6.256 PROJECT: CONVEXITY AND SUM OF SQUARES

YUZHOU GU

1. Introduction

Let $H_{n,k}$ be the space of homogeneous real polynomials (forms) in n variables and of degree k. Several convex subcones of $H_{n,2d}$ have received considerable study:

- nonnegative forms
$$P_{n,2d} := \{p \in H_{n,2d} : p(x) \geq 0 \forall x \in \mathbb{R}^n\},$$
- sum-of-squares (SOS) forms
$$\Sigma_{n,2d} := \{p \in H_{n,2d} : p = \sum_{i \in I} q_i^2 \text{ where } q_i \in H_{n,d}\},$$
- convex forms
$$C_{n,2d} := \{p(x) \in H_{n,2d} : \nabla^2 p(x) \succeq 0 \forall x \in \mathbb{R}^n\},$$
- SOS-convex forms,
$$\Sigma C_{n,2d} := \{p \in H_{n,2d} : y^\top \nabla^2 p(x) y \text{ is SOS}\}.$$

Known results on relationship between these cones are summarized as following.

- Trivial: $\Sigma_{n,2d} \subseteq P_{n,2d}$, $C_{n,2d} \subseteq P_{n,2d}$, $\Sigma C_{n,2d} \subseteq C_{n,2d}$;
- Hilbert [Hil88]: $\Sigma_{n,2d} = P_{n,2d}$ if and only if $n = 2$ or $2d = 2$ or $(n, 2d) = (3, 4)$;
- Helton-Nie [HN10]: $\Sigma C_{n,2d} \subseteq \Sigma_{n,2d}$;
- Ahmadi-Parrilo [AP13]: $\Sigma C_{n,2d} = C_{n,2d}$ if and only if $n = 2$ or $2d = 2$ or $(n, 2d) = (3, 4)$;
- Trivial: $\Sigma_{n,2d} \subseteq C_{n,2d}$ if and only if $2d = 2$;
- Blekherman [Ble09]: For fixed $2d \geq 4$, for n large enough, $C_{n,2d} \not\subseteq \Sigma_{n,2d}$;
- El Khadir [EK20]: $C_{4,4} \subseteq \Sigma_{4,4}$.

In this project I studied problems related to convex forms. Our new result is a tight generalized Cauchy-Schwarz (GCS) inequality for $d = 5$ (Theorem 2). GCS inequalities were introduced in El Khadir [EK20] and its low degree cases were used in the proof of $C_{4,4} \subseteq \Sigma_{4,4}$. I do not know of any application of higher degree GCS inequalities, but they are interesting on their own. In Section 2 we discuss GCS inequalities. We also explain why our method for $d = 5$ cannot be generalized to odd $d \geq 7$ (Theorem 3).

The other sections of this report are a summary of two papers I read during this project. In Section 3, we discuss El Khadir [EK20]'s proof that convex quaternary quartics are SOS, which is one motivation for this project.

During this project, Prof. Parrilo pointed out to me that Saunderson [Sau21] found an explicit example of a convex but non-SOS form. We discuss Saunderson's construction in Section 4.

2. Generalized Cauchy-Schwarz inequalities

GCS inequalities are introduced by El Khadir [EK20] as a tool to prove that convex quaternary quartics are SOS. Recall the Cauchy-Schwarz inequality, which can be stated as

$$Q(x,y) \leq \sqrt{Q(x,x)Q(y,y)}$$

for all positive semidefinite quadratic forms Q. Generalized Cauchy-Schwarz generalizes the usual Cauchy-Schwarz inequality to higher degree forms. Given $p \in H_{n,2d}$, we can define a binary form $Q_p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ as

$$Q_p(x,y) = \frac{1}{(2d)!}\partial_x^d \partial_y^d p.$$

(In [EK20] Q_p is defined via symmetric tensors.) It is homogeneous in (x,y) of degree (d,d) and satisfies the property that $Q_p(x,x) = p(x)$. So when $\deg p = 2$, we get quadratic forms.

Theorem 1 ([EK20, Theorem 3.1])**.** *For any integer $d \geq 1$, there exists a constant A_d such that*

$$Q_p(x,y) \leq A_d\sqrt{p(x)p(y)} \tag{1}$$

for all $p \in C_{n,2d}$, $x, y \in \mathbb{R}^n$.

Proof. Fix $p \in C_{n,2d}$.

Step I: We prove that (1) holds for all $x, y \in \mathbb{R}^n$ if and only if

$$2Q_p(x,y) \leq A_d(p(x) + p(y)) \tag{2}$$

for all $x, y \in \mathbb{R}^n$. The $\Rightarrow$ implication is by AM-GM. The $\Leftarrow$ implication is by observing that for $x, y \in \mathbb{R}^n$ with $Q_p(x,y) > 0$, we have

$$\frac{\sqrt{p(x)p(y)}}{Q_p(x,y)} = \inf_{\lambda>0} \frac{p(x) + \lambda^2 p(y)}{2\lambda Q_p(x,y)} = \inf_{\lambda>0} \frac{p(x) + p(\lambda^{\frac{1}{d}} y)}{2Q_p(x, \lambda^{\frac{1}{d}} y)}.$$

Step II: If x and y are colinear, then there is nothing to prove ($A_d \geq 1$). Otherwise, x and y span a subspace of dimension 2, and by performing a coordinate change, we reduce to the case where $n = 2$, $x = e_1$, $y = e_2$, where e_1 and e_2 are coordinate vectors.

Step III: The set

$$L := \{p \in C_{2,2d} : p(e_1) + p(e_2) = 2\}$$

is compact. So $\sup_{p \in L} Q_p(e_1, e_2)$ is finite, and we can take A_d to be this value. □

Let A_d^* denote the best possible constant such that (1) holds. Let us discuss how to compute A_d^*. As shown in the proof, it suffices to compute

$$A_d^* = \sup_{\substack{p \in C_{2,2d} \\ p(e_1)+p(e_2)=2}} Q_p(e_1, e_2). \tag{3}$$

Let

$$p = \sum_{0 \leq i \leq 2d} p_i x^i y^{2d-i}.$$

d	1	2	3	4	5	6
A_d^*	1.000	1.000	1.000	1.011	1.000	1.061
d	7	8	9	10	11	12
A_d^*	1.000	1.048	1.000	1.153	1.000	1.115

Table 1. Numerical values of A_d^*. First 8 values are from [EK20, Table 1].

Then we can compute that

$$p(e_1) = p_0, \qquad p(e_2) = p_{2d}, \qquad Q_p(e_1, e_2) = \binom{2d}{d}^{-1} p_d.$$

So the objective function is linear, and the constraint $p(e_1) + p(e_2) = 2$ is linear. For the constraint $p \in C_{2,2d}$, we use [AP13, Theorem 5.2], which says that $C_{2,2d} = \Sigma C_{2,2d}$. So $p \in C_{2,2d}$ if and only if $u^\top \nabla p(x) u$ is SOS. So (3) can be solved using semidefinite programming.

$$\begin{aligned} \max \ & \binom{2d}{d}^{-1} p_d \\ \text{s.t.} \quad & p_0 + p_{2d} = 2, \\ & u^\top \nabla^2 p(x) u \text{ is SOS.} \end{aligned}$$

We discuss several simplifications of the program.

- If $p(x, y)$ is a feasible solution, then $\frac{1}{2}(p(x, y) + p(y, x))$ also satisfies the constraints and achieves the same objective value. So we could WLOG assume that $p_i = p_{2d-i}$ for $0 \le i \le d$. Then $p_0 = 1$. So we are left with d variables $p_1, \ldots, p_d$.
- For the case of even d, if $p(x, y)$ is a feasible solution, then $\frac{1}{2}(p(x, y) + p(x, -y))$ also satisfies the constraints and achieve the same objective value. So in this case we could WLOG assume that $p_i = 0$ for odd i, and we are left with $\frac{d}{2}$ variables.
- There are other symmetries of the problem. For example,

  $$f(x, y, u, v) := \begin{bmatrix} u & v \end{bmatrix} \nabla^2 p(x, y) \begin{bmatrix} u \\ v \end{bmatrix}$$

  satisfies

  $$f(x, y, u, v) = f(y, x, v, u) = f(-x, -y, u, v) = f(-x, y, -u, v)$$

  when d is even and

  $$f(x, y, u, v) = f(y, x, v, u) = f(-x, -y, -u, -v)$$

  when d is odd. These symmetries generate a non-trivial group. So we could use the symmetry reduction framework of Gatermann-Parrilo [GP04] to simplify the problem.

Using these simplifications, we can numerically compute A_d^*. I implemented the SDP and its results are summarized in Table 1. MOSEK reports a small duality gap, so these values should be reliable.

El Khadir [EK20] uses symmetry reduction and KKT conditions to algebraically compute A_4^*, and showed that it is equal to the largest root of

$$x^3 - \frac{33}{35}x^2 - \frac{17}{245}x + \frac{13}{42785}.$$

I tried to do the same for A_6^*, by modifying El Khadir's code. However, in the end, I needed to compute the Groebner basis of an ideal defined by 38 quadratic equations in 25 variables, which my computer was not able to solve.

There is another way to compute A_d^*. By definition, A_d^* is the smallest number A such that

$$A(p(x) + p(y)) - 2Q_p(x, y) \geq 0$$

for all $p \in C_{2,2d}$. Note that

$$p(x) = \frac{1}{(2d)!}\partial_x^{2d}, \qquad p(y) = \frac{1}{(2d)!}\partial_y^{2d}, \qquad Q_p(x, y) = \frac{1}{(2d)!}\partial_x^d \partial_y^d.$$

So we in fact want to compute the smallest number A such that

$$\ell_A := A(\partial_x^{2d} + \partial_y^{2d}) - 2\partial_x^d \partial_y^d \in C_{2,2d}^*$$

where

$$C_{n,2d}^* := \{\ell : \ell(p) \geq 0 \forall p \in C_{n,2d}\}.$$

The dual cone $C_{n,2d}^*$ has a nice characterization from Reznick [Rez11]:

$$C_{n,2d}^* = \text{cone}\{\partial_x^2 \partial_y^{2d-2} : x, y \in \mathbb{R}^n\}.$$

So by decomposing $A(\partial_x^{2d} + \partial_y^{2d}) - 2\partial_x^d \partial_y^d$ we can derive upper bounds for A_d^*. For example, for $d = 1, 2, 3$ we have

$$\begin{aligned}
&\partial_x^2 + \partial_y^2 - 2\partial_x \partial_y = (\partial_x - \partial_y)^2, \\
&\partial_x^4 + \partial_y^4 - 2\partial_x^2 \partial_y^2 = (\partial_x - \partial_y)^2 (\partial_x + \partial_y)^2, \\
&\partial_x^6 + \partial_y^6 - 2\partial_x^3 \partial_y^3 = (\partial_x - \partial_y)^2 (\frac{1}{2}(\partial_x + \partial_y)^4 + \frac{1}{2}\partial_x^4 + \frac{1}{2}\partial_y^4).
\end{aligned}$$

So $A_1^* = A_2^* = A_3^* = 1$. Now we can state our result.

Theorem 2. *For all $p \in C_{n,10}$ (convex forms in n variables of degree 10) and all $x, y \in \mathbb{R}^n$, we have*

$$Q_p(x, y) \leq \sqrt{p(x)p(y)}.$$

In other words, $A_5^* = 1$.

Proof.

$$\begin{aligned}
(4) \qquad &\partial_x^{10} + \partial_y^{10} - 2\partial_x^5 \partial_y^5 \\
&= (\partial_x - \partial_y)^2 (\frac{1}{28}\left(\partial_x^8 + \partial_y^8\right) + \frac{1}{392}(\partial_x - \partial_y)^8 + \frac{11}{168}(\partial_x + \partial_y)^8 \\
&+ \frac{115\sqrt{21} + 527}{1176}\left(\left(\frac{5 - \sqrt{21}}{2}\partial_x + \partial_y\right)^8 + \left(\partial_x + \frac{5 - \sqrt{21}}{2}\partial_y\right)^8\right)).
\end{aligned}$$

$\square$

Let us discuss how we found (4). Because

$$\partial_x^{2d} + \partial_y^{2d} - 2\partial_x^d\partial_y^d = (\partial_x - \partial_y)^2 (\sum_{0\le i\le d} \partial_x^i \partial_y^{d-1-i})^2,$$

it is tempting to include $(\partial_x - \partial_y)^2$ in all terms of the decomposition. So we reduce to decompose $\ell := (\sum_{0\le i\le d} \partial_x^i\partial_y^{d-1-i})^2$ as a convex combination of $(\alpha\partial_x + \beta\partial_y)^{2d-2}$. Note that ℓ is invariant under swapping ∂_x and ∂_y. So wlog we could assume the decomposition is a convex combination of

$$(\alpha\partial_x + \beta\partial_y)^{2d-2} + (\beta\partial_x + \alpha\partial_y)^{2d-2}.$$

We select d tuples of (α_i, β_i) and solve for coefficients γ_i such that

$$\sum_{1\le i\le d} \gamma_i((\alpha_i\partial_x + \beta_i\partial_y)^{2d-2} + (\beta_i\partial_x + \alpha_i\partial_y)^{2d-2}) = \ell.$$

If $\gamma_i \ge 0$ for all $1 \le i \le d$ then we succeed. The (α_i, γ_i) in the proof of Theorem 2 are chosen so that the decomposition is particularly simple. A lot of other decompositions exist. For example, we can get different decompositions if we choose (α_i, β_i) randomly.

I tried to do the same for A_7^*. Namely, look for tuples (α_i, β_i) such that γ_i are all non-negative. However, I was not able to find a valid decomposition. It turns out that this has a reason.

Theorem 3. *For all odd $d \ge 7$,*

$$\partial_x^{2d} + \partial_y^{2d} - 2\partial_x^d\partial_y^d \notin \operatorname{cone}\{(\partial_x - \partial_y)^2(\alpha\partial_x + \beta\partial_y)^{2d-2} : \alpha, \beta \in \mathbb{R}\}.$$

Proof. It suffices to prove that

$$(\sum_{0\le i\le d} \partial_x^i\partial_y^{d-1-i})^2 \notin \operatorname{cone}\{(\alpha\partial_x + \beta\partial_y)^{2d-2} : \alpha, \beta \in \mathbb{R}\}.$$

Note that RHS is exactly $P^*_{2,2d-2}$, dual of the cone of non-negative forms in two variables of degree $2d-2$. Because we have two variables, $P_{2,2d-2} = \Sigma_{2,2d-2}$. So a form $\ell \in P^*_{2d-2}$ if and only if $\ell(p^2) \ge 0$ for all $p \in H_{2,d-1}$.

If we write

$$\ell = \sum_{0\le i\le 2d-2} a_i \partial_x^i \partial_y^{2d-2-i},$$
$$p = \sum_{0\le i\le d-1} b_i x^i y^{d-1-i},$$

then

$$\ell(p^2) = \sum_{0\le k\le 2d-2} a_k k!(2d-2-k)! \sum_{\substack{i+j=k \\ 0\le i,j\le d-1}} b_i b_j.$$

In our case, $a_k = \min\{k, 2d-2-k\} + 1$. So $\ell(p^2)$ can be seen as a degree 2 form in d variables $b_0, \dots, b_{d-1}$. Note that $P_{d,2} = \Sigma_{d,2}$. So $\ell(p^2) \ge 0$ for all p if and only if the Hankel matrix

$$H_{i,j} = h_{i+j}, \qquad 0 \le i, j \le d-1,$$
$$h_k = a_k k!(2d-2-k)!$$

is PSD.

In the following we prove that for all odd $d \geq 7$, H is not PSD. Let $d = 2k+1$. We consider the minor $M := H_{(k-2,\dots,k+2),(k-2,\dots,k+2)}$. Then M is a Hankel matrix corresponding to sequence

$$(d-4)!(d+3)!, (d-3)!(d+2)!, (d-2)!(d+1)!, (d-1)!d!, d!(d-1)!,$$
$$d!(d-1)!, (d+1)!(d-2)!, (d+2)!(d-3)!, (d+3)!(d-4)!$$

and $M = (d-4)!^2 M'$ where entries of M' are degree 7 polynomials in d. We can compute that

$$\det M' = p(d)$$

where

$$\begin{aligned} p(x) = &-48(x+1)(x-1)^5 x^5 (x-2)^6 (x-3)^8 \\ &\cdot (x^2-6x+4)(x^4+6x^3-33x^2+70x-36). \end{aligned}$$

The largest real root of $p(x)$ is at ≈ 5.24, so $p(x) < 0$ for all $x \geq 6$. This finishes the proof. $\square$

Theorem 3 tells us that for odd $d \geq 7$, we have to use terms other than $(\partial_x - \partial_y)^2(\alpha\partial_x + \beta\partial_y)^{2d-2}$.

The statement of Theorem 3 is also valid for even $d \geq 6$, but it is not meaningful because [EK20, Appendix B] proves that $A_d^* > 1$ for even $d \geq 4$.

3. El Khadir: Convex quaternary quartics are SOS

In this section we explain El Khadir [EK20]'s proof that convex quaternary quartics are SOS.

Theorem 4 ([EK20, Theorem 1.1]). *Every convex quaternary quartic is SOS, i.e.,* $C_{4,4} \subseteq \Sigma_{4,4}$.

The proof uses Theorem 1 and another type of generalized Cauchy-Schwarz inequalities.

Theorem 5 ([EK20, Theorem 3.1]). *For any integer* $d \geq 1$*, there exists a constant* B_d *such that*

$$|p(z)| \leq B_d Q_p(z, \overline{z}) \tag{5}$$

for all $p \in C_{n,2d}$ *and* $z \in \mathbb{C}^n$.

The proof is entirely similar to the proof for A_d we presented. Let B_d^* denote the best possible constant such that (5) holds. Unlike A_d^* which turns out to be hard to compute, [EK20] proves that

$$B_d^* = \frac{\binom{2(d-1)}{d-1}}{d}.$$

Proof of Theorem 4 only uses the degree-2 GCS inequalities, $A_2^* = B_2^* = 1$. The proof uses a characterization of extreme rays of $\Sigma_{4,4}^*$ by Blekherman [Ble12].

Theorem 6 ([Ble12, Theorem 1.2], [EK20, Theorem 4.2]). *A non-negative quaternary quartic form* p *is SOS if and only if both of the following conditions hold:*

- *For every $v_1, \dots, v_8 \in \mathbb{R}^4$ and $\alpha_2, \dots, \alpha_8 \in \{\pm 1\}$ such that*

$$v_1 v_1^\top = \sum_{2 \le i \le 8} \alpha_i v_i v_i^\top, \tag{6}$$

  we have

$$p(v_1) \le \Big(\sum_{2 \le i \le 8} \sqrt{p(v_i)}\Big)^2. \tag{7}$$

- *For every $z \in \mathbb{C}^4$, $v_3, \dots, v_8 \in \mathbb{R}^4$, and $\alpha_3, \dots, \alpha_8 \in \{\pm 1\}$ such that*

$$z z^\top + \overline{z}\overline{z}^\top = \sum_{3 \le i \le 8} \alpha_i v_i v_i^\top, \tag{8}$$

  we have

$$2(|p(z)| + \Re(p(z))) \le \Big(\sum_{3 \le i \le 8} \sqrt{p(v_i)}\Big)^2. \tag{9}$$

The proof is now a direct application of the GCS inequalities.

Proof of Theorem 4. Let p be a convex quaternary quartic form. Then p satisfies Theorem 1 and 1 with $A_2^* = B_2^* = 1$. We prove that p satisfies both requirements in Theorem 6.

First requirement: Let $v_1, \dots, v_8, \alpha_2, \dots, \alpha_8$ be as in the condition. Squaring both sides of (6) we get

$$v_1 \otimes v_1 \otimes v_1 \otimes v_1 = \sum_{2 \le i,j \le 8} \alpha_i \alpha_j v_i \otimes v_i \otimes v_j \otimes v_j \tag{10}$$

as symmetric tensors. Let T_p be the symmetric tensor corresponding to Q_p, i.e.,

$$Q_p(x, y) = T_p(x \otimes x \otimes y \otimes y).$$

Applying T_p to both sides of (10), we get

$$p(v_1) = \sum_{2 \le i,j \le 8} \alpha_i \alpha_j Q_p(v_i, v_j) \le \sum_{2 \le i,j \le 8} \sqrt{p(v_i) p(v_j)} = \Big(\sum_{2 \le i \le 8} \sqrt{p(v_i)}\Big)^2$$

where the second step is Theorem 1. So we have (7).

Second requirement: Let $z, v_3, \dots, v_8, \alpha_3, \dots, \alpha_8$ be as in the condition. Squaring both sides of (8) we get

$$\begin{aligned} &z \otimes z \otimes z \otimes z + \overline{z} \otimes \overline{z} \otimes \overline{z} \otimes \overline{z} + 2 z \otimes z \otimes \overline{z} \otimes \overline{z} \\ &= \sum_{3 \le i,j \le 8} \alpha_i \alpha_j v_i \otimes v_i \otimes v_j \otimes v_j \end{aligned} \tag{11}$$

as symmetric tensors. Applying T_p to both sides of (11), we get

$$p(z) + p(\overline{z}) + 2 Q_p(z, \overline{z}) = \sum_{3 \le i,j \le 8} \alpha_i \alpha_j Q_p(v_i, v_j).$$

Note that

$$p(z) + p(\overline{z}) = 2\Re(p(z)), \qquad Q_p(z, \overline{z}) \ge |p(z)|$$

where the inequality is by Theorem 5. So

$$2(|p(z)| + \Re(p(z))) \le \sum_{3 \le i,j \le 8} \alpha_i \alpha_j Q_p(v_i, v_j) \le \Big(\sum_{3 \le i \le 8} \sqrt{p(v_i)}\Big)^2.$$

where the second step is Theorem 1. Therefore (9) holds. □

Blekherman [Ble12] also studied extreme rays of $\Sigma_{3,6}$, which gives a characterization very similar to Theorem 6.

Theorem 7 ([Ble12, Theorem 1.1], [EK20, Theorem 6.1]). *A non-negative ternary sextic form p is SOS if and only if both of the following conditions hold:*

- *For every $v_1, \dots, v_9 \in \mathbb{R}^3$ and $\alpha_2, \dots, \alpha_9 \in \{\pm 1\}$ such that*

$$v_1 \otimes v_1 \otimes v_1 = \sum_{2 \le i \le 9} \alpha_i v_i \otimes v_i \otimes v_i, \tag{12}$$

  we have

$$p(v_1) \le \Big(\sum_{2 \le i \le 9} \sqrt{p(v_i)}\Big)^2. \tag{13}$$

- *For every $z \in \mathbb{C}^3$, $v_3, \dots, v_9 \in \mathbb{R}^3$, and $\alpha_3, \dots, \alpha_9 \in \{\pm 1\}$ such that*

$$z \otimes z \otimes z + \overline{z} \otimes \overline{z} \otimes \overline{z} = \sum_{3 \le i \le 9} \alpha_i v_i \otimes v_i \otimes v_i, \tag{14}$$

  we have

$$2(|p(z)| + \Re(p(z))) \le \Big(\sum_{3 \le i \le 9} \sqrt{p(v_i)}\Big)^2. \tag{15}$$

The first requirement is satisfied by all $p \in C_{3,6}$ because $A_3^* = 1$. However, $B_3^* > 1$, so the same proof does not work for the second requirement. Blekherman [Ble12] conjectures that in both Theorem 7 and 6, the second requirement is not needed. If the conjecture is true, then we get that $C_{3,6} \subseteq \Sigma_{3,6}$.

4. Saunderson: Explicit examples of non-SOS convex forms

In this section we discuss Saunderson [Sau21]'s construction of a family of convex non-SOS forms. Let us describe the construction. Let $\mathbb{O}$ denote the normed division algebra of octonions. Let $x, y \in \mathbb{O}^k$ be two octonion vectors, which can be viewed as two $8k$-dimensional real vectors. We define forms

$$\mathrm{cs}_k(x, y) := ||x||^2 ||y||^2 - |(x, y)|^2,$$
$$q_k(x, y) := \mathrm{cs}_k(x, y) + \frac{1}{4}(||x||^2 + ||y||^2)^2$$

where

$$(x, y) := \sum_{1 \le i \le k} \overline{x_i} y_i \in \mathbb{O}.$$

Theorem 8 ([Sau21, Theorem 1.2]). *For $k \ge 17$, the form $q_k(x, y)$ is a convex non-SOS quartic form in $16k$ variables.*

Convexity follows from [Ble13, Theorem 4.75], which says that if $p \in H_{n,2d}$ satisfies

$$|p(v) - 1| \le \frac{1}{2d - 1} \tag{16}$$

for all $||v|| = 1$, then p is convex. This theorem was used in [Ble13] to prove the existence of convex non-SOS forms. By a generalized Cauchy-Schwarz inequality for octonions [Kra98] (which has nothing to do with Theorem 1 and 5), $\mathrm{cs}_k(x,y)$ is non-negative. So

$$\inf_{||x||^2+||y||^2=1} q_k(x,y) = \inf_{||x||^2+||y||^2=1} \frac{1}{4}(||x||^2+||y||^2)^2 = \frac{1}{4},$$
$$\sup_{||x||^2+||y||^2=1} q_k(x,y) = \inf_{||x||^2+||y||^2=1} (||x||^2||y||^2 + \frac{1}{4}(||x||^2+||y||^2)^2) = \frac{1}{2}.$$

So $\frac{4}{3}q_k$ satisfies (16) and q_k is convex.

The proof that q_k is not SOS is by solving the corresponding SDP using symmetry reduction. Let us describe the SDP, which is standard. Let $V = H_{16k,2}$, $W = H_{16k,4}$, $\mathcal{S}^V$ be the space of self-adjoint endomorphisms of V, $\mathcal{S}_+^V$ be the subset of PSD operators in $\mathcal{S}^V$. Let $\mathcal{A} : \mathcal{S}^V \to W$ be the map induced by $\mathcal{A}(cc^\top) = c^2$ for all $c \in V$. The SDP is

$$\text{find } Y \in \mathcal{S}_+^V \text{ s.t. } \mathcal{A}(Y) = q_k.$$

We can observe that q_k has a large symmetry group. In fact, it is invariant under an action of $\mathrm{Spin}(9) \times O(k)$. If we denote a pair $(x,y) \in \mathbb{O}^k \times \mathbb{O}^k$ as an $\mathbb{R}^{16\times k}$ real matrix

$$X = \begin{bmatrix} [x_1] & \cdots & [x_k] \\ [y_1] & \cdots & [x_k] \end{bmatrix},$$

then $\mathrm{Spin}(9)$ acts by left multiplication on $\mathbb{R}^{16}$ via [Har90, Lemma 14.77] and $O(k)$ acts by right multiplication on $\mathbb{R}^k$.

The action of $\mathrm{Spin}(9) \times O(k)$ on $\mathbb{R}^{16}$ naturally upgrades to V and W. It is easy to verify that q_k is invariant under the action, i.e., for any $(g,h) \in \mathrm{Spin}(9) \times O(k)$, we have $q_k(gXh^\top) = q_k(X)$. So for the SDP, we can assume that Y is invariant under the group action, i.e.,

$$\text{find } Y \in \mathcal{S}_+^V \cap \mathrm{End}_{\mathrm{Spin}(9)\times O(k)}(V) \text{ s.t. } \mathcal{A}(Y) = b.$$

We now need to decompose V into a direct sum of irreducible representations. We embed V into a larger representation $\mathbb{R}^{16k\times 16k} = \mathbb{R}^{16\times 16} \otimes \mathbb{R}^{k\times k}$, where $\mathrm{Spin}(9)$ acts on $\mathbb{R}^{16\times 16}$ as $X \mapsto gXg^\top$ and $O(k)$ acts on $\mathbb{R}^{k\times k}$ as $X \mapsto hXh^\top$.

Decomposition of $\mathbb{R}^{16\times 16}$ under $\mathrm{Spin}(9)$ action and decomposition of $\mathbb{R}^{k\times k}$ under $O(k)$ action are both standard. Omitting details,

$$\mathbb{R}^{16\times 16} = \bigoplus_{0\le i\le 4} V_i,$$

where V_i's are non-isomorphic irreducible $\mathrm{Spin}(9)$-representations;

$$\mathbb{R}^{k\times k} = \bigoplus_{j=-1,0,1} U_j,$$

where U_j are non-isomorphic irreducible $O(k)$-representations. So

$$\mathbb{R}^{16\times 16} \otimes \mathbb{R}^{k\times k} = \bigoplus_{\substack{0\le i\le 4 \\ j=-1,0,1}} (V_i \otimes U_j),$$

where $V_i \otimes U_j$ are non-isomorphic irreducible $\mathrm{Spin}(9)\times O(k)$-representations. Among these representations, V_0, V_1, V_4, U_0, U_1 are symmetric and V_2, V_3, U_{-1} are anti-symmetric. A tensor product $V_i \otimes U_j$ is symmetric if and only if both V_i and U_j

are symmetric, or both V_i and U_j are anti-symmetric. Because V is the subspace of symmetric tensors of $\mathbb{R}^{16\times 16} \otimes \mathbb{R}^{k\times k}$, V decomposes as

$$V = \bigoplus_{(i,j)\in\Lambda} (V_i \otimes U_j)$$

where

$$\Lambda = \{(0,0),(0,1),(1,0),(1,1),(4,0),(4,1),(2,-1),(3,-1)\}.$$

The fact that V decomposes into a direct sum of non-isomorphic irreducible representations is very helpful, because

$$\mathcal{S}_+^V \cap \mathrm{End}_{\mathrm{Spin}(9)\times O(k)}(V) = \mathrm{cone}\{P_{V_i\otimes U_j} : (i,j) \in \Lambda\},$$

where P_A is the orthogonal projection onto A. So the SDP becomes a linear program

$$\text{find } \lambda \in \mathbb{R}^V_{\geq 0} \text{ s.t. } \sum_{(i,j)\in\Lambda} \lambda_i \mathcal{A}(P_{V_i\otimes U_j}) = q_k.$$

A linear program is easy to solve. One can verify that the linear program is feasible for $k = 16$ and infeasible for $k = 17$. So q_k is SOS for $k \leq 16$ and is not SOS for $k \geq 17$. This finishes the proof.

5. Conclusion

In this project we studied problems related to convex forms. Our new result is a tight generalized Cauchy-Schwarz inequality for convex forms of degree 10 (Theorem 2). We also explained why a similar construction cannot work for forms of degree $2d$ for odd $d \geq 7$ (Theorem 3). In Section 3 and Section 4, we explained the main proof ideas and steps of El Khadir [EK20]'s result that convex quaternary quartics are SOS, and Saunderson [Sau21]'s construction of a convex and non-SOS quartic form in 272 variables.

There are still a lot of open questions. We list a few below.

- [EK20]: Numerical computation suggests that $A_d^* = 1$ (in Theorem 1) for all odd d. In this project we resolved the case $d = 5$, but the case $d \geq 7$ is still open.
- [Sau21]: Saunderson's construction gives a convex and non-quartic form. However, it is not obvious how one can construct convex and non-SOS forms of higher degree. We either need a way to increase the degree, or need some construction like Ahmadi-Parrilo [AP13, Theorem 5.10] which constructs an element in $C_{n,2d+2}\backslash\Sigma C_{n,2d+2}$ given an element in $P_{n,2d}\backslash C_{n,2d}$.
- [Sau21]: Saunderson's construction has 272 variables, and is likely not the smallest possible. What is the smallest n such that $C_{n,4} \not\subseteq \Sigma_{n,4}$?
- [Ble12]: Can we remove the second requirement in Theorem 7? Is it true that $C_{3,6} \subseteq \Sigma_{3,6}$?
- [Sau21]: Are all convex symmetric forms SOS? It is known that ther are symmetric non-negative forms that are not SOS [GKR16].

References

[AP13] Amir Ali Ahmadi and Pablo A. Parrilo. A complete characterization of the gap between convexity and SOS-convexity. *SIAM Journal on Optimization*, 23(2):811–833, 2013.

[Ble09] Grigoriy Blekherman. Convex forms that are not sums of squares. *arXiv preprint arXiv:0910.0656*, 2009.

[Ble12] Grigoriy Blekherman. Nonnegative polynomials and sums of squares. *Journal of the American Mathematical Society*, 25(3):617–635, 2012.

[Ble13] Grigoriy Blekherman. Nonnegative polynomials and sums of squares. In *Semidefinite optimization and convex algebraic geometry*, volume 13 of *MOS-SIAM Ser. Optim.*, pages 159–202. SIAM, Philadelphia, PA, 2013.

[EK20] Bachir El Khadir. On sum of squares representation of convex forms and generalized Cauchy–Schwarz inequalities. *SIAM Journal on Applied Algebra and Geometry*, 4(2):377–400, 2020.

[GKR16] Charu Goel, Salma Kuhlmann, and Bruce Reznick. On the Choi–Lam analogue of Hilbert's 1888 theorem for symmetric forms. *Linear Algebra and its Applications*, 496:114–120, 2016.

[GP04] Karin Gatermann and Pablo A. Parrilo. Symmetry groups, semidefinite programs, and sums of squares. *Journal of Pure and Applied Algebra*, 192(1-3):95–128, 2004.

[Har90] F. Reese Harvey. *Spinors and calibrations*. Elsevier, 1990.

[Hil88] David Hilbert. Über die darstellung definiter formen als summe von formenquadraten. *Mathematische Annalen*, 32(3):342–350, 1888.

[HN10] J. William Helton and Jiawang Nie. Semidefinite representation of convex sets. *Mathematical Programming*, 122(1):21–64, 2010.

[Kra98] Linus Kramer. Octonion hermitian quadrangles. *Bulletin of the Belgian Mathematical Society-Simon Stevin*, 5(2/3):353–362, 1998.

[Rez11] Bruce Reznick. Blenders. In *Notions of Positivity and the Geometry of Polynomials*, pages 345–373. Springer, 2011.

[Sau21] James Saunderson. A convex form that is not a sum of squares. *arXiv preprint arXiv:2105.08432*, 2021.