Mean Testing under Truncation beyond Gaussian

Yuhao Wang
Amazon
Tokyo, Japan
wanyuhao@amazon.co.jp

Roberto Imbuzeiro Oliveira
IMPA
Rio de Janeiro, RJ, Brazil 22240-006
rimfo@impa.br

Themis Gouleakis
Nanyang Technological University
Singapore
themis.gouleakis@ntu.edu.sg

Abstract

We characterize the fundamental limits of high-dimensional mean testing under arbitrary truncation, where samples are drawn from the conditional distribution $P(\cdot \mid S)$ for an unknown truncation set S that may hide up to an ε-fraction of the probability mass. For distributions with p-th directional moments of magnitude at most $\nu_{P,p}$, truncation induces a bias of order $\mathcal{O}(\nu_{P,p}\varepsilon^{1-1/p})$. This bias creates a sharp information-theoretic detectability floor: when the signal α falls below this threshold, the null and alternative hypotheses are indistinguishable even with infinite data. Above this floor, we prove that a simple second-order test achieving near-optimal sample complexity

$$n = \mathcal{O}\left(\frac{\|\Sigma_P\|}{(\alpha - 4\nu_{P,p}\varepsilon^{1-1/p})^2}\sqrt{d}\right).$$

We further identify a structural escape from this finite-moment bias barrier. Under a *directional median regularity* assumption, truncation bias improves to linear order $\mathcal{O}(\varepsilon)$. This reveals an intermediate regime in which estimation requires $\Theta(d)$ samples for uniform recovery, while testing recovers the classical $\Theta(\sqrt{d})$ rate once truncation bias is eliminated. Together, our results provide a unified framework for mean testing under truncation, connecting finite-moment, sub-Gaussian, and median-regular structural regimes.

1 Introduction

Robust mean estimation has been extensively studied in statistics and machine learning, with a rich literature on contamination and heavy-tailed noise [15, 33, 16, 17, 23]. By contrast, the problem of *mean testing* under arbitrary truncation, where samples are observed only if they fall in an unknown set, remains largely unexplored [3].

Truncation arises naturally in settings such as privacy-preserving data collection, controlled sensing, and heavy-tailed environments [20, 30, 37], where observations outside an unknown acceptance region are systematically discarded. Whereas classical contamination models were primarily designed with outliers in mind [27, 24], truncation induces structured missingness that can remove ε-fraction of the probability mass anywhere in the distribution. This difference has sharp consequences for hypothesis testing, which exhibits detectability thresholds that separate irreducible truncation bias from statistical fluctuation. This raises a fundamental question: *to what extent can truncation obscure the mean, and when is reliable testing possible?*

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

Table 1: Phase diagram for high-dimensional mean testing under arbitrary truncation. For each row, regimes are ordered by increasing signal strength α. Under finite p-moment assumptions, truncation induces a detectability floor $b_{\mathrm{mom}}(\varepsilon) = \Theta(\nu_{P,p}\varepsilon^{1-1/p})$, below which testing is information-theoretically impossible. Under structural regularity, this floor improves to $b_{\mathrm{reg}}(\varepsilon) = \tilde{O}(\varepsilon)$, revealing an intermediate *learning-hard* regime in which testing requires $\Theta(d)$ samples before recovering the optimal $\Theta(\sqrt{d})$ rate. *Gray* indicates impossibility, *red* indicates $\Theta(d)$ sample complexity, and *green* indicates $\Theta(\sqrt{d})$ complexity (all up to logarithmic factors).

	Impossible $\alpha \le b_{\mathrm{reg}}(\varepsilon)$	Learning-hard $b_{\mathrm{reg}}(\varepsilon) < \alpha \le b_{\mathrm{mom}}(\varepsilon)$	Near floor $\alpha \gtrsim b_{\mathrm{mom}}(\varepsilon)$	Strong signal $\alpha \gg b_{\mathrm{mom}}(\varepsilon)$
Moment-only (finite p-moments)	∞	∞	$\Theta(\sqrt{d})$	$\Theta(\sqrt{d})$
Structural regularity	∞	$\Theta(d)$	$\Theta(\sqrt{d})$	$\Theta(\sqrt{d})$

We study the hypothesis testing problem

$$H_0 : \mu_P = 0 \quad \text{versus} \quad H_1 : \|\mu_P\| \ge \alpha, \tag{1}$$

given i.i.d. samples from the truncated law $P(\cdot|S)$, where the truncation set $S \subseteq \mathbb{R}^d$ is unknown and may be chosen adversarially, subject only to $P(\cdot|S) \ge 1-\varepsilon$. The main challenge is that truncation can induce a systematic bias in the observed mean, potentially masking the separation between null and alternative hypotheses. Quantifying this bias is therefore essential for characterizing the detectability limits of the testing problem. Our first contribution is a sharp characterization of truncation-induced bias under finite moment assumptions. For distributions with directional p-th moments upper bounded by $\nu_{P,p}$, we show in Lemma 1 that truncation can shift the mean by at most $\gamma = 2\nu_{P,p}\varepsilon^{1-1/p}$, and that this dependence is information-theoretically unavoidable. Building on this result, we design a simple test that achieves the optimal $\sqrt{d}$ dependence with sample complexity $n = \mathcal{O}\Big(\frac{\|\Sigma_P\|}{(\alpha-2\gamma)^2}\sqrt{d}\Big)$, whenever $\alpha > 2\gamma$. Conversely, we prove a matching impossibility result: when the signal lies below this bias floor, the null and alternative hypotheses become indistinguishable, even with infinitely many samples.

Informal Theorem (Finite-Moment Barrier). *Consider testing $H_0 : \mu_P = 0$ versus $H_1 : \|\mu_P\| \ge \alpha$ from ε-truncated samples of a distribution with finite directional p-th moments upper bounded by $\nu_{P,p}$. There exists a critical scale $\alpha^\star = \Theta(\nu_{P,p}\varepsilon^{1-1/p})$ such that: (i) if $\alpha \ll \alpha^\star$, the hypotheses are information-theoretically indistinguishable, regardless of sample size; (ii) if $\alpha \gg \alpha^\star$, reliable testing is possible with optimal $\sqrt{d}$ dependence in sample complexity (see Theorem 3 in section 4.5).*

Beyond this moment-limited regime, we identify a structural condition that allows one to escape the truncation bias barrier. Under an additional *directional median regularity* assumption — requiring each one-dimensional projection to be median-centered with non-vanishing local density — the truncation bias improves to linear order $\mathcal{O}(\varepsilon)$. This improvement is not a consequence of stronger tail concentration, but of local identifiability of the center in every direction, which limits how much probability mass can be removed near the center without altering the observed distribution. In particular, Gaussian distribution satisfy this condition, and the near-linear truncation bias established in [3] arises as a special case of our analysis. As a result, stable center estimation and reliable testing become possible even for heavy-tailed distributions.

Informal Theorem (Structural Escape). *If the distribution P satisfies directional median regularity (Assumption 1), then truncation induces at most a linear $\mathcal{O}(\varepsilon)$ bias in the center. Consequently, reliable mean testing is possible using $\Theta(d)$ samples whenever $\alpha \gtrsim \varepsilon$, matching the bias scaling of the sub-Gaussian regime even for heavy-tailed distributions (see Theorem 4 in section 5).*

Together, these results provide a unified framework for mean testing under truncation, interpolating between the sub-Gaussian regime, the finite-moment impossibility barrier, and the median-regular setting where robustness is restored. As summarized in Table 1, the $\sqrt{d}$ dependence characteristic of high-dimensional testing is preserved across all feasible regimes, but the robustness (bias scale) varies fundamentally with the structural assumptions.

Our work positions truncation as a distinct and fundamental obstacle in high-dimensional testing, whose difficulty is governed by truncation-induced bias rather than statistical variance. By precisely characterizing detectability limits under finite-moment assumptions and identifying structural regularity conditions that eliminate this bias, we bridge classical robust statistics with modern distribution property testing, and provide principled testing procedures for truncated data.

Remark 1 (Technical contribution)**.** Our proofs use classical tools, including Hölder's inequality, the Efron–Stein variance argument for U-statistics, and uniform concentration over halfspaces. Our contribution is the characterization of detectability under arbitrary unknown truncation beyond Gaussian distributions: a moment-dependent bias bound, an indistinguishable null–alternative construction showing that its scale is unavoidable, and a testing guarantee above this threshold. We further identify directional median regularity as a sufficient condition for stable recovery of the original mean under truncation without requiring finite variance. The Gaussian testing lower bound used in Theorem 5 is taken from [3].

2 Related Work

The study of statistical inference under truncation has a long history, often distinguishing between settings where the truncation mechanism is known and those where it is unknown or adversarial. Our work bridges learning from truncated data and modern high-dimensional hypothesis testing, focusing on detectability limits rather than estimation accuracy. Additional related work is in Appendix A.

Mean testing under truncation and detecting truncation. Most of the above literature focuses on learning/estimation, whereas our focus is hypothesis testing. Closest in spirit, [3] studies Gaussian mean testing under *arbitrary* truncation, identifying sharp limits and a near-linear $\mathcal{O}(\varepsilon\sqrt{\log(1/\varepsilon)})$ truncation bias barrier in the Gaussian/sub-Gaussian regime. Our work generalizes the testing perspective beyond Gaussian assumptions and shows that for general distributions with finite moments, truncation induces a polynomial bias barrier of order $\mathcal{O}(\varepsilon^{1-1/p})$, leading to an information-theoretic detectability floor. Orthogonally, [12, 11] consider problems of *detecting* whether samples have been truncated under structural promises on the truncation set or mechanism; these are complementary to our goal, which assumes truncation may be present and seeks to test the underlying mean.

Heavy-tailed mean estimation and testing. There is a rich literature on robust mean estimation under finite-moment or finite-covariance assumptions, including algorithmic approaches achieving sub-Gaussian-type performance and sharp uniform guarantees [26, 23, 34, 25]. Recent work also highlights qualitative separations between heavy-tailed estimation and adversarial contamination [5]. Our results identify a testing-specific "moment barrier" under truncation: even when accurate estimation is possible (in the absence of truncation) under heavy tails, truncation induces a systematic bias of order $\varepsilon^{1-1/p}$ that renders mean testing information-theoretically impossible below the corresponding detectability floor unless additional structural regularity is imposed.

Geometric and median-based regularity. Median-based stability and geometric regularity conditions are widely used in robust statistics to overcome moment limitations (see, e.g., surveys and references in [33, 17]). Our *directional median regularity* assumption is motivated by this line of work; it imposes a local density near the center in all directions, thereby ruling out adversarial removal of probability mass near the center. We show that such structural regularity improves the truncation-induced bias from $\mathcal{O}(\varepsilon^{1-1/p})$ to $\mathcal{O}(\varepsilon)$ and reveals an intermediate regime where learning remains hard while testing becomes tractable.

3 Notations and Preliminaries

This section introduces the mathematical framework and distributional quantities required to analyze mean testing under truncation.

Notations We consider families of probability distributions $\mathcal{P}$ over $\mathbb{R}^d$. For any $P \in \mathcal{P}$, $\mu_P := \mathbb{E}_{X\sim P}[X]$ denotes the mean vector and $\Sigma_P := \mathrm{Cov}_{X\sim P}(X)$ denotes the covariance matrix. We denote both the Euclidean norm of vectors and the spectral (operator) norm of a matrix by $\|\cdot\|$. Samples $X_1, \ldots, X_n$ are assumed to be independently and identically distributed (i.i.d.). For a real

parameter p and threshold κ, we write $p \uparrow \kappa$ to denote the limit as p increases to κ from below. Accordingly, statements of the form "$\nu_{P,p}$ diverges as $p \uparrow \kappa$" mean that $\nu_{P,p} < \infty$ for all $p < \kappa$ and $\nu_{P,p} = \infty$ for $p \geq \kappa$.

The Truncation Model Let $\varepsilon \in [0, 1/2]$ for all $P \in \mathcal{P}$. We say that samples are subject to an ε-truncation if there exits a measurable set $S \in \mathbb{R}^d$ satisfying $P(S) \geq 1 - \varepsilon$. We define the *truncated law* $P_S := P(\cdot \mid S)$, representing the conditional distribution of P restricted to S. Throughout the paper, we observe i.i.d. samples $X_1, \ldots, X_n \sim P_S$, while the goal is to test

$$H_0 : \mu_P = 0 \qquad \text{versus} \qquad H_1 : \|\mu_P\| \geq \alpha,$$

for some separation parameter $\alpha > 0$.

Directional Moments and Tail Geometry Since truncation may remove mass along arbitrary direction in $\mathbb{R}^d$, we characterize tail behavior via worst-case one-dimensional projection.

Definition 1 (Directional p-th Moment)**.** For $p \geq 1$, the directional p-th moment bound of a distribution P on $\mathbb{R}^d$ is defined as:

$$\nu_{P,p} := \sup_{v \in \mathbb{R}^d : \|v\| = 1} \left(\mathbb{E}_{X \sim P} |\langle v, X - \mu_P \rangle|^p\right)^{1/p}, \tag{2}$$

that is, the worst-case p-th moment over all one-dimensional projections.

The growth of $\nu_{P,p}$ as a function of p characterizes the tail geometry of P. For sub-Gaussian distributions with parameter σ, there exists a constant $C > 0$ such that $\nu_{P,p} \leq C\sigma\sqrt{p}$ for all $p \geq 1$. In contrast, for heavy-tailed distributions with polynomially decaying tails, directional moments typically exist only up to some finite order κ: the quantity $\nu_{P,p}$ remains bounded for $p < \kappa$ but diverges as $p \uparrow \kappa$.

Appendix B discusses the behavior of $\nu_{P,p}$ for sub-Gaussian and heavy-tailed distributions.

Fundamental Limits of Truncation Bias The following lemma establishes the technical foundation for our testing results by quantifying how much an unknown truncation set S can shift the population parameters.

Lemma 1 (Truncated vs. Non-truncated Parameters)**.** *Let P be a distribution on $\mathbb{R}^d$ and S be a measurable set with $P(S) \geq 1 - \varepsilon$ for $\varepsilon \in [0, 1/2]$. Then for any $p \geq 1$:*

1. ***Mean Shift:*** $\|\mu_P - \mu_{P_S}\| \leq 2\nu_{P,p}\varepsilon^{1-1/p}$.
2. ***Covariance Bound:*** $\|\Sigma_{P_S}\| \leq (1 - \varepsilon)^{-2}\|\Sigma_P\| \leq 4\|\Sigma_P\|$.

The $\mathcal{O}(\varepsilon^{1-1/p})$ shift in the mean represents the systematic bias that our test must overcome. Moreover, if P is sub-Gaussian so that $\nu_{P,p} \lesssim \sigma\sqrt{p}$ for all p, optimizing the bound over p yields the nearly linear truncation bias $\mathcal{O}(\varepsilon\sqrt{\log(1/\varepsilon)})$.

4 Mean Testing under Finite-Moment Truncation

This section investigates the fundamental limits of high-dimensional mean testing when samples are subject to arbitrary truncation and the underlying distribution P possesses only finite directional moments. Our analysis demonstrates that truncation induces a systematic detectability floor that is information-theoretically unavoidable under moment-only constraints.

4.1 Effective Signal and the Detectability Floor

The central challenge in mean testing under truncation is that, the learner observes samples from the conditional law P_S rather than the target distribution P. As shown in Lemma 1, for distributions with finite directional p-th moments, a worst-case truncation set S can induce a systematic shift of the population mean of order $\gamma := 2\nu_{P,p}\varepsilon^{1-1/p}$. This motivates the definition of the *effective signal strength*

$$\beta := \max(0, \alpha - \gamma). \tag{3}$$

which quantifies the residual separation between the alternative hypothesis under truncation and the mean 0 case (i.e null hypothesis without truncation). Under the alternative hypothesis H_1 for distribution P, the observed mean of distribution P_S satisfies $\|\mu_{P_S}\| \geq \beta$, whereas under the null hypothesis H_0, truncation alone may induce an observed mean of magnitude $\|\mu_{P_S}\| \leq \gamma$. Consequently, testing performance is governed by whether the effective signal β is above this threshold (i.e $\beta > \gamma$).

This identifies a truncation-induced detectability floor at scale $b(\varepsilon) = \Theta(\gamma)$. Indeed, by Lemma 1, truncation can shift the population mean by at most $b(\varepsilon)$ in an arbitrary direction. Under the null hypothesis H_0, this implies $\|\mu_{P_S}\| \leq b(\varepsilon)$, while under the alternative H_1 we have the deterministic lower bound $\|\mu_{P_S}\| \geq \|\mu_P\| - b(\varepsilon) \geq \beta$ by triangle inequality. When $\alpha \leq c\, b(\varepsilon)$ for some universal constant $c > 0$, the null and alternative become information-theoretically indistinguishable even with infinite samples (Theorem 3), whereas for sufficiently large effective signal β reliable testing is possible.

By centering the analysis around the parameter $\beta - \gamma = \alpha - 2\gamma$, which is required to be positive, we separate systematic bias due to truncation from the statistical fluctuations that determine variance and finite-sample behavior in the subsequent sections.

4.2 A Natural Second-Order Test Statistic

To distinguish between the null hypothesis $H_0 : \|\mu_P\| = 0$ and alternative hypothesis $H_1 : \|\mu_P\| \geq \alpha$, we consider the order-2 U-statistic:

$$F := \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} \langle X_i, X_j \rangle, \tag{4}$$

where $X_1, \ldots, X_n \sim P_S$ are i.i.d.

This statistic admits a simple interpretation. For any distribution Q on $\mathbb{R}^d$ with finite second moments,

$$\mathbb{E}_{X,X' \sim Q}[\langle X, X' \rangle] = \|\mu_Q\|^2. \tag{5}$$

Consequently, F is an unbiased estimator of $\|\mu_Q\|^2$ when $X_i \sim Q$. In contrast to the plug-in estimator $\left\|\frac{1}{n}\sum_{i=1}^n X_i\right\|^2$, whose expectation incurs an $\mathcal{O}(d/n)$ bias that dominates the signal in high dimensions, the U-statistic F avoids this bias and isolates the second-order signal $\|\mu_Q\|^2$. This property is essential for achieving meaningful testing guarantees when d is large.

Under truncation, the observed samples follow $Q = P_S$, and hence $\mathbb{E}[F] = \|\mu_{P_S}\|^2$. In view of the effective signal definition (3), this implies

$$\mathbb{E}[F] \geq \beta^2 \quad \text{under } H_1, \qquad \mathbb{E}[F] \leq \gamma^2 \quad \text{under } H_0.$$

Therefore, a necessary condition for the hypothesis testing to succeed is the effective signal strength β to be at least γ and thus $\alpha > 2\gamma$.

We emphasize that the specific choice of the statistic (4) is not crucial: any second-order statistic estimating $\|\mu_{P_S}\|^2$ would lead to the same detectability behavior. The fundamental difficulty of the problem arises from truncation-induced bias rather than the choice of estimator; statistical fluctuations and confidence guarantees are analyzed in the subsequent sections.

4.3 Variance Control under Finite Moments

We now analyze the variance of the test statistic F defined in (4) under the truncated law P_S. In contrast to the systematic bias induced by truncation, we show that truncation does not fundamentally worsen the variance of second-order statistics, provided the underlying distribution has finite second moments. The following lemma bounds the variance of the order-2 U-statistic in (4).

Lemma 2 (Variance Bound for U-statistic)**.** *Fix a distribution Q and let $X_1, \ldots, X_n \sim Q$ be i.i.d. Then the variance of the U-statistic $F := \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} \langle X_i, X_j \rangle$ satisfies*

$$\mathrm{Var}_Q(F) \;\leq\; \frac{4}{n} \langle \mu_Q, \Sigma_Q \mu_Q \rangle \;+\; \frac{4}{n(n-1)}\, \mathrm{tr}\big(\Sigma_Q^2\big). \tag{6}$$

The proof follows from a standard Efron–Stein inequality for U-statistics. Importantly, the bound (6) depends only on second-order moments of the truncated distribution.

We apply Lemma 2 for $Q = P_S$, which denotes sampling from the truncated distribution P with truncation set S and relate this variance bound to the parameters of the original distribution P as follows:

$$\mathrm{Var}_{P_S}(F) \lesssim \frac{\|\Sigma_P\|\,\|\mu_{P_S}\|^2}{n} + \frac{d\,\|\Sigma_{P_S}\|^2}{n^2} \lesssim \frac{\|\Sigma_P\|\,\|\mu_{P_S}\|^2}{n} + \frac{d\,\|\Sigma_P\|^2}{n^2}, \tag{7}$$

where the last inequality follows by invoking Lemma 1, which yields $\|\Sigma_{P_S}\| \le 4\|\Sigma_P\|$, as well as the bound $\mathrm{tr}\big(\Sigma_{P_S}^2\big) \le d\,\|\Sigma_{P_S}\|^2$.

This variance bound highlights a key contrast between bias and variance under truncation. While truncation can induce a mean shift of order $\varepsilon^{1-1/p}$, it does not affect the variance upper bound. The requirement $p \ge 2$ is essential here, as it guarantees the existence of second moments needed to control (6); when $p < 2$, even the variance of F may be ill-defined.

A detailed proof of Lemma 2 is deferred to Appendix C.

4.4 Achievable Testing Rates

The variance bound for the U-statistic F implies that a simple test achieves a constant probability of success, once the sample size is large enough.

Theorem 1 (Mean Testing under truncation - constant error prob.)**.** *Let $S \subset \mathbb{R}^d$ satisfy $P(S) \ge 1-\varepsilon$ with $\varepsilon \in [0, 1/2]$, and let $X_1, \dots, X_n \overset{iid}{\sim} P_S := P(\cdot \mid S)$. Let $\gamma = 2\,\nu_{P,p}\,\varepsilon^{1-1/p}$ and $\alpha > 2\gamma$. Consider F as above and the test $T = \mathbf{1}\Big\{ F > \frac{\alpha^2}{4} \Big\}$. There exists a universal constant $C > 0$ such that, if*

$$n \ge \max\left\{ 2, C\,\frac{\|\Sigma_P\|\sqrt{d}}{(\alpha - 2\gamma)^2} \right\},$$

then T separates $\|\mu_P\| = 0$ from $\|\mu_P\| \ge \alpha$ with probability at least $2/3$.

The proof of this result uses Chebyshev's inequality to bound the fluctuations of F around $\|\mu_{P_S}\|^2$. Then the standard deviation and the norm of μ_{P_S} are compared under both H_0 and H_1 to obtain the result.

From Theorem 1, one can easily boost the separation probability to any value $1-\delta$ by performing the test T several times over independent samples. This is the content of the next result.

Theorem 2 (Mean Testing under Truncation)**.** *In the same setting as Theorem 1, given $\delta \in (0, 1/3)$, there exists a test T that distinguishes $H_0 : \mu_P = 0$ from $H_1 : \|\mu_P\| \ge \alpha$ with probability at least $1-\delta$ whenever $\alpha > 2\gamma$ and the sample size satisfies*

$$n \ge C\left\{ 2, \frac{\|\Sigma_P\|}{(\alpha - 2\gamma)^2}\sqrt{d} \right\} \log(1/\delta), \tag{8}$$

where $C > 0$ is a universal constant.

The proofs of these results are provided in Appendix C.

Uniformity and interpretation of the rates For fixed $p \ge 2$ and $\nu > 0$, Theorem 2 yields a uniform guarantee over distributions satisfying $\nu_{P,p} \le \nu$, by substituting $\gamma = 2\nu\varepsilon^{1-1/p}$ and using $\|\Sigma_P\| \le \nu^2$ in its conditions. Under the resulting signal and sample-size conditions, the majority-vote test, using threshold $\alpha^2/4$ in each block, has worst-case error at most δ over both hypotheses and all admissible distribution–truncation pairs. The sample-complexity upper bound in (8) scales as $\sqrt{d}$ when the other parameters are fixed, preserving the dimensional dependence of classical mean testing despite arbitrary unknown truncation. The block statistic and threshold do not require knowledge of p or ν, which enter the sufficient signal and sample-size conditions. For sub-Gaussian distributions with $\nu_{P,p} \lesssim \sigma\sqrt{p}$, choosing $p \asymp \log(1/\varepsilon)$ gives the bias bound $O(\sigma\varepsilon\sqrt{\log(1/\varepsilon)})$. The lower-bound construction establishes that the moment-based detectability threshold is sharp up to constant factors. However, matching finite-sample minimax rates throughout the detectable region are not established here. We defer a more comprehensive minimax treatment of bounded moment distribution classes and classes with common median-regularity parameters (c, r) to a future journal version.

4.5 Information-Theoretic Impossibility

We now show that the effective signal threshold identified in Section 4.1 is unavoidable under finite-moment assumptions. In particular, when the truncation-induced bias dominates the signal, no test can reliably distinguish the null and alternative hypotheses, regardless of sample size.

Theorem 3 (Impossibility under severe truncation)**.** *Fix $p \geq 2$ such that $\nu_{P,p} < \infty$. There exists a universal constant $C > 0$ such that if $\alpha \leq C\, \nu_{P,p} \varepsilon^{1-1/p}$, then there exists a distribution P on $\mathbb{R}^d$ with mean vector μ_P satisfying $\|\mu_P\| \geq \alpha$, a distribution P_0 with $\mu_{P_0} = 0$, and a measurable set $S \subseteq \mathbb{R}^d$ with $P(S) \geq 1 - \varepsilon$ such that*

$$P(\cdot \mid S) = P_0.$$

Consequently, no (possibly randomized) test based on i.i.d. samples from the truncated distribution $P(\cdot \mid S)$ can distinguish the hypotheses $\|\mu_P\| = 0$ versus $\|\mu_P\| \geq \alpha$ with error probability bounded away from $1/2$, even as $n \to \infty$.

Proof. Fix $p \geq 2$. Let δ_x denote the Dirac measure at $x \in \mathbb{R}$ and $P = (1-\varepsilon)\delta_0 + \varepsilon \delta_{\varepsilon^{-1/p}}$ be a distribution supported on $\{\varepsilon^{-1/p}, 0\}$. We have that $\mu_P = \varepsilon^{1-1/p}$ and $\nu_{P,p} \leq 2$. Letting $P_0 = \delta_0$ and $S = (-\infty, 0]$ completes the proof, since $P(\cdot|S) = P_0 = \delta_0$ and $P(S) = 1 - \varepsilon$. □

This result establishes the sharp boundary between the achievable and impossible regimes: when the hidden mass carries as much mean energy as the target signal, truncation obscures it.

4.6 Discussion: The Moment Barrier

Sections 4.1–4.5 establish a sharp detectability threshold for mean testing under truncation. Under finite-moment assumptions alone, truncation induces an unavoidable bias of order $\mathcal{O}(\varepsilon^{1-1/p})$, yielding an information-theoretic barrier that persists even with infinitely many samples.

This phenomenon arises from the local asymmetry permitted by moment constraints. An adversary may remove an ε-fraction of mass in a highly asymmetric manner near the center of the distribution, thereby shifting the mean while leaving the observed truncated distribution unchanged. Because the mean is the minimizer of a global L_2 objective, it is inherently sensitive to such localized mass removal. As a result, truncation can selectively eliminate the portion of the distribution carrying the signal, rendering the null and alternative hypotheses indistinguishable. Overcoming this barrier requires structural assumptions that constrain how probability mass is arranged locally near the center, preventing truncation from completely obscuring the signal.

5 Breaking the Moment Barrier via Median Regularity

Section 4 showed that under moment-only assumptions, truncation induces an unavoidable detectability floor: an ε-fraction removal of mass can shift the mean by $\Theta(\varepsilon^{1-1/p})$, rendering inference impossible below this scale. This barrier arises because moment assumptions impose only global tail control and do not constrain the local geometry of the distribution near its center.

In this section, we show that this limitation is not fundamental. By imposing a mild *local* structural condition, *directional median regularity*, we rule out the local hollowing phenomenon responsible for the moment barrier. As a result, truncation induces only linear $\mathcal{O}(\varepsilon)$ bias at the population level. However, this improvement comes at a cost: recovering the center uniformly over all directions requires sample complexity linear in the ambient dimension.

5.1 Directional Median Regularity

To bypass the moment-induced barrier identified in Section 4, we introduce a local structural assumption on the distribution near its center. This assumption is designed to capture the minimal regularity needed for robust inference under truncation, and does not rely on finite variance or higher-order moment bounds.

Assumption 1 (Directional Median Regularity)**.** *Let P be a distribution on $\mathbb{R}^d$ with mean μ_P. We assume that for every unit vector $v \in \mathbb{S}^{d-1}$:*

- *(Directional centering).* $\mathrm{median}(\langle X - \mu_P, v\rangle) = 0$.
- *(Local mass). The projection* $Z_v := \langle X - \mu_P, v\rangle$ *admits a density* f_v *satisfying*

$$f_v(t) \geq c \qquad \text{for all } |t| \leq r,$$

where $c, r > 0$ *do not depend on* v.

Remark 2 (Scope, uniformity, and optimality)**.** Our moment-based results cover distributions with bounded directional p-th moments, including asymmetric and heavy-tailed distributions. For fixed $p \geq 2$ and $\nu > 0$, substituting $\gamma = 2\nu\varepsilon^{1-1/p}$ and $\|\Sigma_P\| \leq \nu^2$ in Theorem 2 gives a uniform guarantee over $\nu_{P,p} \leq \nu$: under its signal and sample-size conditions, the same majority-vote test has error at most δ for every admissible distribution and truncation set under either hypothesis. Together with Theorem 3, this identifies the detectability threshold up to constant factors; matching finite-sample minimax rates throughout the detectable region are not established here.

Assumption 1 requires each projected mean to be a median, with density uniformly bounded below nearby; it does not require finite variance or unimodality. For example, $\mathcal{N}(\mu, I_d)$ satisfies it with $r = 1$ and $c = 1/\sqrt{2\pi e}$. The one-dimensional symmetric three-interval distribution in the proof of Lemma 4 also satisfies it, with $c = 1$ and $r = \xi$, despite its separated modes. These examples illustrate that the assumption concerns mass near the center, while allowing separated components elsewhere. For common c, r, Theorem 5 likewise gives a uniform guarantee under its stated conditions. We defer a fuller minimax treatment to a future journal version.

Assumption 1 enforces a uniform lower bound on the one-dimensional densities of all projections in a neighborhood of the center. As a consequence, for each direction v, small perturbations of the projected distribution function translate into proportionally small shifts of the directional median. This property rules out distributions for which an ε-fraction of mass can be removed near the center without noticeably affecting the observed law, which is precisely the mechanism underlying the moment-induced bias barrier. Under this assumption, truncation-induced bias improves from polynomial to linear order. The following lemma formalizes this stability at the population level.

Lemma 3 (Directional Median Stability)**.** *Suppose that the pair* (P, S) *is such that* P *satisfies Assumption 1 and* $P(S) \geq 1 - \varepsilon$ *for some* $\varepsilon < cr$*. Then for every* $v \in \mathbb{S}^{d-1}$*,*

$$\left|\mathrm{median}(\langle X, v\rangle \mid X \in S) - \langle \mu_P, v\rangle\right| \;\leq\; \frac{\varepsilon}{c}.$$

It is instructive to consider the above lemma when $d = 1$. In this case, Assumption 1 applied to the density f of $X - \mu_P$ gives:

$$2cr \leq \int_{-r}^{r} f(t)\, dt \leq 1,$$

so that $cr \leq 1/2$, with equality if and only if f is uniform over an interval of length $2r$ (and then $c = 1/2r$). The next result shows that, if we are not in this degenerate case — that is, if $cr < 1/2$ — then no analogue of Lemma 3 can hold with $\varepsilon > 2cr$.

Lemma 4 (Sharpness of median stability condition)**.** *Given* $0 < \xi < 1/2$*,* $\varepsilon > 2\xi$*, and* $R > 0$*, there exists a distribution* P *over* $\mathbb{R}$ *and a set* $S \subset \mathbb{R}$ *with the following properties.*

1. P *satisfies Assumption 1 with constants* $c, r > 0$ *such that* $cr = \xi$*;*
2. $P(S) \geq 1 - \varepsilon$*;*
3. $\left|\mathrm{median}(\langle X, v\rangle \mid X \in S) - \langle \mu_P, v\rangle\right| \;\geq\; R.$

Moreover, the truncated distribution P_S *satisfies the second part of Assumption 1 for certain constants* $c_0, r_0 > 0$*.*

5.2 Population Stability under Truncation

Lemma 3 shows that under directional median regularity, truncation induces at most a linear $\mathcal{O}(\varepsilon)$ displacement of the center in every direction. This behavior is fundamentally different from the moment-limited regime studied in Section 4, where truncation can induce a bias of order $\varepsilon^{1-1/p}$.

The underlying mechanism is purely one-dimensional. Conditioning on a truncation set S with $P(S) \geq 1 - \varepsilon$ perturbs the distribution function of each projection $\langle X, v \rangle$ by at most ε uniformly. When combined with a uniform lower bound on the density near the median, this perturbation translates directly into a bound on the shift of the directional median. In this sense, the directional median acts as a *stable pivot* whose location is controlled by the total mass removed, rather than by the behavior of the tails.

This stability property rules out the failure mode responsible for the moment-induced bias barrier. Under moment-only assumptions, the distribution may place arbitrarily little mass near its center, allowing truncation to remove signal while leaving the observed truncated distribution essentially unchanged. Directional median regularity excludes this possibility by enforcing local identifiability of the center in every direction.

5.3 Robust Estimation and Dimensional Sample Complexity

We now turn to estimation from finite samples drawn from the truncated law P_S. Lemma 3 implies that the population center remains $\mathcal{O}(\varepsilon)$-close to μ_P in every direction. To exploit this stability algorithmically, it suffices to estimate directional medians uniformly over all directions.

Given i.i.d. samples $X_1, \ldots, X_n \sim P_S$, let $\hat{m}_v$ denote the empirical median of $\langle X_1, v \rangle, \ldots, \langle X_n, v \rangle$. We define

$$\hat{\mu} \;:=\; \arg\min_{u \in \mathbb{R}^d} \sup_{v \in \mathbb{S}^{d-1}} \big| \langle u, v \rangle - \hat{m}_v \big|.$$

Theorem 4 (Robust Recovery under Median Regularity)**.** *There exists a universal $C_0 > 0$ such that the following holds. Let P satisfy Assumption 1 with a constant $c > 0$ and let S be any truncation set with $P(S) \geq 1 - \varepsilon$, where $0 \leq \varepsilon < cr$. Given n i.i.d. samples from P_S, where n satisfies*

$$n > \frac{C_0}{(cr - \varepsilon)^2} \left(d + \log(1/\delta)\right)$$

the estimator $\hat{\mu}$ satisfies, with probability at least $1 - \delta$,

$$\|\hat{\mu} - \mu_P\| \;\leq\; \frac{2}{c} \left(\varepsilon + C \sqrt{\frac{d + \log(1/\delta)}{n}} \right).$$

Theorem 4 shows that directional median regularity restores near-optimal robustness under truncation. However, achieving $\mathcal{O}(\varepsilon)$ accuracy requires $n = \Theta(d)$ samples. This dependence is unavoidable and reflects the VC dimension of the class of directional halfspaces, which must be controlled to ensure uniform stability across all projections.

Remark 3. We also note that some restriction relating ε to $< cr$ is necessary for our estimator to work. By Lemma 4, $\varepsilon > 2cr$ allows truncated distributions whose medians are arbitrarily far from the un-truncated median. However, the "bad truncated distribution" P_S produced by Lemma 4 does not satisfy Assumption 1 because its median and mean are not equal.

5.4 Testing versus Estimation under Structural Regularity

Directional median regularity eliminates truncation-induced bias for estimation, but it does not remove the intrinsic dimensional difficulty of hypothesis testing. Even in the absence of truncation, testing a d-dimensional mean is known to require $\mathcal{O}(\sqrt{d}/\beta^2)$ samples, where β denotes the effective signal strength.

Theorem 5 (Testing under Structural Regularity)**.** *Fix $d \geq 1$, $\alpha > 0$, and $\delta \in (0, 1/3)$. Consider the problem of testing $H_0 : \mu_P = 0$ versus $H_1 : \|\mu_P\| \geq \alpha > 0$ from n i.i.d. samples drawn from a truncated distribution P_S over $\mathbb{R}^d$, where $P(S) \geq 1 - \varepsilon$. Then there exist constants $c, K > 0$ such that:*

- *If P satisfies Assumption 1 with constants $c, r > 0$, $\varepsilon < \min\{cr, c\alpha/8\}$ and*

$$n \geq K \left(d + \log(1/\delta)\right) \max\{(cr - \varepsilon)^{-2}, (c\alpha)^{-2}\},$$

there exists a test with probability of error at most δ, uniformly over H_0 and H_1.

- *If P is a normal distribution over $\mathbb{R}^d$ with unknown mean μ_P and known covariance $\Sigma_P = I_d$ (which in particular satisfies Assumption 1 for $c = \sqrt{e/2\pi}$, $r = 1$), then for any $\varepsilon \in (0,1)$ and $n \leq c\, d/\varepsilon$, there exists a choice of $\alpha = \Theta(\varepsilon\sqrt{\log(1/\varepsilon)})$ such that no test can achieve error smaller than $1/3$ uniformly over both H_0 and H_1.*

The theorem shows that, for $\varepsilon \lesssim \alpha$ and suitable parameters c, r for P, the sample complexity of testing H_0 versus H_1 over directional median regular distributions is $\Theta_\alpha(d)$. This generalizes the results for [3] for Gaussian distributions; in fact, the second part of the theorem follows from Lemma 3.5 in that paper. Theorem 4 highlights a fundamental separation: directional mean regularity removes the truncation bias barrier $\varepsilon^{1-1/p} \lesssim \alpha$ for learning, but does not circumvent the inherent dimensional cost $\Theta(d)$ of high-dimensional testing.

6 Conclusion and Open Directions

We studied high-dimensional mean testing under arbitrary truncation and characterized the fundamental limits imposed by tail behavior and structural regularity. Our results identify truncation-induced bias as the primary obstacle to testing: under bounded directional p-th moment assumptions, truncation induces an unavoidable polynomial bias of order $\mathcal{O}(\nu_{P,p}\varepsilon^{1-1/p})$, where $\nu_{P,p}$ is the directional moment upper bound, yielding a sharp detectability floor below which the null and alternative hypotheses are information-theoretically indistinguishable, regardless of sample size.

Above this floor, we showed that simple second-order U-statistics achieve near-optimal testing rates, preserving the classical $\Theta(\sqrt{d})$ high-dimensional sample complexity. This separation highlights a key distinction between truncation and variance-driven noise: while truncation can systematically distort the mean, second-order concentration remains stable under finite second moments.

We further demonstrated that the moment-induced barrier is not absolute. Under a directional median regularity assumption, geometric stability restores linear $\mathcal{O}(\varepsilon)$ robustness even for heavy-tailed distributions, revealing an intermediate learning-hard regime before optimal testing rates are recovered. These findings position truncation as a distinct and fundamentally challenging corruption model, where detectability hinges on both tail decay and geometric structure.

Natural extensions include testing and estimation under more general observation mechanisms such as censoring or interval truncation, developing truncation-robust covariance and regression procedures, and understanding computational–statistical tradeoffs in the learning-hard regime identified by median regularity.

7 Acknowledgements

R. I. Oliveira was supported by a *Bolsa de Produtividade em Pesquisa* from CNPq, Brazil (# 305765/2023-0) and a *Cientista do Nosso Estado* grant from Faperj, Rio de Janeiro, Brazil (# E-26/200.485/2023). Themis Gouleakis was supported by a start up grant at Nanyang Technological University, Singapore.

References

[1] Yannick Baraud. Non-asymptotic minimax rates of testing in signal detection. *Bernoulli*, pages 577–606, 2002.

[2] Clément L Canonne, Samuel B Hopkins, Jerry Li, Allen Liu, and Shyam Narayanan. The full landscape of robust mean testing: Sharp separations between oblivious and adaptive contamination. *arXiv preprint arXiv:2307.10273*, 2023.

[3] Clement Louis Canonne, Themis Gouleakis, Yuhao Wang, and Qiping Yang. Gaussian mean testing under truncation. In *The 28th International Conference on Artificial Intelligence and Statistics*, 2025.

[4] Yeshwanth Cherapanamjeri, Constantinos Daskalakis, Andrew Ilyas, and Manolis Zampetakis. What makes a good fisherman? linear regression under self-selection bias. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, pages 1699–1712, 2023.

[5] Yeshwanth Cherapanamjeri and Daniel Lee. Heavy-tailed estimation is easier than adversarial contamination. *Proceedings of Machine Learning Research vol*, 291:1–31, 2025.

[6] A Clifford Cohen. On the solution of estimating equations for truncated and censored samples from normal populations. *Biometrika*, 44(1/2):225–236, 1957.

[7] A Clifford Cohen. *Truncated and censored samples: theory and applications*. CRC press, 1991.

[8] Constantinos Daskalakis, Themis Gouleakis, Chistos Tzamos, and Manolis Zampetakis. Efficient statistics, in high dimensions, from truncated samples. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 639–649. IEEE, 2018.

[9] Constantinos Daskalakis, Themis Gouleakis, Christos Tzamos, and Manolis Zampetakis. Computationally and statistically efficient truncated regression. In *Conference on Learning Theory*, pages 955–960. PMLR, 2019.

[10] Constantinos Daskalakis, Dhruv Rohatgi, and Emmanouil Zampetakis. Truncated linear regression in high dimensions. *Advances in Neural Information Processing Systems*, 33:10338–10347, 2020.

[11] Anindya De, Huan Li, Shivam Nadimpalli, and Rocco A. Servedio. Detecting low-degree truncation. In *STOC*, pages 1027–1038. ACM, 2024.

[12] Anindya De, Shivam Nadimpalli, and Rocco A. Servedio. Testing convex truncation. In *SODA*, pages 4050–4082. SIAM, 2023.

[13] Walter L Deemer Jr and David F Votaw Jr. Estimation of parameters of truncated or censored exponential distributions. *The Annals of Mathematical Statistics*, 26(3):498–504, 1955.

[14] Jules Depersin and Guillaume Lecué. Robust sub-gaussian estimation of a mean vector in nearly linear time. *The Annals of Statistics*, 50(1):511–536, 2022.

[15] Ilias Diakonikolas, Gautam Kamath, Daniel M. Kane, Jerry Li, Ankur Moitra, and Alistair Stewart. Robust estimation without computational intractability. In *Proceedings of the 57th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 655–664. IEEE, 2016.

[16] Ilias Diakonikolas and Daniel M Kane. Recent advances in algorithmic high-dimensional robust statistics. *arXiv preprint arXiv:1911.05911*, 2019.

[17] Ilias Diakonikolas and Daniel M Kane. *Algorithmic high-dimensional robust statistics*. Cambridge university press, 2023.

[18] Ilias Diakonikolas, Daniel M Kane, and Ankit Pensia. Outlier robust mean estimation with subgaussian rates via stability. *Advances in Neural Information Processing Systems*, 33:1830–1840, 2020.

[19] Wilfrid J Dixon. Simplified estimation from censored normal samples. *The Annals of Mathematical Statistics*, pages 385–391, 1960.

[20] Cynthia Dwork, Aaron Roth, et al. The algorithmic foundations of differential privacy. *Foundations and trends® in theoretical computer science*, 9(3–4):211–407, 2014.

[21] William Feller. *An Introduction to Probability Theory and Its Applications, Vol. 2*. John Wiley & Sons, 2nd edition, 1971. Section 7.1.

[22] RA Fisher. Properties and applications of hh functions. *Mathematical tables*, 1:815–852, 1931.

[23] Shivam Gupta, Samuel Hopkins, and Eric Price. Beyond catoni: Sharper rates for heavy-tailed and robust mean estimation. In *The Thirty Seventh Annual Conference on Learning Theory*, pages 2232–2269. PMLR, 2024.

[24] Frank R Hampel. The influence curve and its role in robust estimation. *Journal of the american statistical association*, 69(346):383–393, 1974.

[25] Mikael Møller Høgsgaard and Andrea Paudice. Uniform mean estimation for heavy-tailed distributions via median-of-means. *arXiv preprint arXiv:2506.14673*, 2025.

[26] Sam Hopkins, Jerry Li, and Fred Zhang. Robust and heavy-tailed mean estimation made simple, via regret minimization. *Advances in Neural Information Processing Systems*, 33:11902–11912, 2020.

[27] Peter J Huber and Elvezio M Ronchetti. *Robust statistics*. John Wiley & Sons, 2011.

[28] Andrew Ilyas, Emmanouil Zampetakis, and Constantinos Daskalakis. A theoretical and practical framework for regression and classification from truncated samples. In *International Conference on Artificial Intelligence and Statistics*, pages 4463–4473. PMLR, 2020.

[29] Yuri Ingster. *Nonparametric goodness-of-fit testing under Gaussian models*, volume 169. Springer Science & Business Media, 2012.

[30] Gautam Kamath, Vikrant Singhal, and Jonathan Ullman. Private mean estimation of heavy-tailed distributions. In *Conference on Learning Theory*, pages 2204–2235. PMLR, 2020.

[31] Vasilis Kontonis, Christos Tzamos, and Manolis Zampetakis. Efficient truncated statistics with unknown truncation. In *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1578–1595. IEEE, 2019.

[32] Jane H. Lee, Anay Mehrotra, and Manolis Zampetakis. Efficient statistics with unknown truncation, polynomial time algorithms, beyond gaussians. In *FOCS*, pages 988–1006. IEEE, 2024.

[33] Gábor Lugosi and Shahar Mendelson. Mean estimation and regression under heavy-tailed distributions: A survey. *Foundations of Computational Mathematics*, 19(5):1145–1190, 2019.

[34] Stanislav Minsker. Uniform bounds for robust mean estimators. *Stochastic Processes and their Applications*, page 104724, 2025.

[35] Roman Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge University Press, 2018. Appendix A.2.

[36] Shanshan Wu, Alexandros G Dimakis, and Sujay Sanghavi. Learning distributions generated by one-layer relu networks. *Advances in neural information processing systems*, 32, 2019.

[37] Manolis Zampetakis and Felix Zhou. Private statistical estimation via truncation. *arXiv preprint arXiv:2505.12541*, 2025.

A Additional Related Work

This section provides a more detailed discussion of prior work related to truncation, robust inference, and high-dimensional testing.

The landscape of minimax signal detection and property testing is significantly shaped by several fundamental works that characterize the limits of distinguishing a signal from noise. [1] established sharp non-asymptotic minimax rates of testing for Gaussian sequence models over various constraint sets, including ellipsoids and ℓ_p-bodies for $p \in [0, 2]$. Complementing this, [29] provides a comprehensive treatment of nonparametric goodness-of-fit testing under Gaussian models, defining the minimax separation rates that separate detectable signals from indistinguishable noise in the asymptotic regime. While these classical results define the information-theoretic limits for testing under standard Gaussian noise, they generally assume the observations follow a known distribution without missingness or adversarial intervention. In contrast, the truncation model studied here introduces a systematic bias barrier of order $\mathcal{O}(\nu_{P,p}\varepsilon^{1-1/p})$ that persists even with infinite data, requiring structural regularity to overcome.

Mean estimation and learning from truncated samples. Classical work on truncated and censored samples focuses on parametric estimation (e.g., for exponential or normal families), typically assuming a known truncation mechanism and developing method-of-moments or maximum-likelihood procedures [13, 6, 19, 22, 7]. More recently, algorithmic work studied high-dimensional learning tasks from truncated data under known truncation sets, including efficient algorithms for truncated statistics and truncated regression [8, 9, 10], as well as structured truncation arising from generative models such as one-layer ReLU networks [36]. Beyond known-truncation settings, efficient algorithms for unknown truncation have also been developed for certain parametric families [31, 32], and for regression / classification under truncation mechanisms motivated by self-selection [28, 4].

Robustness: truncation vs. contamination. Robust statistics traditionally studies contamination models (e.g., ε-contamination), where an adversary replaces a fraction of samples with arbitrary outliers [27, 17, 14, 18]. While truncation can be viewed as a restricted form of adaptive (strong) contamination—since an adversary can remove informative mass—it is fundamentally different: truncation creates systematic missingness and can "hollow out" the distribution near its center. For robust mean testing under contamination, [2] characterizes sharp separations between oblivious and adaptive contamination, including regimes where testing is strictly easier than learning. In contrast, under truncation we show that testability is governed by truncation-induced bias floors and, under structural regularity, an intermediate learning-hard regime appears before the optimal $\Theta(\sqrt{d})$ testing rate is recovered. Other recent works have established powerful methods for robust mean estimation—such as the nearly linear-time algorithms by Depersin and Lecué [14] and the stability-based framework by Diakonikolas et al. [18] —in the aforementioned adversarial contamination model.

B Deferred Proof in Section 3

This appendix collects the proofs of the results in Section 3. Lemma 1 establishes basic bounds on parameter distortion under truncation, which are then specialized in Corollaries 6 and 7.

B.1 Sub-Gaussian and Finite Moment Heavy Tailed Distributions

Light-tailed distributions If $X \sim \mathcal{N}(\mu_P, \Sigma_P)$, then for any unit vector v, $\langle v, X - \mu_P\rangle \sim \mathcal{N}(0, v^\top \Sigma_P v)$. Hence, letting $Z \sim \mathcal{N}(0, 1)$,

$$\mathbb{E}|\langle v, X - \mu_P\rangle|^p = (v^\top \Sigma_P v)^{p/2}\, \mathbb{E}|Z|^p = (v^\top \Sigma_P v)^{p/2}\, 2^{p/2} \frac{\Gamma((p+1)/2)}{\sqrt{\pi}}.$$

Using Stirling's approximation, $(\Gamma((p+1)/2))^{1/p} \asymp \sqrt{p}$ as $p \to \infty$ (see, e.g., [35, App. A.2] or [21, Sec. 7.1]). Hence,

$$\nu_{P,p} = (\mathbb{E}|Z|^p)^{1/p} \|\Sigma_P\|^{1/2} \asymp \sqrt{p}\, \|\Sigma_P\|^{1/2}.$$

More generally (see also [35, Theorem 2.5.2]), if X is sub-Gaussian with parameter σ, i.e.,

$$(\mathbb{E}|\langle v, X - \mu_P\rangle|^p)^{1/p} \le C\sigma\sqrt{p}$$

for all $p \geq 1$ and $\|v\| = 1$, then $\nu_{P,p} \leq C\sigma\sqrt{p}$.

Heavy-tailed distributions. For heavy-tailed distributions, higher moments may grow faster than $\sqrt{p}$ or are even unbounded beyond some finite order. Intuitively, $\nu_{P,p}$ measures how large the p-th directional moment can be; it grows slowly (as $\sqrt{p}$) for light-tailed data but much faster when tails are heavy.

To make this precise, suppose that for some unit vector v, the projection $\langle v, X - \mu_P \rangle$ has a polynomially decaying tail:

$$\Pr(|\langle v, X - \mu_P \rangle| > t) \asymp t^{-\kappa}, \qquad t \to \infty,$$

for some tail exponent $\kappa > 2$. Then

$$\mathbb{E}|\langle v, X - \mu_P \rangle|^p < \infty$$

only for $p < \kappa$, then $\nu_{P,p}$ remains bounded for $p < p_0 = \kappa$, but diverges as $p \uparrow \kappa$.

For Gaussian or sub-Gaussian distributions, the truncation bias and covariance distortion both decay nearly linearly in ε.

Corollary 6 (Light-tailed specialization: Gaussian and sub-Gaussian). Suppose P has sub-Gaussian one-dimensional marginals with parameter $\sigma > 0$, that is

$$(\mathbb{E}|\langle v, X - \mu_P \rangle|^p)^{1/p} \leq C_0\, \sigma\sqrt{p} \quad \text{for all } p \geq 1,\ \|v\| = 1.$$

Then for any measurable $S \subset \mathbb{R}^d$ with $P(S) \geq 1 - \varepsilon$ and $\varepsilon \in (0, 1/2]$,

$$\|\mu_P - \mu_{P_S}\| \ \leq\ C_1\, \sigma\, \varepsilon\, \sqrt{\log(1/\varepsilon)},$$

for universal constants $C_1 > 0$.

(Gaussian distribution.) If $X \sim \mathcal{N}(\mu_P, \Sigma_P)$, then the same bounds hold with $\sigma^2 = \|\Sigma_P\|$.

Proof. This follows from Lemma 1 by plugging in the sub-Gaussian moment growth $\nu_{P,p} \leq C_0\sigma\sqrt{p}$ and optimizing over p, which yields $p \asymp \log(1/\varepsilon)$. □

Intuitively, sub-Gaussian moment control implies that, along any direction, the magnitude of typical fluctuations is at scale at most σ, with large deviations becoming exponentially unlikely. Therefore, even if an adversary removes an ε-fraction of the probability mass by truncation, the maximum resulting shift in the mean, which is governed by the largest values that can be removed at ε-quantile scale, is nearly linear on $\sigma\,\varepsilon$ (with an additional $\sqrt{\log(1/\varepsilon)}$ factor) as the bulk of the mass is removed from points at distance $\Theta(\sigma)$ from the mean.

This behavior contrasts sharply with the heavy-tailed setting, where removing an ε-fraction of mass can excise regions that carry a non-negligible share of the total variance. The next corollary quantifies this degradation in robustness as the tail index decreases.

Corollary 7 (Heavy-tailed specialization: finite-p_0 moments). Assume there exist $p_0 \in [2, \infty)$ and $M > 0$ such that

$$\sup_{\|v\|=1} \left(\mathbb{E}\, |\langle v, X - \mu_P \rangle|^{p_0} \right)^{1/p_0} \ \leq\ M.$$

Then for any measurable $S \subset \mathbb{R}^d$ with $P(S) \geq 1 - \varepsilon$ and $\varepsilon \in [0, 1/2]$,

$$\|\mu_P - \mu_{P_S}\| \leq 2\, M\, \varepsilon^{\,1 - \frac{1}{p_0}}.$$

Proof. By assumption, $\nu_{P,p_0} \leq M$. Apply Lemma 1 with $p = p_0$. □

Examples.

- **Student-t tails.** If each one-dimensional projection $\langle v, X-\mu_P\rangle$ has a Student-t_κ distribution (up to scale) with degrees of freedom $\kappa > 2$, then moments exist for $p < \kappa$. Taking any $p_0 \in [2, \kappa)$ and $M = \sup_{\|v\|=1} \left(\mathbb{E}|\langle v, X-\mu_P\rangle|^{p_0}\right)^{1/p_0}$,
$$\|\mu_P - \mu_{P_S}\| \lesssim M\, \varepsilon^{1-\frac{1}{p_0}}.$$
As κ decreases (heavier tails), the best admissible p_0 decreases and the exponents $1-\frac{1}{p_0}$ correspondingly worsen.
- **Pareto tails.** If $\Pr(|\langle v, X-\mu_P\rangle| > t) \asymp t^{-\alpha}$ for some $\alpha > 2$ (uniformly over unit v up to scale), then moments exist for $p < \alpha$. Choosing any $p_0 \in [2, \alpha)$ and the corresponding M yields the same bounds with exponents $1-\frac{1}{p_0}$. In particular, taking p_0 close to α gives the rate $\varepsilon^{1-\frac{1}{\alpha}}$ for the mean deviation (up to constants).

B.2 Truncation-Induced Parameter Distortion

We first prove Lemma 1, which quantifies the bias induced by truncation under finite moment assumptions.

Lemma 2 (Truncated vs. Non-truncated Parameters). Let P be a distribution and S a measurable set with $P(S) \geq 1-\varepsilon$ for $\varepsilon \in [0, 1/2]$. Then for any $p \geq 1$:

1. *Mean Shift:* $\|\mu_P - \mu_{P_S}\| \leq 2\nu_{P,p}\varepsilon^{1-1/p}$.
2. *Covariance Bound:* $\|\Sigma_{P_S}\| \leq (1-\varepsilon)^{-2}\|\Sigma_P\| \leq 4\|\Sigma_P\|$.

Proof. **Mean difference bound:** Let $v \in \mathbb{R}^d$ be any unit vector, i.e. $\|v\| = 1$. We start by expressing the directional difference of the means:
$$|\langle \mu_P - \mu_{P_S}, v\rangle| = \frac{1}{P(S)}|\mathbb{E}_{X\sim P}[\mathbf{1}_{X\in S}\langle v, \mu_P - X\rangle]|.$$
Since $P(S) \geq 1-\varepsilon \geq 1/2$, we have $P(S)^{-1} \leq 2$, and hence
$$|\langle \mu_P - \mu_{P_S}, v\rangle| \leq 2|\mathbb{E}_{X\sim P}[\mathbf{1}_{X\in S}\langle v, \mu_P - X\rangle]|.$$
Next, note that
$$\mathbb{E}_{X\sim P}[\langle v, \mu_P - X\rangle] = 0,$$
so we can rewrite
$$\mathbb{E}_{X\sim P}[\mathbf{1}_{X\in S}\langle v, \mu_P - X\rangle] = \mathbb{E}_{X\sim P}[\langle v, \mu_P - X\rangle] - \mathbb{E}_{X\sim P}[\mathbf{1}_{X\notin S}\langle v, \mu_P - X\rangle] = -\mathbb{E}_{X\sim P}[\mathbf{1}_{X\notin S}\langle v, \mu_P - X\rangle],$$
Applying Hölder's inequality with conjugate exponents p and $q = \frac{p}{q-1}$ gives:
$$|\mathbb{E}_{X\sim P}[\mathbf{1}_{X\notin S}\langle v, \mu_P - X\rangle]| \leq (1-P(S))^{1-1/p}\left(\mathbb{E}_{X\sim P}|\langle v, X-\mu_P\rangle|^p\right)^{1/p}.$$
Since $1 - P(S) \leq \varepsilon$ and by the definition of $\nu_{P,p}$,
$$|\mathbb{E}_{X\sim P}[\mathbf{1}_{X\notin S}\langle v, \mu_P - X\rangle]| \leq \varepsilon^{1-1/p}\nu_{P,p}.$$
Substituting this into the previous inequality yields,
$$|\langle \mu_P - \mu_{P_S}, v\rangle| \leq 2\varepsilon^{1-1/p}\nu_{P,p},$$
Finally, taking the supremum over all unit vectors v gives
$$\|\mu_P - \mu_{P_S}\| \leq 2\varepsilon^{1-1/p}\nu_{P,p}.$$

Covariance difference bound: To control the deviation between Σ_P and Σ_{P_S}, let $X, Y \sim P$ be independent. Then
$$\Sigma_P = \frac{1}{2}\mathbb{E}[(X-Y)^{\otimes 2}],$$
where $a^{\otimes 2} \stackrel{\text{def.}}{=} aa^\top$. Since $(X-Y)^{\otimes 2}$ is positive semidefinite,
$$\Sigma_P \succeq \frac{1}{2}\mathbb{E}[(X-Y)^{\otimes 2}\,\mathbf{1}_{X\in S}\mathbf{1}_{Y\in S}] = P(S)^2\Sigma_{P_S},$$
and thus $\|\Sigma_{P_S}\| \leq P(S)^{-2}\|\Sigma_P\| \leq (1-\varepsilon)^{-2}\|\Sigma_P\| \leq 4\|\Sigma_P\|$, where the last inequality is due to the fact that $\varepsilon \leq \frac{1}{2}$ □

C Deferred Proof in Section 4

This appendix contains the technical proofs for the testing results in Section 4. We first establish a variance bound for the U-statistic used in our test, and then complete the proof of the median-of-means testing procedure under truncation.

C.1 Lemma 3 (Variance Bound for U-statistic).

The following lemma bounds the variance of the U-statistic used to estimate $\|\mu_{P_S}\|^2$, and is the main concentration ingredient in the proof of the testing upper bound.

Lemma 3 (Variance Bound for U-statistic). Fix a distribution Q and let $X_1, \dots, X_n \sim Q$ be i.i.d. Then the variance of the U-statistic

$$F := \binom{n}{2}^{-1} \sum_{1 \le i < j \le n} \langle X_i, X_j \rangle$$

satisfies

$$\mathrm{Var}_Q(F) \;\le\; \frac{4}{n} \langle \mu_Q, \Sigma_Q \mu_Q \rangle \;+\; \frac{4}{n(n-1)} \operatorname{tr}\big(\Sigma_Q^2\big). \tag{9}$$

Proof. The reliability of F as a test statistic depends on its concentration. We bound $\mathrm{Var}(F)$ via Efron-Stein inequality:

$$\mathrm{Var}(F) \le \frac{1}{2} \sum_{i=1}^{n} \mathbb{E}[(F - F^{(i)})^2],$$

where, for each $i \in [n]$, $F^{(i)} = F(X_1, \dots, X_{i-1}, X_i', X_{i+1}, \dots, X_n)$ replaces X_i by an independent copy $X_i' \sim Q$.

Conditioning on X_i, X_i',

$$F - F^{(i)} = \frac{1}{\binom{n}{2}} \sum_{j \ne i} \langle X_i - X_i', X_j \rangle$$

Since the $\{X_j\}_{j \ne i}$ are i.i.d. with mean μ_Q and covariance Σ_Q,

$$\mathbb{E}[F - F^{(i)} \mid X_i, X_i'] = \tfrac{2}{n} \langle X_i - X_i', \mu_Q \rangle, \qquad \mathrm{Var}[F - F^{(i)} \mid X_i, X_i'] = \tfrac{4}{n^2(n-1)} (X_i - X_i')^\top \Sigma_Q (X_i - X_i').$$

It follows that

$$\begin{aligned}
\mathbb{E}[(F - F^{(i)})^2 \mid X_i, X_i'] &= (\mathbb{E}[F - F^{(i)} \mid X_i, X_i'])^2 + \mathrm{Var}[F - F^{(i)} \mid X_i, X_i'] \\
&= \frac{4}{n^2} \langle X_i - X_i', \mu_P \rangle^2 + \frac{n-1}{\binom{n}{2}^2} \langle X_i - X_i', \Sigma_Q (X_i - X_i') \rangle \\
&= \frac{4}{n^2} \langle X_i - X_i', \mu_P \rangle^2 + \frac{4}{n^2(n-1)} \operatorname{tr}(\Sigma_Q (X_i - X_i')^{\otimes 2}).
\end{aligned}$$

Taking expectations over X_i, X_i' and summing over i yields

$$\mathrm{Var}(F) \;\le\; \frac{4}{n} \langle \mu_Q, \Sigma_Q \mu_Q \rangle + \frac{4}{n(n-1)} \operatorname{tr}(\Sigma_Q^2).$$

□

C.2 Median-of-means estimator

In this subsection, we construct a simple test for mean detection under truncation with constant error probability. We next show how to amplify the constant success probability in Theorem 1 to an arbitrary confidence level by standard probability boosting in Theorem 2. This yields the final achievable testing rates under truncation.

Theorem 1 (Mean Testing under truncation - constant error prob.). *Let $S \subset \mathbb{R}^d$ satisfy $P(S) \geq 1-\varepsilon$ with $\varepsilon \in [0, 1/2]$, and let $X_1, \ldots, X_n \overset{iid}{\sim} P_S := P(\cdot \,|\, S)$. Let $\gamma = 2\,\nu_{P,p}\,\varepsilon^{1-1/p}$ and $\alpha > 2\gamma$. Consider F as above and the test $T = \mathbf{1}\left\{ F > \frac{\alpha^2}{4} \right\}$. There exists a universal constant $C > 0$ such that, if*

$$n \;\geq\; \max\left\{ 2, C\,\frac{\|\Sigma_P\|\sqrt{d}}{(\alpha - 2\gamma)^2} \right\},$$

then T separates $\|\mu_P\| = 0$ from $\|\mu_P\| \geq \alpha$ with probability at least $2/3$.

Proof. The fact that the statistic F is an unbiased estimator of the squared Euclidean norm of the truncated mean was noted in the main text: $\mathbb{E}_{P_S}[F] = \|\mu_{P_S}\|^2$. Furthermore, by Lemma 1, we have that $\|\mu_P - \mu_{P_S}\| \leq \gamma$.

Combining the variance bound for F in (7) with Chebyshev's inequality, with obtain that there exists a universal $c > 1$ such that, for any $n \geq 2$:

$$\mathbb{P}\left[|F - \|\mu_{P_S}\|^2| \leq c\,\left(\|\mu_{P_S}\| \sqrt{\frac{\|\Sigma_P\|}{n}} + \frac{\sqrt{d}\|\Sigma_P\|}{n} \right) \right] \geq 2/3. \tag{10}$$

In the remainder of the proof, we set $\Delta := \alpha/2 - \gamma > 0$. By taking the constant $C > 0$ in the assumptions of the theorem to be large enough, we may assume that:

$$n \geq \max\left\{ 2, \frac{\|\Sigma_P\|(c^2 + 4\sqrt{d})}{\Delta^2} \right\}, \tag{11}$$

In the remainder of the proof, we will combine (11) with (10), to prove that our test works as expected. We now distinguish the 2 cases:

(Soundness): Under the alternative hypothesis $\|\mu_P\| \geq \alpha$ we have $\|\mu_{P_S}\| \geq (\alpha - \gamma)$ due to the triangle inequality.

We *claim* that we have $F > \alpha^2/4$ whenever the event in (10) holds (and n is suitably large). Notice that this claim implies $\mathbb{P}(T = 1) = \mathbb{P}(F > \alpha^2/4) \geq 2/3$, so that soundness ensues.

To prove the claim, note that, whenever the event in (10) holds, we have the following lower bound on F:

$$\begin{aligned} F &\geq \|\mu_{P_S}\|^2 - c\,\left(\|\mu_{P_S}\| \sqrt{\frac{\|\Sigma_P\|}{n}} + \frac{\sqrt{d}\|\Sigma_P\|}{n} \right) \\ &= \left(\|\mu_{P_S}\| - \frac{c}{2}\sqrt{\frac{\|\Sigma_P\|}{n}} \right)^2 - \frac{\|\Sigma_P\|(c^2 + 4c\sqrt{d})}{4n}. \end{aligned} \tag{12}$$

Using (11), we obtain

$$\|\mu_{P_S}\| - \frac{c}{2}\sqrt{\frac{\|\Sigma_P\|}{n}} \geq \alpha - \gamma - \frac{\Delta}{2} = \frac{\alpha + \Delta}{2} > 0$$

and

$$\frac{\|\Sigma_P\|(c^2 + 4c\sqrt{d})}{4n} \leq \frac{\Delta^2}{4}.$$

Plugging this back into (12), we obtain that

$$F \geq \left(\frac{\alpha + \Delta}{2} \right)^2 - \frac{\Delta^2}{4} = \frac{\alpha^2}{4} + \frac{\alpha\Delta}{2} > \frac{\alpha^2}{4}$$

whenever the event in (10) holds. This is precisely our claim, and finishes the proof of soundness.

(Completeness): Under the null hypothesis $\|\mu_P\| = 0$ we have $\|\mu_{P_S}\| \leq \gamma$.

We *claim* that, in the present case, we have $F \le \alpha^2/4$ whenever the event in (10) holds (and n is suitably large). Notice that this implies $\mathbb{P}(T = 0) = \mathbb{P}(F \le \alpha^2/4) \ge 2/3$.

To prove the claim, we observe that, when the event (10) holds, we can upper bound:

$$\begin{aligned} F &\le \|\mu_{P_S}\|^2 + c\left(\|\mu_{P_S}\|\sqrt{\frac{\|\Sigma_P\|}{n}} + \frac{\sqrt{d}\|\Sigma_P\|}{n}\right) \\ &= \left(\|\mu_{P_S}\| + \frac{c}{2}\sqrt{\frac{\|\Sigma_P\|}{n}}\right)^2 + \frac{\|\Sigma_P\|(4c\sqrt{d} - c^2)}{4n}. \end{aligned} \tag{13}$$

Using $\|\mu_{P_S}\| \le \gamma$ and (11), we obtain:

$$\|\mu_{P_S}\| + \frac{c}{2}\sqrt{\frac{\|\Sigma_P\|}{n}} \le \gamma + \frac{\Delta}{2} \text{ and } \frac{\|\Sigma_P\|(4c\sqrt{d} - c^2)}{4n} \le \frac{\Delta^2}{4}.$$

Plugging this back into (13) gives

$$F \le \left(\gamma + \frac{\Delta}{2}\right)^2 + \frac{\Delta^2}{4} \le (\gamma + \Delta)^2 = \frac{\alpha^2}{4}$$

whenever the event in (10) holds. This proves our claim and finishes the proof of completeness. □

Theorem 2 (Mean Testing under Truncation)**.** *In the same setting as Theorem 1, given $\delta \in (0, 1/3)$, there exists a test T that distinguishes $H_0 : \mu_P = 0$ from $H_1 : \|\mu_P\| \ge \alpha$ with probability at least $1 - \delta$ whenever $\alpha > 2\gamma$ and the sample size satisfies*

$$n \;\ge\; C\left\{2, \frac{\|\Sigma_P\|}{(\alpha - 2\gamma)^2}\sqrt{d}\right\}\log(1/\delta), \tag{8}$$

where $C > 0$ is a universal constant.

Proof sketch. The result follows from splitting the samples into $C\log(1/\delta)$ blocks, performing the test from Theorem 1 on each block, and then taking a majority vote. □

D Proofs for Section 5

This appendix contains the proofs of the results in Section 5. We first establish stability properties of directional medians under truncation, then show how these properties enable robust mean estimation and improved testing rates under structural regularity assumptions.

D.1 Proof of Lemma 3

Lemma 6 (Directional Median Stability). Suppose that the pair (P, S) is such that P satisfies Assumption 1 and $P(S) \ge 1 - \varepsilon$ for some $\varepsilon < cr$. Then for every $v \in \mathbb{S}^{d-1}$,

$$\left|\text{median}(\langle X, v\rangle \mid X \in S) - \langle \mu_P, v\rangle\right| \;\le\; \frac{\varepsilon}{c}.$$

Proof. Fix a direction $v \in \mathbb{S}^{d-1}$ and define $Z_v = \langle X - \mu_P, v\rangle$. By Assumption 1, Z has median 0 and density $f_v(t) \ge c$ for all $|t| \le r$.

For any $t \in (0, r]$,

$$\mathbb{P}(Z_v \ge t) \;\le\; \frac{1}{2} - \int_0^t f_v(s)\,ds \;\le\; \frac{1}{2} - ct.$$

Conditioning on a truncation set S with $P(S) \ge 1 - \varepsilon$ removes at most ε mass, hence

$$\mathbb{P}_S(Z_v \ge t) \;\le\; \mathbb{P}(Z_v \ge t) + \varepsilon \;\le\; \frac{1}{2} - ct + \varepsilon.$$

In particular, if $t \in (\varepsilon/c, r]$ (a non-vacuous condition since $\varepsilon < cr$), $\mathbb{P}_S(Z_v \geq t) < 1/2$. This implies that the median of Z_v is $< t$ for all $t \in (\varepsilon/c, r]$, which implies that

$$\text{median}(Z_v | X \in S) = \text{median}(\langle X, v\rangle | X \in S) - \langle \mu_P, v\rangle \leq \frac{\varepsilon}{c}.$$

A similar argument with $-Z_v = Z_{-v}$ shows that

$$\text{median}(Z_v | X \in S) = \text{median}(\langle X, v\rangle | X \in S) - \langle \mu_P, v\rangle \geq -\frac{\varepsilon}{c},$$

which proves the Lemma. □

D.2 Proof of Lemma 4

Lemma 7 (Sharpness of median stability condition). Given $0 < \xi < 1/2$, $\varepsilon > 2\xi$, and $R > 0$, there exists a distribution P over $\mathbb{R}$ and a set $S \subset \mathbb{R}$ with the following properties.

1. P satisfies Assumption 1 with constants $c, r > 0$ satisfying $cr = \xi$;
2. $P(S) \geq 1 - \varepsilon$;
3. $\left|\text{median}(\langle X, v\rangle \mid X \in S) - \langle \mu_P, v\rangle\right| \geq R$.

Moreover, the truncated distribution P_S satisfies the second part of Assumption 1 for certain constants $c_0, r_0 > 0$.

Proof. Recall that $0 < \xi < \varepsilon < 1/2$. For a non-degenerate closed interval $I \subset \mathbb{R}$, let L^I denote (unnormalized) Lebesgue measure over this interval. Now assume without loss of generality that $R > 1$. Set $\eta := 1/2 - \xi \in (0, 1/2)$ and define:

$$P^R := L^{[-R-\eta, -R]} + L^{[R, R+\eta]} + L^{[-\xi, \xi]}.$$

This is a compactly supported, absolutely continuous probability measure that is symmetric around 0, therefore the mean and median of P^R are both 0. One can check that P^R satisfies Assumption 1 with $c = 1$, $r = \xi$, so that $cr = \xi$.

Take $\alpha := \min\{(\varepsilon - 2\xi), \eta\} > 0$. Now set

$$S_R := [-R - \eta + \alpha, -R] \cup [1, +\infty).$$

Then $P^R(S^R) = 2\eta - \alpha = 1 - \alpha - 2\xi \geq 1 - \varepsilon$. Moreover, the truncated measure $P^R_{S_R}$ satisfies:

$$P^R_{S_R} := \frac{1}{1 - \alpha - 2\xi}\left(L^{[-R-\eta+\alpha, -R]} + L^{[R, R+\eta]}\right).$$

Since $L^{[R,R+\eta]}((-\infty, R]) = 0$ and $L^{[-R-\eta+\alpha,-R]}((-\infty, R]) = L^{[-R-\eta+\alpha,-R]}(\mathbb{R}) < L^{[R,R+\eta]}(\mathbb{R})$, we see that:

$$P^R_{S_R}((-\infty, R]) = \frac{L^{[-R-\eta+\alpha,-R]}((-\infty, R])}{L^{[-R-\eta+\alpha,-R]}(\mathbb{R}) + L^{[R,R+\eta]}(\mathbb{R})} < 1/2$$

and the median of $P^R(S_R)$ is $> R$. This finishes the proof. □

D.3 Proof of Theorem 4

Theorem 4 (Robust Recovery under Median Regularity)**.** *There exists a universal $C_0 > 0$ such that the following holds. Let P satisfy Assumption 1 with a constant $c > 0$ and let S be any truncation set with $P(S) \geq 1 - \varepsilon$, where $0 \leq \varepsilon < cr$. Given n i.i.d. samples from P_S, where n satisfies*

$$n > \frac{C_0}{(cr - \varepsilon)^2}\left(d + \log(1/\delta)\right)$$

the estimator $\hat{\mu}$ satisfies, with probability at least $1 - \delta$,

$$\|\hat{\mu} - \mu_P\| \leq \frac{2}{c}\left(\varepsilon + C\sqrt{\frac{d + \log(1/\delta)}{n}}\right).$$

Proof. Define $F : \mathbb{R}^d \to \mathbb{R}$ via the following recipe:

$$F(u) := \sup_{v \in \mathbb{S}^{d-1}} |\hat{m}_v - \langle u, v \rangle| \ \ (u \in \mathbb{R}^d).$$

Our estimator for μ_P is $\hat{\mu} \in \operatorname{argmin}_{u \in \mathbb{R}^d} F(u)$. Therefore,

$$\begin{aligned}
\|\hat{\mu} - \mu_P\| &= \sup_{v \in \mathbb{S}^{d-1}} |\langle v, \hat{\mu} - \mu_P \rangle| \\
&\leq \sup_{v \in \mathbb{S}^{d-1}} [|\langle v, \hat{\mu} \rangle - \hat{m}_v| + |\hat{m}_v - \langle v, \mu_P \rangle|] \\
&\leq F(\hat{\mu}) + F(\mu_P) \\
&\leq 2F(\mu_P).
\end{aligned}$$

Therefore, Theorem 4 follows if we can show that, the following inequality holds with probability $\geq 1 - \delta$:

$$F(\mu_P) = \sup_{v \in \mathbb{S}^{d-1}} |\hat{m}_v - \langle \mu_P, v \rangle| \leq \frac{1}{c} \left(\varepsilon + C \sqrt{\frac{d + \log(1/\delta)}{n}} \right) =: t_*. \tag{14}$$

Here, $C > 0$ is a universal constant chosen so that (15) and (16) below hold with suitably high probablity.

To prove this, we first note that, by the same argument as in the proof of Lemma 3, the following holds for any $0 < t \leq r$ and $v \in \mathbb{S}^{d-1}$:

$$\begin{aligned}
\mathbb{P}_S(\langle X, v \rangle \leq \langle \mu_P, v \rangle - t) &\leq \frac{1}{2} + \varepsilon - ct; \\
\mathbb{P}_S(\langle X, v \rangle \geq \langle \mu_P, v \rangle + t) &\leq \frac{1}{2} + \varepsilon - ct.
\end{aligned}$$

A standard VC concentration argument for halfspaces (VC dimension $\Theta(d)$) implies that, with a suitable universal choice of $C > 0$, the following inequalities hold with probability $\geq 1 - \delta$ simultaneously for all $0 < t \leq c$ and $v \in \mathbb{S}^{d-1}$:

$$\frac{1}{n} \sum_{i=1}^{n} \mathbf{1}\{\langle X_i, v \rangle \leq \langle \mu_P, v \rangle - t\} \leq \frac{1}{2} - ct + \varepsilon + C \sqrt{\frac{d + \log(1/\delta)}{n}}; \tag{15}$$

$$\frac{1}{n} \sum_{i=1}^{n} \mathbf{1}\{\langle X_i, v \rangle \leq \langle \mu_P, v \rangle + t\} \leq \frac{1}{2} - ct + \varepsilon + C \sqrt{\frac{d + \log(1/\delta)}{n}}. \tag{16}$$

Recall that t_* is the RHS of (14). Our assumptions on n and ε (with a suitable choice of C_0) guarantee that $t_* < r$. This means that we can take any $t \in (t_*, r]$ in inequalities (15) and (16), and any such t makes the RHS of these inequalities strictly smaller than $1/2$. Recalling that $\hat{m}_v$ is the sample median of $\langle X_i, v \rangle$, we obtain that the following inequalities hold with probability at least $1 - \delta$, simultaneously for all $v \in \mathbb{S}^{d-1}$ and $t \in (t_*, r]$:

$$|\hat{m}_v - \langle \mu_P, v \rangle| < t.$$

Letting $t \searrow t_*$, we obtain that (14) holds with probability $\geq 1 - \delta$, as desired. □

D.4 Proof of Theorem 5

Theorem 5 (Testing under Structural Regularity). *Fix $d \geq 1$, $\alpha > 0$, and $\delta \in (0, 1/3)$. Consider the problem of testing $H_0 : \mu_P = 0$ versus $H_1 : \|\mu_P\| \geq \alpha > 0$ from n i.i.d. samples drawn from a truncated distribution P_S over $\mathbb{R}^d$, where $P(S) \geq 1 - \varepsilon$. Then there exist constants $c, K > 0$ such that:*

- *If P satisfies Assumption 1 with constants $c, r > 0$, $\varepsilon < \min\{cr, c\alpha/8\}$ and*

$$n \geq K \left(d + \log(1/\delta)\right) \max\{(cr - \varepsilon)^{-2}, (c\alpha)^{-2}\},$$

there exists a test with probability of error at most δ, uniformly over H_0 and H_1.

- *If P is a normal distribution over $\mathbb{R}^d$ with unknown mean μ_P and known covariance $\Sigma_P = I_d$ (which in particular satisfies Assumption 1 for $c = \sqrt{e/2\pi}$, $r = 1$), then for any $\varepsilon \in (0, 1)$ and $n \leq c\, d/\varepsilon$, there exists a choice of $\alpha = \Theta(\varepsilon\sqrt{\log(1/\varepsilon)})$ such that no test can achieve error smaller than $1/3$ uniformly over both H_0 and H_1.*

Proof. The second statement in the Theorem follows from [3, Lemma 3.5], so we focus on the first statement. Let $X_1, \ldots, X_n \sim P_S$ be i.i.d. samples, and let $\hat{\mu}$ be the estimator from Theorem 4. Our test takes the form:

$$T := \mathbf{1}\left\{\|\hat{\mu}\| > \alpha/2\right\}.$$

A sufficient condition for T to give the right answer, under H_0 and H_1 both, is that

$$\|\hat{\mu} - \mu_P\| < \frac{\alpha}{2}.$$

Choose a universal $K > C_0$ with $K \geq 64C^2$, where C_0, C are from Theorem 4. The first sample-size term ensures $n > C_0 D/(cr - \varepsilon)^2$. Also, let $D = d + \log(1/\delta)$. The second term ensures $C\sqrt{D/n} \leq c\alpha/8$. Since $\varepsilon < c\alpha/8$, Theorem 4 gives, with probability at least $1 - \delta$,

$$\|\widehat{\mu} - \mu_P\| \leq (2/c)(\varepsilon + C_1\sqrt{D/n}) < \alpha/2.$$

On this event $T = 0$ under the null. Under every alternative, $\|\widehat{\mu}\| \geq \|\mu_P\| - \|\widehat{\mu} - \mu_P\| > \alpha/2$, so $T = 1$.

This finishes the proof. □

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract and introduction state the main theoretical contributions, including the finite-moment truncation bias floor and the structural median-regular regime.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: The paper discusses limitations in Sections 4–6.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

Justification: All theoretical results are formally stated with assumptions and proofs. Proofs are provided either in the main text or in the appendix.

Guidelines:

- The answer [N/A] means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [N/A]

   Justification: The paper is entirely theoretical and does not include empirical experiments.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [N/A]

Justification: The paper does not include experiments, datasets, or code-based empirical evaluations.

Guidelines:

- The answer [N/A] means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (https://neurips.cc/public/guides/CodeSubmissionPolicy) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (https://neurips.cc/public/guides/CodeSubmissionPolicy) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [N/A]

   Justification: The paper does not involve training procedures or empirical experiments.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [N/A]

   Justification: The paper contains no empirical experiments.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
   - The assumptions made should be given (e.g., Normally distributed errors).
   - It should be clear whether the error bar is the standard deviation or the standard error of the mean.

- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [N/A]

   Justification: The paper does not contain computational experiments.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: The research is theoretical in nature and complies with the NeurIPS Code of Ethics.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [N/A]

    Justification: This work is foundational theoretical research in statistical hypothesis testing. The paper does not present direct societal deployment considerations.

    Guidelines:

    - The answer [N/A] means that there is no societal impact of the work performed.
    - If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.

- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [N/A]

    Justification: The paper does not release models, datasets, or software assets that pose misuse risks.

    Guidelines:

    - The answer [N/A] means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [N/A]

    Justification: The paper does not use external datasets, codebases, or pretrained models.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.

- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [N/A]

    Justification: No new assets are relased.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: No human subjects are involved.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: The paper does not involve human subjects research.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

Answer: [N/A]

Justification: LLMs are not part of the scientific methodology or theoretical contributions of this work.

Guidelines:

- The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.