# New Lower Bounds for Private Estimation and a Generalized Fingerprinting Lemma

**Gautam Kamath**
Cheriton School of Computer Science
University of Waterloo
Waterloo, ON N2L 3G1, Canada
g@csail.mit.edu

**Argyris Mouzakis**
Cheriton School of Computer Science
University of Waterloo
Waterloo, ON N2L 3G1, Canada
amouzaki@uwaterloo.ca

**Vikrant Singhal**
Cheriton School of Computer Science
University of Waterloo
Waterloo, ON N2L 3G1, Canada
vikrant.singhal@uwaterloo.ca

## Abstract

We prove new lower bounds for statistical estimation tasks under the constraint of $(\varepsilon, \delta)$-differential privacy. First, we provide tight lower bounds for private covariance estimation of Gaussian distributions. We show that estimating the covariance matrix in Frobenius norm requires $\Omega(d^2)$ samples, and in spectral norm requires $\Omega(d^{3/2})$ samples, both matching upper bounds up to logarithmic factors. We prove these bounds via our main technical contribution, a broad generalization of the fingerprinting method [BUV14] to exponential families. Additionally, using the private Assouad method of Acharya, Sun, and Zhang [ASZ21], we show a tight $\Omega(d/(\alpha^2\varepsilon))$ lower bound for estimating the mean of a distribution with bounded covariance to $\alpha$-error in $\ell_2$-distance. Prior known lower bounds for all these problems were either polynomially weaker or held under the stricter condition of $(\varepsilon, 0)$-differential privacy.

## 1 Introduction

The last several years have seen a surge of interest in algorithms for statistical estimation under the constraint of *Differential Privacy* (DP) [DMNS06]. We now have a rich algorithmic toolbox for private estimation of mean, covariance, and entire distributions, in a variety of settings.

While the community has enjoyed much success designing *algorithms* for estimation tasks, *lower bounds* have been much harder to come by, and consequently, the existing literature lacks a rigorous understanding of some core problems. Lower bounds against the strong privacy constraint of *pure* DP (i.e., $(\varepsilon, 0)$-DP) are generally not too challenging to prove, most often relying upon the well-known packing technique [HT10]. Such packing lower bounds are optimal in a broad range of settings (see, e.g., [BKSW19]), and thus are frequently effective at providing tight minimax sample complexity lower bounds.

However, pure DP is a very strong privacy constraint: in most practical circumstances, it suffices to only require the weaker privacy notion of *approximate* DP (i.e., $(\varepsilon, \delta)$-DP) [DKM$^+$06], in which case packing lower bounds are no longer applicable. Proving lower bounds under approximate DP is much more challenging, and accordingly, the state of affairs is less satisfying. There exist only a couple of techniques which apply in this setting, including the fingerprinting method [BUV14] and

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

differentially private analogues of Le Cam's method and Assouad's lemma [ASZ18, ASZ21]. These techniques can be quite brittle and thus unable to prove lower bounds for some fairly basic settings. For example, both techniques generally require that the underlying distribution has independent marginals and are thus ineffective at proving tight lower bounds for problems involving correlations such as private covariance estimation. Furthermore, the fingerprinting method often needs the distribution's marginals to have a precise functional form, such as Gaussian or Bernoulli. These restrictions make the approaches somewhat brittle for proving lower bounds which differ too much from existing results. Consequently, we lack approximate DP lower bounds for some fundamental settings which seem to qualitatively differ, including Gaussian covariance estimation and mean estimation of heavy-tailed distributions.

### 1.1 Our Results

We circumvent these barriers and fill a number of gaps in the literature by providing tight lower bounds for several estimation tasks under the constraint of approximate differential privacy. Specifically, we address problems including covariance estimation for Gaussians and mean estimation with heavy-tailed data. Our main technical tool for covariance estimation is a novel generalization of the fingerprinting method to exponential families, while we prove our mean estimation lower bound via DP Assouad's lemma.

Our first result is a lower bound for covariance estimation in Frobenius norm.[1]

**Theorem 1.1** (Informal)**.** *For any $\alpha = \mathcal{O}(1)$, any $(\varepsilon, \delta)$-DP mechanism with $\varepsilon, \delta \in [0, 1]$, and $\delta \leq \mathcal{O}(\min\left\{1/n, d^2/(n\log(n/d^2))\right\})$ that takes $n$ samples from an arbitrary $d$-dimensional Gaussian $\mathcal{N}(0, \Sigma)$ and outputs $M(X)$ satisfying $\mathop{\mathbb{E}}_{X,M}\left[\|M(X) - \Sigma\|_F^2\right] \leq \alpha^2$ requires $n \geq \Omega\left(\frac{d^2}{\alpha\varepsilon}\right)$.*

We point out that the above lower bound also implies a lower bound for density estimation of Gaussians with known mean and unknown covariance (see Theorem 1.1 of [DMR18]). This nearly matches the $\widetilde{\mathcal{O}}\left(d^2\right)$ upper bound of [KLSU19] up to polylogarithmic factors.[2] Previously, $\Omega\left(d^2\right)$ lower bounds for Gaussian covariance estimation were only known under the more restrictive constraint of pure differential privacy [KLSU19, BKSW19].

Our result for Gaussian covariance estimation in Frobenius norm implies the following result for spectral estimation.

**Theorem 1.2** (Informal)**.** *For any $\alpha = \mathcal{O}(1/\sqrt{d})$, any $(\varepsilon, \delta)$-DP mechanism with $\varepsilon, \delta \in [0, 1]$, and $\delta \leq \mathcal{O}(\min\left\{1/n, d^2/(n\log(n/d^2))\right\})$ that takes $n$ samples from an arbitrary $d$-dimensional Gaussian $\mathcal{N}(0, \Sigma)$ and outputs $M(X)$ satisfying $\mathop{\mathbb{E}}_{X,M}\left[\|M(X) - \Sigma\|_2^2\right] \leq \alpha^2$ requires $n \geq \Omega\left(\frac{d^{1.5}}{\alpha\varepsilon}\right)$.*

In Theorem 3.8 of [KLSU19], the authors give guarantees for an individual iteration of a private recursive preconditioner. In each iteration, the data is rescaled, resulting in constant-factor progress in privately reducing the condition number of the underlying covariance matrix to a constant. Identifying the preconditioning matrix necessitates performing spectral estimation of the underlying covariance matrix. Thus, the sample complexity upper bound of $\widetilde{\mathcal{O}}(d^{3/2})$ which comes up in the guarantees of the preconditioning steps is also an upper bound for spectral estimation for $\alpha = \mathcal{O}(1)$. This implies that our lower bound is tight up to polylogarithmic factors in the regime $\alpha = \mathcal{O}(1/\sqrt{d})$.

Note that the sample complexity of the same problem sans privacy constraints is $\Theta(d)$, and thus the cost of privacy is polynomial in the dimension $d$. This is in contrast to Gaussian mean estimation

[1] Recall that the Mahalanobis norm of $M$ with respect to $\Sigma$ is $\|M\|_\Sigma = \left\|\Sigma^{-1/2} M \Sigma^{-1/2}\right\|_F$. For the lower bound constructions we consider, $\Omega(1)\,\mathbb{I} \preceq \Sigma \preceq \mathcal{O}(1)\,\mathbb{I}$. In this regime, the Mahalanobis and Frobenius distances are equivalent up to a constant factor. Thus, we will often use them interchangeably and our lower bound for Frobenius estimation also implies a lower bound for Mahalanobis estimation.

[2] We note that all our lower bounds in this work are stated in terms of mean squared error, while most of the upper bounds guarantee error at most $\alpha$ with probability $1 - \beta$ for some $\beta > 0$. In many natural cases (such as when the estimator's range is naturally bounded), lower bounds against MSE can be converted to constant probability statements via a boosting argument (outlined, e.g., in the proof of Theorem 6.1 of [KLSU19]). The details of such a conversion generally standard when applicable and we do not discuss it further here.

in $\ell_2$-norm and covariance estimation in Frobenius norm, which maintain their non-private sample complexities of $\mathcal{O}(d)$ and $\mathcal{O}(d^2)$, respectively.

$\Omega(d^{3/2})$-sample lower bounds for approximate DP covariance estimation in spectral norm were previously known for *worst-case* distributions [DTTZ14]. Our work relaxes this assumption significantly to the Gaussian case while maintaining the same lower bound. A priori, it was not clear that such a result was even true, as there are frequently gaps between the sample complexity of private estimation for worst-case versus well-behaved distributions for even basic settings, including distributions over the hypercube (see, e.g., Remark 6.4 of [BKSW19]).

This lower bound has implications for estimation tasks with scale-dependent error. For example, consider Gaussian mean estimation in Mahalanobis distance. A natural way to approach this problem is to first estimate the covariance matrix spectrally, and then estimate the mean after an appropriate rescaling (see, e.g., [KLSU19]). Our lower bound shows that any such approach must incur the $\Omega(d^{3/2})$ cost of spectral estimation, which is greater than the $\mathcal{O}(d)$ minimax sample complexity of the problem. Indeed, some recent works [BGS+21, LKO21] manage to circumvent this roadblock by adopting more direct (but computationally inefficient) methods.

Finally, we prove new lower bounds for mean estimation of distributions with bounded second moments.[3]

**Theorem 1.3** (Informal)**.** *For any $\alpha \leq 1$, any $(\varepsilon, \delta)$-DP mechanism with $\delta \leq \varepsilon$ that takes $n$ samples from an arbitrary distribution over $\mathbb{R}^d$ with second moments bounded by $1$ and outputs $M(X)$ satisfying $\mathop{\mathbb{E}}_{X,M}\left[\|M(X) - \mu\|_2^2\right] \leq \alpha^2$ requires $n \geq \Omega\left(\frac{d}{\alpha^2 \varepsilon}\right)$.*

Similar $\Theta(d/(\alpha^2\varepsilon))$ upper and lower bounds were previously known for distributions with bounded second moments under the stricter constraint of pure differential privacy [BD14, KSU20, HKM21]. Also, a similar bound is shown in [KLZ21] for concentrated DP, but under a weaker moment assumption which involves only coordinate-wise projections. Our result shows that the same bound holds under the weaker constraint of approximate DP for distributions satisfying the stronger moment assumption, and thus no savings can be obtained by relaxing the privacy notion. In contrast to our other results, we show this lower bound via the DP Assouad method of [ASZ21], thus demonstrating the promise of different approaches for proving lower bounds for differentially private estimation.

### 1.2 Related Work

The fingerprinting technique for proving lower bounds was introduced by Bun, Ullman, and Vadhan [BUV14], which relied on the existence of *fingerprinting codes* [BS98, Tar08]. Since then, the technique has been significantly simplified and refined in a number of ways [SU17a, DSS+15, BSU17, SU17b, KLSU19, CWZ19], including the removal of fingerprinting codes and distilling the main technical component into the *fingerprinting lemma*. Beyond the aforementioned settings of mean estimation of Gaussians and distributions over the hypercube [BUV14, SU15, DSS+15, KLSU19], fingerprinting lower bounds have also been applied in settings including private empirical risk minimization [BST14] and private spectral estimation [DTTZ14]. Differentially private analogues of Fano, Le Cam, and Assouad were first considered in the local DP setting [DJW13], and more recently under the constraint of central DP [ASZ18, ASZ21].

On the upper bounds side, the most relevant body of work is that which focuses on mean and covariance estimation for distributions satisfying a moment bound, see, e.g., [BD14, KV18, BKSW19, BS19, KLSU19, CWZ19, KSU20, WXDX20, DFM+20, BDKU20, AAK21, BGS+21, KLZ21, HLY21, KMS+21, AL21]. Our lower bounds nearly match these upper bounds for Gaussian covariance estimation [KLSU19] and heavy-tailed mean estimation [KSU20, HKM21]. Previously, these bounds were conjectured to be optimal, albeit on somewhat shaky grounds including evidence from lower bounds proven under pure DP or non-rigorous connections with lower bounds in related settings. Our results rigorously prove optimality and confirm these conjectures. Other works study private statistical estimation in different and more general settings, including mixtures of Gaussians [KSSU19, AAL21], graphical models [ZKKW20], discrete distributions [DHS15], and median estimation [AMB19, TVGZ20]. Some recent directions involve guaranteeing user-level

[3] The definition we use for bounded moments involves looking at the central moments of single-dimensional projections of our random vectors.

privacy [LSY+20, LSA+21], or a combination of local and central DP for different users [ADK19]. See [KU20] for further coverage of DP statistical estimation.

### 1.3 Organization

We start by providing our generalized fingerprinting technique in Section 2. Then we show its novel application for proving the lower bound for covariance estimation of Gaussians in Mahalanobis norm, followed by the lower bound for spectral norm estimation in Section 3. Finally, we sketch the proof of the lower bound for mean-estimation of heavy-tailed distributions in Section 4.

## 2 Fingerprinting Proofs for Exponential Families

This section is split into three parts. Section 2.1 provides an overview of existing fingerprinting proofs and motivation for our result, and introduces all the necessary notation. Then, in Section 2.2, our new fingerprinting lemma is stated. Finally, in Section 2.3, we give a general recipe of applying it.

### 2.1 Overview of Fingerprinting Proofs and Intuition for Our Result

At the heart of fingerprinting proofs lies the trade-off between accuracy and privacy. For simplicity, assume we have a univariate data-generating distribution and coming from a family parameterized by its mean $\mu$, which we want to estimate. In contrast to minimax lower bounds in classical statistics (which are commonly obtained by constructing hard testing instances), fingerprinting proofs assume the existence of a subset of the domain of $\mu$ (commonly an interval $[\pm R]$) and that a value is drawn $\mu \sim \mathcal{U}([\pm R])$ . Then, $n$ independent samples $X_1, \ldots, X_n$ are drawn from $\mathcal{D}_\mu$. Now, given an estimator $f : \mathbb{R}^n \to \{\pm R\}$ for $\mu$, we consider the quantity $Z_i \coloneqq c(R, \mu)\left(f\left(X\right) - \mu\right)\left(X_i - \mu\right), \forall i \in [n]$, which measures the correlation between the output of $f$ and the $i$-th sample ($c(R, \mu) > 0$ is a scaling factor that exists for purely technical reasons). Defining $Z \coloneqq \sum_{i \in [n]} Z_i$, all proofs first show a lower bound $L$ on $\mathbb{E}[Z + \left(f\left(X\right) - \mu\right)^2]$, implying that our estimator cannot at the same time be accurate (with respect to the Mean-Squared-Error (MSE)) and exhibit low correlation with individual samples. This component of the proof is traditionally called the *fingerprinting lemma*.

Privacy comes up in the second step of the proof, where we upper-bound the correlation terms $\mathbb{E}\left[Z_i\right]$.

We assume that $f$ is a private mechanism with MSE at most $L/2$. Due to privacy, all correlation terms $Z_i$ should be upper-bounded by the same $U = U\left(\varepsilon, \delta\right)$. Applying this for each sample leads to an inequality of the form $nU \geq L/2 \iff n \geq L/(2U)$, so the only thing that remains is to identify the appropriate parameter range for $\delta$ to prove the desired bound.

Our work is motivated by a fundamental weakness of existing fingerprinting proofs. To prove lower bounds for high-dimensional distributions via fingerprinting, past works have assumed that the underlying distribution is a product distribution. This made it possible to work component-wise using techniques for single-dimensional distributions. This assumption is somewhat restrictive and makes it difficult to generalize the approach to settings with a richer correlation structure.

Our main technical contribution is providing a broad generalization of the celebrated fingerprinting method to exponential families (Definition A.5). Distributions belonging in the same exponential family are parameterized by their *natural parameter vector* $\eta$, instead of their mean $\mu$, for which it is uncommon to have a general closed-form expression. Given $n$ samples $X_1, \ldots, X_n$ from $p_\eta$, it suffices to know the values of the *sufficient statistics* $T\left(X_1\right), \ldots, T\left(X_n\right)$ instead of the samples themselves in order to estimate $\eta$ (see, e.g., [Jor10]). We denote the mean and covariance of $T$ by $\mu_T$ and $\Sigma_T$, respectively.

Based on the above remarks, our fingerprinting lemma provides lower bounds for estimating $\eta$ instead of $\mu$. We assume the existence of vectors $\eta^{(1)}, \eta^{(2)}$ with $\eta_j^{(1)} \leq \eta_j^{(2)}, \forall j \in [k]$, and define the intervals $I_j \coloneqq \left[\eta_j^{(1)}, \eta_j^{(2)}\right]$ for each coordinate $j \in [k]$. Assuming that the hyperrectangle $\bigotimes_{j \in [k]} I_j$ is a subset of $\mathcal{H}$, we draw a natural parameter vector $\eta$ uniformly at random from it. Observe that $I_j$ may not be centered around the origin, which poses a minor technical difficulty. To deal with this, we assume that our estimator does not estimate $\eta$ itself, but rather the deviation of $\eta$ from the *midpoint* $m \coloneqq (\eta^{(1)} + \eta^{(2)})/2$. The two problems are equivalent because any estimator

$f : S^n \to \bigotimes_{j\in[k]}[-(\eta_j^{(2)} - \eta_j^{(1)})/2, (\eta_j^{(2)} - \eta_j^{(1)})/2]$ for the latter problem can be transformed to an estimator for the former, simply by adding $m$ to it (and vice-versa). Defining $R$ to be the *width vector* $\eta^{(2)} - \eta^{(1)}$, the correlation between the estimate for $\eta_j$ and the sample $X_i$ becomes $Z_i^j \coloneqq [R_j^2/4 - (\eta_j - m_j)^2][f_j(X) - (\eta_j - m_j)](T_j(X_i) - \mu_{T,j}), \forall i \in [n], j \in [k]$.

Observe that now, $Z_i^j$ measures the correlation between $f$ and the sufficient statistics of the samples. This is in contrast to older fingerprinting proofs, in which the $Z_i^j$ measure correlation between $f$ and the samples themselves. We consider this to be natural in the context of exponential families. Indeed, we know that, if we want to estimate the natural parameter vector of a distribution that belongs to an exponential family given a dataset $X = (X_1, \ldots, X_n)$, it suffices to know the values $T(X_1), \ldots, T(X_n)$ instead of the samples themselves (see [Jor10]). Thus, despite $f$ receiving $X$, correlation ends up being measured with respect to $T(X_i), i \in [n]$.

In addition to the above, we introduce three more correlation terms:

$$Z_i \coloneqq \sum_{j\in[k]} Z_i^j = \sum_{j\in[k]} [R_j^2/4 - (\eta_j - m_j)^2]\,[f_j(X) - (\eta_j - m_j)]\,(T_j(X_i) - \mu_{T,j}),$$

$$Z^j \coloneqq \sum_{i\in[n]} Z_i^j = [R_j^2/4 - (\eta_j - m_j)^2]\,[f_j(X) - (\eta_j - m_j)] \sum_{i\in[n]} (T_j(X_i) - \mu_{T,j}),$$

$$Z \coloneqq \sum_{i\in[n]} Z_i = \sum_{j\in[k]} Z^j.$$

We motivate each of the above. $Z_i$ describes the correlation between $X_i$ and our estimates for $\eta$. Moreover, $Z^j$ describes the correlation between our estimate for $\eta_j$ and $X$. Finally, $Z$ describes the total correlation between the output of $f$ and $X$. Our fingerprinting lemma for exponential families will involve upper and lower-bounding the quantity $\mathop{\mathbb{E}}_{\eta,X}\left[Z + \|f(X) - (\eta - m)\|_2^2\right]$.

### 2.2 The Main Lemma and Our Approach to Lower Bounds

We prove our new fingerprinting lemma in two parts. First, we assume we have a distribution $p_\eta$ that belongs to an exponential family and we want to estimate $\eta_j$ whereas $\eta_{-j}$ is fixed (Lemma 2.1). This setting is similar to the ones considered in [BSU17, KLSU19], which feature similar lemmas. We then leverage Lemma 2.1 to get a result when the goal is to estimate all of $\eta$ (Lemma 2.2).

**Lemma 2.1.** *Let $p_\eta$ be a distribution over $S \subseteq \mathbb{R}^d$ belonging to an exponential family $\mathcal{E}(T, h)$ with natural parameter vector $\eta \in \mathcal{H} \subseteq \mathbb{R}^k$. We assume that $I_j$ is non-degenerate, and that $\eta_j$ is randomly generated by drawing $\eta_j \sim \mathcal{U}(I_j)$, whereas $\eta_{-j}$ is fixed. Then, for any function $f_j \colon S^n \to [\pm R_j/2]$ that takes as input a dataset $X \coloneqq (X_1, \ldots, X_n) \sim p_\eta^{\bigotimes n}$, we have:*

$$\mathop{\mathbb{E}}_{\eta_j,X}\left[Z^j + [f_j(X) - (\eta_j - m_j)]^2\right] \geq R_j^2/12.$$

*Proof.* Let $g_j \colon \bigotimes_{j\in[k]} I_j \to [\pm R_j/2]$ with $g_j(\eta) \coloneqq \mathop{\mathbb{E}}_{X}[f_j(X)]$. $g_j$ can equivalently be written as:

$$g_j(\eta) = \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} \left(\prod_{i\in[n]} p_\eta(x_i)\right) f_j(x_1, \ldots, x_n)\, dx_1 \ldots dx_n.$$

Observe that:

$$\left(\prod_{i\in[n]} p_\eta(x_i)\right) = \left(\prod_{i\in[n]} h(x_i)\right) \exp\left(\left\langle \eta, \sum_{i\in[n]} T(x_i)\right\rangle - nZ(\eta)\right)$$

$$\implies \frac{\partial}{\partial\eta_j}\left(\prod_{i\in[n]} p_\eta(x_i)\right) = \left(\prod_{i\in[n]} p_\eta(x_i)\right) \sum_{i\in[n]} \left(T_j(x_i) - \frac{\partial}{\partial\eta_j}(Z(\eta))\right)$$

$$= \left(\prod_{i\in[n]} p_\eta(x_i)\right) \sum_{i\in[n]} (T_j(x_i) - \mu_{T,j}).$$

Calculating the partial derivative of $g_j$ with respect to $\eta_j$ and changing the order of differentiation and integration, we get:

$$\begin{aligned}
\frac{\partial}{\partial \eta_j}\left(g_j\left(\eta\right)\right) &= \int_{\mathbb{R}^d} \dots \int_{\mathbb{R}^d} \left(\prod_{i\in[n]} p_\eta\left(x_i\right)\right) f_j\left(x_1,\dots,x_n\right) \sum_{i\in[n]} \left(T_j\left(x_i\right) - \mu_{T,j}\right)\, dx_1 \dots dx_n \\
&= \mathop{\mathbb{E}}_{X}\left[f_j\left(X\right) \sum_{i\in[n]} \left(T_j\left(X_i\right) - \mu_{T,j}\right)\right] \\
&\overset{(a)}{=} \mathop{\mathbb{E}}_{X}\left[f_j\left(X\right) \sum_{i\in[n]} \left(T_j\left(X_i\right) - \mu_{T,j}\right)\right] - \mathop{\mathbb{E}}_{X}\left[\left(\eta_j - m_j\right) \sum_{i\in[n]} \left(T_j\left(X_i\right) - \mu_{T,j}\right)\right] \\
&= \mathop{\mathbb{E}}_{X}\left[\left[f_j\left(X\right) - \left(\eta_j - m_j\right)\right] \sum_{i\in[n]} \left(T_j\left(X_i\right) - \mu_{T,j}\right)\right],
\end{aligned}$$

where $(a)$ relies on the observation that, conditioned on $\eta_j$, $\eta_j - m_j$ and $\sum_{i\in[n]} \left(T_j\left(X_i\right) - \mu_{T,j}\right)$ are independent and $\mathop{\mathbb{E}}_{X}\left[\sum_{i\in[n]} \left(T_j\left(X_i\right) - \mu_{T,j}\right)\right] = 0$.

We proceed to define the function $G_j\colon \bigotimes_{j\in[k]} I_j \to \mathbb{R}$ as:

$$\begin{aligned}
G_j\left(\eta\right) \coloneqq \mathop{\mathbb{E}}_{X}\left[Z^j\right] &= \left[R_j^2/4 - \left(\eta_j - m_j\right)^2\right] \mathop{\mathbb{E}}_{X}\left[\left[f_j\left(X\right) - \left(\eta_j - m_j\right)\right] \sum_{i\in[n]} \left(T_j\left(X_i\right) - \mu_{T,j}\right)\right] \\
&= \left[R_j^2/4 - \left(\eta_j - m_j\right)^2\right] \frac{\partial g_j}{\partial \eta_j}\left(\eta\right).
\end{aligned}$$

Calculating the expectation of $G_j$ with respect to $\eta_j$ yields:

$$\begin{aligned}
\mathop{\mathbb{E}}_{\eta_j}\left[G_j\left(\eta\right)\right] &= \mathop{\mathbb{E}}_{\eta_j}\left[\left[R_j^2/4 - \left(\eta_j - m_j\right)^2\right] \frac{\partial g_j}{\partial \eta_j}\left(\eta\right)\right] \\
&= \frac{1}{|I_j|} \int_{\eta_j^{(1)}}^{\eta_j^{(2)}} \left[R_j^2/4 - \left(\eta_j - m_j\right)^2\right] \frac{\partial g_j}{\partial \eta_j}\left(\eta_1,\dots,\eta_j,\dots,\eta_k\right)\, d\eta_j \\
&\overset{(a)}{=} \cancelto{0}{\frac{1}{|I_j|}\left[R_j^2/4 - \left(\eta_j - m_j\right)^2\right] g_j\left(\eta_1,\dots,\eta_j,\dots,\eta_k\right)\Big|_{\eta_j=\eta_j^{(1)}}^{\eta_j=\eta_j^{(2)}}} \\
&\quad + 2\frac{1}{|I_j|} \int_{\eta_j^{(1)}}^{\eta_j^{(2)}} \left(\eta_j - m_j\right) g_j\left(\eta_1,\dots,\eta_j,\dots,\eta_k\right)\, dt \\
&= 2\mathop{\mathbb{E}}_{\eta_j}\left[\left(\eta_j - m_j\right) g_j\left(\eta\right)\right].
\end{aligned}$$

where $(a)$ uses integration by parts that $m_j$ has been defined to be equal to $(\eta_j^{(1)} + \eta_j^{(2)})/2$.

Based on this, we remark that the inequality we wish to prove can be written in the form:

$$\mathop{\mathbb{E}}_{\eta_j, X}\left[2\left(\eta_j - m_j\right) f_j\left(X\right) + \left[f_j\left(X\right) - \left(\eta_j - m_j\right)\right]^2\right] \geq R_j^2/12. \tag{1}$$

Appealing to linearity of expectation, we focus on the second term and get:

$$\begin{aligned}
\mathop{\mathbb{E}}_{\eta_j,X}\left[\left[f_j\left(X\right)-\left(\eta_j-m_j\right)\right]^2\right] &= \mathop{\mathbb{E}}_{\eta_j,X}\left[f_j^2\left(X\right)-2\left(\eta_j-m_j\right)f_j\left(X\right)+\left(\eta_j-m_j\right)^2\right] \\
&\geq -2\mathop{\mathbb{E}}_{\eta_j,X}\left[\left(\eta_j-m_j\right)f_j\left(X\right)\right]+\mathop{\mathbb{E}}_{\eta_j}\left[\left(\eta_j-m_j\right)^2\right] \\
&= -2\mathop{\mathbb{E}}_{\eta_j,X}\left[\left(\eta_j-m_j\right)f_j\left(X\right)\right]+\mathop{\mathrm{Var}}_{\eta_j}\left(\eta_j\right) \\
&= -2\mathop{\mathbb{E}}_{\eta_j,X}\left[\left(\eta_j-m_j\right)f_j\left(X\right)\right]+R_j^2/12.
\end{aligned}$$

Substituting this to (1) yields the desired result. ■

We now generalize the previous lemma in the setting where all the components of $\eta$ are drawn independently and uniformly at random from intervals.

**Lemma 2.2.** *Let $p_\eta$ be a distribution over $S \subseteq \mathbb{R}^d$ belonging to an exponential family $\mathcal{E}\left(T,h\right)$ with natural parameter vector $\eta \in \mathcal{H} \subseteq \mathbb{R}^k$. We assume that $\eta$ is randomly generated by drawing independently $\eta_j \sim \mathcal{U}\left(I_j\right), \forall j \in [k]$. Then, for any function $f\colon S^n \to \bigotimes_{j\in[k]}\left[\pm R_j/2\right]$ that takes as input a dataset $X \coloneqq \left(X_1,\ldots,X_n\right) \sim p_\eta^{\bigotimes n}$, we have:*

$$\mathop{\mathbb{E}}_{\eta,X}\left[Z+\left\|f\left(X\right)-\left(\eta-m\right)\right\|_2^2\right] \geq \left\|R\right\|_2^2/12.$$

The proof is merely a component-wise application of Lemma 2.1.

### 2.3 From the Lemma to Lower Bounds

We leverage Lemma 2.2 to give a general recipe for proving lower bounds for private estimation. The overall structure of our argument is similar to that of past fingerprinting proofs but, along the way, uses specific properties of exponential families to deal with the fact that we do not have a product distribution.

We assume the existence of an $(\varepsilon,\delta)$-DP mechanism $M\colon S^n \to \bigotimes_{j\in[k]}\left[\pm R_j/2\right]$ with $\varepsilon \in [0,1]$ and $\delta \geq 0$ such that, for any distribution $p_\eta$ with $\eta \in \bigotimes_{j\in[k]} I_j$, we have for $X \sim p_\eta^{\bigotimes n}$:

$$\mathop{\mathbb{E}}_{X,M}\left[\left\|M\left(X\right)-\left(\eta-m\right)\right\|_2^2\right] \leq \alpha^2 \leq \left\|R\right\|_2^2/24. \tag{2}$$

For the above, Lemma 2.2 with $f \equiv M$ yields:

$$\mathop{\mathbb{E}}_{\eta,X,M}\left[Z\right] = \mathop{\mathbb{E}}_{M}\left[\mathop{\mathbb{E}}_{\eta,X}\left[Z\right]\right] \geq \left\|R\right\|_2^2/12 - \mathop{\mathbb{E}}_{\eta,X,M}\left[\left\|M\left(X\right)-\left(\eta-m\right)\right\|_2^2\right] \geq \left\|R\right\|_2^2/24. \tag{3}$$

Thus, our intent is to upper-bound the LHS by a function of $n$, which will allow us to identify a lower bound to get the desired accuracy. To do so, we first need to upper-bound $\mathop{\mathbb{E}}_{\eta,X,M}\left[Z_i^2\right]$.

**Lemma 2.3.** *Conditioning on $\eta$, for any $\varepsilon \in [0,1], \delta \geq 0$, and $T > 0$, it holds that:*

$$\mathop{\mathbb{E}}_{X,M}\left[Z_i\right] \leq 2\delta T + 2\varepsilon\sqrt{\mathop{\mathbb{E}}_{X_{\sim i},M}\left[s^\top \Sigma_T s\right]} + 2\int_T^\infty \mathop{\mathbb{P}}_{X_i}\left[\left\|T\left(X_i\right)-\mu_T\right\|_2 > (4t)/(\left\|R\right\|_\infty^3\sqrt{k})\right]\,dt,$$

*where $s \in \mathbb{R}^k$ with $s_j \coloneqq \left[R_j^2/4-(\eta_j-m_j)^2\right]\left[M_j\left(X_{\sim i}\right)-\left(\eta_j-m_j\right)\right], \forall j \in [k]$.*

The proof involves splitting $Z_i$ in its positive and negative components and applying the definition of DP to each of them. The argument is standard and has appeared in previous works (e.g., see [SU17b]), but we need to apply specific properties of exponential families to deal with the fact that we don't have a product distribution. The proof is lengthy so we omit it.

Combining Lemma 2.3 with (3) yields the main result of this section:

**Theorem 2.4.** *Let $p_\eta$ be a distribution over $S \subseteq \mathbb{R}^d$ belonging to an exponential family $\mathcal{E}(T, h)$ with natural parameter vector $\eta \in \mathcal{H} \subseteq \mathbb{R}^k$. Also, let $\eta^{(1)}, \eta^{(2)} \in \mathcal{H}$ and let $I_j := \left[\eta_j^{(1)}, \eta_j^{(2)}\right], \forall j \in [k]$ be a collection of intervals and $R := \eta^{(2)} - \eta^{(1)}, m := (\eta^{(1)} + \eta^{(2)})/2$ be the corresponding width and midpoint vectors, respectively. Assume that $\bigotimes_{j\in[k]} I_j \subseteq \mathcal{H}$ and that $\eta$ is drawn from the distribution $\mathcal{U}(\bigotimes_{j\in[k]} I_j)$. Moreover, assume that we have a dataset $X \sim p_\eta^{\bigotimes n}$ and an independently drawn point $X_i' \sim p_\eta$ and $X_{\sim i}$ denotes the dataset where $X_i$ has been replaced with $X_i'$. Finally, let $M\colon S^n \to \bigotimes_{j\in[k]} [\pm R_j/2]$ be an $(\varepsilon, \delta)$-DP mechanism with $\varepsilon \in [0, 1], \delta \geq 0$ with $\mathop{\mathbb{E}}_{X,M}\left[\|M(X) - (\eta - m)\|_2^2\right] \leq \alpha^2 \leq \|R\|_2^2/24$. Then, for any $T > 0$, it holds that:*

$$n\left\{2\delta T + 2\varepsilon \mathop{\mathbb{E}}_{\eta}\left[\sqrt{\mathop{\mathbb{E}}_{X_{\sim i},M}\left[s^\top \Sigma_T s\right]}\right] + 2\int_T^\infty \mathop{\mathbb{P}}_{X_i}\left[\|T(X_i) - \mu_T\|_2 > \frac{4t}{\|R\|_\infty^3 \sqrt{k}}\right] dt\right\} \geq \frac{\|R\|_2^2}{24}.$$

In order to apply the above theorem, the main idea is to identify values for $T$ and $\delta$ such that $\delta T \geq 2\int_T^\infty \mathop{\mathbb{P}}_{X_i}\left[\|T(X_i) - \mu_T\|_2 > (4t)/(\|R\|_\infty^3 \sqrt{k})\right] dt$, and $3\delta T n \leq \frac{\|R\|_2^2}{48}$.

The reasoning behind this is that the first and third terms of (2.4) both depend on $T$, with the latter term becoming smaller as $T$ increases. Thus, balancing these two terms involves identifying a value for $T$ that is as small as possible and results in $\delta T$ dominating the other term. At the same time, we want the sum of those two terms to be smaller than the half of the RHS, so that we get a lower bound that is as tight as possible, leading to the constraint on $\delta$.

The above result in the inequality $n\mathop{\mathbb{E}}_{\eta}\left[\sqrt{\mathop{\mathbb{E}}_{X_{\sim i},M}\left[s^\top \Sigma_T s\right]}\right] \geq \|R\|_2^2/(96\varepsilon)$.

Consequently, obtaining the desired sample complexity lower bounds boils down to upper-bounding the term $\mathop{\mathbb{E}}_{X_{\sim i},M}\left[s^\top \Sigma_T s\right]$. In the applications we will see, the upper bounds will be worst-case in terms of $\eta$, so the outer expectation $\mathop{\mathbb{E}}_{\eta}[\cdot]$ will be largely ignored.

# 3 Lower Bounds for Private Gaussian Covariance Estimation

Here, we provide lower bounds for covariance estimation of high-dimensional Gaussians with respect to the Mahalanobis and spectral norms under the constraint of $(\varepsilon, \delta)$-DP.

## 3.1 Estimation with Respect to the Mahalanobis Norm

In this section, we characterize the sample complexity of private Gaussian covariance estimation with respect to the Mahalanobis norm under $(\varepsilon, \delta)$-DP. Previous lower bounds either assumed a stricter notion of privacy (e.g., $\varepsilon$-DP - see [KLSU19]) or exhibited significant gaps with known upper bounds (see Section 1.1.1 of [AAK21] for further discussion). We start by stating our main result.

**Theorem 3.1.** *There exists a distribution $\mathcal{D}$ over covariance matrices $\Sigma \in \mathbb{R}^{d\times d}$ with $\mathbb{I} \preceq \Sigma \preceq 2\mathbb{I}$ such that, given $\Sigma \sim \mathcal{D}$ and $X \sim \mathcal{N}(0, \Sigma)^{\bigotimes n}$, for any $\alpha = \mathcal{O}(1)$ and any $(\varepsilon, \delta)$-DP mechanism $M\colon \mathbb{R}^{n\times d} \to \mathbb{R}^{d\times d}$ with $\varepsilon, \delta \in [0, 1]$, and $\delta \leq \mathcal{O}(\min\left\{1/n, d^2/(n\log(n/d^2))\right\})$ that satisfies $\mathop{\mathbb{E}}_{X,M}\left[\|M(X) - \Sigma\|_\Sigma^2\right] \leq \alpha^2$, it must hold that $n \geq \Omega\left(\frac{d^2}{\alpha\varepsilon}\right)$.*

*Proof Sketch.* The outline of the proof involves (1) writing the Gaussian distribution $\mathcal{N}(0, \Sigma)$ as an exponential family, (2) defining a distribution from which $\eta$ is drawn, (3) showing that estimating the deviation of the natural parameter vector from the midpoint $m$ reduces to covariance estimation (so lower bounds for the former also apply for the latter), (4) upper-bounding the term $\mathbb{E}\left[s^\top \Sigma_T s\right]$, and (5) applying Theorem 2.4 as described at the end of the previous section.

Step (1) is non-trivial because the standard way to write $\mathcal{N}(0, \Sigma)$ as an exponential family involves setting $\eta$ to be equal to the *canonical flattening* of the *precision matrix* $\left(\Sigma^{-1}\right)^\flat$. However, we cannot use this representation, because $\Sigma^{-1}$ is symmetric, preventing the components of $\eta$ from

being generated independently, as is required by Theorem 2.4. Thus, we consider an alternative parameterization. This involves identifying an upper-triangular matrix $U$ such that $\Sigma^{-1} = U + U^\top$ and setting the parameter vector to be equal to $2U^\flat$. The two parameterizations are equivalent.

Step (2) employs a variant of a construction that has appeared in [KLSU19]. The process consists of sampling the elements of the precision matrix uniformly at random from intervals of width $1/(2d)$. For the diagonal elements, these intervals are centered at $3/4$, whereas for the non-diagonal elements they are centered at the origin.

Step (3) requires us to describe an algorithm which, given oracle access to $(\varepsilon, \delta)$-DP estimate $\widehat{\Sigma}$ of $\Sigma$ that is $\alpha$-accurate in Mahalanobis norm, produces an $(\varepsilon, \delta)$-DP estimate of the natural parameter vector that is $\mathcal{O}(a)$-accurate in $\ell_2$-norm. The idea is to calculate $\widehat{\Sigma}^{-1}$, obtain a projection $\widetilde{\Sigma}^{-1}$ of it onto the support of the output of the algorithm described in step (2), find an upper-triangular matrix $\widetilde{U}$ such that $\widetilde{\Sigma}^{-1} = U + U^\top$ and output $2\widetilde{U}^\flat - m$.

Steps (4) and (5) are mostly calculations, so we omit them from here. ■

### 3.2 Estimation with Respect to the Spectral Norm

Here, we prove a lower bound for covariance estimation of high-dimensional Gaussians in spectral norm under the constraint of approximate DP. Employing a reduction-based approach from Mahalanobis estimation to spectral estimation, we directly leverage the results of Section 3.1.

**Theorem 3.2.** *There exists a distribution $\mathcal{D}$ over covariance matrices $\Sigma \in \mathbb{R}^{d\times d}$ with $\mathbb{I} \preceq \Sigma \preceq 2\mathbb{I}$ such that, given $\Sigma \sim \mathcal{D}$ and $X \sim \mathcal{N}(0, \Sigma)^{\otimes n}$, for any $\alpha = \mathcal{O}(1/\sqrt{d})$ and any $(\varepsilon, \delta)$-DP mechanism $M\colon \mathbb{R}^{n\times d} \to \mathbb{R}^{d\times d}$ with $\varepsilon, \delta \in [0, 1]$, and $\delta \le \mathcal{O}(\min\{1/n, d^2/(n\log(n/d^2))\})$ that satisfies $\mathop{\mathbb{E}}_{X,M}\left[\left\|\Sigma^{-\frac{1}{2}}\left(M(X) - \Sigma\right)\Sigma^{-\frac{1}{2}}\right\|_2^2\right] \le \alpha^2$, it must hold that $n \ge \Omega\left(\frac{d^{1.5}}{\alpha\varepsilon}\right)$.*

The proof is a reduction via the property $\|A\|_F \le \sqrt{d}\,\|A\|_2$.

## 4 Lower Bound for Private Heavy-Tailed Mean Estimation

In this section, we state the lower bound for mean-estimation of distributions with bounded second moment with respect to the $\ell_2$ norm and sketch its proof.

**Theorem 4.1.** *Let $\mathcal{D}$ be a distribution over $\mathbb{R}^d$ with second moments bounded by $1$ and unknown mean $\mu$. Then, for any $\alpha \le 1$, any $(\varepsilon, \delta)$-DP mechanism with $\delta \le \varepsilon$ that takes $X \sim \mathcal{D}^{\otimes n}$ as input and outputs $M(X)$ satisfying $\mathop{\mathbb{E}}_{X,M}\left[\|M(X) - \mu\|_2^2\right] \le \alpha^2$ requires $n \ge \Omega\left(\frac{d}{\alpha^2\varepsilon}\right)$.*

*Proof Sketch.* We use the DP Assouad lemma of [ASZ21]. To apply the method, we construct a family of distributions indexed by the points of the binary hypercube $\mathcal{E}_d \coloneqq \{-1, 1\}^d$. Define $p = (2\alpha^2)/d$ and $t = \sqrt{d}/(\sqrt{2}\alpha)$. Then for each $v \in \mathcal{E}_d$, we define a distribution $\mathcal{D}_v$ over $\mathbb{R}^d$, such that for $X \sim \mathcal{D}_v$, the following holds.

$$X_i = \begin{cases} v_i t, & \text{with probability } p \\ 0, & \text{with probability } 1 - p \end{cases}, \forall i \in [d].$$

It can be shown that the second moment of $\mathcal{D}_v$ is at most $pt^2 = 1$.

For $u, v \in \mathcal{E}_d$, we define the loss function between $\mathcal{D}_u$ and $\mathcal{D}_v$ to be $\ell(\theta(\mathcal{D}_u), \theta(\mathcal{D}_v)) = \|\theta(\mathcal{D}_u) - \theta(\mathcal{D}_v)\|_2^2$. This implies that (as defined in Lemma A.4) $\tau = 2p^2t^2 = 2p$. On rearranging the terms in Lemma A.4, the above gives us $D \ge 0.04/(\varepsilon + \delta)$.

The final task is to show that there exists a coupling $(X, Y)$ between the mixture distributions $p_{+i}$ and $p_{-i}$ (as defined in Lemma A.4) with the appropriate value of $D = \mathbb{E}[d_{\mathrm{Ham}}(X, Y)]$. Here, we just construct such a coupling, but verify the bound on $D$ in the appendix. Let $(X, Y)$ be as follows. For every row $X_j$ in $X$, if $X_j^i \neq 0$, then $Y_j^i = -X_j^i$, otherwise $Y_j^i = X_j^i$. For all other coordinates $k \neq i$, $Y_j^k = X_j^k$. We show that in this case, $D = np$. This gives us $D/p = n \ge d/[50(\varepsilon + \delta)\alpha^2)]$, which proves the theorem. ■

## 5 Conclusions and Open Problems

In this work, we gave nearly tight lower bounds for private covariance estimation with respect to the Frobenius and spectral norms under the Gaussian distribution and for mean estimation with respect to the $\ell_2$-norm under distributions with bounded second moments. This constitutes important progress in the lower bounds literature under approximate differential privacy. On a technical level, we significantly generalized the fingerprinting method. An obvious question is whether the fingerprinting method can be further generalized and what are its limitations. A second obvious question is to generalize Theorem 4.1 for higher-order moments.

Another question involves expanding upon the results of Theorems 3.1 and 3.2 to hold in the regimes $\alpha = \mathcal{O}(\sqrt{d})$ and $\alpha = \mathcal{O}(1)$, respectively. For Frobenius estimation, the condition $\alpha = \mathcal{O}(1)$ is a consequence of the distribution over covariance matrices that we used to prove the result, whereas for spectral estimation the condition $\alpha = \mathcal{O}(1/\sqrt{d})$ is an artifact of our reduction-based approach. Especially in the latter case, it would be interesting to come up with a "direct" way to obtain the lower bound, instead of reducing from Frobenius estimation.

Finally, an interesting open question may be to obtain lower bounds for the problems considered here in the *high-probability regime*. Specifically, the guarantees in our lower bounds are expressed in terms of the MSE and in Section 1.1 we sketched the argument to convert them to constant probability guarantees. It would be interesting to come up with techniques to prove tight lower bounds when the probability of success is $1 - \beta$ for $\beta \in (0, 1)$. This was done in the context of pure differential privacy in [HKM21].

## Acknowledgments

The authors would like to thank Jonathan Ullman for helpful feedback on the manuscript, Haoshu Xu for identifying a bug in the proof of Lemma 2.3 which affected how the range of $\delta$ is identified in Theorems 3.1, as well as the anonymous reviewers at NeurIPS for their comments.

GK was supported by an NSERC Discovery Grant, an unrestricted gift from Google, and a University of Waterloo startup grant. AM was supported by an NSERC Discovery Grant and a David R. Cheriton Graduate Scholarship. VS was supported by an NSERC Discovery Grant.

## References

[AAK21] 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.

[AAL21] Ishaq Aden-Ali, Hassan Ashtiani, and Christopher Liaw. Privately learning mixtures of axis-aligned gaussians. In *Advances in Neural Information Processing Systems 34*, NeurIPS '21. Curran Associates, Inc., 2021.

[ADK19] Brendan Avent, Yatharth Dubey, and Aleksandra Korolova. The power of the hybrid model for mean estimation. *Proceedings on Privacy Enhancing Technologies*, 2020(4):48–68, 2019.

[AL21] Hassan Ashtiani and Christopher Liaw. Private and polynomial time algorithms for learning Gaussians and beyond. *arXiv preprint arXiv:2111.11320*, 2021.

[AMB19] Marco Avella-Medina and Victor-Emmanuel Brunel. Differentially private sub-Gaussian location estimators. *arXiv preprint arXiv:1906.11923*, 2019.

[ASZ18] Jayadev Acharya, Ziteng Sun, and Huanyu Zhang. Differentially private testing of identity and closeness of discrete distributions. In *Advances in Neural Information Processing Systems 31*, NeurIPS '18, pages 6878–6891. Curran Associates, Inc., 2018.

[ASZ21] Jayadev Acharya, Ziteng Sun, and Huanyu Zhang. Differentially private Assouad, Fano, and Le Cam. In *Proceedings of the 32nd International Conference on Algorithmic Learning Theory*, ALT '21, pages 48–78. JMLR, Inc., 2021.

[BD14] Rina Foygel Barber and John C Duchi. Privacy and statistical risk: Formalisms and minimax bounds. *arXiv preprint arXiv:1412.4451*, 2014.

[BDKU20] Sourav Biswas, Yihe Dong, Gautam Kamath, and Jonathan Ullman. Coinpress: Practical private mean and covariance estimation. In *Advances in Neural Information Processing Systems 33*, NeurIPS '20, pages 14475–14485. Curran Associates, Inc., 2020.

[BGS+21] Gavin Brown, Marco Gaboardi, Adam Smith, Jonathan Ullman, and Lydia Zakynthinou. Covariance-aware private mean estimation without private covariance estimation. *arXiv preprint arXiv:2106.13329*, 2021.

[BKSW19] Mark Bun, Gautam Kamath, Thomas Steinke, and Zhiwei Steven Wu. Private hypothesis selection. In *Advances in Neural Information Processing Systems 32*, NeurIPS '19, pages 156–167. Curran Associates, Inc., 2019.

[BS98] Dan Boneh and James Shaw. Collusion-secure fingerprinting for digital data. *IEEE Transactions on Information Theory*, 44(5):1897–1905, 1998.

[BS19] Mark Bun and Thomas Steinke. Average-case averages: Private algorithms for smooth sensitivity and mean estimation. In *Advances in Neural Information Processing Systems 32*, NeurIPS '19, pages 181–191. Curran Associates, Inc., 2019.

[BST14] Raef Bassily, Adam Smith, and Abhradeep Thakurta. Private empirical risk minimization: Efficient algorithms and tight error bounds. In *Proceedings of the 55th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '14, pages 464–473, Washington, DC, USA, 2014. IEEE Computer Society.

[BSU17] Mark Bun, Thomas Steinke, and Jonathan Ullman. Make up your mind: The price of online queries in differential privacy. In *Proceedings of the 28th Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '17, pages 1306–1325, Philadelphia, PA, USA, 2017. SIAM.

[BUV14] Mark Bun, Jonathan Ullman, and Salil Vadhan. Fingerprinting codes and the price of approximate differential privacy. In *Proceedings of the 46th Annual ACM Symposium on the Theory of Computing*, STOC '14, pages 1–10, New York, NY, USA, 2014. ACM.

[CWZ19] T. Tony Cai, Yichen Wang, and Linjun Zhang. The cost of privacy: Optimal rates of convergence for parameter estimation with differential privacy. *arXiv preprint arXiv:1902.04495*, 2019.

[DFM+20] Wenxin Du, Canyon Foot, Monica Moniot, Andrew Bray, and Adam Groce. Differentially private confidence intervals. *arXiv preprint arXiv:2001.02285*, 2020.

[DHS15] Ilias Diakonikolas, Moritz Hardt, and Ludwig Schmidt. Differentially private learning of structured discrete distributions. In *Advances in Neural Information Processing Systems 28*, NIPS '15, pages 2566–2574. Curran Associates, Inc., 2015.

[DJW13] John C. Duchi, Michael I. Jordan, and Martin J. Wainwright. Local privacy and statistical minimax rates. In *Proceedings of the 54th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '13, pages 429–438, Washington, DC, USA, 2013. IEEE Computer Society.

[DKK+16] Ilias Diakonikolas, Gautam Kamath, Daniel M. Kane, Jerry Li, Ankur Moitra, and Alistair Stewart. Robust estimators in high dimensions without the computational intractability. In *Proceedings of the 57th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '16, pages 655–664, Washington, DC, USA, 2016. IEEE Computer Society.

[DKM+06] Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor. Our data, ourselves: Privacy via distributed noise generation. In *Proceedings of the 24th Annual International Conference on the Theory and Applications of Cryptographic Techniques*, EUROCRYPT '06, pages 486–503, Berlin, Heidelberg, 2006. Springer.

[DMNS06] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In *Proceedings of the 3rd Conference on Theory of Cryptography*, TCC '06, pages 265–284, Berlin, Heidelberg, 2006. Springer.

[DMR18] Luc Devroye, Abbas Mehrabian, and Tommy Reddad. The total variation distance between high-dimensional Gaussians. *arXiv preprint arXiv:1810.08693*, 2018.

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

[DTTZ14] Cynthia Dwork, Kunal Talwar, Abhradeep Thakurta, and Li Zhang. Analyze Gauss: Optimal bounds for privacy-preserving principal component analysis. In *Proceedings of the 46th Annual ACM Symposium on the Theory of Computing*, STOC '14, pages 11–20, New York, NY, USA, 2014. ACM.

[Ger31] S. Gershgorin. Über die abgrenzung der Eigenwerte einer matrix. *Izv. Akad. Nauk. SSSR Ser. Mat.*, 1:749–754, 1931.

[HKM21] Samuel B Hopkins, Gautam Kamath, and Mahbod Majid. Efficient mean estimation with pure differential privacy via a sum-of-squares exponential mechanism. *arXiv preprint arXiv:2111.12981*, 2021.

[HLY21] Ziyue Huang, Yuting Liang, and Ke Yi. Instance-optimal mean estimation under differential privacy. In *Advances in Neural Information Processing Systems 34*, NeurIPS '21. Curran Associates, Inc., 2021.

[HT10] Moritz Hardt and Kunal Talwar. On the geometry of differential privacy. In *Proceedings of the 42nd Annual ACM Symposium on the Theory of Computing*, STOC '10, pages 705–714, New York, NY, USA, 2010. ACM.

[Jor10] Michael Jordan. Lecture notes for Bayesian modeling and inference, 2010.

[KLSU19] 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.

[KLZ21] Gautam Kamath, Xingtu Liu, and Huanyu Zhang. Improved rates for differentially private stochastic convex optimization with heavy-tailed data. *arXiv preprint arXiv:2106.01336*, 2021.

[KMS+21] Gautam Kamath, Argyris Mouzakis, Vikrant Singhal, Thomas Steinke, and Jonathan Ullman. A private and computationally-efficient estimator for unbounded gaussians. *arXiv preprint arXiv:2111.04609*, 2021.

[KSSU19] Gautam Kamath, Or Sheffet, Vikrant Singhal, and Jonathan Ullman. Differentially private algorithms for learning mixtures of separated Gaussians. In *Advances in Neural Information Processing Systems 32*, NeurIPS '19, pages 168–180. Curran Associates, Inc., 2019.

[KSU20] 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.

[KU20] Gautam Kamath and Jonathan Ullman. A primer on private statistics. *arXiv preprint arXiv:2005.00010*, 2020.

[KV18] 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.

[Li19] Jerry Li. Lecture 18: Private covariance estimation in high dimensions. `https://jerryzli.github.io/robust-ml-fall19/lec18.pdf`, December 2019. Lecture notes.

[LKO21] Xiyang Liu, Weihao Kong, and Sewoong Oh. Differential privacy and robust statistics in high dimensions. *arXiv preprint arXiv:2111.06578*, 2021.

[LM00] Beatrice Laurent and Pascal Massart. Adaptive estimation of a quadratic functional by model selection. *The Annals of Statistics*, 28(5):1302–1338, 2000.

[LSA+21] 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 34*, NeurIPS '21. Curran Associates, Inc., 2021.

[LSY+20] Yuhan Liu, Ananda Theertha Suresh, Felix Yu, Sanjiv Kumar, and Michael Riley. Learning discrete distributions: User vs item-level privacy. In *Advances in Neural Information Processing Systems 33*, NeurIPS '20. Curran Associates, Inc., 2020.

[Mos21] Kamyar Moshksar. On the absolute constant in hanson-wright inequality. *arXiv preprint arXiv:2111.00557*, 2021.

[SU15] Thomas Steinke and Jonathan Ullman. Interactive fingerprinting codes and the hardness of preventing false discovery. In *Proceedings of the 28th Annual Conference on Learning Theory*, COLT '15, pages 1588–1628, 2015.

[SU17a] Thomas Steinke and Jonathan Ullman. Between pure and approximate differential privacy. *The Journal of Privacy and Confidentiality*, 7(2):3–22, 2017.

[SU17b] Thomas Steinke and Jonathan Ullman. Tight lower bounds for differentially private selection. In *Proceedings of the 58th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '17, pages 552–563, Washington, DC, USA, 2017. IEEE Computer Society.

[Tar08] Gabor Tardos. Optimal probabilistic fingerprint codes. *Journal of the ACM*, 55(2), 2008.

[TVGZ20] Christos Tzamos, Emmanouil-Vasileios Vlatakis-Gkaragkounis, and Ilias Zadik. Optimal private median estimation under minimal distributional assumptions. In *Advances in Neural Information Processing Systems 33*, NeurIPS '20, pages 3301–3311. Curran Associates, Inc., 2020.

[Ver18] Roman Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge University Press, 2018.

[WXDX20] Di Wang, Hanshen Xiao, Srinivas Devadas, and Jinhui Xu. On differentially private stochastic convex optimization with heavy-tailed data. In *Proceedings of the 37th International Conference on Machine Learning*, ICML '20, pages 10081–10091. JMLR, Inc., 2020.

[Yu97] Bin Yu. Assouad, Fano, and Le Cam. In *Festschrift for Lucien Le Cam: Research Papers in Probability and Statistics*, pages 423–435. Springer New York, 1997.

[ZKKW20] Huanyu Zhang, Gautam Kamath, Janardhan Kulkarni, and Zhiwei Steven Wu. Privately learning Markov random fields. In *Proceedings of the 37th International Conference on Machine Learning*, ICML '20, pages 11129–11140. JMLR, Inc., 2020.

## Checklist

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes]
   (b) Did you describe the limitations of your work? [Yes] See Section 5.
   (c) Did you discuss any potential negative societal impacts of your work? [N/A]
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [Yes]
   (b) Did you include complete proofs of all theoretical results? [Yes] See Appendices A-G.
3. If you ran experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [N/A]
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [N/A]
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [N/A]
   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [N/A]
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
   (a) If your work uses existing assets, did you cite the creators? [N/A]
   (b) Did you mention the license of the assets? [N/A]
   (c) Did you include any new assets either in the supplemental material or as a URL? [N/A]
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [N/A]

(e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A]

5. If you used crowdsourcing or conducted research with human subjects...

(a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]

(b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]

(c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]

# A Preliminaries

**General Notation.** We use the following notations: $[n] \coloneqq \{1, 2, \ldots, n\}$, $[a \pm R] \coloneqq [a - R, a + R]$, and $\mathbb{R}_+ \coloneqq [0, \infty)$. Given any distribution $\mathcal{D}$, $\mathcal{D}^{\bigotimes n}$ denotes the *product measure* where each *marginal distribution* is $\mathcal{D}$. Thus, if we are given $n$ independent samples from $\mathcal{D}$, we write $(X_1, \ldots, X_n) \sim \mathcal{D}^{\bigotimes n}$. Also, depending on the context, we may use capital Latin characters like $X$ to denote either an individual sample from a distribution or a collection of samples $X \coloneqq (X_1, \ldots, X_n)$. To denote the $j$-th component of a vector, we will use either a subscript (e.g., $X_j$, if the vector is $X$) or a superscript (e.g., $X_i^j$, if the vector is $X_i$). The convention used will be clear depending on the context. The *unit $\ell_2$-sphere* in $d$-dimensions that is centered at the origin is denoted by $\mathbb{S}^{d-1}$. Finally, we give the following definition for distributions with bounded $k$-th moments:

**Definition A.1.** Let $\mathcal{D}$ be a distribution over $\mathbb{R}^d$. We say that $\mathcal{D}$ has $k$-th moments bounded by $C$ if $\mathop{\mathbb{E}}_{X \sim \mathcal{D}} \left[ \left| \left\langle X - \mathop{\mathbb{E}}_{X \sim \mathcal{D}} [X], v \right\rangle \right|^k \right] \leq C, \forall v \in \mathbb{S}^{d-1}$.

**Linear Algebra Preliminaries.** For a vector $v \in \mathbb{R}^d$, $v_i$ refers to its $i$-th component and $v_{-i} \in \mathbb{R}^{d-1}$ describes the vector one gets by removing the $i$-th component from $v$. We use $\vec{1}_d$ to refer to the all 1s vector in $\mathbb{R}^d$. For a pair of vectors $x, y \in \mathbb{R}^d$, we write $x \leq y$ if and only if $x_i \leq y_i, \forall i \in [n]$. Given a matrix $M \in \mathbb{R}^{d \times d}$, we denote the element at the $i$-th row and the $j$-th column by $M_{ij}$ or $(M)_{ij}$. Also, we denote the *canonical flattening* operation of $M$ into a vector by $v = M^\flat \in \mathbb{R}^{d^2}$ and the inverse operation by $M = v^\# \in \mathbb{R}^{d \times d}$. We denote its *spectral norm* by $\|M\|_2 \coloneqq \sup_{v \in \mathbb{S}^{d-1}} \|Mv\|_2$. For symmetric matrices $M$, this reduces to $\|M\|_2 = \sup_{v \in \mathbb{S}^{d-1}} \left| v^\top M v \right|$, which is known to be equal to the largest eigenvalue of $M$ (in absolute value). Also, we denote the *trace* of $M$ by $\operatorname{tr}(M) \coloneqq \sum_{i \in [d]} M_{ii}$, and its *Frobenius norm* by $\|M\|_F \coloneqq \sqrt{\operatorname{tr}(M^\top M)} = \sqrt{\sum_{i=1}^{d} \sum_{j=1}^{d} M_{ij}^2} = \left\| M^\flat \right\|_2$. For symmetric matrices, the latter is equal to the square root of the sum of the squares of its eigenvalues. Moreover, we consider the *Mahalanobis norm* (with respect to a non-degenerate covariance matrix $\Sigma$), both for vectors $\left( \|x\|_\Sigma \coloneqq \left\| \Sigma^{-\frac{1}{2}} x \right\|_2 \right)$ and matrices $\left( \|M\|_\Sigma \coloneqq \left\| \Sigma^{-\frac{1}{2}} M \Sigma^{-\frac{1}{2}} \right\|_F \right)$. Furthermore, we use the notation $\langle \cdot, \cdot \rangle$ to denote *inner products*, either between vectors $x, y \in \mathbb{R}^d$ $\left( \langle x, y \rangle \coloneqq x^\top y \right)$ or between matrices $A, B \in \mathbb{R}^{d \times d}$ $\left( \langle A, B \rangle \coloneqq \operatorname{tr}\left( A B^\top \right) \right)$. The inner product of a pair of $d \times d$ matrices $A, B$ can be reduced to an inner product between vectors in $\mathbb{R}^{d^2}$ via the identity $\langle A, B \rangle = \left\langle A^\flat, B^\flat \right\rangle$. Additionally, given a pair of symmetric matrices $A, B \in \mathbb{R}^d$, we write $A \succeq B$ if and only if $x^\top (A - B) x \geq 0, \forall x \in \mathbb{R}^d$. Finally, given matrices $A \in \mathbb{R}^{n \times m}$ and $B \in \mathbb{R}^{k \times \ell}$, the *Kronecker product* of $A$ with $B$ is denoted by $A \otimes B \in \mathbb{R}^{(nk) \times (m\ell)}$. More linear algebra facts appear in Appendix E.

## A.1 Privacy Preliminaries

We define differential privacy and state its closure under post-processing property.

**Definition A.2** (Differential Privacy (DP) [DMNS06])**.** A randomized algorithm $M\colon \mathcal{X}^n \to \mathcal{Y}$ satisfies $(\varepsilon, \delta)$-differential privacy ($(\varepsilon, \delta)$-DP) if for every pair of neighboring datasets $X, X' \in \mathcal{X}^n$ (i.e., datasets that differ in exactly one entry), we have:

$$\mathbb{P}\left[M\left(X\right) \in Y\right] \leq e^{\varepsilon} \cdot \mathbb{P}\left[M\left(X'\right) \in Y\right] + \delta, \forall Y \subseteq \mathcal{Y}.$$

When $\delta = 0$, we say that $M$ satisfies $\varepsilon$-differential privacy or pure differential privacy.

**Lemma A.3** (Post Processing [DMNS06])**.** *If $M\colon \mathcal{X}^n \to \mathcal{Y}$ is $(\varepsilon, \delta)$-DP, and $P\colon \mathcal{Y} \to \mathcal{Z}$ is any randomized function, then the algorithm $P \circ M$ is $(\varepsilon, \delta)$-DP.*

Finally, we discuss a recent technique from [ASZ21] for proving lower bounds for estimation problems under approximate DP. It is a private version of the well-known Assouad Lemma [Yu97]. Let $\mathcal{P}$ be a family of distributions over $\mathcal{X}^n$, where $\mathcal{X}$ is the data universe and $n$ is the number of samples. Let $\theta\colon \mathcal{P} \to \Theta$ be the quantity associated with the distribution that we want to estimate, and let $\ell\colon \Theta \times \Theta \to \mathbb{R}_+$ be the pseudo-metric (loss function) for estimating $\theta$. We define the risk of an estimator $\hat{\theta}\colon \mathcal{X}^n \to \Theta$ as $\max\limits_{p \in \mathcal{P}} \mathop{\mathbb{E}}\limits_{X \sim p^{\otimes n}} \left[\ell\left(\hat{\theta}\left(X\right), \theta\left(p\right)\right)\right]$. The minimax risk of $(\varepsilon, \delta)$-DP estimators for a statistical task is defined as:

$$R\left(\mathcal{P}, \ell, \varepsilon, \delta\right) \coloneqq \min_{\hat{\theta} \text{ is } (\varepsilon,\delta)\text{-DP}} \max_{p \in \mathcal{P}} \mathop{\mathbb{E}}_{X \sim P^{\otimes n}} \left[\ell\left(\hat{\theta}\left(X\right), \theta\left(p\right)\right)\right].$$

Let $\mathcal{V} \subseteq \mathcal{P}$ be a subset of distributions indexed by the points in the hypercube $\mathcal{E}_d = \{-1, 1\}^d$. Suppose there exists a $\tau \in \mathbb{R}$, such that for all $u, v \in \mathcal{E}_d$, $\ell\left(\theta\left(p_u\right), \theta\left(p_v\right)\right) \geq 2\tau \cdot \sum\limits_{i=1}^{d} \mathbb{1}\left\{u_i \neq v_i\right\}$. For every coordinate $i \in [d]$, define the following mixture distributions:

$$p_{+i} = \frac{2}{|\mathcal{E}_d|} \sum_{e \in \mathcal{E}_d\colon e_i = 1} p_e \text{ and } p_{-i} = \frac{2}{|\mathcal{E}_d|} \sum_{e \in \mathcal{E}_d\colon e_i = -1} p_e.$$

The DP Assouad's method provides a lower bound for the minimax risk by considering the problem of distinguishing between $p_{+i}$ and $p_{-i}$.

**Lemma A.4** (DP Assouad's Method [ASZ21])**.** *For all $i \in [d]$, let $\phi_i\colon \mathcal{X}^n \to \{-1, 1\}$ be a binary classifier. Then we have the following:*

$$R\left(\mathcal{P}, \ell, \varepsilon, \delta\right) \geq \frac{\tau}{2} \cdot \sum_{i=1}^{d} \min_{\phi_i \text{ is } (\varepsilon,\delta)\text{-DP}} \left( \mathop{\mathbb{P}}_{X \sim p_{+i}^{\otimes n}} \left[\phi_i\left(X\right) \neq 1\right] + \mathop{\mathbb{P}}_{X \sim p_{-i}^{\otimes n}} \left[\phi_i\left(X\right) \neq -1\right] \right).$$

*Moreover, if $\forall i \in [d]$, there exists a coupling $(X, Y)$ between $p_{+i}$ and $p_{-i}$ with $\mathbb{E}\left[d_{\mathrm{Ham}}\left(X, Y\right)\right] \leq D$, then:*

$$R\left(\mathcal{P}, \ell, \varepsilon, \delta\right) \geq \frac{d\tau}{2} \cdot \left(0.9e^{-10\varepsilon D} - 10\delta D\right).$$

### A.2 Exponential Families

In this work, we make frequent use of exponential families due to their expressive power. This gives us a unified way of dealing with multiple commonly used distributions (e.g., Gaussians, Bernoullis). We first state the definition of exponential families.

**Definition A.5** (Exponential Family)**.** Let $S \subseteq \mathbb{R}^d$ and $h\colon S \to \mathbb{R}_+, T\colon S \to \mathbb{R}^k$. Given $\eta \in \mathbb{R}^k$, we say that the density function $p_\eta$ belongs to the *exponential family* $\mathcal{E}\left(T, h\right)$ if its density can be written in the form:

$$p_\eta\left(x\right) = h\left(x\right) \exp\left(\eta^\top T\left(x\right) - Z\left(\eta\right)\right), \forall x \in S,$$

where:

$$Z\left(\eta\right) = \ln\left(\int_S h\left(x\right) \exp\left(\eta^\top T\left(x\right)\right) dx\right).$$

The functions $h$ and $T$ are referred to as the *carrier measure* and the *sufficient statistics* of the family, respectively. Additionally, $Z$ is known as the *log-partition function* and $\eta$ is the *natural parameter vector*. Finally, we denote the *range of natural parameters* by $\mathcal{H} \subseteq \mathbb{R}^k$, which is the set of values of $\eta$ for which the log-partition function is well-defined ($Z\left(\eta\right) < \infty$).

We appeal to the following properties of exponential families.

**Proposition A.6** (see [Jor10]). *Let $\mathcal{E}(T, h)$ be an exponential family parameterized by $\eta \in \mathbb{R}^k$. Then, the following hold:*

- *For all $\eta \in \mathcal{H}$, it holds that:*
$$\mu_T \coloneqq \mathop{\mathbb{E}}_{X \sim p_\eta}[T(X)] = \nabla_\eta Z(\eta).$$
- *For all $\eta \in \mathcal{H}$, it holds that:*
$$\Sigma_T \coloneqq \mathop{\mathbb{E}}_{X \sim p_\eta}\left[(T(X) - \mu_T)(T(X) - \mu_T)^\top\right] = \nabla_\eta^2 Z(\eta).$$
- *For all $\eta \in \mathcal{H}$ and $s \in \mathbb{R}^k$ such that $\eta + s \in \mathcal{H}$, it holds that:*
$$\mathop{\mathbb{E}}_{X \sim p_\eta}\left[\exp\left(s^\top T(X)\right)\right] = \exp\left(Z(\eta + s) - Z(\eta)\right).$$

For a number of facts from probability and statistics (both related and unrelated to exponential families), we refer the reader to Appendix F.

# B Omitted Proofs from Section 2

## B.1 Omitted Proofs from Section 2.2

*Proof of Lemma 2.2.* Without loss of generality, we assume that all the intervals $I_j$ are non-degenerate. We have:

$$\begin{aligned}
\mathop{\mathbb{E}}_{\eta, X}\left[Z + \|f(X) - (\eta - m)\|_2^2\right] &= \mathop{\mathbb{E}}_{\eta, X}\left[\sum_{j \in [k]}\left\{Z^j + [f_j(X) - (\eta_j - m_j)]^2\right\}\right] \\
&= \sum_{j \in [k]} \mathop{\mathbb{E}}_{\eta, X}\left[Z^j + [f_j(X) - (\eta_j - m_j)]^2\right] \\
&= \sum_{j \in [k]} \mathop{\mathbb{E}}_{\eta_{-j}}\left[\mathop{\mathbb{E}}_{\eta_j, X}\left[Z^j + [f_j(X) - (\eta_j - m_j)]^2\right]\right] \\
&\overset{(a)}{\geq} \sum_{j \in [k]} \mathop{\mathbb{E}}_{\eta_{-j}}\left[\frac{R_j^2}{12}\right] \\
&= \frac{\|R\|_2^2}{12},
\end{aligned}$$

where $(a)$ is by a direct application of Lemma 2.1.

If there exists a degenerate interval $I_j$, we have $Z^j = f_j(X) - (\eta_j - m_j) = R_j = 0$. Thus, it suffices to apply the same argument as before for the non-degenerate intervals. ■

## B.2 Omitted Proofs from Section 2.3

This appendix is dedicated to proving Lemma 2.3 and Theorem 2.4.

To prove the former, in addition to the correlation terms defined in Section 2.1, we introduce one more correlation term. Let $X_i'$ be an independently drawn sample from $p_\eta$ and let $X_{\sim i}$ be the dataset that is the same as $X$ with $X_i$ replaced by $X_i'$. We define:

$$s_j \coloneqq \left[\frac{R_j^2}{4} - (\eta_j - m_j)^2\right]\left[M_j(X_{\sim i}) - (\eta_j - m_j)\right], \forall j \in [k], \tag{4}$$

$$\widetilde{Z}_i \coloneqq \sum_{j \in [k]} s_j\left(T_j(X_i) - \mu_{T,j}\right). \tag{5}$$

This measures the effect that replacing an individual sample has on the output of the estimator. Observe that, conditioning on $\eta$, $X_{\sim i}$ and $X_i$ are independent. This implies that $\mathop{\mathbb{E}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i\right] = 0$ and $\mathop{\mathrm{Var}}_{X_{\sim i},X_i,M}\left(\widetilde{Z}_i\right) = \mathop{\mathbb{E}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i^2\right]$. Additionally, calculating the variance of $\widetilde{Z}_i$ requires us to reason about the covariance matrix of the sufficient statistics $\Sigma_T$, since the calculation involves terms of the form $\left(T_{j_1}\left(X_i\right) - \mu_{T,j_1}\right)\left(T_{j_2}\left(X_i\right) - \mu_{T,j_2}\right), j_1, j_2 \in [k]$.

The previous remarks are summarized in the following lemma.

**Lemma B.1.** *Conditioning on $\eta$, it holds that:*

$$\begin{aligned}
&\mathop{\mathbb{E}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i\right] = 0,\\
&\mathop{\mathrm{Var}}_{X_{\sim i},X_i,M}\left(\widetilde{Z}_i\right) = \mathop{\mathbb{E}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i^2\right] = \mathop{\mathbb{E}}_{X_{\sim i},M}\left[s^\top \Sigma_T s\right].
\end{aligned}$$

*Proof.* We start by computing the mean. We have:

$$\begin{aligned}
\mathop{\mathbb{E}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i\right] &= \mathop{\mathbb{E}}_{X_{\sim i},X_i,M}\left[\sum_{j\in[k]} s_j\left(T_j\left(X_i\right) - \mu_{T,j}\right)\right]\\
&= \sum_{j\in[k]} \mathop{\mathbb{E}}_{X_{\sim i},X_i,M}\left[s_j\left(T_j\left(X_i\right) - \mu_{T,j}\right)\right]\\
&\overset{(a)}{=} \sum_{j\in[k]} \mathop{\mathbb{E}}_{X_{\sim i},M}\left[s_j\right] \cancelto{0}{\mathop{\mathbb{E}}_{X_i}\left[T_j\left(X_i\right) - \mu_{T,j}\right]}\\
&= 0,
\end{aligned}$$

where $(a)$ relies on the independence of $s_j = \left[\frac{R_j^2}{4} - \left(\eta_j - m_j\right)^2\right]\left[M_j\left(X_{\sim i}\right) - \left(\eta_j - m_j\right)\right]$ and $T_j\left(X_i\right) - \mu_{T,j}$ conditioned on $\eta$.

We now proceed with the variance:

$$\begin{aligned}
\mathop{\mathrm{Var}}_{X_{\sim i},X_i,M}\left(\widetilde{Z}_i\right) &= \mathop{\mathbb{E}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i^2\right]\\
&= \mathop{\mathbb{E}}_{X_{\sim i},X_i,M}\left[\sum_{j_1\in[k]}\sum_{j_2\in[k]} s_{j_1}s_{j_2}\left(T_{j_1}\left(X_i\right) - \mu_{T,j_1}\right)\left(T_{j_2}\left(X_i\right) - \mu_{T,j_2}\right)\right]\\
&= \sum_{j_1\in[k]}\sum_{j_2\in[k]} \mathop{\mathbb{E}}_{X_{\sim i},M}\left[s_{j_1}s_{j_2}\right] \mathop{\mathbb{E}}_{X_i}\left[\left(T_{j_1}\left(X_i\right) - \mu_{T,j_1}\right)\left(T_{j_2}\left(X_i\right) - \mu_{T,j_2}\right)\right]\\
&\overset{(a)}{=} \sum_{j_1\in[k]}\sum_{j_2\in[k]} \mathop{\mathbb{E}}_{X_{\sim i},M}\left[s_{j_1}s_{j_2}\right]\Sigma_{T,j_1j_2}\\
&= \mathop{\mathbb{E}}_{X_{\sim i},M}\left[\sum_{j_1\in[k]}\sum_{j_2\in[k]} s_{j_1}\Sigma_{T,j_1j_2}s_{j_2}\right]\\
&= \mathop{\mathbb{E}}_{X_{\sim i},M}\left[s^\top\Sigma_T s\right].
\end{aligned}$$

where $(a)$ is by the second part of Proposition A.6. ■

Now, we show how to upper-bound $\mathop{\mathbb{E}}_{X,M}\left[Z_i\right]$ using properties of differential privacy. We appeal to the definition of DP to relate $\mathop{\mathbb{E}}_{X,M}\left[Z_i\right]$ with $\mathop{\mathbb{E}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i\right]$ and $\mathop{\mathbb{E}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i^2\right]$. The proof involves splitting $Z_i$ in its positive and negative components and applying the definition of DP to them. The main steps of the proof are standard (e.g., see [SU17b]), but we include everything for completeness.

*Proof of Lemma 2.3.* Consider the random variables $Z_{i,+} = \max\{Z_i, 0\}$ and $Z_{i,-} = -\min\{Z_i, 0\}$. These random variables are non-negative, so for any $T > 0$, we have:

$$\begin{aligned}
\mathop{\mathbb{E}}_{X,M}[Z_i] &= \mathop{\mathbb{E}}_{X,M}[Z_{i,+}] - \mathop{\mathbb{E}}_{\eta,X,M}[Z_{i,-}] \\
&= \int_0^\infty \mathop{\mathbb{P}}_{X,M}[Z_{i,+} > t]\, dt - \int_0^\infty \mathop{\mathbb{P}}_{X,M}[Z_{i,-} > t]\, dt \\
&= \int_0^\infty \mathop{\mathbb{P}}_{X,M}[Z_i > t]\, dt - \int_0^\infty \mathop{\mathbb{P}}_{X,M}[Z_i < -t]\, dt \\
&= \int_0^T \mathop{\mathbb{P}}_{X,M}[Z_i > t]\, dt - \int_0^T \mathop{\mathbb{P}}_{X,M}[Z_i < -t]\, dt \\
&\quad + \int_T^\infty \mathop{\mathbb{P}}_{X,M}[Z_i > t]\, dt - \int_T^\infty \mathop{\mathbb{P}}_{X,M}[Z_i < -t]\, dt \qquad (6)
\end{aligned}$$

We focus our attention on the first two terms of (6). We consider the sets $S_{+,X_i}, S_{-,X_i} \subseteq \mathbb{R}^k$:

$$S_{+,X_i} := \left\{ Y \in \mathbb{R}^k \colon \sum_{j\in[k]} \left[ \frac{R_j^2}{4} - (\eta_j - m_j)^2 \right] [Y_j - (\eta_j - m_j)] \left(T_j(X_i) - \mu_{T,j}\right) \geq t \right\},$$

$$S_{-,X_i} := \left\{ Y \in \mathbb{R}^k \colon \sum_{j\in[k]} \left[ \frac{R_j^2}{4} - (\eta_j - m_j)^2 \right] [Y_j - (\eta_j - m_j)] \left(T_j(X_i) - \mu_{T,j}\right) \leq -t \right\}.$$

By a direct application of the definition of DP on $S_{+,X_i}$ and $S_{-,X_i}$[4] for the neighboring datasets $X$ and $X_{\sim i}$ we have:

$$\begin{aligned}
&\int_0^T \mathop{\mathbb{P}}_{X,M}[Z_i > t]\, dt - \int_0^T \mathop{\mathbb{P}}_{X,M}[Z_i < -t]\, dt \\
&= \int_0^T \mathop{\mathbb{P}}_{X,M}[M(X) \in S_{+,X_i}]\, dt - \int_0^T \mathop{\mathbb{P}}_{X,M}[M(X) \in S_{-,X_i}]\, dt \\
&\leq \int_0^T \left( e^{\varepsilon} \mathop{\mathbb{P}}_{X_{\sim i},X_i,M}[M(X_{\sim i}) \in S_{+,X_i}] + \delta \right) dt \\
&\quad - \int_0^T e^{-\varepsilon} \left( \mathop{\mathbb{P}}_{X_{\sim i},X_i,M}[M(X_{\sim i}) \in S_{-,X_i}] - \delta \right) dt \\
&\leq \int_0^T \left( e^{\varepsilon} \mathop{\mathbb{P}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i > t\right] + \delta \right) dt - \int_0^T e^{-\varepsilon} \left( \mathop{\mathbb{P}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i < -t\right] - \delta \right) dt \\
&= \left(1 + e^{-\varepsilon}\right) \delta T + e^{\varepsilon} \int_0^T \mathop{\mathbb{P}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i > t\right] dt - e^{-\varepsilon} \int_0^T \mathop{\mathbb{P}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i < -t\right] dt.
\end{aligned}$$

[4]These are random subsets of $\mathbb{R}^k$, since they depend on the random vector $X_i$ ($\eta$ is assumed to be fixed). This does not affect the way the definition of privacy is applied.

Under the assumption that $\varepsilon \in [0,1]$, it holds that $e^{\varepsilon} \leq 1 + 2\varepsilon$ and $1 \geq e^{-\varepsilon} \geq 1 - \varepsilon \geq 1 - 2\varepsilon$. Combined with the above, this yields:

$$
\begin{aligned}
&\int_0^T \mathop{\mathbb{P}}_{X,M}\left[Z_i > t\right]\, dt - \int_0^T \mathop{\mathbb{P}}_{X,M}\left[Z_i < -t\right]\, dt \\
&\leq 2\delta T + (2\varepsilon + 1)\int_0^T \mathop{\mathbb{P}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i > t\right]\, dt + (2\varepsilon - 1)\int_0^T \mathop{\mathbb{P}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i < -t\right]\, dt \\
&= 2\delta T + 2\varepsilon \int_0^T \left(\mathop{\mathbb{P}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i > t\right] + \mathop{\mathbb{P}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i < -t\right]\right) dt \\
&\quad + \int_0^T \left(\mathop{\mathbb{P}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i > t\right] - \mathop{\mathbb{P}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i < -t\right]\right) dt \\
&= 2\delta T + 2\varepsilon \int_0^T \mathop{\mathbb{P}}_{X_{\sim i},X_i,M}\left[\left|\widetilde{Z}_i\right| > t\right]\, dt + \int_0^T \left(\mathop{\mathbb{P}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i > t\right] - \mathop{\mathbb{P}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i < -t\right]\right) dt \\
&\leq 2\delta T + 2\varepsilon \int_0^\infty \mathop{\mathbb{P}}_{X_{\sim i},X_i,M}\left[\left|\widetilde{Z}_i\right| > t\right]\, dt + \int_0^T \left(\mathop{\mathbb{P}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i > t\right] - \mathop{\mathbb{P}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i < -t\right]\right) dt \\
&= 2\delta T + 2\varepsilon \mathop{\mathbb{E}}_{X_{\sim i},X_i,M}\left[\left|\widetilde{Z}_i\right|\right] + \int_0^T \left(\mathop{\mathbb{P}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i > t\right] - \mathop{\mathbb{P}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i < -t\right]\right) dt \\
&= 2\delta T + 2\varepsilon \mathop{\mathbb{E}}_{X_{\sim i},X_i,M}\left[\left|\widetilde{Z}_i\right|\right] + \int_0^\infty \left(\mathop{\mathbb{P}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i > t\right] - \mathop{\mathbb{P}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i < -t\right]\right) dt \\
&\quad - \int_T^\infty \left(\mathop{\mathbb{P}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i > t\right] - \mathop{\mathbb{P}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i < -t\right]\right) dt.
\end{aligned}
$$

Defining $\widetilde{Z}_{i,+} = \max\left\{\widetilde{Z}_i, 0\right\}$ and $\widetilde{Z}_{i,-} = -\min\left\{\widetilde{Z}_i, 0\right\}$, we get:

$$
\begin{aligned}
&\int_0^T \mathop{\mathbb{P}}_{X,M}\left[Z_i > t\right]\, dt - \int_0^T \mathop{\mathbb{P}}_{X,M}\left[Z_i < -t\right]\, dt \\
&\leq 2\delta T + 2\varepsilon \mathop{\mathbb{E}}_{X_{\sim i},X_i,M}\left[\left|\widetilde{Z}_i\right|\right] + \left(\mathop{\mathbb{E}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_{i,+}\right] - \mathop{\mathbb{E}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_{i,-}\right]\right) \\
&\quad - \int_T^\infty \left(\mathop{\mathbb{P}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i > t\right] - \mathop{\mathbb{P}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i < -t\right]\right) dt \\
&\leq 2\delta T + 2\varepsilon \mathop{\mathbb{E}}_{X_{\sim i},X_i,M}\left[\left|\widetilde{Z}_i\right|\right] + \cancelto{0}{\mathop{\mathbb{E}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i\right]} \\
&\quad - \int_T^\infty \left(\mathop{\mathbb{P}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i > t\right] - \mathop{\mathbb{P}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i < -t\right]\right) dt \\
&= 2\delta T + 2\varepsilon \mathop{\mathbb{E}}_{X_{\sim i},X_i,M}\left[\left|\widetilde{Z}_i\right|\right] - \int_T^\infty \left(\mathop{\mathbb{P}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i > t\right] - \mathop{\mathbb{P}}_{X_{\sim i},X_i,M}\left[\widetilde{Z}_i < -t\right]\right) dt, \qquad (7)
\end{aligned}
$$

where we appealed to Lemma B.1 along the way.

The second version of Fact F.4 for $X = \left|\widetilde{Z}_i\right|, Y = 1$, and Lemma B.1 yield:

$$\mathop{\mathbb{E}}_{X_{\sim i}, X_i, M}\left[\left|\widetilde{Z}_i\right|\right] \leq \sqrt{\mathop{\mathbb{E}}_{X_{\sim i}, X_i, M}\left[\widetilde{Z}_i^2\right]} = \sqrt{\mathop{\mathbb{E}}_{X_{\sim i}, M}\left[s^\top \Sigma_T s\right]}. \tag{8}$$

Substituting (7) and (8) into (6), we get:

$$\begin{aligned}
\mathop{\mathbb{E}}_{X,M}\left[Z_i\right] &\leq 2\delta T + 2\varepsilon\sqrt{\mathop{\mathbb{E}}_{X_{\sim i}, M}\left[s^\top \Sigma_T s\right]} + \int_T^\infty \left(\mathop{\mathbb{P}}_{X,M}\left[Z_i > t\right] - \mathop{\mathbb{P}}_{X,M}\left[Z_i < -t\right]\right) dt \\
&\quad + \int_T^\infty \left(\mathop{\mathbb{P}}_{X_{\sim i}, X_i, M}\left[\widetilde{Z}_i < -t\right] - \mathop{\mathbb{P}}_{X_{\sim i}, X_i, M}\left[\widetilde{Z}_i > t\right]\right) dt \\
&\leq 2\delta T + 2\varepsilon\sqrt{\mathop{\mathbb{E}}_{X_{\sim i}, M}\left[s^\top \Sigma_T s\right]} + \int_T^\infty \left(\mathop{\mathbb{P}}_{X,M}\left[\left|Z_i\right| > t\right] + \mathop{\mathbb{P}}_{X_{\sim i}, X_i, M}\left[\left|\widetilde{Z}_i\right| > t\right]\right) dt.
\end{aligned} \tag{9}$$

We now show how to bound each of the tail probabilities involving $Z_i$ and $\widetilde{Z}_i$. For $Z_i$, we have:

$$\begin{aligned}
\left|Z_i\right| &= \left|\sum_{j\in[k]}\left[\frac{R_j^2}{4} - \left(\eta_j - m_j\right)^2\right]\left[M_j\left(X\right) - \left(\eta_j - m_j\right)\right]\left(T_j\left(X_i\right) - \mu_{T,j}\right)\right| \\
&\leq \sum_{j\in[k]}\left[\frac{R_j^2}{4} - \left(\eta_j - m_j\right)^2\right]\left|M_j\left(X\right) - \left(\eta_j - m_j\right)\right|\left|T_j\left(X_i\right) - \mu_{T,j}\right| \\
&\overset{(a)}{\leq} \frac{\left\|R\right\|_\infty^3}{4}\sum_{j\in[k]}\left|T_j\left(X_i\right) - \mu_{T,j}\right| \\
&\overset{(b)}{\leq} \frac{\left\|R\right\|_\infty^3}{4}\sqrt{k}\left\|T\left(X_i\right) - \mu_T\right\|_2,
\end{aligned}$$

where $(a)$ relies on the assumption $\left|M_j\left(X_i\right) - \left(\eta_j - m_j\right)\right| \leq R_j, \forall j \in [k]$, and $(b)$ uses the first version of Fact F.4 for $x = \vec{1}_k, y_j = \left|T_j\left(X_i\right) - \mu_{T,j}\right|, \forall j \in [k]$.

Thus, conditioning on the dataset $X$, $Z_i$ is a bounded random variable supported on a 0-centered interval. Consequently, the worst-case upper bound we have for $\int_T^\infty \mathop{\mathbb{P}}_{X,M}\left[\left|Z_i\right| > t\right] dt$ is attained when all the mass of $Z_i$ is concentrated on $\left\{\pm\frac{\left\|R\right\|_\infty^3}{4}\sqrt{k}\left\|T\left(X_i\right) - \mu_T\right\|_2\right\}$ for every dataset $X$. Using this observation, we get:

$$\begin{aligned}
\int_T^\infty \mathop{\mathbb{P}}_{X,M}\left[\left|Z_i\right| > t\right] dt &\leq \int_T^\infty \left[\int_{\mathbb{R}^d}\cdots\int_{\mathbb{R}^d}\left(\prod_{\ell\in[n]} p_\eta\left(x_\ell\right)\right)\mathbb{1}\left\{\left\|T\left(X_i\right) - \mu_T\right\|_2 > \frac{4t}{\left\|R\right\|_\infty^3\sqrt{k}}\right\} dx_1\ldots dx_n\right] dt \\
&= \int_T^\infty \int_{\mathbb{R}^d} p_\eta\left(x_i\right)\mathbb{1}\left\{\left\|T\left(X_i\right) - \mu_T\right\|_2 > \frac{4t}{\left\|R\right\|_\infty^3}\sqrt{k}\right\} dx_i dt \\
&= \int_T^\infty \mathop{\mathbb{P}}_{X_i}\left[\left\|T\left(X_i\right) - \mu_T\right\|_2 > \frac{4t}{\left\|R\right\|_\infty^3\sqrt{k}}\right] dt.
\end{aligned} \tag{10}$$

Upper-bounding the term involving $\widetilde{Z}_i$ in similar fashion and substituting (10) to (9), we get:

$$\mathop{\mathbb{E}}_{X,M}\left[Z_i\right] \leq 2\delta T + 2\varepsilon\sqrt{\mathop{\mathbb{E}}_{X_{\sim i}, M}\left[s^\top \Sigma_T s\right]} + 2\int_T^\infty \mathop{\mathbb{P}}_{X_i}\left[\left\|T\left(X_i\right) - \mu_T\right\|_2 > \frac{4t}{\left\|R\right\|_\infty^3\sqrt{k}}\right] dt.$$

■

**Remark B.2.** The last term in the upper bound of Lemma 2.3 involves the tail of the random variable $\|T(X_i) - \mu_T\|_2$. The obvious way to bound that term is:

$$\int_T^\infty \mathop{\mathbb{P}}_{X_i}\left[\|T(X_i) - \mu_T\|_2^2 > \frac{16t^2}{\|R\|_\infty^6 k}\right] dt \leq \int_T^\infty \frac{\mathop{\mathbb{E}}_{X_i}\left[\|T(X_i) - \mu_T\|_2^2\right]}{\frac{16t^2}{\|R\|_\infty^6 k}} dt = \frac{\|R\|_\infty^6 k \mathrm{tr}(\Sigma_T)}{16} \int_T^\infty \frac{dt}{t^2}$$
$$= \frac{\|R\|_\infty^6 k \mathrm{tr}(\Sigma_T)}{16T}.$$

Following the steps we sketched previously, this approach leads to a range for $\delta$ of the form $\delta \leq \widetilde{\mathcal{O}}\left(\frac{1}{n^2}\right)$. However, when reasoning about the cost of approximate DP for various statistical tasks, in order to claim that a full characterization has been achieved, it is generally the case that lower bounds must be shown for $\delta$ as large as $\widetilde{\mathcal{O}}\left(\frac{1}{n}\right)$. To achieve this, we will apply the following strategy. First, observe that we can approximate $\|T(X) - \mu_T\|_2$ using an $\varepsilon$-net (see Definition E.1). In particular, if $\mathcal{N}_{\frac{1}{4}}$ is a $\frac{1}{4}$-net of $\mathbb{S}^{k-1}$ of minimum cardinality, using Facts E.3, E.4, and a union bound, we get that:

$$\int_T^\infty \mathop{\mathbb{P}}_{X_i}\left[\|T(X_i) - \mu_T\|_2 > \frac{4t}{\|R\|_\infty^3 \sqrt{k}}\right] dt \leq \int_T^\infty \mathop{\mathbb{P}}_{X_i}\left[\frac{1}{1-\frac{1}{4}} \cdot \sup_{v \in \mathcal{N}_{\frac{1}{4}}} |\langle v, T(X_i) - \mu_T\rangle| > \frac{4t}{\|R\|_\infty^3 \sqrt{k}}\right] dt$$
$$= \int_T^\infty \mathop{\mathbb{P}}_{X_i}\left[\sup_{v \in \mathcal{N}_{\frac{1}{4}}} |\langle v, T(X_i) - \mu_T\rangle| > \frac{3t}{\|R\|_\infty^3 \sqrt{k}}\right] dt$$
$$\leq \left|\mathcal{N}_{\frac{1}{4}}\right| \int_T^\infty \mathop{\mathbb{P}}_{X_i}\left[|\langle v, T(X_i) - \mu_T\rangle| > \frac{3t}{\|R\|_\infty^3 \sqrt{k}}\right] dt$$
$$\leq 9^k \int_T^\infty \mathop{\mathbb{P}}_{X_i}\left[|\langle v, T(X_i) - \mu_T\rangle| > \frac{3t}{\|R\|_\infty^3 \sqrt{k}}\right] dt.$$

Now, to reason about the tail probabilities of $|\langle v, T(X_i) - \mu_T\rangle|$, it suffices to leverage the fact that the MGFs of univariate projections of the sufficient statistics of exponential families are well-defined (see the third part of Proposition A.6). This renders it possible to prove Chernoff-type bounds in the context that we are considering. In practice, we will not have to prove our own bounds, instead appealing to various known bounds from the literature.

We conclude this appendix with the complete proof of Theorem 2.4.

*Proof of Theorem 2.4.* By Lemma 2.2 and (3), we have:

$$\mathop{\mathbb{E}}_{\eta,X,M}[Z] \geq \frac{\|R\|_2^2}{24}.$$

Additionally, by Lemma 2.3, we have:

$$\mathop{\mathbb{E}}_{\eta,X,M}[Z] = \sum_{i\in[n]} \mathop{\mathbb{E}}_{\eta,X,M}[Z_i]$$
$$= \sum_{i\in[n]} \mathop{\mathbb{E}}_{\eta}\left[\mathop{\mathbb{E}}_{X,M}[Z_i]\right]$$
$$\leq \sum_{i\in[n]} \mathop{\mathbb{E}}_{\eta}\left[2\delta T + 2\varepsilon\sqrt{\mathop{\mathbb{E}}_{X_{\sim i},M}[s^\top \Sigma_T s]} + 2\int_T^\infty \mathop{\mathbb{P}}_{X_i}\left[\|T(X_i) - \mu_T\|_2 > \frac{4t}{\|R\|_\infty^3 \sqrt{k}}\right] dt\right]$$
$$\leq n\left\{2\delta T + 2\varepsilon \mathop{\mathbb{E}}_{\eta}\left[\sqrt{\mathop{\mathbb{E}}_{X_{\sim i},M}[s^\top \Sigma_T s]}\right] + 2\int_T^\infty \mathop{\mathbb{P}}_{X_i}\left[\|T(X_i) - \mu_T\|_2 > \frac{4t}{\|R\|_\infty^3 \sqrt{k}}\right] dt\right\}.$$

Combining the two, we get the desired result. ■

# C Omitted Proofs from Section 3

## C.1 Omitted Proofs from Section 3.1

The first step of the proof is to write the Gaussian $\mathcal{N}(0,\Sigma)$ as an exponential family. By Fact F.3, we have that the density can be written in the form:

$$\begin{aligned} p_\eta(x) &= h(x)\exp\left(\eta^\top T(x) - Z(\eta)\right), \\ h(x) &= 1, \\ T(x) &= -\frac{1}{2}\left(xx^\top\right)^\flat, \\ \eta &= \left(\Sigma^{-1}\right)^\flat, \\ Z(\eta) &= \frac{d}{2}\ln(2\pi) - \frac{1}{2}\ln\left(\det\left(\eta^\#\right)\right). \end{aligned}$$

Observe that the natural parameter vector $\eta$ is the flattening of the *precision matrix* $\Sigma^{-1}$. However, we cannot use this representation, because $\Sigma^{-1}$ is symmetric. Indeed, this implies that the components of $\eta$ cannot be generated independently, as is required by Theorem 2.4. For that reason, we consider an alternative parameterization.

**Lemma C.1.** *Let $\mathcal{N}(0,\Sigma)$ be a $d$-dimensional Gaussian with unknown covariance $\Sigma$. Then, it belongs to an exponential family with $h_0 \equiv h, \eta_0 = 2U^\flat, T_0 \equiv T, Z_0(\eta_0) = \frac{d}{2}\ln(2\pi) - \frac{1}{2}\ln\left(\det\left(\frac{\eta_0^\# + \left(\eta_0^\#\right)^\top}{2}\right)\right)$ where $U \in \mathbb{R}^{d\times d}$ is an upper triangular matrix with:*

$$U_{ij} := \begin{cases} \frac{1}{2}\left(\Sigma^{-1}\right)_{ii}, & i = j \\ \left(\Sigma^{-1}\right)_{ij}, & j > i \\ 0, & i > j \end{cases}.$$

*In particular, it holds that $Z(\eta) = Z_0(\eta_0)$ and $\eta^\top T(x) = \eta_0^\top T_0(x)$.*

*Proof.* Observe that $\eta^\# = \Sigma^{-1} = U + U^\top = \frac{\eta_0^\# + \left(\eta_0^\#\right)^\top}{2} \implies Z(\eta) = Z_0(\eta_0)$. Additionally, we have:

$$\begin{aligned} \eta^\top T(x) = \left\langle \left(\Sigma^{-1}\right)^\flat, -\frac{1}{2}\left(xx^\top\right)^\flat \right\rangle = -\frac{1}{2}\left\langle \Sigma^{-1}, xx^\top \right\rangle &= -\frac{1}{2}x^\top\Sigma^{-1}x \\ &= -\frac{1}{2}x^\top\left(U + U^\top\right)x \\ &= -\frac{1}{2}x^\top Ux - \frac{1}{2}x^\top U^\top x \\ &\overset{(a)}{=} -x^\top Ux \\ &= -\left\langle U, xx^\top \right\rangle \\ &= \left\langle 2U^\flat, -\frac{1}{2}\left(xx^\top\right)^\flat \right\rangle \\ &= \eta_0^\top T_0(x), \end{aligned}$$

where $(a)$ used the fact that $x^\top U^\top x \in \mathbb{R} \implies x^\top U^\top x = \left(x^\top U^\top x\right)^\top = x^\top Ux$.

Consequently, we have $p_\eta(x) = h(x)\exp\left(\eta^\top T(x) - Z(\eta)\right) = h_0(x)\exp\left(\eta_0^\top T_0(x) - Z_0(\eta_0)\right)$, verifying the equivalence of the two parameterizations. ■

We now consider the following process to generate $\eta_0$. Roughly speaking, modulo the parameterization issues mentioned above, the process consists of sampling the elements of the precision matrix uniformly at random from intervals of width $\frac{1}{2d}$. For the diagonal elements, these intervals are centered at $\frac{3}{4}$, whereas for the non-diagonal elements they are centered at the origin. Similar constructions have been employed previously, e.g., in [KLSU19].

**Algorithm 1** Covariance Matrix Sampling

**Input:** $d \in \mathbb{N}$.
**Output:** A pair $(\Sigma, \eta_0)$.

**procedure** COVSAMP($d$)
  **for** $i \in [d]$ **do**
    **for** $j \in [d] \setminus [i-1]$ **do**
      **if** $i = j$ **then**
        Draw an independent sample $\left(\Sigma^{-1}\right)_{ii} \sim \mathcal{U}\left[\frac{3}{4} \pm \frac{1}{4d}\right]$.
        Let $U_{ii} = \frac{1}{2}\left(\Sigma^{-1}\right)_{ii}$.
      **else**
        Draw an independent sample $\left(\Sigma^{-1}\right)_{ij} \sim \mathcal{U}\left[\pm\frac{1}{4d}\right]$.
        Let $\left(\Sigma^{-1}\right)_{ji} = \left(\Sigma^{-1}\right)_{ij}$.
        Let $U_{ij} = \left(\Sigma^{-1}\right)_{ij}$ and $U_{ji} = 0$.
      **end if**
    **end for**
  **end for**
  Let $\eta_0 = 2U^\flat$.
  **return** $(\Sigma, \eta_0)$.
**end procedure**

We proceed to identify some properties of the matrices generated by the above process that will help us to prove Theorem 3.1.

**Lemma C.2.** *Let $(\Sigma, \eta_0) \in \mathbb{R}^{d\times d} \times \mathbb{R}^{d^2}$ be the output of Algorithm 1. We have that:*

1. $\mathbb{I} \preceq \Sigma \preceq 2\mathbb{I}$.
2. *The components of $\eta_0$ are supported on intervals of the form:*
$$I_\ell = \begin{cases} \left[\frac{3}{4} \pm \frac{1}{4d}\right], & (\ell-1) \bmod d = (\ell-1) \operatorname{div} d \\ \left[\pm\frac{1}{2d}\right], & (\ell-1) \bmod d > (\ell-1) \operatorname{div} d \\ \{0\}, & (\ell-1) \bmod d < (\ell-1) \operatorname{div} d \end{cases}.$$
3. *The components of $\eta_0$ are independent and it holds that $\|R\|_\infty = \frac{1}{d}$ and $\|R\|_2^2 = \frac{1}{2}\left(1 - \frac{1}{2d}\right) \geq \frac{1}{4}$.*
4. *The vector $m \in \mathbb{R}^{d^2}$ is equal to:*
$$m_\ell = \begin{cases} \frac{3}{4}, & \ell = (k-1)\,d + k, \text{ for some } k \in [d] \\ 0, & \text{otherwise} \end{cases}.$$

*Proof.*

1. By a direct application of Theorem E.10, we get that the eigenvalues of $\Sigma^{-1}$ all lie in the interval $\left[\frac{1}{2}, 1\right]$. Thus, we have $\frac{1}{2}\mathbb{I} \preceq \Sigma^{-1} \preceq \mathbb{I} \iff \mathbb{I} \preceq \Sigma \preceq 2\mathbb{I}$.
2. Observe that, the first of the three branches in the definition corresponds to the diagonal elements of $\eta_0^\#$, the second corresponds to those above the diagonal, and the third corresponds to those below the diagonal. The result follows directly from this remark.
3. The non-constant components of $\eta_0$ are all independent draws from a uniform distribution, which immediately yields the independence property. In addition to that, we have $d$ non-constant components taking values in an interval of length $\frac{1}{2d}$, and $\frac{d(d-1)}{2}$ non-constant components take values in an interval of length $\frac{1}{d}$. This implies that $\|R\|_\infty = \frac{1}{d}$ and $\|R\|_2^2 = d \cdot \frac{1}{4d^2} + \frac{d(d-1)}{2} \cdot \frac{1}{d^2} = \frac{1}{2}\left(1 - \frac{1}{2d}\right)$.
4. Observe that the non-diagonal elements of $\eta_0^\# = 2U$ are either constant and equal to $0$ or take values in a $0$-centered interval. This implies that the only non-zero elements of $m$ are those corresponding to the diagonal elements of $\left(\Sigma^{-1}\right)_{ii}$. This yields the desired result.

■

**Remark C.3.** Note that Theorem 3.1 holds only for $\alpha = \mathcal{O}(1)$. In contrast, prior fingerprinting lower bounds for mean estimation of product distributions and Gaussians gave non-trivial results in the low-accuracy regime $\alpha = \mathcal{O}\left(\sqrt{d}\right)$ (see, e.g., the results in Section G). In principle, it may be able to achieve similar lower bounds for covariance estimation in Frobenius norm, as the Frobenius diameter of the set of matrices that satisfy $\mathbb{I} \preceq \Sigma \preceq 2\mathbb{I}$ is equal to $\sqrt{d}$. However, by the third part of the above lemma, our construction only has an upper bound on this diameter of $\mathcal{O}(1)$. This makes estimation trivial for any $\alpha$ which is larger, as one could output an arbitrary parameter vector within the set. Thus, to prove a lower bound for all $\alpha = \mathcal{O}\left(\sqrt{d}\right)$, one must consider a different construction, and it is not merely a deficiency of our analysis.

We show that, if one has an estimator for the covariance matrix $\Sigma$, then this implies an estimator for the natural parameter vector $\eta_0$. Therefore, a lower bound for the latter problem implies a lower bound for the former, allowing us to focus on lower bounds for estimating the natural parameter vector. More precisely, we present an $(\varepsilon, \delta)$-DP mechanism $T_M \colon \mathbb{R}^{n\times d} \to \bigotimes_{j\in[d^2]} I_j$ that satisfies the guarantee $\mathop{\mathbb{E}}_{X,T_M}\left[\|T_M(X) - (\eta_0 - m)\|_2^2\right] \leq \alpha^2$. $T_M$ assumes the existence of an $(\varepsilon, \delta)$-DP mechanism $M \colon \mathbb{R}^{n\times d} \to \mathbb{R}^{d\times d}$ such that $\mathop{\mathbb{E}}_{X,M}\left[\|M(X) - \Sigma\|_\Sigma^2\right] \leq \frac{\alpha^2}{8}$.

**Algorithm 2** From Covariance Estimation to Natural Parameter Estimation

**Input:** $X = (X_1, \ldots, X_n) \sim \mathcal{N}(0, \Sigma)^{\otimes n}$ and a mechanism $M \colon \mathbb{R}^{n\times d} \to \mathbb{R}^{d\times d}$.
**Output:** $T_M(X) \in \bigotimes_{j\in[d^2]} I_j$.

**procedure** $T_M(X)$
  Let $\widehat{\Sigma} = M(X)$.
  **for** $i \in [d]$ **do** ▷ $\widehat{\Sigma}^{-1}$ may not be equal to any possible $\Sigma^{-1}$ generated by Algorithm 1.
    **for** $j \in [d]$ **do**
      Let $\widetilde{\Sigma}_{ij}^{-1}$ be the projection of $\left(\widehat{\Sigma}^{-1}\right)_{ij}$ onto the support of $\Sigma^{-1}$.
    **end for**
  **end for**
  Let $\widetilde{U} \in \mathbb{R}^{d\times d}$ be an upper triangular matrix such that $\widetilde{\Sigma}^{-1} = \widetilde{U} + \widetilde{U}^\top$.
  Let $T_M(X) = 2\widetilde{U}^\flat - m$.
  **return** $T_M(X)$.
**end procedure**

**Lemma C.4.** *Let $\Sigma \in \mathbb{R}^{d\times d}$ be a covariance generated by Algorithm 1, and $X \sim \mathcal{N}(0, \Sigma)^{\otimes n}$ be a dataset. If $M \colon \mathbb{R}^{n\times d} \to \mathbb{R}^{d\times d}$ is an $(\varepsilon, \delta)$-DP mechanism with $\mathop{\mathbb{E}}_{X,M}\left[\|M(X) - \Sigma\|_\Sigma^2\right] \leq \frac{\alpha^2}{8} \leq \frac{1}{32}$, then $T_M \colon \mathbb{R}^{n\times d} \to \bigotimes_{j\in[d^2]} I_j$ is also $(\varepsilon, \delta)$-DP and satisfies $\mathop{\mathbb{E}}_{X,T_M}\left[\|T_M(X) - (\eta_0 - m)\|_2^2\right] \leq \alpha^2 \leq \frac{1}{4}$.*

*Proof.* The privacy guarantee of $T_M$ is an immediate consequence of the guarantee of $M$ and Lemma A.3. Additionally, the only randomness used by $T_M$ is that used by $M$. Having established that, we focus our attention on the accuracy guarantee. We have:

$$\mathop{\mathbb{E}}_{X,T_M}\left[\|T_M(X) - (\eta_0 - m)\|_2^2\right] = 4 \mathop{\mathbb{E}}_{X,M}\left[\left\|\widetilde{U} - U\right\|_F^2\right].$$

We remark now that:

$$\left\|\widetilde{U} - U\right\|_F^2 = \sum_{i\in[d]}\sum_{j\in[d]}\left(\widetilde{U}_{ij} - U_{ij}\right)^2 = \sum_{j>i}\left(\widetilde{U}_{ij} - U_{ij}\right)^2 = \sum_{j>i}\left(\widetilde{\Sigma}_{ij}^{-1} - \left(\Sigma^{-1}\right)_{ij}\right)^2,$$

where we relied on the fact that the diagonal elements of $U$ can take only a single value.

Consequently, we get:

$$\begin{aligned}
\mathop{\mathbb{E}}_{X,T_M}\left[\left\|T_M\left(X\right)-\left(\eta_0-m\right)\right\|_2^2\right] &= 4\mathop{\mathbb{E}}_{X,M}\left[\sum_{j>i}\left(\widetilde{\Sigma}_{ij}^{-1}-\left(\Sigma^{-1}\right)_{ij}\right)^2\right] \\
&= 2\mathop{\mathbb{E}}_{X,M}\left[\sum_{i\neq j}\left(\widetilde{\Sigma}_{ij}^{-1}-\left(\Sigma^{-1}\right)_{ij}\right)^2\right] \\
&\overset{(a)}{\leq} 2\mathop{\mathbb{E}}_{X,M}\left[\sum_{i=1}^{d}\sum_{j=1}^{d}\left(\left(\widehat{\Sigma}^{-1}\right)_{ij}-\left(\Sigma^{-1}\right)_{ij}\right)^2\right] \\
&= 2\mathop{\mathbb{E}}_{X,M}\left[\left\|\widehat{\Sigma}^{-1}-\Sigma^{-1}\right\|_F^2\right], \qquad (11)
\end{aligned}$$

where $(a)$ uses the fact that the projection we perform in Line 5 ensures that $\left|\widetilde{\Sigma}_{ij}^{-1}-\left(\Sigma^{-1}\right)_{ij}\right| \leq \left|\left(\widehat{\Sigma}^{-1}\right)_{ij}-\left(\Sigma^{-1}\right)_{ij}\right|, \forall i,j\in[d]$.

Now, observe that:

$$\begin{aligned}
\left\|\widehat{\Sigma}^{-1}-\Sigma^{-1}\right\|_F = \left\|\Sigma^{-\frac{1}{2}}\left(\mathbb{I}-\Sigma^{\frac{1}{2}}\widehat{\Sigma}^{-1}\Sigma^{\frac{1}{2}}\right)\Sigma^{-\frac{1}{2}}\right\|_F &\overset{(a)}{\leq} \left\|\Sigma^{-\frac{1}{2}}\right\|_2^2\left\|\mathbb{I}-\Sigma^{\frac{1}{2}}\widehat{\Sigma}^{-1}\Sigma^{\frac{1}{2}}\right\|_F \\
&\overset{(b)}{\leq} \left\|\Sigma^{-1}-\widehat{\Sigma}^{-1}\right\|_{\Sigma^{-1}} \\
&= \left\|\Sigma^{-1}-\widehat{\Sigma}^{-1}\right\|_{\Sigma^{-\frac{1}{2}}\widehat{\Sigma}^{-\frac{1}{2}}\widehat{\Sigma}\widehat{\Sigma}^{-\frac{1}{2}}\Sigma^{-\frac{1}{2}}} \\
&\overset{(c)}{=} \left\|\widehat{\Sigma}^{\frac{1}{2}}\Sigma^{\frac{1}{2}}\left(\Sigma^{-1}-\widehat{\Sigma}^{-1}\right)\Sigma^{\frac{1}{2}}\widehat{\Sigma}^{\frac{1}{2}}\right\|_{\widehat{\Sigma}} \\
&= \left\|\widehat{\Sigma}-\widehat{\Sigma}^{\frac{1}{2}}\Sigma^{\frac{1}{2}}\widehat{\Sigma}^{-1}\Sigma^{\frac{1}{2}}\widehat{\Sigma}^{\frac{1}{2}}\right\|_{\widehat{\Sigma}} \\
&\overset{(d)}{=} \left\|\widehat{\Sigma}-\Sigma\right\|_{\widehat{\Sigma}} \\
&\overset{(e)}{\leq} 2\left\|\widehat{\Sigma}-\Sigma\right\|_{\Sigma}, \\
&= 2\left\|T_M\left(X\right)-\Sigma\right\|_{\Sigma}, \qquad (12)
\end{aligned}$$

where $(a)$ uses Fact E.6, $(b)$ uses the fact that $\Sigma^{-1}\preceq\mathbb{I}$ (which holds by the first part of Lemma C.4), $(c)$ uses Lemma E.7 for $A=\Sigma^{-\frac{1}{2}}\widehat{\Sigma}^{-\frac{1}{2}}$, $(d)$ is by Lemma E.8, and $(e)$ is by Lemma E.9.

Substituting based on (12) into (11), we get:

$$\mathop{\mathbb{E}}_{X,T_M}\left[\left\|T_M\left(X\right)-\left(\eta_0-m\right)\right\|_2^2\right] \leq 8\mathop{\mathbb{E}}_{X,M}\left[\left\|M\left(X\right)-\Sigma\right\|_{\Sigma}^2\right],$$

which yields the desired result. ■

An immediate consequence of the above is the following:

**Corollary C.5.** *Let $\Sigma\in\mathbb{R}^{d\times d}$ be a covariance generated by Algorithm 1, and let $X\sim\mathcal{N}\left(0,\Sigma\right)^{\bigotimes n}$. If any $(\varepsilon,\delta)$-DP mechanism $T\colon\mathbb{R}^{n\times d}\to\bigotimes_{j\in[d^2]}I_j$ with $\mathop{\mathbb{E}}_{X,T}\left[\left\|T\left(X\right)-\left(\eta_0-m\right)\right\|_2^2\right]\leq 8\alpha^2\leq\frac{1}{4}$ requires at least $n\geq n_{\eta_0}$ samples, the same sample complexity lower bound holds for any $(\varepsilon,\delta)$-DP mechanism $M\colon\mathbb{R}^{n\times d}\to\mathbb{R}^{d\times d}$ that satisfies $\mathop{\mathbb{E}}_{X,M}\left[\left\|M\left(X\right)-\Sigma\right\|_{\Sigma}^2\right]\leq\alpha^2$.*

Having established the reduction from estimating $\eta_0$ with respect to the $\ell_2$-norm to estimating $\Sigma$ with respect to the Mahalanobis norm, it remains now to apply Theorem 2.4 to the former problem. For that reason, we work towards bounding the various quantities involved in it.

First, in accordance with Theorem 2.4, we assume the existence of an $(\varepsilon, \delta)$-DP mechanism $M \colon \mathbb{R}^{n \times d} \to \bigotimes_{j \in [d^2]} I_j$ such that:

$$\mathop{\mathbb{E}}_{X,M} \left[ \left\| M\left(X\right) - \left(\eta_0 - m\right) \right\|_2^2 \right] \leq 8\alpha^2 \leq \frac{1}{96} \leq \frac{\left\| R \right\|_2^2}{24}, \tag{13}$$

where we appealed to the third part of Lemma C.2.

The next step to establishing our lower bound is reasoning about the quantity $\mathop{\mathbb{E}}_{X_{\sim i}, M} \left[ s^\top \Sigma_{T_0} s \right]$. To do that, we identify an expression for $\Sigma_{T_0}$ for $\mathcal{N}\left(0, \Sigma\right)$ and use a result from [DKK+16]. By Lemma C.1, we have $T_0 \equiv T \implies \Sigma_{T_0} = \Sigma_T$. This yields:

$$\begin{aligned}
\Sigma_T &= \mathop{\mathbb{E}}_{Y \sim \mathcal{N}(0,\Sigma)} \left[ T\left(Y\right) T\left(Y\right)^\top \right] - \mu_T \mu_T^\top \\
&= \frac{1}{4} \mathop{\mathbb{E}}_{Y \sim \mathcal{N}(0,\Sigma)} \left[ \left(YY^\top\right)^\flat \left( \left(YY^\top\right)^\flat \right)^\top \right] - \frac{1}{4} \Sigma^\flat \left( \Sigma^\flat \right)^\top \\
&= \frac{1}{4} \mathop{\mathbb{E}}_{Y \sim \mathcal{N}(0,\Sigma)} \left[ \left(Y \otimes Y\right) \left(Y \otimes Y\right)^\top \right] - \frac{1}{4} \Sigma^\flat \left( \Sigma^\flat \right)^\top
\end{aligned} \tag{14}$$

Observe now that, based on the definition we gave for $s$ in Section 2 and the characteristics of the particular exponential family that we are considering here, the matrix $s^\# \in \mathbb{R}^{d \times d}$ is symmetric. We use this observation to connect our problem with the following result:

**Proposition C.6.** *[Theorem 4.2 from [DKK+16]] Let $\mathcal{S}_{\mathrm{sym}} = \left\{ M^\flat \in \mathbb{R}^{d^2} \colon M = M^\top \right\}$ and $X \sim \mathcal{N}\left(0, \Sigma\right)$. Let $M$ be the $d^2 \times d^2$ matrix given by $M = \mathbb{E}\left[ \left(X \otimes X\right) \left(X \otimes X\right)^\top \right]$. Then, as an operator on $\mathcal{S}_{\mathrm{sym}}$, we have:*

$$M = 2\Sigma^{\otimes 2} + \left(\Sigma^\flat\right) \left(\Sigma^\flat\right)^\top.$$

Using the above, we prove the following lemma:

**Lemma C.7.** *Let $\Sigma \in \mathbb{R}^{d \times d}$ be a covariance matrix generated by Algorithm 1, and let $X \sim \mathcal{N}\left(0, \Sigma\right)^{\otimes n}$ be a dataset. Also, let $M \colon \mathbb{R}^{n \times d} \to \bigotimes_{j \in [d^2]} I_j$ be an $(\varepsilon, \delta)$-DP mechanism such that $\mathop{\mathbb{E}}_{X,M} \left[ \left\| M\left(X\right) - \left(\eta_0 - m\right) \right\|_2^2 \right] \leq 8\alpha^2$. Then, it holds that:*

$$\mathop{\mathbb{E}}_{X_{\sim i}, M} \left[ s^\top \Sigma_T s \right] \leq \frac{\alpha^2}{d^4}.$$

*Proof.* By Proposition C.6 and (14), we get that:

$$
\begin{aligned}
\mathop{\mathbb{E}}_{X_{\sim i},M}\left[s^\top \Sigma_T s\right] &= \frac{1}{4}\mathop{\mathbb{E}}_{X_{\sim i},M}\left[s^\top \mathop{\mathbb{E}}_{Y\sim\mathcal{N}(0,\Sigma)}\left[(Y\otimes Y)(Y\otimes Y)^\top\right] s\right] - \frac{1}{4}\mathop{\mathbb{E}}_{X_{\sim i},M}\left[s^\top \Sigma^\flat \left(\Sigma^\flat\right)^\top s\right] \\
&= \frac{1}{4}\mathop{\mathbb{E}}_{X_{\sim i},M}\left[s^\top \left(2\Sigma^{\otimes 2} + \left(\Sigma^\flat\right)\left(\Sigma^\flat\right)^\top\right) s\right] - \frac{1}{4}\mathop{\mathbb{E}}_{X_{\sim i},M}\left[s^\top \Sigma^\flat \left(\Sigma^\flat\right)^\top s\right] \\
&= \frac{1}{2}\mathop{\mathbb{E}}_{X_{\sim i},M}\left[s^\top \Sigma^{\otimes 2} s\right] \\
&= \frac{1}{2}\mathop{\mathbb{E}}_{X_{\sim i},M}\left[\|s\|_2^2 \frac{s^\top}{\|s\|_2}\Sigma^{\otimes 2}\frac{s}{\|s\|_2}\right] \\
&\overset{(a)}{\leq} \frac{1}{2}\left\|\Sigma^{\otimes 2}\right\|_2 \mathop{\mathbb{E}}_{X_{\sim i},M}\left[\|s\|_2^2\right] \\
&\overset{(b)}{=} \frac{1}{2}\|\Sigma\|_2^2 \mathop{\mathbb{E}}_{X_{\sim i},M}\left[\|s\|_2^2\right] \\
&\overset{(c)}{\leq} 2\mathop{\mathbb{E}}_{X_{\sim i},M}\left[\sum_{j\in[d^2]}\left[\frac{R_j^2}{4} - (\eta_{0,j}-m_j)^2\right]^2 \left[M_j(X_{\sim i}) - (\eta_{0,j}-m_j)\right]^2\right] \\
&\leq \frac{\|R\|_\infty^4}{16} 2\mathop{\mathbb{E}}_{X_{\sim i},M}\left[\|M(X_{\sim i}) - (\eta_0 - m)\|_2^2\right] \\
&\overset{(d)}{\leq} \frac{\|R\|_\infty^4}{8}\cdot 8\alpha^2 \\
&= \|R\|_\infty^4 \alpha^2 \\
&= \frac{\alpha^2}{d^4},
\end{aligned}
$$

where $(a)$ is by the definition of the spectral norm, $(b)$ is by Fact E.5, $(c)$ is by the first part of Lemma C.2, $(d)$ is because $\mathop{\mathbb{E}}_{X,M}\left[\|M(X) - (\eta_0 - m)\|_2^2\right] \leq 8\alpha^2$, and the last equality is by the third part of Lemma C.2. ■

Now, it remains to reason about the terms involving the tail probabilities of $\|T(X_i) - \mu_T\|_2$. Our approach closely follows the strategy sketched in Remark B.2. The following lemma is devoted to implementing this.

**Lemma C.8.** *Let $\Sigma \in \mathbb{R}^{d\times d}$ be a covariance matrix generated by Algorithm 1, and $X_i \sim \mathcal{N}(0,\Sigma)$. For $T \geq \frac{1}{4d^2}$, we have:*

$$
\int_T^\infty \mathop{\mathbb{P}}_{X_i}\left[\|T(X_i) - \mu_T\|_2 > 4d^2 t\right] dt \leq \frac{1}{2cd^2} e^{-2d^2(2cT - \ln(3))},
$$

*where $c \approx 0.036425$.*

*Proof.* Setting $Y := \Sigma^{-\frac{1}{2}} X_i \sim \mathcal{N}(0,\mathbb{I})$, we observe that:

$$
\begin{aligned}
\|T(X_i) - \mu_T\|_2 = \frac{1}{2}\left\|\left(X_i X_i^\top\right)^\flat - \Sigma^\flat\right\|_2 = \frac{1}{2}\left\|X_i X_i^\top - \Sigma\right\|_F &= \frac{1}{2}\left\|\Sigma^{\frac{1}{2}}\left[\left(\Sigma^{-\frac{1}{2}}X_i\right)\left(\Sigma^{-\frac{1}{2}}X_i\right)^\top - \mathbb{I}\right]\Sigma^{\frac{1}{2}}\right\|_F \\
&\overset{(a)}{\leq} \frac{\left\|\Sigma^{\frac{1}{2}}\right\|_2^2}{2}\left\|YY^\top - \mathbb{I}\right\|_F \\
&\overset{(b)}{\leq} \left\|YY^\top - \mathbb{E}\left[YY^\top\right]\right\|_F,
\end{aligned}
$$

where $(a)$ uses Fact E.6, and $(b)$ uses the assumption that $\Sigma \preceq 2\mathbb{I}$ and the remark that, for symmetric PSD matrices, $\left\|\Sigma^{\frac{1}{2}}\right\|_2 = \sqrt{\|\Sigma\|_2}$.

Thus, we have:

$$\int_T^\infty \mathop{\mathbb{P}}_{X_i}\left[\|T(X_i)-\mu_T\|_2 > 4d^2t\right]\,dt \le \int_T^\infty \mathbb{P}\left[\left\|YY^\top - \mathbb{E}\left[YY^\top\right]\right\|_F > 4d^2t\right]\,dt.$$

Working as we described in Remark B.2, we get:

$$\begin{aligned}
\int_T^\infty \mathop{\mathbb{P}}_{X_i}\left[\|T(X_i)-\mu_T\|_2 > 4d^2t\right]\,dt &\le \int_T^\infty \mathbb{P}\left[\left\|\left(YY^\top - \mathbb{E}\left[YY^\top\right]\right)^\flat\right\|_2 > 4d^2t\right]\,dt \\
&\le 9^{d^2}\int_T^\infty \mathbb{P}\left[\left|\left\langle M^\flat, \left(YY^\top - \mathbb{E}\left[YY^\top\right]\right)^\flat\right\rangle\right| > 3d^2t\right]\,dt \\
&\le 9^{d^2}\int_T^\infty \mathbb{P}\left[\left|\left\langle M, YY^\top - \mathbb{E}\left[YY^\top\right]\right\rangle\right| > 3d^2t\right]\,dt \\
&= 9^{d^2}\int_T^\infty \mathbb{P}\left[\left|Y^\top MY - \mathbb{E}\left[Y^\top MY\right]\right| > 3d^2t\right]\,dt \\
&= 9^{d^2}\int_T^\infty \mathbb{P}\left[\left|Y^\top \frac{M+M^\top}{2}Y - \mathbb{E}\left[Y^\top \frac{M+M^\top}{2}Y\right]\right| > 3d^2t\right]\,dt
\end{aligned} \tag{15}$$

where $M$ is a matrix with $\|M\|_F = 1$.

We have that $Y \sim \mathcal{N}(0,\mathbb{I})$, and that $\frac{M+M^\top}{2}$ is symmetric, so we can apply Fact F.7 to (15). Since $\left\|\frac{M+M^\top}{2}\right\|_F \le \frac{\|M\|_F + \left\|M^\top\right\|_F}{2} \le 1$ and $\left\|\frac{M+M^\top}{2}\right\|_2 \le \left\|\frac{M+M^\top}{2}\right\|_F \le 1$, we get:

$$\int_T^\infty \mathop{\mathbb{P}}_{X_i}\left[\|T(X_i)-\mu_T\|_2 > 4d^2t\right]\,dt \le 2\cdot 9^{d^2}\int_T^\infty \exp\left(-c\min\left\{16d^4t^2, 4d^2t\right\}\right)\,dt.$$

By our assumption that on $T$, we get $\min\left\{16d^4t^2, 4d^2t\right\} = 4d^2t, \forall t \ge T$. This leads to:

$$\begin{aligned}
\int_T^\infty \mathop{\mathbb{P}}_{X_i}\left[\|T(X_i)-\mu_T\|_2 > 4d^2t\right]\,dt \le 2\cdot 9^{d^2}\int_T^\infty \exp\left(-4cd^2t\right)\,dt &= \frac{9^{d^2}}{2cd^2}e^{-4cd^2T} \\
&= \frac{1}{2cd^2}e^{-2d^2(2cT-\ln(3))}.
\end{aligned}$$

■

We are now ready to prove the main result of this section.

*Proof of Theorem 3.1.* We assume that the process generating $\Sigma$ is that of Algorithm 1. By the first part of Lemma C.2, the condition $\mathbb{I} \preceq \Sigma \preceq 2\mathbb{I}$ is satisfied. We assume that there exists an $(\varepsilon,\delta)$-DP mechanism $M\colon \mathbb{R}^{n\times d} \to \bigotimes_{j\in[d^2]} I_j$ that satisfies $\mathop{\mathbb{E}}_{X,M}\left[\|M(X)-(\eta_0-m)\|_2^2\right] \le 8\alpha^2 \le \frac{1}{96}$. Combining the results of Lemmas C.7, and C.8, with Theorem 2.4, we get:

$$n\left(2\delta T + 2\frac{\alpha\varepsilon}{d^2} + \frac{1}{cd^2}e^{-2d^2(2cT-\ln(3))}\right) \ge \frac{1}{48}\left(1-\frac{1}{2d}\right) \ge \frac{1}{96}. \tag{16}$$

It remains to set the values of $T$ and $\delta$ appropriately so that:

$$\delta T \ge \frac{1}{cd^2}e^{-2d^2(2cT-\ln(3))} \text{ and } 3n\delta T \le \frac{1}{192}.$$

The first of these two conditions yields:

$$\delta T \geq \frac{1}{cd^2} e^{-2d^2(2cT-\ln(3))} \iff T e^{2d^2(2cT-\ln(3))} \geq \frac{1}{cd^2} \cdot \frac{1}{\delta}.$$

We set $T = \frac{1}{2c}\left(\ln(3) + \frac{1}{2d^2}\ln\left(\frac{1}{\delta}\right)\right)$. This satisfies the constraint of Lemma C.8, which says that it must hold that $T \geq \frac{1}{4d^2}$. Then, the above becomes:

$$\frac{1}{2c}\left(\ln(3) + \frac{1}{2d^2}\ln\left(\frac{1}{\delta}\right)\right)\frac{1}{\delta} \geq \frac{1}{cd^2} \cdot \frac{1}{\delta} \iff \ln\left(\frac{1}{\delta}\right) \geq 2\left(2 - \ln(3)\, d^2\right).$$

The RHS of the last inequality is $< 0$ for $d \geq 2$. For $d = 1$, we get the constraint $\delta \leq e^{2\ln(3)-4} \approx 0.16$. Respecting this constraint, we proceed to identify a range of values for $\delta$ such that $n\delta T \leq \frac{1}{576}$. This is equivalent to:

$$\frac{\delta}{2c}\left(\ln(3) + \frac{1}{2d^2}\ln\left(\frac{1}{\delta}\right)\right) \leq \frac{1}{576n} \iff \delta\left(1 + \frac{\ln\left(\frac{1}{\delta}\right)}{2\ln(3)\, d^2}\right) \leq \frac{c}{288\ln(3)\, n}. \tag{17}$$

The above inequality yields a constraint on $\delta$. To ensure that $\ln\left(\frac{1}{\delta}\right) \geq 0$, we need $\frac{c}{288\ln(3)n\delta} - 1 \geq 0 \iff \delta \leq \frac{c}{288\ln(3)n}$. We know from Lemma C.8 value of $c$ is $c \approx 0.036425$. Thus, this last constraint is stricter than $\delta \leq 0.16$.

For (17) to hold, we will show that it suffices to have:

$$\delta \leq \min\left\{\frac{c}{576\ln(3)\, n}, \frac{cd^2}{576n\ln\left(\frac{288n}{cd^2}\right)}\right\},$$

which trivially satisfies the previous constraint.

For convenience, we set $\phi := \frac{cd^2}{288n}$, so the proposed range for $\delta$ becomes $\delta \leq \min\left\{\frac{\phi}{2\ln(3)d^2}, \frac{\phi}{2\ln\left(\frac{1}{\phi}\right)}\right\}$. Additionally, (17) can be equivalently written as:

$$\delta\left(1 + \frac{\ln\left(\frac{1}{\delta}\right)}{2\ln(3)\, d^2}\right) \leq \frac{\phi}{\ln(3)\, d^2}. \tag{18}$$

To verify our claim, we consider the following two cases:

1. $\min\left\{\frac{\phi}{2\ln(3)d^2}, \frac{\phi}{2\ln\left(\frac{1}{\phi}\right)}\right\} \leq e^{-2\ln(3)d^2}$. We start by arguing that:

$$\min\left\{\frac{\phi}{2\ln(3)\, d^2}, \frac{\phi}{2\ln\left(\frac{1}{\phi}\right)}\right\} \leq e^{-2\ln(3)d^2} \iff \frac{\phi}{2\ln\left(\frac{1}{\phi}\right)} \leq e^{-2\ln(3)d^2}. \tag{19}$$

The $\Longleftarrow$ direction is trivial, so we focus on the $\Longrightarrow$ direction.

We assume that $\min\left\{\frac{\phi}{2\ln(3)d^2}, \frac{\phi}{2\ln\left(\frac{1}{\phi}\right)}\right\} \leq e^{-2\ln(3)d^2}$ but $\frac{\phi}{2\ln\left(\frac{1}{\phi}\right)} > e^{-2\ln(3)d^2}$. Then, it must be the case that $\frac{\phi}{2\ln(3)d^2} \leq e^{-2\ln(3)d^2}$. Setting $y := \frac{1}{\phi}$, the above system of inequalities can be written in the form:

$$\begin{cases} \frac{e^{2\ln(3)d^2}}{2} > y\ln(y) \\ y \geq \frac{e^{2\ln(3)d^2}}{2\ln(3)d^2} \end{cases}.$$

The function $y\ln(y)$ is increasing for $y \geq e^{-1}$. We have $\frac{e^{2\ln(3)d^2}}{2\ln(3)d^2} \geq e^{-1}, \forall d \geq 1$. Thus, for $y \geq \frac{e^{2\ln(3)d^2}}{2\ln(3)d^2}$, we get:

$$\frac{e^{2\ln(3)d^2}}{2} > y\ln(y) \geq \frac{e^{2\ln(3)d^2}}{2\ln(3)\, d^2}\ln\left(\frac{e^{2\ln(3)d^2}}{2\ln(3)\, d^2}\right) \implies \ln(3)\, d^2 < \ln\left(2\ln(3)\, d^2\right).$$

However, the above inequality cannot be satisfied, which leads to a contradiction.

Consequently, we have shown (19). Since we have assumed that $\min\left\{\frac{\phi}{2\ln(3)d^2}, \frac{\phi}{2\ln\left(\frac{1}{\phi}\right)}\right\} \leq e^{-2\ln(3)d^2}$, it must be the case that $\frac{\phi}{2\ln\left(\frac{1}{\phi}\right)} \leq e^{-2\ln(3)d^2}$.

We now turn our attention again to (18). We remark that, for $\delta \leq e^{-2\ln(3)d^2}$, we get $\frac{\ln\left(\frac{1}{\delta}\right)}{2\ln(3)d^2} \geq 1$. As a result, in order to satisfy (18), it suffices to have:

$$2\delta\frac{\ln\left(\frac{1}{\delta}\right)}{2\ln(3)\,d^2} \leq \frac{\phi}{\ln(3)\,d^2} \iff \delta\ln\left(\frac{1}{\delta}\right) \leq \phi. \tag{20}$$

Observe that the function $\delta\ln\left(\frac{1}{\delta}\right)$ is increasing for $\delta \leq e^{-2\ln(3)d^2} < e^{-1}$. Combined with our previous remarks, this implies that, in order to verify (20) for the values of $\delta$ we picked, it suffices to show that it holds for $\delta = \frac{\phi}{2\ln\left(\frac{1}{\phi}\right)}$. So, we have to verify the inequality:

$$\left(\frac{\phi}{2\ln\left(\frac{1}{\phi}\right)}\right)\ln\left(\frac{2\ln\left(\frac{1}{\phi}\right)}{\phi}\right) \leq \phi \iff \frac{\ln\left(2\ln\left(\frac{1}{\phi}\right)\right)}{\ln\left(\frac{1}{\phi}\right)} \leq 1.$$

Setting $z := \ln\left(\frac{1}{\phi}\right)$, this becomes equivalent to $\frac{\ln(2z)}{z} \leq 1$, which holds for all $z > 0 \iff \ln\left(\frac{1}{\phi}\right) > 0 \iff \phi < 1$. This last condition is satisfied by all $\phi$ such that $\frac{\phi}{2\ln\left(\frac{1}{\phi}\right)} \leq e^{-2\ln(3)d^2}$, so the desired result has been established.

2. $\min\left\{\frac{\phi}{2\ln(3)d^2}, \frac{\phi}{2\ln\left(\frac{1}{\phi}\right)}\right\} > e^{-2\ln(3)d^2}$. For this case, we consider two sub-cases, depending on how $\delta$ compares with $e^{-2\ln(3)d^2}$. We have:

   - $\delta \leq e^{-2\ln(3)d^2}$. As before, we argue that it suffices to have $\delta\ln\left(\frac{1}{\delta}\right) \leq \phi$. Since $\delta\ln\left(\frac{1}{\delta}\right)$ is increasing for $\delta \leq e^{-2\ln(3)d^2} < e^{-1}$, all we have to do is show that $e^{-2\ln(3)d^2}2\ln(3)\,d^2 \leq \phi$. This is an immediate consequence of our assumption that $\min\left\{\frac{\phi}{2\ln(3)d^2}, \frac{\phi}{2\ln\left(\frac{1}{\phi}\right)}\right\} > e^{-2\ln(3)d^2}$.
   - $\delta > e^{-2\ln(3)d^2}$. This implies $\frac{\ln\left(\frac{1}{\delta}\right)}{2\ln(3)d^2} < 1$. Thus, for (18) to hold, it suffices that:

     $$2\delta \leq \frac{\phi}{\ln(3)\,d^2} \iff \delta \leq \frac{\phi}{2\ln(3)\,d^2}.$$

     This is satisfied, because our proposed range for $\delta$ is $\delta \leq \min\left\{\frac{\phi}{2\ln(3)d^2}, \frac{\phi}{2\ln\left(\frac{1}{\phi}\right)}\right\}$.

We have established that:

$$T = \frac{1}{2c}\left(\ln(3) + \frac{1}{2d^2}\ln\left(\frac{1}{\delta}\right)\right) \text{ and } \delta \leq \mathcal{O}\left(\min\left\{\frac{1}{n}, \frac{d^2}{n\log\left(\frac{n}{d^2}\right)}\right\}\right),$$

imply:

$$\delta T \geq \frac{1}{cd^2}e^{-2d^2(2cT-\ln(3))} \text{ and } 3n\delta T \leq \frac{1}{192},$$

while respecting all the constraints.

We substitute this to (16) and get $n \geq \frac{d^2}{384\alpha\varepsilon}$. Appealing to Corollary C.5 completes the proof. ■

### C.2 Omitted Proofs from Section 3.2

*Proof of Theorem 3.2.* First, assume that $\Sigma$ is generated by Algorithm 1 (as was assumed in all of Section 3.1). Additionally, assume that there exists an $(\varepsilon, \delta)$-DP mechanism $M : \mathbb{R}^{n\times d} \to \mathbb{R}^{d\times d}$ with $\varepsilon \in [0,1]$ and $\delta \leq \mathcal{O}\left(\min\left\{\frac{1}{n}, \frac{d^2}{n\log\left(\frac{n}{d^2}\right)}\right\}\right)$ that satisfies $\mathop{\mathbb{E}}_{X,M}\left[\left\|\Sigma^{-\frac{1}{2}}\left(M\left(X\right)-\Sigma\right)\Sigma^{-\frac{1}{2}}\right\|_2^2\right] \leq \alpha^2$ and uses $n = o\left(\frac{d^{1.5}}{\alpha\varepsilon}\right)$ samples. Using the standard matrix-norm inequality $\|A\|_F \leq \sqrt{d}\,\|A\|_2, \forall A \in \mathbb{R}^{k\times\ell}$, we get:

$$\begin{aligned}
\mathop{\mathbb{E}}_{X,M}\left[\|M\left(X\right)-\Sigma\|_\Sigma^2\right] &= \mathop{\mathbb{E}}_{X,M}\left[\left\|\Sigma^{-\frac{1}{2}}\left(M\left(X\right)-\Sigma\right)\Sigma^{-\frac{1}{2}}\right\|_F^2\right] \\
&\leq d\mathop{\mathbb{E}}_{X,M}\left[\left\|\Sigma^{-\frac{1}{2}}\left(M\left(X\right)-\Sigma\right)\Sigma^{-\frac{1}{2}}\right\|_2^2\right] \\
&\leq d\alpha^2.
\end{aligned}$$

By Theorem 3.1 for $\alpha \to \sqrt{d}\alpha$, we get that any mechanism with the characteristics of $M$ that estimates a Gaussian covariance matrix in Mahalanobis error $\sqrt{d}\alpha$ requires $n \geq \Omega\left(\frac{d^{1.5}}{\alpha\varepsilon}\right)$ samples. This leads to a contradiction, implying the desired sample complexity lower bound. ■

## D Omitted Proof from Section 4

*Proof of Theorem 4.1.* We define the following family of distributions, such that the second moment of each distribution is upper-bounded by 1. The loss function between the true distribution and the estimate is the squared $\ell_2$-distance between their means. For each $v \in \mathcal{E}_d = \{-1,1\}^d$, we define the distribution $\mathcal{D}_v$ over $\mathbb{R}^d$ as follows. Let $t > 0$ and $0 < p < 1$. For any $X \sim \mathcal{D}_v$, we have:

$$X_i = \begin{cases} v_i t, & \text{with probability } p \\ 0, & \text{with probability } 1-p \end{cases}, \forall i \in [d].$$

Then $\mathop{\mathbb{E}}_{X\sim\mathcal{D}_v}[X] = ptv$. We show that the second moment of this distribution is upper-bounded by $pt^2$, that is, $\mathop{\mathbb{E}}_{X\sim\mathcal{D}_v}\left[\langle X - ptv, u\rangle^2\right] \leq pt^2, \forall u \in \mathbb{S}^{d-1}$. Note that, for any $i \in [d]$, we have $\mathbb{E}\left[\left(X_i - ptv_i\right)^2\right] = p\left(1-p\right)t^2$. We omit the subscript of the expectation when the context is clear. We have:

$$\begin{aligned}
\mathbb{E}\left[\langle X - ptv, u\rangle^2\right] &= \mathbb{E}\left[\left(\sum_{i=1}^d \left(X_i - ptv_i\right)u_i\right)^2\right] \\
&= \sum_{i=1}^d \mathbb{E}\left[\left(X_i - ptv_i\right)^2\right]u_i^2 + \sum_{i\neq j}\mathbb{E}\left[X_i - ptv_i\right]\mathbb{E}\left[X_j - ptv_j\right]u_iu_j \\
&= \sum_{i=1}^d p\left(1-p\right)t^2u_i^2 \\
&= p\left(1-p\right)t^2 \\
&\leq pt^2.
\end{aligned}$$

We want this to be upper-bounded by 1:

$$pt^2 = 1. \tag{21}$$

Next, for any $u, v \in \mathcal{E}_d$, we bound the loss function between $\mathcal{D}_{v_1}$ and $\mathcal{D}_{v_2}$. As mentioned before, we define $\ell\left(\theta\left(\mathcal{D}_u\right), \theta\left(\mathcal{D}_v\right)\right) = \left\|\theta\left(\mathcal{D}_u\right) - \theta\left(\mathcal{D}_v\right)\right\|_2^2$. We have the following:

$$\begin{aligned}
\left\|\theta\left(\mathcal{D}_u\right) - \theta\left(\mathcal{D}_v\right)\right\|_2^2 &= \sum_{i=1}^{d} p^2 t^2 \left(v_i - u_i\right)^2 \\
&= p^2 t^2 \sum_{i:\ u_i \neq v_i} \left(v_i - u_i\right)^2 \\
&= 4p^2 t^2 \sum_{i=1}^{d} \mathbb{1}\left\{v_i \neq u_i\right\}.
\end{aligned}$$

Using the notation in Lemma A.4, we have $\tau = 2p^2t^2$. Using the same lemma, we have:

$$\begin{aligned}
R\left(\mathcal{P}, \ell, \varepsilon, \delta\right) &\geq \frac{d\tau}{2} \cdot \left(0.9e^{-10\varepsilon D} - 10\delta D\right) \\
&\geq \frac{d\tau}{2} \cdot \left[0.9\left(1 - 10\varepsilon D\right) - 10\delta D\right] \\
&= \frac{d\tau}{2} \cdot \left[0.9 - D\left(9\varepsilon + 10\delta\right)\right] \\
&\geq \frac{d\tau}{2} \cdot \left[0.9 - 10D\left(\varepsilon + \delta\right)\right].
\end{aligned}$$

Setting $R\left(\mathcal{P}, \ell, \varepsilon, \delta\right) \leq \alpha^2$ and rearranging the above, we get:

$$D \geq \frac{1}{10\left(\varepsilon + \delta\right)} \cdot \left(0.9 - \frac{2\alpha^2}{d\tau}\right). \tag{22}$$

Now, from Equation (21), $\tau = 2p^2t^2 = 2p$. Using this and Inequality (22), and setting $p = \frac{2\alpha^2}{d}$, we have the following.

$$D \geq \frac{0.04}{\varepsilon + \delta}. \tag{23}$$

For $i \in [d]$, we define the mixture distributions $p_{+i}$ and $p_{-i}$ as in Lemma A.4. We define a coupling $(X, Y)$ between $p_{+i}$ and $p_{-i}$ as follows. For every row $X_j$ in $X$, if $X_j^i \neq 0$, then $Y_j^i = -X_j^i$, otherwise $Y_j^i = X_j^i$. For all other coordinates $k \neq i$, $Y_j^k = X_j^k$. Notice that, for all $k \neq i$, the distributions for coordinate $k$ for both $p_{+i}$ and $p_{-i}$ are identical because all coordinates are independent. Therefore, by setting $Y_j^k = X_j^k$, these coordinates of $X_j$ and $Y_j$ have identical distributions for all $j \in [n]$. The coordinate $i$ of $X_j$ is $t$ with probability $p$, and 0 with probability $1 - p$. Hence, the coordinate $i$ of $Y_j$ is $-t$ with probability $p$, and 0 with probability $1 - p$. This is true for all $j \in [n]$. Since all $n$ rows in both $X$ (hence, in $Y$) are chosen independently at random, the distribution of $X$ is $p_{+i}$ and the distribution of $Y$ is $p_{-i}$. Therefore, $(X, Y)$ is a valid coupling of $p_{+i}$ and $p_{-i}$. Next, we have to determine the value of $D = \mathbb{E}\left[d_{\mathrm{Ham}}\left(X, Y\right)\right]$. Note that for each $j \in [n]$, $X_j$ can only differ from $Y_j$ when $X_j^i \neq 0$. This can only happen with probability $p$. Thus, by linearity of expectation, $D = np$.

Finally, using the fact that $p = \frac{2\alpha^2}{d}$ and $t = \frac{\sqrt{d}}{\sqrt{2}\alpha}$, and substituting this in Inequality (23), we have:

$$np \geq \frac{0.04}{\varepsilon + \delta} \implies n \geq \frac{d}{50\left(\varepsilon + \delta\right)\alpha^2}.$$

When $\varepsilon \geq \delta$, this gives us the desired lower bound. ■

# E Facts from Linear Algebra and Geometry

In this appendix we collect some results from linear algebra and geometry which are referenced throughout this work.

**Definition E.1** ($\varepsilon$-net)**.** Let $(T, d)$ be a metric space. Consider a subset $K \subset T$ and let $\varepsilon > 0$. A subset $\mathcal{N}_\varepsilon \subseteq K$ is called an *$\varepsilon$-net* of $K$ if the following holds:

$$\forall x \in K, \exists y \in \mathcal{N}_\varepsilon \colon d(x, x_0) \leq \varepsilon.$$

**Definition E.2** (Covering Number)**.** The smallest possible cardinality of an $\varepsilon$-net of $K$ is called the *covering number* of $K$ and denoted by $\mathcal{N}(K, d, \varepsilon)$. Equivalently, $\mathcal{N}(K, d, \varepsilon)$ is the smallest number of closed balls with centers in $K$ and radii $\varepsilon$ whose union covers $K$.

**Fact E.3** (Covering numbers of $\mathbb{S}^{d-1}$)**.** *The covering numbers of the unit Euclidean sphere $\mathbb{S}^{d-1}$ satisfy the following for any $\varepsilon > 0$:*

$$\left(\frac{1}{\varepsilon}\right)^d \leq \mathcal{N}\left(\mathbb{S}^{d-1}, \varepsilon\right) \leq \left(1 + \frac{1}{\varepsilon}\right)^d.$$

**Fact E.4** (Exercise 4.4.2 from [Ver18])**.** *Let $x \in \mathbb{R}^d$ and $\mathcal{N}_\varepsilon$ be an $\varepsilon$-net on $\mathbb{S}^{d-1}$. It holds that:*

$$\sup_{y \in \mathcal{N}_\varepsilon} |\langle x, y \rangle| \leq \|x\|_2 \leq \frac{1}{1-\varepsilon} \cdot \sup_{y \in \mathcal{N}_\varepsilon} |\langle x, y \rangle|.$$

**Fact E.5.** *Let $A \in \mathbb{R}^{n \times m}$ and $B \in \mathbb{R}^{k \times \ell}$. Then, it holds that $\|A \otimes B\|_2 = \|A\|_2 \|B\|_2$.*

**Fact E.6.** *Let $M, N \in \mathbb{R}^{d \times d}$. Then, $\|MN\|_F \leq \|M\|_2 \|N\|_F$, and $\|MN\|_F \leq \|N\|_2 \|M\|_F$.*

For a proof of the first of the two inequalities, see Lemma 4.26 of [DKK+16]. The proof for the second one is analogous but we have to work with the rows of $M$ instead of the columns of $N$.

**Lemma E.7.** *Let $A, \Sigma \in \mathbb{R}^{d \times d}$ be non-singular matrices and let $X \in \mathbb{R}^{d \times d}$. Then, it holds that $\|X\|_{A\Sigma A^\top} = \left\|A^{-1} X \left(A^\top\right)^{-1}\right\|_\Sigma$.*

For a proof of this statement for symmetric $A$ (the proof does not change), we refer the reader to Section 2.2 of [Li19].

**Lemma E.8.** *Let $A, B \in \mathbb{R}^{d \times d}$ be symmetric positive definite matrices. Then, it holds that:*

$$\left\|A - A^{\frac{1}{2}} B^{\frac{1}{2}} A^{-1} B^{\frac{1}{2}} A^{\frac{1}{2}}\right\|_A = \|A - B\|_A.$$

*Proof.* We have that:

$$\begin{aligned}
\left\|A - A^{\frac{1}{2}} B^{\frac{1}{2}} A^{-1} B^{\frac{1}{2}} A^{\frac{1}{2}}\right\|_A^2 &= \left\|\mathbb{I} - B^{\frac{1}{2}} A^{-1} B^{\frac{1}{2}}\right\|_F^2 \\
&= \operatorname{tr}\left(\left(\mathbb{I} - B^{\frac{1}{2}} A^{-1} B^{\frac{1}{2}}\right)^2\right) \\
&= \operatorname{tr}\left(\mathbb{I} - 2B^{\frac{1}{2}} A^{-1} B^{\frac{1}{2}} + B^{\frac{1}{2}} A^{-1} B A^{-1} B^{\frac{1}{2}}\right) \\
&\overset{(a)}{=} \operatorname{tr}\left(\mathbb{I} - 2A^{-\frac{1}{2}} B A^{-\frac{1}{2}} + \left(A^{-\frac{1}{2}} B A^{-\frac{1}{2}}\right)^2\right) \\
&= \operatorname{tr}\left(\left(\mathbb{I} - A^{-\frac{1}{2}} B A^{-\frac{1}{2}}\right)^2\right) \\
&= \left\|\mathbb{I} - A^{-\frac{1}{2}} B A^{-\frac{1}{2}}\right\|_F^2 \\
&= \|A - B\|_A^2,
\end{aligned}$$

where in $(a)$ we use the cyclic property of the trace in individual terms. ■

**Lemma E.9.** *Let $\Sigma_1, \Sigma_0 \in \mathbb{R}^{d \times d}$ be a pair of symmetric positive definite matrices. If $\|\Sigma_1 - \Sigma_0\|_{\Sigma_0} < \frac{1}{2}$, it holds that $\|\Sigma_1 - \Sigma_0\|_{\Sigma_1} \leq 2 \|\Sigma_1 - \Sigma_0\|_{\Sigma_0}$.*

*Proof.* By the definition of the Mahalanobis norm, we have:

$$\begin{aligned}\|\Sigma_1-\Sigma_0\|_{\Sigma_0}^2 &= \left\|\mathbb{I}-\Sigma_0^{-\frac{1}{2}}\Sigma_1\Sigma_0^{-\frac{1}{2}}\right\|_F \\ &= \operatorname{tr}\left(\left(\mathbb{I}-\Sigma_0^{-\frac{1}{2}}\Sigma_1\Sigma_0^{-\frac{1}{2}}\right)^2\right) \\ &= \operatorname{tr}\left(\mathbb{I}-2\Sigma_0^{-\frac{1}{2}}\Sigma_1\Sigma_0^{-\frac{1}{2}}+\Sigma_0^{-\frac{1}{2}}\Sigma_1\Sigma_0^{-1}\Sigma_1\Sigma_0^{-\frac{1}{2}}\right) \\ &= d-2\operatorname{tr}\left(\Sigma_0^{-1}\Sigma_1\right)+\operatorname{tr}\left(\left(\Sigma_0^{-1}\Sigma_1\right)^2\right). \end{aligned} \tag{24}$$

Similarly, we have:

$$\|\Sigma_1-\Sigma_0\|_{\Sigma_1}^2 = d-2\operatorname{tr}\left(\Sigma_1^{-1}\Sigma_0\right)+\operatorname{tr}\left(\left(\Sigma_1^{-1}\Sigma_0\right)^2\right). \tag{25}$$

The eigenvalues of the matrix $\Sigma_0^{-1}\Sigma_1$ are positive. Indeed, the previous matrix can be written as $\Sigma_0^{-\frac{1}{2}}\left(\Sigma_0^{-\frac{1}{2}}\Sigma_1\Sigma_0^{-\frac{1}{2}}\right)\Sigma_0^{\frac{1}{2}}$, implying that it similar to the symmetric matrix $\Sigma_0^{-\frac{1}{2}}\Sigma_1\Sigma_0^{-\frac{1}{2}}$. This last matrix is positive definite, since $u^\top\Sigma_0^{-\frac{1}{2}}\Sigma_1\Sigma_0^{-\frac{1}{2}}u=\left(\Sigma_0^{-\frac{1}{2}}u\right)^\top\Sigma_1\left(\Sigma_0^{-\frac{1}{2}}u\right)$ and $\Sigma_1\succ 0$. Thus, if $\lambda_1,\ldots,\lambda_d>0$ are the eigenvalues of $\Sigma_0^{-1}\Sigma_1$, (24) and (25) can be written as:

$$\begin{aligned}\|\Sigma_1-\Sigma_0\|_{\Sigma_0}^2 &= d-2\cdot\sum_{i=1}^d\lambda_i+\sum_{i=1}^d\lambda_i^2=\sum_{i=1}^d\left(\lambda_i-1\right)^2, \\ \|\Sigma_1-\Sigma_0\|_{\Sigma_1}^2 &= d-2\cdot\sum_{i=1}^d\frac{1}{\lambda_i}+\sum_{i=1}^d\frac{1}{\lambda_i^2}=\sum_{i=1}^d\left(\frac{1}{\lambda_i}-1\right)^2.\end{aligned}$$

Therefore, proving the desired upper bound reduces to the following maximization problem:

$$\begin{aligned}\max_{\lambda_i}\quad & \sum_{i=1}^d\left(\frac{1}{\lambda_i}-1\right)^2 \\ \text{s.t.}\quad & \sum_{i=1}^d\left(\lambda_i-1\right)^2=\|\Sigma_1-\Sigma_0\|_{\Sigma_0}^2<\frac{1}{4},\end{aligned}$$

where we have omitted the constraint $\lambda_i>0$, since we necessarily have $\lambda_i\geq 1-\|\Sigma_1-\Sigma_0\|_{\Sigma_0}>\frac{1}{2},\forall i\in[d]$. The above problem is solvable exactly via the KKT conditions, yielding:

$$\|\Sigma_1-\Sigma_0\|_{\Sigma_1}\leq\frac{\|\Sigma_1-\Sigma_0\|_{\Sigma_0}\sqrt{d}}{\sqrt{d}-\|\Sigma_1-\Sigma_0\|_{\Sigma_0}}\leq 2\|\Sigma_1-\Sigma_0\|_{\Sigma_0},$$

for $\|\Sigma_1-\Sigma_0\|_{\Sigma_0}<\frac{1}{2}$ and $d\geq 1$. ■

**Theorem E.10** (Gershgorin Circle Theorem (see [Ger31]))**.** *Let $A\in\mathbb{R}^{d\times d}$. For any eigenvalue $\lambda$ of $A$, there exists an $i\in[d]$, such that:*

$$|\lambda-a_{ii}|\leq\sum_{j\neq i}|a_{ij}|.$$

# F Facts from Probability & Statistics

In this appendix we collect some results from probability and statistics which are referenced throughout this work.

**Fact F.1.** *The Bernoulli distribution $\mathrm{Be}(p)$ is an exponential family with support $S=\{0,1\}$ and $h(x)=1,\eta=\ln\left(\frac{1-p}{p}\right),T(x)=x,Z(\eta)=\ln(1+e^\eta)$.*

**Fact F.2.** *The $d$-dimensional Gaussian distribution with known covariance and unknown mean $\mathcal{N}(\mu, \mathbb{I})$ is an exponential family with support $S = \mathbb{R}^d$ and $h(x) = \frac{1}{(2\pi)^{\frac{d}{2}}} e^{-\frac{\|x\|_2^2}{2}}, \eta = \mu, T(x) = x, Z(\eta) = \frac{\|\eta\|_2^2}{2}$.*

**Fact F.3.** *The $d$-dimensional Gaussian distribution with known mean and unknown covariance $\mathcal{N}(0, \Sigma)$ is an exponential family with support $S = \mathbb{R}^d$ and $h(x) = 1, \eta = \left(\Sigma^{-1}\right)^\flat, T(x) = -\frac{1}{2}\left(xx^\top\right)^\flat, Z(\eta) = \frac{d}{2}\ln(2\pi) - \frac{1}{2}\ln\left(\det\left(\eta^\#\right)\right)$.*

**Fact F.4** (Cauchy-Schwarz)**.** *We consider the following variants of the Cauchy-Schwarz inequality:*

1. *For every $x, y \in \mathbb{R}^d$, we have $|\langle x, y\rangle| \leq \|x\|_2 \|y\|_2$.*
2. *Let $X, Y$ be random variables over $\mathbb{R}$. Then:*

$$\mathbb{E}[|XY|] \leq \sqrt{\mathbb{E}[X^2]}\sqrt{\mathbb{E}[Y^2]},$$

*assuming all the quantities involved are well-defined.*

**Fact F.5.** *We denote the* complementary error function *by:*

$$\operatorname{erfc}(x) := \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2}\, dt, x \geq 0.$$

*It holds that* $\operatorname{erfc}(x) \leq e^{-x^2}, \forall x \geq 0$.

**Fact F.6** (Lemma 1 of [LM00])**.** *If $X \sim \chi^2(k)$, and $\beta \in [0, 1]$, then:*

$$\mathbb{P}\left[X - k \geq 2\sqrt{k \log\left(\frac{1}{\beta}\right)} + 2\log\left(\frac{1}{\beta}\right)\right] \leq \beta,$$

*and:*

$$\mathbb{P}\left[k - X \geq 2\sqrt{k \log\left(\frac{1}{\beta}\right)}\right] \leq \beta.$$

*Equivalently, the above can be written as:*

$$\mathbb{P}[X \geq t] \leq e^{-\frac{\left(\sqrt{2t-k}-\sqrt{k}\right)^2}{4}}, \forall t \geq k,$$

*and:*

$$\mathbb{P}[X \leq t] \leq e^{-\frac{(k-t)^2}{4}}, \forall t \leq k.$$

*Thus, if $Y \sim \mathcal{N}(0, \mathbb{I})$, then* $\mathbb{P}\left[\|Y\|_2^2 \geq d + 2\sqrt{d \log\left(\frac{1}{\beta}\right)} + 2\log\left(\frac{1}{\beta}\right)\right] \leq \beta$.

**Fact F.7** (Gaussian Hanson-Wright Inequality - Theorem 1 and Proposition 1 from [Mos21])**.** *Let $X \in \mathbb{R}^d$ be a random vector with $X \sim \mathcal{N}(0, \mathbb{I})$. If $A \in \mathbb{R}^{d \times d}$ is a non-zero symmetric matrix, we have:*

$$\mathbb{P}\left[\left|X^\top AX - \mathbb{E}\left[X^\top AX\right]\right| \geq t\right] \leq 2\exp\left(-c\min\left\{\frac{t^2}{\|A\|_F^2}, \frac{t}{\|A\|_2}\right\}\right), \forall t \geq 0,$$

*where $c$ is an absolute constant with $c \approx 0.036425$.*

## G Basic Applications: Recovering Existing Lower Bounds

We show how to use Theorem 2.4 to recover existing lower bounds from [KLSU19]: (1) mean estimation of binary product distributions, and (2) mean estimation of high-dimensional Gaussians. Note that both classes of distributions are exponential families. The error metric for these lower bounds would be the mean-squared-error (MSE), as opposed to constant-probability-error, as in [KLSU19], but the bounds could be converted to constant-probability bounds at the cost of a $\log(d)$ factor in the sample complexity (as was sketched in Section 1.1).

### G.1 Mean Estimation of Binary Product Distributions

We start by stating the theorem for mean estimation of binary product distributions over $\{0,1\}^d$.

**Theorem G.1** (Product Distribution Mean Estimation)**.** *There exists a distribution $\mathcal{D}$ over vectors $p \in \left[\frac{1}{3}, \frac{2}{3}\right]^d$ such that, given $p \sim \mathcal{D}$ and $n$ independent samples $X \coloneqq (X_1, \ldots, X_n)$ from a binary product distribution $P$ over $\{0,1\}^d$ with mean $p$, for any $\alpha = \mathcal{O}\left(\sqrt{d}\right)$ and any $(\varepsilon, \delta)$-DP mechanism $M\colon \{0,1\}^{n\times d} \to \left[\frac{1}{3}, \frac{2}{3}\right]^d$ with $\varepsilon, \delta \in [0,1]$, and $\delta = \mathcal{O}\left(\frac{1}{n}\right)$ that satisfies $\mathop{\mathbb{E}}_{X,M}\left[\|M(X) - p\|_2^2\right] \leq \alpha^2$, it holds that $n = \Omega\left(\frac{d}{\alpha\varepsilon}\right)$.*

Fact F.1 establishes that Bernoulli distributions are an exponential family. Using that, we have that the probability mass function of binary product distributions can be written as $p_\eta = h(x)\exp\left(\eta^\top T(x) - Z(\eta)\right), \forall x \in \{0,1\}^d$ with:

$$\begin{aligned} h(x) &= 1, \\ T(x) &= x, \\ \eta &= \left(\ln\left(\frac{1-p_1}{p_1}\right), \ldots, \ln\left(\frac{1-p_d}{p_d}\right)\right)^\top, \\ Z(\eta) &= \prod_{j\in[d]} \ln\left(1 + e^{\eta_j}\right). \end{aligned}$$

Our process to generate $\eta$ involves drawing independently $\eta_j \sim \mathcal{U}[\pm\ln(2)]$. We have that $\eta_j = \ln\left(\frac{1}{p_j} - 1\right) \iff p_j = \frac{1}{1+e^{\eta_j}}$, yielding $\eta_j \in [\pm\ln(2)] \iff p_j \in \left[\frac{1}{3}, \frac{2}{3}\right]$. Thus, we have $I_j = [\pm\ln(2)], R_j = 2\ln(2), \forall j \in [d] \implies \|R\|_2^2 = 4\ln^2(2)\, d, \|R\|_\infty = 2\ln(2)$, and $m = 0$.

We now show how to reduce estimating $\eta$ with an $\ell_2$-guarantee under $(\varepsilon, \delta)$-DP to estimating $p$ with an $\ell_2$-guarantee under the same constraint.

**Lemma G.2.** *Let $p \in \left[\frac{1}{3}, \frac{2}{3}\right]^d$ be a randomly generated vector and let $X \sim P^{\bigotimes n}$ be a dataset drawn from a binary product distribution over $\{0,1\}^d$ with mean $p$. If $M\colon \{0,1\}^{n\times d} \to \left[\frac{1}{3}, \frac{2}{3}\right]^d$ is an $(\varepsilon, \delta)$-DP mechanism satisfying $\mathop{\mathbb{E}}_{X,M}\left[\|M(X) - p\|_2^2\right] \leq \alpha^2 \leq \frac{d}{9}$, then there exists a $T_M\colon \{0,1\}^{n\times d} \to [\pm\ln(2)]^d$ that is also $(\varepsilon, \delta)$-DP and satisfies $\mathop{\mathbb{E}}_{X,T_M}\left[\|T_M(X) - \eta\|_2^2\right] \leq \frac{81}{4}\alpha^2$.*

*Proof.* We define $T_M$ to be $T_{M,j}(X) = \ln\left(\frac{1-M_j(X)}{M_j(X)}\right), \forall j \in [d]$. $T_M$ satisfies $(\varepsilon, \delta)$-DP by Lemma A.3 and the only randomness employed by $T_M$ is that of $M$. Consider now the function $g\colon (0,1) \to \mathbb{R}$ with $g(x) = \ln\left(\frac{1}{x} - 1\right)$ and $g'(x) = -\frac{1}{x(1-x)}$. We have $\eta_j = g(p_j)$ and $T_{M,j}(X) = g(T_j(X))$. Observe now that, by the Mean Value Theorem, we have for some $\xi_j$ between $\eta_j$ and $M_j(X)$ (implying that $\xi_j \in \left[\frac{1}{3}, \frac{2}{3}\right]$):

$$|g(M_j(X)) - g(p_j)| = |g'(\xi_j)|\,|M_j(X) - p_j| \leq \frac{9}{2}|M_j(X) - p_j|.$$

Applying this coordinate-wise, we get $\mathop{\mathbb{E}}_{X,T_M}\left[\|T_M(X) - \eta\|_2^2\right] \leq \frac{81}{4}\mathop{\mathbb{E}}_{X,M}\left[\|M(X) - p\|_2^2\right]$, yielding the desired result. ■

We now present a restatement of the previous lemma from the point of view of lower bounds.

**Corollary G.3.** *Let $p \in \left[\frac{1}{3}, \frac{2}{3}\right]^d$ be a randomly generated vector and let $X \sim P^{\bigotimes n}$ be a dataset drawn from a binary product distribution over $\{0,1\}^d$ with mean $p$. If any $(\varepsilon, \delta)$-DP mechanism $T\colon \{0,1\}^{n\times d} \to [\pm\ln(2)]^d$ satisfying $\mathop{\mathbb{E}}_{X,T}\left[\|T(X) - \eta\|_2^2\right] \leq \frac{81}{4}\alpha^2 \leq \frac{9}{4}d$ requires at least $n \geq n_\eta$ samples, the same sample complexity lower bound holds for any $(\varepsilon, \delta)$-DP mechanism $M\colon \{0,1\}^{n\times d} \to \left[\frac{1}{3}, \frac{2}{3}\right]^d$ that satisfies $\mathop{\mathbb{E}}_{X,M}\left[\|M(X) - p\|_2^2\right] \leq \alpha^2$.*

In contrast to other lower bounds proven in this work using Theorem 2.4, we do not need to study the concentration properties of $\|T(X_i) - \mu_T\|_2 = \|X_i - p\|_2$. Indeed, since $T(X) = X \in \{0,1\}^d$, it suffices to pick a large enough threshold $T > 0$ so that $\mathop{\mathbb{P}}_{X_i}\left[\|X_i - p\|_2 > \frac{t}{2\ln^3(2)\sqrt{d}}\right] = 0, \forall t \geq T$.

Based on the above, we are now ready to prove the main theorem of this section.

*Proof of Theorem G.1.* Let $M\colon \{0,1\}^{n\times d} \to [\pm\ln(2)]^d$ be an $(\varepsilon,\delta)$-DP mechanism with:

$$\mathop{\mathbb{E}}_{X,M}\left[\|M(X) - \eta\|_2^2\right] \leq \frac{81}{4}\alpha^2 \leq \frac{\ln^2(2)\,d}{6} = \frac{\|R\|_2^2}{24}.$$

By Theorem 2.4 for product distributions with $\eta_j \in [\pm\ln(2)], \forall j \in [d]$, we get:

$$n\left\{2\delta T + 2\varepsilon\mathop{\mathbb{E}}_{\eta}\left[\sqrt{\mathop{\mathbb{E}}_{X_{\sim i},M}\left[s^\top\Sigma_T s\right]}\right] + 2\int_T^\infty \mathop{\mathbb{P}}_{X_i}\left[\|X_i - p\|_2 > \frac{t}{2\ln^3(2)\sqrt{d}}\right]dt\right\} \geq \frac{\ln^2(2)\,d}{6}. \tag{26}$$

Observe that, since we have a product distribution and $T(x) = x$, it holds that $\mu_T = p$ and:

$$\Sigma_T = \Sigma = \operatorname{diag}\{p_1(1-p_1), \ldots, p_d(1-p_d)\} \preceq \frac{1}{4}\mathbb{I},$$

for $p_i \in \left[\frac{1}{3}, \frac{2}{3}\right], \forall i \in [d]$.

Thus, we have:

$$\mathop{\mathbb{E}}_{X_{\sim i},M}\left[s^\top\Sigma_T s\right] = \frac{1}{4}\mathop{\mathbb{E}}_{X_{\sim i},M}\left[\|s\|_2^2\right] \leq \frac{\|R\|_\infty^4}{64}\alpha^2 = \frac{\ln^4(2)\,\alpha^2}{4}. \tag{27}$$

Furthermore, we remark that $\|X_i - p\|_2 \leq \frac{2}{3}\sqrt{d}$. Thus, setting $T = \frac{4\ln^3(2)d}{3}$ yields:

$$\int_T^\infty \mathop{\mathbb{P}}_{X_i}\left[\|X_i - p\|_2 > \frac{t}{2\ln^3(2)\sqrt{d}}\right]dt = 0. \tag{28}$$

Substituting (27) and (28) into (26), we get:

$$n\left(\frac{8\ln^3(2)\,d\delta}{3} + \frac{\ln^2(2)\,\alpha\varepsilon}{2}\right) \geq \frac{\ln^2(2)\,d}{6}.$$

Setting $\delta \leq \frac{1}{32\ln(2)n}$ results in $n \geq \frac{d}{12\alpha\varepsilon}$, so appealing to Corollary G.3 completes the proof. ■

### G.2 Mean Estimation of High-Dimensional Gaussians

Our techniques also apply to recovering lower bounds for mean estimation of a Gaussian with known covariance. By Fact F.2, we have that this class of distributions is an exponential family with $\eta = \mu$ and $T(x) = x$ (implying that $\mu_T = \mu, \Sigma_T = \Sigma$). Therefore, we do not need to resort to a reduction-based approach (as we did with binary product distributions).

**Theorem G.4** (Gaussian Mean Estimation)**.** *Given $\mu \sim \mathcal{U}\left([\pm 1]^d\right)$ and $X \sim \mathcal{N}(\mu, \mathbb{I})^{\otimes n}$, for any $\alpha = \mathcal{O}\left(\sqrt{d}\right)$ and any $(\varepsilon,\delta)$-DP mechanism $M\colon \mathbb{R}^{n\times d} \to \mathbb{R}^d$ with $\varepsilon, \delta \in [0,1]$, and $\delta \leq \mathcal{O}\left(\min\left\{\frac{1}{n}, \frac{\sqrt{d}}{n\sqrt{\log\left(\frac{n}{\sqrt{d}}\right)}}\right\}\right)$ that satisfies $\mathop{\mathbb{E}}_{X,M}\left[\|M(X) - \mu\|_2^2\right] \leq \alpha^2$, it holds that $n = \Omega\left(\frac{d}{\alpha\varepsilon}\right)$.*

Before proving Theorem G.4, we start with a lemma that is concerned with upper-bounding the term involving the tail probability of $\|T(X_i) - \mu_T\|_2$. The lemma does not rely on the approach of Remark B.2, instead using the concentration properties of the $\chi^2$-distribution (Fact F.6).

**Lemma G.5.** *Let $\mu \in [\pm 1]^d$, and $X_i \sim \mathcal{N}(\mu, \mathbb{I})$. Assuming that $T \geq 2d$, we have:*

$$\int_T^\infty \mathop{\mathbb{P}}_{X_i}\left[\|T(X_i) - \mu_T\|_2 > \frac{t}{2\sqrt{d}}\right] dt \leq \sqrt{2\pi d} e^{-\frac{\left(\sqrt{T^2 - 2d^2} - \sqrt{2}d\right)^2}{8d}}.$$

*Proof.* Since $X_i \sim \mathcal{N}(\mu, \mathbb{I})$, we have $T(X_i) - \mu_T = X_i - \mu$. We observe that $X_i - \mu \sim \mathcal{N}(0, \mathbb{I}) \implies \|X_i - \mu\|_2^2 \sim \chi^2(d)$. By our assumption that $T \geq 2d \implies \frac{t^2}{4d} \geq \frac{T^2}{4d} \geq d$, we can apply Fact F.6 with $k = d$. This yields:

$$\begin{aligned}
\int_T^\infty \mathop{\mathbb{P}}_{X_i}\left[\|X_i - \mu\|_2^2 > \frac{t^2}{4d}\right] dt &\leq \int_T^\infty \exp\left[-\frac{1}{4}\left(\sqrt{\frac{t^2}{2d} - d} - \sqrt{d}\right)^2\right] dt \\
&\overset{z^2 = \frac{1}{4}\left(\sqrt{\frac{t^2}{2d} - d} - \sqrt{d}\right)^2}{=} 2\sqrt{d} \int_{\frac{\sqrt{\frac{T^2}{2d} - d} - \sqrt{d}}{2}}^\infty \frac{2z + \sqrt{d}}{\sqrt{z^2 + \left(z + \sqrt{d}\right)^2}} e^{-z^2} dz \\
&\overset{(a)}{\leq} 2\sqrt{2d} \int_{\frac{\sqrt{\frac{T^2}{2d} - d} - \sqrt{d}}{2}}^\infty e^{-z^2} dz \\
&= \sqrt{2\pi d}\,\mathrm{erfc}\left(\frac{\sqrt{\frac{T^2}{2d} - d} - \sqrt{d}}{2}\right) \\
&\overset{(b)}{\leq} \sqrt{2\pi d} e^{-\frac{\left(\sqrt{\frac{T^2}{2d} - d} - \sqrt{d}\right)^2}{4}} \\
&= \sqrt{2\pi d} e^{-\frac{\left(\sqrt{T^2 - 2d^2} - \sqrt{2}d\right)^2}{8d}},
\end{aligned}$$

where $(a)$ is by the inequality $a + b \leq \sqrt{2}\sqrt{a^2 + b^2}$ for $a = z, b = z + \sqrt{d}$, and $(b)$ is by Fact F.5. ■

We are now ready to prove the main theorem of this section.

*Proof of Theorem G.4.* Let $M \colon \mathbb{R}^{n \times d} \to [\pm 1]^d$ be an $(\varepsilon, \delta)$-DP mechanism with:

$$\mathop{\mathbb{E}}_{X, M}\left[\|M(X) - \mu\|_2^2\right] \leq \alpha^2 \leq \frac{d}{6} = \frac{\|R\|_2^2}{24}.$$

By Theorem 2.4 for Gaussians $\mathcal{N}(\mu, \mathbb{I})$ with $\mu^j \in [\pm 1], \forall j \in [d]$, we get:

$$n\left\{2\delta T + 2\varepsilon \mathop{\mathbb{E}}_{\mu}\left[\sqrt{\mathop{\mathbb{E}}_{X_{\sim i}, M}\left[\|s\|_2^2\right]}\right] + 2\int_T^\infty \mathop{\mathbb{P}}_{X_i}\left[\|X_i - \mu\|_2 > \frac{t}{2\sqrt{d}}\right] dt\right\} \geq \frac{d}{6}. \tag{29}$$

We have:

$$\mathop{\mathbb{E}}_{X_{\sim i}, M}\left[\|s\|_2^2\right] \leq \frac{\|R\|_\infty^4}{16}\alpha^2 = \alpha^2. \tag{30}$$

We assume that $T \geq 2d$, as is required by Lemma G.5. Later in the proof, when we pick a specific value for $T$, we will verify that this condition is satisfied. Substituting based on that lemma and (30) into (29), we get:

$$n\left[2\delta T + 2\alpha\varepsilon + 2\sqrt{2\pi d} e^{-\frac{\left(\sqrt{T^2 - 2d^2} - \sqrt{2}d\right)^2}{8d}}\right] \geq \frac{d}{6}. \tag{31}$$

It remains to set the values of $T$ and $\delta$ appropriately so that:

$$\delta T \geq 2\sqrt{2\pi d} e^{-\frac{\left(\sqrt{T^2-2d^2}-\sqrt{2}d\right)^2}{8d}} \text{ and } 3n\delta T \leq \frac{d}{12}.$$

The first of these two conditions yields:

$$Te^{\frac{\left(\sqrt{T^2-2d^2}-\sqrt{2}d\right)^2}{8d}} \geq \frac{2\sqrt{2\pi d}}{\delta}.$$

We set $T = 2\sqrt{d}\sqrt{\ln\left(\frac{1}{\delta}\right) + \left(\sqrt{\ln\left(\frac{1}{\delta}\right)} + \sqrt{d}\right)^2}$. Since $\delta \leq 1$, we have $\ln\left(\frac{1}{\delta}\right) \geq 0 \implies T \geq 2d$, thus satisfying the condition of Lemma G.5. Substituting to the previous yields:

$$\begin{aligned}
&2\sqrt{d}\sqrt{\ln\left(\frac{1}{\delta}\right) + \left(\sqrt{\ln\left(\frac{1}{\delta}\right)} + \sqrt{d}\right)^2} \cdot \frac{1}{\delta} \geq \frac{2\sqrt{2\pi d}}{\delta} \\
\iff &\sqrt{\ln\left(\frac{1}{\delta}\right) + \left(\sqrt{\ln\left(\frac{1}{\delta}\right)} + \sqrt{d}\right)^2} \geq \sqrt{2\pi} \\
\iff &\ln\left(\frac{1}{\delta}\right) + \sqrt{d}\sqrt{\ln\left(\frac{1}{\delta}\right)} + \frac{1}{2}(d - 2\pi) \geq 0.
\end{aligned}$$

Setting $x := \sqrt{\ln\left(\frac{1}{\delta}\right)} \geq 0$, the above is a quadratic inequality with respect to $x$. Depending on whether the trinomial in the LHS has any positive roots, this might imply a constraint on $\delta$. The discriminant is $\Delta = 4\pi - d$, so taking cases based on $d$, the strictest constraint we get is $\delta \leq e^{-\pi + \frac{\sqrt{4\pi - 1}}{2}} \approx 0.23$, which is for $d = 1$. Respecting this constraint, we now must identify a range of values for $\delta$ so that we have $3n\delta T \leq \frac{d}{12}$. We have:

$$3n\delta T \leq \frac{d}{12} \iff 6n\delta\sqrt{d}\sqrt{\ln\left(\frac{1}{\delta}\right) + \left(\sqrt{\ln\left(\frac{1}{\delta}\right)} + \sqrt{d}\right)^2} \leq \frac{d}{12}.$$

It holds that $\ln\left(\frac{1}{\delta}\right) + \left(\sqrt{\ln\left(\frac{1}{\delta}\right)} + \sqrt{d}\right)^2 \leq 2\left(\sqrt{\ln\left(\frac{1}{\delta}\right)} + \sqrt{d}\right)^2$, so it suffices to satisfy:

$$6\sqrt{2}n\delta\sqrt{d}\left(\sqrt{\ln\left(\frac{1}{\delta}\right)} + \sqrt{d}\right) \leq \frac{d}{12} \iff \delta\left(1 + \sqrt{\frac{\ln\left(\frac{1}{\delta}\right)}{d}}\right) \leq \frac{1}{72\sqrt{2}n}. \tag{32}$$

The previous inequality imposes a condition on $\delta$. Indeed, to ensure that $\sqrt{\frac{\ln\left(\frac{1}{\delta}\right)}{d}} \geq 0$, it is necessary that $\frac{1}{72\sqrt{2}n\delta} - 1 \geq 0 \iff \delta \leq \frac{1}{72\sqrt{2}n}$, which is stronger than $\delta \leq 0.23$.

For (32) to hold, we will show that it suffices to have:

$$\delta \leq \min\left\{\frac{1}{144\sqrt{2}n}, \frac{\sqrt{d}}{288\sqrt{2}n\sqrt{\ln\left(\frac{144\sqrt{2}n}{\sqrt{d}}\right)}}\right\},$$

which trivially satisfies the previous constraint.

For convenience, we set $\phi := \frac{\sqrt{d}}{144\sqrt{2}n}$, so the proposed range for $\delta$ becomes $\delta \leq \min\left\{\frac{\phi}{\sqrt{d}}, \frac{\phi}{2\sqrt{\ln\left(\frac{1}{\phi}\right)}}\right\}$. Additionally, (32) can be equivalently written as:

$$\delta\left(1 + \sqrt{\frac{\ln\left(\frac{1}{\delta}\right)}{d}}\right) \leq \frac{2\phi}{\sqrt{d}}. \tag{33}$$

To verify our claim, we consider the following two cases:

1. $\min\left\{\frac{\phi}{\sqrt{d}}, \frac{\phi}{2\sqrt{\ln\left(\frac{1}{\phi}\right)}}\right\} \leq e^{-d}$. We start by arguing that:

$$\min\left\{\frac{\phi}{\sqrt{d}}, \frac{\phi}{2\sqrt{\ln\left(\frac{1}{\phi}\right)}}\right\} \leq e^{-d} \iff \frac{\phi}{2\sqrt{\ln\left(\frac{1}{\phi}\right)}} \leq e^{-d}. \tag{34}$$

The $\Longleftarrow$ direction is trivial, so we focus on the $\Longrightarrow$ direction.

We assume that $\min\left\{\frac{\phi}{\sqrt{d}}, \frac{\phi}{2\sqrt{\ln\left(\frac{1}{\phi}\right)}}\right\} \leq e^{-d}$ but $\frac{\phi}{2\sqrt{\ln\left(\frac{1}{\phi}\right)}} > e^{-d}$. Then, it must be the case that $\frac{\phi}{\sqrt{d}} \leq e^{-d}$. Setting $y := \frac{1}{\phi}$, the above system of inequalities can be written in the form:

$$\begin{cases} \frac{e^{2d}}{4} > y^2 \ln(y) \\ y \geq \frac{e^d}{\sqrt{d}} \end{cases}.$$

The function $y^2 \ln(y)$ is increasing for $y \geq e^{-\frac{1}{2}}$. We have $\frac{e^d}{\sqrt{d}} > e^{-\frac{1}{2}}, \forall d \geq 1$. Thus, for $y \geq \frac{e^d}{\sqrt{d}}$, we get:

$$\frac{e^{2d}}{4} > y^2 \ln(y) \geq \frac{e^{2d}}{d} \ln\left(\frac{e^d}{\sqrt{d}}\right) \implies 3d < 2\ln(d).$$

However, this leads to a contradiction, because $3d < 2\ln(d) \leq 2d - 2 \implies d < -2$.

Consequently, we have shown (34). Since we have assumed that $\min\left\{\frac{\phi}{\sqrt{d}}, \frac{\phi}{2\sqrt{\ln\left(\frac{1}{\phi}\right)}}\right\} \leq e^{-d}$, it must be the case that $\frac{\phi}{2\sqrt{\ln\left(\frac{1}{\phi}\right)}} \leq e^{-d}$.

We now turn our attention again to (33). We remark that, for $\delta \leq e^{-d}$, we get $\sqrt{\frac{\ln\left(\frac{1}{\delta}\right)}{d}} \geq 1$. As a result, in order to satisfy (33), it suffices to have:

$$2\delta\sqrt{\frac{\ln\left(\frac{1}{\delta}\right)}{d}} \leq \frac{2\phi}{\sqrt{d}} \iff \delta\sqrt{\ln\left(\frac{1}{\delta}\right)} \leq \phi. \tag{35}$$

Observe that the function $\delta\sqrt{\ln\left(\frac{1}{\delta}\right)}$ is increasing for $\delta \leq e^{-d} < e^{-\frac{1}{2}}$. Combined with our previous remarks, this implies that, in order to verify (35) for the values of $\delta$ we picked, it suffices to show that it holds for $\delta = \frac{\phi}{2\sqrt{\ln\left(\frac{1}{\phi}\right)}}$. So, we have to verify the inequality:

$$\left(\frac{\phi}{2\sqrt{\ln\left(\frac{1}{\phi}\right)}}\right)\sqrt{\ln\left(\frac{2\sqrt{\ln\left(\frac{1}{\phi}\right)}}{\phi}\right)} \leq \phi \iff \frac{\ln\left(2\sqrt{\ln\left(\frac{1}{\phi}\right)}\right)}{\ln\left(\frac{1}{\phi}\right)} \leq 3.$$

Setting $z := \ln\left(\frac{1}{\phi}\right)$, this becomes equivalent to $\frac{\ln\left(2\sqrt{z}\right)}{z} \leq 3$, which holds for all $z > 0 \iff \ln\left(\frac{1}{\phi}\right) > 0 \iff \phi < 1$. This last condition is satisfied by all $\phi$ such that $\frac{\phi}{2\sqrt{\ln\left(\frac{1}{\phi}\right)}} \leq e^{-d}$, so the desired result has been established.

2. $\min\left\{\frac{\phi}{\sqrt{d}}, \frac{\phi}{2\sqrt{\ln\left(\frac{1}{\phi}\right)}}\right\} > e^{-d}$. For this case, we consider two sub-cases, depending on how $\delta$ compares with $e^{-d}$. We have:

- $\delta \leq e^{-d}$. As before, we argue that it suffices to have $\delta\sqrt{\ln\left(\frac{1}{\delta}\right)} \leq \phi$. Since $\delta\sqrt{\ln\left(\frac{1}{\delta}\right)}$ is increasing for $\delta \leq e^{-d} < e^{-\frac{1}{2}}$, all we have to do is show that $e^{-d}\sqrt{d} \leq \phi$. This, however, is an immediate consequence of our assumption that $\min\left\{\frac{\phi}{\sqrt{d}}, \frac{\phi}{2\sqrt{\ln\left(\frac{1}{\phi}\right)}}\right\} > e^{-d}$.
- $\delta > e^{-d}$. This implies $\frac{\ln\left(\frac{1}{\delta}\right)}{d} < 1$. Thus, for (33) to hold, it suffices that:

$$2\delta \leq \frac{2\phi}{\sqrt{d}} \iff \delta \leq \frac{\phi}{\sqrt{d}}.$$

  This is satisfied, because our proposed range for $\delta$ is $\delta \leq \min\left\{\frac{\phi}{\sqrt{d}}, \frac{\phi}{2\sqrt{\ln\left(\frac{1}{\phi}\right)}}\right\}$.

We have established that:

$$T = 2\sqrt{d}\sqrt{\ln\left(\frac{1}{\delta}\right) + \left(\sqrt{\ln\left(\frac{1}{\delta}\right)} + \sqrt{d}\right)^2} \text{ and } \delta \leq \mathcal{O}\left(\min\left\{\frac{1}{n}, \frac{\sqrt{d}}{n\sqrt{\log\left(\frac{n}{\sqrt{d}}\right)}}\right\}\right),$$

imply:

$$\delta T \geq 2\sqrt{2\pi d}e^{-\frac{\left(\sqrt{T^2-2d^2}-\sqrt{2}d\right)^2}{8d}} \text{ and } 3n\delta T \leq \frac{d}{12},$$

while respecting all the constraints.

We substitute this to (31) and get $n \geq \frac{d}{24\alpha\varepsilon}$, which yields the desired result. ■