INSTITUTE of MATHEMATICS

ACADEMY of SCIENCES of the CZECH REPUBLIC

On families of anticommuting matrices

Pavel Hrubeš

Preprint No. 41-2015

PRAHA 2015

On families of anticommuting matrices

Pavel Hrubeš[*]

December 19, 2014

Abstract

Let $e_1, \ldots, e_k$ be complex $n \times n$ matrices such that $e_i e_j = -e_j e_i$ whenever $i \neq j$. We conjecture that

- $\mathrm{rk}(e_1^2) + \mathrm{rk}(e_2^2) + \cdots + \mathrm{rk}(e_k^2) \leq O(n \log n)$.

We show that

(i). $\mathrm{rk}(e_1^n) + \mathrm{rk}(e_2^n) + \cdots + \mathrm{rk}(e_k^n) \leq O(n \log n)$,

(ii). if $e_1^2, \ldots, e_k^2 \neq 0$ then $k \leq O(n)$,

(iii). if $e_1, \ldots, e_k$ have full rank, or at least $n - O(n/\log n)$, then $k = O(\log n)$.

(i) implies that the conjecture holds if $e_1^2, \ldots, e_k^2$ are diagonalizable (or if $e_1, \ldots, e_k$ are). (ii) and (iii) show it holds when their rank is sufficiently large or sufficiently small.

1 Introduction

Consider a family $e_1, \ldots, e_k$ of complex $n \times n$ matrices which pairwise anticommute; i.e., $e_i e_j = -e_j e_i$ whenever $i \neq j$. A standard example is a representation of a Clifford algebra, which gives an anticommuting family of $2\log_2 n + 1$ invertible matrices, if n is a power of two (see Example 1). This is known to be tight: if all the matrices $e_1, \ldots, e_k$ are invertible then k is at most $2\log_2 n + 1$. (see [10] and Theorem 1). However, the situation is much less understood when the matrices are singular. As an example, take the following problem:

Question 1. *Assume that every e_i has rank at least $2n/3$. Is k at most $O(\log n)$?*

We expect the answer should be positive, though we can show only that $k \leq O(n)$. Such a problem can be solved under some extra assumptions. In [6], it was shown that an anticommuting family of diagonalisable matrices can be "decomposed" into representations of Clifford algebras. This indeed answer Question 1 if the e_i's are diagonalisable. In this paper, we formulate a conjecture which relates the size of an anticommuting family with the rank of matrices in the family. We prove some partial results in this direction. In sum, the situation is clear when the matrices are diagonalisable, or their squares are diagonalisable, or even $\mathrm{rk}(e_i^2) = \mathrm{rk}(e_i^3)$. However, we can say very little about the case when the matrices are nilpotent.

One motivation for this study is to understand sum-of-squares composition formulas. A sum-of-squares formula is an identity

$$(x_1^2 + x_2^2 + \cdots + x_k^2) \cdot (y_1^2 + y_2^2 + \cdots + y_k^2) = f_1^2 + f_2^2 + \cdots + f_n^2\,, \tag{1}$$

where $f_1, \ldots, f_n$ are bilinear complex[1] polynomials. We want to know how large must n be in terms of k so that such an identity exists. This problem has a very interesting history, and we refer the reader to the the monograph [10] for details. A classical result of Hurwitz [3] states that $n = k$ can be achieved only for $k \in \{1, 2, 4, 8\}$. Hence, n is strictly larger than k for most values of k, but it is not known how much

[*] `pahrubes@gmail.com`. Supported by ERC grant FEALORA 339691.

[1] The problem is often phrased over $\mathbb{R}$ when the bilinearity condition is automatic.

larger. In particular, we do not known whether $n \geq \Omega(k^{1+\epsilon})$ for some $\epsilon > 0$. In [1], it was shown that such a lower bound would imply an exponential lower bound in a certain circuit model (while the authors obtained an $\Omega(n^{7/6})$ lower bound on *integer* composition formulas in [2]). We point out that our conjecture about anticommuting families implies $n \geq \Omega(k^2/\log k)$, which would be tight. This connection is hardly surprising: already Hurwitz's theorem, as well as the more general Hurwitz-Radon theorem [4, 9], can be proved by reduction to an anticommuting system.

2 The expected rank of anticommuting families

A family $e_1, \dots, e_k$ of $n \times n$ complex matrices will be called *anticommuting* if $e_i e_j = -e_j e_i$ holds for every *distinct* $i, j \in \{1, \dots, k\}$. We conjecture that the following holds ($\mathrm{rk}(A)$ is the rank of the matrix A):

Conjecture 1. *Let $e_1, \dots, e_k$ be an anticommuting family of $n \times n$ matrices. Then*

$$\sum_{i=1}^{k} \mathrm{rk}(e_i^2) \leq O(n \log n)\,.$$

The main motivation is the following theorem:

Theorem 1. *[10] Let $e_1, \dots, e_k$ be an anticommuting family of $n \times n$ invertible matrices. Then $k \leq 2\log_2 n + 1$. The bound is achieved if n is a power of two.*

Under the assumption that e_i^2 are scalar diagonal matrices, this appears in [7] (though it may have been known already to Hurwitz). As stated, it can be found in [10] (Proposition 1.11 and Exercise 12, Chapter 1). There, an exact bound is given

$$k \leq 2q + 1\,, \ \text{ if } n = m2^q \text{ with } m \text{ odd}\,. \tag{2}$$

Theorem 1 shows, first, that the Conjecture holds for invertible matrices and, second, that the purported upper bound cannot be improved: taking $2\log_2 n + 1$ full rank matrices gives $\sum \mathrm{rk}(e_i^2) = (2\log_2 + 1)n$.

A key aspect of Conjecture 1 is that $\sum \mathrm{rk}(e_i^2)$ is bounded in terms of a function of n only. This would fail, had we counted $\sum \mathrm{rk}(e_i)$ instead. For consider 2×2 matrices

$$e_i = \begin{pmatrix} 0 & a_i \\ 0 & 0 \end{pmatrix}, \ a_i \neq 0\,.$$

They trivially anticommute (as $e_i e_j = e_j e_i = 0$), but $\sum_{i=1}^{k} \mathrm{rk}(e_i) = k$, which can be arbitrarily large. However, we also have $e_i^2 = 0$ and this example is vacuous when counting $\sum \mathrm{rk}(e_i^2)$. The minimum requirement of the Conjecture is that every anticommuting family with non-zero squares is finite. This is indeed the case:

Theorem 2. *Let $e_1, \dots, e_k$ be an anticommuting family of $n \times n$ matrices with $e_1^2, \dots, e_k^2 \neq 0$. Then $k \leq O(n)$*

In Theorem 14, we will show that $k \leq 2n - 3$ if n is sufficiently large, which is tight.

Corollary 3. $\sum_{i=1}^{k} \mathrm{rk}(e_i^2) \leq O(n^2)$

We will also show:

Theorem 4. *Let $e_1, \dots, e_k$ be an anticommuting family of $n \times n$ matrices. Then*

$$\sum_{i=1}^{k} \mathrm{rk}(e_i^n) \leq (2\log_2 n + 1)n\,.$$

This implies:

Corollary 5. *Conjecture 1 holds whenever* $\mathrm{rk}(e_i^2) = \mathrm{rk}(e_i^3)$ *for every* e_i *(this is guaranteed if* e_i^2 *is diagonalisable).*

Note that if already $e_1, \ldots, e_k$ are diagonalisable, we obtain $\sum_{i=1}^k \mathrm{rk}(e_i) \le (2\log_2 n + 1)n$.

We will also generalise Theorem 1. In Theorem 7, we show that the assumption that e_i have full rank can be replaced by the assumption that they have almost full rank. This, together with Theorem 2. shows that Conjecture 1 holds if the e_i^2 have either rank close to n or close to $\log n$. Finally, note that the Conjecture implies positive answer to Question 1: if $\mathrm{rk}(e_i) \le 2n/3$ then $\mathrm{rk}(e_i^2) \ge n/3$ and so we must have $k \le O(\log n)$.

Notation and organisation $[k] := \{1, \ldots, k\}$. $\mathbb{C}^{n\times m}$ will denote the set of $n \times m$ complex matrices. For a matrix A, $\mathrm{rk}(A)$ is its the rank. Spectrum of a square matrix A, $\sigma(A)$, is the set of its eigenvalues. A is nilpotent if $A^r = 0$ for some r (or equivalently, $A^n = 0$, or $\sigma(A) = \{0\}$).

In Section 3, we give examples of anticommuting families. In Section 4, we prove Theorems 2, 4 and 7. In Section 5, we prove (2) and determine the bound from Theorem 2 exactly. In Section 6, we outline the connection between our conjecture and the sums-of-squares problem.

We note that our results hold in any field of characteristic different from two.

3 Examples of anticommuting families

We give two examples of anticommuting families. They achieve optimal parameters within its class. Example 1 gives the largest anticommuting family of invertible matrices (Theorem 1), Example 2 the largest family of anticommuting matrices with non-zero squares if $n > 4$ (Theorem 14).

Example 1 - invertible matrices Suppose that $e_1, \ldots, e_k \in \mathbb{C}^{n\times n}$ are anticommuting matrices. Then the following is a family of $k + 2$ anticommuting matrices of dimension $2n \times 2n$:

$$\begin{pmatrix} I_n & 0 \\ 0 & -I_n \end{pmatrix}, \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}, \begin{pmatrix} 0 & e_1 \\ e_1 & 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 & e_k \\ e_k & 0 \end{pmatrix}. \tag{3}$$

Starting with a single non-zero 1×1 matrix, this construction can be applied iteratively to construct a family of $2\log_2 n + 1$ anticommuting invertible $n \times n$ matrices whenever n is a power of two. Moreover, each matrix is diagonalizable. If n is not a power of two but rather of the form $m2^q$ with m odd, we instead obtain $2q + 1$ such matrices.

Example 2 - nilpotent matrices, plus one If $n \ge 2$, consider $n \times n$ matrices of the form

$$e_i = \begin{pmatrix} 0 & u_i & 0 \\ & & v_i^t \\ & & 0 \end{pmatrix},$$

where $u_i, v_i \in \mathbb{C}^{n-2}$ are row-vectors. Then

$$e_i e_j = \begin{pmatrix} 0 & 0 & u_i v_j^t \\ & & 0 \\ & & 0 \end{pmatrix},$$

and so $e_i e_j = -e_j e_i$ iff $u_i v_j^t = -u_j v_i^t$ and $e_i^2 \ne 0$ iff $u_i v_i^t \ne 0$. Setting $r := n - 2$, it is easy to construct row vectors $u_1, \ldots, u_{2r}$, $v_1, \ldots, v_{2r} \in \mathbb{C}^r$ such that for every $i, j \in [2r]$

$$u_i v_i^t \ne 0\,,\ u_i v_j^t = -u_j v_i^t \text{ if } i \ne j\,.$$

This gives an anticommutung family

$$e_1, \ldots, e_{2n-4} \in \mathbb{C}^{n\times n},$$

where every e_i is nilpotent but satisfies $e_i^2 \neq 0$. Note that one can add one more matrix to the family: the diagonal matrix

$$e_0 := \begin{pmatrix} -1 & & \\ & I_{n-2} & \\ & & -1 \end{pmatrix} .$$

This gives $2n-3$ anticommuting matrices with non-zero squares.

4 Lower bounds on family size

In this section, we prove our main theorems. A first observation to make is the following:

Remark 6. *If $e_1, \dots, e_k$ anticommute and $e_1^2, \dots, e_k^2 \neq 0$ then they are linearly independent.*

To see this, assume that $e_1 = \sum_{j>1}^k a_j e_j$. Since e_1 anticommutes with every $e_j, j > 1$, we have $e_1^2 = e_1(\sum a_j e_j) = -(\sum a_j e_j)e_1 = -e_1^2$ and hence $e_1^2 = 0$.

This means that $k \leq n^2$ if $e_1, \dots, e_k \in \mathbb{C}^{n \times n}$. We first show that k must actually be smaller.

Theorem 7. *[Theorem 2 restated] Let $e_1, \dots, e_k \in \mathbb{C}^{n \times n}$ be an anticommuting family with $e_1^2, \dots . e_k^2 \neq 0$. Then $k \leq O(n)$*

In Theorem 14, we will see that the correct bound is $2n-3$ if n is sufficiently large.

Proof. First, there exist row-vectors $u, v \in \mathbb{C}^n$ such that $ue_i^2v^t \neq 0 \in \mathbb{C}$ for every $i \in [k]$. This is because we can view $ue_i^2v^t$ as a polynomial in the $2n$-coordinates of u and v. If $e_i^2 \neq 0$, the polynomial is non-trivial, and so a generic u, v satisfies $ue_i^2v^t \neq 0$ for every $i \in [k]$.

Let us define the $k \times k$ matrix M by

$$M_{ij} := \{ue_ie_jv^t\}_{i,j \in [k]} .$$

Then $\mathrm{rk}(M) \leq n$. This is because M can be factored as $M = L \cdot R$, where L is $k \times n$ matrix with i-th row equal to ue_i and R is $n \times k$ with j-th column equal to e_jv^t. On the other hand, we have $\mathrm{rk}(M) \geq k/2$. This is because $M_{ii} \neq 0$ and, since $e_ie_j = -e_je_i$, $M_{ij} = -M_{ji}$ whenever $j \neq i$. Hence $M + M^t$ is a diagonal matrix with non-zero entries on the diagonal, $\mathrm{rk}(M + M^t) = k$ and so $\mathrm{rk}(M) \geq k/2$. This gives $k/2 \leq \mathrm{rk}(M) \leq n$ and so $k \leq 2n$. □

Remark 6 can be generalised. For $A = \{i_1, \dots, i_r\} \subseteq [k]$ with $i_1 < \dots < i_r$, let e_A be the matrix $e_{i_1}e_{i_2}\cdots e_{i_r}$.

Lemma 8. *Let $e_1, \dots, e_k$ be anticommuting matrices. For $p \leq k$, assume that for every $A \subseteq \{1, \dots, k\}$ with $|A| \leq p$ we have $\prod_{i \in A} e_i^2 \neq 0$. Then the matrices e_A, with $|A| \leq p$ and $|A|$ even, are linearly independent (similarly with odd $|A|$).*

Proof. Suppose that we have a non-trivial linear combination $\sum_{A \text{ even}} a_A e_A = 0$. Let A_0 be a largest A with $a_A \neq 0$. We will show that $\prod_{i \in A_0} e_i^2 = 0$ holds. This implies the statement of the lemma for even A's; the odd case is analogous. The proof is based on the following observations. First, e_i and e_j^2 always commute. Second, if $i \notin A$ then $e_ie_A = (-1)^{|A|}e_Ae_i$, i.e., e_A and e_i commute or anticommute depending on the parity of $|A|$.

Without loss of generality, assume that $A_0 = \{1, \dots, q\}$. For $r \leq q$ and $z \in \mathbb{N}$ let $S_r(z) := \{A \subseteq \{r+1, \dots, k\} : |A| = z \bmod 2\}$. We will show that for every $0 \leq r \leq q$,

$$e_1^2 \cdots e_r^2 \left(\sum_{A \in S_r(r)} a_{[r] \cup A} e_A \right) = 0 . \tag{4}$$

If $r=0$, (4) is just the equality $\sum_{A \text{ even}} a_A e_A = 0$. Assume (4) holds for some $r<q$, and we want to show it holds for $r+1$. Collecting terms that contain e_{r+1} and those that do not, (4) can be rewritten as where

$$e_1^2 \cdots e_r^2 e_{r+1} \left(\sum_{A \in S_{r+1}(r+1)} a_{[r+1]\cup A} e_A \right) = -e_1^2 \cdots e_r^2 \left(\sum_{B \in S_{r+1}(r)} a_{[r]\cup B} e_B \right) .$$

Let f and g be the left and right hand side of the last equality. Since A range over sets of parity $(r+1)$ mod2 and B over sets with parity r mod2, we have $e_{r+1}f = (-1)^{r+1} f e_{r+1}$ and $e_{r+1}g = (-1)^r g e_{r+1}$. Since $f=g$, this gives $e_{r+1}f = -fe_{r+1}0$ and so $e_{r+1}f = 0$. Hence,

$$e_1^2 \cdots e_r^2 e_{r+1}^2 \sum_{A \in S_{r+1}(r+1)} a_{[r+1]\cup A} e_A ,$$

as required in (4). Finally, if we set $r := q$ in (4), we obtain $e_1^2 \cdots e_q^2 \cdot a_{A_0} = 0$ (recall that A_0 is maximal) and so $e_1^2 \cdots e_q^2 = 0$, as required. □

Part (ii) of the following theorem is a generalisation of Theorem 1. Note that part (i) gives $k \le O(\log n)$ whenever $r \ge n - O(n/\log n)$.

Theorem 9. *Let $e_1, \dots, e_k$ be anticommuting matrices in $\mathbb{C}^{n\times n}$ and $r := \min_{i\in[k]} \mathrm{rk}(e_i^2)$.*

(i). If $r > n(1-1/c)$ with $c \in \mathbb{N}$ then $k \le cn^{2/c}$.

(ii). If $r > n\left(1 - \frac{1}{2(\log_2 n+1)}\right)$ then $k \le 2\log_2 n + 1$.

Proof. (i). By Sylvester's inequality, we have $\mathrm{rk}(\prod_{i\in A} e_i^2) > n - |A|n/c$. Hence $\prod_{i\in A} e_i^2 \neq 0$ whenever $|A| \le c$. By Lemma 8, the matrices e_A, $A \subseteq [k]$, $|A| = c$, are linearly independent. Hence $\binom{k}{c} \le n^2$ and the statement follows from the estimate $\binom{k}{c} \ge (k/c)^c$.

In (ii), assume that $k > 2\log_2 n + 1$ and, without loss of generality, $k \le 2\log_2 n + 2$. As above, we conclude $e_1^2 \cdots e_k^2 \neq 0$. The lemma shows that the products e_A, with $|A|$ even, are linearly independent. This gives $2^{k-1} \le n^2$ and so $k \le 2\log_2 n + 1$, a contradiction. □

Before proving Theorem 4, we discuss general structure of anticommuting families. One way to obtain such a family is via a direct sum of simpler families. A family which cannot be so decomposed will be called irreducible. In Proposition 11, we will state some properties of irreducible families which allow to conclude the theorem.

If $A_1 \in \mathbb{C}^{r_1\times r_1}$ and $A_2 \in \mathbb{C}^{r_2\times r_2}$, let $A_1 \oplus A_2$ be the $(r_1+r_2)\times(r_1+r_2)$ matrix

$$A_1 \oplus A_2 = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} .$$

A family $e_1, \dots, e_k \in C^{n\times n}$ will be called *reducible*, if there exists an invertible V such that

$$V e_i V^{-1} = e_i(1) \oplus e_i(2) ,\ i \in [k] \tag{5}$$

where $e_1(1), \dots, e_k(1) \in \mathbb{C}^{r_1\times r_1}$, $e_1(2), \dots, e_k(2) \in \mathbb{C}^{r_2\times r_2}$, with $0 < r_1 < n$ and $r_1 + r_2 = n$. If no such decomposition exists, the family will be called *irreducible.*

Note that the similarity transformation $Ve_1V^{-1}, \dots, Ve_kV^{-1}$ preserves anticommutativity (and rank), and that $e_1, \dots, e_k$ anticommutes iff both $e_1(1), \dots, e_k(1)$ and $e_1(2), \dots, e_k(2)$ do.

Lemma 10. *Let A and B be square matrices of the form*

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} ,\ B = \begin{pmatrix} B_1 & B_3 \\ B_4 & B_2 \end{pmatrix} ,$$

where $A_1, B_1 \in \mathbb{C}^{n\times n}$, $A_2, B_2 \in \mathbb{C}^{m\times m}$. If $AB = -BA$, the following hold:

(i). if there is no λ such that $\lambda \in \sigma(A_1)$ and $-\lambda \in \sigma(A_2)$ then $B_3 = 0$ and $B_4 = 0$,

(ii). if $\sigma(A_1) = \{\lambda_1\}$ and $\sigma(A_2) = \{\lambda_2\}$ for some $\lambda_1, \lambda_2 \neq 0$ then $B_1, B_2 = 0$.

Proof. We first note the folowing:

Claim. *Let $X \in C^{p\times p}$, $Y \in \mathbb{C}^{q\times q}$ and $Z \in \mathbb{C}^{p\times q}$ be such that $XZ = ZY$. If $\sigma(X) \cap \sigma(Y) = \emptyset$ then $Z = 0$.*

Proof. Without loss of generality, we can assume that Y is upper triangular with its eigenvalues $\lambda_1, \ldots, \lambda_r$ on the diagonal. Let $v_1, \ldots, v_q$ be the columns of Z, and assume that some v_i is non-zero. Taking the first such v_i gives $Xv_i = \lambda_i v_i$ – contradiction with $\lambda_i \notin \sigma(X)$. □

Anticommutativity of A and B gives $A_1B_3 = -B_3A_2$ and $A_2B_4 = -B_4A_1$. If A_1, A_2 satisfy the assumption of (i), we have $\sigma(A_1) \cap \sigma(-A_2) = \emptyset$ and so $B_3, B_4 = 0$ by the Claim. We also have $A_1B_1 = -A_1B_1$. If A_1 is as in (ii), we have $\sigma(A_1) \cap \sigma(-A_1) = \emptyset$ and so $B_1 = 0$; similarly for B_2. □

Given A in Jordan normal form, Lemma 8 determines block-structure of B. For example, if A is block-diagonal

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & A_3 & \\ & & & A_4 \end{pmatrix},$$

where $\sigma(A_1) = \{1\}, \sigma(A_2) = \{-1\}$, $\sigma(A_3) = \{0\}$ and $\sigma(A_4) = \{2\}$. Then

$$B = \begin{pmatrix} 0 & B_1 & & \\ B_2 & 0 & & \\ & & B_3 & \\ & & & 0 \end{pmatrix}.$$

Proposition 11. *Let $e_1, \ldots, e_k \in \mathbb{C}^{n\times n}$ be an irreducible anticommuting family. Then every e_i is either invertible or nilpotent. Moreover,*

(i). for every e_i, $\sigma(e_i) \subseteq \{\lambda_i, -\lambda_i\}$ for some $\lambda_i \in \mathbb{C}$,

(ii). if at least two of the matrices are invertible then n is even and the multiplicity of λ_i is exactly $n/2$ in an invertible e_i.

Proof. (i). Assume that there is some e_i with eigenvalues λ_1, λ_2 with $\lambda_1 \neq -\lambda_2$. After a suitable similarity transformation, we can assume that

$$e_i = \begin{pmatrix} e'_i & 0 \\ 0 & e''_i \end{pmatrix},$$

where $e'_i \in \mathbb{C}^{r\times r}$ $e''_i \in \mathbb{C}^{(n-r)\times(n-r)}$ are such that $\sigma(e'_i) \subseteq \{\lambda_1, -\lambda_1\}$ and $\sigma(e''_i) \cap \{\lambda_1, -\lambda_1\} = \emptyset$, for some $0 < r < n$. Lemma 10 part (i) gives that every e_j is of the form

$$e_j = \begin{pmatrix} e'_j & 0 \\ 0 & e''_j \end{pmatrix}$$

and hence the family is reducible.

(i) implies that every e_i is either invertible or nilpotent. For (ii), assume that e_i is non-singular. By (i), we have $\sigma(e_i) \subseteq \{\lambda_i, -\lambda_i\}$ for some $\lambda_i \neq 0$. Decompose e_i as above, but with $\sigma(e'_i) = \{\lambda_1\}$ and $\sigma(e''_i) = \{-\lambda_i\}$. Hence r is the multiplicity of λ_i. The previous lemma part (ii) shows that every e_j, $j \neq i$, is of the form

$$e_j = \begin{pmatrix} 0 & e'_j \\ e''_j & 0 \end{pmatrix},$$

where e'_j is $r \times (n-r)$ and e''_j is $(n-r) \times r$. Hence e_j has rank at most $2r$ and also at most $2(n-r)$. If some e_j is invertible, we must have $r = n/2$. □

Theorem 12. *[Theorem 4 restated] Let $e_1,\dots,e_k \in \mathbb{C}^{n\times n}$ be an anticommutative family. Then $\sum_{i=1}^k \mathrm{rk}(e_i^n) \leq (2\log_2 n+1)n$.*

Proof. Argue by induction on n. If $n=1$, the statement is clear. If $n>1$, assume first that the family is irreducible. By Proposition 11, every e_i is either invertible or nilpotent. If e_i is nilpotent then $e_i^n=0$ and it contributes nothing to the rank. On the other hand, Theorem 1 asserts that there can be at most $2\log_2 n+1$ anticommuting invertible matrices and so indeed $\sum_{i=1}^k \mathrm{rk}(e_i^n) \leq (2\log_2 n+1)n$.

If the family is reducible, consider the decomposition in (5). By the inductive assumption, $\sum \mathrm{rk}(e_i(z)^n) \leq \sum \mathrm{rk}(e_i(z)^{r_z}) \leq (2\log_2 r_z+1)r_z$ for both $z\in\{1,2\}$. Since $\mathrm{rk}(e_i^n)=\mathrm{rk}(e_i(1)^n)+\mathrm{rk}(e_i(2)^n)$, we obtain

$$\begin{aligned}\sum_{i=1}^k \mathrm{rk}(e_i^n) \leq &\sum_{i=1}^k \mathrm{rk}(e_i(1)^{r_1}) + \sum_{i=1}^k \mathrm{rk}(e_i(2)^{r_2}) \leq\\ \leq&(2\log_2 r_1+1)r_1+(2\log_2 r_2+1)r_2 \leq (2\log_2 n+1)(r_1+r_2)=\\ =&(2\log_2 n+1)n\,.\end{aligned}$$

□

5 Some exact bounds

For completeness, we now sketch a proof of (2) from Section 2. We then prove the exact bound in Theorem 2.

Proposition 13. *Let $e_1,\dots,e_k$ be an anticommutative family of invertible $n\times n$ matrices, where $n=m2^q$ with m is odd. Then $k\leq 2q+1$.*

The bound is achieved by Example 1

Proof sketch. Argue by induction on n. If $n>1$, the non-trivial case is when the family is irreducible. If $k>1$, we can assume that

$$e_1=\begin{pmatrix} e_1' & 0\\ 0 & e_1''\end{pmatrix},\; e_j=\begin{pmatrix} 0 & e_j'\\ e_j'' & 0\end{pmatrix},\;\text{ if } j>1. \tag{6}$$

where $e_i',e_i''\in\mathbb{C}^{n/2\times n/2}$ are invertible. This is because, by Proposition 11, we can write e_1 as in (6) with $\sigma(e_1')=\{\lambda\}$, $\sigma(e_1'')=\{-\lambda\}$, $\lambda\neq 0$. Lemma 10 part (ii) gives that every $e_j, j>1$ must indeed be of the form required in (6). If $e_2,\dots,e_k$ anticommute then so do the $k-2$ matrices $e_2e_3,e_2e_4,\dots,e_2e_k$. If $j>1$,

$$e_2e_j=\begin{pmatrix} e_2'e_j'' & 0\\ 0 & e_2''e_j'\end{pmatrix},$$

and so $e_2'e_3'',\dots,e_2'e_k''$ is a family of $k-2$ invertible anticommuting matrices in $\mathbb{C}^{n/2\times n/2}$. The inductive assumption gives $k-2\leq 2(q-1)+1$ and so $k\leq 2q+1$ as required. □

For a natural number n, let $\alpha(n)$ denote the largest k so that there exists an anticommuting family $e_1,\dots,e_k\in\mathbb{C}^{n\times n}$ with $e_1^2,\dots,e_k^2\neq 0$.

Theorem 14.

$$\alpha(n)=\begin{cases} 2n-1\,, & \text{if } n\in\{1,2\}\\ 2n-2\,, & \text{if } n\in\{3,4\}\\ 2n-3\,, & \text{if } n>4\end{cases}$$

The rest of this section is devoted to proving the theorem.

Lemma 15. *If $n>1$, $\alpha(n)$ equals the maximum of the following quantities: a) $2n-3$, b) $\max_{0<r<n}(\alpha(r)+\alpha(n-r))$, c) $2+\alpha(n/2)$ (where we set $\alpha(n/2):=-1$ if n is odd).*

Proof. That $\alpha(n)$ is at least the maximum is seen as follows. $\alpha(n) \geq$ a) is Example 2. $\alpha(n) \geq 2 + \alpha(n/2)$ is seen from (3) in Example 1. For b), suppose we have two anticommuting families $e_1(z), \ldots, e_{k_z}(z) \in \mathbb{C}^{r_z \times r_z}$, $z \in \{1, 2\}$. Then the following is an anticommuting family of $(r_1 + r_2) \times (r_1 + r_2)$ matrices: $e_1(1) \oplus 0, \ldots, e_{k_1} \oplus 0, 0 \oplus e_1(2), \ldots, 0 \oplus e_{k_2}(2)$ (with $0 \in \mathbb{C}^{r_1 \times r_1}, \mathbb{C}^{r_2 \times r_2}$ respectively).

We now prove the opposite inequality. Let $e_1, \ldots, e_k \in \mathbb{C}^{n \times n}$ be an anticommuting family with $e_1^2, \ldots, e_k^2 \neq 0$. We first prove two claims.

Claim 1. *If all the e_i's are nilpotent then $k \leq 2(n-2)$.*

Proof. By a theorem of Jacobson [5], see also [8], a family of anticommuting nilpotent matrices is simultaneously upper triangularisable. So let assume that $e_1, \ldots, e_k$ are upper triangular with zero diagonal, and proceed as in the proof of Theorem 7. For M as defined in the proof, it is enough to show that $\mathrm{rk}(M) \leq n-2$, which gives $k \leq 2(n-2)$. If the e_i's are upper triangular with zero diagonal, we can see that the first column of L and the last row of R are zero. This means $\mathrm{rk}(M) = \mathrm{rk}(LR) \leq n-2$. □

Claim 2. *If e_1, e_2 are invertible then $k \leq 2 + \alpha(n/2)$.*

Proof. As in the proof of Theorem 13, we can assume that the matrices have the form (6). Note that e_2', e_2'' are invertible and $e_2'e_3'', \ldots, e_2'e_k''$ is an anticommuting family of $k-2$ matrices in $\mathbb{C}^{n/2 \times n/2}$. If we show that $(e_2'e_j'')^2 \neq 0$ for every $j \in \{3, \ldots, k\}$, we obtain $k - 2 \leq \alpha(n/2)$ as required.

Let $j \in \{3, \ldots, k\}$. Anticommutativity of e_2 and e_j gives $e_2'e_j'' = -e_j'e_2''$ and $e_2''e_j' = -e_j''e_2'$. Hence

$$\begin{aligned}(e_2'e_j'')^2 = e_2'e_j''e_2'e_j'' &= e_2'(e_j''e_2')e_j'' = -e_2'e_2''e_j'e_j'' \,, \\ &= e_2'e_j''(e_2'e_j'') = -e_2'e_j''e_j'e_2'' \,.\end{aligned}$$

If $(e_2'e_j'')^2 = 0$, the first equality gives $e_j'e_j'' = 0$ and the second $e_j''e_j' = 0$ (recall that e_2', e_2'' are invertible). But since $e_j^2 = e_j'e_j'' \oplus e_j''e_j'$, this gives $e_j^2 = 0$ – contrary to the assumption $e_j^2 \neq 0$. □

To prove the Lemma, assume first that $e_1, \ldots, e_k$ is irreducible. Then the e_i's are either invertible or nilpotent. If there is at most one invertible e_i, Claim 1 gives $k - 1 \leq 2(n-2)$, as in a). If at least two e_i's are invertible, Claim 2 gives $k \leq 2 + \alpha(n/2)$, as in b). If the family is reducible, write it as in (5). For $z \in \{1, 2\}$, let $A_z := \{i \in [k] : e_i(z)^2 \neq 0\}$. Then $A_1 \cup A_2 = [k]$ and so $k \leq \alpha(r_1) + \alpha(r_2)$, as in c). □

Proof of Theorem 14. Using the Lemma, it is easy to verify that the theorem holds for $n \leq 4$. If $n > 4$, the lemma gives $\alpha(n) \geq 2n-3$ and it suffices to prove the opposite inequality. Assume that n is the smallest $n > 4$ such that $\alpha(n) > 2n - 3$. This means that for every $n' < n$, $\alpha(n') = 2n' - \epsilon(n')$ where $\epsilon(n') = 1$ if $n' \in \{1, 2\}$ and $\epsilon(n') > 1$ otherwise. Then either $\alpha(r) + \alpha(n-r) > 2n - 3$ for some $0 < r < n$, or $2 + \alpha(n/2) > 2n - 3$. The first case is impossible: we have $\alpha(r) + \alpha(n-r) = 2n - \epsilon(r) - \epsilon(n-r)$. But $\epsilon(r) + \epsilon(n-r) < 3$ implies $r, (n-r) \in \{1, 2\}$ and so $n \leq 4$. If $2 + \alpha(n/2) > 2n - 3$ we have $2 + 2(n/2) - 2\epsilon(n/2) > 2n - 3$ and so $n < 5 - 2\epsilon(n/2) \leq 3$. □

6 Sum-of-squares formulas

We now briefly discuss the sum-of-squares problem. Let $\sigma(k)$ be the smallest n so that there exists a sum-of-squares formula as in (1) from the Introduction. The following can be found in Chapter 0 of [10]:

Lemma 16. *$\sigma(k)$ is the smallest n such that there exists $k \times n$ matrices $A_1, \ldots A_k$ which satisfy*

$$A_i A_i^t = I_k \,,\ A_i A_j^t = -A_j A_i^t \,,\ \text{ if } i \neq j \,,$$

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

The matrices from the lemma can be converted to anticommuting matrices, which provides a connection between the sum-of-squares problem and Conjecture 1, as follows.

Proposition 17. *(i). If $\sigma(k) = n$, there exists an anticommuting family $e_1, \ldots, e_k \in \mathbb{C}^{(n+2k)\times(n+2k)}$ such that* $\mathrm{rk}(e_1^2), \ldots, \mathrm{rk}(e_k^2) = k$*. (Moreover, we have $e_1^2 = e_2^2 \cdots = e_k^2$ and $e_1^3, \ldots, e_k^3 = 0$.)*

(ii). Hence, Conjecture 1 implies $\sigma(k) = \Omega(k^2/\log k)$.

Proof. Take the $(2k+n)\times(2k+n)$ matrices (with $0 \in \mathbb{C}^{k\times k}$)

$$e_i := \begin{pmatrix} 0 & A_i & 0 \\ & & A_i^t \\ & & 0 \end{pmatrix}, \, i \in [k] \, .$$

The matrices have the required properties as seen from

$$e_i e_j = \begin{pmatrix} 0 & 0 & A_i A_j^t \\ & & 0 \\ & & 0 \end{pmatrix}.$$

We have $\sum_{i=1}^k \mathrm{rk}(e_i^2) = k^2$. As $2k+n \leq 3n$, the Conjecture gives $k^2 = \sum_{i=1}^k \mathrm{rk}(e_i^2) \leq O(3n\log(3n))$ and so $n \geq \Omega(k^2/\log k)$. □

We can see that the matrices obtained in (i) are nilpotent, which is exactly the case of Conjecture 1 we do not know how to handle. Finally, let us note that part (i) is too generous if $\sigma(k) = k$. In this case, we can actually obtain $k-1$ invertible anticommuting matrices in $\mathbb{C}^{k\times k}$. Again following [10], let

$$e_1 := A_1 A_k^t \, , \; e_2 := A_2 A_k^t \, , \ldots \, , \; e_{k-1} := A_{k-1} A_k^t \, .$$

They anticommute, as seen from $A_i A_k^t A_j A_k^t = -A_i A_k^t A_k A_j^t = -A_i A_j^t$ (note that $A_k A_k^t = I$ implies $A_k^t A_k = I$ for square matrices). This is one way how to obtain Hurwitz's $\{1,2,4,8\}$-theorem: if $\sigma(k) = k$, we have $k-1$ invertible anticommuting matrices in $\mathbb{C}^{k\times k}$. By Theorem 7, this gives $k-1 \leq 2\log_2 k + 1$ and hence $k \leq 8$. Furthermore, the precise bound in (2) rules out the k's which are not a power of two.

References

[1] P. Hrubeš, A. Wigderson, and A. Yehudayoff. Non-commutative circuits and the sum of squares problem. *J. Amer. Math. Soc.*, 24:871–898, 2011.

[2] 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.

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

[4] A. Hurwitz. Über die Komposition der quadratischen Formen. *Math. Ann.*, 88:1–25, 1923.

[5] N. Jacobson. *Lie Algebras.* Interscience, New York, 1962.

[6] Y. Kumbasar and A. H. Bilge. Canonical forms for families of anti-commuting diagonalizable operators. *ArXiv*, 2011.

[7] M. H. A. Newman. Note on an algebraic theorem of Eddington. *J. London Math. Soc*, 7:93–99, 1932.

[8] H. Radjavi. The Engel-Jacobson theorem revisited. *J.Algebra*, 111:427–430, 1987.

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

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