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

Oren Mangoubi
Worcester Polytechnic Institute

Nisheeth K. Vishnoi
Yale University

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 $\lambda = (\lambda_1, \ldots, \lambda_d)$ [37, 28, 36]. 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, 45, 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 $\mathbb{P}(\mathcal{A}(M) \in S) \leq e^\varepsilon \mathbb{P}(\mathcal{A}(M') \in S) + \delta$.

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

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 i.i.d. 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 (3) and (4), 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 top $k+1$ eigenvalues of the input matrix are $\geq \Omega(\sqrt{d})$ (Assumption 2.1). Empirically, we observe that Assumption 2.1 is satisfied for covariance matrices of many real-world datasets (see Appendix J), as well as on Wishart random matrices $W = A^\top A$, where A is an $m \times d$ matrix of i.i.d. Gaussian entries, for sufficiently large m (see Appendix I). 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 $i \in [k]$.*

We observe empirically that Assumption 2.1 is satisfied on a number of real-world datasets which were previously used as benchmarks in the differentially private matrix approximation literature [11, 2] (see Appendix J). Assumption 2.1 is also satisfied, for instance, by random Wishart matrices $W = A^\top A$, where A is an $m \times d$ matrix of i.i.d. Gaussian entries, which are a popular model for sample covariance matrices [47]. 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 i.i.d. $N(0,1)$ 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 [39]) 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 (See Appendix K for additional comparison to previous approaches). 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 i.i.d. $N(0,1)$ 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 i.i.d. $N(0,1)$ 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 $\|\hat{V}_k\hat{V}_k^\top - V_kV_k^\top\|_F \leq O\left(\frac{\sqrt{kd}}{\sigma_k-\sigma_{k+1}} \times \frac{\log^{\frac{1}{2}}(\frac{1}{\delta})}{\varepsilon}\right)$ 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 $\|\hat{V}_k\hat{V}_k^\top - V_kV_k^\top\|_F$ 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 $\mathbb{E}[\|\hat{V}_k\hat{V}_k^\top - V_kV_k^\top\|_F]$.

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, integral form with no drift; Theorem 3.7.1 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, and let $X(t) \in \mathbb{R}^n$ be an Itô diffusion process with mean zero defined by the following stochastic differential equation:*

$$\mathrm{d}X_j(t) = \sum_{i=1}^{d} R_{ij}(t)\mathrm{d}W_i(t), \tag{1}$$

for some Itô diffusion $R(t) \in \mathbb{R}^{n\times n}$ adapted to the filtration generated by the Brownian motion $W(t)$. Then for any $T \geq 0$,

$$\begin{aligned} f(X(T)) - f(X(0)) = &\int_0^T \sum_{i=1}^{n}\sum_{\ell=1}^{n}\left(\frac{\partial}{\partial X_\ell} f(X(t))\right) R_{i\ell}(t)\mathrm{d}W_i(t) \\ &+ \frac{1}{2}\int_0^T \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{\ell=1}^{n}\left(\frac{\partial^2}{\partial X_j \partial X_\ell} f(X(t))\right) R_{ij}(t)R_{i\ell}(t)\mathrm{d}t. \end{aligned}$$

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:

$$\Phi(t) := M + B(t) \qquad \forall t \geq 0. \tag{2}$$

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 (2) 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 (3). The (eigenvalue) Dyson Brownian motion is an Itô diffusion and can be expressed can be expressed by the following stochastic differential equation [20]:

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

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 (3), can be expressed by the following stochastic differential equation (see e.g., [3]):

$$\mathrm{d}u_i(t) = \sum_{j \neq i} \frac{\mathrm{d}B_{ij}(t)}{\gamma_i(t) - \gamma_j(t)} u_j(t) - \frac{1}{2} \sum_{j \neq i} \frac{\mathrm{d}t}{(\gamma_i(t) - \gamma_j(t))^2} u_i(t) \qquad \forall i \in [d], t > 0. \quad (4)$$

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

Lemma 3.2 (Theorem 4.4.5 of [43], 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 3.3 (**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 $\sigma_i(A) + \sigma_d(B) \leq \sigma_i(A + B) \leq \sigma_i(A) + \sigma_1(B)$.*

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 (2) 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 3.3) 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 $t \in [0, T]$ (Lemma 4.4),

[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.

provided that the initial gaps are sufficiently large (Assumption 2.1) (See Appendix L for a discussion on why we need this assumption for our proof to work).

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 Itô's Lemma (Lemma 3.1) to obtain a formula for this integral. Roughly speaking, the formula we obtain (Lemma 4.5) is

$$\mathbb{E}\left[\|\Psi(T)-\Psi(0)\|_F^2\right] \approx \int_0^T \mathbb{E}\left[\sum_{i=1}^d \sum_{j\neq i} \frac{(\lambda_i-\lambda_j)^2}{\Delta_{ij}^2(t)}\right] \mathrm{d}t + T\int_0^T \mathbb{E}\left[\sum_{i=1}^d \left(\sum_{j\neq i} \frac{\lambda_i-\lambda_j}{\Delta_{ij}^2(t)}\right)^2\right] \mathrm{d}t \tag{5}$$

6. **Step 6: Completing the proof.** Plugging the bound $\Delta_{ij}(t) \geq \Omega(\sigma_i - \sigma_j)$ into (5), and noting that the first term on the r.h.s. of (5) 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 3 and 4 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

$$\mathrm{d}\Psi(t) = \sum_{i=1}^d \lambda_i \mathrm{d}(u_i(t)u_i^\top(t)). \tag{6}$$

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 (4), together with Itô'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 (6), 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 3.2), 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 3.3), 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 $\mathrm{d}Z_\eta(t)$:

$$\begin{aligned}\mathrm{d}Z_\eta(t) \quad &:= \tfrac{1}{2}\textstyle\sum_{i=1}^d \sum_{j\neq i} |\lambda_i - \lambda_j| \frac{\mathrm{d}B_{ij}(t)}{\max(|\Delta_{ij}(t)|, \eta_{ij})}(u_i(t)u_j^\top(t) + u_j(t)u_i^\top(t)) \\ &\quad + \textstyle\sum_{i=1}^d \sum_{j\neq i} (\lambda_i - \lambda_j) \frac{\mathrm{d}t}{\max(\Delta_{ij}^2(t), \eta_{ij}^2)} u_i(t)u_i^\top(t), \end{aligned} \tag{7}$$

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 $\mathrm{d}Z_\eta(t)$ over the time interval $[0, T]$, and apply Itô's Lemma (Lemma 3.1) to obtain an expression for the Frobenius norm of this integral:

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

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

To prove Lemma 4.5, we write

$$\begin{aligned} Z_\eta(T) - Z_\eta(0) = &\frac{1}{2}\int_0^T \sum_{i=1}^d \sum_{j\neq i} |\lambda_i - \lambda_j| \frac{\mathrm{d}B_{ij}(t)}{\max(|\Delta_{ij}(t)|, \eta_{ij})}(u_i(t)u_j^\top(t) + u_j(t)u_i^\top(t)) \\ &- \int_0^T \sum_{i=1}^d \sum_{j\neq i} (\lambda_i - \lambda_j) \frac{\mathrm{d}t}{\max(\Delta_{ij}^2(t), \eta_{ij}^2)} u_i(t)u_i^\top(t). \end{aligned} \tag{8}$$

To compute the Frobenius norm of the first term on the r.h.s. of (8), we use Itô's Lemma (Lemma 3.1), with $X(t) := \int_0^t \sum_{i=1}^d \sum_{j\neq i} |\lambda_i - \lambda_j| \frac{\mathrm{d}B_{ij}(s)}{\max(|\Delta_{ij}(s)|, \eta_{ij})}(u_i(s)u_j^\top(s) + u_j(s)u_i^\top(s))$ and the function $f(X) := \|X\|_F^2 = \sum_{i=1}^d \sum_{j=1}^d X_{ij}^2$. By Itô's Lemma, we have

$$\begin{aligned} \mathbb{E}[\|X(T)\|_F^2 - \|X(0)\|_F^2] = \mathbb{E}[&\frac{1}{2}\int_0^t \sum_{\ell,r} \sum_{\alpha,\beta} (\frac{\partial}{\partial X_{\alpha\beta}} f(X(t)))R_{(\ell r)(\alpha\beta)}(t)\mathrm{d}B_{\ell r}(t)] \\ &+ \mathbb{E}\left[\frac{1}{2}\int_0^t \sum_{\ell,r} \sum_{i,j} \sum_{\alpha,\beta} \left(\frac{\partial^2}{\partial X_{ij}\partial X_{\alpha\beta}} f(X(t))\right) R_{(\ell r)(ij)}(t)R_{(\ell r)(\alpha\beta)}(t)\mathrm{d}t\right], \end{aligned} \tag{9}$$

where $R_{(\ell r)(ij)}(t) := \left(\frac{|\lambda_i - \lambda_j|}{\max(|\Delta_{ij}(t)|, \eta_{ij})}(u_i(t)u_j^\top(t) + u_j(t)u_i^\top(t))\right)[\ell, r]$, and where we denote by either $H_{\ell r}$ or $H[\ell, r]$ the (ℓ, r)'th entry of any matrix H.

The first term on the r.h.s. of (9) is equal to zero since $\mathrm{d}B_{\ell r}(s)$ is independent of both $X(t)$ and $R(t)$ for all $s \geq t$ and the time-integral of each Brownian motion increment $\mathrm{d}B_{\alpha\beta}(s)$ has zero mean. To compute the second term on the r.h.s. of (9), we use the fact that $\frac{\partial^2}{\partial X_{ij}\partial X_{\alpha\beta}} f(X)$ is equal to 2 for $i = j$ and 0 for $i \neq j$

To compute the Frobenius norm of the second term on the r.h.s. of (8), 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

$$\begin{aligned}
&\mathbb{E}\left[\|\hat{V}\Lambda\hat{V}^\top - V\Lambda V^\top\|_F^2\right] = \mathbb{E}\left[\|\Psi(T) - \Psi(0)\|_F^2\right] \\
&\leq 2\int_0^T \mathbb{E}\left[\sum_{i=1}^d \sum_{j\neq i} \frac{(\lambda_i - \lambda_j)^2}{(\sigma_i - \sigma_j)^2}\right] \mathrm{d}t + T\int_0^T \mathbb{E}\left[\sum_{i=1}^d \left(\sum_{j\neq i} \frac{\lambda_i - \lambda_j}{(\sigma_i - \sigma_j)^2}\right)^2\right] \mathrm{d}t \\
&\leq T\sum_{i=1}^k \sum_{j=i+1}^d \frac{(\lambda_i - \lambda_j)^2}{(\sigma_i - \max(\sigma_j, \sigma_{k+1}))^2} + T^2 \sum_{i=1}^k \left(\sum_{j=i+1}^d \frac{\lambda_i - \lambda_j}{(\sigma_i - \max(\sigma_j, \sigma_{k+1}))^2}\right)^2. \qquad (10)
\end{aligned}$$

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 (10), 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 $\Delta_{ij}(t)$ 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 $\geq \Omega(\sqrt{d})$ (Assumption 2.1).

While we observe that our assumption on the top-k+1 eigenvalue gaps holds on multiple real-world datasets, in practice one may need to apply differentially private matrix approximation on any matrix where the "effective rank" of the matrix is k— that is, on any matrix where the k'th eigenvalue gap $\sigma_k - \sigma_{k+1}$ is large— including on matrices where the gaps in the other eigenvalues may not be large and may even be zero. Unfortunately, for matrices with initial gaps in the top-k eigenvalues smaller than $O(\sqrt{d})$, the gaps $\Delta_{ij}(t)$ in the eigenvalues of the Dyson Brownian motion become small enough that the expectation of the (inverse) second moment term $\frac{1}{\Delta_{ij}^2(t)}$ appearing in the Itô integral (Lemma 4.5) in our analysis may be very large or even infinite. Thus, the main question that remains open is whether one can obtain similar bounds on the utility for differentially private matrix approximation for any initial matrix M where the k'th gap $\sigma_k - \sigma_{k+1}$ is large, without any assumption on the gaps between the other eigenvalues of M.

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.

Acknowledgments and Disclosure of Funding

This research was supported in part by NSF CCF-2104528 and CCF-2112665 awards.

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

[36] 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.

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

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

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

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

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

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

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

[44] 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.

[45] 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.

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

[47] 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]

Contents

1 Introduction **1**

2 Results **4**

3 Preliminaries **6**

4 Proof of Theorem 2.2 **7**

4.1 Outline of proof 7

4.2 Step 3: Computing the stochastic derivative $\mathrm{d}\Psi(t)$ 8

4.3 Step 4: Bounding the eigenvalue gaps 8

4.4 Step 5: Integrating the stochastic differential equation 9

4.5 Step 6: Completing the proof 10

5 Conclusion and Future Work **10**

A Proof of Lemma 4.1 **17**

B Proof of Lemma 4.2 **18**

C Proof of Lemma 4.3 **18**

D Proof of Lemma 4.4 **20**

E Proof of Lemma 4.5 **20**

F Proof of Theorem 2.2 **23**

G Proof of Corollary 2.3 **27**

H Proof of Corollary 2.4 **29**

I Eigenvalue Gaps of Wishart Matrices **30**

J Eigenvalue gaps in real datasets **31**

K Challenges in using previous approaches **33**

L Necessity of Assumption 2.1 in our proof **35**

M High probability bounds **35**

A Proof of Lemma 4.1

Proof of Lemma 4.1. We compute the stochastic derivative $\mathrm{d}(u_i(t)u_i^\top(t))$ by applying the formula (4) for the stochastic derivative $\mathrm{d}u_i(t)$ of the eigenvector $u_i(t)$ in Dyson Brownian motion. For any $t \in [0,T]$, we have that the stochastic derivative $\mathrm{d}(u_i(t)u_i^\top(t))$ satisfies

$$
\begin{aligned}
&\mathrm{d}(u_i(t)u_i^\top(t)) = (u_i(t)+\mathrm{d}u_i(t))(u_i(t)+\mathrm{d}u_i(t))^\top - u_i(t)u_i(t)^\top \\
&= \left(u_i(t) + \sum_{j\neq i}\frac{\mathrm{d}B_{ij}(t)}{\gamma_i(t)-\gamma_j(t)}u_j(t) - \frac{1}{2}\sum_{j\neq i}\frac{\mathrm{d}t}{(\gamma_i(t)-\gamma_j(t))^2}u_i(t)\right) \\
&\qquad \times \left(u_i(t) + \sum_{j\neq i}\frac{\mathrm{d}B_{ij}(t)}{\gamma_i(t)-\gamma_j(t)}u_j(t) - \frac{1}{2}\sum_{j\neq i}\frac{\mathrm{d}t}{(\gamma_i(t)-\gamma_j(t))^2}u_i(t)\right)^\top - u_i(t)u_i(t)^\top \\
&= u_i(t)u_i^\top(t) + \sum_{j\neq i}\frac{\mathrm{d}B_{ij}(t)}{\gamma_i(t)-\gamma_j(t)}(u_i(t)u_j^\top(t)+u_j(t)u_i^\top(t)) - \sum_{j\neq i}\frac{\mathrm{d}t}{(\gamma_i(t)-\gamma_j(t))^2}u_i(t)u_i^\top(t) \\
&\qquad + \sum_{j\neq i}\sum_{\ell\neq i}\frac{\mathrm{d}B_{ij}(t)\mathrm{d}B_{i\ell}(t)}{(\gamma_i(t)-\gamma_j(t))(\gamma_i(t)-\gamma_\ell(t))}u_j(t)u_\ell^\top(t) \\
&\qquad\qquad - \varphi_1(t)\varphi_2(t)^\top - \varphi_2(t)\varphi_1(t)^\top + -\varphi_2(t)\varphi_2(t)^\top - u_i(t)u_i(t)^\top
\end{aligned} \tag{11}
$$

where we define $\varphi_1(t) := \sum_{j\neq i}\frac{\mathrm{d}B_{ij}(t)}{\gamma_i(t)-\gamma_j(t)}u_j(t)$ and $\varphi_2(t) := \sum_{j\neq i}\frac{\mathrm{d}t}{(\gamma_i(t)-\gamma_j(t))^2}u_i(t)$. The terms $\varphi_1(t)\varphi_2(t)^\top$ and $\varphi_2(t)\varphi_1(t)^\top$ have differentials $O(\mathrm{d}B_{ij}\mathrm{d}t)$, and $\varphi_2(t)\varphi_2(t)^\top$ has differentials $O(\mathrm{d}t^2)$; thus, all three terms vanish in the stochastic derivative by Lemma 3.1. Therefore, (11) implies that the stochastic derivative $\mathrm{d}(u_i(t)u_i^\top(t))$ satisfies

$$
\begin{aligned}
\mathrm{d}(u_i(t)u_i^\top(t)) &= \sum_{j\neq i}\frac{\mathrm{d}B_{ij}(t)}{\gamma_i(t)-\gamma_j(t)}(u_i(t)u_j^\top(t)+u_j(t)u_i^\top(t)) - \sum_{j\neq i}\frac{\mathrm{d}t}{(\gamma_i(t)-\gamma_j(t))^2}u_i(t)u_i^\top(t) \\
&\quad + \sum_{j\neq i}\sum_{\ell\neq i}\frac{\mathrm{d}B_{ij}(t)\mathrm{d}B_{i\ell}(t)}{(\gamma_i(t)-\gamma_j(t))(\gamma_i(t)-\gamma_\ell(t))}u_j(t)u_\ell^\top(t) \\
&= \sum_{j\neq i}\frac{\mathrm{d}B_{ij}(t)}{\gamma_i(t)-\gamma_j(t)}(u_i(t)u_j^\top(t)+u_j(t)u_i^\top(t)) - \sum_{j\neq i}\frac{\mathrm{d}t}{(\gamma_i(t)-\gamma_j(t))^2}u_i(t)u_i^\top(t) \\
&\quad + \sum_{j\neq i}\left(\frac{\mathrm{d}B_{ij}(t)}{\gamma_i(t)-\gamma_j(t)}\right)^2 u_j(t)u_j^\top(t) \\
&= \sum_{j\neq i}\frac{\mathrm{d}B_{ij}(t)}{\gamma_i(t)-\gamma_j(t)}(u_i(t)u_j^\top(t)+u_j(t)u_i^\top(t)) - \sum_{j\neq i}\frac{\mathrm{d}t}{(\gamma_i(t)-\gamma_j(t))^2}u_i(t)u_i^\top(t) \\
&\quad + \sum_{j\neq i}\frac{\mathrm{d}t}{(\gamma_i(t)-\gamma_j(t))^2}u_j(t)u_j^\top(t),
\end{aligned} \tag{12}
$$

where the second-to-last equality holds since all terms $\mathrm{d}B_{ij}(t)\mathrm{d}B_{i\ell}(t)$ with $j \neq \ell$ in the sum $\sum_{j\neq i}\sum_{\ell\neq i}\frac{\mathrm{d}B_{ij}(t)\mathrm{d}B_{i\ell}(t)}{(\gamma_i(t)-\gamma_j(t))(\gamma_i(t)-\gamma_\ell(t))}u_j(t)u_\ell^\top(t)$ vanish by Itô's Lemma (Lemma 3.1) since they have mean 0 and are $O(\mathrm{d}B_{ij}(t)\mathrm{d}B_{i\ell}(t))$; we are therefore left only with the terms $j = \ell$ in the sum which have differential terms $(\mathrm{d}B_{ij}(t))^2$ which have mean $\mathrm{d}t$ plus higher-order terms which vanish by Itô's Lemma. Therefore (12) implies that

$$
\begin{aligned}
&\mathrm{d}(u_i(t)u_i^\top(t)) \\
&= \sum_{j\neq i}\frac{\mathrm{d}B_{ij}(t)}{\gamma_i(t)-\gamma_j(t)}(u_i(t)u_j^\top(t)+u_j(t)u_i^\top(t)) - \sum_{j\neq i}\frac{\mathrm{d}t}{(\gamma_i(t)-\gamma_j(t))^2}(u_i(t)u_i^\top(t)-u_j(t)u_j^\top(t)).
\end{aligned}
$$

■

B Proof of Lemma 4.2

Proof of Lemma 4.2. To compute the stochastic derivative of $\Psi(t)$, we would like to apply our formula for the stochastic derivative of the projection matrix $u_i(t)u_i^\top(t)$ for each eigenvector $u_i(t)$ (Lemma 4.1). Towards this end, we first decompose the matrix $\Psi(t)$ as a sum of these projection matrices $u_i(t)u_i^\top(t)$:

$$\Psi(t) = \textstyle\sum_{i=1}^d \lambda_i(u_i(t)u_i^\top(t)). \tag{13}$$

Taking the derivative on both sides of (13), we have

$$\mathrm{d}\Psi(t) = \textstyle\sum_{i=1}^d \lambda_i \mathrm{d}(u_i(t)u_i^\top(t)). \tag{14}$$

Thus, plugging in Lemma 4.1 for each $i \in [d]$ into (14), we have that

$$\begin{aligned}
\mathrm{d}\Psi(t) &= \sum_{i=1}^d \lambda_i \mathrm{d}(u_i(t)u_i^\top(t)) \\
&\overset{\text{Lemma 4.1}}{=} \sum_{i=1}^d \lambda_i \left(\sum_{j\neq i} \frac{\mathrm{d}B_{ij}(t)}{\gamma_i(t)-\gamma_j(t)}(u_i(t)u_j^\top(t) + u_j(t)u_i^\top(t)) \right. \\
&\quad \left. - \sum_{j\neq i} \frac{\mathrm{d}t}{(\gamma_i(t)-\gamma_j(t))^2}(u_i(t)u_i^\top(t) - u_j(t)u_j^\top(t)) \right) \\
&= \frac{1}{2}\sum_{i=1}^d \sum_{j\neq i} (\lambda_i - \lambda_j) \frac{\mathrm{d}B_{ij}(t)}{\gamma_i(t)-\gamma_j(t)}(u_i(t)u_j^\top(t) + u_j(t)u_i^\top(t)) \\
&\quad - \frac{1}{2}\sum_{i=1}^d \sum_{j\neq i} (\lambda_i - \lambda_j) \frac{\mathrm{d}t}{(\gamma_i(t)-\gamma_j(t))^2}(u_i(t)u_i^\top(t) - u_j(t)u_j^\top(t)) \\
&= \frac{1}{2}\sum_{i=1}^d \sum_{j\neq i} (\lambda_i - \lambda_j) \frac{\mathrm{d}B_{ij}(t)}{\gamma_i(t)-\gamma_j(t)}(u_i(t)u_j^\top(t) + u_j(t)u_i^\top(t)) \\
&\quad - \sum_{i=1}^d \sum_{j\neq i} (\lambda_i - \lambda_j) \frac{\mathrm{d}t}{(\gamma_i(t)-\gamma_j(t))^2} u_i(t)u_i^\top(t),
\end{aligned}$$

where the second equality holds by Lemma 4.1. To see why the third equality holds, for the first term inside the summation, note that $\mathrm{d}B(t)$ is a symmetric matrix of differentials which means that $\mathrm{d}B_{ij}(t) = \mathrm{d}B_{ji}(t)$ for all i, j, and hence that $\frac{\mathrm{d}B_{ij}(t)}{\gamma_i(t)-\gamma_j(t)}(u_i(t)u_j^\top(t)+u_j(t)u_i^\top(t)) = -\frac{\mathrm{d}B_{ji}(t)}{\gamma_j(t)-\gamma_i(t)}(u_j(t)u_i^\top(t)+u_i(t)u_j^\top(t))$ for all $i \neq j$. For the second term inside the summation, note that $\frac{\mathrm{d}t}{(\gamma_i(t)-\gamma_j(t))^2}(u_i(t)u_i^\top(t)-u_j(t)u_j^\top(t)) = -\frac{\mathrm{d}t}{(\gamma_j(t)-\gamma_i(t))^2}(u_j(t)u_j^\top(t)-u_i(t)u_i^\top(t))$ for all $i \neq j$. ■

C Proof of Lemma 4.3

Proof of Lemma 4.3. To prove Lemma 4.3 we will use Doob's submartingale inequality. Towards this end, let $\mathcal{F}_s$ be the filtration generated by $B(s)$. First, we note that $\exp(\|B(t)\|_2)$ is a submartingale for all $t \geq 0$; that is, $\mathbb{E}[\exp(\|B(t)\|_2)|\mathcal{F}_s] \geq \exp(\|B(s)\|_2)$ for all $0 \leq s \leq t$.

This is because for all $s \le t$, we have

$$\begin{aligned}
\mathbb{E}[\exp(\|B(t)\|_2)|\mathcal{F}_s] &= \mathbb{E}\left[\exp\left(\sup_{v\in\mathbb{R}^d:\|v\|_2=1} v^\top B(t)v\right) \;\middle|\; \mathcal{F}_s\right] \\
&\ge \exp\left(\mathbb{E}\left[\sup_{v\in\mathbb{R}^d:\|v\|_2=1} v^\top B(t)v \;\middle|\; \mathcal{F}_s\right]\right) \\
&\ge \exp\left(\sup_{v\in\mathbb{R}^d:\|v\|_2=1} \mathbb{E}\left[v^\top B(t)v \mid \mathcal{F}_s\right]\right) \\
&= \exp\left(\sup_{v\in\mathbb{R}^d:\|v\|_2=1} \mathbb{E}\left[v^\top (B(t)-B(s))v + v^\top B(s)v \mid \mathcal{F}_s\right]\right) \\
&= \exp\left(\sup_{v\in\mathbb{R}^d:\|v\|_2=1} \mathbb{E}\left[v^\top (B(t)-B(s))v| \mathcal{F}_s\right] + \mathbb{E}\left[v^\top B(s)v \mid \mathcal{F}_s\right]\right) \\
&= \exp\left(\sup_{v\in\mathbb{R}^d:\|v\|_2=1} \mathbb{E}\left[v^\top B(s)v \mid \mathcal{F}_s\right]\right) \\
&= \exp\left(\sup_{v\in\mathbb{R}^d:\|v\|_2=1} v^\top B(s)v\right), \\
&= \exp(\|B(s)\|_2),
\end{aligned}$$

where the first inequality holds by Jensen's inequality since $\exp(\cdot)$ is convex, and the third equality holds since $v^\top(B(t)-B(s))v$ is independent of $\mathcal{F}_s$ and is distributed as $N(0, 2(t-s))$.

Thus, by Doob's submartingale inequality, for any $\beta > 0$ (we will choose the value of β later to optimize our bound) we have,

$$\begin{aligned}
\mathbb{P}\left(\sup_{t\in[0,T]} \|B(t)\|_2 > 2T(\sqrt{d}+\alpha)\right) &= \mathbb{P}\left(\sup_{t\in[0,T]} \frac{\beta}{2T}\|B(t)\|_2 - \beta\sqrt{d} > \beta\alpha\right) \\
&= \mathbb{P}\left(\sup_{t\in[0,T]} \exp\left(\frac{\beta}{2T}\|B(t)\|_2 - \beta\sqrt{d}\right) > \exp(\beta\alpha)\right) \\
&\le \frac{\mathbb{E}[\exp(\frac{\beta}{2T}\|B(t)\|_2 - \beta\sqrt{d})]}{\exp(\beta\alpha)} \\
&= \frac{\int_0^\infty \mathbb{P}[\exp(\frac{\beta}{2T}\|B(t)\|_2 - \beta\sqrt{d}) > x]\mathrm{d}x}{\exp(\beta\alpha)} \\
&= \frac{\int_0^\infty \mathbb{P}[\frac{1}{2}\|B(t)\|_2 - \sqrt{d} > \beta^{-1}\log(x)]\mathrm{d}x}{\exp(\beta\alpha)} \\
&\le \frac{\int_0^\infty 2e^{-\beta^{-2}\log^2(x)}\mathrm{d}x}{\exp(\beta\alpha)} \\
&= \frac{2\sqrt{\pi}\beta e^{\frac{1}{4}\beta^2}}{\exp(\beta\alpha)} \\
&\le \frac{2\sqrt{\pi}e^{\frac{1}{2}\beta^2}}{\exp(\beta\alpha)} \\
&= 2\sqrt{\pi}e^{\frac{1}{2}\beta^2-\beta\alpha},
\end{aligned}$$

where the first inequality holds by Doob's submartingale inequality, and the second inequality holds by Lemma 3.2. Setting $\beta = \alpha$, we have

$$\mathbb{P}\left(\sup_{t\in[0,T]} \|B(t)\|_2 > T(\sqrt{d}+\alpha)\right) \le 2\sqrt{\pi}e^{-\frac{1}{2}\alpha^2}.$$

■

D Proof of Lemma 4.4

Proof of Lemma 4.4. To prove Lemma 4.4, we plug our high-probability concentration bound for $\sup_{t\in[0,T]} \|B(t)\|_2$ (Lemma 4.3) into Weyl's Inequality (Lemma 3.3). Since, at every time t, $\Phi(t) = M + B(t)$ and $\gamma_1(t) \geq \cdots \geq \gamma_d(t)$ are the eigenvalues of $\Phi(t)$, Weyl's Inequality implies that

$$\gamma_i(t) - \gamma_{i+1}(t) \geq \gamma_i(0) - \gamma_{i+1}(0) - \|B(t)\|_2, \qquad \forall t \in [0,T], i \in [d]. \tag{15}$$

Therefore, plugging Lemma 4.3 into (15) we have that

$$\begin{aligned}
&\mathbb{P}\left(\bigcup_{i\in S}\left\{\inf_{t\in[0,T]} \gamma_i(t) - \gamma_{i+1}(t) < \frac{1}{2}(\gamma_i(0) - \gamma_{i+1}(0)) - \alpha)\right\}\right) \\
&\overset{\text{Eq. (15)}}{\leq} \mathbb{P}\left(\bigcup_{i\in S}\left\{\gamma_i(0) - \gamma_{i+1}(0) - \sup_{t\in[0,T]} 2\|B(t)\|_2 < \frac{1}{2}(\gamma_i(0) - \gamma_{i+1}(0)) - \alpha)\right\}\right) \\
&= \mathbb{P}\left(\bigcup_{i\in S}\left\{\sup_{t\in[0,T]} \|B(t)\|_2 > \frac{1}{4}(\gamma_i(0) - \gamma_{i+1}(0)) + \frac{1}{2}\alpha)\right\}\right) \\
&\overset{\text{Assumption 2.1}}{\leq} \mathbb{P}\left(\bigcup_{i\in S}\left\{\sup_{t\in[0,T]} \|B(t)\|_2 > 2T\sqrt{d} + \frac{1}{2}\alpha)\right\}\right) \\
&= \mathbb{P}\left(\sup_{t\in[0,T]} \|B(t)\|_2 > 2T\sqrt{d} + \frac{1}{2}\alpha)\right) \\
&\overset{\text{Lemma 4.3}}{\leq} 2\sqrt{\pi}e^{-\frac{1}{32}\alpha^2},
\end{aligned}$$

The first inequality holds by (15), and the second inequality holds by Assumption 2.1 since $\gamma_i(0) = \sigma_i$ for each $i \in [d]$ because $\Phi(0) = M$. The last inequality holds by the high-probability concentration bound for $\sup_{t\in[0,T]} \|B(t)\|_2$ (Lemma 4.3). ■

E Proof of Lemma 4.5

Proof of Lemma 4.5. By the definition of $Z_\eta(t)$ we have that

$$\begin{aligned}
Z_\eta(T) - Z_\eta(0) &= \int_0^T \mathrm{d}Z_\eta(t) \\
&= \frac{1}{2}\int_0^T \sum_{i=1}^d \sum_{j\neq i} |\lambda_i - \lambda_j| \frac{\mathrm{d}B_{ij}(t)}{\max(|\Delta_{ij}(t)|, \eta_{ij})}(u_i(t)u_j^\top(t) + u_j(t)u_i^\top(t)) \\
&\quad - \int_0^T \sum_{i=1}^d \sum_{j\neq i} (\lambda_i - \lambda_j) \frac{\mathrm{d}t}{\max(\Delta_{ij}^2(t), \eta_{ij}^2)} u_i(t)u_i^\top(t).
\end{aligned}$$

Therefore, we have that

$$\begin{aligned}
\|Z_\eta(T) - Z_\eta(0)\|_F^2 \leq \frac{1}{2}&\left\|\int_0^T \sum_{i=1}^d \sum_{j\neq i} |\lambda_i - \lambda_j| \frac{\mathrm{d}B_{ij}(t)}{\max(|\Delta_{ij}(t)|, \eta_{ij})}(u_i(t)u_j^\top(t) + u_j(t)u_i^\top(t))\right\|_F^2 \\
&+ \left\|\int_0^T \sum_{i=1}^d \sum_{j\neq i} (\lambda_i - \lambda_j) \frac{\mathrm{d}t}{\max(\Delta_{ij}^2(t), \eta_{ij}^2)} u_i(t)u_i^\top(t)\right\|_F^2.
\end{aligned} \tag{16}$$

The first term on the r.h.s. of (16) (inside its Frobenius norm) is a "diffusion" term–that is, the integral has mean 0 and Brownian motion differentials $\mathrm{d}B_{ij}(t)$ inside the integral. The second term on the r.h.s. (inside its Frobenius norm) is a "drift" term– that is, the integral has non-zero mean and deterministic differentials $\mathrm{d}t$ inside the integral. We bound the diffusion and drift terms separately.

Bounding the diffusion term: We first use Itô's Lemma (Lemma 3.1) to bound the diffusion term in (16). Towards this end, let $f : \mathbb{R}^{d\times d} :\to \mathbb{R}$ be the function which takes as input a $d \times d$ matrix and outputs the square of its Frobenius norm: $f(X) := \|X\|_F^2 = \sum_{i=1}^d \sum_{j=1}^d X_{ij}^2$ for every $X \in \mathbb{R}^{d\times d}$. Then

$$\frac{\partial^2}{\partial X_{ij}\partial X_{\alpha\beta}} f(X) = \begin{cases} 2 & \text{if} \quad (i,j) = (\alpha,\beta) \\ 0 & \text{otherwise.} \end{cases} \tag{17}$$

Define $X(t) := \int_0^t \sum_{i=1}^d \sum_{j\neq i} |\lambda_i - \lambda_j| \frac{\mathrm{d}B_{ij}(s)}{\max(|\Delta_{ij}(s)|,\eta_{ij})}(u_i(s)u_j^\top(s) + u_j(s)u_i^\top(s))$ for all $t \geq 0$. Then

$$\mathrm{d}X_{\ell r}(t) = \sum_{j=1}^d R_{(\ell r)(ij)}(t)\mathrm{d}B_{(ij)}(t) \qquad \forall t \geq 0,$$

where $R_{(\ell r)(ij)}(t) := \left(\frac{|\lambda_i-\lambda_j|}{\max(|\Delta_{ij}(t)|,\eta_{ij})}(u_i(t)u_j^\top(t) + u_j(t)u_i^\top(t))\right)[\ell, r]$, and where we denote by either $H_{\ell r}$ or $H[\ell, r]$ the (ℓ, r)'th entry of any matrix H.

Then we have

$$\begin{aligned}
&\mathbb{E}\left[\left\|\int_0^T \sum_{i=1}^d \sum_{j\neq i} |\lambda_i - \lambda_j| \frac{\mathrm{d}B_{ij}(t)}{\max(|\Delta_{ij}(t)|,\eta_{ij})}(u_i(t)u_j^\top(t) + u_j(t)u_i^\top(t))\right\|_F^2\right] \\
&= \mathbb{E}[f(X(T))] \\
&= \mathbb{E}[f(X(T)) - f(X(0))] \\
&\overset{\text{Itô's Lemma (Lemma 3.1)}}{=} \mathbb{E}\left[\frac{1}{2}\int_0^t \sum_{\ell,r}\sum_{\alpha,\beta}\left(\frac{\partial}{\partial X_{\alpha\beta}} f(X(t))\right) R_{(\ell r)(\alpha\beta)}(t)\mathrm{d}B_{\ell r}(t)\right] \\
&\qquad + \mathbb{E}\left[\frac{1}{2}\int_0^t \sum_{\ell,r}\sum_{i,j}\sum_{\alpha,\beta}\left(\frac{\partial^2}{\partial X_{ij}\partial X_{\alpha\beta}} f(X(t))\right) R_{(\ell r)(ij)}(t)R_{(\ell r)(\alpha\beta)}(t)\mathrm{d}t\right] \\
&= 0 + \mathbb{E}\left[\frac{1}{2}\int_0^t \sum_{\ell,r}\sum_{i,j}\sum_{\alpha,\beta}\left(\frac{\partial^2}{\partial X_{ij}\partial X_{\alpha\beta}} f(X(t))\right) R_{(\ell r)(ij)}(t)R_{(\ell r)(\alpha\beta)}(t)\mathrm{d}t\right],
\end{aligned} \tag{18}$$

where the third equality is Itô's Lemma (Lemma 3.1), and the last equality holds since $\mathbb{E}\left[\int_0^T \left(\frac{\partial}{\partial X_{\alpha\beta}} f(X(t))\right) R_{(\ell r)(\alpha\beta)}(t)\mathrm{d}B_{\ell r}(t)\right] = 0$ for each $\ell, r, \alpha, \beta \in [d]$ because $\mathrm{d}B_{\ell r}(s)$ is independent of both $X(t)$ and $R(t)$ for all $s \geq t$ and the Brownian motion increments $\mathrm{d}B_{\alpha\beta}(s)$ satisfy $\mathbb{E}[\int_t^\tau \mathrm{d}B_{\alpha\beta}(s)] = \mathbb{E}[B_{\alpha\beta}(\tau) - B_{\alpha\beta}(t)] = 0$ for any $\tau \geq t$.

Thus, plugging (17) into (18), we have

$$
\begin{aligned}
&\mathbb{E}\left[\left\|\int_0^T \sum_{i=1}^d \sum_{j\neq i} |\lambda_i-\lambda_j| \frac{\mathrm{d}B_{ij}(t)}{\max(|\Delta_{ij}(t)|,\eta_{ij})}(u_i(t)u_j^\top(t)+u_j(t)u_i^\top(t))\right\|_F^2\right] \\
&\overset{\text{Eq.(17),(18)}}{=} \mathbb{E}\left[\frac{1}{2}\int_0^t \sum_{\ell,r}\sum_{i,j} 2R^2_{(\ell r)(ij)}(t)\mathrm{d}t\right] \\
&= \mathbb{E}\left[\int_0^t \sum_{\ell,r}\sum_{i,j}\left(\left(\frac{|\lambda_i-\lambda_j|}{\max(|\Delta_{ij}(t)|,\eta_{ij})^2}(u_i(t)u_j^\top(t)+u_j(t)u_i^\top(t))\right)[\ell,r]\right)^2\mathrm{d}t\right] \\
&= \mathbb{E}\left[\int_0^t \sum_{i,j}\sum_{\ell,r}\left(\left(\frac{|\lambda_i-\lambda_j|}{\max(|\Delta_{ij}(t)|,\eta_{ij})^2}(u_i(t)u_j^\top(t)+u_j(t)u_i^\top(t))\right)[\ell,r]\right)^2\mathrm{d}t\right] \\
&= \mathbb{E}\left[\int_0^t \sum_{i,j}\left\|\frac{|\lambda_i-\lambda_j|}{\max(|\Delta_{ij}(t)|,\eta_{ij})}(u_i(t)u_j^\top(t)+u_j(t)u_i^\top(t))\right\|_F^2\mathrm{d}t\right] \\
&= 2\int_0^T \mathbb{E}\left[\sum_{i=1}^d\sum_{j\neq i}\frac{(\lambda_i-\lambda_j)^2}{\max(\Delta_{ij}^2(t),\eta_{ij}^2)}\|u_i(t)u_j^\top(t)+u_j(t)u_i^\top(t)\|_F^2\mathrm{d}t\right] \\
&= 4\int_0^T \mathbb{E}\left[\sum_{i=1}^d\sum_{j\neq i}\frac{(\lambda_i-\lambda_j)^2}{\max(\Delta_{ij}^2(t),\eta_{ij}^2)}\mathrm{d}t\right], \qquad (19)
\end{aligned}
$$

where the fifth equality holds because $\langle u_i(t)u_j^\top(t), u_\ell(t)u_h^\top(t)\rangle = 0$ for all $(i,j)\neq(\ell,h)$, and the last equality holds because $\|u_i(t)u_j^\top(t)+u_j(t)u_i^\top(t)\|_F^2 = 2$ for all t with probability 1.

Bounding the drift term: To bound the drift term in (16), we use the Cauchy-Shwarz inequality:

$$
\begin{aligned}
&\left\| \int_0^T \sum_{i=1}^d \sum_{j\neq i} (\lambda_i - \lambda_j) \frac{\mathrm{d}t}{\max(\Delta_{ij}^2(t), \eta_{ij}^2)} u_i(t) u_i^\top(t) \right\|_F^2 \\
&= \left\| \int_0^T \sum_{i=1}^d \sum_{j\neq i} \frac{\lambda_i - \lambda_j}{\max(\Delta_{ij}^2(t), \eta_{ij}^2)} u_i(t) u_i^\top(t) \times 1 \mathrm{d}t \right\|_F^2 \\
&\overset{\text{Cauchy-Schwarz Inequality}}{\leq} \int_0^T \left\| \sum_{i=1}^d \sum_{j\neq i} \frac{\lambda_i - \lambda_j}{\max(\Delta_{ij}^2(t), \eta_{ij}^2)} u_i(t) u_i^\top(t) \right\|_F^2 \mathrm{d}t \times \int_0^T 1^2 \mathrm{d}t \\
&= T \int_0^T \left\| \sum_{i=1}^d \sum_{j\neq i} \frac{\lambda_i - \lambda_j}{\max(\Delta_{ij}^2(t), \eta_{ij}^2)} u_i(t) u_i^\top(t) \right\|_F^2 \mathrm{d}t \\
&= T \int_0^T \sum_{i=1}^d \left\| \sum_{j\neq i} \frac{\lambda_i - \lambda_j}{\max(\Delta_{ij}^2(t), \eta_{ij}^2)} u_i(t) u_i^\top(t) \right\|_F^2 \mathrm{d}t \\
&= T \int_0^T \sum_{i=1}^d \left\| \left(\sum_{j\neq i} \frac{\lambda_i - \lambda_j}{\max(\Delta_{ij}^2(t), \eta_{ij}^2)} \right) u_i(t) u_i^\top(t) \right\|_F^2 \mathrm{d}t \\
&= T \int_0^T \sum_{i=1}^d \left(\sum_{j\neq i} \frac{\lambda_i - \lambda_j}{\max(\Delta_{ij}^2(t), \eta_{ij}^2)} \right)^2 \left\| u_i(t) u_i^\top(t) \right\|_F^2 \mathrm{d}t \\
&= T \int_0^T \sum_{i=1}^d \left(\sum_{j\neq i} \frac{\lambda_i - \lambda_j}{\max(\Delta_{ij}^2(t), \eta_{ij}^2)} \right)^2 \times 1 \mathrm{d}t, \qquad (20)
\end{aligned}
$$

where the first inequality is by the Cauchy-Schwarz inequality for integrals (applied to each entry of the matrix-valued integral). The third equality holds since $\langle u_i(t)u_i^\top(t), u_j(t)u_j^\top(t)\rangle = 0$ for all $i \neq j$. The last equality holds since $\|u_i(t)u_i^\top(t)\|_F^2 = 1$ with probability 1. Therefore, taking the expectation on both sides of (16), and plugging (19) and (20) into (16), we have

$$
\begin{aligned}
\mathbb{E}\left[\|Z_\eta(T) - Z_\eta(0)\|_F^2 \right] \leq 2 \int_0^T \mathbb{E}\left[\sum_{i=1}^d \sum_{j\neq i} \frac{(\lambda_i - \lambda_j)^2}{\max(\Delta_{ij}^2(t), \eta_{ij}^2)} \right] \mathrm{d}t \\
+ T \int_0^T \mathbb{E}\left[\sum_{i=1}^d \left(\sum_{j\neq i} \frac{\lambda_i - \lambda_j}{\max(\Delta_{ij}^2(t), \eta_{ij}^2)} \right)^2 \right] \mathrm{d}t. \qquad (21)
\end{aligned}
$$

■

F Proof of Theorem 2.2

Proof of Theorem 2.2. To complete the proof of Theorem 2.2, we plug in the high-probability concentration bounds on the eigenvalue gaps $\Delta_{ij}(t) = \gamma_i(t) - \gamma_j(t)$ (Lemma 4.4) 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 $\eta_{ij} = \frac{1}{4}(\sigma_i - \max(\sigma_j, \sigma_{k+1}))$, by Lemma 4.2 we have that the derivative $\mathrm{d}\Psi(t)$ satisfies $\mathrm{d}\Psi(t) = \mathrm{d}Z_\eta(t)$ for all $t \in [0, T]$ w.h.p. and hence that $Z_\eta(t) = \Psi(t)$ for all $t \in [0, T]$ w.h.p. Plugging in the high-probability bounds on the gaps $\Delta_{ij}(t)$

(Lemma 4.4) into the bound on $\mathbb{E}\left[\|Z_\eta(T) - Z_\eta(0)\|_F^2\right]$ from Lemma 4.5 therefore allows us to obtain a bound for $\mathbb{E}\left[\|\Psi(T) - \Psi(0)\|_F^2\right]$. To obtain a bound for the utility $\mathbb{E}\left[\|\hat{V}\Lambda\hat{V}^\top - V\Lambda V^\top\|_F^2\right]$ we set $T = \frac{\sqrt{2\log(\frac{1.25}{\delta})}}{\varepsilon}$, in which case we have $\Psi(T) = \hat{V}\Lambda\hat{V}^\top$ and hence that $\mathbb{E}\left[\|\hat{V}\Lambda\hat{V}^\top - V\Lambda V^\top\|_F^2\right] = \mathbb{E}\left[\|\Psi(T) - \Psi(0)\|_F^2\right]$.

Towards this end, for all $i \neq j$ we define η_{ij} as follows: Let $\eta_{ij} = \frac{1}{4}(\sigma_i - \max(\sigma_j, \sigma_{k+1}))$ for $0 < i < j \leq d$ and $i \leq k$, $\eta_{ij} = 0$ if $0 < i < j \leq d$ and $i > k$, and $\eta_{ij} = \eta_{ji}$ otherwise.

Define the event $E = \cap_{i,j\in[d],i\neq j}\{\inf_{t\in[0,T]} \Delta_{ij}(t) \geq \eta_{ij}\}$. And define the event $\hat{E} := \bigcap_{i\in[k]}\left\{\inf_{t\in[0,T]} \gamma_i(t) - \gamma_{i+1}(t) \geq \frac{1}{4}(\sigma_i - \sigma_{i+1}))\right\}$. Then $\hat{E} \subseteq E$. In particular, whenever the event E occurs, by Lemma 4.2 we have that the derivative $\mathrm{d}\Psi(t)$ satisfies $\mathrm{d}\Psi(t) = \mathrm{d}Z_\eta(t)$ for all $t \in [0, T]$ and hence that

$$\Psi(t) = \Psi(0) + \int_0^t \mathrm{d}\Psi(s) = Z_\eta(0) + \int_0^t \mathrm{d}Z_\eta(s) = Z_\eta(t) \qquad \forall t \in [0, T],$$

whenever the event E occurs, since $Z_\eta(0) = \Psi(0)$ by definition. Thus we have that, conditioning $\Psi(t)$ and $Z_\eta(t)$ on the event E,

$$\Psi(t)|E = Z_\eta(t)|E \qquad \forall t \in [0, T]. \tag{22}$$

To bound the utility $\mathbb{E}\left[\|\hat{V}\Lambda\hat{V}^\top - V\Lambda V^\top\|_F^2\right]$, we first separate $\mathbb{E}\left[\|\hat{V}\Lambda\hat{V}^\top - V\Lambda V^\top\|_F^2\right]$ into a sum of terms conditioned on the event E and its complement E^c. By Lemma 4.5 we have

$$\begin{aligned}
&\mathbb{E}\left[\|\hat{V}\Lambda\hat{V}^\top - V\Lambda V^\top\|_F^2\right] \\
&= \mathbb{E}\left[\|\hat{V}\Lambda\hat{V}^\top - V\Lambda V^\top\|_F^2 \middle| E\right] \times \mathbb{P}(E) + \mathbb{E}\left[\|\hat{V}\Lambda\hat{V}^\top - V\Lambda V^\top\|_F^2 \middle| E^c\right] \times \mathbb{P}(E^c) \\
&\leq \mathbb{E}\left[\|\hat{V}\Lambda\hat{V}^\top - V\Lambda V^\top\|_F^2 \middle| E\right] \times \mathbb{P}(E) \\
&+ 4\left(\mathbb{E}\left[\|(\hat{V} - V)\Lambda\hat{V}^\top\|_F^2 \middle| E^c\right] + \mathbb{E}\left[\|V\Lambda(\hat{V} - V)^\top\|_F^2 \middle| E^c\right]\right) \times \mathbb{P}(E^c) \\
&\leq \mathbb{E}\left[\|\hat{V}\Lambda\hat{V}^\top - V\Lambda V^\top\|_F^2 | E\right] \times \mathbb{P}(E) + 8\mathbb{E}\left[\|\hat{V} - V\|_2^2 \times \|\Lambda\|_F^2 \middle| E^c\right] \times \mathbb{P}(E^c) \\
&\leq \mathbb{E}\left[\|\hat{V}\Lambda\hat{V}^\top - V\Lambda V^\top\|_F^2 \middle| E\right] \times \mathbb{P}(E) + 32\|\Lambda\|_F^2 \times \mathbb{P}(E^c) \\
&= \mathbb{E}\left[\|\Psi(T) - \Psi(0)\|_F^2 \middle| E\right] \times \mathbb{P}(E) + 32\|\Lambda\|_F^2 \times \mathbb{P}(E^c) \\
&\leq \mathbb{E}\left[\|\Psi(T) - \Psi(0)\|_F^2 \middle| E\right] \times \mathbb{P}(E) + 32\lambda_1^2 k \times \mathbb{P}(\hat{E}^c),
\end{aligned} \tag{23}$$

where the second inequality holds by the sub-multiplicative property of the Frobenius norm which says that $\|XY\|_F \leq \|X\|_2 \times \|Y\|_F$ for any $X, Y \in \mathbb{R}^{d\times d}$. The third inequality holds since $\|V\|_2 = \|\hat{V}\|_2 = 1$ since $V, \hat{V}$ are orthogonal matrices.

To bound the first term in (23), we use the fact that $\Psi(t)|E = Z_\eta(t)|E$ (Equation (22)) and apply Lemma 4.5 to bound $\mathbb{E}\left[\|Z_\eta(T) - Z_\eta(0)\|_F^2\right]$. Thus we have,

$$
\begin{aligned}
&\mathbb{E}\left[\|\Psi(T) - \Psi(0)\|_F^2 \,\middle|\, E\right] \times \mathbb{P}(E) \\
&\overset{\text{Eq. (22)}}{=} \mathbb{E}\left[\|Z_\eta(T) - Z_\eta(0)\|_F^2 \,\middle|\, E\right] \times \mathbb{P}(E) \\
&\leq \mathbb{E}\left[\|Z_\eta(T) - Z_\eta(0)\|_F^2\right] \\
&\overset{\text{Lemma 4.5}}{=} 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)}\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}{\max(\Delta_{ij}^2(t), \eta_{ij}^2)}\right)^2\right] \mathrm{d}t \\
&\leq 4\int_0^T \mathbb{E}\left[\sum_{i=1}^d \sum_{j=i+1}^d \frac{(\lambda_i - \lambda_j)^2}{\max(\Delta_{ij}^2(t), \eta_{ij}^2)}\right] \mathrm{d}t + 2T\int_0^T \mathbb{E}\left[\sum_{i=1}^d \left(\sum_{j=i+1}^d \frac{|\lambda_i - \lambda_j|}{\max(\Delta_{ij}^2(t), \eta_{ij}^2)}\right)^2\right] \mathrm{d}t \\
&= 4\int_0^T \mathbb{E}\left[\sum_{i=1}^k \sum_{j=i+1}^d \frac{(\lambda_i - \lambda_j)^2}{\max(\Delta_{ij}^2(t), \eta_{ij}^2)}\right] \mathrm{d}t + 2T\int_0^T \mathbb{E}\left[\sum_{i=1}^k \left(\sum_{j=i+1}^d \frac{|\lambda_i - \lambda_j|}{\max(\Delta_{ij}^2(t), \eta_{ij}^2)}\right)^2\right] \mathrm{d}t \\
&\leq 64\int_0^T \mathbb{E}\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] \mathrm{d}t \\
&+ 32T\int_0^T \mathbb{E}\left[\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\right] \mathrm{d}t \\
&= 64T\mathbb{E}\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] + 32T^2\mathbb{E}\left[\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\right] \\
&= 64T\sum_{i=1}^k \sum_{j=i+1}^d \frac{(\lambda_i - \lambda_j)^2}{(\sigma_i - \max(\sigma_j, \sigma_{k+1}))^2} + 32T^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,
\end{aligned}
\tag{24}
$$

where the first equality holds since $\Psi(t)|E = Z_\eta(t)|E$ by (22), and the second equality holds by Lemma 4.5. The second inequality holds since, $\frac{(\lambda_i-\lambda_j)^2}{\max(\Delta_{ij}^2(t),\eta_{ij}^2)} = \frac{(\lambda_j-\lambda_i)^2}{\max(\Delta_{ji}^2(t),\eta_{ji}^2)}$ and $\frac{|\lambda_i-\lambda_j|}{\max(\Delta_{ij}^2(t),\eta_{ij}^2)} = \frac{|\lambda_j-\lambda_i|}{\max(\Delta_{ji}^2(t),\eta_{ji}^2)}$ for all $i, j \in [d]$. The third equality holds since $\lambda_i = 0$ for all $i \geq k+1$. The third inequality holds since $\eta_{ij} = \frac{1}{4}(\sigma_i - \max(\sigma_j, \sigma_{k+1}))$ for all $0 < i < j \leq d$.

To bound the second term in (23), we have by Lemma 4.4 that

$$\begin{aligned}
\mathbb{P}(\hat{E}^c) &= \mathbb{P}\left(\left(\bigcap_{i\in[k]}\left\{\inf_{t\in[0,T]}\gamma_i(t)-\gamma_{i+1}(t)\geq\frac{1}{4}(\sigma_i-\sigma_{i+1}))\right\}\right)^c\right)\\
&= \mathbb{P}\left(\bigcup_{i\in[k]}\left\{\inf_{t\in[0,T]}\gamma_i(t)-\gamma_{i+1}(t)<\frac{1}{4}(\sigma_i-\sigma_{i+1}))\right\}\right)\\
&\overset{\text{Assumption2.1}}{\leq} \mathbb{P}\left(\bigcup_{i\in[k]}\left\{\inf_{t\in[0,T]}\gamma_i(t)-\gamma_{i+1}(t)<\frac{1}{2}(\sigma_i-\sigma_{i+1})-3\log^{\frac{1}{2}}(\lambda_1 k))\right\}\right)\\
&\overset{\text{Lemma4.4}}{\leq} \min(e^{-\log(\lambda_1^2 k))},1)\\
&= \min\left(\frac{1}{\lambda_1^2 k},1\right),
\end{aligned} \tag{25}$$

where the first inequality holds by Assumption 2.1, and the second inequality holds by Lemma 4.4.

Therefore, plugging (24) and (25) into (23), we have

$$\begin{aligned}
&\mathbb{E}\left[\|\hat{V}\Lambda\hat{V}^\top - V\Lambda V^\top\|_F^2\right]\\
&\overset{\text{Eq. (23),(24),(25)}}{\leq} 64T\sum_{i=1}^{k}\sum_{i=j+1}^{d}\frac{(\lambda_i-\lambda_j)^2}{(\sigma_i-\max(\sigma_j,\sigma_{k+1}))^2}+32T^2\sum_{i=1}^{k}\left(\sum_{i=j+1}^{d}\frac{|\lambda_i-\lambda_j|}{(\sigma_i-\max(\sigma_j,\sigma_{k+1}))^2}\right)^2\\
&+\min(32,\ 32\lambda_1^2 k)\\
&\leq O\left(\sum_{i=1}^{k}\sum_{j=i+1}^{d}\frac{(\lambda_i-\lambda_j)^2}{(\sigma_i-\max(\sigma_j,\sigma_{k+1}))^2}+\left(\sum_{j=i+1}^{d}\frac{|\lambda_i-\lambda_j|}{(\sigma_i-\max(\sigma_j,\sigma_{k+1}))^2}\right)^2\right)\frac{\log(\frac{1}{\delta})}{\varepsilon^2},
\end{aligned} \tag{26}$$

where the last inequality holds since $T=\frac{\sqrt{2\log(\frac{1.25}{\delta})}}{\varepsilon}$.

Finally, we have by the Cauchy-Schwarz inequality and Assumption 2.1 that

$$\begin{aligned}
&\left(\sum_{j=i+1}^{d}\frac{|\lambda_i-\lambda_j|}{(\sigma_i-\max(\sigma_j,\sigma_{k+1}))^2}\right)^2=\left(\sum_{j=i+1}^{d}\frac{1}{|\sigma_i-\max(\sigma_j,\sigma_{k+1})|}\times\frac{|\lambda_i-\lambda_j|}{|\sigma_i-\max(\sigma_j,\sigma_{k+1})|}\right)^2\\
&\overset{\text{Cauchy-Schwarz inequality}}{\leq}\left(\sum_{j=i+1}^{d}\frac{1}{(\sigma_i-\max(\sigma_j,\sigma_{k+1}))^2}\right)\times\left(\sum_{j=i+1}^{d}\frac{(\lambda_i-\lambda_j)^2}{(\sigma_i-\max(\sigma_j,\sigma_{k+1}))^2}\right)\\
&\overset{\text{Assumption2.1}}{\leq}\left(\sum_{j=i+1}^{d}\frac{1}{(\sqrt{d})^2}\right)\times\left(\sum_{j=i+1}^{d}\frac{(\lambda_i-\lambda_j)^2}{(\sigma_i-\max(\sigma_j,\sigma_{k+1}))^2}\right)\\
&\leq\sum_{j=i+1}^{d}\frac{(\lambda_i-\lambda_j)^2}{(\sigma_i-\max(\sigma_j,\sigma_{k+1}))^2}.
\end{aligned} \tag{27}$$

In other words, (27) says that the first term inside the outer summation on the r.h.s. of (26) is at least as large as the second term. Therefore, plugging in (27) into (26), we have that

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

■

G Proof of Corollary 2.3

Proof of Corollary 2.3. To prove Corollary 2.3, we must bound the utility $\mathbb{E}[\|\hat{V}\hat{\Sigma}_k\hat{V}^\top - V\Sigma_k V^\top\|_F]$ of the post-processing of the Gaussian mechanism for the rank-k covariance matrix estimation problem. Towards this end, we first plug in $\lambda_i = \sigma_i$ for $i \leq k$ and $\lambda_i = 0$ for $i > k$ into Theorem 2.2 to obtain a bound for $\mathbb{E}[\|\hat{V}\Sigma_k\hat{V}^\top - V\Sigma_k V^\top\|_F]$ (Inequality (30)). We then apply Weyl's inequality (Lemma 3.3) together with a concentration bound for $\|B(T)\|_2$ (Lemma 4.3) to bound the perturbation to the eigenvalues of M when the Gaussian noise matrix $B(T)$ is added to M by the Gaussian mechanism. This implies a bound on $\mathbb{E}[\|\hat{V}\Sigma_k\hat{V}^\top - \hat{V}\hat{\Sigma}_k\hat{V}^\top\|_F]$ (Inequality (32)). Combining these two bounds (30) and (32), implies a bound on the utility $\mathbb{E}[\|\hat{V}\hat{\Sigma}_k\hat{V}^\top - V\Sigma_k V^\top\|_F]$ for the post-processing of the Gaussian mechanism (Inequality (33)).

Bounding the quanitity $\mathbb{E}[\|\hat{V}\Sigma_k\hat{V}^\top - V\Sigma_k V^\top\|_F]$. Let $\lambda_i = \sigma_i$ for $i \leq k$ and $\lambda_i = 0$ for $i > k$, and let $\Lambda := \mathrm{diag}(\lambda_1, \cdots, \lambda_d)$. Then by Assumption 2.1 we have that $\sigma_i - \sigma_{i+1} \geq \frac{8\sqrt{\log(\frac{1.25}{\delta})}}{\varepsilon}\sqrt{d} + c\log^{\frac{1}{2}}(\sigma_1 k)$. By Theorem 2.2, we have

$$\mathbb{E}\left[\|\hat{V}\Lambda\hat{V}^\top - V\Lambda V^\top\|_F^2\right] \leq O\left(\sum_{i=1}^{k}\sum_{j=i+1}^{d}\frac{(\lambda_i-\lambda_j)^2}{(\sigma_i - \max(\sigma_j, \sigma_{k+1}))^2}\right)\frac{\log(\frac{1}{\delta})}{\varepsilon^2}. \tag{28}$$

First, we note that $\lambda_i - \lambda_j \leq \sigma_i - \sigma_j$ for all $i \leq j \leq k$. Then for all $i < j \leq k$, we have

$$\frac{\lambda_i - \lambda_j}{\sigma_i - \max(\sigma_j, \sigma_{k+1})} \leq \frac{\sigma_i - \sigma_k}{\sigma_i - \max(\sigma_j, \sigma_{k+1})} \leq 1.$$

And for all $i \leq k < j \leq d$ we have

$$\frac{\lambda_i - \lambda_j}{\sigma_i - \max(\sigma_j, \sigma_{k+1})} = \frac{\sigma_i}{\sigma_i - \sigma_{k+1}} = \frac{\sigma_i - \sigma_k}{\sigma_i - \sigma_{k+1}} + \frac{\sigma_k}{\sigma_i - \sigma_{k+1}} \leq 1 + \frac{\sigma_k}{\sigma_k - \sigma_{k+1}}.$$

Thus, the summation term in (28) simplifies to

$$\sum_{i=1}^{k}\sum_{j=i+1}^{d}\frac{(\lambda_i-\lambda_j)^2}{(\sigma_i - \max(\sigma_j, \sigma_{k+1}))^2} \leq kd\left(1 + \frac{\sigma_k}{\sigma_k - \sigma_{k+1}}\right)^2 \leq 4kd\left(\frac{\sigma_k}{\sigma_k - \sigma_{k+1}}\right)^2. \tag{29}$$

Therefore, plugging (29) into (28), we have

$$\begin{aligned}\mathbb{E}[\|\hat{V}\Sigma_k\hat{V}^\top - V\Sigma_k V^\top\|_F] &\leq \sqrt{\mathbb{E}[\|\hat{V}\Sigma_k\hat{V}^\top - V\Sigma_k V^\top\|_F^2]} \\ &\overset{\text{Eq.(29),(28)}}{\leq} O\left(\sqrt{kd} \times \frac{\sigma_k}{\sigma_k - \sigma_{k+1}}\right)\frac{\log^{\frac{1}{2}}(\frac{1}{\delta})}{\varepsilon},\end{aligned} \tag{30}$$

where the first inequality holds by Jensen's inequality.

Bounding the perturbation to the eigenvalues. By Weyl's inequality (Lemma 3.3), we have that for every $i \in [d]$

$$\begin{aligned}
\mathbb{E}[(\hat{\sigma}_i - \sigma_i)^2] &\leq \mathbb{E}[(\|B(T)\|_2)^2] \\
&\leq 4\mathbb{E}[((T\sqrt{d})^2] + 4\mathbb{E}\left[\left(\|B(T)\|_2 - T\sqrt{d}\right)^2\right] \\
&\leq 4T^2 d + 4\int_0^\infty \mathbb{P}\left(\left(\|B(T)\|_2 - T\sqrt{d}\right)^2 > \alpha)\right) \mathrm{d}\alpha \\
&= 4T^2 d + 4\int_0^\infty \mathbb{P}\left(\|B(T)\|_2 - T\sqrt{d} > \sqrt{\alpha})\right) \mathrm{d}\alpha \\
&\overset{\text{Lemma(4.3)}}{\leq} 4T^2 d + 8\sqrt{\pi}\int_0^\infty e^{-\frac{1}{8}\frac{\alpha}{T^2}} \mathrm{d}\alpha \\
&= 4T^2 d + 64\sqrt{\pi}T^2 e^{-\frac{1}{8}\frac{\alpha}{T^2}}\Big|_{\alpha=0}^{\infty} \\
&= 4T^2 d + 64\sqrt{\pi}T^2 \\
&\leq 64\sqrt{\pi}\frac{\log(\frac{1}{\delta})}{\varepsilon^2} d. \qquad (31)
\end{aligned}$$

The first inequality holds by Weyl's inequality (Lemma 3.3), and the fourth inequality holds by Lemma 4.3. Therefore, (31) implies that,

$$\begin{aligned}
\mathbb{E}[\|\hat{V}\Sigma_k\hat{V}^\top - \hat{V}\hat{\Sigma}_k\hat{V}^\top\|_F] &= \mathbb{E}[\|\Sigma_k - \hat{\Sigma}_k\|_F] \\
&\leq \sqrt{\mathbb{E}[\|\Sigma_k - \hat{\Sigma}_k\|_F^2]} \\
&\overset{\text{Eq.(31)}}{\leq} O\left(\sqrt{kd}\frac{\log^{\frac{1}{2}}(\frac{1}{\delta})}{\varepsilon}\right), \qquad (32)
\end{aligned}$$

where the first inequality holds by Jensen's inequality, and the second inequality holds by (31). Thus, plugging (32) into (30), we have that

$$\begin{aligned}
\mathbb{E}[\|\hat{V}\hat{\Sigma}_k\hat{V}^\top - V\Sigma_k V^\top\|_F] &\leq \mathbb{E}[\|\hat{V}\Sigma_k\hat{V}^\top - V\Sigma_k V^\top\|_F] + \mathbb{E}[\|\hat{V}\Sigma_k\hat{V}^\top - \hat{V}\hat{\Sigma}_k\hat{V}^\top\|_F] \\
&\overset{\text{Eq.(30),(32)}}{\leq} O\left(\sqrt{kd} \times \frac{\sigma_k}{\sigma_k - \sigma_{k+1}}\right)\frac{\log^{\frac{1}{2}}(\frac{1}{\delta})}{\varepsilon}. \qquad (33)
\end{aligned}$$

Privacy: *Privacy of perturbed covariance matrix $\hat{M}$:* Recall that two matrices $M = A^\top A$ and $M' = A'^\top A'$ are said to be *neighbors* if they arise from $A, A' \in \mathbb{R}^{d\times n}$ which differ by at most one row, and that each row of the datasets A, A' has norm at most 1. In other words, we have that $M - M' = xx^\top$ for some $x \in \mathbb{R}^d$ such that $\|x\| \leq 1$. Define the sensitivity $S := \max_{M,M'\text{neighbors}} \|M - M'\|_{\ell_2}$, where $\|X\|_{\ell_2}$ denotes the Euclidean norm of the upper triangular entries of X (including the diagonal entries). Then we have

$$S = \max_{M,M'\text{neighbors}} \|M - M'\|_{\ell_2} \leq \|M - M'\|_F \leq \max_{\|x\|\leq 1} \|xx^\top\|_F = 1.$$

Then by standard results for the Gaussian Mechansim (e.g., by Theorem A.1 of [18]), we have that the Gaussian mechanism which outputs the upper triangular matrix $\hat{M}_{\text{upper}}$ with the same upper triangular entries as $\hat{M} = M + \frac{S\sqrt{2\log(\frac{1.25}{\delta})}}{\varepsilon}(G + G^\top) = M + \frac{\sqrt{2\log(\frac{1.25}{\delta})}}{\varepsilon}(G + G^\top)$, where G has i.i.d. $N(0,1)$ entries, is (ε,δ)-differentially private. However, since the perturbed matrix $\hat{M}$ is symmetric, it can be obtained from its upper triangular entries $\hat{M}_{\text{upper}}$ without accessing the original matrix M. Thus, the mechanism which outputs $\hat{M} = M + \frac{\sqrt{2\log(\frac{1.25}{\delta})}}{\varepsilon}(G + G^\top)$ must also be (ε,δ)-differentially private.

Privacy of rank-k approximation $\hat{V}\hat{\Sigma}_k\hat{V}^\top$: The mechnaism which outputs the rank-k approximation $\hat{V}\hat{\Sigma}_k\hat{V}^\top$ is (ε,δ)-differentially private, since $\hat{V}\hat{\Sigma}_k\hat{V}^\top$ is obtained by post-processing the perturbed matrix $\hat{M}$ without any additional access to the matrix M.

Namely, to obtain $\hat{V}\hat{\Sigma}_k\hat{V}^\top$, we first (i) compute the spectral decomposition $\hat{M} = \hat{V}\hat{\Sigma}\hat{V}^\top$. Next, (ii) we take the top-k eigenvalues $\hat{\sigma}_1, \ldots, \hat{\sigma}_k$ of $\hat{M}$, and set $\hat{\Sigma}_k = \mathrm{diag}(\hat{\sigma}_1, \ldots, \hat{\sigma}_k, 0, \ldots, 0)$. Finally, we output $\hat{M}_k := \hat{V}\hat{\Sigma}_k\hat{V}^\top$. Both of these steps (i) and (ii) are post-processing of $\hat{M}$ and do not require additional access to the matrix M. In particular, the eigenvalues $\hat{\sigma}_1, \ldots, \hat{\sigma}_k$ of $\hat{M}_k$ are obtained from the perturbed matrix $\hat{M}$, and thus do not compromise privacy. Therefore, the mechanism which outputs the rank-k approximation $\hat{M}_k := \hat{V}\hat{\Sigma}\hat{V}^\top$ must also be (ε, δ)-differentially private.

■

H Proof of Corollary 2.4

Proof of Corollary 2.4. To prove Corollary 2.4, we plug in $\lambda_1 = \cdots = \lambda_k = 1$ and $\lambda_{k+1} = \cdots = \lambda_d = 0$ to Theorem 2.2. Corollary 2.4 considers two cases. In the first case (referred to here as Case I), the eigenvalues σ of the input matrix M satisfies Assumption 2.1 In the second case (referred to here as Case II) the eigenvalues σ of M also satisfy both Assumption 2.1 as well as the lower bound $\sigma_i - \sigma_{i+1} \geq \Omega(\sigma_k - \sigma_{k+1})$ for all $i \leq k$. We derive a bound on the utility $\mathbb{E}[\|\hat{V}_k\hat{V}_k^\top - V_kV_k^\top\|_F]$ in each case separately.

Case I: M satisfies Assumption 2.1.

Plugging in $\lambda_1 = \cdots = \lambda_k = 1$ and $\lambda_{k+1} = \cdots = \lambda_d = 0$ to Theorem 2.2 we get that, since M satisfies Assumption 2.1 for $(M, k, 2, \varepsilon, \delta)$,

$$\begin{aligned}
\mathbb{E}\left[\|\hat{V}_k\hat{V}_k^\top - V_kV_k^\top\|_F^2\right] &= \mathbb{E}\left[\|\hat{V}\Lambda\hat{V}^\top - V\Lambda V^\top\|_F^2\right] \\
&\overset{\text{Theorem 2.2}}{\leq} O\left(\sum_{i=1}^{k}\sum_{j=i+1}^{d}\frac{(\lambda_i - \lambda_j)^2}{(\sigma_i - \max(\sigma_j, \sigma_{k+1}))^2}\right)\frac{\log(\frac{1}{\delta})}{\varepsilon^2} \\
&= O\left(\sum_{i=1}^{k}\sum_{j=k+1}^{d}\frac{1}{(\sigma_i - \max(\sigma_j, \sigma_{k+1}))^2}\right)\frac{\log(\frac{1}{\delta})}{\varepsilon^2} \\
&\leq O\left(\sum_{i=1}^{k}\sum_{j=k+1}^{d}\frac{1}{(\sigma_k - \sigma_{k+1})^2}\right)\frac{\log(\frac{1}{\delta})}{\varepsilon^2} \\
&= O\left(\frac{kd}{(\sigma_k - \sigma_{k+1})^2}\frac{\log(\frac{1}{\delta})}{\varepsilon^2}\right), \qquad (34)
\end{aligned}$$

where the first inequality holds by Theorem 2.2, and the second equality holds since $\lambda_1 = \cdots = \lambda_k = 1$ and $\lambda_{k+1} = \cdots = \lambda_d = 0$.

Thus, applying Jensen's Inequality to Inequality (34), we have that

$$\mathbb{E}[\|\hat{V}_k\hat{V}_k^\top - V_kV_k^\top\|_F] \leq O\left(\frac{\sqrt{kd}}{(\sigma_k - \sigma_{k+1})}\frac{\log^{\frac{1}{2}}(\frac{1}{\delta})}{\varepsilon}\right).$$

Case II: M satisfies Assumption 2.1 and $\sigma_i - \sigma_{i+1} \geq \Omega(\sigma_k - \sigma_{k+1})$ for all $i \leq k$.

Plugging in $\lambda_1 = \cdots = \lambda_k = 1$ and $\lambda_{k+1} = \cdots = \lambda_d = 0$ to Theorem 2.2 we get that, since M satisfies Assumption 2.1 for $(M, k, 2, \varepsilon, \delta)$,

$$\begin{aligned}
\mathbb{E}\left[\|\hat{V}_k\hat{V}_k^\top - V_kV_k^\top\|_F^2\right] &= \mathbb{E}\left[\|\hat{V}\Lambda\hat{V}^\top - V\Lambda V^\top\|_F^2\right] \\
&\overset{\text{Theorem 2.2}}{\leq} O\left(\sum_{i=1}^{k}\sum_{j=i+1}^{d}\frac{(\lambda_i-\lambda_j)^2}{(\sigma_i-\max(\sigma_j,\sigma_{k+1}))^2}\right)\frac{\log(\frac{1}{\delta})}{\varepsilon^2} \\
&= O\left(\sum_{i=1}^{k}\sum_{j=k+1}^{d}\frac{1}{(\sigma_i-\max(\sigma_j,\sigma_{k+1}))^2}\right)\frac{\log(\frac{1}{\delta})}{\varepsilon^2} \\
&\leq O\left(\sum_{i=1}^{k}\sum_{j=k+1}^{d}\frac{1}{(i-k-1)^2(\sigma_k-\sigma_{k+1})^2}\right)\frac{\log(\frac{1}{\delta})}{\varepsilon^2} \\
&\leq O\left(\sum_{i=1}^{k}\frac{d}{(i-k-1)^2(\sigma_k-\sigma_{k+1})^2}\right)\frac{\log(\frac{1}{\delta})}{\varepsilon^2} \\
&\leq O\left(\frac{d}{(\sigma_k-\sigma_{k+1})^2}\frac{\log(\frac{1}{\delta})}{\varepsilon^2}\sum_{i=1}^{k}\frac{1}{i^2}\right) \\
&\leq O\left(\frac{d}{(\sigma_k-\sigma_{k+1})^2}\frac{\log(\frac{1}{\delta})}{\varepsilon^2}\right), \qquad (35)
\end{aligned}$$

where the first inequality holds by Theorem 2.2, the second equality holds since $\lambda_1 = \cdots = \lambda_k = 1$ and $\lambda_{k+1} = \cdots = \lambda_d = 0$, the second inequality holds since $\sigma_i - \sigma_{i+1} \geq \Omega(\sigma_k - \sigma_{k+1})$ for all $i \leq k$, and the last inequality holds since $\sum_{i=1}^{k}\frac{1}{i^2} \leq \sum_{i=1}^{\infty}\frac{1}{i^2} = O(1)$.

Thus, applying Jensen's Inequality to Inequality (35), we have that

$$\mathbb{E}[\|\hat{V}_k\hat{V}_k^\top - V_kV_k^\top\|_F] \leq O\left(\frac{\sqrt{d}}{(\sigma_k-\sigma_{k+1})}\frac{\log^{\frac{1}{2}}(\frac{1}{\delta})}{\varepsilon}\right).$$

■

I Eigenvalue Gaps of Wishart Matrices

In this section, we provide the results of numerical simulations where we compute the minimum eigenvalue gap, $\min_{i\in[d-1]}\sigma_i - \sigma_{i+1}$ of Wishart random matrices $W = A^\top A$, where A is an $m \times d$ matrix with i.i.d. $N(0,1)$ Gaussian entries, for various values of d and m. The goal of these simulations is to evaluate for what values of m, d the eigenvalue gaps in a Wishart random matrix W satisfy Assumption 2.1$(W, k, \lambda_1, \varepsilon, \delta)$, which requires that the size of the top-$k+1$ eigenvalues of the matrix W satisfy $\sigma_i - \sigma_{i+1} \geq \frac{8\sqrt{\log(\frac{1.25}{\delta})}}{\varepsilon}\sqrt{d} + c\log^{\frac{1}{2}}(\lambda_1 k)$ for every $i \in [k]$.

We observe that, when d is held constant and m is increased, the minimum eigenvalue gap size $\min_{i\in[d-1]}\sigma_i - \sigma_{i+1}$ grows roughly proportional to $\sqrt{m}$ (Figure 1). Moreover, if we fix m to be $m = d^3$ and increase d, we observe that the minimum eigenvalue gap is at least as large as $\sqrt{d}$ with high probability and grows roughly proportional to $\sqrt{d}$ (Figure 2). Thus, we expect Wishart random matrices to satisfy Assumption 2.1$(W, k, \lambda_1, \varepsilon, \delta)$ with high probability as long as $m \geq \Omega(\log(\frac{1.25}{\delta})\varepsilon^2 \times d^3)$, for any $k \leq d$ and $\varepsilon, \delta > 0$ where, e.g., $\lambda_1 \leq O(d)$. In particular, we note that in the application of our main result to subspace recovery we set $\lambda_1 = 1$ (Corollary 2.4), and in the application of our main result to rank-k covariance matrix approximation we set $\lambda_1 = \sigma_1$ (Corollary 2.3) and thus have that $\lambda_1 \leq O(\sqrt{d})$ with high probability by the concentration bounds in Lemma 3.2. All simulations were run on Matlab.

Finally, we note that there is a long line of work in random matrix theory which provides results about the distributions of the eigenvalues of random matrix ensembles (see e.g.

[35, 20, 4, 42, 40, 21, 27]), including for Wishart random matrices [38, 44, 41]. For instance, results are given in [41, 46] for the local eigenvalue statistics of families of Wishart matrices where $\frac{d}{m} \to y$ as $d \to \infty$ for any fixed constant $0 < y \leq 1$. In particular, these works include results which give high-probability bounds on the minimum eigenvalue gap of this class of Wishart matrices (see Theorems 16 and 18 in [41], and also Theorem 1.7 of [46] who extend results of [41] to the edge of the spectrum). Results for eigenvalue gap probabilities of other random matrix ensembles are given, e.g., in [40, 23, 22, 9]. However, to the best of our knowledge, we are not aware of any bounds for the minimum eigenvalue gap of families of Wishart matrices where $\frac{m}{d}$ does not converge to a constant as $d \to \infty$. While it may be possible to extend the analysis given in [41] to obtain bounds for the minimum eigenvalue gap for families of Wishart matrices where $\frac{m}{d}$ is polynomial in d, this analysis would be beyond the scope of our paper.

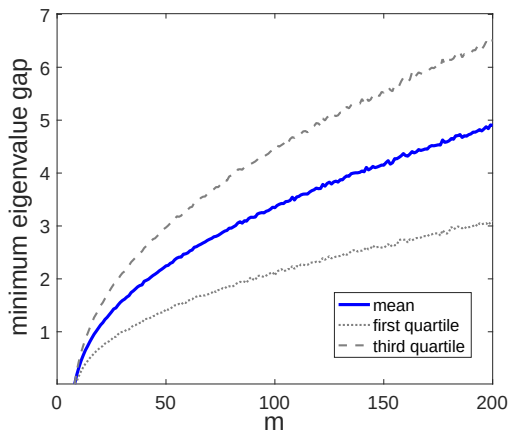

Figure 1: The minimum eigenvalue gap, $\min_{i \in [d-1]} \sigma_i - \sigma_{i+1}$ of Wishart random matrices $W = A^\top A$, where A is an $m \times d$ matrix with i.i.d. $N(0,1)$ Gaussian entries, for various values of m and $d = 10$, averaged over 10,000 trials for each m (blue curve), with first quartile (dashed grey curve) and third quartile (dotted grey curve) also displayed. We observe that the minimum eigenvalue gap size grows roughly proportional to $\sqrt{m}$.

J Eigenvalue gaps in real datasets

In this section we compute the eigenvalues of covariance matrices M of standard real-world datasets, and determine the values of k for which our Assumption 2.1$(M, k, \lambda_1, \varepsilon, \delta)$ holds on these datasets (for $\lambda_1 = \sigma_1$, $\varepsilon = 1$, and $\delta = \frac{1}{100}$) . We consider three standard datsets from the UCI Machine Learning Repository [14]: the 1990 US Census dataset ($d = 124$, $n = 2458285$), the KDD Cup dataset ($d = 36$, $n = 494020$), and the Adult dataset ($d = 6$, $n = 48842$). All three of these datasets were previously used as benchmarks in the differentially private matrix approximation and PCA literature (Census and KDD Cup in e.g. [11], and Adult in e.g. [2]).

As is standard, we pre-process each dataset to ensure that all entries are real-valued and to normalize the range of the measurements used for the different features. Specifically, we remove categorical features and apply min-max normalization to the remaining real-valued features. We then multiply the data matrix by a constant to ensure that all its rows have magnitude at most 1 (to ensure the sensitivity bounds which imply privacy of the Gaussian mechanism [18] hold on the dataset), and subtract the mean of each column. We then compute the eigenvalues of the covariance matrix $M = A^\top A$ of the pre-processed data matrix A.

Our Assumption 2.1$(M, k, \lambda_1, \varepsilon, \delta)$ requires that $\sigma_i - \sigma_{i+1} \geq \frac{8\sqrt{\log(\frac{1.25}{\delta})}}{\varepsilon}\sqrt{d} + 3\log^{\frac{1}{2}}(\lambda_1 k)$ for all $i \leq k$.

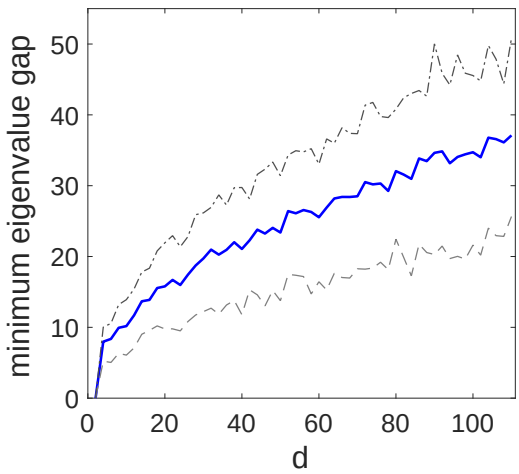

Figure 2: The minimum eigenvalue gap, $\min_{i\in[d-1]} \sigma_i - \sigma_{i+1}$ of Wishart random matrices $W = A^\top A$, where A is an $m \times d$ matrix with i.i.d. $N(0,1)$ Gaussian entries, for various values of d and $m = d^3$, averaged over 200 trials for each d (blue curve), with first quartile (dashed grey curve) and third quartile (dotted grey curve) also displayed. We observe that, if $m = d^3$, the minimum gap size is at least as large as $\sqrt{d}$ with high probability, and grows roughly proportional to $\sqrt{d}$.

As is done in prior works, we consider values of k such that the (non-private) rank-k approximation of M has Frobenius norm which is a large percentage of $\|M\|_F$ (e.g., in [11] they choose values of k such that the Frobenius norm of the low-rank approximation is at least 80% of the original matrix). We will verify that, on the above-mentioned datasets, our Assumption 2.1 holds for values of k large enough such that the (non-private) rank-k approximation contains at least 99% of the Frobenius norm of M.

We first compute the eigenvalues of the covariance matrix of the Census dataset (Figure 3, left). On this dataset, the Frobenius norm of the (non-private) rank-k approximation for $k = 10$ contains $> 99\%$ of the Frobenius norm of M; thus we would like our assumption to hold for values of k for $k \leq 10$. For the census dataset we have $\sigma_1 = 93730$, and thus, for any $k \leq 11$, the r.h.s. of Assumption 2.1 is at most 442. We observe that the eigenvalue gaps satisfy $\sigma_i - \sigma_{i+1} \geq 442$ for all $i \leq 11$ (Figure 3, right); thus, our Assumption 2.1 is satisfied for all $k \leq 11$.

Computing the eigenvalues of the covariance matrix of the KDD Cup dataset (Figure 4, left) we observe that the Frobenius norm of the (non-private) rank-k approximation for $k = 3$ contains $> 99\%$ of the Frobenius norm of M; thus we would like our assumption to hold for values of k at least 3. On this dataset we have $\sigma_1 = 72670$, and thus, for any $k \leq 7$, the r.h.s. of Assumption 2.1 is at most 250. We observe that the eigenvalue gaps satisfy $\sigma_i - \sigma_{i+1} \geq 250$ for $i \leq 7$ (Figure 4, right); thus, Assumption 2.1 is satisfied for all $k \leq 7$ on this dataset.

Computing the eigenvalues of the Adult dataset (Figure 5, left) we observe that the Frobenius norm of the (non-private) rank-k approximation for $k = 4$ contains 99% of the Frobenius norm of M. On this dataset we have that $\sigma_1 = 1195$, and thus, for any $k \leq 4$, the r.h.s. of Assumption 2.1 is at most 103.4. We observe that the eigenvalue gaps satisfy $\sigma_i - \sigma_{i+1} \geq 103.4$ for $i \leq 4$ (Figure 5, right); thus, Assumption 2.1 is satisfied for all $k \leq 4$ on this dataset.

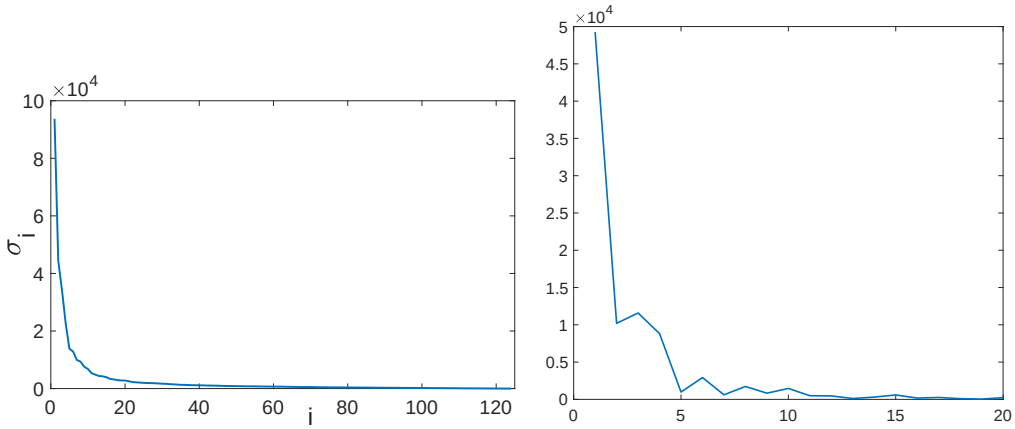

Figure 3: Eigenvalues (left) and eigenvalue gaps (right) of the covariance matrix of the Census dataset. The gaps in the eigenvalues satisfy Assumption 2.1 for any $k \leq 11$, and $\lambda_1 = \sigma_1$, $\varepsilon = 1$, $\delta = \frac{1}{100}$, as for these values Assumption 2.1 requires that $\sigma_i - \sigma_{i+1} \geq 442$ for all $i \leq k$.

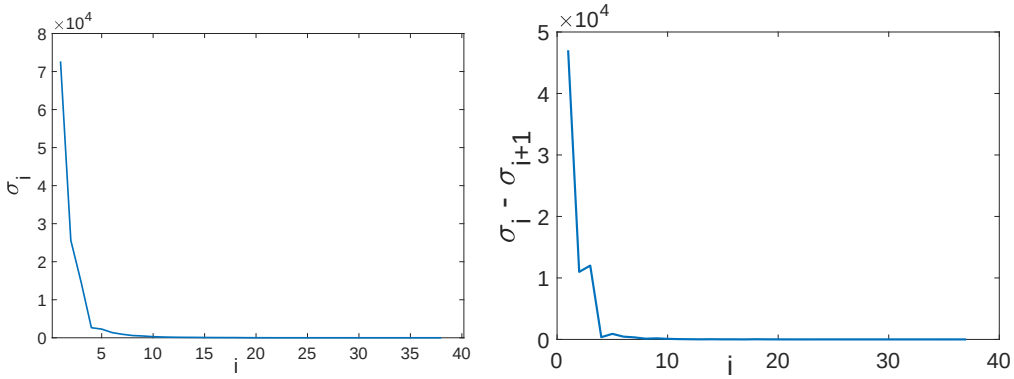

Figure 4: Eigenvalues (left) and eigenvalue gaps (right) of the covariance matrix of the KDD Cup dataset. The gaps in the eigenvalues satisfy Assumption 2.1 for any $k \leq 7$, and $\lambda_1 = \sigma_1$, $\varepsilon = 1$, $\delta = \frac{1}{100}$, as for these values Assumption 2.1 is satisfied when $\sigma_i - \sigma_{i+1} \geq 250$ for all $i \leq k$.

K Challenges in using previous approaches

In the special case of covariance matrix approximation, it is possible to use trace inequalities to bound the quantity $\|M - \hat{V}_k \hat{\Sigma}_k \hat{V}_k^\top\|_F - \|M - V_k \Sigma_k V_k^\top\|_F$ (which is bounded above by the quantity $\|\hat{V}\hat{\Sigma}_k\hat{V}^\top - V\Sigma_k V^\top\|_F$ we bound). This is the approach taken in [19], who apply the fact that

$$\mathrm{tr}(X) \leq \mathrm{rank}(X)\|X\|_2 \qquad \forall X \in \mathbb{R}^{d\times d} \tag{36}$$

to show that

$$\|M - \hat{V}_k \hat{\Sigma}_k \hat{V}_k^\top\|_F^2 - \|M - V_k \Sigma_k V_k^\top\|_F^2 \leq O(k\|M - V_k\Sigma_k V_k^\top\|_2\|E\|_2 + k\|E\|_2^2).$$

The r.h.s. depends on $\sigma_{k+1} = \|M - V_k\Sigma_k V_k^\top\|_2$, and is therefore not invariant to scalar multiplications of M. However, one can obtain a scalar-invariant bound on the quantity $\|M - \hat{V}_k \hat{\Sigma}_k \hat{V}_k^\top\|_F - \|M - V_k \Sigma_k V_k^\top\|_F$ by plugging in $\|M - V_k\Sigma_k V_k^\top\|_2 \leq \|M - V_k\Sigma_k V_k^\top\|_F$ and plugging in the high-probability bound $\|E\|_2 = O(\sqrt{d})$. This leads to a bound of $\|M - \hat{V}_k \hat{\Sigma}_k \hat{V}_k^\top\|_F - \|M - V_k \Sigma_k V_k^\top\|_F \leq O(k\sqrt{d})$. In the special case where $k = d$, this bound is $O(d^{1.5})$, and thus is not tight since we have $\|M - \hat{V}_k \hat{\Sigma}_k \hat{V}_k^\top\|_F - \|M - V_k \Sigma_k V_k^\top\|_F = \|\hat{M} - M\|_F = \|E\|_F = O(d)$ w.h.p. Roughly, the additional factor of $\sqrt{k} = \sqrt{d}$ incurred in

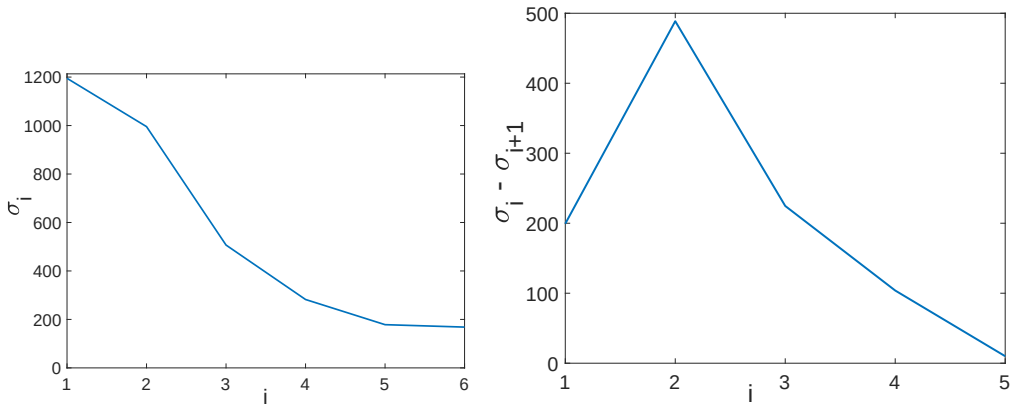

Figure 5: Eigenvalues (left) and eigenvalue gaps (right) of the covariance matrix of the Adult dataset. The gaps in the eigenvalues satisfy Assumption 2.1 for any $k \leq 4$, and $\lambda_1 = \sigma_1$, $\varepsilon = 1$, $\delta = \frac{1}{100}$, as for these values Assumption 2.1 is satisfied when $\sigma_i - \sigma_{i+1} \geq 103.4$ for all $i \leq k$.

their bound is due to the fact that the matrix trace inequality (36) their analysis relies on gives a bound in terms of the spectral norm, even though they only need a bound in terms of the Frobenius norm– which can (in the worst case) be larger than the spectral norm by a factor of $\sqrt{k}$.

Another issue is that the quantity $\|\hat{V}\Sigma_k\hat{V}^\top - V\Sigma_k V^\top\|_F$ we wish to bound can be very sensitive to perturbations to V_k, since $\|\hat{V}\Lambda\hat{V}^\top - V\Lambda V^\top\|_F \geq \lambda_k\|\hat{V}_k\hat{V}_k^\top - V_kV_k^\top\|_F$. Thus, any bound on $\|\hat{V}\Lambda\hat{V}^\top - V\Lambda V^\top\|_F$ must (at the very least) also bound the distance $\|\hat{V}_k\hat{V}_k^\top - V_kV_k^\top\|_F$ between the projection matrices onto the subspace $\mathcal{V}_k$ spanned by the top-k eigenvectors of M. One approach to bounding $\|\hat{V}_k\hat{V}_k^\top - V_kV_k^\top\|_F$, is to use an eigenvector perturbation theorem, such as the Davis-Kahan theorem [13], which says, roughly, that

$$\|\hat{V}_k\hat{V}_k^\top - V_kV_k^\top\|_2 \leq \frac{\|E\|_2}{\sigma_k - \sigma_{k+1}} \tag{37}$$

(this is the approach taken by [19] when proving their utility bounds for subspace recovery). Plugging in the high-probability bound $\|E\|_2 = O(\sqrt{d})$, and using the fact that $\|\hat{V}_k\hat{V}_k^\top - V_kV_k^\top\|_F \leq \sqrt{k}\|\hat{V}_k\hat{V}_k^\top - V_kV_k^\top\|_2$, gives $\|\hat{V}_k\hat{V}_k^\top - V_kV_k^\top\|_F \leq \frac{\sqrt{k}\sqrt{d}}{\sigma_k - \sigma_{k+1}}$ with high probability. To obtain bounds for the utility $\|\hat{V}\Sigma_k\hat{V}^\top - V\Sigma_kV^\top\|_F$ of the covariance matrix approximation, we can decompose $V\Sigma_kV^\top = \sum_{i=1}^k(\sigma_i - \sigma_{i+1})V_iV_i^\top$, and apply the Davis-Kahan theorem to each projection matrix $V_iV_i^\top$:

$$\begin{aligned}
\|\hat{V}\Sigma_k\hat{V}^\top - V\Sigma_kV^\top\|_F &= \|\hat{V}\Sigma_k\hat{V}^\top - V\Sigma_kV^\top\|_F && (38)\\
&= \|\sum_{i=1}^{k-1}(\sigma_i - \sigma_{i+1})(\hat{V}_i\hat{V}_i^\top - V_iV_i^\top) + \sigma_k(\hat{V}_k\hat{V}_k^\top - V_kV_k^\top)\|_F && (39)\\
&\leq \sum_{i=1}^{k-1}(\sigma_i - \sigma_{i+1})\|\hat{V}_i\hat{V}_i^\top - V_iV_i^\top\|_F + \sigma_k\|\hat{V}_k\hat{V}_k^\top - V_kV_k^\top\|_F && (40)\\
&\leq \sum_{i=1}^{k-1}(\sigma_i - \sigma_{i+1})\frac{\sqrt{i}\sqrt{d}}{\sigma_i - \sigma_{i+1}} + \sigma_k\frac{\sqrt{k}\sqrt{d}}{\sigma_k - \sigma_{k+1}} && (41)\\
&= O\left(k^{1.5}\sqrt{d} + \frac{\sigma_k}{\sigma_k - \sigma_{k+1}}\sqrt{k}\sqrt{d}\right), && (42)
\end{aligned}$$

where we define $\sigma_{d+1} := \sigma_d$. Unfortunately, this bound is not tight up to a factor of k, at least in the special case where $k = d$. As a first step to obtaining a tighter bound, we would ideally

like to add up the Frobenius norm of the summands $(\sigma_i - \sigma_{i+1})(\hat{V}_i\hat{V}_i^\top - V_iV_i^\top)$ as a sum-of-squares rather than as a simple sum, in order to decrease the r.h.s. by a factor of $\sqrt{k}$. However, to do so we would need to bound the cross-terms $\mathrm{tr}\left((\hat{V}_i\hat{V}_i^\top - V_iV_i^\top)(\hat{V}_j\hat{V}_j^\top - V_jV_j^\top)\right)$ for $i \neq j$. To bound each of these cross-terms we need to carefully track the interactions between the eigenvectors in the subspaces $\mathcal{V}_i$ and $\mathcal{V}_j$ as the noise E is added to the input matrix M.

We handle such cross-terms by viewing the addition of noise as a continuous-time matrix diffusion $\Psi(t) = M + B(t)$, whose eigenvalues $\gamma_i(t)$ and eigenvectors $u_i(t)$, $i \in [d]$, evolve over time. This allows us to "add up" contributions of different eigenvectors to the Frobenius distance as a stochastic integral,

$$||\hat{V}\Sigma_k\hat{V}^\top - V\Sigma_kV^\top||_F^2 = ||\int_0^T \sum_{i=1}^{d}\sum_{j\neq i} |\sigma_i - \sigma_j| \frac{\mathrm{d}B_{ij}(t)}{\gamma_i(t) - \gamma_j(t)}(u_i(t)u_j^\top(t) + u_j(t)u_i^\top(t))||_F^2,$$

where, roughly, each differential cross term $\frac{\mathrm{d}B_{ij}(t)}{\gamma_i(t)-\gamma_j(t)}(u_i(t)u_j^\top(t) + u_j(t)u_i^\top(t))$ adds noise to the matrix $V\Sigma_kV^\top$ independently of the other terms since the Brownian motion differentials $\mathrm{d}B_{ij}(t)$ are independent for all i, j, t. (16). Roughly, this allows us to add up the contributions of these terms as a sum of squares using Itô's Lemma from stochastic calculus (19).

L Necessity of Assumption 2.1 in our proof

For simplicity, assume that $\varepsilon = \delta = O(1)$. Our proof uses Weyl's inequality to bound the gaps in the eigenvalues $\gamma_i(t) - \gamma_{i+1}(t)$ of the perturbed matrix $M + G(t) + G^\top(t)$ at every time $t \in [0, 1]$, where $G(t)$ has i.i.d. $N(0, t)$ entries.

Weyl's inequality says that for every $i \in d$, $\sigma_i - \|G(t) + G^\top(t)\|_2 \leq \gamma_i(t) \leq \sigma_i + \|G(t) + G^\top(t)\|_2$. Thus, plugging in the high-probability bound $\|G(t)\|_2 \leq 2\sqrt{d}$, we have that

$$\gamma_i(t) - \gamma_{i+1}(t) \geq \sigma_i - \sigma_{i+1} - 4\sqrt{d}. \tag{43}$$

If $\sigma_i - \sigma_{i+1} < \sqrt{d}$, Weyl's inequality does not give *any* bound on the gaps since then the r.h.s. of (43) is $\sigma_i - \sigma_{i+1} - 4\sqrt{d}$ is negative. Thus, to apply Weyl's inequality to bound the eigenvalue gaps $\gamma_i(t) - \gamma_{i+1}(t)$, we require that $\sigma_i - \sigma_{i+1} \geq \Omega(\sqrt{d})$, which is roughly Assumption 2.1.

On the other hand, we note that Weyl's inequality is a worst-case deterministic inequality– it says that $\sigma_i - \|G(t) + G^\top(t)\|_2 \leq \gamma_i(t) \leq \sigma_i + \|G(t) + G^\top(t)\|_2$ with probability 1. However, $G(t)$ is a random matrix, and the Dyson Brownian motion equations (3) which govern the evolution of the eigenvalues of the perturbed matrix $M + G(t) + G^\top(t)$ say that the eigenvalues $\gamma_i(t)$ and $\gamma_j(t)$ repel each other with a "force" of magnitude $\frac{1}{\gamma_i(t)-\gamma_j(t)}$. Thus, on average, we conjecture that the gaps $\gamma_i(t) - \gamma_j(t)$ between any two eigenvalues becomes larger over time, and hence that $\mathbb{E}[\gamma_i(t) - \gamma_j(t)] \geq \sigma_i - \sigma_j$ regardless of the size of the initial gap $\sigma_i - \sigma_j$. This suggests that it may be possible to weaken (or eliminate) Assumption 2.1, while still recovering the same bound in our main result Theorem 2.2.

M High probability bounds

While our current result holds in expectation, it is possible to use our techniques to prove high-probability bounds.

The simplest approach is to plug in the expectation bound in our main result (Theorem 2.2) into Markov's inequality, which says that $P(\|\hat{V}\Lambda\hat{V}^\top - V\Lambda V^\top\|_F^2 \geq s) \leq \frac{E(\|\hat{V}\Lambda\hat{V}^\top - V\Lambda V^\top\|_F^2)}{s}$ for all $s > 0$.

While Markov's inequality gives a high-probability bound, this bound decays as $\frac{1}{s}$. One approach to obtaining high-probability bounds which decay with rate exponential in s might

be to apply concentration inequalities to the part of our proof where we currently use expectations. Namely, our proof of Lemma 4.5 in the appendix uses Itô's Lemma (restated as Lemma 3.1 in the appendix) to show that, roughly,

$$\|\Psi(T)-\Psi(0)\|_F^2 = 4\int_0^T \sum_{i=1}^d \sum_{j\neq i} \frac{(\lambda_i-\lambda_j)^2}{(\gamma_i(t)-\gamma_j(t))^2} + \frac{1}{2}\int_0^t \sum_{\ell,r}\sum_{\alpha,\beta}\left(\frac{\partial}{\partial X_{\alpha\beta}} f(X(t))\right) R_{(\ell r)(\alpha\beta)}(t)\mathrm{d}B_{\ell r}(t),$$

where we define $f(X) := \|X\|_F^2$, $X(t) := \int_0^t \sum_{i=1}^d \sum_{j\neq i} |\lambda_i - \lambda_j| \frac{\mathrm{d}B_{ij}(s)}{|\gamma_i(s)-\gamma_j(s)|}(u_i(s)u_j^\top(s) + u_j(s)u_i^\top(s))$, and $R_{(\ell r)(ij)}(t) := \left(\frac{|\lambda_i-\lambda_j|}{|\gamma_i(t)-\gamma_j(t)|}(u_i(t)u_j^\top(t) + u_j(t)u_i^\top(t))\right)[\ell, r]$.

The two integrals on the r.h.s. are both random variables. For simplicity, our current proof bounds these random variables by taking the expectation of both sides of the equation. In particular, the second term on the r.h.s. has mean 0, and thus vanishes when we apply the expectation.

We do not have to do any additional work to bound the first term on the r.h.s. with high probability, since the only random variables appearing in that term are the eigenvalue gaps $\gamma_i(t) - \gamma_j(t)$, and we have already shown a high-probability bound on these gaps (Lemma 4.4). To bound the second term on the r.h.s. with high probability, in addition to the gaps $\gamma_i(t) - \gamma_j(t)$, we would also need to deal with the Gaussian random variables $B_{\ell r}(t)$ appearing inside the integral. One approach to bounding these random variables $B_{\ell r}(t)$ with high probability might be to apply standard Gaussian concentration inequalities, and it would be interesting to see whether this leads to high-probability bounds which are similar to the expectation bounds we obtain.