Counterexamples to Strassen's direct sum conjecture

by

Yaroslav Shitov

Moscow, Russia

The multiplicative complexity of systems of bilinear forms (and, in particular, the famous question of fast matrix multiplication) is an important area of research in modern theory of computation. One of the foundational papers on the topic is the work [42] by Strassen, who presented an $O(n^{\ln 7/\ln 2})$ algorithm for the multiplication of two $n \times n$ matrices. In his subsequent paper [43] published in 1973, Strassen asked whether the multiplicative complexity of the union of two bilinear systems depending on different variables is equal to the sum of the multiplicative complexities of both systems. A stronger version of this problem was proposed in the 1981 paper [19] by Feig and Winograd, who asked whether any optimal algorithm that computes such a pair of bilinear systems must compute each system separately. These questions became known as the *direct sum conjecture* and *strong direct sum conjecture*, respectively, and they were attracting a notable amount of attention during the four decades. As Feig and Winograd wrote, "*either a proof of, or a counterexample to, the direct sum conjecture will be a major step forward in our understanding of complexity of systems of bilinear forms.*"

A common framework to study the direct sum conjecture is the language of *tensors*, so let us give the precise formulation of the problem and then proceed with our short survey of its background. We recall that a bilinear system can be naturally represented as a 3-dimensional *tensor*, that is, an array of elements $T(i|j|k)$ taken from a field $\mathcal{F}$, where the triples (i, j, k) run over the Cartesian product of finite *indexing sets* I, J, and K. A tensor T is called *decomposable* if there exist vectors $a \in \mathcal{F}^I$, $b \in \mathcal{F}^J$, and $c \in \mathcal{F}^K$ satisfying $T = a \otimes b \otimes c$, which means that $T(i|j|k)$ equals $a_i b_j c_k$ for all $(i, j, k) \in I \times J \times K$. The *rank* of a tensor T, or the *multiplicative complexity* of the corresponding bilinear system, is the smallest r for which T can be written as a sum of r decomposable tensors with entries in $\mathcal{F}$. This quantity is denoted by $\mathrm{rank}_{\mathcal{F}} T$, and we remark that the rank of a tensor may change if one allows to take the entries of decomposable tensors as above from an

extension of $\mathcal{F}$; see [6]. Taking the union of two bilinear systems depending on disjoint sets of variables corresponds to the *direct sum* operation on tensors. More precisely, if T and T' are tensors with disjoint indexing sets I, I', J, J', K, and K', then we can define the direct sum $T \oplus T'$ as a tensor with indexing sets $I \cup I'$, $J \cup J'$, and $K \cup K'$ such that the $(I|J|K)$ block equals T and $(I'|J'|K')$ block equals T', and all entries outside of these blocks are zero. In other words, the direct sums of tensors are a multidimensional analogue of block-diagonal matrices; a basic result of linear algebra says that the ranks of such matrices are equal to the sums of the ranks of their diagonal blocks. Our paper is devoted to the 3-dimensional version of this statement, which has become known as *Strassen's direct sum conjecture*.

Conjecture 1. (See [43].) If T_1 and T_2 are tensors over an infinite field $\mathcal{F}$, then

$$\mathrm{rank}_{\mathcal{F}}(T_1 \oplus T_2) = \mathrm{rank}_{\mathcal{F}} \ T_1 + \mathrm{rank}_{\mathcal{F}} \ T_2.$$

The "$\leqslant$" inequality in Conjecture 1 is obvious because the direct sum of decompositions of T_1 and T_2 is a decomposition of $T_1 \oplus T_2$. If every optimal decomposition of $T_1 \oplus T_2$ arises in this way, then (T_1, T_2) are said to satisfy the *strong direct sum conjecture*; see [3], [9], [10]. Despite having been consistently mentioned and studied during the last four decades by many researchers representing different branches of mathematics, including computer science, algebraic geometry, and combinatorics, both Conjecture 1 and its strong form remained completely open to this date. Namely, its status remained unknown over every field, including the most important and widely studied special case of fields with zero characteristic.

As said above, the interest of computer scientists to Conjecture 1 stems from its relation to the complexity of the fundamental operation of the multiplication of matrices. It became known quite early that the analogue of Conjecture 1 cannot be true for the *border rank*, which is defined in the complex case as the smallest r such that a given tensor is the limit of a sequence of tensors of rank r. A counterexample is due to Schönhage, and the idea of his construction turned out to be exceptionally fruitful in further algorithms for matrix multiplication [40], [44]. Later on, the direct sum conjecture has been referred to and widely studied by the pioneers of the matrix multiplication topic; see the above mentioned paper [19] by Feig and Winograd, and a monograph [48] of Winograd on the arithmetic complexity of computation. In his seminal paper [34], Pan refers to Conjecture 1 as a "famous" problem and disproves its generalization to a more general class of so-called λ-*algorithms*; he mentions the direct sum problems in other earlier [33], [36], [37] and subsequent work [35]. The current research on matrix multiplication does still employ the ideas similar to the direct sum approach, as can be seen from a discussion of Conjecture 1 in a recent survey [46] by Vassilevska Williams.

In algebraic complexity theory, one of the major problems is to find an explicit family of polynomials that require formulas of super-polynomial size [8], [23], [39]. The connection to tensor ranks was established by Raz in [39], who proved that an explicit sequence of d-way $n \times ... \times n$ tensors whose rank grows at least as the $(1-o(1))d$-th power of n allows one to produce a family of polynomials needed to solve the above mentioned problem. However, the current techniques are not sufficient to get such a strong lower bound—no explicit family of $n \times n \times n$ tensors has ever been shown to have rank at least $3n$; see [2], [7], [26], [50]. Conjecture 1, if it was true, would allow one not only to break the $3n$ threshold but also to construct an explicit super-linear bound on the rank of an $n \times n \times n$ tensor, but in reality the question of lower bounds remains a notorious wide open problem.

In algebraic geometry, the concepts of rank and border rank admit a natural description via the secant varieties of the *Segre variety*; see [1], [4], [17], [28], [29]. This approach has led to interesting generalizations and relaxations of Conjecture 1, to further sufficient conditions for tensors to satisfy it, and, therefore, to many new classes of tensors for which the conjecture holds. Let us mention the paper [14] on the so-called *e-computable* tensors, the work [13] devoted to symmetric tensors, the paper [45] dealing with the cactus rank and catalecticant bound, the work [27] proving Conjecture 1 for tensors whose ranks can be computed by a particular adaptation of the so-called *substitution method*, the paper [15] studying the spaces of feasible rank decompositions in context of the direct sum conjecture, the work [12] containing further reformulations and generalizations of Conjecture 1 in terms of secant varieties, the monograph [25] containing a detailed discussion of this conjecture and its consequences for algebraic geometry, and a recent survey paper [16] on the topic.

Different rank-decomposition problems do also take an important place in linear algebra and combinatorics. The most famous of these is the matrix rank, the one this paper is devoted to is the tensor rank, and we can also mention the *non-negative matrix rank*, whose importance has been proved in many valuable recent publications [20], [24], [30], [49]. A version of Conjecture 1 that lies on a borderline of these approaches, that is, the additivity of the *non-negative tensor rank*, has been proved by Qi, Comon, and Lim in [38].

Concerning the positive results towards Conjecture 1, an early progress came from Winograd [47], who needed to prove the rank additivity for a specific class of bilinear systems, and his results have been extended by Auslander, Feig, and himself in [3]. In 1986, Ja'Ja' and Takche [22] proved Conjecture 1 in the case when one of the dimensions of either T_1 or T_2 is at most 2, and this result was subsequently generalized and discussed from the point of view of algebraic geometry by Landsberg and Michałek in [27]. The

best currently known positive result on Conjecture 1 seems to be that it is true if one of the tensors T_1 and T_2 has rank at most 6 or satisfies $|I|=|J|=3$; see [11].

1. Preliminaries

In our paper, we construct a counterexample to Conjecture 1. We begin with several additional definitions, known results, and notational conventions.

1.1. Notation.

Let $\mathcal{F}$ be a field; we denote by $\overline{\mathcal{F}}$ the algebraic closure of $\mathcal{F}$. For a tensor T in $\mathcal{F}^{I \times J \times K}$ and $k \in K$, we define the kth 3-*slice* as a matrix in $\mathcal{F}^{I \times J}$ whose (i,j) entry equals $T(i|j|k)$. For all $i \in I$ and $j \in J$, we can define the ith 1-slice and jth 2-slice of T in a similar way. The *support* of T is the smallest set $I_0 \times J_0 \times K_0 \subset I \times J \times K$ containing all the non-zero entries of T, and the sets I_0, J_0, and K_0 are called the 1-, 2-, and 3-supports of T. Two tensors are called *equivalent* if they become equal when restricted to their supports. If Σ is a non-empty finite set, then we define the Σ-*clone* of T as the tensor T_Σ obtained from T by taking $|\Sigma|$ copies of every element in every indexing set. Namely, we define the tensor T_Σ with indexing sets $\Sigma \times I$, $\Sigma \times J$, and $\Sigma \times K$ as $T_\Sigma(s_1,i|s_2,j|s_3,k)=T(i|j|k)$ for all $s_1,s_2,s_3 \in \Sigma$. Clearly, taking the clone of a tensor does not change its rank. All these notions can be defined for matrices instead of tensors in an analogous way.

1.2. Eliminating rank-one slices.

Now let $V_1 \subset \mathcal{F}^{J \times K}$, $V_2 \subset \mathcal{F}^{I \times K}$, and $V_3 \subset \mathcal{F}^{I \times J}$ be $\mathcal{F}$-linear spaces consisting of matrices. By $T \bmod (V_1,V_2,V_3)$ we denote the set of all tensors that can be obtained from T by adding elements of V_1 to the 1-slices of T, elements of V_2 to the 2-slices of the resulting tensor, and elements of V_3 to the 3-slices of what we obtained after adding the 1- and 2-slices. The following statement is well known; see e.g. Lemma 2 in [21] and Proposition 3.1 in [27].

LEMMA 2. *Let* $T \in \mathcal{F}^{I \times J \times K}$ *and* $K=\{1,...,k\} \cup \{1',...,k'\}$. *Denote an i-th 3-slice of* T *by* S_i *and let* V *be the* $\mathcal{F}$-*linear span of* $S_{1'},...,S_{k'}$. *Then,*

$$\operatorname{rank}_{\mathcal{F}} T \geqslant \min \operatorname{rank}_{\mathcal{F}} T \bmod (0,0,V) + \dim V,$$

and if $S_{1'},...,S_{k'}$ *are also rank-1, then the equality holds.*

Now let finite sets $\mathcal{M}_1$, $\mathcal{M}_2$, and $\mathcal{M}_3$ be bases of the linear spaces V_1, V_2, and V_3 as above. We define the tensor $\mathcal{T}$ with indexing sets $I \cup \mathcal{M}_1$, $J \cup \mathcal{M}_2$, and $K \cup \mathcal{M}_3$ as

(1) $\mathcal{T}(\alpha|\beta|\gamma) = T(\alpha|\beta|\gamma)$ if $(\alpha, \beta, \gamma) \in I \times J \times K$;

(2) for any $\chi \in \{1, 2, 3\}$ and any $m \in \mathcal{M}_\chi$, the mth χ-slice of $\mathcal{T}$ is equivalent to m (that is, coincides with m up to adding zero rows and columns).

We say that $\mathcal{T}$ is obtained from T by *adjoining* the 1-slices $\mathcal{M}_1$, the 2-slices $\mathcal{M}_2$, and the 3-slices $\mathcal{M}_3$, or simply by adjoining $(\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3)$. The result below follows from Lemma 2.

LEMMA 3. *Let T, $\mathcal{T}$, V_1, V_2, and V_3 be as above in this subsection. Then ,*

$$\operatorname{rank}_{\mathcal{F}} \mathcal{T} \geqslant \min \operatorname{rank}_{\mathcal{F}} T \bmod (V_1, V_2, V_3) + \dim V_1 + \dim V_2 + \dim V_3,$$

and if the matrices in the $\mathcal{M}_i$'s are rank-1, then the equality holds.

1.3. The main result and our strategy.

We are going to refute Conjecture 1 by proving the following result.

THEOREM 4. *For any infinite field $\mathcal{F}$, there exist tensors T_1 and T_2 over $\mathcal{F}$ such that*

$$\operatorname{rank}_{\mathcal{F}}(T_1 \oplus T_2) < \operatorname{rank}_{\overline{\mathcal{F}}} T_1 + \operatorname{rank}_{\overline{\mathcal{F}}} T_2.$$

Of course, the assertion with $\overline{\mathcal{F}}$ replaced by $\mathcal{F}$ in the summands of the right-hand side can only be weaker than the initial one, so Theorem 7 disproves Strassen's conjecture in the original formulation, that is, over any infinite field. Our counterexample rests on the following two claims, which we prove in subsequent sections.

CLAIM 5. *Let $T \in \mathcal{F}^{I \times J \times K}$ and assume that W_1, W_2, and W_3 are $\overline{\mathcal{F}}$-linear subspaces of $J \times K$, $I \times K$, and $I \times J$ matrices, respectively. We assume that W_1, W_2, and W_3 have bases consisting of matrices with entries in $\mathcal{F}$. Then, there are a finite set Σ and*

$$\mathcal{M}_1 \subset \mathcal{F}^{(\Sigma \times J) \times (\Sigma \times K)}, \quad \mathcal{M}_2 \subset \mathcal{F}^{(\Sigma \times I) \times (\Sigma \times K)}, \quad and \quad \mathcal{M}_3 \subset \mathcal{F}^{(\Sigma \times I) \times (\Sigma \times J)}$$

consisting of finitely many rank-1 matrices such that

(1) *$\operatorname{span} \mathcal{M}_\delta$ contains the Σ-clone of every $w_\delta \in W_\delta$, and*

(2) *the tensor $\mathcal{T}$ obtained by adjoining $(\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3)$ to the Σ-clone of T satisfies*

$$\operatorname{rank}_{\overline{\mathcal{F}}} \mathcal{T} = \min \operatorname{rank}_{\overline{\mathcal{F}}} T \bmod (W_1, W_2, W_3) + \sum_{\delta=1}^{3} \dim \operatorname{span} \mathcal{M}_\delta.$$

CLAIM 6. *For any infinite field $\mathcal{F}$, there is $T \in \mathcal{F}^{(I_1 \cup I_2) \times (J_1 \cup J_2) \times (K_1 \cup K_2)}$ such that*

$$\mathrm{rank}_{\mathcal{F}}\, T < \min \mathrm{rank}_{\overline{\mathcal{F}}}\, T_{111} \bmod (U_1, U_2, U_3) + \min \mathrm{rank}_{\overline{\mathcal{F}}}\, T_{222} \bmod (V_1, V_2, V_3),$$

where the indexing sets I_1, I_2, J_1, J_2, K_1, and K_2 are disjoint, the tensor T_{ijk} is the $(I_i|J_j|K_k)$ block of T, and U_1, U_2, U_3, V_1, V_2, and V_3 are the $\overline{\mathcal{F}}$-linear spaces spanned, respectively, by the 1-slices of T_{211}, by the 2-slices of T_{121}, by the 3-slices of T_{112}, by the 1-slices of T_{122}, by the 2-slices of T_{212}, and by the 3-slices of T_{221}.

THEOREM 7. *Claims 5 and 6 imply Theorem 4.*

Proof. We set $\mathcal{T}_1$ and M_δ^1 to be, respectively, the tensor and matrix sets obtained by the application of Claim 5 to the tensor T_{111} and linear spaces U_1, U_2, and U_3, as in Claim 6. We define $\mathcal{T}_2$ and M_δ^2 similarly, but with T_{222}, V_1, V_2, and V_3 taken instead of T_{111}, U_1, U_2, and U_3, and we assume without loss of generality that the sets Σ arisen from the applications of Claim 5 to T_{111} and T_{222} are equal. We write $\varrho = \mathrm{rank}_{\mathcal{F}}(\mathcal{T}_1 \oplus \mathcal{T}_2)$, and we denote by $\mathcal{W}_\delta$ the $\mathcal{F}$-linear space spanned by the set of all matrices equivalent to those in M_δ^1 and M_δ^2 of sizes corresponding to the δ-slices of $\mathcal{T}_1 \oplus \mathcal{T}_2$.

We prove Theorem 4 by checking that $\varrho < \mathrm{rank}_{\overline{\mathcal{F}}}\, \mathcal{T}_1 + \mathrm{rank}_{\overline{\mathcal{F}}}\, \mathcal{T}_2$, or, equivalently,

$$\varrho - D < \min \mathrm{rank}_{\overline{\mathcal{F}}}\, T_{111} \bmod (U_1, U_2, U_3) + \min \mathrm{rank}_{\overline{\mathcal{F}}}\, T_{222} \bmod (V_1, V_2, V_3), \qquad (1.1)$$

where $D = \sum_{\delta=1}^3 \dim \mathcal{W}_\delta$. Since $\mathcal{T}_1 \oplus \mathcal{T}_2$ is obtained from $T_{\Sigma 1} \oplus T_{\Sigma 2}$ by adjoining the bases of linear spaces $(\mathcal{W}_1, \mathcal{W}_2, \mathcal{W}_3)$, the left-hand side of (1.1) equals

$$\min \mathrm{rank}_{\mathcal{F}}(T_{\Sigma 1} \oplus T_{\Sigma 2}) \bmod (\mathcal{W}_1, \mathcal{W}_2, \mathcal{W}_3) \qquad (1.2)$$

by Lemma 3 (where $T_{\Sigma \chi}$ denotes the Σ-clone of $T_{\chi \chi \chi}$).

According to item (1) of Claim 5, the matrices equivalent to the Σ-clones of the 1-slices of T_{122} and T_{211} belong to $\mathcal{W}_1$, the matrices equivalent to the Σ-clones of the 2-slices of T_{121} and T_{212} belong to $\mathcal{W}_2$, and the matrices equivalent to the Σ-clones of the 3-slices of T_{112} and T_{221} belong to $\mathcal{W}_3$, so we see that the Σ-clone of T belongs to $(T_{\Sigma 1} \oplus T_{\Sigma 2}) \bmod (\mathcal{W}_1, \mathcal{W}_2, \mathcal{W}_3)$. In other words, the value (1.2) does not exceed $\mathrm{rank}_{\mathcal{F}}\, T$, and since this value equals $\varrho - D$, we complete the proof of (1.1) by applying the conclusion of Claim 6. $\square$

Now we need to check the validity of Claims 5 and 6; in the following section, we prove the former of them by developing the construction as in [41]. In §3, we prove Claim 6 by using results on ranks of generic tensors and sufficient conditions of algebraic independence in fields.

2. The proof of Claim 5

In this section, we follow the notation of Claim 5 and work over a field $\mathcal{F}$. Let $T \in \mathcal{F}^{I \times J \times K}$ be a tensor as in Claim 5; we begin by constructing the corresponding tensor $\mathcal{T}$ in the special case when $W_1 = W_2 = 0$ and W_3 is the 1-dimensional subspace spanned by a rank-r matrix W which has ones at the positions $(a_1, b_1), ..., (a_r, b_r) \in I \times J$ and zeros everywhere else. We set $\varrho = 2|I \times J \times K| + 1$, $\theta = \lceil \log_2 r \rceil$, and $\sigma = 2\varrho^2 r$. We define the trivial partition of a $\sigma \times \sigma$ matrix into σ^2 submatrices of size 1×1 as follows. Any number $s \in \{0, ..., \sigma^2 - 1\}$ can be written as $s = u_{1s}\sigma + u_{2s}$ with $u_{1s}, u_{2s} \in \{0, ..., \sigma - 1\}$. We will write $u(1, s) = \{u_{1s}\}$ and $u(2, s) = \{u_{2s}\}$. Clearly, the σ^2 sets $u(1, t) \times u(2, t)$ are a trivial partition of a $\sigma \times \sigma$ matrix. We proceed with the definition of the sets Σ and $\mathcal{M}_\delta$ as in Claim 5.

Definition 8. We set $\Sigma = \{0, ..., \sigma - 1\}^\theta$; we set π^i to be a function from $\{1, ..., \theta\}$ to $\{1, 2\}$ and assume that these functions are pairwise different for $i \in \{1, ..., r\}$. For any function Φ from $\{1, ..., \theta\}$ to $\{0, ..., \sigma^2 - 1\}$, we define the rank-1 zero-one matrix $M_\Phi \in \mathcal{F}^{(I \times \Sigma) \times (J \times \Sigma)}$ with 1-support equal to

$$\bigcup_{i=1}^{r} \{a_i\} \times u(\pi^i_1, \Phi_1) \times ... \times u(\pi^i_\theta, \Phi_\theta),$$

and 2-support equal to

$$\bigcup_{j=1}^{r} \{b_j\} \times u(3 - \pi^j_1, \Phi_1) \times ... \times u(3 - \pi^j_\theta, \Phi_\theta).$$

Here, the notation π^i_t and Φ_t stands for the images of t under π^i and Φ, respectively. Also, we call an $(a_i \times \Sigma) \times (b_j \times \Sigma)$ block of M_Φ diagonal if $i = j$. We define $\mathcal{M}$ as the set of all such M_Φ and those matrix units that correspond to entries which are non-zero in at least one of the non-diagonal blocks in any of the M_Φ.

Observation 9. The diagonal blocks of $\sum_\Phi M_\Phi$ are matrices of all ones.

Proof. If $\{S_1^1, ..., S_k^1\}, ..., \{S_1^q, ..., S_k^q\}$ are partitions of a set S, then

$$\{S_1^1, ..., S_k^1\} \times ... \times \{S_1^q, ..., S_k^q\}$$

is a partition of S^q. $\qquad \square$

Observation 10. The Σ-clone of W lies in the span of $\mathcal{M}$.

Proof. Follows from Observation 9 because we can turn the entries in the non-diagonal blocks to zeros thanks to the matrix units as in Definition 8. $\qquad \square$

Let us look at the structure of linear combinations of matrices in $\mathcal{M}$.

LEMMA 11. *For any $M \in \operatorname{span} \mathcal{M}$ and $i \neq j$, there is a permutation of rows and columns that sends the submatrix $M(a_i \times \Sigma | b_j \times \Sigma)$ to a $\sigma \times \sigma$ block-diagonal matrix with blocks of size $\sigma^{\theta-1} \times \sigma^{\theta-1}$.*

Proof. We can consider linear combinations of the matrices M_Φ only, because the union of their supports contains the support of any other matrix in $\mathcal{M}$. Since the functions π^i and π^j as in Definition 8 are different, we have $\pi^i_\tau = 3 - \pi^j_\tau$ for some $\tau \in \{1, ..., \theta\}$. Therefore, the τth element in the tuple of I-coordinates of any non-zero entry of $M(a_i \times \Sigma | b_j \times \Sigma)$ should coincide with the τth element in the tuple of its J-coordinates. This τth element can take σ different values, which determine a desired partition of the indexing set. $\square$

Observation 12. Let $S_1, ..., S_r$ be disjoint sets of cardinality c, and let their union be partitioned into c sets $D_1, ..., D_c$ each of which has exactly one element in every of the S_i's. Let us assume that a subset of less than c/r elements was removed from every S_i. Then there is a δ such that none of the elements of D_δ were removed.

Proof. Trivial. $\square$

LEMMA 13. *Let $\mathcal{M}_0$ be a set of $|K|$ elements in $\operatorname{span} \mathcal{M}$. Then either*
(1) *$\mathcal{M}_0$ contains a matrix of rank at least ϱ, or*
(2) *there are $\varphi_1, ..., \varphi_r, \psi_1, ..., \psi_r \in \Sigma$ such that, for every $M \in \mathcal{M}_0$, the matrix*

$$M(a_1 \times \varphi_1, ..., a_r \times \varphi_r \,|\, b_1 \times \psi_1, ..., b_r \times \psi_r)$$

is scalar.

Proof. Assuming that (1) is false, we apply Lemma 11 to a non-diagonal block $(a_i \times \Sigma | b_j \times \Sigma)$ of a matrix in $\mathcal{M}_0$, and we conclude that at most $\varrho \sigma^{\theta-1}$ rows are non-zero in this block. For every fixed i, we get an upper bound of

$$r \varrho \sigma^{\theta-1} |K| < \varrho^2 \sigma^{\theta-1} = \frac{|\Sigma|}{2r}$$

for the quantity of those elements $\xi \in a_i \times \Sigma$ such that the ξth row is non-zero in at least one of the non-diagonal (a_i, b_j) blocks in at least one of the matrices in $\mathcal{M}_0$. In other words, all the matrices obtained from those in $\mathcal{M}_0$ by removing at most $|\Sigma|/2r$ elements from every $a_i \times \Sigma$ will have zero non-diagonal blocks.

Now, we are going to find an $M_{\bar{\Phi}}$ whose non-zero entries of the diagonal blocks do not fall into the rows removed in the previous paragraph, and then we will take the values $\varphi_1, ..., \varphi_r, \psi_1, ..., \psi_r \in \Sigma$ which correspond to every $(a_i \times \varphi_i | b_i \times \psi_i)$ being the non-zero entry of the ith diagonal block of $M_{\bar{\Phi}}$. This will complete the proof, because these

are the M_Φ matrices only among those in $\mathcal{M}$ that can contribute to the diagonal blocks, and the support of any matrix M_Φ with $\Phi \neq \bar{\Phi}$ does not intersect with that of $M_{\bar{\Phi}}$ at a diagonal block. An appropriate $\bar{\Phi}$ can be found with the use of Observation 12, because the intersections of the supports of the $c := |\Sigma|^2$ matrices M_Φ with the set

$$S_i := (a_i \times \Sigma) \times (b_i \times \Sigma)$$

are singletons and partition S_i, and the removed rows contain at most

$$\frac{|\Sigma|^2}{2r} < \frac{c}{r}$$

entries in S_i. □

COROLLARY 14. $\min \operatorname{rank}_{\overline{\mathcal{F}}} T_\Sigma \bmod (0, 0, \operatorname{span} \mathcal{M}) = \min \operatorname{rank}_{\overline{\mathcal{F}}} T \bmod (0, 0, W)$.

Proof. The "$\leqslant$" inequality follows from Observation 10. In order to prove the "$\geqslant$" part, we note that adding a matrix $M \in \operatorname{span} \mathcal{M}$ to a 3-slice of T_Σ will only increase the rank of the tensor provided that the rank of M is at least ϱ. Therefore, we can use Lemma 13 and find, for any tensor T_0 of the minimal rank in $T_\Sigma \bmod (0, 0, \operatorname{span} \mathcal{M})$ and any $i \in I$ and $j \in J$, elements $\alpha(i) \in i \times \Sigma$ and $\beta(j) \in j \times \Sigma$ such that the restriction of T_0 to $(\alpha(I) | \beta(J) | K)$ belongs to $T \bmod (0, 0, W)$. □

Let us now see how the construction of $\mathcal{M}$ allows us to prove Claim 5. In particular, we can define the tensor $\mathcal{T}$ in the case when $W_1 = W_2 = 0$ and $\dim W_3 = 1$. In other words, this is the case when W_3 is spanned by a single matrix W' which can be written as PWQ, where W is the matrix as in the beginning of the section, and P and Q are invertible $I \times I$ and $J \times J$ matrices with entries in $\mathcal{F}$. Namely, we set $\mathcal{T}(T, 0, 0, W')$ to be the tensor obtained from the Σ-clone of T by adjoining the matrices $P^\Sigma M Q^\Sigma$ as 3-slices, where M runs over $\mathcal{M}$ and P^Σ stands for the Kronecker product of P and the $\Sigma \times \Sigma$ unity matrix.[1]

We now assume that (W_1, W_2, U_3) are as in Claim 5 and $W \in \mathcal{F}^{I \times J} \setminus U_3$; we denote the $\overline{\mathcal{F}}$-linear span of $U_3 \cup \{W\}$ by W_3. We set $\mathcal{T}(T, 0, 0, 0) = T$ and recursively define $\mathcal{T}(T, W_1, W_2, W_3)$ as $\mathcal{T}(T_0, W_{1\Sigma}, W_{2\Sigma}, U_{3\Sigma})$, where $T_0 = \mathcal{T}(T, 0, 0, W)$ is the tensor as above, and $W_{1\Sigma}$, $W_{2\Sigma}$, and $U_{3\Sigma}$ are the linear spaces consisting of the Σ-clones of matrices in W_1, W_2, and U_3, respectively. As we see, the tensor $\mathcal{T}(T, W_1, W_2, W_3)$ is obtained from a clone T_Ξ of T by adjoining matrices generating certain linear spaces $(\mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_3)$, and each of these linear spaces contains the Ξ-clones of matrices in the corresponding W_δ by Observation 10. This proves item (1) in Claim 5, and item (2) follows from Lemma 2 because the minimal rank of $T_\Xi \bmod (\mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_3)$ equals the minimal rank of $T \bmod (W_1, W_2, W_3)$ by Corollary 14.

---

[1] This definition corresponds to writing T with respect to the basis in which W' has the form W, then adjoining the matrices in $\mathcal{M}$, and then going back to the initial basis.

3. The proof of Claim 6

Throughout this section, we assume that $\mathcal{F}$ is an infinite field; we also fix a purely transcendental extension of $\mathcal{F}$ with an infinite basis. The elements of this basis are to be called *variables*, and one may actually think of them as variables changing in $\mathcal{F}$. We denote by n a sufficiently large integer, and we write $r(n) = \lfloor 0.34n^2 \rfloor$. We say that a tensor T with entries in an extension of a field F has τ *degrees of freedom over F* if the transcendence degree of the extension obtained from F by adjoining the entries of T equals τ. If τ equals the total number of entries of T, then T is called *generic over F*. We define the $2n \times 2n \times 2n$ tensor as

$$T = \sum_{\alpha=1}^{r(n)} (x^\alpha, \xi^\alpha) \otimes (y^\alpha, \gamma^\alpha) \otimes (z^\alpha, \zeta^\alpha),$$

where x^α, ξ^α, y^α, γ^α, z^α, and ζ^α are n-vectors whose entries are pairwise different variables denoted by x_i^α, ξ_i^α, y_i^α, γ_i^α, z_i^α, and ζ_i^α—and we denote the algebraic closure of the field obtained from $\mathcal{F}$ by adjoining all these variables by $\mathbb{F}$. We define the n-element indexing sets I_1, J_1, and K_1 corresponding to $x_1^\alpha, ..., x_n^\alpha$, to $y_1^\alpha, ..., y_n^\alpha$, and to $z_1^\alpha, ..., z_n^\alpha$, respectively. Similarly, the sets I_2, J_2, and K_2 correspond to $\xi_1^\alpha, ..., \xi_n^\alpha$, to $\gamma_1^\alpha, ..., \gamma_n^\alpha$, and to $\zeta_1^\alpha, ..., \zeta_n^\alpha$. The rank of T does not exceed $r(n)$ for any assignment of values in $\mathcal{F}$ to the variables x_i^α, ξ_i^α, y_i^α, γ_i^α, z_i^α, and ζ_i^α. So, we will conclude the proof of Claim 6 if we produce an assignment satisfying

$$r(n) < \min \operatorname{rank}_\mathbb{F} T_{111} \bmod (U_1, U_2, U_3) + \min \operatorname{rank}_\mathbb{F} T_{222} \bmod (V_1, V_2, V_3), \qquad (3.1)$$

where T_{ijk} denotes $T(I_i | J_j | K_k)$, and U_1, U_2, U_3, V_1, V_2, and V_3 are the $\mathbb{F}$-linear spaces spanned, respectively, by the 1-slices of T_{211}, by the 2-slices of T_{121}, by the 3-slices of T_{112}, by the 1-slices of T_{122}, by the 2-slices of T_{212}, and by the 3-slices of T_{221}.

We keep working with generic tensors and consider $\Phi \in T_{111} \bmod (U_1, U_2, U_3)$. Every such Φ can be obtained by adding $\mathbb{F}$-linear combinations of the 3-slices of T_{112} to every of the 3-slices of T_{111}, then $\mathbb{F}$-linear combinations of the 2-slices of T_{121} to the 2-slices of the resulting tensor, and finally $\mathbb{F}$-linear combinations of the 1-slices of T_{211} to the 1-slices of what was obtained. Denoting by A_i^τ, B_j^τ, and C_k^τ the coefficients of these linear combinations, we get the expression

$$\varphi_{ijk} = \sum_{\alpha=1}^{r(n)} (x_i^\alpha y_j^\alpha z_k^\alpha + a_i^\alpha y_j^\alpha z_k^\alpha + x_i^\alpha b_j^\alpha z_k^\alpha + x_i^\alpha y_j^\alpha c_k^\alpha)$$

for the (i, j, k) entry of Φ. Here, x_i^α, y_j^α, and z_k^α are variables defined above, and

$$a_i^\alpha = \sum_{\tau=1}^n A_i^\tau \xi_\tau^\alpha, \quad b_j^\alpha = \sum_{\tau=1}^n B_j^\tau \gamma_\tau^\alpha, \quad \text{and} \quad c_k^\alpha = \sum_{\tau=1}^n C_k^\tau \zeta_\tau^\alpha$$

are elements of $\mathbb{F}$. We write

$$\mathbb{K} = \mathcal{F}(A_i^\tau, B_i^\tau, C_i^\tau, \xi_i^\alpha, \gamma_i^\alpha, \zeta_i^\alpha), \tag{3.2}$$

and we are going to prove the following.

PROPOSITION 15. *The tensor Φ has at least $n^3 - 6\sqrt{3}n^{2.5}$ degrees of freedom over $\mathbb{K}$.*

Let us see that Proposition 15 implies Claim 6 before we proceed with its proof. We say that a property Π holds for *almost all m-vectors* over an infinite field F, if there is a non-zero polynomial ψ with m variables and coefficients in an extension of F such that $\psi(v) = 0$ is a necessary condition for v not to possess Π.

LEMMA 16. *Proposition 15 implies Claim 6.*

Proof. We write (3.1) as a first-order formula Ψ in the language of fields with free variables $\mathcal{V} = (x_i^\alpha, \xi_i^\alpha, y_i^\alpha, \gamma_i^\alpha, z_i^\alpha, \zeta_i^\alpha)$. The quantifier elimination (see Theorem 1.23 in [5]) allows us to assume that Ψ is a conjunction of clauses $(f_1 *_1 0) \vee ... \vee (f_k *_k 0)$, where $*_k$ is either $=$ or $\neq$, and f_i is a polynomial that depends on $\mathcal{V}$ and has integral coefficients. Therefore, inequality (3.1) does either hold for almost all assignments of $\mathcal{V}$ or fail for almost all such assignments. Now, it is enough to prove (3.1) for generic $\mathcal{V}$, that is, for the tensor T defined in the beginning of this section. Note that the summands in the right-hand side of (3.1) are equal, so it remains to prove that $\operatorname{rank}_\mathbb{F} \Phi > \frac{1}{2}r(n)$. Since the entries of any rank-1 $n \times n \times n$ tensor are products of $3n$ elements, the quantity as in Proposition 15 cannot be greater than $3n \operatorname{rank}_\mathbb{F} \Phi$. Therefore, Proposition 15 implies that $\operatorname{rank}_\mathbb{F} \Phi$ is at least $\frac{1}{3}n^2 - 2\sqrt{3}n^{1.5}$, which is greater than $\frac{1}{2}r(n)$ for large n. $\square$

The following is a very well-known result in algebraic complexity theory; see Theorem 2.2 in [18] for a simple elementary proof in characteristic zero and also Theorem 22 and Remark 23 in [32] for a more advanced treatment for positive characteristic. In particular, the first two paragraphs of the proof of Theorem 2.2 in [18] remain valid even if one replaces $\mathbb{C}$ in the formulation by arbitrary field F; this is so because it is sufficient to consider the case when F is algebraically closed, and in this case the polynomial G as in their proof should be separable.

LEMMA 17. (See Theorem 2.2 in [18] and Remark 23 in [32].) *Let F be a field, and let $p_1, ..., p_q$ be elements of the algebraic closure of $F(s_1, ..., s_q)$. Looking at $p_1, ..., p_q$ as algebraic functions of variables $s_1, ..., s_q$, we define their Jacobian to be the $q \times q$ matrix with (i, j) entry given by $\partial p_i / \partial s_j$. If this Jacobian has non-zero determinant, then $p_1, ..., p_q$ are algebraically independent over F.*

LEMMA 18. *Let F be an infinite field, let $q<n$ be positive integers, and let μ be an n-vector that has at least q degrees of freedom over F. Then the q-vector $Q\mu$ is generic over F for almost all $q \times n$ matrices Q with entries in F.*

Proof. We pick a transcendence basis $\mu_{s_1}, ..., \mu_{s_k}$ of the entries of μ. Every coordinate of μ is now an algebraic function of this basis, so we can define the Jacobian as the $q \times k$ matrix with (i,j) entry given by $\partial(Q\mu)_i/\partial(\mu_{s_j})$. According to Lemma 17, the Jacobian having rank q is a sufficient condition for $Q\mu$ to be generic. This condition is valid when Q is the matrix with ones at the positions (t, s_t) and zeros everywhere else, so one of the $q \times q$ minors of the Jacobian is really a non-zero polynomial in the entries of Q. □

LEMMA 19. *Let F be an infinite field, let $d<n$ be positive integers, and let M be an $n \times m$ matrix that has at least $mn - \delta$ degrees of freedom over F. Let Q be an almost arbitrary $(n-d) \times n$ matrix over F. Then, there are at most δ/d columns in QM whose removal leaves a matrix that is generic with respect to F.*

Proof. First, we assume $m=1$. If $d \leqslant \delta$, then there is nothing to prove because the assumption of the lemma allows us to remove the only column of QM. Otherwise, we apply Lemma 18 and conclude the consideration of the case $m=1$.

Now, we assume $m>1$ and proceed by induction; we define F_j and F_j' as the fields obtained from F by adjoining the entries of the first j columns of M and QM, respectively. A $\hat{\jmath}$th column is called *weak* if the transcendence degree of $F_{\hat{\jmath}}$ over $F_{\hat{\jmath}-1}$ is at most $n-d$. The removal of any weak column of M leaves us a matrix having at least $mn - \delta - n + d$ degrees of freedom, and then we can complete the proof by induction. If no weak columns exist, then Lemma 18 shows that the jth column of QM is generic with respect to F_{j-1} (and consequently with respect to the field $F_{j-1}' \subseteq F_{j-1}$). In particular, the extension $F_j' \supset F_{j-1}'$ has transcendence degree $n-d$, so F_m' has transcendence degree $m(n-d)$ over $F=F_0'$, which means that QM is generic. □

THEOREM 20. *(See Theorem 4.4 in [31].) Let $u^1, ..., u^k$, $v^1, ..., v^k$, and $w^1, ..., w^k$ be n-vectors whose coefficients are algebraically independent over a field F in common. If $k > n^3/(3n-2)$, then the tensor $\sum_{\alpha=1}^{k} u^\alpha \otimes v^\alpha \otimes w^\alpha$ is generic over F.*

Observation 21. Let $p_1, ..., p_q$ be polynomials of the same total degree d over a field F. If $(p_1^{in}, ..., p_q^{in})$ are generic over F, then $(p_1, ..., p_q)$ are generic over F as well, where p^{in} denotes the sum of all monomials of p whose degree is maximal.

Proof. If $h(p_1, ..., p_q)=0$, then $h^{in}(p_1^{in}, ..., p_q^{in})=0$. □

Let $P=\sum_\alpha u^\alpha \otimes v^\alpha \otimes w^\alpha$ be an $n \times n \times n$ tensor over a field F, and let Q be a $q \times n$

matrix over a subfield $K \subset F$. We define

$$P_Q = \sum_\alpha (Qu^\alpha) \otimes (Qv^\alpha) \otimes (Qw^\alpha),$$

and we note that P_Q is obtained from T by a sequence of K-linear transformations of its 1-slices, 2-slices, and 3-slices. In particular, P has at least as many degrees of freedom over K as P_Q does have. We are now ready to prove Proposition 15.

Proof of Proposition 15. Let M be the $n \times 3r(n)$ matrix formed by the vectors x^α, y^α, and z^α as in the definition of T. Since the field $\mathbb{K}$ defined in (3.2) has at most $3nr(n) + 3n^2$ degrees of freedom over $\mathcal{F}$, and since $\mathbb{K}(x^\alpha, y^\alpha, z^\alpha)$ contains

$$\mathcal{F}(x^\alpha, y^\alpha, z^\alpha, \xi^\alpha, \gamma^\alpha, \zeta^\alpha),$$

the matrix M has at least

$$6nr(n) - (3nr(n) + 3n^2) = 3nr(n) - 3n^2$$

degrees of freedom over $\mathbb{K}$. Now, we take $d = \lfloor \sqrt{3n} \rfloor$ and apply Lemma 19 to the matrix M and field $\mathbb{K}$; we denote the vector $Q\pi^\alpha$ by $\bar{\pi}^\alpha$, where π may stand for a, b, c, x, y, or z. The $(n-d) \times (n-d) \times (n-d)$ tensor Φ_Q equals

$$\sum_{\alpha=1}^{r(n)} (\bar{x}^\alpha \bar{y}^\alpha \bar{z}^\alpha + \bar{a}^\alpha \bar{y}^\alpha \bar{z}^\alpha + \bar{x}^\alpha \bar{b}^\alpha \bar{z}^\alpha + \bar{x}^\alpha \bar{y}^\alpha \bar{c}^\alpha). \tag{3.3}$$

We denote by $\mathcal{A}$ the set of all α such that none of the three columns x^α, y^α, and z^α were identified as "weak" and removed while applying Lemma 19. According to this lemma, the union of the vectors $\bar{x}^\alpha$, $\bar{y}^\alpha$, and $\bar{z}^\alpha$, where $\alpha \in \mathcal{A}$ is generic over $\mathbb{K}$. The number of weak columns is at most $3n^2/d$, so that $|\mathcal{A}| > n^2/(3n-2)$, and the $\mathcal{A}$-part of the sum in (3.3) is a tensor generic over $\mathbb{K}$, according to Theorem 20 and Observation 21. (In order to use these, we look at the coordinates of $\bar{x}^\alpha$, $\bar{y}^\alpha$, and $\bar{z}^\alpha$, with $\alpha \in \mathcal{A}$, as variables, and note that $\bar{a}^\alpha$, $\bar{b}^\alpha$, and $\bar{c}^\alpha$ have coordinates in $\mathbb{K}$.) Therefore, Φ_Q is the sum of a tensor generic over $\mathbb{K}$ and a tensor of rank at most $3n^2/d$, so it has at least

$$(n-d)^3 - 3n\frac{3n^2}{d} \geqslant n^3 - 6\sqrt{3}n^{2.5}$$

degrees of freedom over $\mathbb{K}$. $\qquad \square$

Therefore, we completed the proof of Proposition 15, which implies Claim 6, according to Lemma 16. Since Claim 5 was proved in the previous section, we can use Theorem 7 and complete the proof of Theorem 4.

4. Concluding remarks

As said above, the author disproved the direct sum conjecture in the original formulation, that is, over any infinite field. In fact, his proof is going to work for sufficiently large finite fields as well (although he did not try to give an effective upper bound on the size of the smallest field for which it may not work). The author thinks that the direct sum conjecture is false over any field, but in order to construct a counterexample in the way similar to this paper, one would need to estimate the cardinalities of the arising sets of matrices instead of the corresponding dimensions, as it is done in the current proof. Actually, it seems possible to achieve a counterexample in this way, despite a much greater amount of technical difficulties.

A multidimensional analogue of the direct sum conjecture states that the rank of d-dimensional tensors is additive with respect to direct sums. This problem remains open for $d \geqslant 4$, and the author is not sure that the present approach can lead to a progress on the multidimensional version. The analogous statement but restricted to symmetric tensors is open already for $d \geqslant 3$; see [13], [14], [45] for a review of the current state of art and new results on this problem. As said above, the *border rank* version of the direct sum conjecture is known to fail for d-tensors at least when $d=3$; see also a discussion in Chapter 11 of [25].

5. Acknowledgements

References

[1] ABO, H., OTTAVIANI, G. & PETERSON, C., Induction for secant varieties of Segre varieties. *Trans. Amer. Math. Soc.*, 361 (2009), 767–792.

[2] ALEXEEV, B., FORBES, M. A. & TSIMERMAN, J., Tensor rank: some lower and upper bounds, in 26*th Annual IEEE Conference on Computational Complexity*, pp. 283–291. IEEE Computer Soc., Los Alamitos, CA, 2011.

[3] AUSLANDER, L., FEIG, E. & WINOGRAD, S., Direct sums of bilinear algorithms. *Linear Algebra Appl.*, 38 (1981), 175–192.

[4] BALLICO, E. & BERNARDI, A., Tensor ranks on tangent developable of Segre varieties. *Linear Multilinear Algebra*, 61 (2013), 881–894.

[5] BASU, S., POLLACK, R. & ROY, M. F., *Algorithms in Real Algebraic Geometry*. Algorithms and Computation in Mathematics, 10. Springer, Berlin–Heidelberg, 2003.

[6] BERGMAN, G. M., Ranks of tensors and change of base field. *J. Algebra*, 11 (1969), 613–621.

[7] BLÄSER, M., Improvements of the Alder–Strassen bound: algebras with nonzero radical, in *Automata, Languages and Programming*, Lecture Notes in Comput. Sci., 2076, pp. 79–91. Springer, Berlin–Heidelberg, 2001.

[8] — Explicit tensors, in *Perspectives in Computational Complexity*, Progr. Comput. Sci. Appl. Logic, 26, pp. 117–130. Birkhäuser/Springer, Cham, 2014.

[9] BSHOUTY, N. H., On the extended direct sum conjecture, in *Proc. 21st Annual ACM Symposium on Theory of Computing*, pp. 177–185. ACM, New York, NY, 1989.

[10] — On the direct sum conjecture in the straight line model. *J. Complexity*, 14 (1998), 49–62.

[11] BUCZYŃSKI, J., POSTINGHEL, E. & RUPNIEWSKI, F., On Strassen's rank additivity for small three-way tensors. Preprint, 2019. [arXiv:1902.06582 \[math.AG\]](arXiv:1902.06582).

[12] BUCZYŃSKI, J. A., GINENSKY, A. & LANDSBERG, J. M., Determinantal equations for secant varieties and the Eisenbud–Koh–Stillman conjecture. *J. Lond. Math. Soc.*, 88 (2013), 1–24.

[13] CARLINI, E., CATALISANO, M. V. & CHIANTINI, L., Progress on the symmetric Strassen conjecture. *J. Pure Appl. Algebra*, 219 (2015), 3149–3157.

[14] CARLINI, E., CATALISANO, M. V., CHIANTINI, L., GERAMITA, A. V. & WOO, Y., Symmetric tensors: rank, Strassen's conjecture and *e*-computability. *Ann. Sc. Norm. Super. Pisa Cl. Sci.*, 18 (2018), 363–390.

[15] CARLINI, E., CATALISANO, M. V. & ONETO, A., Waring loci and the Strassen conjecture. *Adv. Math.*, 314 (2017), 630–662.

[16] CASAROTTI, A., MASSARENTI, A. & MELLA, M., On Comon's and Strassen's conjectures. *Mathematics*, 6 (2018), 217.

[17] CATALISANO, M. V., GERAMITA, A. V. & GIMIGLIANO, A., Ranks of tensors, secant varieties of Segre varieties and fat points. *Linear Algebra Appl.*, 355 (2002), 263–285. Erratum in *Linear Algebra Appl.*, 367 (2003), 347–348.

[18] EHRENBORG, R. & ROTA, G.-C., Apolarity and canonical forms for homogeneous polynomials. *European J. Combin.*, 14 (1993), 157–181.

[19] FEIG, E. & WINOGRAD, S., On the direct sum conjecture, in *Proc. 22nd Annual Symposium on Foundations of Computer Science*, pp. 91–94. IEEE Computer Soc., Nashville, TN, 1981.

[20] GILLIS, N., The why and how of nonnegative matrix factorization, in *Regularization, Optimization, Kernels, and Support Vector Machines*, Chapman & Hall/CRC Mach. Learn. Pattern Recogn. Ser., pp. 257–291. CRC Press, Boca Raton, FL, 2015.

[21] HOPCROFT, J. E. & KERR, L. R., On minimizing the number of multiplications necessary for matrix multiplication. *SIAM J. Appl. Math.*, 20 (1971), 30–36.

[22] Ja'Ja', J. & Takche, J., On the validity of the direct sum conjecture. *SIAM J. Comput.*, 15 (1986), 1004–1020.

[23] Kalorkoti, K. A., A lower bound for the formula size of rational functions. *SIAM J. Comput.*, 14 (1985), 678–687.

[24] Kubjas, K., Robeva, E. & Sturmfels, B., Fixed points EM algorithm and nonnegative rank boundaries. *Ann. Statist.*, 43 (2015), 422–461.

[25] Landsberg, J. M., *Tensors*: *Geometry and Applications*. Graduate Studies in Mathematics, 128. Amer. Math. Soc., Providence, RI, 2012.

[26] — New lower bounds for the rank of matrix multiplication. *SIAM J. Comput.*, 43 (2014), 144–149.

[27] Landsberg, J. M. & Michałek, M., Abelian tensors. *J. Math. Pure Appl.*, 108 (2017), 333–371.

[28] — On the geometry of border rank decompositions for matrix multiplication and other tensors with symmetry. *SIAM J. Appl. Algebra Geom.*, 1 (2017), 2–19.

[29] Landsberg, J. M. & Teitler, Z., On the ranks and border ranks of symmetric tensors. *Found. Comput. Math.*, 10 (2010), 339–366.

[30] Lee, D. D. & Seung, H. S., Learning the parts of objects by non-negative matrix factorization. *Nature*, 401 (1999), 788–791.

[31] Lickteig, T., Typical tensorial rank. *Linear Algebra Appl.*, 69 (1985), 95–120.

[32] Mittmann, J., Saxena, N. & Scheiblechner, P., Algebraic independence in positive characteristic: a *p*-adic calculus. *Trans. Amer. Math. Soc.*, 366 (2014), 3425–3450.

[33] Pan, V., A unified approach to the analysis of bilinear algorithms. *J. Algorithms*, 2 (1981), 301–310.

[34] — How can we speed up matrix multiplication? *SIAM Rev.*, 26 (1984), 393–415.

[35] — Fast feasible and unfeasible matrix multiplication. Preprint, 2018. arXiv:1804.04102 [cs.DS].

[36] Pan, V. Y., New combinations of methods for the acceleration of matrix multiplication. *Comput. Math. Appl.*, 7 (1981), 73–125.

[37] — The techniques of trilinear aggregating and the recent progress in the asymptotic acceleration of matrix operations. *Theoret. Comput. Sci.*, 33 (1984), 117–138.

[38] Qi, Y., Comon, P. & Lim, L. H., Semialgebraic geometry of nonnegative tensor rank. *SIAM J. Matrix Anal. Appl.*, 37 (2016), 1556–1580.

[39] Raz, R., Tensor-rank and lower bounds for arithmetic formulas. *J. ACM*, 60 (2013), Art. 40, 15 pp.

[40] Schönhage, A., Partial and total matrix multiplication. *SIAM J. Comput.*, 10 (1981), 434–455.

[41] Shitov, Y., A counterexample to Comon's conjecture. *SIAM J. Appl. Algebra Geom.*, 2 (2018), 428–443.

[42] Strassen, V., Gaussian elimination is not optimal. *Numer. Math.*, 13 (1969), 354–356.

[43] — Vermeidung von Divisionen. *J. Reine Angew. Math.*, 264 (1973), 184–202.

[44] — The asymptotic spectrum of tensors and the exponent of matrix multiplication, in *Proc.* 27*th Annual Symposium on Foundations of Computer Science*, pp. 49–54. IEEE Computer Soc., Washington, DC, 1986.

[45] Teitler, Z., Sufficient conditions for Strassen's additivity conjecture. *Illinois J. Math.*, 59 (2015), 1071–1085.

[46] Vassilevska Williams, V., An overview of the recent progress on matrix multiplication. *ACM SIGACT News*, 43 (2012), 57–69.

[47] Winograd, S., Some bilinear forms whose multiplicative complexity depends on the field of constants. *Math. Systems Theory*, 10 (1976/77), 169–180.

[48] — *Arithmetic Complexity of Computations.* CBMS-NSF Regional Conference Series in Applied Mathematics, 33. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1980.

[49] YANNAKAKIS, M., Expressing combinatorial optimization problems by linear programs. *J. Comput. System Sci.*, 43 (1991), 441–466.

[50] ZUIDDAM, J., A note on the gap between rank and border rank. *Linear Algebra Appl.*, 525 (2017), 33–44.

YAROSLAV SHITOV
Izumrudnaya ulitsa 65
kvartira 4
Moscow 129346
Russia
yaroslav-shitov@yandex.ru