Better and Simpler Fingerprinting Lower Bounds for Differentially Private Estimation

Shyam Narayanan*

Abstract

We provide improved lower bounds for two well-known high-dimensional private estimation tasks. First, we prove that for estimating the covariance of a Gaussian up to spectral error α with approximate differential privacy, one needs $\tilde{\Omega}\left(\frac{d^{3/2}}{\alpha\varepsilon} + \frac{d}{\alpha^2}\right)$ samples for any $\alpha \leq O(1)$, which is tight up to logarithmic factors [20, 7]. This improves over previous work [21] which established this for $\alpha \leq O\left(\frac{1}{\sqrt{d}}\right)$, and is also significantly simpler. Next, we prove that for estimating the mean of a heavy-tailed distribution with bounded kth moments with approximate differential privacy, one needs $\tilde{\Omega}\left(\frac{d}{\alpha^{k/(k-1)}\varepsilon} + \frac{d}{\alpha^2}\right)$ samples. This matches known upper bounds [23] and improves over their lower bound [5, 23] which only holds for pure differential privacy.

1 Introduction

Mean and covariance estimation are two of the most fundamental tasks in algorithmic statistics. Simply put, the goal of these tasks, respectively, are: given i.i.d. samples $X_1, \ldots, X_n$ from an unknown distribution $\mathcal{D}$, can we estimate the mean (resp., covariance) of the distribution? This question is especially worthy of investigation for data in high-dimensional Euclidean space, as this setting not only captures many real-world data problems, but also has led to numerous theoretically and practically interesting algorithms.

From the perspective of differential privacy, algorithmic statistics has enjoyed significant work over the past several years, with numerous papers studying differentially private mean [25, 20, 8, 23, 1, 28, 6, 27, 19, 17, 30, 35, 11, 18, 13, 7] and covariance [25, 3, 20, 8, 1, 29, 22, 26, 4, 35, 21, 12, 18, 2] estimation in high dimensions. Much of this work has focused on the setting where the samples are drawn i.i.d. from a Gaussian distribution. This has led to optimal sample complexity bounds for estimating both identity-covariance Gaussian and arbitrary Gaussians [20, 1, 21] in total variation distance, as well as matching polynomial-time algorithms [18]. Recently, there has also been work on "covariance-aware mean estimation", where one wishes to understand the sample complexity of estimating the mean of an unknown-covariance Gaussian, and has led to optimal sample complexity bounds [6] and nearly matching efficient algorithms [13, 7]. Other problems that have been studied include private mean estimation for heavy tailed distributions [20, 17] and private mean/covariance estimation for arbitrary bounded data [3, 27, 19, 30, 12].

Despite this large body of work, we still do not have a full understanding of private estimation for several problems. One such problem is heavy tailed mean estimation with bounded kth moments. Namely, we are promised that for some fixed constant $k \geq 2$, the (high-dimensional) data comes from a distribution $\mathcal{D}$ with unknown mean μ, but with bounded kth moment around μ in every

*shyamsn@mit.edu. Supported by an NSF Graduate Fellowship and a Google Fellowship.

direction, i.e., $\mathbb{E}_{X\sim\mathcal{D}} |\langle X-\mu, v\rangle|^k \le O(1)$ for every unit vector $v \in \mathbb{R}^d$. We wish to privately learn $\hat{\mu}$ such that $\|\hat{\mu}-\mu\|_2 \le \alpha$. The second is that while we understand the complexity of private Gaussian covariance estimation up to Frobenius error (which corresponds to the notation of total variation distance), we do not yet understand complexity for estimation up to Spectral error. In Frobenius error, given samples from $\mathcal{N}(\mu,\Sigma)$, we wish to privately learn some $\hat{\Sigma}$ such that $\|\Sigma^{-1/2}\hat{\Sigma}\Sigma^{-1/2} - I\|_F \le \alpha$, whereas in spectral error, we wish to privately learn $\hat{\Sigma}$ such that $\|\Sigma^{-1/2}\hat{\Sigma}\Sigma^{-1/2} - I\|_{op} \le \alpha$, or equivalently, $(1-\alpha)\Sigma \preccurlyeq \hat{\Sigma} \preccurlyeq (1+\alpha)\Sigma$, where $\preccurlyeq$ represents the Loewner ordering.

1.1 This work

In this work, we prove optimal lower bounds for both private heavy-tailed mean estimation and private Gaussian covariance estimation in spectral norm, matching known upper bounds.

We now state our lower bounds, starting with our result on heavy-tailed mean estimation.

Theorem 1. *For some $\delta = (\frac{\alpha\cdot\varepsilon}{d})^{O(1)}$ and any $\alpha \le O(1)$, any (ε,δ)-DP algorithm that solves mean estimation up to error α for heavy-tailed distributions with bounded kth moment in d dimensions requires sample complexity*

$$n \ge \tilde{\Omega}\Bigg(\underbrace{\frac{d}{\alpha^2}}_{\text{required even for non-private algorithms}} + \frac{d}{\alpha^{k/(k-1)}\varepsilon}\Bigg).$$

Theorem 1 improves on the best-known lower bound, which had a matching sample complexity bound but only held for pure-DP algorithms [5, 23]. As pure-DP is more stringent than approximate-DP, it is more difficult to prove approximate-DP lower bounds: a matching lower bound is only known for Gaussians and when $k = 2$ [20, 21]. Moreover, it matches a known algorithm (upper bound) of [23], up to logarithmic factors in $d, \frac{1}{\alpha}, \frac{1}{\varepsilon}, \frac{1}{\delta}$. Hence, up to logarithmic factors this essentially completes the picture for private heavy-tailed mean estimation.

Next, we state our lower bound for spectral covariance estimation.

Theorem 2. *For some $\delta = (\frac{\alpha\cdot\varepsilon}{d})^{O(1)}$ and any $\alpha \le O(1)$, any (ε,δ)-DP algorithm that solves covariance estimation up to spectral error α for Gaussians in d dimensions requires sample complexity*

$$n \ge \tilde{\Omega}\Bigg(\underbrace{\frac{d}{\alpha^2}}_{\text{required even for non-private algorithms}} + \frac{d^{3/2}}{\alpha\varepsilon}\Bigg).$$

Theorem 2 improves over the best-known lower bound of [21], which had a matching sample complexity bound but only held for $\alpha \le O(\frac{1}{\sqrt{d}})$. Moreover, it matches known algorithms of [20, 7], up to logarithmic factors in $d, \frac{1}{\alpha}, \frac{1}{\varepsilon}, \frac{1}{\delta}$. Hence, up to logarithmic factors this essentially completes the picture for private Gaussian covariance estimation up to spectral error. We also remark that our proof generalizes Theorem 1.1 in [21] while being considerably simpler.

Finally, Theorem 2 also leads to improved lower bounds for private empirical covariance estimation of arbitrary bounded data. Given data points $X_1,\dots,X_n$ that are promised to lie in a d-dimensional ball of radius 1, there is an algorithm that can provide an estimate $\hat{\Sigma}$ for the empirical covariance Σ, up to error $\|\hat{\Sigma}-\Sigma\|_F \lesssim \min\left(\frac{d}{n}, \frac{d^{1/4}}{\sqrt{n}}\right)$, ignoring polynomial factors in $\varepsilon, \log\frac{1}{\delta}$ [31, 14, 12]. The best corresponding lower bound is a matching $\Omega\left(\frac{d}{n}\right)$ when $n \ge d^2$, but is a worse $\max\left(\frac{1}{\sqrt{n}}, \frac{\sqrt{d}}{n}\right)$ for $d^{1/2} \le n \le d^2$ [20, 21, 12]. Our proof of Theorem 2 can be used to improve the lower bound to be a matching $\Omega\left(\frac{d}{n}\right)$ for all $n \ge d^{3/2}$, and $\max\left(\frac{1}{n^{1/3}}, \frac{\sqrt{d}}{n}\right)$ for $d^{1/2} \le n \le d^{3/2}$.

2 Proof Overview

Fingerprinting Oveview: Our covariance estimation lower bound is based on the technique of fingerprinting, used in many other works for privacy lower bounds [10, 16, 32, 15, 33, 34, 9, 20, 21, 24]. We first describe a general approach explaining fingerprinting lower bounds.

Suppose we are trying to estimate a parameter θ that characterizes a distribution $\mathcal{D}_\theta$. (For covariance estimation, $\theta = \Sigma$ and $\mathcal{D}_\theta = \mathcal{N}(0, \Sigma)$.) We fix a (ε, δ)-DP mechanism M with input $X_1, \ldots, X_n \sim \mathcal{D}_\theta$ and with output some estimate $\hat{\theta}$. Consider drawing i.i.d. samples $X_1, \ldots, X_n \sim \mathcal{D}_\theta$ and fresh i.i.d. samples $X'_1, \ldots, X'_n \sim \mathcal{D}_\theta$, and for each index $i \in [n]$, define the statistics

$$Z_i := \langle f(X_1, \ldots, X_i, \ldots, X_n, \theta), g(X_i, \theta) \rangle \quad \text{and} \quad Z'_i := \langle f(X_1, \ldots, X'_i, \ldots, X_n, \theta), g(X_i, \theta) \rangle, \quad (1)$$

for some fixed functions f, g, where f will depend only on $M(X_1, \ldots, X_n)$ and θ. The idea is that Z'_i is the inner product of two independent quantities (since X_i is not in the set $\{X_1, \ldots, X'_i, \ldots, X_n\}$), which makes it easier to bound the mean and variance of Z'_i. Moreover, if M is a private algorithm, then the distribution of Z_i and Z'_i, even for *fixed* samples $\{X_i\}, \{X'_i\}$ and θ, are close, which means the overall distribution of Z_i and Z'_i are similar after removing the conditioning on $\{X_i\}, \{X'_i\}$ and θ. Hence, we can also bound the distribution of Z_i, and thus bound $\mathbb{E}[Z_i]$.

Conversely, we will show that if M is a reasonably accurate estimator, then $\mathbb{E}\left[\sum_{i=1}^n Z_i\right]$ will have to be large compared to our bounds on each $\mathbb{E}[Z_i]$, unless n is sufficiently large. To actually prove this, we first carefully choose the functions f and g as well as the prior distribution on the parameter θ. Then, we prove a "fingerprinting" lemma, which proves if $X_1, \ldots, X_n \sim \mathcal{D}_\theta$, then either $M(X_1, \ldots, X_n)$ is not a good estimate for θ with reasonable probability, or $\mathbb{E}\left[\sum_{i=1}^n Z_i\right]$ is large. The main technical difficulties lie in choosing the functions and distributions, and then proving the fingerprinting lemma.

Covariance Estimation Lower Bound. We will prove a stronger statement: namely, for any $\alpha \leq O(\sqrt{d})$, there exists a distribution $\mathcal{P}$ on the covariance Σ with the following two properties.

1. With very high probability, $\Sigma \sim \mathcal{P}$ has all eigenvalues $\Theta(1)$.
2. For any (ε, δ)-DP algorithm $M(X_1, \ldots, X_n)$, if $\Sigma \sim \mathcal{P}$ and $X_1, \ldots, X_n \sim \mathcal{N}(0, \Sigma)$, where $\mathbb{E}[\|M(X_1, \ldots, X_n) - \Sigma\|_F^2] \leq \alpha$, then we must have $n \geq \Omega\left(\frac{d^2}{\alpha\varepsilon}\right)$.

By setting $\alpha' = \frac{\alpha}{\sqrt{d}}$ (so $\frac{d^2}{\alpha\varepsilon} = \frac{d^{3/2}}{\alpha'\varepsilon}$), this implies that we cannot have $\|M(X_1, \ldots, X_n) - \Sigma\|_{op} \leq \alpha'$ with very high probability, and since all eigenvalues of Σ are $\Theta(1)$, this implies our desired result. This holds for any $\alpha \leq O(\sqrt{d})$, and hence for any $\alpha' \leq O(1)$.

The choices of f, g in (1) will be quite simple: we choose $f(X_1, \ldots, X_n, \Sigma) = M(X_1, \ldots, X_n) - \Sigma$ and $g(X_i) = X_i X_i^\top - \Sigma$, so

$$Z_i := \langle M(X_1, \ldots, X_n) - \Sigma, X_i X_i^\top - \Sigma \rangle \quad \text{and} \quad Z'_i = \langle M(X_1, \ldots, X'_i, \ldots, X_n) - \Sigma, X_i X_i^\top - \Sigma \rangle.$$

Using the fact that $X_i X_i^\top$ is an unbiased estimator for Σ, a simple calculation shows that $\mathbb{E}[Z'_i] = 0$. Moreover, assuming M is a reasonably good estimator of Σ, we can show $Var(Z'_i) \leq 4\alpha^2$. Next, (ε, δ)-DP will imply for reasonably small δ that $\mathbb{E}[Z_i] \leq O(\alpha\varepsilon)$ for all i. Hence, $\mathbb{E}[\sum_{i=1}^n Z_i] \leq O(n \cdot \alpha\varepsilon)$ if M is differentially private and reasonably accurate.

Next, we show a lower bound on $\mathbb{E}[\sum_{i=1}^n Z_i]$, assuming M is a sufficiently accurate estimator. This lower bound does not utilize any privacy constraints. Note that $\mathbb{E}[\sum_{i=1}^n Z_i] =$

$n \cdot \mathbb{E}[\langle M(X_1, \ldots, X_n) - \Sigma, \bar{\Sigma} - \Sigma \rangle]$, where $\bar{\Sigma} = \frac{1}{n}\sum_{i=1}^{n} X_i X_i^\top$ is the empirical covariance. So, we want to show this quantity is larger than $O(n \cdot \alpha\varepsilon)$, which contradicts our above bound, unless either $n \geq \Omega\left(\frac{d^2}{\alpha\varepsilon}\right)$ or $\|M(X_1, \ldots, X_n) - \Sigma\|_F > \alpha$ holds with reasonable probability.

We can rewrite our desired quantity as

$$\mathbb{E}\left[\sum_{i=1}^{n} Z_i\right] = n \cdot \left(\mathbb{E}\left\langle M(X_1, \ldots, X_n) - \bar{\Sigma}, \bar{\Sigma} - \Sigma\right\rangle + \mathbb{E}\left[\|\bar{\Sigma} - \Sigma\|_F^2\right]\right). \tag{2}$$

It is well-known that $\mathbb{E}\left[\|\bar{\Sigma} - \Sigma\|_F^2\right] = \Theta(\frac{d^2}{n})$, and we can also write

$$\begin{aligned}\left|\mathbb{E}\left\langle M(X_1, \ldots, X_n) - \bar{\Sigma}, \bar{\Sigma} - \Sigma\right\rangle\right| &= \left|\mathbb{E}\left\langle M(X_1, \ldots, X_n) - \bar{\Sigma}, \bar{\Sigma} - \mathbb{E}[\Sigma|X_1, \ldots, X_n]\right\rangle\right| \\ &\leq \sqrt{\mathbb{E}\|M(X_1, \ldots, X_n) - \bar{\Sigma}\|_F^2 \cdot \mathbb{E}\|\bar{\Sigma} - \mathbb{E}[\Sigma|X_1, \ldots, X_n]\|_F^2}. \end{aligned} \tag{3}$$

Assuming that $M(X_1, \ldots, X_n)$ is a good estimator of Σ, it will also be a good estimator of $\bar{\Sigma}$, and $\mathbb{E}\|M(X_1, \ldots, X_n) - \bar{\Sigma}\|_F^2 \leq \alpha^2$. We have avoided discussing the prior distribution of Σ, but to bound $\mathbb{E}\|\bar{\Sigma} - \mathbb{E}[\Sigma|X_1, \ldots, X_n]\|_F^2$, we need to define the prior. The prior that we choose will be an *Inverse Wishart* distribution, which is known to be the classic *conjugate prior* of the Multivariate Gaussian. What this means is that if the prior distribution of Σ follows an Inverse Wishart distribution and we sample $X_1, \ldots, X_n \sim \mathcal{N}(0, \Sigma)$, the posterior distribution of Σ given $X_1, \ldots, X_n$ also follows an Inverse Wishart distribution (with a different parameter setting). This will make it easy to compute $\mathbb{E}[\Sigma|X_1, \ldots, X_n]$. We will choose Σ to be a (scaled) Inverse Wishart distribution with $C \cdot d$ degrees of freedom for a sufficiently large constant C. With a proper scaling, all of the eigenvalues of Σ will be between 0.5 and 1.5, and the posterior distribution will have expectation $\left(1 + O(\frac{d}{n})\right) \cdot \bar{\Sigma}$. From this, it is not hard to bound $\mathbb{E}\|\bar{\Sigma} - \mathbb{E}[\Sigma|X_1, \ldots, X_n]\|_F^2 \leq O(\frac{d^2}{n^2}) \cdot \mathbb{E}[\|\bar{\Sigma}\|_F^2] = O\left(\frac{d^3}{n^2}\right)$. Combining this with Equations (2) and (3) and our bound $\mathbb{E}\|M(X_1, \ldots, X_n) - \bar{\Sigma}\|_F^2 \leq \alpha^2$, this implies that

$$\mathbb{E}\left[\sum_{i=1}^{n} Z_i\right] = n \cdot \left[\Theta\left(\frac{d^2}{n}\right) \pm O\left(\sqrt{\alpha^2 \cdot \frac{d^3}{n^2}}\right)\right].$$

As long as $\alpha \leq c\sqrt{d}$ for some small constant c, this implies $\mathbb{E}[\sum_{i=1}^{n} Z_i] \geq \Omega(d^2)$. As we already explained why $\mathbb{E}[\sum_{i=1}^{n} Z_i] \leq O(n \cdot \alpha\varepsilon)$, this implies that as long as $\alpha \leq c\sqrt{d}$, any (ε, δ)-DP algorithm that can estimate Σ up to Frobenius error α needs $O(n \cdot \alpha\varepsilon) \geq \Omega(d^2)$, or $n \geq \Omega\left(\frac{d^2}{\alpha\varepsilon}\right)$.

Heavy Tailed Mean Estimation: This result will follow from a simple application of the fact that privately learning μ up to error α requires $\Omega(\frac{d}{\alpha\varepsilon})$ samples from $\mathcal{N}(\mu, I)$ [20]. Specifically, we will draw a distribution that, with probability $\alpha^{k/(k-1)}$ is drawn as $\mathcal{N}(\mu', \alpha^{-2/(k-1)} \cdot I)$ for some unknown μ' with $\|\mu'\| \leq O(\alpha^{-1/(k-1)})$. It is straightforward to check that this distribution has bounded kth moment, and the actual mean, $\mu = \alpha^{k/(k-1)} \cdot \mu'$, has norm $O(\alpha)$. However, to learn μ to error α, one must learn μ' up to error exactly $\alpha^{-1/(k-1)}$, not just $O(\alpha^{-1/(k-1)})$. A minor modification of the lower bound in [20] can show that learning μ' is essentially equivalent to learning the mean of identity covariance Gaussian up to error 1. This requires $\Omega\left(\frac{d}{\varepsilon}\right)$ samples. However, because only an $\alpha^{k/(k-1)}$ fraction of the points were actually from the Gaussian, we need $\Omega\left(\frac{d}{\varepsilon \cdot \alpha^{k/(k-1)}}\right)$ samples in total. This argument can be made formal by converting an instance of Gaussian estimation into this distribution by padding.

References

[1] Ishaq Aden-Ali, Hassan Ashtiani, and Gautam Kamath. On the sample complexity of privately learning unbounded high-dimensional gaussians. In *Proceedings of the 32nd International Conference on Algorithmic Learning Theory*, ALT '21, pages 185–216. JMLR, Inc., 2021.

[2] Daniel Alabi, Pravesh K Kothari, Pranay Tankala, Prayaag Venkat, and Fred Zhang. Privately estimating a Gaussian: Efficient, robust and optimal. In *Proceedings of the 55th Annual ACM Symposium on the Theory of Computing*, STOC '23, New York, NY, USA, 2023. ACM.

[3] Kareem Amin, Travis Dick, Alex Kulesza, Andres Munoz, and Sergei Vassilvitskii. Differentially private covariance estimation. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.

[4] Hassan Ashtiani and Christopher Liaw. Private and polynomial time algorithms for learning gaussians and beyond. In *Conference on Learning Theory*, pages 1075–1076. PMLR, 2022.

[5] Rina Foygel Barber and John C. Duchi. Privacy and statistical risk: Formalisms and minimax bounds. *CoRR*, abs/1412.4451, 2014.

[6] Gavin Brown, Marco Gaboardi, Adam D. Smith, Jonathan R. Ullman, and Lydia Zakynthinou. Covariance-aware private mean estimation without private covariance estimation. In *Advances in Neural Information Processing Systems*, pages 7950–7964, 2021.

[7] Gavin Brown, Samuel B Hopkins, and Adam Smith. Fast, sample-efficient, affine-invariant private mean and covariance estimation for subgaussian distributions. *arXiv preprint arXiv:2301.12250*, 2023.

[8] Mark Bun, Gautam Kamath, Thomas Steinke, and Zhiwei Steven Wu. Private hypothesis selection. *IEEE Trans. Inf. Theory*, 67(3):1981–2000, 2021.

[9] Mark Bun, Thomas Steinke, and Jonathan R. Ullman. Make up your mind: The price of online queries in differential privacy. *J. Priv. Confidentiality*, 9(1), 2019.

[10] Mark Bun, Jonathan R. Ullman, and Salil P. Vadhan. Fingerprinting codes and the price of approximate differential privacy. In *Symposium on Theory of Computing*, pages 1–10. ACM, 2014.

[11] Rachel Cummings, Vitaly Feldman, Audra McMillan, and Kunal Talwar. Mean estimation with user-level privacy under data heterogeneity. In *Advances in Neural Information Processing Systems*, volume 35, pages 29139–29151, 2022.

[12] Wei Dong, Yuting Liang, and Ke Yi. Differentially private covariance revisited. In *Advances in Neural Information Processing Systems*, volume 35, pages 850–861, 2022.

[13] John Duchi, Saminul Haque, and Rohith Kuditipudi. A fast algorithm for adaptive private mean estimation. *arXiv preprint arXiv:2301.07078*, 2023.

[14] Cynthia Dwork, Aleksandar Nikolov, and Kunal Talwar. Efficient algorithms for privately releasing marginals via convex relaxations. *Discret. Comput. Geom.*, 53(3):650–673, 2015.

[15] Cynthia Dwork, Adam D. Smith, Thomas Steinke, Jonathan R. Ullman, and Salil P. Vadhan. Robust traceability from trace amounts. In *IEEE 56th Annual Symposium on Foundations of Computer Science*, FOCS '15, pages 650–669. IEEE Computer Society, 2015.

[16] Moritz Hardt and Jonathan R. Ullman. Preventing false discovery in interactive data analysis is hard. In *55th IEEE Annual Symposium on Foundations of Computer Science*, FOCS '14, pages 454–463. IEEE Computer Society, 2014.

[17] Samuel B Hopkins, Gautam Kamath, and Mahbod Majid. Efficient mean estimation with pure differential privacy via a sum-of-squares exponential mechanism. In *Proceedings of the 54th Annual ACM Symposium on the Theory of Computing*, STOC '22, New York, NY, USA, 2022. ACM.

[18] Samuel B. Hopkins, Gautam Kamath, Mahbod Majid, and Shyam Narayanan. Robustness implies privacy in statistical estimation. In *Proceedings of the 55th Annual ACM Symposium on the Theory of Computing*, STOC '23, New York, NY, USA, 2023. ACM.

[19] Ziyue Huang, Yuting Liang, and Ke Yi. Instance-optimal mean estimation under differential privacy. In *Advances in Neural Information Processing Systems*, pages 25993–26004, 2021.

[20] Gautam Kamath, Jerry Li, Vikrant Singhal, and Jonathan Ullman. Privately learning high-dimensional distributions. In *Proceedings of the 32nd Annual Conference on Learning Theory*, COLT '19, pages 1853–1902, 2019.

[21] Gautam Kamath, Argyris Mouzakis, and Vikrant Singhal. New lower bounds for private estimation and a generalized fingerprinting lemma. In *Advances in Neural Information Processing Systems 35*, NeurIPS '22, 2022.

[22] Gautam Kamath, Argyris Mouzakis, Vikrant Singhal, Thomas Steinke, and Jonathan Ullman. A private and computationally-efficient estimator for unbounded gaussians. In *Proceedings of the 35th Annual Conference on Learning Theory*, COLT '22, pages 544–572, 2022.

[23] Gautam Kamath, Vikrant Singhal, and Jonathan Ullman. Private mean estimation of heavy-tailed distributions. In *Proceedings of the 33rd Annual Conference on Learning Theory*, COLT '20, pages 2204–2235, 2020.

[24] Gautam Kamath and Jonathan Ullman. A primer on private statistics. *CoRR*, abs/2005.00010, 2020.

[25] Vishesh Karwa and Salil Vadhan. Finite sample differentially private confidence intervals. In *Proceedings of the 9th Conference on Innovations in Theoretical Computer Science*, ITCS '18, pages 44:1–44:9, Dagstuhl, Germany, 2018. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[26] Pravesh K Kothari, Pasin Manurangsi, and Ameya Velingker. Private robust estimation by stabilizing convex relaxations. In *Proceedings of the 35th Annual Conference on Learning Theory*, COLT '22, pages 723–777, 2022.

[27] Daniel Levy, Ziteng Sun, Kareem Amin, Satyen Kale, Alex Kulesza, Mehryar Mohri, and Ananda Theertha Suresh. Learning with user-level privacy. In *Advances in Neural Information Processing Systems*, pages 12466–12479, 2021.

[28] Xiyang Liu, Weihao Kong, Sham Kakade, and Sewoong Oh. Robust and differentially private mean estimation. In *Advances in Neural Information Processing Systems 34*, NeurIPS '21. Curran Associates, Inc., 2021.

[29] Xiyang Liu, Weihao Kong, and Sewoong Oh. Differential privacy and robust statistics in high dimensions. In *Proceedings of the 35th Annual Conference on Learning Theory*, COLT '22, pages 1167–1246, 2022.

[30] Shyam Narayanan, Vahab Mirrokni, and Hossein Esfandiari. Tight and robust private mean estimation with few users. In *International Conference on Machine Learning*, pages 16383–16412. PMLR, 2022.

[31] Aleksandar Nikolov, Kunal Talwar, and Li Zhang. The geometry of differential privacy: the sparse and approximate cases. In *Symposium on Theory of Computing Conference, STOC'13*, pages 351–360. ACM, 2013.

[32] Thomas Steinke and Jonathan R. Ullman. Interactive fingerprinting codes and the hardness of preventing false discovery. In *Proceedings of The 28th Conference on Learning Theory*, volume 40 of *COLT '15*, pages 1588–1628. JMLR.org, 2015.

[33] Thomas Steinke and Jonathan R. Ullman. Between pure and approximate differential privacy. *J. Priv. Confidentiality*, 7(2), 2016.

[34] Thomas Steinke and Jonathan R. Ullman. Tight lower bounds for differentially private selection. In *58th IEEE Annual Symposium on Foundations of Computer Science*, FOCS '17, pages 552–563. IEEE Computer Society, 2017.

[35] Eliad Tsfadia, Edith Cohen, Haim Kaplan, Yishay Mansour, and Uri Stemmer. Friendlycore: Practical differentially private aggregation. In *Proceedings of the 39th International Conference on Machine Learning*, ICML '22, pages 21828–21863. JMLR, Inc., 2022.