

SIMPLICIAL FACES OF THE SET OF CORRELATION MATRICES

JOEL A. TROPP

ABSTRACT. This paper concerns the facial geometry of the set of $n \times n$ correlation matrices. The main result states that almost every set of r vertices generates a simplicial face, provided that $r \leq \sqrt{cn}$, where c is an absolute constant. This bound is qualitatively sharp because the set of correlation matrices has no simplicial face generated by more than $\sqrt{2n}$ vertices.

1. MOTIVATION

A *correlation matrix* is a positive-semidefinite (psd) matrix whose diagonal entries are identically equal to one. The set of all correlation matrices with a fixed dimension is called the *elliptope*. As we will explain, the elliptope arises naturally in combinatorial optimization as an approximation to the *cut polytope*. Motivated by this application, we may ask how well the elliptope approximates the cut polytope. In particular, it is valuable to understand what faces the elliptope and the cut polytope have in common. The purpose of this paper is to investigate this question. We will demonstrate that the elliptope and the cut polytope share an enormous number of low-dimensional simplicial faces.

1.1. Graphs and Cuts. Let $G := (V, E)$ be an undirected graph with vertex set $V = \{1, \dots, n\}$. To each subset S of vertices, we associate the vector $\mathbf{c}_S \in \mathbb{R}^n$ whose entries are given by

$$(\mathbf{c}_S)_i := \begin{cases} +1, & i \in S; \\ -1, & i \in \bar{S}. \end{cases}$$

We have written $\bar{S} := V \setminus S$ for the set complement. The *Laplacian* of the graph is the $n \times n$ psd matrix

$$\mathbf{L} := \mathbf{L}_G := \frac{1}{4} \sum_{\{i,j\} \in E} (\mathbf{e}_i - \mathbf{e}_j)(\mathbf{e}_i - \mathbf{e}_j)^\dagger \quad (1.1)$$

where the vector $\mathbf{e}_i \in \mathbb{R}^n$ has a one in the i th coordinate and zeros elsewhere. For a subset S of vertices, we can easily evaluate the quadratic form defined by the Laplacian \mathbf{L} at the vector \mathbf{c}_S :

$$\mathbf{c}_S^\dagger \mathbf{L} \mathbf{c}_S = \frac{1}{4} \sum_{\{i,j\} \in E} ((\mathbf{c}_S)_i - (\mathbf{c}_S)_j)^2 = \#\{(i, j) \in S \times \bar{S} : \{i, j\} \in E\}. \quad (1.2)$$

In words, the value $\mathbf{c}_S^\dagger \mathbf{L} \mathbf{c}_S$ of the quadratic form equals the number of edges that connect S and its complement \bar{S} , which is called the *weight* of the graph cut induced by S .

1.2. Combinatorial and Semidefinite Formulations of the Maximum Cut. This discussion suggests that we can use a mathematical program to optimize the weight of a cut. Let us introduce some definitions.

Definition 1.1 (Cuts). Let n be a natural number. An n -dimensional *cut vector* is a member of set $\{\pm 1\}^n$. An $n \times n$ *cut matrix* takes the form $\mathbf{c}\mathbf{c}^\dagger$ where \mathbf{c} is an n -dimensional cut vector. The *cut polytope* \mathcal{C}_n is the convex hull of the $n \times n$ cut matrices:

$$\mathcal{C}_n := \text{conv}\{\mathbf{c}\mathbf{c}^\dagger : \mathbf{c} \in \{\pm 1\}^n\}. \quad (1.3)$$

These objects are sometimes called *signed* cut vectors, matrices, and polytopes.

Date: 5 December 2016. Revised 10 November 2017.

2010 Mathematics Subject Classification. Primary: 52A20, 15B48. Secondary: 52B12, 90C27.

Key words and phrases. Correlation matrix, cut polytope, elliptope, face, matrix concentration, probabilistic method.

Let A be an $n \times n$ real psd matrix, and consider three equivalent mathematical programs:

$$\begin{aligned} & \text{maximize} && \mathbf{x}^\top A \mathbf{x} && \text{subject to} && \mathbf{x} \text{ is a cut vector.} \\ & \text{maximize} && \text{trace}(A\mathbf{X}) && \text{subject to} && \mathbf{X} \text{ is a cut matrix.} \\ & \text{maximize} && \text{trace}(A\mathbf{X}) && \text{subject to} && \mathbf{X} \in \mathcal{C}_n. \end{aligned} \tag{1.4}$$

To see that the first two are equivalent, write $\mathbf{X} = \mathbf{x}\mathbf{x}^\top$ for a cut vector \mathbf{x} and cycle the trace. As for the third, the construction (1.3) implies that the extreme points of the cut polytope \mathcal{C}_n are precisely the $n \times n$ cut matrices. Therefore, the third program attains its optimal value at a cut matrix.

In view of (1.2), we can try to find the maximum weight of a cut in the graph G by solving (1.4) with $A = L_G$, where L_G is the graph Laplacian (1.1). Finding the maximum weight of a cut in a general graph is NP-hard [Kar72], so we cannot accomplish this task with a polynomial-time algorithm unless $P = NP$.

One remedy is to relax (1.4) to reach a tractable computational problem. To do so, notice that each cut matrix is a real psd matrix whose diagonal entries are equal to one. This motivates another definition.

Definition 1.2 (Elliptope). Let n be a natural number. The *elliptope* \mathcal{E}_n is the convex set

$$\mathcal{E}_n := \{\mathbf{X} \in \mathbb{S}_+^n : \text{diag}(\mathbf{X}) = \mathbf{1}\}. \tag{1.5}$$

The set \mathbb{S}_+^n comprises the $n \times n$ real psd matrices, and each entry of the vector $\mathbf{1} \in \mathbb{R}^n$ equals one. The members of the elliptope are called *correlation matrices*.

The elliptope \mathcal{E}_n contains every $n \times n$ cut matrix, so it also contains the cut polytope \mathcal{C}_n . Therefore, we may attempt to approximate the value of (1.4) by means of the semidefinite programming problem

$$\text{maximize} \quad \text{trace}(A\mathbf{X}) \quad \text{subject to} \quad \mathbf{X} \in \mathcal{E}_n. \tag{1.6}$$

The elliptope \mathcal{E}_n is an affine slice of the psd cone \mathbb{S}_+^n , so there are polynomial-time algorithms for completing the optimization (1.6) to a fixed accuracy in the real arithmetic model [BTN01, Lec. 5].

1.3. Analysis of the Semidefinite Relaxation. How well does the tractable formulation (1.6) work? For psd A , a randomized rounding argument [Nes98] shows that the optimal values of (1.4) and (1.6) satisfy

$$\frac{2}{\pi} \cdot \text{val}(1.6) \leq \text{val}(1.4) \leq \text{val}(1.6). \tag{1.7}$$

The constant $2/\pi$ cannot be improved. But it has been observed empirically that the value of (1.6) is usually within a few percent of the value of (1.4). For example, see [DP93, PR95, GW95, MT11].

Results like (1.7) are often attributed to Goemans & Williamson [GW95] or to Nesterov [Nes98], but the provenance is longer. Indeed, the bound (1.7) is equivalent to the “little” Grothendieck theorem [Gro53, Thm. 4]. See the surveys [KN12, Pis12] for a modern introduction to Grothendieck’s work.

1.4. Facial Geometry of the Elliptope and Optimization. From the bound (1.7), we learn that the elliptope \mathcal{E}_n is a uniformly good relaxation of the cut polytope \mathcal{C}_n . Yet the elliptope approximates the cut polytope far more accurately than (1.7) suggests.

Indeed, it is fruitful to think of the elliptope as a “shrink-wrapped” cut polytope. By this, we mean that the elliptope \mathcal{E}_n adheres to low-dimensional faces of the cut polytope \mathcal{C}_n , while it curves away from the higher-dimensional faces. Figure 1.1 illustrates this claim for \mathcal{C}_3 and \mathcal{E}_3 . We see that both \mathcal{C}_3 and \mathcal{E}_3 have the same vertices and edges, while the facets of \mathcal{C}_3 are not preserved in the relaxation \mathcal{E}_3 .

This paper demonstrates that a similar phenomenon holds more broadly. We will prove that many low-dimensional simplicial faces of the cut polytope \mathcal{C}_n are also faces of the elliptope \mathcal{E}_n .

This geometric observation provides a heuristic understanding of why the semidefinite relaxation (1.6) often performs better than (1.7) suggests.

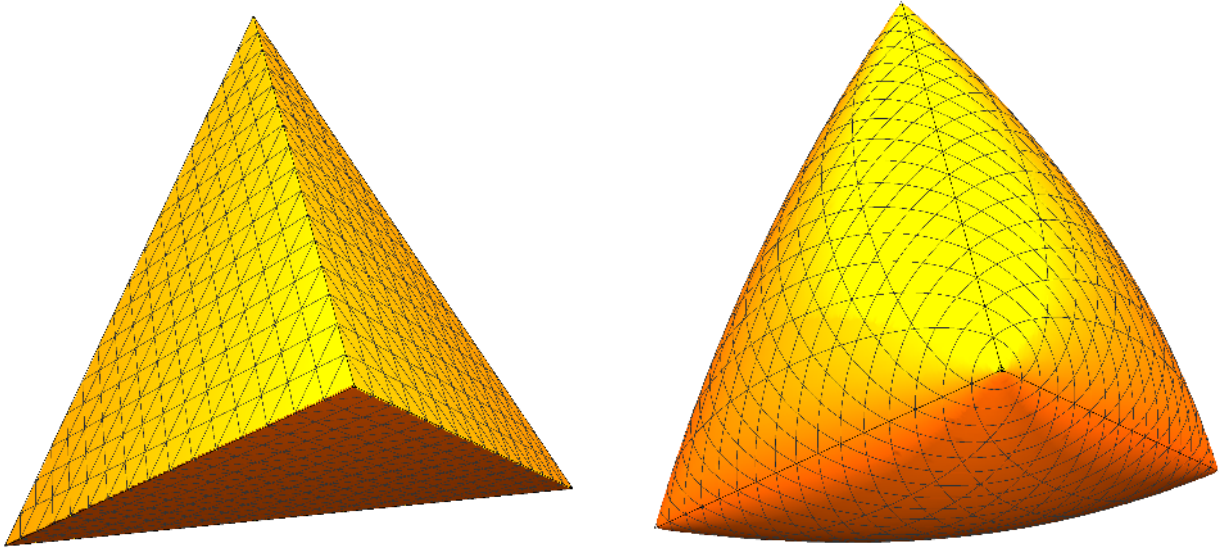


FIGURE 1.1: *The Cut Polytope and the Elliptope.* Define the linear map $T : \mathbb{S}^3 \rightarrow \mathbb{R}^3$ that extracts the strict lower triangle of a 3×3 symmetric matrix as a 3-dimensional vector. **[left]** The image $T(\mathcal{C}_3)$ of the cut polytope \mathcal{C}_3 . The closest vertex is the image of the cut matrix $\mathbf{1}\mathbf{1}^\top \in \mathbb{S}_+^3$. **[right]** The image $T(\mathcal{E}_3)$ of the elliptope \mathcal{E}_3 , seen from the same vantage.

1.5. Notation. We write $\|\cdot\|$ for the Euclidean norm on \mathbb{R}^d . The standard basis vector \mathbf{e}_i has a one in the i th coordinate and zeros elsewhere. The symbol $\mathbf{1}$ refers to a vector whose entries are identically equal to one. The dimensions of these special vectors are determined by context. The notation $^\top$ represents the transpose of a vector. We frequently use the componentwise product \odot of two vectors, which is also known as the Hadamard or Schur product.

The set \mathbb{S}^n contains the $n \times n$ real symmetric matrices, and \mathbb{S}_+^n is the subset of $n \times n$ real psd matrices. The map $\lambda_{\min} : \mathbb{S}^n \rightarrow \mathbb{R}$ computes the smallest eigenvalue of a symmetric matrix. The letter \mathbf{I} refers to the identity matrix, and the letter $\mathbf{J} := \mathbf{1}\mathbf{1}^\top$ denotes a square matrix of ones. The dimensions of these special matrices are determined by context. The symbol \vee refers to the symmetric tensor product of vectors or matrices; see [Bha97, Chap. I] for an overview of multilinear algebra.

The operator \mathbb{P} returns the probability of an event, while \mathbb{E} computes the expectation of a random variable, a random vector, or a random matrix. The abbreviation *iid* means independent and identically distributed. Small capitals (e.g., SBERN) are used for the names of probability distributions. The symbol \sim means “has the distribution.”

2. BACKGROUND AND RESULTS

This section outlines our results on the facial structure of the elliptope (1.5). We begin with a review of the definition of a simplicial face of a convex set, and we summarize known results about the simplicial faces of the elliptope. Next, we describe a random model that generates candidates for simplicial faces. The main results delineate situations where this construction is likely to be successful.

2.1. Facial Geometry of Convex Sets. We begin with a reminder about some relevant definitions from convex geometry. For further background, see [Roc70, HUL01].

Definition 2.1 (Dimension). Let K be a convex set in \mathbb{R}^d . The *dimension* of K is defined as the dimension of the affine hull of K .

Definition 2.2 (Face). Let K be a convex set in \mathbb{R}^d . A *face* F of K is a convex subset of K for which

$$\mathbf{x}, \mathbf{y} \in K \quad \text{and} \quad \theta \mathbf{x} + (1 - \theta) \mathbf{y} \in F \quad \text{for some } \theta \in (0, 1) \quad \text{imply that} \quad \mathbf{x}, \mathbf{y} \in F.$$

A face is also called an *extreme set*. A 0-dimensional face is commonly called an *extreme point*.

Definition 2.3 (Simplicial Face). A k -dimensional face F of a convex set is *simplicial* if F is the convex hull of an affinely independent family of $k + 1$ points.

Definition 2.4 (Vertex). Let K be a convex set in \mathbb{R}^d . A point $\mathbf{x} \in K$ is a *vertex* of K if the normal cone $\mathcal{N}(\mathbf{x}; K)$ has dimension d . For reference, $\mathcal{N}(\mathbf{x}; K) := \{\mathbf{z} \in \mathbb{R}^d : \mathbf{z}^\top(\mathbf{y} - \mathbf{x}) \leq 0 \text{ for all } \mathbf{y} \in K\}$.

Heuristically, a vertex is a sharp corner of a convex set. Vertices are always extreme points, but extreme points need not be vertices!

2.2. Facial Geometry of the Elliptope. The literature contains a lot of information about the facial geometry of the elliptope. Let us present some key results, which are due to Laurent & Poljak [LP95, LP96].

Fact 2.5 (Vertices). *The elliptope \mathcal{E}_n has 2^{n-1} vertices. These vertices are precisely the $n \times n$ cut matrices $\mathbf{c}\mathbf{c}^\top$, where $\mathbf{c} \in \{\pm 1\}^n$ is a cut vector.*

The most natural candidate for a face of the elliptope is the convex hull of a set of vertices. This construction does not always yield a face, but—when it does—that face is always simplicial.

Fact 2.6 (Faces Generated by Vertices are Simplicial). *Let $\mathbf{c}_1, \dots, \mathbf{c}_r \in \{\pm 1\}^n$ be cut vectors. Consider the set*

$$F := \text{conv}\{\mathbf{c}_1 \mathbf{c}_1^\top, \dots, \mathbf{c}_r \mathbf{c}_r^\top\} \subset \mathcal{E}_n.$$

If F is a face of the elliptope \mathcal{E}_n , then F is a simplicial face of \mathcal{E}_n .

Fact 2.6 does *not* assert that every simplicial face of the elliptope is generated by vertices. Even so, we can bound the possible dimension of a simplicial face, regardless of its structure.

Fact 2.7 (Dimension of Simplicial Faces). *The elliptope \mathcal{E}_n has a simplicial face of dimension k if and only if $k(k + 1) \leq 2(n - 1)$. In particular, it is necessary that the dimension $k < \sqrt{2(n - 1)}$.*

Our interest in the facial structure of the elliptope is motivated by its connection with the facial structure of the cut polytope (1.3).

Fact 2.8 (Coincidental Faces). *If F is a face of the elliptope \mathcal{E}_n generated by vertices, then F is also a face of the cut polytope \mathcal{C}_n .*

Indeed, Fact 2.8 follows directly from the definition of a face, the fact that the elliptope contains the cut polytope, and the fact that the vertices of the elliptope are elements of the cut polytope.

2.3. A Random Model for Simplicial Faces. We will study the extent to which the elliptope \mathcal{E}_n approximates the cut polytope \mathcal{C}_n by identifying a large number of simplicial faces of the elliptope. Our approach is based on the probabilistic method. In view of Fact 2.6, we can attempt to construct simplicial faces of \mathcal{E}_n by drawing a collection of random vertices and forming its convex hull.

For a parameter $p \in [0, 1]$, we define the *signed Bernoulli distribution*:

$$\text{SBERN}(p) := \begin{cases} +1, & \text{with probability } p; \\ -1, & \text{with probability } 1 - p. \end{cases}$$

We extend this distribution to vectors of length n by taking a direct product:

$$\text{SBERN}(p, n) := \text{SBERN}(p) \times \dots \times \text{SBERN}(p) \in \{\pm 1\}^n.$$

That is, a random vector from $\text{SBERN}(p, n)$ has n entries, each drawn independently from $\text{SBERN}(p)$.

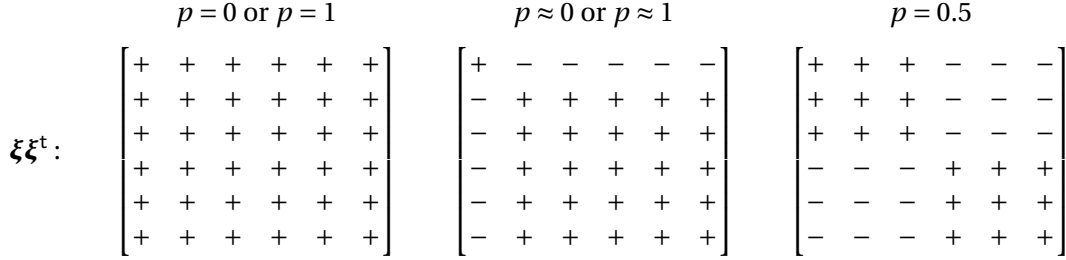


FIGURE 2.1: *Balance of Random Cut Matrices.* Let $\xi \sim \text{SBERN}(p, n)$. The balance parameter p controls the (average) proportion of entries in the random cut matrix $\xi\xi^t$ that are positive and negative. This display gives a nominal illustration of the effect. We have abbreviated the numbers ± 1 by their signs \pm .

Fix the parameter $p \in (0, 1)$, the number r of vertices, and the dimension n . Let us present a random model $\text{FACE}(p, r, n)$ for a prospective face F of the ellipsope \mathcal{E}_n . Draw random vectors ξ_1, \dots, ξ_r independently from the distribution $\text{SBERN}(p, n)$. Construct the random convex set

$$F := \text{conv}\{\xi_1\xi_1^t, \dots, \xi_r\xi_r^t\} \subset \mathcal{E}_n. \quad (2.1)$$

Our goal is to understand when F is likely to be a simplicial face of \mathcal{E}_n . The parameter p controls the typical “balance” of positive and negative entries that appear in the random vertices $\xi_i\xi_i^t$. See Figure 2.1 for a simple illustration.

2.4. Simplicial Faces: Quantitative Results. Our first set of results gives quantitative bounds on the probability that the random model $\text{FACE}(p, r, n)$ generates a simplicial face of \mathcal{E}_n . By this, we mean that the bounds contain explicit and reasonable constants, but the form of the bounds is suboptimal. The first statement concerns the special case where the balance parameter $p = 0.5$.

Theorem 2.9 (Simplicial Faces I). *Fix the balance parameter $p = 0.5$, the number r of vertices, and the dimension n . Draw a random set F from the distribution $\text{FACE}(p, r, n)$ described in (2.1). Then*

$$\mathbb{P}\{F \text{ is a simplicial face of } \mathcal{E}_n \text{ with dimension } r-1\} \geq 1 - r^2 \exp\left(\frac{-n}{r^2}\right).$$

The proof of Theorem 2.9 begins in Section 3 and continues in Section 4.

Here is the most noteworthy consequence of Theorem 2.9. When $r^2 \log r^2 < n$, there is a positive probability that a random set $F \sim \text{FACE}(0.5, r, n)$ is a simplicial face of \mathcal{E}_n with dimension $r-1$. The stricter bound $r \leq \sqrt{n/\log n}$ is also sufficient for F to be a simplicial face. It follows from Fact 2.8 that \mathcal{E}_n and \mathcal{C}_n share a large number of simplicial faces with dimension up to $\sqrt{n/\log n}$.

Next, we generalize to the case where the balance parameter p is general. This result has slightly larger constants than Theorem 2.9.

Theorem 2.10 (Simplicial Faces II). *Fix the balance parameter $p \in (0, 1)$, the number r of vertices, and the dimension n . Draw a random set F from the distribution $\text{FACE}(p, r, n)$ described in (2.1). Then*

$$\mathbb{P}\{F \text{ is a simplicial face of } \mathcal{E}_n \text{ with dimension } r-1\} \geq 1 - r^2 \exp\left(\frac{-4p^2(1-p)^2 n}{r^2}\right).$$

The proof of Theorem 2.10 begins in Section 3 and continues in Section 4.

Theorem 2.10 yields simplicial faces when the number r satisfies $r \leq 2p(1-p)\sqrt{n/\log n}$. In particular, for any fixed choice of the balance parameter p and the number r of vertices, the model $\text{FACE}(p, r, n)$ produces simplicial faces whenever the dimension n is sufficiently large.

2.5. Simplicial Faces: Qualitative Result. Our last result gives a qualitative bound on the probability that the random model $\text{FACE}(p, r, n)$ generates a simplicial face of \mathcal{E}_n . By this, we mean that the form of the bound improves on Theorem 2.10, but the analysis incurs large constant factors.

Theorem 2.11 (Simplicial Faces III). *Fix the balance parameter $p \in (0, 1)$, the number r of vertices, and the dimension n . Draw a random set F from the distribution $\text{FACE}(p, r, n)$ described in (2.1). Then*

$$\mathbb{P}\{F \text{ is a simplicial face of } \mathcal{E}_n \text{ with dimension } r-1\} \geq 1 - 4 \exp\left(\frac{r^2 - cp^2(1-p)^2n}{4}\right).$$

The number c is a positive, absolute constant that satisfies $c \geq 0.0003$.

The proof of Theorem 2.11 begins in Section 3 and continues in Section 5.

Here is the key consequence of Theorem 2.11. When $r+3 \leq p(1-p)\sqrt{cn}$, there is a positive probability that a random set $F \sim \text{FACE}(p, r, n)$ is a simplicial face of \mathcal{E}_n with dimension $r-1$. Fact 2.7 shows that this bound is qualitatively optimal when $p = 0.5$.

We see that Theorem 2.11 removes a parasitic logarithmic term from Theorem 2.9. In addition, the probability bound in Theorem 2.11 is significantly stronger. On the other hand, the relative size of the constants ensures that Theorem 2.9 gives a quantitative benefit for any realistic dimension n .

2.6. Counting Faces. Theorems 2.9, 2.10, and 2.11 also have an enumerative interpretation. Draw a random vector $\xi \sim \text{SBERN}(0.5, n)$. Then ξ is uniformly distributed over the set $\{\pm 1\}^n$ of cut vectors, and Fact 2.5 implies that $\xi\xi^\dagger$ is a uniformly random vertex of the elliptope \mathcal{E}_n . These observations yield an alternative procedure for drawing a random set F from the model $\text{FACE}(0.5, r, n)$: Let F be the convex hull of r vertices of \mathcal{E}_n , chosen uniformly at random, with replacement.¹ We obtain roughly $(e2^{n-1}/r)^r$ different sets in this manner. When $r \ll \sqrt{n}$, most of these sets are simplicial faces of \mathcal{E}_n .

2.7. Related Work. We can articulate the heuristic that a “small” collection of “generic” vertices of the elliptope induces a simplicial face. Theorems 2.9, 2.10, and 2.11 are all instantiations of this principle. In fact, this idea is already visible in a combinatorial construction of Laurent & Poljak [LP96, Cor. 4.5].

Fact 2.12 (Combinatorial Construction of Simplicial Faces). *Let $\mathbf{c}_1, \dots, \mathbf{c}_r \in \{\pm 1\}^n$ be cut vectors in general position. That is,*

$$\left[\bigodot_{i \in I} (\mathbf{1} + \mathbf{c}_i) \right] \odot \left[\bigodot_{i \notin I} (\mathbf{1} - \mathbf{c}_i) \right] \neq \mathbf{0} \quad \text{for each subset } I \text{ of } \{1, \dots, r\}.$$

Then $\text{conv}\{\mathbf{c}_1 \mathbf{c}_1^\dagger, \dots, \mathbf{c}_r \mathbf{c}_r^\dagger\}$ is a simplicial face of the elliptope \mathcal{E}_n .

Fact 2.12 only has content when the number r of cut vectors satisfies $r \leq \log_2 n$. In contrast, our probabilistic argument is valid in a wider parameter range. Theorem 2.9 operates in the regime $r \leq \sqrt{n/\log n}$, and Theorem 2.11 has implications when $r \leq \sqrt{cn}$. Even so, our analysis depends on the same sufficient condition (Fact 3.1, below) that Laurent & Poljak use to establish Fact 2.12.

3. PROOF STRATEGY

This section outlines our technique for proving Theorems 2.9, 2.10, and 2.11. The argument begins with a sufficient condition, due to Laurent & Poljak [LP96], for a family of cut vectors to generate a simplicial face. The challenge is to understand the probability that a collection of *random* cut vectors satisfies the sufficient condition. We explain how to reduce this question to a problem that can be addressed using matrix concentration inequalities. In Sections 4 and 5, we carry out this program.

¹In our parameter regime, it is unlikely that any vertex of \mathcal{E}_n is chosen more than once, so this model is not substantially different from drawing vertices without replacement.

3.1. Deterministic Condition for a Simplicial Face. The first ingredient in our argument is a sufficient condition [LP96, Thm. 4.2] for a family of cut vectors to generate a simplicial face of the elliptope.

Fact 3.1 (Sufficient Condition for a Simplicial Face). *Assume that $r \geq 2$. Let $\mathbf{c}_1, \dots, \mathbf{c}_r \in \{\pm 1\}^n$ be cut vectors, and consider the set*

$$F := \text{conv}\{\mathbf{c}_1 \mathbf{c}_1^\top, \dots, \mathbf{c}_r \mathbf{c}_r^\top\} \subset \mathcal{E}_n.$$

Define $R := 1 + r(r-1)/2$, and introduce the two matrices

$$\begin{aligned} \mathbf{W} &:= \mathbf{W}(\mathbf{c}_1, \dots, \mathbf{c}_r) := [\mathbf{c}_1 \ \dots \ \mathbf{c}_r] \in \mathbb{R}^{n \times r}; \\ \mathbf{Z} &:= \mathbf{Z}(\mathbf{c}_1, \dots, \mathbf{c}_r) := [\mathbf{1} \mid \mathbf{c}_1 \odot \mathbf{c}_2 \ \dots \ \mathbf{c}_i \odot \mathbf{c}_j \ \dots \ \mathbf{c}_{r-1} \odot \mathbf{c}_r] \in \mathbb{R}^{n \times R} \quad \text{where } 1 \leq i < j \leq r. \end{aligned} \quad (3.1)$$

If \mathbf{W} and \mathbf{Z} both have full column rank, then F is a simplicial face of \mathcal{E}_n with dimension $r-1$.

For our purposes, it is more natural to consider the dual form of the condition that \mathbf{W} and \mathbf{Z} have full column rank. Express these two matrices in terms of their rows:

$$\mathbf{W} = \begin{bmatrix} \mathbf{w}_1^\top \\ \vdots \\ \mathbf{w}_n^\top \end{bmatrix} \in \mathbb{R}^{n \times r} \quad \text{and} \quad \mathbf{Z} = \begin{bmatrix} \mathbf{z}_1^\top \\ \vdots \\ \mathbf{z}_n^\top \end{bmatrix} \in \mathbb{R}^{n \times R}.$$

Suppose that

$$\text{span}\{\mathbf{w}_1, \dots, \mathbf{w}_n\} = \mathbb{R}^r \quad \text{and} \quad \text{span}\{\mathbf{z}_1, \dots, \mathbf{z}_n\} = \mathbb{R}^R. \quad (3.2)$$

Then Fact 3.1 implies that the F is a simplicial face of \mathcal{E}_n .

Remark 3.2 (Variant Sufficient Condition). Let $R' := r(r-1)/2$. Using the same notation as in Fact 3.1, we define the matrix

$$\mathbf{Y} := \mathbf{Y}(\mathbf{c}_1, \dots, \mathbf{c}_r) := [\mathbf{1} - \mathbf{c}_1 \odot \mathbf{c}_2 \ \dots \ \mathbf{1} - \mathbf{c}_i \odot \mathbf{c}_j \ \dots \ \mathbf{1} - \mathbf{c}_{r-1} \odot \mathbf{c}_r] \in \mathbb{R}^{n \times R'}.$$

The indices lie in the range $1 \leq i < j \leq r$. Using the fact that the \mathbf{c}_i are cut vectors, Laurent & Poljak [LP96, Condition (iii), p. 540] show that \mathbf{Y} has full column rank if and only if \mathbf{Z} has full column rank.

3.2. Sufficient Condition for the Random Model. Fix the balance parameter $p \in (0, 1)$, the number r of vertices, and the ambient dimension n . Draw independent random vectors ξ_1, \dots, ξ_r from the distribution $\text{SBERN}(p, n)$, and construct the random set

$$F := \text{conv}\{\xi_1 \xi_1^\top, \dots, \xi_r \xi_r^\top\} \sim \text{FACE}(p, r, n).$$

We need to determine the probability that ξ_1, \dots, ξ_r satisfy the sufficient condition from Fact 3.1. This gives a lower bound on the probability that F is a simplicial face of \mathcal{E}_n .

We can check the dual form (3.2) of the sufficient condition. Consider the matrix $\mathbf{W} = \mathbf{W}(\xi_1, \dots, \xi_r) \in \mathbb{R}^{n \times r}$, defined in (3.1). Observe that the coordinates of the ξ_i are iid, so the matrix \mathbf{W} has iid rows. More precisely, the n rows of \mathbf{W} are iid copies of a random vector $\mathbf{w} \in \mathbb{R}^r$ where $\mathbf{w} \sim \text{SBERN}(p, r)$.

In a similar vein, consider the matrix $\mathbf{Z} = \mathbf{Z}(\xi_1, \dots, \xi_r) \in \mathbb{R}^{n \times R}$, defined in (3.1). Introduce a random vector $\mathbf{z} \in \mathbb{R}^R$, whose entries are derived from the random vector $\mathbf{w} \in \mathbb{R}^r$ as follows.

$$z_0 := 1 \quad \text{and} \quad z_{ij} := w_i w_j \quad \text{for } 1 \leq i < j \leq r. \quad (3.3)$$

Then the n rows of \mathbf{Z} are iid copies of the random vector \mathbf{z} .

Therefore, to verify (3.2), we must compute the probability that n iid copies of the random vector \mathbf{w} span the space \mathbb{R}^r and that n iid copies of the random vector \mathbf{z} span the space \mathbb{R}^R . The following proposition summarizes this discussion.

Proposition 3.3 (Sufficient Condition for Random Model). *Fix the balance parameter $p \in (0, 1)$, the number r of vertices where $r \geq 2$, and the dimension n . Define $R := 1 + r(r-1)/2$. Introduce a random vector $\mathbf{w} \sim \text{SBERN}(p, r)$, and define $\mathbf{z} \in \mathbb{R}^R$ by the formula (3.3).*

Draw iid copies $\mathbf{w}_1, \dots, \mathbf{w}_n$ of the random vector $\mathbf{w} \in \mathbb{R}^r$ and iid copies $\mathbf{z}_1, \dots, \mathbf{z}_n$ of the random vector $\mathbf{z} \in \mathbb{R}^R$. Then a random set $F \sim \text{FACE}(p, r, n)$ satisfies

$$\mathbb{P}\{F \text{ is not a simplicial face of } \mathcal{E}_n\} \leq \mathbb{P}\{\text{span}\{\mathbf{w}_1, \dots, \mathbf{w}_n\} \neq \mathbb{R}^r\} + \mathbb{P}\{\text{span}\{\mathbf{z}_1, \dots, \mathbf{z}_n\} \neq \mathbb{R}^R\}.$$

The easiest way to complete the calculations required by Proposition 3.3 is to invoke methods from the field of matrix concentration inequalities [Tro15b]. Among other things, this theory gives practical estimates for the minimum singular value of a random matrix with iid rows. This type of result leads directly to a bound on the probability that an iid family of random vectors spans a linear space. In the next two sections, we complete our program by combining Proposition 3.3 with two different types of matrix concentration.

Remark 3.4 (Variant Sufficient Condition for Random Model). Instate the notation from Proposition 3.3. Define $R' := r(r-1)/2$, and derive a random vector $\mathbf{y} \in \mathbb{R}^{R'}$ from the vector $\mathbf{w} \in \mathbb{R}^r$ as follows.

$$y_{ij} := 1 - w_i w_j \quad \text{for } 1 \leq i < j \leq r. \quad (3.4)$$

In view of Fact 3.1, Remark 3.2, and Proposition 3.3,

$$\mathbb{P}\{F \text{ is not a simplicial face of } \mathcal{E}_n\} \leq \mathbb{P}\{\text{span}\{\mathbf{w}_1, \dots, \mathbf{w}_n\} \neq \mathbb{R}^r\} + \mathbb{P}\{\text{span}\{\mathbf{y}_1, \dots, \mathbf{y}_n\} \neq \mathbb{R}^{R'}\}.$$

4. METHOD 1: THE MATRIX CHERNOFF INEQUALITY

This section contains the proofs of Theorem 2.9 and 2.10. The approach relies on a well-known consequence of the matrix Chernoff inequality [AW02, Tro12].

4.1. Tools. Let us present a specialization of the lower tail bound from the matrix Chernoff inequality. This result was first obtained by Ahlswede & Winter [AW02, Thm. 19], and it later received a significant upgrade [Tro12, Thm. 1.1].

Fact 4.1 (Ahlswede & Winter; Tropp). Consider a random vector $\mathbf{x} \in \mathbb{R}^d$ with second-moment matrix $\Sigma := \mathbb{E}[\mathbf{x}\mathbf{x}^\dagger]$. Assume that

$$\lambda := \lambda_{\min}(\Sigma) \quad \text{and} \quad \|\mathbf{x}\|^2 \leq B \quad \text{almost surely.}$$

Draw iid copies $\mathbf{x}_1, \dots, \mathbf{x}_s$ of the random vector \mathbf{x} . Then

$$\mathbb{P}\{\text{span}\{\mathbf{x}_1, \dots, \mathbf{x}_s\} \neq \mathbb{R}^d\} \leq d \cdot \exp\left(\frac{-\lambda s}{2B}\right).$$

Proof Sketch. This statement follows immediately by applying the simplified form [Tro12, Rem. 5.3] of the matrix Chernoff inequality for the minimum eigenvalue to the random matrices $\mathbf{X}_i = \mathbf{x}_i \mathbf{x}_i^\dagger$. We set the tail parameter $t = 0$. \square

Remark 4.2 (Refined Probability Bounds). Fact 4.1 gives a good estimate for how large s must be to make the probability bound nontrivial. To obtain more accurate bounds when s is larger, one must combine the matrix Chernoff bound with a scalar concentration inequality. We omit these developments.

4.2. Proof of Theorem 2.9. For this result, the balance parameter $p = 0.5$. We may also assume that $r \geq 2$, or else the result holds trivially because of Fact 2.5. We instate the notation of Proposition 3.3.

Let $\mathbf{w}_1, \dots, \mathbf{w}_n$ be iid copies of $\mathbf{w} \in \mathbb{R}^r$. Fact 4.1 readily implies that

$$\mathbb{P}\{\text{span}\{\mathbf{w}_1, \dots, \mathbf{w}_n\} \neq \mathbb{R}^r\} \leq r \cdot \exp\left(\frac{-n}{2r}\right). \quad (4.1)$$

Indeed, since $\mathbf{w} \sim \text{SBERN}(0.5, r)$, we quickly determine that $\|\mathbf{w}\|^2 = r$ and that $\mathbb{E}[\mathbf{w}\mathbf{w}^\dagger] = \mathbf{I}$.

Now, let $\mathbf{z}_1, \dots, \mathbf{z}_n$ be iid copies of $\mathbf{z} \in \mathbb{R}^R$. In this case, Fact 4.1 yields

$$\mathbb{P}\{\text{span}\{\mathbf{z}_1, \dots, \mathbf{z}_n\} \neq \mathbb{R}^R\} \leq R \cdot \exp\left(\frac{-n}{2R}\right). \quad (4.2)$$

To establish this point, recall that the random vector $\mathbf{z} \in \mathbb{R}^R$ is derived from \mathbf{w} via the formula (3.3). Therefore, $\|\mathbf{z}\|^2 = R$ and a short calculation yields $\mathbb{E}[\mathbf{z}\mathbf{z}^\dagger] = \mathbf{I}$.

To complete the proof, combine the bounds (4.1) and (4.2):

$$\mathbb{P}\{\text{span}\{\mathbf{w}_1, \dots, \mathbf{w}_n\} \neq \mathbb{R}^r\} + \mathbb{P}\{\text{span}\{\mathbf{z}_1, \dots, \mathbf{z}_n\} \neq \mathbb{R}^{R'}\} \leq r^2 \cdot \exp\left(\frac{-n}{r^2}\right).$$

We have used the relations $2r \leq r + R \leq 2R \leq r^2$, which are valid because $R = 1 + r(r-1)/2$ and $r \geq 2$. An application of Proposition 3.3 completes the proof of Theorem 2.9.

4.3. Proof of Theorem 2.10. As before, we may assume that $r \geq 2$. This time, we need the alternative sufficient condition from Remark 3.4, and we instate the notation from this remark. It is also productive to abbreviate $\alpha := (2p-1)^2$, which is the squared expectation of an SBERN(p) random variable.

Let $\mathbf{w} \sim \text{SBERN}(p, r)$. It is immediate that $\|\mathbf{w}\|^2 = r$. By direct calculation, the second-moment matrix of the random vector $\mathbf{w} \in \mathbb{R}^r$ takes the form

$$\mathbf{M}_r := \mathbb{E}[\mathbf{w}\mathbf{w}^\dagger] = (1-\alpha) \cdot \mathbf{I} + \alpha \cdot \mathbf{J} \in \mathbb{S}_+^r. \quad (4.3)$$

Recall that \mathbf{J} is the matrix of ones. It follows immediately that $\lambda_{\min}(\mathbf{M}_r) = 1 - \alpha$. Fact 4.1 delivers

$$\mathbb{P}\{\text{span}\{\mathbf{w}_1, \dots, \mathbf{w}_n\} \neq \mathbb{R}^r\} \leq r \cdot \exp\left(-\frac{(1-\alpha)n}{2r}\right), \quad (4.4)$$

where $\mathbf{w}_1, \dots, \mathbf{w}_n$ are iid copies of the random vector \mathbf{w} .

Next, consider the random vector $\mathbf{y} \in \mathbb{R}^{R'}$, derived from $\mathbf{w} \in \mathbb{R}^r$ via the formula (3.4). Note that $\|\mathbf{y}\|^2 \leq 4R'$ because the entries of \mathbf{y} takes values in the set $\{0, 2\}$. We assert the following bound on the minimum eigenvalue of the second-moment matrix of \mathbf{y} .

Claim 4.3. *Let $\Sigma := \mathbb{E}[\mathbf{y}\mathbf{y}^\dagger]$ be the second-moment matrix of \mathbf{y} . Then $\lambda_{\min}(\Sigma) \geq (1-\alpha)^2$.*

We will verify Claim 4.3 in Section 4.4. Granted this result, Fact 4.1 provides

$$\mathbb{P}\{\text{span}\{\mathbf{y}_1, \dots, \mathbf{y}_n\} \neq \mathbb{R}^{R'}\} \leq R' \cdot \exp\left(-\frac{(1-\alpha)^2 n}{8R'}\right), \quad (4.5)$$

where $\mathbf{y}_1, \dots, \mathbf{y}_n$ are iid copies of the random vector \mathbf{y} .

Combine (4.4) and (4.5) to reach

$$\mathbb{P}\{\text{span}\{\mathbf{w}_1, \dots, \mathbf{w}_n\} \neq \mathbb{R}^r\} + \mathbb{P}\{\text{span}\{\mathbf{y}_1, \dots, \mathbf{y}_n\} \neq \mathbb{R}^{R'}\} \leq r^2 \cdot \exp\left(\frac{-(1-\alpha)^2 n}{4r^2}\right).$$

We have also used the relations $2r \leq r + R' \leq 2R' \leq r^2$, which hold because $r \geq 2$ and $R' = r(r-1)/2$. This calculation also depends on the bound $(1-\alpha)^2 \leq (1-\alpha)$, which is valid because $\alpha \in (0, 1)$. Finally, note that $1-\alpha = 4p(1-p)$. In view of Remark 3.4, we arrive at Theorem 2.10.

4.4. Proof of Claim 4.3. The argument is expressed most easily in the language of multilinear algebra; see [Bha97, Chap. I] for more background. This approach was inspired by conversations with Richard Küng.

Introduce the linear space $V := \mathbb{R}^r \vee \mathbb{R}^r$ of symmetric tensors, equipped with the real inner product $\langle \cdot, \cdot \rangle$. Let $\vee : \mathbb{R}^r \times \mathbb{R}^r \rightarrow V$ be the symmetric bilinear map that constructs an elementary symmetric tensor $\mathbf{u} \vee \mathbf{v} \in V$ from two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^r$. We always have the relation $\mathbf{u} \vee \mathbf{v} = \mathbf{v} \vee \mathbf{u}$. The space V admits the orthonormal basis $\{\mathbf{e}_i \vee \mathbf{e}_j : 1 \leq i \leq j \leq r\}$, where the \mathbf{e}_i are the standard basis vectors in \mathbb{R}^r . For linear operators \mathbf{A}, \mathbf{B} acting on \mathbb{R}^r , we can define a linear operator $\mathbf{A} \vee \mathbf{B}$ acting on V by the rule

$$\mathbf{A} \vee \mathbf{B} : \mathbf{e}_i \vee \mathbf{e}_j \longmapsto (\mathbf{A}\mathbf{e}_i) \vee (\mathbf{B}\mathbf{e}_j) \quad \text{for } 1 \leq i \leq j \leq r.$$

If \mathbf{A} and \mathbf{B} are both psd, then $\mathbf{A} \vee \mathbf{B}$ is a psd operator on V .

Consider the subspace $W := \text{span}\{\mathbf{e}_i \vee \mathbf{e}_j : 1 \leq i < j \leq r\}$ of the inner product space V . It is natural to treat the random vector $\mathbf{y} \in \mathbb{R}^{r(r-1)/2}$ as an element of W by identifying $\langle \mathbf{y}, \mathbf{e}_i \vee \mathbf{e}_j \rangle = y_{ij}$ for $1 \leq i < j \leq r$. Similarly, $\Sigma = \mathbb{E}[\mathbf{y}\mathbf{y}^\dagger]$ is the linear operator on W given by

$$\langle \mathbf{e}_{i'} \vee \mathbf{e}_{j'}, \Sigma(\mathbf{e}_i \vee \mathbf{e}_j) \rangle = \mathbb{E}[\langle \mathbf{e}_{i'} \vee \mathbf{e}_{j'}, \mathbf{y} \rangle \langle \mathbf{y}, \mathbf{e}_i \vee \mathbf{e}_j \rangle] \quad \text{where} \quad \begin{array}{l} 1 \leq i < j \leq r; \\ 1 \leq i' < j' \leq r. \end{array}$$

Using the definition (3.4) of \mathbf{y} and the fact that $\mathbf{w} \sim \text{SBERN}(p, r)$, we quickly compute that

$$\langle \mathbf{e}_{i'} \vee \mathbf{e}_{j'}, \boldsymbol{\Sigma}(\mathbf{e}_i \vee \mathbf{e}_j) \rangle = \begin{cases} (1-\alpha)^2 + (1-\alpha^2), & i = i' \text{ and } j = j' \\ (1-\alpha)^2 + \alpha(1-\alpha), & i = i' \text{ xor } j = j' \\ (1-\alpha)^2 + \alpha(1-\alpha), & i = j' \text{ xor } j = i' \\ (1-\alpha)^2, & i \neq i' \text{ and } j \neq j' \end{cases} \quad \text{where} \quad \begin{cases} 1 \leq i < j \leq r; \\ 1 \leq i' < j' \leq r. \end{cases}$$

This expression is useful, but it takes more work to expose the spectral properties of $\boldsymbol{\Sigma}$.

The key idea is to identify the operator $\boldsymbol{\Sigma}$ on W as the restriction of an operator $\check{\boldsymbol{\Sigma}}$ on V . Define

$$\check{\boldsymbol{\Sigma}} := (1-\alpha)^2 \cdot (\mathbf{I} \vee \mathbf{I}) + \alpha(1-\alpha) \cdot (\mathbf{I} \vee \mathbf{J} + \mathbf{J} \vee \mathbf{I}) + \frac{1}{2}(1-\alpha)^2 \cdot (\mathbf{J} \vee \mathbf{J}). \quad (4.6)$$

Using the fact that $\mathbf{J}\mathbf{e}_i = \mathbf{1} = \sum_{k=1}^r \mathbf{e}_k$ for each index $1 \leq i \leq r$, we can check that

$$\langle \mathbf{e}_{i'} \vee \mathbf{e}_{j'}, \check{\boldsymbol{\Sigma}}(\mathbf{e}_i \vee \mathbf{e}_j) \rangle = \langle \mathbf{e}_{i'} \vee \mathbf{e}_{j'}, \boldsymbol{\Sigma}(\mathbf{e}_i \vee \mathbf{e}_j) \rangle \quad \text{where} \quad \begin{cases} 1 \leq i < j \leq r; \\ 1 \leq i' < j' \leq r. \end{cases}$$

As promised, $\boldsymbol{\Sigma}$ is the restriction of $\check{\boldsymbol{\Sigma}}$ to W .

The representation (4.6) of the operator $\check{\boldsymbol{\Sigma}}$ allows us to determine its spectrum with ease. The operators \mathbf{I} and \mathbf{J} are both psd, so the operators $\mathbf{I} \vee \mathbf{J}$ and $\mathbf{J} \vee \mathbf{I}$ and $\mathbf{J} \vee \mathbf{J}$ are also psd. Weyl's monotonicity principle implies that

$$\lambda_{\min}(\check{\boldsymbol{\Sigma}}) \geq \lambda_{\min}((1-\alpha)^2 \cdot (\mathbf{I} \vee \mathbf{I})) = (1-\alpha)^2.$$

The first inequality depends on the fact that $\alpha \in [0, 1]$, and the second relation holds because $\mathbf{I} \vee \mathbf{I}$ is the identity operator on V . Finally, the operator $\boldsymbol{\Sigma}$ is a restriction of $\check{\boldsymbol{\Sigma}}$, so we conclude that

$$\lambda_{\min}(\boldsymbol{\Sigma}) \geq \lambda_{\min}(\check{\boldsymbol{\Sigma}}) \geq (1-\alpha)^2.$$

This establishes Claim 4.3.

5. METHOD 2: OLIVEIRA'S LOWER TAIL INEQUALITY

This section contains the proof of Theorem 2.11. The argument depends on a recent matrix concentration inequality due to Oliveira [Oli16].

5.1. Tools. We begin with a summary of the technical tools that we require. The key result is a specialization of Oliveira's lower tail inequality [Oli16, Thm. 1.1].

Fact 5.1 (Oliveira). *Consider a random vector $\mathbf{x} \in \mathbb{R}^d$ whose second-moment matrix $\mathbb{E}[\mathbf{x}\mathbf{x}^\dagger]$ is nonsingular. Compute the hypercontractive parameter*

$$h := h(\mathbf{x}) := \max_{\mathbf{u} \neq \mathbf{0}} \frac{\mathbb{E}(\mathbf{x}^\dagger \mathbf{u})^4}{[\mathbb{E}(\mathbf{x}^\dagger \mathbf{u})^2]^2}.$$

Draw iid copies $\mathbf{x}_1, \dots, \mathbf{x}_s$ of the random vector \mathbf{x} . Then

$$\mathbb{P} \left\{ \text{span} \{ \mathbf{x}_1, \dots, \mathbf{x}_s \} \neq \mathbb{R}^d \right\} \leq 2 \exp \left(\frac{d}{2} - \frac{s}{162h} \right).$$

To bound the parameter h that appears in the last result, we need the following hypercontractive inequality. For example, see [O'D14, Thm. 10.21].

Fact 5.2 (Hypercontractivity). *Consider a polynomial $q : \{\pm 1\}^n \rightarrow \mathbb{R}$ with real coefficients and degree k . Draw a random vector $\boldsymbol{\xi} \sim \text{SBERN}(p, n)$ where $p \in (0, 1)$. Then the random variable $X := q(\boldsymbol{\xi})$ satisfies*

$$\frac{\mathbb{E} X^4}{(\mathbb{E} X^2)^2} \leq \left[\frac{9}{p(1-p)} \right]^k.$$

5.2. Proof of Theorem 2.11. As usual, assume that $r \geq 2$. We also instate the notation from Proposition 3.3. Recall that the random vector $\mathbf{w} \sim \text{SBERN}(p, r)$. It is easy to verify that the second-moment matrix $\mathbb{E}[\mathbf{w}\mathbf{w}^\top]$ is nonsingular; see (4.3). The entries of the random vector $\mathbf{z} \in \mathbb{R}^R$ are derived from \mathbf{w} by means of the formula (3.3). The second-moment matrix $\mathbb{E}[\mathbf{z}\mathbf{z}^\top]$ is also nonsingular; indeed, for fixed $\mathbf{u} \in \mathbb{R}^R$, the random variable $(\mathbf{z}^\top \mathbf{u})^2$ is identically zero only if $\mathbf{u} = \mathbf{0}$.

Let us begin with the probability that n iid copies $\mathbf{w}_1, \dots, \mathbf{w}_n$ of the random vector $\mathbf{w} \in \mathbb{R}^r$ span all of \mathbb{R}^r . We will use Oliveira's result, Fact 5.1, to establish that

$$\mathbb{P}\{\text{span}\{\mathbf{w}_1, \dots, \mathbf{w}_n\} \neq \mathbb{R}^r\} \leq 2 \exp\left(\frac{r}{2} - \frac{p(1-p)n}{1458}\right). \quad (5.1)$$

Observe that, for any vector $\mathbf{u} \in \mathbb{R}^r$, the linear form $\mathbf{w}^\top \mathbf{u} = \sum_{i=1}^r w_i u_i$ is a polynomial of degree one in the entries of \mathbf{w} . Fact 5.2 implies that

$$\frac{\mathbb{E}(\mathbf{w}^\top \mathbf{u})^4}{[\mathbb{E}(\mathbf{w}^\top \mathbf{u})^2]^2} \leq \frac{9}{p(1-p)}.$$

Therefore, the hypercontractive parameter $h(\mathbf{w}) \leq 9p^{-1}(1-p)^{-1}$. The claim (5.1) now follows from Fact 5.1.

Second, we study the probability that iid copies $\mathbf{z}_1, \dots, \mathbf{z}_n$ of the random vector $\mathbf{z} \in \mathbb{R}^R$ span all of \mathbb{R}^R . We will apply Fact 5.1 to obtain

$$\mathbb{P}\{\text{span}\{\mathbf{z}_1, \dots, \mathbf{z}_n\} \neq \mathbb{R}^R\} \leq 2 \exp\left(\frac{R}{2} - \frac{p^2(1-p)^2 n}{13122}\right). \quad (5.2)$$

For any vector $\mathbf{u} \in \mathbb{R}^R$, we can express $\mathbf{z}^\top \mathbf{u} = u_0 + \sum_{i < j} w_i w_j u_{ij}$. This is a polynomial of degree two in the entries of \mathbf{w} . Fact 5.2 implies that

$$\frac{\mathbb{E}(\mathbf{z}^\top \mathbf{u})^4}{[\mathbb{E}(\mathbf{z}^\top \mathbf{u})^2]^2} \leq \left[\frac{9}{p(1-p)}\right]^2.$$

Therefore, the parameter $h(\mathbf{z}) \leq 81p^{-2}(1-p)^{-2}$, and the claim (5.2) follows from Fact 5.1.

To complete the argument, combine the inequalities (5.1) and (5.2) to arrive at the estimate

$$\mathbb{P}\{\text{span}\{\mathbf{w}_1, \dots, \mathbf{w}_n\} \neq \mathbb{R}^r\} + \mathbb{P}\{\text{span}\{\mathbf{z}_1, \dots, \mathbf{z}_n\} \neq \mathbb{R}^R\} \leq 4 \exp\left(\frac{r^2}{4} - \frac{p^2(1-p)^2 n}{13122}\right).$$

We have used the bounds $r \leq R \leq r^2/2$, which are valid because $r \geq 2$ and $R = 1 + r(r-1)/2$. Introduce this inequality into Proposition 3.3 to complete the proof of Theorem 2.11.

Remark 5.3 (Alternative Proof). Let us mention an alternative approach to Theorem 2.11 based on Mendelson's Small Ball Method [KM15, Tro15a]. This technique yields a comparable outcome, but it takes more steps to apply.

ACKNOWLEDGMENTS

The author thanks Richard Küng and Benjamin Recht for helpful conversations related to this work. This research was partially supported by ONR award N00014-11-1002 and the Gordon & Betty Moore Foundation.

REFERENCES

- [AW02] R. Ahlswede and A. Winter. Strong converse for identification via quantum channels. *IEEE Trans. Inform. Theory*, 48(3):569–579, 2002.
- [Bha97] R. Bhatia. *Matrix analysis*. Springer-Verlag, New York, 1997.
- [BTN01] A. Ben-Tal and A. Nemirovski. *Lectures on modern convex optimization*. MPS/SIAM Series on Optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Programming Society (MPS), Philadelphia, PA, 2001. Analysis, algorithms, and engineering applications.
- [DP93] C. Delorme and S. Poljak. The performance of an eigenvalue bound on the max-cut problem in some classes of graphs. *Discrete Math.*, 111(1-3):145–156, 1993. Graph theory and combinatorics (Marseille-Luminy, 1990).
- [Gro53] A. Grothendieck. Résumé de la théorie métrique des produits tensoriels topologiques. *Bol. Soc. Mat. São Paulo*, 8:1–79, 1953.

- [GW95] M. X. Goemans and D. P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. Assoc. Comput. Mach.*, 42(6):1115–1145, 1995.
- [HUL01] J.-B. Hiriart-Urruty and C. Lemaréchal. *Fundamentals of convex analysis*. Grundlehren Text Editions. Springer-Verlag, Berlin, 2001. Abridged version of it Convex analysis and minimization algorithms. I [Springer, Berlin, 1993; MR1261420 (95m:90001)] and it II [ibid.; MR1295240 (95m:90002)].
- [Kar72] R. M. Karp. Reducibility among combinatorial problems. In *Complexity of computer computations (Proc. Sympos., IBM Thomas J. Watson Res. Center, Yorktown Heights, N.Y., 1972)*, pages 85–103. Plenum, New York, 1972.
- [KM15] V. Koltchinskii and S. Mendelson. Bounding the smallest singular value of a random matrix without concentration. *Int. Math. Res. Not. IMRN*, (23):12991–13008, 2015.
- [KN12] S. Khot and A. Naor. Grothendieck-type inequalities in combinatorial optimization. *Comm. Pure Appl. Math.*, 65(7):992–1035, 2012.
- [LP95] M. Laurent and S. Poljak. On a positive semidefinite relaxation of the cut polytope. *Linear Algebra Appl.*, 223/224:439–461, 1995. Special issue honoring Miroslav Fiedler and Vlastimil Pták.
- [LP96] M. Laurent and S. Poljak. On the facial structure of the set of correlation matrices. *SIAM J. Matrix Anal. Appl.*, 17(3):530–547, 1996.
- [MT11] M. McCoy and J. A. Tropp. Two proposals for robust PCA using semidefinite programming. *Electron. J. Stat.*, 5:1123–1160, 2011.
- [Nes98] Y. Nesterov. Semidefinite relaxation and nonconvex quadratic optimization. *Optim. Methods Softw.*, 9(1-3):141–160, 1998.
- [O’D14] R. O’Donnell. *Analysis of Boolean functions*. Cambridge University Press, New York, 2014.
- [Oli16] R. I. Oliveira. The lower tail of random quadratic forms with applications to ordinary least squares. *Probab. Theory Related Fields*, 166(3-4):1175–1194, 2016.
- [Pis12] G. Pisier. Grothendieck’s theorem, past and present. *Bull. Amer. Math. Soc. (N.S.)*, 49(2):237–323, 2012.
- [PR95] S. Poljak and F. Rendl. Solving the max-cut problem using eigenvalues. *Discrete Appl. Math.*, 62(1-3):249–278, 1995. Partitioning and decomposition in combinatorial optimization.
- [Roc70] R. T. Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.
- [Tro12] J. A. Tropp. User-friendly tail bounds for sums of random matrices. *Found. Comput. Math.*, 12(4):389–434, 2012.
- [Tro15a] J. A. Tropp. Convex recovery of a structured signal from independent random linear measurements. In *Sampling theory, a renaissance*, Appl. Numer. Harmon. Anal., pages 67–101. Birkhäuser/Springer, Cham, 2015.
- [Tro15b] J. A. Tropp. An introduction to matrix concentration inequalities. *Foundations and Trends in Machine Learning*, 8(1-2):1–230, 2015.