Optimal Bounds for Tyler's M-Estimator for Elliptical Distributions

Akshay Ramachandran ARAMACH@CS.UBC.CA

Lap Chi Lau LAPCHI@UWATERLOO.CA

Editors: Matus Telgarsky and Jonathan Ullman

Abstract

A fundamental problem in statistics is estimating the shape matrix of an Elliptical distribution. This generalizes the familiar problem of Gaussian covariance estimation, for which the sample covariance achieves optimal estimation error. For Elliptical distributions, Tyler proposed a natural M-estimator and showed strong statistical properties in the asymptotic regime, independent of the underlying distribution. Numerical experiments show that this estimator performs very well, and that Tyler's iterative procedure converges quickly to the estimator. Franks and Moitra recently provided the first distribution-free error bounds in the finite sample setting, as well as the first rigorous convergence analysis of Tyler's iterative procedure. However, their results exceed the sample complexity of the Gaussian setting by a $\log^2 d$ factor. We close this gap by proving optimal sample threshold and error bounds for Tyler's M-estimator for all Elliptical distributions, fully matching the Gaussian result. Moreover, we recover the algorithmic convergence even at this lower sample threshold. Our approach builds on the operator scaling connection of Franks and Moitra by introducing a novel 'pseudorandom' condition, which we call ∞-expansion. We show that Elliptical distributions satisfy ∞-expansion at the optimal sample threshold, and then prove a novel scaling result for inputs satisfying this condition.

Keywords: Elliptical distribution, Robust statistics, Frame and Operator Scaling

1. Introduction

The covariance matrix of random variable is a natural and useful statistic of high dimensional distributions, as it gives insight into the geometry of the data. Estimation of the covariance is therefore a fundamental task in data analysis. For sufficiently nice distributions, such as multivariate Gaussians, the sample covariance is a very accurate estimator that is easy to compute. However in many practical situations, the underlying distribution is less well-behaved, so the sample covariance is less accurate. In fact, for sufficiently heavy-tailed distributions such as the multivariate t-distribution, the covariance matrix may not even exist Wainwright (2019).

Elliptical distributions are a well-studied model class of heavy tailed distributions (see Kelker (1970), Gupta et al. (2013)), and include important special cases such as multivariate Gaussian and t-distributions. While Elliptical distributions do not always have a well-defined covariance matrix, they are parameterized by a 'shape matrix' which captures similar geometric properties. Tyler (1987) proposed a natural estimator for the shape matrix, along with a simple iterative procedure to compute the estimator. He was able to prove strong asymptotic guarantees for this estimator, essentially recovering some of the desirable statistical properties of the Gaussian setting. Importantly, the asymptotic guarantees shown by Tyler (1987) are *distribution-free*, in that they are independent

of the underlying distribution or shape matrix. Both the estimator and the iterative procedure have been shown to perform well in numerical experiments.

Following this seminal result, there have been attempts to show estimation guarantees for Elliptical distributions in the finite sample regime. Soloveychik and Wiesel (2014) showed that Tyler's M-Estimator achieved optimal error in the Frobenius norm, but with an additional factor that depends on the condition number of the shape matrix. Regularized estimators have also been proposed, which can be computed efficiently but do not have the same provable statistical properties as Tyler's estimator. For a thorough discussion of these results, see the survey of Wiesel et al. (2015).

The first distribution-free guarantees for Tyler's M-Estimator were proven recently in Franks and Moitra (2020); they also gave the first rigorous analysis of Tyler's iterative procedure, showing linear convergence to the estimator. These results are nearly optimal, but have sample threshold and error results that exceed the Gaussian setting by $\log d$ factors. It is natural to ask whether Tyler's estimator can be shown to match this optimal guarantee, or whether shape estimation for Elliptical distributions is strictly more difficult than Gaussian covariance estimation.

In this work, we show optimal sample complexity and error guarantees for Tyler's M-Estimator for the shape matrix of Elliptical distributions. These results are tight, as they match the known lower bounds for the special case of Gaussians. We also recover the algorithmic analysis of Franks and Moitra (2020) with fewer samples, showing the same linear convergence of Tyler's iterative procedure at the optimal sample threshold.

1.1. Our Results

Our formal sample complexity result is as follows:

Theorem 1 *Given $n \gtrsim \frac{d}{\varepsilon^2}$ samples from an elliptical distribution with shape $\Sigma \in \mathrm{PD}(d)$, where ε is at most a small constant, Tyler's M-Estimator $\hat{\Sigma}$ satisfies*

$$\|I_d - \Sigma^{1/2}\hat{\Sigma}^{-1}\Sigma^{1/2}\|_{\mathrm{op}} \leq \varepsilon, \quad \textit{and} \quad \|I_d - \Sigma^{1/2}\hat{\Sigma}^{-1}\Sigma^{1/2}\|_F \leq \varepsilon\sqrt{d}$$

with probability $\geq 1 - \exp(-\Omega(\varepsilon^2 n))$.

This improves on Theorem 1.1 of Franks and Moitra (2020) by removing $\log d$ factors from the sample threshold and error bound. It also matches the error guarantee of Theorem 1.2 of Franks and Moitra (2020) in the Frobenius norm, while improving the sample threshold from $n \gtrsim d^2$ to $n \gtrsim d$. Both the sample threshold and error rate of our result are optimal up to constant factors, as they match the lower bound for the special case of covariance estimation for multivariate Gaussians.

Our measure of error in Theorem 1 is known as the relative operator norm. This is the relevant measure for statistical applications: Arbas et al. (2023) show strong bounds on the relative operator norm bounds between $\Sigma, \hat{\Sigma}$ imply similar strong bounds on KL-divergence and total variation distance between the corresponding Gaussian distributions $N(0, \Sigma), N(0, \hat{\Sigma})$.

For our second main result, we recover the algorithmic convergence of Franks and Moitra (2020) with fewer samples.

Theorem 2 *Given $n \gtrsim d$ samples from an elliptical distribution with shape $\Sigma \in \mathrm{PD}(d)$, with probability $\geq 1 - \exp(-\Omega(n))$ the T-th iterate $\overline{\Sigma}_{(T)}$ of Tyler's procedure approximates the M-Estimator $\hat{\Sigma}$ with error*

$$\|I_d - \hat{\Sigma}^{1/2}\overline{\Sigma}_{(T)}^{-1}\hat{\Sigma}^{1/2}\|_F \leq \delta$$

in $T \lesssim |\log\det\Sigma| + d + \log(1/\delta)$ *iterations.*

This improves on the sample requirement of Theorem 1.3 in Franks and Moitra (2020) by $\log^2 d$. Further, this requirement is optimal, as the estimator is not even well-defined for $n < d$.

In the following subsection, we discuss our techniques, showing how they build on and improve the previous approach of Franks and Moitra (2020).

1.2. Techniques

An important observation of Franks and Moitra (2020) was to show a connection between Tyler's M-Estimator and *frame and operator scaling*. This is an optimization problem over matrices arising in the context of geometric invariant theory, and has recently attracted much interest due to its connections to problems in algebraic complexity (see Garg et al. (2015), Bürgisser et al. (2019)). The key technical contribution in Franks and Moitra (2020) is to show that, for sufficiently large sample size $n \gtrsim d\log d$, the data from Elliptical distributions satisfy a 'quantum expansion' property. They can then appeal to sophisticated scaling results of Kwok et al. (2021) for quantum expanders to prove their results for Tyler's M-Estimator.

In this work, we follow a similar approach, improving both parts of the argument to prove optimal guarantees. We first identify a stronger 'pseudorandom' condition, ∞-expansion (see Definition 13), and prove that $n \gtrsim d$ samples from an Elliptical distribution satisfies this condition with high probability. Then we give an improved analysis of frame scaling, showing the ∞-expansion condition implies stronger operator norm bounds than the results of Kwok et al. (2021). We believe this result is of independent interest, and in Section 7 we discuss potential future applications to the Paulsen problem in frame theory and the tensor normal model in statistics. Our sample complexity results follow by combining the above two steps.

Our proof of the algorithmic result in Theorem 2 also crucially uses the connection to scaling. In Franks and Moitra (2020), the authors study a *geodesically convex* optimization formulation for Tyler's M-Estimator. They use this perspective to show: (1) Tyler's iterative procedure can be seen as a natural descent method; (2) quantum expansion is related to a geodesic version of *strong convexity*. The convergence follows by standard convex analysis applied in this geodesic setting.

In our work, we use a different 'pseudorandom' condition, ∞-expansion, instead of quantum expansion. Nevertheless we show that the convergence analysis of Franks and Moitra (2020) still follows from our results. Concretely, in Appendix C we show that our ∞-expansion condition implies quantum expansion, which allows us to apply the same argument as in Franks and Moitra (2020) to prove fast convergence of Tyler's iterative procedure at the optimal sample threshold.

2. Preliminaries and Technical Overview

In this section we formally define the statistical estimation problem considered in this work. We also formally define frame scaling and show its connection to Tyler's M-Estimator. We end this section with a proof outline of our results, including the main technical ingredients.

Notation: we use $f \lesssim g$ and $f \leq O(g)$ interchangeably to mean that there is a universal constant $C > 0$ such that $f \leq C \cdot g$, and we use $f \gtrsim g$ and $f \geq \Omega(g)$ for the opposite inequality. $S^{d-1} \subseteq \mathbb{R}^d$ is the set of unit vectors; $\mathrm{PD}(d)$ is the set of positive definite matrices in $\mathbb{R}^{d\times d}$; and $\mathrm{diag}(n)$ is the set of diagonal matrices in $\mathbb{R}^{n\times n}$. For vector $y \in \mathbb{R}^n$, $\mathrm{diag}(y)$ is the diagonal matrix

with entries y_j; and by abuse of notation, for matrix $F \in \mathbb{R}^{n\times n}$, $\mathrm{diag}(F)$ is the diagonal matrix with the same entries as F on the diagonal and remaining entries zero.

2.1. Elliptical Distributions and Tyler's M-Estimator

We begin with the formal definition of our statistical model.

Definition 3 (Elliptical Distribution) *For fixed 'shape' matrix $\Sigma \in \mathrm{PD}(d)$ and scalar random variable $u \in \mathbb{R}$, the Elliptical random variable $X \sim E(\Sigma, u)$ is distributed as*

$$X := \Sigma^{1/2} V \cdot u$$

where $V \sim S^{d-1}$ is a uniformly random unit vector, and u is independent of V.

These generalize the family of Gaussian distributions: $N(0,\Sigma)$ is equivalent to $E(\Sigma, u)$ as the norm $u := \|g\|_2$ and direction $g/\|g\|_2 \sim S^{d-1}$ of standard Gaussian $g \sim N(0, I_d)$ are independent. By computing the second moment, we see that if the covariance matrix of an Elliptical distribution exists, then it must be proportional to the shape matrix.

In this work, we study the following estimator for the shape matrix Σ.

Definition 4 (Tyler's M-Estimator) *Given $x_1, ..., x_n \in \mathbb{R}^d$, consider the following equations:*

$$\frac{d}{n}\sum_{j=1}^{n} \frac{x_j x_j^T}{x_j^T \hat{\Sigma}^{-1} x_j} = \hat{\Sigma}, \qquad \textit{and} \qquad \mathrm{Tr}[\hat{\Sigma}] = d.$$

If the above equations have a unique *solution in $\hat{\Sigma} \in \mathrm{PD}(d)$, then Tyler's M-estimator is defined to be that solution; otherwise it is not well-defined.*

2.2. Frame Scaling

In this subsection, we define frame scaling and describe the connection to Tyler's M-Estimator. Frames are linear algebraic primitives with applications to a variety of fields including coding theory Casazza and Kutyniok (2013), learning theory Diakonikolas et al. (2021), and communication complexity Förster (2002). Formally, they are spanning sets of vectors $V := \{v_1, ..., v_n\} \in \mathbb{R}^{d\times n}$. Note that we can represent frames as matrix $V \in \mathbb{R}^{d\times n}$ or tuple $\{v_1, ..., v_n\} \subseteq \mathbb{R}^d$, and we use these interchangeably depending on the context. Frames can be thought of as overcomplete bases: $x \in \mathbb{R}^d$ can be encoded in a redundant manner in terms of its 'frame coefficients' $\{\langle x, v_j\rangle\}_{j\in[n]}$. Two basic and desirable properties of frames are: (1) isotropy condition $VV^T = I_d$, which allows for easy reconstruction from frame coefficients $x = \sum_{j\in[n]} \langle x, v_j\rangle v_j$; and (2) equal-norm condition $\|v_j\|_2^2 = \frac{d}{n}$, which ensures 'balance' in the sense that no coefficient is too important on average. The following are used to measure the quality of a frame:

Definition 5 *Given frame $V \in \mathbb{R}^{d\times n}$, its size is $s(V) := \|V\|_F^2$, and its error is*

$$E(V) := d \cdot VV^T - s(V) I_d, \qquad F(V) := \mathrm{diag}(n \cdot V^T V - s(V) I_n).$$

Observe that $E(V)$ measures distance from the isotropy condition, and $F(V)$ measures distance from the equal-norm condition. The goal of frame scaling is to transform a given frame to satisfy these two conditions simultaneously. We will use the following measures of error:

Definition 6 *For frame $V \in \mathbb{R}^{d\times n}$, the ℓ_2 and ℓ_∞ error measures are*

$$\Delta(V) := \frac{1}{d}\|E(V)\|_F^2 + \frac{1}{n}\|F(V)\|_F^2, \qquad \|(E,F)\|_{\mathrm{op}} := \max\{\|E\|_{\mathrm{op}}, \|F\|_{\mathrm{op}}\}.$$

V is ε-doubly balanced if $\|(E,F)\|_{\mathrm{op}} \le s(V)\cdot\varepsilon$, and is doubly balanced if $\varepsilon = 0$.

We note the following simple relation for later use.

Lemma 7 (Lemma 2.15 in Kwok et al. (2021)) *For frame $V \in \mathbb{R}^{d\times n}$, $\Delta(V) \le \|E(V)\|_{\mathrm{op}}^2 + \|F(V)\|_{\mathrm{op}}^2$. In particular if V is ε-doubly balanced then $\Delta(V) \le 2s(V)^2\cdot\varepsilon^2$.*

We can now formally define frame scaling.

Definition 8 (Frame Scaling Problem) *Given frame $U \in \mathbb{R}^{d\times n}$, find left/right scalings $L \in \mathbb{R}^{d\times d}$ and $R \in \mathrm{diag}(n)$ such that $V := LUR$ is doubly balanced:*

$$VV^T = \frac{s(V)}{d} I_d, \qquad \forall j \in [n] : \|v_j\|_2^2 = \frac{s(V)}{n}.$$

The key insight in Franks and Moitra (2020) is the following connection between frame scaling and Tyler's M-Estimator, which can be verified directly by comparing definitions.

Lemma 9 (Example 2.3 in Franks and Moitra (2020)) *Consider input $X = \{x_1, ..., x_n\} \in \mathbb{R}^{d\times n}$.*

- *If $\hat{\Sigma}$ is the M-Estimator for input $\{x_1, ..., x_n\}$ according to Definition 4, then the following defines a scaling solution for frame X according to Definition 8:*

$$L := \hat{\Sigma}^{-1/2} \quad , \quad R_{jj} := (\langle x_j, \hat{\Sigma}^{-1} x_j\rangle)^{-1/2}.$$

- *Conversely, if (L,R) is a frame scaling solution for X, then the following satisfies the equations in Definition 4 for Tyler's M-Estimator:*

$$\hat{\Sigma} = \frac{d\cdot(L^TL)^{-1}}{\mathrm{Tr}[(L^TL)^{-1}]}.$$

This insight allows us to analyze Tyler's M-Estimator for input $\{x_1, ..., x_n\}$ by studying the scaling solution for frame X.

2.3. Technical Overview

In this subsection, we give an outline of our proof, including our main technical contributions.

The first step, following Franks and Moitra (2020), is to notice some useful invariance properties: the equations for Tyler's M-Estimator in Definition 4 are unaffected by scalar transformation $x_j \to \lambda x_j$ for $\lambda \in \mathbb{R}$; similarly, a direct computation shows that if $\hat{\Sigma}$ satisfies the equations in Definition 4 for input $\{x_j\}$, then for any invertible $A \in \mathbb{R}^{d\times d}$, $A\hat{\Sigma}A^T$ satisfies the equations for transformed input $\{Ax_j\}$; finally, the relative operator error considered in Theorem 1 is invariant under linear transformations:

$$(A\Sigma A^T)^{1/2}(A\hat{\Sigma}A^T)^{-1}(A\Sigma A^T)^{1/2} = \Sigma^{1/2}\hat{\Sigma}^{-1}\Sigma^{1/2}.$$

Combining these gives a reduction from general Elliptical distributions to a simpler special case:

Lemma 10 (Observation 3.1 in Franks and Moitra (2020)) *Consider samples $x_1, ..., x_n \sim E(\Sigma, u)$ from the elliptical distribution with shape matrix Σ. Then the 'normalized' data $v_j := \frac{\Sigma^{-1/2} v_j}{\|\Sigma^{-1/2} v_j\|_2}$ follows Elliptical distribution $E(I_d, 1)$, where 1 denotes the deterministic random variable that always takes value 1. Further, if $\hat{\Sigma}_X, \hat{\Sigma}_V$ are M-estimators for X, V according to Definition 4, then*

$$\hat{\Sigma}_V = \Sigma^{-1/2} \hat{\Sigma}_X \Sigma^{-1/2} \quad \textit{and} \qquad \|I_d - \Sigma^{1/2} \hat{\Sigma}_X^{-1} \Sigma^{1/2}\|_{\mathrm{op}} = \|I_d - \hat{\Sigma}_V^{-1}\|_{\mathrm{op}}.$$

This shows, in order to prove error bounds for Tyler's M-Estimator in the relative operator norm, it suffices to study the special case of $E(I_d, 1)$. We note that this reduction is only for the sake of analysis, as we do not have knowledge of the true shape matrix when computing the estimator. The distribution $E(I_d, 1)$ is equivalent to the uniform distribution on the sphere S^{d-1}, and we will use concentration properties of this simple distribution for our results.

Next, we use the connection between Tyler's M-Estimator and frame scaling as described in Lemma 9. If we can prove strong bounds on the left scaling solution for frame $V \in \mathbb{R}^{d \times n}$ with columns $v_j \sim S^{d-1}$, then this transfers directly to error bounds on the estimator.

In Franks and Moitra (2020), the authors used the following property to analyze frame scaling.

Definition 11 (Quantum Expansion) *Frame $V \in \mathbb{R}^{d \times n}$ is a $(1 - \lambda)$ quantum expander if*

$$\forall y \perp 1_n, \|y\|_2 \leq 1 : \quad \left\| \sum_{j \in [n]} y_j v_j v_j^* \right\|_F \leq \frac{s(V)(1 - \lambda)}{\sqrt{dn}}.$$

This is also called the 'spectral gap condition' in the work of Kwok et al. (2021), where the authors show the following strong scaling bound:

Theorem 12 (Theorem 1.7 in Kwok et al. (2021)) *For ε-doubly balanced frame $V \in \mathbb{R}^{d \times n}$, if V satisfies $(1 - \lambda)$-quantum expansion for $\lambda^2 \gtrsim \varepsilon \log d$, then the scaling solution satisfies*

$$\|L - I_d\|_{\mathrm{op}} \lesssim \frac{\varepsilon \log d}{\lambda}.$$

The main technical contribution of Franks and Moitra (2020) is to show that random instances $v_1, ..., v_n \sim S^{d-1}$ are quantum expanders with high probability when $n \gtrsim d \log d$. This allows them to apply the above scaling result to show near-optimal bounds for Tyler's M-Estimator.

Our key insight for proving optimal bounds involves the following novel notion of expansion.

Definition 13 (∞-Expansion) *Frame $V \in \mathbb{R}^{d \times n}$ of size $s(V)$ satisfies $(1 - \lambda)$-∞-expansion if*

$$\forall y \perp 1_n, \|y\|_\infty \leq 1 : \quad \left\| \sum_{j \in [n]} y_j v_j v_j^* \right\|_{\mathrm{op}} \leq \frac{s(V)(1 - \lambda)}{d}.$$

As compared to Definition 11, this condition involves all $\|y\|_\infty \leq 1$ as opposed to all $\|y\|_2 \leq 1$, while the output is now bounded in terms of the operator norm instead of the Frobenius norm. This is in fact a stronger property, as in Theorem 26 we show ∞-expansion implies quantum expansion. We use this stronger condition to give an improved analysis of frame scaling.

Theorem 14 *For ε-doubly balanced frame $V \in \mathbb{R}^{d\times n}$, if V satisfies $(1-\lambda)$-∞-expansion for $\lambda^2 \gtrsim \varepsilon$, then the scaling solution satisfies*

$$\|L - I_d\|_{\mathrm{op}} \lesssim \frac{\varepsilon}{\lambda}.$$

As compared to Theorem 12, our result improves the scaling bound and expansion requirement, both by a $\log d$ factor, but we use the stronger ∞-expansion condition instead of quantum expansion. These quantitative improvements are the key to our optimal sample complexity bounds.

In order to apply the above scaling result, we show that random instances satisfy ∞-expansion.

Theorem 15 *Given $v_1, ..., v_n \sim S^{d-1}$ for $n \gtrsim d$ large enough, the input V satisfies $(1-\lambda)$-∞-expansion for $\lambda \geq \Omega(1)$ with probability at least $1 - \exp(-\Omega(n))$.*

This improves the sample requirement of Theorem 2.10 in Franks and Moitra (2020) by a $\log d$ factor while showing the stronger ∞-expansion condition, as opposed to quantum expansion. Our main sample complexity result, Theorem 1, follows by combining the above ingredients.

Our approach to the algorithmic convergence result Theorem 2 also relies on ∞-expansion. Franks and Moitra (2020) showed that quantum expansion is essentially equivalent to strong convexity of a natural potential function for frame scaling. Further, Tyler's iterative approach can be seen as a natural descent method for this potential function, so the convergence bound follows by standard arguments from convex analysis.

In Theorem 26 we show that ∞-expansion implies quantum expansion. This directly implies that the convergence analysis of Franks and Moitra (2020) for Tyler's iterative procedure holds at the optimal sample threshold $n \gtrsim d$.

2.4. Organization

We first prove our sample complexity result Theorem 1 in Section 3, assuming our two main technical ingredients. In Section 4, we prove Theorem 14 which gives a novel scaling result for inputs satisfying ∞-expansion. In Section 5, we prove Theorem 15, showing random random frames satisfy ∞-expansion. In Section 6, we prove Theorem 2, showing convergence of Tyler's iterative procedure. We defer some of the technical details from these proofs to Appendix A, B, and C, respectively. We conclude in Section 7 with some related problems.

3. Optimal Sample Complexity via ∞-Expansion

In this section, we show Theorem 1 as a straightforward consequence of our two main technical ingredients, which are proved in the following two sections. In order to apply our new frame scaling result in Theorem 14, we need the input to be ε-doubly balanced. For this we apply the following standard concentration result for random unit vectors:

Theorem 16 (Theorem 5.14 in Franks and Moitra (2020), Theorem 5.39 in Vershynin (2010)) *Given $n \gtrsim \frac{d}{\varepsilon^2}$ uniformly random unit vectors $v_1, ..., v_n \sim S^{d-1}$, where ε is at most a small constant,*

$$\left\| \frac{d}{n} \sum_{j\in[n]} v_j v_j^T - I_d \right\|_{\mathrm{op}} \leq \varepsilon$$

with probability at least $1 - \exp(-\Omega(\varepsilon^2 n))$.

We can now carry out the argument presented in Section 2.3 for our sample complexity bound.

Proof [Proof of Theorem 1] By Lemma 10, we can assume the underlying distribution is $E(I_d, 1)$, so our input frame $V \in \mathbb{R}^{d\times n}$ has columns distributed according to $v_j \sim S^{d-1}$. For $n \gtrsim \frac{d}{\varepsilon^2}$ large enough, Theorem 16 implies V is ε-doubly balanced, and Theorem 15 implies V satisfies $(1-\lambda)$-∞-expansion for $\lambda^2 \geq \Omega(1) \gtrsim \varepsilon$, both occurring with probability $\geq 1-\exp(-\Omega(\varepsilon^2 n))$. Therefore, we can apply Theorem 14 to input V and bound the frame scaling solution L

$$\|L - I_d\|_{\mathrm{op}} \lesssim \frac{\varepsilon}{\lambda} \lesssim \varepsilon,$$

as $\lambda \geq \Omega(1)$. Finally, we can use Lemma 9 to bound the M-Estimator $\hat{\Sigma} = \frac{d(L^TL)^{-1}}{\mathrm{Tr}[(L^TL)^{-1}]}$:

$$\|I_d - \Sigma^{1/2}\hat{\Sigma}^{-1}\Sigma^{1/2}\|_{\mathrm{op}} \lesssim \|L - I_d\|_{\mathrm{op}} \lesssim \varepsilon,$$

where we used $\Sigma = I_d$, as well as Taylor approximation bounds $\frac{1}{d}\mathrm{Tr}[(L^TL)^{-1}] = 1 + O(\|L - I_d\|_{\mathrm{op}})$ and $\|I_d - L^TL\|_{\mathrm{op}} \lesssim \|L - I_d\|_{\mathrm{op}}$. ■

4. Frame Scaling with ∞-Expansion

The goal of this section is to prove Theorem 14, which shows optimal bounds for frame scaling when the input satisfies the ∞-expansion condition in Definition 13. We follow the approach of Kwok et al. (2021), which proved a slightly weaker bound using the weaker quantum expansion condition. Our analysis lifts to the more general operator scaling problem, but we omit this as our setting only involves frames.

We first collect useful facts about frame scaling from Kwok et al. (2021). Then, we state the main technical lemma where we use ∞-expansion. The final ingredient is a robustness result, showing ∞-expansion is preserved under small scalings. The full proof is given in Appendix A.

The main character of our analysis is a dynamical system that converges to the frame scaling solution. It can be interpreted as a continuous version of the natural Flip-Flop algorithm, so we begin by motivating and defining this procedure. Frame scaling (Definition 8) involves two conditions, isotropy and equal-norm, each of which is easy to satisfy individually:

$$L := (VV^T)^{-1/2} \implies LVV^TL^T = I_d; \quad R := \mathrm{diag}(V^TV)^{-1/2} \implies \mathrm{diag}(R^TV^TVR) = I_n.$$

This suggests the following procedure, devised by Gurvits (2004) for operator scaling:

$$V_{t+1} \leftarrow (V_tV_t)^{-1/2}V_t, \qquad V_{t+2} \leftarrow V_{t+1}\,\mathrm{diag}(V_{t+1}^TV_{t+1})^{-1/2}. \tag{1}$$

We note that this procedure is intimately tied to Tyler's iteration, and we discuss this connection in more detail in Section 6. The seminal work of Garg et al. (2015) showed that this simple algorithm converges to the scaling solution (even for the more general operator scaling problem). Their goal was to prove worst-case time complexity bounds, and they were able to use deep results from geometric invariant theory to show polynomial time convergence of the above discrete algorithm.

In our setting, as well as in Kwok et al. (2021), we want to prove beyond worst-case bounds for scaling when the input is already nearly doubly balanced. Therefore we study the following continuous process which allows for more refined control of the trajectory, as opposed to the the Flip-Flop algorithm which may take large steps in each iteration.

Definition 17 (Dynamical System, Def 2.16 in Kwok et al. (2021)) *For frame $V \in \mathbb{R}^{d\times n}$ with size $s(V)$ and error matrices $E(V), F(V)$ according to Definition 5, V_t is the solution to the following differential equation:*

$$-\partial_t V_t = E(V_t)V_t + V_t F(V_t) = (dV_tV_t^T - s(V_t)I_d)V_t + V_t \operatorname{diag}(nV_t^TV_t - s(V_t)I_n), \qquad V_0 = V.$$

Intuitively, the first term $E(V)V$ improves the isotropy condition, and the second $VF(V)$ drives towards equal norms. This can be seen as a continuous and simultaneous version of Flip-Flop.

It turns out that the above dynamical system automatically produces a frame scaling solution! The following result allows us to control the scaling solutions throughout the trajectory.

Lemma 18 (Corollary 3.12 and 3.15 in Kwok et al. (2021)) *The solution to the dynamical system in Definition 17 produces a frame scaling of the form $V_t = L_tVR_t$, where $L_t \in \mathbb{R}^{d\times d}, R_t \in \operatorname{diag}(n)$. Further, these scalings can be bounded as*

$$\|L_T - I_d\|_{\mathrm{op}} \le \exp\left(\int_0^T \|E(V_t)\|_{\mathrm{op}}\right) - 1; \quad \|R_T - I_d\|_{\mathrm{op}} \le \exp\left(\int_0^T \|F(V_t)\|_{\mathrm{op}}\right) - 1.$$

Therefore, in order to prove strong scaling bounds, it suffices to show that the error $(E(V), F(V))$ decreases quickly. The following lemma is the key new step in our analysis, where we use ∞-expansion condition to show convergence of the error.

Lemma 19 (Informal) *For ε-doubly balanced $V \in \mathbb{R}^{d\times n}$ satisfying $(1-\lambda)$-∞-expansion for $\lambda \gtrsim \varepsilon$,*

$$-\partial_{t=0}\|(E_t, F_t)\|_{\mathrm{op}} \gtrsim s(V)\cdot\lambda\cdot\|(E_t, F_t)\|_{\mathrm{op}}.$$

We defer the formal statement and its proof to Appendix A. The takeaway is that ∞-expansion implies operator norm error decreases at an exponential rate. This is analogous to Proposition 3.9 in Kwok et al. (2021) which shows exponential convergence of Δ using quantum expansion. This justifies Definition 13 as we are able to use the stronger operator norm condition to prove stronger operator norm bounds.

The final step is to show that ∞-expansion is maintained throughout the dynamical system:

Lemma 20 (Analogue of Lemma 3.20 in Kwok et al. (2021)) *Let $V \in \mathbb{R}^{d\times n}$ be ε-doubly balanced for $\varepsilon \le 1/2$, and consider scaling $U = LVR$ with $\delta := \max\{\|L - I_d\|_{\mathrm{op}}, \|R - I_n\|_{\mathrm{op}}\} \le 1/2$. If V satisfies $(1-\lambda)$-∞-expansion, then U satisfies $(1-\lambda')$-∞-expansion with*

$$s(U)(1-\lambda') \le s(V)(1-\lambda) + O(\delta) \implies \lambda' \ge 1 - \frac{s(V)(1-\lambda) + O(\delta)}{s(U)}.$$

We omit the proof as it is identical to that of Lemma 3.20 in Kwok et al. (2021), replacing $\|\cdot\|_F$ for quantum expansion with $\|\cdot\|_{\mathrm{op}}$ for ∞-expansion.

Theorem 14 now follows by combining these ingredients. Namely, Lemma 19 shows fast convergence of the error; by Lemma 18, this implies the scaling cannot become too large; finally, Lemma 20 implies that this maintains ∞-expansion. Therefore, we have fast convergence for all time, which implies strong scaling bounds by Lemma 18. The formal proof is given in Appendix A.

5. ∞-Expansion for Random Frames

In this section, we show random unit vector frames satisfy ∞-expansion with high probability. Our proof strategy is as follows: first, we reduce ∞-expansion to a simpler 'pseudorandom' condition involving column subsets; next, for technical reasons, we first show the pseudorandom condition for Gaussian frames; finally, we show that this property is maintained after normalization $v_j = \frac{g_j}{\|g_j\|_2}$, which proves pseudorandomness for random unit vectors.

We begin by reducing our expansion condition to a simpler condition involving column subsets. Recall that Definition 13 involves a bound on $\|\sum_j y_j v_j v_j^T\|_{\text{op}}$ for all $y \perp 1_n, \|y\|_\infty \leq 1$. We instead consider a simpler condition involving only column subsets, i.e. $y \in \{0,1\}^n$.

Definition 21 (Pseudorandom condition) *Frame $V \in \mathbb{R}^{d\times n}$ is $(\alpha_{\min}, \alpha_{\max}, \beta)$-pseudorandom if*

$$\forall B \subseteq [n], |B| = \beta n : \quad \beta\frac{\alpha_{\min}}{d} I_d \preceq V_B V_B^T \preceq \beta\frac{\alpha_{\max}}{d} I_d.$$

We use $(\alpha_{\min}, \beta)$-pseudorandomness to denote that only the lower bound condition is satisfied.

This condition is related to, and slightly weaker than, the pseudorandom condition defined in Kwok et al. (2017), which was used in the solution of the Paulsen problem in frame theory.

We next show that $\beta = 1/2$ pseudorandomness and ∞-expansion are essentially equivalent. We defer the proof to Appendix B.

Lemma 22 *Let $V \in \mathbb{R}^{d\times n}$ be an ε-doubly balanced frame. If V $(\alpha_{\min}, \alpha_{\max}, \frac{1}{2})$-pseudorandom, then V satisfies $(1-\lambda)$-∞-expansion with*

$$s(V)(1-\lambda) \leq \min\{s(V)(1+\varepsilon) - \alpha_{\min} \quad , \quad \alpha_{\max} - s(V)(1-\varepsilon)\}.$$

Conversely, if V satisfies $(1-\lambda)$-∞-expansion, then it is $(\alpha_{\min}, \alpha_{\max}, \frac{1}{2})$-pseudorandom with

$$s(V)(\lambda - \varepsilon) \leq \alpha_{\min} \leq \alpha_{\max} \leq s(V)(2 - (\lambda - \varepsilon)).$$

We would like to apply concentration to prove the pseudorandom condition for random unit vectors. This would require a union bound over all subsets $\binom{[n]}{\beta n}$, which for $\beta = 1/2$ has cardinality $\exp(\Omega(n))$. The failure probability $\exp(-\varepsilon^2 n)$ in Theorem 16 is too high for this approach.

As a workaround, we note that the uniform distribution $v \sim S^{d-1}$ is equivalent to first sampling a random Gaussian vector $g \sim N(0, I_d)$, and then normalizing $v \leftarrow g/\|g\|_2$. It turns out that we can use a slightly stronger concentration bound to show the pseudorandom condition for Gaussians.

Theorem 23 *Gaussian frame $G \in \mathbb{R}^{d\times n}$, with entries $G_{ij} \sim N(0, \frac{1}{nd})$, is $(\alpha_{\min} \geq \Omega(1), \alpha_{\max} \leq O(1), \beta = \frac{1}{4})$-pseudorandom according to Definition 21.*

We also have that normalization does not affect the pseudorandom property too much.

Lemma 24 *For frame G that is $(\alpha_{\min}, \alpha_{\max}, \beta)$-pseudorandom, normalized frame $v_j = \frac{g_j}{\|g_j\|_2}$ has size $s(V) = n$ and is $(\alpha_{\min}(V), 2\beta)$-pseudorandom for*

$$\alpha_{\min}(V) \geq s(V)\frac{\alpha_{\min}}{2\alpha_{\max}}.$$

We prove these results in Appendix B. We can now combine the above ingredients to prove ∞-expansion for random unit vectors.

Proof [Proof of Theorem 15] Theorem 35 implies $(\Omega(1), O(1), \beta = 1/4)$-pseudorandomness for Gaussian frame $G \sim N(0, \frac{1}{nd} I_d \otimes I_n)$ with high probability $\geq 1 - \exp(-\Omega(n))$. V is equivalently distributed as the normalized Gaussian frame $v_j \leftarrow \frac{g_j}{\|g_j\|_2}$, so Lemma 24 implies that V is $(\alpha_{\min} \geq \Omega(n), \beta = \frac{1}{2})$-pseudorandom. Also, for ε a small constant, Theorem 16 implies V is ε-doubly balanced with probability $\geq 1 - \exp(-\Omega(n))$. Therefore we can apply Lemma 22 to show V satisfies $(1-\lambda)$-∞-expansion for

$$\lambda \geq \frac{\alpha_{\min}(V)}{s(V)} - \varepsilon \geq \Omega(1),$$

where we used $\alpha_{\min} \geq \Omega(n)$, the size is $s(V) = n$, and ε is a small enough constant. ■

6. Convergence of Tyler's Iterative Procedure

In this section we prove Theorem 2, showing Tyler's iterative procedure has linear convergence to the estimator as soon as $n \gtrsim d$. This improves upon the sample requirement of Theorem 1.3 in Franks and Moitra (2020) by a $\log^2 d$ factor. Concretely, we show that the key property used in the analysis of Franks and Moitra (2020), quantum expansion, already holds at this lower sample threshold. This result is optimal up to constants, as the M-estimator does not even exist for $n < d$.

We begin by motivating and defining Tyler's iterative procedure. In the setting of Elliptical distributions, the goal is to estimate the shape matrix using Tyler's M-Estimator. By Lemma 9 this corresponds to the left scaling solution for the sample data. Since the Flip-Flop procedure in Equation (1) computes a sequence of iterates converging to the scaling solution, we could use the above correspondence to compute a sequence of estimators converging to the M-Estimator.

Tyler's iterative procedure exactly follows two iterations of Flip-Flop while keeping the right scaling implicit. Recall two iterations of Flip-Flop gives

$$L_{t+1} \leftarrow (L_t V R_t^2 V^T L_t^T)^{-1/2} L_t \quad , \quad R_{t+1} \leftarrow R_t \operatorname{diag}(R_t V^T L_t^T L_t V R_t)^{-1/2}.$$

Now we observe that for any left scaling L, there is a naturally induced right scaling $R_{jj} := 1/\|Lv_j\|_2$ which fixes the equal-norm condition. Indeed this is equivalent to the second step of Flip-Flop above. Therefore we can keep this step implicit and update the estimator as

$$\overline{\Sigma}_{t+1} \leftarrow \sum_{j \in [n]} \frac{\overline{\Sigma}_t^{-1/2} v_j v_j^T \overline{\Sigma}_t^{-1/2}}{v_j^T \overline{\Sigma}_t^{-1} v_j}.$$

The main algorithmic result of Franks and Moitra (2020) uses the perspective of *geodesic convexity* to analyze Tyler's iterative procedure. In particular, they show that the quantum expansion condition of Kwok et al. (2021) is intimately tied to geodesic *strong convexity*.

The result below follows from the arguments in Franks and Moitra (2020):

Theorem 25 (Section 4 of Franks and Moitra (2020)) *Consider input $x_1, ..., x_n \sim E(\Sigma, u)$ with Tyler's M-Estimator $\hat{\Sigma}$. Let $U := \{u_1, ..., u_n\}$ be the 'normalized' frame, $u_j := \frac{\hat{\Sigma}^{-1/2} x_j}{\|\hat{\Sigma}^{-1/2} x_j\|_2}$. If U is*

a $(1-\Omega(1))$*-quantum expander according to Definition 11, then Tyler's iterative procedure satisfies*

$$\|I_d - \hat{\Sigma}^{1/2}\overline{\Sigma}_{(T)}^{-1}\hat{\Sigma}^{-1/2}\|_F \leq \delta$$

within iteration $T \lesssim |\log\det(\hat{\Sigma})| + d + \log(1/\delta)$.

Proof [Proof Sketch] For input X, the inverse of Tyler's M-estimator $\hat{\Sigma}^{-1}$ is the optimizer of the following capacity function

$$f_X(Z) := \frac{d}{n}\sum_{j\in[n]} \log\langle v_j, Zv_j\rangle - \log\det(Z).$$

By Theorem 4.1 in Franks and Moitra (2020), if U is a $(1-\Omega(1))$-quantum expander, then f_X is $\Omega(1)$ geodesically strongly convex on a large neighborhood of the optimizer $\hat{\Sigma}^{-1}$. Lemma 4.2 in Franks and Moitra (2020) shows that Tyler's iterative procedure is a descent method for f_X which starts at $\overline{\Sigma}_{(0)} := I_d$ and reaches this strongly convex neighborhood within $O(|\log\det(\hat{\Sigma})| + d)$ iterations. From this point, Lemma 4.3 in Franks and Moitra (2020) shows linear convergence to the optimum via geodesic strong convexity. ■

The key technical contribution in Theorem 2.9 of Franks and Moitra (2020) is to show random unit vector frames are quantum expanders above sample threshold $n \gtrsim d\log d$. In Theorem 15, we showed that random frames satisfy the ∞-expansion condition. Our refined result in Theorem 27 in fact shows that the frame scaling solution also satisfies ∞-expansion. In Appendix C we show that ∞-expansion implies quantum expansion.

Theorem 26 *For doubly balanced frame* $V \in \mathbb{R}^{d\times n}$*, if* V *satisfies* $(1-\lambda)$*-*∞*-expansion, then it is also a* $(1-\Omega(\lambda^2))$*-quantum expander.*

We can now combine this result with the arguments of Franks and Moitra (2020) to prove linear convergence of Tyler's iterative procedure at the optimal sample threshold.

Proof [Proof of Theorem 2] Let $X \in \mathbb{R}^{d\times n}$ be generated according to $E(\Sigma, u)$. By Lemma 10, frame $V \in \mathbb{R}^{d\times n}$ has columns $v_j := \frac{\Sigma^{-1/2}x_j}{\|\Sigma^{-1/2}x_j\|_2}$ distributed according to $E(I_d, 1)$. Similarly, by Lemma 9, scalings $L := \hat{\Sigma}^{-1/2}, R_{jj} := 1/\|\hat{\Sigma}^{-1/2}x_j\|_2$ gives frame scaling solution $U := LVR$ for input V. Therefore, in order to apply Theorem 25, we need to show that the frame scaling solution for random unit vector input satisfies quantum expansion.

Following the proof of Theorem 1, by Theorem 16 and Theorem 15, we have that V is ε-doubly balanced and satisfies $(1-\lambda)$-∞-expansion for $\lambda \geq \Omega(1)$. Therefore, by Theorem 27(3), the frame scaling solution U satisfies $(1-\Omega(1))$-∞-expansion. By Theorem 26, this implies U is also a $(1-\Omega(1))$-quantum expander. Also note $|\log\det(\hat{\Sigma})| \approx |\log\det(\Sigma)|$ by the correspondence to the scaling solution (Lemma 9) and Theorem 27(2). The iteration bound follows by Theorem 25. ■

7. Conclusion

In this work, we showed optimal sample complexity bounds for Tyler's M-Estimator for the shape matrix of Elliptical distributions. These bounds are tight (up to constant factors), as they match known lower bounds for the special case of Gaussian covariance estimation. We also recover the algorithmic analysis of Franks and Moitra (2020), showing linear convergence of Tyler's iterative procedure with optimal sample threshold $n \gtrsim d$.

Our proof follows the connection to frame scaling, as described in Franks and Moitra (2020). We identify a new 'pseudorandom' condition for frames, ∞-expansion, and show it is satisfied by random frames with high probability. Next, we use the ∞-expansion condition to prove a new result for frame scaling. We believe this result could be of independent interest. We highlight two problems which could potentially benefit from similar techniques.

In upcoming work, we use techniques similar to this paper to show optimal distance bounds for the Paulsen problem. This was a major problem in frame theory Cahill and Casazza (2013), open for nearly two decades despite significant attention. Formally, given ε-doubly balanced frame $U \in \mathbb{R}^{d\times n}$, we want to find a nearby frame $V \in \mathbb{R}^{d\times n}$ that is doubly balanced. The main conjecture was to show a distance bound that is independent of n. In Kwok et al. (2017), the authors were able to affirm this conjecture, showing a distance bound of $\text{poly}(d) \cdot \varepsilon$ using a smoothed analysis approach to frame scaling. Indeed, their analysis involved a pseudorandom property similar to our results in Section 4. Hamilton and Moitra (2019) improved the distance bound to $d \cdot \varepsilon$, which is a factor d greater than the known lower bound ε. By sharpening the techniques in this paper, we are able to close this gap and show optimal distance ε for the Paulsen problem.

The second problem is the matrix and tensor normal model in statistics. We are given tensor data $X \in \mathbb{R}^{d_1} \otimes ... \otimes \mathbb{R}^{d_k}$ from a Gaussian distribution $N(0, \Sigma)$, with the promise that the covariance matrix has a matching form $\Sigma = \Sigma_1 \otimes ... \otimes \Sigma_k$. In Franks et al. (2021), the authors show strong sample complexity bounds for this problem. We believe that techniques related to ∞-expansion can be used to give optimal sample complexity and error bounds for the tensor normal model.

References

Jamil Arbas, Hassan Ashtiani, and Christopher Liaw. Polynomial time and private learning of unbounded gaussian mixture models. In *International Conference on Machine Learning*, pages 1018–1040. PMLR, 2023.

Peter Bürgisser, Cole Franks, Ankit Garg, Rafael Oliveira, Michael Walter, and Avi Wigderson. Towards a theory of non-commutative optimization: geodesic 1st and 2nd order methods for moment maps and polytopes. In *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 845–861. IEEE, 2019.

Jameson Cahill and Peter Casazza. The Paulsen problem in operator theory. *Operators and Matrices*, 2013.

Peter G. Casazza and Gitta Kutyniok, editors. *Finite Frames: Theory and Applications*. Birkhauser Basel, 2013.

Ilias Diakonikolas, Daniel Kane, and Christos Tzamos. Forster decomposition and learning halfspaces with noise. *Advances in Neural Information Processing Systems*, 34:7732–7744, 2021.

Jürgen Förster. A linear lower bound on the unbounded error probabilistic communication complexity. *Journal of Computer and System Sciences*, 65, 2002.

Cole Franks, Rafael Oliveira, Akshay Ramachandran, and Michael Walter. Near optimal sample complexity for matrix and tensor normal models via geodesic convexity. *arXiv preprint arXiv:2110.07583*, 2021.

William Cole Franks and Ankur Moitra. Rigorous guarantees for tyler's m-estimator via quantum expansion. In *Conference on Learning Theory*, pages 1601–1632. PMLR, 2020.

Ankit Garg, Leonid Gurvits, Rafael Mendes de Oliveira, and Avi Wigderson. A deterministic polynomial time algorithm for non-commutative rational identity testing. *arXiv preprint arXiv:1511.03730*, 2015.

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

Leonid Gurvits. Classical complexity and quantum entanglement. *Journal of Computer and System Sciences*, 2004.

Linus Hamilton and Ankur Moitra. The Paulsen problem made simple. In *Innovations in Theoretical Computer Science (ITCS)*, 2019.

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.

T.C. Kwok, L.C. Lau, Y.T. Lee, and A. Ramachandran. The Paulsen problem, continuous operator scaling, and smoothed analysis. In *Symposium on Theory of Computing (STOC)*. ACM, 2017.

T.C. Kwok, L.C. Lau, Y.T. Lee, and A. Ramachandran. Spectral analysis of matrix scaling and operator scaling. *SIAM Journal of Computing*, 50, 2021.

Paul Milgrom and Ilya Segal. Envelope theorems for arbitrary choice sets. *Econometrica*, 70(2): 583–601, 2002.

Gilles Pisier. *The volume of convex bodies and Banach space geometry*. Cambridge University Press, 1989.

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

Terence Tao. *Topics in random matrix theory*. American Mathematical Society, 2012.

David E Tyler. A distribution-free m-estimator of multivariate scatter. *The annals of Statistics*, pages 234–251, 1987.

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

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

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

Appendix A. Proof of Theorem 14

Here we prove Theorem 14 using ∞-expansion. We show the following more precise result:

Theorem 27 *Consider frame $V \in \mathbb{R}^{d\times n}$ that is ε-doubly balanced and satisfies $(1-\lambda)$-∞-expander with $\lambda^2 \gtrsim \varepsilon$. If (V_t, L_t, R_t) is the dynamical system given in Definition 17 and Lemma 18,*

1. *$(L_\infty, R_\infty) := \lim_{t\to\infty}(L_t, R_t)$ is the frame scaling solution for V according to Definition 8;*
2. *The scaling solution satisfies $\|L_\infty - I_d\|_{\mathrm{op}}, \|R_\infty - I_n\|_{\mathrm{op}} \lesssim \frac{\varepsilon}{\lambda}$;*
3. *$V_\infty := \lim_{t\to\infty} V_t = L_\infty V R_\infty$ satisfies $(1-\lambda/2)$-∞-expansion, and has size $s(V_\infty) \geq s(V)(1-O(\frac{\varepsilon^2}{\lambda}))$.*

Our plan is to show exponential convergence of the error matrices through gradient flow. We begin by explicitly calculating how the error matrices change in the dynamical system. The following expression is implicit in the proof of Proposition 3.2 in Kwok et al. (2021).

Lemma 28 *For frame $V \in \mathbb{R}^{d\times n}$, let V_t be the solution to the dynamical system as given in Definition 17. Then for any fixed $x \in \mathbb{R}^d$ and $j \in [n]$,*

$$-\partial_{t=0}\langle xx^T, VV^T\rangle = 2\langle xx^T, EVV^T + VFV^T\rangle; \quad -\partial_{t=0}\|v_j\|_2^2 = 2(F_{jj}\|v_j\|_2^2 + \langle E, v_j v_j^T\rangle).$$

Proof For the first expression, we compute

$$-\partial_{t=0}\langle xx^T, V_tV_t^T\rangle = \langle xx^T, (EV+VF)V^T + V(EV+VF)^T\rangle = 2\langle xx^T, EVV^T + VFV^T\rangle,$$

where we used the product rule using Definition 17 for the derivative, and in the last step we used that $E = d\cdot VV^T - s(V)I_d$ commutes with VV^T. For the second, a similar calculation gives

$$-\partial_{t=0}V_t^TV_t = (EV+VF)^TV + V^T(EV+VF) = 2V^TEV + FV^TV + V^TVF.$$

The second expression follows by considering the j-th diagonal entry of the above expression. ■

Recall $E = dVV^T - sI_d$ and $F = \mathrm{diag}(nV^TV - sI_n)$ by Definition 5. We want to show that these errors decrease under ∞-expansion. For illustration, let $x \in S^{d-1}$ be the top eigenvector $Ex = \|E\|_{\mathrm{op}}x$, and note this implies $d\cdot VV^Tx = (s+\|E\|_{\mathrm{op}})x$. Then the first term above is

$$d\cdot\langle xx^T, EVV^T\rangle = \|E\|_{\mathrm{op}}\langle xx^T, dVV^T\rangle = \|E\|_{\mathrm{op}}(s+\|E\|_{\mathrm{op}}) \approx s\|E\|_{\mathrm{op}},$$

assuming $\|E\|_{\mathrm{op}} \ll s$ for simplicity. This tells us that for top eigenvector x, the E term in Lemma 28 decreases the quadratic form $\langle xx^T, VV^T\rangle$. Intuitively, this pushes the error towards balanced.

The other term involves VFV^T for $\mathrm{Tr}[F] = 0$. We can bound this using ∞-expansion:

$$|d\langle xx^T, VFV^T\rangle| \leq \|x\|_2^2\|dVFV^T\|_{\mathrm{op}} \leq s(1-\lambda)\|F\|_{\mathrm{op}},$$

where the first step was by definition of operator norm, and in the second step we applied Definition 13 to F. This implies that this second term contributes only a small competing force for the change in Lemma 28, and so intuitively the error should decrease quickly. We formalize this intuition below. We will also need the following lemma to control how the size changes.

Lemma 29 (Lemma 2.17 in Kwok et al. (2021)) *The change in size of V for the dynamical system in Definition 17 is*

$$-\partial_t s(V_t) = 2\Delta(V) = \frac{2}{d}\|E(V)\|_F^2 + \frac{2}{n}\|F(V)\|_F^2.$$

Lemma 30 *Given frame $V \in \mathbb{R}^{d\times n}$ with V_t the solution to the dynamical system in Definition 17, if V satisfies $(1-\lambda)$-∞-expansion at time $t = 0$, then*

$$-\frac{1}{2}\partial_{t=0}\|E(V_t)\|_{\text{op}} \geq s(V)(\|E\|_{\text{op}} - (1-\lambda)\|F\|_{\text{op}}) - \|(E,F)\|_{\text{op}}^2;$$
$$-\frac{1}{2}\partial_{t=0}\|F(V_t)\|_{\text{op}} \geq s(V)(\|F\|_{\text{op}} - \|E\|_{\text{op}}) - \|(E,F)\|_{\text{op}}(\|E\|_{\text{op}} + \|F\|_{\text{op}}).$$

Proof We begin with the E term. First assume $\|E\|_{\text{op}} = \langle xx^T, E\rangle$ for $x \in S^{d-1}$, so we compute

$$-\partial_{t=0}\|E(V_t)\|_{\text{op}} = -\partial_{t=0}\langle xx^T, dV_tV_t^T - s(V_t)I_d\rangle = 2d\langle xx^T, EVV^T + VFV^T\rangle - 2\Delta(V),$$

where in the first step we used the generalized Envelope Theorem of Milgrom and Segal (2002) to write the derivative of $\|E\|_{\text{op}} = \max_{x\in S^{d-1}}|\langle xx^T, E\rangle|$, and in the final step we used Lemma 28 for the first term and Lemma 29 for the change in size. We continue

$$\begin{aligned}-\frac{1}{2}\partial_{t=0}\|E(V_t)\|_{\text{op}} &\geq \|E\|_{\text{op}}(s(V) + \|E\|_{\text{op}}) - d\|VFV^T\|_{\text{op}} - (\|E\|_{\text{op}}^2 + \|F\|_{\text{op}}^2)\\ &\geq s(V)\|E\|_{\text{op}} - s(V)(1-\lambda)\|F\|_{\text{op}}^2 - \|F\|_{\text{op}}^2,\end{aligned}$$

where we used that x is a unit eigenvector of E and dVV^T, with eigenvalue $\|E\|_{\text{op}}$ and $s(V) + \|E\|_{\text{op}}$ respectively, and we bounded $\Delta(V) \leq \|E\|_{\text{op}}^2 + \|F\|_{\text{op}}^2$ using Lemma 7.

For the other case $\|E\|_{\text{op}} = -\langle xx^T, E\rangle$ with $x \in S^{d-1}$, we can use the same argument to bound

$$\begin{aligned}-\frac{1}{2}\partial_{t=0}\|E(V_t)\|_{\text{op}} &= \frac{1}{2}\partial_{t=0}\langle xx^T, dV_tV_t^T - s(V_t)I_d\rangle = -d\langle xx^T, EVV^T + VFV^T\rangle + \Delta(V)\\ &\geq \|E\|_{\text{op}}(s(V) - \|E\|_{\text{op}}) - d\|VFV^T\|_{\text{op}} \geq s(V)(\|E\|_{\text{op}} - (1-\lambda)\|F\|_{\text{op}}) - \|E\|_{\text{op}}^2,\end{aligned}$$

where we used that x is an eigenvector of dVV^T with eigenvalue $s(V) - \|E\|_{\text{op}}$, and $\Delta(V) \geq 0$. The result follows by combining the two cases.

For the F-bound, we follow the same steps as above without applying ∞-expansion. So first, if $\|F\|_{\text{op}} = n\|v_j\|_2^2 - s(V)$ for some $j \in [n]$, then

$$-\partial_{t=0}\|F(V_t)\|_{\text{op}} = -\partial_{t=0}(n\|V_te_j\|_2^2 - s(V_t)) = 2n(F_{jj}\|v_j\|_2^2 + \langle v_j, Ev_j\rangle) - 2\Delta(V),$$

where the first step is again by the Envelope Theorem of Milgrom and Segal (2002) applied to $\|F\|_{\text{op}} = \max_{j\in[n]}|F_{jj}|$ for diagonal F, and in the final step we used Lemma 28 for the first term and Lemma 29 for the change in size. We continue

$$\begin{aligned}-\frac{1}{2}\partial_{t=0}\|F(V_t)\|_{\text{op}} &\geq n\|v_j\|_2^2(\|F\|_{\text{op}} - \|E\|_{\text{op}}) - (\|E\|_{\text{op}}^2 + \|F\|_{\text{op}}^2)\\ &\geq s(V)(\|F\|_{\text{op}} - \|E\|_{\text{op}}) - \|E\|_{\text{op}}(\|E\|_{\text{op}} + \|F\|_{\text{op}}),\end{aligned}$$

where we used $n\|v_j\|_2^2 = s(V) + \|F\|_{\mathrm{op}}$ by assumption on j, bounded $\langle v_j, Ev_j\rangle \le \|E\|_{\mathrm{op}}\|v_j\|_2^2$, and used Lemma 7 to bound $\Delta(V)$.

For the other case, $\|F\|_{\mathrm{op}} = -(n\|v_j\|_2^2 - s(V))$, we can use the same argument to show

$$\begin{aligned}-\frac{1}{2}\partial_{t=0}\|F(V_t)\|_{\mathrm{op}} = \frac{1}{2}\partial_{t=0}(n\|V_t e_j\|_2^2 - s(V_t)) &= -n(F_{jj}\|v_j\|_2^2 + \langle v_j, Ev_j\rangle) + \Delta(V)\\ &\ge n\|v_j\|_2^2(\|F\|_{\mathrm{op}} - \|E\|_{\mathrm{op}}) \ge (s(V) - \|F\|_{\mathrm{op}})(\|F\|_{\mathrm{op}} - \|E\|_{\mathrm{op}}).\end{aligned}$$

By Definition 6 $\|(E,F)\|_{\mathrm{op}} = \max\{\|E\|_{\mathrm{op}}, \|F\|_{\mathrm{op}}\}$, so the result follows by combining both cases. ■

The above lemma shows that the convergence of E and F depend on the relative sizes of $\|E\|_{\mathrm{op}}, \|F\|_{\mathrm{op}}$. By some case analysis, we can show one of the errors always decreases.

Lemma 31 *Let frame $V \in \mathbb{R}^{d\times n}$ be ε-doubly balanced and satisfy $(1-\lambda)$-∞-expansion for $1 \ge \lambda \ge \varepsilon$. Then for any $\delta \in [0,1)$:*

$$\begin{aligned}(1+\delta)\|E\|_{\mathrm{op}} \ge \|F\|_{\mathrm{op}} \quad &\Longrightarrow \quad -\partial_{t=0}\log\|E(V_t)\|_{\mathrm{op}} \ge 2s(V)(\lambda - \delta - \varepsilon);\\ \|E\|_{\mathrm{op}} \le (1-\delta)\|F\|_{\mathrm{op}} \quad &\Longrightarrow \quad -\partial_{t=0}\log\|F(V_t)\|_{\mathrm{op}} \ge 2s(V)(\delta - 2\varepsilon).\end{aligned}$$

Proof In the first case $(1+\delta)\|E\|_{\mathrm{op}} \ge \|F\|_{\mathrm{op}}$, we can bound the change in E as

$$\begin{aligned}-\frac{1}{2}\partial_{t=0}\log\|E(V_t)\|_{\mathrm{op}} = \frac{-\partial_{t=0}\|E(V_t)\|_{\mathrm{op}}}{2\|E(V)\|_{\mathrm{op}}} &\ge s(V)\left(1 - (1-\lambda)\frac{\|F\|_{\mathrm{op}}}{\|E\|_{\mathrm{op}}}\right) - \frac{\|(E,F)\|_{\mathrm{op}}^2}{\|E\|_{\mathrm{op}}}\\ &\ge s(V)(1-(1-\lambda)(1+\delta)) - s(V)\varepsilon(1+\delta) \ge s(V)(\lambda - \delta - \varepsilon),\end{aligned}$$

where we used the E-bound from Lemma 30, for the third step we used the case assumption to bound $\|F\|_{\mathrm{op}} \le \|(E,F)\|_{\mathrm{op}} \le (1+\delta)\|E\|_{\mathrm{op}}$, and the ε-doubly balanced condition to bound $\|(E,F)\|_{\mathrm{op}} \le s(V)\varepsilon$ according to Definition 6, and in the last step we used $1 \ge \lambda \ge \varepsilon$.

Similarly, in the other case $\|E\|_{\mathrm{op}} \le (1-\delta)\|F\|_{\mathrm{op}}$, we can bound F as

$$\begin{aligned}-\frac{1}{2}\partial_{t=0}\log\|F(V_t)\|_{\mathrm{op}} = \frac{-\partial_{t=0}\|F(V_t)\|_{\mathrm{op}}}{2\|F\|_{\mathrm{op}}} &\ge s(V)\left(1 - \frac{\|E\|_{\mathrm{op}}}{\|F\|_{\mathrm{op}}}\right) - \frac{\|(E,F)\|_{\mathrm{op}}}{\|F\|_{\mathrm{op}}}(\|E\|_{\mathrm{op}} + \|F\|_{\mathrm{op}})\\ &\ge s(V)(1-(1-\delta)) - s(V)\varepsilon(1+(1-\delta)) \ge s(V)(\delta - \varepsilon(2-\delta)),\end{aligned}$$

where we used the F-bound from Lemma 30, and for the third step we used the case assumption $\|E\|_{\mathrm{op}} \le (1-\delta)\|F\|_{\mathrm{op}}$ and the ε-doubly balanced condition. ■

Choosing δ appropriately and combining this analysis with Lemma 20 gives our result.

Proof [Proof of Theorem 27] We assume $s(V) = 1$ since the doubly balanced and ∞-expansion condition are homogenous. We require the following technical assumptions:

(a) $1 \ge s(V_t) \ge 1 - \varepsilon$;

(b) V_t is $\frac{5}{3}\varepsilon$-doubly balanced.

(c) V_t is $(1-\lambda/2)$-∞-expander

Note that all three conditions are strictly satisfied at time $t = 0$. Let T be the first time some assumption fails. Our plan is to show exponential convergence of the error

$$\forall t \in [0, T]: \qquad \|(E_t, F_t)\|_{\text{op}} \lesssim \varepsilon e^{-\Omega(\lambda t)}.$$

We use this convergence to show $T = \infty$, and the theorem follows.

For $\delta \in [0, 1)$ chosen later, consider potential function

$$h(t) := \max\{(1+\delta)\|E_t\|_{\text{op}}, \|F_t\|_{\text{op}}\},$$

where we use shorthand $E_t := E(V_t), F_t := F(V_t)$. For a given time $t \in [0, T]$, if $(1+\delta)\|E_t\|_{\text{op}} \geq \|F_t\|_{\text{op}}$, then we can apply the first line of Lemma 31:

$$-\partial_t \log h(t) = -\frac{\partial_t \|E(V_t)\|_{\text{op}}}{\|E(V_t)\|_{\text{op}}} \geq 2s(V_t)\left(\frac{\lambda}{2} - \delta - \frac{5}{3}\varepsilon\right),$$

where in the last step we used assumption (b) and (c) that V_t is $\frac{5}{3}\varepsilon$-doubly balanced and satisfies $(1 - \lambda/2)$-∞-expansion conditions with $1 \geq \frac{\lambda}{2} \geq \frac{5}{3}\varepsilon$.

Otherwise, we have $(1+\delta)\|E\|_{\text{op}} \leq \|F\|_{\text{op}} \implies \frac{\|E\|_{\text{op}}}{\|F\|_{\text{op}}} \leq (1+\delta)^{-1} = 1 - \frac{\delta}{1+\delta}$. Therefore we can apply the second line of Lemma 31:

$$-\partial_t \log h(t) = -\frac{\partial_t \|F(V_t)\|_{\text{op}}}{\|F(V_t)\|_{\text{op}}} \geq 2s(V_t)\left(\frac{\delta}{1+\delta} - 2\frac{5}{3}\varepsilon\right),$$

where again we used assumption (c) that V_t is $\frac{5}{3}\varepsilon$-doubly balanced.

Choosing $\delta = \lambda/4 \leq 1/4$ to balance the two cases, we can bound the potential function

$$-\partial_t \log h(t) \geq 2s(V_t) \min\left\{\frac{\lambda}{2} - \frac{\lambda}{4} - 2\varepsilon, \frac{\lambda}{4(1+\delta)} - \frac{10}{3}\varepsilon\right\} \geq \frac{\lambda}{3},$$

where we used assumption (a) $s(V_t) \geq 1 - \frac{\varepsilon}{4}$, as well as $\delta \leq 1/4$, and $\lambda \gtrsim \varepsilon$ large enough.

Therefore, we have exponential convergence for the error:

$$\log \max\{(1+\delta)\|E_t\|_{\text{op}}, \|F_t\|_{\text{op}}\} = \log h(t) = \log h(0) + \int_{\tilde{t}=0}^{t} \partial_{\tilde{t}} \log h(\tilde{t}) \leq \log(1+\delta)\varepsilon - \lambda t/3,$$

where the first step was by definition of $h(t)$, the second was by the fundamental theorem of calculus, and in the final step we bounded the first term $h(0) \leq \varepsilon(1+\delta)$ as V is ε-doubly balanced with $s(V) = 1$, and for the integral we used that $\partial_t \log h(t) \leq -\lambda/3$ as shown above. Rearranging gives

$$\|E_t\|_{\text{op}} \leq \varepsilon e^{-\lambda t/3}, \quad \|F_t\|_{\text{op}} \leq (1+\delta)\varepsilon e^{-\lambda t/3} \leq \frac{5}{4}\varepsilon e^{-\lambda t/3}. \tag{2}$$

We can now use this exponential convergence to prove the conclusions of the theorem. We begin by bounding the scaling up to time T:

$$\|L_T - I_d\|_{\text{op}} \leq \exp\left(\int_0^T \|E_t\|_{\text{op}}\right) - 1 \leq \exp\left(\varepsilon \int_0^T e^{-\lambda t/3}\right) - 1 \leq \exp(3\varepsilon/\lambda) - 1 \leq \frac{4\varepsilon}{\lambda}, \tag{3}$$

where the first step was by Lemma 18, in the second step we plugged in the convergence bound of Equation (2), and the final steps were by Taylor approximation $e^x = 1 + O(x)$ along with the assumption $\lambda^2 \gtrsim \varepsilon$ large enough. The same argument gives the bound on right scaling

$$\|R_T - I_n\|_{\text{op}} \leq \exp\left(\int_0^T \|F_t\|_{\text{op}}\right) - 1 \leq \frac{5\varepsilon}{\lambda}.$$

By Lemma 20, this implies V_T satisfies $(1-\lambda')$-∞-expansion with

$$\lambda' \geq 1 - \frac{s(V)(1-\lambda) + O(\varepsilon/\lambda)}{s(V_T)} \geq \lambda - O(\lambda + \varepsilon) > \lambda/2, \tag{4}$$

where we used assumptions (a) $s(V_T) \geq 1 - \varepsilon$ in the second step, and $\lambda^2 \gtrsim \varepsilon$ in the final step.

We can also bound the change in size up to time T:

$$s(V) - s(V_T) = -\int_0^T \partial_t s(V_t) = 2\int_0^T \Delta(V_t) \leq 2\int_0^T (\|E_t\|_{\text{op}}^2 + \|F_t\|_{\text{op}}^2) \tag{5}$$

$$\leq 6\varepsilon^2 \int_0^T e^{-2\lambda t/3} < \frac{6\varepsilon^2}{2\lambda/3} < \varepsilon, \tag{6}$$

where the first step was by fundamental theorem of calculus, in the second step we used Lemma 29 for $\partial_t s$, in the third we used Lemma 7 to bound Δ, in the fourth we applied convergence from Equation (2), and in the final inequality we used our assumption that $\lambda \gtrsim \varepsilon$ large enough.

Now Assume for contradiction that $T < \infty$, so one of the assumptions (a,b,c) fails at time T. Assumption (a) $1 \geq s(V_T) > 1 - \varepsilon$ is strictly satisfied by Equation (6). We can also bound

$$\frac{\|(E_T, F_T)\|_{\text{op}}}{s(V_T)} < \frac{5\varepsilon/4}{1 - 9\varepsilon^2/\lambda} < \frac{5}{3}\varepsilon,$$

where we used Equation (2) to upper bound the error and Equation (6) to lower bound the size. By Definition 6, this implies V_T is ε'-doubly balanced for $\varepsilon' < \frac{5}{3}\varepsilon$, so assumption (b) is strictly satisfied. Finally, Equation (3) shows that V_T satisfies $(1-\lambda')$-∞-expansion for $\lambda' > \lambda/2$, so assumption (c) is also strictly satisfied. By continuity, (a,b,c) must still be satisfied for some $T' > T$. But this contradicts maximality of T, so the above analysis holds for all time.

We can now show the conclusions of the theorem. For item (1), note

$$\lim_{t\to\infty} \|(E_t, F_t)\|_{\text{op}} \lesssim \lim_{t\to\infty} \varepsilon e^{-\lambda t/3} = 0,$$

i.e. V_∞ is doubly balanced. By Lemma 18, this implies the scaling limits (L_∞, R_∞) exist and are the frame scaling solution for V according to Definition 8. Item (2) follows from the calculation in Equation (3) applied at $T = \infty$, and similarly item (3) follows from Equation (4), (6). ∎

Appendix B. Proof of ∞-Expansion for Random Frames

In this section, we prove the technical details required for our ∞-expansion result in Theorem 15.

We first restate and prove Lemma 22, showing $\beta = 1/2$ pseudorandomness and ∞-expansion are essentially equivalent. We will consider pseudorandomness for other β later in this section.

Lemma 32 *Let $V \in \mathbb{R}^{d\times n}$ be an ε-doubly balanced frame. If V $(\alpha_{\min}, \alpha_{\max}, \frac{1}{2})$-pseudorandom, then V satisfies $(1-\lambda)$-∞-expansion with*

$$s(V)(1-\lambda) \le \min\{s(V)(1+\varepsilon) - \alpha_{\min} \quad , \quad \alpha_{\max} - s(V)(1-\varepsilon)\}.$$

Conversely, if V satisfies $(1-\lambda)$-∞-expansion, then it is $(\alpha_{\min}, \alpha_{\max}, \frac{1}{2})$-pseudorandom with

$$s(V)(\lambda-\varepsilon) \le \alpha_{\min} \le \alpha_{\max} \le s(V)(2-(\lambda-\varepsilon)).$$

Remark 33 *We always assume βn is an integer for simplicity. Technically, the condition for non-integer βn should involve a fractional coordinate. For example if $\beta n = k + z$ where $z \in (0,1)$ then β-pseudorandomness should be the condition*

$$\beta\frac{\alpha_{\min}}{d} I_d \preceq V_B V_B^T + z v_j v_j^T \preceq \beta\frac{\alpha_{\max}}{d} I_d$$

for all $|B| = k, j \notin B$. For simplicity we will focus on the integer case as the fractional definition only differs by lower order terms.

Proof [Proof of Lemma 32] It can be shown by majorization that the vertices of the polytope

$$P := 1_n^{\perp} \cap B_\infty = \{y \in \mathbb{R}^n \mid \langle y, 1_n\rangle = 0, \|y\|_\infty \le 1\}$$

are of the form $1_A - 1_B$ for disjoint sets $|A| = |B| = \lfloor n/2 \rfloor$. For simplicity, we assume n is even so these vertices can be rewritten as $1_n - 21_B = 21_A - 1_n$. This implies that for any $a \in \mathbb{R}^n$,

$$\sup_{y\in P}\langle a, y\rangle = 2\max_A \langle a, 1_A\rangle - \langle a, 1_n\rangle = \langle a, 1_n\rangle - 2\min_B \langle a, 1_B\rangle \tag{7}$$

where $|A| = |B| = n/2$. We relate ∞-expansion and pseudorandomness by applying the above for $a_j := \langle x, v_j\rangle^2$ with $x \in S^{d-1}$.

In particular, let $x \in S^{d-1}, y \in P$ satisfy $\frac{s(V)(1-\lambda)}{d} = |\langle xx^T, VYV^T\rangle|$. If V is $(\alpha_{\min}, \alpha_{\max}, \frac{1}{2})$-pseudorandom according to Definition 21 and ε-doubly balanced according to Definition 6, then

$$\frac{s(V)(1-\lambda)}{d} \le 2\max_A \langle xx^T, V_A V_A^T\rangle - \langle xx^T, VV^T\rangle \le \frac{\alpha_{\max}}{d} - \frac{s(V)(1-\varepsilon)}{d},$$

where we used Equation (7) in the first step, and in the last step we upper bounded $V_A V_A^T$ using $\alpha_{\max}$ and the other term using the ε-doubly balanced condition. By the same argument we get

$$\frac{s(V)(1-\lambda)}{d} \le \langle xx^T, VV^T\rangle - 2\min_B \langle xx^T, V_B V_B^T\rangle \le \frac{s(V)(1+\varepsilon)}{d} - \frac{\alpha_{\min}}{d},$$

where we instead lower bounded $V_B V_B^T$ using $\alpha_{\min}$.

Conversely, let $x \in S^{d-1}, |B| = n/2$ satisfy $2\langle xx^T, V_B V_B^T\rangle = \frac{\alpha_{\max}}{d}$. If V satisfies $(1-\lambda)$-∞-expansion, then for $Y := \mathrm{diag}(21_B - 1_n)$

$$\frac{\alpha_{\max}}{d} = \langle xx^T, (2V_B V_B^T - VV^T) + VV^T\rangle \le \|VYV^T\|_{\mathrm{op}} + \|VV^T\|_{\mathrm{op}} \le \frac{s(V)(1-\lambda)}{d} + \frac{s(V)(1+\varepsilon)}{d},$$

where in the second step we substituted $Y := \mathrm{diag}(21_B - 1_n)$, and in the final step we used ∞-expansion Definition 13 to bound $\|VYV^T\|_{\mathrm{op}}$ and the ε-doubly balanced condition Definition 6 to bound $\|VV^T\|_{\mathrm{op}}$. By the same argument, we can lower bound

$$\frac{\alpha_{\min}}{d} = \min_{|B|=n/2} \langle xx^T, VV^T - (VV^T - 2V_BV_B^T)\rangle \geq \frac{s(V)(1-\varepsilon)}{d} - \frac{s(V)(1-\lambda)}{d}.$$

■

Our plan is to show pseudorandomness for Gaussian frames, and then show that this implies pseudorandomness for random unit vectors. For this, we require that the normalization step does not affect the pseudorandom property too much.

Lemma 34 *For frame G that is $(\alpha_{\min}, \alpha_{\max}, \beta)$-pseudorandom, normalized frame $v_j = \frac{g_j}{\|g_j\|_2}$ has size $s(V) = n$ and is $(\alpha_{\min}(V), 2\beta)$-pseudorandom for*

$$\alpha_{\min}(V) \geq s(V)\frac{\alpha_{\min}}{2\alpha_{\max}}.$$

Proof Consider $B \subseteq [n]$ with $|B| = 2\beta n$. By relabeling, we assume $B = \{1, ..., 2\beta n\}$ and the norms are in increasing order $\|g_1\|_2 \leq ... \leq \|g_{|B|}\|_2$. Letting $S := \{1, ..., \beta n\}$ be the half containing $\|g_j\|_2$ of smaller norm, we can lower bound

$$V_BV_B^T = \sum_{j\in B} \frac{g_jg_j^T}{\|g_j\|_2^2} \succeq \sum_{j\in S} \frac{g_jg_j^T}{\|g_j\|_2^2} \succeq \frac{1}{\|g_{|S|}\|_2^2} G_SG_S^T, \tag{8}$$

where the first step is by definition $v_j := \frac{g_j}{\|g_j\|_2}$, in the second step we consider the smaller subset $S \subseteq B$, and in the final step we use that the norms are in increasing order so $\|g_{j\in S}\|_2 \geq \|g_{|S|}\|_2$. We can lower bound $G_SG_S^T$ using pseudorandomness, so we require an upper bound for $\|g_{|S|}\|_2^2$:

$$\|g_{|S|}\|_2^2 \leq \mathbb{E}_{j\in B-S}\|g_j\|_2^2 = \frac{1}{\beta n}\langle I_d, G_{B-S}G_{B-S}^T\rangle \leq \frac{1}{\beta n}\alpha_{\max}\beta = \frac{\alpha_{\max}}{n},$$

where in the first step we used that the norms are in increasing order so $\|g_{|S|}\|_2 \leq \|g_{j\in B-S}\|_2$, and in the third step we used the pseudorandomness upper bound according to Definition 21 for G with subset $|B - S| = \beta n$. Plugging this into the previous calculation gives

$$V_BV_B^T \succeq \frac{1}{\|g_{|S|}\|_2^2} G_SG_S^T \succeq \frac{n}{\alpha_{\max}}\frac{\alpha_{\min}\beta}{d} I_d = n\frac{\alpha_{\min}}{2\alpha_{\max}}\frac{|B|}{n}\frac{1}{d} I_d, \tag{9}$$

where the first step was from Equation (8), in the second step we used the upper bound $\|g_{|S|}\|_2^2 \leq \frac{\alpha_{\max}}{n}$ and the lower bound $G_SG_S^T \succeq \frac{\alpha_{\min}\beta}{d} I_d$ by pseudorandomness of G for $|S| = \beta n$, and in the final step we used that $|B| = 2\beta n$. Since $B \subseteq [n]$ was arbitrary, and noting $s(V) = n$ by normalization, this verifies $\alpha_{\min}$-pseudorandomness according to Definition 21. ■

In the remainder, we prove Theorem 23 showing Gaussian frames satisfy the pseudorandom property with high probability.

B.1. Preliminaries: Concentration and Approximation

We will use Gaussian concentration and some standard approximation arguments.

Theorem 35 (Corollary 5.35 of Vershynin (2010)) *For $d \le m$, let $G \in \mathbb{R}^{d\times m}$ be a random matrix whose entries are independent standard normal random variables $G_{ij} \sim N(0,1)$ for $i \in [d], j \in [m]$. Then for any $\theta > 0$,*

$$\sqrt{m} - \sqrt{d} - \theta \le \sigma_{\min}(G) \le \sigma_{\max}(G) \le \sqrt{m} + \sqrt{d} + \theta$$

with probability at least $1 - 2e^{-\theta^2/2}$.

$\sigma_{\min}(G) \ge 0$ always, so the lower tail bound becomes trivial for $\theta > \sqrt{m}$. If we require non-trivial guarantees with failure probability better than $\exp(-m/2)$, we can use the following stronger concentration for the lower tail.

Lemma 36 (Fact 4.5.7(3) in Kwok et al. (2017)) *For $g \sim N(0, I_m)$ and any $c \ge 2$,*

$$Pr[\|g\|_2^2 \le e^{-c}m] \le \exp(-cm/4).$$

We also require the following standard approximation arguments.

Lemma 37 (see e.g. Tao (2012)) *For $0 \le \beta \le 1/2$, $\binom{n}{\beta n} \le 2^{\beta n(1+\log_2(1/\beta))}$.*

Lemma 38 (Lemma 4.10 in Pisier (1989)) *$\mathcal{N}$ is called an η-net of S^{d-1} if, for any $x \in S^{d-1}$ there is $y \in \mathcal{N}$ such that $\|y - x\|_2 \le \eta$. For any $\eta > 0$, there is an η-net $\mathcal{N} \subseteq S^{d-1}$ with cardinality*

$$|\mathcal{N}| \le (1 + 2/\eta)^d.$$

Lemma 39 *For $A \in \mathbb{R}^{d\times m}$, if $\mathcal{N}$ is an η-net of S^{d-1}, then*

$$\sigma_{\min}(A) \ge \inf_{x\in\mathcal{N}} \|x^T A\|_2 - \eta\|A\|_{\text{op}}.$$

We omit the proof of the above as it is standard. Note that $\sigma_{\min} \ge 0$ always, so the above bound is only non-trivial when $\eta < \inf_{x\in\mathcal{N}} \|x^T A\|_2/\|A\|_{\text{op}}$.

B.2. Proof of Gaussian Pseudorandomness

In this subsection, we use the Gaussian concentration to prove pseudorandomness.

Theorem 40 *Fix $\beta \in (0, 1/2]$, and let $G \in \mathbb{R}^{d\times n}$ be a Gaussian frame with entries $G_{ij} \sim N(0,1)$ for $n \gtrsim d/\beta$. Then with probability $\ge 1 - \exp(-\beta n)$, G is $(\alpha_{\min}, \alpha_{\max}, \beta)$-pseudorandom according to Definition 21 for*

$$\alpha_{\min} \gtrsim nd \cdot \beta^{O(1)}, \quad \alpha_{\max} \lesssim nd(1 + \log(1/\beta)),$$

We first prove spectral upper and lower bounds for a fixed $B \subseteq [n]$, and then show pseudorandomness by a union bound over all subsets.

Lemma 41 *Let $G \in \mathbb{R}^{d\times m}$ be a random matrix standard Gaussian entries $G_{ij} \sim N(0,1)$. If $m \gtrsim d$ is large enough, for any $c \geq 4$,*

$$\frac{e^{-c}}{9} m \cdot I_d \preceq GG^T \preceq (2+c)m \cdot I_d$$

with probability $\geq 1 - 2\exp(-cm/8)$.

Proof Applying Theorem 35 with $\theta = \sqrt{cm}$ for some $c \geq 4$ chosen later, we have

$$\sigma_{\max}(G) \leq \sqrt{m} + \sqrt{d} + \theta = \sqrt{m}(1 + \sqrt{d/m} + \sqrt{c}) \leq \sqrt{m(2+c)}, \tag{10}$$

with failure probability $\leq \exp(-cm)$, where in the last step we used $m \gtrsim d$ and $c \geq 4$.

Next, we want to apply a net argument to lower bound $\sigma_{\min}$. For a fixed $x \in S^{d-1}$, we can apply Lemma 36 to show

$$Pr[\|x^T G\|_2^2 \leq e^{-c}m] \leq \exp(-cm/4), \tag{11}$$

where we used that $x^T G$ is distributed according to $N(0, I_m)$ by orthogonal invariance of G and the fact that $x \in S^{d-1}$. Now let $\mathcal{N} \subseteq S^{d-1}$ be an η-net with $\eta := \frac{2}{3}\sqrt{e^{-c}/(2+c)}$ for the same c as above. By Lemma 38 we can bound the cardinality as

$$|\mathcal{N}| \leq (1 + 2/\eta)^d = (1 + 3\sqrt{e^c(2+c)})^d \leq e^{cd},$$

where in the last step we use the assumption $c \geq 4$. Applying union bound, we have that the lower bound in Equation (11) holds simultaneously for all elements of $\mathcal{N}$ with failure probability at most

$$|\mathcal{N}| \exp(-cm/4) \leq \exp(cd - cm/4) \leq \exp(-cm/8),$$

where we used the cardinality bound for $|\mathcal{N}|$, and in the last step we used the assumption $m \gtrsim d$.

Finally, we can lower bound $\sigma_{\min}$ by approximation:

$$\sigma_{\min}(G) \geq \min_{x\in\mathcal{N}} \|x^T G\|_2 - \eta\|G\|_{\text{op}} \geq \sqrt{e^{-c}m} - \frac{2}{3}\sqrt{e^{-c}/(2+c)}\sqrt{m(2+c)} = \frac{1}{3}\sqrt{e^{-c}m},$$

where the first step was by Lemma 39, and in the second step we used the lower bound over the net from Equation (11) and the upper bound for $\|G\|_{\text{op}}$ from Equation (10), as well as our choice of $\eta = \frac{2}{3}\sqrt{e^{-c}(2+c)}$. The spectral bounds for GG^T follow from these singular value bounds, and the failure probability can be bounded by a union bound for the upper and lower bounds. ■

We can now prove the pseudorandom property for Gaussian frames.

Proof [Proof of Theorem 40] We want to apply the above to every subset $B \subseteq [n]$ with $|B| = \beta n$. Therefore we need a union bound over $\binom{n}{\beta n} \leq 2^{\beta n(1-\log_2\beta)}$ sets (Lemma 37). Fixing $B \subseteq [n]$ with $|B| = \beta n$ and applying Lemma 41 with $c = 8(2 - \log\beta)$ gives

$$\frac{(\beta/e^2)^8}{9}\frac{\beta nd}{d} \cdot I_d = \frac{e^{-c}}{9}|B| \cdot I_d \preceq G_B G_B^T \preceq (2+c)|B| \cdot I_d = (18 + 8\log(1/\beta))\frac{\beta nd}{d} \cdot I_d.$$

By union bound, we get spectral bounds for all subsets $|B| = \beta n$ with total failure probability

$$\leq 2^{\beta n(1-\log_2\beta)} \cdot 2\exp(-c\beta n/8) \leq \exp(-\beta n),$$

as $c = 8(2 - \log\beta)$. By Definition 21, this implies $(\alpha_{\min}, \alpha_{\max}, \beta)$-pseudorandomness with

$$\alpha_{\min} \gtrsim nd \cdot \beta^8, \qquad \alpha_{\max} \lesssim nd(1 + \log_2(1/\beta)).$$

■

Appendix C. Proof of ∞-expansion implies Quantum Expansion

In this section, we prove Theorem 26 showing ∞-expansion implies quantum expansion. We will use the following simple consequence of the pseudorandom property.

Lemma 42 *If $V \in \mathbb{R}^{d\times n}$ is $(\alpha_{\min}, \alpha_{\max}, \beta)$-pseudorandom according to Definition 21, then for all subspaces $A \subseteq \mathbb{R}^d$ and subsets $B \subseteq [n]$ with $|B| \geq \beta n$,*

$$\alpha_{\min}\frac{\dim(A)}{d}\frac{|B|}{n} \leq \|P_A V_B\|_F^2 = \langle P_A, V_B V_B^T\rangle \leq \alpha_{\max}\frac{\dim(A)}{d}\frac{|B|}{n}$$

Proof This follows by a simple averaging: for $|B| \geq \beta n$,

$$\frac{1}{|B|}V_B V_B^T = \mathbb{E}_T \frac{1}{|T|}V_T V_T^T$$

where $T \subseteq B$ is a uniformly random subset of size $|T| = \beta n$. Using Definition 21 of pseudorandomness for these smaller set and rearranging gives:

$$\alpha_{\min}\frac{|B|}{nd}I_d = \frac{|B|}{\beta n}\frac{\alpha_{\min}\beta}{d}I_d \preceq V_B V_B^T \preceq \frac{|B|}{\beta n}\frac{\alpha_{\max}\beta}{d}I_d = \alpha_{\max}\frac{|B|}{nd}I_d.$$

The bounds for inner product $\langle P_A, V_B V_B^T\rangle$ follow directly from these spectral bounds. ■

We can now show that ∞-expansion implies quantum expansion. Our strategy will be to follow the Cheeger style argument of Franks and Moitra (2020).

Theorem 43 *For doubly balanced frame $V \in \mathbb{R}^{d\times n}$, if V satisfies $(1-\lambda)$-∞-expansion according to Definition 13, then it satisfies $(1-\Omega(\lambda^2))$-quantum expansion according to Definition 11.*

Proof Corollary A.4 in Franks and Moitra (2020) shows that doubly balanced V satisfies $(1-\mathrm{ch}(V)^2)$-quantum expansion for

$$\mathrm{ch}(V) := \min_{A\subseteq\mathbb{R}^d, B\subseteq[n]} \frac{\|(I_d - P_A)V_B\|_F^2 + \|P_A V_{\overline{B}}\|_F^2}{\|P_A V\|_F^2 + \|V_B\|_F^2},$$

where the minimum is over all subspaces $A \subseteq \mathbb{R}^d$ and subsets $B \subseteq [n]$ satisfying $\frac{\dim(A)}{d} + \frac{|B|}{n} \leq 1$.

The result follows if we can show $(1-\lambda)$-∞-expansion implies $\mathrm{ch}(V) \gtrsim \lambda$. Fix subspace $A \subseteq \mathbb{R}^d$ and subset $B \subseteq [n]$, let $a := \frac{\dim(A)}{d}, b := \frac{|B|}{n}$ and assume $a + b \leq 1$. The goal is to show

$$\|(I_d - P_A)V_B\|_F^2 + \|P_A V_{\overline{B}}\|_F^2 \gtrsim s(V)\lambda(a+b). \tag{12}$$

Note that ∞-expansion and quantum expansion are both homogeneous conditions, so we assume $s(V) = 1$ for simplicity. By Lemma 22, $(1-\lambda)$-∞-expansion implies that V is $(\alpha_{\min} \geq \lambda, \frac{1}{2})$-pseudorandom. By Lemma 42, this implies, for all subspaces $A \subseteq \mathbb{R}^d$ and subsets $|B| \geq n/2$,

$$\|P_A V_B\|_F^2 = \langle P_A, V_B V_B^T\rangle \geq \alpha_{\min}\frac{\dim(A)}{d}\frac{|B|}{n} \geq \lambda\frac{\dim(A)}{d}\frac{|B|}{n}. \tag{13}$$

We can use this to lower bound Equation (12) by some case analysis:

- If $b \geq 1/2$: we use the pseudorandom condition for subset B to lower bound

$$\|(I_d - P_A)V_B\|_F^2 \geq \alpha_{\min}(1-a)b \geq \lambda \frac{(1-a)b}{a+b}(a+b) \geq \frac{\lambda}{4}(a+b),$$

  where the first step is by pseudorandomness Equation (13) for $b \geq 1/2$, and in the final step we use the assumptions $a + b \leq 1, b \geq 1/2$ so $1 - a \geq 1/2$ and $b \geq a$.

- If $b \leq 1/2$ and $b \leq 2a$: we use the pseudorandom property for complement subset $\overline{B}$

$$\|P_A V_{\overline{B}}\|_F^2 \geq \alpha_{\min} a(1-b) \geq \lambda \frac{a(1-b)}{a+b}(a+b) \geq \frac{\lambda}{6}(a+b),$$

  where in the last step we used our case assumptions to bound $1 - b \geq 1/2$ and $6a \geq a + b$.

- Finally, if $b \leq 1/2$ and $b \geq 2a$: we use the doubly balanced property to lower bound

$$\|(I_d - P_A)V_B\|_F^2 \geq \|(I_d - P_A)V\|_F^2 - \|V_{\overline{B}}\|_F^2 \geq (1-a) - (1-b) = \frac{b-a}{a+b}(a+b) \geq \frac{a+b}{3},$$

  where in the first step we used $\|(I_d - P_A)V\|_F^2 = \|(I_d - P_A)V_B\|_F^2 + \|(I_d - P_A)V_{\overline{B}}\|_F^2$ and $\|(I_d - P_A)V_{\overline{B}}\|_F \leq \|V_{\overline{B}}\|_F$ since $I_d - P_A$ is a projection, the second step was by the doubly balanced condition, and in the final step we used $b \geq 2a$ so $\frac{b-a}{a+b} \geq \frac{1}{3}$.

Combining the above three cases gives Equation (12) for arbitrary A, B with $a + b \leq 1$. This implies $\mathrm{ch}(V) \geq \frac{\lambda}{6}$ according to the definition, which proves quantum expansion via Corollary A.4 of Franks and Moitra (2020) as stated above. ■