Sum-of-Squares Optimization & Sparsity Structure of Equiangular Tight Frames

Tim Kunisky

(joint work with Afonso Bandeira)

Courant Institute of Mathematical Sciences

SampTA 2019 : July 9, 2019

The Two Topics:

Sum-of-Squares Relaxations (SOS)	Equiangular Tight Frames (ETFs)
continuous optimization	discrete geometry
semidefinite programming	harmonic analysis
algebraic proofs	combinatorial designs

The Basic Connection:

ETFs are feasible points for SOS relaxations of an optimization problem.

The Basic Connection:

ETFs are feasible points for SOS relaxations of an optimization problem.

Therefore, ETFs have to satisfy some (new!) inequalities.

Part 1: Motivation

A Gramian description of the degree 4 generalized elliptope (2018)

Initial Motivation: Relaxing MaxCut

Are **relaxations** good for problems like this?

$$\mathrm{M}(\boldsymbol{A}) := \max_{\boldsymbol{x} \in \{\pm 1\}^N} \boldsymbol{x}^\top \boldsymbol{A} \boldsymbol{x}$$

Initial Motivation: Relaxing MaxCut

Are **relaxations** good for problems like this?

$$\mathrm{M}(\boldsymbol{A}) := \max_{\boldsymbol{x} \in \{\pm 1\}^N} \boldsymbol{x}^\top \boldsymbol{A} \boldsymbol{x}$$

This is the usual one:

$$\mathrm{M}(\boldsymbol{A}) = \max_{\boldsymbol{x} \in \{\pm 1\}^N} \langle \boldsymbol{A}, \boldsymbol{x}\boldsymbol{x}^\top \rangle \le \max_{\substack{\boldsymbol{X} \succeq \boldsymbol{0} \\ X_{ii} = 1}} \langle \boldsymbol{A}, \boldsymbol{X} \rangle.$$

We call this a relaxation because it is based on an inclusion between the *cut polytope* and the *elliptope*,

$$\mathscr{C}^N := \mathrm{conv}(\{\boldsymbol{x}\boldsymbol{x}^\top : \boldsymbol{x} \in \{\pm 1\}^N\})$$

$$\bigcap$$

$$\{\boldsymbol{X} \in \mathbb{R}^{N \times N}_{\mathrm{sym}} : \boldsymbol{X} \succeq \boldsymbol{0}, X_{ii} = 1\} =: \mathscr{E}_2^N.$$

Sum-of-Squares: A Recipe for Improvement

To get a better algorithm, find a tighter inclusion!

Look at $\boldsymbol{Y} = (\boldsymbol{x} \otimes \boldsymbol{x})(\boldsymbol{x} \otimes \boldsymbol{x})^\top \in \mathbb{R}^{N^2 \times N^2}_{\mathrm{sym}}$, and write down linear and psd inequalities it must satisfy for $\boldsymbol{x} \in \{\pm 1\}^N$:

- Positivity: $\boldsymbol{Y} \succeq \boldsymbol{0}$.
- Normalization: $Y_{(ii)(jj)} = 1$.
- "$x_i^2 x_j x_k = x_j x_k$": $Y_{(ii)(jk)} = Y_{(i'i')(jk)}$.
- Symmetry: $Y_{(ij)(k\ell)} = Y_{(\pi(i)\pi(j))(\pi(k)\pi(\ell))}$.

Sum-of-Squares: A Recipe for Improvement

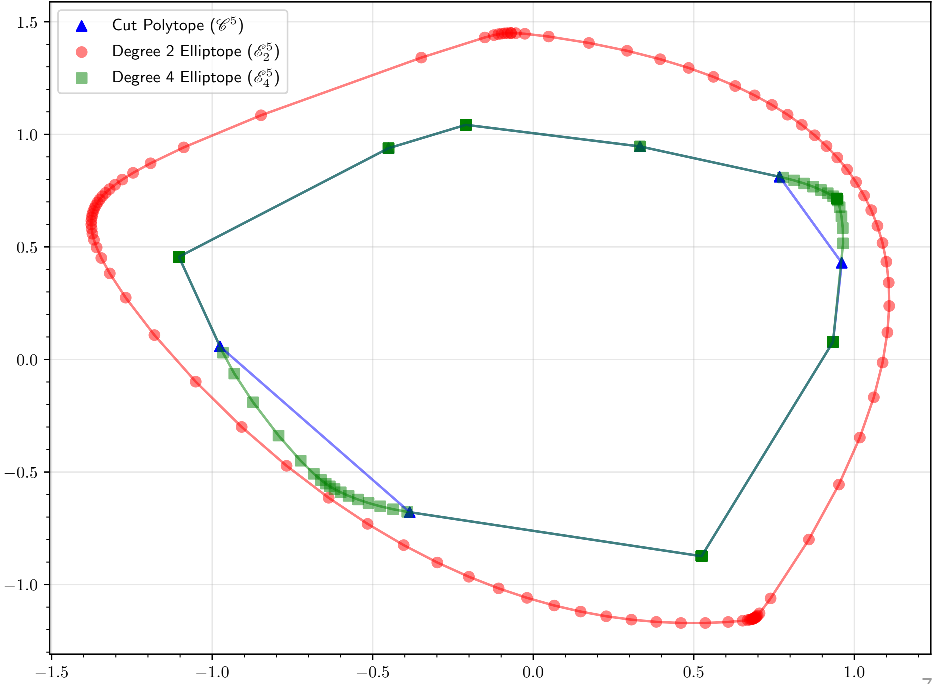
To get a better algorithm, find a tighter inclusion!

Look at $\boldsymbol{Y} = (\boldsymbol{x} \otimes \boldsymbol{x})(\boldsymbol{x} \otimes \boldsymbol{x})^\top \in \mathbb{R}^{N^2 \times N^2}_{\mathrm{sym}}$, and write down linear and psd inequalities it must satisfy for $\boldsymbol{x} \in \{\pm 1\}^N$:

- Positivity: $\boldsymbol{Y} \succeq \boldsymbol{0}$.
- Normalization: $Y_{(ii)(jj)} = 1$.
- "$x_i^2 x_j x_k = x_j x_k$": $Y_{(ii)(jk)} = Y_{(i'i')(jk)}$.
- Symmetry: $Y_{(ij)(k\ell)} = Y_{(\pi(i)\pi(j))(\pi(k)\pi(\ell))}$.

Let $\mathscr{E}_4^N$ be the $\boldsymbol{X} \in \mathbb{R}^{N \times N}_{\mathrm{sym}}$ that are *extendable* to such $\boldsymbol{Y}$ (occurring as $X_{ij} = Y_{(11)(ij)}$, corresponding to $\boldsymbol{x}\boldsymbol{x}^\top$).

Then, $\mathscr{C}^N \subseteq \mathscr{E}_4^N \subseteq \mathscr{E}_2^N$.



Cut Polytope ($\mathscr{C}^5$)
Degree 2 Elliptope ($\mathscr{E}_2^5$)
Degree 4 Elliptope ($\mathscr{E}_4^5$)
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−1.0
−0.5
0.0
0.5
1.0

Vague Problem:

"Understand how to build" an extension of a given $\boldsymbol{X} \in \mathscr{E}_2^N \setminus \mathscr{C}^N$ to degree 4.

Specific Problem:

Explicitly construct some members of $\mathscr{E}_4^N \setminus \mathscr{C}^N$ and their extensions.

The Only Previous Answer

Laurent (2003)[1] showed (something stronger than) that, when $N \geq 4$, then

$$\begin{bmatrix} 1 & -\frac{1}{N-1} & \cdots & -\frac{1}{N-1} \\ -\frac{1}{N-1} & 1 & \cdots & -\frac{1}{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{N-1} & -\frac{1}{N-1} & \cdots & 1 \end{bmatrix} \in \mathscr{E}_4^N \setminus \mathscr{C}^N.$$

[1] Slightly previous work of Grigoriev (2001) and slightly later work of Schoenebeck (2008) did very similar things.

The Only Previous Answer

Laurent (2003)[1] showed (something stronger than) that, when $N \geq 4$, then

$$\begin{bmatrix} 1 & -\frac{1}{N-1} & \cdots & -\frac{1}{N-1} \\ -\frac{1}{N-1} & 1 & \cdots & -\frac{1}{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{N-1} & -\frac{1}{N-1} & \cdots & 1 \end{bmatrix} \in \mathscr{E}_4^N \setminus \mathscr{C}^N.$$

This is the **Gram matrix of the simplex ETF**!

So, what about other (real) ETFs?

[1] Slightly previous work of Grigoriev (2001) and slightly later work of Schoenebeck (2008) did very similar things.

ETFs: A Very Brief Review

An *equiangular tight frame (ETF)* is a set of N vectors $\boldsymbol{v}_1, \dots, \boldsymbol{v}_N$ in $\mathbb{R}^r$ or $\mathbb{C}^r$, such that:

- They are unit norm: $\|\boldsymbol{v}_i\|_2 = 1$.
- They are equiangular: $|\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle| = \mu$ for all $i \neq j$.
- They form a tight frame: $\sum_{i=1}^{N} \boldsymbol{v}_i \boldsymbol{v}_i^* = \frac{N}{r} \boldsymbol{I}_r$.

Most important high-level intuition: broadly speaking, ETFs are **rigid**, **combinatorial**, **rare** objects.

Degree 4 Extensibility of ETF Gram Matrices

Theorem. (Bandeira, K. '18) The Gram matrix of an ETF of N vectors in $\mathbb{R}^r$ is degree 4 extensible if and only if

$$N < \frac{r(r+1)}{2}.$$

If so,

$$Y_{(ij)(k\ell)} := \frac{\frac{r(r-1)}{2}}{\frac{r(r+1)}{2} - N}(X_{ij}X_{k\ell} + X_{ik}X_{j\ell} + X_{i\ell}X_{jk}) - \frac{r^2(1-\frac{1}{N})}{\frac{r(r+1)}{2} - N}\sum_{m=1}^{N} X_{im}X_{jm}X_{km}X_{\ell m}$$

gives a degree 4 extension.

Part 2: Applications to ETFs

Sum-of-squares optimization and the sparsity structure of equiangular tight frames (2019)

Digging Into the Degree 4 Witness

The $\boldsymbol{Y}$ from our theorem has two eigenspaces:

$$\boldsymbol{Y} = \text{vec}(\boldsymbol{X})\text{vec}(\boldsymbol{X})^\top + \lambda \boldsymbol{P},$$

for $\boldsymbol{P}$ an orthogonal projector.

The theorem's proof includes a formula for $\boldsymbol{P} = \boldsymbol{P}_{\boldsymbol{X}} \succeq \boldsymbol{0}$, whose quadratic form we can test to get inequalities on $\boldsymbol{X}$:

$$\langle \boldsymbol{A}, \boldsymbol{P}_{\boldsymbol{X}}[\boldsymbol{A}] \rangle \geq 0.$$

Where Does $\boldsymbol{P}$ Project To?

> **Definition.** For $K \subset \mathbb{R}^d$ a closed convex set and $\boldsymbol{x} \in K$,
>
> $$\mathrm{pert}_K(\boldsymbol{x}) := \{\boldsymbol{y} : \boldsymbol{x} \pm t\boldsymbol{y} \in K \text{ for all } t \text{ suff. small}\}.$$
>
> (Or, the affine hull of the smallest face containing $\boldsymbol{x}$.)

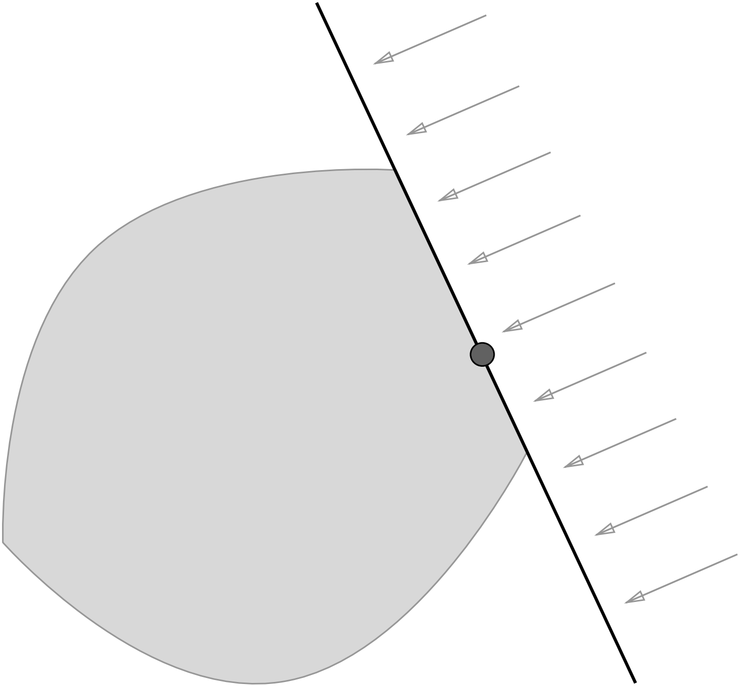
Then, $\boldsymbol{P}$ projects to $\mathrm{vec}(\mathrm{pert}_{\mathscr{E}_2^N}(\boldsymbol{X}))$.

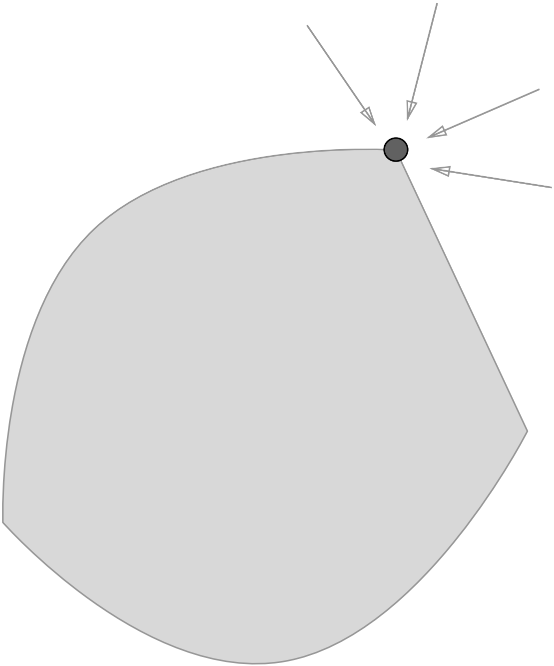
Where Does $\boldsymbol{P}$ Project To?

> **Definition.** For $K \subset \mathbb{R}^d$ a closed convex set and $\boldsymbol{x} \in K$,
>
> $$\mathrm{pert}_K(\boldsymbol{x}) := \{\boldsymbol{y} : \boldsymbol{x} \pm t\boldsymbol{y} \in K \text{ for all } t \text{ suff. small}\}.$$
>
> (Or, the affine hull of the smallest face containing $\boldsymbol{x}$.)

Then, $\boldsymbol{P}$ projects to $\mathrm{vec}(\mathrm{pert}_{\mathscr{E}_2^N}(\boldsymbol{X}))$.

> **Remark.** The same method gives a formula for $\boldsymbol{P}$ when $\boldsymbol{X}$ is the Gram matrix of a **complex** ETF, too, if we use
>
> $$\widetilde{\mathscr{E}}_2^N := \{\boldsymbol{X} \in \mathbb{C}^{N\times N}_{\mathrm{herm}} : \boldsymbol{X} \succeq \boldsymbol{0}, X_{ii} = 1\}.$$
>
> Everything from now on applies to that version.





The Master Matrix Inequality

$\boldsymbol{P} \succeq \boldsymbol{0}$ turns out to be equivalent to this.

> **Corollary.** Let $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N \in \mathbb{C}^r$ for $r > 1$ form an ETF. Let $\boldsymbol{R} \in \mathbb{R}^{r \times r}_{\text{sym}}$ be defined by
>
> $$R_{k\ell} = \sum_{i=1}^{N} |(\boldsymbol{v}_i)_k|^2 |(\boldsymbol{v}_i)_\ell|^2.$$
>
> Then,
>
> $$\boldsymbol{R} \preceq \frac{1 - \frac{1}{r}}{1 - \frac{1}{N}} \boldsymbol{I}_r + \frac{\frac{N}{r} - 1}{r(1 - \frac{1}{N})} \boldsymbol{1}\boldsymbol{1}^\top.$$

For $\boldsymbol{V}$ the ETF's "short fat matrix," the entries of $\boldsymbol{R}$ measure how much magnitudes of $\boldsymbol{V}$'s rows correlate.

Controlling Sparsity

Test the diagonal entries of the master matrix inequality.

> **Corollary.** Let $\boldsymbol{V} \in \mathbb{C}^{r\times N}$ be the short fat matrix of an ETF for $r > 1$. Let $\boldsymbol{w}$ be in the row space of $\boldsymbol{V}$. Then,
>
> $$\|\boldsymbol{w}\|_0 \geq \frac{N}{1 + \frac{(r-1)^2}{N-1}}.$$
>
> *Proof.* The master inequality gives $\|\boldsymbol{w}\|_4^4 \leq C\|\boldsymbol{w}\|_2^4$. By Cauchy-Schwarz, $\|\boldsymbol{w}\|_0 \geq \|\boldsymbol{w}\|_2^4 / \|\boldsymbol{w}\|_4^4 \geq C$.

Controlling Spark

We can also control sparsity of vectors perpendicular to the row space.

> **Corollary.** Let $\boldsymbol{V} \in \mathbb{C}^{r\times N}$ be the short fat matrix of an ETF for $r > 1$. Then,
>
> $$\operatorname{spark}(\boldsymbol{V}) := \min_{\substack{\boldsymbol{w}\in\mathbb{C}^N\setminus\{\boldsymbol{0}\} \\ \boldsymbol{V}\boldsymbol{w}=\boldsymbol{0}}} \|\boldsymbol{w}\|_0 \geq \frac{N}{1+\frac{(N-r-1)^2}{N-1}}.$$
>
> *Proof.* Previous slide on Naimark complement of $\boldsymbol{V}$.

Controlling Sparsity Pattern Overlap

Now, test the 2×2 minors of the master matrix inequality.

> **Corollary.** Let $\boldsymbol{V} \in \mathbb{C}^{r \times N}$ be the short fat matrix of an ETF for $r > 1$. Define
>
> $$D := \frac{N}{r^2}\left(1 + \frac{(r-1)^2}{N-1}\right), \quad E := \frac{\frac{N}{r} - 1}{r(1 - \frac{1}{N})}.$$
>
> Then, for any two distinct rows $\boldsymbol{w}, \boldsymbol{w}'$ of $\boldsymbol{V}$,
>
> $$\left| \sum_{i=1}^{N} |w_i|^2 |w_i'|^2 - E \right|^2 \leq \left(D - \|\boldsymbol{w}\|_4^4\right)\left(D - \|\boldsymbol{w}'\|_4^4\right).$$
>
> *Proof.* Determinant monotonicity on 2×2 minors.

Measuring Tightness

When do we know how good these inequalities are?

> **Proposition.** For any *Steiner ETF*[2] built from a *finite projective plane*, we have equality in the master matrix inequality:
>
> $$\boldsymbol{R} = \frac{1-\frac{1}{r}}{1-\frac{1}{N}}\boldsymbol{I}_r + \frac{\frac{N}{r}-1}{r(1-\frac{1}{N})}\mathbf{1}\mathbf{1}^\top.$$
>
> More generally, the dimension of the "tight subspace" for a Steiner ETF is the number of Steiner system "points," while r is the number of "lines."

[2] An ETF built from a combinatorial design generalizing the incidence structure of a finite geometry.

My Favorite ETF Open Problem (I'm Biased)

For $\boldsymbol{X} \in \mathbb{R}^{N \times N}_{\text{sym}}$ the Gram matrix of a real ETF, let

$$d(\boldsymbol{X}) = \max\{d \in 2\mathbb{N} : \boldsymbol{X} \in \mathscr{E}^N_d\}.$$

What does this number depend on?

Only (N, r)?

Can higher SOS constructions teach us more about sparsity?

Or about other structure in ETFs?

Thank you!