A subquadratic upper bound on sum-of-squares composition formulas

Pavel Hrubeš *

February 20, 2024

Abstract

For every n, we construct a sum-of-squares identitity

$$(\sum_{i=1}^{n} x_i^2)(\sum_{j=1}^{n} y_j^2) = \sum_{k=1}^{s} f_k^2 \,,$$

where f_k are bilinear forms with complex coefficients and $s = O(n^{1.62})$. Previously, such a construction was known with $s = O(n^2/\log n)$. The same bound holds over any field of positive characteristic.

1 Introduction

The problem of Hurwitz [8] asks for which integers n, m, s does there exist a sum-of-squares identity

$$(x_1^2 + \cdots + x_n^2)(y_1^2 + \cdots + y_m^2) = f_1^2 + \cdots + f_s^2 \,, \tag{1}$$

where $f_1, \ldots, f_s$ are bilinear forms in x and y with complex coefficients. Historically, the problem was motivated by existence of non-trivial identities with $n = m = s$. The first one is

$$(x_1^2 + x_2^2)(y_1^2 + y_2^2) = (x_1y_1 - x_2y_2)^2 + (x_1y_2 + x_2y_1)^2 \,.$$

It can be interpreted as asserting multiplicativity of the norm on complex numbers. Euler's 4-square identity is an example with $n, m, s = 4$ which has later been interpreted as multiplicativity of the norm on quaternions. The final one is an 8-square identity which arises in connection to the algebra of octonions.

Let $\sigma(n)$ denote the smallest s such that an identity (1) with $n = m$ exists. For every n, $\sigma(n) \geq n$. The above identities show that $\sigma(n) = n$ if $n \in \{1, 2, 4, 8\}$. A classical result of Hurwitz [8] shows that these are the only cases when equality holds: $\sigma(n) = n$ iff $n \in \{1, 2, 4, 8\}$. An extension of this

*Institute of Mathematics of ASCR, pahrubes@gmail.com. This work was supported by Czech Science Foundation GAČR grant 19-27871X.

result is given by Hurwitz-Radon theorem [11]: an identity (1) exists with $s = n$ iff $m \leq \rho(n)$, where $\rho(n)$ is the Hurwitz-Radon number. The value of $\rho(n)$ is known exactly; if n is a power of 2, $\rho(n)$ lies between $2\log_2 n$ and $2\log_2 n + 2$. As shown in [12], Hurwitz-Radon theorem remains valid over any field of characteristic different from two. Hurwitz's problem is an intriguing question with connections to several branches of mathematics. We recommend D. Shapiro's monograph [13] on this subject.

The asymptotic behavior of $\sigma(n)$ is not known. Trivial bounds are $n \leq \sigma(n) \leq n^2$. Hurwitz's theorem implies that the first inequality is strict if n is sufficiently large. Using Hurwitz-Radon theorem, the trivial upper bound can be improved to

$$\sigma(n) \leq O(n^2/\log n)\,.$$

As far as we are aware, this was the best asymptotic upper bound previously known. In this paper, we will improve it to a truly subquadratic bound

$$\sigma(n) \leq O(n^{1.62})\,. \tag{2}$$

A specific motivation for this problem comes from arithmetic circuit complexity. In [6], Wigderson, Yehudayoff and the current author related the sum-of-squares problem with complexity of non-commutative computations. Non-commutative arithmetic circuit is a model for computing polynomials whose variables do not multiplicatively commute. Since the seminal paper of Nisan [10], it has been an open problem to give a superpolynomial lower bound on circuit size in this model. In [6], it has been shown that a superlinear lower bound of $\Omega(n^{1+\epsilon})$ on $\sigma(n)$ translates to an exponential lower bound in the non-commutative setting. Hence, providing asymptotic lower bounds on Hurwitz's problem can be seen as a concrete approach towards answering Nisan's question. A more general result of this flavor was given by Carmosino et al. in [1]. In an attempt to implement the sum-of-squares approach, the authors from [6] gave an $\Omega(n^{6/5})$ lower bound for sum-of-squares composition formulas over integers [7]. However, the upper bound (2) goes in the opposite direction. Since it is superlinear, it does not immediately frustrate the approach from [6], it merely dampens its optimism.

2 The main result

Let $\mathbb{F}$ be a field. Define $\sigma_{\mathbb{F}}(n,m)$ as the smallest s such that there exist bilienear[1] $f_1, \ldots, f_s \in \mathbb{F}[x_1, \ldots, x_n, y_1, \ldots y_m]$ satisfying (1). Furthermore, let $\sigma_{\mathbb{F}}(n) := \sigma_{\mathbb{F}}(n,n)$.

Theorem 1. *Let $\mathbb{F}$ be either $\mathbb{C}$ or a filed of positive characteristic. Then $\sigma_{\mathbb{F}}(n) \leq O(n^c)$ where $c < 1.62$.*

This will be proved in Section 4. In Section 5.1, we will give a modification of Theorem 1 that applies also to any field.

[1] I.e., of the form $\sum_{i,j} a_{i,j} x_i y_j$.

Remark 2. *If $\mathbb{F}$ has characteristic two, the result is trivial. Since* $(\sum_i x_i^2)(\sum_j y_j^2) = (\sum_{i,j} x_i y_j)^2$, *we have* $\sigma_{\mathbb{F}}(n,m) = 1$.

Notation Given vectors $u, v \in \mathbb{F}^n$, $\langle u, v\rangle := \sum_{i=1}^n u_i v_i$ is their inner product. For a set S, $\binom{S}{k}$ denotes the set of k-element subsets of S and $\binom{S}{\leq k}$ the set of subsets with at most k elements. $\binom{n}{\leq k} := \sum_{i=0}^k \binom{n}{i}$. $[n]$ is the set $\{1, \ldots, n\}$.

3 Hurwitz-Radon conditions

In this section, we give some well-known properties of σ that we will need later.

The definition immediately implies thet $\sigma_{\mathbb{F}}(n,m)$ is symmetric, subadditive, and monotone:

$$\begin{aligned} \sigma_{\mathbb{F}}(n,m) &= \sigma_{\mathbb{F}}(m,n)\,, \\ \sigma_{\mathbb{F}}(n,m_1+m_2) &\leq \sigma_{\mathbb{F}}(n,m_1) + \sigma_{\mathbb{F}}(n,m_2)\,, \\ \sigma_{\mathbb{F}}(n,m) &\leq \sigma_{\mathbb{F}}(n,m')\,,\ m \leq m'\,. \end{aligned} \tag{3}$$

The following lemma gives a characterization of σ in terms of Hurwitz-Radon conditions (4). A proof can be found, e.g., in [13], but we present it for completeness.

Lemma 3. *Let $\mathbb{F}$ be a field of characteristic different from two. Then $\sigma_{\mathbb{F}}(n.m)$ equals the smallest s such that there exist matrices $H_1, \ldots, H_m \in \mathbb{F}^{n\times s}$ satisfying*

$$\begin{aligned} &H_i H_i^t = I_n\,, \\ &H_i H_j^t + H_j H_i^t = 0\,,\ i \neq j\,, \end{aligned} \tag{4}$$

for every $i, j \in [m]$.

Proof. Let $f_1, \ldots, f_s$ be bilinear polynomials in variables $x_1, \ldots, x_n$ and $y_1, \ldots, y_m$. Then the vector $\bar{f} = (f_1, \ldots, f_s)$ can be written as

$$\bar{f} = \sum_{i=1}^n \bar{x} H_i y_i\,,$$

where $\bar{x} = (x_1, \ldots, x_n)$ and $H_i \in \mathbb{F}^{n\times s}$. Hence

$$\sum_{k=1}^s f_k^2 = \bar{f}\bar{f}^t = \sum_i y_i^2 \bar{x} H_i H_i^t \bar{x}^t + \sum_{i<j} y_i y_j \bar{x}(H_i H_j^t + H_j H_i^t)\bar{x}^t\,.$$

If the matrices satisfy (4), this equals $\sum_i y_i^2 \bar{x} I_n \bar{x}^t = (y_1^2 + \cdots + y_m^2)(x_1^2 + \cdots + x_n^2)$, which gives a sum-of-squares identity with s squares. Conversely, if $(y_1^2 + \cdots + y_m^2)(x_1^2 + \cdots + x_n^2) = \sum f_k^2$, we must have $\bar{x} H_i H_i^t \bar{x}^t = x_1^2 + \cdots + x_n^2$ and $\bar{x}(H_i H_j^t + H_j H_i^t)\bar{x}^t = 0$. In characteristic different from 2, this is possible only if the conditions (4) are satisfied. □

Given a natural number of the form $n = 2^k a$ where a is odd, the Hurwitz-Radon number is defined as

$$\rho(n) = \begin{cases} 2k+1\,, & \text{if } k = 0 \\ 2k\,, & \text{if } k = 1 \\ 2k\,, & \text{if } k = 2 \\ 2k+2\,, & \text{if } k = 3 \end{cases} \ \mathsf{mod}\, 4$$

Observe that

$$2\log_2 n \le \rho(n) \le 2\log_2(n) + 2\,,$$

whenever n is a power of two.

Square matrices A_1, A_2 *anticommute* if $A_1A_2 = -A_2A_1$. A family of square matrices $A_1, \ldots, A_t$ will be called *anticommuting* if A_i, A_j anticommute for every $i \neq j$.

The following lemma is a key ingredient in the proof of Hurwitz-Radon theorem. A self-contained construction can be found in [2].

Lemma 4. *For every n, there exists an anticommuting family of $t = \rho(n) - 1$ integer matrices $e_1, \ldots, e_t \in \mathbb{Z}^{n\times n}$ which are orthonormal and antisymmetric (i.e., $e_ie_i^t = I_n$ and $e_i = -e_i^t$).*

Remark 5. *A straightforward construction (see, e.g., [5]) gives an anticommuting family of $t = 2\log_2 n + 1$ integer matrices $e_1, \ldots, e_t \in \mathbb{Z}^{n\times n}$ with $e_i^2 = \pm I_n$ whenever n is a power of two. With minor modifications, these matrices could be used in the subsequent construction instead.*

4 The construction

Let $e_1, \ldots, e_t$ be a set of square matrices. Given $A = \{i_1, \ldots, i_k\} \subseteq [t]$ with $i_1 < \cdots < i_k$, let $e_A := \prod_{j=1}^{k} e_{i_j}$.

Lemma 6. *Let $e_1, \ldots, e_t$ be a set of anticommuting matrices. If $A, B \subseteq [t]$ have even size (resp. odd size) then e_A, e_B anticommute assuming $|A \cap B|$ is odd (resp. even).*

Proof. Since e_i anticommutes with every e_j, $j \neq i$, but commutes with itself, we obtain

$$e_Ae_i = (-1)^{|A\setminus\{i\}|} e_ie_A\,.$$

This implies that

$$e_Ae_B = (-1)^q e_Be_A\,,$$

where $q = |A| \cdot |B| - |A \cap B|$. Hence if A, B are even (resp. odd) and their intersection is odd (resp. even), q is odd and e_A, e_B anticommute. $\square$

Given integers $0 \le k \le t$, a *(k,t)-parity representation* of dimension s over a field $\mathbb{F}$ is a map $\xi : \binom{[t]}{k} \to \mathbb{F}^s$ such that for every $A, B \in \binom{[t]}{k}$

$$\begin{aligned} &\langle \xi(A), \xi(A) \rangle = 1\,, \\ &\langle \xi(A), \xi(B) \rangle = 0\,, \text{ if } A \ne B \text{ and } (|A \cap B| = k \,\mathsf{mod}\, 2)\,. \end{aligned} \tag{5}$$

Lemma 7. *Let $0 \le k \le t$. Over $\mathbb{C}$, there exists a (k,t)-parity representation of dimension $\binom{t}{\le \lfloor k/2 \rfloor}$. If $\mathbb{F}$ is a field of odd characteristic p, there exists a (k,t)-parity representation of dimension $(p-1)\binom{t}{\le \lfloor k/2 \rfloor}$.*

The case of odd characteristic will be proved in the Appendix..

Proof of Lemma 7 over $\mathbb{C}$. Let $0 \le k \le t$ be given and $d := \lfloor k/2 \rfloor$.

For $a \in \{0,1\}^t$, let $|a|$ be the number of ones in a. Recall that a polynomial is multilinear, if every variable in it has individual degree at most one. We first observe:

Claim 8. *There exists a multilinear polynomial $f \in \mathbb{Q}(x_1, \dots, x_t)$ of degree $\le d$ such that for every $a \in \{0,1\}^t$*

$$f(a) = \begin{cases} 1\,, & \text{if } |a| = k \\ 0\,, & \text{if } |a| < k \text{ and } (|a| = k \,\mathsf{mod}\, 2)\,. \end{cases} \tag{6}$$

Proof of Claim. Consider the polynomial

$$g(x_1, \dots, x_t) := c \prod_{0 \le i < k,\, i = k \,\mathsf{mod}\, 2} \Big(\sum_{j=1}^{t} x_j - i\Big)\,.$$

Then g has degree d and we can choose $c \in \mathbb{Q}$ so that g satisfies (6). Since we care about inputs from $\{0,1\}^t$, g can be rewritten as a multilinear polynomial f of degree at most d. $\square$

Since f is multilinear, we can write it as

$$f(x_1, \dots, x_t) = \sum_{C \in \binom{[t]}{\le d}} \alpha_C \prod_{i \in C} x_i\,,$$

where α_C are rational coefficients. Identifying a subset A of $[t]$ with its characteristic vector in $\{0,1\}^t$, we have

$$f(A) = \sum_{C \subseteq A} \alpha_C\,.$$

Let $s := \binom{t}{\le d}$. Given $A \in \binom{[t]}{k}$, let $\xi(A) \in \mathbb{C}^s$ be the vector whose coordinates are indexed by subsets $C \in \binom{[t]}{\le d}$ such that

$$\xi(A)_C = \begin{cases} (\alpha_C)^{1/2}\,, & \text{if } C \subseteq A \\ 0\,, & \text{if } C \not\subseteq A\,. \end{cases}$$

This guarantees

$$\langle \xi(A), \xi(B) \rangle = \sum_C \xi(A)_C \xi(B)_C = \sum_{C \subseteq A \cap B} \alpha_C = f(A \cap B)\,.$$

Hence conditions (6) translate to the desired properties of the map ξ. □

Combining Lemma 6 and 7, we obtain the following bound on σ:

Theorem 9. *Let n be a non-negative integer. Let $0 \leq k \leq \rho(n) - 1$ and $m := \binom{\rho(n)-1}{k}$ Then*

$$\sigma_{\mathbb{C}}(n, m) \leq n \cdot \binom{\rho(n) - 1}{\leq \lfloor k/2 \rfloor}\,.$$

If $\mathbb{F}$ is a field of odd characteristic p then

$$\sigma_{\mathbb{F}}(n, m) \leq (p-1)n \cdot \binom{\rho(n) - 1}{\leq \lfloor k/2 \rfloor}\,.$$

Proof. Let n, k, m be as in the assumption. Let $e_1, \ldots, e_t$ be the matrices from Lemma 4 with $t = \rho(n) - 1$. Let ξ be the (k, t)-parity representation given by the previous lemma. For $A \in \binom{[t]}{k}$, let

$$H_A := e_A \times \xi(A)\,,$$

where e_A is defined as in Lemma 6, and $\xi(A)$ is viewed as a row vector.

Note that each H_A has dimension $n \times (ns)$ where s is the dimension of the parity representation, and there are $m = \binom{t}{k}$ such matrices H_A. By Lemma 3, it is sufficient to show that the system of matrices $H_A, A \in \binom{[t]}{k}$, satisfies Hurwitz-Radon conditions (4).

We have

$$H_A H_B^t = (e_A e_B^t) \times (\xi(A)\xi(B)^t) = \langle \xi(A), \xi(B) \rangle \cdot e_A e_B^t\,.$$

Since every e_i is orthonormal, we have $e_A e_A^t = I_n$. From (5), we have $\langle \xi(A), \xi(A) \rangle = 1$ and hence

$$H_A H_A^t = I_n\,.$$

If $A \neq B$ then

$$H_A H_B^t + H_B H_A^t = \langle \xi(A), \xi(B) \rangle \cdot (e_A e_B^t + e_B e_A^t)\,. \tag{7}$$

If $|A \cap B| = k \bmod 2$ then $\langle \xi(A), \xi(B) \rangle = 0$ by (5) and hence (7) equals zero. If $|A \cap B| \neq k \bmod 2$ then $e_A e_B^t + e_B e_A^t = 0$. This is because $e_A e_B = -e_B e_A$ by Lemma 6 and that, since e_i are antisymmetric, e_A, e_B are either both symmetric or both antisymmetric. Therefore (7) equals zero for every $A \neq B \in \binom{[t]}{k}$. □

Theorem 1 is an application of Theorem 9.

Proof of Theorem 1. Assume first that n is a power of 16. This gves $\rho(n) = 2\log_2(n)+1$. Let k be the smallest integer with $n \leq \binom{2\log_2 n}{k} =: m$. From the previous theorem and monotonicity of σ (cf. (3)), we obtain

$$\sigma_{\mathbb{F}}(n) \leq \sigma_{\mathbb{F}}(n,m) \leq cns\,,$$

where the constant c depends on the field only and $s := \binom{2\log_2 n}{\leq \lfloor k/2 \rfloor}$.

We have $k = 2(\alpha+\epsilon_n)\log_2 n$ where $\alpha \in (0,\frac{1}{2})$ is such that $H(\alpha) = 1/2$ (H is the binary entropy function) and $\epsilon_n \to 0$ as n approaches infinity. We also have

$$s \leq 2^{2H(\frac{\alpha+\epsilon_n}{2})\log_2 n} = n^{2H(\frac{\alpha}{2})+\epsilon'_n}\,,$$

where $\epsilon'_n \to 0$. Hence

$$\sigma_{\mathbb{F}}(n) \leq cn^{1+2H(\frac{\alpha}{2})+\epsilon'_n}\,.$$

The numerical value of α is $0.11\ldots$ which leads to $\sigma_{\mathbb{F}}(n) \leq cn^{1.615+\epsilon'_n} \leq O(n^{1.616})$.

If n is not a power of 16, take n' with $n < n' < 16n$ which is. By monotonicity of σ, we have $\sigma_{\mathbb{F}}(n) \leq \sigma_{\mathbb{F}}(n')$. □

4.1 Comments

Remark 10. *(i). Instead of $\mathbb{C}$, Theorems 9 and 1 apply to any field of characteristic zero where all rationals have a square root.*

(ii). In positive characteristic, the bounds in Lemma 7 and Theorem 9 can sometimes be improved: if $\mathbb{F} \supseteq \mathbb{F}_{p^2}$, the factor $(p-1)$ can be dropped. For certain values of k, $\binom{t}{\leq \lfloor k/2 \rfloor}$ can be replaced with $\binom{n}{\lfloor k/2 \rfloor}$ (cf. Remark 18).

An improvement on the dimension of parity representation in Lemma 7, if possible, will lead to an improvement in Theorem 1. However, this dimension cannot be too small:

Remark 11. *If k is even, every (k,t)-parity representation must have dimension at least $s = \binom{\lfloor t/2 \rfloor}{k/2}$ over any field. This is because there exists a family $\mathcal{A}$ of k-element subsets of $[t]$ whose pairwise intersection is even, and $|\mathcal{A}| = s$. The map ξ must assign linearly independent vectors to elements of $\mathcal{A}$. Similarly for k odd.*

On the other hand, $\binom{t}{\leq \lfloor k/2 \rfloor}$ in Lemma 7 can be replaced with $\binom{t}{\leq \lfloor t-k/2 \rfloor}$ which gives a smaller bound if if $k > t/2$. This is because we can apply the construction of parity representation to complements of $A \in \binom{[t]}{k}$.

The notion of (k,t)-parity representation can be restated in the language of *orthonormal representations* of graphs of Lovász [9]. Given a graph G with vertex set V, its orthonormal representation is a map $\xi(V) :\to \mathbb{F}^s$ such that for every $u, v \in V$

$$\begin{aligned} \langle \xi(u), \xi(u) \rangle &= 1\,, \\ \langle \xi(u), \xi(v) \rangle &= 0\,, \text{ if } u \neq v \text{ are not adjacent in } G. \end{aligned}$$

In this language, (k,t)-parity representation is an orthonormal representation of the following combinatorial Knesser-type graph $G_{k,t}$: vertices of $G_{k,t}$ are k-element subsets of $[t]$. There is an edge between u and v iff $|u \cap v| \neq k \,\mathsf{mod}\, 2$. Orthogonal representations of related graphs have been studied by Haviv in [4, 3].

5 Modifications and extensions

5.1 A sum of bilinear products

Define $\beta_{\mathbb{F}}(n)$ as the smallest s such there exists an identity

$$(x_1^2 + \cdots + x_n^2)(y_1^2 + \cdots + y_n^2) = f_1 f_1' + \cdots + f_s f_s' \,,$$

where $f_1, \ldots, f_s$ and $f_1', \ldots, f_s'$ are bilinear forms with coefficients from $\mathbb{F}$.

In some contexts, β is a more natural quantity than σ. In this section, we give a modification of Theorem 1 in terms of β:

Theorem 12. *Over any field, $\beta_{\mathbb{F}}(n) \leq O(n^c)$ where $c < 1.62$.*

Note that $\beta_{\mathbb{F}}(n) \leq \sigma_{\mathbb{F}}(n)$ over any field. Furthermore, it is easy to see that

$$\begin{aligned} \sigma_{\mathbb{C}}(n) &\leq 2\beta_{\mathbb{C}}(n)\,, \\ \sigma_{\mathbb{F}}(n) &\leq 2(p-1)\beta_{\mathbb{F}}(n)\,, \quad \text{if } \mathbb{F} \text{ has characteristic } p > 0\,. \end{aligned}$$

This means that Theorem 1 can be seen as a consequence of Theorem 12.

The proof of Theorem 12 is a straightforward modification of that of Theorem 1 and we give only a sketch.

The following is an analogy of Lemma 3; we omit the proof.

Lemma 13. *Assume that there are matrices $H_1, \ldots H_m, \tilde{H}_1, \ldots, \tilde{H}_m \in \mathbb{F}^{n\times s}$ satisfying*

$$H_i \tilde{H}_i^t = I_n \,, H_{i_1}\tilde{H}_{i_2}^t + H_{i_2}\tilde{H}_{i_1}^t = 0\,,$$

for every $i \in [m]$ and $i_1 \neq i_2 \in [m]$. Then $\beta_{\mathbb{F}}(n,m) \leq s$.

Lemma 14. *For $0 \leq k \leq t$ and any field $\mathbb{F}$ of characteristic different from two, there exists a pair of maps $\xi, \tilde{\xi} : \binom{[t]}{k} \to \mathbb{F}^s$ with $s \leq \binom{t}{\leq \lfloor k/2 \rfloor}$ such that for every for every $A, B \in \binom{[t]}{k}$*

$$\begin{aligned} \langle \xi(A), \tilde{\xi}(A)\rangle &= 1\,, \\ \langle \xi(A), \tilde{\xi}(B)\rangle &= 0\,, \ \textit{if } A \neq B \textit{ and } (|A\cap B| = k \,\mathsf{mod}\, 2)\,. \end{aligned}$$

Proof. The proof is almost the same as that of Lemma 7. Equipped with the polynomial f from Claim 8 or Lemma 16, it is is sufficient to modify the definition of ξ as follows:

$$\xi(A)_C = \begin{cases} \alpha_C\,, & \text{if } C \subseteq A \\ 0\,, & \text{if } C \not\subseteq A\,. \end{cases}\,, \ \tilde{\xi}(A)_C = \begin{cases} 1\,, & \text{if } C \subseteq A \\ 0\,, & \text{if } C \not\subseteq A\,. \end{cases}$$

□

Proof sketch of Thoreom 12. In Theorem 9, replace the matrices H_A by the pair

$$H_A := e_A \times \xi(A)\,,\ \tilde{H}_A = e_A \times \tilde{\xi}(A)\,.$$

They satisfy the conditions from Lemma 13 and we can proceed as in Theorem 1. □

5.2 A tensor product construction

We now outline an alternative construction of non-trivial sum-of-squares identities. While it gives different types of identities, it does not seem to give better bounds asymptotically.

Instead of the products of anticommuting matrices e_A, one can take the *tensor* product of matrices satisfying Hurwitz-Radon conditions (4). Namely, given such matrices $H_1, \dots, H_m \in \mathbb{F}^{n\times s}$, and $a \in [m]^\ell$, let

$$H_a := H_{a_1} \times H_{a_2} \cdots \times H_{a_\ell}\,.$$

Observe that every H_a satisfies $H_a H_a^t = I$ and that

$$H_a H_b^t + H_b H_a^t = 0\,,$$

whenever a and b have *odd* Hamming distance (i.e., they differ in an odd number of coordinates). As in Lemma 7, we can find a map $\xi : [m]^\ell \to \mathbb{C}^s$ with $s \leq (4m)^{\ell/2}$ such that

$$\begin{aligned} \langle \xi(a), \xi(a)\rangle &= 1\,, \\ \langle \xi(a), \xi(b)\rangle &= 0\,,\ \text{if } a \neq b \text{ have even Hamming distance.} \end{aligned}$$

This gives for every ℓ

$$\sigma_{\mathbb{C}}(n^\ell, m^\ell) \leq \sigma_{\mathbb{C}}(n,m)^\ell (4m)^{\ell/2}$$

For example, starting with $\sigma_{\mathbb{C}}(8,8) = 8$, we have

$$\sigma_{\mathbb{C}}(8^\ell, 8^\ell) \leq 8^{11\ell/6}\,.$$

6 Open problems

Let Even_t denote the set of even-sized subsets of $[t]$. A map $\xi : \mathrm{Even}_t \to \mathbb{F}^s$ will be called a *t-parity representation of dimension s* if for every $A, B \in \mathrm{Even}_t$

$$\begin{aligned} \langle \xi(A), \xi(A)\rangle &= 1\,, \\ \langle \xi(A), \xi(B)\rangle &= 0\,,\ \text{if } A \neq B \text{ and } |A \cap B| \text{ is even.} \end{aligned}$$

Problem 1. *Over $\mathbb{C}$, does there exist a t-parity representation of dimension at most $2^{(0.5+o(1))t}$?*

If this were the case, we could improve the upper bound of Theorem 1 to $\sigma_{\mathbb{C}}(n,n) \leq n^{1.5+o(1)}$. A more surprising consequence would be that $\sigma_C(n,n^2) \leq n^{2+o(1)}$. The constant 0.5 in Problem 1 cannot be improved: since there exists a family of $2^{\lfloor t/2 \rfloor}$ subsets of $[t]$ with pairwise even intersection, every t-parity representation must have dimension at least $2^{\lfloor t/2 \rfloor}$ (cf. Remark 11). On the other hand, Lemma 7 implies that there exists a t-parity representation of dimension at most $2^{H(0.25)+o(1))t} < 2^{0.82t}$.

Our results do not apply to sum-of-squares composition formulas over the real numbers. Since $\mathbb{R}$ is one of the most natural choices of the underlying field in Hurwitz's problem, it is desirable to extend the construction in this direction. This motivates the following:

Problem 2. *Over $\mathbb{R}$, does there exist a t-parity representation of dimension $O(2^{ct})$, where $c < 1$?*

References

[1] Marco L. Carmosino, Russell Impagliazzo, Shachar Lovett, and Ivan Mihajlin. Hardness amplification for non-commutative arithmetic circuits. In *Proceedings of the 33rd Computational Complexity Conference*, CCC '18, Dagstuhl, DEU, 2018. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[2] A. Geramita and N. Pullman. A theorem of Hurwitz and Radon and orthogonal projective modules. *Proceedings of The American Mathematical Society - PROC AMER MATH SOC*, 42:51–51, 01 1974.

[3] I. Haviv. On minrank and the Lovász Theta function. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, 2018.

[4] I. Haviv. Topological bounds on the dimension of orthogonal representations of graphs. *European Journal of Combinatorics*, 81:84–97, 2019.

[5] P. Hrubeš. On families of anticommuting matrices. *Electronic Colloquium in Computational Complexity*, 2014.

[6] P. Hrubeš, A. Wigderson, and A. Yehudayoff. Non-commutative circuits and the sum of squares problem. In *STOC' 10 Proceedings of the 42nd symposium on Theory of Computing*, pages 667–676, 2010.

[7] P. Hrubeš, A. Wigderson, and A. Yehudayoff. An asymptotic bound on the composition number of integer sums of squares formulas. *Canadian Mathematical Bulletin*, 56:70–79, 2013.

[8] A. Hurwitz. Über die Komposition der quadratischen Formen von beliebigvielen Variabeln. *Nach. Ges. der Wiss. Göttingen*, pages 309–316, 1898.

[9] L. Lovász. On the Shannon capacity of a graph. *IEEE Trans. Inform. Theory*, 25(1):1–7, 1979.

[10] N. Nisan. Lower bounds for non-commutative computation. In *Proceeding of the 23th STOC*, pages 410–418, 1991.

[11] J. Radon. Lineare scharen orthogonalen Matrizen. *Abh. Math. Sem. Univ. Hamburg*, 1(2-14), 1922.

[12] D. B. Shapiro. Quadratic forms and similarities. *Bull. Amer. Math. Soc.*, 81(6), 1975.

[13] D. B. Shapiro. *Compositions of quadratic forms.* De Gruyter expositions in mathematics 33, 2000.

A Proof of Lemma 7 in positive characteristic

Given non-negative integers $\bar{n} = (n_1, \ldots, n_d)$ let $B(\bar{n})$ be the $d \times d$ matrix $\{B(\bar{n})_{i,j}\}_{i,j\in[d]}$ with

$$B(\bar{n})_{i,j} = \binom{n_j}{i-1}.$$

We assume that $\binom{n}{k} = 0$ whenever $n < k$; this guarantees $\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!}$.

Lemma 15. *If* $\bar{n} = (r, r+2, \ldots, r+2(d-1))$ *for some non-negative integer* r *then* $\det(B(\bar{n})) = 2^{\binom{d}{2}}$.

Proof. We claim that

$$\det(B(\bar{n})) = (\prod_{i=0}^{d-1} i!)^{-1} \det(V(\bar{n})),$$

where $V(\bar{n})$ is the Vandermonde matrix with entries $V(\bar{n})_{i,j} = n_j^{i-1}$. To see this, multiply every i-th row of $B(\bar{n})$ by $(i-1)!$ to obtain the matrix $B'(\bar{n})$. An i-th row r_i of $B'(\bar{n})$ is of the form $(n_1^i + g_i(n_1), \ldots, n_d^i + g_i(n_d))$ where g_i is a polynomial of degree $< i$. This means that r_i equals the i-th row of $V(\bar{n})$ plus a suitable linear of combination of the first $i-1$ rows of $V(\bar{n})$. Therefore, $\det(B'(\bar{n})) = \det(V(\bar{n}))$.

Given $\bar{n}$ as in the assumption, we obtain

$$\begin{aligned}\det(V(\bar{n})) &= \prod_{1\le j_1<j_2\le d} (n_{j_2} - n_{j_1}) = \prod_{1\le j_1<j_2\le d} (2j_2 - 2j_1) \\ &= 2^{\binom{d}{2}} \prod_{1\le j_1<j_2\le d} (j_2 - j_1) = \prod_{i=1}^{d-1}\prod_{j_1=1}^{d-i} i = \prod_{i=1}^{d-1} i^{(d-i)} = \\ &= \prod_{i=1}^{d-1} i!\,.\end{aligned}$$

This shows that $\det(B(\bar{n})) = 2^{\binom{d}{2}}$. □

Lemma 16. *Let p be an odd prime. Given $0 \leq k \leq t$, there exists a multilinear polynomial $f \in \mathbb{F}_p(x_1, \ldots, x_t)$ of degree at most $d = \lfloor k/2 \rfloor$ such that for every $a \in \{0,1\}^t$*

$$f(a) = \begin{cases} 1\,, & \text{if } |a| = k \\ 0\,, & \text{if } |a| < k \text{ and } (|a| = k \,\mathsf{mod}\, 2)\,. \end{cases}$$

Proof. We look for f of the form $f = \sum_{j=0}^{d} c_j S_t^j$ where S_t^j is the elementary symmetric polynomial $S_t^j = \sum_{|A|=j} \prod_{i \in A} x_i$. Given $a \in \{0,1\}^t$,

$$f(a) = \sum_{j=0}^{d} c_j \binom{|a|}{j} \,\mathsf{mod}\, p\,.$$

We are therefore looking for a solution of the linear system

$$B(\bar{n})\,(c_0 \ldots, c_d)^t = (0, \ldots, 0, 1)^t\ ,$$

where $\bar{n} = (0, 2, \ldots, 2d)$, if k is even, and $\bar{n} = (1, 3, \ldots, 2d+1)$, if k is odd. By the previous lemma, $B(\bar{n})$ is invertible over $\mathbb{F}_p$ and such a solution exists. □

Lemma 17. *If $\mathbb{F}$ is a field of odd characteristic p, there exists a (k,t)-parity representation of dimension $(p-1)\binom{t}{\leq \lfloor k/2 \rfloor}$.*

Proof. If every element of $\mathbb{F}_p$ has a square root in $\mathbb{F}$, the proof is the same as over $\mathbb{C}$. In general, proceed as follows. Since every element of $\mathbb{F}_p$ is a sum of at most $(p-1)$ ones, we can write

$$f(x_1, \ldots, x_t) = \sum_{C \in \mathcal{C}} \prod_{i \in C} x_i\,,$$

where $\mathcal{C}$ is a multiset of $s = (p-1)\binom{t}{\leq d}$ subsets of $[t]$. For $A \in \binom{[t]}{k}$, let $\xi(A) \in \mathbb{F}^s$ be a vector whose coordinates are indexed by elements C of $\mathcal{C}$ so that

$$\xi(A)_C = \begin{cases} 1\,, & \text{if } C \subseteq A \\ 0\,, & \text{if } C \nsubseteq A\,. \end{cases}$$

□

Remark 18. *(i). Over $\mathbb{F}_{p^2}$ or a larger field, the factor of $(p-1)$ in Lemma 17 can be dropped. This is because every element of $\mathbb{F}_p$ has a square root in $\mathbb{F}_{p^2}$.*

(ii). For specific values of k, a stronger bound is possible. For example, if $k = 2p^\ell - 1$, there is a (k,t)-parity representation of dimension $\binom{t}{\lfloor k/2 \rfloor}$. It follows from Lucas' theorem that in this case, f in Lemma 16 can be taken simply as the elementary symmetric polynomial of degree $\lfloor k/2 \rfloor$. This polynomial has only $\binom{t}{\lfloor k/2 \rfloor}$ non-zero monomials.