

An Upper Bound of $9/4$ for the Matrix Multiplication Exponent

OpenAI

October 2, 2026

Abstract

We prove that the exponent of matrix multiplication over the complex numbers is at most $9/4$.

1 Introduction

The exponent ω of matrix multiplication over $\mathbb{C}$ is the infimum of the real numbers τ such that, for every $\varepsilon > 0$, two $n \times n$ matrices can be multiplied in $O_\varepsilon(n^{\tau+\varepsilon})$ scalar arithmetic operations. The dimension n tends to infinity; the algorithm and its constants may depend on ε . We prove the following bound.

Theorem 1.1. *For every $\varepsilon > 0$, two $n \times n$ complex matrices can be multiplied using $O_\varepsilon(n^{9/4+\varepsilon})$ arithmetic operations. In particular, $\omega \leq 9/4$.*

Strassen’s seven-product algorithm for two-by-two matrices showed that recursion can reduce the arithmetic exponent below three [13]. Bilinear algorithms can be represented by trilinear coefficient arrays, or tensors; linear substitutions and tensor products then express reductions and recursive composition. Approximate bilinear algorithms enlarge the available constructions, and Bini’s interpolation argument converts them into exact asymptotic algorithms [3]. Schönhage’s asymptotic sum inequality turns efficient simultaneous computations of independent matrix products into bounds on the exponent [11]. Strassen’s laser method extracts such independent products from tensor powers [14, 15]. His asymptotic spectrum describes asymptotic tensor comparisons through numerical invariants compatible with sums, products, and restrictions [16, 4].

Coppersmith and Winograd combined tensor powers with progression-free sets to extract independent products from an auxiliary form [5]. Analyses of higher tensor powers by Stothers [12] and Davie–Stothers [6], Vassilevska Williams [18], and Le Gall [10] developed this construction; Le Gall formulated its analysis through convex optimization. Alman and Vassilevska Williams refined the extraction of independent components [2]. Duan, Wu, and Zhou analyzed losses caused by combining components across recursion levels [7]. Further asymmetric analyses by Vassilevska Williams, Xu, Xu, and Zhou [19] and Alman et al. [1] improved the resulting bounds. Dupont et al. report $\omega < 2.371177$ by optimizing how components are combined across recursion levels [8].

Our proof combines explicit tensor degenerations with Strassen’s spectral viewpoint [16, 4]. Tensor characters are normalized numerical invariants that add on direct sums, multiply on tensor products, and are monotone under restriction. By Lemma 2.2, it suffices to bound every character on matrix multiplication. Each character has value n^{3t} there for a positive parameter t . We bound t through the growth of the character’s values on polynomial multiplication. As in earlier tensor-power

methods, we use interpolation and count words in which each block label occurs a prescribed number of times.

The three variable groups of a trilinear form are called legs. Useful tensor blocks can share variables, whereas a full direct sum requires independence on all three legs. Our principal construction separates blocks that are independent on two legs and share the third. We give the shared-leg variables tentative block labels, then sum substitutions with root-of-unity coefficients to cancel terms unless their indices satisfy a linear relation. We then assign integer weights to the variables so that the three weights in each surviving term sum to the square of the difference between its tentative label and its actual block label. Retaining the terms of weight zero therefore makes the labels agree. Each resulting independent block is tensored with an auxiliary dot product. When we apply this construction to an N th tensor power, the numbers of source copies and output blocks have a fixed ratio, independent of N . Its N th root tends to one. [Proposition 3.1](#) gives the construction, and [Corollary 3.2](#) uses it to bound character values by counting blocks with prescribed frequencies.

For polynomial multiplication with input coefficient vectors of lengths a, b , the output vector has length $a + b - 1$. We measure this tensor by a normalized geometric mean of a character's values under all six permutations of the legs. The resulting profile $P(a, b)$ is positive and symmetric in the input lengths, and polynomial interpolation gives $P(a, b) \leq (a + b - 1)^{1/t}$.

Two applications of the separation inequality constrain the growth of P . First, tensoring with a two-dimensional dot product enlarges the second-input and output coefficient spaces. We choose a subspace of each so that multiplication by every first-input polynomial maps the input subspace into the output subspace. Multiplication on these subspaces and on their quotients gives the two neighboring input lengths. Bases adapted to these subspaces allow one choice of weights to remove the terms between the two blocks for every first input. The separation inequality then shows that successive increments of P do not increase in either input length; [Section 4.1](#) constructs the subspaces using a determinant.

For the second application, we partition the second-input and output coefficient indices into three matched intervals and retain only terms between matched intervals. The three resulting blocks are copies of polynomial multiplication with input lengths a, h , up to a permutation of the legs. Since P combines all leg permutations, the separation inequality gives

$$P(a, 3h + a - 1) \geq 3P(a, h).$$

Together with $P(1, b) = b$, these properties force $P(a, a) \geq a^{4/3}$. Comparing this growth with the interpolation bound yields $t \leq 3/4$, hence $3t \leq 9/4$. [Theorem 1.1](#) concerns the asymptotic arithmetic exponent; the proof does not specify a competitive finite matrix size.

[Section 2](#) introduces the character framework and states the required existence result. [Section 3](#) proves the separation inequality, and [Section 4](#) derives the two polynomial inequalities. [Section 5](#) proves the discrete growth estimate and completes the arithmetic argument. [Appendix A](#) gives a complete proof of the character-existence result used here.

2 Tensors and their characters

Throughout, tensors are finite trilinear forms over $\mathbb{C}$. Their three variable groups are called the X -, Y -, and Z -legs. A restriction applies an independent linear substitution on each leg. We write $A \geq B$ if B is a restriction of A .

In a full direct sum, the summands have disjoint variables on all three legs. The tensor product multiplies coefficient arrays and pairs the corresponding variable indices on each leg; we write it by juxtaposition or by $\otimes$. These operations make the set S of tensors modulo mutual restriction

a commutative semiring, with the compatible order induced by restriction. The zero tensor is 0, the scalar tensor xyz is 1, and the integer $m \geq 0$ denotes the direct sum of m copies of 1. Thus $mA = A^{\oplus m}$.

The tensor rank $R(A)$ is the least integer m with $m \geq A$. It is finite by expansion in bases. Every nonzero tensor restricts to 1, by evaluating its three arguments on suitable lines, so

$$R(A) \geq A \geq 1 \quad (A \neq 0). \quad (2.1)$$

Rank is monotone, subadditive, and submultiplicative. We do not assume its additivity. A *leg flattening* is the matrix obtained by placing that leg against the pair of other legs. Its matrix rank is restriction-monotone, additive on full direct sums, and multiplicative on tensor products.

The trilinear form encoding $n \times n$ matrix multiplication and its exact-rank exponent are

$$T_n = \sum_{i,j,k=1}^n x_{ij}y_{jk}z_{ki}, \quad \nu = \inf_{n \geq 2} \log_n R(T_n). \quad (2.2)$$

Reindexing identifies $T_n T_m$ with T_{nm} . Flattening and the displayed expansion give $n^2 \leq R(T_n) \leq n^3$, so $2 \leq \nu \leq 3$ and $R(T_n) \geq n^\nu$. At the end of the proof we convert the bound on ν into the asserted arithmetic bound.

Definition 2.1. A *tensor character* is an additive, multiplicative, restriction-monotone map $\lambda : S \rightarrow \mathbb{R}_{\geq 0}$ with $\lambda(1) = 1$.

The three flattening ranks are characters, and every character satisfies $\lambda(m) = m$ and $1 \leq \lambda(A) \leq R(A)$ for $A \neq 0$. Characters are the spectral points of the tensor semiring [16, 4]. We need only the following special case of spectral separation, proved in [Appendix A](#) using finite dimensional separation, compactness, and a fixed-point theorem.

Lemma 2.2 (Detecting characters). *Let $d \geq 2$ and $k \geq 0$ be integers with $k < d^\nu$, where ν is defined by (2.2). There is a tensor character λ with $\lambda(T_d) \geq k$.*

2.1 Dot-product exponents

For $m \geq 1$, define

$$B_X(m) = x \sum_{i=1}^m y_i z_i,$$

and obtain the other two oriented dot products $B_Y(m), B_Z(m)$ by cyclically permuting the legs. For any character there are numbers $p_X, p_Y, p_Z \in [0, 1]$ such that

$$\lambda(B_X(m)) = m^{p_X}, \quad \lambda(B_Y(m)) = m^{p_Y}, \quad \lambda(B_Z(m)) = m^{p_Z}. \quad (2.3)$$

Indeed, $f(m) = \lambda(B_X(m))$ is positive and nondecreasing, satisfies $f(mn) = f(m)f(n)$, and lies between 1 and m . Put $p_X = \log_2 f(2)$. For $r \geq 1$, let $\ell = \lfloor r \log_2 m \rfloor$. Comparison with $2^\ell \leq m^r < 2^{\ell+1}$ gives $f(2)^\ell \leq f(m)^r \leq f(2)^{\ell+1}$. Taking logarithms and letting $r \rightarrow \infty$ proves the first identity; the other two follow in the same way.

The product $B_X(m)B_Y(m)B_Z(m)$ has variables x_{bc}, y_{ac}, z_{ab} and monomials $x_{bc}y_{ac}z_{ab}$, so it is isomorphic to T_m . Consequently

$$\lambda(T_m) = m^{p_X + p_Y + p_Z}. \quad (2.4)$$

Setting all off-diagonal variables of T_m to zero leaves m independent scalar products. Hence $T_m \geq m$, and $p_X + p_Y + p_Z \geq 1$. In particular, the parameter $t = (p_X + p_Y + p_Z)/3$ used below is always positive.

2.2 Degenerations and interpolation

We use only the following elementary form of degeneration. After fixed linear substitutions, give each variable an integer weight, multiply it by the corresponding power of a parameter ε , and retain the terms of minimum total weight. Write $A \rightsquigarrow B$ if this procedure takes A to B , with the convention $0 \rightsquigarrow 0$. Although individual weights may be negative, shifting the minimum total weight to zero gives a polynomial in ε with constant coefficient B .

Lemma 2.3 (Interpolation). *If $A \rightsquigarrow B$, then $\lambda(A) \geq \lambda(B)$ for every tensor character λ .*

Proof. This is the interpolation passage from approximate to exact bilinear algorithms [3], applied to characters. Let $F(\varepsilon) = B + \varepsilon F_1 + \cdots + \varepsilon^L F_L$ be the transformed tensor. For each nonzero ε , it is a restriction of A . The polynomial $F(\varepsilon)^{\otimes j}$ has degree at most jL and constant coefficient $B^{\otimes j}$. Interpolation at $jL + 1$ distinct nonzero points writes this coefficient as a linear combination of specializations. A linear combination of tensors on common spaces is a restriction of their full direct sum: identify the corresponding variables and absorb each scalar coefficient into one leg. Therefore

$$\lambda(B)^j \leq (jL + 1)\lambda(A)^j.$$

Taking j th roots and letting $j \rightarrow \infty$ proves the result. The degeneration, and hence L , is fixed during this limit. $\square$

3 Separating a shared tensor leg

Suppose tensor blocks share their first leg and have disjoint variables on each of the other two legs. We separate the first leg by a degeneration that also tensors each block with a dot product. Each block's character value is therefore multiplied by the value of that dot product. [Corollary 3.2](#) applies this construction to tensor powers and counts the blocks with prescribed frequencies.

Proposition 3.1 (Finite separation). *Let $M \geq 1$ and*

$$A = \sum_{h=1}^M A_h, \quad A_h \in X \otimes Y_h \otimes Z_h,$$

where the Y_h are pairwise disjoint coordinate spaces and so are the Z_h . Then

$$A^{\oplus 5M} \rightsquigarrow \bigoplus_{h=1}^M (A_h \otimes B_X(M)). \quad (3.1)$$

Each A_h on the right has its own copy of X .

Proof. Choose a common first-leg basis x_a and sector bases $y_{h,b}, z_{h,c}$, and write

$$A = \sum_{h,a,b,c} c_{h,a,b,c} x_a y_{h,b} z_{h,c}.$$

The internal index ranges may depend on h . Set $L = 5M$, choose a primitive L th root of unity ζ , and index the source copies by $r = 0, \dots, L - 1$. Introduce target variables $X_{a,g}, Y_{h,b,u}, Z_{h,c,v}$, where

$g, u, v \in \{1, \dots, M\}$. Here g is a tentative sector label, and u, v will supply the dot product. Make the substitutions

$$\begin{aligned} x_a^{(r)} &\mapsto \sum_{g=1}^M \zeta^{2rg} X_{a,g}, \\ y_{h,b}^{(r)} &\mapsto \sum_{u=1}^M \zeta^{r(u-h)} Y_{h,b,u}, \\ z_{h,c}^{(r)} &\mapsto \frac{1}{L} \sum_{v=1}^M \zeta^{r(-v-h)} Z_{h,c,v}. \end{aligned}$$

Within each source copy, each phase depends only on the indices of its target variable. All copies map into common target spaces, so summing over r multiplies each target coefficient by

$$\frac{1}{L} \sum_{r=0}^{L-1} \zeta^{r[u-v+2(g-h)]}.$$

This Fourier sum is 1 when $u-v+2(g-h)$ is divisible by L and 0 otherwise. Since $|u-v+2(g-h)| \leq 3(M-1) < L$, precisely the terms satisfying

$$u - v + 2(g - h) = 0 \tag{3.2}$$

survive.

To force the tentative label g to equal h , give the target variables the weights

$$w(X_{a,g}) = g^2, \quad w(Y_{h,b,u}) = hu - h^2, \quad w(Z_{h,c,v}) = -hv.$$

On every surviving term, (3.2) gives

$$g^2 + hu - h^2 - hv = (g - h)^2 \geq 0.$$

The weight-zero tensor is therefore

$$D = \sum_{h,a,b,c,u} c_{h,a,b,c} X_{a,h} Y_{h,b,u} Z_{h,c,u}.$$

The variables are disjoint across h on all three legs, and the common u index supplies $B_X(M)$ within each block. Thus D is the full direct sum in (3.1). The transformed tensor has degree at most $(M-1)^2$, even though some individual variable weights are negative. $\square$

Corollary 3.2 (Shared-leg entropy inequality). *Let $T = \sum_{i=1}^s T_i$ be a sum of nonzero tensors with a common first space and matched, pairwise disjoint sectors on the other two legs, as in Proposition 3.1. For every tensor character λ and every probability vector $q = (q_1, \dots, q_s)$,*

$$\lambda(T) \geq e^{px H(q)} \prod_{i=1}^s \lambda(T_i)^{q_i}, \quad H(q) = - \sum_{i=1}^s q_i \log q_i. \tag{3.3}$$

Here logarithms are natural and $0 \log 0 = 0$.

Proof. We apply Proposition 3.1 to words of a prescribed frequency, using the counting argument familiar from tensor-power constructions [10, Appendix A of the full version]. First take q rational and let N run through multiples of a common denominator. On the second and third legs of $T^{\otimes N}$,

retain only words with Nq_i occurrences of sector i . Every supported term has the same sector word on these two legs. The retained tensor A therefore has

$$M = \frac{N!}{\prod_i (Nq_i)!}$$

matched sectors, with the first leg still shared. Each sector is a product of the prescribed T_i , up to reordering factors, and hence has character value $C_N = \prod_i \lambda(T_i)^{Nq_i}$.

The separation uses $5M$ source copies and produces M independent blocks, each of character value $M^{p_X} C_N$: the dot-product index lies within each block. Applying [Proposition 3.1](#) and [Lemma 2.3](#) at fixed N gives

$$5M\lambda(T)^N \geq 5M\lambda(A) \geq M^{1+p_X} C_N.$$

The interpolation limit in [Lemma 2.3](#) is taken before varying N , so no degree bound uniform in N is needed. Cancel M and take N th roots. Stirling's formula gives $N^{-1} \log M \rightarrow H(q)$, proving [\(3.3\)](#) for rational q . Finally, $\lambda(T_i) \geq 1$, so the right side of [\(3.3\)](#) is continuous on the probability simplex. Rational approximation proves the result for all q . $\square$

4 Two inequalities for polynomial multiplication

We apply the shared-leg inequality to polynomial multiplication, whose tensor rank is explicit. For positive integers a, b , let

$$C(a, b) = \sum_{i=0}^{a-1} \sum_{j=0}^{b-1} x_i y_j z_{i+j}. \quad (4.1)$$

This is multiplication of homogeneous binary forms of degrees $a-1$ and $b-1$, paired with the dual of the output space. Evaluation at $a+b-1$ distinct points followed by interpolation gives $R(C(a, b)) \leq a+b-1$. Its output flattening has rank $a+b-1$, since every output monomial is a product of input monomials. Hence

$$R(C(a, b)) = a + b - 1. \quad (4.2)$$

Fix a character λ , put $t = (p_X + p_Y + p_Z)/3 > 0$, and let $\lambda_\pi(A)$ denote its value after permuting the legs of A by $\pi \in S_3$. Each λ_π is again a character; denote its first-leg dot-product exponent by $p_X^{(\pi)}$. Each of the three original exponents occurs twice, so

$$\sum_{\pi \in S_3} p_X^{(\pi)} = 2(p_X + p_Y + p_Z) = 6t. \quad (4.3)$$

Products over permuted legs are used in tensor-power arguments [[10](#), Appendix A.3 of the full version]. Here we take all six permutations, so each original leg occurs equally often in each position, and define the *symmetrized profile*

$$P(a, b) = \left(\prod_{\pi \in S_3} \lambda_\pi(C(a, b)) \right)^{1/(6t)}. \quad (4.4)$$

Exchanging the input legs permutes the factors, while $C(1, b) = B_X(b)$ determines the boundary values. These identities, [\(4.2\)](#), and [\(4.3\)](#) give

$$P(a, b) = P(b, a) > 0, \quad P(1, b) = b, \quad P(a, b) \leq (a + b - 1)^{1/t}. \quad (4.5)$$

The next two lemmas show that successive increments of P in either argument are nonincreasing, and that, with the first argument a fixed, replacing the second argument h by $3h + a - 1$ increases the value at least threefold. [Lemma 5.1](#) will use these two constraints to force growth along the diagonal, which can then be compared with the rank bound in [\(4.5\)](#).

4.1 A determinant filtration

Tensoring with $B_X(2)$ doubles the second-input and output spaces. Multiplication on a quotient and a kernel of these spaces will produce the neighboring input lengths $b + 1$ and $b - 1$. To obtain a degeneration of the whole tensor, the bases and weights must work for every first input.

Lemma 4.1 (Discrete concavity). *For $a \geq 1$ and $b \geq 2$,*

$$2P(a, b) \geq P(a, b + 1) + P(a, b - 1). \quad (4.6)$$

By symmetry, the analogous inequality holds in the first argument.

Proof. Let V_e be the degree- e homogeneous forms in u, v , and set $W_1 = \text{span}(s, w)$ for a second pair of variables. The tensor $C(a, b) \otimes B_X(2)$ represents multiplication

$$V_{a-1} \times (V_{b-1} \otimes W_1) \longrightarrow V_{a+b-2} \otimes W_1.$$

Put $D = uw - vs$. For $e \geq 1$, diagonal substitution $(s, w) = (u, v)$ gives the exact sequence

$$0 \longrightarrow DV_{e-1} \longrightarrow V_e \otimes W_1 \longrightarrow V_{e+1} \longrightarrow 0. \quad (4.7)$$

Diagonal substitution is surjective on monomials, so its kernel has dimension e . That kernel contains DV_{e-1} , which also has dimension e because multiplication by D is injective. This proves exactness and gives the determinant exact sequence underlying the degree-one Clebsch–Gordan construction [\[9, Section 2.6 of the author version\]](#).

Choose quotient lifts $u^{e-i}v^is$ for $0 \leq i \leq e$, followed by v^ew , and append the kernel basis $Du^{e-1-j}v^j$ for $0 \leq j < e$. Their diagonal images and their kernel factors are the standard monomial bases. Fix these bases at degrees $b - 1$ and $a + b - 2$, independently of the first input. For every $f \in V_{a-1}$, multiplication commutes with diagonal substitution and sends Dg to $D(fg)$. Thus all slices have the form

	input quotient	input kernel
output quotient	M_f^+	0
output kernel	*	M_f^-

(4.8)

in these fixed bases. Here M_f^+ is multiplication $V_b \rightarrow V_{a+b-1}$, and M_f^- is multiplication $V_{b-2} \rightarrow V_{a+b-3}$. In particular, the two diagonal tensor blocks are exactly $C(a, b + 1)$ and $C(a, b - 1)$.

In these bases, give weight zero to the first leg and to all quotient coordinates, weight -1 to the second-leg kernel coordinates, and weight $+1$ to the output-dual kernel coordinates. The only potentially negative block in [\(4.8\)](#) is zero, while the block marked $*$ has weight $+1$ and disappears. The weight-zero tensor therefore consists of $C(a, b + 1)$ and $C(a, b - 1)$, with a common first leg and disjoint second and third legs. For example, when $a = b = 2$, the identity $u(vw) = v^2s + vD$ shows that multiplication of a quotient lift can have both quotient and kernel components.

Applying degeneration monotonicity and [Corollary 3.2](#) to these two blocks, for any probability pair $q = (q_+, q_-)$ we obtain

$$2^{p_X^{(\pi)}} \lambda_\pi(C(a, b)) \geq e^{p_X^{(\pi)} H(q)} \lambda_\pi(C(a, b + 1))^{q_+} \lambda_\pi(C(a, b - 1))^{q_-}.$$

Multiply all six inequalities at the same q and take the power $1/(6t)$. By (4.3), this gives

$$2P(a, b) \geq e^{H(q)} P(a, b+1)^{q_+} P(a, b-1)^{q_-}.$$

For positive A, B , the maximum of $e^{H(q)} A^{q_+} B^{q_-}$ is $A + B$, attained at $q = (A, B)/(A + B)$. For positive q_+, q_- , this follows from the weighted arithmetic-geometric mean inequality applied to A/q_+ and B/q_- ; the boundary follows by continuity. Taking this common maximizing pair proves (4.6). $\square$

	Z-index					
	0	1	2	3	4	
$y = 0$	x_0	x_1				Shaded: weight 0 Crossed: weight +1 Blank: no term
$y = 1$		~~x_0~~	x_1			
$y = 2$			x_0	~~x_1~~		
$y = 3$				x_0	x_1	

Figure 1: The degeneration of $C(2, 4)$, with $a = 2$, $h = 1$. A cell (y, z) contains x_i when $z = y + i$. The three shaded blocks survive, and the two crossed terms disappear. The blocks have disjoint Y - and Z -coordinates but share x_0, x_1 . The middle block is $C(2, 1)$ with its second and third legs exchanged and its first-leg basis reversed.

4.2 Three sectors

Lemma 4.2 (Shifted tripling). *For all positive integers a, h ,*

$$P(a, 3h + a - 1) \geq 3P(a, h). \quad (4.9)$$

Proof. Put $B = 3h + a - 1$. A supported term of $C(a, B)$ has indices $0 \leq x < a$, $0 \leq y < B$, and $z = x + y$. Partition the Y -indices and Z -indices as follows:

	left	middle	right
Y	$[0, h - 1]$	$[h, 2h + a - 2]$	$[2h + a - 1, 3h + a - 2]$
Z	$[0, h + a - 2]$	$[h + a - 1, 2h + a - 2]$	$[2h + a - 1, 3h + 2a - 3]$

Give weight +1 to middle Y -coordinates, weight -1 to middle Z -coordinates, and zero to all other coordinates. A left Y -index forces a left Z -index, because $x + y \leq h + a - 2$; a right Y -index forces a right Z -index. Thus no supported term has negative weight. The weight-zero part consists exactly of the three matched sectors. Figure 1 shows the smallest nontrivial example.

Translations identify the two outer branches with $C = C(a, h)$. The middle branch is

$$\sum_{u=0}^{a-1} \sum_{r=0}^{h-1} x_{a-1-u} y_{h+u+r} z_{h+a-1+r}.$$

It is C with its second and third legs exchanged and its first-leg basis reversed. The reversal identifies this individual branch; no different first-leg maps are applied to the three branches of the whole tensor.

Let σ exchange the second and third legs. The uniform probability vector on the three branches gives

$$\lambda_\pi(C(a, B)) \geq 3^{p_X^{(\pi)}} \lambda_\pi(C)^{2/3} \lambda_\pi(C^\sigma)^{1/3}.$$

Symmetrization removes the distinction between the ordinary and exchanged branches: composition with σ permutes the six leg permutations. Thus $\prod_\pi \lambda_\pi(C^\sigma) = \prod_\pi \lambda_\pi(C)$. Multiplying the six inequalities and taking the power $1/(6t)$ proves (4.9). $\square$

5 Discrete growth and the arithmetic exponent

Concavity and shifted tripling force a lower bound on the profile along the diagonal.

Lemma 5.1 (Diagonal growth). *Let $P : \mathbb{N}_{>0}^2 \rightarrow \mathbb{R}_{>0}$ be symmetric, with $P(1, b) = b$. Suppose it satisfies (4.6) for $a \geq 1$, $b \geq 2$ and (4.9) for $a, h \geq 1$. Then $P(a, a) \geq a^{4/3}$ for every $a \geq 1$.*

Proof. For fixed a , concavity makes the increments $\Delta_{a,h} = P(a, h+1) - P(a, h)$ nonincreasing in h . Set $D_a = P(a, a)$ and $H_a = 2D_a/(3a-1)$. Applying (4.9) with $h = a$ bounds the first increment after the diagonal:

$$2D_a \leq P(a, 4a-1) - D_a = \sum_{h=a}^{4a-2} \Delta_{a,h} \leq (3a-1)\Delta_{a,a}.$$

Thus $\Delta_{a,a} \geq H_a$. Comparing consecutive diagonal values requires one increment in each coordinate, so symmetry and concavity give, for $a \geq 2$,

$$D_a - D_{a-1} = \Delta_{a-1,a-1} + \Delta_{a,a-1} \geq H_{a-1} + H_a. \quad (5.1)$$

Substituting $D_a = (3a-1)H_a/2$ gives a multiplicative recurrence for H_a . Rearranging and iterating from $H_1 = 1$ yields

$$\begin{aligned} H_a &\geq \left(1 + \frac{1}{3(a-1)}\right) H_{a-1} \quad (a \geq 2), \\ H_a &\geq \prod_{m=1}^{a-1} \left(1 + \frac{1}{3m}\right) \quad (a \geq 1). \end{aligned} \quad (5.2)$$

For $a = 1$ the product is empty and equals 1. For $m \geq 1$, the elementary inequality $(1 + 1/(3m))^3 \geq 1 + 1/m$ now gives

$$H_a^3 \geq \prod_{m=1}^{a-1} \left(1 + \frac{1}{3m}\right)^3 \geq \prod_{m=1}^{a-1} \left(1 + \frac{1}{m}\right) = a.$$

Since $(3a-1)/2 \geq a$, we conclude $D_a \geq aH_a \geq a^{4/3}$. $\square$

Proof of Theorem 1.1. For every character, the profile (4.4) satisfies all hypotheses of Lemma 5.1. Combining it with (4.5) gives

$$a^{4/3} \leq P(a, a) \leq (2a-1)^{1/t}.$$

Letting $a \rightarrow \infty$ forces $t \leq 3/4$. Equation (2.4) therefore gives $\lambda(T_d) \leq d^{9/4}$ for every character.

For each integer $d \geq 2$, take $k_d = \lceil d^\nu \rceil - 1$. Then $d^\nu - 1 \leq k_d < d^\nu$, so Lemma 2.2 supplies a character with $\lambda(T_d) \geq k_d$. This character may depend on d , but the upper bound just proved holds for every character. Hence

$$d^\nu \leq d^{9/4} + 1.$$

Taking logarithms and letting $d \rightarrow \infty$ yields $\nu \leq 9/4$.

To obtain an arithmetic algorithm, fix $\varepsilon > 0$ and choose $0 < \delta < \varepsilon$. The definition of ν supplies a fixed integer $u \geq 2$ and an exact rank decomposition of T_u of length $r < u^{9/4+\delta}$. Its bilinear identities remain valid on matrix blocks: their coefficients are central scalars, and each product has its left-input block before its right-input block. For N a power of u , recursive application therefore has cost

$$S(N) \leq rS(N/u) + C_{u,r}(N/u)^2.$$

Flattening gives $r \geq u^2$. Summing the recursion levels gives $S(N) = O(N^{\log_u r})$ if $r > u^2$, and $S(N) = O(N^2 \log N)$ if $r = u^2$. Both are $O_\varepsilon(N^{9/4+\varepsilon})$. Padding an arbitrary n to the next power of u increases its size by less than the fixed factor u and proves the theorem. $\square$

Remark 5.2. In the complex arithmetic model used here, all coefficients of the recursive algorithm are fixed once ε is chosen. They can also be chosen algebraic: the equations for an exact rank decomposition have rational coefficients, so the Nullstellensatz turns existence of a complex solution into existence of one over $\overline{\mathbb{Q}}$. The finitely many resulting coefficients then lie in one number field.

A Existence of detecting characters

We prove [Lemma 2.2](#) directly within the asymptotic-spectrum framework [\[16, 4\]](#). A *state* is a normalized, additive, restriction-monotone map $S \rightarrow \mathbb{R}_{\geq 0}$; it is a character if it is also multiplicative. Every state satisfies $0 \leq \lambda(s) \leq R(s)$ and $\lambda(s) \geq 1$ for $s \neq 0$.

Fix integers $d \geq 2$ and $k \geq 0$ with $k < d^\nu$. To obtain a character with $\lambda(T_d) \geq k$, we will construct states satisfying the stronger family of inequalities

$$\lambda(T_d s) \geq k\lambda(s) \quad (s \in S). \tag{A.1}$$

These inequalities are preserved by normalized tensor multiplication: for any nonzero z , replacing $\lambda(x)$ by $\lambda(zx)/\lambda(z)$ preserves every inequality by applying [\(A.1\)](#) to zs . The next lemma uses this invariance to obtain multiplicativity in a nonempty compact convex family of states. We then prove nonemptiness by deriving a contradiction to $k < d^\nu$ from an empty family.

A.1 Obtaining multiplicativity

Lemma A.1. *Let $\mathcal{K}$ be a nonempty compact convex set of states in the topology of coordinatewise convergence. Suppose that, for every nonzero $z \in S$, the map*

$$(P_z \lambda)(x) = \frac{\lambda(zx)}{\lambda(z)}$$

takes $\mathcal{K}$ into itself. Then $\mathcal{K}$ contains a multiplicative state.

Proof. Since $\lambda(z) \geq 1$, each P_z is continuous in the product topology. The space $\mathbb{R}^S$ is Hausdorff and locally convex, so the Schauder–Tychonoff fixed-point theorem [\[17\]](#) gives a fixed point μ of P_z . Set $c = \mu(z)$. The subset

$$\mathcal{K}' = \{\lambda \in \mathcal{K} : \lambda(zx) = c\lambda(x) \text{ for every } x \in S\}$$

is nonempty and compact. Because c is fixed, its defining equations are linear in λ , so it is convex. Taking $x = 1$ shows that $\lambda(z) = c$ throughout $\mathcal{K}'$, so all its members are fixed by P_z . Moreover,

every P_w with $w \neq 0$ preserves it:

$$(P_w \lambda)(zx) = \frac{\lambda(zwx)}{\lambda(w)} = c \frac{\lambda(wx)}{\lambda(w)} = c(P_w \lambda)(x).$$

Repeating the construction inside $\mathcal{K}'$ retains all earlier equations, so induction gives a common fixed point in $\mathcal{K}$ for every finite collection of maps P_z . Their fixed sets are closed, and compactness then gives a simultaneous fixed point for all nonzero z . It satisfies $\lambda(zx) = \lambda(z)\lambda(x)$ for $z \neq 0$. Additivity implies $\lambda(0) = 0$, which handles multiplication by zero. $\square$

A.2 Constructing the detecting state space

Proof of Lemma 2.2. For the fixed integers d, k above, let $\mathcal{K}$ be the subset of $\prod_{s \in S} [0, R(s)]$ defined by normalization, additivity, monotonicity, and (A.1). The defining conditions are closed affine equalities and inequalities in a compact product, so $\mathcal{K}$ is compact and convex. It remains to prove nonemptiness before applying Lemma A.1.

Obtaining a catalytic comparison

Suppose $\mathcal{K}$ were empty. By compactness, a finite subsystem of its defining conditions is inconsistent. Let $E \subset S$ contain every coordinate appearing in this subsystem, together with 0 and 1. Adjoin $\lambda(0) = 0$, $\lambda(1) = 1$, and write each coordinate bound homogeneously as

$$\lambda(s) \geq 0, \quad R(s)\lambda(1) - \lambda(s) \geq 0 \quad (s \in E).$$

The resulting finite system remains inconsistent.

To separate the equalities from the inequalities, let e_s , $s \in E$, be the coordinate vectors of $\mathbb{R}^E$. Quotient by the span of the selected additive relation vectors and e_0 , and write u for the image of e_1 . Let C be the cone generated by the images of all inequality vectors. Before taking the quotient, these vectors have the forms

$$e_b - e_a \ (b \geq a), \quad e_{T_d s} - k e_s, \quad e_s, \quad R(s)e_1 - e_s.$$

The finitely generated cone C is closed. If $-u \notin C$, separation gives a linear functional L nonnegative on C with $L(u) > 0$. Dividing by $L(u)$ would give a normalized solution of the finite subsystem, including its bounds. This contradicts inconsistency, so $-u \in C$.

All vectors and relations have integer coefficients. Membership $-u \in C$ is a feasible rational linear system, so there are rational nonnegative cone coefficients and rational coefficients for the relation vectors. Clearing denominators yields a positive integer m and an identity with nonnegative integer coefficients on the inequality vectors.

Map this identity to the additive Grothendieck group of S by sending e_s to $[s]$. This group adjoins formal additive inverses to the commutative monoid $(S, +)$, so all additive relation vectors map to zero. The bound vectors map to the order differences $[y] - [0]$ and $[R(s)] - [s]$. Direct sums combine these and the other order differences into $[y] - [x]$ with $y \geq x$, while distributivity combines the terms from (A.1) into $[T_d s] - k[s]$ for one $s \in S$. We obtain

$$-m[1] = [y] - [x] + [T_d s] - k[s], \quad y \geq x.$$

Equality in this group means equality after adding a common element of the monoid. Thus there is $c \in S$ with

$$m + y + T_d s + c = x + k s + c.$$

Put $D = y + c$. Using $y \geq x$ gives the catalytic comparison

$$D + ks \geq D + m + T_d s. \quad (\text{A.2})$$

Here $s \neq 0$: otherwise $D \geq D + m$, contradicting the additivity and monotonicity of any flattening rank.

Reusing the catalyst

Iterating (A.2) reuses the same tensor D , so its additive cost stays fixed as the number of copies grows. For every integer $n \geq 1$,

$$D + nks \geq D + nm + nT_d s \geq nT_d s.$$

Since $s \geq 1$, we have $R(D)s \geq R(D) \geq D$. With $K = nk + R(D)$, it follows that

$$Ks \geq nT_d s, \quad K^j s \geq n^j T_d^j s \quad (j \geq 1). \quad (\text{A.3})$$

For the induction, multiply the j th comparison by K , then use $Ks \geq nT_d s$:

$$K^{j+1} s \geq n^j T_d^j (Ks) \geq n^{j+1} T_d^{j+1} s.$$

This keeps a single factor s and requires no cancellation. Also $K > 0$, since the first right-hand side in (A.3) is nonzero.

To bound K using ν , we convert the n^j scalar copies in (A.3) into a second matrix-multiplication factor. Fix n and $\delta > 0$. By the definition of ν , choose a fixed integer $u \geq 2$ with $R(T_u) \leq u^{\nu+\delta}$, and set

$$\ell_j = \left\lfloor \frac{j \log_u n}{\nu + \delta} \right\rfloor, \quad g_j = u^{\ell_j}.$$

Submultiplicativity gives

$$R(T_{g_j}) \leq R(T_u)^{\ell_j} \leq u^{(\nu+\delta)\ell_j} \leq n^j.$$

Thus the diagonal tensor n^j restricts to T_{g_j} . Multiplying this restriction by $T_d s$ and using $s \geq 1$, we obtain

$$K^j s \geq n^j T_d s \geq T_{d^j g_j} s \geq T_{d^j g_j}.$$

A direct sum of K^j copies of s has rank at most $K^j R(s)$. Therefore

$$R(s) K^j \geq R(T_{d^j g_j}) \geq (d^j g_j)^\nu.$$

First let $j \rightarrow \infty$, keeping n, δ, u fixed. Since $g_j^{1/j} \rightarrow n^{1/(\nu+\delta)}$, taking j th roots gives

$$nk + R(D) \geq d^\nu n^{\nu/(\nu+\delta)}.$$

The choice of u has disappeared from the inequality, so letting $\delta \downarrow 0$ at fixed n gives $nk + R(D) \geq nd^\nu$. Dividing by n and then letting $n \rightarrow \infty$ gives $k \geq d^\nu$, since the catalyst D is fixed. This contradiction proves that $\mathcal{K}$ is nonempty.

To apply Lemma A.1, it remains to check that each P_z preserves $\mathcal{K}$. For any nonzero z , the map preserves normalization, additivity, and monotonicity. It also preserves (A.1), because

$$(P_z \lambda)(T_d s) = \frac{\lambda(T_d z s)}{\lambda(z)} \geq k \frac{\lambda(z s)}{\lambda(z)} = k(P_z \lambda)(s).$$

Its coordinate bounds follow from $zx \leq R(x)z$. Lemma A.1 supplies a multiplicative state in $\mathcal{K}$, hence a character. Taking $s = 1$ in (A.1) gives $\lambda(T_d) \geq k$. $\square$

References

- [1] J. Alman, R. Duan, V. Vassilevska Williams, Y. Xu, Z. Xu, and R. Zhou, More asymmetry yields faster matrix multiplication, in *Proc. 2025 Annual ACM-SIAM Symposium on Discrete Algorithms*, SIAM, 2025, 2005–2039. doi:10.1137/1.9781611978322.63.
- [2] J. Alman and V. Vassilevska Williams, A refined laser method and faster matrix multiplication, *TheoretCS* **3** (2024), Article 21, 1–32. doi:10.46298/theoretics.24.21. Conference version in SODA 2021, 522–539.
- [3] D. Bini, Relations between exact and approximate bilinear algorithms. Applications, *Calcolo* **17** (1980), 87–97. doi:10.1007/BF02575865.
- [4] M. Christandl, P. Vrana, and J. Zuiddam, Universal points in the asymptotic spectrum of tensors, *J. Amer. Math. Soc.* **36** (2023), 31–79. doi:10.1090/jams/996.
- [5] D. Coppersmith and S. Winograd, Matrix multiplication via arithmetic progressions, *J. Symbolic Comput.* **9** (1990), 251–280. doi:10.1016/S0747-7171(08)80013-2.
- [6] A. M. Davie and A. J. Stothers, Improved bound for complexity of matrix multiplication, *Proc. Roy. Soc. Edinburgh Sect. A* **143** (2013), 351–369. doi:10.1017/S0308210511001648.
- [7] R. Duan, H. Wu, and R. Zhou, Faster matrix multiplication via asymmetric hashing, in *Proc. 64th IEEE Annual Symposium on Foundations of Computer Science*, IEEE, 2023, 2129–2138. doi:10.1109/FOCS57990.2023.00130.
- [8] E. Dupont, M. Eisenberger, B. Kozlovskii, A. Mehrabian, F. J. R. Ruiz, A. See, R. Zhou, J. Alman, V. Vassilevska Williams, and M. Balog, Improving the matrix multiplication exponent with modern optimization and AlphaEvolve, preprint, 2026. arXiv:2608.16884.
- [9] E. Kowalski, *An introduction to the representation theory of groups*, Graduate Studies in Mathematics **155**, American Mathematical Society, 2014. doi:10.1090/gsm/155. Author version: *Representation theory*, February 17, 2017, Section 2.6.
- [10] F. Le Gall, Powers of tensors and fast matrix multiplication, in *Proc. 39th International Symposium on Symbolic and Algebraic Computation*, ACM, 2014, 296–303. Full version: arXiv:1401.7714v1.
- [11] A. Schönhage, Partial and total matrix multiplication, *SIAM J. Comput.* **10** (1981), 434–455. doi:10.1137/0210032.
- [12] A. J. Stothers, *On the complexity of matrix multiplication*, Ph.D. thesis, University of Edinburgh, 2010. hdl:1842/4734.
- [13] V. Strassen, Gaussian elimination is not optimal, *Numer. Math.* **13** (1969), 354–356. doi:10.1007/BF02165411.
- [14] V. Strassen, The asymptotic spectrum of tensors and the exponent of matrix multiplication, in *Proc. 27th Annual Symposium on Foundations of Computer Science*, IEEE, 1986, 49–54. doi:10.1109/SFCS.1986.52.
- [15] V. Strassen, Relative bilinear complexity and matrix multiplication, *J. Reine Angew. Math.* **375/376** (1987), 406–443. doi:10.1515/crll.1987.375-376.406.

- [16] V. Strassen, The asymptotic spectrum of tensors, *J. Reine Angew. Math.* **384** (1988), 102–152. [doi:10.1515/crll.1988.384.102](https://doi.org/10.1515/crll.1988.384.102).
- [17] A. Tychonoff, Ein Fixpunktsatz, *Math. Ann.* **111** (1935), 767–776. [doi:10.1007/BF01472256](https://doi.org/10.1007/BF01472256).
- [18] V. Vassilevska Williams, Multiplying matrices faster than Coppersmith–Winograd, in *Proc. 44th Annual ACM Symposium on Theory of Computing*, ACM, 2012, 887–898. [doi:10.1145/2213977.2214056](https://doi.org/10.1145/2213977.2214056). Expanded corrected author version: *Multiplying matrices in $O(n^{2.373})$ time*, July 1, 2014.
- [19] V. Vassilevska Williams, Y. Xu, Z. Xu, and R. Zhou, New bounds for matrix multiplication: from alpha to omega, in *Proc. 2024 Annual ACM-SIAM Symposium on Discrete Algorithms*, SIAM, 2024, 3792–3835. [doi:10.1137/1.9781611977912.134](https://doi.org/10.1137/1.9781611977912.134).