Sum-of-squares composition formulas of size $N^{\sqrt{2}+o(1)}$ and subquadratic non-commutative circuits in every characteristic

Eurekist*

November 2026

Abstract

We show that over every field the crossing function $[x = v][y = u]$ on $[n] \times [m] \times [m] \times [n]$ is a sum of $\exp\bigl((1 + o(1))\sqrt{\ln^2 n + \ln^2 m}\bigr)$ functions, each of rank one across $(x, y) \,|\, (u, v)$ or across $(x, u) \,|\, (y, v)$. In characteristic zero this is implicit in a Heisenberg vertex-algebra identity from a recent paper on 3-SAT by Claude (Anthropic); we extend it to positive characteristic by working in a quotient of the Fock space. Three consequences follow. First, sum-of-squares composition formulas: $\sigma_{\mathbb{F}}(n) \le n^{\sqrt{2}+o(1)}$ for $\mathbb{F}$ containing $\sqrt{-1}$ or of odd positive characteristic, improving Hrubeš's $O(n^{1.62})$. Second, Strassen's algorithm alone gives non-commutative circuits of size $O(n^{1.985})$ for $\mathsf{ID}_n = \sum_{i,j} x_i y_j x_i y_j$ over every field, which answers Hrubeš's question about characteristic two. Third, over fields where the known rectangular matrix multiplication bounds hold, the size is $O(n^{1.859})$, improving Hrubeš's $O(n^{1.96})$.

1 Introduction

Hurwitz's problem asks for which n, m, s there are bilinear forms $f_1, \dots, f_s$ with

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

Let $\sigma_{\mathbb{F}}(n, m)$ be the least such s over a field $\mathbb{F}$, and $\sigma_{\mathbb{F}}(n) = \sigma_{\mathbb{F}}(n, n)$. The asymptotics of $\sigma_{\mathbb{F}}(n)$ are unknown. Hrubeš, Wigderson and Yehudayoff [4] showed that a superlinear lower bound $\sigma(n) = \Omega(n^{1+\varepsilon})$ would give exponential lower bounds for non-commutative arithmetic circuits, through the degree-four polynomial

$$\mathsf{ID}_{n,m} = \sum_{i \in [n],\, j \in [m]} x_i y_j x_i y_j, \qquad \mathsf{ID}_n = \mathsf{ID}_{n,n}.$$

Hrubeš [3] proved $\sigma_{\mathbb{F}}(n) = O(n^{1.62})$ when $\mathbb{F}$ contains $\sqrt{-1}$ or has positive characteristic, and $\mathrm{size}_{\mathbb{F}}^{(nc)}(\mathsf{ID}_n) = O(n^{1.96})$ when $\operatorname{char}\mathbb{F} \neq 2$. Hrubeš asks whether a bound $O(n^{2-\varepsilon})$ holds in characteristic two [3, Problem 3]. There (1) is trivial, so it gives no information about ID_n.

We work with a quantity that makes sense in every characteristic.

*Corrections: max@eurekist.com. Summary and checking scripts: `https://eurekist.com/findings/sum-of-squares`.

Definition 1. Let $T_{n,m}(x,y,u,v) = [x = v][y = u]$ for $x, v \in [n]$, $y, u \in [m]$. A function g of (x,y,u,v) has *row type* if $g = a(x,y)\,b(u,v)$, and *column type* if $g = c(x,u)\,e(y,v)$. Let $P_{\mathbb{F}}(n,m)$ be the least number of functions of row or column type, with values in $\mathbb{F}$, whose sum is $T_{n,m}$.

For $\mathbb{F} = \mathbb{Q}$ and $n = m$ this is the quantity $P(N)$ of [2, Appendix G]. It is the rank in Naslund's Problem 12 [8], which asks for a lower bound, for the two partitions 12|34 and 13|24. By [4, Lemma 3.5] it is the *width* of ID_n in the sense of that paper, and it is the PT-rank of I_{n^2} of Rossman and Zhu [9]. The quantity $P_{\mathbb{F}}(n,m)$ is non-decreasing in n and in m, by restriction.

Theorem 1. *For every field $\mathbb{F}$, as $\max(n,m) \to \infty$,*

$$\ln P_{\mathbb{F}}(n,m) \le (1+o(1))\sqrt{\ln^2 n + \ln^2 m},$$

where the $o(1)$ depends only on $\operatorname{char}\mathbb{F}$. In particular $P_{\mathbb{F}}(n,n) \le n^{\sqrt{2}+o(1)}$. If $\operatorname{char}\mathbb{F} = 0$ then $P_{\mathbb{F}}(n,n) = O\big(n^{\sqrt{2}}(\ln n)^5\big)$.

Theorem 2. *For every field $\mathbb{F}$ there are bilinear forms with $(\sum_i x_i^2)(\sum_j y_j^2) = \sum_{k\le s} f_k f'_k$ and $s \le P_{\mathbb{F}}(n,m)$. If $\operatorname{char}\mathbb{F} \ne 2$ and -1 is a sum of r squares in $\mathbb{F}$, then $\sigma_{\mathbb{F}}(n,m) \le (\mathsf{r}+1)P_{\mathbb{F}}(n,m)$. So $\sigma_{\mathbb{F}}(n,m) \le 2P_{\mathbb{F}}(n,m)$ if $\sqrt{-1} \in \mathbb{F}$, and $\sigma_{\mathbb{F}}(n,m) \le 3P_{\mathbb{F}}(n,m)$ in odd positive characteristic. In these cases $\sigma_{\mathbb{F}}(n) \le n^{\sqrt{2}+o(1)}$.*

Theorem 3. *(a) Over every field, $\operatorname{size}^{(nc)}_{\mathbb{F}}(\mathsf{ID}_n) = O(n^{1.985})$, using only Strassen's algorithm for fast matrix multiplication. In particular the answer to [3, Problem 3] is yes.*

(b) Let $\omega(1,3,1)$ be the exponent over $\mathbb{F}$ of multiplying an $n \times n^3$ matrix by an $n^3 \times n$ matrix. If $\omega(1,3,1) < 4.1989$ over $\mathbb{F}$, as in [11] (see Remark 1), then $\operatorname{size}^{(nc)}_{\mathbb{F}}(\mathsf{ID}_n) \le n^{(1+\sqrt{5}+\omega(1,3,1))/4+o(1)} = O(n^{1.859})$.

The same bounds hold for the non-commutative elementary symmetric polynomial $S_{4,n}$, by [3, Proposition 26]. Consequently, $\operatorname{size}^{(nc)}_{\mathbb{F}}(S_{k,n}) = O(n^{k-4+c})$ for each fixed $k \ge 4$, with the exponent c from (a) or (b), as in [3, Corollary 27]. The implied constants depend on $\operatorname{char}\mathbb{F}$, and on k for $S_{k,n}$.

Provenance. The identity behind Theorem 1 in characteristic zero is Appendix B of Claude [2], which uses it to build crossing gadgets for a 3-SAT algorithm. Appendix G of that paper records the bound $P(N) = O(N^{11/6})$ from a different lemma and cites $O(N^{1.62})$; to our knowledge the bound $N^{\sqrt{2}+o(1)}$ has not been stated. The algebra we use in positive characteristic is the simple quotient vertex algebra of Li and Mu [6] (with $\lambda = 0$, $\dim\mathfrak{h} = 1$, level one). Our contributions are the short deduction, the rectangular form, the crossing identity in positive characteristic (Section 2), and the circuit bounds, which use rectangular labels and work directly with $P_{\mathbb{F}}$ instead of squares.

Limits. All statements are asymptotic. Table 1 shows that the bound beats $n^{1.62}$ only for astronomically large n: near 2^{68} with exact term counts, and near 2^{104} with the simpler bound displayed in Proposition 4. In characteristic two, a rough estimate puts the point where the circuit of Theorem 3(a) drops below the trivial n^2 at $\log_2 n$ in the thousands. The results say nothing about sums of squares over $\mathbb{R}$, and the problem can depend on the field [10]. And $n^{\sqrt{2}}$ is still superlinear, so the lower-bound programme of [4] survives, though with less room.

2 The identity

Let $R = \mathbb{Z}[t_1, t_2, \dots]$, with t_k of weight k, and let R_E be its part of weight E. For $k \geq 1$ let ∇_{-k} be multiplication by t_k and $\nabla_k = k\,\partial/\partial t_k$. For a monomial $a = \prod_i t_{\mu_i}$ and $r \in \mathbb{Z}$ put

$$a[r] = \sum_{\substack{k_i \neq 0 \\ \sum_i k_i = -r}} \prod_i \binom{-k_i - 1}{\mu_i - 1} \prod_{k_i < 0} \nabla_{k_i} \prod_{k_i > 0} \nabla_{k_i},$$

with differentiations applied first, and extend linearly in a. On each R_E only finitely many terms are non-zero, $a[r]$ maps R_E to R_{E+r}, and it preserves R when $a \in R$. For $h \geq 1$, $b \in R_h$ and $a = \sum_{h'=0}^{h} \bar{a}_{h'}$ with $\bar{a}_{h'} \in R_{h'}$, put $a\langle i \rangle = \sum_{h'} \binom{h'+h}{i} \bar{a}_{h'}$ and

$$Q_s(a,b) = \sum_{i=0}^{2h} (-1)^i\, a\langle i \rangle [s - h + i]\; b[h - s - i].$$

Claude [2, Lemma 9] proves the following, on $R_E \otimes \mathbb{Q}$, for every $E \geq 0$. The fields $Y(a,\omega) = \sum_r a[r]\omega^{r-h'}$ are those of the Heisenberg vertex algebra, and the lemma is an explicit form of their locality.

(i) $Q_s(a,b)$ does not change when the two factors are exchanged in all products.

(ii) $Q_s(a,b) = 0$ for $|s| > E + h$.

(iii) $\sum_{|s| \leq E+h} Q_s(a,b) = (a[-h]b) \cdot \mathrm{Id}$.

Moreover, let $y_n = \sum_{j=1}^{n} \binom{n}{j} t_j$ and, for a partition μ of h, let $a_\mu = \prod_i (-1)^{\mu_i - 1} y_{\mu_i}$. Then [2, proof of Lemma 10], for all $b \in R_h$,

(iv) $a_\mu[-h]\, b = \prod_i \nabla_{\mu_i} b$.

Here a_μ has monomials of weight at most h, as (i)–(iii) require.

For readers who know vertex algebras, (i)–(iii) are standard. For homogeneous a of weight h' and $H = h' + h$, locality says that $(\omega_1 - \omega_2)^H$ times $Y(a,\omega_1)Y(b,\omega_2)$ and times $Y(b,\omega_2)Y(a,\omega_1)$ are the same Laurent polynomial, namely $\sum_{n<H} (\omega_1 - \omega_2)^{H-n-1} Y(a_{(n)}b, \omega_2)$. Setting $\omega_1 = \omega_2$ leaves $Y(a_{(H-1)}b, \omega_2)$, which is a scalar because $a_{(H-1)}b$ has weight zero. Linearity covers mixed weights.

For $a, b \in R$ all operators in (i)–(iv) have integer matrix entries in the monomial basis, so (i)–(iv) are equalities of integer matrices and hold over $\mathbb{Z}$. Hence they hold after tensoring with any field $\mathbb{F}$.

The quotient. Let $p = \mathrm{char}\,\mathbb{F}$. If $p = 0$ let $V = R \otimes \mathbb{F}$. If $p > 0$, let $I \subset R \otimes \mathbb{F}$ be the ideal generated by all t_k with $p \mid k$ and all t_k^p, and let $V = (R \otimes \mathbb{F})/I$. All the operators ∇_k, $k \neq 0$, preserve I: multiplications clearly do, $\nabla_k = 0$ when $p \mid k$, and $\nabla_k(t_k^p g) = t_k^p \nabla_k g$ because $p\, t_k^{p-1} = 0$; for the other generators this is the Leibniz rule. So every $a[r]$ with $a \in R$ acts on V, and (i)–(iv) hold on V. The space V is graded, with V_E having as basis the monomials t^ν for $\nu \in \mathcal{L}(E)$, where $\mathcal{L}(E)$ is the set of partitions of E with no part divisible by p and no part repeated p or more times. For $p = 0$, $\mathcal{L}(E)$ is the set of all partitions of E. Write $q(E) = |\mathcal{L}(E)|$ and $S(E) = \sum_{e \leq E} q(e)$. For a partition ν with m_k parts equal to k, let $z_\nu = \prod_k k^{m_k} m_k!$. Then $z_\nu \neq 0$ in $\mathbb{F}$ for $\nu \in \mathcal{L}(E)$.

Proposition 4. *Let $h \geq 1$, $E \geq 0$. For $x, v \in \mathcal{L}(h)$, $y, u \in \mathcal{L}(E)$ and $|s| \leq E + h$ define*

$$A_s(x, y, u, v) = z_v^{-1}\,[t^u]\,Q_s(a_x, t^v)\,t^y \in \mathbb{F}.$$

Then $\sum_s A_s(x, y, u, v) = [x = v][y = u]$, and each A_s is a sum of at most $S(E+h+s)$ functions of row type, and also a sum of at most $S(E + h - s)$ functions of column type. Consequently

$$P_{\mathbb{F}}\big(q(h), q(E)\big) \leq (2E + 2h + 1)\,S(E + h).$$

Proof. By (iii), $\sum_s Q_s(a_x, t^v)t^y = (a_x[-h]t^v)\,t^y$. By (iv), $a_x[-h]t^v = \prod_i \nabla_{x_i} t^v$. If $x \neq v$ then, since $|x| = |v|$, some k occurs more often in x than in v, and the product vanishes. If $x = v$ it equals z_v. This gives the identity.

For the column form expand the definition, inserting the monomial basis of V after the operator from $b = t^v$ has acted:

$$A_s = \sum_{i=0}^{2h} \sum_{\sigma \in \mathcal{L}(E+h-s-i)} (-1)^i\,[t^u]\big(a_x\langle i\rangle[s - h + i]\,t^\sigma\big) \cdot z_v^{-1}[t^\sigma]\big(t^v[h - s - i]\,t^y\big).$$

The first factor depends on (x, u) and the second on (y, v). The levels $E + h - s - i$ for $0 \leq i \leq 2h$ are distinct and at most $E + h - s$, so there are at most $S(E + h - s)$ terms. For the row form use (i) first:

$$A_s = \sum_{i=0}^{2h} \sum_{\sigma \in \mathcal{L}(E+s-h+i)} (-1)^i\,[t^\sigma]\big(a_x\langle i\rangle[s - h + i]\,t^y\big) \cdot z_v^{-1}[t^u]\big(t^v[h - s - i]\,t^\sigma\big),$$

whose factors depend on (x, y) and on (u, v), with at most $S(E + h + s)$ terms. For $s \leq 0$ use the row form and for $s > 0$ the column form. Each A_s then uses at most $S(E + h)$ terms, and there are $2E + 2h + 1$ values of s. □

Lemma 5. *Let $c_0 = \pi\sqrt{2/3}$ and $c_p = c_0(1-1/p)$ for primes p. Then $\ln S(E) = (c_p + o(1))\sqrt{E}$ as $E \to \infty$.*

Proof. For $p = 0$ this is Hardy–Ramanujan. For a prime p, with $F(w) = \prod_{k\geq 1}(1 - w^k)^{-1}$,

$$\sum_E q(E)w^E = \prod_{p\nmid k} \frac{1 - w^{pk}}{1 - w^k} = \frac{F(w)\,F(w^{p^2})}{F(w^p)^2}.$$

Since $\ln F(e^{-\tau}) = \pi^2/(6\tau) + O(\ln(1/\tau))$ as $\tau \downarrow 0$, the logarithm of this product at $w = e^{-\tau}$ is $(1 + o(1))\frac{\pi^2}{6\tau}(1 - \frac{1}{p})^2$. Kohlbecker's Tauberian theorem [5] (see also [1, Section 4.12]) then gives the claim, as $2\sqrt{\pi^2(1 - 1/p)^2/6} = c_p$. □

Proof of Theorem 1. For $n \geq 2$ let E_1 be the least integer with $S(E_1)/(E_1 + 1) \geq n$, and let E_2 be the same for m. By averaging there are $h \leq E_1$ with $q(h) \geq n$ and $E \leq E_2$ with $q(E) \geq m$, and $h \geq 1$ because $q(0) = 1$. (We may assume $n \geq 2$, since T is symmetric in the roles of n and m and $P_{\mathbb{F}}(1, 1) = 1$.) By monotonicity and Proposition 4,

$$P_{\mathbb{F}}(n, m) \leq (2E_1 + 2E_2 + 1)\,S(E_1 + E_2).$$

By Lemma 5 and the minimality of E_1, E_2: for every $\delta > 0$ there is C_δ with $c_p\sqrt{E_1} \le (1+\delta)\ln n + C_\delta$ and $c_p\sqrt{E_2} \le (1+\delta)\ln m + C_\delta$, hence $c_p\sqrt{E_1+E_2} \le (1+\delta)\sqrt{\ln^2 n + \ln^2 m} + \sqrt{2}C_\delta$. Lemma 5 also gives $\ln S(E_1+E_2) \le (1+\delta)c_p\sqrt{E_1+E_2} + C'_\delta$, and the polynomial factor is negligible. This proves the bound.

For $p = 0$ and $n = m$, let h be least with $p(h) \ge n$, where $p(\cdot)$ is the partition function, and take $E = h$. Then $P_{\mathbb{F}}(n,n) \le (4h+1)S(2h)$ and $S(2h) \le (2h+1)p(2h) = O(e^{c_0\sqrt{2h}})$ by Hardy–Ramanujan. Also $p(h-1) < n$ gives $c_0\sqrt{h} \le \ln n + \ln h + O(1)$. So $P_{\mathbb{F}}(n,n) = O(h\, n^{\sqrt{2}} h^{\sqrt{2}})$ with $h = O(\ln^2 n)$, which is $O(n^{\sqrt{2}}(\ln n)^{2+2\sqrt{2}})$. □

3 Composition formulas

Proof of Theorem 2. Multiply $T_{n,m} = \sum_k g_k$ by $x_x y_y x_v y_u$ and sum over all indices, in commuting variables. The left side becomes $\sum_{x,y} x_x^2 y_y^2$. A term $a(x,y)b(u,v)$ gives the product of bilinear forms $(\sum a(x,y)x_x y_y)(\sum b(u,v)x_v y_u)$, and a term $c(x,u)e(y,v)$ gives $(\sum c(x,u)x_x y_u)(\sum e(y,v)x_v y_y)$. This proves $s \le P_{\mathbb{F}}(n,m)$. If $\operatorname{char}\mathbb{F} \ne 2$ then $ff' = (\frac{f+f'}{2})^2 - (\frac{f-f'}{2})^2$, and if $-1 = \sum_{i\le \mathsf{r}} \alpha_i^2$ the last term is r squares. In odd positive characteristic -1 is a sum of two squares in the prime field. □

The rectangular bound can be compared with [3, Theorem 11]. For $m = n^\beta$ it gives $\sigma_{\mathbb{F}}(n, n^\beta) \le n^{\sqrt{1+\beta^2}+o(1)}$, so the excess over the trivial lower bound n is about $n^{\beta^2/2}$ for small β. Hrubeš's Problem 1 asks for a parity representation that would give $n^{1.5+o(1)}$; the bound here is already below that exponent, by a different route.

4 Non-commutative circuits

We use the circuit model of [3]: $\operatorname{size}_{\mathbb{F}}^{(nc)}$ is the least size of a non-commutative arithmetic circuit computing the given polynomials simultaneously. Let $\mathrm{MM}(a,b,c)$ be the commutative circuit size of the product of an $a \times b$ matrix by a $b \times c$ matrix of variables. We take from [3] two facts, both valid over every field: Baur–Strassen, and the conversion in Lemma 20 and Remark 21 of that paper. A set of bilinear forms in $\bar{x}, \bar{y}$ with a commutative circuit of size C has non-commutative xy-ordered and yx-ordered versions computable in size $O(C)$. The polynomial $\sum_k f_k \star z_k$, where $(\sum a_{ij}x_i y_j) \star z = \sum a_{ij} x_i z y_j$, has non-commutative size $O(C+r)$, where r is the number of z variables.

Lemma 6 (after [3, Lemma 24])**.** *Let Q be a set of bilinear forms in $\bar{x}$ (n variables) and $\bar{y}$ (ℓ variables), let $\bar{y}_1, \dots, \bar{y}_m$ be copies of $\bar{y}$, and let $1 \le K \le |Q|$. All $f(\bar{x}, \bar{y}_t)$, $f \in Q$, $t \in [m]$, can be computed by a commutative circuit of size $O\big(\lceil |Q|/K \rceil (\mathrm{MM}(\ell, m, K) + n\ell K)\big)$.*

Proof. Split Q into groups of K forms. For a group $f_k = \sum a_{ijk} x_i y_j$, $k \in [K]$, introduce variables $z_{t,k}$ and consider

$$\sum_{k,t} f_k(\bar{x}, \bar{y}_t) z_{t,k} = \sum_{j,k} \Big(\sum_i a_{ijk} x_i \Big) W_{jk}, \qquad W_{jk} = \sum_t y_{j,t} z_{t,k}.$$

The W_{jk} are entries of a matrix product. The ℓK linear forms in $\bar{x}$ cost $O(n\ell K)$ together. By Baur–Strassen the partial derivatives with respect to $z_{t,k}$, which are the $f_k(\bar{x}, \bar{y}_t)$, cost a constant factor more. □

Lemma 7. *Suppose $T_{n,\ell}$ is a sum of P functions of row or column type over $\mathbb{F}$. Let Q be the corresponding set of at most $2P$ bilinear forms, as in the proof of Theorem 2, and let C be the commutative size of all $f(\bar{x},\bar{y}_t)$, $f \in Q$, $t \in [m]$. Then $\mathrm{size}_{\mathbb{F}}^{(nc)}(\mathsf{ID}_{n,\ell m}) = O(C + Pm)$.*

Proof. In non-commuting variables,

$$\mathsf{ID}_{n,\ell} = \sum_{x,y,u,v} T_{n,\ell}(x,y,u,v)\, x_x y_y x_v y_u.$$

A row term $a(x,y)b(u,v)$ contributes the product $\big(\sum a(x,y)x_x y_y\big)\big(\sum b(u,v)x_v y_u\big)$ of two xy-ordered forms. A column term $c(x,u)e(y,v)$ contributes $\big(\sum c(x,u)x_x y_u\big) \star \big(\sum e(y,v)y_y x_v\big)$, which is an xy-form with a yx-ordered form substituted for z. Also $\mathsf{ID}_{n,\ell m} = \sum_t \mathsf{ID}_{n,\ell}(\bar{x},\bar{y}_t)$. So we sum at most Pm row products, and compute one polynomial $\sum f \star z$ in at most Pm variables z, then substitute. No division by two is used, so the argument works in every characteristic. □

Proof of Theorem 3. Note first that $\mathsf{ID}_{n'n'',N}$ is a sum of n'' copies of $\mathsf{ID}_{n',N}$ on disjoint blocks of $\bar{x}$, and ID_n is a restriction of ID_N for $n \le N$.

(b) Let $\omega_3 = \omega(1,3,1)$. Take $n = \ell^2$, $m = \ell^3$, $K = \ell$. Theorem 1 gives $P = P_{\mathbb{F}}(\ell^2,\ell) \le \ell^{\sqrt{5}+o(1)}$. Lemma 6 gives $C \le \ell^{\sqrt{5}-1+o(1)}(\ell^{\omega_3+o(1)} + \ell^4) = \ell^{\sqrt{5}-1+\omega_3+o(1)}$, since $\omega_3 \ge 4$. Also $Pm \le \ell^{\sqrt{5}+3+o(1)}$, which is below this bound on C. By Lemma 7, $\mathrm{size}_{\mathbb{F}}^{(nc)}(\mathsf{ID}_{\ell^2,\ell^4}) \le \ell^{\sqrt{5}-1+\omega_3+o(1)}$, and with ℓ^2 blocks of $\bar{x}$, $\mathrm{size}_{\mathbb{F}}^{(nc)}(\mathsf{ID}_{\ell^4}) \le \ell^{1+\sqrt{5}+\omega_3+o(1)}$. With $N = \ell^4$ this is the claimed exponent, and $(1+\sqrt{5}+4.1989)/4 < 1.8588$.

(a) Take $\ell = L^5$, $n = \ell^6$, $m = \ell^{6.2} = L^{31}$, $K = \ell$. Theorem 1 gives $P \le \ell^{\sqrt{37}+o(1)} \le \ell^{6.0828}$ for large ℓ. Strassen's algorithm, on m/ℓ blocks, gives $\mathrm{MM}(\ell,m,\ell) = O((m/\ell)\ell^{\log_2 7}) \le \ell^{8.0075}$. Also $n\ell K = \ell^8$. So $C \le \ell^{5.0828+8.0075}$, and $Pm \le \ell^{12.29}$ is below this bound. Hence $\mathrm{size}_{\mathbb{F}}^{(nc)}(\mathsf{ID}_{\ell^6,\ell^{7.2}}) \le \ell^{13.0903}$. With $\ell^{1.2}$ blocks of $\bar{x}$ we get $\mathrm{size}_{\mathbb{F}}^{(nc)}(\mathsf{ID}_{\ell^{7.2}}) \le \ell^{14.2903}$. Since $14.2903/7.2 < 1.98477$, this proves (a). Strassen's algorithm and all other ingredients work over every field.

The statements for $S_{k,n}$ follow as in [3, Proposition 26, Corollary 27]. □

Remark 1. The value $\omega(1,3,1) < 4.1989$ is from [11]; Hrubeš uses about 4.1997 from [7] (the value there is 4.199712). Laser-method bounds are generally understood to hold over every field, but we did not find this stated for these specific rectangular bounds, so part (b) is conditioned on it. Part (a) does not use them. Optimising the split in the proof of (b) over real parameters gives an exponent near 1.854.

5 Numerics and checks

Lemmas 9 and 10 of [2] are listed as formalized in the Lean development accompanying that paper; we did not build or check it, and nothing in this note has been formalized. We verified Proposition 4 by exact computation with an implementation written from the definitions (scripts `check_pieces.py`, `check_explicit.py`). For $E = h$: over $\mathbb{Z}$ for $h \le 6$, over $\mathbb{F}_3$ for $h \le 10$, and over $\mathbb{F}_2$ for $h \le 16$ ($h = 16$ has five labels), we checked the exchange property (i) for every s, the identity $\sum_s A_s = [x = v][y = u]$ entry by entry, and the rank bounds in both directions. The explicit decomposition into rank-one terms was also checked, including rectangular cases $E \ne h$: over $\mathbb{Z}$ up to $(h,E) = (5,3)$ and $(4,4)$, over $\mathbb{F}_2$ up to $(12,8)$, over $\mathbb{F}_3$ up to $(6,6)$, over $\mathbb{F}_5$ up to $(6,3)$, and over $\mathbb{F}_7$ at $(5,5)$.

h	$\log_2 N$	exponent of $\sum_s \min(R_s, O_s)$	exponent of $p(2h)$
10	5.4	2.639	1.723
40	15.2	2.020	1.574
160	36.6	1.739	1.502
640	81.5	1.594	1.464
2500	170.9	1.516	1.443
10000	354.0	1.471	1.430

Table 1: Characteristic zero, $N = p(h)$: the exponent $\log_N$ of the bound on $P(N)$ from Proposition 4 with exact partition numbers. Here $R_s = \sum_{i=0}^{2h} p(s+i)$ and $O_s = \sum_{i=0}^{2h} p(2h - s - i)$ are the term counts for $E = h$. The limit is $\sqrt{2} \approx 1.4142$. The bound passes Hrubeš's 1.62 near $N = 2^{68}$.

References

[1] N. H. Bingham, C. M. Goldie, J. L. Teugels. *Regular Variation*. Cambridge University Press, 1987.

[2] Claude (Anthropic). A $2^{o(n)}$ time algorithm for 3-SAT. Preprint, October 2026. `https://github.com/anthropics/formal-math/tree/main/eth`.

[3] P. Hrubeš. A subquadratic upper bound on Hurwitz's problem and related non-commutative polynomials. ECCC TR24-099, 2024 (we use its numbering); *SIAM J. Comput.* 55(2):451–467, 2026. Earlier version in CCC 2024.

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

[5] E. E. Kohlbecker. Weak asymptotic properties of partitions. *Trans. Amer. Math. Soc.* 88:346–365, 1958.

[6] H. Li, Q. Mu. Heisenberg VOAs over fields of prime characteristic and their representations. *Trans. Amer. Math. Soc.* 370:1159–1184, 2018. arXiv:1501.04314.

[7] F. Le Gall, F. Urrutia. Improved rectangular matrix multiplication using powers of the Coppersmith–Winograd tensor. SODA 2018.

[8] E. Naslund. The partition rank of a tensor and k-right corners in $\mathbb{F}_q^n$. *J. Combin. Theory Ser. A* 174, 2020.

[9] B. Rossman, D. Zhu. Multiquadratic sum-of-squares lower bounds imply $\mathsf{VNC}^1 \neq \mathsf{VNP}$. arXiv:2512.01227, 2025.

[10] C. Zhang, H. Zhu. The Hurwitz sum-of-squares problem depends on the base field. arXiv:2605.00590, 2026.

[11] V. Vassilevska Williams, Y. Xu, Z. Xu, R. Zhou. New bounds for matrix multiplication: from alpha to omega. SODA 2024.