Connections Between Sum-of-Squares Optimization and Structured Tight Frames

Tim Kunisky

(joint work with Afonso Bandeira)

Courant Institute of Mathematical Sciences

SPIE Wavelets & Sparsity 2019 : August 15, 2019

Unit-Norm Tight Frames

Unit-Norm Tight Frames

Vectors $\boldsymbol{v}_1, \dots, \boldsymbol{v}_N$ such that the Gram matrix $\boldsymbol{X} = (\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle)_{i,j=1}^N$ has:

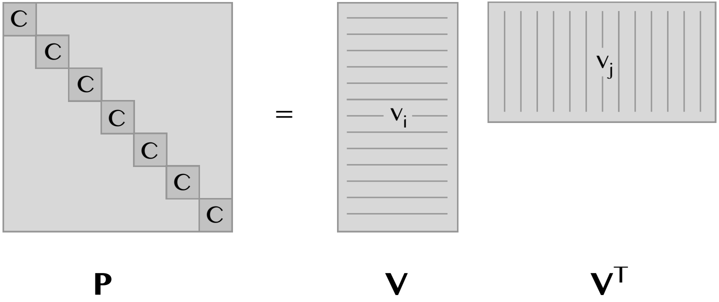
1. $\operatorname{diag}(\boldsymbol{X}) = \mathbf{1}$.
2. $\boldsymbol{X}$ is a constant multiple of a projection matrix.

Unit-Norm Tight Frames

Vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N$ such that the Gram matrix $\boldsymbol{X} = (\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle)_{i,j=1}^N$ has:

1. $\text{diag}(\boldsymbol{X}) = \mathbf{1}$.
2. $\boldsymbol{X}$ is a constant multiple of a projection matrix.

Seen differently, "UNTF-ness" is a property of a **subspace**:

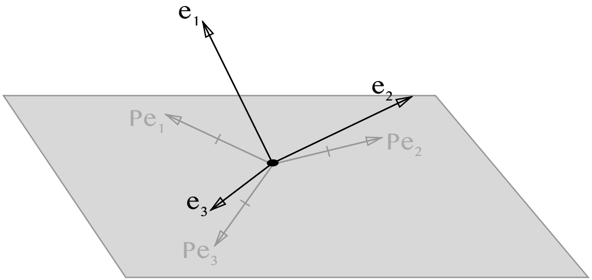


Unit-Norm Tight Frames

Vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N$ such that the Gram matrix $\boldsymbol{X} = (\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle)_{i,j=1}^N$ has:

1. $\operatorname{diag}(\boldsymbol{X}) = \mathbf{1}$.
2. $\boldsymbol{X}$ is a constant multiple of a projection matrix.

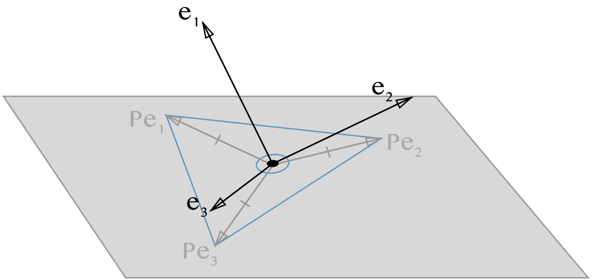
Seen differently, “UNTF-ness” is a property of a **subspace**:



Unit-Norm Tight Frames: Special Kinds

We like UNTFs with only a few different values of P_{ij} (angles or distances among their vectors).

- **Equiangular:** $P_{ij} \in \{\alpha, -\alpha\}$ for all $i \neq j$.
- **Two-distance:** $P_{ij} \in \{\alpha, \beta\}$ for all $i \neq j$.



This Talk:

How to take a very nice UNTF and build a bigger, slightly less nice UNTF.

But mostly, the strange way we found this construction.

Motivation

References (2018-2019):

(1) *A Gramian description of the degree 4 generalized elliptope*

(2) *SoS optimization and the sparsity structure of equiangular tight frames*

(3) *A tight degree 4 SoS lower bound for the Sherrington-Kirkpatrick Hamiltonian*

Relaxations of MaxCut

How to approximate:

$$\max_{\boldsymbol{x}\in\{\pm 1\}^N} \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} \,?$$

Relaxations of MaxCut

How to approximate:

$$\max_{\boldsymbol{x} \in \{\pm 1\}^N} \langle \boldsymbol{A}, \boldsymbol{x}\boldsymbol{x}^\top \rangle \ ?$$

Relaxations of MaxCut

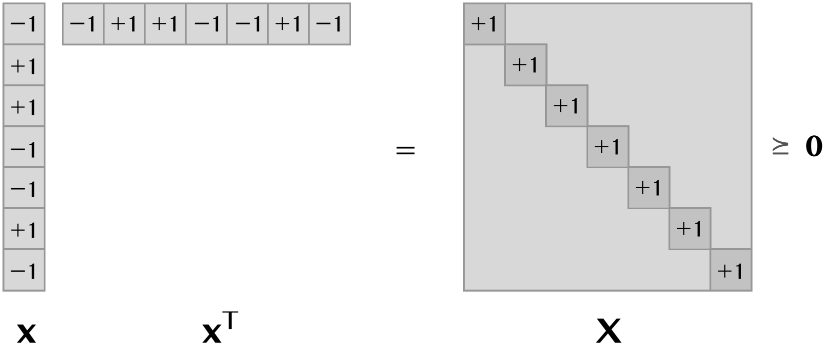
How to approximate:

$$\mathscr{C}^N := \operatorname{conv}\left(\boldsymbol{x}\boldsymbol{x}^\top : \boldsymbol{x} \in \{\pm 1\}^N\right) \ ?$$

Relaxations of MaxCut

How to approximate:

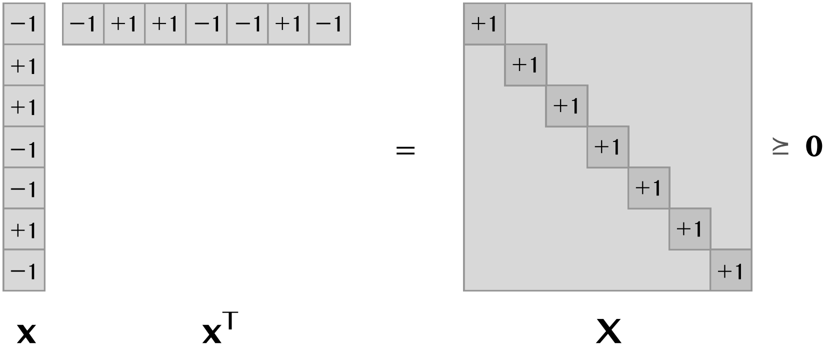
$$\max_{\boldsymbol{x} \in \{\pm 1\}^N} \langle \boldsymbol{A}, \boldsymbol{x}\boldsymbol{x}^\top \rangle \ ?$$



Relaxations of MaxCut

The **degree 2 sum-of-squares (SoS) relaxation**:

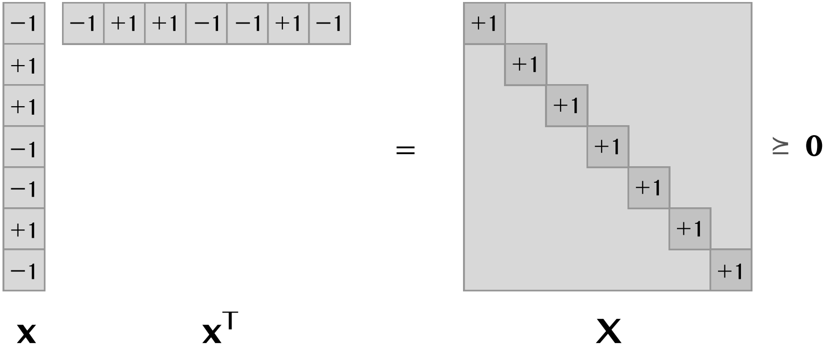
$$\max_{\boldsymbol{x} \in \{\pm 1\}^N} \langle \boldsymbol{A}, \boldsymbol{x}\boldsymbol{x}^\top \rangle \leq \max_{\substack{\boldsymbol{X} \succeq \boldsymbol{0} \\ \mathrm{diag}(\boldsymbol{X}) = \boldsymbol{1}}} \langle \boldsymbol{A}, \boldsymbol{X} \rangle.$$



Relaxations of MaxCut

The **degree 2 sum-of-squares (SoS) relaxation**:

$$\mathscr{E}_2^N := \left\{ \boldsymbol{X} \in \mathbb{R}^{N \times N}_{\mathrm{sym}} : \boldsymbol{X} \succeq \mathbf{0}, \mathrm{diag}(\boldsymbol{X}) = \mathbf{1} \right\}.$$

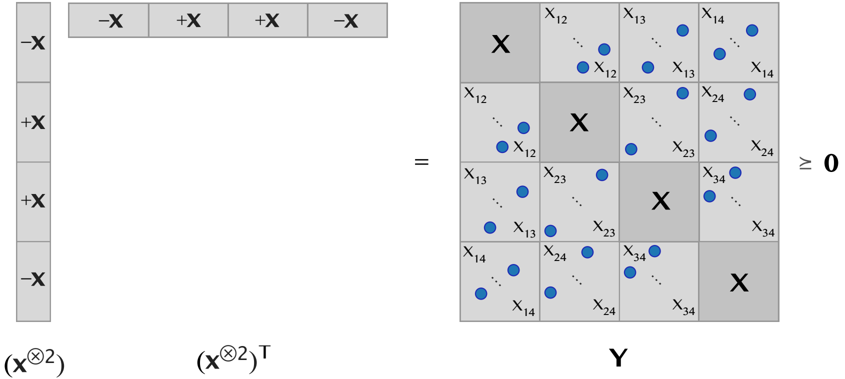


Did You Notice?

The constant multiples of projection matrices in $\mathscr{E}_2^N$ are exactly the Gram matrices of UNTFs.

More on that in a minute...

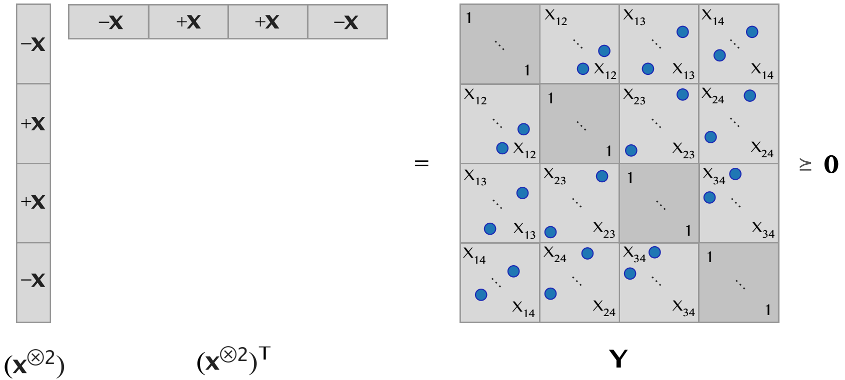
Back to MaxCut...How to Improve?



New constraints:

●: $(x_1x_2)(x_3x_4) =$
$(x_1x_3)(x_2x_4) =$
$(x_1x_4)(x_2x_3)$

Back to MaxCut...How to Improve?



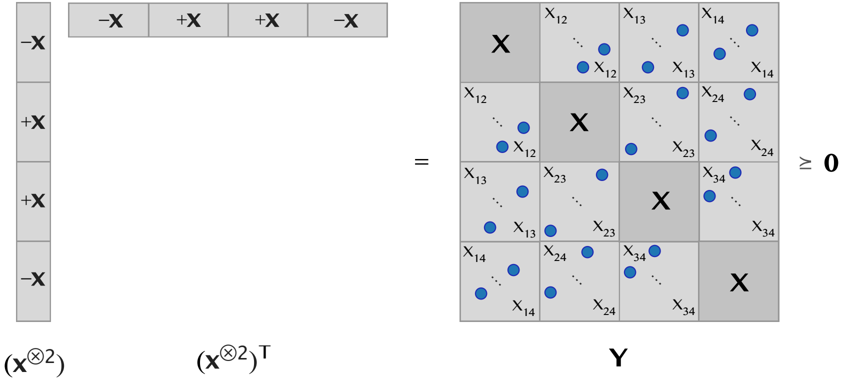
New constraints:

● : $(x_1x_2)(x_3x_4) =$
$(x_1x_3)(x_2x_4) =$
$(x_1x_4)(x_2x_3)$

Back to MaxCut...How to Improve?

The **degree 4 SoS relaxation**:

$$\max_{\boldsymbol{x}\in\{\pm 1\}^N} \langle \boldsymbol{A}, \boldsymbol{x}\boldsymbol{x}^\top \rangle \leq \max_{\boldsymbol{X} \text{ extended by } \boldsymbol{Y}} \langle \boldsymbol{A}, \boldsymbol{X} \rangle.$$



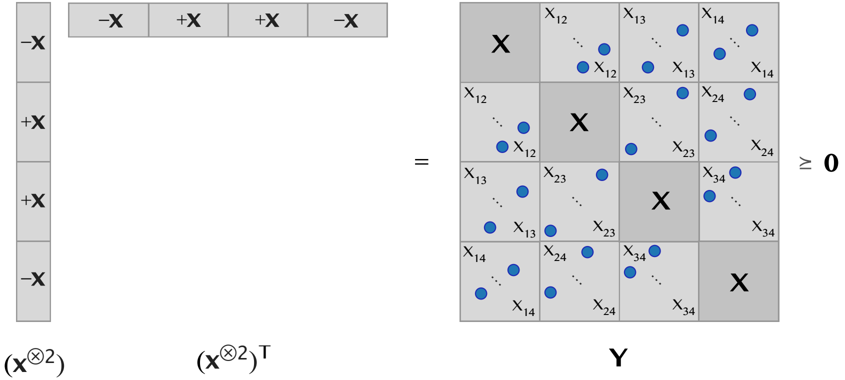
New constraints:

●: $(x_1x_2)(x_3x_4) = (x_1x_3)(x_2x_4) = (x_1x_4)(x_2x_3)$

Back to MaxCut...How to Improve?

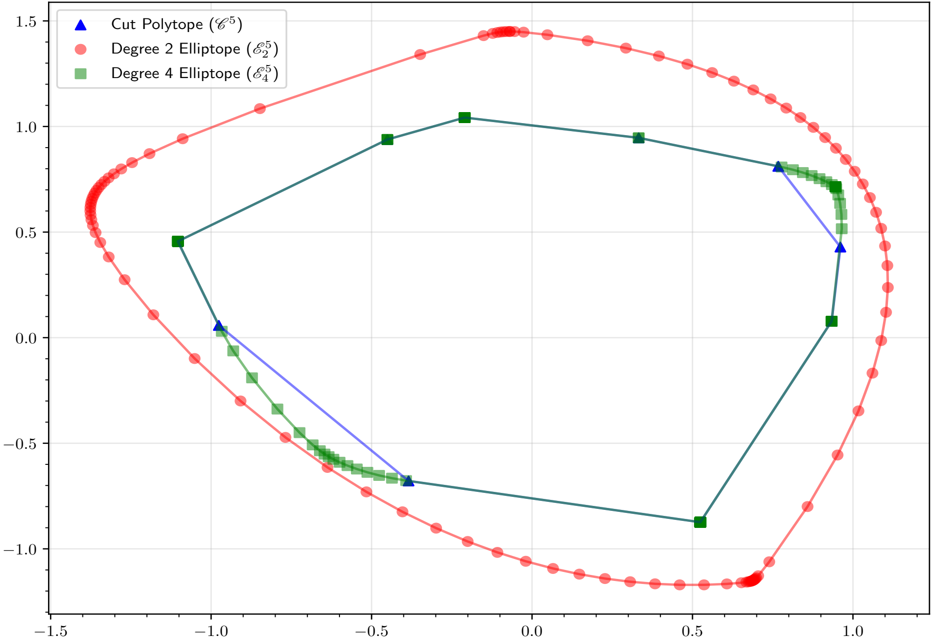
The **degree 4 SoS relaxation**:

$$\mathscr{E}_4^N := \left\{ \boldsymbol{X} \in \mathbb{R}^{N\times N}_{\mathrm{sym}} : \boldsymbol{X} \text{ extendable like this} \right\}.$$



New constraints:

●: $(x_1x_2)(x_3x_4) = (x_1x_3)(x_2x_4) = (x_1x_4)(x_2x_3)$


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

Our Specific Problem

How to approximate:

$$M = \max_{\boldsymbol{x} \in \{\pm 1\}^N} \boldsymbol{x}^\top \boldsymbol{A} \boldsymbol{x} \text{ for } A_{ij} \sim_{\text{iid}} \text{Normal}(0, 2) \text{ ?}$$

Our Specific Problem

How to approximate:

$$M = \max_{\boldsymbol{x} \in \{\pm 1\}^N} \boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x} \text{ for } \boldsymbol{W} = \frac{\boldsymbol{A} + \boldsymbol{A}^\top}{2} \sim \mathrm{GOE}(N) \text{ ?}$$

Our Specific Problem

How to approximate:

$$M = \max_{\boldsymbol{x} \in \{\pm 1\}^N} \boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x} \text{ for } \boldsymbol{W} = \frac{\boldsymbol{A} + \boldsymbol{A}^\top}{2} \sim \mathrm{GOE}(N) \text{ ?}$$

$M = (1.5^+ + o(1))N^{3/2}$ [Parisi 1979].

Our Specific Problem

How to approximate:

$$M = \max_{\boldsymbol{x} \in \{\pm 1\}^N} \boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x} \text{ for } \boldsymbol{W} = \frac{\boldsymbol{A} + \boldsymbol{A}^\top}{2} \sim \mathrm{GOE}(N) \text{ ?}$$

$M = (1.5^+ + o(1))N^{3/2}$ [Parisi 1979]. Algorithmic upper bounds: $M \leq \cdots$

Bound	Value	Reference
$\lambda_{\max}(\boldsymbol{W}) \cdot \|\boldsymbol{x}\|^2$	$(2 + o(1))N^{3/2}$	[Wigner 1955]

Our Specific Problem

How to approximate:

$$M = \max_{\boldsymbol{x} \in \{\pm 1\}^N} \boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x} \text{ for } \boldsymbol{W} = \frac{\boldsymbol{A} + \boldsymbol{A}^\top}{2} \sim \mathrm{GOE}(N) \; ?$$

$M = (1.5^+ + o(1))N^{3/2}$ [Parisi 1979]. Algorithmic upper bounds: $M \leq \cdots$

Bound	Value	Reference
$\lambda_{\max}(\boldsymbol{W}) \cdot \|\boldsymbol{x}\|^2$	$(2 + o(1))N^{3/2}$	[Wigner 1955]
Degree 2 SoS	$(2 + o(1))N^{3/2}$	[Montanari, Sen 2016]

Our Specific Problem

How to approximate:

$$M = \max_{\boldsymbol{x} \in \{\pm 1\}^N} \boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x} \text{ for } \boldsymbol{W} = \frac{\boldsymbol{A} + \boldsymbol{A}^\top}{2} \sim \mathrm{GOE}(N) \text{ ?}$$

$M = (1.5^+ + o(1))N^{3/2}$ [Parisi 1979]. Algorithmic upper bounds: $M \leq \cdots$

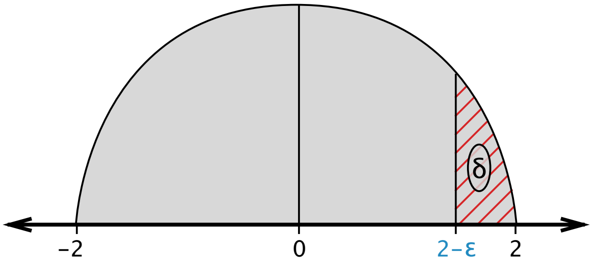
Bound	Value	Reference
$\lambda_{\max}(\boldsymbol{W}) \cdot \|\boldsymbol{x}\|^2$	$(2 + o(1))N^{3/2}$	[Wigner 1955]
Degree 2 SoS	$(2 + o(1))N^{3/2}$	[Montanari, Sen 2016]
Degree 4 SoS	$(2 + o(1))N^{3/2}$	[Bandeira, K. 2019] [Raghavendra et al 2019]

The Degree 2 SoS Construction

[Montanari, Sen 2016] *SDPs on sparse random graphs and their application...*

The Degree 2 SoS Construction

Take $\boldsymbol{X} \approx \delta^{-1}\boldsymbol{P}$ for $\boldsymbol{P}$ the projection matrix to the top δN eigenvectors of $\boldsymbol{W}$.

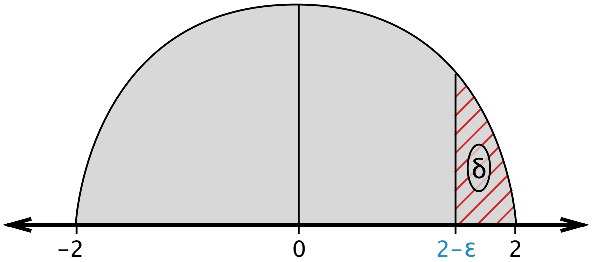


[Montanari, Sen 2016] *SDPs on sparse random graphs and their application...*

The Degree 2 SoS Construction

Take $\boldsymbol{X} \approx \delta^{-1} \boldsymbol{P}$ for $\boldsymbol{P}$ the projection matrix to the top δN eigenvectors of $\boldsymbol{W}$.

Observation: $\boldsymbol{X}$ is "nearly" the Gram matrix of a UNTF!



[Montanari, Sen 2016] *SDPs on sparse random graphs and their application...*

So We Wondered...

What do degree 4 extensions of Gram matrices of UNTFs look like?

The Case of ETFs

Theorem. 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 an extension.

[Bandeira, K. 2018] *A Gramian description of the degree 4 generalized elliptope*
[Bandeira, K. 2019] *SoS optimization and sparsity structure of equiangular tight frames*

Resolving the Gaussian Case

Pretend $\boldsymbol{X} \approx \delta^{-1}\boldsymbol{P}$ is an ETF, and use the same formula. With a small correction, this shows $\boldsymbol{X} \in \mathscr{E}_4^N$, so we get...

> **Theorem.** When $\boldsymbol{W} \sim \mathrm{GOE}(N)$,
>
> $$\max_{\boldsymbol{X} \in \mathscr{E}_4^N} \langle \boldsymbol{W}, \boldsymbol{X} \rangle = (2 + o(1))N^{3/2}$$
>
> with high probability.

[Bandeira, K. 2019] *A tight degree 4 SoS lower bound for the SK Hamiltonian*

New Structured UNTFs

Reference (SPIE 2019):
Connections between SoS optimization and structured tight frames

So...Where Did This Come From?

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

Spectral Constraints

$\boldsymbol{Y}$ is complicated entrywise, but simple spectrally:

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

$\boldsymbol{P}$ is the projector to a subspace of $\mathbb{R}^{N^2}$ where all of the eigenvectors of $\boldsymbol{Y} - \text{vec}(\boldsymbol{X})\text{vec}(\boldsymbol{X})^\top$ *must* lie (for any degree 4 extension).

[Bandeira, K. 2018] *A Gramian description of the degree 4 generalized elliptope*

Spectral Constraints

$\boldsymbol{Y}$ is complicated entrywise, but simple spectrally:

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

$\boldsymbol{P}$ is the projector to a subspace of $\mathbb{R}^{N^2}$ where all of the eigenvectors of $\boldsymbol{Y} - \mathrm{vec}(\boldsymbol{X})\mathrm{vec}(\boldsymbol{X})^\top$ *must* lie (for any degree 4 extension).

(Namely, $\boldsymbol{P}$ projects to the vectorized *perturbation subspace* of $\boldsymbol{X}$ in $\mathscr{E}_2^N$.)

[Bandeira, K. 2018] *A Gramian description of the degree 4 generalized elliptope*

Hang On...

$\boldsymbol{P}$ is a projection matrix. What is its diagonal?

Hang On...

$\boldsymbol{P}$ is a projection matrix. What is its diagonal?

$$\begin{aligned}\operatorname{diag}(\boldsymbol{P}) &= \lambda^{-1}\operatorname{diag}(\boldsymbol{Y} - \operatorname{vec}(\boldsymbol{X})\operatorname{vec}(\boldsymbol{X})^\top) \\ &= \lambda^{-1}(\mathbf{1} - \operatorname{vec}(\boldsymbol{X}) \circ \operatorname{vec}(\boldsymbol{X})).\end{aligned}$$

Hang On...

$\boldsymbol{P}$ is a projection matrix. What is its diagonal?

$$\begin{aligned}\operatorname{diag}(\boldsymbol{P}) &= \lambda^{-1}\operatorname{diag}(\boldsymbol{Y} - \operatorname{vec}(\boldsymbol{X})\operatorname{vec}(\boldsymbol{X})^\top) \\ &= \lambda^{-1}(\mathbf{1} - \operatorname{vec}(\boldsymbol{X}) \circ \operatorname{vec}(\boldsymbol{X})).\end{aligned}$$

But $\boldsymbol{X}$ is the Gram matrix of an ETF:

$$(\operatorname{vec}(\boldsymbol{X}) \circ \operatorname{vec}(\boldsymbol{X}))_{(ij)} = X_{ij}^2 = \begin{cases} 1 & : \quad i = j, \\ \alpha^2 & : \quad i \neq j. \end{cases}$$

Hang On...

$\boldsymbol{P}$ is a projection matrix. What is its diagonal?

$$\begin{aligned}\operatorname{diag}(\boldsymbol{P}) &= \lambda^{-1}\operatorname{diag}(\boldsymbol{Y} - \operatorname{vec}(\boldsymbol{X})\operatorname{vec}(\boldsymbol{X})^\top) \\ &= \lambda^{-1}(\mathbf{1} - \operatorname{vec}(\boldsymbol{X}) \circ \operatorname{vec}(\boldsymbol{X})).\end{aligned}$$

But $\boldsymbol{X}$ is the Gram matrix of an ETF:

$$(\operatorname{vec}(\boldsymbol{X}) \circ \operatorname{vec}(\boldsymbol{X}))_{(ij)} = X_{ij}^2 = \begin{cases} 1 & : \quad i = j, \\ \alpha^2 & : \quad i \neq j. \end{cases}$$

> **Corollary.** Let $\boldsymbol{X}$ be the Gram matrix of a non-maximal ETF. Then, the minor of $\boldsymbol{P}$ indexed by (ij) with $i < j$ is a "UNTF projector." Its UNTF consists of $\frac{r(r+1)}{2} - N$ vectors in $\mathbb{R}^{N(N-1)/2}$.

Example: Simplex ETFs $\rightsquigarrow$ Johnson TDTFs

The simplest ETFs:

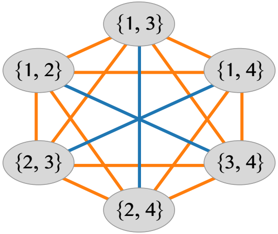
$$\frac{N}{N-1}\Big(\underbrace{\boldsymbol{I}_N - \frac{1}{N}\mathbf{1}\mathbf{1}^\top}_{\text{projector to } \mathbf{1}^\perp}\Big) = \begin{bmatrix} 1 & -\frac{1}{N-1} & \cdots & -\frac{1}{N-1} \\ -\frac{1}{N-1} & 1 & \cdots & -\frac{1}{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{N-1} & -\frac{1}{N-1} & \cdots & 1 \end{bmatrix}.$$

Example: Simplex ETFs ⇝ Johnson TDTFs

The simplest ETFs:

$$\frac{N}{N-1}\Big(\underbrace{I_N - \frac{1}{N}\mathbf{1}\mathbf{1}^\top}_{\text{projector to } \mathbf{1}^\perp}\Big) = \begin{bmatrix} 1 & -\frac{1}{N-1} & \cdots & -\frac{1}{N-1} \\ -\frac{1}{N-1} & 1 & \cdots & -\frac{1}{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{N-1} & -\frac{1}{N-1} & \cdots & 1 \end{bmatrix}.$$

“Degree 4 lifting” ⇝ **two-distance UNTF** of Johnson graph:



C	α	α	β	α	α
α	C	α	α	β	α
α	α	C	α	α	β
β	α	α	C	α	α
α	β	α	α	C	α
α	α	β	α	α	C

Open Problems

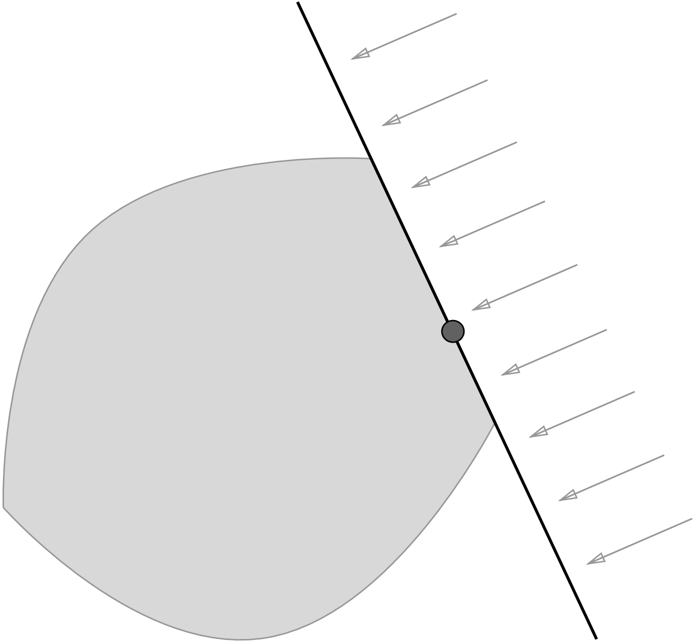
- Counting distances in degree 4 liftings
- Structure of “entry graphs” of resulting few-distance UNTFs
- Relation to line graph construction
- Generalization to higher degree SoS

Thank you!

Where Does $\boldsymbol{P}$ Project?

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

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



(Back)