DIFFERENTIALLY PRIVATE APPROXIMATION OF THE JOHN ELLIPSOID

Anonymous authors
Paper under double-blind review

ABSTRACT

We study the problem of approximating the John ellipsoid (JE) of a given (centrally symmetric) polytope of n constraints in a Euclidean space under differential privacy (DP). We give the first differentially private algorithm for this problem under the standard model, where neighboring datasets may differ arbitrarily in one a single constraint. Our work also extends to the complimentary problem of Minimum Enclosing Ellipsoid of n points in the Euclidean space.

Our approach is based on the recent non-private multiplicative-weights algorithm of Cohen et al. (2019). First we introduce a non-private generalization of the Cohen et al algorithm, yielding a $(1+\gamma)$-approximation of the JE problem while violating at most κn constraints in $O(\log(1/\kappa)/\gamma)$ iterations. This variant works by projecting the intermediate weights assigned to the constraints onto the set of κ-dense distributions, similarly to Bun et al. (2020).

We then design a ρ-zCDP variant of this algorithm by adding Gaussian noise to the weighted covariance matrix aggregated in each step of the algorithm. Under a mild goodness assumption on the data we can assert that the resulting noisy matrix is close to the true matrix, thereby achieving essentially the same guarantee as the non-private algorithm provided sufficiently many input points. Thus our method achieves an efficient DP poly-time algorithm under concrete sample complexity bounds.

1 INTRODUCTION

The John ellipsoid is one of the most fundamental geometric objects associated with a convex body. Given a convex polytope $K \subseteq \mathbb{R}^d$, the John Ellipsoid (JE) is defined as the maximum-volume ellipsoid contained in K. A classical theorem of John (1948) shows that this ellipsoid provides a canonical geometric approximation of the body: if $E \subseteq K$ is the John ellipsoid, then $K \subseteq d \cdot E$, where $d \cdot E$ denotes a dilation of E by a factor of d about its center. Moreover, when K is centrally symmetric (i.e. $\forall x \in K$ its antipode $-x$ is also in K) and the JE must be centered at the origin, this guarantee improves to $K \subseteq \sqrt{d}\, E$.

These structural properties make the John ellipsoid a central tool in high-dimensional geometry and its algorithmic applications. These include sampling and numerical integration (Vempala & Institute, 2005; Chen et al., 2018), linear bandit problems (Bubeck et al., 2012; Hazan & Karnin, 2016), linear programming (Lee & Sidford, 2014), and differential privacy bounds (Nikolov et al., 2013). Beyond these areas, the John ellipsoid also plays an important role in experimental design, a classical topic in statistics (Atwood, 1969). In particular, the D-optimal design problem seeks to maximize the determinant of the Fisher information matrix (Kiefer & Wolfowitz, 1960; Atwood, 1969), which is known to be equivalent to computing the John ellipsoid of a symmetric polytope (see, e.g., Todd (2016)). D-optimal design has attracted renewed interest in the machine learning community (Allen-Zhu et al., 2017; Wang et al., 2017; Lu et al., 2018).

As is common, we consider in this work symmetric polytopes of the form $K = \{x \in \mathbb{R}^d : -\mathbf{1} \leq Px \leq \mathbf{1}\}$, where the $n \times d$ matrix P is composed of n constraints each denote by x_i. In many applications, these n constraints may contain sensitive information—such as locations, personal attributes, or high-dimensional feature vectors—motivating the need for private algorithms. Differential privacy (Dwork et al., 2006) requires that the output of an algorithm is stable under the modification of a single x_i in P, thereby protecting individual-level information. However, this

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

requirement is particularly challenging in geometric settings: core quantities such as covariance matrices and enclosing or inscribed ellipsoids can be highly sensitive, and in some instances a single point may significantly alter the optimal solution. This is particularly true for the JE problem, where by adding to P a new constraint x_{n+1} we may obtain a new polytope K' of significantly smaller volume than K, and therefore a significantly smaller JE. This inherent instability makes the direct computation of the John ellipsoid incompatible with privacy constraints. Consequently, we adopt a *bi-criteria formulation*: we allow a multiplicative approximation of $1+\gamma$ to the JE volume objective and we also allow a relaxation of the containment requirement – it may violate up to a κ-fraction of the constraints.

Our contribution. In this work, we present a new framework for approximating the John ellipsoid under differential privacy, based on multiplicative-weights algorithm and KL-projection. Our starting point is the non-private algorithm of Cohen et al. (2019), which uses a multiplicative update over weights and returns, after $O(\log(n/d)/\gamma)$ iterations, a $(1+\gamma)$-approximation of the input's JE. Our design of a DP-equivalent of this algorithm is composed of two stages.

The first step is to introduce a (non-private) generalization of the algorithm of Cohen et al, in which the weights assigned to constraints in each iteration must be κ-dense — namely, each weight is upper bounded by $\frac{1}{\kappa} \cdot \frac{d}{n}$. We do so by projecting the true weights on the set of κ-dense weights, a projection that is simple to compute. In Section 3 we analyze the resulting algorithm and present the following guarantee.

Theorem 1.1. *Let $P = \{x_1, \ldots, x_n\} \subseteq \mathbb{R}^d$. For any $\gamma, \kappa \in (0,1)$, after $T = O\Big(\frac{\log(1/\kappa)}{\gamma}\Big)$ iterations, Algorithm 1 outputs a positive definite matrix M, defining the ellipsoid*

$$E := \{x \in \mathbb{R}^d : x^\top M^{-1} x \leq 1\},$$

such that: there exists a subset $S \subseteq [n]$ with $|S| \geq (1-\kappa)n$ such that

$$e^{-\gamma/2} E \subseteq K_S,$$

where $K_S := \{x \in \mathbb{R}^d : |x_i^\top x| \leq 1 \ \forall i \in S\}$; and, for the original polytope we have

$$K \subseteq \sqrt{d}\, E,$$

where $K := \{x \in \mathbb{R}^d : |x_i^\top x| \leq 1$ for all $i \in [n]\}$; and moreover, we have

$$\mathrm{vol}(e^{-\gamma/2} E) \ \geq \ e^{-d\gamma/2}\, \mathrm{vol}(E^*),$$

where E^* is the John ellipsoid of K.

Our ellipsoid is γ-close to the John Ellipsoid in any direction, (just like as the approximation of Cohen et al). If we wish to get γ^*-close approximation in volume, we simply set $\gamma = \gamma^*/d$, resulting in an algorithm that runs in $O\Big(\frac{d\log(1/\kappa)}{\gamma^*}\Big)$ many iterations.

The second step is to present a zCDP analogue of the above-mentioned algorithm. We do simply by applying the Gaussian mechanism to the weighted covariance matrix aggregated in each step of the algorithm. Leveraging on the fact that the weight of each x_i is upper bounded due to κ-density, it is easy to upper bound the sensitivity of the weighted covariance matrix and scale the Gaussian noise accordingly. However, in order to assert that the added noise does not interfere with the weights assigned to the constraints we need to rely on a goodness assumption.

Definition 1.2 ((κ,τ)-good input)**.** We say that $P = \{x_1, \ldots, x_n\} \subseteq \mathbb{R}^d$ is (κ,τ)-good if for every unit vector $v \in \mathbb{S}^{d-1}$, defining

$$\mathrm{BAD}_v := \left\{ i \in [n] : \langle x_i, v\rangle^2 < \frac{2\tau}{d} \right\},$$

we have $|\mathrm{BAD}_v| < \frac{\kappa n}{2}$.

For example, when the n constraints are drawn from the uniform distribution over the unit-sphere, then the goodness assumption holds for constant values of κ and τ. For our purposes, the (κ,τ)-

goodness assumption ensures that every κ-dense measure induces a well-conditioned covariance matrix, which is needed to control the effect of Gaussian noise in the private algorithm (See Lemma 4.2). Altogether, in Section 4 we introduce and analyze our ρ-zCDP algorithm, obtaining the following guarantee.

Theorem 1.3. *Let $P = \{x_1, \ldots, x_n\} \subseteq \mathbb{R}^d$ be (κ, τ)-good input with range R. For any ρ and $\gamma, \beta \in (0,1)$, Algorithm 2 is ρ-zCDP, and, after $T = O\Big(\frac{\log(1/\kappa)}{\gamma}\Big)$ iterations, supposing that*

$$n \geq \frac{32dR^2(\sqrt{d} + \sqrt{2\log(\frac{T+1}{\beta})})}{\kappa\gamma\tau\sqrt{2\rho/(T+1)}} = O\left(\frac{dR^2(\sqrt{d} + \sqrt{\log(\frac{\log(1/\kappa)}{\gamma\beta})})\sqrt{\log(1/\kappa)}}{\kappa\gamma^{1.5}\tau\sqrt{\rho}}\right)$$

it outputs with probability $\geq 1 - \beta$ a positive definite matrix M, defining the ellipsoid

$$E := \{x \in \mathbb{R}^d : x^\top M^{-1} x \leq 1\},$$

such that:

- *there exists a subset $S \subseteq [n]$ with $|S| \geq (1-\kappa)n$ such that*

  $$\sqrt{(1-\gamma/8)}e^{-\gamma/2}E \subseteq K_S,$$

  where $K_S := \{x \in \mathbb{R}^d : |x_i^\top x| \leq 1 \; \forall i \in S\}$; and, for the original polytope we have

  $$K \subseteq \sqrt{d(1+\gamma/8)}\, E,$$

  where $K := \{x \in \mathbb{R}^d : |x_i^\top x| \leq 1$ for all $i \in [n]\}$.

- *moreover, we have*

  $$\mathrm{vol}(\sqrt{(1-\gamma/8)}e^{-\gamma/2}E) \;\geq\; e^{-9d\gamma/14}\,\mathrm{vol}(E^*),$$

  where E^* is the John ellipsoid of K.

We comment that our algorithm does not require the goodness condition to be checked before it is executed. The goodness assumption is used only in our utility analysis, guaranteeing that each weighted second-moment matrix the algorithm iterates through has sufficiently large least eigenvalue. Accordingly, our sample-complexity bound should be interpreted as a sufficient guarantee for inputs satisfying this condition, rather than as a characterization of all inputs on which the algorithm succeeds — as our algorithm may succeed even on inputs that are not (κ, τ)-good.

Extensions. Our problem also admits an equivalent interpretation through polar duality. Let $C_P := \mathrm{conv}\{\pm x_1, \ldots, \pm x_n\}$. Then the centrally symmetric polytope considered throughout this work satisfies $K_P = C_P^\circ$, where $(\cdot)^\circ$ denotes the polar body. Since polarity reverses containment and maps origin-centered ellipsoids to origin-centered ellipsoids, the John ellipsoid of K_P is precisely the polar of the minimum-volume enclosing ellipsoid (MVEE) of C_P. Consequently, our private algorithm also yields guarantees for a robust version of the centered MVEE problem in which a κ-fraction of the input points may be discarded as outliers.

This MVEE interpretation further allows us to extend our results beyond centrally symmetric, origin-centered ellipsoids. In particular, the standard homogeneous-coordinate lifting for the MVEE problem (Todd, 2016) embeds each point $y_i \in \mathbb{R}^d$ as $x_i = (y_i, 1) \in \mathbb{R}^{d+1}$, thereby reducing the uncentered MVEE problem in $\mathbb{R}^d$ to a centered problem in one higher dimension. We show that this reduction extends to our trimmed setting and preserves both the approximation factor and the trimming guarantee. Hence, through the MVEE interpretation above, our results extend also to general uncentered ellipsoids. For the private guarantee in the uncentered setting, we additionally require that the lifted points satisfy the corresponding goodness assumption. We discuss this in Section 5.

Related Work. Computing or approximating the John Ellipsoid (JE) has been extensively studied in optimization, computational geometry, and statistics. Classical approaches rely on reductions to the minimum volume enclosing ellipsoid (MVEE) and interior-point methods (Khachiyan & Todd, 1993; Nesterov & Nemirovskii, 1994), with subsequent improvements achieving a $(1+\gamma)$-approximation in time $O(dn^3/\gamma)$ (Kumar & Yildirim, 2005; Todd & Yıldırım, 2007). The problem

is also equivalent to D-optimal design (Atwood, 1969; Todd, 2016), which has motivated first-order and proximal methods (Lu et al., 2018; Gutman & Pena, 2018). More recently, faster algorithms based on iterative updates and sketching techniques have been developed (Cohen et al., 2019; Woodruff & Yasuda, 2024; Cao et al., 2026; Li et al., 2024), with the method of Cohen et al. (2019) achieving near-optimal iteration complexity via a simple multiplicative update.

In the context of differential privacy, much less is known for geometric optimization problems of this type. A closely related line of work studies the private 1-cluster problem, introduced by Nissim et al. (2016), where the goal is to find a small-radius ball containing almost all of a prescribed number of input points. Subsequent work obtained improved guarantees for this problem in both the central and local models (Nissim & Stemmer, 2018), and Ghazi et al. (2020) gave efficient private algorithms achieving essentially the corresponding non-private approximation ratios, up to a small additive loss in the number of covered points. For the closely related Minimum Enclosing Ball (MEB) problem, Mahpud & Sheffet (2022) gave the first linear-time differentially private fPTAS, achieving a $(1+\gamma)$-approximation to the optimal radius while allowing a small fraction of points to remain uncovered. Taken together, these works illustrate a recurring feature of private geometric optimization: under the standard neighboring-dataset model, useful utility guarantees naturally take a bi-criteria form, simultaneously approximating the geometric objective and permitting a small number of constraints or points to be violated.

Gu et al. (2024) proposed a differentially private algorithm for approximating the John ellipsoid under a weaker notion of neighboring inputs, where a single datum can be perturbed only by a small ϵ_0 (i.e., the two neighboring instances differ on a single entry where $\|x_i - x_i'\| \leq \epsilon_0$). Their claim-to-fame is to upper-bound the ℓ_2-sensitivity of the weighted leverage-score vector by $L\epsilon_0$ where L is some factor left unspecified[1]; thus allowing them to revise the algorithm of Cohen et al. so that *all* n constraints are approximately satisfied (a guarantee clearly infeasible under standard DP). We, in contrast, use the the standard definition of DP while introducing concrete privacy-utility trade-off in the form of an explicit sample complexity bound (and also achieving smaller iteration complexity then their algorithm).

Our use of dense measures is motivated by the smooth boosting literature, particularly the lazy Bregman-projection framework of Bun et al. (2020), building on earlier work of Barak et al. (2009). In that line of work, multiplicative weights are combined with projection onto a dense class of measures to keep the influence of every datum controlled in a similar fashion to Hsu et al. (2014). We adapt this mechanism to the John ellipsoid problem, where the maintained measure determines a weighted covariance matrix and the multiplicative update is driven by logarithmic leverage-type scores.

2 PRELIMINARIES

John ellipsoid and geometric formulation. Let $P = \{x_1, \ldots, x_n\} \subseteq \mathbb{R}^d$, where $\operatorname{span}(P) = \mathbb{R}^d$. We consider the centrally symmetric polytope

$$K_P := \left\{x \in \mathbb{R}^d : |x_i^\top x| \leq 1,\ \forall i \in [n]\right\}$$

The John ellipsoid E^* of K_P is the maximum-volume ellipsoid contained in K_P. By symmetry, E^* is centered at the origin.

For a nonnegative measure $\mu \in \mathbb{R}^n_{\geq 0}$ with $\sum_{i=1}^n \mu_i = d$, define

$$\Sigma(\mu) := \sum_{i=1}^n \mu_i x_i x_i^\top, \qquad M(\mu) := \Sigma(\mu)^{-1},$$

whenever $\Sigma(\mu)$ is invertible. The corresponding ellipsoid is

$$E(\mu) := \left\{x \in \mathbb{R}^d : x^\top M(\mu)^{-1} x \leq 1\right\}$$

and its volume is proportional to $\det(M(\mu))^{1/2}$.

[1]They bound L as some (unspecified) polynomial in the numerous parameters introduced by their various assumptions.

The primal-dual formulation of the John ellipsoid provides a geometric interpretation of these measures and the corresponding volume guarantees. We recall this formulation and its consequences in Appendix A.

Definition 2.1 ($(1+\gamma)$-approximate κ-trimmed John solution)**.** Let $\gamma > 0$ and $\kappa \in (0,1)$. A nonnegative measure μ of total mass d is a $(1+\gamma)$-approximate κ-trimmed John solution if

$$\left|\left\{i \in [n] : x_i^\top M(\mu) x_i \le e^\gamma\right\}\right| \ge (1-\kappa)n$$

In particular, a scaling $E(\mu)$ by $e^{-\gamma}$ is contained inside the polytope induced by at least $(1-\kappa)n$ constraints, while its volume approximates that of E^*. The corresponding geometric guarantees are established in Appendix A.

Dense measures and KL projection. For $\kappa \in (0,1)$, define the set of κ-dense measures

$$\Gamma_\kappa := \left\{\mu \in \mathbb{R}^n_{\ge 0} : \sum_{i=1}^n \mu_i = d, \quad \mu_i \le \frac{d}{\kappa n}\ \forall i \in [n]\right\}$$

For two measures μ, ν, their KL divergence is $D_{\mathrm{KL}}(\mu\|\nu) := \sum_{i=1}^n \left[\mu_i \log\left(\frac{\mu_i}{\nu_i}\right) + \nu_i - \mu_i\right]$. The KL projection onto Γ_κ is $\Pi_{\Gamma_\kappa}(\nu) := \arg\min_{\mu\in\Gamma_\kappa} D_{\mathrm{KL}}(\mu\|\nu)$.

By the Pythagorean property of Bregman projections (Bregman, 1967), for every $q \in \Gamma_\kappa$ and measure ν that sums to d

$$D_{\mathrm{KL}}(q\|\Pi_{\Gamma_\kappa}(\nu)) \le D_{\mathrm{KL}}(q\|\nu)$$

The explicit projection and its computation are given in Appendix B.

Leverage scores. For every $i \in [n]$, define $\psi_i(\mu) := x_i^\top M(\mu) x_i$, $\phi_i(\mu) := \log \psi_i(\mu)$. The function ϕ_i is convex in μ whenever $\Sigma(\mu)$ is positive definite (Cohen et al., 2019). Moreover, the weighted leverage scores $\ell_i(\mu) := \mu_i \psi_i(\mu)$ satisfy $0 \le \ell_i(\mu) \le 1$ and $\sum_{i=1}^n \ell_i(\mu) = d$.

Differential privacy (DP). Two datasets are neighboring if they differ in at most one constraint. We use zero-concentrated differential privacy (zCDP) (Bun & Steinke, 2016). A randomized mechanism $\mathcal{M}$ is ρ-zCDP if, for all neighboring datasets P, P' and all $\alpha > 1$,

$$D_\alpha(\mathcal{M}(P)\|\mathcal{M}(P')) \le \alpha\rho$$

where D_α denotes the Rényi divergence of order α.

We use the Gaussian mechanism and the adaptive composition property of zCDP. In particular, the composition of mechanisms with privacy parameters $\rho_1, \dots, \rho_T$ satisfies $(\sum_{t=1}^T \rho_t)$-zCDP. Furthermore, ρ-zCDP implies (ε, δ)-DP for $\varepsilon = \rho + 2\sqrt{\rho \log(1/\delta)}$.

Additional properties of DP and Gaussian concentration facts are recalled in Appendix A.

3 A Non-Private Bi-Criteria Approximating Algorithm for JE

In this section, we present a deterministic algorithm for approximating the John ellipsoid while allowing a κ-fraction of the constraints to be violated. Our algorithm extends the multiplicative-weights approach of Cohen et al. (2019) by projecting the intermediate weights onto the set of κ-dense measures. This yields the following guarantee.

Theorem 3.1. *Let $P = \{x_1, \dots, x_n\} \subseteq \mathbb{R}^d$ span $\mathbb{R}^d$. For any $\gamma, \kappa \in (0,1)$, Algorithm 1 outputs, after $T = O\left(\frac{\log(1/\kappa)}{\gamma}\right)$ iterations, a positive definite matrix M defining the ellipsoid*

$$E := \{x \in \mathbb{R}^d : x^\top M^{-1} x \le 1\},$$

such that (i) there exists $S \subseteq [n]$ with $|S| \ge (1-\kappa)n$ for which

$$e^{-\gamma/2} E \subseteq K_S, \qquad \text{and} \qquad K \subseteq \sqrt{d}\, E,$$

where $K_S := \{x : |x_i^\top x| \le 1,\ \forall i \in S\}$; and moreover (ii) $\mathrm{vol}(e^{-\gamma/2}E) \ge e^{-d\gamma/2}\,\mathrm{vol}(E^*)$,. where E^* is the John ellipsoid of K_P.

The algorithm. We maintain an unprojected positive measure w^t together with its κ-dense KL projection $\mu^t := \Pi_{\Gamma_\kappa}(w^t) \in \Gamma_\kappa$. At each iteration, the projected measure μ^t determines the weighted covariance matrix $\Sigma(\mu^t) = \sum_{i=1}^n \mu_i^t x_i x_i^\top$ and its inverse $M^t = M(\mu^t)$. The unprojected weights are then updated multiplicatively according to the corresponding quadratic scores, $w_i^{t+1} = w_i^t x_i^\top M^t x_i$. Thus, the projection is used to determine the measure supplied to the covariance oracle, while the multiplicative update is applied to the unprojected weights.

Algorithm 1 Non-Private John Ellipsoid Approximation

Input: $P = \{x_i\}_{i=1}^n$, $\gamma, \kappa \in (0,1)$.
$T \leftarrow \lceil \log(1/\kappa)/\gamma \rceil$
$w_i^0 \leftarrow d/n$ for all $i \in [n]$
for $t = 0, \dots, T-1$ **do**
  $\mu^t \leftarrow \Pi_{\Gamma_\kappa}(w^t)$
  $M^t \leftarrow \left(\sum_{i=1}^n \mu_i^t x_i x_i^\top\right)^{-1}$
  for $i = 1, \dots, n$ **do**
    $w_i^{t+1} \leftarrow w_i^t x_i^\top M^t x_i$
$\overline{\mu} \leftarrow \frac{1}{T}\sum_{t=0}^{T-1} \mu^t$
return $\left(\sum_{i=1}^n \overline{\mu}_i x_i x_i^\top\right)^{-1}$

Analysis. The main ingredient of our analysis is a KL-potential inequality that compares the algorithm's iterates with any fixed κ-dense measure. Recall that $\phi_i(\mu) := \log(x_i^\top M(\mu) x_i)$.

Lemma 3.2. Let $\overline{\mu} = \frac{1}{T}\sum_{t=0}^{T-1} \mu^t$. Then, for every $q \in \Gamma_\kappa$,

$$\sum_{i=1}^n q_i \phi_i(\overline{\mu}) \le \frac{D_{\mathrm{KL}}(q\|\mu^0)}{T}$$

Proof. Define $\theta_i^t := \log\left(\frac{w_i^t}{\mu_i^0}\right)$, $F(\theta) := \max_{\mu \in \Gamma_\kappa}\left\{\langle \mu, \theta\rangle - D_{\mathrm{KL}}(\mu\|\mu^0)\right\}$. Since $w_i^t = \mu_i^0 e^{\theta_i^t}$ and $\mu^t = \Pi_{\Gamma_\kappa}(w^t)$, the measure μ^t attains the maximum in

$$F(\theta^t) = \max_{\mu\in\Gamma_\kappa}\left\{\langle \mu, \log\left(\frac{w^t}{\mu^0}\right)\rangle + \sum_i w_i^t - d - \langle \mu, \log\left(\frac{\mu}{\mu^0}\right)\rangle\right\} = \max_{\mu\in\Gamma_\kappa}\left\{\sum_i w_i^t - d - D_{\mathrm{KL}}(\mu||w^t)\right\}$$

By the Pythagorean property of KL projection (Theorem A.12) for every $\nu \in \Gamma_\kappa$,

$$F(\theta^t) - \left(\langle \nu, \theta^t\rangle - D_{\mathrm{KL}}(\nu\|\mu^0)\right) = -D_{\mathrm{KL}}(\mu^t||w^t) + D_{\mathrm{KL}}(\nu||w^t) \ge D_{\mathrm{KL}}(\nu||\mu^t)$$

Now note that the update rule gives $\theta_i^{t+1} = \theta_i^t + \phi_i(\mu^t)$. Hence,

$$\begin{aligned}
F(\theta^{t+1}) &= \max_{\mu\in\Gamma_\kappa}\left\{\langle \mu, \theta^t\rangle + \sum_i \mu_i\phi_i(\mu^t) - D_{\mathrm{KL}}(\mu\|\mu^0)\right\} \le \max_{\mu\in\Gamma_\kappa}\left\{F(\theta^t) - D_{\mathrm{KL}}(\mu\|\mu^t) + \sum_i \mu_i\phi_i(\mu^t)\right\} \\
&= F(\theta^t) + \max_{\mu\in\Gamma_\kappa}\left\{\sum_i \mu_i\left(\log(e^{\phi_i(\mu^t)}) - \log(\frac{\mu_i}{\mu_i^t})\right)\right\} = F(\theta^t) + \max_{\mu\in\Gamma_\kappa}\left\{\sum_i \mu_i\left(-\log(\mu_i/\mu_i^t e^{\phi_i(\mu^t)}\right)\right\} \\
&\overset{(*)}{\le} F(\theta^t) + d\left(\log(d/\sum_i \mu_i^t x_i^\top M(\mu^t) x_i\right) \overset{(**)}{\le} F(\theta^t) + d\log(\frac{d}{d}) = F(\theta^t)
\end{aligned}$$

where inequality $(*)$ follows from the log-sum inequality and the fact that $\sum_i \mu_i = d$, and inequality $(**)$ is due to $\sum_i \mu_i e^{\phi_i(\mu^t)} = \sum_i \mu_i^t x_i^\top M(\mu^t) x_i = d$. As $\theta^0 = 0$, we have $F(\theta^T) \le F(\theta^0) = 0$. Fix any $q \in \Gamma_\kappa$ and note that $\theta_i^T = \sum_t \phi_i(\mu^t)$ so

$$\sum_{t=0}^{T-1}\sum_i q_i\phi_i(\mu^t) = \langle q, \theta^T\rangle \le F(\theta^T) + D_{\mathrm{KL}}(q\|\mu^0) \le D_{\mathrm{KL}}(q\|\mu^0)$$

Dividing by T and applying Jensen's inequality to the convex functions ϕ_i gives

$$\sum_i q_i\phi_i(\overline{\mu}) \le \frac{1}{T}\sum_{t=0}^{T-1}\sum_i q_i\phi_i(\mu^t) \le \frac{D_{\mathrm{KL}}(q\|\mu^0)}{T} \qquad \square$$

We now give the summary of the proof of Theorem 3.1. The uniform initialization and the density constraint imply

$$D_{\mathrm{KL}}(q\|\mu^0) \le d\log(1/\kappa), \qquad \forall q \in \Gamma_\kappa$$

Combining this bound with Lemma 3.2 and our choice of T gives

$$\sup_{q\in\Gamma_\kappa} \sum_{i=1}^{n} \frac{q_i}{d} \log\big(x_i^\top M(\overline{\mu}) x_i\big) \le \frac{1}{d} \cdot \frac{d\log(1/\kappa)}{T} \le \gamma \tag{1}$$

In particular, sort the n constraints $x_{(1)}, x_{(2)}, ..., x_{(n)}$ in descending order according to $\phi_i(\bar{\mu})$, and set q as the uniform (dense) measure over the top κn constraints. Plugging-in this q into Equation (1) above we obtain $\phi_{(\kappa n)}(\bar{\mu}) \le \sum_{i=1}^{\kappa n} \frac{1}{\kappa n}\phi_{(i)}(\bar{\mu}) \le \gamma$, asserting that at least $(1-\kappa)n$ constraints satisfy $x_i^\top M(\overline{\mu}) x_i \le e^\gamma$. The corresponding ellipsoid therefore satisfies the trimmed-containment guarantee after scaling by $e^{-\gamma/2}$.

Finally, the dual formulation of the John ellipsoid yields the volume guarantee of Theorem 3.1, while the normalization $\sum_i \overline{\mu}_i = d$ implies $K \subseteq \sqrt{d}\, E$ as shown in Cohen et al. (2019). The complete arguments are provided in Appendix C.

4 A Differentially Private Algorithm for Approximating the John Ellipsoid

We now extend the projected multiplicative-weights algorithm of Section 3 to the private setting. The algorithm has the same structure as its non-private counterpart, but replaces each exact weighted covariance matrix with a Gaussian- perturbed one before inversion. The projection onto Γ_κ bounds the weight assigned to every constraint, while the (κ,τ)-goodness assumption guarantees that the weighted covariance matrices remain sufficiently well-conditioned for the perturbation to have only a small effect on the update.

We obtain the following utility guarantee.

Theorem 4.1. *Let $P = \{x_1, \dots, x_n\} \subseteq \mathbb{R}^d$ be (κ,τ)-good and suppose that $\|x_i\|_2 \le R$ for every $i \in [n]$. Let $\gamma, \kappa, \beta \in (0,1)$ and $\rho > 0$, and set $T = \left\lceil \frac{2\log(1/\kappa)}{\gamma} \right\rceil$, $\eta = \frac{\gamma}{8}$. Suppose*

$$n \ge \frac{32dR^2\left(\sqrt{d} + \sqrt{2\log((T+1)/\beta)}\right)}{\kappa\gamma\tau\sqrt{2\rho/(T+1)}}$$

Then, with probability at least $1-\beta$, Algorithm 2 does not return $\perp$ and outputs a positive definite matrix $\widehat{M}$, such that by defining the ellipsoid

$$\widehat{E}_0 := \left\{x \in \mathbb{R}^d : x^\top \widehat{M}^{-1} x \le 1\right\}$$

and its scaling $\widehat{E} := \sqrt{1-\eta}\, e^{-\gamma/2}\widehat{E}_0$ we get that there exists $S \subseteq [n]$ with $|S| \ge (1-\kappa)n$ such that

$$\widehat{E} \subseteq K_S, \qquad K \subseteq \sqrt{d(1+\eta)}\,\widehat{E}_0,$$

and moreover

$$\mathrm{vol}(\widehat{E}) \ge e^{-d\gamma/2}\left(\frac{1-\eta}{1+\eta}\right)^{d/2} \mathrm{vol}(E^*),$$

where E^* is the John ellipsoid of K.

The algorithm. As in the non-private algorithm, we maintain an unprojected measure w^t and use its κ-dense KL projection $\mu^t = \Pi_{\Gamma_\kappa}(w^t)$ to define the covariance matrix. The difference is that, at each iteration, we perturb the weighted covariance matrix $\Sigma(\mu^t) = \sum_{i=1}^n \mu_i^t x_i x_i^\top$ with symmetric Gaussian noise and use the inverse of the perturbed matrix in the multiplicative update. A final independent invocation of the private oracle is used to release the ellipsoid.

Privacy. For any $\mu \in \Gamma_\kappa$, changing one constraint changes the weighted covariance matrix by Frobenius norm at most $\frac{4dR^2}{\kappa n}$. Consequently, Algorithm 3, when invoked with a κ-dense measure μ, satisfies ρ_0-zCDP under the noise calibration above. The proof is given in Appendix D.

Algorithm 2 Private John Ellipsoid Approximation

Input: $P = \{x_i\}_{i=1}^n$, $\gamma, \kappa \in (0,1)$, $\rho > 0$, $\tau > 0$, $R > 0$.
$T \leftarrow \lceil 2\log(1/\kappa)/\gamma \rceil$
$\rho_0 \leftarrow \rho/(T+1)$
$w_i^0 \leftarrow d/n$ for all $i \in [n]$
for $t = 0, \ldots, T-1$ **do**
  $\mu^t \leftarrow \Pi_{\Gamma_\kappa}(w^t)$ ▷ Project onto the set of κ-dense measures
  $\widehat{M}^t \leftarrow$ PRIVATE-ORACLE$(P, \mu^t, \rho_0, \kappa, R)$
  if $\widehat{M}^t = \perp$ **then**
    return $\perp$
  for $i = 1, \ldots, n$ **do**
    $w_i^{t+1} \leftarrow w_i^t x_i^\top \widehat{M}^t x_i$
$\overline{\mu} \leftarrow \frac{1}{T}\sum_{t=0}^{T-1} \mu^t$
$\widehat{M} \leftarrow$ PRIVATE-ORACLE$(P, \overline{\mu}, \rho_0, \kappa, R)$
return $\widehat{M}$

Algorithm 3 Private Covariance Oracle

procedure PRIVATE-ORACLE$(P, \mu, \rho_0, \kappa, R)$
  $\Sigma \leftarrow \sum_{i=1}^n \mu_i x_i x_i^\top$
  $\sigma \leftarrow \frac{4dR^2}{\kappa n\sqrt{2\rho_0}}$
  Sample a symmetric Gaussian matrix N with scale σ
  $\widehat{\Sigma} \leftarrow \Sigma + N$
  if $\lambda_{\min}(\widehat{\Sigma}) < \tau/2$ **then** ▷ τ is a goodness assumption parameter
    return $\perp$
  return $\widehat{\Sigma}^{-1}$

Utility analysis. The first ingredient is that goodness prevents any κ-dense weighted covariance matrix from becoming nearly singular.

Lemma 4.2 ((κ,τ)-goodness implies uniform non-degeneracy)**.** If P is (κ,τ)-good, then

$$\lambda_{\min}(\Sigma(\mu)) \geq \tau \qquad \forall \mu \in \Gamma_\kappa$$

Therefore, whenever the Gaussian perturbation satisfies $\|N^t\|_2 \leq \eta\tau$, we have

$$(1-\eta)\Sigma^t \preceq \Sigma^t + N^t \preceq (1+\eta)\Sigma^t$$

and hence

$$\frac{1}{1+\eta}M^t \preceq \widehat{M}^t \preceq \frac{1}{1-\eta}M^t$$

In particular, each noisy logarithmic score differs from its exact counterpart by at most 2η.

The KL-potential argument of Section 3 is stable under these perturbations.

Lemma 4.3 (Perturbed potential inequality)**.** Suppose that $\|N^t\|_2 \leq \eta\tau$ for every $t = 0, \ldots, T-1$, where $\eta \leq 1/2$. Then, for every $q \in \Gamma_\kappa$,

$$\sum_{i=1}^n q_i \log\big(x_i^\top M(\overline{\mu}) x_i\big) \leq \frac{D_{\mathrm{KL}}(q\|\mu^0)}{T} + 4d\eta$$

Combining Lemma 4.3 with $D_{\mathrm{KL}}(q\|\mu^0) \leq d\log(1/\kappa)$, our choice $T \geq 2\log(1/\kappa)/\gamma$, and $\eta = \gamma/8$ gives

$$\sup_{q\in\Gamma_\kappa} \sum_{i=1}^n \frac{q_i}{d} \log\big(x_i^\top M(\overline{\mu}) x_i\big) \leq \gamma$$

Exactly as in the non-private analysis, this implies that at least $(1-\kappa)n$ constraints satisfy $x_i^\top M(\overline{\mu}) x_i \leq e^\gamma$, and therefore $e^{-\gamma/2}E(\overline{\mu}) \subseteq K_S$ for some $|S| \geq (1-\kappa)n$.

Finally, Gaussian matrix concentration and a union bound imply that the sample-size assumption of Theorem 4.1 guarantees $\|N^t\|_2 \le \eta\tau$ simultaneously for all $T+1$ oracle calls with probability at least $1-\beta$. Applying the same matrix-perturbation bound to the final oracle call transfers the containment and volume guarantees from $M(\overline{\mu})$ to the released matrix $\widehat{M}$.

The complete proofs, including the matrix perturbation, perturbed KL argument, Gaussian concentration bound, and final-release analysis, are given in Appendix D.

5 Extensions to Minimum-Volume Enclosing Ellipsoids

Our results extend directly to the centered minimum-volume enclosing ellipsoid (MVEE) problem through the standard polar-duality correspondence between maximum-volume inscribed and minimum-volume enclosing ellipsoids (Schneider, 2013; Todd, 2016). In particular, for $P = \{x_1, \dots, x_n\} \subseteq \mathbb{R}^d$, the polytope $K = \{x : |x_i^\top x| \le 1,\ \forall i \in [n]\}$ is the polar of $\mathrm{conv}\{\pm x_1, \dots, \pm x_n\}$, and the polar of the John ellipsoid of K is the centered MVEE of P. Namely, if $E^* = \{x : \ x^\top M^{-1} x \le 1\}$ is the JE of K, then its polar, $(E^*)^\circ = \{x : \ x^\top M x \le 1\}$, is the MVEE of P. This correspondence also preserves our trimmed formulation: if the John-ellipsoid guarantee holds for a subset $S \subseteq [n]$ with $|S| \ge (1-\kappa)n$, then the polar ellipsoid contains the corresponding $(1-\kappa)$-fraction of the MVEE input points, with the corresponding α^{-1} volume-approximation guarantee for the MVEE, where α denotes the volume-approximation factor for the John ellipsoid. We give the details of this reduction in Appendix A.2. We comment that since taking the polar is deterministic post-processing of the output of Algorithm 2, the reduction incurs no additional privacy loss.

Even though the above-mentioned result deals with *centered* convex bodies (since clearly the origin is the center of $\mathrm{conv}\{\pm x_1, \dots, \pm x_n\}$), it easily extends to the general uncentered MVEE problem using the standard homogeneous-coordinate lifting of Todd (2016). In order to approximate the MVEE of a convex body $\mathrm{conv}\{y_1, y_2, \dots, y_n\}$, We simply map each $y_i \in \mathbb{R}^d$ to $(y_i, 1) \in \mathbb{R}^{d+1}$ and reduce the uncentered problem to the centered problem in one additional dimension. While this reduction is standard, we verify in Lemma F.1 that it continues to hold under our trimmed objective: the same subset of at least $(1-\kappa)n$ points is preserved, together with the approximation factor. Thus, combining the lifting with the polar reduction above extends both our non-private and private guarantees to uncentered MVEEs. We comment that in order to obtain a private guarantee, the *lifted* dataset must satisfy the (κ, τ)-goodness assumption in dimension $d+1$. The full trimmed lifting argument is given in Appendix F.

6 Conclusion

We developed a projected multiplicative-weights framework for approximating the John ellipsoid under a trimmed feasibility criterion. By restricting the dual weights to κ-dense measures, our non-private algorithm produces an ellipsoid satisfying at least a $(1-\kappa)$-fraction of the constraints after an $e^{-\gamma}$ rescaling, while approximately preserving the optimal volume in $O(\log(1/\kappa)/\gamma)$ iterations. We also analyzed a Gaussian-perturbed variant and showed that, under the (κ, τ)-goodness assumption and sufficiently many points, the perturbation required for privacy preservation gives analogous utility guarantees. Through polar duality and homogeneous-coordinate lifting, our results extend to trimmed centered and uncentered minimum-volume enclosing ellipsoids.

Our analysis highlights the role of dense measures and robust non-degeneracy in controlling the effect of perturbations on geometric optimization. An important direction for future work is to determine whether comparable guarantees can be obtained under weaker assumptions than (κ, τ)-goodness, which is sufficient but not necessary for a successful execution of the algorithm. Moreover, while our analysis treats κ and τ as fixed parameters, our population characterization in Appendix E shows that, for any fixed κ, the goodness parameter τ can be related to the lower quantiles of the underlying distribution. Developing a privacy-preserving, data-dependent procedure for selecting κ and estimating an appropriate τ, while maintaining the utility guarantees of our algorithm, remains an interesting direction for future work. More broadly, it would be interesting to investigate the applicability of projected multiplicative-weights methods to other robust geometric optimization problems.

AI Disclosure Statement

In this work, we used generative AI tools (ChatGPT by OpenAI) to assist in the formulation and refinement of mathematical claims, writing of proofs, and the discussion and refinement of the theoretical framework and research methodology. Additionally, we used generative AI tools to assist with drafting and editing portions of the manuscript, improving readability and clarity, organizing the presentation of theoretical results, and suggesting titles and keywords.

All AI-assisted mathematical statements, arguments, and proofs were reviewed and verified by the authors. The authors also reviewed and revised AI-assisted text to ensure its accuracy, consistency with the research contributions, and appropriate attribution of existing work. The remaining tasks requiring disclosure were either not applicable to this work or were performed without the assistance of generative AI tools.

The authors take full responsibility for the final content of this paper, including all mathematical claims, proofs, and text produced with the assistance of generative AI.

References

Zeyuan Allen-Zhu, Yuanzhi Li, Aarti Singh, and Yining Wang. Near-optimal design of experiments via regret minimization. In Doina Precup and Yee Whye Teh (eds.), *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pp. 126–135. PMLR, 06–11 Aug 2017. URL https://proceedings.mlr.press/v70/allen-zhu17e.html.

Corwin L. Atwood. Optimal and efficient designs of experiments. *The Annals of Mathematical Statistics*, 40(5):1570–1602, 1969. ISSN 00034851, 21688990. URL http://www.jstor.org/stable/2239549.

Boaz Barak, Moritz Hardt, and Satyen Kale. The uniform hardcore lemma via approximate bregman projections. In *Proceedings of the twentieth annual ACM-SIAM symposium on Discrete algorithms*, pp. 1193–1200. SIAM, 2009.

Lev M Bregman. The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. *USSR computational mathematics and mathematical physics*, 7(3):200–217, 1967.

Sébastien Bubeck, Nicolò Cesa-Bianchi, and Sham Kakade. Towards minimax policies for online linear optimization with bandit feedback. *Journal of Machine Learning Research*, 23, 02 2012.

Mark Bun and Thomas Steinke. Concentrated differential privacy: Simplifications, extensions, and lower bounds. In *Theory of cryptography conference*, pp. 635–658. Springer, 2016.

Mark Bun, Marco Leandro Carmosino, and Jessica Sorrell. Efficient, noise-tolerant, and private learning via boosting, 2020. URL https://arxiv.org/abs/2002.01100.

Yang Cao, Xiaoyu Li, Zhao Song, Xin Yang, and Tianyi Zhou. Faster algorithms for structured john ellipsoid computation. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2026. URL https://openreview.net/forum?id=7RErGDbO6K.

Yuansi Chen, Raaz Dwivedi, Martin J. Wainwright, and Bin Yu. Fast mcmc sampling algorithms on polytopes. *J. Mach. Learn. Res.*, 19(1):2146–2231, January 2018. ISSN 1532-4435.

Michael B. Cohen, Ben Cousins, Yin Tat Lee, and Xin Yang. A near-optimal algorithm for approximating the john ellipsoid. In Alina Beygelzimer and Daniel Hsu (eds.), *Proceedings of the Thirty-Second Conference on Learning Theory*, volume 99 of *Proceedings of Machine Learning Research*, pp. 849–873. PMLR, 25–28 Jun 2019. URL https://proceedings.mlr.press/v99/cohen19a.html.

Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In *Theory of cryptography conference*, pp. 265–284. Springer, 2006.

Badih Ghazi, Ravi Kumar, and Pasin Manurangsi. Differentially private clustering: tight approximation ratios. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*, NIPS '20, Red Hook, NY, USA, 2020. Curran Associates Inc. ISBN 9781713829546.

Jiuxiang Gu, Xiaoyu Li, Yingyu Liang, Zhenmei Shi, Zhao Song, and Junwei Yu. Fast john ellipsoid computation with differential privacy optimization. In *CPAL*, 2024. URL https://api.semanticscholar.org/CorpusID:271860263.

David Gutman and Javier Pena. A unified framework for bregman proximal methods: subgradient, gradient, and accelerated gradient schemes, 12 2018.

Elad Hazan and Zohar Karnin. Volumetric spanners: An efficient exploration basis for learning. *Journal of Machine Learning Research*, 17(119):1–34, 2016. URL http://jmlr.org/papers/v17/hazan16a.html.

Justin Hsu, Aaron Roth, Tim Roughgarden, and Jonathan Ullman. Privately solving linear programs. In Javier Esparza, Pierre Fraigniaud, Thore Husfeldt, and Elias Koutsoupias (eds.), *Automata, Languages, and Programming*, pp. 612–624, Berlin, Heidelberg, 2014. Springer Berlin Heidelberg. ISBN 978-3-662-43948-7.

Fritz John. Extremum problems with inequalities as subsidiary conditions. In *Studies and Essays Presented to R. Courant on his 60th Birthday, January 8, 1948*, pp. 187–204. Interscience Publishers, Inc., New York, N. Y., 1948.

Leonid Khachiyan and Michael Todd. On the complexity of approximating the maximal inscribed ellipsoid for a polytope. *Math. Program.*, 61:137–159, 08 1993. doi: 10.1007/BF01582144.

J. Kiefer and J. Wolfowitz. The equivalence of two extremum problems. *Canadian Journal of Mathematics*, 12:363–366, 1960. doi: 10.4153/CJM-1960-030-4.

Piyush Kumar and E. Alper Yildirim. Minimum-volume enclosing ellipsoids and core sets. *Journal of Optimization Theory and Applications*, 126(1):1–21, 2005.

Yin Tat Lee and Aaron Sidford. Path finding methods for linear programming: Solving linear programs in $\tilde{O}$(vrank) iterations and faster algorithms for maximum flow. In *2014 IEEE 55th Annual Symposium on Foundations of Computer Science*, pp. 424–433, 2014. doi: 10.1109/FOCS.2014.52.

Xiaoyu Li, Zhao Song, and Junwei Yu. Quantum speedups for approximating the john ellipsoid, 2024. URL https://arxiv.org/abs/2408.14018.

Haihao Lu, Robert M. Freund, and Yurii Nesterov. Relatively smooth convex optimization by first-order methods, and applications. *SIAM Journal on Optimization*, 28(1):333–354, 2018. doi: 10.1137/16M1099546. URL https://doi.org/10.1137/16M1099546.

Bar Mahpud and Or Sheffet. A differentially private linear-time fptas for the minimum enclosing ball problem. In *Advances in Neural Information Processing Systems 35: Annual Conference on Neural Information Processing Systems 2022, NeurIPS 2022*, 2022. URL https://proceedings.neurips.cc/paper_files/paper/2022/file/cd830afc6208a346e4ec5caf1b08b4b4-Paper-Conference.pdf.

Yurii Nesterov and Arkadii Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*, volume 13 of *SIAM Studies in Applied and Numerical Mathematics*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 1994. doi: 10.1137/1.9781611970791.

Aleksandar Nikolov, Kunal Talwar, and Li Zhang. The geometry of differential privacy: the sparse and approximate cases. In *Proceedings of the Forty-Fifth Annual ACM Symposium on Theory of Computing*, STOC '13, pp. 351–360, New York, NY, USA, 2013. Association for Computing Machinery. ISBN 9781450320290. doi: 10.1145/2488608.2488652. URL https://doi.org/10.1145/2488608.2488652.

Kobbi Nissim and Uri Stemmer. Clustering algorithms for the centralized and local models. In Firdaus Janoos, Mehryar Mohri, and Karthik Sridharan (eds.), *Proceedings of Algorithmic Learning Theory*, volume 83 of *Proceedings of Machine Learning Research*, pp. 619–653. PMLR, 07–09 Apr 2018. URL https://proceedings.mlr.press/v83/nissim18a.html.

Kobbi Nissim, Uri Stemmer, and Salil Vadhan. Locating a small cluster privately. In *Proceedings of the 35th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems*, PODS '16, pp. 413–427, New York, NY, USA, 2016. Association for Computing Machinery. ISBN 9781450341912. doi: 10.1145/2902251.2902296. URL https://doi.org/10.1145/2902251.2902296.

Rolf Schneider. *Convex Bodies: The Brunn–Minkowski Theory*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2 edition, 2013.

Michael J. Todd. *Minimum-Volume Ellipsoids*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2016. doi: 10.1137/1.9781611974386. URL https://epubs.siam.org/doi/abs/10.1137/1.9781611974386.

Michael J. Todd and E. Alper Yıldırım. On khachiyan's algorithm for the computation of minimum-volume enclosing ellipsoids. *Discrete Applied Mathematics*, 155(13):1731–1744, 2007. ISSN 0166-218X. doi: https://doi.org/10.1016/j.dam.2007.02.013. URL https://www.sciencedirect.com/science/article/pii/S0166218X07000716.

V. N. Vapnik and A. Ya. Chervonenkis. *On the Uniform Convergence of Relative Frequencies of Events to Their Probabilities*, pp. 11–30. Springer International Publishing, Cham, 2015. ISBN 978-3-319-21852-6. doi: 10.1007/978-3-319-21852-6_3. URL https://doi.org/10.1007/978-3-319-21852-6_3.

Santosh Vempala and Mathematical Sciences Research Institute. *Geometric Random Walks: A Survey*, pp. 577–616. Mathematical Sciences Research Institute Publications. Cambridge University Press, 2005.

Yining Wang, Adams Wei Yu, and Aarti Singh. On computationally tractable selection of experiments in measurement-constrained regression models. *Journal of Machine Learning Research*, 18(143):1–41, 2017. URL http://jmlr.org/papers/v18/17-175.html.

David P. Woodruff and Taisuke Yasuda. John ellipsoids via lazy updates. In A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang (eds.), *Advances in Neural Information Processing Systems*, volume 37, pp. 68239–68257. Curran Associates, Inc., 2024. doi: 10.52202/079017-2179. URL https://proceedings.neurips.cc/paper_files/paper/2024/file/7e1854dc0e983210dc12feb6c2016b98-Paper-Conference.pdf.

A Additional Background

In this section, we recall the geometric, optimization, and privacy tools used in the analysis of our algorithms.

A.1 John Ellipsoid: Primal and Dual Formulations

Primal formulation. Let $P = \{x_1, \ldots, x_n\} \subseteq \mathbb{R}^d$ satisfy $\mathrm{span}(P) = \mathbb{R}^d$. Recall that we consider the centrally symmetric polytope

$$K_P := \left\{x \in \mathbb{R}^d : |x_i^\top x| \leq 1,\ \forall i \in [n]\right\}$$

By symmetry, the John ellipsoid of K_P is centered at the origin and can be represented as

$$E(G) := \left\{x \in \mathbb{R}^d : x^\top G^{-2} x \leq 1\right\}$$

where $G \succ 0$. Its volume is

$$\mathrm{vol}(E(G)) = \mathrm{vol}(B_2^d)\det(G)$$

where B_2^d denotes the Euclidean unit ball.

The containment condition $E(G) \subseteq K_P$ is equivalent to

$$\max_{x \in E(G)} |x_i^\top x| \leq 1 \qquad \forall i \in [n]$$

Writing $x = Gy$ with $\|y\|_2 \leq 1$, we obtain

$$\max_{x \in E(G)} |x_i^\top x| = \max_{\|y\|_2 \leq 1} |x_i^\top G y| = \|G x_i\|_2$$

Consequently, the John ellipsoid is obtained by solving

$$\begin{aligned} \text{maximize} \quad & \log\det(G^2) \\ \text{subject to} \quad & G \succ 0, \\ & \|G x_i\|_2^2 \leq 1, \qquad \forall i \in [n] \end{aligned} \tag{P}$$

Equivalently, defining $B := G^2$, we obtain the convex optimization problem

$$\begin{aligned} \text{maximize} \quad & \log\det B \\ \text{subject to} \quad & B \succ 0, \\ & x_i^\top B x_i \leq 1, \qquad \forall i \in [n] \end{aligned} \tag{P'}$$

The corresponding ellipsoid is

$$E(B) := \{x \in \mathbb{R}^d : x^\top B^{-1} x \leq 1\}$$

Dual formulation. Introduce nonnegative Lagrange multipliers $\mu_1, \ldots, \mu_n$ for the constraints $x_i^\top B x_i \leq 1$. The Lagrangian of (P') is

$$\begin{aligned} L(B, \mu) &= \log\det B + \sum_{i=1}^n \mu_i (1 - x_i^\top B x_i) \\ &= \log\det B + \sum_{i=1}^n \mu_i - \operatorname{tr}(B \Sigma(\mu)) \end{aligned}$$

where

$$\Sigma(\mu) := \sum_{i=1}^n \mu_i x_i x_i^\top.$$

For $\Sigma(\mu) \succ 0$, maximizing the Lagrangian over $B \succ 0$ gives the stationarity condition $B^{-1} - \Sigma(\mu) = 0$ and therefore $B = \Sigma(\mu)^{-1}$. Substituting into the Lagrangian yields

$$\sup_{B \succ 0} L(B, \mu) = -\log\det \Sigma(\mu) + \sum_{i=1}^n \mu_i - d$$

Thus, the dual problem is

$$\begin{aligned} \text{minimize} \quad & D(\mu) := \sum_{i=1}^n \mu_i - \log\det \Sigma(\mu) - d \\ \text{subject to} \quad & \mu_i \geq 0, \qquad \forall i \in [n] \\ & \Sigma(\mu) \succ 0 \end{aligned} \tag{D}$$

We adopt the convention that $D(\mu) = +\infty$ when $\Sigma(\mu)$ is singular.

Since the primal problem satisfies Slater's condition, strong duality holds:

$$\log\det B^* = D(\mu^*)$$

where B^* and μ^* denote optimal primal and dual solutions, respectively.

John decomposition. The optimality conditions imply $B^* = \Sigma(\mu^*)^{-1}$, and consequently $(B^*)^{-1} = \sum_{i=1}^n \mu_i^* x_i x_i^\top$. This representation is commonly referred to as the John decomposition.

Furthermore, complementary slackness gives

$$\mu_i^* \left(x_i^\top B^* x_i - 1\right) = 0 \qquad \forall i \in [n]$$

Taking the trace of $B^*\Sigma(\mu^*) = I_d$, we obtain

$$\begin{aligned} d &= \operatorname{tr}(B^*\Sigma(\mu^*)) \\ &= \sum_{i=1}^n \mu_i^* x_i^\top B^* x_i \\ &= \sum_{i=1}^n \mu_i^* \end{aligned}$$

Hence the optimal dual measure has total mass d, and its objective simplifies to

$$D(\mu^*) = -\log\det\Sigma(\mu^*)$$

Geometric interpretation of trimmed solutions. Recall that a nonnegative measure μ of total mass d is a $(1+\gamma)$-approximate κ-trimmed John solution if

$$\left|\left\{i \in [n] : x_i^\top M(\mu) x_i \le e^\gamma\right\}\right| \ge (1-\kappa)n$$

where $M(\mu) := \Sigma(\mu)^{-1}$.

The following lemma relates this condition to the geometric guarantees of the corresponding ellipsoid.

Lemma A.1 ($(1+\gamma)$-approximate trimmed solution is a good trimmed rounding)**.** Let μ be a $(1+\gamma)$-approximate κ-trimmed John solution, and define

$$E := \left\{x \in \mathbb{R}^d : x^\top M(\mu)^{-1} x \le 1\right\}$$

Let

$$S := \left\{i \in [n] : x_i^\top M(\mu) x_i \le e^\gamma\right\}$$

and

$$K_S := \left\{x \in \mathbb{R}^d : |x_i^\top x| \le 1,\ \forall i \in S\right\}$$

Then $|S| \ge (1-\kappa)n$ and

$$e^{-\gamma/2} E \subseteq K_S$$

Moreover, for the original polytope,

$$K_P \subseteq \sqrt{d}\, E$$

Proof. The bound $|S| \ge (1-\kappa)n$ follows directly from the definition of an approximate trimmed John solution.

To establish trimmed containment, let $x \in e^{-\gamma/2}E$. By definition, $x^\top M(\mu)^{-1} x \le e^{-\gamma}$. For every $i \in S$, we also have $x_i^\top M(\mu) x_i \le e^\gamma$. Applying Cauchy-Schwarz gives

$$\begin{aligned} |x_i^\top x| &= \left|\left\langle M(\mu)^{1/2} x_i, M(\mu)^{-1/2} x \right\rangle\right| \\ &\le \sqrt{x_i^\top M(\mu) x_i}\sqrt{x^\top M(\mu)^{-1} x} \\ &\le e^{\gamma/2} e^{-\gamma/2} = 1 \end{aligned}$$

Therefore, $e^{-\gamma/2}E \subseteq K_S$.

Next, let $x \in K_P$. Then $|x_i^\top x| \le 1$ for every $i \in [n]$. Since $M(\mu)^{-1} = \Sigma(\mu)$,

$$x^\top M(\mu)^{-1} x = x^\top \left(\sum_{i=1}^n \mu_i x_i x_i^\top\right) x = \sum_{i=1}^n \mu_i |x_i^\top x|^2 \le \sum_{i=1}^n \mu_i = d$$

Thus $x \in \sqrt{d}\, E$, proving $K_P \subseteq \sqrt{d}\, E$. □

Volume interpretation via dual feasibility. The dual formulation provides a lower bound on the volume of the ellipsoid associated with any normalized dual measure.

Lemma A.2 (Volume interpretation via dual feasibility)**.** Let $\mu \in \mathbb{R}^n_{\geq 0}$ satisfy

$$\sum_{i=1}^{n} \mu_i = d, \qquad \Sigma(\mu) \succ 0$$

and define

$$E := \left\{ x \in \mathbb{R}^d : x^\top M(\mu)^{-1} x \leq 1 \right\}$$

Let E^* be the John ellipsoid of K_P. For every $\gamma \geq 0$, the scaled ellipsoid $E_\gamma := e^{-\gamma/2} E$ satisfies

$$\mathrm{vol}(E_\gamma) \geq e^{-d\gamma/2} \, \mathrm{vol}(E^*)$$

Proof. Let μ^* denote an optimal dual solution. By weak duality, $D(\mu) \geq D(\mu^*)$. Since both measures have total mass d, the dual objectives simplify to

$$D(\mu) = -\log\det \Sigma(\mu)$$

and

$$D(\mu^*) = -\log\det \Sigma(\mu^*)$$

Consequently,

$$\det \Sigma(\mu) \leq \det \Sigma(\mu^*)$$

The volume of E is

$$\mathrm{vol}(E) = \mathrm{vol}(B_2^d) \det(M(\mu))^{1/2} = \frac{\mathrm{vol}(B_2^d)}{\sqrt{\det \Sigma(\mu)}}$$

Similarly, by strong duality and the John decomposition,

$$\mathrm{vol}(E^*) = \frac{\mathrm{vol}(B_2^d)}{\sqrt{\det \Sigma(\mu^*)}}.$$

Hence $\mathrm{vol}(E) \geq \mathrm{vol}(E^*)$.

Finally, scaling a d-dimensional ellipsoid by $e^{-\gamma/2}$ multiplies its volume by $e^{-d\gamma/2}$. Therefore,

$$\begin{aligned} \mathrm{vol}(E_\gamma) &= e^{-d\gamma/2} \, \mathrm{vol}(E) \\ &\geq e^{-d\gamma/2} \, \mathrm{vol}(E^*) \end{aligned}$$

□

The preceding volume comparison does not by itself imply that E_γ is contained in the original polytope. The corresponding trimmed-containment guarantee follows separately from Lemma A.1.

A.2 Polar Interpretation: MVEE and MVIE

We recall the relationship between polar bodies, minimum-volume enclosing ellipsoids (MVEE), and maximum-volume inscribed ellipsoids (MVIE). For the centrally symmetric polytopes considered in this work, the MVIE coincides with the John ellipsoid.

Definition A.3 (Polar body (Schneider, 2013))**.** Let $K \subseteq \mathbb{R}^d$ be a convex body containing the origin. Its polar body is

$$K^\circ := \left\{ y \in \mathbb{R}^d : \langle x, y \rangle \leq 1 \text{ for every } x \in K \right\}$$

When K is centrally symmetric, this is equivalently

$$K^\circ = \left\{ y \in \mathbb{R}^d : |\langle x, y \rangle| \leq 1 \text{ for every } x \in K \right\}$$

Fact A.4 (Basic properties of polarity (Schneider, 2013))**.** Let K_1, K_2 be convex bodies containing the origin. Then:

1. *Containment reversal:*
$$K_1 \subseteq K_2 \quad \Longrightarrow \quad K_2^\circ \subseteq K_1^\circ$$
2. *Bipolar identity:* If K is closed and convex and contains the origin in its interior, then
$$(K^\circ)^\circ = K$$
3. *Scaling:* For every $\alpha > 0$,
$$(\alpha K)^\circ = \alpha^{-1} K^\circ$$

Fact A.5 (Polarity of centered ellipsoids (Todd, 2016))**.** Let $M \succ 0$ and define
$$E(M) := \left\{x \in \mathbb{R}^d : x^\top M^{-1} x \leq 1\right\}$$
Then
$$E(M)^\circ = \left\{y \in \mathbb{R}^d : y^\top M y \leq 1\right\}$$
In particular, the polar of an origin-centered ellipsoid is again an origin-centered ellipsoid.

John ellipsoid and centered MVEE. Given $P = \{x_1, \ldots, x_n\}$, define the centrally symmetric convex hull
$$C_P := \mathrm{conv}\{\pm x_1, \ldots, \pm x_n\}$$
By the definition of polarity,
$$K_P = C_P^\circ, \qquad K_P^\circ = C_P$$
Let E^* be the John ellipsoid of K_P. Since $E^* \subseteq K_P$, containment reversal gives
$$C_P = K_P^\circ \subseteq (E^*)^\circ$$
Thus $(E^*)^\circ$ is an origin-centered ellipsoid enclosing all the points of P.

Conversely, let C be any origin-centered ellipsoid satisfying $C_P \subseteq C$. Taking polars gives
$$C^\circ \subseteq C_P^\circ = K_P$$
Therefore, C° is an admissible centered inscribed ellipsoid of K_P.

For an origin-centered ellipsoid $E(M)$, we have
$$\mathrm{vol}(E(M)) = \mathrm{vol}(B_2^d) \det(M)^{1/2}$$
and
$$\mathrm{vol}(E(M)^\circ) = \mathrm{vol}(B_2^d) \det(M)^{-1/2}$$
Consequently,
$$\mathrm{vol}(E)\,\mathrm{vol}(E^\circ) = \mathrm{vol}(B_2^d)^2$$

Maximizing the volume of an origin-centered ellipsoid inscribed in K_P is therefore equivalent to minimizing the volume of an origin-centered ellipsoid enclosing C_P.

It follows that $C^* := (E^*)^\circ$ is the centered minimum-volume enclosing ellipsoid of C_P, and equivalently the centered MVEE of P.

Trimmed containment under polarity. Let $S \subseteq [n]$ and define
$$K_S := \left\{x : |x_i^\top x| \leq 1 \quad \forall i \in S\right\}$$
Suppose that an origin-centered ellipsoid $\widehat{E}$ satisfies $\widehat{E} \subseteq K_S$. Taking polars yields
$$\mathrm{conv}\{\pm x_i : i \in S\} = K_S^\circ \subseteq \widehat{E}^\circ$$
In particular,
$$x_i \in \widehat{E}^\circ \qquad \forall i \in S$$
Thus, if $|S| \geq (1-\kappa)n$, the polar ellipsoid encloses at least a $(1-\kappa)$-fraction of the input points.

Moreover, if $\mathrm{vol}(\widehat{E}) \geq \alpha\,\mathrm{vol}(E^*)$, then
$$\frac{\mathrm{vol}(\widehat{E}^\circ)}{\mathrm{vol}(C^*)} = \frac{\mathrm{vol}(E^*)}{\mathrm{vol}(\widehat{E})} \leq \alpha^{-1}$$
Hence, the trimmed-containment and volume guarantees for the John ellipsoid translate directly into corresponding guarantees for the centered MVEE problem.

The extension to general uncentered MVEEs through homogeneous-coordinate lifting is discussed separately in Section F.

A.3 Additional Differential Privacy and Concentration Facts

We recall the standard privacy definitions, composition properties, and Gaussian concentration inequalities used in the analysis of our private algorithm.

Differential privacy.

Definition A.6 (Differential Privacy (Dwork et al., 2006)). A randomized mechanism $\mathcal{M} : \mathcal{X}^n \to \mathcal{Y}$ is (ε, δ)-differentially private if, for every pair of neighboring datasets $P, P' \in \mathcal{X}^n$ and every measurable set $S \subseteq \mathcal{Y}$,

$$\Pr[\mathcal{M}(P) \in S] \leq e^{\varepsilon} \Pr[\mathcal{M}(P') \in S] + \delta$$

When $\delta = 0$, the mechanism is called ε-differentially private.

Definition A.7 (Zero-Concentrated Differential Privacy (zCDP) (Bun & Steinke, 2016)). A randomized mechanism $\mathcal{M}$ satisfies ρ-zCDP if, for every pair of neighboring datasets P, P' and every $\alpha > 1$,

$$D_\alpha\big(\mathcal{M}(P) \| \mathcal{M}(P')\big) \leq \alpha\rho$$

where D_α is the Renyi divergence of order α.

The following standard properties of zCDP are used throughout the paper.

Fact A.8 (Composition and post-processing). Suppose that $\mathcal{M}_t$ satisfies ρ_t-zCDP for $t = 1, \ldots, T$. If each mechanism satisfies its claimed privacy guarantee conditionally on the previously released outputs, then their adaptive composition satisfies $\left(\sum_{t=1}^{T} \rho_t\right)$-zCDP.

Moreover, arbitrary randomized post-processing of the output of a ρ-zCDP mechanism preserves the same privacy guarantee.

Fact A.9 (Conversion to approximate DP). A ρ-zCDP mechanism also satisfies (ε, δ)-DP for every $\delta \in (0, 1)$, where $\varepsilon = \rho + 2\sqrt{\rho \log(1/\delta)}$.

Gaussian mechanism. Let $f : \mathcal{X}^n \to \mathbb{R}^m$ have global ℓ_2-sensitivity

$$\Delta_2(f) := \sup_{P \sim P'} \|f(P) - f(P')\|_2$$

Fact A.10 (Gaussian mechanism (Bun & Steinke, 2016)). Let $Z \sim \mathcal{N}(0, \sigma^2 I_m)$. The mechanism

$$\mathcal{M}(P) := f(P) + Z$$

satisfies ρ-zCDP whenever

$$\sigma^2 \geq \frac{\Delta_2(f)^2}{2\rho}$$

Gaussian perturbation of symmetric matrices. We identify the vector space of real symmetric $d \times d$ matrices with $\mathbb{R}^{d(d+1)/2}$ through an orthonormal basis under the Frobenius inner product

$$\langle A, B \rangle_F := \operatorname{tr}(A^\top B)$$

Let $F_{ii} := e_i e_i^\top$, and, for $i < j$, let $F_{ij} := \frac{e_i e_j^\top + e_j e_i^\top}{\sqrt{2}}$. The collection $\{F_{ii}\}_{i=1}^d \cup \{F_{ij}\}_{i<j}$ is an orthonormal basis of the real symmetric matrices.

A symmetric Gaussian perturbation with scale σ can be written as

$$N = \sum_{i=1}^{d} z_{ii} F_{ii} + \sum_{1 \leq i < j \leq d} z_{ij} F_{ij}$$

where the coefficients z_{ij} are independent $\mathcal{N}(0, \sigma^2)$ random variables.

Equivalently, its diagonal entries have variance σ^2, and its off-diagonal entries have variance $\sigma^2/2$.

Under this normalization, perturbing a symmetric matrix-valued query with N is equivalent to applying the standard Gaussian mechanism to its coordinates in an orthonormal basis.

In particular, a symmetric matrix-valued query with global Frobenius sensitivity Δ_F satisfies ρ-zCDP when perturbed using $\sigma^2 \geq \frac{\Delta_F^2}{2\rho}$.

Sensitivity of a weighted covariance matrix. Consider the weighted covariance query

$$\Sigma(P, \mu) := \sum_{i=1}^{n} \mu_i x_i x_i^\top$$

where $\mu \in \Gamma_\kappa$ is fixed and $\|x_i\|_2 \le R$ for every i.

Let P and P' differ only in their j-th constraint, where x_j is replaced by x'_j. Then

$$\begin{aligned}
&\|\Sigma(P, \mu) - \Sigma(P', \mu)\|_F \\
&\quad = \mu_j \|x_j x_j^\top - x'_j x_j'^\top\|_F \\
&\quad \le \mu_j \left(\|x_j x_j^\top\|_F + \|x'_j x_j'^\top\|_F\right) \\
&\quad \le 2\mu_j R^2 \le \frac{2dR^2}{\kappa n}
\end{aligned}$$

Thus, for a fixed measure μ, the weighted covariance query has global Frobenius sensitivity at most $\Delta_F \le \frac{2dR^2}{\kappa n}$.

Gaussian norm concentration.

Fact A.11 (Gaussian vector concentration)**.** Let $Z \sim \mathcal{N}(0, \sigma^2 I_d)$. Then, for every $u > 0$,

$$\Pr\left[\|Z\|_2 \ge \sigma\left(\sqrt{d} + \sqrt{2u}\right)\right] \le e^{-u}$$

Equivalently, for every $\beta \in (0, 1)$,

$$\Pr\left[\|Z\|_2 \le \sigma\left(\sqrt{d} + \sqrt{2\log(1/\beta)}\right)\right] \ge 1 - \beta$$

This bound follows from the concentration of the Euclidean norm of a standard Gaussian vector, whose squared norm follows a χ^2_d distribution.

For the symmetric Gaussian matrix described above, a corresponding spectral-norm concentration bound takes the form

$$\Pr\left[\|N\|_2 > C\sigma(\sqrt{d} + \sqrt{u})\right] \le e^{-u}$$

where $C > 0$ is an absolute constant.

The precise constant depends on the normalization of the symmetric Gaussian matrix distribution.

In particular, for independent symmetric Gaussian perturbations $N^0, \ldots, N^T$, a union bound gives

$$\Pr\left[\max_{0 \le t \le T} \|N^t\|_2 \le C\sigma\left(\sqrt{d} + \sqrt{\log\frac{T+1}{\beta}}\right)\right] \ge 1 - \beta$$

This concentration inequality is used in Appendix D to obtain the sample-size condition ensuring that the noisy covariance matrices remain well-conditioned.

A.4 Additional Facts on KL Projection and Leverage Scores

Measures and generalized KL divergence. A nonnegative measure over the finite domain $[n]$ is a vector $\mu \in \mathbb{R}^n_{\ge 0}$. Throughout the paper, the iterates belong to the set of measures of total mass d.

For nonnegative measures μ, ν, with $\nu_i > 0$ for all i, we define the generalized KL divergence by

$$D_{\mathrm{KL}}(\mu\|\nu) := \sum_{i=1}^{n} \left[\mu_i \log\left(\frac{\mu_i}{\nu_i}\right) + \nu_i - \mu_i\right]$$

with the convention $0 \log 0 = 0$.

When both measures have total mass d, the linear terms cancel, and

$$D_{\mathrm{KL}}(\mu\|\nu) = \sum_{i=1}^{n} \mu_i \log\left(\frac{\mu_i}{\nu_i}\right)$$

Recall that the set of κ-dense measures is

$$\Gamma_\kappa := \left\{ \mu \in \mathbb{R}^n_{\geq 0} : \sum_{i=1}^{n} \mu_i = d, \quad \mu_i \leq \frac{d}{\kappa n} \ \forall i \in [n] \right\}$$

Bregman projection. Let Γ be a nonempty closed convex set of nonnegative measures with total mass d. The KL projection of a strictly positive measure $\widehat{\mu}$ onto Γ is

$$\Pi_\Gamma(\widehat{\mu}) := \arg\min_{\mu \in \Gamma} D_{\mathrm{KL}}(\mu\|\widehat{\mu})$$

The following standard result gives the property of KL projection used in the analysis of the multiplicative updates.

Theorem A.12 (Bregman Pythagorean inequality (Bregman, 1967))**.** *Let Γ be a nonempty closed convex set of measures of total mass d, and let $\widehat{\mu}$ be a strictly positive measure. For every $q \in \Gamma$,*

$$D_{\mathrm{KL}}(q\|\Pi_\Gamma(\widehat{\mu})) + D_{\mathrm{KL}}(\Pi_\Gamma(\widehat{\mu})\|\widehat{\mu}) \leq D_{\mathrm{KL}}(q\|\widehat{\mu})$$

In particular,

$$D_{\mathrm{KL}}(q\|\Pi_\Gamma(\widehat{\mu})) \leq D_{\mathrm{KL}}(q\|\widehat{\mu})$$

Note that the inequality is (somewhat surprisingly) in the opposite direction to the standard triangle inequality.

Proof. Let $p := \Pi_\Gamma(\widehat{\mu})$. The first-order optimality condition for the convex minimization defining p implies that, for every $q \in \Gamma$,

$$\langle q - p, \nabla F(p) - \nabla F(\widehat{\mu}) \rangle \geq 0$$

where

$$F(\mu) := \sum_{i=1}^{n} (\mu_i \log \mu_i - \mu_i)$$

The three-point identity for the Bregman divergence generated by F gives

$$\begin{aligned} &D_{\mathrm{KL}}(q\|\widehat{\mu}) - D_{\mathrm{KL}}(q\|p) \\ &\quad - D_{\mathrm{KL}}(p\|\widehat{\mu}) \\ &= \langle q - p, \nabla F(p) - \nabla F(\widehat{\mu}) \rangle \geq 0 \end{aligned}$$

Rearranging proves the first inequality. The second follows from the nonnegativity of KL divergence. □

The explicit computation of Π_{Γ_κ} is established in the separate KL-projection appendix. In particular, for strictly positive $\widehat{\mu}$, it takes the form

$$\Pi_{\Gamma_\kappa}(\widehat{\mu})_i = \min\left\{ \frac{d}{\kappa n}, c\widehat{\mu}_i \right\}$$

where $c > 0$ is chosen to ensure that the projected measure has total mass d.

Weighted covariance and leverage scores. For a measure $\mu \in \mathbb{R}^n_{\geq 0}$ such that $\Sigma(\mu) \succ 0$, recall

$$\Sigma(\mu) := \sum_{j=1}^{n} \mu_j x_j x_j^\top, \qquad M(\mu) := \Sigma(\mu)^{-1}$$

The corresponding quadratic and logarithmic scores are

$$\psi_i(\mu) := x_i^\top M(\mu) x_i, \qquad \phi_i(\mu) := \log \psi_i(\mu)$$

The logarithmic scores are considered for nonzero constraint vectors, so that $\psi_i(\mu) > 0$.

The following convexity property is central to our multiplicative-weights analysis.

Lemma A.13 (Convexity (Cohen et al., 2019))**.** For every nonzero x_i, the function

$$\phi_i(\mu) = \log \left(x_i^\top \left(\sum_{j=1}^{n} \mu_j x_j x_j^\top \right)^{-1} x_i \right)$$

is convex on the domain where $\Sigma(\mu) \succ 0$.

Basic properties of leverage scores. Define the weighted leverage score

$$\ell_i(\mu) := \mu_i \psi_i(\mu) = \mu_i x_i^\top \Sigma(\mu)^{-1} x_i$$

Lemma A.14 (Basic properties of weighted leverage scores)**.** For every $\mu \in \mathbb{R}^n_{\geq 0}$ such that $\Sigma(\mu) \succ 0$,

$$0 \leq \ell_i(\mu) \leq 1 \qquad \forall i \in [n]$$

and

$$\sum_{i=1}^{n} \ell_i(\mu) = d$$

Proof. Nonnegativity follows immediately from $\mu_i \geq 0$ and $\Sigma(\mu)^{-1} \succ 0$.

For the upper bound, observe that

$$\Sigma(\mu) - \mu_i x_i x_i^\top = \sum_{j \neq i} \mu_j x_j x_j^\top \succeq 0$$

Therefore, $\mu_i x_i x_i^\top \preceq \Sigma(\mu)$. Multiplying on both sides by $\Sigma(\mu)^{-1/2}$ gives

$$\mu_i \Sigma(\mu)^{-1/2} x_i x_i^\top \Sigma(\mu)^{-1/2} \preceq I_d$$

The matrix on the left has at most one nonzero eigenvalue, equal to

$$\mu_i x_i^\top \Sigma(\mu)^{-1} x_i = \ell_i(\mu)$$

Hence $\ell_i(\mu) \leq 1$.

Finally,

$$\sum_{i=1}^{n} \ell_i(\mu) = \sum_{i=1}^{n} \mu_i x_i^\top \Sigma(\mu)^{-1} x_i = \operatorname{tr}\left(\Sigma(\mu)^{-1} \sum_{i=1}^{n} \mu_i x_i x_i^\top \right) = \operatorname{tr}(I_d) = d$$

□

In particular, the identity $\sum_{i=1}^{n} \mu_i \psi_i(\mu) = d$ will be used to control the potential change of the two-level multiplicative-weights update. Although the unprojected measure w^t need not have total mass d, its projection $\mu^t = \Pi_{\Gamma_\kappa}(w^t)$ does, and therefore $\sum_{i=1}^{n} \mu_i^t \psi_i(\mu^t) = d$ at every iteration.

B KL PROJECTION ONTO Γ_κ

Lemma B.1. Let $\widehat{\mu} \in \mathbb{R}^n_{>0}$ be an arbitrary strictly positive measure, and let

$$\Gamma_\kappa = \left\{ \mu \in \mathbb{R}^n_{\geq 0} : \sum_{i=1}^n \mu_i = d, \quad \mu_i \leq \frac{d}{\kappa n} \ \forall i \right\}$$

Define $u := \frac{d}{\kappa n}$.

Then the Bregman projection of $\widehat{\mu}$ onto Γ_κ with respect to the generalized KL divergence

$$\mathrm{KL}(\mu \| \widehat{\mu}) = \sum_{i=1}^n \left[\mu_i \log\left(\frac{\mu_i}{\widehat{\mu}_i}\right) + \widehat{\mu}_i - \mu_i \right]$$

is uniquely given by

$$\mu_i^\star = \min\{u,\, c\widehat{\mu}_i\}$$

where $c > 0$ is chosen so that $\sum_{i=1}^n \mu_i^\star = d$.

Proof. We consider the convex optimization problem

$$\min_{\mu \in \mathbb{R}^n_{\geq 0}} \mathrm{KL}(\mu \| \widehat{\mu})$$

subject to

$$\sum_{i=1}^n \mu_i = d \qquad \text{and} \qquad \mu_i \leq u \qquad \forall i \in [n]$$

Since $\widehat{\mu}_i > 0$ for every i, the objective is finite and strictly convex on $\mathbb{R}^n_{>0}$. Moreover, Γ_κ is a nonempty compact convex set. Hence the problem admits a unique minimizer.

Ignoring terms that do not depend on μ, the objective can be written as

$$f(\mu) = \sum_{i=1}^n \left[\mu_i \log\left(\frac{\mu_i}{\widehat{\mu}_i}\right) - \mu_i \right]$$

For each coordinate,

$$\frac{\partial f}{\partial \mu_i} = \log\left(\frac{\mu_i}{\widehat{\mu}_i}\right)$$

Introduce a Lagrange multiplier $\lambda \in \mathbb{R}$ for the equality constraint $\sum_i \mu_i = d$, and multipliers $\alpha_i \geq 0$ for the upper-bound constraints $\mu_i - u \leq 0$.

Since the optimum has positive coordinates when $\widehat{\mu}_i > 0$, we may omit the nonnegativity multipliers. The Lagrangian is therefore

$$L(\mu, \lambda, \alpha) = f(\mu) + \lambda \left(\sum_{i=1}^n \mu_i - d \right) + \sum_{i=1}^n \alpha_i (\mu_i - u)$$

The KKT stationarity condition gives

$$\log\left(\frac{\mu_i^\star}{\widehat{\mu}_i}\right) + \lambda + \alpha_i = 0$$

Thus

$$\mu_i^\star = \widehat{\mu}_i e^{-\lambda - \alpha_i}$$

We now distinguish two cases.

- Case 1: the coordinate is not capped. Suppose $\mu_i^\star < u$. By complementary slackness,

  $$\alpha_i(\mu_i^\star - u) = 0$$

  and hence $\alpha_i = 0$. Therefore, $\mu_i^\star = e^{-\lambda}\widehat{\mu}_i$.

  Now define $c := e^{-\lambda} > 0$. Then every uncapped coordinate satisfies

  $$\mu_i^\star = c\widehat{\mu}_i$$

- Case 2: the coordinate is capped. Suppose $\mu_i^\star = u$. The stationarity condition gives

  $$\log\left(\frac{u}{\widehat{\mu}_i}\right) + \lambda + \alpha_i = 0$$

  Since $\alpha_i \geq 0$,

  $$\log\left(\frac{u}{\widehat{\mu}_i}\right) + \lambda \leq 0$$

  Equivalently,

  $$u \leq e^{-\lambda}\widehat{\mu}_i = c\widehat{\mu}_i$$

  Thus a capped coordinate is precisely one for which $c\widehat{\mu}_i \geq u$.

Combining the two cases, we obtain

$$\mu_i^\star = \begin{cases} c\widehat{\mu}_i, & c\widehat{\mu}_i < u \\ u, & c\widehat{\mu}_i \geq u \end{cases}$$

Hence $\mu_i^\star = \min\{u, c\widehat{\mu}_i\}$.

The scalar c is determined by the equality constraint

$$\sum_{i=1}^{n} \mu_i^\star = d$$

Therefore it must satisfy

$$\sum_{i=1}^{n} \min\{u, c\widehat{\mu}_i\} = d$$

It remains to show that such a c exists and is unique.

Define

$$g(c) := \sum_{i=1}^{n} \min\{u, c\widehat{\mu}_i\}$$

Since every $\widehat{\mu}_i > 0$, the function g is continuous and nondecreasing. Moreover, $g(0) = 0$, while

$$\lim_{c\to\infty} g(c) = nu = \frac{d}{\kappa} \geq d$$

Hence, by continuity, there exists $c > 0$ such that $g(c) = d$.

If $\kappa < 1$, then $nu = \frac{d}{\kappa} > d$, and g is strictly increasing at every point for which $g(c) < nu$. Therefore the solution of $g(c) = d$ is unique.

Thus the unique KL projection is

$$\Pi_{\Gamma_\kappa}(\widehat{\mu})_i = \min\left\{\frac{d}{\kappa n}, c\widehat{\mu}_i\right\}$$

where c is chosen so that the projected measure has total mass d.

□

C Proofs for the Non-Private Algorithm

We provide the auxiliary results and the complete proof of Theorem 3.1.

C.1 KL-Divergence and Iteration Bounds

Lemma C.1. Let $\mu_i^0 = d/n$ for all $i \in [n]$. Then

$$\max_{q \in \Gamma_\kappa} D_{\mathrm{KL}}(q \| \mu^0) \le d \log(1/\kappa)$$

Proof. For every $q \in \Gamma_\kappa$, the normalization $\sum_i q_i = d$ gives

$$D_{\mathrm{KL}}(q \| \mu^0) = \sum_{i=1}^n q_i \log\left(\frac{q_i}{d/n}\right)$$

Since $q_i \le d/(\kappa n)$,

$$D_{\mathrm{KL}}(q \| \mu^0) \le \sum_{i=1}^n q_i \log(1/\kappa) = d \log(1/\kappa)$$

□

Corollary C.2. Let $\overline{\mu} = \frac{1}{T}\sum_{t=0}^{T-1} \mu^t$. If $T \ge \frac{\log(1/\kappa)}{\gamma}$, then, for every $q \in \Gamma_\kappa$,

$$\sum_{i=1}^n \frac{q_i}{d} \phi_i(\overline{\mu}) \le \gamma$$

Proof. By Lemmas 3.2 and C.1,

$$\sum_{i=1}^n \frac{q_i}{d} \phi_i(\overline{\mu}) \le \frac{D_{\mathrm{KL}}(q \| \mu^0)}{dT} \le \frac{d \log(1/\kappa)}{dT} \le \gamma$$

□

C.2 Trimmed Containment and Volume

Lemma C.3 (Trimmed containment guarantee)**.** Let μ be a measure of total mass d such that

$$\sum_{i=1}^n q_i \log\left(x_i^\top M(\mu) x_i\right) \le d\gamma \qquad \forall q \in \Gamma_\kappa$$

Then the set

$$S := \left\{ i \in [n] : x_i^\top M(\mu) x_i \le e^\gamma \right\}$$

satisfies $|S| \ge (1-\kappa)n$.

Proof. Suppose, towards a contradiction, that more than κn indices satisfy $x_i^\top M(\mu) x_i > e^\gamma$. Let B denote the set of these indices and define

$$q_i := \frac{d}{|B|} \mathbf{1}_{\{i \in B\}}$$

Since $|B| > \kappa n$, we have $q_i \le d/(\kappa n)$, and hence $q \in \Gamma_\kappa$.

However, for every $i \in B$, $\log(x_i^\top M(\mu) x_i) > \gamma$. Therefore,

$$\sum_{i=1}^n q_i \log\left(x_i^\top M(\mu) x_i\right) = \frac{d}{|B|} \sum_{i \in B} \log\left(x_i^\top M(\mu) x_i\right) > d\gamma$$

contradicting the assumption. Thus $|S| \ge (1-\kappa)n$. □

Corollary C.4 (Approximate κ-trimmed John ellipsoid)**.** Suppose that a measure μ of total mass d satisfies

$$\sum_{i=1}^{n} \frac{q_i}{d} \log\left(x_i^\top M(\mu) x_i\right) \le \gamma \qquad \forall q \in \Gamma_\kappa$$

Let

$$E(\mu) := \{x : x^\top M(\mu)^{-1} x \le 1\}, \qquad E_\gamma := e^{-\gamma/2} E(\mu)$$

Then there exists $S \subseteq [n]$ with $|S| \ge (1-\kappa)n$ such that

$$E_\gamma \subseteq K_S, \qquad K_P \subseteq \sqrt{d}\, E(\mu)$$

and

$$\mathrm{vol}(E_\gamma) \ge e^{-d\gamma/2} \mathrm{vol}(E^*)$$

Proof. By Lemma C.3, there exists $S \subseteq [n]$ with $|S| \ge (1-\kappa)n$ such that

$$x_i^\top M(\mu) x_i \le e^\gamma \qquad \forall i \in S.$$

For every $x \in E_\gamma$,

$$x^\top M(\mu)^{-1} x \le e^{-\gamma}$$

Thus, by Cauchy–Schwarz, for every $i \in S$,

$$|x_i^\top x| = |\langle M(\mu)^{1/2} x_i, M(\mu)^{-1/2} x\rangle| \le \sqrt{x_i^\top M(\mu) x_i}\sqrt{x^\top M(\mu)^{-1} x} \le e^{\gamma/2} e^{-\gamma/2} = 1$$

Hence $E_\gamma \subseteq K_S$.

Next, for any $x \in K_P$,

$$x^\top M(\mu)^{-1} x = \sum_{i=1}^{n} \mu_i |x_i^\top x|^2 \le \sum_{i=1}^{n} \mu_i = d$$

Consequently, $K_P \subseteq \sqrt{d}\, E(\mu)$.

For the volume guarantee, recall that the John-ellipsoid dual objective is

$$D(\mu) = \sum_{i=1}^{n} \mu_i - \log\det \Sigma(\mu) - d$$

Since μ has total mass d,

$$D(\mu) = -\log\det \Sigma(\mu)$$

Let μ^* denote an optimal dual solution. By weak duality and strong duality at the optimum,

$$-\log\det \Sigma(\mu) \ge -\log\det \Sigma(\mu^*)$$

Since $\mathrm{vol}(E(\mu))$ is proportional to $\det(\Sigma(\mu))^{-1/2}$, we obtain

$$\mathrm{vol}(E(\mu)) \ge \mathrm{vol}(E^*)$$

Finally,

$$\mathrm{vol}(E_\gamma) = e^{-d\gamma/2} \mathrm{vol}(E(\mu)) \ge e^{-d\gamma/2} \mathrm{vol}(E^*)$$

□

C.3 Proof of the Main Non-Private Guarantee

Proof of Theorem 3.1. Algorithm 1 initializes $\mu_i^0 = d/n$ and performs $T = \left\lceil \frac{\log(1/\kappa)}{\gamma} \right\rceil$ iterations.

By Corollary C.2, the averaged measure $\overline{\mu}$ satisfies

$$\sum_{i=1}^{n} \frac{q_i}{d} \log\left(x_i^\top M(\overline{\mu}) x_i\right) \le \gamma \qquad \forall q \in \Gamma_\kappa$$

Moreover, each iterate belongs to Γ_κ, which is convex. Hence $\overline{\mu} \in \Gamma_\kappa$, and in particular $\sum_i \overline{\mu}_i = d$.

Applying Corollary C.4 to $\overline{\mu}$ gives the claimed trimmed containment, outer containment, and volume guarantees for the output $M = M(\overline{\mu})$. The claimed iteration complexity follows from the definition of T. □

D PROOFS FOR THE PRIVATE ALGORITHM

We provide the details underlying the privacy and utility analysis of Algorithm 2. We first analyze the sensitivity of the covariance oracle, and then show that sufficiently small Gaussian perturbations preserve the multiplicative-weights potential argument.

D.1 PRIVACY OF THE MAIN ALGORITHM

Lemma D.1 (Privacy)**.** Algorithm 2 is ρ-zCDP.

Proof. We first show that Algorithm 3 is ρ-zCDP. Let P and P' be two neighboring datasets of size n that differ in a single point, with $x_i \in P$ replaced by $x_i' \in P'$. Assuming that $\|x_i\|_2, \|x_i'\|_2 \le R$, the corresponding weighted covariance matrices can differ in spectral norm by at most

$$\begin{aligned}\|\sum_j \mu_j(P)x_jx_j^\top - \sum_j \mu_j(P')x_jx_j^\top\|_F &= \|\sum_j (\mu_j(P)-\mu_j(P'))x_jx_j^\top + \mu_i(P')x_ix_i^\top\|_F \\ &\le R^2\|\mu(P)-\mu(P')\|_1 + \frac{2dR^2}{\kappa n} \le \frac{4dR^2}{\kappa n}\end{aligned}$$

where the last inequality follows from the bound on the TV-distance between the neighboring κ-dense measures μ and μ', as proven in Bun et al. (2020) (Lemma 25). where the last inequality follows from $\mu \in \Gamma_\kappa$, and hence $\mu_i \le d/(\kappa n)$ for every $i \in [n]$.

Algorithm 2 invokes Algorithm 3 a total of $T+1$ times, with each invocation satisfying $\rho/(T+1)$-zCDP. The final invocation uses the averaged measure $\overline{\mu} = \frac{1}{T}\sum_{t=1}^{T}\mu^t$. Since each $\mu^t \in \Gamma_\kappa$ and Γ_κ is convex, we also have $\overline{\mu} \in \Gamma_\kappa$, so the same sensitivity bound applies to the final release. The claimed ρ-zCDP guarantee therefore follows from sequential composition of zCDP. □

D.2 GOODNESS AND MATRIX PERTURBATION

Proof of Lemma 4.2. Fix $\mu \in \Gamma_\kappa$ and a unit vector $v \in \mathbb{S}^{d-1}$. Let

$$p_i := \frac{\mu_i}{d}, \qquad B_v := \left\{ i \in [n] : \langle x_i, v\rangle^2 < \frac{2\tau}{d} \right\}$$

By (κ,τ)-goodness, $|B_v| < \frac{\kappa n}{2}$. Since $p_i \le 1/(\kappa n)$, $\sum_{i\in B_v} p_i < \frac{1}{2}$. Therefore,

$$\begin{aligned}v^\top \Sigma(\mu) v &= d\sum_{i=1}^{n} p_i \langle x_i, v\rangle^2 \\ &\ge 2\tau \sum_{i\notin B_v} p_i > \tau\end{aligned}$$

Taking the infimum over all unit vectors v gives $\lambda_{\min}(\Sigma(\mu)) \ge \tau$. □

Lemma D.2 (Inverse perturbation)**.** Let $\Sigma \succ 0$ and $\widehat{\Sigma} = \Sigma + N$, where N is symmetric. Suppose

$$\|\Sigma^{-1/2} N \Sigma^{-1/2}\|_2 \le \eta \le \frac{1}{2}$$

Writing $M := \Sigma^{-1}$, $\widehat{M} := \widehat{\Sigma}^{-1}$, we have

$$\frac{1}{1+\eta} M \preceq \widehat{M} \preceq \frac{1}{1-\eta} M$$

Consequently, for every nonzero x,

$$\left|\log(x^\top \widehat{M} x) - \log(x^\top M x)\right| \le 2\eta \qquad \text{and} \qquad \left|\operatorname{tr}(\widehat{M}\Sigma) - d\right| \le 2d\eta$$

Proof. The assumption gives

$$-\eta I \preceq \Sigma^{-1/2} N \Sigma^{-1/2} \preceq \eta I$$

Hence

$$(1-\eta)\Sigma \preceq \widehat{\Sigma} \preceq (1+\eta)\Sigma$$

Inverting yields

$$\frac{1}{1+\eta} M \preceq \widehat{M} \preceq \frac{1}{1-\eta} M$$

Therefore,

$$\frac{1}{1+\eta} \leq \frac{x^\top \widehat{M} x}{x^\top M x} \leq \frac{1}{1-\eta}$$

Since $\eta \leq 1/2$,

$$\log(1+\eta) \leq \eta, \qquad -\log(1-\eta) \leq 2\eta$$

which proves the logarithmic-score bound.

Moreover, the eigenvalues of $\Sigma^{1/2}\widehat{M}\Sigma^{1/2}$ belong to $\left[\frac{1}{1+\eta}, \frac{1}{1-\eta}\right]$. Thus

$$\left|\mathrm{tr}(\widehat{M}\Sigma) - d\right| \leq \frac{d\eta}{1-\eta} \leq 2d\eta$$

□

D.3 Stability of the KL-Potential Argument

Proof of Lemma 4.3. For every round t, define

$$\widehat{\phi}_i^t := \log\left(x_i^\top \widehat{M}^t x_i\right), \qquad \phi_i^t := \phi_i(\mu^t) = \log\left(x_i^\top M(\mu^t) x_i\right)$$

and

$$\theta_i^t := \log\left(\frac{w_i^t}{\mu_i^0}\right), \qquad F(\theta) := \max_{\mu \in \Gamma_\kappa} \left\{\langle \mu, \theta\rangle - D_{\mathrm{KL}}(\mu\|\mu^0)\right\}$$

Since $w_i^t = \mu_i^0 e^{\theta_i^t}$, for every $\mu \in \Gamma_\kappa$ we have

$$\langle \mu, \theta^t\rangle - D_{\mathrm{KL}}(\mu\|\mu^0) = -D_{\mathrm{KL}}(\mu\|w^t) + \sum_i w_i^t - d$$

The last term is independent of μ, and since $\mu^t = \Pi_{\Gamma_\kappa}(w^t)$, the measure μ^t attains the maximum in $F(\theta^t)$.

Moreover, by the Pythagorean property of the KL projection (Theorem A.12), for every $\nu \in \Gamma_\kappa$,

$$F(\theta^t) - \left(\langle \nu, \theta^t\rangle - D_{\mathrm{KL}}(\nu\|\mu^0)\right) = -D_{\mathrm{KL}}(\mu^t\|w^t) + D_{\mathrm{KL}}(\nu\|w^t) \geq D_{\mathrm{KL}}(\nu\|\mu^t)$$

Equivalently,

$$\langle \nu, \theta^t\rangle - D_{\mathrm{KL}}(\nu\|\mu^0) \leq F(\theta^t) - D_{\mathrm{KL}}(\nu\|\mu^t)$$

The private multiplicative update gives $w_i^{t+1} = w_i^t e^{\widehat{\phi}_i^t}$, and therefore $\theta_i^{t+1} = \theta_i^t + \widehat{\phi}_i^t$. Hence,

$$\begin{aligned} F(\theta^{t+1}) &= \max_{\mu \in \Gamma_\kappa} \left\{\langle \mu, \theta^t\rangle + \sum_i \mu_i \widehat{\phi}_i^t - D_{\mathrm{KL}}(\mu\|\mu^0)\right\} \\ &\leq F(\theta^t) + \max_{\mu \in \Gamma_\kappa} \left\{\sum_i \mu_i \widehat{\phi}_i^t - D_{\mathrm{KL}}(\mu\|\mu^t)\right\} \end{aligned}$$

Since both μ and μ^t have total mass d,

$$D_{\mathrm{KL}}(\mu\|\mu^t) = \sum_i \mu_i \log\left(\frac{\mu_i}{\mu_i^t}\right)$$

and thus

$$\sum_i \mu_i \widehat{\phi}_i^t - D_{\mathrm{KL}}(\mu\|\mu^t) = -\sum_i \mu_i \log\left(\frac{\mu_i}{\mu_i^t e^{\widehat{\phi}_i^t}}\right) \overset{(*)}{\leq} d\log\left(\frac{\sum_i \mu_i^t e^{\widehat{\phi}_i^t}}{d}\right)$$

where $(*)$ follows from the log-sum inequality and $\sum_i \mu_i = d$. Consequently,

$$F(\theta^{t+1}) \leq F(\theta^t) + d\log\left(\frac{1}{d}\sum_i \mu_i^t e^{\widehat{\phi}_i^t}\right)$$

By Lemma D.2, $|\widehat{\phi}_i^t - \phi_i^t| \leq 2\eta$, and hence $e^{\widehat{\phi}_i^t} \leq e^{2\eta} e^{\phi_i^t}$. Therefore,

$$\sum_i \mu_i^t e^{\widehat{\phi}_i^t} \leq e^{2\eta} \sum_i \mu_i^t e^{\phi_i^t}$$

Recall the leverage-score identity, $\sum_i \mu_i^t e^{\phi_i^t} = \sum_i \mu_i^t x_i^\top M(\mu^t) x_i = d$. It follows that

$$F(\theta^{t+1}) \leq F(\theta^t) + d\log(e^{2\eta}) = F(\theta^t) + 2d\eta$$

Since $\theta^0 = 0$ and $\mu^0 \in \Gamma_\kappa$, we have $F(\theta^0) = 0$. Summing over the T rounds gives $F(\theta^T) \leq 2dT\eta$. Fix any $q \in \Gamma_\kappa$. By the definition of F,

$$\langle q, \theta^T\rangle - D_{\mathrm{KL}}(q\|\mu^0) \leq F(\theta^T)$$

and since $\theta_i^T = \sum_{t=0}^{T-1} \widehat{\phi}_i^t$, we obtain

$$\sum_{t=0}^{T-1}\sum_i q_i \widehat{\phi}_i^t \leq D_{\mathrm{KL}}(q\|\mu^0) + 2dT\eta$$

Again using $|\widehat{\phi}_i^t - \phi_i^t| \leq 2\eta$ and $\sum_i q_i = d$, we have

$$\sum_{t=0}^{T-1}\sum_i q_i \phi_i^t \leq \sum_{t=0}^{T-1}\sum_i q_i \widehat{\phi}_i^t + 2dT\eta \leq D_{\mathrm{KL}}(q\|\mu^0) + 4dT\eta$$

Dividing by T and applying Jensen's inequality to the convex functions ϕ_i gives

$$\sum_i q_i \phi_i(\overline{\mu}) \leq \frac{1}{T}\sum_{t=0}^{T-1}\sum_i q_i \phi_i(\mu^t) \leq \frac{D_{\mathrm{KL}}(q\|\mu^0)}{T} + 4d\eta$$

as required. □

D.4 High-Probability Utility Guarantee

Lemma D.3 (Simultaneous control of the Gaussian perturbations)**.** Suppose that each oracle call uses symmetric Gaussian noise with scale $\sigma = \frac{2dR^2}{\kappa n\sqrt{2\rho/(T+1)}}$. If

$$n \geq \frac{32dR^2\left(\sqrt{d} + \sqrt{2\log((T+1)/\beta)}\right)}{\kappa\gamma\tau\sqrt{2\rho/(T+1)}}$$

then, with probability at least $1-\beta$,

$$\max_{0\leq t\leq T} \|N^t\|_2 \leq \frac{\gamma\tau}{8}$$

Proof. For each symmetric Gaussian perturbation, the standard spectral-norm concentration bound gives

$$\Pr\left[\|N^t\|_2 > \sigma\left(\sqrt{d} + \sqrt{2\log((T+1)/\beta)}\right)\right] \leq \frac{\beta}{T+1}$$

A union bound over the $T+1$ oracle calls gives, with probability at least $1-\beta$,

$$\max_{0\leq t\leq T} \|N^t\|_2 \leq \sigma\left(\sqrt{d} + \sqrt{2\log((T+1)/\beta)}\right)$$

The assumed lower bound on n makes the right-hand side at most $\gamma\tau/8$. □

Lemma D.4 (Utility of the averaged measure)**.** Under the assumptions of Theorem 4.1, on the event of Lemma D.3,

$$\sup_{q\in\Gamma_\kappa}\sum_{i=1}^{n} q_i \log\big(x_i^\top M(\overline{\mu})x_i\big) \le d\gamma$$

Consequently, there exists $S \subseteq [n]$ with $|S| \ge (1-\kappa)n$ such that

$$e^{-\gamma/2}E(\overline{\mu}) \subseteq K_S$$

and

$$\mathrm{vol}\left(e^{-\gamma/2}E(\overline{\mu})\right) \ge e^{-d\gamma/2}\,\mathrm{vol}(E^*)$$

Proof. Set $\eta = \gamma/8$. By Lemma 4.3, for every $q \in \Gamma_\kappa$,

$$\begin{aligned}\sum_i q_i \log\big(x_i^\top M(\overline{\mu})x_i\big) &\le \frac{D_{\mathrm{KL}}(q\|\mu^0)}{T} + 4d\eta \\ &\le \frac{d\log(1/\kappa)}{T} + \frac{d\gamma}{2}\end{aligned}$$

Since $T \ge 2\log(1/\kappa)/\gamma$, the right-hand side is at most $d\gamma$.

The trimmed-containment and volume guarantees then follow exactly as in the non-private case. □

D.5 Utility of the Final Released Ellipsoid

Lemma D.5. Let

$$M := M(\overline{\mu}), \qquad \widehat{M} := \big(\Sigma(\overline{\mu}) + N^T\big)^{-1}$$

On the event of Lemma D.3,

$$\frac{1}{1+\eta}M \preceq \widehat{M} \preceq \frac{1}{1-\eta}M, \qquad \eta = \frac{\gamma}{8}$$

Consequently, the ellipsoid

$$\widehat{E} := \left\{x \in \mathbb{R}^d : x^\top \left((1-\eta)e^{-\gamma}\widehat{M}\right)^{-1} x \le 1\right\}$$

satisfies $\widehat{E} \subseteq K_S$ for some $S \subseteq [n]$ with $|S| \ge (1-\kappa)n$, and

$$\mathrm{vol}(\widehat{E}) \ge e^{-d\gamma/2}\left(\frac{1-\eta}{1+\eta}\right)^{d/2}\mathrm{vol}(E^*)$$

Moreover, $K_P \subseteq \sqrt{d(1+\eta)}\,\widehat{E}_0$, where $\widehat{E}_0 = \{x : x^\top \widehat{M}^{-1}x \le 1\}$.

Proof. The matrix inequalities follow from Lemma D.2.

Let

$$M_\gamma := e^{-\gamma}M, \qquad \widehat{M}_\gamma := (1-\eta)e^{-\gamma}\widehat{M}$$

Since $\widehat{M} \preceq \frac{1}{1-\eta}M$, we have $\widehat{M}_\gamma \preceq M_\gamma$. Therefore, the ellipsoid associated with $\widehat{M}_\gamma$ is contained in the ellipsoid associated with M_γ. By Lemma D.4, the latter is contained in K_S, and hence $\widehat{E} \subseteq K_S$.

For the volume guarantee, $\widehat{M} \succeq \frac{1}{1+\eta}M$, so $\widehat{M}_\gamma \succeq \frac{1-\eta}{1+\eta}M_\gamma$. Taking determinants gives

$$\mathrm{vol}(\widehat{E}) \ge \left(\frac{1-\eta}{1+\eta}\right)^{d/2}\mathrm{vol}\left(e^{-\gamma/2}E(\overline{\mu})\right)$$

and the desired bound follows from Lemma D.4.

Finally, for every $x \in K_P$, $x^\top M^{-1}x \le d$. Since $\widehat{M} \succeq M/(1+\eta)$, we obtain

$$\widehat{M}^{-1} \preceq (1+\eta)M^{-1}$$

and therefore

$$x^\top \widehat{M}^{-1}x \le d(1+\eta)$$

Thus $K_P \subseteq \sqrt{d(1+\eta)}\,\widehat{E}_0$. □

Proof of Theorem 4.1. By Lemma D.3, with probability at least $1-\beta$ all $T+1$ perturbations satisfy

$$\|N^t\|_2 \le \eta\tau, \qquad \eta = \gamma/8$$

Lemma 4.2 implies $\lambda_{\min}(\Sigma(\mu^t)) \ge \tau$, and hence

$$\lambda_{\min}(\Sigma(\mu^t) + N^t) \ge (1-\eta)\tau \ge \tau/2$$

Thus no oracle invocation returns $\perp$.

Lemma D.4 establishes the guarantee for the averaged measure, and Lemma D.5 transfers it to the released matrix $\widehat{M}$. □

E Population Characterization of The Goodness Parameter

The (κ,τ)-goodness condition used in our private analysis is a finite-sample property of the input dataset. In this section, we show that this condition admits a natural population-level characterization for every distribution $\mathcal{P}$ over $\mathbb{R}^d$. In particular, for any $\kappa \in (0,1)$, we define a canonical distribution-dependent parameter $\tau_{\mathcal{P},\kappa}$ in terms of the lower quantiles of the one-dimensional squared projections $\langle X, v\rangle^2$. We then show that, with sufficiently many i.i.d. samples from $\mathcal{P}$, the resulting dataset is $(\kappa, \tau_{\mathcal{P},\kappa})$-good with high probability. Thus, rather than viewing τ as an arbitrary structural parameter of a finite dataset, it can be interpreted as a population quantity measuring how far the distribution is uniformly bounded away from degeneracy across all directions. The proof follows from uniform convergence over the class of centered slabs, whose VC dimension is $O(d)$.

E.1 Preliminaries for this section: VC dimension

Definition E.1 (VC dimension)**.** Let $\mathcal{A}$ be a class of subsets of a domain $\mathcal{X}$. We say that a finite set $S = \{x_1, \dots, x_m\} \subseteq \mathcal{X}$ is *shattered* by $\mathcal{A}$ if, for every subset $T \subseteq S$, there exists a set $A \in \mathcal{A}$ such that

$$A \cap S = T$$

The *Vapnik–Chervonenkis (VC) dimension* of $\mathcal{A}$, denoted by $\mathrm{VCdim}(\mathcal{A})$, is the largest integer m for which there exists a set $S \subseteq \mathcal{X}$ of cardinality m that is shattered by $\mathcal{A}$. If arbitrarily large finite sets can be shattered by $\mathcal{A}$, then $\mathrm{VCdim}(\mathcal{A}) = \infty$.

Theorem E.2 (Uniform convergence for VC classes (Vapnik & Chervonenkis, 2015))**.** *Let $\mathcal{A}$ be a class of measurable subsets of a domain $\mathcal{X}$ with VC dimension*

$$\mathrm{VCdim}(\mathcal{A}) = v$$

Let $X_1, \dots, X_n \overset{\text{i.i.d.}}{\sim} \mathcal{P}$. Then there exists a universal constant $C > 0$ such that, for every $\epsilon, \beta \in (0,1)$, if

$$n \ge C\frac{v\log(1/\epsilon) + \log(1/\beta)}{\epsilon^2}$$

then, with probability at least $1-\beta$,

$$\sup_{A\in\mathcal{A}} \left| \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\{X_i \in A\} - \Pr_{X\sim\mathcal{P}}[X \in A] \right| \le \epsilon$$

Lemma E.3 (Sauer's Lemma)**.** Let $\mathcal{H}$ be a class of subsets of a domain $\mathcal{X}$ with VC dimension $\nu < \infty$. Let $\Pi_{\mathcal{H}}(m)$ denote the growth function of $\mathcal{H}$, defined by

$$\Pi_{\mathcal{H}}(m) := \max_{\substack{S\subseteq\mathcal{X} \\ |S|=m}} |\{H \cap S : H \in \mathcal{H}\}|$$

Then, for every integer $m \ge 0$,

$$\Pi_{\mathcal{H}}(m) \le \sum_{j=0}^{\nu} \binom{m}{j}$$

In particular, if $m \ge \nu \ge 1$, then

$$\Pi_{\mathcal{H}}(m) \le \sum_{j=0}^{\nu} \binom{m}{j} \le \left(\frac{em}{\nu}\right)^{\nu}$$

E.2 Population Characterization of The Goodness Parameter

Lemma E.4. Let $\mathcal{P}$ be a distribution over $\mathbb{R}^d$, let $\kappa, \beta \in (0,1)$. For every unit vector $v \in \mathbb{S}^{d-1}$, define the cumulative distribution function

$$F_v(t) := \Pr_{x\sim\mathcal{P}}\left[\langle x, v\rangle^2 \le t\right]$$

and its generalized inverse

$$F_v^{-1}(p) := \inf\{t \ge 0 : F_v(t) \ge p\}$$

Define

$$\tau_{\mathcal{P},\kappa} := \frac{d}{2}\inf_{\|v\|_2=1} F_v^{-1}\left(\frac{\kappa}{4}\right)$$

Then there exists a universal constant $C > 0$ such that, if

$$n \ge C\frac{d\log(1/\kappa) + \log(1/\beta)}{\kappa^2}$$

then, with probability at least $1-\beta$, the sample $P = \{x_1, \dots, x_n\} \sim \mathcal{P}^n$ is $(\kappa, \tau_{\mathcal{P},\kappa})$-good. That is, simultaneously for every unit vector $v \in \mathbb{S}^{d-1}$,

$$\left|\left\{i \in [n] : \langle x_i, v\rangle^2 < \frac{2\tau_{\mathcal{P},\kappa}}{d}\right\}\right| < \frac{\kappa n}{2}$$

Proof. Set

$$a := \frac{2\tau_{\mathcal{P},\kappa}}{d} = \inf_{\|v\|_2=1} F_v^{-1}\left(\frac{\kappa}{4}\right)$$

For every unit vector v, define

$$A_v := \left\{x \in \mathbb{R}^d : \langle x, v\rangle^2 < a\right\}$$

By the definition of a,

$$a \le F_v^{-1}\left(\frac{\kappa}{4}\right) \qquad \text{for every } v \in \mathbb{S}^{d-1}$$

Hence

$$\Pr_{x\sim\mathcal{P}}[x \in A_v] = \Pr_{x\sim\mathcal{P}}\left[\langle x, v\rangle^2 < a\right] \le \frac{\kappa}{4}$$

Indeed, by the definition of the generalized inverse,

$$\Pr\left[\langle x, v\rangle^2 < F_v^{-1}\left(\frac{\kappa}{4}\right)\right] \le \frac{\kappa}{4}$$

Now consider the class of sets

$$\mathcal{A} := \left\{A_v : v \in \mathbb{S}^{d-1}\right\}$$

Each A_v is a centered slab of the form

$$A_v = \left\{x : -\sqrt{a} < \langle x, v\rangle < \sqrt{a}\right\}$$

We now use the following lemma.

Lemma E.5 (VC dimension of centered slabs)**.** Fix $a > 0$, and consider the class

$$\mathcal{A} = \left\{A_v = \left\{x \in \mathbb{R}^d : -\sqrt{a} < \langle x, v\rangle < \sqrt{a}\right\} : v \in \mathbb{S}^{d-1}\right\}$$

Then

$$\mathrm{VCdim}(\mathcal{A}) = O(d)$$

More precisely,

$$\mathrm{VCdim}(\mathcal{A}) < 16(d+1)$$

Proof. Let $\mathcal{H}_d$ denote the class of affine halfspaces in $\mathbb{R}^d$. It is well known that

$$\mathrm{VCdim}(\mathcal{H}_d) = d + 1$$

For every $v \in \mathbb{S}^{d-1}$, we can write

$$A_v = \{x : \langle x, v\rangle < \sqrt{a}\} \cap \{x : -\langle x, v\rangle < \sqrt{a}\}$$

Thus every set in $\mathcal{A}$ is the intersection of two affine halfspaces. Therefore,

$$\mathcal{A} \subseteq \mathcal{H}_d^{\cap 2}$$

where

$$\mathcal{H}_d^{\cap 2} := \{H_1 \cap H_2 : H_1, H_2 \in \mathcal{H}_d\}$$

Let $\Pi_{\mathcal{C}}(m)$ denote the growth function of a class $\mathcal{C}$, namely the maximum number of distinct subsets that $\mathcal{C}$ can induce on a set of m points.

For any m-point set S, each set of the form

$$(H_1 \cap H_2) \cap S$$

is determined by the pair

$$(H_1 \cap S, H_2 \cap S)$$

Hence

$$\Pi_{\mathcal{A}}(m) \le \Pi_{\mathcal{H}_d^{\cap 2}}(m) \le \Pi_{\mathcal{H}_d}(m)^2$$

Since $\mathrm{VCdim}(\mathcal{H}_d) = d + 1$, Sauer's lemma (Lemma E.3) gives, for $m \ge d + 1$,

$$\Pi_{\mathcal{H}_d}(m) \le \left(\frac{em}{d+1}\right)^{d+1}$$

Therefore,

$$\Pi_{\mathcal{A}}(m) \le \left(\frac{em}{d+1}\right)^{2(d+1)}$$

Suppose that $\mathcal{A}$ shatters a set of size m. Then

$$2^m \le \Pi_{\mathcal{A}}(m)$$

and consequently

$$2^m \le \left(\frac{em}{d+1}\right)^{2(d+1)}$$

Now take $m = 16(d + 1)$. Then the above inequality would imply

$$2^{16(d+1)} \le (16e)^{2(d+1)}$$

However, $2^{16} > (16e)^2$, and therefore $2^{16(d+1)} > (16e)^{2(d+1)}$. This is a contradiction.

Hence no set of size $16(d + 1)$ can be shattered by $\mathcal{A}$, so

$$\mathrm{VCdim}(\mathcal{A}) < 16(d + 1)$$

In particular,

$$\mathrm{VCdim}(\mathcal{A}) = O(d)$$

□

By Lemma E.5, $\mathrm{VCdim}(\mathcal{A}) = O(d)$. Therefore, by setting $\nu = O(d)$ and $\epsilon = \kappa/4$ in Theorem E.2, there exists a universal constant $C > 0$ such that, if

$$n \ge C\frac{d\log(1/\kappa) + \log(1/\beta)}{\kappa^2}$$

then, with probability at least $1-\beta$,

$$\sup_{A\in\mathcal{A}}\left|\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}\{x_i\in A\}-\Pr_{x\sim\mathcal{P}}[x\in A]\right|<\frac{\kappa}{4}$$

Condition on this event. Then, simultaneously for every unit vector v,

$$\begin{aligned}\frac{1}{n}\left|\left\{i\in[n]:\langle x_i,v\rangle^2<\frac{2\tau_{\mathcal{P},\kappa}}{d}\right\}\right| &= \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}\{x_i\in A_v\}\\ &< \Pr_{x\sim\mathcal{P}}[x\in A_v]+\frac{\kappa}{4}\le\frac{\kappa}{4}+\frac{\kappa}{4}=\frac{\kappa}{2}.\end{aligned}$$

Thus,

$$\left|\left\{i\in[n]:\langle x_i,v\rangle^2<\frac{2\tau_{\mathcal{P},\kappa}}{d}\right\}\right|<\frac{\kappa n}{2}$$

for every unit vector v, and hence P is $(\kappa,\tau_{\mathcal{P},\kappa})$-good. □

F Reduction from uncentered to centered trimmed MVEE

Lemma F.1. Let $Y=\{y_1,\dots,y_n\}\subseteq\mathbb{R}^d$ and define the lifted points

$$x_i:=\begin{pmatrix}y_i\\1\end{pmatrix}\in\mathbb{R}^{d+1},\qquad i\in[n]$$

Let $X:=\{x_1,\dots,x_n\}$.

Suppose that a centered MVEE algorithm applied to X outputs the ellipsoid

$$\widehat{E}_X=\left\{x\in\mathbb{R}^{d+1}:x^\top\widehat{M}x\le 1\right\},\qquad \widehat{M}\succ 0$$

which contains at least a $(1-\kappa)$-fraction of the lifted points, denoted $S^\star$, and whose volume is at most e^γ times that of the optimal centered ellipsoid containing X.

Write

$$\widehat{M}=\begin{pmatrix}A & b\\ b^\top & c\end{pmatrix}$$

and define

$$\widehat{y}:=-A^{-1}b,\qquad s:=c-b^\top A^{-1}b$$

Then

$$\widehat{E}_Y:=\left\{y\in\mathbb{R}^d:(y-\widehat{y})^\top\widehat{M}_Y(y-\widehat{y})\le 1\right\},\qquad \widehat{M}_Y:=\frac{A}{1-s}$$

contains $S^\star$, the same subset of original points as $\widehat{E}_X$ contains among the lifted points. In particular,

$$\left|\left\{i:y_i\in\widehat{E}_Y\right\}\right|\ge(1-\kappa)n$$

Moreover,

$$\mathrm{vol}_d(\widehat{E}_Y)\le e^\gamma\mathrm{OPT}^{\mathrm{unc}}(Y)$$

where $\mathrm{OPT}^{\mathrm{unc}}(Y)$ is the minimum volume of an uncentered ellipsoid containing Y.

Proof. For every $y\in\mathbb{R}^d$,

$$\begin{aligned}\begin{pmatrix}y\\1\end{pmatrix}^\top\widehat{M}\begin{pmatrix}y\\1\end{pmatrix} &= y^\top Ay+2b^\top y+c\\ &= (y-\widehat{y})^\top A(y-\widehat{y})+c-b^\top A^{-1}b\\ &= (y-\widehat{y})^\top A(y-\widehat{y})+s\end{aligned}$$

Hence

$$\begin{pmatrix} y \\ 1 \end{pmatrix} \in \widehat{E}_X \quad \Longleftrightarrow \quad (y - \widehat{y})^\top A (y - \widehat{y}) \le 1 - s$$

Since $\widehat{M} \succ 0$, its Schur complement satisfies $s > 0$; moreover, since $\widehat{E}_X$ contains at least one lifted point, we have $s < 1$. Therefore

$$\widehat{M}_Y = \frac{A}{1 - s}$$

is well defined and

$$\begin{pmatrix} y \\ 1 \end{pmatrix} \in \widehat{E}_X \quad \Longleftrightarrow \quad (y - \widehat{y})^\top \widehat{M}_Y (y - \widehat{y}) \le 1$$

Consequently, for every $i \in [n]$,

$$x_i \in \widehat{E}_X \quad \Longleftrightarrow \quad y_i \in \widehat{E}_Y$$

and thus

$$\left| \left\{ i : y_i \in \widehat{E}_Y \right\} \right| = \left| \left\{ i : x_i \in \widehat{E}_X \right\} \right| \ge (1 - \kappa) n$$

It remains to establish the approximation guarantee. By the Schur complement identity,

$$\det(\widehat{M}) = \det(A) \cdot s$$

whereas

$$\det(\widehat{M}_Y) = \frac{\det(A)}{(1 - s)^d}$$

Therefore

$$\frac{\det(\widehat{M})}{\det(\widehat{M}_Y)} = s(1 - s)^d$$

The function

$$f(s) = s(1 - s)^d, \qquad s \in (0, 1)$$

is maximized at $s = \frac{1}{d+1}$, and hence

$$s(1 - s)^d \le \frac{d^d}{(d + 1)^{d+1}}$$

Denoting $C_d := \frac{d^d}{(d+1)^{d+1}}$ we obtain

$$\det(\widehat{M}_Y) \ge \frac{\det(\widehat{M})}{C_d} \tag{1}$$

Now fix any subset $S \subseteq [n]$. Let $M_S^\star$ be the matrix of the minimum-volume centered ellipsoid of the lifted points $\{x_i : i \in S\}$, written as

$$\{x : x^\top M_S^\star x \le 1\}$$

and let $M_{Y,S}^\star$ be the matrix of the minimum-volume uncentered ellipsoid of the original points $\{y_i : i \in S\}$, written as

$$\{y : (y - y_S^\star)^\top M_{Y,S}^\star (y - y_S^\star) \le 1\}$$

By the standard lifting reduction from uncentered to centered MVEE, the Schur complement parameter of the optimal lifted solution is

$$s^\star = \frac{1}{d + 1}$$

and its leading block satisfies

$$A^\star = \frac{d}{d + 1} M_{Y,S}^\star$$

Consequently,

$$\det(M_S^\star) = \frac{d^d}{(d + 1)^{d+1}} \det(M_{Y,S}^\star) = C_d \det(M_{Y,S}^\star) \tag{2}$$

Since the relation (2) holds for every subset S, it also holds for $S^\star$. Thus, if $M^\star$ and $M_Y^\star$ denote optimal matrices for the lifted centered and original uncentered trimmed problems, respectively, then

$$\det(M^\star) = C_d \det(M_Y^\star) \tag{3}$$

Using (1) and (3), we obtain

$$\begin{aligned}
\frac{\mathrm{vol}_d(\widehat{E}_Y)}{\mathrm{OPT}^{\mathrm{unc}}(Y)} &\overset{(*)}{=} \sqrt{\frac{\det(M_Y^\star)}{\det(\widehat{M}_Y)}} \\
&\leq \sqrt{\frac{\det(M^\star)/C_d}{\det(\widehat{M})/C_d}} \\
&= \sqrt{\frac{\det(M^\star)}{\det(\widehat{M})}} \overset{(*)}{=} \frac{\mathrm{vol}_{d+1}(\widehat{E}_X)}{\mathrm{OPT}^{\mathrm{cen}}(X)} \leq e^\gamma
\end{aligned}$$

where the inequalities marked with $(*)$ follow from the fact that an ellipsoid $\{x : \ x^T M x \leq 1\}$ has volume proportional to $\det(M)^{-1/2}$. Therefore

$$\mathrm{vol}_d(\widehat{E}_Y) \leq e^\gamma \mathrm{OPT}^{\mathrm{unc}}(Y)$$

as claimed. □