Common systems of two equations over the binary field

Daniel Král'[*1], Ander Lamaison[1], and Péter Pál Pach[†2]

[1]Faculty of Informatics, Masaryk University, Botanická 68A, 602 00 Brno, Czech Republic.
[2]Department of Computer Science and Information Theory, Budapest University of Technology and Economics, Műegyetem rkp. 3., H-1111 Budapest, Hungary; MTA-BME Lendület Arithmetic Combinatorics Research Group, ELKH, Műegyetem rkp. 3., H-1111 Budapest, Hungary.

Abstract

A system of linear equations over a finite field $\mathbb{F}_q$ is said to be common if, among all two-colorings of $\mathbb{F}_q^n$, the uniform random coloring minimizes the number of monochromatic solutions asymptotically. The notion of common systems of linear equations was introduced by Saad and Wolf, as an analogue to the well-studied notion of common graphs.

Fox, Pham and Zhao characterized the common systems consisting of one equation. We study systems consisting of two equations over the binary field $\mathbb{F}_2$. We characterize, up to a finite number of cases, which systems with an odd number of variables are common. Our characterization answers a question by Kamčev, Liebenau and Morrison in the affirmative way whether there exist common systems of equations that are not translation invariant.

1 Introduction

The study of monochromatic solutions of systems of equations in colored arithmetic structures is one of the central topics in additive combinatorics. An 1892 result of this type by Hilbert [6] is often cited as the very first result in Ramsey theory, preceding Ramsey's Theorem [11] itself. Theorems by Van der Waerden [15] and Rado [10] are other famous examples of statements of this kind. However, these results only concern the existence of a monochromatic solutions, and they do not deal with the number of solutions, or with the structure of the solution set.

The problem that we study here originates from the quantitative question concerning Ramsey type results in graph theory: the notion of common graphs. This notion is closely related to one of the most well-known open problems in extremal combinatorics—Sidorenko's Conjecture. A graph H is common if, whenever the edges of K_n are colored in two colors, the number of monochromatic copies of H is asymptotically minimized for the uniformly random coloring. In other words, as n tends to infinity, the proportion of morphisms from H to K_n which are monochromatic tends to at least $2^{|E(H)|-1}$ regardless of a two-coloring of the edges of K_n. Goodman [5] proved that the triangle is common and Erdős [3] conjectured that all cliques are common. Burr and Rosta [1] extended the conjecture of Erdős and conjectured that every graph is common. Sidorenko [13] disproved the latter conjecture, by proving that a triangle with a pendant edge is not common. The conjecture of Erdős' was disproved by Thomason [14], who proved that no clique on at least four vertices is common. More generally, any

[*]E-mail: {dkral,lamaison}@fi.muni.cz. The work of D.K. and A.L. has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 648509). This publication reflects only its authors' view; the European Research Council Executive Agency is not responsible for any use that may be made of the information it contains. The authors were also supported by the MUNI Award in Science and Humanities of the Grant Agency of Masaryk University.

[†]P.P.P. was supported by the Lendület program of the Hungarian Academy of Sciences (MTA) and by the National Research, Development and Innovation Office NKFIH (Grant Nr. K124171 and K129335).

graph containing K_4 is not common [7]. We note that Sidorenko's Conjecture asserts that quasirandom graphs minimizes densities of bipartite graphs, and so if true, it would imply that all bipartite graphs are common.

We study the notion of common systems of linear equations over $\mathbb{F}_q$, which was introduced by Saad and Wolf [12]. Given a system L and a subset $A \subseteq \mathbb{F}_q^n$, we denote by $\mathrm{sol}(L; A)$ the set of elements $\mathbf{x} \in A^k$ with $L(\mathbf{x}) = 0$. The set A should be understood as one of the color classes while the other color class is $\mathbb{F}_q^n \setminus A$. We say that a system L is non-degenerate if each variable is constrained by at least one of the equations.

Definition 1. *Let L be a non-degenerate, full rank system of m linear equations over k variables, on $\mathbb{F}_q$. We say that L is common if, for any positive integer n and any $A \subseteq \mathbb{F}_q^n$ we have*

$$|\mathrm{sol}(L; A)| + |\mathrm{sol}(L; \mathbb{F}_2^n \setminus A)| \geq \frac{|\mathrm{sol}(L; \mathbb{F}_2^n)|}{2^{k-1}}. \tag{1}$$

Otherwise, we say that L is uncommon.

Cameron, Cilleruelo and Serra [2] proved that every linear equation with non-zero coefficients on an odd number of variables is common. Saad and Wolf [12] proved that any equation with an even number of variables such that the coefficients can be split into pairs adding up to zero is common. The characterization of the common systems of a single linear equation was given by Fox, Pham and Zhao [4], who showed that an equation is common if and only if it belongs to one of the classes previously described. An immediate corollary of the just mentioned results is that every linear equation over the two element field $\mathbb{F}_2$ is common.

Kamčev, Liebenau and Morrison [8, 9] gave several results concerning common systems L of two or more linear equations. Some of these results give conditions on subsystems induced by L, which imply that L is uncommon; a system L induces a subsystem if the subsystem is implied by L. For example, Kamčev et al. [9] showed that if L induces a subsystem of two equations on four variables, then L is uncommon. Consequently, if every solution of L contains an arithmetic progression of length four formed by the same variables, then L is uncommon. Kamčev et al. [9] also posed several open questions. In particular, they asked whether there exists a system of linear equations of rank at least two that is common but not translation invariant.

2 Our results

Our results concern systems of two linear equations over the binary field $\mathbb{F}_2$. Since the only non-zero coefficient in $\mathbb{F}_2$ is 1, every linear system of two equations can be written in the form

$$x_1 + x_2 + \cdots + x_r = x'_1 + x'_2 + \cdots + x'_s = x''_1 + x''_2 + \cdots + x''_t$$

for some values of r, s, t. We denote this system by $L_{r,s,t}$.

In this work, we consider the case that the number $r + s + t$ of variables is odd only. While we have also obtained results for the case when the sum is even, the analysis used in the proof is more complicated and we intend to complete them and included to the journal version of this work. We discuss the differences between the even and the odd cases when we sketch the proof of our main result in Section 3. In the odd case, our main result is a characterization of common systems up to a finite number of triples (r, s, t):

Theorem 2. *(a) If all r, s and t are odd, then $L_{r,s,t}$ is common.*

(b) If r is odd, s and t are even and $t \geq 2r + s$, then $L_{r,s,t}$ is uncommon.

(c) There exists a constant C such that if r is odd, s and t are even, $s \leq t < 2r + s$ and $t \geq C$, then $L_{r,s,t}$ is common.

The remaining cases can be obtained by permuting r, s and t.

The constant C in Theorem 2(c) arises from an estimate in a certain bound in the proof, and could potentially be just an artifact of the proof. In fact, numerical computations for small cases suggest that the statement holds with $C = 0$.

The significance of the condition $t \geq 2r + s$ is that it is the threshold at which the random coloring becomes locally uncommon. In Case (b), the coloring that violates (1) can be obtained by changing the color of relatively few elements of the uniformly random coloring.

The systems described in Case (c) are common, have rank two and are not translation invariant (since it contains the equation $x_1 + \cdots + x_r + x'_1 + \cdots + x'_s = 0$, with an odd number of variables, adding the same non-zero vector to all variables of a solution yields a non-solution). Hence, this answers the aforementioned question of Kamčev, Liebenau and Morrison. In fact, we can show that $L_{1,2,2}$ is common using a Cauchy-Schwarz argument.

3 Sketch of the proof

Similarly to [4], the two main techniques employed in our arguments are using a Fourier transform and replacing the original integer optimization problem by its linear relaxation.

Let $L = L_{r,s,t}$ be a system of two linear equations over the binary filed, and let $k = r+s+t$. Further, let $\mathbb{F}_2^n = R \cup B$ be a coloring of the elements of $\mathbb{F}_2^n$, i.e., the elements of one of the colors form the set R and of the other color the set B. For simplicity, we denote the variables as $\mathbf{x} = (x_1, x_2, \ldots, x_k)$. The number of monochromatic solutions can be written as

$$|\mathrm{sol}(L, R)| + |\mathrm{sol}(L, B)| = \sum_{\mathbf{x} \in \mathrm{sol}(L, \mathbb{F}_2^n)} \left(\prod_{i=1}^{k} 1_R(x_i) + \prod_{i=1}^{k} 1_B(x_i) \right) \tag{2}$$

where 1_R and 1_B are the indicator functions for the sets R and B.

A system L is common if and only if the equation (2) has a value at least $2^{n(k-2)-(k-1)}$ for any n and any function $1_R : \mathbb{F}_2^n \to \{0, 1\}$ and $1_B = 1 - 1_R$, which is equivalent to (1) with $A = R$ (if (2) drops below $2^{n(k-2)-(k-1)}$ for one n, it fails for infinitely many n's by a multiplicative factor). It can be observed that, if the system L is common, then the same inequality also holds if 1_R is replaced by any function $f : \mathbb{F}_2^n \to [0, 1]$, and 1_B by $1 - f$, i.e., it holds that

$$\sum_{\mathbf{x} \in \mathrm{sol}(L, \mathbb{F}_2^n)} \left(\prod_{i=1}^{k} f(x_i) + \prod_{i=1}^{k} (1 - f)(x_i) \right) \geq 2^{n(k-2)-(k-1)} \tag{3}$$

for all functions $f : \mathbb{F}_2^n \to [0, 1]$. The reason is as follows. If the function f is a counterexample for (3), then it is possible to obtain a partition $\mathbb{F}_2^N = R \cup B$ with $N \gg n$ by adding each $(y_1, y_2, \ldots, y_N) \in \mathbb{F}_2^N$ to R independently with probability $f(y_1, y_2, \ldots, y_n)$. With high probability, this partition will not satisfy (1).

A key step in deciding whether all functions f satisfy (3) is to consider the Fourier transform of (3). For each $y \in \mathbb{F}_2^n$, we define the Fourier coefficient

$$\hat{f}(y) = 2^{-n} \sum_{x \in \mathbb{F}_2^n} (-1)^{\langle x, y \rangle} f(x).$$

With this transformation, the condition (3) for L being common is transformed to

$$\begin{aligned} 2^{1-r-s-t} \leq & \hat{f}(0)^{r+s+t} + (1-\hat{f}(0))^{r+s+t} \\ & + \left(\hat{f}(0)^r - (\hat{f}(0)-1)^r\right) \sum_{y \in \mathbb{F}_2^n \setminus \{0\}} \hat{f}(y)^{s+t} \\ & + \left(\hat{f}(0)^s - (\hat{f}(0)-1)^s\right) \sum_{y \in \mathbb{F}_2^n \setminus \{0\}} \hat{f}(y)^{t+r} \\ & + \left(\hat{f}(0)^t - (\hat{f}(0)-1)^t\right) \sum_{y \in \mathbb{F}_2^n \setminus \{0\}} \hat{f}(y)^{r+s}, \end{aligned} \tag{4}$$

where $\hat{f}$ is the Fourier transform of any function $f : \mathbb{F}_2^n \to [0,1]$. This is the crucial point at which the parity of $r+s+t$ becomes important: the Fourier transform of the terms corresponding to the product of f contain a term, $\sum_{y \neq 0} \sum_{z \notin \{0,y\}} \hat{f}(y)^r \hat{f}(z)^s \hat{f}(y+z)^t$, which cancels out with a similar term coming from the product of $1-f$. When $r+s+t$ is even, these terms add up rather than cancel out, bringing an additional layer of complexity to the equation.

Using the transformation to (4), the proof of Theorem 2(a) is easy: the last three lines of (4) are trivially non-negative and the inequality in the first line holds by convexity of x^{r+s+t}. Theorem 2(b) is proven by giving an explicit description of a function that violates our inequality. Finally, Theorem 2(c), which is the most complex case, is established using Parseval's identity in combination with a fine case analysis.

References

[1] S.A. Burr and V. Rosta, On the Ramsey multiplicities of graphs - problems and recent results. *Journal of Graph Theory* 4 (1980), 347–361.

[2] P. Cameron, J. Cilleruelo, and O. Serra, On monochromatic solutions of equations in groups. *Revista Matemática Iberoamericana* 23 (2007), 385–395.

[3] P. Erdős, On the number of complete subgraphs contained in certain graphs. *A Magyar Tudományos Akadémia. Matematikai Kutató Intézetének Közleményei* 7 (1962), 459–464.

[4] J. Fox, H.T. Pham and Y. Zhao, Common and Sidorenko linear equations. *Quarterly Journal of Mathematics* 72 (2021), 1223–1234.

[5] A.W. Goodman, On sets of acquaintances and strangers at any party. *American Mathematical Monthly* 66 (1959), 778–783.

[6] D. Hilbert, Über die Irerducibilität ganzer rationaler Functionen it ganzzahligen Coefficienten. *Journal für die Reine und Angewandte Mathematik* 110 (1892), 104–129.

[7] C. Jagger, P. Šťovíček and A. Thomason, Multiplicities of subgraphs, Combinatorica 16 (1996), 123–141.

[8] N. Kamčev, A. Liebenau and N. Morrison, On uncommon systems of equations (2021). ArXiv:2106.08986.

[9] N. Kamčev, A. Liebenau and N. Morrison, Towards a characterisation of Sidorenko systems (2021). ArXiv:2107.14413.

[10] R. Rado, Studien zur Kombinatorik. *Mathematische Zeitschrift* 36 (1933), 424–470.

[11] F.P. Ramsey, On a problem of formal logic. *Proceedings of the London Mathematical Society* 30 (1930), 264–286.

[12] A. Saad and J. Wolf, Ramsey multiplicity of linear patterns in certain finite abelian groups. *Quarterly Journal of Mathematics* 68 (2017), 125–140.

[13] A.F. Sidorenko. Cycles in graphs and functional inequalities. *Matematicheskie Zametki* 46 (1989), 72–79.

[14] A. Thomason, A disproof of a conjecture of Erdös in Ramsey theory. *Journal of the London Mathematical Society* 2 (1989), 246–255.

[15] B. van der Waerden, Beweis einer baudetschen Vermutung. *Nieuw Archief voor Wiskunde* 15 (1927), 212–216.