Rigorous Guarantees for Tyler's M-Estimator via Quantum Expansion

W. Cole Franks   FRANKS@MIT.EDU

Ankur Moitra   MOITRA@MIT.EDU

Editors: Jacob Abernethy and Shivani Agarwal

Abstract

Estimating the shape of an elliptical distribution is a fundamental problem in statistics. One estimator for the shape matrix, Tyler's M-estimator, has been shown to have many appealing asymptotic properties. It performs well in numerical experiments and can be quickly computed in practice by a simple iterative procedure. Despite the many years the estimator has been studied in the statistics community, there was neither a non-asymptotic bound on the rate of the estimator nor a proof that the iterative procedure converges in polynomially many steps.

Here we observe a surprising connection between Tyler's M-estimator and operator scaling, which has been intensively studied in recent years in part because of its connections to the Brascamp-Lieb inequality in analysis. We use this connection, together with novel results on quantum expanders, to show that Tyler's M-estimator has the optimal rate up to factors logarithmic in the dimension, and that in the generative model the iterative procedure has a linear convergence rate even without regularization.

Keywords: Elliptical distribution, operator scaling, M-estimator

1. Introduction

The covariance matrix Σ of a joint random variable X is a fundamental object in statistics which encodes useful information about the geometry of the distribution of X. Estimation of the covariance matrix is a central task in data analysis, and in many situations the *sample covariance matrix* is a good estimator. However, heavy-tailed random variables need not have a covariance matrix at all, and even when the covariance matrix exists the sample covariance matrix need not converge at all to the true one.

Elliptical distributions Kelker (1970); Cambanis et al. (1981) are a well-studied class of random variables used to model heavy-tailed data Gupta et al. (2013). Though elliptical distributions need not have covariance matrices, they are characterized by parameter called the *shape matrix* with a similar geometric interpretation. Tyler Tyler (1987) defined an estimator, known as *Tyler's M-estimator*, for the shape matrix of an elliptical distribution and proposed an iterative procedure to compute it. Furthermore, he established many powerful and surprising statistical properties for it. First, it is strongly consistent and asymptotically normal. In fact, its asymptotic distribution is *distribution-free* in the sense that it that it does not depend on which elliptical distribution is generating the data. Second, it is the most robust estimator for the shape of an elliptical distribution in the sense that it minimizes the maximum asymptotic variance.

There are some nonasymptotic bounds on the sample complexity Soloveychik and Wiesel (2014), but these bounds are only comparable to the Gaussian case when error is measured in Frobenius norm. Moreover, the error has additional factors depending on the condition number of the shape matrix. From the computational standpoint, there are provably efficient algorithms for *regularized* versions of the estimator Goes et al. (2017), but the regularized versions do not inherit the appealing statistical properties of Tyler's M-estimator. See the survey Wiesel and Zhang (2015) for more information on Tyler's M-estimator and its regularized variants. Despite the significant attention the problem has received, the following are absent from the existing literature:

1. Necessary and sufficient conditions for the existence and uniqueness of Tyler's M-estimator.

2. A non-asymptotic upper bound on the sample complexity for estimation of the shape matrix in spectral norm comparable to the Gaussian case.

3. A rigorous proof that Tyler's iterative procedure converges at a linear rate.

Here we close all these gaps simultaneously by making a new connection, which surprisingly has gone unnoticed for decades, between Tyler's M-estimator and *operator scaling* Gurvits (2003). We describe the setup for operator scaling in Section 2.2, but the name roughly refers to a group of problems generalizing Sinkhorn's classic "matrix scaling" problem in which one seeks to obtain a doubly stochastic matrix by rescaling the rows and columns of a given nonnegative matrix. Tyler's M-estimator arises in operator scaling because it is precisely the "rescaling" for a certain operator constructed from the samples.

Though the name is fairly new, operator scaling was studied far earlier in the context of *geometric invariant theory* in algebra. As we will discuss in Section 2.1, necessary and sufficient conditions for the existence of Tyler's M-estimator follow from the work King (1994) in this field. Later, the authors of Gurvits and Samorodnitsky (2002); Forster (2002); Barthe (1998) independently studied Eq. (4) for applications in convex geometry, communication complexity, and real analysis, respectively. It was shown in Hardt and Moitra (2013), and rather implicitly in Gurvits and Samorodnitsky (2002), that there is an algorithm to solve Eq. (4) up to error ε in time polynomial in $\log(1/\varepsilon)$ and the bit-length of the samples, though both are very slow due to their use of the ellipsoid algorithm. Next, an iterative procedure for operator scaling was proposed in Gurvits (2003), thus implicitly showing that Tyler's iterative procedure converges in time polynomial in $1/\varepsilon$ and the bit-length of the samples. Garg et al. (2016) proved the same guarantees in a significantly more general setting, and used them to obtain new upper bounds in algebraic complexity and then in Garg et al. (2017a) to compute the Brascamp-Lieb constant in analysis.

Clearly a great deal of information about Tyler's M-estimator can be gleaned from the existing operator scaling literature, but there remain a few hurdles to Items 2 and 3. Firstly, it is unclear how well Tyler's estimator in a statistical sense in terms of proving finite sample guarantees. Secondly, results implying *linear* $(\log \varepsilon^{-1})$ convergence of iterative procedures like that of Tyler are rare, and the existing results are not explicit enough to produce polynomial time algorithms Soules (1991); Knight (2008) in the sense that they do not specify the norm in which they obtain convergence. We clear these remaining hurdles through a somewhat surprising and subtle application of *quantum expansion*, a tool from quantum information theory introduced recently to operator scaling in Kwok et al. (2019). Moreover, we significantly sharpen the bounds in Kwok et al. (2019) in order to obtain an inverse exponential failure probability and an optimal rate of convergence up to logarithmic factors.

We state our main theorems, slightly loosened for readability, here. Our first theorem plays the role in shape estimation analogous to the role of the matrix Chernoff theorem in covariance estimation, in the sense that it shows how well an "empirical" version $\widehat{\Sigma}$ of the shape approximates the "population" shape Σ.

Theorem 1.1 (Sample complexity) *For $n \geq Cp \log^2 p/\varepsilon^2$ samples from an elliptical distribution of shape Σ, Tyler's M-estimator $\widehat{\Sigma}$ satisfies*

$$\|I_p - \Sigma^{1/2}\widehat{\Sigma}^{-1}\Sigma^{1/2}\|_{op} \leq \varepsilon$$

with probabilty $1 - \exp(-\Omega(n\varepsilon^2/\log^2 p))$ provided ε is a small enough constant. Here $C > 0$ is an absolute constant.

This theorem provides guarantees for estimating the shape in the spectral norm, which was left as an open question in Soloveychik and Wiesel (2014). Up to logarithmic factors, Theorem 1.1 recovers the bound of Soloveychik and Wiesel (2014) on the Frobenius norm $\|\hat{\Sigma}^{-1} - \Sigma^{-1}\|_F$. They showed $\|\hat{\Sigma}^{-1} - \Sigma^{-1}\|_F \leq O(p/\gamma\sqrt{n})$ with high probability, where γ is a parameter depending on the spectrum of Σ. We also prove guarantees for the Frobenius norm that, when the number of samples is large, are sharper than naively using Theorem 1.1 to bound the Frobenius norm.

Theorem 1.2 (Sample complexity in Frobenius norm) *For $n \geq Cp^2/\varepsilon^2$ samples from an elliptical distribution of shape Σ, Tyler's M-estimator $\widehat{\Sigma}$ satisfies*

$$\|I_p - \Sigma^{1/2}\widehat{\Sigma}^{-1}\Sigma^{1/2}\|_F \leq \varepsilon$$

with probabilty $1 - \exp(-\Omega(n\varepsilon^2/p^2))$ provided ε is a small enough constant. Here $C > 0$ is an absolute constant.

Finally, we show that Tyler's iterative procedure converges quickly under mild conditions.

Theorem 1.3 (Iterative procedure) *For $n \geq Cp \log^2 p$ samples from an elliptical distribution of shape Σ, Tyler's iterative procedure computes $\overline{\Sigma}$ satisfying*

$$\|I_p - \widehat{\Sigma}^{1/2}\overline{\Sigma}^{-1}\widehat{\Sigma}^{1/2}\|_F \leq \varepsilon$$

in

$$O(|\log \det \Sigma| + p + \log(1/\varepsilon)) \tag{1}$$

iterations with probability $1 - \exp(-\Omega(n/p \log^2 p))$.

Moreover, in Appendix C we show that the above theorems hold even with access to finitely many bits of the samples. In light of the many recent results on operator scaling and its generalizations Straszak and Vishnoi (2019); Bürgisser et al. (2017, 2018, 2019); Kwok et al. (2019), we hope the connection we make here will shed light on other problems in statistical estimation.

Remark 1.4 (Error metric) *One notes that we use the following measure of error:*

$$\|I_p - \Sigma^{1/2}\widehat{\Sigma}^{-1}\Sigma^{1/2}\|_{op}. \tag{2}$$

We use this error because the Mahalanobis distance *between* $\widehat{\Sigma}, \Sigma$, *given by*

$$\|I_p - \Sigma^{1/2}\widehat{\Sigma}^{-1}\Sigma^{1/2}\|_F, \tag{3}$$

exceeds it by at most a factor $\sqrt{p}$. *The Mahalanobis distance is a natural metric to study, because in the special case when* X *is a Gaussian with covariance* Σ, *if the Mahalanobis distance is at most a small constant then it is on the order of the total variation distance between* X *and the Gaussian with covariance* $\widehat{\Sigma}$ *Barsov and Ul'yanov (1987). Because* $\Omega(p^2)$ *samples are required to estimate the covariance of a Gaussian to constant Mahalanobis distance, Theorems 1.1 and 1.3 are tight up to logarithmic factors. Moreover, the error metric Eq. (2) satisfies an approximate version of the triangle inequality (Lemma D.1), so approximating the estimator in this metric suffices to approximate the shape in the metric.*

2. Tyler's M-estimator and operator scaling

In this section we outline the connection between Tyler's M-estimator and operator scaling. As mentioned in the introduction, Tyler's M-estimator is precisely the "rescaling" for a certain operator constructed from the samples. In some sense, the size of the rescaling governs the nearness of the estimator to the truth. Bounding the sizes of such scalings is a problem that arises naturally in operator scaling, and as shown by Kwok et al. (2019), the size can be bounded by showing that the constructed operator is an *approximate quantum expander*. In Section 4 we make this observation precise, and to prove Theorems 1.1 and 1.3 we improve bounds on the probability that the operator is a quantum expander and show that the general form of Tyler's iterative procedure, known as *Sinkhorn scaling*, converges in time $\text{poly}(\log(1/\varepsilon))$ on approximate quantum expanders.

2.1. Elliptical distributions

Consider X drawn from a centered elliptical distribution $u\Sigma^{1/2}V$ on $\mathbb{R}^p$, in which u is random scalar independent of V, Σ (known as the *shape matrix*) is a fixed $p \times p$ positive-semidefinite matrix with $\det \Sigma = 1$, and V is a uniformly random element of $\mathbb{S}^{p-1}$.

Our task is to find $\widehat{\Sigma}$ estimating Σ. If $x_i, i \in [n]$ are drawn i.i.d from X, then Tyler's M-estimator $\widehat{\Sigma}$ for Σ is the defined to be the solution $\widehat{\Sigma}$ to the two equations

$$\frac{p}{n}\sum_{i=1}^n \frac{x_i x_i^\dagger}{x_i^\dagger \widehat{\Sigma}^{-1} x_i} = \widehat{\Sigma} \tag{4}$$
$$\text{tr}\,\widehat{\Sigma} = p.$$

when the solution exists and is unique.

It is known that if every k-dimensional subspace contains *strictly less* than kn/p vectors x_i, then Tyler's M-estimator exists and is unique, and also that if Tyler's M-estimator exists then every k-dimensional subspace contains *at most* kn/p vectors Tyler (1987). Note the subtle difference between the two conditions. In fact, as shown by King in 1994 in a different language and without any apparent knowledge of Tyler's M estimator, the above sufficient condition is actually necessary. Below we include King's theorem stated in the language of Tyler's M estimator. We remark that Item 3 in what follows does not require the full sophistication of King's theorem, so we provide a short proof in appendix Appendix E for completeness.

Theorem 2.1 (Existence of Tyler's M-estimator King (1994)) *Let $S = \{x_1, \ldots, x_n\} \subset \mathbb{R}^p$.*

1. *Eq. (4) can be solved in $\mathrm{PD}(p)$ up to error ε for all $\varepsilon > 0$ if and only if every k-dimensional subspace of $\mathbb{R}^p$ contains at most kn/p elements of S.*

2. *Eq. (4) has a solution in $\mathrm{PD}(p)$ if and only*

   (a) *Every k-dimensional subspace of $\mathbb{R}^p$ contains at most kn/p elements of S, and*

   (b) *For every k-dimensional subspace L of $\mathbb{R}^p$ containing exactly kn/p elements of S, there is a subspace L' complementary to L in $\mathbb{R}^p$ containing the other $n - kn/p$ elements of S.*

3. *Eq. (4) has a unique solution in $\mathrm{PD}(p)$, or equivalently, Tyler's M-estimator exists, if and only if every k-dimensional subspace of $\mathbb{R}^p$ contains strictly less than kn/p elements of S.*

2.2. Operator scaling

The objects of study are *completely positive maps*, maps $\Phi : \mathrm{Mat}(p) \to \mathrm{Mat}(n)$ between matrix spaces such that there exist $A_1, \ldots, A_r \in \mathrm{Mat}(n \times p)$ such that $\Phi(Y) = \sum_{i=1}^r A_i Y A_i^\dagger$. Completely positive maps arose in the study of $C^\star$ algebras, and play a role in quantum mechanics analogous to the role played by nonnegative matrices in classical probability. We will need some terminology. Let $\mathrm{PD}(p)$ denote the set of positive-definite $p \times p$ matrices.

Definition 2.2 (Completely positive maps) *Let $\Phi(X) = \sum_{i=1}^r A_i X A_i^\dagger$ be a completely positive map.*

- *The size of Φ, denoted $s(\Phi)$, is given by $\mathrm{tr}\,\Phi(I_p)$.*

- *The dual $\Phi^* : \mathrm{Mat}(n) \to \mathrm{Mat}(p)$ of Φ is its Hermitian adjoint and is given by $\Phi^*(X) = \sum_{i=1}^r A_i X A_i^\dagger$.*

- *Say Φ is ε-doubly balanced if*

$$\|\Phi(I_p) - s(\Phi)I_n/n\|_{op} \le s(\Phi)\varepsilon/n \text{ and } \|\Phi^*(I_n) - s(\Phi)I_p/p\|_{op} \le s(\Phi)\varepsilon/p,$$

   and balanced if it is 0-doubly balanced.

- *If $L \in \mathrm{GL}(p), R \in \mathrm{GL}(n)$, let $\Phi_{L,R}$, called a scaling of Φ by L, R, denote the completely positive map given by*

$$\Phi_{L,R} : X \mapsto R\Phi(L^\dagger X L)R^\dagger.$$

   Note that $(\Phi_{L,R})^* = \Phi_{R,L}$.

A central problem in operator scaling is the the existence, and computational efficiency of finding, doubly balanced scalings of a given operator. In the context of Tyler's M-estimator, the relevant example is a completely positive map $\Phi_{\boldsymbol{x}}$ constructed from a tuple $\boldsymbol{x}$ of vectors. This construction arises in the context of the Brascamp-Lieb inequalities in real analysis Garg et al. (2017a). We next see that finding a balanced scaling of the operator solves Eq. (4) and vice versa.

Example 2.3 (The map $\Phi_{\boldsymbol{x}}$.) *For $\boldsymbol{x} \in (\mathbb{C}^p)^n$, let*

$$\Phi_{\boldsymbol{x}} : X \mapsto \mathrm{diag}(\langle x_i, X x_i \rangle : i \in [n]).$$

Then $\Phi_{\boldsymbol{x}}^*(I_n) = \sum_{i=1}^n x_i x_i^\dagger$, and so if $\hat{\Sigma}$ is a solution to Eq. (4) then $\Phi_{\boldsymbol{x}}$ is balanced upon scaling by

$$\hat{\Sigma}^{-1/2} \text{ and } \Phi(\Sigma^{-1})^{-1/2} = \mathrm{diag}\left(x_i^\dagger \widehat{\Sigma}^{-1} x_i : i \in [n]\right)^{-1/2},$$

and if $\Phi_{\boldsymbol{x}}$ is balanced upon scaling by $A, \mathrm{diag}(b)$ we obtain a solution to Eq. (4) by setting $\widehat{\Sigma} = (A^\dagger A)^{-1}$.

It is not hard to see that if $\Phi_{\boldsymbol{x}}$ has a balanced scaling at all, then it has one of the form $A, \mathrm{diag}(b)$. Thus, $\Phi_{\boldsymbol{x}}$ has a balanced scaling if and only if Tyler's M-estimator exists for the samples $x_1, \ldots, x_n$.

As discussed earlier, the size of the scalings control the accuracy of the estimator, and we may control the size by showing that the operator is an approximate quantum expander. Roughly, a completely positive map is an approximate quantum expander if, as a linear map, its first singular vector is close to the identity matrix and its second singular value is strictly less than the first.

Definition 2.4 (Appoximate quantum expanders) *Let Φ be a completely positive map. Then Φ is a $(\varepsilon, 1 - \lambda)$-quantum expander if it is ε-doubly balanced and*

$$\sup_{X \in \mathrm{Herm}(p):\mathrm{tr}\,X=0} \frac{\|\Phi(X)\|_F}{\|X\|_F} \le (1 - \lambda)\frac{s(\Phi)}{\sqrt{np}}.$$

A $(0, 1 - \lambda)$-quantum expander is called a $(1 - \lambda)$-quantum expander.

We use a result of Kwok et al. (2019) relating expansion to the condition numbers of the scaling factors of an approximate quantum expander.

Theorem 2.5 (Follows from Theorem 1.7 of Kwok et al. (2019)) *If Φ is an $(\varepsilon, 1 - \lambda)$ quantum expander and $\varepsilon \log p/\lambda^2$ is at most a small enough constant, then there are L, R with $\det L = \det R = 1$ satisfying*

$$\|I_p - L^\dagger L\|_{op}, \|I_n - R^\dagger R\|_{op} = O(\varepsilon \log p/\lambda)$$

such that $\Phi_{L,R}$ is doubly balanced.

Kwok et al. (2019) actually studied a relaxed notion of approximate quantum expansion, called the *spectral gap*. We discuss the relationship between spectral gap and approximate quantum expansion in Appendix B, and show in particular that the hypotheses of Theorem 1.7 of Kwok et al. (2019) are equivalent to those of Theorem 2.5.

Next we review Sinkhorn's algorithm for completely positive maps from Gurvits (2003); the below presentation is slightly unusual because we only wish to keep track of one of the scalings.

Definition 2.6 (Sinkhorn's algorithm) *Let $Z_0 \in \mathrm{PD}(p)$. For $i \in [T]$,*

- *Set $X = \Phi(Z_{i-1})^{-1}$.*

- *Set $Z_i = \left(\frac{p}{n}\Phi^*(X)\right)^{-1}$.*

We say Z_i is the ith Sinkhorn iterate starting at Z_0.

As the procedure of Tyler is computing $\widehat{\Sigma}$, the output of the ith step of Tyler's iterative procedure is actually $pZ_i^{-1}/\operatorname{tr} Z_i^{-1}$ where Z_i is the ith Sinkhorn iterate for $\Phi_{\boldsymbol{x}}$ starting at I_p. Next we define a function that arises as a progress measure for the Sinkhorn scaling algorithm.

Definition 2.7 (The capacity) *Let* $\Phi : \operatorname{Mat}(p) \to \operatorname{Mat}(n)$ *be a completely positive map. Define the function* $f^\Phi : \operatorname{PD}(p) \to \mathbb{R}$ *by*

$$f^\Phi : \frac{p}{n} \log \det \Phi(Z) - \log \det Z.$$

The quantity $\inf_{Z \succ 0} f^\Phi$ *is the logarithm of the* capacity, *denoted* $\operatorname{cap}(\Phi)$, *defined in* Gurvits (2003).

The link between capacity and scalability is a central result in operator scaling.

Theorem 2.8 (Gurvits (2003)) $\operatorname{cap}(\Phi) > 0$ *if and only if* Φ *has an* ε*-doubly balanced scaling for every* $\varepsilon > 0$.

This result suggests that to find Tyler's M-estimator to accuracy ε in time $\operatorname{poly}(p, n, \log(1/\varepsilon)$, it suffices to find an ε-minimizer of f^Φ in time $\operatorname{poly}(p, n, \log(1/\varepsilon)$, and indeed this was also shown in Gurvits (2003). We next record a few properties of the function f^Φ.

- f^Φ is invariant under multiplication of the input by a positive scalar, i.e.

$$f^\Phi(Z) = f^\Phi(\alpha Z) \text{ for all } \alpha > 0, Z \in \operatorname{PD}(p) \tag{5}$$

  This shows that the progress measure doesn't change whether we use normalized or unnormalized Sinkhorn iterates.

- f^Φ has a property we call *geodesic convexity*: $f^\Phi(A^\dagger e^{tX} A)$ is convex in t for all $X \in \operatorname{Herm}(p)$ and $A \in \operatorname{GL}(p)$.

For a certain choice of metric on the manifold of positive definite matrices $\operatorname{PD}(p)$, this definition of geodesic convexity does match the usual one, in which a real valued function on a Riemmannian manifold is defined to be geodesically convex if it is convex along geodesics. We make use of the *geodesic gradient* of a function $f : \operatorname{PD}(p) \to \mathbb{R}$ at a point Z, given by

$$\nabla f(Z) := \nabla_X f(\sqrt{Z} e^X \sqrt{Z})|_{X=0} \tag{6}$$

Note that also $\nabla f^\Phi(Z) = \nabla f^\Phi(\alpha Z)$ for all $\alpha, Z \in \operatorname{PD}(p)$. For example, for $f = f^\Phi$ as defined in Definition 2.7 we have

$$\nabla f^\Phi(Z) = \frac{p}{n}\sqrt{Z}\Phi^*(\Phi(Z)^{-1})\sqrt{Z} - I_p. \tag{7}$$

This identity gives intuition for Theorem 2.8: if f^Φ has a local minimum, then it is a critical point at which ∇f^Φ vanishes. But if the right-hand side of Eq. (7) is zero, one readily checks that $\Phi_{Z^{1/2},\Phi(Z)^{-1/2}}$ is doubly balanced. In particular, if $\Phi = \Phi_{\boldsymbol{v}}$ then we obtain

$$\nabla f^{\Phi_{\boldsymbol{v}}}(\Sigma) = \frac{p}{n}\sum_{i=1}^n \frac{\Sigma^{1/2}v_i v_i^\dagger \Sigma^{1/2}}{\|\Sigma v_i\|^2} - I_p,$$

which shows that if $f^{\Phi_{\boldsymbol{v}}}$ has a local minimum then Eq. (4) has a solution.

2.3. Technical contributions

Having established the link between scalings and Tyler's M-estimator, we now state our result on quantum expansion, our result on iterative scaling of quantum expanders, and their corollary for Tyler's M-estimator.

We now state the results on expansion of Φ_v which, in conjunction with Theorem 2.5, allows us to conclude Theorems 1.1 and 1.3. The following improves on Kwok et al. (2019) which showed the same but with failure probability $O(p/n^{3/4})$.

Theorem 2.9 *There are absolute constants C, c, λ such that the following holds. Let $v_1, \ldots, v_n$ be Haar random unit vectors from the sphere $\mathbb{S}^{p-1}$. Then Φ_v is an $(\varepsilon, 1 - \lambda)$ quantum expander with probability at least $1 - O(e^{-q(p,n,\varepsilon)})$, where*

$$q(p, n, \varepsilon) = \min\{(c\sqrt{n}\varepsilon - C\sqrt{p})^2, cn - Cp\log p\}.$$

We prove the theorem in Appendix A. We combine it with Theorem 2.5 in Section 3 to prove Theorem 1.1.

We next turn to algorithmic considerations. By the connection between optimization and scaling in Section 2.2, Tyler's M-estimator can be computed by finding an ε-approximate minimizer to the convex function $f^{\Phi_x^*}$. The works Gurvits and Samorodnitsky (2002); Hardt and Moitra (2013) use the ellipsoid method to accomplish this, showing Tyler's M-estimator can be computed up to accuracy ε in time $\text{poly}(p, n, \log(1/\varepsilon))$. Later in Allen-Zhu et al. (2018) it was shown that the optimization problem can be solved by second-order "trust region" methods.

Though the algorithms of Gurvits and Samorodnitsky (2002); Hardt and Moitra (2013); Allen-Zhu et al. (2018) have polynomial time guarantees, both the ellipsoid method and trust regions tend to be slow. Tyler's iterative procedure, on the other hand, is very simple and fast in practice. We next discuss our results in this direction, which will be proven in Section 4.2. Firstly we show that Sinkhorn's algorithm converges in time $O(\log(1/\varepsilon))$ for quantum expanders.

Theorem 2.10 *If $\Phi : \text{Mat}(p) \to \text{Mat}(n)$ is a $1 - \lambda$-quantum expander and $\|\nabla f^{\Phi}(Z_0)\|_F \leq c\lambda^2$, then the Tth Sinkhorn iterate Z starting at Z_0 satifies*

$$\|\nabla f^{\Phi}\|_F = \exp(-O(\lambda T))$$

for some small enough constant c.

In Section 4.3 we straightforwardly combine this theorem with Theorem 2.9 to show Theorem 1.3 on the fast rate of convergence for Tyler's iterative procedure.

Finally, we prove a version of Theorem 2.5 adapted for the Frobenius norm. In some cases, as for Tyler's M estimator, one wishes to obtain Frobenius bounds on $L^{\dagger}L$. One can simply combine the bound $\|\cdot\|_F \leq \sqrt{p}\|\cdot\|_{op}$ with the previous theorem, but our techniques yield a bound that is sharper by a logarithmic factor. Our result is incomparable to Theorem 2.5 and appeals only to geodesic convexity rather than the more sophisticated methods of Kwok et al. (2019). The theorem is proven in Appendix C.

Theorem 2.11 (Frobenius version of Kwok et al. (2019)) *Suppose Φ is a $(\varepsilon, 1 - \lambda)$-quantum expander and $\varepsilon\sqrt{p}/\lambda$ is at most a small enough constant. Then there are L, R with $\det L = \det R = 1$ such that $\Phi_{L,R}$ is doubly balanced and*

$$\|I_p - L^{\dagger}L\|_F, \|I_p - R^{\dagger}R\|_{op} = O(\varepsilon\sqrt{p}/\lambda).$$

2.4. Organization of the paper

In Section 3 we prove Theorems 1.1 and 1.2 on the convergence of Tyler's M-estimator using quantum expansion. In Section 4 we show Theorem 2.10 on the fast convergence of Sinkhorn's algorithm for quantum expanders, and use this to show that Tyler's iterative procedure converges quickly (Theorem 1.3). Both Section 3 and Section 4 rely on Theorem 2.9 on the quantum expansion of the operator $\Phi_{\boldsymbol{v}}$, the proof of which we delay until Appendix A. In the appendix we show relationships between approximate quantum expansion and the spectral gap defined in Kwok et al. (2019), prove Theorem 2.11 on condition number of scalings, and extend Theorem 1.3 and Theorem 1.1 to the finite precision setting.

3. Sample complexity bounds via quantum expansion

First, we observe that we only need to consider the accuracy of Tyler's M-estimator when the shape is the identity.

Observation 3.1 *Suppose $x_i = u_i \Sigma^{1/2} v_i$. Then Y solves Eq. (4) for $\boldsymbol{x}$ if and only if $Z = \Sigma^{-1/2} Y \Sigma^{-1/2}$ solves Eq. (4) for $\boldsymbol{v}$, and clearly*

$$\|I_p - Z^{-1}\|_{op} = \|I_p - \Sigma^{1/2} Y^{-1} \Sigma^{1/2}\|_{op}.$$

This means we need only show that Tyler's M-estimator is accurate for the elliptical distribution in which $u_i = 1$ and $\Sigma = I_p$. This does not violate our assumption that we do not know Σ, because will never need *access* to the v_i's.

Proof [Proof of Theorem 1.1 and Theorem 1.2] We first show Theorem 1.1. As X is a centered elliptical distribution, $x_i = u_i \Sigma v_i$ for $v_1, \ldots, v_n$ equal to Haar random unit vectors. Let λ be a small constant as in Theorem 2.9. By Theorem 2.9, with probability at least $1 - O(e^{-q(p,n,\varepsilon/\log p)}) = 1 - \exp(-\Omega(n\varepsilon^2/\log^2 p))$ for $n \geq Cp \log^2 p/\varepsilon^2$ the operator $\Phi_{\boldsymbol{v}}$ is an $(\varepsilon/\log p, 1 - \lambda)$-quantum expander.

By Theorem 2.5, there is a solution Z to Eq. (4) for $\boldsymbol{v}$ that satisfies $\|Z^{-1} - I_p\|_{op} = C\varepsilon' \log p$. By Observation 3.1, $Z = \Sigma^{-1/2} Y \Sigma^{-1/2}$ where Y solves Eq. (4) for $\boldsymbol{x}$ and in particular

$$\|I_p - \Sigma^{1/2} Y^{-1} \Sigma^{1/2}\|_{op} = O(\varepsilon). \tag{8}$$

Tyler's M-estimator is $\widehat{\Sigma} = pY/\operatorname{tr} Y$. From Eq. (8), we have $Y \in (1 + O(\varepsilon))\Sigma$ and as $\operatorname{tr} \Sigma = p$ we have $\operatorname{tr} Y = p(1 + O(\varepsilon))$ so that

$$\|I_p - \Sigma^{1/2} \widehat{\Sigma}^{-1} \Sigma^{1/2}\|_{op} = O(\varepsilon).$$

By replacing ε by a suitable constant multiple we may assume the error bound is ε rather than $O(\varepsilon)$. To show Theorem 1.2, apply Theorem 2.9 with $\varepsilon/\sqrt{p}$ rather than $\varepsilon/\log p$ and use Theorem 2.11 in place of Theorem 2.5. ∎

4. Fast Sinkhorn for quantum expanders

As discussed in the introduction, quantum expansion is a natural way to bound sizes of scalings. Because scalings are optimizers of the objective function f^Φ of Definition 2.7 in Section 2.2, and bounding the size of optimizers is frequently accomplished through strong convexity, one is led to investigate the relationship between quantum expansion and strong convexity.

In this section we prove such a relationship and use it to show that Tyler's iterative procedure converges linearly. Because Tyler's iterative procedure can be considered a descent method which makes progress proportional to the gradient of f^Φ, it's straightforward to verify (see Section 4.2) that Sinkhorn scaling converges linearly provided f^Φ is strongly convex along the *entire trajectory* of the procedure. We show that this is the case in Section 4.1 by showing that f^Φ is strongly convex in a suitably large *sublevel set*. Finally in Section 4.3 we specialize to Tyler's iterative procedure.

4.1. Strong convexity from quantum expansion

By Eq. (5) of f^Φ, we can only hope for for strong geodesic convexity on the manifold of positive definite matrices with determinant 1, which we call $\mathrm{PD}_1(p)$. The main technical result of this section is proved in Appendix C.

Theorem 4.1 *Suppose Φ is a $(1-\lambda)$-quantum expander. For some absolute constants $C, c > 0$, the following hold.*

1. *$f^\Phi : \mathrm{PD}_1(p) \to \mathbb{R}$ is geodesically $\lambda/2$-strongly convex on the geodesic ball of radius $c\lambda$ about I_p. In particular, for all $Z \in \mathrm{PD}_1(p)$ we have*

$$\left\| \nabla f^\Phi(Z) \right\|_F \geq \frac{\lambda}{2} \min\{c\lambda, \|\log Z\|_F\}.$$

2. *If $\|\nabla f^\Phi(Z)\|_F \leq c\lambda^2$, then $f^\Phi : \mathrm{PD}_1(p) \to \mathbb{R}$ is geodesically $\lambda/2$-strongly convex on a sublevel set S of f^Φ containing Z.*

4.2. Fast Sinkhorn from strong convexity

The goal of this section is to prove Theorem 2.10. In Section 4.3, we will use Theorem 2.10 in conjunction with Theorem 2.9 to show that Tyler's iterative procedure converges linearly.

The theorem follows immediately from Theorem 4.1 and Lemma 4.3, which we will prove shortly. The next lemma is the classic, and completely standard, analysis of progress in the Sinkhorn scaling algorithm. For completeness we include the proof in Appendix C.

Lemma 4.2 *Suppose Z_0, Z are successive Sinkhorn scaling iterates. If $\|\nabla f^\Phi(Z_0)\|_F^2$ is at most 1, then we have*

$$f^\Phi(Z) \leq f^\Phi(Z_0) - \frac{1}{6} \|\nabla f^\Phi(Z_0)\|_F^2.$$

Lemma 4.3 *Suppose $\|\nabla f^\Phi(Z)\|_F \leq 1$. If $f^\Phi : \mathrm{PD}_1(p) \to \mathbb{R}$ is geodesically λ-strongly convex in a sublevel set of f containing Z_0, the Tth Sinkhorn iterate Z_T starting at Z_0 satisfies*

$$\|\nabla f^\Phi(Z_T)\|_F^2 \leq 2^{-\lambda T/24}.$$

10

Proof Define $f := f^\Phi$. If f is geodesically λ-strongly convex in a sublevel set $S \subset \mathrm{PD}_1(p)$ of f containing Z then by Eq. (25),

$$f(Z^*) \geq f(Z_0) - \frac{1}{2\lambda}\|\nabla f(Z_0)\|_F^2.$$

By Lemma 4.2, each step of Sinkhorn iteration decreases f. Because S is a sublevel set, the normalized Sinkhorn iterates Z_i remain in S. Recall that f and its gradient are the same at the normalized and unnormalized Sinkhorn iterates. Moreover, if $\|\nabla f(Z_0)\|_F^2 = \varepsilon > 0$, then by Lemma 4.2 the number of steps T before $\|\nabla f(Z)\|_F^2 \leq \varepsilon/2$ must satisfy

$$f(Z_0) - T\varepsilon/12 \geq f(Z^*) \geq f(Z_0) - \frac{1}{\lambda}\varepsilon,$$

or $T \leq 12/\lambda$. Repeating this argument for any power of 2 tells us that after T steps $\|\nabla f(Z_T)\|_F^2 \leq 2^{-\lambda T/24}$. ∎

Proof [Proof of Theorem 2.10] By Item 2 of Theorem 4.1, f^Φ is geodesically convex on a level set containing Z_0. By Lemma 4.3, the Tth Sinkhorn iterate Z starting at Z_0 satifies $\|\nabla f^\Phi(Z)\|_F = \exp(-O(\lambda T))$. ∎

Remark 4.4 *Lemma 4.3 applies for any descent method, such as geodesic gradient descent, satisfying the conclusion of Lemma 4.2.*

4.3. Fast Sinkhorn for elliptical distributions.

Here we prove Theorem 1.3. First we show that the gradient quickly becomes less than a small constant on the iterates of Sinkhorn's algorithm. For short-hand, define

$$f_{\boldsymbol{v}} := f^{\Phi_{\boldsymbol{v}}}.$$

Lemma 4.5 *Suppose Φ is a $1 - \lambda$-quantum expander. Then for some*

$$T = O\left(\frac{1}{\varepsilon^2}(f^\Phi(Z_0) - p\log(s(\Phi)/n))\right),$$

the output Z_T of Sinkhorn's algorithm starting at Z_0 satisfies $\|\nabla f^\Phi(Z)\|_F \leq \varepsilon$.

Proof By standard analyses of Sinkhorn's algorithm Garg et al. (2017b), T can be taken to be on the order of

$$\frac{1}{\varepsilon^2}\left(f^\Phi(Z_0) - \inf_Z f^\Phi(Z)\right). \tag{9}$$

Because Φ is doubly balanced, the infimum occurs at $Z = I_p$ and we have

$$\inf_Z f^\Phi(Z) = \frac{p}{n}\log\det s(\Phi)I_n/n = p\log(s(\Phi)/n).$$

∎

The next lemma, proven in Appendix C, asserts that a doubly balanced scaling of a sufficiently good approximate quantum expander is a quantum expander. This is the expander to which we will be applying Theorem 2.10.

Lemma 4.6 *Suppose $\delta \leq c\lambda^2/\log p$ and that Φ is a $(\delta, 1 - \lambda)$-quantum expander. Then Φ is scalable and for any L, R such that $\Phi_{L,R}$ is doubly balanced, $\Phi_{L,R}$ is a $1 - \lambda + O(\delta \log p/\lambda)$ quantum expander.*

Corollary 4.7 *Let λ be a small constant as in Theorem 2.9. With probability $1 - O(e^{-q(p,n,c\lambda^2/\log p)})$,*

$$(\Phi_{\boldsymbol{x}})_{\widehat{\Sigma}^{-1/2}, \Phi_{\boldsymbol{x}}(\widehat{\Sigma}^{-1})^{-1/2}} \tag{10}$$

is a $1 - .5\lambda$-quantum expander. In other words, the operator $\Phi_{\boldsymbol{z}}$ where

$$z_i = \widehat{\Sigma}^{-1/2} x_i / \|\widehat{\Sigma}^{-1/2} x_i\|$$

is a $1 - .5\lambda$ quantum expander.

Proof Let λ be a small constant as in Theorem 2.9. With probability at least $1 - O(e^{-q(p,n,c\lambda^2/\log p)}) = 1 - O(e^{-\Omega(n/(p\log^2 p))})$ for $n \geq Cp \log^2 p$, the operator $\Phi_{\boldsymbol{v}}$ is an $(c\lambda^2/\log p, 1 - \lambda)$ quantum expander. As the operator in Eq. (10) is a doubly balanced scaling of $\Phi_{\boldsymbol{v}}$, Lemma 4.6 implies the claim. ■

We can now prove Theorem 1.3.

Proof Let λ be a small constant as in Theorem 2.9. By Corollary 4.7, with probability at least $1 - O(e^{-q(p,n,c\lambda^2/\log p)}) = 1 - O(e^{-\Omega(n/(p\log^2 p))})$ for $n \geq Cp \log^2 p$, the operator

$$\Psi := (\Phi_{\boldsymbol{x}})_{\widehat{\Sigma}^{-1/2}, \Phi_{\boldsymbol{x}}(\widehat{\Sigma}^{-1})^{-1/2}}$$

is a $(1 - .5\lambda)$-quantum expander. Moreover, as reasoned in the proof of Theorem 1.1, $\|\Sigma^{1/2}\widehat{\Sigma}^{-1}\Sigma^{1/2} - I_p\|_{op} = O(c\lambda^2)$. Because Ψ is a suitable scaling of $\Phi_{\boldsymbol{x}}$, if the normalized Sinkhorn iterates of $\Phi_{\boldsymbol{x}}$ starting at I_p are Z_i then the normalized Sinkhorn iterates of Ψ starting at $\widehat{\Sigma}$ are $Z_i' = \widehat{\Sigma}^{1/2} Z_i \widehat{\Sigma}^{1/2}$. By Lemma 4.5, for some T as in Lemma 4.5 we have $\|\nabla f^{\Psi}(Z_T')\|_F \leq \delta_0$ where $\delta_0 = c\lambda^2$.

Because Ψ is doubly balanced and $\operatorname{tr}\widehat{\Sigma} = p$,

$$f^{\Psi}(\widehat{\Sigma}) - \log(s(\Phi)/n) \leq p \log\left(\frac{1}{n} \operatorname{tr}\Psi(\widehat{\Sigma})\right) - \log\det(\widehat{\Sigma}) - \log(s(\Phi)/n)$$

$$= -\log\det(\widehat{\Sigma}),$$

where the inequality is Jensen's. Next, using that $\widehat{\Sigma} \in \Sigma(1 + O(1))$ we have $|\log\det(\widehat{\Sigma})| = O(p) + |\log\det(\Sigma)|$, so $T = O((p + |\log\det(\widehat{\Sigma})|/\delta_0^2) = O(p + |\log\det(\Sigma)|)$.

By Theorem 2.10, the number of further steps T to obtain a normalized Sinkhorn iterate Z' with $\|\nabla f^{\Psi}(Z')\|_F \leq \varepsilon$ is $O(\log(1/\varepsilon))$. We may assume $\varepsilon \leq c\lambda$, because otherwise $\log(1/\varepsilon) = \Theta(1)$. Next we apply Item 1 of Lemma C.9 to Ψ at I_p to see that $\|\log(Z')\|_F = 2\varepsilon/\lambda = O(\varepsilon)$, so the corresponding normalized Sinkhorn iterate Z satisfies

$$\|\log \widehat{\Sigma}_1^{1/2} Z \widehat{\Sigma}_1^{1/2}\|_F = O(\varepsilon).$$

Applying Lemma D.2 completes the proof. ■

References

Zeyuan Allen-Zhu, Ankit Garg, Yuanzhi Li, Rafael Oliveira, and Avi Wigderson. Operator scaling via geodesically convex optimization, invariant theory and polynomial identity testing. In *Proceedings of the Symposium on the Theory of Computing (STOC 2018)*, pages 172–181, 2018.

Alexander Barg and D Yu Nogin. Bounds on packings of spheres in the Grassmann manifold. *IEEE Transactions on Information Theory*, 48(9):2450–2454, 2002.

SS Barsov and VV Ul'yanov. Estimates of the proximity of Gaussian measures. In *Sov. Math., Dokl*, volume 34, pages 462–466, 1987.

Franck Barthe. On a reverse form of the Brascamp-Lieb inequality. *Inventiones mathematicae*, 134 (2):335–361, 1998.

Peter Bürgisser, Ankit Garg, Rafael Oliveira, Michael Walter, and Avi Wigderson. Alternating minimization, scaling algorithms, and the null-cone problem from invariant theory. In *Proceedings of Innovations in Theoretical Computer Science (ITCS 2018)*, 2017. doi: 10.4230/LIPIcs.ITCS. 2018.24.

Peter Bürgisser, Cole Franks, Ankit Garg, Rafael Oliveira, Michael Walter, and Avi Wigderson. Efficient algorithms for tensor scaling, quantum marginals, and moment polytopes. In *Proceedings of the Symposium on Foundations of Computer Science (FOCS 2018)*, pages 883–897. IEEE, 2018.

Peter Bürgisser, Cole Franks, Ankit Garg, Rafael Oliveira, Michael Walter, and Avi Wigderson. Towards a theory of non-commutative optimization: geodesic first and second order methods for moment maps and polytopes. *arXiv preprint arXiv:1910.12375*, 2019.

Stamatis Cambanis, Steel Huang, and Gordon Simons. On the theory of elliptically contoured distributions. *Journal of Multivariate Analysis*, 11(3):368–385, 1981.

Jürgen Forster. A linear lower bound on the unbounded error probabilistic communication complexity. *Journal of Computer and System Sciences*, 65(4):612–625, 2002.

Ankit Garg, Leonid Gurvits, Rafael Oliveira, and Avi Wigderson. A deterministic polynomial time algorithm for non-commutative rational identity testing. In *Proceedings of the Symposium on Foundations of Computer Science (FOCS 2016)*, pages 109–117. IEEE, 2016. doi: 10.1109/ FOCS.2016.95.

Ankit Garg, Leonid Gurvits, Rafael Oliveira, and Avi Wigderson. Algorithmic and optimization aspects of Brascamp-Lieb inequalities, via operator scaling. In *Proceedings of the Symposium on the Theory of Computing (STOC 2017)*, pages 397–409. ACM, 2017a.

Ankit Garg, Leonid Gurvits, Rafael Oliveira, and Avi Wigderson. Operator scaling: theory and applications. *Foundations of Computational Mathematics*, pages 1–68, 2017b.

John Goes, Gilad Lerman, and Boaz Nadler. Robust sparse covariance estimation by thresholding Tyler's M-estimator. *arXiv preprint arXiv:1706.08020*, 2017.

Arjun K Gupta, Tamas Varga, and Taras Bodnar. *Elliptically contoured models in statistics and portfolio theory*. Springer, 2013.

Leonid Gurvits. Classical deterministic complexity of Edmonds' problem and quantum entanglement. In *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 10–19, 2003.

Leonid Gurvits and Alex Samorodnitsky. A deterministic algorithm for approximating the mixed discriminant and mixed volume, and a combinatorial corollary. *Discrete & Computational Geometry*, 27(4):531–550, 2002.

Moritz Hardt and Ankur Moitra. Algorithms and hardness for robust subspace recovery. In *Conference on Learning Theory*, pages 354–375, 2013.

Douglas Kelker. Distribution theory of spherical distributions and a location-scale parameter generalization. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 419–430, 1970.

Alastair D King. Moduli of representations of finite dimensional algebras. *The Quarterly Journal of Mathematics*, 45(4):515–530, 1994.

Philip A Knight. The Sinkhorn–Knopp algorithm: convergence and applications. *SIAM Journal on Matrix Analysis and Applications*, 30(1):261–275, 2008.

Tsz Chiu Kwok, Lap Chi Lau, and Akshay Ramachandran. Spectral analysis of matrix scaling and operator scaling. In *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1184–1204. IEEE, 2019.

Olivier Marchal, Julyan Arbel, et al. On the sub-Gaussianity of the Beta and Dirichlet distributions. *Electronic Communications in Probability*, 22, 2017.

Ilya Soloveychik and Ami Wiesel. Performance analysis of tyler's covariance estimator. *IEEE Transactions on Signal Processing*, 63(2):418–426, 2014.

George W Soules. The rate of convergence of Sinkhorn balancing. *Linear algebra and its applications*, 150:3–40, 1991.

Damian Straszak and Nisheeth K Vishnoi. Computing maximum entropy distributions everywhere. In *Proceedings of Machine Learning Research, 32nd Annual Conference on Learning Theory*, 2019.

David E. Tyler. A distribution-free M-estimator of multivariate scatter. *The Annals of Statistics*, 15(1):234–251, Mar 1987. ISSN 0090-5364. doi: 10.1214/aos/1176350263. URL http://dx.doi.org/10.1214/aos/1176350263.

Roman Vershynin. Introduction to the non-asymptotic analysis of random matrices. *arXiv preprint arXiv:1011.3027*, 2010.

Martin J Wainwright. *High-dimensional statistics: A non-asymptotic viewpoint*, volume 48. Cambridge University Press, 2019.

Ami Wiesel and Teng Zhang. Structured robust covariance estimation. *Foundations and Trends® in Signal Processing*, 8(3):127–216, 2015.

Appendix A. Quantum expansion for random unit vectors

Here we prove Theorem 2.9. We prove the theorem in Appendix A.3 after stating the main technical ingredient of the proof, Lemma A.14.

A.1. A Cheeger constant for vectors.

Lemma A.1 *Let $v = (v_1, \ldots, v_n) \in (\mathbb{C}^p)^n$ be a tuple of vectors. Then Φ_v is an (ε, λ)-quantum expander if and only if Φ_v is ε-doubly balanced and for all unitaries U, the matrix $B^U \in \mathrm{Mat}(p, n)$ given by*

$$B^U_{i,j} = B^U_{i,j} = |(Uv_i)_j|^2$$

satisfies

$$\sup_{x \in \mathbb{R}^n, \langle x, \mathbf{1}_p \rangle = 0} \frac{\|B^U x\|}{\|x\|} \leq (1 - \lambda) s(\Phi_v)/\sqrt{np}.$$

Proof Write

$$
\begin{aligned}
\sup_{X \in \mathrm{Herm}(p), \mathrm{tr}\, X = 0} \frac{\|\Phi(X)\|_F^2}{\|X\|_F^2} &= \sup_{U \in U(p)} \sup_{x \in \mathbb{R}^n : \langle x, \mathbf{1}_p \rangle = 0} \frac{\|\Phi(U^\dagger \mathrm{diag}(x) U)\|_F^2}{\|x\|^2} \\
&= \sup_{U \in U(p)} \sup_{x \in \mathbb{R}^n : \langle x, \mathbf{1}_p \rangle = 0} \frac{\sum_{i=1}^n \langle Uv_i, \mathrm{diag}(x) Uv_i \rangle^2}{\|x\|^2} \\
&= \sup_{U \in U(p)} \sup_{x \in \mathbb{R}^n : \langle x, \mathbf{1}_p \rangle = 0} \frac{\sum_{i=1}^n |(B^U x)_i|^2}{\|x\|^2}.
\end{aligned}
$$

∎

Definition A.2 *Let $B \in \mathrm{Mat}(p, n)$ be a nonnegative matrix. The Cheeger constant $\mathrm{ch}(B)$ of the weighted bipartite graph associated to B is given by*

$$\mathrm{ch}(B) := \min_{T \subset [p], |T| \leq p/2, S \subset [n]} \phi(S, T)$$

where

$$\phi(S, T) := \frac{\mathrm{cut}(S, T)}{\min\{\mathrm{vol}(S, T), \mathrm{vol}(\overline{S}, \overline{T})\}}$$

where

$$\mathrm{vol}(S, T) := \sum_{i \in T, j \in [n]} b_{ij} + \sum_{i \in [p], j \in S} b_{ij} \text{ and } \mathrm{cut}(S, T) := \sum_{i \notin T, j \in S} b_{ij} + \sum_{i \in T, j \notin S} b_{ij}.$$

Lemma A.3 *Let $B \in \mathrm{Mat}(p, n)$ be a nonnegative matrix. Suppose B is ε-doubly balanced. Let $\mathrm{ch}(B)$ denote the Cheeger constant of the weighted bipartite graph associated to B. Then for $\varepsilon < c\,\mathrm{ch}(B)^2$,*

$$\sup_{x \in \mathbb{R}^n, \langle x, \mathbf{1}_p \rangle = 0} \frac{\|Bx\|_2}{\|x\|_2} \leq \max\left\{1/2, 1 - \mathrm{ch}(B)^2 + C \frac{\varepsilon}{\mathrm{ch}(B)^2}\right\} \frac{s(B)}{\sqrt{np}}.$$

Here C, c are absolute constants.

15

Proof By Kwok et al. (2019), the spectral gap of B is at most $\mathrm{ch}(B)^2 - 3\varepsilon$. By Lemma B.3, B is a quantum expander with the desired parameters. ∎

Lemma A.1 and Lemma A.3 lead us to consider the *minimum* value

$$\mathrm{ch}(\boldsymbol{v}) := \min_U \mathrm{ch}(B^U),$$

and immediately imply the following:

Corollary A.4 *Let* $\varepsilon < c\,\mathrm{ch}(\boldsymbol{v})^2$. *If the operator* $\Phi_{\boldsymbol{v}}$ *is* ε-*doubly balanced, then* $\Phi_{\boldsymbol{v}}$ *is an*

$$\left(\varepsilon, \max\left\{ \frac{1}{2}, 1 - \mathrm{ch}(\boldsymbol{v})^2 + C\frac{\varepsilon}{\mathrm{ch}(\boldsymbol{v})^2} \right\} \right)$$

quantum expander. Here C, c' *are absolute constants.*

We next show how to express $\mathrm{ch}(\boldsymbol{v})$ as an infimum over projections rather than unitaries. This will be easier to control in the random setting.

Definition A.5 (Conductance for projections) *Suppose* $\boldsymbol{v} \in (\mathbb{R}^p)^n$, $\pi : \mathbb{R}^p \to \mathbb{R}^p$ *is an orthogonal projection, and* $S \subset [n]$. *We think of* S, π *as a cut, and the quantity* ϕ *defined below as the conductance of the cut. Define*

$$\phi(S, \pi) = \frac{\mathrm{cut}(S, \pi)}{\min\{\mathrm{vol}(S, \pi), \mathrm{vol}(\overline{S}, I_p - \pi)\}}.$$

where

$$\mathrm{vol}(S, \pi) = \sum_{i \in n} \|\pi v_i\|^2 + \sum_{i \in S} \|v_i\|^2, \text{ and } \mathrm{cut}(S, \pi) = \sum_{i \in S} \|(I_p - \pi)v_i\|^2 + \sum_{i \notin S} \|\pi v_i\|^2.$$

Lemma A.6 $\mathrm{ch}(\boldsymbol{v}) = \min_{S, \pi:\mathrm{rank}\,\pi \le p/2} \phi(S, \pi).$

Proof The conductance $\phi(S, T)$ the cut S, T of B^U is exactly the same as the conductance of $\phi(S, \pi)$ of $\boldsymbol{v}$ where π is the orthogonal projection to the span $\{Ue_i : i \in T\}$, which ranges over all orthogonal projections as T, U range over $\binom{[p]}{\le p/2} \times U(p)$. Hence,

$$\min_U \min_{S,T:|T| \le p/2} \phi(S, T) = \min_{S,\pi:\mathrm{rank}\,\pi \le p/2} \phi(S, \pi).$$

∎

A.2. Probabilistic preliminaries

Recall the Bernstein condition, which can be used to show from the moments of a random variable that it is subexponential.

Definition A.7 (Subexponential variables and Bernstein's condition) *Recall that random variable X with mean μ is said to be (ν, b)-subexponential if*

$$\mathbb{E}[e^{\lambda(X-\mu)}] \leq e^{\frac{1}{2}\nu^2\lambda^2} \text{ for all } |\lambda| \leq 1/b.$$

Say a random variable X with mean μ and variance σ^2 satisfies the Bernstein condition with parameter b *if for all integers $k \geq 3$ we have*

$$|\mathbb{E}[(X-\mu)^k]| \leq \frac{1}{2}k!\sigma^2 b^{k-2}. \tag{11}$$

It is known that a random variable satisfying the Bernstein condition with parameter b is $(2\sigma, 2b)$-subexponential.

Recall that the $\mathrm{Beta}(\alpha, \beta)$ distribution has mean $\mu = \alpha/(\alpha + \beta)$, variance $\sigma^2 = \alpha\beta/(\alpha + \beta)^2(\alpha + \beta + 1)$, and k^{th} raw moment

$$\prod_{r=0}^{k-1} \frac{\alpha + r}{\alpha + \beta + r}.$$

We will be concerned with the setting when $2\alpha \leq 2\beta$ are positive integers adding to p.

Lemma A.8 *If X is $\mathrm{Beta}(1/2, p - 1/2)$, then X is (ν, β)-sub-exponential where $\nu, b = O(p^{-1})$.*

Proof As X has $\sigma^2 = O(p^{-2})$, it suffices to show that X satisfies the Bernstein property with parameter $b = O(1/\beta)$.

Because μ and X are nonnegative, $-\mu^k \leq (X - \mu)^k \leq X^k$. Thus,

$$-\mu^k \leq \mathbb{E}[(X - \mu)^k] \leq \mathbb{E}[X^k].$$

Thus, $|\mathbb{E}[(X - \mu)^k]| \leq \max\{\mu^k, \mathbb{E}[X^k]\} = \mathbb{E}[X^k]$. The Bernstein property is satisfied if for $k \geq 3$ we have

$$\begin{aligned}
b^{k-2} &\geq \frac{2}{k!}\mathbb{E}[X^k]/\sigma^2 \\
&= \frac{2}{k!}\frac{(\alpha + \beta)^2(\alpha + \beta + 1)}{\alpha\beta}\prod_{r=0}^{k-1}\frac{\alpha + r}{\alpha + \beta + r} \\
&= \frac{2}{k!}\frac{(\alpha + \beta)(\alpha + 1)}{\beta}\prod_{r=2}^{k-1}\frac{\alpha + r}{\alpha + \beta + r} \\
&= O\left(\frac{1}{k!}\prod_{r=2}^{k-1}\frac{\frac{1}{2} + r}{p + r}\right) \\
&= O\left(\frac{1}{k!}\frac{1}{p^{k-2}}\prod_{r=2}^{k-1}(r + 1)\right) \\
&= O(1/p^{k-2}),
\end{aligned}$$

so we may take $b = O(1/p)$. ∎

Lemma A.9 *Let v be a uniformly random element of $\mathbb{S}^{p-1}$, and define*

$$X_k = \sum_{i=1}^{k} |v_i|^2 \text{ and } Y = \sum_{i=k+1}^{l} |v_i|^2$$

for $1 \leq k < l \leq p$. That is, X_k and Y are sums of squares of disjoint sets of coordinates of v. Then X_k and Y are negatively associated.

Proof Recall that it suffices to show $\Pr[X \geq a \wedge Y \geq b] \leq \Pr[X \geq a]\Pr[Y \geq b]$ for all $a, b \in (0, 1)$, or equivalently that $\Pr[Y \geq b | X \geq a] \leq \Pr[Y \geq b]$ for all $a, b \in (0, 1)$.

We may sample X and Y by sampling independent χ-squared random variables $Z_1, \ldots, Z_p$ and setting $X = X'/(X'+Y'+Z'), Y = Y'/(X'+Y'+Z')$ where $X' = \sum_{i=1}^{k} Z_i, Y' = \sum_{i=k+1}^{l} Z_i$, and $Z' = \sum_{i=l+1}^{p} Z_i$. In particular, X', Y', Z' are independent.

Now $X \geq a$ if and only if $X' \leq a(X'+Y'+Z')$, or $Y' \leq f(X', Z') := -Z' + (1-a)X'/a$, and $Y' \geq b$ if and only if $Y' \leq b(X'+Y'+Z')$ or $Y' \geq g(X', Z') := b(X'+Z')/(1-b) \geq 0$. Thus, we seek to show that

$$\Pr[Y' \geq g(X', Z') | Y' \leq f(X', Z')] \leq \Pr[Y' \geq g(X', Z')].$$

To prove this, it suffices to show that $\Pr[Y' \geq \alpha | Y' \leq \beta] \leq \Pr[Y' \geq \alpha]$ for all numbers $0 \leq \alpha < \beta$. This is true because the event $Y' \geq \alpha$ contains the complement of $Y' \leq \beta$. ∎

Lemma A.10 *Let X_k be as in Lemma A.9, i.e. $X_k = \text{Beta}(k - 1/2, p - k - 1/2)$. Then X_k is $(\nu = O(k^{1/2}p^{-1}), b = O(p^{-1}))$-subexponential.*

Proof By Lemma A.8, it suffices to show $\mathbb{E}[e^{\lambda(X_k-\mu_k)}] \leq \mathbb{E}[e^{\lambda(X_1-\mu_k)}]^k$, for then

$$\mathbb{E}[e^{\lambda(X_k-\mu_k)}] \leq e^{k\lambda^2\nu^2/2}.$$

This we show by induction. Clearly the claim holds for $k = 1$. For $k \geq 2$, note that $X_k = X_{k-1}+Y$ where Y is as in the previous lemma. Note that the marginal distribution of Y is that of X_1.

Let $\mathbb{E}[X_k] = \mu_k$. The function $e^{\lambda(X-\mu_k)}$ is either monotone increasing or decreasing in X; by Lemma A.9, X_{k-1} and Y are negatively associated and so

$$\mathbb{E}[e^{\lambda(X_k-\mu_k)}] = \mathbb{E}[e^{\lambda(X_{k-1}-\mu_{k-1})}e^{\lambda(Y-\mu_1)}] \leq \mathbb{E}[e^{\lambda(X_{k-1}-\mu_{k-1})}]\mathbb{E}[e^{\lambda(X_1-\mu_1)}].$$

This completes the proof of the inductive hypothesis. ∎

Theorem A.11 (Subexponential tail bound, Wainwright (2019)) *If Y_i are i.i.d (ν_i, b) subexponential, and $Y = \sum_{i=1}^{n} Y_i$, then*

$$\Pr[|Y - \mu| \geq nt] \leq \begin{cases} 2e^{-nt^2/\nu_*^2} & \text{for } 0 \leq t \leq b/\nu_* \\ 2e^{-nt/2b} & \text{for } t > b/\nu_* \end{cases}$$

for $\nu_* = \sqrt{\sum_{i=1}^{n} \nu_i^2/n}$.

By applying the previous theorem with $t = \varepsilon k/p$, $\nu_i = \nu = O(\sqrt{k}/p)$ and $b = O(1/p)$ we have the following corollary.

Corollary A.12 *Let $Y_1, \ldots, Y_n$ be independently distributed with each $Y_i \sim X_k$ for X_k as in Lemma A.10. Then $Y = Y_1 + \cdots + Y_n$ satisfies*

$$\Pr[(1-\varepsilon)nk/p \leq Y \leq (1+\varepsilon)nk/p] \leq 2e^{-\Omega(kn\varepsilon^2)}$$

for all $\varepsilon \leq c$, where c is a small enough constant.

Finally, we use a standard result in random matrix theory. If V is a Haar-random element of $\mathbb{S}^{p-1}$, then the distribution $\sqrt{p}V$ is in isotropic position and is $O(1)$ subgaussian Marchal et al. (2017), so the following theorem follows from Theorem 5.39 of (as rewritten in Equation 5.25) of Vershynin (2010).

Theorem A.13 (Vershynin (2010)) *Let $v_1, \ldots, v_n$ independent, Haar random unit vectors. There are absolute constants $C, c > 0$ such that for all $C\sqrt{p/n} \leq c\delta$, $\delta \leq 1$ we have*

$$\left\| \frac{p}{n} \sum_{i=1}^{n} v_i v_i^\dagger - I_p \right\|_{op} \leq \delta.$$

with probability at least

$$1 - 2\exp\left(-(c\sqrt{n}\delta - C\sqrt{p})^2\right).$$

A.3. The Cheeger constant of random unit vectors

The purpose of this section is to prove the following lemma, which in conjunction with Corollary A.4 implies Theorem 2.9.

Lemma A.14 *For λ at most some absolute constant, if $v_1, \ldots, v_n$ are sampled i.i.d. from $\mathbb{S}^{p-1}$, then $\mathrm{ch}(\boldsymbol{v}) \geq \lambda$ with probability at least $1 - e^{-\Omega(n)+O(p\log p)}$.*

The proof, which is merely a more involved version of the proof of Theorem A.13 in Vershynin (2010), combines concentration of the random variable $\phi(S, \pi)$ for fixed S, π with a union bound over $2^{[n]} \times N$ where N is an ε-net for the projections. We proceed to prove concentration.

Before showing the detailed proof of the lemma, we use the statement to complete the proof of Theorem 2.9.

Proof [Proof of Theorem 2.9] By Theorem A.13 applied to the distribution $\sqrt{p}V$ where V is a Haar random unit vector, the operator $\Phi_{\boldsymbol{v}}$ is ε-doubly balanced with probability $1 - 2e^{-(c\sqrt{n}\varepsilon-C\sqrt{p})^2}$. By Lemma A.14, for some small enough absolute constant $\lambda_0 > 0$ the quantity $\mathrm{ch}(\boldsymbol{v})$ is at least λ_0 with probability at least $1 - O(e^{-cn+Cp\log p})$. Both events occur with probability $1 - O(\exp(-\min\{(c\sqrt{n}\varepsilon - C\sqrt{p})^2, cn - Cp\log p\}))$. By Corollary A.4, there is an absolute constant c_0 such that if $\varepsilon < c_0\lambda_0^2$ then $\Phi_{\boldsymbol{v}}$ is a $(\varepsilon, 1-\lambda)$-quantum expander with probability $1 - O(\exp(-\min\{(c\sqrt{n}\varepsilon - C\sqrt{p})^2, cn - Cp\log p\}))$ for $\lambda > 0$ another absolute constant. ∎

Lemma A.15 *Suppose π is of rank $k \leq p/2$. Then with probability at least $1 - e^{-\Omega(kn)}$,*

$$\min_{S \subset [n]} \Phi(S, \pi) \geq c \tag{12}$$

$$\text{and } .5kn/p \leq \sum_{i=1}^{n} \|\pi' v_i\|^2 \leq n - .5kn/p \tag{13}$$

Here c is some small enough constant.

Proof We prove that the failure probability of Eqs. (12) and (13) are individually $O(e^{nkp})$; the claim will then follow by the union bound. We first bound the failure probability of Eq. (13). Note that $\sum_{i=1}^{n} \|\pi v_i\|^2$ is distributed as $W = \sum_{i=1}^{n} W_i$ for $W_i \sim X_k$. By Corollary A.12, we have $\Pr[|W - nk/p| \geq .5nk/p] = O(e^{-nkp})$. Hence $\min\{W, n - W\} \geq .5nk/p$ with probability at least $1 - e^{-\Omega(nk)}$

We now bound the failure probabilty of Eq. (12). For fixed S, π, observe that ϕ is distributed as

$$\phi(S, \pi) \sim \frac{Y + Z}{\min\{\ell + W, 2n - \ell - W\}}$$

where $Y = \sum_{i=1}^{\ell} Y_i$ for $Y \sim X_{p-k}$ and $Z = \sum_{i=1}^{n-\ell} Z_i$ for $Z_i \sim X_k$ with Y and Z independent, and $W = \sum_{i=1}^{n} W_i$ for $W_i \sim X_k$.

Let α be a constant that we will make small compared to c.

Case 1: $\ell \geq \alpha n$.
Here we use the bound $\phi(S, \pi) \geq Y/n$. By Corollary A.12, with probability at least $1 - O(e^{-\Omega(\ell(p-k))}) = 1 - O(e^{-\Omega(np)})$, Y is at least $.5\ell(p - k)/p \geq .25\alpha n$. The number of S is at most 2^n, so with probability at least $1 - O(2^n e^{-\Omega(pn)}) = 1 - O(e^{-\Omega(pn)})$ there exists no S with $\ell \geq \alpha n$ such that $\phi(S, \pi) \leq .25\alpha$.

Case 2: $\ell \leq \alpha n$.
We claim two events A and B hold simultaneously with probability at least $1 - O(e^{-\Omega(kn)})$. Let A be the event that $W + \ell \leq \ell + 1.5nk/p = O(\max\{\ell, nk/p\})$. Let B be the event that $Y + Z$ is at least $\max\{.25\ell, .5nk/p\} = \Omega(\max\{\ell, nk/p\})$. By Corollary A.12 A holds with probability at least $1 - O(e^{-\Omega(nk)})$. We now bound the failure probability of B. If $.25\ell \geq .5nk/p$, then Y is at least $.25\ell$ with probability $1 - O(e^{-\Omega(lp)}) = 1 - O(e^{-\Omega(nk)})$, and $Z \geq .5nk/p$ with probability $1 - O(e^{-\Omega(nk)})$. This shows that B holds with probability $1 - O(e^{-\Omega(nk)})$, and by the union bound A and B hold simultaneously with probability $1 - O(e^{-\Omega(nk)})$.

Condition on A and B. For $\alpha \leq 1/8$, A implies that $W \leq n - \ell$ and so $\min\{\ell + W, 2n - \ell - W\} = \ell + W$. Now A and B imply

$$\phi(S, \pi) = \Omega\left(\frac{\max\{\ell, nk/p\}}{\max\{\ell, nk/p\}}\right) = \Omega(1).$$

Now, A and B fail for some S with $\ell \leq \alpha n$ with probability on the order of

$$e^{-ckn} \binom{n}{\leq \alpha n} \leq e^{-\Omega(kn) + (\ln 2)H(\alpha)n} \leq e^{-\Omega(kn)}$$

20

provided $H(\alpha)$, the binary entropy of α, is a small enough constant. Hence, with probability $1 - e^{-\Omega(kn)}$, $\Phi(S, \pi) \geq c$ for all S with $\ell \leq \alpha n$. ∎

We now prove lemmas allowing us to construct and apply δ-nets for the set of rank-k orthogonal projections. Recall that a δ-net for the projections of rank k in the operator norm is a subset N of projections such that for all projections $\pi : \mathbb{R}^d \to \mathbb{R}^d$ of rank k there exists $\pi' \in N$ such that $\|\pi - \pi\|_{op} \leq \delta$.

Lemma A.16 *Suppose that a subset N of the projections of rank k is a δ-net, and that for all $\pi' \in N$ we have*

$$\min_{S \subset [n]} \phi(S, \pi') \geq \beta. \tag{14}$$

$$and \ .5kn/p \leq \sum_{i=1}^{n} \|\pi' v_i\|^2 \leq n - .5kn/p \tag{15}$$

Then for $\delta \leq \frac{ck}{p}$, we have $\min_{S \subset [n]} \phi(S, \pi) \geq \beta - O(p\delta/k)$ for all projections π of rank k. Here c is some small enough constant.

Proof Consider a rank k-projection π, and let $\pi \in N$ such that $\|\pi - \pi'\|_{op} \leq \delta$. Fix $S \subset [n]$.

Observe that

$$\min\{\text{vol}(S, \pi'), \text{vol}(\overline{S}, I_p - \pi')\} \geq \min\{\sum_{i=1}^{n} \|\pi' v_i\|^2, n - \sum_{i=1}^{n} \|\pi' v_i\|^2\} \geq .5kn/p.$$

Because $\phi(S, \pi') \geq \beta$, we have

$$\text{cut}(S, \pi') \geq .5\beta kn/p.$$

Because $\|\pi v_i\|^2 = \langle v_i, \pi v_i \rangle$ is linear in π, we have that $\|\pi' v_i\|^2 - \|\pi v_i\|^2 \leq \delta$ and similarly $\|(I_p - \pi')v_i\|^2 - \|(I_p - \pi)v_i\|^2\| \leq \delta$. We may then write

$$\text{cut}(S, \pi) \geq \sum_{i \in S} \|(I_p - \pi')v_i\|^2 + \sum_{i \notin S} \|\pi' v_i\|^2 - n\delta$$

$$\geq \left(1 - \delta \frac{p}{.5\beta k}\right) \text{cut}(S, \pi').$$

On the other hand,

$$\text{vol}(S, \pi) = \sum_{i \in n} \|\pi v_i\|^2 + \ell$$

$$\leq \sum_{i \in n} \|\pi' v_i\|^2 + \ell + n\delta$$

$$\leq \text{vol}(S, \pi') \left(1 + \delta \frac{p}{.5k}\right)$$

and similarly $\text{vol}(\overline{S}, I_p - \pi) \leq \text{vol}(\overline{S}, I_p - \pi') \left(1 + \delta \frac{p}{.5k}\right)$. Thus,

$$\phi(S, \pi) \geq \phi(S, \pi') \frac{1 - \delta \frac{p}{.5\beta k}}{1 + \delta \frac{p}{.5k}} \geq \beta - O(p\delta/k).$$

∎

Following a standard method to prove the existence of δ-nets, we consider a maximal code N distance $\delta/2$ in the set X of rank k projections. By the triangle inequality, there can be no point at distance at least δ from every point of N, else it could be added to the packing. We then use the Hamming bound to bound $|N|$.

Lemma A.17 *There is a δ-net N of the rank k orthogonal projections with $|N| = \exp(O(pk|\ln \delta|))$.*

Proof Let the subset N of the rank k-projections be a maximal code of distance $\delta/2$ in the Frobenius norm. By the discussion preceding the statement of the theorem, N is a δ net in the Frobenius norm and hence also in the operator norm. By Barg and Nogin (2002), $|N| = \exp(O(pk|\ln \delta|))$. ∎

Finally we prove the main result of the section.

Proof [Proof of Lemma A.14] Fix $k \leq p/2$. By Lemma A.17, there is a $c\lambda/p$ net N with $|N| = \exp(O(pk \log p))$ for absolute constants c, λ. By Lemma A.15 and the union bound, the conditions Eqs. (14) and (15) in Lemma A.16 hold with $\beta = 2\lambda$ for all $\pi \in N$ with probability $1 - e^{-\Omega(nk)}e^{O(pk \log p)}$. By Lemma A.16, $\min_{S \subset [n]} \phi(S, \pi) \geq \lambda$ for all projections π of rank k assuming c is small enough compared to λ.

By the union bound over $k \in [p/2]$, $\mathrm{ch}(v) \geq \lambda$ with probability at least $1 - O(pe^{-\Omega(n)+O(p \log p)}) = 1 - O(e^{-\Omega(n)+O(p \log p)})$. ∎

Appendix B. Spectral gap

Definition B.1 (Spectral gap) *Let $\Phi : \mathrm{Mat}(p) \to \mathrm{Mat}(n)$ be a completely positive map. Say Φ has spectral gap λ if its second singular value is at most $(1 - \lambda)\frac{s(\Phi)}{\sqrt{np}}$.*

Here is the theorem appearing in Kwok et al. (2019) from which we can show Theorem 2.5.

Theorem B.2 *Let $p \leq n$. Suppose that the completey positive map $\Phi : \mathrm{Mat}(n) \to \mathrm{Mat}(p)$ is ε-doubly balanced and has spectral gap λ with $\varepsilon \log p/\lambda^2$ is at most a small enough constant. Then the condition number of the scaling solutions L and R such that $\Phi_{L,R}$ is doubly stochastic satisfy*

$$\kappa(L), \kappa(R) \leq 1 + O(\varepsilon \log p/\lambda).$$

In the next lemma, which is straightforward, we prove a relationship between the spectral gap and quantum expansion.

Lemma B.3 *Let Φ be a completely positive map. There are constants $C, c > 0$ such that the following holds for $\varepsilon \leq c\lambda$.*

- *If Φ is a $(\varepsilon, 1 - \lambda)$-quantum expander then it has spectral gap $\lambda - O(\varepsilon)$, and*

- *If Φ has spectral gap λ and is ε-doubly balanced then it is an (ε, α)-quantum expander for*

$$\alpha = \max\{1/2, 1 - \lambda\} + O(\varepsilon/\lambda).$$

Before proving Lemma B.3 let us use it to prove Theorem 2.5.

Proof [Proof of Theorem 2.5] By Lemma B.3, Φ has spectral gap $\lambda - O(\varepsilon) = \Omega(\lambda)$. Apply Theorem B.2 to $\Phi^* : \mathrm{Mat}(n) \to \mathrm{Mat}(p)$; because Φ^* is the adjoint of Φ, the two have the same spectral gap. Hence the scaling factors R and L that make $(\Phi^*)_{R,L}$ doubly balanced satisfy $\kappa(L), \kappa(R) \leq 1 + O(\varepsilon \log p/\lambda)$. By scaling R and L by the respective geometric means of their singular values, we may assume that $\|I_p - L^\dagger L\|_{op}, \|I_p - R^\dagger R\|_{op} = O(\varepsilon \log p/\lambda)$ and $\det L = \det R = 1$. ∎

Lemma B.3 is essentially an elementary statement about linear maps. Its proof uses the following technical linear algebra lemmas as well as one from Kwok et al. (2019).

Lemma B.4 *Suppose $A \in \mathrm{Mat}(m, n)$ has spectral gap λ. Then there are constants $C, c > 0$ such that for every unit vector x such that $\|Ax\| > (1 - \delta)\|A\|_{op}$ with $\delta/\lambda < c$, we have*

$$\|Ay\| \leq (\max\{1/2, 1 - \lambda\} + C\delta/\lambda)\|A\|_{op}$$

for all unit vectors $y \in x^\perp$.

Proof It is enough to prove the claim when A has operator norm 1. Let v denote a top right singular vector of A. We assume x is a unit vector. First we claim that $|\langle x, v \rangle| \geq \sqrt{1 - 2\delta/\lambda}$. Write $x = \sqrt{(1 - \alpha)}uv + \sqrt{\alpha}w$ where $w \in v^\perp$ is a unit vector, $\alpha \in [0, 1]$, and u is in the complex unit circle. Because Aw is in the orthogonal complement of Av, we have

$$(1 - \delta)^2 \leq \|Ax\|_2^2 = (1 - \alpha)\|Auv\|_2^2 + \alpha\|Aw\|_2^2 \leq (1 - \alpha)1 + \alpha(1 - \lambda)^2$$
$$= 1 - \alpha(1 - (1 - \lambda)^2)$$

Hence, $\alpha \leq (1 - (1 - \delta)^2)/(1 - (1 - \lambda)^2) \leq 2\delta/\lambda$. Next, for a unit vector $y \in x^\perp$ write $y = \sqrt{\beta}u'v + \sqrt{1 - \beta}w'$ where $w' \in v^\perp$ is a unit vector and $\beta \in [0, 1]$. Because $\langle x, y \rangle = 0$,

$$\sqrt{\beta}\sqrt{1 - \alpha} = \sqrt{1 - \beta}\sqrt{\alpha}|\langle w, w' \rangle| \leq \sqrt{1 - \beta}\sqrt{\alpha}.$$

Hence $\sqrt{\beta}/\sqrt{1 - \beta} \leq \sqrt{\alpha}/\sqrt{1 - \alpha}$ and so $\beta \leq \alpha \leq 2\delta/\lambda$. Take $\lambda' = \min\{\lambda, 1/2\}$. Now

$$\|Ay\|_2^2 \leq \beta + (1 - \beta)(1 - \lambda')^2 \leq 2\delta/\lambda + (1 - 2\delta/\lambda)(1 - \lambda')^2.$$

Hence $\|Ay\|_2 \leq (1 - \lambda')\sqrt{1 - 2\delta/\lambda + 2\delta/\lambda(1 - \lambda')^2} \leq (1 - \lambda')\sqrt{1 + 6\delta/\lambda}$, which completes the proof because we have assumed δ/λ is a small enough constant. ∎

Lemma B.5 *Suppose $A \in \mathrm{Mat}(m, n)$ is such that $\|A\|_{op} = 1$ and there exists a vector x such that for all vectors $y \in x^\perp$, we have $\|Ay\| \leq (1 - \lambda)\|y\|$. Then A has spectral gap λ, i.e. $\sigma(A) \leq 1 - \lambda$.*

Proof We use the Rayleigh trace. Write

$$1 + \sigma_2(A)^2 = \sigma_1^2(A) + \sigma_2^2(A) = \sup_{\dim L = 2} \mathrm{tr}\, \pi_L A^\dagger A.$$

For any L, we may take u_1, u_2 to be an orthonormal basis spanning L with $u_1 \in x^\perp$. Then

$$\mathrm{tr}\, \pi_L A^\dagger A = \|Au_1\|_F^2 + \|Au_2\|_F^2 \leq (1 - \lambda)^2 + 1.$$

Thus, $\sigma_2(A) \leq 1 - \lambda$. ∎

Lemma B.6 (Lemma 3.6 of [Kwok et al. (2019)](#)) *Let $\Phi : \mathrm{Mat}(p) \to \mathrm{Mat}(n)$ be an ε-doubly balanced completely positive map. The first singular value of the linear map Φ is at most $(1 + \varepsilon)s(\Phi)/\sqrt{np}$.*

Proof [Proof of Lemma B.3] Suppose $\Phi : \mathrm{Mat}(n) \to \mathrm{Mat}(m)$ is ε-doubly balanced. By Lemma B.6, the operator norm of Φ is at most $(1 + \varepsilon)s(\Phi)/\sqrt{mn}$. Let E denote the unit vector $I_n/\sqrt{n}$.

First assume that Φ has spectral gap λ. Because Φ is ε-doubly balanced,

$$\|\Phi(E)\|_F \geq \|(1 - \varepsilon)s(\Phi)I_m/m\sqrt{n}\|_F \geq (1 - \varepsilon)s(\Phi)/\sqrt{mn}.$$

Hence, $\|\Phi(E)\|_F \geq ((1 - \varepsilon)/(1 + \varepsilon))\|\Phi\|_{op}$. Applying Lemma B.4 to Φ, $x = E$, and $\delta = 1 - (1 - \varepsilon)/(1 + \varepsilon) \leq 2\varepsilon$, we have that $\|\Phi(Y)\| \leq (\max\{1/2, 1 - \lambda\} + C\varepsilon/\lambda)\|Y\|_F$ for all Y orthogonal to E and hence Φ is an $(\varepsilon, \max\{1/2, 1 - \lambda\} + C\varepsilon/\lambda)$-quantum expander.

On the other hand, if Φ is a $(\varepsilon, 1 - \lambda)$-quantum expander, then $\|\Phi(Y)\|_F \leq \|Y\|_F(1 - \lambda)s(\Phi)/\sqrt{mn}$ for all Y orthogonal to E, and hence $\|\Phi(Y)\|_F \leq \|Y\|_F(1 - \lambda)\|\Phi\|_{op}/(1 - \varepsilon)$ for all Y orthogonal to E. By Lemma B.5 applied with Φ normalized to have operator norm 1 we have that Φ has spectral gap $1 - (1 - \lambda)/(1 - \varepsilon) \geq \lambda - C\varepsilon$. ∎

Appendix C. Finite precision

Here we show that even with access to only finitely many bits of $v_1, \ldots, v_n$, Tyler's M-estimator $\widehat{\Sigma}$ still exists with high probability and remains close to the true shape Σ. Let v' denote the result of rounding the entries of v.

Recall the proof of Theorem 1.1. To show that $\widehat{\Sigma}$ is close to Σ, it sufficed to show that Tyler's M-estimator on v is close to the identity. Similarly, to show $\widehat{\Sigma}$ is close to Σ even after rounding, it suffices to show that Tyler's M-estimator on v' is close to the identity. This would follow from Theorem 2.9 with v replaced by v', or more precisely that if v is chosen at random then $\phi_{v'}$ is a $(\varepsilon, 1 - \lambda)$ quantum expander with high probability. As Theorem 2.9 already shows ϕ_v is an $(\varepsilon, 1 - \lambda)$ quantum expander with the desired probability, the following lemma is enough.

Lemma C.1 *Suppose $p \leq n$, that v, v' are such that $\|v_i - v_i'\| \leq \delta$ for all $i \in [n]$ and that Φ_v is an $(\varepsilon, 1 - \lambda)$ quantum expander. Then $\Phi_{v'}$ is an*

$$(\varepsilon - \delta n, 1 - \lambda' + \delta\sqrt{pn})$$

quantum expander.

Proof If for some $\alpha \leq \max\{c\lambda, 1/p\}$ we can show

$$\|\Phi_v - \Phi_{v'}\|_{op} \leq \alpha s(\Phi_v), \tag{16}$$

then $s(\Phi_v)(1 - \alpha\sqrt{p}) \leq s(\Phi_{v'}) \leq (1 + \alpha\sqrt{p})s(\Phi_v)$, and so $\|\Phi_{v'}(I_p) - \Phi_v(I_p)\|_{op} \leq \alpha\sqrt{p}\,s(\Phi)$, $\|\Phi_{v'}^*(I_n) - \Phi_v^*(I_n)\|_{op} \leq \alpha\sqrt{n}\,s(\Phi_{v'})$, and finally

$$\sup_{\mathrm{tr}\,X=0, X\in\mathrm{Herm}(p)} \frac{\|\Phi_{v'}(X)\|_F}{\|X\|_F} \leq (\lambda + \alpha)s(\Phi_v)$$

$$\leq \frac{(\lambda + \alpha)}{(1 - \alpha\sqrt{p})}s(\Phi_v).$$

Plugging in these parameters and using $p \leq n$ tells us $\Phi_{\boldsymbol{v}'}$ is an

$$(\varepsilon + O(\alpha\sqrt{n}), \lambda - O(\alpha\sqrt{p}))$$

quantum expander. We now upper bound α in Eq. (16). We may write $\Phi_{\boldsymbol{v}'} = \Phi_{\boldsymbol{v}} + T + S + \Phi_{\boldsymbol{v}'-\boldsymbol{v}}$ where the operators S, T are given by $S(X)_{i,i} = v_i^{\dagger} X(v_i' - v)$ and $T(X)_{i,i} = (v_i' - v)^{\dagger} X v_i$ and $\Phi_{\boldsymbol{v}'-\boldsymbol{v}}$. Because $\|S\|_{op} = \|T\|_{op} \leq \delta\sqrt{n}$ and $\|\Phi_{\boldsymbol{v}-\boldsymbol{v}'}\|_{op} \leq \delta^2\sqrt{n}$, we may take $\alpha = \delta\sqrt{n}$. ∎

Now let $\boldsymbol{x}'$ denote the result of rounding $\boldsymbol{x}$ to b bits after the decimal place so that each entry of $\boldsymbol{x}' - \boldsymbol{x}$ is at most 2^{-b}. From the discussion at the beginning of this section, we have the following results.

Theorem C.2 *Let $M \geq 1$ be a bound on the condition number of Σ, let $\boldsymbol{x} = x_1, \ldots, x_n$ be drawn from an elliptical distribution with shape Σ, and let $\boldsymbol{x}' = x_1', \colon x_n'$ be the result of rounding x_i to*

$$b_i \geq C \log \left(\frac{Mnp \log p}{\varepsilon \|x_i\|} \right)$$

bits after the decimal place. Let $c > 0$ be an absolute constant.

1. *With probability $1 - O(e^{-q(p,n,\varepsilon/\log p)})$ for $q(p,n,\varepsilon)$ is as in Theorem 2.9, the estimator $\tilde{\Sigma} = \widehat{\Sigma}(\boldsymbol{x}')$ on the rounded vectors satisfies*

$$\|I_p - \Sigma^{1/2}\tilde{\Sigma}^{-1}\Sigma^{1/2}\|_{op} = O(\varepsilon) \tag{17}$$

   provided $\delta \leq c/\sqrt{p}$.

2. *With probability $1 - O(e^{-q(p,n,c/\log p)})$, Sinkhorn scaling on $\boldsymbol{x}'$ outputs $\overline{\Sigma}$ satisfying*

$$\|I_p - \widehat{\Sigma}^{1/2}\overline{\Sigma}^{-1}\widehat{\Sigma}^{1/2}\|_{op} = O(\varepsilon) \tag{18}$$

   in

$$O(|\log \det \Sigma| + p + \log(1/\varepsilon))$$

   iterations.

Proof Let λ be a small constant as in Theorem 2.9. By Theorem 2.9, with probability at least $1 - O(e^{-q(p,n,\varepsilon/\log p)})$, the operator $\Phi_{\boldsymbol{v}}$ is an $(\varepsilon/\log p, 1 - \lambda)$-quantum expander. Condition on this event. The proof of Theorem 1.1 shows that $\|I_p - \Sigma^{1/2}\widehat{\Sigma}^{-1}\Sigma^{1/2}\|_{op} \leq \varepsilon$. Let $\tilde{\Sigma}$ denote Tyler's M-estimator for the rounded vectors $\boldsymbol{x}'$. To prove Item 1, by Lemma D.1 it suffices to show that $\tilde{\Sigma}$ is close to $\widehat{\Sigma}$. First note that $\widehat{\Sigma}$ has condition number $O(M)$ because $\|I_p - \Sigma^{1/2}\widehat{\Sigma}^{-1}\Sigma^{1/2}\|_{op} \leq 1$.

By Corollary 4.7,

$$\Psi := (\Phi_{\boldsymbol{x}})_{\widehat{\Sigma}^{-1/2}, \Phi_{\boldsymbol{x}}(\widehat{\Sigma}^{-1})^{-1/2}}$$

is a $1 - .5\lambda$ quantum expander. Note that $\Psi = \Phi_{\boldsymbol{z}}$ for $z_i = \sqrt{Z}x_i/\|\sqrt{Z}x_i\|$, $Z := \widehat{\Sigma}^{-1}$. Now if we let $z_i' = \sqrt{Z}x_i'/\|\sqrt{Z}x_i'\|$, observe that z_i, z_i' are scale-invariant in x_i, x_i', and set $x = x_i/\|x_i\| = $

$1, x' = x'_i/\|x'_i\|$ so that $\|x - x'\| \le 2^{-b}/\|x_i\|$, then $\|z_i - z'_i\|$ is

$$\left\| \frac{\sqrt{Z}xx^\dagger\sqrt{Z}}{\langle x, Zx\rangle} - \frac{\sqrt{Z}x'(x')^\dagger\sqrt{Z}}{\langle x', Zx'\rangle} \right\|_F$$

$$\le \|Z\|_{op} \frac{\|\langle x', Zx'\rangle xx^\dagger - \langle x, Zx\rangle x'(x')^\dagger\|_F}{\langle x, Zx\rangle\langle x', Zx'\rangle}.$$

$$\le \|Z\|_{op}\|Z^{-1}\|_{op}^2\|\langle x', Zx'\rangle xx^\dagger - \langle x, Zx\rangle x'(x')^\dagger\|_F$$

$$\le \|Z\|_{op}\|Z^{-1}\|_{op}^2(\langle x', Zx'\rangle\|xx^\dagger - x'(x')^\dagger\|_F + \|x\|^2|\operatorname{tr} Z(xx^\dagger - x(x')^\dagger)|)$$

$$\le (CM)^2 \cdot O(\|xx^\dagger - x'(x')^\dagger\|_F) = O((CM)^2\sqrt{p}2^{-b}/\|x_i\|),$$

By our choice of b_i and Lemma C.1, the operator

$$\Psi' := (\Phi_{\boldsymbol{x}'})_{\widehat{\Sigma}^{-1/2}, \Phi_{\boldsymbol{x}'}(\widehat{\Sigma}^{-1})^{-1/2}}$$

is a $(\varepsilon/\log p, 1 - .25\lambda)$ quantum expander. By Theorem 2.5, the left scaling $Y^{1/2} \in \mathrm{PD}_1(p)$ for Ψ' satisfies $\|I_p - Y\|_{op} = O(\varepsilon)$. Then $Y^{1/2}\widehat{\Sigma}^{-1/2}$ is a left scaling for $\Phi_{\boldsymbol{x}'}$, i.e. $\tilde{\Sigma}_1^{-1} = \widehat{\Sigma}_1^{-1/2}Y\widehat{\Sigma}_1^{-1/2}$. We conclude that

$$\|\widehat{\Sigma}_1^{-1/2}\tilde{\Sigma}_1\widehat{\Sigma}_1^{-1/2} - I_p\|_{op} = O(\varepsilon). \tag{19}$$

Eq. (17) follows from Eq. (19) and Lemma D.2.

To prove Item 2, condition on the operator $\Phi_{\boldsymbol{v}}$ being an $(c/\log p, 1 - \lambda)$-quantum expander for c a small enough constant. This happens with probability $1 - O(e^{-q(p,n,c/\log p)})$. Let Z_T denote the result of running normalized Sinkhorn iteration on $\Phi_{\boldsymbol{x}'}$. Let T be large enough that $\|I_p - \tilde{\Sigma}_1^{1/2}Z_T\tilde{\Sigma}_1^{1/2}\|_{op} \le \varepsilon$; by the proof of Theorem 1.3, T satisfies Eq. (1) in Theorem 1.3. Eq. (18) follows from Eq. (19), Lemma D.2, and Lemma D.1. $\blacksquare$

Proofs for Section 4

Remark C.3 (Normalization) *The usual convention is $\operatorname{tr}\Sigma = \operatorname{tr}\widehat{\Sigma} = p$, but to avoid more calculations like those at the end of the proof of Theorem 1.1 it will more convenient to assume $\Sigma, \widehat{\Sigma}$ and the Sinkhorn iterates are in $\mathrm{PD}_1(p)$. For this reason if $A \in \mathrm{PD}(p)$ we let $A_1 := \det(A)^{-1/p}A \in \mathrm{PD}_1(p)$. An elementary calculation (Lemma D.2) shows that switching between the two normalizations for $\Sigma, \widehat{\Sigma}$ only incurs a constant factor error in Eq. (2) provided it is at most a small constant. If Z_i are a sequence of Sinkhorn iterates, then the elements of the sequence $(Z_i)_1$ are called the normalized Sinkhorn iterates.*

We say a function $f : \mathrm{PD}_1(p) \to \mathbb{R}$ is geodesically γ-strongly convex at Z if for all Hermitian X with $\operatorname{tr} X = 0$, we have

$$\partial_t^2 f\left(\sqrt{Z}e^{tX}\sqrt{Z}\right) \ge \gamma\|X\|_F^2.$$

A key feature of this definition is that for $f = f^\Phi$, different points Z correspond to scalings of Φ. More precisely, if f^Φ is γ strongly convex at Z if and only if $f^{\Phi_{Z,I_n}}$ is strongly convex at I_p.

Lemma C.4 (Strong convexity from expansion) *There are constants $C, c > 0$ such that for all $\varepsilon < c$, if $\Phi : \mathrm{Mat}(p) \to \mathrm{Mat}(n)$ is a $(\varepsilon, 1 - \lambda)$-quantum expander, then the function $X \mapsto \log \det(\Phi(X))$ from $\mathrm{PD}_1(p)$ to $\mathbb{R}$ is geodesically*

$$\frac{n}{p} \left(2\lambda - \lambda^2 - C\varepsilon \right)$$

strongly convex at the origin. Note that $2\lambda - \lambda^2 \geq \lambda$.

Proof Consider X Hermitian with $\mathrm{tr}\, X = 0$. Observe that

$$\partial_t^2 \log \det(\Phi(e^{tX}))|_{t=0} = \mathrm{tr}\, \Phi(I_p)^{-1}\Phi(X^2) - \mathrm{tr}\, \Phi(I_p)^{-1}\Phi(X)\Phi(I_p)^{-1}\Phi(X).$$

We lower bound the first term and upper bound the second. Because Φ is ε-doubly balanced, $\Phi(I_p) \succeq (1-\varepsilon)(s(\Phi)/n)I_n$, so using the monotonicity of the trace inner product in each argument under the Loewner ordering we have

$$\mathrm{tr}\, \Phi(I_p)^{-1}\Phi(X)\Phi(I_p)^{-1}\Phi(X) \leq (s(\Phi)/n)^{-2}(1 - \varepsilon)^{-1} \mathrm{tr}\, \Phi(X)^2.$$

on the other hand, $\Phi^*(I_n) \succeq (1 - \varepsilon)(s(\Phi)/p)I_p$ and $\Phi(I_p) \preceq (1 + \varepsilon)(s(\Phi)/n)I_p$, so

$$
\begin{aligned}
\mathrm{tr}\, \Phi(I_p)^{-1}\Phi(X^2) &\geq (1 + \varepsilon)^{-1}(s(\Phi)/n)^{-1} \mathrm{tr}\, \Phi(X^2) \\
&= (1 + \varepsilon)^{-1}(s(\Phi)/n)^{-1} \mathrm{tr}\, \Phi(X^2)I_n \\
&= (1 + \varepsilon)^{-1}(s(\Phi)/n)^{-1} \mathrm{tr}\, X^2\Phi^*(I_n) \\
&\geq (1 + \varepsilon)^{-1}\frac{n}{p}(1 - \varepsilon) \mathrm{tr}\, X^2.
\end{aligned}
$$

Combining these two bounds and using the quantum expansion of Φ and that $\varepsilon < c$, we have

$$
\begin{aligned}
\partial_t^2 \log \det(\Phi(e^{tX})) &\geq \frac{n}{p}\|X\|_F^2 \left((1 + \varepsilon)^{-1}(1 - \varepsilon) - (1 - \lambda)^2(1 - \varepsilon)^{-1} \right). \\
&\geq \frac{n}{p}\|X\|_F^2 \left((2\lambda - \lambda^2)(1 - C\varepsilon) - C\varepsilon \right).
\end{aligned}
$$

$\blacksquare$

Because the function $Z \mapsto \log \det Z$ is geodesically linear, we have the following corollary.

Corollary C.5 *If $\varepsilon < c$ and Φ is a (ε, λ)-quantum expander, then the function $f^\Phi : \mathrm{PD}_1(p) \to \mathbb{R}$ is geodesically $\lambda - C\varepsilon$-strongly convex at the origin. Here $C, c > 0$ are absolute constants.*

While it is nice to know that f^Φ is geodesically strongly convex at I_p, to deduce bounds on the sizes of scalings we need to show that the function is strongly convex *near* I_p. The next lemma is the first step in this direction.

Lemma C.6 (Expansion under scaling) *If Φ is an $(\varepsilon < 1, 1 - \lambda)$-quantum expander and $\|L^\dagger L - I_p\|_{op} \leq \delta$ and $\|R^\dagger R - I_n\|_{op} \leq \delta$ for $\delta < c$, then $\Phi_{L,R}$ is an $(\varepsilon + C\delta, 1 - \lambda + C\delta)$-quantum expander. Here $C, c > 0$ are absolute constants.*

Proof We first show the balanced-ness condition. Because $(1 - \delta)I_p \preceq L^\dagger L \preceq (1 + \delta)I_p$ and $(1 - \delta)I_p \preceq RR^\dagger \preceq (1 + \delta)I_n$, we have

$$(1 - \delta)^2 s(\Phi) \leq s(\Phi_{L,R}) \leq (1 + \delta)^2 s(\Phi). \tag{20}$$

Furthermore,

$$R\Phi(L^\dagger I_p L)R^\dagger \preceq (1 + \delta)^2(1 + \varepsilon)\frac{s(\Phi)}{n}I_n.$$
$$\preceq \frac{(1 + \delta)^2}{(1 - \delta)^2}(1 + \varepsilon)\frac{s(\Phi_{L,R})}{n}I_n$$

The other condition is similar.

We next verify the expansion condition. We seek to bound the operator norm of $\Phi_{L,R}$ restricted to the traceless Hermitians. First note that the map $Y \to R^\dagger Y R$ has operator norm at most $(1+\delta)$, so the desired bound will be the operator norm of $X \mapsto \Phi(L^\dagger X L)$ restricted to the traceless Hermitians multiplied by $(1 + \delta)$. Let $\pi(Z) = Z - (\operatorname{tr} Z)I_p/p$ denote the projection of Z to its traceless part, and let $\Psi : X \mapsto \pi(L^\dagger X L)$. Note that Ψ maps the traceless Hermitians to the traceless Hermitians, and Ψ has operator norm at most $(1 + \delta)$. Then we may write

$$\Phi(L^\dagger X L) = \Phi \circ \Psi(X) + \Phi(I_p) \operatorname{tr}(L^\dagger X L)/p.$$

By the triangle inequality, the operator norm of $X \mapsto \Phi(L^\dagger X L)$ restricted to the traceless Hermitian matrices is at most the sum of that of $\Phi \circ \Psi$ and that of $\Gamma : X \mapsto \Phi(I_p) \operatorname{tr}(L^\dagger X L)/p$. The former is immediately seen to be at most $(1 - \lambda)(1 + \delta)s(\Phi)/\sqrt{np}$, and because $\operatorname{tr} X = 0$ we have

$$
\begin{aligned}
\|\Gamma\|_{op} &\leq \frac{1}{p}\|\Phi(I_p)\|_F \sup_{\operatorname{tr} X=0} \frac{\operatorname{tr} L^\dagger X L}{\|X\|_F} \\
&= \frac{1}{p}\|\Phi(I_p)\|_F \sup_{\operatorname{tr} X=0} \frac{\operatorname{tr} LL^\dagger X}{\|X\|_F} \\
&\leq \frac{1}{p}\|\Phi(I_p)\|_F \sup_{\operatorname{tr} X=0} \frac{\operatorname{tr}(L^\dagger L - I_p)X}{\|X\|_F} \\
&\leq \frac{1}{p}\|\Phi(I_p)\|_F \sup_{\operatorname{tr} X=0} \delta\sqrt{p} \\
&\leq (1 + \varepsilon)s(\Phi)\delta/\sqrt{np},
\end{aligned}
$$

where in the last condition we used that Φ is ε-doubly balanced. Using Eq. (20) again, we have

$$\sup_{X \in \operatorname{Herm}(n), \operatorname{tr} X=0} \frac{\|\Phi_{L,R}(X)\|_F^2}{\|X\|_F^2} \leq \frac{s(\Phi_{L,R})}{\sqrt{np}(1 - \delta)^2} ((1 - \lambda)(1 + \delta) + (1 + \varepsilon)\delta).$$

Because $\delta \leq c$ and $\varepsilon < 1$, we have the claim. $\blacksquare$

Corollary C.7 *There is an absolute constant c such that the following holds. Suppose Φ is a $(\varepsilon, 1 - \lambda)$-quantum expander and $\varepsilon \leq c\lambda$. Then f^Φ is geodesically $\lambda/2$-strongly convex on the geodesic ball of radius $c\lambda$ about I_p.*

Proof The strong convexity of f^Φ at a point Z equals the strong convexity of $g := f^{\Phi\sqrt{Z},I_n}$ at I_p. For $\kappa > 0$ to be determined shortly, by Lemma C.6, if $\|Z - I_p\|_{op} \leq \kappa$, then $\Phi_{\sqrt{Z},I_n}$ is an $(\varepsilon + C\kappa, 1 - \lambda + C\kappa)$-quantum expander. By Corollary C.5, for C' a large enough absolute constant, f^Φ is $\lambda - C'\varepsilon - C'\kappa$-strongly convex at Z provided $\|Z - I_p\|_{op} \leq \kappa$, which is implied by $\|\log Z\|_F \leq c'\kappa$ for κ, c' at most a small absolute constant.

If we take $\kappa = c\lambda$ for c a small enough constant we conclude that f^Φ is $\lambda/2$ strongly convex on the geodesic ball of radius $c\lambda$ about I_p. ∎

We now include two lemmas with standard results from geodesically convex optimization. The first will imply a sharper bound for estimating the shape matrix in Frobenius norm.

Lemma C.8 *Suppose that $f : \mathrm{PD}_1(p) \to \mathbb{R}$ is geodesically λ-strongly convex on the geodesic ball of radius κ about Z_0 and that $\|\nabla f(Z_0)\|_F < \lambda\kappa$. Then there is a unique optimizer Z^* of f satisfying*

$$\|\log Z_0^{-1/2} Z^* Z_0^{-1/2}\|_F \leq \frac{\|\nabla f(Z_0)\|_F}{\lambda} \tag{21}$$

and

$$f(Z^*) \geq f(Z_0) - \frac{1}{2\lambda}\|\nabla f(Z_0)\|_F^2. \tag{22}$$

Proof By replacing the function with $X \mapsto f(\sqrt{Z_0} X \sqrt{Z_0})$, it is enough to prove the assertions in the case $Z_0 = I_p$. Let B be the geodesic ball of radius κ about I_p. By geodesic convexity on B, we have

$$\frac{d^2}{dt^2} f(e^{tH}) = \frac{d^2}{ds^2} f(e^{\frac{1}{2}tH} e^{sH} e^{\frac{1}{2}tH}) \geq \lambda. \tag{23}$$

for all $t \leq \kappa, \|H\|_F = 1$. Integrating Eq. (23) once, we have

$$\frac{d}{dt} f(e^{tH}) \geq \frac{d}{dt} f(e^{tH})|_{t=0} + \lambda t \geq -\|\nabla f(I_p)\|_F + \lambda t, \tag{24}$$

for $0 \leq t \leq \kappa, \|H\|_F = 1$. The last inequality is by Cauchy-Schwarz. Integrating again, we find that

$$f(e^{tH}) \geq -t\|\nabla f(I_p)\|_F + \frac{\lambda}{2}t^2. \tag{25}$$

$0 \leq t \leq \kappa, \|H\|_F = 1$. Evaluating Eq. (25) with $t = \kappa$ shows that on the boundary of B, the function value is at least $f(I_p)$. As a consequence, a local minimum Z^* must occur in B, and by strong convexity it must be unique. Solving Eq. (24) for t with $e^{tH} = Z^*$, and hence left-hand side equal to zero, shows $t \leq \|\nabla f(I_p)\|_F/\lambda \leq \kappa$, proving Eq. (21). Eq. (25) is minimum at this value of t; evaluating at the minimum implies Eq. (22). ∎

We now apply the previous lemma to f^Φ to conclude Theorem 2.11.

Proof [Proof of Theorem 2.11] We first show, using the assumption that Φ is ε-doubly balanced, that $\|\nabla f^\Phi(I_p)\|_F = O(\varepsilon\sqrt{p})$. Recall that $\nabla f^\Phi(I_p) = \frac{p}{n}\Phi^*(\Phi(I_p)^{-1}) - I_n$. By ε-balancedness and

$\varepsilon \leq c$, $\|(n\Phi(I_p)/s(\Phi))^{-1} - I_n\|_{op} = O(\varepsilon)$, so $\|A\|_{op} := \|\Phi(I_p)^{-1} - nI_p/s(\Phi)\|_{op} = O(n\varepsilon/s(\Phi))$. Now

$$\|\frac{p}{n}\Phi^*(\Phi(I_p)^{-1}) - I_n\|_{op} \leq \left\|\frac{p}{n}\Phi^*(nI_p/s(\Phi)) - I_n\right\|_{op} + \|A\|_{op}$$
$$= O(\varepsilon).$$

Thus $\|\nabla f^\Phi\|_F \leq \sqrt{p}\|\nabla f^\Phi\|_{op} = O(\varepsilon\sqrt{p})$. By Corollary C.7, we may now apply Lemma C.8 with $\kappa = c\lambda$ and $Z_0 = I_p$ to conclude that there is a unique optimizer Z^* for f^Φ satisfying

$$\|\log Z^*\|_F = O(\sqrt{p}\varepsilon/\lambda). \tag{26}$$

Because $\sqrt{p}\varepsilon/\lambda \leq c$, $\|I - Z\|_F = O(\sqrt{p}\varepsilon/\lambda)$. We may take $L = \sqrt{Z^*}$. It is a straightforward calculation to show that, by ε-balancedness of Φ, the scaling R may be taken as some multiple of $\Phi(Z^*)^{-1/2}$. ∎

The next lemma us useful for showing that the iterative algorithm converges quickly. In what follows we assume that the identity is the optimizer. The results are sharper in this case, and we may assume this is the case by using Kwok et al. (2019) to "translate" the optimizer to the identity.

Lemma C.9 *Suppose that $f : \mathrm{PD}_1(p) \to \mathbb{R}$ is geodesically λ-strongly convex on the geodesic ball of radius κ about Z_0, and that $\nabla f(Z_0) = 0$. If $\|\nabla f(Y)\|_F < \lambda\kappa/8$, then Y is contained in a sublevel set of f on which f is geodesically λ-strongly convex.*

Proof By replacing the function with $X \mapsto f(\sqrt{Z_0}X\sqrt{Z_0})$, it is enough to prove the lemma when $Z_0 = I_p$. Let S be the largest level set contained in the geodesic ball B of radius κ about I_p. That is, if $h := \min_{X \in \partial B} f(X)$, then
$$S = \{X : f(X) \leq h\}.$$

It suffices to show that if Y is outside S, then $\|\nabla f(Y)\|_F \geq \kappa/8\lambda$. Let X be a point in ∂B, i.e. a point with $\|\log X\|_F = \kappa$, with $f(X) = h$. Let $H = \log X/\|\log X\|_F$. By Eq. (25), we have $f(X) = h \geq \lambda\kappa^2/2$. Now let $Y \notin S$. By continuity, there is a value t_0 between 0 and κ such that $f(e^{tK}) = h$ where $K = \log Y/\|\log Y\|_F$. By the mean value theorem, there is a value t between 0 and κ such that $\frac{d}{dt}f(e^{tX}) \geq h/t_0 \geq \kappa/2$. By convexity and Cauchy-Schwarz, $\|\nabla f(Y)\|_F \geq \frac{d}{dt}f(e^{tX})|_{t=\kappa} \geq h/t_0 \geq \kappa/2$. ∎

Finally we may prove Theorem 4.1.

Proof [Proof of Theorem 4.1] By Corollary C.7, we may apply Lemma C.9 with $\kappa = c\lambda$ and $Z_0 = I_p$ to conclude Item 2. Applying Lemma C.8 with $\kappa = c\lambda$ and $Z_0 = I_p$ and using geodesic convexity yields Item 1. ∎

Proof [Proof of Lemma 4.2]

Recall that in Sinkhorn's algorithm, Z is chosen so that $\frac{p}{n} Z \Phi^*(X) = I_p$, where X was chosen such that $X \Phi(Z_0) = I_n$. Now

$$
\begin{aligned}
f^\Phi(Z) &= \frac{p}{n} \log \det \Phi(Z) - \log \det(Z) \\
&= \frac{p}{n} \left(\log \det(\Phi(Z)X) - \log \det(X) \right) - \log \det(Z) \\
&\leq p \log \left(\frac{1}{n} \operatorname{tr} \Phi(Z)X \right) - \frac{p}{n} \log \det(X) - \log \det(Z) \\
&= p \log \left(\frac{1}{n} \operatorname{tr} Z \Phi^*(X) \right) - \frac{p}{n} \log \det(X) - \log \det(Z) \\
&= f^\Phi(Z_0) + \log \det(W),
\end{aligned}
$$

for $W := \sqrt{Z_0} Z^{-1} \sqrt{Z_0}$. Here the inequality is Jensen's inequality and the last equality is by our choice of Z_0 and Z. Compute

$$
\nabla f^\Phi(Z_0) = \frac{p}{n} \sqrt{Z_0} \Phi^* (\Phi(Z_0)^{-1}) \sqrt{Z_0} - I_p = \frac{p}{n} \sqrt{Z_0} \Phi^*(X) \sqrt{Z_0} - I_p = W - I_p.
$$

Note that $\operatorname{tr} W = \operatorname{tr} Z_0 Z^{-1} = \frac{p}{n} \operatorname{tr} Z_0 \Phi^*(X) = \frac{p}{n} \operatorname{tr} \Phi(Z_0)X = p$. By Lemma 5.1 of Garg et al. (2017b) (a robust version of the AM-GM inequality)

$$
\log \det(W) \leq -\frac{1}{6} \|W - I_p\|_F = -\frac{1}{6} \|\nabla f^\Phi(Z_0)\|_F^2.
$$

∎

Proof [Proof of Lemma 4.6] By Theorem 2.5, a unique minimizer $Y \in \mathrm{PD}_1(p)$ to f^Φ exists (i.e., a matrix Y such that $Y^{1/2}$ and $\Phi(Y)^{-1/2}$ are scaling solutions) satisfying $\|I_p - Y\|_{op}, \|I_p - \Phi(Y)^{-1}\|_{op} = O(\delta \log p / \lambda)$. By Lemma C.6, the map $\Psi := \Phi_{\sqrt{Y}, \Phi(Y)^{-1/2}}$ is a $(0, 1 - \lambda + O(\delta \log p / \lambda))$-quantum expander. By uniqueness of Y, $L^\dagger L, R^\dagger R$ match $Y, \Phi(Y)^{-1}$ up to scalar multiples, and it is easy to verify that expansion of $\Phi_{L,R}$ depends only on the equivalence class of $L^\dagger L, R^\dagger R$ under scalar multiplication. ∎

Appendix D. Normalization conventions and error metric

Define $d(A, B) := \|I_p - A^{1/2} B^{-1} A^{1/2}\|_{op}$ for $A, B \in \mathrm{PD}(p)$. Note that

$$
d(A, B) \leq \varepsilon \iff B^{-1} = A^{-1}(1 + \varepsilon),
$$

where for $X \in \mathrm{Herm}(p)$ the short-hand $X(1 \pm \varepsilon)$ denotes the interval of $Y \in \mathrm{Herm}(p)$ such that $X(1 - \varepsilon) \preceq Y \preceq X(1 + \varepsilon)$.

Lemma D.1 *Let $A, B, C \in \mathrm{PD}(p)$.*

- *provided $d(A, B), d(B, C)$ are small enough constants,*

$$
d(A, C) = O(D(A, B) + d(B, C)).
$$

- *provided $d(A, B)$ is at most a small enough constant,*

$$d(B, A) = O(d(A, B)).$$

- *provided $d(A, B)$ is at most a small enough constant,*

$$d(A^{-1}, B^{-1}) = O(d(A, B)).$$

Proof *The first item follows because*

$$\begin{aligned} C^{-1} &\in B^{-1}(1 \pm d(B, C)) \\ &\subset A^{-1}(1 \pm d(A, B))(1 \pm d(B, C)) \\ &\subset A^{-1}(1 \pm O(d(A, B) + d(B, C))). \end{aligned}$$

The second item follows because $B^{-1} \in A^{-1}(1 \pm d(A, B))$ implies $A^{-1} \in B^{-1}(1 \pm O(d(A, B)))$ provided ε is small enough. The third item follows because for $A, B \in \mathrm{PD}(p)$ we have $A \preceq B$ if and only if $A^{-1} \succeq B^{-1}$. ∎

Lemma D.2 *Suppose $A, B \in \mathrm{PD}(p)$ satisfy $\mathrm{tr}\, A = \mathrm{tr}\, B = 1$, and let $A_1 = \det(A)^{-1/p}A$, $B_1 = \det(B)^{-1/p}B \in \mathrm{PD}_1(p)$. Then if $d(A_1, B_1)$ is at most a small constant,*

$$d(A, B) = O(d(A_1, B_1)).$$

Proof Note that $A = pA_1/\mathrm{tr}\, A_1, B = pB_1/\mathrm{tr}\, B_1$. Because for $A, B \in \mathrm{PD}(p)$ we have $A \preceq B$ if and only if $A^{-1} \succeq B^{-1}$, we have that $A \in B(1 \pm O(d(A_1, B_1)))$. This tells us $\mathrm{tr}\, A \in (1 \pm O(d(A_1, B_1)))\,\mathrm{tr}\, B$, so

$$A^{-1/2}BA^{-1/2} \in (1 \pm O(d(A_1, B_1)))A_1^{-1/2}B_1A_1^{-1/2} \subset (1 \pm O(d(A_1, B_1)))I_p.$$

This completes the proof. ∎

Appendix E. Existence and uniqueness

Here we prove Item 3 of Theorem 2.1.

Proof For the sufficiency of the condition we refer to Tyler (1987). It remains to prove that it is necessary. Suppose Eq. (4) has a solution $A \in \mathrm{PD}_p$ and let $y_i = A^{-1/2}x_i/\|A^{-1/2}x_i\|$ so that y_i are in radial isotropic position, i.e. $\|y_i\| = 1$ and $\frac{p}{n}\sum_{i=1}^n y_i y_i^\dagger = I_p$. Suppose the condition in Item 3 is violated. Clearly, because A is a linear transformation, the condition is violated for y_i also, i.e. there is a k-dimensional subspace $0 \subsetneq W \subsetneq \mathbb{R}^p$ containing at least nk/p points y_i. Let π_W denote the orthogonal projection to W. Then

$$\mathrm{tr}\, \pi_W = \frac{p}{n}\sum_{i=1}^n \mathrm{tr}\, \pi_W y_i y_i^\dagger \geq \frac{p}{n}(nk/p),$$

but both sides are equal to k and so the inequality must be an equality. Equality can hold if and only if $\pi_W y_i = 0$ for all $y_i \notin W$. That is, for every i we have that $y_i \in W$ or $y_i \in W^\perp$. Then for any $a > 0$, the matrix $B = (a\pi_W + a^{-1}\pi_{W^\perp})$ satisfies $By_i/\|By_i\| = y_i$. Thus $A^{1/2}B^{-1}A^{1/2}$, renormalized if necessary, also solves Eq. (4). By the invertibility of A, $A^{1/2}B^{-1}A^{1/2}$ constitutes a family of distinct solutions to Eq. (4), none of which is a constant multiple of the other. ∎