Finding solutions with distinct variables to systems of linear equations over $$\mathbb {F}_p$$ F p

Mathematische Annalen
https://doi.org/10.1007/s00208-022-02391-y

Mathematische Annalen

Finding solutions with distinct variables to systems of linear equations over $\mathbb{F}_p$

Lisa Sauermann[1]

Received: 9 June 2021 / Revised: 29 January 2022 / Accepted: 5 March 2022

Abstract

Let us fix a prime p and a homogeneous system of m linear equations $a_{j,1}x_1 + \cdots + a_{j,k}x_k = 0$ for $j = 1, \ldots, m$ with coefficients $a_{j,i} \in \mathbb{F}_p$. Suppose that $k \geq 3m$, that $a_{j,1} + \cdots + a_{j,k} = 0$ for $j = 1, \ldots, m$ and that every $m \times m$ minor of the $m \times k$ matrix $(a_{j,i})_{j,i}$ is non-singular. Then we prove that for any (large) n, any subset $A \subseteq \mathbb{F}_p^n$ of size $|A| > C \cdot \Gamma^n$ contains a solution $(x_1, \ldots, x_k) \in A^k$ to the given system of equations such that the vectors $x_1, \ldots, x_k \in A$ are all distinct. Here, C and Γ are constants only depending on p, m and k such that $\Gamma < p$. The crucial point here is the condition for the vectors $x_1, \ldots, x_k$ in the solution $(x_1, \ldots, x_k) \in A^k$ to be distinct. If we relax this condition and only demand that $x_1, \ldots, x_k$ are not all equal, then the statement would follow easily from Tao's slice rank polynomial method. However, handling the distinctness condition is much harder, and requires a new approach. While all previous combinatorial applications of the slice rank polynomial method have relied on the slice rank of diagonal tensors, we use a slice rank argument for a non-diagonal tensor in combination with combinatorial and probabilistic arguments.

1 Introduction

Given a linear system of equations with coefficients in $\mathbb{F}_p$ for some fixed prime p, what is the largest size of a subset $A \subseteq \mathbb{F}_p^n$ which does not contain a (non-trivial) solution to the given linear system of equations? This is a fundamental question in additive combinatorics, and it can be viewed as the finite field analog of similar questions for subsets $A \subseteq \{1, \ldots, N\}$ whose history dates back many decades (see e.g. [9,

Communicated by Kannan Soundararajan.

Research supported by NSF Grant CCF-1900460 and NSF Award DMS-2100157.

✉ Lisa Sauermann
lsauerma@mit.edu

[1] Mathematics Department, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Published online: 01 April 2022

15]). For example, a k-term arithmetic progression can be described by a system of $k-2$ linear equations (with k variables), and bounding the largest possible size of a k-term-progression-free subset in $\mathbb{F}_p^n$ is an intensively studied and still wide open problem [1, 5–7, 11–13, 17, 18]. Another well-studied special case is the system consisting of the single equation $x_1+x_2-x_3-x_4=0$, and subsets $A\subseteq\mathbb{F}_p^n$ without non-trivial solutions to this equation are called *Sidon sets* (see e.g. [3]). More generally, for any integer $h\geq 2$, subsets without a non-trivial solution to the equation $x_1+\cdots+x_h-x_{h+1}-\cdots-x_{2h}=0$ are called *B_h-sets* and the problem of bounding the largest possible size of a B_h-set in $\mathbb{F}_p^n$ is still open (see e.g. [20]).

Let us from now on fix a prime p and consider a linear system of m equations in k variables of the form

$$\begin{aligned} a_{1,1}x_1+\cdots+a_{1,k}x_k&=0 \\ &\vdots \\ a_{m,1}x_1+\cdots+a_{m,k}x_k&=0 \end{aligned} \qquad (\star)$$

with coefficients $a_{j,i}\in\mathbb{F}_p$ for $j=1,\dots,m$ and $i=1,\dots,k$. For large n, we are interested in the largest possible size of a subset $A\subseteq\mathbb{F}_p^n$ such that there is no (non-trivial) solution $(x_1,\dots,x_k)\in A^k$ to $(\star)$.

If we have $a_{j,1}+\cdots+a_{j,k}\neq 0$ for some $j\in\{1,\dots,m\}$ (i.e. if for one of the m equations the coefficients do not sum up to zero), then it is easy to see that there exists a subset $A\subseteq\mathbb{F}_p^n$ of size $p^{n-1}=(1/p)\cdot p^n$ such that the system $(\star)$ does not have any solutions $(x_1,\dots,x_k)\in A^k$ (indeed, we can take A to be the set of all vectors in $\mathbb{F}_p^n$ whose first coordinate is 1). For fixed p, this means that up to constant factors A can be as large as the entire space $\mathbb{F}_p^n$. However, the problem becomes much more interesting when $a_{j,1}+\cdots+a_{j,k}=0$ for $j=1,\dots,m$.

So let us from now on assume that $a_{j,1}+\cdots+a_{j,k}=0$ for $j=1,\dots,m$. Note that then $(x,\dots,x)$ is a solution to the system $(\star)$ for every $x\in A$. However, we can ask for the largest size of a subset $A\subseteq\mathbb{F}_p^n$ without a solution $(x_1,\dots,x_k)\in A^k$ to $(\star)$ where $x_1,\dots,x_k$ are not all equal. It is a highly non-trivial result to show that we must have $|A|=o(p^n)$ as $n\to\infty$ (this can for example be deduced from [16,Theorem 1] or [27,Theorem 2.2], both of which rely on deep results on hypergraph regularity, and the actual quantitative bounds obtained this way are very poor).

Building upon a breakthrough of Ellenberg and Gijswijt [6] on bounding the size of 3-term-progression-free subsets of $\mathbb{F}_p^n$ as well as on prior work of Croot, Lev, and Pach [4], Tao [28] introduced a new polynomial method, which is now called the *slice rank polynomial method*. This method immediately gives much stronger upper bounds on the size of A in the question above if we assume that the number k of variables in the system $(\star)$ is sufficiently large with respect to the number m of equations. In fact, assuming that $k\geq 2m+1$, one can prove that the size of A needs to be *exponentially smaller* than p^n if there is no solution $(x_1,\dots,x_k)\in A^k$ to $(\star)$ where $x_1,\dots,x_k$ are not all equal.

Theorem 1.1 *(Tao) For any fixed integers $m\geq 1$ and $k\geq 2m+1$ and a fixed prime p, there exists a constant $1\leq\Gamma_{p,m,k}<p$ such that the following holds: For any*

coefficients $a_{j,i} \in \mathbb{F}_p$ *for* $j = 1, \dots, m$ *and* $i = 1, \dots, k$ *with* $a_{j,1} + \cdots + a_{j,k} = 0$ *for* $j = 1, \dots, m$, *for any non-negative integer* n *and any subset* $A \subseteq \mathbb{F}_p^n$ *of size* $|A| > (\Gamma_{p,m,k})^n$, *the system* ($\star$) *has a solution* $(x_1, \dots, x_k) \in A^k$ *such that the vectors* $x_1, \dots, x_k \in A$ *are not all equal.*

As mentioned above, Theorem 1.1 follows immediately from Tao's slice rank polynomial method [28]. For the reader's convenience we will present the proof in Sect. 3.

Recall that our original question was to bound the size of a subset $A \subseteq \mathbb{F}_p^n$ which does not contain a non-trivial solution to the system ($\star$). Obviously, it needs to be specified what we mean by "non-trivial" here. If a solution $(x_1, \dots, x_k) \in A^k$ is considered non-trivial as soon as $x_1, \dots, x_k \in A$ are not all equal, then Theorem 1.1 provides a strong bound for $|A|$ as long as $k \geq 2m + 1$. However, there are also several other notions of "non-trivial" solutions to a linear equation or a system of linear equations (see for example Ruzsa's work [23, 24]), and a particularly natural such notion is to demand for a "non-trivial" solution $(x_1, \dots, x_k) \in A^k$ to consist of *distinct* $x_1, \dots, x_k$.

In the integer setting, the problem of bounding the largest size of a subset $A \subseteq \{1, \dots, N\}$ with no solution $(x_1, \dots, x_k) \in A^k$ to a given linear system of equations where $x_1, \dots, x_k$ are distinct has already been considered almost fifty years ago [15]. Here, we consider the same problem in the setting of $\mathbb{F}_p^n$, i.e. we are asking for the largest size of a subset $A \subseteq \mathbb{F}_p^n$ which does not contain a solution $(x_1, \dots, x_k) \in A^k$ to ($\star$) with distinct $x_1, \dots, x_k$.

Similarly to Theorem 1.1 above, our main result states that if the number k of variables in the system ($\star$) is sufficiently large with respect to the number m of equations and if the system ($\star$) is reasonably generic, then the size of A must be exponentially smaller than p^n.

Theorem 1.2 *For any fixed integers* $m \geq 1$ *and* $k \geq 3m$ *and a fixed prime* p, *there exist constants* $C_{p,m,k} \geq 1$ *and* $1 \leq \Gamma^*_{p,m,k} < p$ *such that the following holds: Let* $a_{j,i} \in \mathbb{F}_p$ *for* $j = 1, \dots, m$ *and* $i = 1, \dots, k$ *be coefficients with* $a_{j,1} + \cdots + a_{j,k} = 0$ *for* $j = 1, \dots, m$ *such that every* $m \times m$ *minor of the* $m \times k$ *matrix* $(a_{j,i})_{j,i}$ *is non-singular. Then for any non-negative integer* n *and any subset* $A \subseteq \mathbb{F}_p^n$ *of size* $|A| > C_{p,m,k} \cdot (\Gamma^*_{p,m,k})^n$, *the system* ($\star$) *has a solution* $(x_1, \dots, x_k) \in A^k$ *such that the vectors* $x_1, \dots, x_k \in A$ *are all distinct.*

While it may seem at first that it should not make much of a difference whether one demands $x_1, \dots, x_k \in A$ to be distinct or to be not all equal, it is in fact much more difficult to prove the existence of a solution $(x_1, \dots, x_k) \in A^k$ with distinct $x_1, \dots, x_k$ as in Theorem 1.2. In fact, in some cases demanding such a distinct solution can require qualitatively different bounds for the size of A (see Theorems 2 and 4 in [20], and see also Theorems 3.2 and 3.3 in [23]). The slice rank polynomial method argument proving Theorem 1.1 completely breaks down when requiring $x_1, \dots, x_k$ to be distinct, and new ideas are required in order to prove Theorem 1.2. In particular, our proof uses a variant of the slice rank polynomial method that has not appeared in combinatorial applications before.

The problem of bounding the size of a subset of $A \subseteq \mathbb{F}_p^n$ not containing a solution with distinct variables to a given linear equation or system of linear equations has been studied by various authors in the past. In the case of the single equation $x_1 + \dots + x_p = 0$, which is very closely related to the Erdős-Ginzburg-Ziv problem in discrete geometry, Naslund [22] proved a bound of the form $|A| \le C_p \cdot (\Gamma_p)^n$ for some Γ_p between $0.84p$ and $0.92p$ (the constant C_p was later improved in [10]). The best current bound in this case is $|A| \le C_p \cdot (2\sqrt{p})^n$ due to the author [25].

In the case of a general single linear equation $a_1x_1 + \dots + a_kx_k = 0$ with $a_1 + \dots + a_k = 0$ (i.e. in the case of $m = 1$), Theorem 1.2 has been proved by Mimura and Tokushige [21,Theorem 1], and this case also follows from [25,Theorem 5.1] with significantly better bounds for $\Gamma^*_{p,1,k}$. Different specific examples of linear systems of multiple equations have been considered in [19–21], but Theorem 1.2 is the first relatively general result for systems of multiple equations.

We remark that for fixed m and k (with $k \ge 3m$) and a prime p that is large with respect to m and k, most systems ($\star$) satisfy the condition in Theorem 1.2 requiring that every $m \times m$ minor of the $m \times k$ matrix $(a_{j,i})_{j,i}$ is non-singular. More precisely, for $p \to \infty$ (with m and k fixed) the fraction of choices for the coefficients $a_{j,i} \in \mathbb{F}_p$ (with $a_{j,1} + \dots + a_{j,k} = 0$ for all j) that violate this non-singularity condition tends to zero. In this sense, for large p, and fixed $k \ge 3m$, Theorem 1.2 applies to *almost all* systems ($\star$) where the problem is of interest (i.e. where $a_{j,1} + \dots + a_{j,k} = 0$ for $j = 1, \dots, m$).

It is worth noting that one cannot hope to remove the assumption in Theorem 1.2 on the non-singularity of the $m \times m$ minors of the matrix $(a_{j,i})_{j,i}$. Indeed, suppose that the equation $x_1 - x_2 = 0$ was part of the system ($\star$) or could be obtained as a linear combination of the equations in ($\star$). Then clearly ($\star$) does not have any solutions $(x_1, \dots, x_k)$ with distinct $x_1, \dots, x_k$. It may potentially be possible to weaken the non-singularity assumption in a way that excludes such situations (see also the discussion in Sect. 5), but this example shows that some assumption is necessary.

The proof of Theorem 1.2 relies on probabilistic subspace sampling arguments and combinatorial ideas, as well as on a new way to apply the slice rank polynomial method. A key step for proving Theorem 1.2 will be to show the following theorem.

Theorem 1.3 *For any fixed integers $m \ge 1$, $r \ge 2$ and $k \ge 2m + r - 1$ and a fixed prime p, there exist constants $C^{\text{rank}}_{p,m,k,r} \ge 1$ and $1 \le \Gamma^{\text{rank}}_{p,m,k,r} < p$ such that the following holds: For any coefficients $a_{j,i} \in \mathbb{F}_p$ for $j = 1, \dots, m$ and $i = 1, \dots, k$ with $a_{j,1} + \dots + a_{j,k} = 0$ for $j = 1, \dots, m$, for any non-negative integer n and any subset $A \subseteq \mathbb{F}_p^n$ of size $|A| > C^{\text{rank}}_{p,m,k,r} \cdot (\Gamma^{\text{rank}}_{p,m,k,r})^n$, the system ($\star$) has a solution $(x_1, \dots, x_k) \in A^k$ such that the subspace* $\operatorname{span}(x_1, \dots, x_k) \subseteq \mathbb{F}_p^n$ *spanned by the vectors $x_1, \dots, x_k$ has dimension* $\dim \operatorname{span}(x_1, \dots, x_k) \ge r$.

Theorem 1.3 can be viewed as a “high-rank” generalization of Tao’s slice rank method result in Theorem 1.1 (where one wants to find a solution $(x_1, \dots, x_k) \in A^k$ of high rank). Indeed, taking $r = 2$ in Theorem 1.3 implies Theorem 1.1. Also note that unlike in Theorem 1.2, in Theorem 1.3 we do not require any genericity assumption on the matrix $(a_{j,i})_{j,i}$.

Organization In the next section we will deduce Theorem 1.2 from Theorem 1.3. Section 3 contains some background on the slice rank polynomial method (including a proof of Theorem 1.1), and discusses the new way in which it will be used in the proof of Theorem 1.3. The proof of Theorem 1.3 is in Sect. 4. We finish with some concluding remarks in Sect. 5.

Notation. As usual, we write $[k] = \{1, \dots, k\}$. For a subset $I \subseteq [k]$, an I*-tuple* of vectors in $\mathbb{F}_p^n$ is a tuple $(x_i \mid i \in I)$ indexed by the set I (with $x_i \in \mathbb{F}_p^n$ for all $i \in I$). For example, a $[k]$-tuple is simply a k-tuple $(x_1, \dots, x_k)$ and a $\{1, 2, 4\}$-tuple is a tuple (x_1, x_2, x_4).

For a vector space V and a subspace U, we can consider the quotient space V/U. We denote the projection of a vector $x \in V$ onto this quotient space by $\operatorname{proj}_{V/U}(x)$. Note that $\operatorname{proj}_{V/U}(x)$ is a vector in V/U and it is non-zero if and only if $x \notin U$.

2 Proof of Theorem 1.2 assuming Theorem 1.3

We deduce Theorem 1.2 from Theorem 1.3 using an inductive argument. More precisely, we will show the following theorem by induction on ℓ (where the base case $\ell = m + 1$ will be obtained from Theorem 1.3, and the final case $\ell = k$ will give the desired statement in Theorem 1.2).

Theorem 2.1 *For any fixed integers $m \geq 1$ and $k \geq 3m$ and $m + 1 \leq \ell \leq k$ as well as a fixed prime p, there exist constants $C \geq 1$ and $0 < c \leq 1$ such that the following holds: Let $a_{j,i} \in \mathbb{F}_p$ for $j = 1, \dots, m$ and $i = 1, \dots, k$ be coefficients with $a_{j,1} + \cdots + a_{j,k} = 0$ for $j = 1, \dots, m$ such that every $m \times m$ minor of the $m \times k$ matrix $(a_{j,i})_{j,i}$ is non-singular. Then for any non-negative integer n and any subset $A \subseteq \mathbb{F}_p^n$ of size $|A| > C \cdot p^{(1-c)n}$, the system $(\star)$ has a solution $(x_1, \dots, x_k) \in A^k$ such that* $\dim \operatorname{span}(x_1, \dots, x_k) \geq m + 1$ *and such that among $x_1, \dots, x_k \in A$ there are at least ℓ distinct vectors.*

Note that Theorem 1.2 follows from Theorem 2.1 for $\ell = k$. Indeed, taking $C_{p,m,k} = C$ and $\Gamma^*_{p,m,k} = p^{1-c} < p$ (for the constants C and c obtained in Theorem 2.1 for $\ell = k$), we can see that every subset $A \subseteq \mathbb{F}_p^n$ of size $|A| > C_{p,m,k} \cdot (\Gamma^*_{p,m,k})^n = C \cdot p^{(1-c)n}$ contains a solution $(x_1, \dots, x_k) \in A^k$ to the system $(\star)$ such that among $x_1, \dots, x_k \in A$ there are at least k distinct vectors (meaning that the vectors $x_1, \dots, x_k \in A$ are all distinct).

In the proof of Theorem 2.1 we will use a probabilistic subspace sampling argument. This will require the following easy lemma.

Lemma 2.2 *Let $n > 0$ and $0 \leq d \leq n$ be integers, and let $y_1, \dots, y_s \in \mathbb{F}_p^n$ be linearly independent vectors in $\mathbb{F}_p^n$. Then for a uniformly random d-dimensional subspace $V \subseteq \mathbb{F}_p^n$, we have*

$$\mathbb{P}[y_1, \dots, y_s \in V] = \frac{p^d - 1}{p^n - 1} \cdot \frac{p^d - p}{p^n - p} \cdots \frac{p^d - p^{s-1}}{p^n - p^{s-1}} \leq \left(\frac{p^d}{p^n}\right)^s$$

Proof Let us first check the inequality between the term in the middle and the term on the right-hand side. If $s > d$, then the middle term is zero and therefore the inequality is true. If $s \leq d$, then the s factors of the middle term are all positive and each of them is at most $(p^d - 1)/(p^n - 1) \leq p^d/p^n$, which also implies the inequality.

Let us now prove by induction on s that the probability on the left-hand side equals the term in the middle. For the base case $s = 1$, note that $y_1 \neq 0$ and that each of the $p^n - 1$ non-zero vectors in $\mathbb{F}_p^n$ is equally likely to be contained in V. Since V always contains exactly $p^d - 1$ non-zero vectors, we can conclude that $\mathbb{P}[y_1 \in V] = (p^d - 1)/(p^n - 1)$ as desired.

Now suppose that $s \geq 2$ and that we have already proved

$$\mathbb{P}[y_1, \dots, y_{s-1} \in V] = \frac{p^d - 1}{p^n - 1} \cdot \frac{p^d - p}{p^n - p} \cdots \frac{p^d - p^{s-2}}{p^n - p^{s-2}}.$$

Then it suffices to show that $\mathbb{P}[y_s \in V \mid y_1, \dots, y_{s-1} \in V] = (p^d - p^{s-1})/(p^n - p^{s-1})$. Note that by the assumption on $y_1, \dots, y_s$ being linearly independent, the vector y_s does not lie in the span of $y_1, \dots, y_{s-1}$. Furthermore, this span is $(s-1)$-dimensional and therefore consists of exactly p^{s-1} vectors. Hence there are exactly $p^n - p^{s-1}$ vectors outside $\operatorname{span}(y_1, \dots, y_s)$ and each of these vectors is equally likely to be contained in V when conditioning on the event $y_1, \dots, y_{s-1} \in V$. Since under this conditioning, V always contains exactly $p^d - p^{s-1}$ vectors outside $\operatorname{span}(y_1, \dots, y_s)$, we can conclude that $\mathbb{P}[y_s \in V \mid y_1, \dots, y_{s-1} \in V] = (p^d - p^{s-1})/(p^n - p^{s-1})$ as desired.

Let us now prove Theorem 2.1 by induction on ℓ, assuming that Theorem 1.3 is true.

Proof of Theorem 2.1 The base case $\ell = m + 1$ follows easily from Theorem 1.3 for $r = m + 1$. Indeed, we have $k \geq 3m = 2m + r - 1$, so there are constants $C^{\text{rank}}_{p,m,k,m+1} \geq 1$ and $1 \leq \Gamma^{\text{rank}}_{p,m,k,m+1} < p$ such that the statement in Theorem 1.3 holds. Now, let $C = C^{\text{rank}}_{p,m,k,m+1}$ and let $0 < c \leq 1$ be such that $p^{1-c} = \Gamma^{\text{rank}}_{p,m,k,m+1}$. Then for any subset $A \subseteq \mathbb{F}_p^n$ of size $|A| > C \cdot p^{(1-c)n} = C^{\text{rank}}_{p,m,k,m+1} \cdot (\Gamma^{\text{rank}}_{p,m,k,m+1})^n$, the system $(\star)$ has a solution $(x_1, \dots, x_k) \in A^k$ such that $\dim \operatorname{span}(x_1, \dots, x_k) \geq m + 1$. Note that the condition $\dim \operatorname{span}(x_1, \dots, x_k) \geq m+1$ automatically implies that there must be at least $\ell = m + 1$ distinct vectors among $x_1, \dots, x_k$. This proves Theorem 2.1 for $\ell = m + 1$.

Now let us assume that $m + 2 \leq \ell \leq k$ and that Theorem 2.1 holds for $\ell - 1$ with constants $C' \geq 1$ and $0 < c' \leq 1$. We will prove that then Theorem 2.1 also holds for ℓ. Let us take

$$c = \frac{1}{(m/c') + 1}$$

and note that

$$\begin{aligned} 1-(m+1)c &= \frac{(m/c')+1-(m+1)}{(m/c')+1} \\ &= m\cdot\frac{(1/c')-1}{(m/c')+1} = (1-c')\cdot\frac{(m/c')}{(m/c')+1} = (1-c')(1-c). \end{aligned} \tag{2.1}$$

Our goal is to show that the statement in Theorem 2.1 holds for some constant $C \geq 1$ only depending on m, k, ℓ and p. By making the constant C sufficiently large we may assume that $n \geq m/(1-c)$. We may furthermore assume that $0 \notin A$ (otherwise we can delete the zero-vector from A and account for that by increasing the constant C by 1).

Hence it suffices to prove that for any $n \geq m/(1-c)$ and any subset $A \subseteq \mathbb{F}_p^n$ of size $|A| > 2p(C' + k^2 2^k) \cdot p^{(1-c)n}$ with $0 \notin A$, the system ($\star$) has a solution $(x_1, \ldots, x_k) \in A^k$ satisfying the conditions in Theorem 2.1.

So let us assume for contradiction that $n \geq m/(1-c)$ and that $A \subseteq \mathbb{F}_p^n$ is a subset of size

$$|A| > 2p(C' + k^2 2^k) \cdot p^{(1-c)n}$$

with $0 \notin A$, in which we cannot find a solution $(x_1, \ldots, x_k) \in A^k$ to the system ($\star$) with $\dim \operatorname{span}(x_1, \ldots, x_k) \geq m+1$ and such that among $x_1, \ldots, x_k \in A$ there are at least ℓ distinct vectors.

The following definition of an I-interesting I-tuple plays a crucial role for the proof. Roughly speaking, our strategy will be to sample a random subspace $V \subseteq \mathbb{F}_p^n$ of a suitably chosen dimension and to find a fairly large subset $A^* \subseteq A \cap V$ which does not contain any I-interesting I-tuple. On the other hand, given that $A^* \subseteq V$ is a large subset of the vector space V, by the induction hypothesis for $\ell - 1$, the set A^* needs to contain a solution $(x_1, \ldots, x_k) \in (A^*)^k$ to the system ($\star$) with $\dim \operatorname{span}(x_1, \ldots, x_k) \geq m+1$ and such that among $x_1, \ldots, x_k \in A$ there are at least $\ell - 1$ distinct vectors. We will see that having such a solution $(x_1, \ldots, x_k) \in (A^*)^k$ will actually force the set A^* to contain some I-interesting I-tuple. This contradiction will finish the proof of the induction step. □

Definition 2.3 For a subset $I \subseteq [k]$ of size $|I| = m+1$, let us say that an I-tuple $(x_i \mid i \in I)$ of vectors $x_i \in A$ for $i \in I$ is *I-interesting* if the vectors x_i for $i \in I$ are linearly independent and the I-tuple $(x_i \mid i \in I)$ can be extended to a solution $(x_1, \ldots, x_k) \in A^k$ to the system ($\star$) such that among the $k - m - 1$ vectors x_j for $j \in [k] \setminus I$ there are at least $\ell - m - 1$ distinct vectors.

The following claim states, roughly speaking, that every solution $(x_1, \ldots, x_k) \in A^k$ satisfying the conditions in the induction hypothesis for $\ell - 1$ must contain an I-interesting I-tuple.

Claim 2.4 *Let $(x_1, \ldots, x_k) \in A^k$ be a solution to the system ($\star$) with* $\dim \operatorname{span}$ $(x_1, \ldots, x_k) \geq m+1$ *and such that among $x_1, \ldots, x_k \in A$ there are exactly $\ell - 1$ distinct vectors. Then there is a subset $I \subseteq [k]$ of size $|I| = m+1$ such that the I-tuple $(x_i \mid i \in I)$ is I-interesting.*

Proof Since $\ell - 1 \leq k - 1$, at least one vector must be repeated among $x_1, \dots, x_k$. So let $a, b \in [k]$ be distinct indices such that $x_a = x_b$. Note that $x_a \neq 0$ by our assumption that $0 \notin A$. Hence, since $\dim \operatorname{span}(x_1, \dots, x_k) \geq m + 1$, we can extend the single vector x_a to a list of $m + 1$ linearly independent vectors chosen among $x_1, \dots, x_k$. In other words, we can find a subset $I \subseteq [k]$ of size $|I| = m + 1$ with $a \in I$ such that the vectors x_i for $i \in I$ are linearly independent (and therefore in particular distinct). It now suffices to check that among the $k - m - 1$ vectors x_j for $j \in [k]\backslash I$ there are at least $\ell - m - 1$ distinct vectors. Since $x_a = x_b$ and the vectors x_i for $i \in I$ are distinct, we have $b \in [k]\backslash I$. Recall that among $x_1, \dots, x_k \in A$ there are exactly $\ell - 1$ distinct vectors. When omitting the $m + 1$ vectors x_i with $i \in I$, at most m of these $\ell - 1$ distinct vectors disappear (since the vector $x_a = x_b$ remains even though it is deleted once, and at most m other vectors get deleted). Hence there are indeed at least $\ell - m - 1$ distinct vectors among the vectors x_j for $j \in [k]\backslash I$. □

On the other hand, our assumption on the set A implies the following structural property for certain solutions $(x_1, \dots, x_k) \in A^k$ to $(\star)$ containing an I-interesting I-tuple.

Claim 2.5 *Let $I \subseteq [k]$ be a subset of size $|I| = m + 1$, and suppose that $(x_1, \dots, x_k) \in A^k$ is a solution to the system $(\star)$ such that the I-tuple $(x_i \mid i \in I)$ is I-interesting and such that there are at least $\ell - m - 1$ distinct vectors among x_j for $j \in [k]\backslash I$. Then at least one of the vectors x_i for $i \in I$ also occurs among the vectors x_j for $j \in [k]\backslash I$.*

Proof By Definition 2.3, the vectors x_i for $i \in I$ must be linearly independent (and therefore in particular distinct). Now, $\dim \operatorname{span}(x_1, \dots, x_k) \geq \dim \operatorname{span}(x_i \mid i \in I) = m + 1$ and by our assumption on the set A this means that there cannot be ℓ distinct vectors among $x_1, \dots, x_k$. Recall that there are at least $\ell - m - 1$ distinct vectors among x_j for $j \in [k]\backslash I$ and exactly $m + 1$ distinct vectors among x_i for $i \in I$. If these two lists of vectors were disjoint, then we would obtain at least ℓ distinct vectors among $x_1, \dots, x_k$. Hence some vector must occur both among x_i for $i \in I$ and among x_j for $j \in [k]\backslash I$. □

Using Claim 2.5, we can show the following upper bound on the number of I-interesting I-tuples for any subset I. This bound will enable us to perform the desired subspace sampling argument, obtaining a subset $A^* \subseteq A$ without any I-interesting I-tuples.

Lemma 2.6 *For each subset $I \subseteq [k]$ of size $|I| = m + 1$, there are at most $k^2 \cdot p^{mn}$ different I-interesting I-tuples $(x_i \mid i \in I)$.*

Proof Let us fix a subset $I \subseteq [k]$ of size $|I| = m + 1$. For each I-interesting I-tuple $(x_i \mid i \in I)$, let us consider the sums $\sum_{i \in I} a_{t,i} x_i$ for $t = 1, \dots, m$ (in other words, we consider the contributions of the vectors x_i with $i \in I$ to the left-hand sides of the equations in the system $(\star)$). It suffices to prove that for any choice of $b_1, \dots, b_m \in \mathbb{F}_p^n$ there can be at most k^2 different I-interesting I-tuples $(x_i \mid i \in I)$ with $\sum_{i \in I} a_{t,i} x_i = b_t$ for $t = 1, \dots, m$. Indeed, summing over all $(p^n)^m = p^{mn}$ choices for $b_1, \dots, b_m$ would then give the desired bound on the total number of I-interesting I-tuples.

So let us fix $b_1, \dots, b_m \in \mathbb{F}_p^n$, and suppose that there are more than k^2 different I-interesting I-tuples $(x_i \mid i \in I)$ with $\sum_{i \in I} a_{t,i} x_i = b_t$ for $t = 1, \dots, m$. Let $\mathcal{T}$ be the set of all such I-interesting I-tuples $(x_i \mid i \in I)$, then $|\mathcal{T}| > k^2$.

We claim that for any two distinct $(x_i \mid i \in I), (x_i' \mid i \in I) \in \mathcal{T}$ we must have $x_h \neq x_h'$ for all $h \in I$. Suppose that we had $x_h = x_h'$ for some $h \in I$. Note that then

$$\sum_{i \in I \setminus \{h\}} a_{t,i} x_i = b_t - a_{t,h} x_h = b_t - a_{t,h} x_h' = \sum_{i \in I \setminus \{h\}} a_{t,i} x_i'$$

for $t = 1, \dots, m$. However, the $m \times m$ matrix $(a_{t,i})_{t \in [m], i \in I \setminus \{h\}}$ is non-singular by the assumption in Theorem 2.1. Hence we can conclude that $x_i = x_i'$ for all $i \in I \setminus \{h\}$ and hence $(x_i \mid i \in I) = (x_i' \mid i \in I)$, which is a contradiction. So for any two distinct $(x_i \mid i \in I), (x_i \mid i \in I) \in \mathcal{T}$ we must indeed have $x_h \neq x_h'$ for all $h \in I$.

Next, we claim that we can find vectors $x_j \in A$ for $j \in [k] \setminus I$ such that $\sum_{j \in [k] \setminus I} a_{t,j} x_j = -b_t$ for $t = 1, \dots, m$ and such that there are at least $\ell - m - 1$ distinct vectors among x_j for $j \in [k] \setminus I$. Indeed, consider any $(x_i \mid i \in I) \in \mathcal{T}$. By Definition 2.3 we can extend $(x_i \mid i \in I)$ to a solution $(x_1, \dots, x_k) \in A^k$ to the system $(\star)$ such that there are at least $\ell - m - 1$ distinct vectors among x_j for $j \in [k] \setminus I$. Now, for every $t = 1, \dots, m$ we have $\sum_{j \in [k] \setminus I} a_{t,j} x_j = -\sum_{i \in I} a_{t,i} x_i = -b_t$, as desired.

So let us now fix a choice of vectors $x_j \in A$ for $j \in [k] \setminus I$ such that $\sum_{j \in [k] \setminus I} a_{t,j} x_j = -b_t$ for $t = 1, \dots, m$ and such that there are at least $\ell - m - 1$ distinct vectors among x_j for $j \in [k] \setminus I$. Then for every $(x_i \mid i \in I) \in \mathcal{T}$, the k-tuple $(x_1, \dots, x_k) \in A^k$ satisfies

$$x_{t,1} x_1 + \dots + x_{t,k} x_k = \sum_{i \in I} a_{t,i} x_i + \sum_{j \in [k] \setminus I} a_{t,j} x_j = b_t + (-b_t) = 0$$

for $t = 1, \dots, m$. In other words, for every $(x_i \mid i \in I) \in \mathcal{T}$, the k-tuple $(x_1, \dots, x_k) \in A^k$ is a solution to the system $(\star)$. By Claim 2.5 this means that at least one of the vectors x_i for $i \in I$ must also occur among the vectors x_j for $j \in [k] \setminus I$. Hence for every $(x_i \mid i \in I) \in \mathcal{T}$ one of the $|I| = m + 1 \leq k$ vectors in the I-tuple $(x_i \mid i \in I)$ must be one of the fixed $k - |I| \leq k$ vectors in the set $\{x_j \mid j \in [k] \setminus I\}$. As $|\mathcal{T}| > k^2$, by the pigeonhole principle this implies that there must be two distinct I-tuples $(x_i \mid i \in I), (x_i' \mid i \in I) \in \mathcal{T}$ with $x_h = x_h'$ for some index $h \in I$. But this is a contradiction to what we showed above. This completes the proof of the lemma. □

Now, let

$$d = \lfloor (1 - c) n / m \rfloor,$$

and note that $d \geq 1$ by our assumption that $n \geq m/(1 - c)$.

Let us consider a uniformly random d-dimensional subspace $V \subseteq \mathbb{F}_p^n$. The following two claims give useful bounds for the expected number of vectors in $A \cap V$ and the expected number of I-interesting I-tuples in V.

Claim 2.7 $\mathbb{E}[|A \cap V|] > (C' + k^2 2^k) \cdot p^{(1-(m+1)c)n/m}$.

Proof Recall that we assumed $0 \notin A$. So by Lemma 2.2 (applied with $s = 1$), for each vector $x \in A$ we have

$$\mathbb{P}[x \in V] = \frac{p^d - 1}{p^n - 1} \geq \frac{p^d/2}{p^n} \geq \frac{1}{2p} \cdot \frac{p^{(1-c)n/m}}{p^n}.$$

Using $|A| > 2p(C' + k^2 2^k) \cdot p^{(1-c)n}$, we obtain

$$\begin{aligned}\mathbb{E}[|A \cap V|] \geq \frac{1}{2p} \cdot \frac{p^{(1-c)n/m}}{p^n} \cdot |A| &> (C' + k^2 2^k) \cdot p^{((1-c)n/m)-cn}\\ &= (C' + k^2 2^k) \cdot p^{(1-(m+1)c)n/m},\end{aligned}$$

as desired. □

Claim 2.8 *For each subset $I \subseteq [k]$ of size $|I| = m + 1$, the expected number of I-interesting I-tuples $(x_i \mid i \in I)$ with $x_i \in V$ for all $i \in I$ is at most $k^2 \cdot p^{(1-(m+1)c)n/m}$.*

Proof Recall from Lemma 2.6 that the number of I-interesting I-tuples $(x_i \mid i \in I)$ is at most $k^2 \cdot p^{mn}$. For each of these tuples $(x_i \mid i \in I)$ the $m + 1$ vectors x_i with $i \in I$ are linearly independent (see Definition 2.3), so by Lemma 2.2 we have

$$\mathbb{P}[x_i \in V \text{ for all } i \in I] \leq \left(\frac{p^d}{p^n}\right)^{m+1} \leq \left(\frac{p^{(1-c)n/m}}{p^n}\right)^{m+1}.$$

Thus, all in all the expected number of I-interesting I-tuples $(x_i \mid i \in I)$ with $x_i \in V$ for all $i \in I$ is at most

$$\left(\frac{p^{(1-c)n/m}}{p^n}\right)^{m+1} \cdot k^2 \cdot p^{mn} = k^2 \cdot \frac{p^{(1-c)n(m+1)/m}}{p^n} = k^2 \cdot p^{(1-(m+1)c)n/m},$$

as claimed. □

Let Z be the total number of I-interesting I-tuples $(x_i \mid i \in I)$ with $x_i \in V$ for all $i \in I$, summed over all subsets $I \subseteq [k]$ of size $|I| = m + 1$. Since there are only $\binom{k}{m+1} \leq 2^k$ choices for I, Claim 2.8 implies that $\mathbb{E}[Z] \leq 2^k \cdot k^2 \cdot p^{(1-(m+1)c)n/m}$. Hence together with Claim 2.7 we obtain

$$\begin{aligned}\mathbb{E}[|A \cap V| - Z] &> (C' + k^2 2^k) \cdot p^{(1-(m+1)c)n/m} - 2^k \cdot k^2 \cdot p^{(1-(m+1)c)n/m}\\ &= C' \cdot p^{(1-(m+1)c)n/m}.\end{aligned}$$

Thus, for the rest of this proof, we can fix an outcome for the random d-dimensional subspace $V \subseteq \mathbb{F}_p^n$ for which we have $|A \cap V| - Z > C' \cdot p^{(1-(m+1)c)n/m}$.

We can now define a subset $A^* \subseteq A \cap V$ by deleting one vector from each I-interesting I-tuple $(x_i \mid i \in I)$ with $x_i \in V$ for all $i \in I$ (for all subsets $I \subseteq [k]$ of size $|I| = m + 1$). Then

$$|A^*| \geq |A \cap V| - Z > C' \cdot p^{(1-(m+1)c)n/m}$$

and there do not exist any I-interesting I-tuples $(x_i \mid i \in I)$ with $x_i \in A^*$ for all $i \in I$ (for any $I \subseteq [k]$ of size $|I| = m + 1$).

On the other hand, (2.1) yields

$$|A^*| > C' \cdot p^{(1-(m+1)c)n/m} = C' \cdot p^{(1-c')(1-c)n/m} \geq C' \cdot p^{(1-c')d}.$$

As $V \cong \mathbb{F}_p^d$, we can interpret $A^* \subseteq V$ as a subset of $\mathbb{F}_p^d$ of size $|A^*| > C' \cdot p^{(1-c')d}$. By the induction hypothesis (which states that Theorem 2.1 holds for $\ell - 1$ with the constants C' and c'), we can conclude that there must be a solution $(x_1, \dots, x_k) \in (A^*)^k \subseteq A^k$ to the system $(\star)$ such that $\dim \operatorname{span}(x_1, \dots, x_k) \geq m + 1$ and such that among $x_1, \dots, x_k \in A$ there are at least $\ell - 1$ distinct vectors. By our assumption on A, there must be exactly $\ell - 1$ distinct vectors among $x_1, \dots, x_k \in A$. But now Claim 2.4 implies that there exists a subset $I \subseteq [k]$ of size $|I| = m + 1$ such that the I-tuple $(x_i \mid i \in I)$ is I-interesting. Since $x_i \in A^*$ for all $i \in I$, this is a contradiction. This finally finishes the proof of Theorem 2.1. □

3 Background on the slice rank polynomial method

In this section, we will give some background on Tao's slice rank polynomial method, and explain the way in which we will use it in the proof of Theorem 1.3. This involves utilizing a lemma due to Sawin and Tao [26] (see Lemma 3.5 below), which gives lower bounds on the slice rank of non-diagonal tensors of a certain form (see Corollary 3.6 below). This leads to Corollary 3.7 below, which will then be used as a black box in the proof of Theorem 1.3.

Statements like Lemma 3.5 and Corollary 3.6 are hard to use in combinatorial applications, since in practice it is often difficult to find a (non-diagonal) tensor with the particular structure required in these statements. In fact, to the author's knowledge, all previous combinatorial applications of the slice rank polynomial method relied on diagonal tensors, meaning that this paper is the first to find a way to exploit the strength of Lemma 3.5 for non-diagonal tensors. Doing so requires a careful combinatorial setup and analysis, see Sect. 4.

Let us start by defining the notion of slice rank, which was introduced by Tao [28] in a blog post and later given this name. In our setting it will be most convenient to think of k-dimensional tensors as functions $f : [L]^k \to \mathbb{F}$ for some integer $L \geq 0$ and some field $\mathbb{F}$. By considering a vector space V over $\mathbb{F}$ with basis $v_1, \dots, v_L$, one can identify such a function with the element $\sum_{\ell_1,\dots,\ell_k \in [L]} f(\ell_1, \dots, \ell_k)\, v_{\ell_1} \otimes \cdots \otimes v_{\ell_k}$ in the k-fold tensor product $V \otimes \cdots \otimes V$. Yet another way to think of such a function $f : [L]^k \to \mathbb{F}$ is to think of a k-dimensional hypermatrix of size $L \times \cdots \times L$ with entries in $\mathbb{F}$. However, in our discussion we will formulate everything just in terms of functions $f : [L]^k \to \mathbb{F}$.

Definition 3.1 (Tao) Let $L \geq 1$ and $k \geq 2$ be integers, and let $\mathbb{F}$ be a field. A function $f : [L]^k \to \mathbb{F}$ has *slice rank 1*, if it can be expressed in the form

$$f(\ell_1, \dots, \ell_k) = g(\ell_i) \cdot h(\ell_1, \dots, \ell_{i-1}, \ell_{i+1}, \dots, \ell_k)$$

for some $i \in [k]$ and non-zero functions $g : [L] \to \mathbb{F}$ and $h : [L]^{k-1} \to \mathbb{F}$. The *slice rank* of an arbitrary function $f : [L]^k \to \mathbb{F}$ is defined to be the minimum number r such that f can be written as the sum of r functions of slice rank 1.

In other words, a function $f : [L]^k \to \mathbb{F}$ is defined to have slice rank 1, if it can be written as the product of a function depending on just one of the k variables and a function depending on the remaining $k-1$ variables. Note that this notion differs from the standard definition of the rank of a tensor, where a function of rank 1 is required to be a product of k functions depending on one variable each. The slice rank of a function is always at most as large as its rank according to the standard definition. Also note that the slice rank of any function $f : [L]^k \to \mathbb{F}$ is at most L. Indeed, for each $\ell \in [L]$, we can take g_ℓ to be the indicator function of ℓ and define $h_\ell(\ell_2, \dots, \ell_k) = f(\ell, \ell_2, \dots, \ell_k)$. Then we have $f(\ell_1, \dots, \ell_k) = \sum_{\ell \in [L]} g_\ell(\ell_i) \cdot h_\ell(\ell_2, \dots, \ell_k)$, which shows that $f : [L]^k \to \mathbb{F}$ has slice rank at most L.

Tao's slice rank polynomial method [28] combines his notion of slice rank as in Definition 3.1 with an easy but very powerful polynomial factoring argument. This argument first appeared in work of Croot, Lev, and Pach [4] on subsets $\mathbb{Z}_4^n$ without 3-term arithmetic progressions, and was then used again in Ellenberg and Gijswijt's breakthrough [6] on the cap-set problem bounding the size of 3-term-progression-free subsets of $\mathbb{F}_3^n$ (and more generally $\mathbb{F}_p^n$ for any fixed p). On his blog, Tao [28] gave a reformulation of the proof of Ellenberg and Gijswijt using the notion of the slice rank (which he introduced in this proof). His reformulation still uses the same polynomial factoring argument that is also at the heart of the proofs of Ellenberg-Gijswijt and of Croot-Lev-Pach in the $\mathbb{Z}_4^n$ setting.

In our context of studying solutions to linear systems of equations, this polynomial factoring argument gives the following lemma. For integers $m \geq 1$ and $k \geq 2m+1$ and a prime p, let us define

$$\Gamma_{p,m,k} = \min_{0<z\leq 1} \frac{1 + z + \cdots + z^{p-1}}{z^{(p-1)m/k}} < p. \tag{3.1}$$

It is not hard to see that this minimum exists. Furthermore, at $z = 1$ the function on the left-hand size has value p and positive derivative (since $k \geq 2m+1$), which implies that indeed $\Gamma_{p,m,k} < p$. It is also easy to see that $\Gamma_{p,m,k} \geq 1$, so we have $1 \leq \Gamma_{p,m,k} < p$.

Lemma 3.2 *(Croot-Lev-Pach, Tao) Suppose we are given a linear system of equations with coefficients in $\mathbb{F}_p$ and constant terms in $\mathbb{F}_p^n$, consisting of $m \geq 1$ equations* *in $k \geq 2m+1$ variables. Then for any integer L and any vectors $x_i^{(\ell)} \in \mathbb{F}_p^n$ for $i = 1, \dots, k$ and $\ell = 1, \dots, L$, the function $f : [L]^k \to \mathbb{F}_p$ defined by*

$$f(\ell_1, \dots, \ell_k) = \begin{cases} 1 & \textit{if } (x_1^{(\ell_1)}, \dots, x_k^{(\ell_k)}) \textit{ is a solution to the given system of equations} \\ 0 & \textit{otherwise.} \end{cases}$$

has slice rank at most $k \cdot (\Gamma_{p,m,k})^n$.

Proof Let the given linear system of equations be of the form

$$\begin{aligned} a_{1,1}x_1 + \cdots + a_{1,k}x_k &= b_1 \\ &\vdots \\ a_{m,1}x_1 + \cdots + a_{m,k}x_k &= b_m, \end{aligned}$$

where $a_{j,i} \in \mathbb{F}_p$ for $j = 1, \dots, m$ and $i = 1, \dots, k$ and $b_j \in \mathbb{F}_p^n$ for $j = 1, \dots, m$. Furthermore, for any vector $x \in \mathbb{F}_p^n$, let us write $x(1), \dots, x(n)$ for the coordinates of x (which are elements of $\mathbb{F}_p$). In particular, $x_i^{(\ell)}(1), \dots, x_i^{(\ell)}(n)$ are the coordinates of $x_i^{(\ell)} \in \mathbb{F}_p^n$ for $i = 1, \dots, k$ and $\ell = 1, \dots, L$ and $b_j(1), \dots, b_j(n)$ are the coordinates of $b_j \in \mathbb{F}_p^n$ for $j = 1, \dots, m$.

Now, we claim that

$$f(\ell_1, \dots, \ell_k) = \prod_{j=1}^{m} \prod_{s=1}^{n} \left(1 - (a_{j,1}x_1^{(\ell_1)}(s) + \cdots + a_{j,k}x_k^{(\ell_k)}(s) - b_j(s))^{p-1}\right) \tag{3.2}$$

for all $\ell_1, \dots, \ell_k \in [L]$. Indeed, if $\ell_1, \dots, \ell_k \in [L]$ are such that $(x_1^{(\ell_1)}, \dots, x_k^{(\ell_k)})$ is a solution to the system of equations above, then we have $a_{j,1}x_1^{(\ell_1)} + \cdots + a_{j,k}x_k^{(\ell_k)} - b_j = 0$ for all $j = 1, \dots, m$ and consequently $a_{j,1}x_1^{(\ell_1)}(s) + \cdots + a_{j,k}x_k^{(\ell_k)}(s) - b_j(s) = 0$ for all $j = 1, \dots, m$ and $s = 1, \dots, n$. This means that all factors on the right-hand side of (3.2) are equal to 1, and therefore the product is indeed equal to $f(\ell_1, \dots, \ell_k) = 1$. In the other case, where $(x_1^{(\ell_1)}, \dots, x_k^{(\ell_k)})$ is not a solution to the system of equations, there must be some $j \in \{1, \dots, m\}$ and $s \in \{1, \dots, n\}$ such that $a_{j,1}x_1^{(\ell_1)}(s) + \cdots + a_{j,k}x_k^{(\ell_k)}(s) - b_j(s) \neq 0$. But then we have $(a_{j,1}x_1^{(\ell_1)}(s) + \cdots + a_{j,k}x_k^{(\ell_k)}(s) - b_j(s))^{p-1} = 1$, and so the factor $1 - (a_{j,1}x_1^{(\ell_1)}(s) + \cdots + a_{j,k}x_k^{(\ell_k)}(s) - b_j(s))^{p-1}$ on the right-hand side of (3.2) is 0. Hence the entire product on the right-hand side is 0 and therefore equal to $f(\ell_1, \dots, \ell_k) = 0$. Thus, (3.2) is indeed true.

We can now use the polynomial representation of f in (3.2) to show the desired upper bound on the slice rank of f. Note that the right-hand side of (3.2) is a polynomial of degree $mn(p-1)$ in the kn variables $x_i^{(\ell_i)}(1), \dots, x_i^{(\ell_i)}(n)$ for $i = 1, \dots, k$. Let us imagine that we multiply out the product on the right-hand side of (3.2). Then we can write $f(\ell_1, \dots, \ell_k)$ as a linear combination of monomials in the variables

$x_i^{(\ell_i)}(1), \dots, x_i^{(\ell_i)}(n)$ for $i = 1, \dots, k$, where each monomial has degree at most $mn(p-1)$.

Since in $\mathbb{F}_p$ we have $y^p = y$ for all $y \in \mathbb{F}_p$, we can replace each higher power $(x_i^{(\ell_i)}(s))^d$ with $d \geq p$ by a power $(x_i^{(\ell_i)}(s))^{d'}$ with $d' \in \{1, \dots, p-1\}$ (and $d' \equiv d \mod p-1$). This way we can represent $f(\ell_1, \dots, \ell_k)$ as a linear combination of monomials in the variables $x_i^{(\ell_i)}(1), \dots, x_i^{(\ell_i)}(n)$ for $i = 1, \dots, k$, where each monomial has degree at most $mn(p-1)$ and each individual variable appears with degree at most $p-1$.

For each of the monomials in this representation the total degree is at most $mn(p-1)$, so there must be some $i \in [k]$ such that the monomial has degree at most $mn(p-1)/k$ in the variables $x_i^{(\ell_i)}(1), \dots, x_i^{(\ell_i)}(n)$. Hence each monomial is of the form

$$(x_i^{(\ell_i)}(1))^{d_1} \cdots (x_i^{(\ell_i)}(n))^{d_n} \cdot g(\ell_1, \dots, \ell_{i-1}, \ell_{i+1}, \dots, \ell_k)$$

for some $i \in [k]$, some $d_1, \dots, d_n \in \{0, \dots, p-1\}$ with $d_1 + \dots + d_n \leq mn(p-1)/k$ and some function $g : [L]^{k-1} \to \mathbb{F}_p$ (where g is a monomial in the variables $x_{i'}^{(\ell_{i'})}(1), \dots, x_{i'}^{(\ell_{i'})}(n)$ for $i \in [k]\setminus\{i\}$).

By grouping together the monomials with the same i and the same $d_1, \dots, d_n$ in the above representation, we now obtain a representation of $f(\ell_1, \dots, \ell_k)$ as a sum of terms of the form

$$(x_i^{(\ell_i)}(1))^{d_1} \cdots (x_i^{(\ell_i)}(n))^{d_n} \cdot h(\ell_1, \dots, \ell_{i-1}, \ell_{i+1}, \dots, \ell_k)$$

for some $i \in [k]$, some $d_1, \dots, d_n \in \{0, \dots, p-1\}$ with $d_1 + \dots + d_n \leq mn(p-1)/k$ and some function $h : [L]^{k-1} \to \mathbb{F}_p$ (here, the function h is obtained as a linear combination of the functions g that we previously considered). Note that each such term is a slice rank 1 function. Hence the slice rank of f is at most

$$k \cdot |\{(d_1, \dots, d_n) \in \{0, \dots, p-1\}^n \mid d_1 + \dots + d_n \leq mn(p-1)/k\}|.$$

Thus, using the following claim, we obtain the desired bound for the slice rank of f.

Claim 3.3 $|\{(d_1, \dots, d_n) \in \{0, \dots, p-1\}^n \mid d_1 + \dots + d_n \leq mn(p-1)/k\}| \leq (\Gamma_{p,m,k})^n$.

Proof We need to prove that for a uniformly random choice of $(d_1, \dots, d_n) \in \{0, \dots, p-1\}^n$ we have $\mathbb{P}[d_1 + \dots + d_n \leq mn(p-1)/k] \leq (\Gamma_{p,m,k})^n \cdot p^{-n}$. Indeed, for any $0 < z \leq 1$, by Markov's inequality we have

$$\begin{aligned}
\mathbb{P}[d_1 + \dots + d_n \leq mn(p-1)/k] &\leq \mathbb{P}[z^{d_1+\dots+d_n} \geq z^{mn(p-1)/k}] \\
&\leq \frac{\mathbb{E}[z^{d_1+\dots+d_n}]}{z^{mn(p-1)/k}} = \frac{\mathbb{E}[z^{d_1}] \cdots \mathbb{E}[z^{d_n}]}{z^{mn(p-1)/k}} = \frac{((1+z+\dots+z^{p-1})/p)^n}{z^{mn(p-1)/k}} \\
&= \left(\frac{1+z+\dots+z^{p-1}}{z^{(p-1)m/k}}\right)^n \cdot p^{-n},
\end{aligned}$$

where we used that $d_1, \dots, d_n$ can be viewed as independent uniformly random elements of $\{0, \dots, p-1\}$. Hence

$$\begin{aligned}\mathbb{P}[d_1 + \cdots + d_n \le mn(p-1)/k] &\le \left(\min_{0<z\le 1} \frac{1+z+\cdots+z^{p-1}}{z^{(p-1)m/k}} \right)^n \cdot p^{-n} \\ &= (\Gamma_{p,m,k})^n \cdot p^{-n},\end{aligned}$$

as desired. □

This finishes the proof of Lemma 3.2. □

We remark that the proof of Lemma 3.2 is a straightforward generalization of the corresponding arguments in Tao's blog post [28] (which are for the case $m = 1$ and $k = 3$), apart from the bound for the quantity in Claim 3.3. This bound as well as the proof of Claim 3.3 appeared for example in [2,Proposition 4.12].

As an example of a typical application of Tao's slice rank polynomial method, let us now show how Theorem 1.1 can be deduced from Lemma 3.2 and the following lemma due to Tao [28,Lemma 1] stating that diagonal tensors have large slice rank.

Lemma 3.4 *(Tao) Let $L \ge 1$ and $k \ge 2$ be integers, and let $\mathbb{F}$ be a field. Suppose that $f : [L]^k \to \mathbb{F}$ is a function such that $f(\ell_1, \dots, \ell_k) \ne 0$ whenever $\ell_1 = \cdots = \ell_k$ and such that $f(\ell_1, \dots, \ell_k) = 0$ whenever $\ell_1, \dots, \ell_k \in [L]$ are not all equal. Then the slice rank of f is equal to L.*

By combining Lemmas 3.2 and 3.4 one can obtain Theorem 1.1. This proof appeared in Tao's blog post [28] in the special case of $m = 1$, $k = 3$ and ($\star$) being the single equation $x_1 - 2x_2 + x_3 = 0$, which corresponds to 3-term arithmetic progressions (in the same blog post he also introduced the notion of slice rank and proved Lemma 3.4).

Proof Theorem 1.1 Let $A \subseteq \mathbb{F}_p^n$ be such that every solution $(x_1, \dots, x_k) \in A^k$ of the system ($\star$) satisfies $x_1 = \cdots = x_k$. Let $L = |A|$, and let $A = \{x^{(1)}, \dots, x^{(L)}\}$. Now, define the function $f : [L]^k \to \mathbb{F}_p$ by setting

$$f(\ell_1, \dots, \ell_k) = \begin{cases} 1 & \text{if } (x^{(\ell_1)}, \dots, x^{(\ell_k)}) \text{ is a solution to } (\star) \\ 0 & \text{otherwise.} \end{cases}.$$

Note that by Lemma 3.2 the slice rank of f is at most $k \cdot (\Gamma_{p,m,k})^n$.

On the other hand, whenever $\ell_1, \dots, \ell_k$ are such that $(x^{(\ell_1)}, \dots, x^{(\ell_k)})$ is a solution to ($\star$), then by our assumption on A we have $x^{(\ell_1)} = \cdots = x^{(\ell_k)}$ and therefore $\ell_1 = \cdots = \ell_k$. Thus, we have $f(\ell_1, \dots, \ell_k) = 0$ whenever $\ell_1, \dots, \ell_k \in [L]$ are not all equal. Furthermore, for all $\ell \in L$, the k-tuple $(x^{(\ell)}, \dots, x^{(\ell)})$ is a solution to ($\star$), since we assumed that $a_{j,1} + \cdots + a_{j,k} = 0$ for $j = 1, \dots, m$. Hence whenever $\ell_1 = \cdots = \ell_k$, we have $f(\ell_1, \dots, \ell_k) = 1 \ne 0$. Thus, by Lemma 3.4 the slice rank of f equals L.

All in all, we obtain $|A| = L \le k \cdot (\Gamma_{p,m,k})^n$. So we have shown that every set $A \subseteq \mathbb{F}_p^n$ with the property that every solution $(x_1, \dots, x_k) \in A^k$ of the system ($\star$) satisfies $x_1 = \cdots = x_k$ must have size $|A| \le k \cdot (\Gamma_{p,m,k})^n$.

This almost gives Theorem 1.1. In order to prove the theorem, we need to prove $|A| \leq (\Gamma_{p,m,k})^n$ instead of the weaker bound $|A| \leq k \cdot (\Gamma_{p,m,k})^n$. We can use a power trick, as follows.

Suppose that $A \subseteq \mathbb{F}_p^n$ is a set with the property that every solution $(x_1, \dots, x_k) \in A^k$ of the system $(\star)$ satisfies $x_1 = \dots = x_k$. Note that for every positive integer m, the set $A^m = A \times \dots \times A \subseteq \mathbb{F}_p^n \times \dots \times \mathbb{F}_p^n = \mathbb{F}_p^{nm}$ also has this property. Therefore, by what we proved above, we must have $|A|^m = |A^m| \leq k \cdot (\Gamma_{p,m,k})^{nm}$ and therefore $|A| \leq k^{1/m} \cdot (\Gamma_{p,m,k})^n$ for every positive integer m. Hence $|A| \leq (\Gamma_{p,m,k})^n$, as desired. □

The proof of Theorem 1.1 follows the typical pattern of applications of the slice rank polynomial method in combinatorics: If one wants to bound the size of a combinatorial structure satisfying certain conditions (here the set $A = \{x^{(1)}, \dots, x^{(L)}\}$), one defines a function f in terms of this structure in such a way that a polynomial factoring argument as in the proof of Lemma 3.2 gives an upper bound on the slice rank of f. On the other hand, the conditions on the combinatorial structure imply that the function f must be of a special form, giving a lower bound for its slice rank in terms of the size of the structure. Combining both bounds, one then obtains an upper bound for the size of the combinatorial structure, as desired.

As mentioned above, so far in all combinatorial applications of the slice rank polynomial method, the "special form" of the function f under consideration was a diagonal tensor as in Lemma 3.4. Not long after Tao's original blog post [28], another post by Sawin and Tao [26] appeared on Tao's blog, proving the following statement [26,Proposition 4] concerning the slice rank of certain non-diagonal tensors.

Lemma 3.5 *(Sawin-Tao) Let $L \geq 1$ and $k \geq 2$ be integers, and let $\mathbb{F}$ be a field. Fix k total orderings $\preceq^1, \dots, \preceq^k$ on the set $[L] = \{1, \dots, L\}$, and consider the resulting product partial order $\preceq$ on $[L]^k$. Suppose that $f : [L]^k \to \mathbb{F}$ is a function such that the set $S = \{(\ell_1, \dots, \ell_k) \in [L]^k \mid f(\ell_1, \dots, \ell_k) \neq 0\} \subseteq [L]^k$ is an antichain with respect to the partial order $\preceq$. Then the slice rank of f equals*

$$\min_{S = S_1 \cup \dots \cup S_k} (|\pi_1(S_1)| + \dots + |\pi_k(S_k)|),$$

where the minimum is taken over all partitions $S = S_1 \cup \dots \cup S_k$ and where $\pi_i : [L]^k \to L$ denotes the projection to the i-th factor for $i = 1, \dots, k$.

The following corollary of Lemma 3.5 gives a slightly more concrete statement about certain tensors having large slice rank.

Corollary 3.6 *Let $L \geq 1$ and $k \geq 2$ be integers, and let $\mathbb{F}$ be a field. Let $[k] = J_1 \cup \dots \cup J_t$ be a partition such that $|J_h| \geq 2$ for $h = 1, \dots, t$. Now, suppose that $f : [L]^k \to \mathbb{F}$ is a function such that $f(\ell, \dots, \ell) \neq 0$ for all $\ell \in [L]$ and such that for any choice of $\ell_1, \dots, \ell_k \in [L]$ with $f(\ell_1, \dots, \ell_k) \neq 0$ we have $|\{\ell_j \mid j \in J_h\}| = 1$ for $h = 1, \dots, t$. Then the slice rank of f is equal to L.*

Proof Let us define total orderings $\preceq^1, \dots, \preceq^k$ on the set $[L]$ as follows. For each $h = 1, \dots, t$, choose one element $j_h \in J_h$ and define $\preceq^{j_h}$ to be the canonical increasing ordering $1 \preceq^{j_h} 2 \preceq^{j_h} \cdots \preceq^{j_h} L$ on $[L]$. Furthermore, choose a different element $i_h \in J_h$ (recall that $|J_h| \geq 2$) and define the total ordering $\preceq^{i_h}$ to be the opposite total ordering on $[L]$, i.e. the ordering $L \preceq^{i_h} L-1 \preceq^{i_h} \cdots \preceq^{i_h} 1$. For all remaining elements $j \in J_h \backslash \{j_h, i_h\}$, define $\preceq^j$ to be an arbitrary total ordering on $[L]$.

We claim that the function $f : [L]^k \to \mathbb{F}$ satisfies the assumption in Lemma 3.5 with respect to the total orderings $\preceq^1, \dots, \preceq^k$ we just defined. Indeed, suppose that $(\ell_1, \dots, \ell_k), (\ell'_1, \dots, \ell'_k) \in [L]^k$ with $f(\ell_1, \dots, \ell_k) \neq 0$ and $f(\ell'_1, \dots, \ell'_k) \neq 0$ are such that $(\ell_1, \dots, \ell_k) \preceq (\ell'_1, \dots, \ell'_k)$ in the product partial order $\preceq$. Then $\ell_i \preceq^i \ell'_i$ for all $i \in [k]$. We will show that we actually have $\ell_i = \ell'_i$ for all $i \in [k]$. So let $i \in [k]$, and let $h \in [t]$ be such that $i \in J_h$. By our assumption on f we have $|\{\ell_j \mid j \in J_h\}| = 1$ and $|\{\ell'_j \mid j \in J_h\}| = 1$, which means that $\ell_i = \ell_{j_h} = \ell_{i_h}$ and $\ell'_i = \ell'_{j_h} = \ell'_{i_h}$ for the elements $j_h, i_h \in J_h$ that we fixed above. Hence $\ell_i = \ell_{j_h} \preceq^{j_h} \ell'_{j_h} = \ell'_i$ in the total ordering $\preceq^{j_h}$, which by the definition of $\preceq^{j_h}$ means that $\ell_i \leq \ell'_i$. On the other hand we have $\ell_i = \ell_{i_h} \preceq^{i_h} \ell'_{i_h} = \ell'_i$ in the total ordering $\preceq^{i_h}$, which by the definition of $\preceq^{i_h}$ means that $\ell_i \geq \ell'_i$. Thus, we can conclude that $\ell_i = \ell'_i$ for all $i \in [k]$, so $(\ell_1, \dots, \ell_k) = (\ell'_1, \dots, \ell'_k)$.

Hence the set $S = \{(\ell_1, \dots, \ell_k) \in [L]^k \mid f(\ell_1, \dots, \ell_k) \neq 0\} \subseteq [L]^k$ is indeed an antichain with respect to the partial order $\preceq$. So Lemma 3.5 applies and we can conclude that the slice rank of f equals

$$\min_{S=S_1 \cup \cdots \cup S_k} \left(|\pi_1(S_1)| + \cdots + |\pi_k(S_k)|\right),$$

where the minimum is taken over all partitions $S = S_1 \cup \cdots \cup S_k$.

We claim that for any partition $S = S_1 \cup \cdots \cup S_k$ we must have $\pi_1(S_1) \cup \cdots \cup \pi_k(S_k) = [L]$. Indeed, for any $\ell \in [L]$, by the assumption $f(\ell, \dots, \ell) \neq 0$ we have $(\ell, \dots, \ell) \in S$ and therefore $(\ell, \dots, \ell) \in S_i$ for some $i \in [k]$. This means that $\ell \in \pi_i(S_i) \subseteq \pi_1(S_1) \cup \cdots \cup \pi_k(S_k)$. Hence $\pi_1(S_1) \cup \cdots \cup \pi_k(S_k) = [L]$ and therefore $|\pi_1(S_1)| + \cdots + |\pi_k(S_k)| \geq |\pi_1(S_1) \cup \cdots \cup \pi_k(S_k)| = L$ for any partition $S = S_1 \cup \cdots \cup S_k$.

Thus, the slice rank of f is at least L. On the other hand, the slice rank of $f : [L]^k \to \mathbb{F}$ is also at most L. Hence the slice rank of f equals L. □

Note that Corollary 3.6 can be viewed as a generalization of Lemma 3.4 for diagonal tensors by considering the partition of $[k]$ into a single set $J_1 = [k]$. The slice rank polynomial method has had many interesting applications in combinatorics (see for example Grochow's survey [14]), but all of them rely on using a diagonal tensor (meaning that they proceed via Lemma 3.4). This paper gives the first combinatorial application that uses the additional strength of Corollary 3.6 for non-diagonal tensors.

Being able to use Corollary 3.6 in the proof of Theorem 1.3 requires new combinatorial ideas and a rather technical inductive setup. In order to not further complicate the presentation of the proof of Theorem 1.3 in Sect. 4, we record the following statement here, which combines Corollary 3.6 with Lemma 3.2. This statement will then be used in the proof of Theorem 1.3 as a black box, so that no discussion of slice rank arguments is required in Sect. 4 anymore.

Corollary 3.7 *Suppose we are given a linear system of equations with coefficients in $\mathbb{F}_p$ and constant terms in $\mathbb{F}_p^n$, consisting of $m \geq 1$ equations in $k \geq 2m+1$ variables. Let $(x_1^{(\ell)}, \dots, x_k^{(\ell)}) \in (\mathbb{F}_p^n)^k$ for $\ell = 1, \dots, L$ be solutions in $\mathbb{F}_p^n$ to this system of equations. Suppose that there exists a partition $[k] = J_1 \cup \dots \cup J_t$ with $|J_h| \geq 2$ for $h = 1, \dots, t$ such that the following condition holds: For any choice of $\ell_1, \dots, \ell_k \in [L]$ such that $(x_1^{(\ell_1)}, \dots, x_k^{(\ell_k)})$ is a solution to the given system of equations, we have $|\{\ell_j \mid j \in J_h\}| = 1$ for all $h = 1, \dots, t$. Then we must have $L \leq k \cdot (\Gamma_{p,m,k})^n$.*

Proof Let us define the function $f : [L]^k \to \mathbb{F}_p$ as in Lemma 3.2, then the slice rank of f is at most $k \cdot (\Gamma_{p,m,k})^n$. On the other hand, the function f satisfies the assumptions in Corollary 3.6, so the slice rank of f is equal to L. Hence $L \leq k \cdot (\Gamma_{p,m,k})^n$, as desired. □

Remark 3.8 We remark that the constant factor k can be removed from the bound $L \leq k \cdot (\Gamma_{p,m,k})^n$ in Corollary 3.7 by a power trick similar to the one in the proof of Theorem 1.1. However, since this power trick is notationally cumbersome in this setting and since the bound $L \leq k \cdot (\Gamma_{p,m,k})^n$ suffices for our purposes, we stated Corollary 3.7 with this weaker bound.

4 Proof of Theorem 1.3

4.1 Inductive setup

Somewhat similar to our approach in Sect. 2, we will also use an inductive argument to prove Theorem 1.3. However, this induction requires a somewhat technical setup. We will associate a certain weight to each solution $(x_1, \dots, x_k)$ to the system $(\star)$, and then induct on this quantity.

Let us fix positive integers m and k, a prime p and coefficients $a_{j,i} \in \mathbb{F}_p$ for the system $(\star)$ such that the assumption $a_{j,1} + \dots + a_{j,k} = 0$ for $j = 1, \dots, m$ in Theorem 1.3 is satisfied. We start by making the following definitions.

Definition 4.1 Let V be a vector space over $\mathbb{F}_p$. For a given solution $(x_1, \dots, x_k) \in (V \setminus \{0\})^k$ to the system $(\star)$, let us say that a subset $I \subseteq [k]$ is *admissible* if it satisfies the following two conditions:

(i) The vectors x_i for $i \in I$ are linearly independent.
(ii) For all $j \in [k] \setminus I$ we have $x_j \notin \operatorname{span}(x_i \mid i \in I)$.

Note that the empty set $I = \emptyset$ always satisfies the conditions for being admissible (here, we use that the vectors $x_1, \dots, x_k$ are all non-zero). In particular, for every solution $(x_1, \dots, x_k) \in (V \setminus \{0\})^k$ to the system $(\star)$ there exists some admissible subset $I \subseteq [k]$.

Definition 4.2 Let V be a vector space over $\mathbb{F}_p$. For a given solution $(x_1, \dots, x_k) \in (V \setminus \{0\})^k$ to the system $(\star)$, let us define the *weight* of an admissible subset $I \subseteq [k]$

as follows: Letting $U = \operatorname{span}(x_i \mid i \in I)$, we define the weight of I with respect to $(x_1, \dots, x_k)$ to be

$$(k+1) \cdot |I| + \left| \left\{ \operatorname{span}(\operatorname{proj}_{V/U}(x_j)) \;\middle|\; j \in [k] \backslash I \right\} \right|.$$

The second summand in the expression above for the weight of I counts the number of different subspaces of the quotient space V/U that are of the form $\operatorname{span}(\operatorname{proj}_{V/U}(x_j))$ for some $j \in [k]\backslash I$. Note that we clearly have $|\{\operatorname{span}(\operatorname{proj}_{V/U}(x_j)) \mid j \in [k]\backslash I\}| \le k - |I| \le k$.

Definition 4.3 Let V be a vector space over $\mathbb{F}_p$. For a solution $(x_1, \dots, x_k) \in (V\backslash\{0\})^k$ to the system $(\star)$, let us define the *weight* $\omega(x_1, \dots, x_k)$ of $(x_1, \dots, x_k)$ to be the maximum weight of any admissible subset $I \subseteq [k]$ with respect to $(x_1, \dots, x_k)$. Furthermore, let us define $\mathcal{I}(x_1, \dots, x_k) \subseteq [k]$ to be an admissible subset $I \subseteq [k]$ for $(x_1, \dots, x_k)$ where this maximum weight is attained (if there are several choices for I attaining the maximum weight, we choose one arbitrarily).

Note that the weight $\omega(x_1, \dots, x_n)$ is clearly a non-negative integer. Also note that for a vector space $V \subseteq V'$ over $\mathbb{F}_p$, and a solution $(x_1, \dots, x_k) \in (V\backslash\{0\})^k \subseteq (V'\backslash\{0\})^k$ to the system $(\star)$, the weight $\omega(x_1, \dots, x_k)$ of $(x_1, \dots, x_k)$ does not depend on whether $x_1, \dots, x_k$ are interpreted as vectors in V or in V'.

Claim 4.4 *Let V be a vector space over $\mathbb{F}_p$, and let $(x_1, \dots, x_k) \in (V\backslash\{0\})^k$ be a solution to the system $(\star)$. Then we have $\omega(x_1, \dots, x_k) \notin \{0, (k+1), 2 \cdot (k+1), \dots, (k-1) \cdot (k+1)\}$,*

$$|\mathcal{I}(x_1, \dots, x_k)| = \lfloor \omega(x_1, \dots, x_k)/(k+1) \rfloor,$$

and

$$\dim \operatorname{span}(x_1, \dots, x_k) \ge \omega(x_1, \dots, x_k)/(k+1).$$

Proof Let $I = \mathcal{I}(x_1, \dots, x_k)$, and let $U = \operatorname{span}(x_i \mid i \in I)$. Then $\omega(x_1, \dots, x_k)$ equals the weight of I with respect to $(x_1, \dots, x_k)$. In other words,

$$\omega(x_1, \dots, x_k) = (k+1) \cdot |I| + \left| \left\{ \operatorname{span}(\operatorname{proj}_{V/U}(x_j)) \;\middle|\; j \in [k] \backslash I \right\} \right|. \tag{4.1}$$

First, let us consider the case $I = [k]$. Then we have $|\mathcal{I}(x_1, \dots, x_k)| = |I| = k$ and $\omega(x_1, \dots, x_k) = (k+1) \cdot k$, and $\dim \operatorname{span}(x_1, \dots, x_k) \ge \dim \operatorname{span}(x_i \mid i \in I) = |I| = k$ (here, we used condition (i) in Definition 4.1). Hence the statements in the claim are satisfied in the case $I = [k]$.

Now let us assume that $I \ne [k]$. Then there is at least one vector x_j with $j \in [k]\backslash I$, and so from (4.1) we obtain

$$(k+1) \cdot |I| + 1 \le \omega(x_1, \dots, x_k) \le (k+1) \cdot |I| + k.$$

This in particular implies $\lfloor \omega(x_1,\dots,x_k)/(k+1)\rfloor = |I| = |\mathcal{I}(x_1,\dots,x_k)|$. It furthermore implies that $\omega(x_1,\dots,x_k)$ is not divisible by $k+1$ and hence $\omega(x_1,\dots,x_k) \notin \{0,(k+1),2\cdot(k+1),\dots,(k-1)\cdot(k+1)\}$.

Finally, recall from condition (i) in Definition 4.1 that the vectors x_i for $i \in I$ are linearly independent. Furthermore, there is some $j \in [k]\setminus I$ and by condition (ii) in Definition 4.1 we have $x_j \notin \operatorname{span}(x_i \mid i \in I)$. Thus, $\dim\operatorname{span}(x_1,\dots,x_k) \ge |I|+1 \ge \omega(x_1,\dots,x_k)/(k+1)$. □

The following proposition is similar to Theorem 1.3, but instead of finding a solution $(x_1,\dots,x_k) \in A^k$ to $(\star)$ with $\dim\operatorname{span}(x_1,\dots,x_k) \ge r$, it aims to find a solution $(x_1,\dots,x_k) \in A^k$ with large weight $\omega(x_1,\dots,x_k)$. We will prove Proposition 4.5 by induction on w, and we will deduce Theorem 1.3 by considering $w = (r-1)(k+1)-1$. Recall that we are assuming $a_{j,1}+\dots+a_{j,k} = 0$ for $j = 1,\dots,m$.

Proposition 4.5 *Fix a non-negative integer w and assume that $k \ge 2m+1+\lfloor w/(k+1)\rfloor$. Then there exist constants $C \ge 1$ and $0 < c \le 1$ such that the following holds: For any non-negative integer n and any subset $A \subseteq \mathbb{F}_p^n\setminus\{0\}$ of size $|A| > C\cdot p^{(1-c)n}$, the system $(\star)$ has a solution $(x_1,\dots,x_k) \in A^k$ with weight $\omega(x_1,\dots,x_k) > w$.*

Let us now deduce Theorem 1.3 from Proposition 4.5.

Proof of Theorem 1.3 As in the theorem statement, let $r \ge 2$ and assume that $k \ge 2m+r-1$. Now, let $w = (r-1)(k+1)-1$, and note that then $\lfloor w/(k+1)\rfloor = r-2$ and therefore $k \ge 2m+1+r-2 = 2m+1+\lfloor w/(k+1)\rfloor$. Let us define $C^{\text{rank}}_{p,m,k,r} = C+1$ and $\Gamma^{\text{rank}}_{p,m,k,r} = p^{1-c} < p$, where $C \ge 1$ and $0 < c \le 1$ are the constants obtained from Proposition 4.5.

Now, let n be a non-negative integer and let $A \subseteq \mathbb{F}_p^n$ be a subset of size $|A| > C^{\text{rank}}_{p,m,k,r}\cdot(\Gamma^{\text{rank}}_{p,m,k,r})^n$. Then we have

$$|A\setminus\{0\}| \ge |A|-1 > (C^{\text{rank}}_{p,m,k,r}-1)\cdot(\Gamma^{\text{rank}}_{p,m,k,r})^n = C\cdot p^{(1-c)n}.$$

Thus, by Proposition 4.5 applied to $A\setminus\{0\}$, there exists a solution $(x_1,\dots,x_k) \in (A\setminus\{0\})^k \subseteq A^k$ to the system $(\star)$ with weight $\omega(x_1,\dots,x_k) > w = (r-1)(k+1)-1$, Note that this means that $\omega(x_1,\dots,x_k) \ge (r-1)(k+1)$. However, as $2 \le r \le k-1$, the first statement in Claim 4.4 implies that $\omega(x_1,\dots,x_k) \ne (r-1)(k+1)$. Hence $\omega(x_1,\dots,x_k) > (r-1)(k+1)$ and consequently by the last statement in Claim 4.4 we have $\dim\operatorname{span}(x_1,\dots,x_k) \ge \omega(x_1,\dots,x_k)/(k+1) > r-1$. This means that $\dim\operatorname{span}(x_1,\dots,x_k) \ge r$, as desired. □

We will prove Proposition 4.5 in the following subsection. In the proof, we will use the following lemma, giving a structural property for the spans $\operatorname{span}(\operatorname{proj}_{V/U}(x_j))$ appearing in Definition 4.2.

Lemma 4.6 *Let V be a vector space over $\mathbb{F}_p$, and let $(x_1,\dots,x_k) \in (V\setminus\{0\})^k$ be a solution to the system $(\star)$. Furthermore, let $I = \mathcal{I}(x_1,\dots,x_k)$ and $U = \operatorname{span}(x_i \mid i \in I)$. Then there exist a partition $[k]\setminus I = J_1 \cup \dots \cup J_t$ with $|J_h| \ge 2$ for $h = 1,\dots,t$ and distinct one-dimensional subspaces $W_1,\dots,W_t \subseteq V/U$ such that $\operatorname{span}(\operatorname{proj}_{V/U}(x_j)) = W_h$ for all $h = 1,\dots,t$ and all $j \in J_h$.*

Proof Recall that the set $I = \mathcal{I}(x_1, \dots, x_k)$ is admissible for $(x_1, \dots, x_k)$, and that therefore by condition (ii) in Definition 4.1 we have $x_j \notin \operatorname{span}(x_i \mid i \in I) = U$ for all $j \in [k] \backslash I$. Hence for every $j \in [k] \backslash I$, the projection $\operatorname{proj}_{V/U}(x_j)$ is a non-zero vector in V/U and therefore $\operatorname{span}(\operatorname{proj}_{V/U}(x_j))$ is a one-dimensional subspace of V/U.

Let $W_1, \dots, W_t \subseteq V/U$ be the list of all distinct one-dimensional subspaces of V/U that are of the form $\operatorname{span}(\operatorname{proj}_{V/U}(x_j))$ for some $j \in [k] \backslash I$. Furthermore, for each $h = 1, \dots, t$, let $J_h \subseteq [k] \backslash I$ be the set of indices $j \in [k] \backslash I$ such that $\operatorname{span}(\operatorname{proj}_{V/U}(x_j)) = W_h$. Then $[k] \backslash I = J_1 \cup \dots \cup J_t$ is a partition.

It remains to show that we have $|J_h| \geq 2$ for $h = 1, \dots, t$. Note that this is equivalent to showing that for each $j \in [k] \backslash I$ there exists some $j' \in [k] \backslash I$ with $j' \neq j$ and $\operatorname{span}(\operatorname{proj}_{V/U}(x_j)) = \operatorname{span}(\operatorname{proj}_{V/U}(x_{j'}))$.

So let us fix some $j \in [k] \backslash I$, and suppose for contradiction that $\operatorname{span}(\operatorname{proj}_{V/U}(x_{j'})) \neq \operatorname{span}(\operatorname{proj}_{V/U}(x_j))$ for all $j' \in [k] \backslash (I \cup \{j\})$. As both $\operatorname{span}(\operatorname{proj}_{V/U}(x_{j'}))$ and $\operatorname{span}(\operatorname{proj}_{V/U}(x_j))$ are one-dimensional subspaces of V/U, this means that $\operatorname{proj}_{V/U}(x_{j'}) \notin \operatorname{span}(\operatorname{proj}_{V/U}(x_j))$ for all $j' \in [k] \backslash (I \cup \{j\})$. In other words, we have $x_{j'} \notin U + \operatorname{span}(x_j) = \operatorname{span}(x_i \mid i \in I \cup \{j\})$ for all $j' \in [k] \backslash (I \cup \{j\})$. Hence the set $I \cup \{j\}$ satisfies condition (ii) in Definition 4.1.

Furthermore, recall that the vectors x_i for $i \in I$ are linearly independent by condition (i) in Definition 4.1. Now, $x_j \notin U = \operatorname{span}(x_i \mid i \in I)$ implies that x_j is also linearly independent from these vectors. In other words, the vectors x_i for $i \in I \cup \{j\}$ are linearly independent and the set $I \cup \{j\}$ satisfied condition (i) in Definition 4.1. Thus, we can conclude that the set $I \cup \{j\}$ is admissible for $(x_1, \dots, x_k)$.

Let us now compare the weights of the admissible sets I and $I \cup \{j\}$ with respect to $(x_1, \dots, x_k)$. Since $I = \mathcal{I}(x_1, \dots, x_k)$ must have the maximum weight among all admissible sets with respect to $(x_1, \dots, x_k)$, the weight of $I \cup \{j\}$ can be at most as large as the weight of I. However, the weight of I is

$$\begin{aligned}(k+1) \cdot |I| + \left| \left\{ \operatorname{span}(\operatorname{proj}_{V/U}(x_j)) \,\middle|\, j \in [k] \backslash I \right\} \right| &\leq (k+1) \cdot |I| + k \\ &< (k+1) \cdot |I \cup \{j\}|,\end{aligned}$$

which means that the weight of $I \cup \{j\}$ is actually larger than the weight of I. This is a contradiction. □

4.2 Proof of Proposition 4.5

Let us now prove Proposition 4.5 by induction on w.

For the base case $w = 0$, we can take $C = 1$ and $c = 1$. Then for any subset $A \subseteq \mathbb{F}_p^n \backslash \{0\}$ of size $|A| > C \cdot p^{(1-c)n} = 1$, we can pick some vector $x \in A$ and consider the k-tuple $(x_1, \dots, x_k) = (x, \dots, x) \in A^k$. Note that by our assumption $a_{j,1} + \dots + a_{j,k} = 0$ for $j = 1, \dots, m$ made at the beginning of this section, this k-tuple $(x_1, \dots, x_k)$ is a solution to the system $(\star)$. Furthermore, by the first statement in Claim 4.4 we have $\omega(x_1, \dots, x_k) \neq 0$. Hence $\omega(x_1, \dots, x_k) > 0$ and we have proved the base case $w = 0$.

Now let us assume that $w \geq 1$ and that Proposition 4.5 holds for $w-1$ with constants $C' \geq 1$ and $0 < c' \leq 1$. Also recall that we made the assumption $k \geq 2m+1+\lfloor w/(k+1)\rfloor$ in the statement of Proposition 4.5.

First, consider the case that $w \in \{0, (k+1), 2\cdot(k+1), \dots, (k-1)\cdot(k+1)\}$. In this case, we can take $C = C'$ and $c = c'$. Indeed, for any subset $A \subseteq \mathbb{F}_p^n \backslash \{0\}$ of size $|A| > C \cdot p^{(1-c)n} = C' \cdot p^{(1-c')n}$, by the induction hypothesis the system $(\star)$ has a solution $(x_1, \dots, x_k) \in A^k$ with weight $\omega(x_1, \dots, x_k) > w-1$. Since $\omega(x_1, \dots, x_k) \notin \{0, (k+1), 2\cdot(k+1), \dots, (k-1)\cdot(k+1)\}$ by Claim 4.4, we have $\omega(x_1, \dots, x_k) \neq w$ and therefore $\omega(x_1, \dots, x_k) > w$. This shows the desired statement in Proposition 4.5 for w.

We may therefore from now on assume that $w \notin \{0, (k+1), 2\cdot(k+1), \dots, (k-1)\cdot(k+1)\}$. By the assumption $k \geq 2m+1+\lfloor w/(k+1)\rfloor$, we have $w < k\cdot(k+1)$. Hence w is not divisible by $k+1$.

Since $k \geq 2m+1+\lfloor w/(k+1)\rfloor$, we have $k - \lfloor w/(k+1)\rfloor \geq 2m+1$, so the constant $\Gamma_{p,m,k-\lfloor w/(k+1)\rfloor} < p$ considered in Sect. 3 is well-defined. Let us write $\Gamma = \Gamma_{p,m,k-\lfloor w/(k+1)\rfloor}$, then $1 \leq \Gamma < p$.

Now, $p/\Gamma > 1$ and $c' > 0$, so there is some $c > 0$ such that $p^c = (p/\Gamma)^{c'/(k-1)}$. Note that $c \leq 1$ since $c' \leq 1$ and $\Gamma \geq 1$. So $0 < c \leq 1$, as desired.

We need to prove that there exists a constant $C \geq 1$ such that for any non-negative integer n and any subset $A \subseteq \mathbb{F}_p^n \backslash \{0\}$ of size $|A| > C \cdot p^{(1-c)n}$, the system $(\star)$ has a solution $(x_1, \dots, x_k) \in A^k$ with weight $\omega(x_1, \dots, x_k) > w$.

Note that by making the constant C large enough, we may assume that n is sufficiently large such that $(p/\Gamma)^n > (2kp^2)^{(2k+1)(k-1)}$.

Hence it suffices to prove that for any n with $(p/\Gamma)^n > (2kp^2)^{(2k+1)(k-1)}$ and any subset $A \subseteq \mathbb{F}_p^n \backslash \{0\}$ of size $|A| > 4C'(2kp)^{2k+1} \cdot p^{(1-c)n}$, there is a solution $(x_1, \dots, x_k) \in A^k$ to $(\star)$ with $\omega(x_1, \dots, x_k) > w$.

So let us assume for contradiction that $(p/\Gamma)^n > (2kp^2)^{(2k+1)(k-1)}$ and $|A| > 4C'(2kp)^{2k+1} \cdot p^{(1-c)n}$ and that every solution $(x_1, \dots, x_k) \in A^k$ of the system $(\star)$ of has weight $\omega(x_1, \dots, x_k) \leq w$.

Since $p^c = (p/\Gamma)^{c'/(k-1)}$ by the definition of c, we have

$$|A| > 4C'(2kp)^{2k+1} \cdot p^{(1-c)n} = 4C'(2kp)^{2k+1} \cdot p^n \cdot \left(\frac{\Gamma}{p}\right)^{c'n/(k-1)}. \tag{4.2}$$

Using that $0 < c' \leq 1$ and $\Gamma < p$ and $C' \geq 1$, this furthermore implies

$$\begin{aligned} |A| > 4C'(2kp)^{2k+1} \cdot p^n \cdot \left(\frac{\Gamma}{p}\right)^{c'n/(k-1)} &\geq 4C'(2kp)^{2k+1} \cdot p^n \cdot \left(\frac{\Gamma}{p}\right)^{n} \\ &\geq (2kp)^{2k+1} \cdot \Gamma^n. \end{aligned} \tag{4.3}$$

Our proof strategy is somewhat similar to the proof of Theorem 2.1 in Sect. 2, again using a probabilistic subspace sampling argument. We will again consider a random subspace $V \subseteq \mathbb{F}_p^n$ of a suitably chosen dimension. This time, our goal is to find a fairly large subset $A^* \subseteq A \cap V$ such that there are no solutions $(x_1, \dots, x_k) \in (A^*)^k$ to the system $(\star)$ of weight $\omega(x_1,\dots,x_k) = w$. By our assumption on A, this means that every solution $(x_1,\dots,x_k) \in (A^*)^k$ to the system $(\star)$ must satisfy $\omega(x_1,\dots,x_k) \le w-1$. We will then obtain a contradiction by applying the induction hypothesis for $w-1$ to the set A^*.

In order to be able to obtain the desired set $A^* \subseteq A \cap V$ not containing any solutions $(x_1,\dots,x_k) \in (A^*)^k$ to $(\star)$ with $\omega(x_1,\dots,x_k) = w$, we will bound the total number of solutions $(x_1,\dots,x_k) \in A^k$ to $(\star)$ in the set A with $\omega(x_1,\dots,x_k) = w$. More precisely, since the relevant probabilities for the subspace sampling argument depend on the dimension $\dim \operatorname{span}(x_1,\dots,x_k)$ for each solution $(x_1,\dots,x_k) \in A^k$, we will actually bound the number of solutions $(x_1,\dots,x_k) \in A^k$ to $(\star)$ with $\omega(x_1,\dots,x_k) = w$ and $\dim \operatorname{span}(x_1,\dots,x_k) = r$ for every possible value of this dimension r. The next claim analyzes these possible values.

Claim 4.7 *For every solution $(x_1,\dots,x_k) \in A^k$ to the system $(\star)$ of weight $\omega(x_1,\dots,x_k) = w$, we have $|\mathcal{I}(x_1,\dots,x_k)| = \lfloor w/(k+1)\rfloor$ and $\lfloor w/(k+1)\rfloor + 1 \le \dim \operatorname{span}(x_1,\dots,x_k) \le k$.*

Proof Fix a solution $(x_1,\dots,x_k) \in A^k$ to $(\star)$ with $\omega(x_1,\dots,x_k) = w$. By Claim 4.4 we have $|\mathcal{I}(x_1,\dots,x_k)| = \lfloor \omega(x_1,\dots,x_k)/(k+1)\rfloor = \lfloor w/(k+1)\rfloor$.

For the second part of the claim, the upper bound $\dim \operatorname{span}(x_1,\dots,x_k) \le k$ is clear. For the lower bound, note that by Claim 4.4 we have

$$\dim \operatorname{span}(x_1,\dots,x_k) \ge \omega(x_1,\dots,x_k)/(k+1) = w/(k+1) > \lfloor w/(k+1)\rfloor,$$

where in the last step we used that w is not divisible by $k+1$. Hence we must have $\dim \operatorname{span}(x_1,\dots,x_k) \ge \lfloor w/(k+1)\rfloor + 1$, as desired. □

We remark that one can actually obtain a stronger upper bound for $\dim \operatorname{span}$ $(x_1,\dots,x_k)$ than the bound in the claim above by taking into account the linear relations imposed by the system $(\star)$. However, the trivial upper bound $\dim \operatorname{span}(x_1,\dots,x_k) \le k$ suffices for our argument.

In light of Claim 4.7, for every $r = \lfloor w/(k+1)\rfloor + 1,\dots,k$, we need to bound the number of solutions $(x_1,\dots,x_k) \in A^k$ to $(\star)$ with $\omega(x_1,\dots,x_k) = w$ and $\dim \operatorname{span}(x_1,\dots,x_k) = r$. We will count these solutions by distinguishing all possibilities for the set $\mathcal{I}(x_1,\dots,x_k)$, noting that by Claim 4.7, $\mathcal{I}(x_1,\dots,x_k) \subseteq [k]$ is always a subset of size $\lfloor w/(k+1)\rfloor$.

The following lemma is the key step in order to obtain the desired bound. It gives a structural property for solutions $(x_1,\dots,x_k) \in A^k$ to $(\star)$ with $\omega(x_1,\dots,x_k) = w$ where $\mathcal{I}(x_1,\dots,x_k) = I$ for some fixed set $I \subseteq [k]$ (of size $|I| = \lfloor w/(k+1)\rfloor$) and where the vectors $x_i \in A$ for $i \in I$ are fixed. In the proof of this lemma, we will use the slice rank polynomial method in the form of Corollary 3.7. Recall that we defined $\Gamma = \Gamma_{p,m,k-\lfloor w/(k+1)\rfloor} < p$.

Lemma 4.8 *Fix a subset $I \subseteq [k]$ of size $|I| = \lfloor w/(k+1)\rfloor$, and fix vectors $x_i \in A$ for $i \in I$. Let $U = \operatorname{span}(x_i \mid i \in I)$. Suppose that $(x_1^{(\ell)},\dots,x_k^{(\ell)}) \in A^k$ for $\ell = 1,\dots,L$ is a list of solutions to the system $(\star)$ such that for all $\ell = 1,\dots,L$ we have $\omega(x_1^{(\ell)},\dots,x_k^{(\ell)}) = w$ and $\mathcal{I}(x_1^{(\ell)},\dots,x_k^{(\ell)}) = I$ and $x_i^{(\ell)} = x_i$ for all $i \in I$.*

Furthermore, suppose that the sets $\{\mathrm{span}(\mathrm{proj}_{\mathbb{F}_p^n/U}(x_j^{(\ell)})) \mid j \in [k]\setminus I\}$ *are disjoint for all* $\ell = 1, \dots, L$. *Then we must have* $L \le k^{k+1} \cdot \Gamma^n$.

Proof Suppose for contradiction that $L > k^{k+1} \cdot \Gamma^n$. For each $\ell = 1, \dots, L$ we can apply Lemma 4.6 to $(x_1^{(\ell)}, \dots, x_k^{(\ell)})$ and obtain a partition $[k]\setminus I = J_1^{(\ell)} \cup \dots \cup J_{t_\ell}^{(\ell)}$ with $|J_h^{(\ell)}| \ge 2$ for $h = 1, \dots, t_\ell$ and distinct one-dimensional subspaces $W_1^{(\ell)}, \dots, W_{t_\ell}^{(\ell)} \subseteq \mathbb{F}_p^n/U$ such that $\mathrm{span}(\mathrm{proj}_{\mathbb{F}_p^n/U}(x_j^{(\ell)})) = W_h^{(\ell)}$ for all $h = 1, \dots, t_\ell$ and all $j \in J_h^{(\ell)}$. Note that by the condition $|J_h^{(\ell)}| \ge 2$ for $h = 1, \dots, t_\ell$, the sets $J_h^{(\ell)}$ are all non-empty and we have $t_\ell \le k$ for all $\ell = 1, \dots, L$.

Let us now distinguish all possibilities for the partitions $[k]\setminus I = J_1^{(\ell)} \cup \dots \cup J_{t_\ell}^{(\ell)}$. Since $t_\ell \le k$, there are at most k^k possibilities for such a partition. Hence, as $L > k^{k+1} \cdot \Gamma^n$, for more than $k \cdot \Gamma^n$ different ℓ we must obtain the same partition $[k]\setminus I = J_1^{(\ell)} \cup \dots \cup J_{t_\ell}^{(\ell)}$. In other words, there exists a partition $[k]\setminus I = J_1 \cup \dots \cup J_t$ which occurs for more than $k \cdot \Gamma^n$ different ℓ. By relabeling, we may assume that this partition occurs for $\ell = 1, \dots, L'$ for some $L' > k \cdot \Gamma^n$.

To summarize, we now have a fixed partition $[k]\setminus I = J_1^{(\ell)} \cup \dots \cup J_t^{(\ell)}$ with $|J_h| \ge 2$ for $h = 1, \dots, t$ and for each $\ell = 1, \dots, L'$ (where $L' > k \cdot \Gamma^n$) we have distinct one-dimensional subspaces $W_1^{(\ell)}, \dots, W_t^{(\ell)} \subseteq \mathbb{F}_p^n/U$ such that $\mathrm{span}(\mathrm{proj}_{\mathbb{F}_p^n/U}(x_j^{(\ell)})) = W_h^{(\ell)}$ for all $h = 1, \dots, t$ and all $j \in J_h$. Now, for each $\ell = 1, \dots, L'$ we have $\{\mathrm{span}(\mathrm{proj}_{\mathbb{F}_p^n/U}(x_j^{(\ell)})) \mid j \in [k]\setminus I\} = \{W_1^{(\ell)}, \dots, W_t^{(\ell)}\}$ and by the assumptions in the lemma these sets are disjoint for all $\ell = 1, \dots, L'$. This means that the subspaces $W_h^{(\ell)} \subseteq \mathbb{F}_p^n/U$ are distinct for all $h = 1, \dots, t$ and all $\ell = 1, \dots, L'$.

Recall that for every $\ell = 1, \dots, L'$ we have $\mathcal{I}(x_1^{(\ell)}, \dots, x_k^{(\ell)}) = I$. In particular, this means that the set I is admissible for $(x_1^{(\ell)}, \dots, x_k^{(\ell)})$. Hence, recalling that $x_i^{(\ell)} = x_i$ for all $i \in I$, by condition (ii) in Definition 4.1 we have $x_j^{(\ell)} \notin \mathrm{span}(x_i \mid i \in I) = U$ for all $j \in [k]\setminus I$ and all $\ell = 1, \dots, L'$. Furthermore, using that $L' \ge 1$ as $L' > k \cdot \Gamma^n$, condition (i) in Definition 4.1 implies that the vectors x_i for $i \in I$ are linearly independent.

For the moment, choose any $\ell \in \{1, \dots, L'\}$ (using that $L' \ge 1$). Since we have $\mathcal{I}(x_1^{(\ell)}, \dots, x_k^{(\ell)}) = I$ and $\omega(x_1^{(\ell)}, \dots, x_k^{(\ell)}) = w$, the weight of the admissible set I with respect to $(x_1^{(\ell)}, \dots, x_k^{(\ell)})$ must be equal to w. On the other hand, this weight is

$$\begin{aligned}(k+1) \cdot |I| + \left|\left\{\mathrm{span}(\mathrm{proj}_{\mathbb{F}_p^n/U}(x_j^{(\ell)})) \,\middle|\, j \in [k]\setminus I\right\}\right| &= (k+1) \cdot |I| + \left|\left\{W_1^{(\ell)}, \dots, W_t^{(\ell)}\right\}\right| \\ &= (k+1) \cdot |I| + t.\end{aligned}$$

Hence we can conclude that

$$t = w - (k+1) \cdot |I|. \tag{4.4}$$

For ease of notation, let us assume for the rest of the proof of this lemma that the set $I \subseteq [k]$ consists of the last $|I| = \lfloor w/(k+1) \rfloor$ indices, i.e. $I = \{k - |I| + 1, \dots, k\}$ (otherwise we can just relabel the indices). Given the fixed vectors $x_i \in A$ for $i \in$

$I = \{k-|I|+1, \dots, k\}$, we can now interpret ($\star$) as a (non-homogeneous) system of m linear equations in the $k-|I|$ variables $x_1, \dots, x_{k-|I|}$. For each $\ell = 1, \dots, L'$ the $(k-|I|)$-tuple $(x_1^{(\ell)}, \dots, x_{k-|I|}^{(\ell)})$ is a solution to this system of equations, since $(x_1^{(\ell)}, \dots, x_{k-|I|}^{(\ell)}, x_{k-|I|+1}, \dots, x_k) = (x_1^{(\ell)}, \dots, x_k^{(\ell)})$ is a solution to ($\star$) (recall the assumption that $x_i^{(\ell)} = x_i$ for all $i \in I$).

Our goal is now to apply Corollary 3.7 to this system of equations and the solutions $(x_1^{(\ell)}, \dots, x_{k-|I|}^{(\ell)})$ for $\ell = 1, \dots, L'$. The following claim states that the condition in the statement of Corollary 3.7 is satisfied.

Claim 4.9 *We have $|\{\ell_j \mid j \in J_h\}| = 1$ for $h = 1, \dots, t$ for any choice of $\ell_1, \dots, \ell_{k-|I|} \in [L']$ such that $(x_1^{(\ell_1)}, \dots, x_{k-|I|}^{(\ell_{k-|I|})})$ is a solution to ($\star$) when interpreted as a system in the first $k-|I|$ variables after fixing the given $x_i \in A$ for $i \in I = \{k-|I|+1, \dots, k\}$.*

Before proving this claim, let us first finish the rest of the proof of the lemma. Recall that $k-|I| = k - \lfloor w/(k+1) \rfloor \geq 2m+1$. We can therefore apply Corollary 3.7 to the system of m linear equations in the $k-|I|$ variables $x_1, \dots, x_{k-|I|}$ obtained from ($\star$) after fixing the given $x_i \in A$ for $i \in I$. By Claim 4.9 the solutions $(x_1^{(\ell)}, \dots, x_{k-|I|}^{(\ell)})$ for $\ell = 1, \dots, L'$ to this system and the partition $[k] \setminus I = J_1 \cup \dots \cup J_t$ satisfy the conditions in Corollary 3.7. Hence the corollary implies that $L' \leq k \cdot \Gamma^n_{p,m,k-\lfloor w/(k+1) \rfloor} = k \cdot \Gamma^n$. This contradicts the lower bound $L' > k \cdot \Gamma^n$ from above. This contradiction finishes the proof of the lemma, apart from proving Claim 4.9.

Proof of Claim 4.9 Let us fix $\ell_1, \dots, \ell_{k-|I|} \in [L']$ such that $(x_1^{(\ell_1)}, \dots, x_{k-|I|}^{(\ell_{k-|I|})})$ is a solution to this system, meaning that $(x_1^{(\ell_1)}, \dots, x_{k-|I|}^{(\ell_{k-|I|})}, x_{k-|I|+1}, \dots, x_k)$ is a solution to the original system ($\star$). By our assumption on the set A, this solution $(x_1^{(\ell_1)}, \dots, x_{k-|I|}^{(\ell_{k-|I|})}, x_{k-|I|+1}, \dots, x_k) \in A^k$ to ($\star$) must have weight

$$\omega(x_1^{(\ell_1)}, \dots, x_{k-|I|}^{(\ell_{k-|I|})}, x_{k-|I|+1}, \dots, x_k) \leq w. \tag{4.5}$$

We claim that the set $I = \{k-|I|+1, \dots, k\}$ is admissible for $(x_1^{(\ell_1)}, \dots, x_{k-|I|}^{(\ell_{k-|I|})}, x_{k-|I|+1}, \dots, x_k)$. Indeed, as shown above the vectors x_i for $i \in I$ are linearly independent, so condition (i) in Definition 4.1 is satisfied. We also proved above that $x_j^{(\ell)} \notin \operatorname{span}(x_i \mid i \in I) = U$ for all $j \in [k] \setminus I$ and all $\ell = 1, \dots, L'$. In particular, $x_j^{(\ell_j)} \notin \operatorname{span}(x_i \mid i \in I)$ for all $j \in [k] \setminus I$ and so condition (ii) is satisfied as well. Hence the set $I = \{k-|I|+1, \dots, k\}$ is admissible for $(x_1^{(\ell_1)}, \dots, x_{k-|I|}^{(\ell_{k-|I|})}, x_{k-|I|+1}, \dots, x_k)$.

Now, by (4.5) the weight of the admissible set $I = \{k-|I|+1, \dots, k\}$ with respect to the solution $(x_1^{(\ell_1)}, \dots, x_{k-|I|}^{(\ell_{k-|I|})}, x_{k-|I|+1}, \dots, x_k)$ to ($\star$) is at most w. Hence

$$\begin{aligned} w &\geq (k+1)\cdot|I| + \left|\left\{\operatorname{span}(\operatorname{proj}_{\mathbb{F}_p^n/U}(x_j^{(\ell_j)}))\;\middle|\; j\in[k]\backslash I\right\}\right|\\ &= (k+1)\cdot|I| + \left|\bigcup_{h=1}^{t}\left\{W_h^{(\ell_j)}\;\middle|\; j\in J_h\right\}\right|. \end{aligned}$$

Here, for the second step we used that $[k]\backslash I = J_1\cup\cdots\cup J_t$ is a partition and that $\operatorname{span}(\operatorname{proj}_{\mathbb{F}_p^n/U}(x_j^{(\ell)})) = W_h^{(\ell)}$ for all $h=1,\dots,t$, all $j\in J_h$ and all $\ell=1,\dots,L'$. Together with (4.4) this yields

$$t = w-(k+1)\cdot|I| \geq \left|\bigcup_{h=1}^{t}\left\{W_h^{(\ell_j)}\;\middle|\; j\in J_h\right\}\right| = \sum_{h=1}^{t}|\{\ell_j\mid j\in J_h\}|,$$

where in the last step we used that the spaces $W_h^{(\ell)}\subseteq\mathbb{F}_p^n/U$ are distinct for all $h=1,\dots,t$ and all $\ell=1,\dots,L'$. Since the sets J_h for $h=1,\dots,t$ are non-empty (as $|J_h|\geq 2$), we have $|\{\ell_j\mid j\in J_h\}|\geq 1$ for $h=1,\dots,t$. Hence the previous inequality implies that we must have $|\{\ell_j\mid j\in J_h\}| = 1$ for all $h=1,\dots,t$, as desired. □

This finishes the proof of Lemma 4.8. □

Lemma 4.8 states that for a fixed subset $I\subseteq[k]$ (of size $|I| = \lfloor w/(k+1)\rfloor$) and fixed vectors $x_i\in A$ for $i\in I$, there cannot be too many solutions $(x_1,\dots,x_k)\in A^k$ to ($\star$) with $\omega(x_1,\dots,x_k) = w$ and $\mathcal{I}(x_1,\dots,x_k) = I$ such that the sets $\{\operatorname{span}(\operatorname{proj}_{\mathbb{F}_p^n/U}(x_j))\mid j\in[k]\backslash I\}$ are disjoint for all of these solutions. We will now use this lemma to bound the total number of solutions $(x_1,\dots,x_k)\in A^k$ to ($\star$) with $\omega(x_1,\dots,x_k) = w$ and $\mathcal{I}(x_1,\dots,x_k) = I$ for fixed I and fixed $x_i\in A$ for $i\in I$ (more precisely, we bound the number of such solutions with $\dim\operatorname{span}(x_1,\dots,x_k) = r$ for each fixed r), as stated in the following lemma. Roughly speaking, the proof strategy for this lemma is to choose a maximal collection of solutions $(x_1',\dots,x_k')$ of the desired form for which the sets $\{\operatorname{span}(\operatorname{proj}_{\mathbb{F}_p^n/U}(x_j'))\mid j\in[k]\backslash I\}$ are disjoint, and then to use that for every solution $(x_1,\dots,x_k)$ satisfying the desired conditions we must have $\operatorname{proj}_{\mathbb{F}_p^n/U}(x_t)\in\{\operatorname{span}(\operatorname{proj}_{\mathbb{F}_p^n/U}(x_j'))\mid j\in[k]\backslash I\}$ for some $t\in[k]\backslash I$ and some $(x_1',\dots,x_k')$ in the chosen collection. This means that there are only relatively few possibilities for $x_t\in A$, and we will be able to derive the desired bound on the number of possible $(x_1,\dots,x_k)$.

Lemma 4.10 *Fix a subset $I\subseteq[k]$ of size $|I| = \lfloor w/(k+1)\rfloor$, and fix vectors $x_i\in A$ for $i\in I$. Furthermore, fix an integer r with $\lfloor w/(k+1)\rfloor+1\leq r\leq k$. Then the number of solutions $(x_1,\dots,x_k)\in A^k$ to ($\star$) with $\omega(x_1,\dots,x_k) = w$, $\mathcal{I}(x_1,\dots,x_k) = I$ and* $\dim\operatorname{span}(x_1,\dots,x_k) = r$ *is at most $k^{k+3}2^kp^{rk}\cdot\Gamma^n\cdot|A|^{r-|I|-1}$.*

Proof Let $U = \operatorname{span}(x_i\mid i\in I)$. Furthermore, let us fix a list of solutions $(x_1^{(\ell)},\dots,x_k^{(\ell)})\in A^k$ for $\ell=1,\dots,L$ to ($\star$) of maximum possible length L such that for all $\ell=1,\dots,L$ we have $\omega(x_1^{(\ell)},\dots,x_k^{(\ell)})) = w$ and $\mathcal{I}(x_1^{(\ell)},\dots,x_k^{(\ell)})) = I$ and $x_i^{(\ell)} = x_i$ for all $i\in I$, and such that the sets $\{\operatorname{span}(\operatorname{proj}_{\mathbb{F}_p^n/U}(x_j^{(\ell)}))\mid j\in[k]\backslash I\}$ are disjoint for all $\ell=1,\dots,L$.

Then by Lemma 4.8 we must have $L \leq k^{k+1} \cdot \Gamma^n$ (and in particular, such a list of maximum possible length exists). For every $\ell = 1, \dots, L$ and every $j \in [k] \backslash I$, define

$$W_j^{(\ell)} = \operatorname{span}(x_i^{(\ell)} \mid i \in I \cup \{j\}) = \operatorname{span}(x_i \mid i \in I) + \operatorname{span}(x_j^{(\ell)}) = U + \operatorname{span}(x_j^{(\ell)}),$$

and note that each $W_j^{(\ell)}$ is a subspace of $\mathbb{F}_p^n$ of dimension at most $|I| + 1 = \lfloor w/(k+1) \rfloor + 1 \leq r$. Hence the union $\bigcup_{\ell=1}^{L} \bigcup_{j \in [k] \backslash I} W_j^{(\ell)}$ is a subset of $\mathbb{F}_p^n$ of size

$$\left| \bigcup_{\ell=1}^{L} \bigcup_{j \in [k] \backslash I} W_j^{(\ell)} \right| \leq L \cdot k \cdot p^r \leq k^{k+1} \cdot \Gamma^n \cdot k \cdot p^r = k^{k+2} p^r \cdot \Gamma^n. \tag{4.6}$$

Claim 4.11 *Suppose $x_j \in A$ for $j \in [k] \backslash I$ are vectors such that $(x_1, \dots, x_k) \in A^k$ is a solution to ($\star$) with $\omega(x_1, \dots, x_k) = w$, $\mathcal{I}(x_1, \dots, x_k) = I$ and* $\dim \operatorname{span}(x_1, \dots, x_k) = r$. *Then there exists an index $t \in [k] \backslash I$ and a subset $S \subseteq [k] \backslash (I \cup \{t\})$ of size $|S| = r - |I| - 1$ such that $x_t \in \bigcup_{\ell=1}^{L} \bigcup_{j \in [k] \backslash I} W_j^{(\ell)}$ and such that $x_1, \dots, x_k \in$* $\operatorname{span}(x_i \mid i \in I \cup \{t\} \cup S)$.

Proof By maximality of our chosen list of solutions $(x_1^{(\ell)}, \dots, x_k^{(\ell)})$, it cannot be possible to extend this list by taking $(x_1^{(L+1)}, \dots, x_k^{(L+1)}) = (x_1, \dots, x_k)$. Since $(x_1, \dots, x_k)$ satisfies all of the other conditions, this means that we must have

$$\{\operatorname{span}(\operatorname{proj}_{\mathbb{F}_p^n/U}(x_j)) \mid j \in [k] \backslash I\} \cap \{\operatorname{span}(\operatorname{proj}_{\mathbb{F}_p^n/U}(x_j^{(\ell)})) \mid j \in [k] \backslash I\} \neq \emptyset$$

or some $\ell \in [L]$. Hence we can find $t \in [k] \backslash I$ and $j \in [k] \backslash I$ and $\ell \in [L]$ such thatspan$(\operatorname{proj}_{\mathbb{F}_p^n/U}(x_t)) = \operatorname{span}(\operatorname{proj}_{\mathbb{F}_p^n/U}(x_j^{(\ell)}))$. In particular, we have $\operatorname{proj}_{\mathbb{F}_p^n/U}(x_t) \in \operatorname{span}(\operatorname{proj}_{\mathbb{F}_p^n/U}(x_j^{(\ell)}))$. This means that $x_t \in U + \operatorname{span}(x_j^{(\ell)}) = W_j^{(\ell)}$. Thus, we have found the desired index $t \in [k] \backslash I$ with $x_t \in \bigcup_{\ell=1}^{L} \bigcup_{j \in [k] \backslash I} W_j^{(\ell)}$. It remains to find a set $S \subseteq [k] \backslash (I \cup \{t\})$ with the desired conditions.

Since $\mathcal{I}(x_1, \dots, x_k) = I$, the set I is admissible for $(x_1, \dots, x_k)$, and so by condition (i) in Definition 4.1 the vectors x_i for $i \in I$ are linearly independent. Furthermore, by condition (ii) in Definition 4.1 we have $x_t \notin \operatorname{span}(x_i \mid i \in I) = U$. Hence the vectors x_i for $i \in I \cup \{t\}$ are linearly independent. We can therefore find a subset $S \subseteq [k] \backslash (I \cup \{t\})$ such that the vectors x_i for $i \in I \cup \{t\} \cup S$ form a basis of the space $\operatorname{span}(x_1, \dots, x_k)$. Then we clearly have $x_1, \dots, x_k \in \operatorname{span}(x_i \mid i \in I \cup \{t\} \cup S)$. Furthermore, as $\dim \operatorname{span}(x_1, \dots, x_k) = r$, we also have $|S| = r - |I| - 1$, as desired. □

In order to prove the lemma, we need to show that the number of choices for $(x_j \mid j \in [k] \backslash I)$ satisfying the conditions in Claim 4.11 is at most $k^{k+3} 2^k p^{rk} \cdot \Gamma^n \cdot |A|^{r-|I|-1}$. We will count by distinguishing all possibilities for the index $t \in [k] \backslash S$ and the set $S \subseteq [k] \backslash (I \cup \{t\})$ obtained from Claim 4.11. There are clearly at most k possibilities for t and at most 2^k possibilities for S.

Hence it suffices to prove that for every fixed $t \in [k]\backslash S$ and every fixed set $S \subseteq [k]\backslash(I \cup \{t\})$ of size $|S| = r - |I| - 1$ there are at most $k^{k+2} p^{rk} \cdot \Gamma^n \cdot |A|^{r-|I|-1}$ possibilities for choosing vectors $x_j \in A$ for $j \in [k]\backslash I$ such that $x_t \in \bigcup_{\ell=1}^{L} \bigcup_{j\in[k]\backslash I} W_j^{(\ell)}$ and $x_1, \dots, x_k \in \operatorname{span}(x_i \mid i \in I \cup \{t\} \cup S)$.

By (4.6) the number of possibilities for x_t is at most $k^{k+2} p^r \cdot \Gamma^n$. For every $j \in S$, the number of possibilities for $x_j \in A$ is at most $|A|$. Finally, after making all these choices, for each of the remaining $j \in [k]\backslash(I \cup \{t\} \cup S)$, we have at most p^r choices for x_j, since $x_j \in \operatorname{span}(x_i \mid i \in I \cup \{t\} \cup S)$ and $\dim \operatorname{span}(x_i \mid i \in I \cup \{t\} \cup S) \le |I| + 1 + |S| = r$ (recall that the vectors x_i for $i \in I$ are fixed). Thus, the number of possibilities for choosing $x_j \in A$ for $j \in [k]\backslash I$ with the above properties is indeed at most

$$\begin{aligned} k^{k+2} p^r \cdot \Gamma^n \cdot |A|^{|S|} \cdot (p^r)^{k-|I|-|S|-1} &= k^{k+2} p^r \cdot \Gamma^n \cdot |A|^{r-|I|-1} \cdot (p^r)^{k-r} \\ &\le k^{k+2} p^{rk} \cdot \Gamma^n \cdot |A|^{r-|I|-1}. \end{aligned}$$

This finishes the proof of the lemma. □

By adding up the bound in Lemma 4.10 over all possible choices of the subset $I \subseteq [k]$ and the vectors $x_i \in A$ for $i \in I$, we can show the following corollary.

Corollary 4.12 *For any fixed r with $\lfloor w/(k+1) \rfloor + 1 \le r \le k$, the number of solutions $(x_1, \dots, x_k) \in A^k$ to the system $(\star)$ with $\omega(x_1, \dots, x_k) = w$ and $\dim \operatorname{span}(x_1, \dots, x_k) = r$ is at most $(2k)^{2k} p^{rk} \cdot \Gamma^n \cdot |A|^{r-1}$.*

Proof By Claim 4.7 every such solution $(x_1, \dots, x_k)$ satisfies $|\mathcal{I}(x_1, \dots, x_k)| = \lfloor w/(k+1) \rfloor$.

We claim that for each set $I \subseteq [k]$ with $|I| = \lfloor w/(k+1) \rfloor$, there are at most $k^{k+3} 2^k p^{rk} \cdot \Gamma^n \cdot |A|^{r-1}$ solutions $(x_1, \dots, x_k) \in A^k$ to the system $(\star)$ with $\omega(x_1, \dots, x_k) = w$ and $\dim \operatorname{span}(x_1, \dots, x_k) = r$ and $\mathcal{I}(x_1, \dots, x_k) = I$. Indeed, there are $|A|^{|I|}$ possibilities to choose vectors $x_i \in A$ for $i \in I$, and after fixing these vectors the desired number of solutions $(x_1, \dots, x_k) \in A^k$ is by Lemma 4.10 at most $k^{k+3} 2^k p^{rk} \cdot \Gamma^n \cdot |A|^{r-|I|-1}$. So in total we have indeed at most $|A|^{|I|} \cdot k^{k+3} 2^k p^{rk} \cdot \Gamma^n \cdot |A|^{r-|I|-1} = k^{k+3} 2^k p^{rk} \cdot \Gamma^n \cdot |A|^{r-1}$ solutions for any given I.

Since the number of possible subsets $I \subseteq [k]$ with $|I| = \lfloor w/(k+1) \rfloor$ is clearly at most 2^k, the number of solutions $(x_1, \dots, x_k) \in A^k$ as in the statement of the corollary is at most

$$2^k \cdot k^{k+3} 2^k p^{rk} \cdot \Gamma^n \cdot |A|^{r-1} \le (2k)^{2k} p^{rk} \cdot \Gamma^n \cdot |A|^{r-1}.$$

Here, we used that $k \ge 3$ since $k \ge 2m + 1 + \lfloor w/(k+1) \rfloor$ and $m \ge 1$. □

As mentioned above, we now perform a random subspace sampling argument. Choose the unique integer d such that

$$\frac{1}{(2k)^{2k+1} \cdot p^{2k+1}} \cdot \left(\frac{p}{\Gamma}\right)^{n/(k-1)} < p^d \le \frac{1}{(2k)^{2k+1} \cdot p^{2k}} \cdot \left(\frac{p}{\Gamma}\right)^{n/(k-1)}. \tag{4.7}$$

Claim 4.13 *We have $2 \le d \le n$.*

Proof Recall that we assumed that $(p/\Gamma)^n > (2kp^2)^{(2k+1)(k-1)}$. Hence

$$p^d > \frac{1}{(2k)^{2k+1} \cdot p^{2k+1}} \cdot (2kp^2)^{2k+1} = p^{2k+1},$$

and consequently $d \geq 2k+1 \geq 2$. For the upper bound on d, recall that $1 \leq \Gamma < p$, so $p^d \leq (p/\Gamma)^{n/(k-1)} \leq p^n$ and therefore $d \leq n$. □

Let us now consider a uniformly random d-dimensional subspace $V \subseteq \mathbb{F}_p^n$. The following claims give useful bounds for the expected number of vectors in $A \cap V$ and the expected number of solutions $(x_1, \ldots, x_k) \in A^k$ with $\omega(x_1, \ldots, x_k) = w$ and $x_1, \ldots, x_k \in V$.

Claim 4.14 *We have*

$$\mathbb{E}[|A \cap V|] > \frac{3}{4} \cdot \frac{p^d}{p^n} \cdot |A|.$$

Proof Recall that $0 \notin A$. So by Lemma 2.2 (applied with $s = 1$), for each vector $x \in A$ we have

$$\mathbb{P}[x \in V] = \frac{p^d - 1}{p^n - 1} > \frac{3}{4} \cdot \frac{p^d}{p^n}.$$

Here we used that $p^d - 1 \geq (3/4)p^d$ since $p \geq 2$ and $d \geq 2$ (see Claim 4.13). Now, adding this up over all $x \in A$ gives the desired bound. □

Claim 4.15 *For any r with $\lfloor w/(k+1) \rfloor + 1 \leq r \leq k$, the expected number of solutions $(x_1, \ldots, x_k) \in A^k$ to $(\star)$ with $\omega(x_1, \ldots, x_k) = w$ and* $\dim \operatorname{span}(x_1, \ldots, x_k) = r$ *such that $x_1, \ldots, x_k \in V$ is at most*

$$\frac{1}{2k} \cdot \frac{p^d}{p^n} \cdot |A|.$$

Proof By Corollary 4.12, the total number of solutions $(x_1, \ldots, x_k) \in A^k$ to $(\star)$ with $\omega(x_1, \ldots, x_k) = w$ and $\dim \operatorname{span}(x_1, \ldots, x_k) = r$ is at most $(2k)^{2k} p^{rk} \cdot \Gamma^n \cdot |A|^{r-1}$. For each of these solutions, by Lemma 2.2 (applied to a list of r linearly independent vectors among $x_1, \ldots, x_k$), the probability of having $x_1, \ldots, x_k \in V$ is at most $(p^d/p^n)^r$.

Thus, in the case $r \geq 2$ the expected number of solutions in the claim statement is by $|A| \leq p^n$ and (4.7) at most

$$(2k)^{2k} p^{rk} \cdot \Gamma^n \cdot |A|^{r-1} \cdot \left(\frac{p^d}{p^n}\right)^r \leq \frac{1}{2k} \cdot (2k)^{2k+1} p^{rk} \cdot \Gamma^n \cdot |A| \cdot (p^n)^{r-2} \cdot \left(\frac{p^d}{p^n}\right)^{r-1} \cdot \frac{p^d}{p^n}$$
$$= \frac{1}{2k} \cdot \frac{p^d}{p^n} \cdot |A| \cdot (2k)^{2k+1} p^{rk} \cdot \frac{\Gamma^n}{p^n} \cdot (p^d)^{r-1}$$

$$\leq \frac{1}{2k}\cdot\frac{p^d}{p^n}\cdot|A|\cdot(2k)^{2k+1}p^{rk}\cdot\left(\frac{1}{(2k)^{2k+1}\cdot p^{2k}}\right)^{r-1}\cdot\left(\frac{\Gamma}{p}\right)^{n\cdot\left(1-\frac{r-1}{k-1}\right)}\leq\frac{1}{2k}\cdot\frac{p^d}{p^n}\cdot|A|$$

where in the last step we used that $\Gamma < p$. It remains to consider the case $r = 1$. In this case the expected number of solutions as in the claim statement is at most

$$(2k)^{2k}p^{rk}\cdot\Gamma^n\cdot|A|^{r-1}\cdot\left(\frac{p^d}{p^n}\right)^r = (2k)^{2k}p^k\cdot\Gamma^n\cdot\frac{p^d}{p^n}\leq\frac{1}{2k}\cdot|A|\cdot\frac{p^d}{p^n},$$

where we used that $|A| \geq (2kp)^{2k+1}\cdot\Gamma^n \geq (2k)^{2k+1}p^k\cdot\Gamma^n$ by (4.3). □

Corollary 4.16 *The expected number of solutions $(x_1,\dots,x_k)\in A^k$ to $(\star)$ with $\omega(x_1,\dots,x_k) = w$ and $x_1,\dots,x_k\in V$ is at most*

$$\frac{1}{2}\cdot\frac{p^d}{p^n}\cdot|A|.$$

Proof By Claim 4.7, for every solution $(x_1,\dots,x_k)\in A^k$ to $(\star)$ with $\omega(x_1,\dots,x_k) = w$ we must have $\lfloor w/(k+1)\rfloor + 1 \leq \dim\operatorname{span}(x_1,\dots,x_k) \leq k$. Hence, the claim follows from Claim 4.15 by summing over all integers r with $\lfloor w/(k+1)\rfloor+1 \leq r \leq k$ (noting that there are at most k possibilities for r). □

Let Z be the number of solutions $(x_1,\dots,x_k)\in A^k$ to $(\star)$ with $\omega(x_1,\dots,x_k) = w$ and $x_1,\dots,x_k \in V$. By Corollary 4.16, we have $\mathbb{E}[Z] \leq (1/2)\cdot(p^d/p^n)\cdot|A|$. Together with Claim 4.14 this yields

$$\mathbb{E}[|A\cap V| - Z] > \frac{3}{4}\cdot\frac{p^d}{p^n}\cdot|A| - \frac{1}{2}\cdot\frac{p^d}{p^n}\cdot|A| = \frac{1}{4}\cdot\frac{p^d}{p^n}\cdot|A|.$$

So let us fix an outcome for the d-dimensional subspace $V\subseteq\mathbb{F}_p^n$ such that $|A\cap V| - Z > (1/4)\cdot(p^d/p^n)\cdot|A|$.

We can now define a subset $A^*\subseteq A\cap V$ by deleting one vector from each solution $(x_1,\dots,x_k)\in A^k$ to $(\star)$ with $\omega(x_1,\dots,x_k) = w$ and $x_1,\dots,x_k\in V$. Then

$$|A^*| \geq |A\cap V| - Z > \frac{1}{4}\cdot\frac{p^d}{p^n}\cdot|A|$$

and there do not exist any solutions $(x_1,\dots,x_k)\in(A^*)^k$ to $(\star)$ with $\omega(x_1,\dots,x_k) = w$.

Our goal is now to obtain a contradiction by applying the induction hypothesis for $w-1$ to $A^* \subseteq V \cong \mathbb{F}_p^d$. Note that using (4.2) and (4.7), we have

$$|A^*| > \frac{1}{4} \cdot \frac{p^d}{p^n} \cdot |A| \geq \frac{1}{4} \cdot \frac{p^d}{p^n} \cdot 4C'(2kp)^{2k+1} \cdot p^n \cdot \left(\frac{\Gamma}{p}\right)^{c'n/(k-1)}$$
$$\geq C' \cdot p^d \cdot (2kp)^{c'(2k+1)} \cdot \left(\frac{\Gamma}{p}\right)^{c'n/(k-1)} \geq C' \cdot p^{(1-c')d}.$$

Since $V \cong \mathbb{F}_p^d$, we can interpret $A^* \subseteq V$ as a subset of $\mathbb{F}_p^d$. As $|A^*| > C' \cdot p^{(1-c')d}$, by the induction hypothesis for $w-1$, there must be a solution $(x_1, \dots, x_k) \in (A^*)^k$ to the system $(\star)$ with $\omega(x_1, \dots, x_k) > w-1$. By our construction of A^*, we have $\omega(x_1, \dots, x_k) \neq w$. This means that we must have $\omega(x_1, \dots, x_k) > w$, but since $(x_1, \dots, x_k) \in (A^*)^k \subseteq A^k$ this is a contradiction to our assumption on the set A. This finishes the inductive proof of Proposition 4.5.

5 Concluding remarks

Recall that in our main result, Theorem 1.2, we assumed that every $m \times m$ minor of the $m \times k$ matrix $(a_{j,i})_{j,i}$ is non-singular. As discussed in the introduction, it is not possible to remove this assumption completely, but it may be possible to weaken it in some ways.

In order for the conclusion of Theorem 1.2 to hold, it is certainly necessary to assume that no equation of the form $x_i - x_{i'} = 0$ for distinct $i, i' \in [k]$ is in the span of the equations in the system $(\star)$, since otherwise there do not exist any solution $(x_1, \dots, x_k) \in (\mathbb{F}_p^n)^k$ to $(\star)$ where $x_1, \dots, x_k$ are distinct. In other words (using our other assumption that $a_{j,1} + \dots + a_{j,k} = 0$ for $j = 1, \dots, m$, which is also necessary as discussed in the introduction), in Theorem 1.2 we certainly need to assume that the row-span of the $m \times k$ matrix $(a_{j,i})_{j,i}$ does not contain any vector with exactly two non-zero entries.

If we assume that $a_{j,1} + \dots + a_{j,k} = 0$ for $j = 1, \dots, m$ and that the row-span of the $m \times k$ matrix $(a_{j,i})_{j,i}$ does not contain any vector with exactly two non-zero entries, then every subset $A \subseteq \mathbb{F}_p^n$ not containing solution $(x_1, \dots, x_k) \in A^k$ to $(\star)$ with distinct $x_1, \dots, x_k$ must have size $|A| = o(p^n)$ as $n \to \infty$ (where p, m and k are fixed). This follows from an arithmetic removal lemma for solutions to systems of linear equations due to Král', Serra, and Vena [16,Theorem 1] and independently Shapira [27,Theorem 2.2]. It would be plausible that one also has a bound of the form $|A| \leq C_{p,m,k} \cdot (\Gamma^*_{p,m,k})^n$ with $\Gamma^*_{p,m,k} < p$ under these weaker assumptions, i.e. it would be plausible that Theorem 1.2 also holds with these weaker assumptions.

However, proving Theorem 1.2 under these weaker assumptions is most likely extremely difficult. In fact, proving such a statement (even with an assumption that the number of variables is very large in terms of the number of equations) would imply a bound of the form $|A| \leq C_{p,k} \cdot (\Gamma^*_{p,k})^n$ with $\Gamma^*_{p,k} < p$ for the size of a k-term-progression-free subset $A \subseteq \mathbb{F}_p^n$. Indeed, for any (large) K, one can take $(\star)$

to be the system of $k-1$ equations in $k+K$ variables consisting of the equations $x_i - 2x_{i+1} + x_{i+2} = 0$ for $i = 1, \dots, k-2$ as well as the equation $Kx_k - x_{k+1} - \dots - x_{k+K} = 0$. Then $a_{j,1} + \dots + a_{j,k} = 0$ for $j = 1, \dots, k-1$ and the row-span of the $(k-1) \times (k+K)$ matrix $(a_{j,i})_{j,i}$ does not contain any vector with exactly two non-zero entries. But for any solution $(x_1, \dots, x_{k+K}) \in A^{k+K}$ with distinct $x_1, \dots, x_{k+K}$, the vectors $x_1, \dots, x_k$ form a non-constant k-term arithmetic progression. Hence any k-term-progression-free subset $A \subseteq \mathbb{F}_p^n$ in particular does not contain a solution $(x_1, \dots, x_{k+K}) \in A^{k+K}$ to this system of equations with distinct $x_1, \dots, x_k$.

For $k \geq 4$, proving that a k-term-progression-free subset $A \subseteq \mathbb{F}_p^n$ has size $|A| \leq C_{p,k} \cdot (\Gamma_{p,k}^*)^n$ with $\Gamma_{p,k}^* < p$ is a big open problem, that has received the attention of many researchers, especially after Ellenberg and Gijswijt [6] proved such a statement for $k = 3$. Weakening the assumptions on the matrix $(a_{j,i})_{j,i}$ in Theorem 1.2 in the way discussed above is at least as difficult a problem, and therefore seems to be out of reach of current methods.

A more tractable problem might be to improve upon the value of the constant $\Gamma_{p,m,k}^* < p$ in Theorem 1.2 that our proof obtains. Given the inductive nature of the proof with the repeated subspace sampling arguments, this value is likely not optimal. It would be extremely interesting to determine whether one can take $\Gamma_{p,m,k}^*$ to be equal to the constant $\Gamma_{p,m,k}$ in Theorem 1.1 as defined in (3.1). Even in the case where ($\star$) consists only of one equation (i.e. in the case $m = 1$) this is a widely open problem, and in the special case of the equation $x_1 + \dots + x_p = 0$ it has applications to bounding Erdős-Ginzburg-Ziv constants (see [10, 22, 25]).

Acknowledgements The author would like to thank Max Wenqiang Xu and the referee for helpful comments. This work was completed while the author was at the Institute for Advanced Study in Princeton, NJ.

Funding Open Access funding provided by the MIT Libraries.This research was supported by NSF Grant CCF-1900460 and NSF Award DMS-2100157.

Availability of data and material. Not applicable.

Declarations

Conflict of interest The author states that there is no conflict of interest.

Code availability Not applicable.

References

1. Bateman, M., Katz, N.H.: New bounds on cap sets. J. Am. Math. Soc. **25**, 585–613 (2012)
2. Blasiak, J., Church, T., Cohn, H., Grochow, J.A., Naslund, E., Sawin, W.F., Umans, C.: *On cap sets and the group-theoretic approach to matrix multiplication*, Discrete Anal. 2017, Paper No. 3, 27pp
3. Cilleruelo, J.: Combinatorial problems in finite fields and Sidon sets. Combinatorica **32**, 497–511 (2012)
4. Croot, E., Lev, V.F., Pach, P.P.: Progression-free sets in $\mathbb{Z}_4^n$ are exponentially small. Ann. Math. **185**, 331–337 (2017)
5. Edel, Y.: Extensions of generalized product caps. Des. Codes Cryptogr. **31**, 5–14 (2004)
6. Ellenberg, J.S., Gijswijt, D.: On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression. Ann. Math. **185**, 339–343 (2017)
7. Elsholtz, C., Pach, P.P.: Caps and progression-free sets in $\mathbb{Z}_m^n$. Des. Codes Cryptogr. **88**, 2133–2170 (2020)
8. Erdős, P., Ginzburg, A., Ziv, A.: Theorem in the additive number theory. Bull. Res. Council Israel **10F**, 41–43 (1961)
9. Erdős, P., Turán, P.: On Some Sequences of Integers. J. Lond. Math. Soc. **11**, 261–264 (1936)
10. Fox, J., Sauermann, L.: Erdős-Ginzburg-Ziv constants by avoiding three-term arithmetic progressions. Electron. J. Combin. **25**, 9 (2018). (**Paper 2.14**)
11. Green, B., Tao, T.: An inverse theorem for the Gowers $U^3(G)$- norm, with applications. Proc. Edinb. Math. Soc. **51**, 73–153 (2008)
12. Green, B., Tao, T.: New bounds for Szemerédi's theorem, I: Progressions of length 4 in finite field geometries. Proc. Lond. Math. Soc. **98**, 365–392 (2009)
13. Green, B., Tao, T.: *New bounds for Szemerédi's theorem, Ia: Progressions of length 4 in finite field geometries revisited*, preprint, arXiv:1205.1330 (2012)
14. Grochow, J.A.: New applications of the polynomial method: the cap set conjecture and beyond. Bull. Am. Math. Soc. **56**, 29–64 (2019)
15. Komlós, J., Sulyok, M., Szemerédi, E.: Linear problems in combinatorial number theory. Acta Math. Acad. Sci. Hungar. **26**, 113–121 (1975)
16. Král', D., Serra, O., Vena, L.: A removal lemma for systems of linear equations over finite fields. Israel J. Math. **187**, 193–207 (2012)
17. Lin, Y., Wolf, J.: On subsets of $\mathbb{F}_q^n$ containing no k- term progressions. Eur. J. Combin. **31**, 1398–1403 (2010)
18. Meshulam, R.: On subsets of finite abelian groups with no 3-term arithmetic progressions. J. Combin. Theory Ser. A **71**, 168–172 (1995)
19. Mimura, M., Tokushige, N.: *Avoiding a shape, and the slice rank method for a system of equations*, preprint, (2019) arXiv:1909.10509
20. Mimura, M., Tokushige, N.: *Solving linear equations in a vector space over a finite field*, Discrete Math. **344** (2021), Paper No. 112603, 11 pp
21. Mimura, M., Tokushige, N.: *Solving linear equations in a vector space over a finite field II*, preprint, (2020) http://www.cc.u-ryukyu.ac.jp/~hide/sol2.pdf
22. Naslund, E.: Exponential Bounds for the Erdős-Ginzburg-Ziv Constant. J. Combin. Theory Ser. A **174**, 19 (2020). (**105185**)
23. Ruzsa, I.Z.: Solving a linear equation in a set of integers I. Acta Arith. **65**, 259–282 (1993)
24. Ruzsa, I.Z.: Solving a linear equation in a set of integers II. Acta Arith. **72**, 385–397 (1995)
25. Sauermann, L.: On the size of subsets of $\mathbb{F}_p^n$ without p distinct elements summing to zero. Israel J. Math. **243**, 63–79 (2021)
26. Sawin, W., Tao, T.: *Notes on the "slice rank" of tensors*, blog post, 2016, https://terrytao.wordpress.com/2016/08/24/notes-on-the-slice-rank-of-tensors/
27. Shapira, A.: A proof of Green's conjecture regarding the removal properties of sets of linear equations. J. Lond. Math. Soc. **81**, 355–373 (2010)
28. Tao, T.: *A symmetric formulation of the Croot-Lev-Pach-Ellenberg-Gijswijt capset bound*, blog post, (2016) http://terrytao.wordpress.com/2016/05/18/a

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.