Small Sunflowers and the Structure of Slice Rank Decompositions

Thomas Karam
Mathematical Institute, University of Oxford, UK

Abstract

Let $d \geq 3$ be an integer. We show that whenever an order-d tensor admits $d+1$ decompositions according to Tao's slice rank, if the linear subspaces spanned by their one-variable functions constitute a sunflower for each choice of special coordinate, then the tensor admits a decomposition where these linear subspaces are contained in the centers of these respective sunflowers. As an application, we deduce that for every nonnegative integer k and every finite field $\mathbb{F}$ there exists an integer $C(d,k,|\mathbb{F}|)$ such that every order-d tensor with slice rank k over $\mathbb{F}$ admits at most $C(d,k,|\mathbb{F}|)$ decompositions with length k, up to a class of transformations that can be easily described.

2012 ACM Subject Classification Mathematics of computing → Discrete mathematics; Mathematics of computing → Combinatorics

Keywords and phrases Slice rank, tensors, sunflowers, decompositions

Digital Object Identifier 10.4230/LIPIcs.ITCS.2024.67

Related Version *Full Version*: `https://arxiv.org/abs/2308.07101`

Funding The author is supported by a European Research Council (ERC) Grant, number 883810.

Acknowledgements The author thanks Timothy Gowers for discussions around Gowers's paper [8] that attracted the author's attention to the idea of slice rank decompositions of the zero tensor, which was involved there and plays a role in the present paper as well. The author also thanks Jordan Ellenberg for encouraging him to prove results stating that even if the analogue of a "perfect" property which holds for the matrix rank does not hold for other notions of rank on tensors, some appropriate weakening of it does.

1 Introduction

Throughout we will use the following notations. If k is a positive integer, then $[k]$ will denote the set $\{1,\dots,k\}$ of positive integers up to k. The letter $\mathbb{F}$ will denote a field. All our statements will be uniform with respect to the choice of the field $\mathbb{F}$, unless we explicitly write the assumption that the field $\mathbb{F}$ is finite. Even in that latter case, the dependence will only be on the size of the field $\mathbb{F}$ and not involve its structure.

We will prove various statements involving order-d tensors, where $d \geq 2$ is some integer. We will refer to functions $[n_1] \times \dots \times [n_d] \to \mathbb{F}$ for some positive integers $n_1,\dots,n_d$ as *order-d tensors*. All our statements will be uniform in the integers $n_1,\dots,n_d$ once all other parameters are fixed, and we will use the integers $n_1,\dots,n_d$ without defining them again in our statements.

Let k be a positive integer, and let $\mathbb{F}$ be a field. If $M : [n_1] \times [n_2] \to \mathbb{F}$ is a matrix with rank k, then all decompositions of M as a sum of k rank-1 matrices can be obtained from one another up to changes of bases. In particular, there exist some linear subspaces $A_1 \subset \mathbb{F}^{n_1}, A_2 \subset \mathbb{F}^{n_2}$ such that if

$$M(x,y) = \sum_{i=1}^{k} f_i(x) g_i(y)$$

is a decomposition of M, then $\langle f_1, \dots, f_k \rangle = A_1$ and $\langle g_1, \dots, g_k \rangle = A_2$. If $\mathbb{F}$ is a finite field, then it is well-known that the number of decompositions of M, where two decompositions are viewed as the same if their $2k$-tuples $(f_1, \dots, f_k, g_1, \dots, g_k)$ are the same, is equal to the number

$$(|\mathbb{F}|^k - 1)(|\mathbb{F}|^k - |\mathbb{F}|) \dots (|\mathbb{F}|^k - |\mathbb{F}|^{k-1}) = |\mathbb{F}|^{k^2} \prod_{i=1}^{k} (1 - |\mathbb{F}|^{-i}) \tag{1}$$

of bases of A_1. This number is between $\omega|\mathbb{F}|^{k^2}$ and $|\mathbb{F}|^{k^2}$, where we take ω to be the absolute constant

$$\prod_{i=1}^{\infty} (1 - 2^{-i}) > 0.$$

We note that even from a purely qualitative perspective, the analogous boundedness statement becomes false in general for decompositions of length even one greater than the rank of M. For instance, if $M(x, y) = f(x)g(y)$, then for any functions $f_1, f_2 : [n_1] \to \mathbb{F}$ satisfying $f_2 - f_1 = f$ we have

$$M(x, y) = f_2(x)g(y) - f_1(x)g(y),$$

so as long as the function f_2 is not fixed, the function f_1 can be completely arbitrary.

Our main aim in the present paper is to obtain a comparable statement for the notion of slice rank on higher-order tensors. We have found the task of formulating such a statement to already be challenging, which is why we shall state it precisely only in Section 3, after going through several constructions guiding us towards its formulation and ruling out stronger ones in Section 2. Informally speaking, we will prove that up to a natural class of transformations, the number of minimal-length slice rank decompositions of a tensor over a finite field is bounded above in a way that depends only on the order of the tensor, on its slice rank, and on the size of the field. As we will also discuss, the bound that we obtain cannot be too far from the optimal bound. Let us recall the definition of the slice rank and even before this, the definition of the tensor rank.

► **Definition 1.** *Let $d \geq 2$ be an integer, and let $T : [n_1] \times \dots \times [n_d] \to \mathbb{F}$ be an order-d tensor. We say that the* tensor rank *of T, denoted by $\operatorname{tr} T$, is the smallest nonnegative integer k such that there exist functions $a_{j,i} : [n_j] \to \mathbb{F}$ for each $j \in [d]$ and $i \in [k]$ such that*

$$T(x_1, \dots, x_d) = \sum_{i=1}^{k} a_{1,i}(x_1) a_{2,i}(x_2) \dots a_{d,i}(x_d) \tag{2}$$

is satisfied for all $(x_1, \dots, x_d) \in [n_1] \times \dots \times [n_d]$.

We say that an expression such as (2) *is a* tensor rank decomposition *of T, say that the integer k is the* length *of the decomposition, and say that the decomposition has* minimal length *if its length is equal to the tensor rank of T.*

For every $x \in [n_1] \times \dots \times [n_d]$ and every $j \in [d]$, we write $\overline{x_j}$ for $(x_{j'})_{j' \in [d] \setminus \{j\}}$.

► **Definition 2.** *Let $d \geq 2$ be an integer, and let $T : [n_1] \times \dots \times [n_d] \to \mathbb{F}$ be an order-d tensor. We say that the* slice rank *of T, denoted by $\operatorname{sr} T$, is the smallest nonnegative integer k such that there exist nonnegative integers $r_1, \dots, r_d$ with $r_1 + \dots + r_d = k$ satisfying one of the following two equivalent properties.*

1. *There exist matrices* $M_j : [n_j] \times (\prod_{j' \neq j} [n_{j'}]) \to \mathbb{F}$ *with rank at most* r_j *for each* $j \in [d]$ *such that*

$$T(x_1, \ldots, x_d) = \sum_{j=1}^{d} M_j(x_j, \overline{x_j})$$

is satisfied for all $(x_1, \ldots, x_d) \in [n_1] \times \cdots \times [n_d]$.

2. *There exist functions* $a_{j,i} : [n_j] \to \mathbb{F}$ *and* $b_{j,i} : \prod_{j' \neq j} [n_{j'}] \to \mathbb{F}$ *for each* $j \in [d]$ *and each* $i \in [r_j]$ *such that*

$$T(x_1, \ldots, x_d) = \sum_{j=1}^{d} \sum_{i=1}^{r_j} a_{j,i}(x_j) b_{j,i}(\overline{x_j}) \tag{3}$$

is satisfied for all $(x_1, \ldots, x_d) \in [n_1] \times \cdots \times [n_d]$.

We say that an expression such as (3) *is a* slice rank decomposition *of* T*, say that the integer* $r_1 + \cdots + r_d$ *is the* length *of the decomposition, and say that the decomposition has* minimal length *if its length is equal to the slice rank of* T*.*

Before going further, let us recall some of the history of the slice rank and of the ways in which it has been studied.

The slice rank was originally introduced by Tao [19] in 2016 as a reformulation of a central idea from the breakthrough of Croot, Lev and Pach [6], which led to the solution to the cap-set problem by Ellenberg and Gijswijt [7]. Since then, the slice rank has been used successfully several times as a tool to solve combinatorial problems. For instance, it was also shown by Naslund and Sawin [16] that subsets of the cube $\{0,1\}^n$ containing no 3-sunflower are exponentially sparse, and further properties on the slice rank proved by Sawin and Tao [18] were later used by Sauermann [17] to obtain properties guaranteeing the existence of solutions with pairwise distinct variables to systems of equations in subsets of vector spaces over finite prime fields which are not exponentially sparse. In another direction, a variant of the slice rank, the partition rank, was defined by Naslund in [14], originally to obtain polynomial upper bounds on problems on k-right corners, and was more recently used again by Naslund [15], this time to obtain exponential lower bounds on the chromatic number of $\mathbb{R}^n$ with multiple forbidden distances.

Although all these works have primarily used the slice rank and the partition rank as tools, interest in studying them and their basic properties for their own sake has recently been building up. In the post of Sawin and Tao [18], a characterisation of the slice rank in terms of coverings in the case of tensors supported inside an antichain had already been discussed, and this characterisation later played an important role in the work of Sauermann [17] that we previously mentioned. Later, it was proved by Gowers [8] that the slice rank of a direct sum of two tensors is the sum of their slice ranks.

As another example in this direction, the facts that a high-rank matrix must contain a high-rank submatrix of not too large size and that a matrix with high rank after every modification of the diagonal must have a high rank submatrix for which the sets of rows and columns are disjoint were extended by the author in [11] to a class of notions of rank containing the slice and partition ranks. Still on the topic of subtensors, it was shown by Briët and Castro-Silva [5] that for another wide class of notions of rank, a random subtensor of a high-rank tensor must have high rank for some natural way of choosing the restriction at random, together with some analogous results on random restrictions of polynomials.

The three results from the previous paragraph found applications. The first, conditionally on better bounds, was discussed in [5] as a way to provide an alternate proof of some of the main results of that paper. The second was used by Gowers and the author [9] as a key stepping stone to extend to distributions inside $\{0,1\}^n$ the result of Green and Tao [10] on the equidistribution of high-rank polynomials over finite prime fields. The third was applied (primarily in the case of polynomials, although the case of tensors was involved as well) by Briët, Buhrman, Castro-Silva, and Neumann [4] to prove limitations on the decoding of corrupted classical and quantum error-correcting codes with $\mathrm{NC}^0[\oplus]$ circuits.

The slice and partition ranks have also been studied using more algebraic methods. For instance, a "universal" role that the partition rank plays among notions of rank has been explored by Bik, Draisma and Eggermont [3]. Another universality result was proved by Kazhdan and Ziegler [12] on the strength of polynomials (an analogue of the partition rank of tensors), and later led to an extension by Bik, Danelon, Draisma, and Eggermont [2] to the more abstract setting of arbitrary polynomial functors.

Let us finally mention that in the case of the tensor rank, much effort has gone into understanding the notion of identifiability of a tensor, which is much stronger than the mere boundedness of its set of minimal-length decompositions that we focus on (primarily in the case of the slice rank) in the present paper: if k is a nonnegative integer, then a tensor T with tensor rank k is said to be identifiable if there is only one way of writing it as a sum

$$T_1 + \cdots + T_k$$

where $T_1, \ldots, T_k$ have tensor rank 1, up to permutations of the k tensors $T_1, \ldots, T_k$. This area of research, perhaps started by Kruskal [13] in 1977, is still very active. A recent paper [1] of Ballico, Bernardi and Santarsiero contains, among other interesting things, a wealth of references on identifiability.

We now define some more notations which we shall use throughout the paper. Let $d \geq 2$ be an integer. If x is an element of $[n_1] \times \cdots \times [n_d]$ and J is a subset of $[d]$, then we shall write $x(J)$ for the restriction of x to its coordinates in J, that is, for the element of $\prod_{j \in J} [n_j]$ defined by $x(J)_j = x_j$ for every $j \in J$. If $a, a' : [n_j] \to \mathbb{F}$ are functions for some $j \in [d]$, then we write $a.a'$ for the element of $\mathbb{F}$ defined by

$$\sum_{x_j \in [n_j]} a(x_j) a'(x_j).$$

More generally, if $J' \subset J$ are non-empty subsets of $[d]$, and $U : \prod_{j \in J} [n_j] \to \mathbb{F}$, $U' : \prod_{j \in J'} [n_j] \to \mathbb{F}$ are functions, then $U'.U : \prod_{j \in J \setminus J'} [n_j] \to \mathbb{F}$ will denote the function (or rather element of $\mathbb{F}$ if $J' = J$) defined by

$$(U'.U)(x(J \setminus J')) = \sum_{x(J') \in \prod_{j \in J'} [n_j]} U'(x(J')) U(x(J))$$

for all $x(J \setminus J')$. If s is a positive integer, and $j_1, \ldots, j_s$ are pairwise distinct elements of $[d]$, then we write $\overline{x_{j_1}, \ldots, x_{j_s}}$ for $x([d] \setminus \{j_1, \ldots, j_s\})$.

We will repeatedly use the following observations. If a decomposition (3) has minimal length, then for every $j \in [d]$ its functions $a_{j,i}$ with $i \in [r_j]$ are linearly independent. Even if the decomposition does not have minimal length, each of the d parts of the decomposition and hence the whole decomposition can always be written in a way which ensures that these functions are linearly independent, and we will always assume this to be the case whenever we write any slice rank decomposition of any tensor.

It follows from Gaussian elimination that if for some $j \in [d]$ and some integer $r \geq 1$ some functions $a_1, \dots, a_r : [n_j] \to \mathbb{F}$ are linearly independent, then there exist functions $a_1^*, \dots, a_r^* : [n_j] \to \mathbb{F}$ satisfying $a_i^*.a_{i'} = 1_{i=i'}$ for any $i, i' \in [r]$. We will refer to the family $(a_1^*, \dots, a_r^*)$ as a family of *dual functions* to the family of functions $(a_1, \dots, a_r)$. More generally, if for some non-empty subset J of $[d]$ and some integers $r_j \geq 1$ for each $j \in J$ the functions $a_{j,1}, \dots, a_{j,r_j}$ are linearly independent for every $j \in J$, then the tensor products

$$\otimes_{j \in J} a_{j,i_j} : \prod_{j \in J} [n_j] \to \mathbb{F}$$

with $(i_j)_{j \in J} \in \prod_{j \in J} [r_j]$ are linearly independent, and furthermore

$$(\otimes_{j \in J} a_{j,i_j}^*).(\otimes_{j \in J} a_{j,i'_j}) = \prod_{j \in J} 1_{i_j = i'_j}$$

is equal to 1 if $i_j = i'_j$ for every $j \in J$, and to 0 otherwise.

The remainder of the paper is organised as follows. In Section 2 we describe several rather systematic ways in which we may construct several slice rank decompositions of the same tensor. Against these constructions we then formulate in Section 3 our result on the structure of the set of slice rank decompositions of a tensor, Theorem 12, together with two other results on slice rank decompositions, Proposition 9 and Theorem 10, from which Theorem 12 can be deduced and which also appear to be of intrinsic interest. This deduction is then performed in Section 4. In Section 5, which is independent from the other sections, we give a simple proof of a statement showing that in the case of the tensor rank, the result on decompositions of matrices extends in a rather optimistic way. The proofs of Proposition 9 and Theorem 10 are carried out in the full version of this manuscript. Finally, in Section 6 we conclude by mentioning some questions which remain open.

2 Constructions of different minimal-length slice rank decompositions

We next go over five ways in which several minimal-length slice rank decompositions can arise for the same tensor.

▶ **Construction 3.** Since it is possible to obtain new decompositions of matrices using changes of bases, it is possible to obtain new decompositions of T from an existing decomposition by rewriting the decompositions of any of the individual matrices $M_1, \dots, M_d$ while leaving the d-tuple $(M_1, \dots, M_d)$ unchanged.

However, this d-tuple is not unique in general, and there are other ways of obtaining new decompositions, which brings us to the remaining four constructions.

▶ **Construction 4.** To see that the matrices $M_1, \dots, M_d$ are not individually determined, consider a tensor T for which the slice rank and tensor rank are equal to some common positive integer k. Let

$$T(x_1, \dots, x_d) = \sum_{i=1}^{k} a_{1,i}(x_1) \dots a_{d,i}(x_d)$$

be a tensor rank decomposition of T with length k. Then for every partition $\{I_1, \dots, I_d\}$ of $[k]$ into d sets $I_1, \dots, I_d$ that are possibly empty, the decomposition

$$T(x_1, \dots, x_d) = \sum_{j=1}^{d} \sum_{i \in I_j} a_{j,i}(x_j) (\prod_{j' \neq j} a_{j',i}(x_{j'}))$$

is a slice-rank decomposition of T with length k.

Tensors with equal slice rank and tensor rank exist, since for instance the "identity" tensor $I_{d,k} : [k]^d \to \mathbb{F}$ defined by

$$I_{d,k}(x_1, \dots, x_d) = \sum_{i=1}^{k} 1_{x_1=i} \dots 1_{x_d=i}$$

was shown by Tao ([19], Lemma 1) to have slice rank k. Since the tensor rank of this tensor is at most k by definition, and since the slice rank is always at most the tensor rank (as can be seen from their definitions), it follows that both ranks are equal for this tensor.

For instance, in the case $d = 3$, we obtain a slice rank decomposition

$$I_{3,k}(x, y, z) = \sum_{i \in I_1} 1_{x=i} 1_{y=z=i} + \sum_{i \in I_2} 1_{y=i} 1_{x=z=i} + \sum_{i \in I_3} 1_{z=i} 1_{x=y=i}$$

for every tripartition $\{I_1, I_2, I_3\}$ of $[k]$ (again, with I_1, I_2, I_3 allowed to be empty).

As we can see from Construction 4 not even the sizes of the d parts of the decomposition are determined in general. One comment that can nonetheless be made about Construction 4 is that although the matrices $M_1, \dots, M_d$ are not always the same, the set of all summands is always the same set

$$\{1_{x_1=i} \dots 1_{x_d=i} : i \in [k]\}.$$

However, there is another class of examples where this usually does not hold, even up to changes of bases in the matrices $M_1, \dots, M_d$, and which even encompasses typical tensors.

► **Construction 5.** Let k be a positive integer. A tensor $T : [k]^d \to \mathbb{F}$ can in particular be written in d different ways as a sum of its $(d-1)$-variable slices of each type, i.e. be decomposed as

$$T(x_1, \dots, x_d) = \sum_{i=1}^{k} 1_{x_j=i} T(x_1, \dots, x_{j-1}, i, x_{j+1}, \dots, x_d)$$

for every $j \in [d]$. If T has slice rank k, then these d decompositions are each minimal-length decompositions. If the field $\mathbb{F}$ is finite, then the number of tensors $[k]^d \to \mathbb{F}$ with slice rank at most $k - 1$ is at most

$$k^d |\mathbb{F}|^{(k-1)(k^{d-1}+k)} = k^d |\mathbb{F}|^{k^d - k^{d-1} + k^2 - k}.$$

Provided that $d \geq 3$, we can crudely bound this number above by

$$(k^d / |\mathbb{F}|^k) |\mathbb{F}|^{k^d}.$$

The ratio is less than 1 for k large enough, and tends to 0 as k tends to infinity, which shows that random tensors have slice rank k. Unlike in the $I_{d,k}$ example, the sets of k summands involved in each of the d decompositions are in particular usually very different.

► **Construction 6.** Let k be a positive integer. Provided that a tensor T with slice rank k does have a decomposition with length k that does not have all its terms of the same of the three types, there is a systematic way of changing some of the components of the decomposition and obtain another decomposition with length k. In the $d = 3$ case, for instance, if

$$T(x, y, z) = a(x) b(y, z) + c(y) d(x, z),$$

then for any function $e : [n_3] \to \mathbb{F}$, the functions $b' = b + c \otimes e$ and $d' = d - a \otimes e$ satisfy

$$T(x, y, z) = a(x) b'(y, z) + c(y) d'(x, z).$$

Because the function e is arbitrary, and in particular there are unboundedly many such functions as n tends to infinity, it is never true, for any integers $d \geq 3$, $k \geq 2$ and any field $\mathbb{F}$, that every order-d tensor with slice rank k has a number of minimal-length decompositions which is bounded above by some function of $d, k, \mathbb{F}$. A boundedness statement of this kind may only hold up to this type of transformation.

If T is an order-3 tensor with a more general slice rank decomposition

$$T(x,y,z) = \sum_{i=1}^{r} a_i(x)b_i(y,z) + \sum_{j=1}^{s} c_j(x)d_j(y,z) + \sum_{k=1}^{t} e_k(z)f_k(x,y), \tag{4}$$

then whenever $(i,j) \in [r] \times [s]$ and $e : [n_3] \to \mathbb{F}$ is a function, replacing b_i and d_j by respectively

$$b_i + c_j \otimes e \text{ and } d_j - a_i \otimes e$$

leads to a new decomposition of T.

More generally, if

$$T(x_1,\dots,x_d) = \sum_{j=1}^{d} \sum_{i=1}^{r_j} a_{j,i}(x_j) b_{j,i}(\overline{x_j})$$

is a slice rank decomposition of T, $c : [n_3] \times \dots \times [n_d]$ is a function, and $i_1 \in [r_1]$, $i_2 \in [r_2]$ are two indices, then replacing b_{1,i_1} and b_{2,i_2} by respectively

$$b_{1,i_1} + a_{2,i_2} \otimes c \text{ and } b_{2,i_2} - a_{1,i_1} \otimes c$$

leads to a new decomposition of T. The roles of either of the first two coordinates may of course be exchanged with those of any of the other coordinates.

▶ **Construction 7.** We can generalise Construction 6 further. Let us begin by describing a variant of Construction 6, in the simplest setting. If we have the decomposition

$$T(x,y,z) = a(x)b(y,z) + c(y)d(x,z) + e(z)f(x,y),$$

$\lambda_1, \lambda_2, \lambda_3$ are elements of $\mathbb{F}$ adding up to 0 then the functions

$$\begin{aligned} b' &= b + \lambda_1 c \otimes e \\ d' &= d + \lambda_2 a \otimes e \\ f' &= f + \lambda_3 a \otimes c \end{aligned}$$

lead to the decomposition

$$T(x,y,z) = a(x)b'(y,z) + c(y)d'(x,z) + e(z)f'(x,y).$$

We note that this transformation reduces to successively applying three of the transformations from Construction 6, since we can always find $\mu_{12,1}$, $\mu_{12,2}$, $\mu_{13,1}$, $\mu_{13,3}$, $\mu_{23,2}$, $\mu_{23,3} \in \mathbb{F}$ satisfying

$$\begin{aligned} \mu_{12,1} + \mu_{12,2} &= 0 \qquad & \lambda_1 &= \mu_{12,1} + \mu_{13,1} \\ \mu_{13,1} + \mu_{13,3} &= 0 \qquad & \lambda_2 &= \mu_{12,2} + \mu_{23,2} \\ \mu_{23,2} + \mu_{23,3} &= 0 \qquad & \lambda_3 &= \mu_{13,3} + \mu_{23,3}. \end{aligned}$$

We may for instance choose $\mu_{13,1}$ and $\mu_{13,3}$ to be zero, and $\mu_{12,1}, \mu_{12,2}, \mu_{23,2}, \mu_{23,3}$ are then uniquely determined.

If (4) is a more general decomposition of an order-3 tensor T, then whenever $(i,j,k) \in [r] \times [s] \times [t]$ and $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{F}$ satisfy

$$\lambda_1 + \lambda_2 + \lambda_3 = 0,$$

replacing b_i, d_j, f_k by respectively

$$b_i + \lambda_1 c_j \otimes e_k$$
$$d_j + \lambda_2 a_i \otimes e_k$$
$$f_k + \lambda_3 a_i \otimes c_j$$

leads to a new decomposition of T.

In the case of general $d \geq 3$, if (3) is a decomposition of T, then we can choose some index i_j for each $j \in [d]$, and choose some $\lambda_1, \ldots, \lambda_d \in \mathbb{F}$ satisfying

$$\lambda_1 + \cdots + \lambda_d = 0.$$

Replacing the function b_{j,i_j} for each $j \in [d]$ by the function b'_{j,i_j} defined by

$$b'_{j,i_j}(\overline{x_j}) = b_{j,i_j}(\overline{x_j}) + \lambda_j a_{1,i_1}(x_1) \ldots a_{j-1,i_{j-1}}(x_{j-1}) a_{j+1,i_{j+1}}(x_{j+1}) \ldots a_{d,i_d}(x_d)$$

then provides a new decomposition of T.

However, we can extend this class of transformations yet further: we can choose successively a subset J of $[d]$ with size at least 2, some indices $i_j \in [r_j]$ for each $j \in J$, some functions $c_j : \prod_{j \in [d] \setminus J} [n_j] \to \mathbb{F}$ for each $j \in J$ (instead taken to be a elements of $\mathbb{F}$ if $J = [d]$) satisfying

$$\sum_{j \in J} c_j = 0,$$

and replace for each $j \in J$ the function b_{j,i_j} by the function b'_{j,i_j} defined by

$$b'_{j,i_j}(\overline{x_j}) = b_{j,i_j}(\overline{x_j}) + (\prod_{j' \in J \setminus \{j\}} a_{j',i_{j'}}(x_{j'})) c_j(x([d] \setminus J)).$$

This yet again provides a new decomposition of T.

Just as in the case of order-3 tensors, for general $d \geq 3$ and $2 \leq |J| \leq d$ doing this transformation ultimately reduces to successively applying transformations from Construction 6 several times. For instance, writing $J = \{j_1, \ldots, j_s\}$, there is a unique way of choosing them while only using the $s-1$ pairs of indices $\{j_1, j_2\}, \ldots, \{j_{s-1}, j_s\}$. Nonetheless, we discuss Construction 7 separately, as in our proofs we will think of the corresponding transformation as a single transformation rather than as a succession of several transformations from Construction 6.

Any of the five constructions that we have just described may be combined together: as shown by Gowers [8], the slice rank of a block diagonal tensor is equal to the sum of the slice ranks of the diagonal blocks, so putting together minimal-length decompositions of the five examples each respectively corresponding to a tensor $T_1, \ldots, T_5$ leads to a minimal-length decomposition of the diagonal sum

$$T_1 \oplus \cdots \oplus T_5,$$

and this is still the case after modifying any of the decompositions of $T_1, \ldots, T_5$ using Constructions 3 to 7 respectively.

3 Statements of main results

The previous section was devoted to constructions illustrating how a variety of different slice rank decompositions may arise for the same tensor. In the present section we describe our main results, which will be in the converse direction and show that to some extent, the set of minimal-length slice rank decompositions of the same tensor cannot be too rich. We will do so separately for the $(d-1)$-variable functions and for the one-variable functions of minimal-length decompositions, and will go through the following three steps, corresponding respectively to the proofs of the three main results of this paper, Proposition 9, Theorem 10, and Theorem 12.

1. We will begin by showing that if two decompositions have the same one-variable functions, then any differences between the two decompositions arise from Construction 6 and Construction 7.
2. After this, we will show a statement which in particular implies that no more than d decompositions can have their sets of one-variable functions in the jth coordinate be jointly linearly independent for every $j \in [d]$ simultaneously.
3. This result will allow us to deduce that, in the finite field case, there are only a bounded number of possibilities for the one-variable functions, with a bound depending only on the order of the tensor, on its slice rank, and on the size of the finite field.

In summary, we can partition the decompositions according to their one-variable functions in a bounded number of sets, and inside each of these sets we can completely describe the possibilities for the $(d-1)$-variable functions once we know what these are for one of the decompositions from the set.

Let us now go through these steps more formally. Although the following definition is a bit long, all that it describes is differences between the original and the new $(d-1)$-variable functions before and after successively applying Construction 6 and Construction 7 as much as can be done. If J is a subset of $[d]$, then for simplicity of notation we will denote a sequence $(i_j)_{j \in J}$ by i_J.

▶ **Definition 8.** *Let $d \geq 3$ be an integer, and let $r_1, \dots, r_d$ be nonnegative integers. We say that a decomposition*

$$\sum_{j=1}^{d} \sum_{i=1}^{r_j} a_{j,i}(x_j) b_{j,i}(\overline{x_j})$$

is in zero form *if there exist functions*

$$c_{J,j,i,i_{J \setminus \{j\}}} : \prod_{j'' \in [d] \setminus J} [n_{j''}] \to \mathbb{F}$$

(instead taken to be elements of $\mathbb{F}$ if $J = [d]$) for every subset $J \subset [d]$ with $2 \leq |J| \leq d$, every $j \in [d]$, every $i \in [r_j]$ and every $i_{J \setminus \{j\}}$ satisfying the following two properties.

1. *For every $j \in [d]$ and every $i \in [r_j]$ we can write*

$$b_{j,i}(\overline{x_j}) = \sum_{J \subset [d] : j \in J} \sum_{i_{J \setminus \{j\}}} \left(\prod_{j' \in J \setminus \{j\}} a_{j', i_{j'}}(x_{j'}) \right) c_{J,j,i,i_{J \setminus \{j\}}}(x([d] \setminus J)).$$

2. *For every $J \subset [d]$ with $2 \leq |J| \leq d$ and every $(i_{j'})_{j' \in J}$, we obtain a sum of 0 whenever we take the sum over all functions $c_{J,j,i,i_{J \setminus \{j\}}}$ where the sequence $i_{J \setminus \{j\}}$ completed to a sequence indexed by J by introducing the additional term $i_j = i$ is equal to $(i_{j'})_{j' \in J}$.*

Given this definition, the first component of our results can be easily formulated. It can be checked that any decomposition that is in zero form adds up to the zero tensor, and our first result is a converse of that. We assume as usual that the functions $a_{j,i}$ of decompositions are linearly independent for each $j \in [d]$.

▶ **Proposition 9.** *Let $d \geq 3$ be an integer, and let $r_1, \ldots, r_d$ be nonnegative integers. If a decomposition*

$$\sum_{j=1}^{d} \sum_{i=1}^{r_j} a_{j,i}(x_j) b_{j,i}(\overline{x_j})$$

is equal to the zero tensor, then it is in zero form.

The next theorem corresponds to the second step of the argument that we described at the beginning of this section. The key structure that we will use is that of sunflowers in a linear algebra sense: if $A^0, A^1, \ldots, A^h$ are linear subspaces which are all jointly in direct sum for some positive integer h, then we say that the linear subspaces

$$A^0 \oplus A^1, \ldots, A^0 \oplus A^h$$

constitute a *sunflower* with *center* A^0 and *petals* $A^1, \ldots, A^h$.

What we will show is that if we can find $d+1$ decompositions θ of the same tensor (not necessarily with minimal length) where for each $j \in [d]$ the linear subspaces spanned by the one-variable functions $a_{j,i}$ coming from the respective decompositions constitute a sunflower in the sense that we have just described, then the tensor has a slice rank decomposition where for every $j \in [d]$ the one-variable functions $a_{j,i}$ all belong to the center A^0_j of the sunflower, and hence in particular the tensor has slice rank at most

$$\dim A^0_1 + \cdots + \dim A^0_d.$$

We state this result for any $d \geq 2$: the $d = 2$ case is relevant since we use it in the proof of the $d = 3$ case.

▶ **Theorem 10.** *Let $d \geq 2$ be an integer, let $\mathbb{F}$ be a field, and let T be an order-d tensor over $\mathbb{F}$. Let $r^0_1, \ldots, r^0_d$ be nonnegative integers, and let $a^0_{j,i} : [n_j] \to \mathbb{F}$ be one-variable functions for every $j \in [d]$ and every $i \in [r^0_j]$. Let $h > d$ be an integer, and assume that the following two conditions are satisfied.*

1. *For every $\theta \in [h]$ there exist nonnegative integers $r^\theta_1, \ldots, r^\theta_d$ and a decomposition of T of the type*

$$\sum_{j=1}^{d} \left(\sum_{i=1}^{r^0_j} a^0_{j,i}(x_j) b^{0,\theta}_{j,i}(\overline{x_j}) + \sum_{i=1}^{r^\theta_j} a^\theta_{j,i}(x_j) b^\theta_{j,i}(\overline{x_j}) \right)$$

   for some one-variable functions $a^\theta_{j,i} : [n_j] \to \mathbb{F}$ and for some $(d-1)$-variable functions $b^{0,\theta}_{j,i}, b^\theta_{j,i} : \prod_{j' \neq j} [n_{j'}] \to \mathbb{F}$.
2. *For each $j \in [d]$ the one-variable functions*

$$a^0_{j,1}, \ldots, a^0_{j,r^0_j}, a^1_{j,1}, \ldots, a^1_{j,r^1_j}, \ldots, a^h_{j,1}, \ldots, a^h_{j,r^h_j} : [n_j] \to \mathbb{F}$$

   are linearly independent.

Then there exist $(d-1)$-variable functions $b^0_{j,i}:\prod_{j'\neq j}[n_{j'}]\to\mathbb{F}$ such that

$$\sum_{j=1}^{d}\sum_{i=1}^{r^0_j} a^0_{j,i}(x_j)b^0_{j,i}(\overline{x_j})$$

is a decomposition of T and in particular

$$\operatorname{sr} T \leq r^0_1+\cdots+r^0_d.$$

We stress that in the assumptions of Theorem 10 the decompositions are not required to have minimal length. The inequality $h>d$ is optimal, as can be seen from either Construction 4 or Construction 5.

We note that there is a large class of examples where the subspaces $A^\theta_1,\dots,A^\theta_d$ are respectively the same for every minimal-length decomposition of T. Informally, this is the class of order-d tensors which admit a minimal-length slice rank decomposition where the $(d-1)$-variable functions of each kind are sufficiently separated with respect to the slice rank of order-$(d-1)$ tensors.

▶ **Proposition 11.** *Let $d\geq 3$, $k\geq 1$ be integers, and let T be an order-d tensor with slice rank k. Assume that there exist nonnegative integers $r_1,\dots,r_d$ satisfying*

$$r_1+\cdots+r_d=k,$$

and a decomposition

$$\sum_{j=1}^{d}\sum_{i=1}^{r_j} a_{j,i}(x_j)b_{j,i}(\overline{x_j})$$

of T such that for every $j\in[d]$ the order-$(d-1)$ tensors $b_{j,1},\dots,b_{j,r_j}:\prod_{j'\neq j}[n_j]\to\mathbb{F}$ satisfy

$$\operatorname{sr}\left(\sum_{i=1}^{r_j}\lambda_i b_{j,i}\right)\geq 2k \tag{5}$$

for every $(\lambda_1,\dots,\lambda_{r_j})\in\mathbb{F}^{r_j}\setminus\{0\}$ and for every $j\in[d]$. Then the linear subspaces

$$A_j=\langle a_{j,i}: i\in[r_j]\rangle$$

with $j\in[d]$ are such that whenever

$$\sum_{j=1}^{d}\sum_{i=1}^{r'_j} a'_{j,i}(x_j)b'_{j,i}(\overline{x_j})$$

is a slice rank decomposition of T satisfying $r'_1+\cdots+r'_d=k$, we have $r'_j=r_j$ and

$$A_j=\langle a'_{j,i}: i\in[r'_j]\rangle$$

for every $j\in[d]$.

Proof. For every $j\in[d]$ we write A_j and A'_j for the linear subspaces

$$\langle a_{j,1},\dots,a_{j,r_j}\rangle \text{ and } \langle a'_{j,1},\dots,a'_{j,r_j}\rangle$$

respectively. If for some $j_0 \in [d]$ the subspace A'_{j_0} does not contain A_{j_0}, then there exists a function $u : [n_{j_0}] \to \mathbb{F}$ such that $u.a = 0$ for every $a \in A'_{j_0}$ but $u.a \neq 0$ for some $a \in A_{j_0}$. Applying u to both sides of the equality

$$\sum_{j=1}^{d} \sum_{i=1}^{r_j} a_{j,i}(x_j) b_{j,i}(\overline{x_j}) = \sum_{j=1}^{d} \sum_{i=1}^{r'_j} a'_{j,i}(x_j) b'_{j,i}(\overline{x_j})$$

we obtain

$$\sum_{i=1}^{r_{j_0}} (u.a_{j_0,i}) b_{j_0,i}(\overline{x_{j_0}}) + \sum_{j \neq j_0} \sum_{i=1}^{r_j} a_{j,i}(x_j)(u.b_{j,i})(\overline{x_{j_0}, x_j}) = \sum_{j \neq j_0} \sum_{i=1}^{r'_j} a'_{j,i}(x_j)(u.b'_{j,i})(\overline{x_{j_0}, x_j}).$$

Our assumption (5) shows that

$$(u.a_{j_0,1}, \dots, u.a_{j_0,r_{j_0}}) \neq 0$$

and hence that the first sum of the left-hand side has slice rank at least $2k$. All inner summands of all other sums each have slice rank at most 1, and there are at most $2k - 1$ such summands, which is a contradiction. Therefore, A'_j contains A_j for every $j \in [d]$. Since we have

$$r_1 + \dots + r_d = k = r'_1 + \dots + r'_d,$$

we conclude $A'_j = A_j$ for every $j \in [d]$. ◀

From Proposition 9 and Theorem 10 we then deduce our theorem on the structure of minimal-length slice rank decompositions of tensors, which we now state.

▶ **Theorem 12.** *Let $d \geq 3$, $k \geq 1$ be integers, let $\mathbb{F}$ be a field, and let T be an order-d tensor over $\mathbb{F}$ and with slice rank equal to k. Then we have the following.*

1. *If the field $\mathbb{F}$ is finite, then there exists a set $\mathcal{A}$ of d-tuples $(W_1, \dots, W_d)$ of linear subspaces of $\mathbb{F}^{n_1}, \dots, \mathbb{F}^{n_d}$ respectively, with size $|\mathcal{A}| \leq d^k |\mathbb{F}|^{dk^2}$, such that if $r_1, \dots, r_d$ are nonnegative integers satisfying*

$$r_1 + \dots + r_d = k \tag{6}$$

and $a_{j,i} : [n_j] \to \mathbb{F}$ with $j \in [d]$ and $i \in [r_j]$ are functions satisfying

$$T(x_1, \dots, x_d) = \sum_{j=1}^{d} \sum_{i=1}^{r_j} a_{j,i}(x_j) b_{j,i}(\overline{x_j}) \tag{7}$$

for some $(d-1)$-variable functions $b_{j,i} : \prod_{j' \neq j} [n_{j'}] \to \mathbb{F}$, then

$$(\langle a_{1,1}, \dots, a_{1,r_1} \rangle, \dots, \langle a_{d,1}, \dots, a_{d,r_d} \rangle) \in \mathcal{A}.$$

In particular there are at most $d^k |\mathbb{F}|^{(d+1)k^2}$ possibilities in total for

$$((r_1, \dots, r_d), (a_{1,1}, \dots, a_{1,r_1}), \dots, (a_{d,1}, \dots, a_{d,r_d}))$$

satisfying (6) and (7).

2. *For any field $\mathbb{F}$, if two decompositions*

$$T(x_1,\dots,x_d) = \sum_{j=1}^{d}\sum_{i=1}^{r_j} a_{j,i}(x_j)b^1_{j,i}(\overline{x_j}) = \sum_{j=1}^{d}\sum_{i=1}^{r_j} a_{j,i}(x_j)b^2_{j,i}(\overline{x_j}) \tag{8}$$

of T have the same one-variable functions $a_{j,i}$, then the decomposition

$$\sum_{j=1}^{d}\sum_{i=1}^{r_j} a_{j,i}(x_j)(b^2_{j,i} - b^1_{j,i})(\overline{x_j}) \tag{9}$$

is in zero form.

The decompositions need not be assumed to have minimal length in the second item of Theorem 12, even if like the first item of Theorem 12 it was conceived primarily with this case in mind. The following example shows that the exponents in the two bounds from the first item of Theorem 12 can only be away from the optimal bounds by factors which are linear in d, uniformly in k and $\mathbb{F}$.

▶ **Example 13.** *Let $d \geq 4$ be an integer, let r be a positive integer, let $M : [n_1] \times [n_2] \to \mathbb{F}$ be a rank-$2r$ matrix, let $c : [n_3] \times \dots \times [n_d] \to \mathbb{F}$ be an order-$(d-2)$ tensor with slice rank at least $2r$, and let T be the order-d tensor defined by*

$$T(x_1,\dots,x_d) = M(x_1,x_2)c(x_3,\dots,x_d).$$

Then the tensor T has slice rank $2r$, and there are at least $\omega|\mathbb{F}|^{r^2}$ possibilities for the linear subspace $\langle a^1_{1,1},\dots,a^1_{1,r}\rangle$ of any length-$2r$ slice rank decomposition of T of the type

$$\sum_{i=1}^{r} a_{1,i}(x_1)b_{1,i}(\overline{x_1}) + \sum_{i=r+1}^{2r} a_{2,i}(x_2)b_{2,i}(\overline{x_2}).$$

Proof. Let A be the linear subspace spanned by the functions $f_1,\dots,f_{2r}$ in any rank decomposition

$$M(x_1,x_2) = \sum_{i=1}^{2r} f_i(x_1)g_i(x_2)$$

of M. For any such decomposition of M, the decompositions

$$\sum_{i=1}^{r} f_i(x_1)(g_ic)(x_2,x_3,\dots,x_d) + \sum_{i=r+1}^{2r} g_i(x_2)(f_ic)(x_1,x_3,\dots,x_d)$$

are decompositions of T with length $2r$, and the linear subspace $\langle f_1,\dots,f_r\rangle$ can be any dimension-r linear subspace of A. In turn, there are at least

$$\omega|\mathbb{F}|^{(2r)^2}/|\mathbb{F}|^{2r^2}|\mathbb{F}|^{r^2}$$

possibilities for this linear subspace, which provides the desired bound. It hence suffices to check that these decompositions indeed have minimal length, in other words that T has slice rank no less than $2r$. Assume for contradiction that

$$\sum_{i=1}^{r_1} a_{1,i}(x_1)b_{1,i}(\overline{x_1}) + \sum_{i=1}^{r_2} a_{2,i}(x_2)b_{2,i}(\overline{x_2}) + \sum_{j=3}^{d}\sum_{i=1}^{r_j} a_{j,i}(x_j)b_{j,i}(\overline{x_j})$$

is a slice rank decomposition of T satisfying

$$r_1 + r_2 + r_3 + \cdots + r_d < 2r.$$

The linear subspace

$$U = \{a_1^* \in \mathbb{F}^{n_1} : a_1^*.a_{1,1} = 0, \ldots, a_1^*.a_{1,r_1} = 0\}$$

has dimension $n - r_1$, so since M has rank $2r$, the linear subspace

$$\{a_1^*.M : a_1^* \in U\}$$

has dimension at least $2r - r_1$, so is not contained inside the linear subspace $\langle a_{2,1}, \ldots, a_{2,r_2}\rangle$. Therefore, we can find $a_1^* \in U$ and an element a_2^* of $\mathbb{F}^{n_2}$ such that

$$a_2^*.a_{2,1} = 0, \ldots, a_2^*.a_{2,r_2} = 0$$

but which also satisfies

$$(a_1^* \otimes a_2^*).M = a_2^*.(a_1^*.M) \neq 0.$$

Applying $a_1^* \otimes a_2^*$ to the original decomposition of T then provides a slice rank decomposition

$$\lambda c(x_3, \ldots, x_d) = \sum_{j=3}^{d} \sum_{i=1}^{r_j} a_{j,i}(x_j)((a_1^* \otimes a_2^*).b_{j,i})(\overline{x_1, x_2, x_j})$$

of λc for some $\lambda \neq 0$, which contradicts that $\operatorname{sr} c \geq 2r$. ◀

4 Deduction of the structure theorem on slice rank decompositions

In this section we deduce Theorem 12 from Proposition 9 and Theorem 10.

Proof of Theorem 12. Let us begin by proving the second item. Because the two decompositions (8) are decompositions of the same tensor, their difference (9) is equal to the zero tensor, so it follows from Proposition 9 that the decomposition (9) is in zero form.

There remains to prove the first item. Let k be the slice rank of T. If $k = 0$ then we are done, so let us assume $k \geq 1$. We consider an arbitrary set of minimal-length decompositions

$$\sum_{j=1}^{d} \sum_{i=1}^{r_j^\theta} a_{j,i}^\theta(x_j) b_{j,i}^\theta(\overline{x_j})$$

of T indexed by $\theta \in [h]$ for some positive integer h. For every $\theta \in [h]$ we write $A_1^\theta, \ldots, A_d^\theta$ for the linear subspaces spanned by the functions $a_{1,i}^\theta, \ldots, a_{d,i}^\theta$ respectively. We assume that whenever θ, θ' are distinct elements of $[h]$, the equality $A_j^\theta = A_j^{\theta'}$ fails for at least one $j \in [d]$.

Let M_1 be a maximal subset of $[h]$ such that for each $j \in [d]$ the linear subspaces A_j^θ with $\theta \in M_1$ are all in direct sum. By Theorem 10 the set M_1 must have size at most d, as otherwise we would have $T = 0$. We can hence write

$$[h] = \bigcup_{1 \leq j_1 \leq d} \bigcup_{w_1} \Theta_{(j_1, w_1)}$$

where the union over w_1 is taken over all non-zero vectors of $\oplus_{\theta \in M_1} A_{j_1}^\theta$, and where for every (j_1, w_1) and every $\theta \in \Theta_{(j_1, w_1)}$ the line containing w_1 is contained in $A_{j_1}^\theta$.

We now fix (j_1, w_1), and let M_2 be a maximal subset of $\Theta_{(j_1,w_1)}$ such that the linear subspaces $A^\theta_{j_1}$ with $\theta \in M_2$ have pairwise intersection $\langle w_1 \rangle$, and such that for every $j \neq j_1$ the linear subspaces A^θ_j with $\theta \in M_2$ are in direct sum. Again, by applying Theorem 10 the set M_2 has size at most d, as otherwise we would have $\operatorname{sr} T \leq 1$, and we can further write

$$\Theta_{(j_1,w_1)} = \bigcup_{1 \leq j_2 \leq d} \bigcup_{w_2} \Theta_{(j_1,w_1),(j_2,w_2)}$$

where the union over w_2 is taken over all non-zero vectors of $\oplus_{\theta \in M_1} A^\theta_{j_2}$ if $j_2 \neq j_1$ (resp. taken over all vectors linearly independent from w if $j_2 = j_1$), and where for every $\theta \in \Theta_{(j_2,w_2)}$ the line containing w_2 is contained in $A^\theta_{j_2}$ (resp. the plane containing w_1, w_2 is contained in $A^\theta_{j_2}$).

Iterating further we obtain a tree structure with depth k, and with root $[h]$. Let us describe the inductive step of the process. Once the tree is constructed up to some depth $\kappa \leq k$, a (depth-κ) leaf of the tree constructed at this point is a subset

$$\Theta_{(j_1,w_1),(j_2,w_2),\dots,(j_\kappa,w_\kappa)}$$

of $[h]$ such that there exist linear subspaces $W_1 \subset \mathbb{F}^{n_1}, \dots, W_d \subset \mathbb{F}^{n_d}$ with

$$\dim W_1 + \dots + \dim W_d = \kappa$$

and which satisfy $W_j \subset A^\theta_j$ for every $j \in [d]$ and every $\theta \in \Theta_{(j_1,w_1),(j_2,w_2),\dots,(j_\kappa,w_\kappa)}$. If $\kappa < k$, then let $M_{\kappa+1}$ be a maximal subset of $\Theta_{(j_1,w_1),(j_2,w_2),\dots,(j_\kappa,w_\kappa)}$ such that the linear subspaces A^θ_j with $\theta \in M_{\kappa+1}$ have pairwise intersection W_j for every $j \in [d]$. Theorem 10 then shows that $M_{\kappa+1}$ has size at most d (as otherwise we would have $\operatorname{sr} T \leq \kappa$), and we can hence write

$$\Theta_{(j_1,w_1),(j_2,w_2),\dots,(j_\kappa,w_\kappa)} = \bigcup_{1 \leq j_{\kappa+1} \leq d} \bigcup_{w_{\kappa+1}} \Theta_{(j_1,w_1),(j_2,w_2),\dots,(j_\kappa,w_\kappa),(j_{\kappa+1},w_{\kappa+1})}, \tag{10}$$

where the union over $w_{\kappa+1}$ is over all vectors of

$$(+_{\theta \in M_{\kappa+1}} A^\theta_{j_{\kappa+1}}) \setminus W_{j_{\kappa+1}}.$$

We then take the sets $\Theta_{(j_1,w_1),(j_2,w_2),\dots,(j_\kappa,w_\kappa),(j_{\kappa+1},w_{\kappa+1})}$ to be the immediate descendants of the set $\Theta_{(j_1,w_1),(j_2,w_2),\dots,(j_\kappa,w_\kappa)}$ in the tree.

If instead $\kappa = k$ then $\dim W_1 + \dots + \dim W_d$ and $\dim A^\theta_1 + \dots + \dim A^\theta_d$ are both equal to k, so are equal to one another. Hence, once we reach depth k the linear subspaces $A^\theta_1, \dots, A^\theta_d$ are all completely determined by $W_1, \dots, W_d$. By our initial assumption that for any distinct $\theta, \theta' \in [h]$ we have $A^\theta_j \neq A^{\theta'}_j$ for some $j \in [d]$, the sets $\Theta_{(j_1,w_1),(j_2,w_2),\dots,(j_k,w_k)}$ each have size at most 1.

We have obtained a tree with depth k, and each node of the tree has at most $d|\mathbb{F}|^{dk}$ immediate descendants, so the number of leaves, and hence of d-tuples of linear subspaces $(A^\theta_1, \dots, A^\theta_d)$ is at most $(d|\mathbb{F}|^{dk})^k = d^k |\mathbb{F}|^{dk^2}$. Furthermore, for each choice of a d-tuple of subspaces $(A^\theta_1, \dots, A^\theta_d)$ with respective dimensions $r_1, \dots, r_d$, by the number (1) of decompositions of matrices there are at most

$$|\mathbb{F}|^{r_1^2} \dots |\mathbb{F}|^{r_d^2} \leq |\mathbb{F}|^{k^2}$$

possibilities for

$$((a_{1,1}, \dots, a_{1,r_1}), \dots, (a_{d,1}, \dots, a_{d,r_d})).$$

This finishes the proof of the first item of Theorem 12. ◀

5 Simpler analogue for tensor rank decompositions

In this section we prove a simpler analogue of Theorem 12 in the case of the tensor rank, which is much more similar to the statement for matrices that we discussed in the introduction.

▶ **Proposition 14.** *Let $d \geq 2$ be an integer, let k be a nonnegative integer, and let $\mathbb{F}$ be a field. If $T : [n_1] \times \cdots \times [n_d] \to \mathbb{F}$ is an order-d tensor with tensor rank k, then there exist linear subspaces $A_1 \subset \mathbb{F}^{n_1}, \ldots, A_d \subset \mathbb{F}^{n_d}$ such that if*

$$T(x_1, \ldots, x_d) = \sum_{i=1}^{k} a_{1,i}(x_1) \ldots a_{d,i}(x_d) \tag{11}$$

is a tensor rank decomposition of T with length k, then we have

$$A_1 = \langle a_{1,1}, \ldots, a_{1,k} \rangle, \ldots, A_d = \langle a_{d,1}, \ldots, a_{d,k} \rangle.$$

In particular, if the field $\mathbb{F}$ is finite, then the number of possible tensor rank decompositions of T with length k is at most $|\mathbb{F}|^{(d-1)k^2}$.

Proof. Assume that

$$T(x_1, \ldots, x_d) = \sum_{i=1}^{k} a_{1,i}(x_1) \ldots a_{d,i}(x_d) = \sum_{i=1}^{k} a'_{1,i}(x_1) \ldots a'_{d,i}(x_d)$$

are two decompositions of T with length k. To prove the first conclusion, it suffices to show that for each $j \in [d]$ the linear subspaces $\langle a_{j,1}, \ldots, a_{j,k} \rangle$ and $\langle a'_{j,1}, \ldots, a'_{j,k} \rangle$ are the same. Assume for contradiction that they are not. Then without loss of generality the linear subspace $\langle a_{1,1}, \ldots, a_{1,k} \rangle$ is not contained in the linear subspace $\langle a'_{1,1}, \ldots, a'_{1,k} \rangle$, so there exists $u : [n_1] \to \mathbb{F}$ such that $(u.a'_{1,1}, \ldots, u.a'_{1,k}) = 0$ but $(u.a_{1,1}, \ldots, u.a_{1,k}) \neq 0$. Applying u to both decompositions of T we obtain

$$\sum_{i=1}^{k} (u.a_{1,i}) a_{2,i}(x_2) \ldots a_{d,i}(x_d) = 0,$$

so the products $a_{2,i} \otimes \cdots \otimes a_{d,i}$ with $i \in [k]$ are linearly dependent. Assuming without loss of generality that the product $a_{2,k} \otimes \cdots \otimes a_{d,k}$ is a linear combination of the products $a_{2,i} \otimes \cdots \otimes a_{d,i}$ with $i \in [k-1]$, we can hence write

$$T(x_1, \ldots, x_d) = \sum_{i=1}^{k-1} \alpha_{1,i}(x_1) a_{2,i}(x_2) \ldots a_{d,i}(x_d)$$

for some new functions $\alpha_{1,1}, \ldots, \alpha_{1,k-1} : [n_1] \to \mathbb{F}$. This is a tensor rank decomposition of T with length $k-1$, which contradicts that T has tensor rank k.

If the field $\mathbb{F}$ is finite, then for each of the $k(d-1)$ functions $a_{j,i}$ with $j \in \{2, \ldots, d\}$ and $i \in [k]$ involved in a given decomposition of T, there are at most $|A_j| \leq |\mathbb{F}|^k$ possibilities. Once these functions are fixed, the linear independence of the products $a_{2,i} \otimes \cdots \otimes a_{d,i}$ with $i \in [k]$, shows that the functions $a_{1,1}, \ldots, a_{1,k}$ are determined. The bound follows. ◀

Although the upper bound that we obtain specialises in the case $d = 2$ to the upper bound obtained at the start of the introduction, the number of tensor rank decompositions can be much less than $|\mathbb{F}|^{(d-1)k^2}$ for general d. Indeed if an order-d tensor T with tensor rank k is identifiable in the sense that we mentioned in the introduction, then it can be written in at most $k!$ ways as a sum of k tensors with tensor rank 1 (those being the same, up to permutation); as in turn there are $(|\mathbb{F}|-1)^{d-1}$ ways of rewriting a tensor with tensor rank 1 as a product of d functions, such a tensor has at most $k!(|\mathbb{F}|-1)^{d-1}$ choices for the number of possible tensor rank decompositions in our sense. Nonetheless, the following example shows that the behaviour from the $d=2$ case for k large can carry over to tensors of higher order, and that the exponent in the bound from Proposition 14 cannot be improved by a factor greater than d in general. We recall the absolute constant $\omega = \prod_{i=1}^{\infty}(1-2^{-i}) > 0$ defined at the start of the introduction.

▶ **Example 15.** *Let $d \geq 2$ be an integer, and let k be a nonnegative integer. Let T be the order-d tensor defined by*

$$T(x_1,\dots,x_d) = M(x_1,x_2)a_3(x_3)\dots a_d(x_d) \tag{12}$$

for some matrix $M : [n_1]\times[n_2] \to \mathbb{F}$ with rank k and some non-zero functions $a_3 : [n_3] \to \mathbb{F},\dots,a_d : [n_d] \to \mathbb{F}$. Then T has tensor rank equal to k, and the number of tensor rank decompositions of length k of T is exactly

$$(|\mathbb{F}|-1)^{(d-2)k}|\mathbb{F}|^{k^2}\prod_{i=1}^{k}(1-|\mathbb{F}|^{-i}),$$

which is between $\omega|\mathbb{F}|^{k^2}$ and $|\mathbb{F}|^{k^2+(d-2)k}$.

Proof. In one direction, writing a rank decomposition of the matrix M with length k and plugging it in (12) shows $\operatorname{tr} T \leq k$. In the other direction, if

$$T(x_1,\dots,x_d) = \sum_{i=1}^{k'} a_{1,i}(x_1)a_{2,i}(x_2)a_{3,i}(x_3)\dots a_{d,i}(x_d) \tag{13}$$

is a tensor rank decomposition of T for some nonnegative integer k', then letting $a_3^* : [n_3] \to \mathbb{F},\dots,a_d^* : [n_d] \to \mathbb{F}$ be functions such that $a_j^*.a_j = 1$ for each $j \in \{3,\dots,d\}$ and applying $a_3^*,\dots,a_d^*$ to both decompositions (12) and (13) leads to

$$M(x_1,x_2) = \sum_{i=1}^{k'} C_i a_{1,i}(x_1)a_{2,i}(x_2)$$

for some coefficients $C_i \in \mathbb{F}$, so $k' \geq \operatorname{rk} M$ and hence $\operatorname{tr} T \geq k$.

The decomposition (12) shows that the linear subspaces $A_3,\dots,A_d$ of Proposition 14 each have dimension 1, and there are hence at most $(|\mathbb{F}|-1)^{(d-2)k}$ choices for

$$((a_{3,1},\dots,a_{3,k}),\dots,(a_{d,1},\dots,a_{d,k})) \tag{14}$$

since if one of these functions were zero, then this would contradict $\operatorname{tr} T = k$. Once this choice is made, Proposition 14 shows that there are at most

$$|\mathbb{F}|^{k^2}\prod_{i=1}^{k}(1-|\mathbb{F}|^{-i})$$

choices for $(a_{2,1},\dots,a_{2,k})$, and the linear independence of the products $a_{2,i}\otimes\dots\otimes a_{d,i}$ with $i \in [k]$ involved in a minimal-length tensor rank decomposition then shows that $(a_{1,1},\dots,a_{1,k})$

is completely determined. Conversely, for any given choice of (14) such that $a_{j,i} \in A_j \setminus \{0\}$ for every $j \in \{3, \dots, d\}$ and $i \in [k]$, the decomposition (11) becomes

$$\sum_{i=1}^{k} D_i a_{1,i}(x_1) a_{2,i}(x_2) a_3(x_3) \dots a_d(x_d)$$

for some $D_1, \dots, D_k \in \mathbb{F} \setminus \{0\}$, so whenever $f_1, \dots, f_k : [n_1] \to \mathbb{F}$, $g_1, \dots, g_k : [n_2] \to \mathbb{F}$ provide a decomposition

$$M(x_1, x_2) = \sum_{i=1}^{k} f_i(x_1) g_i(x_2)$$

of M with length k, taking $a_{1,i} = f_i$ and $a_{2,i} = g_i / D_i$ for every $i \in [k]$ provides a decomposition of T with length k. ◀

6 Open questions

Let us finish by discussing a few questions that are left open by our current results and proofs. As we had explained in Section 5, in the tensor rank case the exponent in the upper bound that we show on the number of minimal-length tensor rank decompositions cannot be improved by more than a factor of d, but this nonetheless does not give a complete answer to our first question.

▶ **Question 16.** *What are the optimal bounds in Proposition 14, and for which tensors are they attained ?*

Returning to the slice rank, Example 13 shows that Theorem 12 is false for infinite fields, that in the finite field case the bounds cannot be uniform with respect to $\mathbb{F}$, and furthermore that even for a fixed finite field they grow square-exponentially in the slice rank of the tensor. This leads us instead to the following formulation, which we believe provides a statement of a similar kind and which is uniform over all fields.

▶ **Conjecture 17.** *Let $d \geq 3$, $k \geq 1$ be integers. Then there exists a positive integer $C(d,k)$ such that the following holds. If T is an order-d tensor over some arbitrary field $\mathbb{F}$ and with slice rank equal to k, then there exist linear subspaces*

$$A_1 \subset \mathbb{F}^{n_1}, \dots, A_d \subset \mathbb{F}^{n_d}$$

each with dimension at most $C(d,k)$ such that if $r_1, \dots, r_d$ are nonnegative integers satisfying

$$r_1 + \dots + r_d = k$$

and $a_{j,i} : [n_j] \to \mathbb{F}$ with $j \in [d]$ and $i \in [r_j]$ are functions satisfying

$$T(x_1, \dots, x_d) = \sum_{j=1}^{d} \sum_{i=1}^{r_j} a_{j,i}(x_j) b_{j,i}(\overline{x_j})$$

for some $(d-1)$-variable functions $b_{j,i} : \prod_{j' \neq j} [n_{j'}] \to \mathbb{F}$, then

$$\langle a_{1,1}, \dots, a_{1,r_1} \rangle \subset A_1, \dots, \langle a_{d,1}, \dots, a_{d,r_d} \rangle \subset A_d.$$

We would furthermore not be surprised if for any fixed $d \geq 3$ we could take $C(d,k)$ to be linear in k, as we have not managed to build a counterexample disproving this.

▶ **Question 18.** *Is Conjecture 17 true with furthermore $C(d,k) = O(k)$ for every $d \geq 3$?*

One further direction in which we believe that our results can be taken further is that of formulating and proving suitable generalisations of Theorem 10 and Theorem 12 for other notions of rank. As discussed in the introduction, one such notion that is of interest is the partition rank.

▶ **Question 19.** *Can we prove an analogue of Theorem 12 for the partition rank ?*

As was the case for the slice rank, formulating an adequate statement already appears to be challenging, and it is useful to keep in mind that the five ways in which various slice rank decompositions can arise all extend to the partition rank. Let us finish by describing these extensions. For $d \geq 2$ an integer, we write $\mathcal{B}_2([d])$ for the set of bipartitions $\{J, J^c\}$ of $[d]$ with $J, J^c \neq \emptyset$. We recall that if $d \geq 2$ is an integer, J is a subset of $[d]$, then for every $x \in [n_1] \times \dots \times [n_d]$ we write $x(J)$ for the restriction of x to its coordinates in J. Let us also recall the definition of the partition rank.

▶ **Definition 20.** *Let $d \geq 2$ be an integer, and let $T : [n_1] \times \dots \times [n_d] \to \mathbb{F}$ be an order-d tensor. We say that the* partition rank *of T, denoted by $\operatorname{pr} T$, is the smallest nonnegative integer k such that there exist nonnegative integers $r_{\{J,J^c\}}$ for each bipartition $\{J, J^c\} \in \mathcal{B}_2([d])$, satisfying*

$$\sum_{\{J,J^c\} \in \mathcal{B}_2([d])} r_{\{J,J^c\}} = k$$

and one of the following two equivalent properties.

1. *There exist matrices $M_{\{J,J^c\}} : (\prod_{j \in J} [n_j]) \times (\prod_{j \in J^c} [n_j]) \to \mathbb{F}$ with rank at most $r_{\{J,J^c\}}$ for each $\{J, J^c\} \in \mathcal{B}_2([d])$ such that*

$$T(x_1, \dots, x_d) = \sum_{\{J,J^c\} \in \mathcal{B}_2([d])} M_{\{J,J^c\}}(x(J), x(J^c)) \tag{15}$$

   is satisfied for all $(x_1, \dots, x_d) \in [n_1] \times \dots \times [n_d]$.
2. *For each $\{J, J^c\} \in \mathcal{B}_2([d])$ and each $i \in [r_{\{J,J^c\}}]$ there exist some functions $a_{J,i} : \prod_{j \in J} [n_j] \to \mathbb{F}$ and $a_{J^c,i} : \prod_{j \in J^c} [n_j] \to \mathbb{F}$ such that*

$$T(x_1, \dots, x_d) = \sum_{\{J,J^c\} \in \mathcal{B}_2([d])} \sum_{i=1}^{r_{\{J,J^c\}}} a_{J,i}(x(J)) a_{J^c,i}(x(J^c)) \tag{16}$$

   is satisfied for all $(x_1, \dots, x_d) \in [n_1] \times \dots \times [n_d]$.

We say that an expression such as (16) *is a* partition rank decomposition *of T, say that the integer*

$$\sum_{\{J,J^c\} \in \mathcal{B}_2([d])} r_{\{J,J^c\}}$$

is the length *of the decomposition, and say that the decomposition has* minimal length *if its length is equal to the partition rank of T.*

Constructions 3 to 7 from Section 2 respectively adapt to the following. Let $d \geq 3$, $k \geq 1$ be integers, and let T be an order-d tensor with partition rank k.

▶ **Construction 21.** If $M_{\{J,J^c\}} : (\prod_{j\in J}[n_j]) \times (\prod_{j\in J^c}[n_j]) \to \mathbb{F}$ with $\{J,J^c\} \in \mathcal{B}_2([d])$ are matrices satisfying (15) and

$$\sum_{\{J,J^c\}\in\mathcal{B}_2([d])} \operatorname{rk} M_{\{J,J^c\}} = \operatorname{pr} T,$$

then rewriting any minimal-length matrix decomposition of any of the matrices $M_{\{J,J^c\}}$ leads to a new minimal-length partition rank decomposition of T.

▶ **Construction 22.** If the tensor rank of T is equal to the partition rank k of T, and

$$T(x_1,\dots,x_d) = \sum_{i=1}^{k} a_{1,i}(x_1)\dots a_{d,i}(x_d)$$

is a length-k tensor rank decomposition of T then for any map $\iota : [k] \to \mathcal{B}_2([d])$ the decomposition

$$T(x_1,\dots,x_d) = \sum_{\{J,J^c\}\in\mathcal{B}_2([d])} \sum_{i\in\iota^{-1}(\{J,J^c\})} (\prod_{j\in J} a_{j,i})(x(J))(\prod_{j\in J^c} a_{j,i})(x(J^c))$$

is a minimal-length partition rank decomposition of T. This is merely the most extreme form that this construction can take, and tensors T satisfying much weaker assumptions than $\operatorname{pr} T = \operatorname{tr} T$ will lend themselves to it. For instance if T has partition rank k and can merely be written as

$$T(x_1,\dots,x_d) = \sum_{i=1}^{k} a_{i,J_{i,1}}(x(J_{i,1}))a_{i,J_{i,2}}(x(J_{i,2}))a_{i,J_{i,3}}(x(J_{i,3}))$$

where for every $i \in [k]$ the set $\{J_{i,1}, J_{i,2}, J_{i,3}\}$ is a tripartition of $[d]$ with $J_{i,1}, J_{i,2}, J_{i,3}$ each non-empty and $a_{i,J_{i,s}} : \prod_{j\in J_{i,s}}[n_j] \to \mathbb{F}$ is a function for each $s = 1,2,3$ then we already obtain different minimal-length partition rank decompositions of T depending on how we view each summand as a product of two terms.

▶ **Construction 23.** If $d \geq 3$, then for k large a tensor $[k]^d \to \mathbb{F}$ usually has partition rank equal to k. Indeed, if the field $\mathbb{F}$ is finite then there are at most

$$k^{2^d}|\mathbb{F}|^{(k-1)(k^{d-1}+k)} \leq (k^{2^d}/|\mathbb{F}|^k)|\mathbb{F}|^{k^d}$$

tensors with partition rank at most $k-1$. If T has partition rank equal to k, then writing T as the sum of its $(d-1)$-variable slices as in Construction 5 again provides d minimal-length partition rank decomposition of T. We can also write T as a sum of its d'-variable slices for $d' < d-1$, but the resulting partition rank decomposition then has length $k^{d-d'} > k$, which is not minimal, and we hence have less of a generalisation in this direction.

▶ **Construction 24.** If the decomposition

$$T(x_1,\dots,x_d) = \sum_{\{J,J^c\}\in\mathcal{B}_2([d])} \sum_{i=1}^{r_{\{J,J^c\}}} a_{J,i}(x(J))a_{J^c,i}(x(J^c))$$

is a minimal-length decomposition of T, $\{J_1, J_2, J\}$ is a tripartition of $[d]$ with J_1, J_2 each with size at least 1, $i_1 \in [r_{\{J_1,J_1^c\}}]$ and $i_2 \in [r_{\{J_2,J_2^c\}}]$ are indices, and $c : \prod_{j\in J}[n_j] \to \mathbb{F}$ is a function if J is non-empty and an element of $\mathbb{F}$ otherwise, then replacing $a_{J_1^c,i_1}$ and $a_{J_2^c,i_2}$ by respectively

$$a_{J_1^c,i_1} + a_{J_2,i_2} \otimes c \text{ and } a_{J_2^c,i_2} - a_{J_1,i_1} \otimes c$$

leads to a new minimal-length partition rank decomposition of T.

▶ **Construction 25.** More generally, if $s \in [2, d]$ is an integer, $\{J_1, \dots, J_s, J\}$ is a partition of $[d]$ with $J_1, \dots, J_s$ each with size at least 1, $i_1 \in [r_{\{J_1, J_1^c\}}], \dots, i_s \in [r_{\{J_s, J_s^c\}}]$ are indices, and $c_1, \dots, c_s$ are functions $\prod_{j \in J} [n_j] \to \mathbb{F}$ if J is non-empty and elements of $\mathbb{F}$ otherwise, which satisfy

$$c_1 + \dots + c_s = 0,$$

then replacing $a_{J_t^c, i_t}$ by

$$a_{J_t^c, i_t} + (\otimes_{t' \in [s] \setminus \{t\}} a_{J_{t'}, i_{t'}}) \otimes c_t$$

for each $t \in [s]$ leads to a new minimal-length partition rank decomposition of T.

Construction 24 shows that whenever $|J| \geq 2$, it is never true (for any integers $d \geq 3$, $k \geq 2$, and any field $\mathbb{F}$) that for every order-d tensor T with partition rank k over $\mathbb{F}$ there exists a linear subspace A_J of $\otimes_{j \in J} \mathbb{F}^{n_j}$ with dimension bounded above depending on $d, k, \mathbb{F}$ only such that any functions $a_{J,i}$ from any minimal-length partition rank decomposition of T are contained in A_J.

Moreover, in the case of the partition rank there may be new constructions in addition to the analogues of Constructions 3 to 7 discussed above.

References

1 Edoardo Ballico, Alessandra Bernardi, and Pierpaola Santarsiero. Identifiability of rank-3 tensors, 2021.

2 Arthur Bik, Alessandro Danelon, Jan Draisma, and Rob H. Eggermont. Universality of high-strength tensors. *Vietnam Journal of Mathematics*, 50, 2021.

3 Arthur Bik, Jan Draisma, and Rob H. Eggermont. Polynomials and tensors of bounded strength. *Communications in Contemporary Mathematics*, 21, 2019.

4 Jop Briët, Harry Buhrman, Davi Castro-Silva, and Niels M. P. Neumann. Noisy decoding by shallow circuits with parities: classical and quantum, 2023.

5 Jop Briët and Davi Castro-Silva. Random restrictions of high-rank tensors and polynomial maps, 2022.

6 Ernie Croot, Vsevolod Lev, and Peter Pach. Progression-free sets in $\mathbb{Z}_4^n$ are exponentially small. *Annals of Mathematics*, 185, 2016.

7 Jordan S. Ellenberg and Dion Gijswijt. On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression. *Annals of Mathematics*, 185, 2016.

8 W. T. Gowers. The slice rank of a direct sum, 2021.

9 W. T. Gowers and Thomas Karam. Equidistribution of high-rank polynomials with variables restricted to subsets of $\mathbb{F}_p$, 2022.

10 Ben Green and Terence Tao. The distribution of polynomials over finite fields, with applications to the gowers norms. *Contributions to Discrete Mathematics*, 4, 2007.

11 Thomas Karam. High-rank subtensors of high-rank tensors, 2022.

12 David Kazhdan and Tamar Ziegler. Properties of high rank subvarieties of affine spaces. *Geometric and Functional Analysis*, 30, 2020.

13 Joseph B. Kruskal. Three-way arrays: Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. *Linear Algebra and its Applications*, 18, 1977.

14 Eric Naslund. The partition rank of a tensor and k-right corners in $\mathbb{F}_q^n$. *Journal of Combinatorial Theory, Series A*, 174, 2020.

15 Eric Naslund. The chromatic number of $\mathbb{R}^n$ with multiple forbidden distances, 2023.

16 Eric Naslund and Will Sawin. Upper bounds for sunflower-free sets. *Forum of Mathematics, Sigma*, 5, 2017.

17 Lisa Sauermann. Finding solutions with distinct variables to systems of linear equations over $\mathbb{F}_p$. *Mathematische Annalen*, 386, 2021.

18 Will Sawin and Terence Tao. Notes on the “slice rank” of tensors, 2016.

19 Terence Tao. A symmetric formulation of the croot-lev-pach-ellenberg-gijswijt capset bound, 2016.