

On the size of subsets of F_q^n avoiding solutions to linear systems with repeated columns

de Bruyn, Josse van Dobben; Gijswijt, Dion

DOI

[10.37236/10883](https://doi.org/10.37236/10883)

Publication date

2023

Document Version

Final published version

Published in

Electronic Journal of Combinatorics

Citation (APA)

de Bruyn, J. V. D., & Gijswijt, D. (2023). On the size of subsets of F_q^n avoiding solutions to linear systems with repeated columns. *Electronic Journal of Combinatorics*, 30(4). <https://doi.org/10.37236/10883>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

On the size of subsets of $\mathbb{F}_q^n$ avoiding solutions to linear systems with repeated columns

Josse van Dobben de Bruyn

Dion Gijswijt

Submitted: Nov 19, 2021; Accepted: Aug 28, 2023; Published: Oct 6, 2023

©The authors. Released under the CC BY license (International 4.0).

Abstract

Consider a system of m balanced linear equations in k variables with coefficients in $\mathbb{F}_q$. If $k \geq 2m + 1$, then a routine application of the slice rank method shows that there are constants $\beta, \gamma \geq 1$ with $\gamma < q$ such that, for every subset $S \subseteq \mathbb{F}_q^n$ of size at least $\beta \cdot \gamma^n$, the system has a solution $(x_1, \dots, x_k) \in S^k$ with $x_1, \dots, x_k$ not all equal. Building on a series of papers by Mimura and Tokushige and on a paper by Sauermann, this paper investigates the problem of finding a solution of higher non-degeneracy; that is, a solution where $x_1, \dots, x_k$ are pairwise distinct, or even a solution where $x_1, \dots, x_k$ do not satisfy any balanced linear equation that is not a linear combination of the equations in the system.

In this paper, we focus on linear systems with repeated columns. For a large class of systems of this type, we prove that there are constants $\beta, \gamma \geq 1$ with $\gamma < q$ such that every subset $S \subseteq \mathbb{F}_q^n$ of size at least $\beta \cdot \gamma^n$ contains a solution that is non-degenerate (in one of the two senses described above). This class is disjoint from the class covered by Sauermann's result, and captures the systems studied by Mimura and Tokushige into a single proof. Moreover, a special case of our results shows that, if $S \subseteq \mathbb{F}_p^n$ is a subset such that $S - S$ does not contain a non-trivial k -term arithmetic progression (with p prime and $3 \leq k \leq p$), then S must have exponentially small density.

Mathematics Subject Classifications: 05D40, 11B25

1 Introduction

1.1 Background and prior results

For several decades, one of the major open problems in extremal combinatorics had been to determine whether or not, for a given prime $p \geq 3$, there is a constant $c_p < p$ such that every subset $S \subseteq \mathbb{F}_p^n$ of size $|S| \geq c_p^n$ contains a non-trivial 3-term arithmetic progression;

Delft Institute of Applied Mathematics, Delft University of Technology, The Netherlands
(jdob@dtu.dk, d.c.gijswijt@tudelft.nl).

that is, a solution to the equation $x - 2y + z = 0$ with $x, y, z \in S$ distinct. For $p = 3$, this problem was known as the *cap set problem*.

In 2016, Ellenberg and Gijswijt [EG17] solved this problem for all primes $p \geq 3$, using a new application of the polynomial method developed by Croot, Lev and Pach [CLP17]. The solution was subsequently recast by Tao [Tao16] in terms of the slice rank of tensors (or hypermatrices). Together, these developments have led to a surge of interest in problems related to the cap set problem, using the slice rank polynomial method to attempt to solve other problems.

One of these related open problems is to find the largest size of a subset of $\mathbb{F}_p^n$ without a non-trivial k -term arithmetic progression, where $p \geq k \geq 4$ and $n \rightarrow \infty$. It is not known whether or not there is a constant $c_{p,k} < p$ such that every set $S \subseteq \mathbb{F}_p^n$ with $|S| \geq c_{p,k}^n$ contains a k -term arithmetic progression. This problem is believed to be beyond the reach of current slice rank methods.

Instead, mathematicians have turned their attention to related problems. Recently, Mimura and Tokushige [MT19a, MT19b, MT20] and Sauermann [Sau23] have started developing techniques to bound the maximum size of a subset of $\mathbb{F}_q^n$ which avoids non-degenerate solutions to a given system of linear equations over a finite field $\mathbb{F}_q$. More formally, given a fixed matrix $A = (a_{ij}) \in \mathbb{F}_q^{m \times k}$, we want to bound the maximum size of a subset $S \subseteq \mathbb{F}_q^n$ for which there are no k -tuples $(\mathbf{x}_1, \dots, \mathbf{x}_k) \in S^k$ satisfying

$$\begin{cases} a_{11}\mathbf{x}_1 + \dots + a_{1k}\mathbf{x}_k = 0, \\ \vdots \\ a_{m1}\mathbf{x}_1 + \dots + a_{mk}\mathbf{x}_k = 0; \end{cases} \quad (\star)$$

except possibly trivial/degenerate solutions (more on that later). Note that the variables $\mathbf{x}_1, \dots, \mathbf{x}_k$ are not taken from $\mathbb{F}_q$, but from $\mathbb{F}_q^n$ as $n \rightarrow \infty$.

If $a_{i1} + \dots + a_{ik} \neq 0$ for some i (i.e. the coefficients in one of the rows do not sum to zero), then there are large subsets of $\mathbb{F}_q^n$ with no solutions at all to $(\star)$. Indeed, let $S \subseteq \mathbb{F}_q^n$ be the set of all vectors whose first coordinate is equal to 1. If some row of $(\star)$ does not sum to zero, then S does not contain solutions to $(\star)$, and $|S| = q^{n-1} = \frac{1}{q} \cdot |\mathbb{F}_q^n|$, so S contains a constant proportion of the vectors in $\mathbb{F}_q^n$. (This example is due to Sauermann [Sau23].)

We will henceforth assume that $a_{i1} + \dots + a_{ik} = 0$ for all i . Such equations are called *balanced linear equations* (or *affine dependences*), and the system $(\star)$ is also called *balanced*. Recent results show that the problem becomes much more interesting in this case.

If the system $(\star)$ is balanced, then every set $S \subseteq \mathbb{F}_q^n$ has at least $|S|$ solutions to $(\star)$, namely the solutions of the form $(a, \dots, a)$ for $a \in S$. So the question is: how large does S have to be to guarantee the existence of solutions to $(\star)$ which are somehow non-degenerate? For this we consider three different notions of non-degeneracy:

Definition 1.1. A solution $(\mathbf{x}_1, \dots, \mathbf{x}_k) \in (\mathbb{F}_q^n)^k$ of $(\star)$ is called:

- (a) *non-trivial* if $\mathbf{x}_1, \dots, \mathbf{x}_k$ are not all equal.

- (b) a $(\star)$ -shape¹ if $\mathbf{x}_1, \dots, \mathbf{x}_k$ are pairwise distinct.
- (c) *generic*² if every balanced linear equation (over $\mathbb{F}_q$) satisfied by $(\mathbf{x}_1, \dots, \mathbf{x}_k)$ is a linear combination of the equations in $(\star)$.

The requirements get stronger in each step, moving from (a) to (c). Indeed, it is clear that every $(\star)$ -shape is a non-trivial solution. Furthermore, if the system $(\star)$ does not rule out the existence of $(\star)$ -shapes in $\mathbb{F}_q^n$ (in other words, if no linear combination of the equations in $(\star)$ equals $\mathbf{x}_i - \mathbf{x}_j = 0$ for some $i \neq j$), then every generic solution is a $(\star)$ -shape.

The easiest of these problems is finding a non-trivial solution. If the number of variables is sufficiently large (specifically, if $k \geq 2m + 1$), then this can be done by a routine application of the slice rank method.

Theorem 1.2 ([Tao16], see also [Sau23, Theorem 1.1]³). *If $k \geq 2m + 1$, then there exists a constant $\Gamma_{q,m,k} < q$ such that every subset $S \subseteq \mathbb{F}_q^n$ of size at least $(\Gamma_{q,m,k})^n$ has a non-trivial solution of $(\star)$.*

If $k \leq 2m$, then the problem of finding non-trivial bounds is believed to be beyond the reach of current (slice rank) methods. Accordingly, most⁴ of our results are on systems with $k \geq 2m + 1$; see Remark 2.4(i).

The aim of this paper is to refine Theorem 1.2 to the stronger notions of non-degeneracy from Definition 1.1. For this we use the following terminology:

Definition 1.3. The linear system $(\star)$ is called:

- (a) *moderate*¹ if there exist constants $\beta, \gamma > 0$ with $\gamma < q$ such that every subset $S \subseteq \mathbb{F}_q^n$ of size at least $\beta \cdot \gamma^n$ contains a $(\star)$ -shape;
- (b) *temperate*² if there exist constants $\beta, \gamma > 0$ with $\gamma < q$ such that every subset $S \subseteq \mathbb{F}_q^n$ of size at least $\beta \cdot \gamma^n$ contains a generic solution of $(\star)$.

If $(\star)$ consists of the single equation $\mathbf{x}_1 + \dots + \mathbf{x}_p = 0$ over $\mathbb{F}_p$ (with p prime), then the existence of $(\star)$ -shapes is tightly linked to the Erdős–Ginzburg–Ziv constant of the group $\mathbb{F}_p^n$. If $p \geq 3$, then this system is moderate over $\mathbb{F}_p$; this is implicit in [Nas20] and [Sau21]. Furthermore, the method in [Sau21] can be easily adapted to show that every balanced linear equation with at least 3 variables forms a moderate linear system.

The problem of determining whether or not a system of two or more equations is moderate was first studied by Mimura and Tokushige [MT19a, MT19b, MT20].⁵ They

¹Following terminology from Mimura and Tokushige [MT19a, MT19b, MT20].

²Terminology introduced by the authors.

³To get rid of the constant factor $C_{q,m,k}$ from [Sau23, Theorem 1.1], use the power trick.

⁴The only exception is when we study different, related problems for which the assumption $k \geq 2m + 1$ is not necessary (such as Lemma 4.1, which is one of the main tools in our proofs).

⁵Similar results over the integers had been obtained by Ruzsa in the 1990s [Ruz93, Ruz95], but Mimura and Tokushige were the first to study this problem for vector spaces over a finite field.

showed that several specific linear systems are moderate. Although all of their proofs rely on more or less the same idea, the details of the proofs are so different that a new proof is needed for each new system. We discuss some of their results in more detail in §7.

The first general result in this direction was found by Sauermann [Sau23]. In an elaborate proof, using a new application of the slice rank method and a subspace sampling argument, she showed that $(\star)$ -shapes can always be found if the number of variables is sufficiently large and if the system is very much non-degenerate:

Theorem 1.4 ([Sau23, Theorem 1.2]). *If $k \geq 3m$ and every $m \times m$ submatrix of A is invertible, then $(\star)$ is moderate.*

Despite its generality, this result does not replace the results of Mimura and Tokushige, because the systems they studied have many singular $m \times m$ submatrices (so Theorem 1.4 does not apply).

The third and final problem is that of finding a generic solution. A partial result in this direction was found by Sauermann, who showed that solutions of higher dimension exist as the number of variables becomes larger:

Theorem 1.5 ([Sau23, Theorem 1.3]). *If $r \geq 2$ and $k \geq 2m - 1 + r$, then there are constants $C_{p,m,k,r}^{\text{rank}} \geq 1$ and $\Gamma_{p,m,k,r}^{\text{rank}} < p$ such that every subset $S \subseteq \mathbb{F}_p^n$ of size at least $C_{p,m,k,r}^{\text{rank}} \cdot (\Gamma_{p,m,k,r}^{\text{rank}})^n$ has a solution $(\mathbf{x}_1, \dots, \mathbf{x}_k) \in S^k$ of $(\star)$ satisfying $\dim(\text{span}(\mathbf{x}_1, \dots, \mathbf{x}_k)) \geq r$.*

Finding solutions of high dimension is closely related to finding a generic solution, as we explain in §5.

1.2 Main results of this paper

The main results of this paper are twofold. First, we prove a general result on finding $(\star)$ -shapes, which contains most of the results from [MT19a, MT19b, MT20] as special cases. Second, we prove a general result for finding generic solutions, which we believe to be the first of its kind.

Throughout the paper, we focus on a specific class of systems that is completely different from the class of systems studied by Sauermann. Where Sauermann's result (Theorem 1.4 above) requires every $m \times m$ submatrix to be invertible, we require the opposite: there must be sufficiently many linear dependencies between the columns. Specifically, we focus on the class of ‘type (RC)’ linear systems, which we define as follows:

Definition 1.6. Consider the linear system $(\star)$, whose coefficients are specified by the matrix $A = (a_{ij}) \in \mathbb{F}_q^{m \times k}$.

- (a) We say that two indices in $[k]$ are *equivalent* if the corresponding columns of A are nonzero scalar multiples of one another. This defines an equivalence relation on $[k]$. We will refer to the equivalence classes of this equivalence relation as the *column equivalence classes*.

- (b) We say that $(\star)$ is a *type (RC) linear system*⁶ if it is balanced and has at most one column equivalence class of size 1.
- (c) We say that a column equivalence class *sums to zero* if the columns indexed by that class add up to the zero vector.

Examples of type (RC) linear systems will be given in §7 below. Among these examples are the systems studied by Mimura and Tokushige.

The assumptions made throughout this paper can be summarized as follows:

Situation 1.7. Let $(\star)$ be a type (RC) linear system, given by the coefficient matrix $A = (a_{ij}) \in \mathbb{F}_q^{m \times k}$, with ℓ column equivalence classes. Furthermore, assume that $(\star)$ is non-degenerate and irreducible (see Definition 2.1 below).

In all of our main results below, we assume that $(\star)$ and A are as in Situation 1.7. In particular, we always assume that $(\star)$ is irreducible. However, we note that our results can also be applied to reducible systems. We show in Proposition 2.2 (resp. Proposition 5.4) that a system is moderate (resp. temperate) if and only if every irreducible subsystem is moderate (resp. temperate).

Our first main result is a sufficient condition for a type (RC) linear system to be moderate.

Theorem A. *Let $(\star)$, A , m , k and ℓ be as in Situation 1.7. Suppose that $(\star)$ satisfies at least one of the following additional properties:*

- (i) *none of the column equivalence classes of size 2 sums to zero;*
- (ii) *every column equivalence class sums to zero, and $k \geq 3$.*

Then $(\star)$ is moderate.

This result encompasses most of the systems studied by Mimura and Tokushige, and the rest can be recovered using a slight modification of our proof. See §7 for a detailed discussion.

Our second main result is a sufficient condition for a type (RC) linear system to be temperate.

Theorem B. *Let $(\star)$, A , m , k and ℓ be as in Situation 1.7. Suppose that $(\star)$ satisfies at least one of the following additional properties:*

- (i) *$\ell = m + 1$;*
- (ii) *every column equivalence class sums to zero.*

Then $(\star)$ is temperate.

⁶Terminology introduced by the authors ('RC' stands for 'repeated columns').

The requirements of Theorem B are more restrictive than those of Theorem A (see Remark 2.4(ii)).⁷ In particular, one of the systems studied by Mimura and Tokushige does not meet these requirements (see §7 for a detailed discussion).

We do not know if every irreducible linear system of type (RC) is moderate and/or temperate, but we have the following partial result. We say that a balanced linear equation satisfied by $(\mathbf{x}_1, \dots, \mathbf{x}_k) \in S^k$ preserves the column equivalence classes of $(\star)$ if appending that equation to the system $(\star)$ preserves the column equivalence classes. We prove the following:

Theorem C. *Let $(\star)$, A , m , k and ℓ be as in Situation 1.7. Then there exist constants $\beta, \gamma > 0$ with $\gamma < q$ such that every subset $S \subseteq \mathbb{F}_q^n$ of size at least $\beta \cdot \gamma^n$ has a solution $(\mathbf{x}_1, \dots, \mathbf{x}_k) \in S^k$ of $(\star)$ with the following properties:*

- (i) *every balanced linear equation satisfied by $(\mathbf{x}_1, \dots, \mathbf{x}_k)$ preserves the column equivalence classes of $(\star)$;*
- (ii) $\dim(\text{aff}(\mathbf{x}_1, \dots, \mathbf{x}_k)) \geq \min(k - \ell, k - 2)$.

Theorem C improves upon Theorem 1.5 whenever $2 \leq \ell < 2m$; see Remark 6.8.

Finally, we turn to an application of our techniques and results. In characteristic 0, results like Bourgain's theorem [Bou90] (see also [TV06, Chapter 12]) show that it is substantially easier to find long arithmetic progressions in sum sets than in general sets. Using the techniques from this paper, we establish a similar result in vector spaces over $\mathbb{F}_q$.

Given sets $S_1, \dots, S_l \subseteq \mathbb{F}_q^n$, we define the *affinely independent restricted sum set* (or *AIR-sumset*) as follows:

$$S_1 \overset{\text{aff}}{+} \dots \overset{\text{aff}}{+} S_l := \{\mathbf{x}_1 + \dots + \mathbf{x}_l \mid \mathbf{x}_1 \in S_1, \dots, \mathbf{x}_l \in S_l \text{ affinely independent}\}.$$

Further, if $(\star)$ is linear system which is not necessarily balanced, then we say that a solution $(\mathbf{x}_1, \dots, \mathbf{x}_k) \in (\mathbb{F}_q^n)^k$ is *linearly generic* if every linear equation (over $\mathbb{F}_q$) satisfied by $(\mathbf{x}_1, \dots, \mathbf{x}_k)$ is a linear combination of the equations in $(\star)$. By comparison, the solutions which we call *generic* throughout this paper (see Definition 1.1(c)) only satisfy this property for *balanced* linear equations (so by 'generic' we will always mean 'affinely generic').

Corollary D. *Let $\mathbb{F}_q$ be a finite field, let $(\star)$ be a (not necessarily balanced) linear system over $\mathbb{F}_q$, and let $c_1, \dots, c_l \in \mathbb{F}_q \setminus \{0\}$ with $c_1 + \dots + c_l = 0$. Then there are constants $\beta, \gamma \geq 1$ with $\gamma < q$ such that, for every subset $S \subseteq \mathbb{F}_q^n$ of size at least $\beta \cdot \gamma^n$, the set $(c_1 \cdot S \overset{\text{aff}}{+} \dots \overset{\text{aff}}{+} c_l \cdot S) \cup \{0\}$ contains a linearly generic solution of $(\star)$.*

Note that Corollary D does not impose any restriction on the linear system $(\star)$; that is, the coefficient matrix $A \in \mathbb{F}_q^{m \times k}$ can be arbitrary. This is a significant difference with

⁷Except that Theorem B(ii) does not have the condition $k \geq 3$. That condition is included in Theorem A to rule out the system $\mathbf{x}_1 - \mathbf{x}_2 = 0$. It is not hard to see that this particular system is temperate but not moderate.

our main results and Sauermann's result (Theorem 1.4 above), which only work for very specific classes of linear systems.

In Corollary D, we only need to append 0 to the AIR-sumset when one of the single-variable equations $\mathbf{x}_j = 0$ ($j \in [k]$) can be written as a linear combination of the equations in the linear system $(\star)$. If this is not the case, then a linearly generic solution $(\mathbf{x}_1, \dots, \mathbf{x}_k)$ will satisfy $\mathbf{x}_j \neq 0$ for all $j \in [k]$, so it is not necessary to append 0 to the AIR-sumset.

By letting $(\star)$ be the system that encodes a k -term arithmetic progression, Corollary D contains the following special case:

Corollary E. *Let p be prime, and let $3 \leq k \leq p$. Then, for every subset $S \subseteq \mathbb{F}_p^n$ of size at least $p^{1+(1-\frac{1}{k})n}$, the set $(S - S) \setminus \{0\}$ contains a non-trivial k -term arithmetic progression.*

We note that this special case can be proved without using the slice rank method, using only a simple counting argument (see §7 for details).

1.3 Overview of the main ideas and organization of this paper

Main ideas. There are two new techniques in this paper.

First, the majority of our results depend on a ‘replacement trick’. This trick works roughly as follows. If the j_1 -th and j_2 -th columns of A are non-zero multiples of one another, and if we have a long enough list $\{(\mathbf{x}_1^{(i)}, \dots, \mathbf{x}_k^{(i)})\}_{i=1}^L$ of pairwise disjoint solutions to $(\star)$, then we use tricoloured sum-free sets to recombine these solutions to obtain new solutions of $(\star)$. This is done by taking one of the solutions from this list, say $(\mathbf{x}_1^{(i)}, \dots, \mathbf{x}_k^{(i)})$, and replacing $\mathbf{x}_{j_1}^{(i)}$ and $\mathbf{x}_{j_2}^{(i)}$ by (respectively) $\mathbf{x}_{j_1}^{(i')}$ and $\mathbf{x}_{j_2}^{(i')}$, for some $i', i'' \neq i$. We show in Corollary 3.6 that there exists $i \in [L]$ which admits one such replacement (the ‘single replacement trick’), and in Corollary 6.2 that there exists $i \in [L]$ which admits many replacements (the ‘multiple replacement trick’).

The second main ingredient in our proofs is Lemma 4.1, which shows that, for every subset $S \subseteq \mathbb{F}_q^n$ of size at least $q^{1+(1-\frac{1}{k})n}$, the difference set $S - S$ contains linearly generic solutions to every linear system in k variables (even systems with $k < 2m + 1$). The proof relies only on a simple counting argument, using the pigeonhole principle.

We point out that this paper does not make use of the full strength of Theorem 1.2, as we only use the slice rank method for 3-tensors. Indeed, the replacement trick relies on tricoloured sum-free sets, and Lemma 4.1 does not rely on the slice rank method at all.

The constants. Theorem A(i), Theorem B(i), and Theorem C rely only on the replacement trick. Hence, the base of the exponent in the upper bounds from these theorems⁸ is equal to Γ_q , the constant from the bound on tricoloured sum-free sets (see Theorem 3.2).

Theorem A(ii), Theorem B(ii), and Corollary D rely on a combination of the replacement trick and Lemma 4.1. Hence, the base of the exponent in the upper bounds from these theorems is the maximum of Γ_q and $q^{\frac{k-1}{k}}$.

Corollary E relies solely on Lemma 4.1. The base of the exponent in the upper bound is $p^{1-\frac{1}{k}}$.

⁸By ‘the base of the exponent in the upper bound’, we mean the constant $\gamma < q$ in the upper bound $\beta \cdot \gamma^n$.

Organization of the paper. This paper consists of three parts.

First, in §2–4, we focus on moderate systems. In §2, we discuss the generalities of moderate systems, and we show that we may restrict our attention to irreducible systems. In §3, we establish the ‘single replacement trick’, and use it to prove Theorem A(i). In §4, we establish the other main technique of this paper (Lemma 4.1), and combine it with the replacement trick to prove Theorem A(ii).

Second, in §5–6, we focus on temperate systems. In §5, we discuss the generalities of temperate systems. Here we show how the problem of finding solutions of high rank is related to the problem of finding a generic solution, and we show that we may once again restrict our attention to irreducible systems. In §6, we establish the ‘multiple replacement trick’, and use it to prove Theorem B and Theorem C.

Finally, in §7, we discuss several examples and applications. Here we prove Corollary D and Corollary E, and we recover most of the results from [MT19a, MT19b, MT20] as special cases of our results. Furthermore, we show that the system conjectured to be moderate in [MT20] is indeed moderate.

2 Preliminaries on moderate systems

In this paper, we study linear systems of the form

$$\begin{cases} a_{11}\mathbf{x}_1 + \cdots + a_{1k}\mathbf{x}_k = 0, \\ \vdots \\ a_{m1}\mathbf{x}_1 + \cdots + a_{mk}\mathbf{x}_k = 0; \end{cases} \quad (\star)$$

with coefficient matrix $A = (a_{ij}) \in \mathbb{F}_q^{m \times k}$ and variables $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{F}_q^n$.

Following standard usage, we say that two linear systems $(\star)$ and $(\star')$ are *equivalent* if each equation in $(\star)$ is a linear combination of the equations in $(\star')$ and vice versa. Furthermore, we say that a variable $\mathbf{x}_i$ is *used* by the linear system $(\star)$ if it occurs with non-zero coefficient in at least one equation.

Definition 2.1. The linear system $(\star)$ is said to be:

- (a) *non-degenerate* if the rows of A are linearly independent and every variable is used (equivalently: A has rank m and A has no zero columns);
- (b) *reducible* if it is equivalent to a linear system $(\star')$ with the property that the variables $\mathbf{x}_1, \dots, \mathbf{x}_k$ can be partitioned into two or more classes in such a way that every equation in $(\star')$ only uses variables from one partition class. If this is not the case, then $(\star)$ is said to be *irreducible*.

Passing to an equivalent system or deleting columns with only zeroes does not change the problem of finding a $(\star)$ -shape, so we may assume without loss of generality that $(\star)$ is non-degenerate. The following proposition shows that we can also restrict our attention to irreducible systems.

Proposition 2.2. Suppose that $(\star)$ is equivalent to a linear system $(\star')$ whose coefficient matrix can be written as

$$\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$$

for some $A_1 \in \mathbb{F}_q^{m_1 \times k_1}$ and $A_2 \in \mathbb{F}_q^{m_2 \times k_2}$ with $m_1, m_2, k_1, k_2 \neq 0$. Then $(\star)$ is moderate if and only if the systems given by A_1 and A_2 are moderate.

Proof. If $(\star')$ is moderate with constants $\beta, \gamma > 0$, where $\gamma < q$, then the same holds for the systems given by A_1 and A_2 as every $(\star')$ -shape $(\mathbf{x}_1, \dots, \mathbf{x}_{k_1+k_2})$ yields an A_1 -shape $(\mathbf{x}_1, \dots, \mathbf{x}_{k_1})$ and an A_2 -shape $(\mathbf{x}_{k_1+1}, \dots, \mathbf{x}_{k_1+k_2})$.

Conversely, suppose that for $i = 1, 2$, the system given by A_i is moderate, with constants $\beta_i, \gamma_i > 0$, where $\gamma_i < q$. Let $S \subseteq \mathbb{F}_q^n$ be a set of size at least $\max(\beta_1 \gamma_1^n, k_1 + \beta_2 \gamma_2^n)$. Since $|S| \geq \beta_1 \gamma_1^n$, we may choose an A_1 -shape $(\mathbf{x}_1, \dots, \mathbf{x}_{k_1})$ in S . Since $|S \setminus \{\mathbf{x}_1, \dots, \mathbf{x}_{k_1}\}| \geq \beta_2 \gamma_2^n$, we may choose an A_2 -shape $(\mathbf{y}_1, \dots, \mathbf{y}_{k_2})$ in $S \setminus \{\mathbf{x}_1, \dots, \mathbf{x}_{k_1}\}$. Then $(\mathbf{x}_1, \dots, \mathbf{x}_{k_1}, \mathbf{y}_1, \dots, \mathbf{y}_{k_2})$ is a $(\star')$ -shape. Since $\max(\beta_1 \gamma_1^n, k_1 + \beta_2 \gamma_2^n) \in \mathcal{O}(\max(\gamma_1, \gamma_2)^n)$, this shows that $(\star')$, and therefore $(\star)$, is moderate. $\square$

Therefore we may restrict our attention to irreducible systems, as stipulated in Situation 1.7.

The following proposition will be useful later on.

Proposition 2.3. Let $(\star)$ be a linear system given by the matrix $A = (a_{ij}) \in \mathbb{F}_q^{m \times k}$. If $(\star)$ is non-degenerate and irreducible, and if $m \geq 2$, then every non-zero linear equation implied by $(\star)$ uses at least two column equivalence classes, and $\ell \geq m + 1$.

Proof. Let ℓ be the number of column equivalence classes, and note that $m = \text{rank}(A) \leq \ell$ (recall that the columns with indices in the same column equivalence class are scalar multiples of each other). Suppose for the sake of contradiction that some linear combination of the rows of $(\star)$ uses exactly one column equivalence class. By passing to an equivalent system and permuting the columns, we may assume without loss of generality that the first row of $(\star)$ only uses the column equivalence class $C = \{1, \dots, |C|\} \subseteq [k]$. Since the columns indexed by C are non-zero multiples of one another, we have $a_{1j} \neq 0$ for all $j \in C$.

By Gaussian elimination, we may pass to an equivalent system $(\star')$, given by the matrix $A' = (a'_{ij}) \in \mathbb{F}_q^{m \times k}$, such that $a'_{i1} = 0$ for all $i > 1$. Since elementary row operations preserve the column equivalence classes, we have $a'_{ij} = 0$ for all $(i, j) \in \{2, \dots, m\} \times C$. It follows that every row in $(\star')$ uses variables from either C or $[k] \setminus C$, but not both. Since $\ell \geq m \geq 2$, we have $|C|, |[k] \setminus C| \neq 0$, so it follows that $(\star)$ is reducible. This is a contradiction, so we conclude that every (non-zero) equation implied by $(\star)$ uses at least two column equivalence classes.

To prove that $\ell \geq m + 1$, let A' be the matrix obtained by deleting from A the columns in one column equivalence class. By the above, every non-zero element of the row space of A' uses at least one of the remaining $\ell - 1$ column equivalence classes. It follows that $\text{rank}(A') = m$, so $\ell - 1 \geq m$. $\square$

Remark 2.4. In Theorems A – C we are in Situation 1.7; that is, our system is of type (RC) and is irreducible and non-degenerate. For context, we mention two facts about this situation that are not needed in the proofs, but may be helpful nonetheless.

- (i) If $m \geq 2$, then $k \geq 2m + 1$. This follows directly from Proposition 2.3 and the fact that, in a type (RC) linear system, every column equivalence class except at most one must have size at least two.
- (ii) If $\ell = m + 1$, then either every column class sums to zero, or none of the column classes sums to zero. Indeed, after row operations and permuting columns we may assume that $A = \begin{bmatrix} I & B \end{bmatrix}$, where every column of B is a scalar multiple of one of the vectors in $\{\mathbf{e}_1, \dots, \mathbf{e}_m, \mathbf{b}\}$, and $\mathbf{b}$ has no zero entries. So for every $i \in [m]$, the union of the column equivalence classes of $\mathbf{e}_i$ and $\mathbf{b}$ sums to zero.

3 Proof of Theorem A(i)

In this section, we develop the first main technique (the ‘single replacement trick’, see Corollary 3.6) and use it to prove Theorem A(i).

Definition 3.1. Let G be an abelian group. A sequence $\{(x_i, y_i, z_i)\}_{i=1}^L$ in G^3 is called a *tricoloured sum-free set in G* if for all $i, i', i'' \in [L]$ one has $x_i + y_{i'} + z_{i''} = 0$ if and only if $i = i' = i''$.

Note that the definition implies $|\{x_1, \dots, x_L\}| = |\{y_1, \dots, y_L\}| = |\{z_1, \dots, z_L\}| = L$; that is, in a tricoloured sum-free set there can be no repetitions in each of the coordinates (separately).

For all positive integers $t \geq 2$, define

$$J(t) := \frac{1}{t} \min_{0 < x < 1} \frac{1 + x + \dots + x^{t-1}}{x^{\frac{t-1}{3}}}.$$

It follows from [BCC⁺17, Prop. 4.12] that $J(t)$ is decreasing in t . Hence, for all $t \geq 2$ we have $J(t) \leq J(2) = 3 \cdot 2^{-5/3} < 0.945$. For a prime power q , define $\Gamma_q := qJ(q) < 0.945q$.⁹ By a routine application of the slice rank method, one can prove the following bound on the size of tricoloured sum-free sets.

Theorem 3.2 ([BCC⁺17]). *Let q be a prime power, and let $\{(\mathbf{x}_i, \mathbf{y}_i, \mathbf{z}_i)\}_{i=1}^L$ be a tricoloured sum-free set in $\mathbb{F}_q^n$. Then $L < (\Gamma_q)^n$.*

To recover Theorem 3.2 from known results, one has to proceed in three steps. First, the bound $L \leq 3 \cdot (\Gamma_q)^n$ follows from [BCC⁺17]. Second, to get rid of the additional factor 3, use the ‘power trick’ (a tricoloured sum-free set of size L in $\mathbb{F}_q^n$ gives one of size L^N in $\mathbb{F}_q^{nN}$ for all N). Finally, to get a strict inequality, prove that $(\Gamma_q)^n$ is never an integer

⁹Alternatively, for a prime power $q = p^s$, one could define $\Gamma_q := (pJ(p))^s < 0.945^s q$. This gives a slightly better bound, and Theorem 3.2 remains true as stated, because $\mathbb{F}_q^n \cong \mathbb{F}_p^{sn}$ as groups.

(see for instance [Dob23, Remark 5.11]). Alternatively, the results in this paper can be recovered by passing to a marginally higher constant $\Gamma_q + \varepsilon$ instead of Γ_q , because we have $L \leq (\Gamma_q)^n < (\Gamma_q + \varepsilon)^n$ for all $\varepsilon > 0$.

To prove the ‘single replacement trick’, we start with the following lemma.

Lemma 3.3. *Let q be a prime power, and let Γ_q be as in Theorem 3.2. Let $\alpha, \beta \in \mathbb{F}_q \setminus \{0\}$, let $\mathbf{x}_1, \dots, \mathbf{x}_L \in \mathbb{F}_q^n$ be distinct, and let $\mathbf{y}_1, \dots, \mathbf{y}_L \in \mathbb{F}_q^n$ be distinct. If $L \geq (\Gamma_q)^n$, then there exist $i, i', i'' \in [L]$ with $i \neq i', i''$ and $\alpha \mathbf{x}_i + \beta \mathbf{y}_i = \alpha \mathbf{x}_{i'} + \beta \mathbf{y}_{i''}$.*

Proof. For $i \in [L]$, define $\mathbf{z}_i := \alpha \mathbf{x}_i + \beta \mathbf{y}_i$. Each triple in the sequence $\{(\alpha \mathbf{x}_i, \beta \mathbf{y}_i, -\mathbf{z}_i)\}_{i=1}^L$ sums to zero, but we have $L \geq (\Gamma_q)^n$, so it follows from Theorem 3.2 that this sequence is not a tricoloured sum-free set. Therefore we may choose $i, i', i'' \in [L]$, not all equal, such that $\alpha \mathbf{x}_i + \beta \mathbf{y}_i = \mathbf{z}_i = \alpha \mathbf{x}_{i'} + \beta \mathbf{y}_{i''}$.

Suppose that $i'' = i$. Then we have $\alpha \mathbf{x}_i = \alpha \mathbf{x}_{i'}$, hence $\mathbf{x}_i = \mathbf{x}_{i'}$ (because $\alpha \neq 0$), and therefore $i = i'$ (because $\mathbf{x}_1, \dots, \mathbf{x}_L$ are distinct), contrary to our assumption that i, i' and i'' are not all equal. This is a contradiction, so we must have $i'' \neq i$. An analogous argument shows that $i' \neq i$. $\square$

Remark 3.4. In Lemma 3.3, we do not require that $i' \neq i''$. The case that $i' = i''$ corresponds to the case that $\mathbf{z}_1, \dots, \mathbf{z}_L$ are not all distinct. This does not matter for the rest of the proof.

Definition 3.5. We say that two solutions $\vec{\mathbf{x}} = (\mathbf{x}_1, \dots, \mathbf{x}_k)$ and $\vec{\mathbf{y}} = (\mathbf{y}_1, \dots, \mathbf{y}_k)$ to $(\star)$ are *disjoint* if $\{\mathbf{x}_1, \dots, \mathbf{x}_k\} \cap \{\mathbf{y}_1, \dots, \mathbf{y}_k\} = \emptyset$. Note that we do not require the $\mathbf{x}_j$ (resp. the $\mathbf{y}_j$) to be pairwise distinct.

Corollary 3.6 (‘Single replacement trick’). *Let $\{(\mathbf{x}_1^{(i)}, \dots, \mathbf{x}_k^{(i)})\}_{i=1}^L$ be a list of pairwise disjoint solutions of $(\star)$, and suppose that j_1 and j_2 are distinct indices in the same column equivalence class. If $L \geq (\Gamma_q)^n$, then there exist $i, i', i'' \in [L]$ with $i \neq i', i''$ such that the k -tuple $(\mathbf{y}_1, \dots, \mathbf{y}_k) \in (\mathbb{F}_q^n)^k$ given by*

$$\mathbf{y}_j = \begin{cases} \mathbf{x}_j^{(i)}, & \text{if } j \neq j_1, j_2; \\ \mathbf{x}_j^{(i')}, & \text{if } j = j_1; \\ \mathbf{x}_j^{(i'')}, & \text{if } j = j_2; \end{cases}$$

is also a solution of $(\star)$.

Proof. Since the j_1 -th and j_2 -th column of $(\star)$ are multiples of one another, we may choose a vector $\mathbf{v} \in \mathbb{F}_q^m$ and constants $\alpha, \beta \neq 0$ such that the j_1 -th column is equal to $\alpha \mathbf{v}$ and the j_2 -th column is equal to $\beta \mathbf{v}$.

By assumption, the vectors $\mathbf{x}_{j_1}^{(1)}, \dots, \mathbf{x}_{j_1}^{(L)}$ are distinct, and likewise the vectors $\mathbf{x}_{j_2}^{(1)}, \dots, \mathbf{x}_{j_2}^{(L)}$ are distinct, so it follows from Lemma 3.3 that there exist $i, i', i'' \in [L]$ with $i \neq i', i''$ and $\alpha \mathbf{x}_{j_1}^{(i)} + \beta \mathbf{x}_{j_2}^{(i)} = \alpha \mathbf{x}_{j_1}^{(i')} + \beta \mathbf{x}_{j_2}^{(i')}$. Hence, the total contribution of $\mathbf{x}_{j_1}^{(i)}$ and $\mathbf{x}_{j_2}^{(i)}$ to the equations of $(\star)$ is the same as the contribution of $\mathbf{x}_{j_1}^{(i')}$ and $\mathbf{x}_{j_2}^{(i')}$. Since $(\mathbf{x}_1^{(i)}, \dots, \mathbf{x}_k^{(i)})$ is a solution of $(\star)$, so is $(\mathbf{y}_1, \dots, \mathbf{y}_k)$. $\square$

We now prove the first main result of this paper, using the replacement trick from the preceding corollary.

Proof of Theorem A(i). Let $(\star)$, A , m , k and ℓ be as in Situation 1.7, and suppose that $(\star)$ satisfies property (i) from Theorem A (none of the column equivalence classes of size 2 sums to zero). Furthermore, let Γ_q be the constant from Theorem 3.2.

We prove by induction on λ that, for every $\lambda \in [k]$, there is a constant $\beta_\lambda \geq 1$ such that every subset $S \subseteq \mathbb{F}_q^n$ of size at least $\beta_\lambda \cdot (\Gamma_q)^n$ contains a solution $(\mathbf{x}_1, \dots, \mathbf{x}_k) \in S^k$ of $(\star)$ with at least λ different vectors; that is, $|\{\mathbf{x}_1, \dots, \mathbf{x}_k\}| \geq \lambda$. Setting $\lambda = k$ then proves the theorem.

For $\lambda = 1$, the claim is trivially true with $\beta_1 = 1$, since $(\mathbf{x}, \dots, \mathbf{x})$ is a solution of $(\star)$ for every $\mathbf{x} \in \mathbb{F}_q^n$.

For the induction step, suppose that $\lambda_0 \in [k-1]$ is given such that the statement is true for $\lambda = \lambda_0$. Define $\beta_{\lambda_0+1} := \beta_{\lambda_0} + P(k, \lambda_0) \cdot k$, where $P(k, \lambda_0)$ denotes the number of partitions of a k -element set into λ_0 parts.

Let $S \subseteq \mathbb{F}_q^n$ be a set of size at least $\beta_{\lambda_0+1} \cdot (\Gamma_q)^n = \beta_{\lambda_0} \cdot (\Gamma_q)^n + P(k, \lambda_0) \cdot (\Gamma_q)^n \cdot k$. Create a list of disjoint solutions $\{(\mathbf{x}_1^{(i)}, \dots, \mathbf{x}_k^{(i)})\}_{i=1}^{L_0}$ of $(\star)$ in S , each with at least λ_0 different vectors, by repeatedly finding such a solution in S and removing it from S . By the induction hypothesis, we can find a new solution as long as the remaining set has size at least $\beta_{\lambda_0} \cdot (\Gamma_q)^n$, and in each step we remove at most k vectors from S , so we find a list of length $L_0 \geq P(k, \lambda_0) \cdot (\Gamma_q)^n$.

If one of the solutions in the list has strictly more than λ_0 different vectors, then we are done. So we may assume that every solution in the list has exactly λ_0 different vectors.

We sort the entries in the list according to their partition pattern. We say that a solution $(\mathbf{x}_1^{(i)}, \dots, \mathbf{x}_k^{(i)})$ is *compatible* with a partition $[k] = J_1 \cup \dots \cup J_{\lambda_0}$ if for all $j_1, j_2 \in [k]$ we have: $\mathbf{x}_{j_1}^{(i)} = \mathbf{x}_{j_2}^{(i)}$ if and only if j_1 and j_2 belong to the same partition class. Evidently every solution is compatible with exactly one partition. By the pigeonhole principle, we may choose a partition $[k] = J_1 \cup \dots \cup J_{\lambda_0}$ that occurs at least $(\Gamma_q)^n$ times in our list of solutions. Thus, we obtain a list $\{(\mathbf{y}_1^{(i)}, \dots, \mathbf{y}_k^{(i)})\}_{i=1}^{L_1}$ of solutions of the same partition type, where $L_1 \geq (\Gamma_q)^n$.

Now we have two competing partitions of $[k]$, given by the column equivalence classes and the (now fixed) partition type $[k] = J_1 \cup \dots \cup J_{\lambda_0}$. For $j_1, j_2 \in [k]$, we write $j_1 \parallel j_2$ if j_1 and j_2 are in the same column equivalence class, and $j_1 \equiv j_2$ if j_1 and j_2 belong to the same class in the partition $[k] = J_1 \cup \dots \cup J_{\lambda_0}$ (i.e. if $\mathbf{y}_{j_1}^{(i)} = \mathbf{y}_{j_2}^{(i)}$ for all $i \in [L_1]$).

Since $\lambda_0 < k$, we may choose distinct $j_0, j_1 \in [k]$ with $j_0 \equiv j_1$. Furthermore, since $(\star)$ is a type (RC) linear system (see Definition 1.6(b)), it has at most one column equivalence class of size 1, so we may assume without loss of generality that j_1 belongs to a column equivalence class of size 2 or more. We distinguish two cases, depending on which of the column equivalence classes j_0 and j_1 belong to.

- **Case 1:** $j_0 \not\parallel j_1$ or j_0 and j_1 belong to the same column equivalence class of size at least 3. In this case, we may choose $j_2 \neq j_0, j_1$ such that $j_1 \parallel j_2$. By Corollary 3.6,

there is a solution $(z_1, \dots, z_k)$ of $(\star)$ of the form

$$z_j = \begin{cases} y_j^{(i)}, & \text{if } j \neq j_1, j_2; \\ y_j^{(i')}, & \text{if } j = j_1; \\ y_j^{(i'')}, & \text{if } j = j_2; \end{cases}$$

for some $i, i', i'' \in [L_1]$ with $i \neq i', i''$. In other words, $(z_1, \dots, z_k)$ is obtained by taking the solution $(y_1^{(i)}, \dots, y_k^{(i)})$ and replacing two entries.

We prove that $|\{z_1, \dots, z_k\}| \geq \lambda_0 + 1$. First, note that $\{z_{j_1}, z_{j_2}\} \cap \{z_j \mid j \neq j_1, j_2\} = \emptyset$, since the solutions in the list were disjoint. Now we distinguish two cases.

- If $j_1 \equiv j_2$, then the removal of the j_1 -th and j_2 -th vectors from $(y_1^{(i)}, \dots, y_k^{(i)})$ does not change the number of different vectors, since $y_{j_0}^{(i)} = y_{j_1}^{(i)} = y_{j_2}^{(i)}$. We replace them by two vectors z_{j_1}, z_{j_2} which are distinct from the other vectors in the solution (but possibly $z_{j_1} = z_{j_2}$), so the number of different vectors increases by at least 1.
- If $j_1 \not\equiv j_2$, then the removal of j_1 -th and j_2 -th vectors from $(y_1^{(i)}, \dots, y_k^{(i)})$ decreases the number of different vectors by at most 1, because $y_{j_0}^{(i)} = y_{j_1}^{(i)}$. In this case we are guaranteed to have $z_{j_1} \neq z_{j_2}$: different solutions in the list are disjoint, but even within the same solution the j_1 -th and j_2 -th entry are always different (because $j_1 \not\equiv j_2$). Thus, adding z_{j_1} and z_{j_2} to the solution increases the number of different vectors by 2. The net effect is an increase of at least 1.

This proves our claim that $|\{z_1, \dots, z_k\}| \geq \lambda_0 + 1$.

- **Case 2:** j_0 and j_1 belong to the same column equivalence class of size 2. Then, by assumption (i) from the theorem statement, the j_0 -th and j_1 -th columns of $(\star)$ do not sum to zero.

By Corollary 3.6, there is a solution $(z_1, \dots, z_k)$ of $(\star)$ of the form

$$z_j = \begin{cases} y_j^{(i)}, & \text{if } j \neq j_0, j_1; \\ y_j^{(i')}, & \text{if } j = j_0; \\ y_j^{(i'')}, & \text{if } j = j_1; \end{cases}$$

for some $i, i', i'' \in [L_1]$ with $i \neq i', i''$.

Suppose for the sake of contradiction that $z_{j_0} = z_{j_1}$; that is, $y_{j_0}^{(i')} = y_{j_1}^{(i'')}$. Since the j_0 -th and j_1 -th columns of $(\star)$ do not sum to zero, and since $y_{j_0}^{(i)} = y_{j_1}^{(i)}$, the fact that both $(y_1^{(i)}, \dots, y_k^{(i)})$ and $(z_1, \dots, z_k)$ are solutions of $(\star)$ implies that $y_{j_0}^{(i')} = y_{j_1}^{(i'')} = y_{j_0}^{(i)} = y_{j_1}^{(i)}$. This is a contradiction, because $i \neq i', i''$, and different solutions of the list are disjoint. Therefore we must have $z_{j_0} \neq z_{j_1}$.

The removal of $\mathbf{y}_{j_0}^{(i)}$ and $\mathbf{y}_{j_1}^{(i)}$ from the solution decreases the number of different vectors by at most 1, since $\mathbf{y}_{j_0}^{(i)} = \mathbf{y}_{j_1}^{(i)}$. On the other hand, putting back $\mathbf{z}_{j_0}$ and $\mathbf{z}_{j_1}$ increases the number of different vectors by 2, since we have $\mathbf{z}_{j_0} \neq \mathbf{z}_{j_1}$ and $\{\mathbf{z}_{j_1}, \mathbf{z}_{j_2}\} \cap \{\mathbf{z}_j \mid j \neq j_1, j_2\} = \emptyset$. The net effect is an increase of at least 1, so we have $|\{\mathbf{z}_1, \dots, \mathbf{z}_k\}| \geq \lambda_0 + 1$. $\square$

4 Proof of Theorem A(ii)

In this section, we develop our second main technique and combine it with the techniques from the previous section to prove Theorem A(ii).

Our second main technique is the following lemma, which shows that, for every subset $S \subseteq \mathbb{F}_q^n$ of size at least $q^{1+(1-\frac{1}{k})n}$, the difference set $S - S$ contains linearly generic solutions to every linear system in k variables (including systems with $k < 2m + 1$).¹⁰ The proof uses a simple counting argument and does not rely on the slice rank method at all.

Lemma 4.1. *Let $A = (a_{ij}) \in \mathbb{F}_q^{m \times k}$ be a non-zero matrix and let $S \subseteq \mathbb{F}_q^n$ have size at least $q^{1+(1-\frac{1}{k})n}$. Then there are $(\mathbf{x}_1, \dots, \mathbf{x}_k), (\mathbf{y}_1, \dots, \mathbf{y}_k) \in S^k$ such that, for all $\mathbf{b} = (b_1, \dots, b_k) \in \mathbb{F}_q^k$, one has $b_1\mathbf{x}_1 + \dots + b_k\mathbf{x}_k = b_1\mathbf{y}_1 + \dots + b_k\mathbf{y}_k$ if and only if $\mathbf{b}$ is in the row space of A .*

Proof. By removing redundant rows, we may assume without loss of generality that $\text{rank } A = m$. If $k = m$, then we can take $\mathbf{x} = \mathbf{y} \in S^k$ arbitrary. Hence, we may assume that $k \geq m + 1$. By performing elementary row operations and permuting columns, we may assume without loss of generality that A is of the form $[A' \ I_m]$ for some $A' \in \mathbb{F}_q^{m \times (k-m)}$.

The matrix A defines a function $f : (\mathbb{F}_q^n)^k \rightarrow (\mathbb{F}_q^n)^m$, where $[f(\mathbf{x}_1, \dots, \mathbf{x}_k)]_i = a_{i1}\mathbf{x}_1 + \dots + a_{ik}\mathbf{x}_k$. By the pigeonhole principle, we may choose some $\vec{\mathbf{z}} = (\mathbf{z}_1, \dots, \mathbf{z}_m) \in (\mathbb{F}_q^n)^m$ such that the set $T := f^{-1}(\vec{\mathbf{z}}) \cap S^k$ has size $|T| \geq |S|^k / q^{mn} \geq q^k q^{(k-m-1)n}$.

Let $\pi : (\mathbb{F}_q^n)^k \rightarrow (\mathbb{F}_q^n)^{k-m}$ be the projection onto the first $k - m$ coordinates, let $g : T \rightarrow (\mathbb{F}_q^n)^{k-m}$ be the restriction of π to T , and let $T' := g[T]$. Since A is of the form $[A' \ I_m]$, it is easy to see that for every $(\mathbf{x}_1, \dots, \mathbf{x}_{k-m}) \in (\mathbb{F}_q^n)^{k-m}$ there is exactly one possible choice of $(\mathbf{x}_{k-m+1}, \dots, \mathbf{x}_k) \in (\mathbb{F}_q^n)^m$ such that $f(\mathbf{x}_1, \dots, \mathbf{x}_k) = \vec{\mathbf{z}}$. Therefore g is injective, and it follows that $|T'| = |T|$.

Let $D = \{(\mathbf{w}_1, \dots, \mathbf{w}_{k-m}) \in (\mathbb{F}_q^n)^{k-m} \mid \mathbf{w}_1, \dots, \mathbf{w}_{k-m} \text{ are linearly dependent}\}$. Then $|D| < q^{k-m} q^{(k-m-1)n}$ since there are fewer than q^{k-m} possible linear relations.

Choose some $\vec{\mathbf{y}}' = (\mathbf{y}'_1, \dots, \mathbf{y}'_{k-m}) \in T'$. Since $|T' - \vec{\mathbf{y}}'| = |T'| > |D|$, we have $(T - \vec{\mathbf{y}}') \setminus D \neq \emptyset$, so we may choose $(\mathbf{x}'_1, \dots, \mathbf{x}'_{k-m}) \in T'$ such that $\mathbf{x}'_1 - \mathbf{y}'_1, \dots, \mathbf{x}'_{k-m} - \mathbf{y}'_{k-m}$ are linearly independent. Let $(\mathbf{x}_1, \dots, \mathbf{x}_k), (\mathbf{y}_1, \dots, \mathbf{y}_k) \in T \subseteq S^k$ be the (unique) preimages of $(\mathbf{x}'_1, \dots, \mathbf{x}'_{k-m})$ and $(\mathbf{y}'_1, \dots, \mathbf{y}'_{k-m})$ under g . Note that $(\mathbf{x}_1, \dots, \mathbf{x}_{k-m}) = (\mathbf{x}'_1, \dots, \mathbf{x}'_{k-m})$ and $(\mathbf{y}_1, \dots, \mathbf{y}_{k-m}) = (\mathbf{y}'_1, \dots, \mathbf{y}'_{k-m})$, since g is just a coordinate projection.

¹⁰We stress that the added generality of omitting the assumption that $k \geq 2m + 1$ will be needed in applications of Lemma 4.1 in the proofs of Theorem A(ii), Theorem B(ii) and Theorem C, because there we apply Lemma 4.1 to a system which has one column from each column equivalence class from the original system $(\star)$.

We claim that $(\mathbf{x}_1, \dots, \mathbf{x}_k)$ and $(\mathbf{y}_1, \dots, \mathbf{y}_k)$ satisfy the required property.

Since $f(\mathbf{x}_1, \dots, \mathbf{x}_k) = f(\mathbf{y}_1, \dots, \mathbf{y}_k) = \mathbf{z}$, it is clear that $b_1\mathbf{x}_1 + \dots + b_k\mathbf{x}_k = b_1\mathbf{y}_1 + \dots + b_k\mathbf{y}_k$ whenever $(b_1, \dots, b_k)$ is in the row space of A .

Now let $\mathbf{b} = (b_1, \dots, b_k) \in \mathbb{F}_q^k$ be an arbitrary row vector such that $b_1\mathbf{x}_1 + \dots + b_k\mathbf{x}_k = b_1\mathbf{y}_1 + \dots + b_k\mathbf{y}_k$. Since A is of the form $[A' \ I_m]$, we can add a linear combination of the rows of A to $\mathbf{b}$ to obtain a vector $\mathbf{c} = (c_1, \dots, c_k) \in \mathbb{F}_q^k$ with $c_{k-m+1} = \dots = c_k = 0$. By linearity, we have $c_1\mathbf{x}_1 + \dots + c_k\mathbf{x}_k = c_1\mathbf{y}_1 + \dots + c_k\mathbf{y}_k$, or equivalently,

$$c_1(\mathbf{x}_1 - \mathbf{y}_1) + \dots + c_{k-m}(\mathbf{x}_{k-m} - \mathbf{y}_{k-m}) = 0.$$

Since $\mathbf{x}_1 - \mathbf{y}_1, \dots, \mathbf{x}_{k-m} - \mathbf{y}_{k-m}$ are linearly independent, it follows that $c_1 = \dots = c_{k-m} = 0$, so we have $c_j = 0$ for all $j \in [k]$. This shows that $\mathbf{b}$ is in the row space of A . $\square$

We now come to the proof of Theorem A(ii). The proof is largely analogous to the proof of Theorem A(i) (see §3), the main difference being that we now use Lemma 4.1 to control column equivalence classes that sum to zero.

We prove the following slightly stronger theorem.

Theorem 4.2. *Let $(\star)$, A , m , k and ℓ be as in Situation 1.7. Suppose that there is a partition $[k] = P_1 \cup \dots \cup P_{2s}$ such that:*

- (i) *for all $r \in [s]$, the columns of A indexed by $P_r \cup P_{s+r}$ sum to zero;*
- (ii) *if $(b_1, \dots, b_k) \in \mathbb{F}_q^k \setminus \{0\}$ is a non-zero element in the row space of A , then one has $\sum_{j \in P_r} b_j \neq 0$ for at least two different values of $r \in [s]$.¹¹*
- (iii) *if C is a column equivalence class of size 2 that sums to zero, then there is some $r \in [s]$ such that $C = P_r \cup P_{s+r}$.*

Then $(\star)$ is moderate.

Before we prove Theorem 4.2, we first show how it implies Theorem A(ii).

Proof of Theorem A(ii), assuming Theorem 4.2. Let $C_1, \dots, C_\ell \subseteq [k]$ denote the column equivalence classes of A . We distinguish two cases:

- If $\ell = 1$, then we have $m = \text{rank}(A) \leq \ell = 1$, so we are in the situation with a single equation. Since we assumed $k \geq 3$, there is no column equivalence class of size 2, so it follows from Theorem A(i) that $(\star)$ is moderate.
- Suppose that $\ell \geq 2$. Since A is non-degenerate, every column of A is non-zero. Hence, since the column equivalence classes of A sum to zero, every column equivalence class has size at least 2. For every $r \in [\ell]$, choose $j_r \in C_r$ arbitrary, and set $P_r := \{j_r\}$ and $P_{\ell+r} := C_r \setminus \{j_r\}$.

¹¹Note that we only look at $r \in \{1, \dots, s\}$, and we ignore all $r \in \{s+1, \dots, 2s\}$. This is because it follows from (i) that $\sum_{j \in P_r} b_j \neq 0$ if and only if $\sum_{j \in P_{s+r}} b_j \neq 0$. An equivalent statement is that $\sum_{j \in P_r} b_j \neq 0$ for at least four different values of $r \in [2s]$.

We prove that the partition $[k] = P_1 \cup \dots \cup P_{2\ell}$ satisfies the properties from Theorem 4.2. Property (i) is met because each of the column equivalence classes sums to zero, and property (iii) is met by construction. To see that property (ii) is met, recall that $(\star)$ is irreducible, so it follows from Proposition 2.3 that every non-zero element of the row space of A uses at least two different column equivalence classes. $\square$

Proof of Theorem 4.2. Let Γ_q be the constant from Theorem 3.2. We prove by induction on λ that, for every $\lambda \in [k]$, there is a constant $\beta_\lambda \geq 1$ such that every subset $S \subseteq \mathbb{F}_q^n$ of size at least $\beta_\lambda \cdot (\max(\Gamma_q, q^{\frac{k-1}{k}}))^n$ contains a solution $(\mathbf{x}_1, \dots, \mathbf{x}_k) \in S^k$ of $(\star)$ satisfying the following properties:

- (a) the solution contains at least λ different vectors; that is, $|\{\mathbf{x}_1, \dots, \mathbf{x}_k\}| \geq \lambda$;
- (b) for every column equivalence class of size 2 that sums to zero, the variables $\mathbf{x}_{j_1}, \mathbf{x}_{j_2}$ corresponding to that class are distinct.

Before proving the base case, we first show that the induction step from the proof of Theorem A(i) carries through unchanged. This time, part (b) of the induction hypothesis replaces the assumption (i) from Theorem A. To see that property (b) is automatically maintained by the proof of Theorem A(i), recall that the induction step consists of choosing a column equivalence class C_t and replacing two variables from that class by other values, leaving the other classes unchanged. Since we started and ended with a solution of $(\star)$, the contribution of the variables $\{\mathbf{x}_j \mid j \in C_t\}$ to $(\star)$ must have remained the same. Property (b) is equivalent to saying that the contribution of $\{\mathbf{x}_j \mid j \in C\}$ to $(\star)$ is non-zero for every column equivalence class C of size 2 that sums to zero, so this property is automatically maintained by the proof of Theorem A(i).

It remains to prove the base case. Let $B = (b_{ir}) \in \mathbb{F}_q^{m \times s}$ be the matrix given by

$$b_{ir} := \sum_{j \in P_r} a_{ij} = - \sum_{j \in P_{s+r}} a_{ij}.$$

Suppose that $S \subseteq \mathbb{F}_q^n$ has size at least $q \cdot (\max(\Gamma_q, q^{\frac{k-1}{k}}))^n$. It follows from Lemma 4.1 that there are $(\mathbf{z}_1, \dots, \mathbf{z}_s), (\mathbf{z}_{s+1}, \dots, \mathbf{z}_{2s}) \in S^s$ such that, for all $(c_1, \dots, c_s) \in \mathbb{F}_q^s$, one has $c_1 \mathbf{z}_1 + \dots + c_s \mathbf{z}_s = c_1 \mathbf{z}_{s+1} + \dots + c_s \mathbf{z}_{2s}$ if and only if $(c_1, \dots, c_s)$ is in the row space of B . By assumption (ii), none of the standard unit vectors $\mathbf{e}_1, \dots, \mathbf{e}_s \in \mathbb{F}_q^s$ is in the row space of B , so it follows that $\mathbf{z}_r \neq \mathbf{z}_{s+r}$ for all $r \in [s]$ (since $\mathbf{z}_r = \mathbf{z}_{s+r}$ would imply that $\mathbf{e}_r$ is in the row space of B).

Since $[k] = P_1 \cup \dots \cup P_{2s}$ is a partition, we may define $\mathbf{y}_1, \dots, \mathbf{y}_k \in \{\mathbf{z}_1, \dots, \mathbf{z}_{2s}\} \subseteq S$ in such a way that $\mathbf{y}_j = \mathbf{z}_r$ if and only if $j \in P_r$. Then for all $i \in [m]$ we have

$$\begin{aligned} a_{i1} \mathbf{y}_1 + \dots + a_{ik} \mathbf{y}_k &= \sum_{j \in P_1} a_{ij} \mathbf{z}_1 + \dots + \sum_{j \in P_{2s}} a_{ij} \mathbf{z}_{2s} \\ &= b_{i1} \mathbf{z}_1 + \dots + b_{is} \mathbf{z}_s - b_{i1} \mathbf{z}_{s+1} - \dots - b_{is} \mathbf{z}_{2s} = 0, \end{aligned}$$

so $(\mathbf{y}_1, \dots, \mathbf{y}_k) \in S^k$ is a solution of $(\star)$. Clearly $|\{\mathbf{y}_1, \dots, \mathbf{y}_k\}| \geq 1$. Furthermore, by assumption (iii), for every column equivalence class $C = \{j_1, j_2\}$ of size 2 that sums to zero, there is some $r \in [s]$ such that $P_r = \{j_1\}$ and $P_{s+r} = \{j_2\}$, so it follows that $\mathbf{y}_{j_1} = \mathbf{z}_r \neq \mathbf{z}_{s+r} = \mathbf{y}_{j_2}$. $\square$

5 Preliminaries on temperate systems

We now shift our attention from moderate to temperate systems. We show that the problem of finding a generic solution is closely related to the problem of finding solutions of high dimension, and we show that we may once again restrict our attention to irreducible systems.

For an affine subspace $X \subseteq \mathbb{F}_q^n$ we let $\dim(X)$ denote the dimension of X . So $\dim(X)$ is the maximum number of affinely independent vectors in X minus one. For a set $S \subseteq \mathbb{F}_q^n$, we let $\text{aff}(S)$ denote the affine hull of S .

Definition 5.1. For any given k -tuple $(\mathbf{x}_1, \dots, \mathbf{x}_k) \in (\mathbb{F}_q^n)^k$, let

$$\text{Ann}_{\text{bal}}(\mathbf{x}_1, \dots, \mathbf{x}_k) = \{(b_1, \dots, b_k) \in \mathbb{F}_q^k \mid b_1 \mathbf{x}_1 + \dots + b_k \mathbf{x}_k = 0, \quad b_1 + \dots + b_k = 0\}.$$

So the elements of $\text{Ann}_{\text{bal}}(\mathbf{x}_1, \dots, \mathbf{x}_k)$ correspond to the balanced linear equations satisfied by $(\mathbf{x}_1, \dots, \mathbf{x}_k)$.

Lemma 5.2. For every $(\mathbf{x}_1, \dots, \mathbf{x}_k) \in (\mathbb{F}^n)^k$ we have

$$\dim(\text{aff}(\mathbf{x}_1, \dots, \mathbf{x}_k)) + \dim(\text{Ann}_{\text{bal}}(\mathbf{x}_1, \dots, \mathbf{x}_k)) = k - 1.$$

Proof. Let $A \in \mathbb{F}^{(n+1) \times k}$ be the matrix

$$A = \begin{pmatrix} 1 & \cdots & 1 \\ | & & | \\ \mathbf{x}_1 & \cdots & \mathbf{x}_k \\ | & & | \end{pmatrix}.$$

For $I \subseteq [k]$ the vectors $\mathbf{x}_i$, $i \in I$ are affinely independent if and only if the columns of A indexed by I are linearly independent. So $\text{rank}(A) = \dim(\text{aff}(\mathbf{x}_1, \dots, \mathbf{x}_k)) + 1$.

Evidently, $\ker(A)$ is precisely $\text{Ann}_{\text{bal}}(\mathbf{x}_1, \dots, \mathbf{x}_k)$, so the result follows from the rank-nullity theorem. $\square$

Corollary 5.3. Let $(\star)$ be a balanced linear system of rank m , with coefficient matrix $A \in \mathbb{F}_q^{m \times k}$, and let $(\mathbf{x}_1, \dots, \mathbf{x}_k)$ be a solution of $(\star)$. Then $\dim(\text{aff}(\mathbf{x}_1, \dots, \mathbf{x}_k)) \leq k - m - 1$, with equality if and only if $(\mathbf{x}_1, \dots, \mathbf{x}_k)$ is a generic solution of $(\star)$.

Proof. Since $(\mathbf{x}_1, \dots, \mathbf{x}_k)$ is a solution of the system $(\star)$, the row space of A is contained in $\text{Ann}_{\text{bal}}(\mathbf{x}_1, \dots, \mathbf{x}_k)$. Therefore we have $m = \text{rank}(A) \leq \dim(\text{Ann}_{\text{bal}}(\mathbf{x}_1, \dots, \mathbf{x}_k))$, so it follows from Lemma 5.2 that

$$\dim(\text{aff}(\mathbf{x}_1, \dots, \mathbf{x}_k)) = k - 1 - \dim(\text{Ann}_{\text{bal}}(\mathbf{x}_1, \dots, \mathbf{x}_k)) \leq k - 1 - m.$$

Clearly we have equality if and only if the row space of A is equal to $\text{Ann}_{\text{bal}}(\mathbf{x}_1, \dots, \mathbf{x}_k)$, which is equivalent to saying that all balanced linear equations satisfied by $(\mathbf{x}_1, \dots, \mathbf{x}_k)$ are linear combinations of the equations in $(\star)$. $\square$

Proposition 5.4. Suppose that $(\star)$ is equivalent to a linear system $(\star')$ whose coefficient matrix A' can be written as

$$A' = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$$

for some $A_1 \in \mathbb{F}_q^{m_1 \times k_1}$ and $A_2 \in \mathbb{F}_q^{m_2 \times k_2}$ with $m_1, m_2, k_1, k_2 \neq 0$. Then $(\star)$ is temperate if and only if the systems given by A_1 and A_2 are temperate.

Proof. If $(\star')$ is temperate, then it is easy to see that the same holds for the systems given by A_1 and A_2 .

Suppose that for $i = 1, 2$ the system given by A_i is temperate, with constants $\beta_i, \gamma_i > 0$, where $\gamma_i < q$. Let γ satisfy $\max(\gamma_1, \gamma_2) < \gamma < q$, and choose β such that

$$\beta q^{\gamma n} \geq \max(qn \cdot \beta_1 q^{\gamma_1 n}, nq^{k_1} \cdot \beta_2 q^{\gamma_2 n}) \quad \text{for all } n \in \mathbb{Z}_{\geq 1}.$$

Let $S \subseteq \mathbb{F}_q^n$ have size $|S| \geq \beta q^{\gamma n}$. For $i \in [n]$ and $\alpha \in \mathbb{F}_q$, write $S(i, \alpha) := \{\mathbf{x} \in S \mid x_i = \alpha\}$. We claim that there exist $i \in [n]$ and distinct $\alpha', \alpha'' \in \mathbb{F}_q$ such that $|S(i, \alpha')|, |S(i, \alpha'')| \geq \frac{|S|}{qn}$. For each coordinate $i \in [n]$, let $\alpha_i \in \arg \max_{\alpha \in \mathbb{F}_q} |S(i, \alpha)|$ be a most popular value. Then $S \setminus \{(\alpha_1, \dots, \alpha_n)\} = \cup_{i \in [n]} (S \setminus S(i, \alpha_i))$. So we can choose $i \in [n]$ such that $|S \setminus S(i, \alpha_i)| \geq \frac{|S|-1}{n}$. Then there is an $\alpha'' \neq \alpha_i$ such that $|S(i, \alpha'')| \geq \frac{|S|-1}{n(q-1)} \geq \frac{|S|}{qn}$. Taking $\alpha' = \alpha_i$ proves the claim.

Without loss of generality, we will assume that we can take $i = 1$ in the claim. We denote $S_1 = S(1, \alpha')$ and $S_2 = S(1, \alpha'')$. Since $|S_1| \geq \beta_1 q^{\gamma_1 n}$, there exists a generic solution $\vec{\mathbf{y}} = (\mathbf{y}_1, \dots, \mathbf{y}_{k_1}) \in (S_1)^{k_1}$ to the linear system given by A_1 . We can take $I \subseteq [n]$ with $|I| \leq k_1 - 1$ such that for all $\mathbf{b} = (b_1, \dots, b_{k_1}) \in \mathbb{F}_q^{k_1}$ with $b_1 + \dots + b_{k_1} = 0$ we have:

$$\forall i \in I : (b_1 \mathbf{y}_1 + \dots + b_{k_1} \mathbf{y}_{k_1})_i = 0 \implies b_1 \mathbf{y}_1 + \dots + b_{k_1} \mathbf{y}_{k_1} = 0.$$

Indeed, if $M \in \mathbb{F}_q^{n \times k_1}$ is the matrix with columns $\mathbf{y}_1, \dots, \mathbf{y}_{k_1}$, then we can take $I \subseteq [n]$ of size $|I| \leq k_1 - 1$ such that the rows of M are contained in the span of the rows indexed by I and the row vector $(1, \dots, 1)$. Since $y_{11} = \dots = y_{k_1 1}$ we may assume that $1 \notin I$.

As $\vec{\mathbf{y}}$ is a generic solution to the system given by A_1 , we obtain

$$\forall i \in I : (b_1 \mathbf{y}_1 + \dots + b_{k_1} \mathbf{y}_{k_1})_i = 0 \implies \mathbf{b} \in \text{rowspan}(A_1). \quad (5.5)$$

We can take $\alpha_i \in \mathbb{F}_q$ for each $i \in I$ such that $T = \{\mathbf{x} \in S_2 \mid x_i = \alpha_i \text{ for all } i \in I\}$ has size $|T| \geq |S_2| \cdot q^{1-k_1} \geq \beta_2 q^{\gamma_2 n}$.

It follows that there exists a generic solution $\vec{\mathbf{z}} \in T^{k_2}$ to the system given by A_2 . Now $\vec{\mathbf{x}} = (\vec{\mathbf{y}}, \vec{\mathbf{z}})$ is a generic solution to $(\star')$. Indeed, let $\mathbf{b} = (b_1, \dots, b_k) \in \text{Ann}_{\text{bal}}(\mathbf{x}_1, \dots, \mathbf{x}_k)$. It suffices to show that $\mathbf{b} \in \text{rowspan}(A')$. Looking at the first coordinate and using that $b_1 + \dots + b_k = 0$, we see that

$$0 = (b_1 + \dots + b_{k_1})\alpha' + (b_{k_1+1} + \dots + b_k)\alpha'' = (b_1 + \dots + b_{k_1})(\alpha' - \alpha'').$$

Since $\alpha' \neq \alpha''$, we find that $b_1 + \dots + b_{k_1} = 0 = b_{k_1+1} + \dots + b_k$. Since $\vec{\mathbf{z}} \in T^{k_2}$ it follows that

$$(b_1 \mathbf{y}_1 + \dots + b_{k_1} \mathbf{y}_{k_1})_i = (b_1 \mathbf{x}_1 + \dots + b_k \mathbf{x}_k)_i = 0 \quad (\forall i \in I).$$

It now follows from (5.5) that $(b_1, \dots, b_{k_1}) \in \text{rowspace}(A_1)$. So after modifying $\mathbf{b}$ by an element of $\text{rowspace}(A')$, we may assume that $b_1, \dots, b_{k_1} = 0$. Hence the fact that $\mathbf{b} \in \text{Ann}_{\text{bal}}(\mathbf{x}_1, \dots, \mathbf{x}_k)$ implies that $b_{k_1+1}\mathbf{z}_1 + \dots + b_k\mathbf{z}_{k_2} = 0$. Since $\mathbf{z}$ is generic, we conclude that $(b_{k_1+1}, \dots, b_k) \in \text{rowspace}(A_2)$. Hence, $\mathbf{b} \in \text{rowspace}(A')$. $\square$

6 Proof of Theorem B and Theorem C

In this section, we develop the multiple replacement trick (Corollary 6.2) and use it (in combination with Lemma 4.1) to prove Theorem B and Theorem C.

We start with a many-solutions version of Lemma 3.3.

Lemma 6.1. *Let q be a prime power, let $N_0 = (\Gamma_q)^n$, where Γ_q is as in Theorem 3.2, and let t be a positive integer. Let $\mathbf{x}_1, \dots, \mathbf{x}_L \in \mathbb{F}_q^n$ be distinct, let $\mathbf{y}_1, \dots, \mathbf{y}_L \in \mathbb{F}_q^n$ be distinct, and let $\alpha, \beta \in \mathbb{F}_q \setminus \{0\}$. If $L \geq 4tN_0$, then there exists an $i \in [L]$ such that*

$$\left| \left\{ (i', i'') \in ([L] \setminus \{i\})^2 \mid \alpha \mathbf{x}_{i'} + \beta \mathbf{y}_{i''} = \alpha \mathbf{x}_i + \beta \mathbf{y}_i \right\} \right| \geq t.$$

Proof. Write

$$T := \{(i, i', i'') \in [L]^3 \mid \alpha \mathbf{x}_{i'} + \beta \mathbf{y}_{i''} = \alpha \mathbf{x}_i + \beta \mathbf{y}_i \text{ and } i \neq i', i''\}.$$

By Lemma 3.3, the set $T \cap J^3$ is nonempty for all $J \subseteq [L]$ with $|J| \geq N_0$. We claim that $|T \cap J^3| \geq |J| - N_0$ for all $J \subseteq [L]$. Indeed, suppose that $|T \cap J^3| < |J| - N_0$; then we could delete fewer than $|J| - N_0$ elements from J to obtain a set J' of size $|J'| > N_0$ such that $T \cap (J')^3$ is empty: a contradiction. So $|T \cap J^3| - |J| + N_0 \geq 0$ for all $J \subseteq [L]$.

Let J be the random subset of $[L]$ obtained by independently taking each element of $[L]$ with probability $\frac{1}{2t}$. We have $\mathbb{E}[|J|] = \frac{L}{2t}$ and $\mathbb{E}[|T \cap J^3|] \leq \frac{|T|}{(2t)^2}$ since $|\{i, i', i''\}| \geq 2$ for all $(i, i', i'') \in T$. From $\mathbb{E}[|T \cap J^3| - |J| + N_0] \geq 0$ we obtain $\frac{|T|}{4t^2} \geq \frac{L}{2t} - N_0$, and therefore $\frac{|T|}{L} \geq 2t - \frac{4t^2 N_0}{L} \geq t$. Hence, by the pigeonhole principle, there is an $i \in [L]$ such that $|\{(i', i'') \in [L]^2 \mid (i, i', i'') \in T\}| \geq t$, as required. $\square$

Recall that two solutions $(\mathbf{x}_1, \dots, \mathbf{x}_k)$ and $(\mathbf{y}_1, \dots, \mathbf{y}_k)$ are said to be disjoint if $\{\mathbf{x}_1, \dots, \mathbf{x}_k\} \cap \{\mathbf{y}_1, \dots, \mathbf{y}_k\} = \emptyset$. We obtain a corollary analogous to Corollary 3.6.

Corollary 6.2 ('Multiple replacement trick'). *Let $\{(\mathbf{x}_1^{(i)}, \dots, \mathbf{x}_k^{(i)})\}_{i=1}^L$ be a list of pairwise disjoint solutions of $(\star)$, and suppose that j_1 and j_2 are distinct indices from the same column equivalence class. Suppose that $L \geq 4t \cdot (\Gamma_q)^n$ for some positive integer t . Then there exists $i \in [L]$ and t distinct pairs $(i'_s, i''_s) \in ([L] \setminus \{i\})^2$, $s \in [t]$, such that $(\mathbf{y}_1^{(s)}, \dots, \mathbf{y}_k^{(s)}) \in (\mathbb{F}_q^n)^k$ given by*

$$\mathbf{y}_j^{(s)} = \begin{cases} \mathbf{x}_j^{(i)}, & \text{if } j \neq j_1, j_2; \\ \mathbf{x}_j^{(i'_s)}, & \text{if } j = j_1; \\ \mathbf{x}_j^{(i''_s)}, & \text{if } j = j_2; \end{cases}$$

is also a solution of $(\star)$ for all $s \in [t]$.

Proof. Since the j_1 -th and j_2 -th column of $(\star)$ are nonzero multiples of one another, we may choose a vector $\mathbf{v} \in \mathbb{F}_q^m$ and constants $\alpha, \beta \neq 0$ such that the j_1 -th column is equal to $\alpha \mathbf{v}$ and the j_2 -th column is equal to $\beta \mathbf{v}$.

By assumption, the vectors $\mathbf{x}_{j_1}^{(1)}, \dots, \mathbf{x}_{j_1}^{(L)}$ are pairwise distinct, and likewise the vectors $\mathbf{x}_{j_2}^{(1)}, \dots, \mathbf{x}_{j_2}^{(L)}$ are pairwise distinct, so it follows from Lemma 6.1 that there exist $i \in [L]$ and t distinct pairs $(i'_s, i''_s) \in ([L] \setminus \{i\})^2$, $s \in [t]$, with $\alpha \mathbf{x}_{j_1}^{(i)} + \beta \mathbf{x}_{j_2}^{(i)} = \alpha \mathbf{x}_{j_1}^{(i'_s)} + \beta \mathbf{x}_{j_2}^{(i''_s)}$. Hence, the total contribution of $\mathbf{x}_{j_1}^{(i)}$ and $\mathbf{x}_{j_2}^{(i)}$ to the equations of $(\star)$ is the same as the contribution of $\mathbf{x}_{j_1}^{(i'_s)}$ and $\mathbf{x}_{j_2}^{(i''_s)}$. Since $(\mathbf{x}_1^{(i)}, \dots, \mathbf{x}_k^{(i)})$ is a solution of $(\star)$, so is $(\mathbf{y}_1^{(s)}, \dots, \mathbf{y}_k^{(s)})$. $\square$

Definition 6.3. Let $A \in \mathbb{F}_q^{m \times k}$ be a matrix and let $j_1, j_2 \in [k]$ be distinct elements in the same column equivalence class of A . We say that $(b_1, \dots, b_k) \in \mathbb{F}_q^k$ *breaks* the pair $\{j_1, j_2\}$ if after adding the row $(b_1, \dots, b_k)$ to A , the columns indexed by j_1 and j_2 are no longer scalar multiples of one another.

Lemma 6.4. Let $(\star)$, A , m , k and ℓ be as in Situation 1.7, let $j_1, j_2 \in [k]$ be distinct indices in the same column equivalence class, and let $\{(\mathbf{x}_1^{(i)}, \dots, \mathbf{x}_k^{(i)})\}_{i=1}^L$ be a list of pairwise disjoint solutions to $(\star)$. If $L \geq 4q^k(\Gamma_q)^n$, then there exists $i \in [L]$ and a solution $(\mathbf{y}_1, \dots, \mathbf{y}_k)$ to $(\star)$ such that:

- (i) $\mathbf{y}_j = \mathbf{x}_j^{(i)}$ for all $j \neq j_1, j_2$ and $\mathbf{y}_j \in \{\mathbf{x}_j^{(1)}, \dots, \mathbf{x}_j^{(L)}\}$ for $j \in \{j_1, j_2\}$;
- (ii) $\text{Ann}_{\text{bal}}(\mathbf{y}_1, \dots, \mathbf{y}_k) \subseteq \text{Ann}_{\text{bal}}(\mathbf{x}_1^{(i)}, \dots, \mathbf{x}_k^{(i)})$;
- (iii) no $\mathbf{b} \in \text{Ann}_{\text{bal}}(\mathbf{y}_1, \dots, \mathbf{y}_k)$ breaks the pair $\{j_1, j_2\}$.

Proof. By Corollary 6.2, we may choose $i \in [L]$ and a sequence $\{(i'_s, i''_s)\}_{s=1}^{q^k}$ of q^k pairwise distinct pairs $(i'_s, i''_s) \in ([L] \setminus \{i\})^2$ such that, for all $s \in [q^k]$, the k -tuple $(\mathbf{z}_1^{(s)}, \dots, \mathbf{z}_k^{(s)}) \in S^k$ defined by

$$\mathbf{z}_j^{(s)} = \begin{cases} \mathbf{x}_j^{(i)} & \text{if } j \in [k] \setminus \{j_1, j_2\} \\ \mathbf{x}_j^{(i'_s)} & \text{if } j = j_1 \\ \mathbf{x}_j^{(i''_s)} & \text{if } j = j_2 \end{cases}$$

is a solution to $(\star)$.

If $\mathbf{b} = (b_1, \dots, b_k)$ breaks the pair $\{j_1, j_2\}$, then the contributions $b_{j_1} \mathbf{z}_{j_1}^{(s)} + b_{j_2} \mathbf{z}_{j_2}^{(s)}$ for $s \in [q^k]$ are pairwise distinct. Therefore we can have $\mathbf{b} \in \text{Ann}_{\text{bal}}(\mathbf{z}_1^{(s)}, \dots, \mathbf{z}_k^{(s)})$ for at most one value of s . Since the number of $\mathbf{b} \in \mathbb{F}_q^k$ with $b_1 + \dots + b_k = 0$ is less than q^k , we may choose $s_0 \in [q^k]$ such that no $\mathbf{b} \in \text{Ann}_{\text{bal}}(\mathbf{z}_1^{(s_0)}, \dots, \mathbf{z}_k^{(s_0)})$ breaks the pair $\{j_1, j_2\}$.

Set $\mathbf{y} := \mathbf{z}^{(s_0)}$. Then (i) and (iii) are met. To prove (ii), let $\mathbf{b} \in \text{Ann}_{\text{bal}}(\mathbf{y}_1, \dots, \mathbf{y}_k)$ be given. Since $\mathbf{b}$ does not break the pair $\{j_1, j_2\}$, we have $b_{j_1} \mathbf{z}_{j_1}^{(s_0)} + b_{j_2} \mathbf{z}_{j_2}^{(s_0)} = b_{j_1} \mathbf{x}_{j_1}^{(i)} + b_{j_2} \mathbf{x}_{j_2}^{(i)}$, and therefore $\mathbf{b} \in \text{Ann}_{\text{bal}}(\mathbf{x}_1^{(i)}, \dots, \mathbf{x}_k^{(i)})$, as desired. $\square$

Lemma 6.5. Let $(\star)$, A , m , k and ℓ be as in Situation 1.7. Let $S \subseteq \mathbb{F}_q^n$ have size $|S| \geq q^{1+\frac{\ell-1}{\ell}n}$. Assume that at least one of the following two conditions holds:

- (i) $\ell = m + 1$;
- (ii) every column equivalence class sums to zero.

Then there exists a solution $\vec{x} = (\mathbf{x}_1, \dots, \mathbf{x}_k) \in S^k$ to $(\star)$ with the following property:

$$\text{If } \mathbf{b} \in \text{Ann}_{\text{bal}}(\mathbf{x}_1, \dots, \mathbf{x}_k) \text{ preserves the column} \\ \text{equivalence classes of } (\star), \text{ then } \mathbf{b} \in \text{rowspace}(A). \quad (6.6)$$

Proof. Let $[k] = C_1 \cup \dots \cup C_\ell$ be the partition of $[k]$ into column equivalence classes.

We first consider the case that condition (i) holds. Let $\vec{x} = (\mathbf{x}_1, \dots, \mathbf{x}_k)$ be any solution to $(\star)$. Suppose that $\vec{x}$ satisfies a balanced equation $b_1\mathbf{x}_1 + \dots + b_k\mathbf{x}_k = 0$ that preserves the column equivalence classes of $(\star)$, but $(b_1, \dots, b_k)$ is not in the row space of A . Let A' be the $(m+1) \times k$ matrix obtained by adding the row $(b_1, \dots, b_k)$ to A . Then $\text{rank}(A') = m+1 = \ell$. For $t \in [\ell]$ let $\sigma_t \in \mathbb{F}_q^{m+1}$ be the sum of the columns of A' in class C_t . Since the column rank of A' is ℓ , it follows that if we take one index from each column equivalence class, the corresponding ℓ columns are linearly independent. Let $I = \{t \in [\ell] \mid \sigma_t \neq 0\}$. Then the σ_t , $t \in I$ are linearly independent and $\sum_{t \in I} \sigma_t = \sum_{t \in [\ell]} \sigma_t = 0$. It follows that $I = \emptyset$. So all column equivalence classes of A' (and hence of A) sum to zero, and we are in case (ii).

We now consider the case that condition (ii) holds. Denote by $V \subseteq \mathbb{F}_q^k$ the set of vectors that preserve the column equivalence classes of A . We will assume (by reordering the columns of A) that $C_1 = \{1, \dots, |C_1|\}, \dots, C_\ell = \{k - |C_\ell| + 1, \dots, k\}$. So there are row vectors $\mathbf{v}_t \in \mathbb{F}_q^{|C_t|}$ such that

$$V = \{[c_1\mathbf{v}_1 \ \cdots \ c_\ell\mathbf{v}_\ell] \mid c_1, \dots, c_\ell \in \mathbb{F}_q\}.$$

Since the rows of A belong to V and A has no zero columns, the $\mathbf{v}_t$ have only nonzero entries. By scaling, we may assume that the first entry of $\mathbf{v}_t$ equals 1. For $t \in [\ell]$ let $j_t = |C_1| + \dots + |C_{t-1}| + 1$. So for all $\mathbf{b} \in V$, we have $\mathbf{b} = [b_{j_1}\mathbf{v}_1 \ \cdots \ b_{j_\ell}\mathbf{v}_\ell]$.

Let $A' = (a'_{it}) \in \mathbb{F}_q^{m \times \ell}$ be the submatrix of A induced by columns $j_1, \dots, j_\ell$. Then $(b_1, \dots, b_k)$ is in the row space of A if and only if $(b_{j_1}, \dots, b_{j_\ell})$ is in the row space of A' .

Consider the system

$$\sum_{t=1}^{\ell} a'_{it} \mathbf{y}_t = 0 \text{ for all } i \in [m].$$

Since $|S| \geq q^{1+\frac{\ell-1}{\ell}n}$, it follows by Lemma 4.1 that there are $(\mathbf{y}_1, \dots, \mathbf{y}_\ell)$ and $(\mathbf{z}_1, \dots, \mathbf{z}_\ell)$ in S^ℓ such that for all $(b_1, \dots, b_\ell) \in \mathbb{F}_q^\ell$ one has $b_1(\mathbf{y}_1 - \mathbf{z}_1) + \dots + b_\ell(\mathbf{y}_\ell - \mathbf{z}_\ell) = 0$ if and only if $(b_1, \dots, b_\ell)$ is in the row space of A' . Define $(\mathbf{x}_1, \dots, \mathbf{x}_k) \in S^k$ by setting (for $t \in [\ell]$ and $j \in C_t$)

$$\mathbf{x}_j = \begin{cases} \mathbf{y}_t & \text{if } j = j_t, \\ \mathbf{z}_t & \text{if } j \in C_t \setminus \{j_t\}. \end{cases}$$

Since the entries of each $\mathbf{v}_t$ sum to zero (the column equivalence classes sum to zero by assumption), we have

$$b_{j_1}(\mathbf{y}_1 - \mathbf{z}_1) + \cdots + b_{j_\ell}(\mathbf{y}_\ell - \mathbf{z}_\ell) = 0 \iff b_1\mathbf{x}_1 + \cdots + b_k\mathbf{x}_k = 0 \quad (6.7)$$

for every $\mathbf{b} \in V$.

We now check that $(\mathbf{x}_1, \dots, \mathbf{x}_k)$ satisfies the required properties. To show that it is a solution to $(\star)$, let $\mathbf{b}$ be a row of A . Then the restriction $(b_{j_1}, \dots, b_{j_\ell})$ is in the row space of A' , so $b_{j_1}(\mathbf{y}_1 - \mathbf{z}_1) + \cdots + b_{j_\ell}(\mathbf{y}_\ell - \mathbf{z}_\ell) = 0$. Hence, by (6.7), we have $b_1\mathbf{x}_1 + \cdots + b_k\mathbf{x}_k = 0$ as required.

Let $\mathbf{b} \in V \cap \text{Ann}_{\text{bal}}(\mathbf{x}_1, \dots, \mathbf{x}_k)$. It remains to show that $\mathbf{b}$ is in the row space of A . By (6.7), we have $b_{j_1}(\mathbf{y}_1 - \mathbf{z}_1) + \cdots + b_{j_\ell}(\mathbf{y}_\ell - \mathbf{z}_\ell) = 0$, so $(b_{j_1}, \dots, b_{j_\ell})$ is in the row space of A' . It follows that $\mathbf{b}$ is in the row space of A . $\square$

We are now ready to prove Theorem B and Theorem C.

Proof of Theorem B. Let Γ_q be the constant from Theorem 3.2. For every nonnegative integer t , we define

$$N_t := q^{1+\frac{\ell-1}{\ell}n} + t \cdot (4kq^k(\Gamma_q)^n).$$

Let $[k] = C_1 \cup \cdots \cup C_\ell$ be the partition of $[k]$ into column equivalence classes of A . We will prove by induction on $|P|$ that, for every set $P \subseteq \binom{C_1}{2} \cup \cdots \cup \binom{C_\ell}{2}$ of equivalent pairs and for every set $S \subseteq \mathbb{F}_q^n$ of size $|S| \geq N_{|P|}$, the system $(\star)$ has a solution $\vec{\mathbf{x}} = (\mathbf{x}_1, \dots, \mathbf{x}_k) \in S^k$ that satisfies (6.6) and such that no $(b_1, \dots, b_k) \in \text{Ann}_{\text{bal}}(\mathbf{x}_1, \dots, \mathbf{x}_k)$ breaks a pair in P .

- For $|P| = 0$, the claim follows directly from Lemma 6.5.
- Assume that $|P| \geq 1$ and that the claim holds for all sets of fewer than $|P|$ pairs. Fix some $\{j_1, j_2\} \in P$, write $L = 4q^k(\Gamma_q)^n$, and let $S \subseteq \mathbb{F}_q^n$ be a set of size $|S| \geq N_{|P|}$. Since $|S| \geq N_{|P|} \geq kL + N_{|P|-1}$, it follows from the induction hypothesis that there exist L pairwise disjoint solutions $\vec{\mathbf{x}}^{(1)}, \dots, \vec{\mathbf{x}}^{(L)} \in S^k$ to $(\star)$ that satisfy (6.6) and such that no $(b_1, \dots, b_k) \in \text{Ann}_{\text{bal}}(\mathbf{x}_1^{(i)}, \dots, \mathbf{x}_k^{(i)})$ breaks a pair in $P \setminus \{\{j_1, j_2\}\}$, for all $i \in [L]$.

By Lemma 6.4, we may choose $i_0 \in [L]$ and a solution $\vec{\mathbf{x}} = (\mathbf{x}_1, \dots, \mathbf{x}_k) \in S^k$ to $(\star)$ such that $\text{Ann}_{\text{bal}}(\mathbf{x}_1, \dots, \mathbf{x}_k) \subseteq \text{Ann}_{\text{bal}}(\mathbf{x}_1^{(i_0)}, \dots, \mathbf{x}_k^{(i_0)})$ and no $\mathbf{b} \in \text{Ann}_{\text{bal}}(\mathbf{x}_1, \dots, \mathbf{x}_k)$ breaks the pair $\{j_1, j_2\}$. By construction, $\vec{\mathbf{x}}^{(i_0)}$ satisfies (6.6) and no $\mathbf{b} \in \text{Ann}_{\text{bal}}(\mathbf{x}_1^{(i_0)}, \dots, \mathbf{x}_k^{(i_0)})$ breaks a pair in $P \setminus \{\{j_1, j_2\}\}$, so the same properties are true for $\vec{\mathbf{x}}$, because $\text{Ann}_{\text{bal}}(\mathbf{x}_1, \dots, \mathbf{x}_k) \subseteq \text{Ann}_{\text{bal}}(\mathbf{x}_1^{(i_0)}, \dots, \mathbf{x}_k^{(i_0)})$. We conclude that $\vec{\mathbf{x}}$ satisfies (6.6) and no $\mathbf{b} \in \text{Ann}_{\text{bal}}(\mathbf{x}_1, \dots, \mathbf{x}_k)$ breaks a pair in P .

Letting $P = \binom{C_1}{2} \cup \cdots \cup \binom{C_\ell}{2}$ completes the proof. $\square$

Proof of Theorem C. If all column equivalence classes sum to zero, the result follows directly from Theorem B(ii). Assume therefore that not all column equivalence classes

sum to zero. Let Γ_q be the constant from Theorem 3.2. For every nonnegative integer t we define

$$N_t := t \cdot (4kq^k(\Gamma_q)^n).$$

Let $S \subseteq \mathbb{F}_q^n$ have size $|S| \geq N_{k_2}$. By the same argument as in the proof of Theorem B, we have a solution $\vec{x} = (\mathbf{x}_1, \dots, \mathbf{x}_k) \in S^k$ to $(\star)$ such that no $\mathbf{b} \in \text{Ann}_{\text{bal}}(\mathbf{x}_1, \dots, \mathbf{x}_k)$ breaks a pair from the same column equivalence class. In other words, $\mathbf{b}$ preserves the column equivalence classes, so this proves part (i).

For part (ii), observe that $\text{Ann}_{\text{bal}}(\mathbf{x}_1, \dots, \mathbf{x}_k)$ does not contain all balanced linear equations that preserve the column equivalence classes, for otherwise every column equivalence class must sum to zero, contrary to our assumption. So we have $\dim(\text{Ann}_{\text{bal}}(\mathbf{x}_1, \dots, \mathbf{x}_k)) \leq \ell - 1$, and therefore $\dim(\text{aff}(\mathbf{x}_1, \dots, \mathbf{x}_k)) \geq k - \ell$, by Lemma 5.2. $\square$

Remark 6.8. We compare the rank of the solution $(\mathbf{x}_1, \dots, \mathbf{x}_k)$ in Theorem C to the rank given by Theorem 1.5. Suppose we are in Situation 1.7, and set $r = k - 2m + 1$. Then $k \geq 2m - 1 + r$, so it follows from Theorem 1.5 that we can find a solution with $\dim(\text{span}(\mathbf{x}_1, \dots, \mathbf{x}_k)) \geq r$, and therefore $\dim(\text{aff}(\mathbf{x}_1, \dots, \mathbf{x}_k)) \geq r - 1 = k - 2m$.

So how do these two compare? If $\ell = 1$, then we must have $m = 1$ (because we assume that the rows of A are linearly independent), so in this case the rank from Theorem C and Theorem 1.5 agree. If $\ell \geq 2$, then we see that Theorem C improves upon Theorem 1.5 whenever $m > \frac{\ell}{2}$. Then again, Theorem C only applies to a smaller class of linear systems.

7 Examples and applications

We conclude this paper by looking at a few examples of type (RC) linear systems, to highlight the applications and limitations of the results from this paper. First we will look at an application to sumsets in $\mathbb{F}_q^n$. We show that our results can be used to find non-trivial solutions of an arbitrary linear system in the difference set $S - S$, but not in the sumset $S + S$. After that, we will look at the systems studied by Mimura and Tokushige [MT19a, MT19b, MT20]. We show that our techniques furnish alternative proofs that those systems are moderate, and in many cases we strengthen this to show that the system is also temperate.

7.1 Applications to sum and difference sets

Since this paper studies linear systems with repeated columns, one obvious question is to which extent our results can be applied to the problem of finding solutions to a system of linear equations in sum and difference sets. Throughout this section, let $\mathbb{F}_q$ be a finite field of characteristic p , and let $c_1, \dots, c_l \in \mathbb{F}_q \setminus \{0\}$. We consider the affinely independent sumset (or AIR-sumset)

$$T := c_1 \cdot S \overset{\text{aff}}{+} \dots \overset{\text{aff}}{+} c_l \cdot S = \{c_1 \mathbf{x}_1 + \dots + c_l \mathbf{x}_l \mid \mathbf{x}_1, \dots, \mathbf{x}_l \in S \text{ affinely independent}\}.$$

If $c_1 + \dots + c_l = 0$, then Corollary D states that T contains generic solutions to every linear system $(\star)$, provided that S is sufficiently large. We now prove this statement.

Proof of Corollary D. Let $A = (a_{ij}) \in \mathbb{F}_q^{m \times k}$ be the coefficient matrix of the system $(\star)$. (Recall from the statement of Corollary D that A may be arbitrary.) Let $A' = (a'_{ij}) \in \mathbb{F}_q^{m \times lk}$ be the $m \times lk$ matrix

$$A' = [c_1 A \mid c_2 A \mid \cdots \mid c_l A],$$

and let $(\star')$ be the corresponding linear system. Every column equivalence class of $(\star')$ is the union of sets of the form $\{j, j+k, \dots, j+(l-1)k\}$ (for some $j \in [k]$), so $(\star')$ is of type (RC). Furthermore, the column equivalence classes sum to zero, because $c_1 + \cdots + c_l = 0$. Hence it follows from Theorem B(ii) and Proposition 5.4 that $(\star')$ is temperate. Therefore there are constants $\beta, \gamma \geq 1$ with $\gamma < q$ such that every set $S \subseteq \mathbb{F}_q^n$ with $|S| \geq \beta \cdot \gamma^n$ contains a generic solution of $(\star')$. Choose such a generic solution $(\mathbf{x}_1, \dots, \mathbf{x}_{lk}) \in S^{lk}$, and define $\mathbf{y}_1, \dots, \mathbf{y}_k \in c_1 \cdot S + \cdots + c_l \cdot S$ by

$$\mathbf{y}_j := c_1 \mathbf{x}_j + c_2 \mathbf{x}_{j+k} + \cdots + c_l \mathbf{x}_{j+(l-1)k}.$$

Clearly $(\mathbf{y}_1, \dots, \mathbf{y}_k)$ is a solution of the linear system $(\star)$. We show that $(\mathbf{y}_1, \dots, \mathbf{y}_k)$ is linearly generic and that $\mathbf{y}_1, \dots, \mathbf{y}_k \in (c_1 \cdot S \overset{\text{aff}}{+} \cdots \overset{\text{aff}}{+} c_l \cdot S) \cup \{0\}$.

First, let $\mathbf{b} = (b_1, \dots, b_k) \in \mathbb{F}_q^k$ be such that $b_1 \mathbf{y}_1 + \cdots + b_k \mathbf{y}_k = 0$. Then $(\mathbf{x}_1, \dots, \mathbf{x}_{lk})$ belongs to the kernel of the $1 \times lk$ matrix

$$B' = [c_1 \mathbf{b} \mid c_2 \mathbf{b} \mid \cdots \mid c_l \mathbf{b}].$$

Since $c_1 + \cdots + c_l = 0$, the entries of B' sum to 0, so B' represents a balanced linear equation satisfied by $(\mathbf{x}_1, \dots, \mathbf{x}_{lk})$. Since $(\mathbf{x}_1, \dots, \mathbf{x}_{lk})$ is a generic solution of $(\star')$, it follows that B' is in the row space of A' . Equivalently, $\mathbf{b}$ is in the row space of A . This shows that $(\mathbf{y}_1, \dots, \mathbf{y}_k)$ is linearly generic.

To complete the proof, it suffices to show that $\mathbf{y}_j = 0$ whenever the vectors $\mathbf{x}_j, \mathbf{x}_{j+k}, \dots, \mathbf{x}_{j+(l-1)k}$ are affinely dependent, for every $j \in [k]$. To that end, suppose that $\mathbf{x}_j, \mathbf{x}_{j+k}, \dots, \mathbf{x}_{j+(l-1)k}$ are affinely dependent. Then there is some $\mathbf{b} = (b_1, \dots, b_l) \in \mathbb{F}_q^l \setminus \{0\}$ with $b_1 + \cdots + b_l = 0$ and

$$b_1 \mathbf{x}_j + b_2 \mathbf{x}_{j+k} + \cdots + b_l \mathbf{x}_{j+(l-1)k} = 0. \quad (\mathbf{b}')$$

Since $(\mathbf{x}_1, \dots, \mathbf{x}_{lk})$ is generic, the balanced linear equation $(\mathbf{b}')$ is a linear combination of the equations in $(\star')$. By choosing some $r \in [l]$ such that $b_r \neq 0$ and restricting our attention to the variables $\mathbf{x}_{(r-1)k+1}, \dots, \mathbf{x}_{rk}$ (i.e. the r -th block in the block matrix representation of A'), we see that the equation $\mathbf{y}_j = 0$ is a linear combination of the equations in $(\star)$. $\square$

Corollary E can be deduced from Corollary D by letting $(\star)$ be the linear system that encodes a k -term arithmetic progression and setting $l = 2$ and $(c_1, c_2) = (1, -1)$. We show that Corollary E does not depend on the full strength of Corollary D, as it follows immediately from Lemma 4.1.

Proof of Corollary E. Let $(\star)$ be a linear system which encodes a k -term arithmetic progression, for instance the system given by the matrix

$$A = \begin{pmatrix} 1 & -2 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & -2 & 1 \end{pmatrix} \in \mathbb{F}_p^{(k-2) \times k}.$$

Let $S \subseteq \mathbb{F}_p^n$ with $|S| \geq p^{1+(1-\frac{1}{k})n}$. By Lemma 4.1, there are $(\mathbf{x}_1, \dots, \mathbf{x}_k), (\mathbf{y}_1, \dots, \mathbf{y}_k) \in S^k$ such that $(\mathbf{x}_1 - \mathbf{y}_1, \dots, \mathbf{x}_k - \mathbf{y}_k)$ is a linearly generic solution of $(\star)$.

Since the standard basis vectors $\mathbf{e}_1, \dots, \mathbf{e}_k \in \mathbb{F}_q^k$ are not in the row space of A ,¹² we have $\mathbf{x}_j - \mathbf{y}_j \neq 0$ for all $j \in [k]$. Likewise, since the vectors $\mathbf{e}_j - \mathbf{e}_{j'}$ ($j \neq j'$) are not in the row space of A ,¹² we have $\mathbf{x}_j - \mathbf{y}_j \neq \mathbf{x}_{j'} - \mathbf{y}_{j'}$ whenever $j \neq j'$. It follows that $(\mathbf{x}_1 - \mathbf{y}_1, \dots, \mathbf{x}_k - \mathbf{y}_k)$ is a non-trivial k -AP in $(S - S) \setminus \{0\}$. $\square$

Remark 7.1. The preceding proof carries through unchanged if A is replaced by an arbitrary matrix, and if the difference set $(S - S) \setminus \{0\}$ is replaced by the sum set $c_1 \cdot S + \dots + c_l \cdot S$ with $c_1 + \dots + c_l = 0$ (replace $\mathbf{x}_j - \mathbf{y}_j \in S - S$ by $c_1 \mathbf{x}_j + (c_2 + \dots + c_l) \mathbf{y}_j \in c_1 \cdot S + \dots + c_l \cdot S$). So a weaker version of Corollary D, where the AIR-sumset is replaced by an ordinary sumset, can also be proved by a simple counting argument, without using the slice rank method.

Remark 7.2. Now consider once again the sumset $c_1 \cdot S + \dots + c_l \cdot S$, but this time assume that $c_1 + \dots + c_l \neq 0$. In this case, the techniques from this paper do not say anything non-trivial about the problem of finding a non-trivial k -AP in the sum set $c_1 \cdot S + \dots + c_l \cdot S$.

We explain why the results from this paper do not work when $c_1 + \dots + c_l \neq 0$. It is tempting to try to repeat the proof of Corollary D, but we run into a problem: The column equivalence classes no longer sum to zero, so we have to replace Theorem B(ii) by Theorem B(i). However, this imposes two extra conditions on the original $m \times k$ matrix in the proof of Corollary D, namely that $A\mathbf{1} = 0$ (i.e. $(\star)$ is balanced) and that $k = \text{rank}(A) + 1$. So we can only say something for a very specific class of linear systems. In fact, this class is so specific that the coefficient matrix must satisfy $\ker(A) = \text{span}(\mathbf{1})$, so every solution of the original system must be constant!

Likewise, it is tempting to try to repeat the proof of Corollary D, but this time replacing Theorem B(ii) by Theorem A(i). After all, to find (say) a non-trivial k -AP, it is enough to find a solution with $\mathbf{y}_1, \dots, \mathbf{y}_k$ pairwise distinct instead of a generic solution. Here we run into another problem. In the proof of Corollary D, we can find a solution $(\mathbf{x}_1, \dots, \mathbf{x}_{lk}) \in S^{lk}$ of the extended system $(\star')$ with $\mathbf{x}_1, \dots, \mathbf{x}_{lk}$ pairwise distinct. But when we recombine these to form a solution $(\mathbf{y}_1, \dots, \mathbf{y}_k) \in (c_1 \cdot S + \dots + c_l \cdot S)^k$ of the original system $(\star)$, we may end up with $\mathbf{y}_1 = \dots = \mathbf{y}_k$, since we have no way to avoid these additional equations. In fact, if we use the proof of Theorem A(i) as an algorithm to find the $\mathbf{x}_1, \dots, \mathbf{x}_{lk}$, then this is guaranteed to happen: We start with a solution where

¹²To prove this, it is sufficient to note that there exist non-trivial k -APs in $\mathbb{F}_q^n \setminus \{0\}$.

all variables $\mathbf{x}_1, \dots, \mathbf{x}_{lk}$ are equal, and then modify the variables in such a way that the contribution to each column equivalence class remains the same, so the equation $\mathbf{y}_1 = \dots = \mathbf{y}_k$ is maintained throughout the proof. Once again, the techniques from this paper are unable to say anything non-trivial.

7.2 The systems studied by Mimura and Tokushige

In a series of papers [MT19a, MT19b, MT20], Mimura and Tokushige studied several specific (classes of) linear systems, and showed that each of them is moderate. These were the first results of this type. We show that our results and techniques furnish alternative proofs for all systems studied by Mimura and Tokushige (though our constants might not be as good).

The systems studied by Mimura and Tokushige have integer entries, and can therefore be interpreted as a linear system over $\mathbb{F}_q$ for an arbitrary prime power $q = p^s$. Depending on the system, Mimura and Tokushige sometimes had to assume that $p \neq 2$ or $p \neq 3$, and we shall do the same.

Example 7.3. In [MT19a], Mimura and Tokushige studied a *star of k three-term arithmetic progressions*, given by the linear system $(\mathcal{S}_{*k})$ with coefficient matrix

$$\begin{pmatrix} 1 & 1 & 0 & 0 & \cdots & 0 & 0 & -2 \\ 0 & 0 & 1 & 1 & \cdots & 0 & 0 & -2 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 1 & -2 \end{pmatrix} \in \mathbb{F}_q^{k \times (2k+1)},$$

and proved that this system is moderate whenever $p \geq 3$.

This result can be recovered as a special case of Theorem A, and strengthened to $(\mathcal{S}_{*k})$ being temperate by Theorem B. Indeed, $(\mathcal{S}_{*k})$ is a type (RC) linear system, as it is balanced and there is only one column equivalence class of size 1. If $p \neq 2$, then the system is non-degenerate and irreducible, and all column equivalence classes have sum $\pm 2 \neq 0$, so it follows from Theorem A(i) that $(\mathcal{S}_{*k})$ is moderate. Additionally, since there are k equations and $k + 1$ column equivalence classes, it follows from Theorem B(i) that $(\mathcal{S}_{*k})$ is temperate.

Example 7.4. Also in [MT19a], Mimura and Tokushige point out that their proof also extends to a ‘fan’ of k three-term arithmetic progressions, given by the linear system $(\mathcal{S}'_{*k})$ with coefficient matrix

$$\begin{pmatrix} 1 & -2 & 0 & 0 & \cdots & 0 & 0 & 1 \\ 0 & 0 & 1 & -2 & \cdots & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & -2 & 1 \end{pmatrix} \in \mathbb{F}_q^{k \times (2k+1)}.$$

Analogously to Example 7.3, it follows from Theorem A(i) and Theorem B(i) that $(\mathcal{S}'_{*k})$ is moderate and temperate, provided that $p \neq 2$.

Example 7.5. In [MT19b], Mimura and Tokushige studied the problem of avoiding a ‘W shape’, and showed that the linear system $(\mathcal{W})$ with coefficient matrix

$$\begin{pmatrix} 1 & -1 & -1 & 1 & 0 \\ 1 & 0 & -2 & 0 & 1 \end{pmatrix} \in \mathbb{F}_q^{2 \times 5}$$

is moderate whenever $p \geq 3$.

This is not a type (RC) linear system, since there are 3 column equivalence classes of size 1, so this result cannot be recovered as a special case of Theorem A or Theorem B.

Nevertheless, our techniques from §3 can be adapted to recover this result as well. Indeed, let Γ_q be the constant from Theorem 3.2, and let $S \subseteq \mathbb{F}_q^n$ with $|S| \geq 4 \cdot (\Gamma_q)^n$. By repeatedly finding a non-trivial 3-AP and removing it from S , we can find a list $\{(\mathbf{x}_1^{(i)}, \mathbf{x}_3^{(i)}, \mathbf{x}_5^{(i)})\}_{i=1}^L$ of $L \geq (\Gamma_q)^n$ pairwise disjoint non-trivial 3-APs in S^3 . For all $i \in [L]$, set $\mathbf{x}_2^{(i)} = \mathbf{x}_3^{(i)}$ and $\mathbf{x}_4^{(i)} = \mathbf{x}_5^{(i)}$, so that $(\mathbf{x}_1^{(i)}, \mathbf{x}_2^{(i)}, \mathbf{x}_3^{(i)}, \mathbf{x}_4^{(i)}, \mathbf{x}_5^{(i)}) \in S^5$ is a solution of $(\mathcal{W})$. Since 2 and 4 belong to the same column equivalence class, it follows from Corollary 3.6 that there are $i \neq i', i''$ such that the 5-tuple $(\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \mathbf{y}_4, \mathbf{y}_5) = (\mathbf{x}_1^{(i)}, \mathbf{x}_2^{(i')}, \mathbf{x}_3^{(i)}, \mathbf{x}_4^{(i'')}, \mathbf{x}_5^{(i)}) \in S^5$ is also a solution of $(\mathcal{W})$. Then $\mathbf{y}_1, \mathbf{y}_3, \mathbf{y}_5$ are pairwise distinct because they stem from the same non-trivial 3-AP, and $\{\mathbf{y}_1, \mathbf{y}_3, \mathbf{y}_5\} \cap \{\mathbf{y}_2, \mathbf{y}_4\} = \emptyset$ because they stem from disjoint solutions. Finally, note that $\mathbf{y}_2 \neq \mathbf{y}_4$, for otherwise the first equation of $(\mathcal{W})$ would imply that $\mathbf{y}_1 = \mathbf{y}_3$. This shows that $(\mathcal{W})$ is moderate.

With minor modifications, the preceding argument also shows that $(\mathcal{W})$ is temperate. Indeed, by repeating the argument, but using multiple replacement (Corollary 6.2) instead of single replacement (Corollary 3.6), we can make sure that $\mathbf{x}_2^{(i')}$ is not in the line through $\mathbf{x}_1^{(i)}$, $\mathbf{x}_3^{(i)}$ and $\mathbf{x}_5^{(i)}$. Then $\dim(\text{aff}(\mathbf{x}_1^{(i)}, \mathbf{x}_2^{(i')}, \mathbf{x}_3^{(i)}, \mathbf{x}_4^{(i'')}, \mathbf{x}_5^{(i)})) \geq 2$, so it follows from Corollary 5.3 that this solution is generic.

Example 7.6. In [MT20], Mimura and Tokushige studied the system (T) with coefficient matrix

$$\begin{pmatrix} 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & -2 & 1 & 1 \end{pmatrix} \in \mathbb{F}_q^{2 \times 5},$$

and proved that it is moderate whenever $p \geq 3$.

Once again, this result can be recovered as a special case of Theorem A(i), and strengthened to (T) being temperate by Theorem B(i).

Example 7.7. In [MT20], Mimura and Tokushige studied the class of linear systems (lS_{k+2}) . This class is defined as follows: let $k \geq 1$, and let $a_1, \dots, a_{k+2} \in \mathbb{F}_q$ be non-zero such that $a_1 + \dots + a_{k+2} = 0$. Then (lS_{k+2}) is given by the coefficient matrix

$$\begin{pmatrix} a_1 & \cdots & a_k & a_{k+1} & a_{k+2} & 0 & 0 & \cdots & 0 & 0 \\ a_1 & \cdots & a_k & 0 & 0 & a_{k+1} & a_{k+2} & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_1 & \cdots & a_k & 0 & 0 & 0 & 0 & \cdots & a_{k+1} & a_{k+2} \end{pmatrix} \in \mathbb{F}_q^{l \times (k+2l)}.$$

In [MT20, Thm. 5], Mimura and Tokushige showed that such a system is always moderate. (This contains the linear system (S_1) from [MT20] as a special case.)

This result can be recovered as a special case of Theorem A, and strengthened to (lS_{k+2}) being temperate by Theorem B. Indeed, (lS_{k+2}) is balanced, and it has one column equivalence class of size $k \geq 1$ and l column equivalence classes of size 2, so it is a type (RC) linear system. Furthermore, the system is non-degenerate and irreducible. Note that, if one column equivalence class sums to zero, then all column equivalence classes must sum to zero, so it follows from either Theorem A(i) or Theorem A(ii) that (lS_{k+2}) is moderate. Furthermore, since the number of equations is l and the number of column equivalence classes is $l + 1$, it follows from either Theorem B(i) or Theorem B(ii) that (lS_{k+2}) is temperate.

Example 7.8. In [MT20], Mimura and Tokushige studied the class of linear systems $(2T_{k,l})$. This class is defined as follows: let $k \geq 1$ and $l \geq 2$, and let $a_1, \dots, a_{k+l} \in \mathbb{F}_q$ be non-zero such that $a_1 + \dots + a_{k+l} = 0$. Then $(2T_{k,l})$ is given by the coefficient matrix

$$\begin{pmatrix} a_1 & \cdots & a_k & a_{k+1} & \cdots & a_{k+l} & 0 & \cdots & 0 \\ a_1 & \cdots & a_k & 0 & \cdots & 0 & a_{k+1} & \cdots & a_{k+l} \end{pmatrix} \in \mathbb{F}_q^{2 \times (k+2l)}.$$

In [MT20, Thm. 6], Mimura and Tokushige showed that such a system is always moderate. (This contains the linear system (S_2) from [MT20] as a special case.)

This result can be recovered as a special case of Theorem A, and strengthened to $(2T_{k,l})$ being temperate by Theorem B. The argument is analogous to that of Example 7.7.

Example 7.9. In [MT20], Mimura and Tokushige studied the linear system (S_3^-) with coefficient matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 & -4 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & -4 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & -4 \end{pmatrix} \in \mathbb{F}_q^{3 \times 10},$$

and proved that it is moderate whenever $p \neq 2$.¹³

This result can be recovered as a special case of Theorem A, provided that $p \neq 2, 3$.¹⁴ The results from this paper are insufficient to determine whether (S_3^-) is temperate, because there are not enough equations to apply Theorem B(i).

Example 7.10. Finally, in [MT20, Conjecture 1], Mimura and Tokushige conjectured that the system (S_3) with coefficient matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 & -4 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & -4 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & -4 \end{pmatrix} \in \mathbb{F}_q^{3 \times 11}$$

¹³The authors don't make the assumption $p \neq 2$ explicit in their proof. This assumption is necessary because the sum of the second and third row of the coefficient matrix is congruent to $(0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 0 \ -1 \ 0) \pmod{2}$. So for $p = 2$ the system cannot be moderate because it forces two variables to be equal.

¹⁴If $p = 2$, then there are three column equivalence classes of size 1, so the system is not of type (RC). Furthermore, if $p \in \{2, 3\}$, then there are column equivalence classes of size 2 that sum to zero, but not all column equivalence classes sum to 0, so neither Theorem A(i) nor Theorem A(ii) applies in this case. If $p \notin \{2, 3\}$, then the system is of type (RC) and none of column equivalence classes sums to zero, so Theorem A(i) applies.

is moderate. This is confirmed by our results. If $p \neq 2$, then it follows from Theorem A(i) and Theorem B(i) that (S_3) is moderate and temperate. If $p = 2$, then some of the columns become zero, so they correspond to free variables. After removing those columns, it follows from Theorem A(ii) and Theorem B(ii) that (S_3) is moderate and temperate.

In summary: in all examples except Example 7.9, we were able to prove that the system is moderate and temperate, thereby strengthening prior results (and proving a conjecture) of Mimura and Tokushige. In Example 7.9, we gave an alternative proof of the fact that the system is moderate, but we were unable to determine whether the system is also temperate.

In Example 7.5, we could not apply Theorem A. Instead, we needed a proof that was adapted to this particular system, using results from §3, to furnish an alternative proof that the system is moderate. In all other examples, the fact that the system is moderate follows immediately from Theorem A.

Acknowledgements The first author is partially supported by the Dutch Research Council (NWO), project number 613.009.127.

References

- [BCC⁺17] Jonah Blasiak, Thomas Church, Henry Cohn, Joshua A. Grochow, Eric Naslund, William F. Sawin, and Chris Umans. On cap sets and the group-theoretic approach to matrix multiplication. *Discrete Analysis*, 2017:3, 2017. 27pp. [doi:10.19086/da.1245](https://doi.org/10.19086/da.1245).
- [Bou90] J. Bourgain. On arithmetic progressions in sums of sets of integers. In A. Baker, B. Bollobás, and A. Hajnal, editors, *A tribute to Paul Erdős*, pages 105–110. Cambridge University Press, 1990. [doi:10.1017/CB09780511983917.008](https://doi.org/10.1017/CB09780511983917.008).
- [CLP17] Ernie Croot, Vsevolod L. Lev, and Péter Pál Pach. Progression-free sets in $\mathbb{Z}_4^n$ are exponentially small. *Annals of Mathematics*, 185(1):331–337, 2017. [doi:10.4007/annals.2017.185.1.7](https://doi.org/10.4007/annals.2017.185.1.7).
- [Dob23] Josse van Dobben de Bruyn. *Divisorial gonality of graphs, the slice rank polynomial method, and tensor products of convex cones*. PhD thesis, TU Delft, 2023. [doi:10.4233/uuid:bb2db244-e032-46bd-a9d7-a36b9ce0ce0e](https://doi.org/10.4233/uuid:bb2db244-e032-46bd-a9d7-a36b9ce0ce0e).
- [EG17] Jordan S. Ellenberg and Dion Gijswijt. On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression. *Annals of Mathematics*, 185(1):339–343, 2017. [doi:10.4007/annals.2017.185.1.8](https://doi.org/10.4007/annals.2017.185.1.8).
- [MT19a] Masato Mimura and Norihide Tokushige. *Avoiding a star of three-term arithmetic progressions*. Preprint, 2019. [arXiv:1909.10507](https://arxiv.org/abs/1909.10507).
- [MT19b] Masato Mimura and Norihide Tokushige. *Avoiding a shape, and the slice rank method for a system of equations*. Preprint, 2019. [arXiv:1909.10509](https://arxiv.org/abs/1909.10509).
- [MT20] Masato Mimura and Norihide Tokushige. *Solving linear equations in a vector space over a finite field II*. Preprint, 2020. URL: <http://www.cc.u-ryukyu.ac.jp/~hide/sol2.pdf>.

- [Nas20] Eric Naslund. Exponential bounds for the Erdős–Ginzburg–Ziv constant. *Journal of Combinatorial Theory, Series A*, 174:105185, 2020. [doi:10.1016/j.jcta.2019.105185](https://doi.org/10.1016/j.jcta.2019.105185).
- [Ruz93] Imre Z. Ruzsa. Solving a linear equation in a set of integers I. *Acta Arithmetica*, 65(3):259–282, 1993. [doi:10.4064/aa-65-3-259-282](https://doi.org/10.4064/aa-65-3-259-282).
- [Ruz95] Imre Z. Ruzsa. Solving a linear equation in a set of integers II. *Acta Arithmetica*, 72(4):385–397, 1995. [doi:10.4064/aa-72-4-385-397](https://doi.org/10.4064/aa-72-4-385-397).
- [Sau21] Lisa Sauermann. On the size of subsets of $\mathbb{F}_p^n$ without p distinct elements summing to zero. *Israel Journal of Mathematics*, 243(1):63–79, 2021. [doi:10.1007/s11856-021-2145-x](https://doi.org/10.1007/s11856-021-2145-x).
- [Sau23] Lisa Sauermann. Finding solutions with distinct variables to systems of linear equations over $\mathbb{F}_p^n$. *Mathematische Annalen*, 386(1-2):1–33, 2023. [doi:10.1007/s00208-022-02391-y](https://doi.org/10.1007/s00208-022-02391-y).
- [Tao16] Terence Tao. *Notes on the “slice rank” of tensors*. Blog post, 2016. URL: <https://terrytao.wordpress.com/2016/08/24/notes-on-the-slice-rank-of-tensors/>.
- [TV06] Terence Tao and Van H. Vu. *Additive Combinatorics*, volume 105 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 2006.