Re-Analyze Gauss: Bounds for Private Matrix Approximation via Dyson Brownian Motion

Anonymous Author(s)
Affiliation
Address
email

Abstract

Given a symmetric matrix M and a vector λ, we present new bounds on the Frobenius-distance utility of the Gaussian mechanism for approximating M by a matrix whose spectrum is λ, under (ε, δ)-differential privacy. Our bounds depend on both λ and the gaps in the eigenvalues of M, and hold whenever the top $k+1$ eigenvalues of M have sufficiently large gaps. When applied to the problems of private rank-k covariance matrix approximation and subspace recovery, our bounds yield improvements over previous bounds. Our bounds are obtained by viewing the addition of Gaussian noise as a continuous-time matrix Brownian motion. This viewpoint allows us to track the evolution of eigenvalues and eigenvectors of the matrix, which are governed by stochastic differential equations discovered by Dyson. These equations allow us to bound the utility as the square-root of a sum-of-squares of perturbations to the eigenvectors, as opposed to a sum of perturbation bounds obtained via Davis-Kahan-type theorems.

1 Introduction

Given a dataset $A \in \mathbb{R}^{m \times d}$, which consists of m individuals with d-dimensional features, methods for preprocessing or prediction from A often use the covariance matrix $M := A^\top A$ of A. In many such applications one computes a rank-k approximation to M, or finds a matrix *close* to M with a specified set of eigenvalues $\lambda = (\lambda_1, \ldots, \lambda_d)$ [38, 28, 37]. Examples include the rank-k covariance matrix approximation problem where one seeks to compute a rank-k matrix which minimizes a given distance to M, and the subspace recovery problem where the goal is to compute a rank k-projection matrix $H = V_k V_k^\top$, where V_k is the $d \times k$ matrix whose columns are the top-k eigenvectors of M. These matrix approximation problems are ubiquitous in ML and have a rich algorithmic history; see [29, 46, 10, 8].

In some cases, the rows of A correspond to sensitive features of individuals and the release of solutions to aforementioned matrix approximation problems may reveal their private information, e.g., as in the case of the Netflix prize problem [5]. Differential privacy (DP) has become a popular notion to quantify the extent to which an algorithm preserves privacy of individuals [15]. Algorithms for solving low-rank matrix approximation problems have been widely studied under DP constraints [30, 7, 19, 17]. Notions of DP studied in the literature include (ε, δ)-DP [17, 25, 26, 19] which is the notion we study in this paper, as well as pure $(\varepsilon, 0)$-DP [17, 30, 2, 32]. To define a notion of DP in problems involving covariance matrices, following [7, 17], two matrices $M = A^\top A$ and $M' = A'^\top A'$ are said to be *neighbors* if they arise from A, A' which differ by at most one row and as, is oftentimes done, require that each row of the datasets A, A' has norm at most 1. For any $\varepsilon, \delta \geq 0$, a randomized mechanism $\mathcal{A}$ is (ε, δ)-differentially private if for all neighbors $M, M' \in \mathbb{R}^{d \times d}$, and any measurable subset S of outputs of $\mathcal{A}$, we have $\mathbb{P}(\mathcal{A}(M) \in S) \leq e^\varepsilon \mathbb{P}(\mathcal{A}(M') \in S) + \delta$.

Submitted to 36th Conference on Neural Information Processing Systems (NeurIPS 2022). Do not distribute.

The problem. We consider a class of problems where one wishes to compute an approximation to a symmetric $d \times d$ matrix under (ε, δ)-differential privacy constraints. Specifically, given $M = A^\top A$ for $A \in \mathbb{R}^{m \times d}$, together with a vector λ of target eigenvalues $\lambda_1 \geq \cdots \geq \lambda_d$, the goal is to output a $d \times d$ matrix $\hat{H}$ with eigenvalues λ which minimizes the Frobenius-norm distance $\|\hat{H} - H\|_F$ under (ε, δ)-differential privacy constraints. Here H is the matrix with eigenvalues λ and the same eigenvectors as M. This class of problems includes as a special case the subspace recovery problem if we set $\lambda_1 = \cdots = \lambda_k = 1$ and $\lambda_{k+1} = \cdots = \lambda_d = 0$. It also includes the rank-k covariance approximation problems if we set $\lambda_i = \sigma_i$ for $i \leq k$, where $\sigma_1 \geq \cdots \geq \sigma_d$ are the eigenvalues of M. Since revealing σ_is may violate privacy constraints, the eigenvalues of the output matrix $\hat{H}$ should not be the same as those of H.

Various distance functions have been used in the literature to evaluate the utility of (ε, δ)-DP mechanisms for matrix approximation problems, including the Frobenius-norm distance $\|\hat{H} - H\|_F$ (e.g. [19, 2])and the Frobenius inner product utility $\langle M, H - \hat{H} \rangle$ (e.g. [11, 19, 24]). Note that while a bound $\|H - \hat{H}\|_F \leq b$ implies an upper bound on the inner product utility of $\langle M, H - \hat{H} \rangle \leq \|M\|_F \cdot b$ (by the Cauchy-Schwarz inequality), an upper bound on the inner product utility does not (in general) imply any upper bound on the Frobenius-norm distance. Moreover, the Frobenius-norm distance can be a good utility metric to use if the goal is to recover a low rank matrix H from a dataset of noisy observations (see e.g. [12]). Hence, we use the Frobenius-norm distance to measure the utility of an (ε, δ)-DP mechanism.

Related work. The problem of approximating a matrix under differential privacy constraints has been widely studied. In particular, prior works have provided algorithms for problems where the goal is to approximate a covariance matrix under differential privacy constraints, including rank-k PCA and subspace recovery [7, 30, 19, 33] as well as rank-k covariance matrix approximation [7, 19, 2]. Another set of works have studied the problem of approximating a rectangular data matrix A under DP [7, 1, 25, 26]. We note that upper bounds on the utility of differentially-private mechanisms for rectangular matrix approximation problems can grow with the number of datapoints m, while those for covariance matrix approximation problems oftentimes depend only on the dimension d of the covariance matrix and do not grow with m. Prior works which deal with covariance matrix approximation problems such as rank-k covariance matrix approximation and subspace recovery are the most relevant to our paper. The notion of DP varies among the different works on differentially-private matrix approximation, with many of these works considering the notion (ε, δ)-DP [25, 26, 19], while other works focus on (pure) $(\varepsilon, 0)$-DP [30, 2, 33].

Analysis of the Gaussian mechanism in [19]. [19] analyze a version of the Gaussian mechanism of [16], where one perturbs the entries of M by adding a symmetric matrix E with iid Gaussian entries $N(0, \sqrt{\log(\frac{1}{\delta})}/\varepsilon)$, to obtain an (ε, δ)-differentially private mechanism which outputs a perturbed matrix $\hat{M} = M + E$. One can then post-process this matrix $\hat{M}$ to obtain a rank-k projection matrix which projects onto the subspace spanned by the top-k eigenvectors of $\hat{M}$ (for the rank-k PCA or subspace recovery problem), or a rank-k matrix $\hat{H}$ with the same top-k eigenvectors and eigenvalues as $\hat{M}$ (for the rank-k covariance matrix approximation problem). [19] consider different notions of utility in their results, including the inner product utility (for PCA), and the Frobenius-norm and spectral-norm distance distances (for low-rank approximation and subspace recovery).

In one set of results, [19] give lower utility bounds of $\tilde{\Omega}(k\sqrt{d})$ w.h.p. for the rank-k PCA problem with respect to the inner product utility $\langle M, H \rangle$, together with matching upper bounds provided by a post-processing of the Gaussian mechanism, where $\tilde{\Omega}$ hides polynomial factors of $\frac{1}{\varepsilon}$ and $\log(\frac{1}{\delta})$ (their Theorems 3 and 18). As noted by the authors, their lower bounds are tight for matrices M with the "worst-case" spectral profile σ, but they can obtain improved upper bounds for matrices M where $\sigma_k - \sigma_{k+1} > \tilde{\Omega}(\sqrt{d})$ (Theorem 3 of [19]).

For the subspace recovery problem, [19] obtain a Frobenius-distance bound of $\|\hat{H} - H\|_F \leq \tilde{O}\left(\sqrt{kd}/(\sigma_k - \sigma_{k+1})\right)$ w.h.p. for a post-processing of the Gaussian mechanism whenever $\sigma_k - \sigma_{k+1} > \tilde{\Omega}(\sqrt{d})$ (implied by their Theorem 6, which is stated for the spectral norm). And for the rank-k covariance matrix approximation problem, [19] show a utility bound of

$\|\hat{H} - M\|_F - \|H - M\|_F \leq \tilde{O}(k\sqrt{d})$ w.h.p. for a post-processing of the Gaussian mechanism (Theorem 7 in [19]), and also give related bounds for the spectral norm. While their Frobenius bound for the covariance matrix approximation problem is independent of the number of datapoints m, it may not be tight. For instance, when $k = d$, one can easily obtain a better bound since, by the triangle inequality, $\|\hat{H} - M\|_F - \|H - M\|_F \leq \|\hat{H} - H\|_F = \|\hat{M} - M\|_F = \|E\|_F \leq O(d)$ w.h.p., since $\|E\|_F$ is just the norm of a vector of d^2 Gaussians with variance $\tilde{O}(1)$. Moreover, the bound for the the rank-k covariance approximation problem, $\|\hat{H} - H\|_F \leq \tilde{O}(k\sqrt{d})$, is also a worst-case upper bound for any spectral profile σ as the right hand side of the bound not depend on the eigenvalues σ.

Thus, a question arises of whether the Frobenius-norm utility bounds for the rank-k covariance matrix approximation and subspace recovery problems are tight for all spectral profiles σ, and whether the analysis of the Gaussian mechanism can be improved to achieve better utility bounds. A more general question is to obtain utility bounds for the Gaussian mechanism for the matrix approximation problems for arbitrary λ.

Our contribution. Our main result is a new upper bound on the Frobenius-distance utility of the Gaussian mechanism for the general matrix approximation problem for a given M and λ (Theorem 2.2). Our bound depends on the eigenvalues of M and the entries of λ.

The novel insight is to view the perturbed matrix $M + E$ as a continuous-time symmetric matrix diffusion, where each entry of the matrix $M + E$ is the value reached by a (one-dimensional) Brownian motion after some time $T = \sqrt{\log(\frac{1}{\delta})}/\varepsilon$. This matrix-valued Brownian motion, which we denote by $\Phi(t)$, induces a stochastic process on the eigenvalues $\gamma_1(t) \geq \cdots \geq \gamma_d(t)$ and corresponding eigenvectors $u_1(t), \ldots, u_d(t)$ of $\Phi(t)$ originally discovered by Dyson and now referred to as Dyson Brownian motion, with initial values $\gamma_i(0) = \sigma_i$ and $u_i(0)$ which are the eigenvalues and eigenvectors of the initial matrix M [20].

We then use the stochastic differential equations (2) and (3), which govern the evolution of the eigenvalues and eigenvectors of the Dyson Brownian motion, to track the perturbations to each eigenvector. Roughly speaking, these equations say that, as the Dyson Brownian motion evolves over time, every pair of eigenvalues $\gamma_i(t)$ and $\gamma_j(t)$, and corresponding eigenvectors $u_i(t)$ and $u_j(t)$, interacts with the other eigenvalue/eigenvector with the magnitude of the interaction term proportional to $\frac{1}{\gamma_i(t)-\gamma_j(t)}$ at any given time t. This allows us to bound the perturbation of the eigenvectors at every time t, provided that the initial gaps in the eigenvectors $\sigma_i - \sigma_j$ are sufficiently large. We then derive a stochastic differential equation which tracks how the utility changes as the Dyson Brownian motion evolves over time (Lemma 4.1), and integrate this differential equation over time to obtain a bound on the (expectation of) the utility $\mathbb{E}[\|\hat{H} - H\|_F]$ (Lemma 4.5) as a function of the gaps $\gamma_i(t) - \gamma_j(t)$.

Plugging in basic estimates (Lemma 4.4) for the eigenvalue gaps $\gamma_i(t) - \gamma_j(t)$ to Lemma 4.5, we obtain a bound on the expected utility $\mathbb{E}[\|\hat{H} - H\|_F]$ (Theorem 2.2) for the different matrix approximation problems as a function of the eigenvalue gaps $\sigma_i - \sigma_j$ of the input matrix M. Roughly speaking, our bound is the square-root of a sum-of-squares of the ratios, $\frac{\lambda_i - \lambda_j}{\sigma_i - \max(\sigma_j, \sigma_{k+1})}$, of eigenvalue gaps of the input and output matrices.

When applied to the rank-k covariance matrix approximation problem (Corollary 2.3), Theorem 2.2 implies a bound of $\mathbb{E}[\|\hat{H} - H\|_F] \leq \tilde{O}(\sqrt{kd})$ whenever the eigenvalues σ of the input matrix M satisfy $\sigma_k - \sigma_{k+1} \geq \Omega(\sigma_k)$ and the gaps in top $k + 1$ eigenvalues satisfy $\sigma_i - \sigma_{i+1} \geq \tilde{\Omega}(\sqrt{d})$. Thus, when M satisfies the above condition on σ, our bound improves by a factor of $\sqrt{k}$ on the (expectation of) the previous bound of [19], which says that $\|\hat{H} - M\|_F - \|H - M\|_F \leq \tilde{O}(k\sqrt{d})$ w.h.p., since by the triangle inequality $\|\hat{H} - M\|_F - \|H - M\|_F \leq \|\hat{H} - H\|_F$. This condition on σ is satisfied, e.g., for matrices M whose eigenvalue gaps are at least as large as those of the Wishart random covariance matrices with sufficiently many datapoints m (see Section 2 for details). And, if σ is such that $\sigma_i - \sigma_{i+1} \geq \Omega(\sigma_k - \sigma_{k+1})$ for $i \leq k$, Theorem 2.2 implies a bound of $\mathbb{E}[\|\hat{H} - H\|_F] \leq \tilde{O}(\sqrt{d}/(\sigma_k - \sigma_{k+1}))$ for the subspace recovery problem (Corollary 2.4), improving by a factor of $\sqrt{k}$ (in expectation) on the previous bound of [19], which implies that $\|\hat{H} - M\|_F - \|H - M\|_F \leq \tilde{O}\left(\sqrt{kd}/(\sigma_k - \sigma_{k+1})\right)$ w.h.p.

2 Results

Our main result (Theorem 2.2) gives a new and unified upper bound on the Frobenius-norm utility of a post-processing of the Gaussian mechanism, for the general matrix approximation problem where one is given a symmetric matrix $M \in \mathbb{R}^{d\times d}$ and a vector λ with $\lambda_1 \geq \cdots \geq \lambda_d$, and the goal is to compute a matrix $\hat{H}$ with eigenvalues λ which minimizes the distance $\|\hat{H} - H\|_F$. Here H is the matrix with eigenvalues λ and the same eigenvectors as M. Plugging in different choices of λ to Theorem 2.2, we obtain as corollaries new Frobenius-distance utility bounds for the rank-k covariance matrix approximation problem (Corollary 2.3) and the subspace recovery problem (Corollary 2.4). Our results rely on the following assumption about the eigenvalues of the input matrix M:

Assumption 2.1 ($(M, k, \lambda_1, \varepsilon, \delta)$ Eigenvalue gaps)**.** *The gaps in the top $k+1$ eigenvalues eigenvalues $\sigma_1 \geq \cdots \geq \sigma_d$ of the matrix $M \in \mathbb{R}^{d\times d}$ satisfy $\sigma_i - \sigma_{i+1} \geq \frac{8\sqrt{\log(\frac{1.25}{\delta})}}{\varepsilon}\sqrt{d} + 3\log^{\frac{1}{2}}(\lambda_1 k)$ for every $i \in [k]$.*

Assumption 2.1 is satisfied, for instance, by random Wishart matrices $W = A^\top A$, where A is an $m \times d$ matrix of iid Gaussian entries, which are a popular model for sample covariance matrices [48]. This is because the minimum gap $\sigma_i - \sigma_{i+1}$ of a Wishart matrix grows proportional to $\sqrt{m}$ with high probability; thus for large enough m, Assumption 2.1 holds (see Appendix I for details). Hence, the assumption requires that the gaps in the top $k+1$ eigenvalues of M are at least as large as the gaps in a random Wishart matrix.

Theorem 2.2 (Main result). *Let $\varepsilon, \delta > 0$, and given a symmetric matrix $M \in \mathbb{R}^{d\times d}$ with eigenvalues $\sigma_1 \geq \cdots \geq \sigma_d$ and corresponding orthonormal eigenvectors $v_1, \ldots, v_d$. Let G be a matrix with iid $N(0,1)$ entries, and consider the mechanism that outputs $\hat{M} = M + \frac{\sqrt{2\log(\frac{1.25}{\delta})}}{\varepsilon}(G + G^\top)$. Then such a mechanism is (ε, δ)-differentially private. Moreover, let $\lambda_1 \geq \cdots \geq \lambda_d$ and $k \in [d]$ be any numbers such that $\lambda_i = 0$ for $i > k$, and define $\Lambda := \mathrm{diag}(\lambda_1, \ldots, \lambda_d)$ and $V = [v_1, \ldots, v_d]$, and define $\hat{\sigma}_1 \geq \cdots \geq \hat{\sigma}_d$ to be the eigenvalues of $\hat{M}$ with corresponding orthonormal eigenvectors $\hat{v}_1, \ldots, \hat{v}_d$ and $\hat{V} = [\hat{v}_1, \ldots, \hat{v}_d]$. Then if M satisfies Assumption 2.1 for $(M, k, \lambda_1, \varepsilon, \delta)$, we have*

$$\mathbb{E}\left[\|\hat{V}\Lambda\hat{V}^\top - V\Lambda V^\top\|_F^2\right] \leq O\left(\sum_{i=1}^{k}\sum_{j=i+1}^{d}\frac{(\lambda_i-\lambda_j)^2}{(\sigma_i - \max(\sigma_j, \sigma_{k+1}))^2}\right)\frac{\log(\frac{1}{\delta})}{\varepsilon^2}$$

The fact that the mechanism in this theorem is (ε, δ)-differentially private follows from standard results about the Gaussian mechanism [19]. Given any list of eigenvalues λ, and letting $\Lambda = \mathrm{diag}(\lambda)$, one can post-process the matrix $\hat{M}$ by computing its spectral decomposition $\hat{M} = \hat{V}\hat{\Sigma}\hat{V}^\top$ and replacing its eigenvalues to obtain a matrix $\hat{V}\Lambda\hat{V}^\top$ with eigenvalues λ and eigenvectors $\hat{V}$. Since $\hat{V}\Lambda\hat{V}^\top$ is a post-processing of the Gaussian mechanism, the mechanism which outputs $\hat{V}\Lambda\hat{V}^\top$ is differentially private as well. Theorem 2.2 bounds the excess utility $\mathbb{E}[\|\hat{V}\Lambda\hat{V}^\top - V\Lambda V^\top\|_F^2]$ (whenever the gaps in the eigenvalues $\sigma_1 \geq \cdots \geq \sigma_d$ of the input matrix satisfy Assumption 2.1) as a sum-of-squares of the ratio of the gaps $\lambda_i - \lambda_j$ in the given eigenvalues to the corresponding gaps $\sigma_i - \max(\sigma_j, \sigma_{k+1})$ in the eigenvalues of the input matrix (note that $\lambda_i - \lambda_j = \lambda_i - \max(\lambda_j, \lambda_{k+1})$ since $\lambda_j = 0$ for $j \geq k+1$).

While we do not know if Theorem 2.2 is tight for all choices of λ and k, it does give a tight bound for some problems. Namely, when applied to the covariance matrix estimation problem, in the special case where $k = d$ Theorem 2.2 implies a bound of $\mathbb{E}[\|\hat{M} - M\|_F] \leq \tilde{O}(\sqrt{kd}) = O(d)$ (see Corollary 2.3). Since $\hat{M} - M = \frac{\sqrt{2\log(\frac{1.25}{\delta})}}{\varepsilon}(G + G^\top)$, the matrix $\hat{M} - M$ has independent Gaussian entries with mean zero and variance $\tilde{O}(1)$, and we have from concentration results for Gaussian random matrices (see e.g. Theorem 2.3.6 of [40]) that $\mathbb{E}[\|\hat{M} - M\|_F] = \tilde{\Omega}(d)$, implying that the bound in Theorem 2.2 is tight in this case.

The proof of Theorem 2.2 differs from prior works, including that of [19] which use Davis-Kahan-type theorems [13] and trace inequalities, and instead relies on an interpretation of the Gaussian mechanism as a diffusion process which may be of independent interest. This connection allows us to use sophisticated tools from stochastic differential equations and random matrix theory. We present an outline of the proof in Section 4.

Application to covariance matrix approximation: Plugging $\lambda_i = \sigma_i$ for $i \leq k$ and $\lambda_i = 0$ for $i > k$ into Theorem 2.2, and plugging in concentration bounds for the perturbation to the eigenvalues σ_i, we obtain utility bounds for covariance matrix approximation:

Corollary 2.3 (**Rank-k covariance matrix approximation**). *Let $\varepsilon, \delta > 0$, and given a symmetric matrix $M \in \mathbb{R}^{d\times d}$ with eigenvalues $\sigma_1 \geq \cdots \geq \sigma_d$ and corresponding orthonormal eigenvectors $v_1, \ldots, v_d$. Let G be a matrix with iid $N(0,1)$ entries, and consider the mechanism that outputs $\hat{M} = M + \frac{\sqrt{2\log(\frac{1.25}{\delta})}}{\varepsilon}(G + G^\top)$. Then such a mechanism is (ε, δ)-differentially private. Moreover, for any $k \in [d]$, define $\Sigma_k := \mathrm{diag}(\sigma_1, \ldots, \sigma_k, 0 \ldots, 0)$ and $V = [v_1, \ldots, v_d]$, and define $\hat{\sigma}_1 \geq \cdots \geq \hat{\sigma}_d$ to be the eigenvalues of $\hat{M}$ with corresponding orthonormal eigenvectors $\hat{v}_1, \ldots, \hat{v}_d$, and define $\hat{\Sigma}_k := \mathrm{diag}(\hat{\sigma}_1, \ldots, \hat{\sigma}_k, 0 \ldots, 0)$ and $\hat{V} := [\hat{v}_1, \ldots, \hat{v}_d]$. Then if M satisfies Assumption 2.1 for $(M, k, \sigma_1, \varepsilon, \delta)$, and defining $\sigma_{d+1} := 0$, we have $\mathbb{E}\left[\|\hat{V}\hat{\Sigma}_k\hat{V}^\top - V\Sigma_k V^\top\|_F\right] \leq O\left(\sqrt{kd} \times \frac{\sigma_k}{\sigma_k - \sigma_{k+1}}\right)\frac{\log^{\frac{1}{2}}(\frac{1}{\delta})}{\varepsilon}$.*

The proof appears in Appendix G. If $\sigma_k - \sigma_{k+1} = \Omega(\sigma_k)$, then Corollary 2.3 implies that $\mathbb{E}\left[\|\hat{V}\hat{\Sigma}_k\hat{V}^\top - V\Sigma_k V^\top\|_F\right] \leq O\left(\sqrt{kd}\frac{\log^{\frac{1}{2}}(\frac{1}{\delta})}{\varepsilon}\right)$. Thus, for matrices M with eigenvalues satisfying Assumption 2.1 and where $\sigma_k - \sigma_{k+1} = \Omega(\sigma_k)$, Corollary 2.3 improves by a factor of $\sqrt{k}$ on the bound in Theorem 7 of [19] which says $\|\hat{V}\hat{\Sigma}_k\hat{V}^\top - M\|_F - \|V\Sigma_k V^\top - M\|_F = \tilde{O}(k\sqrt{d})$ w.h.p.. This is because an upper bound on $\|\hat{V}\hat{\Sigma}_k\hat{V}^\top - V\Sigma_k V^\top\|_F$ implies an upper bound on $\|\hat{V}\hat{\Sigma}_k\hat{V}^\top - M\|_F - \|V\Sigma_k V^\top - M\|_F$ by the triangle inequality. On the other hand, while their result does not require a bound on the gaps in the eigenvalue of M and bounds their utility w.h.p., our Corollary 2.4 requires a bound on the gaps of the top $k+1$ eigenvalues of M and bounds the expected utility $\mathbb{E}[\|\hat{V}\hat{\Sigma}_k\hat{V}^\top - V\Sigma_k V^\top\|_F]$.

Application to subspace recovery: Plugging in $\lambda_1 = \cdots = \lambda_k = 1$ and $\lambda_{k+1} = \cdots = \lambda_d = 0$, the post-processing step in Theorem 2.2 outputs a projection matrix, and we obtain utility bounds for the subspace recovery problem.

Corollary 2.4 (**Subspace recovery**). *Let $\varepsilon, \delta > 0$, and given a symmetric matrix $M \in \mathbb{R}^{d\times d}$ with eigenvalues $\sigma_1 \geq \cdots \geq \sigma_d$ and corresponding orthonormal eigenvectors $v_1, \ldots, v_d$. Let G be a matrix with iid $N(0,1)$ entries, and consider the mechanism that outputs $\hat{M} = M + \frac{\sqrt{2\log(\frac{1.25}{\delta})}}{\varepsilon}(G + G^\top)$. Then such a mechanism is (ε, δ)-differentially private. Moreover, for any $k \in [d]$, define the $d \times k$ matrices $V_k = [v_1, \ldots, v_k]$ and $\hat{V}_k = [\hat{v}_1, \ldots, \hat{v}_k]$, where $\hat{\sigma}_1 \geq \cdots \geq \hat{\sigma}_d$ denote the eigenvalues of $\hat{M}$ with corresponding orthonormal eigenvectors $\hat{v}_1, \ldots, \hat{v}_d$. Then if M satisfies Assumption 2.1 for $(M, k, 2, \varepsilon, \delta)$, we have $\mathbb{E}\left[\|\hat{V}_k\hat{V}_k^\top - V_k V_k^\top\|_F\right] \leq O\left(\frac{\sqrt{kd}}{\sigma_k - \sigma_{k+1}} \times \frac{\log^{\frac{1}{2}}(\frac{1}{\delta})}{\varepsilon}\right)$. Moreover, if we also have that $\sigma_i - \sigma_{i+1} \geq \Omega(\sigma_k - \sigma_{k+1})$ for all $i \leq k$, then $\mathbb{E}\left[\|\hat{V}_k\hat{V}_k^\top - V_k V_k^\top\|_F\right] \leq O\left(\frac{\sqrt{d}}{\sigma_k - \sigma_{k+1}} \times \frac{\log^{\frac{1}{2}}(\frac{1}{\delta})}{\varepsilon}\right)$.*

The proof appears in Appendix H. For matrices M satisfying Assumption 2.1, the first inequality of Corollary 2.4 recovers (in expectation) the Frobenius-norm utility bound implied by Theorem 6 of [19], which states that $\|\hat{V}_k\hat{V}_k^\top - V_k V_k^\top\|_F \leq O\left(\frac{\sqrt{kd}}{\sigma_k - \sigma_{k+1}} \times \frac{\log^{\frac{1}{2}}(\frac{1}{\delta})}{\varepsilon}\right)$ w.h.p. Moreover, for many input matrices M with spectral profiles $\sigma_1 \geq \cdots \geq \sigma_d$ satisfying Assumption 2.1, Theorem 2.2 implies stronger bounds than those in [19] for the subspace recovery problem. For instance, if we also have that $\sigma_i - \sigma_{i+1} \geq \Omega(\sigma_k - \sigma_{k+1})$ for all $i \leq k$, the bound given in the second inequality of Corollary 2.4 improves on the bound of [19] by a factor of $\sqrt{k}$. On the other hand, while their result only requires that $\sigma_k - \sigma_{k+1} \geq \sqrt{d}$ and bounds the Frobenius distance $\|\hat{V}_k\hat{V}_k^\top - V_k V_k^\top\|_F$ w.h.p., our Corollary 2.4 requires a bound on the gaps of the top $k+1$ eigenvalues of M and bounds the expected Frobenius distance $\mathbb{E}[\|\hat{V}_k\hat{V}_k^\top - V_k V_k^\top\|_F]$.

3 Preliminaries

Brownian motion and stochastic calculus. A Brownian motion $W(t)$ in $\mathbb{R}$ is a continuous process that has stationary independent increments (see e.g., [34]). In a multi-dimensional Brownian motion, each coordinate is an independent and identical Brownian motion. The filtration $\mathcal{F}_t$ generated by $W(t)$ is defined as $\sigma\left(\cup_{s\leq t}\sigma(W(s))\right)$, where $\sigma(\Omega)$ is the σ-algebra generated by Ω. $W(t)$ is a martingale with respect to $\mathcal{F}_t$.

Definition 3.1 (**Itô Integral**). *Let $W(t)$ be a Brownian motion for $t \geq 0$, let $\mathcal{F}_t$ be the filtration generated by $W(t)$, and let $z(t) : \mathcal{F}_t \to \mathbb{R}$ be a stochastic process adapted to $\mathcal{F}_t$. The Itô integral is defined as $\int_0^T z(t)\mathrm{d}W(t) := \lim_{\omega\to 0}\sum_{i=1}^{\frac{T}{\omega}} z(i\omega)\times[W((i+1)\omega) - W(i\omega)]$.*

Lemma 3.1 (**Itô's Lemma, special case; Theorem 3.7.2 of [31]**). *Let $f : \mathbb{R}^n \to \mathbb{R}$ be any twice-differentiable function. Let $W(t) \in \mathbb{R}^n$, be a Brownian motion. Then $\mathrm{d}f(W(t)) = \nabla f(W(t))^\top \mathrm{d}W(t) + \frac{1}{2}\Delta f(W(t))\mathrm{d}t$, where $\Delta f := \sum_{i=1}^n \frac{\partial^2}{\partial x_i \partial x_i}$ is the Laplacian.*

Lemma 3.2 (**Itô Isometry (one-dimension); Lemma 3.1.5 of [35]**). *Let $W(t)$ be a Brownian motion for $t \geq 0$, let $\mathcal{F}_t$ be the filtration generated by $W(t)$, and $z(t) : \mathcal{F}_t \to \mathbb{R}$ a stochastic process adapted to $\mathcal{F}_t$. Then we have $\mathbb{E}\left[\left(\int_0^T z(t)\mathrm{d}W(t)\right)^2\right] = \mathbb{E}\left[\int_0^T z(t)^2\mathrm{d}t\right]$.*

Corollary 3.3 (**Itô Isometry (matrix-valued stochastic processes)**). *Let $W(t)$ be a $d\times d$-dimensional Brownian motion for $t \geq 0$, let $\mathcal{F}_t$ be the filtration generated by $W(t)$, and let $Z(t) : \mathcal{F}_t \to \mathbb{R}^{d\times d}$ be a stochastic process adapted to $\mathcal{F}_t$. Then $\mathbb{E}\left[\left\|\int_0^T Z(t)\odot \mathrm{d}W(t)\right\|_F^2\right] = \mathbb{E}\left[\int_0^T \|Z(t)\|_F^2\mathrm{d}t\right]$, where the Hadamard product $\odot$ denotes entry-wise multiplication*

Dyson Brownian motion. Let $W(t) \in \mathbb{R}^{d\times d}$ be a matrix where each entry is an independent standard Brownian motion with distribution $N(0, tI_d)$ at time t, and let $B(t) = W(t) + W^\top(t)$. Define the symmetric-matrix valued stochastic process $\Phi(t)$ as follows:

$$\Phi(t) := M + B(t) \qquad \forall t \geq 0. \tag{1}$$

The process $\Phi(t)$ is referred to as (matrix) Dyson Brownian motion. At every time $t > 0$ the eigenvalues $\gamma_1(t), \ldots, \gamma_d(t)$ of $\Phi(t)$ are distinct with probability 1, and (1) induces a stochastic process on the eigenvalues and eigenvectors. The process on the eigenvalues and eigenvectors can be expressed via the following diffusion equations. The eigenvalue diffusion process, which is also referred to as (eigenvalue) "Dyson Brownian motion", is defined by the stochastic differential equation (2). The (eigenvalue) Dyson Brownian motion is an Itô diffusion and can be expressed can be expressed by the following stochastic differential equation [20]:

$$\mathrm{d}\gamma_i(t) = \mathrm{d}B_{ii}(t) + \sum_{j\neq i}\frac{1}{\gamma_i(t)-\gamma_j(t)}\mathrm{d}t \qquad \forall i \in [d], t > 0. \tag{2}$$

The corresponding eigenvector process $v_1(t), \ldots, v_d(t)$, referred to as the Dyson vector flow, is also an Itô diffusion and, conditional on the eigenvalue process (2), can be expressed by the following stochastic differential equation (see e.g., [3]):

$$\mathrm{d}u_i(t) = \sum_{j\neq i}\frac{\mathrm{d}B_{ij}(t)}{\gamma_i(t)-\gamma_j(t)}u_j(t) - \frac{1}{2}\sum_{j\neq i}\frac{\mathrm{d}t}{(\gamma_i(t)-\gamma_j(t))^2}u_i(t) \qquad \forall i \in [d], t > 0. \tag{3}$$

Eigenvalue bounds. The following two Lemmas will help us bound the gaps in the eigenvalues of the Dyson Brownian motion:

Lemma 3.4 (Theorem 4.4.5 of [44], special case [1]). *Let $W \in \mathbb{R}^{d\times d}$ with i.i.d. $N(0,1)$ entries. Then $\mathbb{P}(\|W\|_2 > 2(\sqrt{d}+s) < 2e^{-s^2}$ for any $s > 0$.*

Lemma 3.5 (**Weyl's Inequality; [6]**). *If $A, B \in \mathbb{R}^{d\times d}$ are two symmetric matrices, and denoting the i'th-largest eigenvalue of any symmetric matrix M by $\sigma_i(M)$, we have $\sigma_i(A) + \sigma_d(B) \leq \sigma_i(A+B) \leq \sigma_i(A) + \sigma_1(B)$.*

[1] The theorem is stated for sub-Gaussian entries in terms of a constant C; this constant is $C = 2$ in the special case where the entries are $N(0,1)$ Gaussian.

4 Proof of Theorem 2.2

We give an overview of the proof of Theorem 2.2, along with the main technical lemmas used to prove this result. Section 4.1 outlines the different steps in our proof. In Steps 1 and 2 we construct the matrix-valued diffusion used in our proof. Steps 3,4, and 5 present the main technical lemmas, and in step 6 we explain how to complete the proof. The statements of the lemmas and the highlights of their proofs, are given in Sections 4.2, 4.3, 4.4. In section 4.5 we explain how to complete the proof. The full proofs are deferred to the appendix.

4.1 Outline of proof

1. **Step 1: Expressing the Gaussian Mechanism as a Dyson Brownian Motion.** To obtain our utility bound, we view the Gaussian mechanism as a matrix-valued Brownian motion (1) initialized at the input matrix M: $\Phi(t) := M + B(t) \qquad \forall t \geq 0$. If we run this Brownian motion for time $T = \sqrt{2\log(\frac{1.25}{\delta})}/\varepsilon$ we have that $\Phi(T) = (\sqrt{2\log(\frac{1.25}{\delta})}/\varepsilon)(G+G^\top)$, recovering the output of the Gaussian mechanism. In other words, the input to the Gaussian mechanism is $M = \Phi(0)$, and the output is $\hat{M} = \Phi(T)$.
2. **Step 2: Expressing the post-processed mechanism as a matrix diffusion $\Psi(t)$.** Our goal is to bound $\|\hat{V}\Lambda\hat{V}^\top - V\Lambda V^\top\|_F$, where $M = V\Sigma V^\top$ and $\hat{M} = \hat{V}\hat{\Sigma}\hat{V}^\top$ are spectral decompositions of M and $\hat{M}$. To bound the error $\|\hat{V}\Lambda\hat{V}^\top - V\Lambda V^\top\|_F$ we will define a stochastic process $\Psi(t)$ such that $\Psi(0) = V\Lambda V^\top$ and $\Psi(t) = \hat{V}\Lambda\hat{V}^\top$, and then bound the Frobenius distance $\|\Psi(T) - \Psi(0)\|_F$ by integrating the (stochastic) derivative of $\Psi(t)$ over the time interval $[0, T]$.

   Towards this end, at every time t, let $\Phi(t) = U(t)\Gamma(t)U(t)^\top$ be a spectral decomposition of the symmetric matrix $\Phi(t)$, where $\Gamma(t)$ is a diagonal matrix with diagonal entries $\gamma_1(t) \geq \cdots \geq \gamma_d(t)$ that are the eigenvalues of $\Phi(t)$, and $U(t) = [u_1(t), \cdots, u_d(t)]$ is a $d \times d$ orthogonal matrix whose columns $u_1(t), \cdots, u_d(t)$ are an orthonormal basis of eigenvectors of $\Phi(t)$. At every time t, define $\Psi(t)$ to be the symmetric matrix with eigenvalues Λ and eigenvectors given by the columns of $U(t)$: $\Psi(t) := U(t)\Lambda U(t)^\top \ \forall t \in [0, T]$.
3. **Step 3: Computing the stochastic derivative $\mathrm{d}\Psi(t)$.** To bound the expected squared Frobenius distance $\mathbb{E}[\|\Psi(T) - \Psi(0)\|_F^2]$, we first compute the stochastic derivative $\mathrm{d}\Psi(t)$ of the matrix diffusion $\Psi(T)$ (Lemma 4.2).
4. **Step 4: Bounding the eigenvalue gaps.** The equation for the derivative $\mathrm{d}\Psi(t)$ includes terms with magnitude proportional to the inverse of the eigenvalue gaps $\Delta_{ij}(t) := \gamma_i(t) - \gamma_j(t)$ for each $i, j \in [d]$, which evolve over time. In order to bound these terms, we use Weyl's inequality (Lemma 3.5) to show that w.h.p. the gaps in the top $k+1$ eigenvalues $\Delta_{ij}(t)$ satisfy $\Delta_{ij}(t) \geq \Omega(\sigma_i - \sigma_j)$ for every time $t \in [0, T]$, provided that the initial gaps are sufficiently large (Lemma 4.4).
5. **Step 5: Integrating the stochastic differential equation.** Next, we express the expected squared Frobenius distance $\mathbb{E}[\|\Psi(T) - \Psi(0)\|_F^2]$ as an integral $\|\Psi(T) - \Psi(0)\|_F^2] = \mathbb{E}\left[\left\|\int_0^T \mathrm{d}\Psi(t)\right\|_F^2\right]$. We then apply the Itô Isometry for matrices (Corollary 3.3) to obtain a formula for this integral. Roughly speaking, the formula we obtain (Lemma 4.5) is

$$\mathbb{E}\left[\|\Psi(T) - \Psi(0)\|_F^2\right] \approx \int_0^T \mathbb{E}\left[\sum_{i=1}^d \sum_{j\neq i} \frac{(\lambda_i - \lambda_j)^2}{\Delta_{ij}^2(t)}\right] \mathrm{d}t + T\int_0^T \mathbb{E}\left[\sum_{i=1}^d \left(\sum_{j\neq i} \frac{\lambda_i - \lambda_j}{\Delta_{ij}^2(t)}\right)^2\right] \mathrm{d}t \tag{4}$$

6. **Step 6: Completing the proof.** Plugging the bound $\Delta_{ij}(t) \geq \Omega(\sigma_i - \sigma_j)$ into (4), and noting that the first term on the RHS of (4) is at least as large as the second term since $\sigma_i - \sigma_j \geq \sqrt{d}$, we obtain the bound in Theorem 2.2.

4.2 Step 3: Computing the stochastic derivative $\mathrm{d}\Psi(t)$

$\Psi(t)$ is itself a matrix-valued diffusion. We use the eigenvalue and eigenvector dynamics 2 and 3 together with Itô's Lemma (Lemma 3.1) to compute the Itô derivative of this diffusion. Towards this end, we first decompose the matrix $\Psi(t)$ as a sum of its eigenvectors: $\Psi(t) = \sum_{i=1}^d \lambda_i u_i(t) u_i^\top(t)$. Thus, we have

$$\mathrm{d}\Psi(t) = \sum_{i=1}^d \lambda_i \mathrm{d}(u_i(t)u_i^\top(t)). \tag{5}$$

We begin by computing the stochastic derivative $\mathrm{d}(u_i(t)u_i^\top(t))$ for each $i \in [d]$, by applying the formula for the derivative of $u_i(t)$ in (3), together with Ito's Lemma (Lemma 3.1):

Lemma 4.1 (Stochastic derivative of $u_i(t)u_j^\top(t)$). *For all $t \in [0,T]$, $\mathrm{d}(u_i(t)u_i^\top(t)) = \sum_{j\neq i} \frac{\mathrm{d}B_{ij}(t)}{\gamma_i(t)-\gamma_j(t)}(u_i(t)u_j^\top(t) + u_j(t)u_i^\top(t)) + \sum_{j\neq i} \frac{\mathrm{d}t}{(\gamma_i(t)-\gamma_j(t))^2}(u_i(t)u_i^\top(t) - u_j(t)u_j^\top(t))$.*

The proof is in Appendix A. Plugging Lemma 4.1 into (5), we get an expression for $\mathrm{d}\Psi(t)$:

Lemma 4.2 (Stochastic derivative of $\Psi(t)$; see Appendix B for proof). *For all $t \in [0,T]$ we have that $\mathrm{d}\Psi(t) = \frac{1}{2}\sum_{i=1}^d \sum_{j\neq i}(\lambda_i - \lambda_j)\frac{\mathrm{d}B_{ij}(t)}{\gamma_i(t)-\gamma_j(t)}(u_i(t)u_j^\top(t) + u_j(t)u_i^\top(t)) + \sum_{i=1}^d \sum_{j\neq i}(\lambda_i - \lambda_j)\frac{\mathrm{d}t}{(\gamma_i(t)-\gamma_j(t))^2}u_i(t)u_i^\top(t)$.*

4.3 Step 4: Bounding the eigenvalue gaps

The derivative in Lemma 4.2 contains terms with magnitude proportional to the inverse of the eigenvalue gaps $\Delta_{ij}(t) := \gamma_i(t) - \gamma_j(t)$. To bound these terms, we would like to show that $\inf_{t\in[0,T]} \Delta_{ij}(t) \geq \Omega(\sigma_i - \sigma_j)$ for each $i < j \leq k+1$ with high probability. Towards this end, we first apply the spectral norm concentration bound for Gaussian random matrices (Lemma 3.4), which provides a high-probability bound for $\|B(t)\|_2$ at any time t, together with Doob's submartingale inequality, to show that the spectral norm of the matrix-valued Brownian motion $B(t)$ does not exceed $T\sqrt{d}$ at any time $t \in [0,T]$ w.h.p.:

Lemma 4.3 (Spectral norm bound). *For every $T > 0$, we have, $\mathbb{P}\left(\sup_{t\in[0,T]} \|B(t)\|_2 > 2T\sqrt{d} + \alpha)\right) \leq 2\sqrt{\pi}e^{-\frac{1}{8}\frac{\alpha^2}{T^2}}$.*

The proof appears in Appendix C. Next, we use Lemma 4.3 to bound the eigenvalue gaps:

Lemma 4.4 (Eigenvalue gap bound). *Whenever $\gamma_i(0) - \gamma_{i+1}(0) \geq 4T\sqrt{d}$ for every $i \in S$ and $T > 0$ and some subset $S \subset [d-1]$, we have $\mathbb{P}\left(\bigcup_{i\in S}\left\{\inf_{t\in[0,T]} \gamma_i(t) - \gamma_{i+1}(t) < \frac{1}{2}(\gamma_i(0) - \gamma_{i+1}(0)) - \alpha)\right\}\right) \leq 2\sqrt{\pi}e^{-\frac{1}{32}\alpha^2}$.*

To prove Lemma 4.4, we plug Lemma 4.3 into Weyl's Inequality (Lemma 3.5), to show that

$$\gamma_i(t) - \gamma_{i+1}(t) \geq \sigma_i - \sigma_{i+1} - \|B(t)\|_2 \geq \Omega(\sigma_i - \sigma_{i+1} - T\sqrt{d}) \geq \Omega(\sigma_i - \sigma_{i+1}),$$

with high probability for each $i \leq k$ (Lemma 4.4). The last inequality holds since Assumption 2.1 ensures $\sigma_i - \sigma_{i+1} \geq \frac{1}{2}T\sqrt{d}$ for $i \leq k$. The full proof is in Appendix D.

4.4 Step 5: Integrating the stochastic differential equation

Next, we would like to integrate the derivative $\mathrm{d}\Psi(t)$ to obtain an expression for $\mathbb{E}[\|\Psi(T) - \Psi(0)\|_F^2]$, and to then plug in our high-probability bounds (Lemma 4.4) for the gaps $\Delta_{ij}(t)$. To allow us to later plug in these high-probability bounds after we integrate and take the expectation, we define a new diffusion process $Z_\eta(t)$ which has nearly the same stochastic differential equation as 4.2, except that each eigenvalue gap $\Delta_{ij}(t)$ is not permitted to become smaller than the value $\eta_{ij} = \frac{1}{4}(\sigma_i - \max(\sigma_j, \sigma_{k+1}))$ for each $i < j$.

Towards this end, fix any $\eta \in \mathbb{R}^{d\times d}$, define the following matrix-valued Itô diffusion $Z_\eta(t)$ via its Itô derivative $\mathrm{d}Z_\eta(t)$:

$$\begin{aligned}\mathrm{d}Z_\eta(t) \quad &:= \tfrac{1}{2}\textstyle\sum_{i=1}^d \sum_{j\neq i} |\lambda_i - \lambda_j| \frac{\mathrm{d}B_{ij}(t)}{\max(|\Delta_{ij}(t)|, \eta_{ij})}(u_i(t)u_j^\top(t) + u_j(t)u_i^\top(t)) \\ &\quad + \textstyle\sum_{i=1}^d \sum_{j\neq i}(\lambda_i - \lambda_j)\frac{\mathrm{d}t}{\max(\Delta_{ij}^2(t), \eta_{ij}^2)}u_i(t)u_i^\top(t), \end{aligned} \tag{6}$$

with initial condition $Z_\eta(0) := \Psi(0)$. Thus, $Z_\eta(t) = \Psi(0) + \int_0^t \mathrm{d}Z_\eta(s)$ for all $t \geq 0$. We then integrate the $\mathrm{d}Z_\eta(t)$ over the time interval $[0,T]$, and apply the Itô isometry (Corollary 3.3) to obtain an expression for the Frobenius norm of this integral:

Lemma 4.5 (Frobenius distance integral). *For any $T > 0$, $\mathbb{E}\left[\|Z_\eta(T) - Z_\eta(0)\|_F^2\right] =$*

$$2\int_0^T \mathbb{E}\left[\sum_{i=1}^d \sum_{j\neq i} \frac{(\lambda_i - \lambda_j)^2}{\max(\Delta_{ij}^2(t), \eta_{ij}^2)}\mathrm{d}t\right] + T\int_0^T \mathbb{E}\left[\sum_{i=1}^d \left(\sum_{j\neq i} \frac{\lambda_i - \lambda_j}{\max(\Delta_{ij}^2(t), \eta_{ij}^2)}\right)^2\right]\mathrm{d}t.$$

To prove Lemma 4.5, we write $\mathbb{E}[\|Z(T) - Z(0)\|_F^2] =$

$$\frac{1}{2}\mathbb{E}\left[\left\|\int_0^T \left(\sum_{i=1}^d \sum_{j\neq i} \frac{|\lambda_i - \lambda_j|}{\max(|\Delta_{ij}(t)|, \eta)}(u_i(t)u_j^\top(t) + u_j(t)u_i^\top(t))\right) \odot \mathrm{d}B(t)\right\|_F^2\right]$$
$$+ \mathbb{E}\left[\left\|\int_0^T \sum_{i=1}^d \sum_{j\neq i} \frac{\lambda_i - \lambda_j}{\max(\Delta_{ij}^2(t), \eta^2)} u_i(t)u_i^\top(t)\mathrm{d}t\right\|_F^2\right]. \tag{7}$$

To compute the first term on the RHS, we apply the Itô isometry (Corollary 3.3). The key step is to notice that, since iid Gaussian matrices are invariant (in distribution) to orthogonal transformations, we have that $U(t)\mathrm{d}B(t)U^\top(t)$ has the same distribution as $\mathrm{d}B(t)$. Moreover, since $u_i(t), u_j(t), \Delta_{ij}(t)$ are all diffusion processes w.r.t. the filtration $\mathcal{F}_t$ generated by the Brownian motion $B(t)$, $\mathrm{d}B(t)$ is independent of the value of the eigenvectors $u_i(\tau), u_j(\tau)$ and eigenvalue gaps $\Delta_{ij}(\tau)$ at all times $\tau \leq t$ up to the present time t. Thus, we can replace the term $\varphi(t) \odot \mathrm{d}B(t)$ inside the integral, where $\varphi(t) := \sum_{i=1}^d \sum_{j\neq i} \frac{|\lambda_i - \lambda_j|}{\max(|\Delta_{ij}(t)|, \eta)}(u_i(t)u_j^\top(t) + u_j(t)u_i^\top(t))$, with $U(t)\varphi(t)U^\top(t) \odot \mathrm{d}B(t) = \left(\sum_{i=1}^d \sum_{j\neq i} \frac{|\lambda_i - \lambda_j|}{\max(|\Delta_{ij}(t)|, \eta_{ij})} e_j^\top e_i^\top\right) \odot \mathrm{d}B(t)$, where $e_1, \ldots, e_d$ are the standard basis vectors. This allows us to apply the Itô isometry to each entry of the matrix individually to compute the first term on the RHS. To compute the second term on the RHS, we use the Cauchy-Schwarz inequality. The full proof appears in Appendix E.

4.5 Step 6: Completing the proof

To complete the proof, we plug in the high-probability bounds on the eigenvalue gaps from Section 4.3 into Lemma 4.5. Since by Lemma 4.4 $\Delta_{ij}(t) \geq \frac{1}{2}(\sigma_i - \sigma_j)$ w.h.p. for each $i, j \leq k+1$, and $\eta_{ij} = \frac{1}{4}(\sigma_i - \max(\sigma_j, \sigma_{k+1}))$, we must also have that $Z_\eta(t) = \Psi(t)$ for all $t \in [0, T]$ w.h.p. Plugging in the high-probability bounds $\Delta_{ij}(t) \geq \frac{1}{2}(\sigma_i - \sigma_j)$ for each $i, j \geq k+1$, and noting that $\lambda_i - \lambda_j = 0$ for all $i, j > k$, we get that

$$\mathbb{E}\left[\|\hat{V}\Lambda\hat{V}^\top - V\Lambda V^\top\|_F^2\right] = \mathbb{E}\left[\|\Psi(T) - \Psi(0)\|_F^2\right]$$
$$\leq 2\int_0^T \mathbb{E}\left[\sum_{i=1}^d \sum_{j\neq i} \frac{(\lambda_i - \lambda_j)^2}{(\sigma_i - \sigma_j)^2}\right]\mathrm{d}t + T\int_0^T \mathbb{E}\left[\sum_{i=1}^d \left(\sum_{j\neq i} \frac{\lambda_i - \lambda_j}{(\sigma_i - \sigma_j)^2}\right)^2\right]\mathrm{d}t$$
$$\leq T\sum_{i=1}^k \sum_{j=i+1}^d \frac{(\lambda_i - \lambda_j)^2}{(\sigma_i - \max(\sigma_j, \sigma_{k+1}))^2} + T^2 \sum_{i=1}^k \left(\sum_{j=i+1}^d \frac{\lambda_i - \lambda_j}{(\sigma_i - \max(\sigma_j, \sigma_{k+1}))^2}\right)^2. \tag{8}$$

Since $(\sigma_i - \max(\sigma_j, \sigma_{k+1}) \geq \Omega(\sqrt{d})$ for all $i \leq k$ and $j \in [d]$, we can use the Cauchy-Schwarz inequality to show that the second term is (up to a factor of T) smaller than the first term: $\sum_{i=1}^k \left(\sum_{j=i+1}^d \frac{\lambda_i - \lambda_j}{(\sigma_i - \max(\sigma_j . \sigma_{k+1}))^2}\right)^2 \leq \sum_{i=1}^k \sum_{j=i+1}^d \frac{(\lambda_i - \lambda_j)^2}{(\sigma_i - \max(\sigma_j, \sigma_{k+1})}$. Plugging $T = \frac{\sqrt{2\log(\frac{1.25}{\delta})}}{\varepsilon}$ into (8), we obtain the bound in Theorem 2.2. For the full proof of Theorem 2.2, see Appendix F.

5 Conclusion and Future Work

We present a new analysis of the Gaussian mechanism for a large class of symmetric matrix approximation problems, by viewing this mechanism as a Dyson Brownian motion initialized at the input matrix M. This viewpoint allows us to leverage the stochastic differential equations which govern the evolution of the eigenvalues and eigenvectors of Dyson Brownian motion to obtain new utility bounds for the Gaussian mechanism. To obtain our utility bounds, we show that the gaps in the eigenvalues of the Dyson Brownian motion stay at least as large as the initial gap sizes (up to a constant factor), as long as the initial gaps in the top $k+1$ eigenvalues of the input matrix are $\geq \Omega(\sqrt{d})$. The main question that remains open is whether for other initial matrices M the gaps in the eigenvalues of Dyson Brownian motion grow over time as well, and, if true, whether one can use this fact to obtain similar utility bounds for input matrices M with eigenvalue gaps that are smaller than $\sqrt{d}$. Finally, this paper analyzes a mechanism in differential privacy, which has many implications for preserving sensitive information of individuals. Thus, we believe our work will have positive societal impacts and do not foresee any negative impacts to society.

References

[1] Dimitris Achlioptas and Frank McSherry. Fast computation of low-rank matrix approximations. *Journal of the ACM (JACM)*, 54(2):1–19, 2007.

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

[3] Greg W Anderson, Alice Guionnet, and Ofer Zeitouni. *An introduction to random matrices*. Number 118. Cambridge university press, 2010.

[4] Ludwig Arnold. On wigner's semicircle law for the eigenvalues of random matrices. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 19(3):191–198, 1971.

[5] James Bennett and Stan Lanning. The Netflix Prize. In *Proceedings of KDD cup and workshop*, volume 2007, page 35. New York, NY, USA., 2007.

[6] Rajendra Bhatia. *Matrix analysis*, volume 169. Springer Science & Business Media, 2013.

[7] Avrim Blum, Cynthia Dwork, Frank McSherry, and Kobbi Nissim. Practical privacy: the sulq framework. In *Proceedings of the twenty-fourth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, pages 128–138, 2005.

[8] Avrim Blum, John Hopcroft, and Ravindran Kannan. *Foundations of data science*. Cambridge University Press, 2020.

[9] Paul Bourgade. Extreme gaps between eigenvalues of wigner matrices. *Journal of the European Mathematical Society*, 2021.

[10] Emmanuel J Candes and Yaniv Plan. Matrix completion with noise. *Proceedings of the IEEE*, 98(6):925–936, 2010.

[11] Kamalika Chaudhuri, Anand Sarwate, and Kaushik Sinha. Near-optimal differentially private principal components. *Advances in neural information processing systems*, 25, 2012.

[12] Mark A Davenport and Justin Romberg. An overview of low-rank matrix recovery from incomplete observations. *IEEE Journal of Selected Topics in Signal Processing*, 10(4):608–622, 2016.

[13] Chandler Davis and William Morton Kahan. The rotation of eigenvectors by a perturbation. iii. *SIAM Journal on Numerical Analysis*, 7(1):1–46, 1970.

[14] Dheeru Dua and Casey Graff. UCI machine learning repository, 2017.

[15] Cynthia Dwork. Differential Privacy. In *ICALP (2)*, volume 4052 of *Lecture Notes in Computer Science*, pages 1–12. Springer, 2006.

[16] Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor. Our data, ourselves: Privacy via distributed noise generation. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 486–503. Springer, 2006.

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

[18] Cynthia Dwork and Aaron Roth. The Algorithmic Foundations of Differential Privacy. *Foundations and Trends® in Theoretical Computer Science*, 9(3–4):211–407, 2014.

[19] Cynthia Dwork, Kunal Talwar, Abhradeep Thakurta, and Li Zhang. Analyze gauss: optimal bounds for privacy-preserving principal component analysis. In *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, pages 11–20, 2014.

[20] Freeman J Dyson. A brownian-motion model for the eigenvalues of a random matrix. *Journal of Mathematical Physics*, 3(6):1191–1198, 1962.

[21] László Erdős and Horng-Tzer Yau. Universality of local spectral statistics of random matrices. *Bulletin of the American Mathematical Society*, 49(3):377–414, 2012.

[22] Renjie Feng, Gang Tian, and Dongyi Wei. Small gaps of goe. *Geometric and Functional Analysis*, 29(6):1794–1827, 2019.

[23] Alessio Figalli and Alice Guionnet. Universality in several-matrix models via approximate transport maps. *Acta mathematica*, 217(1):81–176, 2016.

[24] Alon Gonem and Ram Gilad-Bachrach. Smooth sensitivity based approach for differentially private pca. In *Algorithmic Learning Theory*, pages 438–450. PMLR, 2018.

[25] Moritz Hardt and Aaron Roth. Beating randomized response on incoherent matrices. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, pages 1255–1268, 2012.

[26] Moritz Hardt and Aaron Roth. Beyond worst-case analysis in private singular vector computation. In *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*, pages 331–340, 2013.

[27] Jiaoyang Huang, Benjamin Landon, and Horng-Tzer Yau. Bulk universality of sparse random matrices. *Journal of Mathematical Physics*, 56(12):123301, 2015.

[28] Mia Hubert and Sanne Engelen. Robust pca and classification in biosciences. *Bioinformatics*, 20(11):1728–1736, 2004.

[29] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An introduction to statistical learning*, volume 112. Springer, 2013.

[30] Michael Kapralov and Kunal Talwar. On differentially private low rank approximation. In *Proceedings of the twenty-fourth annual ACM-SIAM symposium on Discrete algorithms*, pages 1395–1414. SIAM, 2013.

[31] Gregory F Lawler. Stochastic calculus: An introduction with applications. *American Mathematical Society*, 2010.

[32] Jonathan Leake, Colin S. McSwiggen, and Nisheeth K. Vishnoi. Sampling matrices from Harish-Chandra-Itzykson-Zuber densities with applications to quantum inference and differential privacy. In *STOC*, pages 1384–1397. ACM, 2021.

[33] Jonathan Leake and Nisheeth K. Vishnoi. On the computability of continuous maximum entropy distributions with applications. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, pages 930–943, 2020.

[34] Peter Mörters and Yuval Peres. *Brownian motion*, volume 30. Cambridge University Press, 2010.

[35] Bernt Oksendal. *Stochastic differential equations: an introduction with applications*. Springer Science & Business Media, 2013.

[36] LA Pastur and VA Martchenko. The distribution of eigenvalues in certain sets of random matrices. *Math. USSR-Sbornik*, 1(4):457–483, 1967.

[37] Haipeng Shen and Jianhua Z Huang. Sparse principal component analysis via regularized low rank matrix approximation. *Journal of multivariate analysis*, 99(6):1015–1034, 2008.

[38] Azer P Shikhaliev, Lee C Potter, and Yuejie Chi. Low-rank structured covariance matrix estimation. *IEEE Signal Processing Letters*, 26(5):700–704, 2019.

[39] Jack W Silverstein. The smallest eigenvalue of a large dimensional wishart matrix. *The Annals of Probability*, pages 1364–1368, 1985.

[40] Terence Tao. *Topics in random matrix theory*, volume 132. American Mathematical Soc., 2012.

[41] Terence Tao and Van Vu. Random matrices: universality of local eigenvalue statistics. *Acta mathematica*, 206(1):127–204, 2011.

[42] Terence Tao and Van Vu. Random covariance matrices: Universality of local statistics of eigenvalues. *The Annals of Probability*, 40(3):1285–1315, 2012.

[43] Craig A Tracy and Harold Widom. Level-spacing distributions and the airy kernel. *Communications in Mathematical Physics*, 159(1):151–174, 1994.

[44] Roman Vershynin. *High-dimensional probability: An introduction with applications in data science*, volume 47. Cambridge university press, 2018.

[45] Pierpaolo Vivo, Satya N Majumdar, and Oriol Bohigas. Large deviations of the maximum eigenvalue in wishart random matrices. *Journal of Physics A: Mathematical and Theoretical*, 40(16):4317, 2007.

[46] Thijs Vogels, Sai Praneeth Karimireddy, and Martin Jaggi. Powersgd: Practical low-rank gradient compression for distributed optimization. *Advances in Neural Information Processing Systems*, 32, 2019.

[47] Ke Wang. Random covariance matrices: Universality of local statistics of eigenvalues up to the edge. *Random Matrices: Theory and Applications*, 1(01):1150005, 2012.

[48] John Wishart. The generalised product moment distribution in samples from a normal multivariate population. *Biometrika*, pages 32–52, 1928.

Checklist

The checklist follows the references. Please read the checklist guidelines carefully for information on how to answer these questions. For each question, change the default **[TODO]** to [Yes] , [No] , or [N/A] . You are strongly encouraged to include a **justification to your answer**, either by referencing the appropriate section of your paper or providing a brief inline description. For example:

- Did you include the license to the code and datasets? [N/A]
- Did you include the license to the code and datasets? [N/A]
- Did you include the license to the code and datasets? [N/A]

Please do not modify the questions and only use the provided macros for your answers. Note that the Checklist section does not count towards the page limit. In your paper, please delete this instructions block and only keep the Checklist section heading above along with the questions/answers below.

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes]
   (b) Did you describe the limitations of your work? [Yes]
   (c) Did you discuss any potential negative societal impacts of your work? [Yes] (See Section 5)
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [Yes] (See Section 2 and Assumption 2.1)
   (b) Did you include complete proofs of all theoretical results? [Yes] (See Section 4 and Appendix A, B, C, D, E,F, G,H)
3. If you ran experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [Yes] (See Appendix I)
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes] (See Appendix I)
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [Yes] (See Appendix I)
   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [Yes] (See Appendix I)
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
   (a) If your work uses existing assets, did you cite the creators? [N/A]
   (b) Did you mention the license of the assets? [N/A]
   (c) Did you include any new assets either in the supplemental material or as a URL? [N/A]
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [N/A]
   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A]
5. If you used crowdsourcing or conducted research with human subjects...
   (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]

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

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