Numerische Mathematik (2019) 143:905–922
https://doi.org/10.1007/s00211-019-01070-6

Numerische Mathematik

Grothendieck constant is norm of Strassen matrix multiplication tensor

Shmuel Friedland[1] · Lek-Heng Lim[2] · Jinjie Zhang[3]

Received: 3 January 2019 / Revised: 10 May 2019 / Published online: 22 August 2019

Abstract

We show that two important quantities from two disparate areas of complexity theory—Strassen's exponent of matrix multiplication ω and Grothendieck's constant K_G —are different measures of size for the same underlying object: the matrix multiplication tensor, i.e., the 3-tensor or bilinear operator $\mu_{l,m,n} : \mathbb{F}^{l\times m} \times \mathbb{F}^{m\times n} \to \mathbb{F}^{l\times n}$, $(A, B) \mapsto AB$ defined by matrix-matrix product over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. It is well-known that Strassen's exponent of matrix multiplication is the greatest lower bound on (the log of) the *tensor rank* of $\mu_{l,m,n}$. We will show that Grothendieck's constant is the least upper bound on a *tensor norm* of $\mu_{l,m,n}$, taken over all $l, m, n \in \mathbb{N}$. Aside from relating the two celebrated quantities, this insight allows us to rewrite Grothendieck's inequality as a norm inequality

$$\|\mu_{l,m,n}\|_{1,2,\infty} = \max_{X,Y,M\neq 0} \frac{|\operatorname{tr}(XMY)|}{\|X\|_{1,2}\|Y\|_{2,\infty}\|M\|_{\infty,1}} \leqslant K_G.$$

We prove that Grothendieck's inequality is unique in the sense that if we generalize the $(1, 2, \infty)$-norm to arbitrary $p, q, r \in [1, \infty]$,

$$\|\mu_{l,m,n}\|_{p,q,r} = \max_{X,Y,M\neq 0} \frac{|\operatorname{tr}(XMY)|}{\|X\|_{p,q}\|Y\|_{q,r}\|M\|_{r,p}},$$

then $(p, q, r) = (1, 2, \infty)$ is, up to cyclic permutations, the only choice for which $\|\mu_{l,m,n}\|_{p,q,r}$ is uniformly bounded by a constant independent of l, m, n.

Mathematics Subject Classification 15A60 · 46B28 · 46B85 · 47A07 · 65Y20 · 68Q17 · 68Q25

The work in this article is supported by DARPA D15AP00109 and NSF IIS 1546413. LHL is supported by a DARPA Director's Fellowship and the Eckhardt Faculty Fund.

Extended author information available on the last page of the article

1 Introduction

Fifty years ago in this journal, Volker Strassen [46] announced an astounding result—the product of a pair of 2×2 matrices may be obtained with just seven multiplications:

$$\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix} = \begin{bmatrix} a_1 b_1 + a_2 b_2 & \beta + \gamma + (a_1 + a_2 - a_3 - a_4) b_4 \\ \alpha + \gamma + a_4(b_2 + b_3 - b_1 - b_4) & \alpha + \beta + \gamma \end{bmatrix},$$

where $\alpha = (a_3 - a_1)(b_3 - b_4)$, $\beta = (a_3 + a_4)(b_3 - b_1)$, $\gamma = a_1 b_1 + (a_3 + a_4 - a_1)(b_1 + b_4 - b_3)$. Applied recursively, this gives an algorithm for forming the product of a pair of $n \times n$ matrices with just $O(n^{\log_2 7})$ multiplications, as opposed to $O(n^3)$ using the usual formula for matrix-matrix product. In addition, Strassen also showed that: (i) the number of additions may be bounded by a constant times the number of multiplications; (ii) matrix inversion may be achieved with the same complexity as matrix multiplication. In short, if there is an algorithm that forms matrix product in $O(n^\omega)$ multiplication, then it yields an $O(n^\omega)$ algorithm that would solve n linear equations in n unknowns, which is by far the most ubiquitous problem in all of scientific and engineering computing. The smallest possible ω became known as the *exponent of matrix multiplication*.

Strassen's pioneering work on of complexity of matrix inversion, or equivalently, matrix multiplication, [46–49] captured the interests of numerical analysts and theoretical computer scientists alike. It launched a new field: A systematic study of this and other related problems has blossomed into what is now often called algebraic computational complexity [10]. The value of ω was gradually lowered over the years. Some milestones include the Coppersmith–Winograd [12] bound $O(n^{2.375477})$ that resisted progress for more than two decades until Vassilevska-Williams's improvement [51] to $O(n^{2.3728642})$; the current record, due to Le Gall [36], is $O(n^{2.3728639})$. Strassen showed [47] that the best possible ω is in fact given by

$$\omega = \inf_{n \in \mathbb{N}} \log_n \bigl(\operatorname{rank}(\mu_{n,n,n})\bigr),$$

where $\mu_{n,n,n}$ is the *Strassen matrix multiplication tensor*—the 3-tensor in $(\mathbb{F}^{n\times n})^* \otimes (\mathbb{F}^{n\times n})^* \otimes \mathbb{F}^{n\times n}$ associated with matrix-matrix product, i.e., the bilinear operator

$$\mathbb{F}^{n\times n} \times \mathbb{F}^{n\times n} \to \mathbb{F}^{n\times n}, \quad (A, B) \mapsto AB,$$

where $\mathbb{F}$ is any field but we will assume $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ throughout this article.

Those unfamiliar with multilinear algebra [35] may regard the 3-tensor $\mu_{n,n,n}$ and the bilinear operator as the same object. If we choose a basis on $\mathbb{F}^{n\times n}$ (or three different bases, one on each copy of $\mathbb{F}^{n\times n}$), then $\mu_{n,n,n}$ may be represented as a 3-dimensional hypermatrix in $\mathbb{F}^{n^2\times n^2\times n^2}$. Over any $\mathbb{F}$-vector spaces $\mathbb{U}$, $\mathbb{V}$, $\mathbb{W}$, one may define *tensor rank* [26] for 3-tensors $\tau \in \mathbb{U}\otimes\mathbb{V}\otimes\mathbb{W}$ by

$$\operatorname{rank}(\tau) = \min\left\{r : \tau = \sum\nolimits_{i=1}^{r} u_i \otimes v_i \otimes w_i\right\}.$$

In fact, Strassen showed that the tensor rank of a 3-tensor $\mu_\beta \in \mathbb{U}^*\otimes\mathbb{V}^*\otimes\mathbb{W}$ associated with a bilinear operator $\beta : \mathbb{U}\times\mathbb{V}\to\mathbb{W}$ gives the least number of multiplications required to compute β. The value of ω is in general dependent on the choice of $\mathbb{F}$, as tensor rank is well-known to be field dependent [37]; we will write $\omega^{\mathbb{F}}$ when we need to emphasize this.

A world apart from Strassen's work on fast matrix multiplication/inversion is the *Grothendieck inequality*, originally established to relate fundamental norms on tensor products of Banach spaces [21]. The *Grothendieck constant* K_G is the smallest constant such that for every $l, m, n \in \mathbb{N}$ and every matrix $M = (M_{ij}) \in \mathbb{F}^{m\times n}$,

$$\max_{\|x_i\|=\|y_j\|=1}\left|\sum_{i=1}^{m}\sum_{j=1}^{n} M_{ij}\langle x_i, y_j\rangle\right| \leqslant K_G \max_{|\varepsilon_i|=|\delta_j|=1}\left|\sum_{i=1}^{m}\sum_{j=1}^{n} M_{ij}\varepsilon_i\delta_j\right| \tag{1}$$

where the maximum on the left is taken over all $x_i, y_j \in \mathbb{F}^l$ of unit 2-norm, and the maximum on the right is taken over all $\varepsilon_i, \delta_j \in \mathbb{F}$ of unit absolute value (so over $\mathbb{R}$, $\varepsilon_i = \pm 1$ and $\delta_j = \pm 1$; over $\mathbb{C}$, $\varepsilon_i = e^{i\theta_i}$ and $\delta_j = e^{i\phi_j}$). The value on the left side of (1) is the same for all $l \geqslant m+n$ and as such some authors restrict themselves to $l = m+n$.

The existence of such a constant independent of l, m and n was discovered by Alexandre Grothendieck in 1953. Alternative proofs via factorization of linear operators, geometry of Banach spaces, absolutely p-summing operators, etc, may be found in [28,38,40,41] and references therein. In particular, the formulation in (1) was due to Lindenstrauss and Pełczyński [38]. Like ω, K_G is also dependent on the choice of $\mathbb{F}$ (we proffer an explanation in Sect. 1.2); we will write $K_G^{\mathbb{F}}$ when we need to emphasize this. Henceforth we will assume that $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

The inequality (1) has found applications in numerous areas, including Banach space theory, C^* algebra, harmonic analysis, operator theory, quantum mechanics, and most recently, computer science. In theoretical computer science, Grothendieck's inequality has notably appeared in studies of unique games conjecture [29–31,42,43] and SDP relaxations of NP-hard combinatorial problems [2–5,11]. In quantum information theory, Grothendieck's inequality arises unexpectedly in Bell inequalities [17,24,50] and in XOR games [7–9], among several other areas; Grothendieck constants of specific orders, e.g., $K_G^{\mathbb{C}}(3)$ and $K_G^{\mathbb{C}}(4)$, also play important roles in quantum information theory [1,15,25]. The inequality has even been applied to some rather surprising areas, e.g., to communication complexity [39,44,45] and to privacy-preserving data analysis [16].

Although the Grothendieck constant appears in numerous mathematical statements and has many equivalent interpretations in physics and computer science, its exact value remains unknown and estimating increasingly sharper bounds for $K_G^{\mathbb{F}}$ has been a major undertaking. The current best known bounds are $K_G^{\mathbb{R}} \in [1.676, 1.782]$, established in [13] (lower) and [33] (upper); and $K_G^{\mathbb{C}} \in (1.338, 1.404]$, established in [14] (lower) and [22] (upper). A major recent breakthrough [6] established that Krivine's upper bound $\pi/\big(2\log(1+\sqrt{2})\big) \approx 1.782$ for $K_G^{\mathbb{R}}$ is not sharp. There have also been efforts in approximating Grothendieck's constants of specific orders, e.g., see [15,25] for recent results on $K_G^{\mathbb{C}}(3)$ and $K_G^{\mathbb{C}}(4)$, defined below in (5).

1.1 Strassen's exponent and Grothendieck's constant

What exactly is Strassen's exponent ω? The discussion in the previous page shows that it is the greatest lower bound for the (log of the) tensor rank of the Strassen matrix multiplication tensor:

$$\log_n\big(\operatorname{rank}(\mu_{n,n,n})\big) \geqslant \omega \qquad \text{for all } n \in \mathbb{N}. \tag{2}$$

What exactly is Grothendieck's constant K_G? We will show that it is the least upper bound for the tensor $(1, 2, \infty)$-norm (see (6) below) of the Strassen matrix multiplication tensor:

$$\|\mu_{l,m,n}\|_{1,2,\infty} \leqslant K_G \qquad \text{for all } l, m, n \in \mathbb{N}. \tag{3}$$

If we desire a greater parallel to (2), we may drop l and m in (3)—there is no loss of generality in assuming that $l = 2n$ and $m = n$, i.e., K_G is also the least upper bound so that

$$\|\mu_{2n,n,n}\|_{1,2,\infty} \leqslant K_G \qquad \text{for all } n \in \mathbb{N}. \tag{4}$$

In other words, ω and K_G are just different measures of size: rank and norm respectively, for the same underlying object, Strassen's matrix multiplication tensor $\mu_{l,m,n}$. Note that both ω and K_G are universal constants—uniform bounds independent of the dimension n in (2) and (4).

In addition, the *Grothendieck constant of order* $l \in \mathbb{N}$, a popular notion in quantum information theory (e.g., [1,15,25]), is given by a simple variation, namely, the least upper bound $K_G(l)$ in

$$\|\mu_{l,m,n}\|_{1,2,\infty} \leqslant K_G(l) \qquad \text{for all } m, n \in \mathbb{N}. \tag{5}$$

We will define the $(1, 2, \infty)$-norm for an arbitrary 3-tensor formally in Sect. 4 but at this point it suffices to know its value for $\mu_{l,m,n}$, namely,

$$\|\mu_{l,m,n}\|_{1,2,\infty} = \max_{X,Y,M\neq 0} \frac{|\operatorname{tr}(XMY)|}{\|X\|_{1,2}\|Y\|_{2,\infty}\|M\|_{\infty,1}} \tag{6}$$

where $X \in \mathbb{F}^{l\times m}$, $M \in \mathbb{F}^{m\times n}$, $Y \in \mathbb{F}^{n\times l}$, and $\|M\|_{p,q} := \max_{x\neq 0}\|Mx\|_q/\|x\|_p$ denotes the matrix (p, q)-norm.

1.2 Grothendieck's inequality is unique

The inequality (3) is in fact just Grothendieck's inequality, as we will see later in Sect. 3. The characterizations of ω and K_G in (2) and (3) hold over both $\mathbb{R}$ and $\mathbb{C}$ although their values are field dependent. Incidentally the fact that Grothendieck's constant is essentially a tensor norm immediately explains why it is field dependent—because, like tensor rank, tensor norms are also field dependent [19].

An advantage of the formulation in (3) is that we obtain a natural family of (p, q, r)-norms on $\mu_{l,m,n}$ given by

$$\|\mu_{l,m,n}\|_{p,q,r} := \max_{X,Y,M\neq 0} \frac{|\operatorname{tr}(XMY)|}{\|X\|_{p,q}\|Y\|_{q,r}\|M\|_{r,p}}$$

for any triple $1 \leqslant p, q, r \leqslant \infty$. This family of norms will serve as a platform for us to better comprehend Grothendieck's constant and Grothendieck's inequality. The study of the general (p, q, r)-case shows why the $(1, 2, \infty)$-case is extraordinary. We will deduce a generalization of Grothendieck's inequality and show that the case $(p, q, r) = (1, 2, \infty)$, i.e., Grothendieck's inequality, is the only one up to trivial cyclic permutations[1] where there is a universal upper bound, i.e., Grothendieck's constant, that holds for all $l, m, n \in \mathbb{N}$.

Theorem 1 (Grothendieck–Hölder inequality) *Let $1 \leqslant p, q, r \leqslant \infty$ and $l, m, n \in \mathbb{N}$. Then*

$$\frac{1}{l^{|1/q-1/2|}\cdot m^{|1/p-1/2|}\cdot n^{|1/r-1/2|}} \leqslant \|\mu_{l,m,n}\|_{p,q,r} \leqslant K_G^{\mathbb{F}}\cdot l^{|1/q-1/2|}\cdot m^{1-1/p}\cdot n^{1/r}.$$

When $p = 1$, $q = 2$, and $r = \infty$, the upper bound gives Grothendieck's inequality (1).

Theorem 2 (Uniqueness of Grothendieck's inequality) *Let $1 \leqslant p, q, r \leqslant \infty$ and $l, m, n \in \mathbb{N}$. Then $\|\mu_{l,m,n}\|_{p,q,r}$ is uniformly bounded for all $l, m, n \in \mathbb{N}$ if and only if*

$$(p, q, r) \in \{(1, 2, \infty),\ (\infty, 1, 2),\ (2, \infty, 1)\}.$$

Theorem 1 follows from Theorems 4 and 5. Theorem 2 is just Theorem 6.

2 Strassen matrix multiplication tensor

An important observation, straightforward for those acquainted with tensors [34,35,37] but perhaps less so for those accustomed to viewing them as "multiway arrays," is that the bilinear operator

$$\beta \in \mathbb{F}^{l\times m} \times \mathbb{F}^{m\times n} \to \mathbb{F}^{l\times n}, \quad (X, Y) \mapsto XY, \tag{7}$$

[1] This is unavoidable as (p, q, r)-norms of $\mu_{l,m,n}$ are invariant under cyclic permutations of p, q, r. See Lemma 1(i).

and the trilinear functional

$$\tau : \mathbb{F}^{l\times m} \times \mathbb{F}^{m\times n} \times \mathbb{F}^{n\times l} \to \mathbb{F}, \quad (X, Y, Z) \mapsto \operatorname{tr}(XYZ), \tag{8}$$

are given by[2] the same 3-tensor in

$$(\mathbb{F}^{l\times m})^* \otimes (\mathbb{F}^{m\times n})^* \otimes \mathbb{F}^{l\times n} \cong (\mathbb{F}^{l\times m})^* \otimes (\mathbb{F}^{m\times n})^* \otimes (\mathbb{F}^{n\times l})^*.$$

In other words, as 3-tensors, there is no difference between the product of two matrices and the trace of product of three matrices.

To see this, let $E_{ij} \in \mathbb{F}^{m\times n}$ denote the matrix with 1 in its (i, j)th entry and zeros everywhere else, so that $\{E_{ij} : i = 1, \dots, m;\ \ j = 1, \dots, n\}$ is the standard basis for $\mathbb{F}^{m\times n}$. Its dual basis for the dual space of linear functionals

$$(\mathbb{F}^{m\times n})^* := \{\varphi : \mathbb{F}^{m\times n} \to \mathbb{F} : \varphi(\alpha X + \beta Y) = \alpha\varphi(X) + \beta\varphi(Y)\}$$

is then given by $\{\varepsilon_{ij} : i = 1, \dots, m;\ \ j = 1, \dots, n\}$ where $\varepsilon_{ij} : \mathbb{F}^{m\times n} \to \mathbb{F}, X \mapsto x_{ij}$, is the linear functional that takes an $m \times n$ matrix to its (i, j)th entry. Now choose the standard inner product on $\mathbb{F}^{m\times n}$, i.e., $\langle X, Y\rangle = \operatorname{tr}(X^{\mathsf{T}}Y)$. Then $\varepsilon_{ij}(X) = \langle E_{ij}, X\rangle$ for all $X \in \mathbb{F}^{m\times n}$, which allows us to identify $(\mathbb{F}^{m\times n})^*$ with $\mathbb{F}^{n\times m}$ and linear functional $\varepsilon_{ij} \in (\mathbb{F}^{m\times n})^*$ with the matrix $E_{ji} \in \mathbb{F}^{n\times m}$.

It remains to observe that the usual formula for matrix-matrix product gives

$$\begin{aligned}
\beta(X, Y) &= \sum\nolimits_{i,k=1}^{l,n} \left(\sum_{j=1}^{m} x_{ij} y_{jk}\right) E_{ik} \\
&= \sum_{i,k=1}^{l,n} \left(\sum\nolimits_{j=1}^{m} \varepsilon_{ij}(X)\varepsilon_{jk}(Y)\right) E_{ik} \\
&= \sum\nolimits_{i,k=1}^{l,n} \left(\sum\nolimits_{j=1}^{m} (\varepsilon_{ij} \otimes \varepsilon_{jk})(X, Y)\right) E_{ik} \\
&= \left(\sum\nolimits_{i,j,k=1}^{l,m,n} \varepsilon_{ij} \otimes \varepsilon_{jk} \otimes E_{ik}\right)(X, Y)
\end{aligned}$$

for all $X \in \mathbb{F}^{l\times m}$, $Y \in \mathbb{F}^{m\times n}$, and therefore

$$\beta = \sum\nolimits_{i,j,k=1}^{l,m,n} \varepsilon_{ij} \otimes \varepsilon_{jk} \otimes E_{ik} \in (\mathbb{F}^{l\times m})^* \otimes (\mathbb{F}^{m\times n})^* \otimes \mathbb{F}^{l\times n}. \tag{9}$$

[2] To be more precise, by the universal property of tensor products [35, Chapter XVI, §1], β induces a linear map $\beta_* : \mathbb{F}^{l\times m} \otimes \mathbb{F}^{m\times n} \to \mathbb{F}^{l\times n}$ and τ induces a linear map $\tau_* : \mathbb{F}^{l\times m} \otimes \mathbb{F}^{m\times n} \otimes \mathbb{F}^{n\times l} \to \mathbb{F}$, i.e., $\beta_* \in (\mathbb{F}^{l\times m})^* \otimes (\mathbb{F}^{m\times n})^* \otimes \mathbb{F}^{l\times n}$ and $\tau_* \in (\mathbb{F}^{l\times m})^* \otimes (\mathbb{F}^{m\times n})^* \otimes (\mathbb{F}^{n\times l})^*$. We identify β, τ with the linear maps β_*, τ_* they induce.

A similar simple calculation,

$$\begin{aligned}\tau(X, Y, Z) &= \sum_{i,j,k=1}^{l,m,n} x_{ij} y_{jk} z_{ki}\\ &= \sum_{i,j,k=1}^{l,m,n} \varepsilon_{ij}(X)\varepsilon_{jk}(Y)\varepsilon_{ki}(Z)\\ &= \sum_{i,j,k=1}^{l,m,n} (\varepsilon_{ij} \otimes \varepsilon_{jk} \otimes \varepsilon_{ki})(X, Y, Z)\\ &= \Big(\sum_{i,j,k=1}^{l,m,n} \varepsilon_{ij} \otimes \varepsilon_{jk} \otimes \varepsilon_{ki}\Big)(X, Y, Z)\end{aligned}$$

for all $X \in \mathbb{F}^{l\times m}$, $Y \in \mathbb{F}^{m\times n}$, $Z \in \mathbb{F}^{n\times l}$, and therefore

$$\tau = \sum_{i,j,k=1}^{l,m,n} \varepsilon_{ij} \otimes \varepsilon_{jk} \otimes \varepsilon_{ki} \in (\mathbb{F}^{l\times m})^* \otimes (\mathbb{F}^{m\times n})^* \otimes (\mathbb{F}^{n\times l})^*. \tag{10}$$

By our identification, $(\mathbb{F}^{m\times n})^* = \mathbb{F}^{n\times m}$ and $\varepsilon_{ki} = E_{ik}$. So we see from (9) and (10) that indeed $\beta = \tau$ as 3-tensors. We denote this tensor by $\mu_{l,m,n}$. This has been variously called the *Strassen matrix multiplication tensor* or the *structure tensor for matrix-matrix product* [10,34,37,52].

3 Grothendieck's constant and Strassen's tensor

Let l, m, n be positive integers and let $M = (M_{ij}) \in \mathbb{F}^{m\times n}$. Let $x_1, \dots, x_m$ and $y_1, \dots, y_n \in \mathbb{F}^l$ be vectors of unit 2-norm. We will regard $x_1, \dots, x_m$ as columns of a matrix $X \in \mathbb{F}^{l\times m}$ and $y_1^{\mathsf{T}}, \dots, y_n^{\mathsf{T}}$ as rows of a matrix $Y \in \mathbb{F}^{n\times l}$.

Recall that for any $p \geqslant 1$ with Hölder conjugate p^*, i.e., $1/p + 1/p^* = 1$, we have

$$\begin{aligned}\|X\|_{1,p} &:= \max_{z\neq 0} \frac{\|Xz\|_p}{\|z\|_1} = \max_{i=1,\dots,m} \|x_i\|_p,\\ \|Y\|_{p,\infty} &:= \max_{z\neq 0} \frac{\|Yz\|_\infty}{\|z\|_p} = \max_{i=1,\dots,n} \|y_i\|_{p^*},\end{aligned} \tag{11}$$

and

$$\begin{aligned}\|M\|_{\infty,1} &:= \max_{z\neq 0} \frac{\|Mz\|_1}{\|z\|_\infty} = \max_{|\delta_j|=1} \sum_{i=1}^{m} \Big|\sum_{j=1}^{n} M_{ij}\delta_j\Big|\\ &= \max_{|\varepsilon_i|=1,\ |\delta_j|=1} \Big|\sum_{i=1}^{m} \sum_{j=1}^{n} M_{ij}\varepsilon_i\delta_j\Big|,\end{aligned}$$

which may be further simplified for $\mathbb{F} = \mathbb{R}$ as

$$\|M\|_{\infty,1} = \max_{\varepsilon_i=\pm 1,\ \delta_j=\pm 1} \Big|\sum_{i=1}^{m} \sum_{j=1}^{n} M_{ij}\varepsilon_i\delta_j\Big| = \max_{\varepsilon,\delta\in\{\pm 1\}^n} |\varepsilon^{\mathsf{T}} M\delta|. \tag{12}$$

We refer the reader to [19] for a proof that

$$\|\tau\|_{1,2,\infty} := \max_{X,Y,M\neq 0} \frac{|\tau(X,M,Y)|}{\|X\|_{1,2}\|Y\|_{2,\infty}\|M\|_{\infty,1}} \tag{13}$$

defines a norm for any tensor $\tau \in (\mathbb{F}^{l\times m})^* \otimes (\mathbb{F}^{m\times n})^* \otimes (\mathbb{F}^{n\times l})^*$, regarded as a trilinear functional.

Theorem 3 *Grothendieck's inequality (1) may be stated as*

$$\max_{X,Y,M\neq 0} \frac{|\mathrm{tr}(XMY)|}{\|X\|_{1,2}\|Y\|_{2,\infty}\|M\|_{\infty,1}} \leqslant K_G^{\mathbb{F}}, \tag{14}$$

or more succinctly as

$$\|\mu_{l,m,n}\|_{1,2,\infty} \leqslant K_G^{\mathbb{F}}, \tag{15}$$

where $\mu_{l,m,n}$ is the Strassen matrix multiplication tensor for the product of $l \times m$ and $m \times n$ matrices, i.e.,

$$\mu_{l,m,n} = \sum_{i=1}^{l}\sum_{j=1}^{m}\sum_{k=1}^{n} \varepsilon_{ij} \otimes \varepsilon_{jk} \otimes \varepsilon_{ki} \in (\mathbb{F}^{l\times m})^* \otimes (\mathbb{F}^{m\times n})^* \otimes (\mathbb{F}^{n\times l})^*. \tag{16}$$

Proof We first make the observation that the maximum over $\|x_i\|_2 = 1$, $\|y_j\|_2 = 1$ in (1) is equivalent to a maximum over $\|x_i\|_2 \leqslant 1$, $\|y_j\|_2 \leqslant 1$ because the latter maximum is attained at the boundary [20, Lemma 2.2]. Also, by (11) with $p = 2$, $\|X\|_{1,2} = \|Y\|_{2,\infty} = 1$ is equivalent to $\|x_i\|_2 \leqslant 1$, $\|y_j\|_2 \leqslant 1$ for all $i = 1, \dots, m$, $j = 1, \dots, n$. Since

$$\sum\nolimits_{i=1}^{m} \sum\nolimits_{j=1}^{n} M_{ij}\langle x_i, y_j\rangle = \begin{cases} \mathrm{tr}(XMY) & \text{if } \mathbb{F} = \mathbb{R}, \\ \mathrm{tr}(XM\overline{Y}) & \text{if } \mathbb{F} = \mathbb{C}, \end{cases}$$

we see that Grothendieck's inequality (1) may be stated as

$$\max_{X,Y,M\neq 0} \frac{|\mathrm{tr}(XMY)|}{\|X\|_{1,2}\|Y\|_{2,\infty}\|M\|_{\infty,1}} \leqslant K_G^{\mathbb{F}},$$

when $\mathbb{F} = \mathbb{R}$ and as

$$\max_{X,Y,M\neq 0} \frac{|\mathrm{tr}(XM\overline{Y})|}{\|X\|_{1,2}\|Y\|_{2,\infty}\|M\|_{\infty,1}} \leqslant K_G^{\mathbb{F}},$$

when $\mathbb{F} = \mathbb{C}$. However, as matrix (p, q)-norms are invariant under complex conjugation,

$$\max_{X,Y,M\neq 0} \frac{|\operatorname{tr}(XM\overline{Y})|}{\|X\|_{1,2}\|Y\|_{2,\infty}\|M\|_{\infty,1}} = \max_{X,\overline{Y},M\neq 0} \frac{|\operatorname{tr}(XM\overline{Y})|}{\|X\|_{1,2}\|\overline{Y}\|_{2,\infty}\|M\|_{\infty,1}}$$
$$= \max_{X,Y,M\neq 0} \frac{|\operatorname{tr}(XMY)|}{\|X\|_{1,2}\|Y\|_{2,\infty}\|M\|_{\infty,1}}.$$

Hence (14) in fact gives Grothendieck's inequality for both $\mathbb{F} = \mathbb{R}$ and $\mathbb{C}$. By our discussion in Sect. 2 and our norm in (13), (14) is just (15). □

This allows us to define Grothendieck's constant in terms of tensor norms: For $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$,

$$K_G^{\mathbb{F}} = \sup_{l,m,n\in\mathbb{N}} \|\mu_{l,m,n}\|_{1,2,\infty}. \tag{17}$$

Since $\|\mu_{l,m,n}\|_{1,2,\infty} = \|\mu_{m+n,m,n}\|_{1,2,\infty}$ for all $l \geqslant m+n$,

$$K_G^{\mathbb{F}} = \sup_{m,n\in\mathbb{N}} \|\mu_{m+n,m,n}\|_{1,2,\infty} = \sup_{n\in\mathbb{N}} \|\mu_{2n,n,n}\|_{1,2,\infty}.$$

In addition, the Grothendieck constant of order $l \in \mathbb{N}$ [1,15,25] may be defined as

$$K_G^{\mathbb{F}}(l) = \sup_{m,n\in\mathbb{N}} \|\mu_{l,m,n}\|_{1,2,\infty}.$$

4 Grothendieck–Hölder inequality

The norm in (13) admits a natural generalization to arbitrary $p, q, r \in [1, \infty]$ as

$$\|\tau\|_{p,q,r} := \max_{X,Y,M\neq 0} \frac{|\tau(X, M, Y)|}{\|X\|_{p,q}\|Y\|_{q,r}\|M\|_{r,p}}$$

defined for any $\tau \in (\mathbb{F}^{l\times m})^* \otimes (\mathbb{F}^{m\times n})^* \otimes (\mathbb{F}^{n\times l})^*$, regarded as a trilinear functional

$$\tau : \mathbb{F}^{l\times m} \times \mathbb{F}^{m\times n} \times \mathbb{F}^{n\times l} \to \mathbb{F}.$$

In this article, we will only be interested in $\tau = \mu_{l,m,n}$, the Strassen tensor. We first state some simple observations that will be useful later.

Lemma 1 *Let $p, q, r \in [1, \infty]$. Then the (p, q, r)-norm of $\mu_{l,m,n}$*

(i) *is invariant under cyclic permutation of p, q, r,*

$$\|\mu_{l,m,n}\|_{p,q,r} = \|\mu_{l,m,n}\|_{r,p,q} = \|\mu_{l,m,n}\|_{q,r,p};$$

(ii) *transforms under Hölder conjugation as*

$$\|\mu_{l,m,n}\|_{p,q,r} = \|\mu_{l,m,n}\|_{r^*,q^*,p^*}.$$

Recall that p^* *is the Hölder conjugate of* p, *i.e.,* $1/p + 1/p^* = 1$.

Proof Since the numerator $\operatorname{tr}(XMY) = \operatorname{tr}(MYX) = \operatorname{tr}(YXM)$ and the denominator is the product $\|X\|_{p,q}\|M\|_{r,p}\|Y\|_{q,r}$, cyclic permutations of $(p,q), (r,p), (q,r)$ leave the quotient and thus the maximum

$$\|\mu_{l,m,n}\|_{p,q,r} = \max_{X,Y,M\neq 0} \frac{|\operatorname{tr}(XMY)|}{\|X\|_{p,q}\|M\|_{r,p}\|Y\|_{q,r}}$$

invariant. Now just observe that the cyclic permutations

$$(p,q),\ (r,p),\ (q,r) \ \to\ (q,r),\ (p,q),\ (r,p) \ \to\ (r,p),\ (q,r),\ (p,q)$$

correspond to the following permutations

$$(p,q,r) \ \to\ (q,r,p) \ \to\ (r,p,q).$$

Let X^{H} denote the conjugate transpose of X. Since $\|X\|_{p,q} = \|X^{\mathsf{H}}\|_{q^*,p^*}$ and $|\operatorname{tr}(XMY)| = |\operatorname{tr}(Y^{\mathsf{H}}M^{\mathsf{H}}X^{\mathsf{H}})|$, we have

$$\frac{|\operatorname{tr}(XMY)|}{\|X\|_{p,q}\|Y\|_{q,r}\|M\|_{r,p}} = \frac{|\operatorname{tr}(Y^{\mathsf{H}}M^{\mathsf{H}}X^{\mathsf{H}})|}{\|Y^{\mathsf{H}}\|_{r^*,q^*}\|X^{\mathsf{H}}\|_{q^*,p^*}\|M^{\mathsf{H}}\|_{p^*,r^*}}.$$

Taking maximum over all nonzero X, Y, M yields the required equality. Note that the proof works over both $\mathbb{R}$ and $\mathbb{C}$. □

A straightforward application of Hölder's inequality yields an upper bound for $\|\mu_{l,m,n}\|_{p,q,r}$.

Theorem 4 *Let* $p,q,r \in [1,\infty]$ *and* $l,m,n \in \mathbb{N}$. *For any nonzero matrices* $X \in \mathbb{F}^{l\times m}$, $Y \in \mathbb{F}^{n\times l}$ *and* $M \in \mathbb{F}^{m\times n}$, *the following inequality is sharp:*

$$\frac{|\operatorname{tr}(XMY)|}{\|X\|_{p,q}\|Y\|_{q,r}\|M\|_{r,p}} \leqslant \frac{|\operatorname{tr}(XMY)|}{\|X\|_{1,2}\|Y\|_{2,\infty}\|M\|_{\infty,1}} \cdot l^{|1/q-1/2|} \cdot m^{1-1/p} \cdot n^{1/r}. \tag{18}$$

Furthermore, we have a generalization of Grothendieck's inequality:

$$\begin{aligned}\|\mu_{l,m,n}\|_{p,q,r} &= \max_{X,Y,M\neq 0} \frac{|\operatorname{tr}(XMY)|}{\|X\|_{p,q}\|Y\|_{q,r}\|M\|_{r,p}} \\ &\leqslant K_G^{\mathbb{F}} \cdot l^{|1/q-1/2|} \cdot m^{1-1/p} \cdot n^{1/r}.\end{aligned} \tag{19}$$

Proof First let $1 \leqslant q \leqslant 2$. Hölder's inequality gives $\|x\|_q \leqslant l^{1/q-1/2}\|x\|_2$; taken together with the fact that $\|x\|_p \leqslant \|x\|_q$ whenever $q \leqslant p$, we get

$$\|X\|_{1,2} \leqslant \|X\|_{1,q} \leqslant \|X\|_{p,q}, \qquad \|Y\|_{2,\infty} \leqslant \|Y\|_{2,r} \leqslant l^{1/q-1/2}\|Y\|_{q,r}. \tag{20}$$

The same argument also gives $\|M\|_{\infty,p} \leqslant \|M\|_{\infty,1} \leqslant m^{1-1/p}\|M\|_{\infty,p}$ for $1 \leqslant p \leqslant \infty$ and thus

$$\|M\|_{\infty,1} \leqslant m^{1-1/p}\|M\|_{\infty,p} \leqslant n^{1/r} \cdot m^{1-1/p}\|M\|_{r,p}. \tag{21}$$

The inequality (18) then follows from (20) and (21). To see that it is sharp, we use the following $m \times n$ rank-one matrices:

$$E_{m,n} := \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}, \quad C_{m,n} := \begin{bmatrix} 1 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 1 & 0 & \dots & 0 \end{bmatrix},$$
$$R_{m,n} := \begin{bmatrix} 1 & 1 & \dots & 1 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}, \quad J_{m,n} := \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \dots & 1 \end{bmatrix}.$$

It is easy to check that

$$\begin{aligned} &\|E_{l,m}\|_{p,q} = 1, && \|R_{n,l}\|_{q,r} = l^{1-1/q}, && \|J_{m,n}\|_{r,p} = m^{1/p} \cdot n^{1-1/r}, \\ &\|E_{l,m}\|_{1,2} = 1, && \|R_{n,l}\|_{2,\infty} = l^{1/2}, && \|J_{m,n}\|_{\infty,1} = mn. \end{aligned}$$

Since (18) becomes an equality when $X = E_{l,m}$, $Y = R_{n,l}$, and $M = J_{m,n}$, it is sharp for $1 \leqslant q \leqslant 2$.

Next let $2 < q \leqslant \infty$. Similarly, we have

$$l^{1/q-1/2}\|X\|_{1,2} \leqslant \|X\|_{1,q} \leqslant \|X\|_{p,q} \qquad \text{and} \qquad \|Y\|_{2,\infty} \leqslant \|Y\|_{2,r} \leqslant \|Y\|_{q,r},$$

which together with (21) give us (18). In this case the sharpness follows from selecting $X = C_{l,m}$, $Y = E_{n,l}$, $M = J_{m,n}$, and noting that

$$\|C_{l,m}\|_{p,q} = l^{1/q}, \quad \|C_{l,m}\|_{1,2} = l^{1/2}, \quad \|E_{n,l}\|_{q,r} = 1, \quad \|E_{n,l}\|_{2,\infty} = 1.$$

The inequality (19) follows from taking maximum over nonzero X, M, Y and supremum over l, m, n. When $(p, q, r) = (1, 2, \infty)$, it yields Grothendieck's inequality (14). □

The upper bound in (19) depends on l, m, n except when (p, q, r) is $(1, 2, \infty)$ or a cyclic permutation (by Lemma 1(i)). An immediate question is whether a uniform bound independent of l, m, n might perhaps also exist for some other values of (p, q, r), i.e.,

$$K_{p,q,r} := \sup_{l,m,n \in \mathbb{N}} \|\mu_{l,m,n}\|_{p,q,r} < \infty? \tag{22}$$

In Sect. 5, we will see that $K_{p,q,r} = \infty$ for all

$$(p, q, r) \notin \{(1, 2, \infty),\ (\infty, 1, 2),\ (2, \infty, 1)\}.$$

Nevertheless, we stress that while the $(1, 2, \infty)$-norm is unique up to cyclic permutation among the (p, q, r)-norms in (13) for uniformly bounding $\mu_{l,m,n}$ over all $l, m, n \in \mathbb{N}$, there are other tensor norms with this property. For example, we may consider the tensor *spectral norm* [19] of $\mu_{l,m,n}$,

$$\|\mu_{l,m,n}\|_\sigma := \max_{X,Y,M \neq 0} \frac{|\mathrm{tr}(XMY)|}{\|X\|_F \|Y\|_F \|M\|_F}$$

where the norm on X, Y, M is the matrix Frobenius (i.e., Hilbert–Schmidt) norm. In this case,

$$\|\mu_{l,m,n}\|_\sigma = 1, \quad \text{for all } l, m, n \in \mathbb{N}, \tag{23}$$

since, by Cauchy–Schwartz and the submultiplicativity of the Frobenius norm,

$$|\mathrm{tr}(XMY)| \leqslant \|X\|_F \|MY\|_F \leqslant \|M\|_F \|X\|_F \|Y\|_F,$$

and equality is attained by choosing M, X, Y with 1 in the $(1, 1)$th entry and 0 everywhere else.

We will use (23) to obtain lower bounds on $\|\mu_{l,m,n}\|_{p,q,r}$ below. The inequalities (19) and (24) will collectively be referred to as the *Grothendieck–Hölder inequality*.

Theorem 5 *Let $p, q, r \in [1, \infty]$ and $l, m, n \in \mathbb{N}$. Then*

$$\frac{1}{l^{|1/q-1/2|} \cdot m^{|1/p-1/2|} \cdot n^{|1/r-1/2|}} \leqslant \|\mu_{l,m,n}\|_{p,q,r}. \tag{24}$$

Proof For $n \in \mathbb{N}$ and $p, q \in [1, \infty]$, let

$$c_{p,q}(n) := n^{\max\{0, 1/p - 1/q\}}.$$

Then for any $M \in \mathbb{F}^{m \times n}$, the following sharp inequality holds [32, Theorem 4.3],

$$\|M\|_{p,q} \leqslant c_{q,2}(m) c_{2,p}(n) \|M\|_F.$$

It follows that

$$\|X\|_{p,q} \leqslant c_{q,2}(l) c_{2,p}(m) \|X\|_F, \qquad \|Y\|_{q,r} \leqslant c_{r,2}(n) c_{2,q}(l) \|Y\|_F,$$
$$\|M\|_{r,p} \leqslant c_{p,2}(m) c_{2,r}(n) \|M\|_F,$$

and for any tensor $\tau \in (\mathbb{F}^{l\times m})^* \otimes (\mathbb{F}^{m\times n})^* \otimes (\mathbb{F}^{n\times l})^*$, we have

$$\|\tau\|_\sigma \leqslant \|\tau\|_{p,q,r} \cdot l^{|1/q-1/2|} \cdot m^{|1/p-1/2|} \cdot n^{|1/r-1/2|}.$$

Plugging in $\tau = \mu_{l,m,n}$ and using (23), we obtain (24). □

A practical implication of (22) is that if

$$(p,q) \text{ and } (q,r) \in \{(1,1),\ (2,2),\ (\infty,\infty),\ (1,q),\ (q,\infty)\}, \tag{25}$$

then $\|X\|_{p,q}$ and $\|Y\|_{q,r}$ can be computed in polynomial time (to arbitrary precision) and

$$\max_{X,Y\neq 0} \frac{|\operatorname{tr}(XMY)|}{\|X\|_{p,q}\|Y\|_{q,r}} \leqslant K_{p,q,r}\|M\|_{r,p}$$

in principle gives a polynomial-time approximation of $\|M\|_{r,p}$, which is NP-hard [23] if (r,p) is not one of the special cases in (25). Unfortunately, we now know that as $K_{p,q,r} = \infty$ in all other cases, this only works when $(p,q,r) = (1,2,\infty), (\infty,1,2)$, or $(2,\infty,1)$, all three are equivalent to Grothendieck's inequality.

5 The Grothendieck inequality is unique

We show that $(p,q,r) = (1,2,\infty)$ is, up to a cyclic permutation, the only case for which (22) holds. We will first rule out a large number of cases with the following proposition.

Proposition 1 *Let $p,q,r \in [1,\infty]$. If there exists a finite constant $K_{p,q,r} > 0$ such that $\|\mu_{l,m,n}\|_{p,q,r} \leqslant K_{p,q,r}$ for all $l,m,n \in \mathbb{N}$, then*

$$\min(p,q,r) = 1 \qquad \text{and} \qquad \max(p,q,r) = \infty.$$

Proof Let $I_{m,n} \in \mathbb{F}^{m\times n}$ be the matrix obtained by appending zero rows or columns to the identity matrix[3] I_n or I_m,

$$I_{m,n} := \begin{cases} [I_n, 0_{m-n}]^{\mathsf{T}} & \text{if } m \geqslant n, \\ [I_m, 0_{n-m}] & \text{if } m < n. \end{cases}$$

Then its matrix (p,q)-norm is

$$\|I_{m,n}\|_{p,q} = \begin{cases} \min\{m,n\}^{1/q-1/p} & \text{if } p \geqslant q, \\ 1 & \text{if } p < q. \end{cases} \tag{26}$$

[3] Note that $I_{n,n} = I_n$. For consistency, we will always use the latter notation when it is a square matrix.

This follows from an easy calculation using Hölder inequality: For $m \geqslant n$,

$$\|I_{m,n}\|_{p,q} = \max_{z\neq 0} \frac{\|I_{m,n}z\|_q}{\|z\|_p} = \max_{z\neq 0} \frac{\|z\|_q}{\|z\|_p} = \begin{cases} n^{1/q-1/p} & \text{if } p \geqslant q, \\ 1 & \text{if } p < q, \end{cases}$$

and for $m < n$,

$$\begin{aligned} \|I_{m,n}\|_{p,q} &= \max_{z\neq 0} \frac{\|I_{m,n}z\|_q}{\|z\|_p} = \max_{z\neq 0} \frac{\|z_m\|_q}{\|z\|_p} = \max_{z_m\neq 0} \frac{\|z_m\|_q}{\|z_m\|_p} \\ &= \begin{cases} m^{1/q-1/p} & \text{if } p \geqslant q, \\ 1 & \text{if } p < q, \end{cases} \end{aligned}$$

where $z_m = [z_1, \ldots, z_m] \in \mathbb{F}^m$ is the vector comprising the first m entries of z.

Set $X = I_{l,m}$, $Y = I_{n,l}$, and $M = I_{m,n}$. Then $\operatorname{tr}(XMY) = \min\{l, m, n\}$, and by (26), we obtain

$$\begin{aligned} &\frac{|\operatorname{tr}(XMY)|}{\|X\|_{p,q}\|Y\|_{q,r}\|M\|_{r,p}} \\ &= \begin{cases} \min\{l,m,n\} \min\{m,n\}^{1/r-1/p} & \text{if } p \leqslant q \leqslant r, \\ \min\{l,m,n\} \min\{l,n\}^{1/q-1/r} \min\{m,n\}^{1/r-1/p} & \text{if } p \leqslant r \leqslant q, \\ \min\{l,m,n\} \min\{l,m\}^{1/p-1/q} \min\{m,n\}^{1/r-1/p} & \text{if } q \leqslant p \leqslant r, \\ \min\{l,m,n\} \min\{l,m\}^{1/p-1/q} & \text{if } q \leqslant r \leqslant p, \\ \min\{l,m,n\} \min\{l,n\}^{1/q-1/r} & \text{if } r \leqslant p \leqslant q, \\ \min\{l,m,n\} \min\{l,m\}^{1/p-1/q} \min\{l,n\}^{1/q-1/r} & \text{if } r \leqslant q \leqslant p. \end{cases} \end{aligned}$$

Suppose $l = 2n$, $m = n$ and $p \leqslant q \leqslant r$, then

$$\lim_{n\to\infty} \frac{|\operatorname{tr}(XMY)|}{\|X\|_{p,q}\|Y\|_{q,r}\|M\|_{r,p}} = \lim_{n\to\infty} n^{1/r-1/p+1} = \infty$$

unless $p = 1$ and $r = \infty$. Repeating the argument for all possible permutations of (p, q, r) and taking advantage of Lemma 1(i), we conclude that $\min(p, q, r) = 1$ and $\max(p, q, r) = \infty$ is necessary for the uniform boundedness of $\|\mu_{l,m,m}\|_{p,q,r}$. □

We will next eliminate the remaining possibilities. Recall that an $n \times n$ *Hadamard matrix* $H_n \in \mathbb{R}^{n\times n}$ is one with entries ± 1 such that $H_n^{\mathsf{T}} H_n = nI_n$ [27, Section 2.1]. The simplest example is $H_2 = \left[\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right]$. It is still unknown if $n \times n$ Hadamard matrices exist for every $n = 4m$, $m \in \mathbb{N}$, but it is well-known that the Kronecker product of H_2 with itself k times gives a $2^k \times 2^k$ Hadamard matrix for any $k \in \mathbb{N}$. The bottom line for us is that there exist Hadamard matrices of arbitrarily large dimensions, a fact that we use in our proof below whenever we take $n \to \infty$ with Hadamard matrices.

Theorem 6 (Uniqueness of Grothendieck's inequality) *Let* $1 \leqslant p, q, r \leqslant \infty$ *and* $l, m, n \in \mathbb{N}$. *Then* $\|\mu_{l,m,n}\|_{p,q,r}$ *is uniformly bounded for all* $l, m, n \in \mathbb{N}$ *if and only if*

$$(p, q, r) \in \{(1, 2, \infty),\ (\infty, 1, 2),\ (2, \infty, 1)\}.$$

Proof We will see that it suffices to take $l = m = n$ throughout this proof. By Lemma 1(i) and Proposition 1, we may assume that $p = 1$ and either $q = \infty$ or $r = \infty$. We will show that $\operatorname{tr}(XMY)$ is unbounded for judiciously chosen $n \times n$ real matrices X, M, and Y as $n \to \infty$.

Case I: $(1, q, \infty)$, $1 \leqslant q \leqslant \infty$. Suppose $2 < q \leqslant \infty$. Let $X_0 = n^{-1/q}\Delta$ for some arbitrary $\Delta = (\delta_{ij}) \in \{\pm 1\}^{n\times n}$ and let $Y_0 = I_n$. Then $\|X_0\|_{1,q} = n^{-1/q}\|\Delta\|_{1,q} = 1$ and $\|Y_0\|_{q,\infty} = \|I_n\|_{q,\infty} = 1$ by (11). For any $M = (M_{ij}) \in \mathbb{R}^{n\times n}$,

$$\begin{aligned}\max_{\|X\|_{1,q},\ \|Y\|_{q,\infty}\leqslant 1} |\operatorname{tr}(XMY)| \geqslant |\operatorname{tr}(X_0 M Y_0)| &= n^{-1/q}|\operatorname{tr}(\Delta M)| \\ &= n^{-1/q}\left|\sum\nolimits_{i,j=1}^{n} \delta_{ij} M_{ij}\right|.\end{aligned}$$

Since $\Delta \in \{\pm 1\}^{n\times n}$ is arbitrary, we will choose δ_{ij} so that $\delta_{ij} M_{ij}$ is nonnegative. Thus

$$\max_{\|X\|_{1,q},\ \|Y\|_{q,\infty}\leqslant 1} |\operatorname{tr}(XMY)| \geqslant n^{-1/q}\sum\nolimits_{i,j=1}^{n} |M_{ij}|. \tag{27}$$

Let $H_n \in \{\pm 1\}^{n\times n}$ be a Hadamard matrix. So $H_n H_n^{\mathsf{T}} = nI_n$ and all singular values of H_n are $\sqrt{n}$ [18]. Therefore, by (12),

$$\|H_n\|_{\infty,1} = \max_{\varepsilon,\delta\in\{\pm 1\}^n} |\varepsilon^{\mathsf{T}} H_n \delta| \leqslant \sigma_{\max}(H_n)\|\varepsilon\|_2\|\delta\|_2 = n^{3/2}. \tag{28}$$

Let $M = n^{-3/2}H_n$. Then $\|M\|_{\infty,1} \leqslant 1$ and by (27),

$$\max_{\|X\|_{1,q},\ \|Y\|_{q,\infty},\ \|M\|_{\infty,1}\leqslant 1} |\operatorname{tr}(XMY)| \geqslant n^{-1/q} \times n^{-3/2} \times n^2 = n^{1/2-1/q} \to \infty$$

as $n \to \infty$.

Suppose $1 \leqslant q < 2$. Since the Hölder conjugates are $r^* = 1$, $2 < q^* \leqslant \infty$, and $p^* = \infty$, by Lemma 1(ii), this reduces to the case we just treated.

Case II: $(1, \infty, r)$, $1 \leqslant r \leqslant \infty$. For $r = \infty$, we have $(1, \infty, \infty)$, which is same as the $q = \infty$ case in Case I. For $r = 1$, we have $(1, \infty, 1)$, but by Lemma 1(i), this is equivalent to $(1, 1, \infty)$, which is same as the $q = 1$ case in Case I. So we may assume $1 < r < \infty$.

Suppose $1 < r < 2$. Let $M = n^{1/r-1}I_n$ and $Y = n^{-3/2}H_n$ where $H_n \in \{\pm 1\}^{n\times n}$ is a Hadamard matrix. Then $\|M\|_{r,1} = n^{1/r-1}\|I_n\|_{r,1} = 1$ by (26), and $\|Y\|_{\infty,r} \leqslant \|Y\|_{\infty,1} = n^{-3/2}\|H_n\|_{\infty,1} \leqslant 1$ by (28). We choose $X \in \{\pm 1\}^{n\times n}$ such that $\operatorname{tr}(XH_n) = n^2$ and thus $\operatorname{tr}(XY) = n^{-1/2}$. Clearly $\|X\|_{1,\infty} = 1$ by (11). Hence

$$\operatorname{tr}(XMY) = n^{1/r-1}\operatorname{tr}(XY) = n^{1/r-1/2} \to \infty$$

as $n \to \infty$.

Suppose $2 < r < \infty$. Since the Hölder conjugates are $1 < r^* < 2$, $q^* = 1$, and $p^* = \infty$, by Lemma 1(ii) , this is equivalent to the case $(r^*, 1, \infty)$. Now by Lemma 1(i), this is in turn equivalent to the case $(1, \infty, r^*)$ with $1 < r^* < 2$, which is the case we just treated.

Suppose $r = 2$. Let $Y = n^{-1}H_n$ where $H_n \in \{\pm 1\}^{n\times n}$ is again a Hadamard matrix. Then

$$\|H_n\|_{\infty,2} = \max_{x\in\{\pm 1\}^n} \|H_n x\|_2 \leqslant \sigma_{\max}(H_n)\sqrt{n} = n.$$

So $\|Y\|_{\infty,2} \leqslant 1$. Let $M = n^{-1/2}I_n$. Then $\|M\|_{2,1} = 1$ by (26). Let $X \in \{\pm 1\}^{n\times n}$ be such that $\operatorname{tr}(XH_n) = n^2$ and thus $\operatorname{tr}(XY) = n$. Clearly $\|X\|_{1,\infty} = 1$ by (11). We have

$$\operatorname{tr}(XMY) = n^{-1/2}\operatorname{tr}(XY) = n^{1/2} \to \infty$$

as $n \to \infty$. □

6 Conclusion

We hope our characterization of Grothendieck's constant as a norm of the central object in the study of fast matrix multiplications would spur interactions between the two areas and perhaps even facilitate the determination of its exact value. Knowing that Grothendieck's inequality is a unique instance within a family of natural norm inequalities may help us better understand its ubiquity and utility.

References

1. Acín, A., Gisin, N., Toner, B.: Grothendieck's constant and local models for noisy entangled quantum states. Phys. Rev. A **73**(6), 0621055 (2006)
2. Alon, N., Berger, E.: The Grothendieck constant of random and pseudo-random graphs. Discrete Optim. **5**(2), 323–327 (2008)
3. Alon, N., Makarychev, K., Makarychev, Y., Naor, A.: Quadratic forms on graphs. Invent. Math. **163**(3), 499–522 (2006)
4. Alon, N., Naor, A.: Approximating the cut-norm via Grothendieck's inequality. SIAM J. Comput. **35**(4), 787–803 (2006)
5. Arora, S., Berger, E., Hazan, E., Kindler, G., Safra, M.: On non-approximability for quadratic programs. In: Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science, pp. 206–215 (2005)
6. Braverman, M., Makarychev, K., Makarychev, Y., Naor, A.: The Grothendieck constant is strictly smaller than Krivine's bound. Forum Math. Pi **1**, e442 (2013)
7. Briët, J., Buhrman, H., Toner, B.: A generalized Grothendieck inequality and nonlocal correlations that require high entanglement. Commun. Math. Phys. **305**(3), 827–843 (2011)
8. Briët, J., de Oliveira Filho, F.M., Vallentin, F.: The positive semidefinite Grothendieck problem with rank constraint. In: Automata, languages and programming. Part I, Lecture Notes in Comput. Sci., vol. 6198, pp. 31–42. Springer, Berlin (2010)

9. Briët, J., de Oliveira Filho, F.M., Vallentin, F.: Grothendieck inequalities for semidefinite programs with rank constraint. Theory Comput. **10**, 77–105 (2014)
10. Bürgisser, P., Clausen, M., Shokrollahi, A.: Algebraic Complexity Theory, vol. 315. Grundlehren der Mathematischen Wissenschaften, Springer, Berlin (1997)
11. Charikar, M., Wirth, A.: Maximizing quadratic programs: extending Grothendieck's inequality. In: Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science, pp. 54–60 (2004)
12. Coppersmith, D., Winograd, S.: Matrix multiplication via arithmetic progressions. J. Symbolic Comput. **9**(3), 251–280 (1990)
13. Davie, A.M.: Lower bound for k_g. Unpublished note (1984)
14. Davie, A.M.: Matrix norms related to Grothendieck's inequality. In: Banach spaces (Columbia, Mo., 1984), Lecture Notes in Math., vol. 1166, pp. 22–26. Springer, Berlin (1985)
15. Diviánszky, P., Bene, E., Vértesi, T.: Qutrit witness from the Grothendieck constant of order four. Phys. Rev., A(96) (2017)
16. Dwork, C., Nikolov, A., Talwar, K.: Efficient algorithms for privately releasing marginals via convex relaxations. Discrete Comput. Geom. **53**(3), 650–673 (2015)
17. Fishburn, P.C., Reeds, J.A.: Bell inequalities, Grothendieck's constant, and root two. SIAM J. Discrete Math. **7**(1), 48–56 (1994)
18. Friedland, S., Aliabadi, M.: Linear algebra and matrices. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (2018)
19. Friedland, S., Lim, L.-H.: Nuclear norm of higher-order tensors. Math. Comp. **87**(311), 1255–1281 (2018)
20. Friedland, S., Lim, L.-H., Zhang, J.: An elementary proof of Grothendieck's inequalty. Enseign. Math. **64**(3/4), 327–351 (2018)
21. Grothendieck, A.: Résumé de la théorie métrique des produits tensoriels topologiques. Bol. Soc. Mat. São Paulo **8**, 1–79 (1953)
22. Haagerup, U.: A new upper bound for the complex Grothendieck constant. Israel J. Math. **60**(2), 199–224 (1987)
23. Hendrickx, J.M., Olshevsky, A.: Matrix p-norms are NP-hard to approximate if $p \neq 1, 2, \infty$. SIAM J. Matrix Anal. Appl. **31**(5), 2802–2812 (2010)
24. Heydari, H.: Quantum correlation and Grothendieck's constant. J. Phys. A **39**(38), 11869–11875 (2006)
25. Hirsch, F., Quintino, M.T., Vértesi, T., Navascués, M., Brunner, N.: Better local hidden variable models for two-qubit werner states and an upper bound on the Grothendieck constant $K_G(3)$. Quantum **1**, 3 (2017)
26. Hitchcock, F.L.: The expression of a tensor or a polyadic as a sum of products. J. Math. Phys. **6**(1), 164–189 (1927)
27. Horadam, K.J.: Hadamard Matrices and their Applications. Princeton University Press, Princeton (2007)
28. Jameson, G.J.O.: Summing and Nuclear Norms in Banach Space Theory, vol. 8. London Mathematical Society Student Texts, Cambridge University Press, Cambridge (1987)
29. Khot, S., Naor, A.: Grothendieck-type inequalities in combinatorial optimization. Commun. Pure Appl. Math. **65**(7), 992–1035 (2012)
30. Khot, S., Naor, A.: Sharp kernel clustering algorithms and their associated Grothendieck inequalities. Random Struct. Algorithms **42**(3), 269–300 (2013)
31. Kindler, G., Naor, A., Schechtman, G.: The UGC hardness threshold of the L_p Grothendieck problem. Math. Oper. Res. **35**(2), 267–283 (2010)
32. Klaus, A.-L., Li, C.-K.: Isometries for the vector (p, q) norm and the induced (p, q) norm. Linear Multilinear Algebra **38**(4), 315–332 (1995)
33. Krivine, J.-L.: Constantes de Grothendieck et fonctions de type positif sur les sphères. Adv. Math. **31**(1), 16–30 (1979)
34. Landsberg, J.M.: Tensors: Geometry and Applications. Graduate Studies in Mathematics, American Mathematical Society, Providence (2012)
35. Lang, S.: Algebra, Graduate Texts in Mathematics, vol. 211, 3rd edn. Springer, New York (2002)
36. Le Gall, F.: Powers of tensors and fast matrix multiplication. In: ISSAC 2014—Proceedings of the 39th International Symposium on Symbolic and Algebraic Computation, pp. 296–303. ACM, New York (2014)

37. Lim, L.-H.: Tensors and Hypermatrices. Handbook of Linear Algebra, vol. 211, 2nd edn. CRC Press, Boca Raton (2013)
38. Lindenstrauss, J., Pełczyński, A.: Absolutely summing operators in L_p-spaces and their applications. Studia Math. **29**, 275–326 (1968)
39. Linial, N., Shraibman, A.: Lower bounds in communication complexity based on factorization norms. Random Struct. Algorithms **34**(3), 368–394 (2009)
40. Pisier, G.: Factorization of linear operators and geometry of Banach spaces. In: CBMS Regional Conference Series in Mathematics. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, vol. 60 (1986)
41. Pisier, G.: Grothendieck's theorem, past and present. Bull. Am. Math. Soc. (N.S.) **49**(2), 237–323 (2012)
42. Raghavendra, P.: Optimal algorithms and inapproximability results for every CSP? [extended abstract]. In: STOC'08, pp. 245–254. ACM, New York (2008)
43. Raghavendra, P., Steurer, D.: Towards computing the Grothendieck constant. In: Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 525–534. SIAM, Philadelphia, PA (2009)
44. Regev, O.: Bell violations through independent bases games. Quantum Inf. Comput. **12**(1–2), 9–20 (2012)
45. Regev, O., Toner, B.: Simulating quantum correlations with finite communication. SIAM J. Comput., 39(4):1562–1580(2009/10)
46. Strassen, V.: Gaussian elimination is not optimal. Numer. Math. **13**(4), 354–356 (1969)
47. Strassen, V.: Vermeidung von Divisionen. J. Reine Angew. Math. **264**, 184–202 (1973)
48. Strassen, V.: Rank and optimal computation of generic tensors. Linear Algebra Appl. **52**(53), 645–685 (1983)
49. Strassen, V.: Relative bilinear complexity and matrix multiplication. J. Reine Angew. Math. **375**(376), 406–443 (1987)
50. Tsirelson, B.S.: Quantum generalizations of Bell's inequality. Lett. Math. Phys. **4**(2), 93–100 (1980)
51. Williams, V.V.: Multiplying matrices faster than Coppersmith–Winograd [extended abstract]. In: STOC'12—Proceedings of the 2012 ACM Symposium on Theory of Computing, pp. 887–898. ACM, New York (2012)
52. Ye, K., Lim, L.-H.: Fast structured matrix computations: tensor rank and Cohn-Umans method. Found. Comput. Math. **18**(1), 45–95 (2018)

Affiliations

Shmuel Friedland[1] · Lek-Heng Lim[2] · Jinjie Zhang[3]

✉ Lek-Heng Lim
lekheng@galton.uchicago.edu

Shmuel Friedland
friedlan@uic.edu

Jinjie Zhang
jiz003@ucsd.edu

[1] Department of Mathematics, Statistics and Computer Science, University of Illinois, Chicago, IL, USA

[2] Computational and Applied Mathematics Initiative, Department of Statistics, University of Chicago, Chicago, IL, USA

[3] Department of Mathematics, University of California, San Diego, CA, USA