Sum-of-Squares Optimization and the Sparsity Structure of Equiangular Tight Frames

Afonso S. Bandeira*† and Dmitriy Kunisky*
*Courant Institute of Mathematical Sciences, New York University, NY 10012
†Center for Data Science, New York University, NY 10012

***Abstract*—Equiangular tight frames (ETFs) may be used to construct examples of feasible points for semidefinite programs arising in sum-of-squares (SOS) optimization. We show how generalizing the calculations in a recent work of the authors' that explored this connection also yields new bounds on the sparsity of (both real and complex) ETFs. One corollary shows that Steiner ETFs corresponding to finite projective planes are optimally sparse in the sense of achieving tightness in a matrix inequality controlling overlaps between sparsity patterns of distinct rows of the synthesis matrix. We also formulate several natural open problems concerning further generalizations of our technique.**

I. Introduction

One of the most important objects in combinatorial optimization is the *cut polytope*, the convex set of matrices

$$\mathscr{C}^N := \mathsf{conv}\left(\left\{\boldsymbol{x}\boldsymbol{x}^\top : \boldsymbol{x} \in \{\pm 1\}^N\right\}\right) \subset \mathbb{R}^{N\times N}_{\mathsf{sym}}. \tag{1}$$

The cut polytope has a rich discrete geometry [6] describing the solution space of the problem of finding the largest cut in a graph, which amounts to maximizing a linear function over $\mathscr{C}^N$. By the classical result of Karp [15], this problem cannot be solved in polynomial time unless $\mathsf{P} = \mathsf{NP}$. Therefore, *relaxations* of $\mathscr{C}^N$ to larger but more algorithmically tractable convex sets have been proposed, perhaps the best-known of which is the relaxation to the *(real) elliptope*

$$\mathscr{E}^N := \left\{\boldsymbol{X} \in \mathbb{R}^{N\times N}_{\mathsf{sym}} : \boldsymbol{X} \succeq \boldsymbol{0}, \mathsf{diag}(\boldsymbol{X}) = \boldsymbol{1}\right\} \supseteq \mathscr{C}^N. \tag{2}$$

There is extensive literature both on the geometry of $\mathscr{E}^N$ (thoroughly described in [6]), and on approximating optimization over $\mathscr{C}^N$ by optimization over $\mathscr{E}^N$ (e.g. [14], [8], [16]).

Every $\boldsymbol{X} \in \mathscr{E}^N$ is the Gram matrix of unit vectors, $X_{ij} = \langle \boldsymbol{v}_i, \boldsymbol{v}_j\rangle$ for some $\boldsymbol{v}_1, \dots, \boldsymbol{v}_N \in \mathbb{R}^r$ with $r := \mathsf{rank}(\boldsymbol{X}) \le N$. The boundary of $\mathscr{E}^N$ consists of $\boldsymbol{X}$ having $r < N$. In the language of frame theory, these boundary points are the Gram matrices of *overcomplete unit norm frames*. We will explore applications of geometric results about $\mathscr{E}^N$ to the following types of structured frames.

Definition I.1. *Unit vectors $\boldsymbol{v}_1, \dots, \boldsymbol{v}_N \in \mathbb{C}^r$ with $\boldsymbol{X} = (\langle \boldsymbol{v}_i, \boldsymbol{v}_j\rangle)_{i,j=1}^N$ form a* unit norm tight frame (UNTF) *if any of the following equivalent conditions hold.*

1) $\sum_{i=1}^N \boldsymbol{v}_i\boldsymbol{v}_i^* = \frac{N}{r}\boldsymbol{I}_r$.
2) *The nonzero eigenvalues of $\boldsymbol{X}$ all equal $\frac{N}{r}$.*
3) $\|\boldsymbol{X}\|_F^2 = \sum_{i=1}^N\sum_{j=1}^N |X_{ij}|^2 = \frac{N^2}{r}$.

The $\boldsymbol{v}_i$ form an equiangular tight frame (ETF) *if moreover there exists $\alpha \in [0,1]$ such that $|X_{ij}| = |\langle \boldsymbol{v}_i, \boldsymbol{v}_j\rangle| = \alpha$ whenever $i \ne j$.*

When moreover $\boldsymbol{v}_i \in \mathbb{R}^r$, we have $\boldsymbol{X} \in \mathscr{E}^N$, and such points form an interesting subset of the elliptope's boundary. In both the real and complex cases, UNTFs and ETFs have been studied in great detail previously (e.g. [20], [4], [3], [11]).

The set $\mathscr{E}^N$, however, is only the first of a sequence of tightening relaxations of $\mathscr{C}^N$, described by the *sum-of-squares (SOS)* hierarchy (see [17], [18]). In the recent work [1], the authors studied membership of the Gram matrices of real ETFs in one such tighter relaxation through a connection with the following general concept from convex geometry (see Section IV for further details).

Definition I.2. *For a closed convex set $K \subseteq \mathbb{R}^d$ and $\boldsymbol{X} \in K$, the* perturbation of $\boldsymbol{X}$ in K *is the subspace*

$$\mathsf{pert}_K(\boldsymbol{X}) := \left\{\boldsymbol{A} \in \mathbb{R}^d : \exists\, t > 0 \text{ with } \boldsymbol{X} \pm t\boldsymbol{A} \in K\right\}. \tag{3}$$

To that end, in [1] we computed the projection operator to $\mathsf{pert}_{\mathscr{E}_2^N}(\boldsymbol{X})$ for $\boldsymbol{X}$ the Gram matrix of any real ETF.

In this paper, we give a self-contained presentation of this calculation for perturbation subspaces in the complex generalization of the elliptope,

$$\widetilde{\mathscr{E}}^N := \left\{\boldsymbol{X} \in \mathbb{C}^{N\times N}_{\mathsf{herm}} : \boldsymbol{X} \succeq \boldsymbol{0}, \mathsf{diag}(\boldsymbol{X}) = \boldsymbol{1}\right\}. \tag{4}$$

Using that this projection operator is positive semidefinite (psd), we derive inequalities in degree 4 polynomials of the entries of (real or complex) ETF vectors, which translate to new inequalities controlling the sparsity and spark of ETF vectors. Further background and details of abbreviated proofs are available in the extended version [2] of this paper.

II. Projecting to the Perturbation Subspace

In this section, we compute the orthogonal projection operator to $\mathsf{pert}_{\widetilde{\mathscr{E}}^N}(\boldsymbol{X})$,

$$\mathcal{P}\boldsymbol{A} := \underset{\boldsymbol{B} \in \mathsf{pert}_{\widetilde{\mathscr{E}}^N}(\boldsymbol{X})}{\mathrm{argmin}} \frac{1}{2}\|\boldsymbol{A} - \boldsymbol{B}\|_F^2. \tag{5}$$

The key tool is the following classical result describing perturbation subspaces for the elliptope.

Proposition II.1 (Theorem 1(a) of [19]). *Let $\boldsymbol{X} \in \widetilde{\mathscr{E}}^N$ with $\mathsf{rank}(\boldsymbol{X}) = r$, and let $\boldsymbol{v}_1, \dots, \boldsymbol{v}_N \in \mathbb{C}^r$ such that $X_{ij} = \langle \boldsymbol{v}_i, \boldsymbol{v}_j\rangle$ for $i, j \in [N]$. Write $\boldsymbol{V} \in \mathbb{C}^{r\times N}$ for the matrix with the $\boldsymbol{v}_i$ as its columns. Then,*

$$\mathsf{pert}_{\widetilde{\mathscr{E}}^N}(\boldsymbol{X}) = \left\{\boldsymbol{V}^*\boldsymbol{H}\boldsymbol{V} : \begin{array}{c}\boldsymbol{H} \in \mathbb{C}^{r\times r}_{\mathsf{herm}}, \\ \boldsymbol{v}_i^*\boldsymbol{H}\boldsymbol{v}_i = 0\ \forall\, i \in [N]\end{array}\right\}. \tag{6}$$

Our formula is then the following (retaining the notation from Proposition II.1).

Lemma II.2. *Let $\boldsymbol{v}_1, \dots, \boldsymbol{v}_N \in \mathbb{C}^r$ form a UNTF. Suppose the $\boldsymbol{v}_i\boldsymbol{v}_i^*$ are linearly independent, or equivalently that $|\boldsymbol{X}|^{\odot 2}$ is non-singular. Let $\boldsymbol{x}_1, \dots, \boldsymbol{x}_N$ be the columns of $\boldsymbol{X}$. Then,*

$$\mathcal{P}\boldsymbol{A} = \frac{r^2}{N^2}\left(\boldsymbol{XAX} - \sum_{i,j=1}^{N}(|\boldsymbol{X}|^{\odot 2})^{-1}_{ij}(\boldsymbol{x}_i^*\boldsymbol{A}\boldsymbol{x}_i)\boldsymbol{x}_j\boldsymbol{x}_j^*\right). \quad (7)$$

(For any $\boldsymbol{M} \in \mathbb{C}^{m\times n}$, we denote by $|\boldsymbol{M}|^{\odot 2}$ the matrix with entries $|M_{ij}|^2$, the Schur product of $\boldsymbol{M}$ with its conjugate.)

Proof. By Proposition II.1 and the variational characterization (5), we have

$$\mathcal{P}\boldsymbol{A} = \boldsymbol{V}^*\boldsymbol{H}^\star(\boldsymbol{A})\boldsymbol{V}, \quad (8)$$

$$\boldsymbol{H}^\star(\boldsymbol{A}) := \operatorname*{argmin}_{\substack{\boldsymbol{H}\in\mathbb{C}^{r\times r}_{\mathsf{herm}} \\ \boldsymbol{v}_i^*\boldsymbol{H}\boldsymbol{v}_i = 0 \text{ for } i\in[N]}} \mathsf{obj}(\boldsymbol{H};\boldsymbol{A}), \quad (9)$$

$$\begin{aligned}\mathsf{obj}(\boldsymbol{H};\boldsymbol{A}) &:= \frac{1}{2}\|\boldsymbol{A} - \boldsymbol{V}^*\boldsymbol{H}\boldsymbol{V}\|_F^2 \\ &= \frac{1}{2}\|\boldsymbol{A}\|_F^2 + \frac{N^2}{2r^2}\|\boldsymbol{H}\|_F^2 - \langle \boldsymbol{VAV}^*, \boldsymbol{H}\rangle. \quad (10)\end{aligned}$$

(In the final equation we use the UNTF property.) Introducing a vector of Lagrange multipliers $\boldsymbol{\gamma} \in \mathbb{R}^N$ for the constraints in the optimization defining $\boldsymbol{H}^\star$, we obtain the Lagrangian

$$\begin{aligned}L(\boldsymbol{H},\boldsymbol{\gamma};\boldsymbol{A}) &:= \mathsf{obj}(\boldsymbol{H};\boldsymbol{A}) - \sum_{i=1}^{N}\gamma_i\boldsymbol{v}_i^*\boldsymbol{H}\boldsymbol{v}_i \\ &= \frac{1}{2}\|\boldsymbol{A}\|_F^2 + \frac{N^2}{2r^2}\|\boldsymbol{H}\|_F^2 \\ &\quad - \left\langle \boldsymbol{VAV}^* + \sum_{i=1}^{N}\gamma_i\boldsymbol{v}_i\boldsymbol{v}_i^*, \boldsymbol{H}\right\rangle. \quad (11)\end{aligned}$$

The first-order condition for optimality then implies that

$$\boldsymbol{H}^\star(\boldsymbol{A}) = \boldsymbol{VAV}^* + \sum_{i=1}^{N}\gamma_i(\boldsymbol{A})\boldsymbol{v}_i\boldsymbol{v}_i^* \quad (12)$$

for some $\boldsymbol{\gamma}(\boldsymbol{A})$ such that $\boldsymbol{v}_i^*\boldsymbol{H}^\star(\boldsymbol{A})\boldsymbol{v}_i = 0$ for $i \in [N]$. These constraints may be written as the system

$$\sum_{j=1}^{N}|\langle\boldsymbol{v}_i,\boldsymbol{v}_j\rangle|^2\gamma_j(\boldsymbol{A}) = -\boldsymbol{v}_i^*\boldsymbol{VAV}^*\boldsymbol{v}_i \text{ for } i\in[N], \quad (13)$$

which is a linear system in $\boldsymbol{\gamma}(\boldsymbol{A})$ with matrix $|\boldsymbol{X}|^{\odot 2}$. Since this matrix is invertible by assumption, there is a unique solution

$$\gamma_i(\boldsymbol{A}) = -\sum_{j=1}^{N}(|\boldsymbol{X}|^{\odot 2})^{-1}_{ij}\boldsymbol{v}_j^*\boldsymbol{VAV}^*\boldsymbol{v}_j. \quad (14)$$

The result follows by substituting into (12) and then (8). □

The result we will use to obtain sparsity inequalities follows from manipulations of the fact that $\mathcal{P}$ is psd, whereby $\langle\boldsymbol{A},\mathcal{P}\boldsymbol{A}\rangle \geq 0$ for any $\boldsymbol{A} \in \mathbb{C}^{N\times N}_{\mathsf{herm}}$. Surprisingly, this fact is equivalent to the following simpler matrix inequality.

Lemma II.3. *Let $\boldsymbol{v}_1, \dots, \boldsymbol{v}_N \in \mathbb{C}^r$ form a UNTF. Suppose the $\boldsymbol{v}_i\boldsymbol{v}_i^*$ are linearly independent, or equivalently that $|\boldsymbol{X}|^{\odot 2}$ is non-singular. Then,*

$$(|\boldsymbol{V}|^{\odot 2})(|\boldsymbol{X}|^{\odot 2})^{-1}(|\boldsymbol{V}|^{\odot 2})^\top \preceq \boldsymbol{I}_r, \quad (15)$$

or equivalently

$$(|\boldsymbol{V}|^{\odot 2})^\top(|\boldsymbol{V}|^{\odot 2}) \preceq |\boldsymbol{X}|^{\odot 2}. \quad (16)$$

Remark. *After obtaining Lemma II.3, we discovered the reference [21], which gives a general result of the form* (16) *not depending on the $\boldsymbol{v}_i$ forming a tight frame. The proof of [21] builds a psd block matrix through the Schur product theorem and uses that its Schur complement remains psd. Our proof gives a more direct geometric argument, and, as we discuss in Section IV, may generalize to higher-degree inequalities.*

Proof. Let us write

$$\mathcal{P}\boldsymbol{A} := \mathcal{P}_1\boldsymbol{A} - \mathcal{P}_2\boldsymbol{A}, \quad (17)$$

$$\mathcal{P}_1\boldsymbol{A} := \frac{r^2}{N^2}\boldsymbol{XAX}, \quad (18)$$

$$\mathcal{P}_2\boldsymbol{A} := \frac{r^2}{N^2}\sum_{i,j=1}^{N}(|\boldsymbol{X}|^{\odot 2})^{-1}_{ij}(\boldsymbol{x}_i^*\boldsymbol{A}\boldsymbol{x}_i)\boldsymbol{x}_j\boldsymbol{x}_j^*. \quad (19)$$

Writing $\widehat{\boldsymbol{V}} := \sqrt{\frac{r}{N}}\boldsymbol{V}$, we have

$$\langle\boldsymbol{A},\mathcal{P}_1\boldsymbol{A}\rangle = \left\|\widehat{\boldsymbol{V}}\boldsymbol{A}\widehat{\boldsymbol{V}}^*\right\|_F^2, \quad (20)$$

$$\langle\boldsymbol{A},\mathcal{P}_2\boldsymbol{A}\rangle = \sum_{i,j=1}^{N}\Big[(|\boldsymbol{X}|^{\odot 2})^{-1}_{ij} \left(\boldsymbol{v}_i^*\widehat{\boldsymbol{V}}\boldsymbol{A}\widehat{\boldsymbol{V}}^*\boldsymbol{v}_i\right)\left(\boldsymbol{v}_j^*\widehat{\boldsymbol{V}}\boldsymbol{A}\widehat{\boldsymbol{V}}^*\boldsymbol{v}_j\right)\Big]. \quad (21)$$

Any $\boldsymbol{H} \in \mathbb{C}^{r\times r}_{\mathsf{herm}}$ may be expressed in the form $\frac{r}{N}\widehat{\boldsymbol{V}}\boldsymbol{A}\widehat{\boldsymbol{V}}^*$ by taking $\boldsymbol{A} = \widehat{\boldsymbol{V}}^*\boldsymbol{H}\widehat{\boldsymbol{V}}$. Therefore, the inequality $\langle\boldsymbol{A},\mathcal{P}\boldsymbol{A}\rangle = \langle\boldsymbol{A},\mathcal{P}_1\boldsymbol{A}\rangle - \langle\boldsymbol{A},\mathcal{P}_2\boldsymbol{A}\rangle \geq 0$ holding for $\boldsymbol{A} \in \mathbb{C}^{N\times N}_{\mathsf{herm}}$ is equivalent to the following inequality holding for $\boldsymbol{H} \in \mathbb{C}^{r\times r}_{\mathsf{herm}}$:

$$\sum_{i=1}^{N}\sum_{j=1}^{N}(|\boldsymbol{X}|^{\odot 2})^{-1}_{ij}\left(\boldsymbol{v}_i^*\boldsymbol{H}\boldsymbol{v}_i\right)\left(\boldsymbol{v}_j^*\boldsymbol{H}\boldsymbol{v}_j\right) \leq \|\boldsymbol{H}\|_F^2. \quad (22)$$

Since applying a unitary transformation to a UNTF produces another UNTF having the same Gram matrix $\boldsymbol{X}$, we may assume that $\boldsymbol{H}$ is diagonal and real-valued. Rewriting (22) in the diagonal entries of $\boldsymbol{H}$ then gives the result. □

III. Sparsity Inequalities for ETFs

Our remaining results take advantage of the fact that the matrix $|\boldsymbol{X}|^{\odot 2}$ is very simple for an ETF:

$$|\boldsymbol{X}|^{\odot 2} = (1-\alpha^2)\boldsymbol{I}_r + \alpha^2\boldsymbol{1}\boldsymbol{1}^\top. \quad (23)$$

Moreover, by the well-known *Welch bound* [22], α depends only on the dimension parameters N and r:

$$\alpha = \sqrt{\frac{N-r}{r(N-1)}}. \quad (24)$$

In this case, Lemma II.3 gives the following.

Theorem III.1. *Let* $\boldsymbol{v}_1, \dots, \boldsymbol{v}_N \in \mathbb{C}^r$ *for* $r > 1$ *form an ETF. Define* $\boldsymbol{R} := (|\boldsymbol{V}|^{\odot 2})(|\boldsymbol{V}|^{\odot 2})^\top \in \mathbb{R}^{r \times r}$*, with entries* $R_{k\ell} = \sum_{i=1}^N |(\boldsymbol{v}_i)_k|^2 |(\boldsymbol{v}_i)_\ell|^2$*. Then,*

$$\boldsymbol{R} \preceq \frac{1-\frac{1}{r}}{1-\frac{1}{N}} \boldsymbol{I}_r + \frac{\frac{N}{r}-1}{r(1-\frac{1}{N})} \boldsymbol{1}\boldsymbol{1}^\top. \tag{25}$$

Proof. When $r > 1$ then $\alpha < 1$, so $|\boldsymbol{X}|^{\odot 2}$ is invertible and Lemma II.3 applies. From (23) and (24), we find

$$(|\boldsymbol{X}|^{\odot 2})^{-1} = \frac{1-\frac{1}{N}}{1-\frac{1}{r}} \left(\boldsymbol{I}_N - \frac{N-r}{N(N-1)} \boldsymbol{1}\boldsymbol{1}^\top \right). \tag{26}$$

We have

$$(|\boldsymbol{V}|^{\odot 2}) \boldsymbol{1}\boldsymbol{1}^\top (|\boldsymbol{V}|^{\odot 2})^\top = \frac{N^2}{r^2} \boldsymbol{1}\boldsymbol{1}^\top, \tag{27}$$

and substituting into (15) gives the result. □

The next result gives the exact dimension of the subspace on which the inequality (25) is sharp for any *Steiner ETF* [13], a general combinatorial construction of an ETF as a Kronecker-like product of the incidence matrix of a Steiner system with a Hadamard matrix. For the sake of brevity we assume all necessary terminology and basic combinatorial results, and refer the reader to the work [13] or the general reference [5] for further information.

Proposition III.2. *Let* $\boldsymbol{v}_1, \dots, \boldsymbol{v}_N \in \mathbb{R}^r$ *be a Steiner ETF constructed from a Steiner system with parameters* $(2, k, v)$ *and a Hadamard matrix of suitable size. Let* $\boldsymbol{R} \in \mathbb{R}^{r \times r}$ *have entries* $R_{k\ell} = \sum_{i=1}^N |(\boldsymbol{v}_i)_k|^2 |(\boldsymbol{v}_i)_\ell|^2$*. Then,*

$$\mathsf{dim}\left(\mathsf{ker}\left(\frac{1-\frac{1}{r}}{1-\frac{1}{N}} \boldsymbol{I}_r + \frac{\frac{N}{r}-1}{r(1-\frac{1}{N})} \boldsymbol{1}\boldsymbol{1}^\top - \boldsymbol{R} \right) \right) = v. \tag{28}$$

When the Steiner system is a finite projective plane, then $r = v$*, so the inequality* (25) *is an equality of matrices.*

Proof. Let $b = \frac{v(v-1)}{k(k-1)}$ be the number of blocks in the underlying Steiner system and $\rho = \frac{v-1}{k-1}$ be the number of blocks in which each point lies (typically this is denoted r in combinatorics, which we adjust to avoid conflict with our notation for ETF dimensions). Then, per the construction of [13], $r = b$ and $N = v(1+\rho)$.

We first compute $\boldsymbol{R}$: letting $\boldsymbol{N} \in \mathbb{R}^{v \times b}$ be the incidence matrix of points and blocks of the Steiner system, we find

$$\boldsymbol{R} = \frac{\rho+1}{\rho^2} \boldsymbol{N}^\top \boldsymbol{N} = k \boldsymbol{I}_b + \boldsymbol{A}_G, \tag{29}$$

where $\boldsymbol{A}_G$ is the adjacency matrix of the block intersection graph G. G is a strongly regular graph, admitting a spectral expansion

$$\boldsymbol{A}_G = k(\rho-1) \widehat{\boldsymbol{1}}_b \widehat{\boldsymbol{1}}_b^\top + (\rho - 1 - k) \boldsymbol{P}_{U_+} - k \boldsymbol{P}_{U_-}, \tag{30}$$

where $\widehat{\boldsymbol{1}}_b = \frac{1}{\sqrt{b}} \boldsymbol{1}_b$, $U_\pm$ are eigenspaces orthogonal to one another and to the vector $\boldsymbol{1}$ and satisfying $U_+ \oplus U_- \oplus \boldsymbol{1} = \mathbb{R}^b$, and $\boldsymbol{P}_{U_\pm}$ are the projectors onto these subspaces. The corresponding dimensions are

$$\mathsf{dim}(U_+) = v - 1, \tag{31}$$

$$\mathsf{dim}(U_-) = b - v. \tag{32}$$

We thus obtain the spectral expansion of $\boldsymbol{R}$,

$$\boldsymbol{R} = k\left(1 + \frac{1}{\rho}\right) \widehat{\boldsymbol{1}}_r \widehat{\boldsymbol{1}}_r^\top + \left(1 - \frac{1}{\rho^2}\right) \boldsymbol{P}_{U_+}. \tag{33}$$

Some algebraic manipulations show that the following identities hold between the eigenvalues of the right-hand side of (25) and those of $\boldsymbol{R}$:

$$k\left(1 + \frac{1}{\rho}\right) = \frac{1-\frac{1}{r}}{1-\frac{1}{N}} + \frac{N-r}{r(1-\frac{1}{N})}, \tag{34}$$

$$1 - \frac{1}{\rho^2} = \frac{1-\frac{1}{r}}{1-\frac{1}{N}}, \tag{35}$$

and therefore in fact

$$\frac{1-\frac{1}{r}}{1-\frac{1}{N}} \boldsymbol{I}_r + \frac{\frac{N}{r}-1}{r(1-\frac{1}{N})} \boldsymbol{1}\boldsymbol{1}^\top - \boldsymbol{R} = \frac{1-\frac{1}{r}}{1-\frac{1}{N}} \boldsymbol{P}_{U_-}. \tag{36}$$

The result then follows from the dimension formula (31). □

Finally, to illustrate how some more concrete results may be obtained from Theorem III.1, we give corollaries controlling ETF sparsity, spark, and the overlap of rows of the synthesis matrix $\boldsymbol{V}$. (We outline the proofs and leave further details to the extended paper [2].) Recall that the *spark* is defined as

$$\mathsf{spark}(\boldsymbol{V}) := \min_{\substack{\boldsymbol{x} \in \mathbb{C}^N \setminus \{\boldsymbol{0}\} \\ \boldsymbol{V}\boldsymbol{x} = \boldsymbol{0}}} \|\boldsymbol{x}\|_0 = \min_{\boldsymbol{x} \in \mathsf{row}(\boldsymbol{V})^\perp \setminus \{\boldsymbol{0}\}} \|\boldsymbol{x}\|_0. \tag{37}$$

The natural dual measure of sparsity, sometimes called *cospark*, is

$$\mathsf{sparsity}(\boldsymbol{V}) := \min_{\boldsymbol{x} \in \mathsf{row}(\boldsymbol{V}) \setminus \{\boldsymbol{0}\}} \|\boldsymbol{x}\|_0, \tag{38}$$

which gives control of the sparsity of the entire matrix $\boldsymbol{V}$ by controlling each row.

Corollary III.3. *Let* $\boldsymbol{v}_1, \dots, \boldsymbol{v}_N \in \mathbb{C}^r$ *form an ETF, and let* $\boldsymbol{V} \in \mathbb{C}^{r \times N}$ *have the* $\boldsymbol{v}_i$ *as its columns. Then,*

$$\mathsf{sparsity}(\boldsymbol{V}) \geq N \left(1 + \frac{(r-1)^2}{N-1}\right)^{-1}, \tag{39}$$

$$\mathsf{spark}(\boldsymbol{V}) \geq N \left(1 + \frac{(N-r-1)^2}{N-1}\right)^{-1}. \tag{40}$$

Proof Sketch. By applying a suitable rotation and testing the diagonal entries of (25), we obtain an upper bound on $\|\boldsymbol{V}^\top \boldsymbol{y}\|_4^4$ for any $\boldsymbol{y} \in \mathbb{C}^r$ with $\|\boldsymbol{y}\|_2 = 1$. Applying the Cauchy-Schwarz inequality $\|\boldsymbol{x}\|_0 \geq \|\boldsymbol{x}\|_2^4 / \|\boldsymbol{x}\|_4^4$ then gives the sparsity result. Applying the same argument to the Naimark complement ETF gives the spark result. □

This improves on a similar bound from Theorem 5 of [10]; on the Naimark complements of the infinite families of ETFs constructed by [7], [13], [12], [9] with dimensions scaling as $N \sim r^{3/2}$, our spark bound gives an asymptotic improvement of sub-leading order in the dimension.

The final corollary we mention shows that the overlap between sparsity patterns of distinct rows of an ETF in fact has a certain "typical" value for a given pair of dimensions r and N, from which its possible deviations are bounded.

Corollary III.4. *Let $\boldsymbol{v}_1, \dots, \boldsymbol{v}_N \in \mathbb{C}^r$ form an ETF, and let $\boldsymbol{V} \in \mathbb{C}^{r \times N}$ have the $\boldsymbol{v}_i$ as its columns. Let $\boldsymbol{a}, \boldsymbol{b} \in \mathsf{row}(\boldsymbol{V})$ with $\langle \boldsymbol{a}, \boldsymbol{b} \rangle = 0$ and $\|\boldsymbol{a}\|_2^2 = \|\boldsymbol{b}\|_2^2 = \frac{N}{r}$ (for instance, two distinct rows of $\boldsymbol{V}$). Let $D := \frac{N}{r^2}(1 + \frac{(r-1)^2}{N-1})$ and $E := \frac{\frac{N}{r} - 1}{r(1 - \frac{1}{N})}$. Then,*

$$\left|\langle |\boldsymbol{a}|^{\odot 2}, |\boldsymbol{b}|^{\odot 2} \rangle - E\right|^2 \leq \left(D - \|\boldsymbol{a}\|_4^4\right)\left(D - \|\boldsymbol{b}\|_4^4\right). \qquad (41)$$

Proof Sketch. D and E equal the diagonal and off-diagonal entries respectively of the right-hand side of (25). After rotating to an ETF in which $\boldsymbol{a}$ and $\boldsymbol{b}$ occur as rows, the result follows by taking the determinant of a 2×2 minor of the difference between the right- and left-hand sides of (25). □

IV. Open Problems

The results in this paper arose from studying sum-of-squares relaxations of the cut polytope, which are defined as follows.

Definition IV.1. *$\mathscr{E}_d^N$ is the set of $\boldsymbol{X} \in \mathbb{R}_{\mathsf{sym}}^{N \times N}$ for which there exists $\boldsymbol{Y} \in \mathbb{R}_{\mathsf{sym}}^{N^{d/2} \times N^{d/2}}$ where, identifying indices of $\boldsymbol{Y}$ with elements of $[N]^{d/2}$, the following conditions hold.*

1) *$Y_{\boldsymbol{ij}} = 1$ whenever all indices occur an even number of times across $\boldsymbol{i}$ and $\boldsymbol{j}$.*
2) *$Y_{\boldsymbol{ij}}$ depends only on the set of indices occuring an odd number of times across $\boldsymbol{i}$ and $\boldsymbol{j}$.*
3) *$Y_{(1,\dots,1,i)(1,\dots,1,j)} = X_{ij}$ for all $i, j \in [N]$.*
4) *$\boldsymbol{Y} \succeq \boldsymbol{0}$.*

The following inclusions hold among these sets:

$$\mathscr{E}^N = \mathscr{E}_2^N \supsetneq \mathscr{E}_4^N \supsetneq \cdots \supsetneq \mathscr{E}_{N+\mathbb{1}\{N \text{ odd}\}}^N = \mathscr{C}^N. \qquad (42)$$

In [1], we showed that if $\boldsymbol{X}$ is the Gram matrix of a real ETF of N vectors in $\mathbb{R}^r$, then $\boldsymbol{X} \in \mathscr{E}_4^N$ if and only if $N < \frac{r(r+1)}{2}$, in which case the subspace $\mathsf{vec}(\mathsf{pert}_{\mathscr{E}_2^N}(\boldsymbol{X}))$ forms an eigenspace of the associated $\boldsymbol{Y} \in \mathbb{R}^{N^2 \times N^2}$. These origins motivate the following question for future investigation.

Question IV.2. *Given the Gram matrix $\boldsymbol{X}$ of a real ETF of N vectors in $\mathbb{R}^r$, what is $d(\boldsymbol{X}) = \max\{d : \boldsymbol{X} \in \mathscr{E}_d^N\}$? Does the answer depend only on N and r, i.e. $d(\boldsymbol{X}) = d(N, \mathsf{rank}(\boldsymbol{X}))$?*

More specifically, we are interested in the details of the construction that would underlie such a result.

Question IV.3. *When $\boldsymbol{X}$ is the Gram matrix of a real ETF and $\boldsymbol{X} \in \mathscr{E}_d^N$ with a "witness" $\boldsymbol{Y}$ per Definition IV.1, is there a tractable description of the eigenspaces of $\boldsymbol{Y}$?*

If this is the case, then we may hope to imitate the present approach: compute projectors to the analogous subspaces for complex ETFs, write down the positivity relation for those operators, and derive polynomial inequalities in the ETF entries. It remains to be seen, however, whether such inequalities could still be interpreted as giving information about sparsity of ETFs, or whether they would control new quantities for higher degrees.

Acknowledgements

We thank Dustin Mixon for many useful discussions and references. Afonso S. Bandeira was partially supported by NSF grants DMS-1712730 and DMS-1719545, and by a grant from the Sloan Foundation. Dmitriy Kunisky was partially supported by NSF grants DMS-1712730 and DMS-1719545.

References

[1] Afonso S Bandeira and Dmitriy Kunisky. A Gramian description of the degree 4 generalized elliptope. *arXiv preprint arXiv:1812.11583*, 2018.
[2] Afonso S Bandeira and Dmitriy Kunisky. Sum-of-squares optimization and the sparsity structure of equiangular tight frames. *arXiv preprint arXiv:1901.10697*, 2019.
[3] Peter G Casazza and Gitta Kutyniok. *Finite frames: Theory and applications*. Springer, 2012.
[4] Peter G Casazza, Dan Redmond, and Janet C Tremain. Real equiangular frames. In *Information Sciences and Systems, 2008. CISS 2008. 42nd Annual Conference on*, pages 715–720. IEEE, 2008.
[5] Charles J Colbourn and Jeffrey H Dinitz. *Handbook of combinatorial designs*. CRC press, 2006.
[6] M.M. Deza and M. Laurent. *Geometry of Cuts and Metrics*. Algorithms and Combinatorics. Springer Berlin Heidelberg, 2009.
[7] Cunsheng Ding and Tao Feng. A generic construction of complex codebooks meeting the welch bound. *IEEE transactions on information theory*, 53(11):4245–4250, 2007.
[8] Uriel Feige and Gideon Schechtman. On the optimality of the random hyperplane rounding technique for MAX CUT. *Random Structures & Algorithms*, 20(3):403–440, 2002.
[9] Matthew Fickus, John Jasper, Dustin G Mixon, Jesse D Peterson, and Cody E Watson. Polyphase equiangular tight frames and abelian generalized quadrangles. *Applied and Computational Harmonic Analysis*, 2017.
[10] Matthew Fickus and Dustin G Mixon. Numerically erasure-robust frames. *Linear Algebra and its Applications*, 437(6):1394–1407, 2012.
[11] Matthew Fickus and Dustin G Mixon. Tables of the existence of equiangular tight frames. *arXiv preprint arXiv:1504.00253*, 2015.
[12] Matthew Fickus, Dustin G Mixon, and John Jasper. Equiangular tight frames from hyperovals. *IEEE Transactions on Information Theory*, 62(9):5225–5236, 2016.
[13] Matthew Fickus, Dustin G Mixon, and Janet C Tremain. Steiner equiangular tight frames. *Linear algebra and its applications*, 436(5):1014–1027, 2012.
[14] Michel X Goemans and David P Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM (JACM)*, 42(6):1115–1145, 1995.
[15] Richard M Karp. Reducibility among combinatorial problems. In *Complexity of computer computations*, pages 85–103. Springer, 1972.
[16] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan ODonnell. Optimal inapproximability results for max-cut and other 2-variable csps? *SIAM Journal on Computing*, 37(1):319–357, 2007.
[17] Jean B Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3):796–817, 2001.
[18] Monique Laurent. Sums of squares, moment matrices and optimization over polynomials. In *Emerging applications of algebraic geometry*, pages 157–270. Springer, 2009.
[19] Chi-Kwong Li and Bit-Shun Tam. A note on extreme correlation matrices. *SIAM Journal on Matrix Analysis and Applications*, 15(3):903–908, 1994.
[20] Mátyás A Sustik, Joel A Tropp, Inderjit S Dhillon, and Robert W Heath Jr. On the existence of equiangular tight frames. *Linear Algebra and its applications*, 426(2-3):619–635, 2007.
[21] Bo-Ying Wang and Fuzhen Zhang. Schur complements and matrix inequalities of Hadamard products. *Linear and Multilinear Algebra*, 43(1-3):315–326, 1997.
[22] Lloyd Welch. Lower bounds on the maximum cross correlation of signals (corresp.). *IEEE Transactions on Information theory*, 20(3):397–399, 1974.