}(0,\sigma^2)) &\le \mathrm{D_{KL}}(F_2\|F_1) \le \chi^2(F_1\|F_2)\\ &= \int_1^2 \frac{(p_{F_1}(x) - p_{F_2}(x))^2}{p_{F_1}(x)}dx\\ &\le (b')^2 \int_{-\sqrt{\varepsilon_1}}^{\sqrt{\varepsilon_1}} h(x)^2 dx\\ &\le (b')^2 \cdot \int_{-\sqrt{\varepsilon_1}}^{\sqrt{\varepsilon_1}} (b\sqrt{\varepsilon_1})^2 dx = O(\varepsilon^{1.5}).\end{aligned}$$

For the lower bound, we note that $F_2(1+\alpha) = \alpha - 4\varepsilon = F_1(1+\alpha) - 4\varepsilon$. Similar to the proof of Lemma 5, we have

$$F_1 * \mathcal{N}(0,\sigma^2)(1+\alpha) - F_2 * \mathcal{N}(0,\sigma^2)(1+\alpha) = \Omega(\varepsilon).$$

Analogously, we can derive the same lower bound

$$\mathrm{D_{KL}}(F_2 * \mathcal{N}(0,\sigma^2) \| F_1 * \mathcal{N}(0,\sigma^2)) \ge \Omega(\varepsilon^{3/2}).$$

### G.3 Proof of Theorem 4

By applying the Le Cam's two point lower bound, it is sufficient to show that the following proposition holds.

**Proposition 12.** *Suppose that $\varepsilon > 0$. Denote $\mathcal{F}$ as the set the set of distributions satisfying Assumption 4. Consider an online algorithm $\pi$ with a fixed budget $t$ which outputs $\hat{G}$. Given the distribution with CDF $F \in \mathcal{F}$ of the underlying arms and the algorithm $\pi$, let $p_{\pi,F}(\{(a_i, y_i)\}_{i=1}^t)$ denote the distribution of the action-observation pairs up to the $t$-th iteration. Then, for any $\theta \in (0, 1/4)$, there exist $F_1, F_2 \in \mathcal{F}$ with trimmed means $g(F_1)$ and $g(F_2)$ such that $|g(F_1) - g(F_2)| \ge \varepsilon$ and*

$$\mathrm{D_{KL}}(p_{\pi,F_1} \| p_{\pi,F_2}) \le O(\varepsilon^{2.5-2\theta} t).$$

Similar to the proof of Proposition 10, we start with the following proposition.

**Proposition 13.** *Suppose that $\theta > 0$ is a given constant. Let $\mathcal{F}$ denote the set of distributions satisfying Assumption 4. Then, there exists two distribution $F_1, F_2 \in \mathcal{F}$ with trimmed mean $g(F_1)$ and $g(F_2)$ such that $g(F_2) \le g(F_1) - \varepsilon$ and they satisfy that*

$$\mathrm{D_{KL}}(F_1\|F_2) = O(\varepsilon^{1.5-\theta}),$$

*Denote $\varphi_{\sigma^2}$ as the pdf of $\mathcal{N}(0,\sigma^2)$. For sufficiently small $\varepsilon$ and for $\sigma$ satisfying $\sigma^2 \ge \varepsilon^{1-\theta}$, we further have*

$$\left|\log \frac{p_{F_1} * \varphi_{\sigma^2}(x)}{p_{F_2} * \varphi_{\sigma^2}(x)}\right| \le O(\varepsilon^3).$$

Then, we present the proof of Proposition 12.

*Proof.* Consider $F_1$ and $F_2$ as distributions constructed in Proposition 13. Based on Proposition 13 and Lemma 22, analogously, we have

$$\mathrm{D}_{\mathrm{KL}}(p_{\pi,F_1}\|p_{\pi,F_2}) \le \sum_{i\in\mathbb{N}} \left(O(\varepsilon^{1.5-\theta})\mathbb{I}(n_i \ge \varepsilon^{\theta-1}) + O(\varepsilon^3)\mathbb{I}(n_i \le \varepsilon^{\theta-1})\right).$$

Note that

$$t = \sum_{i\in\mathbb{N}} n_i \ge \varepsilon^{\theta-1}\sum_{i\in\mathbb{N}} \mathbb{I}(n_i \ge \varepsilon^{\theta-1}), \quad t \ge \sum_{i\in\mathbb{N}} \mathbb{I}(n_i \le \varepsilon^{\theta-1}).$$

This implies that

$$\mathrm{D}_{\mathrm{KL}}(p_{\pi,F_1}\|p_{\pi,F_2}) \le O(\varepsilon^{2.5-2\theta}t).$$

This completes the proof. □

### G.4 Proof of Proposition 13

Consider two distributions constructed in Lemma 6. It is sufficient to show the bound on KL divergence and the pointwise bound. Firstly, we note that $F_1$ is uniform on $[1,2]$ and the density of $F_2$ only differs from the density of $F_1$ in $[1+\alpha-\sqrt{\varepsilon_1}, 1+\alpha+\sqrt{\varepsilon_1}]$. The difference is upper bounded by $b\sqrt{\varepsilon_1}$. We note that for sufficiently small $\varepsilon$, we have $\varepsilon_1 \le \varepsilon^{1-\theta/2}$. Therefore, we have

$$\begin{aligned}
\mathrm{D}_{\mathrm{KL}}(F_1\|F_2) &= -\int_{1+\alpha-\sqrt{\varepsilon_1}}^{1+\alpha+\sqrt{\varepsilon_1}} p_{F_1}\log\frac{p_{F_2}}{p_{F_1}}dx \\
&= -\int_{1+\alpha-\sqrt{\varepsilon_1}}^{1+\alpha+\sqrt{\varepsilon_1}} p_{F_1}\left(\left(\frac{p_{F_2}}{p_{F_1}}-1\right) - \frac{1}{2}\left(\frac{p_{F_2}}{p_{F_1}}-1\right)^2 + O(\varepsilon_1^{3/2})\right)dx \\
&= \frac{1}{2}\int_{1+\alpha-\sqrt{\varepsilon_1}}^{1+\alpha+\sqrt{\varepsilon_1}} \left(\frac{p_{F_2}}{p_{F_1}}-1\right)^2 dx + O(\varepsilon^2) \\
&= O(\varepsilon_1^{\frac{3}{2}}) = O(\varepsilon^{\frac{3}{2}(1-\theta/2)}) = O(\varepsilon^{1.5-\theta}).
\end{aligned}$$

We note that $F_1$ and $F_2$ have $2k$ matched moments, where $k > 6/\theta$. Analogous to the result in Lemma 20, for all $x \in \mathbb{R}$,

$$\left|\log\frac{p_{F_1} * \varphi_{\sigma^2}(x)}{p_{F_2} * \varphi_{\sigma^2}(x)}\right| \le C\frac{b\sqrt{\varepsilon_1}\sqrt{\varepsilon_1}^{2k+2}}{\sigma^{2k+2}} \le C\cdot\frac{b\varepsilon^{(k+1)(1-\theta/2)}}{\varepsilon^{(k+1)(1-\theta)}} = O(\varepsilon^{(k+1)\theta/2}) \le O(\varepsilon^3).$$

Here $C > 0$ is an absolute constant. This completes the proof.