Equivariant semidefinite lifts and sum of squares hierarchies

Pablo A. Parrilo

Laboratory for Information and Decision Systems
Electrical Engineering and Computer Science
Massachusetts Institute of Technology

Joint work with Hamza Fawzi and James Saunderson

Cargese 2014
Question: representability of convex sets

Existence and efficiency:

- When is a convex set representable by conic optimization?
- How to quantify the number of additional variables that are needed?

Given a convex set $C$, is it possible to represent it as

$$C = \pi(K \cap L)$$

where $K$ is a cone, $L$ is an affine subspace, and $\pi$ is a linear map?
SDP representations

In full generality, difficult to understand (but we’re making progress!)

- Characterized by a Yannakakis-like theorem
- Set $C$ may have many “inequivalent” PSD lifts
- For nonpolyhedral sets, continuity considerations arise
- Constructive techniques (e.g., SOS) have additional properties

Our starting point: “symmetric” (equivariant) lifts.
Lifts and symmetries

A natural requirement: lift should be “symmetric”.

- Informally: lift “respects” the symmetries of the convex body $C$.
- Basic idea: symmetries of $C$ “lift” to symmetries upstairs in $K \cap L$.

(Formal definition will follow, examples first!).

Lifts and symmetries

A natural requirement: lift should be “symmetric”.

- Informally: lift “respects” the symmetries of the convex body $C$.
- Basic idea: symmetries of $C$ “lift” to symmetries upstairs in $K \cap L$.

(Formal definition will follow, examples first!).

Long history: Yannakakis’91, Kaibel-Pashkovich-Theis’10 (”symmetry matters”), Lee-Raghavendra-Steurer-Tan’14 ...
Lifts and symmetries

A natural requirement: lift should be “symmetric”.

- Informally: lift “respects” the symmetries of the convex body $C$.
- Basic idea: symmetries of $C$ “lift” to symmetries upstairs in $K \cap L$

(Formal definition will follow, examples first!).

Examples and non-examples (I)

An equivariant psd lift of the square $[-1, 1]^2$:

$$[-1, 1]^2 = \left\{ (x_1, x_2) \in \mathbb{R}^2 : \exists u \in \mathbb{R} \begin{bmatrix} 1 & x_1 & x_2 \\ x_1 & 1 & u \\ x_2 & u & 1 \end{bmatrix} \succeq 0 \right\}. \quad (1)$$

Square as a projection of the elliptope:
Examples and non-examples (II)

A 3-dimensional hyperboloid:

\[ H = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1, x_2, x_3 \geq 0 \text{ and } x_1 x_2 x_3 \geq 1\}. \]

A non-equivariant psd lift of $H$ of size 6:

\[ H = \{(x_1, x_2, x_3) : \exists y, z \geq 0 \quad x_1 x_2 \geq y^2, \quad x_3 \geq z^2, \quad yz \geq 1\} \]

\[ = \{(x_1, x_2, x_3) : \exists y, z \quad \begin{bmatrix} x_1 & y \\ y & x_2 \end{bmatrix} \succeq 0, \quad \begin{bmatrix} x_3 & z \\ z & 1 \end{bmatrix} \succeq 0, \quad \begin{bmatrix} y & 1 \\ 1 & z \end{bmatrix} \succeq 0\}. \]

$H$ is invariant under permutation of coordinates, but the lift does not respect this symmetry (role of variables is different).
Examples and non-examples (II)

A 3-dimensional hyperboloid:

\[ H = \{ (x_1, x_2, x_3) \in \mathbb{R}^3 : x_1, x_2, x_3 \geq 0 \text{ and } x_1 x_2 x_3 \geq 1 \}. \]

A non-equivariant psd lift of \( H \) of size 6:

\[ H = \{ (x_1, x_2, x_3) : \exists y, z \geq 0 \quad x_1 x_2 \geq y^2, \quad x_3 \geq z^2, \quad yz \geq 1 \} \]

\[ = \{ (x_1, x_2, x_3) : \exists y, z \quad \begin{bmatrix} x_1 & y \\ y & x_2 \end{bmatrix} \succeq 0, \quad \begin{bmatrix} x_3 & z \\ z & 1 \end{bmatrix} \succeq 0, \quad \begin{bmatrix} y & 1 \\ 1 & z \end{bmatrix} \succeq 0 \} . \]

\( H \) is invariant under permutation of coordinates, but the lift does not respect this symmetry (role of variables is different).
Equivariant lifts

Let $P \subset \mathbb{R}^n$ be a polytope invariant under the action of a group $G \subset GL(\mathbb{R}^n)$, with a lift $P = \pi(S^d_+ \cap L)$.

Definition: The lift is $G$-equivariant if there is a group homomorphism $\rho : G \rightarrow GL(\mathbb{R}^d)$ such that:

1. **Subspace $L$ is invariant under conjugation by $\rho$:**
   
   
   
   $Y \in L \quad \implies \quad \rho(T)Y\rho(T)^T \in L \quad \forall T \in G.$

2. **$\rho$ “intertwines” the lift map**

   
   
   
   $\pi(\rho(T)Y\rho(T)^T) = T\pi(Y), \quad \forall T \in G, \forall Y \in S^d_+ \cap L.$
Unlike in the LP case, several slightly different definitions are possible (mainly, affine-equivariance vs. projective-equivariance).

We prefer this one, for a few reasons:

- More natural in affine setting
- Sum of squares hierarchies are intrinsically affine-equivariant
- Consistent with symmetry-reduction techniques for SDP/SOS (e.g., Gatermann-P.’04)
Orbitopes

Special class of convex bodies: orbitopes

\[ C = \{ \text{conv}(g \cdot x_0) : g \in G \}, \]

where \( G \) is a compact group.

Many important examples: hypercubes, hyperspheres, Grassmannians, Birkhoff polytope, permutahedra, parity polytope, cut polytope, ...
Orbitopes

Special class of convex bodies: orbitopes

\[ C = \{ \text{conv}(g \cdot x_0) : g \in G \}, \]

where \( G \) is a compact group.

Many important examples: hypercubes, hyperspheres, Grassmannians, Birkhoff polytope, permutahedra, parity polytope, cut polytope, \ldots

SDP aspects analyzed in Sanyal-Sottile-Sturmfels’11, earlier appearances in Barvinok-Vershik’88, Barvinok-Blekherman’05, etc.
Example: $SO(n)$-orbitope

Consider $SO(n)$, the group of $n \times n$ matrices with determinant one. This is the orbit of $I$ under $O(n)$ action.

Convex hull is of interest in optimization problems involving rotation matrices.

$SO(n)$ orbitope has an SDP representation!

Explicit construction based on the double cover of $SO(n)$ with spin group. (Saunderson-P.-Willsky, arXiv:1403:4914)
Regular orbitopes

Convex hull of a group orbit

\[ C = \{ \text{conv}(g \cdot x_0) : g \in G \}, \]

An orbitope is regular if the stabilizer of a point is the trivial subgroup. Equivalently, a bijection between group elements and extreme points.

E.g., for the symmetric group \( S_n \) (permutahedron), if all entries of \( x_0 \) are distinct, then \( C \) is regular.
A structure theorem for equivariant lifts

Equivariant lifts of orbitopes are particularly nice.

**Why?**: Every *equivariant* SDP lift is of *sum of squares* type.

More formally:

**Theorem [FSP 13]**: Let $P$ be a $G$-regular orbitope, with a $G$-equivariant lift of size $d$. Then for any linear form $\ell$, there exist functions $f_j \in V$ such that

$$\ell_{\max} - \ell(x) = \sum_j f_j(x)^2 \quad \forall x \in X$$

where $X = \text{ext}(P)$, and $V$ is a $G$-invariant subspace of $\mathcal{F}(X)$, where $\mathcal{F}(X)$ is the space of real-valued functions on $X$. 
A structure theorem for equivariant lifts

Equivariant lifts of orbitopes are particularly nice.

**Why?** Every *equivariant* SDP lift is of *sum of squares* type.

More formally:

**Theorem [FSP 13]:** Let $P$ be a $G$-regular orbitope, with a $G$-equivariant lift of size $d$. Then for any linear form $\ell$, there exist functions $f_j \in V$ such that

$$\ell_{\max} - \ell(x) = \sum_j f_j(x)^2 \quad \forall x \in X$$

where $X = \text{ext}(P)$, and $V$ is a $G$-invariant subspace of $\mathcal{F}(X)$, where $\mathcal{F}(X)$ is the space of real-valued functions on $X$. 
Factorization theorem

Let $P$ be a polytope, with extreme points $X = \text{ext}(P)$ and a PSD lift $P = \pi(S^d_+ \cap L)$.

Recall the generalization of Yannakakis’ theorem, characterizing the existence of SDP lifts:

**Theorem** [GPT11]: There exists a map $A : X \to S^d_+$, such that for any facet-defining inequality $\ell(x) \leq \ell_{\text{max}}$, there is $B(\ell) \in S^d_+$ with

$$\ell_{\text{max}} - \ell(x) = \langle A(x), B(\ell) \rangle \quad \forall x \in X.$$
Proof sketch

Since orbitope is regular, we can associate a group element \( \iota(x) \in G \) to every extreme point. We have then:

\[
\ell_{\text{max}} - \ell(x) = \langle A(x)B(\ell) \rangle \\
= \langle A(\iota(x) \cdot x_0)B(\ell) \rangle \\
= \langle \rho(\iota(x))A(x_0)\rho(\iota(x))^T B(\ell) \rangle \\
= \text{vec}(\rho(\iota(x)))^T (A(x_0) \otimes B(\ell)) \text{vec}(\rho(\iota(x)))
\]

and \( \rho(\iota(x)) \) defines a \( G \)-invariant subspace of functions on \( X \).
Equivariant lifts are of SOS-type

**Theorem [FSP 13]**: Let $P$ be a $G$-regular orbitope, with a $G$-equivariant lift of size $d$. Then for any linear form $\ell$, there exist functions $f_j \in V$ such that

$$\ell_{\text{max}} - \ell(x) = \sum_j f_j(x)^2 \quad \forall x \in X$$

where $X = \text{ext}(P)$, and $V$ is a $G$-invariant subspace of $\mathcal{F}(X)$.

Why is this useful?

Can use **representation theory** to understand invariant subspaces of $\mathcal{F}(X)$ (isotypic decomposition).

For polytopes, these are finite-dimensional subspaces of polynomials.
Regular polygons as regular orbitopes

A regular polygon in the plane. Invariant under dihedral group (rotations/flips).

Functions on vertices $X$ can be represented by

$$V = V_0 \oplus V_1 \oplus \cdots \oplus V_d$$

where $V_k$ are subspaces of polynomials.

For the case of the square, these invariant subspaces are

$$\{1\}, \{x, y\}, \{xy\}$$

i.e.,

$$\langle (1, 1, 1, 1) \rangle, \quad \langle (1, 1, -1, -1), (1, -1, 1, -1) \rangle, \quad \langle (1, -1, -1, 1) \rangle,$$
Invariant subspaces

Every invariant subspace of $\mathcal{F}(X)$ is a sum of (possibly many) of the $V_i$.

Thus, the size of every equivariant lift of an orbitope corresponds to the sum of dimensions of the subspaces $V_i$ that appear in $V$.

Understanding which $V_i$ can (or cannot) appear in an SOS representation, will allow us to produce (or bound) equivariant representations.

Example: For the square, this argument easily yields that no equivariant lift of size 2 can exist.
Regular polygons

For instance, for the regular hexagon we have the invariant subspaces

\[ \{ \cos t, \sin t \}, \ \{ \cos 2t, \sin 2t \}, \ \{ \cos 3t, \sin 3t \}. \]

Picking the first two, we obtain the SDP lift:

\[
\begin{bmatrix}
1 & x & y & t \\
x & (1 + r)/2 & s/2 & r \\
y & s/2 & (1 - r)/2 & -s \\
t & r & -s & 1
\end{bmatrix} \succeq 0
\]

In general, carefully choosing the subspaces \( V_i \) can yield equivariant lifts that are exponentially better than “naive” SOS. (Fawzi-Saunderson-P.’14, Hamza’s talk later in the week?)
Can use *representation theory* to understand invariant subspaces of $\mathcal{F}(X)$. For polytopes, these are finite-dimensional subspaces of polynomials. Computing their dimensions, we obtain lower bounds on symmetric representations.

Next, two examples of important polytopes in combinatorial optimization, and a nonpolyhedral orbitope.
Parity polytope

The \textit{parity polytope} \( \text{PAR}_n \) is the convex hull of all points \( x \in \{-1, 1\}^n \) that have an even number of \(-1\).

**Theorem [FSP13]:** Let \( \text{PAR}_n \) be the parity polytope. Then, any \( \Gamma_{\text{parity}} \)-equivariant psd lift of \( \text{PAR}_n \) must have size \( \geq \left( \left\lceil \frac{n}{4} \right\rceil \right) \).

Remark: Weakening the symmetry requirements (e.g., only permutations, or only even sign-flips), \( \text{PAR}_n \) has polynomial-size LP/SDP lifts.
Cut polytope

The cut polytope is defined as

\[ \text{CUT}_n = \text{conv}(xx^T : x \in \{-1, 1\}^n). \]

Theorem [FSP13]: Any psd lift of \( \text{CUT}_n \) that is equivariant with respect to the cube (hyperoctahedral) group must have size \( \geq \binom{n}{\lceil n/4 \rceil} \).

Related work in:
Example: $\text{SO}(n)$-orbitope

Recall the $\text{SO}(n)$ orbitope (convex hull of rotation matrices).

Diagonal slice is the parity polytope $\text{PAR}_n$, and can show that we inherit its lower bounds.

As a consequence, the spin-based construction (which is equivariant, and exponential-sized) is optimal!

(Saunderson-P.-Willsky, arXiv:1403:4914)
Summary

- Equivariant lifts of regular orbitopes can be understood
- Structure theorem: all equivariant lifts are of SOS type
- Lower bounds from representation theory

If you want to know more:


Thanks for your attention!
Summary

- Equivariant lifts of regular orbitopes can be understood
- Structure theorem: all equivariant lifts are of SOS type
- Lower bounds from representation theory

If you want to know more:


Thanks for your attention!
Summary

- Equivariant lifts of regular orbitopes can be understood
- Structure theorem: all equivariant lifts are of SOS type
- Lower bounds from representation theory

If you want to know more:


Thanks for your attention!