COMMON AND SIDORENKO LINEAR EQUATIONS

JACOB FOX, HUY TUAN PHAM, AND YUFEI ZHAO

Abstract. A linear equation with coefficients in $\mathbb{F}_q$ is common if the number of monochromatic solutions in any two-coloring of $\mathbb{F}_q^n$ is (as $n \to \infty$) asymptotically at least the expected number in a random two-coloring. The linear equation is Sidorenko if the number of solutions in any dense subset of $\mathbb{F}_q^n$ is asymptotically at least the expected number in a random set of the same density.

In this paper, we characterize those linear equations which are common, and those which are Sidorenko. The main novelty is a construction based on choosing random Fourier coefficients that show that certain linear equations do not have these properties. This solves problems posed in the paper of Saad and Wolf.

1. INTRODUCTION

Sidorenko’s conjecture [20] (also conjectured earlier in a stronger form by Erdős and Simonovits [22]) is a major open problem in extremal graph theory. We say that a bipartite graph $H$ is Sidorenko if the density of copies of $H$ in a graph with fixed edge density is asymptotically minimized by the random graph with the same edge density. Sidorenko’s conjecture says that all bipartite graphs $H$ are Sidorenko. Many graphs are known to have the Sidorenko property, including bipartite graphs with a vertex complete to the other part, see [4, 5, 6, 7, 16, 17, 18, 23]. A coloring variant of Sidorenko’s conjecture, the Burr-Rosta conjecture [2] (extending an earlier conjecture of Erdős [8]), claims that the density of monochromatic copies of any fixed graph $H$ in any coloring of $K_n$ is asymptotically minimized by the random coloring. While the Burr-Rosta conjecture was disproved by Thomason [24] and Sidorenko [21], many graphs $H$ are known to satisfy the Burr-Rosta conjecture, such graphs are called common graphs.

Saad and Wolf [19] explored various analogues of Sidorenko and common graphs in the arithmetic setting. In this setting, we consider the minimum density of solutions to a system of linear equations in a subset of given density in a finite abelian group, or the minimum density of monochromatic solutions to a system of linear equations in a two-coloring of the abelian group. In particular, we consider the setting of a fixed linear system $L$ with coefficients in some finite field $\mathbb{F}_q$. We say that a system of linear equations $L = 0$ is common if the density of monochromatic solutions to this system in any two-coloring of $\mathbb{F}_q^n$ is asymptotically at least what we expect from a random coloring. Likewise, we say that a linear system $L = 0$ is Sidorenko if the density of solutions to this system in any dense subset of $\mathbb{F}_q^n$ is asymptotically at least what we expect from a random set with the same density. The formal definitions for single linear homogeneous equations are given below. While the above arithmetic problems do not directly
Conjecture 1.2. Is it true that adding sufficiently many free variables makes any linear system common as well as those that are Sidorenko, resolving the above conjecture and question.

Theorem 1.4. Let \( L(x_1, \ldots, x_k) = a_1x_1 + \cdots + a_kx_k \) be a linear form with \( a_1, \ldots, a_k \in \mathbb{F}_q \setminus \{0\} \).

(a) If \( a_1, \ldots, a_k \) can be partitioned into pairs each summing to zero, then the equation \( L = 0 \) is Sidorenko and common.

(b) If \( k \) is even and \( a_1, \ldots, a_k \) cannot be partitioned into pairs each summing to zero, then the equation \( L = 0 \) is not common and not Sidorenko.

(c) If \( k \) is odd, then the equation \( L = 0 \) is common but not Sidorenko.
A neat part of the proof is that of Theorem 1.4(b), which constructs a random function by selecting random value of its Fourier transform in order to identify the canceling pairs of coefficients $a_i$.

Adding a free variable to a linear equation is the same as adding a variable with coefficient 0 to the equation. Hence, the density of solutions to the new equation (with $l$ coefficients being 0) is simply the density of solutions to the original equation (with only nonzero coefficients), multiplied by the set density to the power $l$. The following theorem answers Question 1.3.

**Theorem 1.5.** Let $L(x_1, \ldots, x_k) = a_1 x_1 + \cdots + a_k x_k$ be a linear form with $a_1, \ldots, a_k \in \mathbb{F}_q \setminus \{0\}$. Let $L'$ be the linear form obtained by adding $\ell \geq 1$ free variables to $L$ (equivalently, $L'(x_1, \ldots, x_{k+l}) = a_1 x_1 + \cdots + a_k x_k + 0x_{k+1} + \cdots + 0x_{k+l}$). If $a_1, \ldots, a_k$ can be partitioned into pairs each summing to zero, the equation $L' = 0$ is Sidorenko and common, and otherwise $L' = 0$ is not common and not Sidorenko.

It is natural to study the same problems for inhomogeneous linear equations.

**Definition 1.6.** Let $L(x_1, \ldots, x_k) = a_1 x_1 + \cdots + a_k x_k$ be a linear form with $a_1, \ldots, a_k \in \mathbb{F}_q \setminus \{0\}$. We say that the linear form $L$ is **inhomogeneous-Sidorenko** if for every $n$, every nonzero $b \in \mathbb{F}_q^n$, and every $A \subseteq \mathbb{F}_q^n$, the number of solutions to $L(x_1, \ldots, x_k) = b$ with $x_1, \ldots, x_k \in A$ is at least $|A|^k/q^n$. We say that the linear form $L$ is **inhomogeneous-common** if for every $n$, every nonzero $b \in \mathbb{F}_q^n$, and every coloring of $\mathbb{F}_q^n$ with two colors, the number of monochromatic solutions $(x_1, \ldots, x_k) \in (\mathbb{F}_q^n)^k$ to $L(x_1, \ldots, x_k) = b$ is at least $2^{1-k}q^{n(k-1)}$.

We remark that the choice of nonzero $b$ above is inconsequential, since for any $n$, if the properties in Definition 1.6 are satisfied for any nonzero $b \in \mathbb{F}_q^n$, then they are satisfied for every nonzero $b \in \mathbb{F}_q^n$. Indeed, for any nonzero $b, b' \in \mathbb{F}_q^n$, we can find an invertible linear transformation $C$ such that $Cb = b'$, then $a_1 x_1 + \cdots + a_k x_k = b$ if and only if $a_1 C x_1 + \cdots + a_k C x_k = b'$, so the solutions to $L = b$ in $A$ and the solutions to $L = b'$ in $CA$ are in one-to-one correspondence via the invertible transformation $C$. The next result gives a simple characterization of the inhomogeneous-common linear forms, and shows that no linear form is inhomogeneous-Sidorenko.

**Theorem 1.7.** Let $L(x_1, \ldots, x_k) = a_1 x_1 + \cdots + a_k x_k$ be a linear form with $a_1, \ldots, a_k \in \mathbb{F}_q \setminus \{0\}$. Let $L'$ be the linear form obtained by adding $\ell \geq 0$ free variables to $L$.

Then $L'$ is never inhomogeneous-Sidorenko, and $L'$ is inhomogeneous-common if and only if $k$ is odd and $\ell = 0$.

## 2. Proofs

We first note that the Sidorenko property is equivalent to its functional version, where we replace the subset $A \subseteq \mathbb{F}_q^n$ by a function $f : \mathbb{F}_q^n \to [0,1]$. Indeed, given a linear form $L(x_1, \ldots, x_k) = a_1 x_1 + \cdots + a_k x_k$ and a function $f : \mathbb{F}_q^n \to [0,1]$, we write

$$L = 0 \iff \Lambda_{L=0}(f) := E_{x=(x_1, \ldots, x_k) \in (\mathbb{F}_q^n)^k : L(x)=0} [f(x_1) \cdots f(x_k)].$$

Then the equation $L = 0$ is Sidorenko if and only if for every $n$ and every $f : \mathbb{F}_q^n \to [0,1]$,

$$\Lambda_{L=0}(f) \geq (Ef)^k.$$  \hfill (1)

Indeed, being Sidorenko in the sense of Definition 1.1 is equivalent to (1) for all $f$ of the form $1_A$ with $A \subseteq \mathbb{F}_q^n$. Conversely, suppose (1) fails for some $f : \mathbb{F}_q^n \to [0,1]$. We can extend
It suffices to show that equation is automatically common. It suffices to show that there exists $f$ given $L$ such that $\hat{f}$ exists where $\hat{t}$ the convexity of $n$ coordinates distinct, which goes to 0 as $n' \to \infty$. Since $\mathbb{E}_A[(1] = \mathbb{E}_A[|f|] = \mathbb{E}_A$ by the convexity of $t \mapsto t^k$, it follows that if (1) fails for some $f$, then for $n'$ large enough, there exists $A \subseteq \mathbb{F}_q^{n+n'}$ such that (1) fails for $f = 1_A$. Thus the set version and the functional formulations of the property of being Sidorenko are equivalent. The same argument shows that if $\Lambda_{L=0}(A) \geq (|A|/q)^k + o_n(1)$ for every $A$, where the $o_n(1)$ term tends to 0 as $n \to \infty$ uniformly in $A$, then in fact $\Lambda_{L=0}(A) \geq (|A|/q)^k$.

Likewise, the property of being common also has an equivalent functional formulation: given a $k$-variable linear form $L$ over $\mathbb{F}_q$, the equation $L = 0$ is common if and only if for every $f: \mathbb{F}_q^n \to [0,1]$, \[
\Lambda_{L=0}(f) + \Lambda_{L=0}(1-f) \geq 2^{1-k}.
\]

We denote the Fourier transform of a function $f: \mathbb{F}_q^n \to \mathbb{C}$ by \[
\hat{f}(r) = \mathbb{E}_{x \in \mathbb{F}_q^n} f(x)r(x), \quad r \in \mathbb{F}_q^n
\]
where $\hat{G}$ is the group of characters of $G$, i.e., homomorphisms $r: G \to \mathbb{C}^\times$. Though, except for the evaluation $r(x)$, we write $\hat{G}$ additively, so that we can write, e.g., $\hat{f}(2r)$, and also use $r \neq 0$ to mean a character other than the identity.

Recall the following identity which relates a twisted convolution with the Fourier transform: given $L(x_1, \ldots, x_k) = a_1x_1 + \cdots + a_kx_k$, \[
\Lambda_{L=0}(f) = \sum_{r \in \mathbb{F}_q^n} \hat{f}(a_1r) \cdots \hat{f}(a_kr). \tag{2}
\]

Since $\hat{f}(0) = Ef$ and $(1-f)(r) = -\hat{f}(r)$ for all $r \neq 0$, we have \[
\Lambda_{L=0}(f) + \Lambda_{L=0}(1-f) = (Ef)^k + (1-Ef)^k + (1+(-1)^k) \sum_{r \neq 0} \hat{f}(a_1r) \cdots \hat{f}(a_kr). \tag{3}
\]

**Proof of Theorem 1.4.** (a) Write $m = k/2$ and $L(x_1, x'_1, \ldots, x_m, x'_m) = a_1(x_1 - x'_1) + \cdots + a_m(x_m - x'_m)$. For $f: \mathbb{F}_q^n \to [0,1]$, we have \[
\Lambda_{L=0}(f) = \mathbb{E}_z (\mathbb{E}_{x_1, \ldots, x_m} \mathbb{1}_{a_1x_1 + \cdots + a_mx_m} f(x_1) \cdots f(x_m))^2 \geq (\mathbb{E}_z \mathbb{E}_{x_1, \ldots, x_m} \mathbb{1}_{a_1x_1 + \cdots + a_mx_m} f(x_1) \cdots f(x_m))^2 = (\mathbb{E}_z \mathbb{E}_{x_1, x_m} f(x_1) \cdots f(x_m))^2 = (Ef)^{2m}.
\]

Thus $L = 0$ is Sidorenko, which then implies that it must be common, since $\Lambda_{L=0}(f) + \Lambda_{L=0}(1-f) \geq (Ef)^k + (1-Ef)^k \geq 2^{1-k}$ by the convexity of $t \mapsto t^k$.

(b) It suffices to show that $L = 0$ is not common, since as we just noted, every Sidorenko equation is automatically common. It suffices to show that there exists $f: \mathbb{F}_q \to [0,1]$ such that \[
\Lambda_{L=0}(f) + \Lambda_{L=0}(1-f) < 2^{1-k}.
\]
Since \( k \) is even, it remains to exhibit some \( f : \mathbb{F}_q \to [0, 1] \) such that \( \mathbb{E} f = 1/2 \) and such that the sum in the right-hand side of (3) above is negative. We shall do it by choosing random values for \( \hat{f} \).

For each \( r \in \mathbb{F}_q \setminus \{0\} \), let \( \xi_r \) be a random unit complex number. If \( q \) is odd, we choose the \( \xi_r \)'s subject to \( \xi_r = \overline{\xi_{-r}} \) but i.i.d. uniform otherwise. If \( q \) is even, we choose each \( \xi_r \in \{-1, 1\} \) uniformly i.i.d. Construct a random function \( f : \mathbb{F}_q \to [0, 1] \) by setting \( \hat{f}(0) = 1/2 \) and \( \hat{f}(r) = \xi_r/(2q) \) for every \( r \in \mathbb{F}_q \setminus \{0\} \). By the inverse transform \( f(x) = \sum_{r \in \mathbb{F}_q} \hat{f}(r)r(x) \), we see that such \( f \) takes real values in \([0, 1]\).

Then, for every \( r \in \mathbb{F}_q \setminus \{0\} \), the product \( \hat{f}(a_1r) \cdots \hat{f}(a_kr) \) has expectation zero over this random \( f \), since \( a_1, \ldots, a_k \) cannot be partitioned into “canceling” pairs each summing to zero.

Thus, the expectation of (3) over this random \( f \) is equal to \( 2^{1-k} \). However, if all \( \xi_r \) are sufficiently close to 1 (in the case of even \( q \), when all \( \xi_r \) are equal to 1), then the value of (3) is greater than \( 2^{1-k} \). Thus, it follows that there is some \( f \) such that the value of (3) is strictly less than \( 2^{1-k} \), as desired.

(c) Since \( k \) is odd, by (3) and the convexity of \( t \mapsto t^k \),

\[
\Lambda_{L=0}(f) + \Lambda_{L=0}(1 - f) = (\mathbb{E} f)^k + (1 - \mathbb{E} f)^k \geq 2^{1-k},
\]

and thus \( L = 0 \) is common.

On the other hand, setting \( \hat{f}(0) = 1/2 \) and \( \hat{f}(r) = -1/(2q) \) for all \( r \neq 0 \), we see from (2) that \( \Lambda_{L=0}(f) < (\mathbb{E} f)^k \), so that \( L = 0 \) is not Sidorenko. (Alternatively, we can set \( A = \mathbb{F}_q \setminus \{0\} \) and deduce by the inclusion-exclusion principle that \( \Lambda_{L=0}(1_A) = (1 - 1/q)^k + (-1/q)^k(q - 1) \).)

Proof of Theorem 1.5. When \( a_1, \ldots, a_k \) can be partitioned into pairs each summing to zero, the proof that \( \Lambda' = 0 \) is Sidorenko and common follows from 1.4(a) since in this case \( \Lambda_{L'=0}(f) = (\mathbb{E} f)^k \Lambda_{L=0}(f) \geq (\mathbb{E} f)^{k+\ell} \).

Now consider the case when \( k \) is odd. It suffices to exhibit a function \( f : \mathbb{F}_q \to [0, 1] \) such that \( \Lambda_{L=0}(f) + \Lambda_{L'=0}(f) < 2^{1-k-\ell} \). Consider the function \( f : \mathbb{F}_q \to [0, 1] \) defined by

\[
f(x) = \begin{cases} \frac{1}{2} + c - (q - 1)\beta & \text{if } x = 0, \\ \frac{1}{2} + c + \beta & \text{if } x \neq 0. \end{cases}
\]

(4)

where \( \beta = 1/(2q) \) is a constant and \( c > 0 \) is sufficiently small. Its Fourier transform is given by

\[
\hat{f}(r) = \begin{cases} \frac{1}{2} + c & \text{if } r = 0, \\ -\beta & \text{if } r \neq 0, \end{cases}
\]

and \( (1 - f)(r) = \begin{cases} \frac{1}{2} - c & \text{if } r = 0, \\ \beta & \text{if } r \neq 0. \end{cases} \)

(5)

From (2), we find that

\[
\Lambda_{L=0}(f) = \left(\frac{1}{2} + c\right)^k + (q - 1)(-\beta)^k
\]

and

\[
\Lambda_{L=0}(1 - f) = \left(\frac{1}{2} - c\right)^k + (q - 1)\beta^k.
\]
Since $\Lambda_{L'=0}(f) = (Ef)^{\ell} \Lambda_{L=0}(f)$ and $\Lambda_{L'=0}(1-f) = (1 - Ef)^{\ell} \Lambda_{L=0}(1-f)$, and recall that $k$ is odd, we have

$$\Lambda_{L'=0}(f) + \Lambda_{L'=0}(1-f) = \left( \frac{1}{2} + c \right)^{\ell} \left( \left( \frac{1}{2} + c \right)^{k} - (q-1)\beta^{k} \right) + \left( \frac{1}{2} - c \right)^{\ell} \left( \left( \frac{1}{2} - c \right)^{k} + (q-1)\beta^{k} \right)$$

$$= \left( \frac{1}{2} + c \right)^{k+\ell} + \left( \frac{1}{2} - c \right)^{k+\ell} - (q-1)\beta^{k} \left( \left( \frac{1}{2} + c \right)^{\ell} - \left( \frac{1}{2} - c \right)^{\ell} \right)$$

$$= 2^{1-k-\ell} - (q-1)\beta^{k} 2^{-\ell+1} \beta + O(c^2),$$

which is less than $2^{-k-\ell+1}$ as long as $c$ is small enough.

Finally, suppose $k$ is even but $a_1, \ldots, a_k$ cannot be partitioned into pairs summing to zero. In the proof of Theorem 1.4(b) we constructed an $f: \mathbb{F}_q \to [0, 1]$ with $Ef = 1/2$ such that $\Lambda_{L=0}(f) + \Lambda_{L=0}(f) < 2^{1-k}$. Then this $f$ has $\Lambda_{L=0}(f) = (Ef)^{\ell} \Lambda_{L=0}(f) = 2^{-\ell} \Lambda_{L=0}(f)$ and likewise $\Lambda_{L=0}(f) = 2^{-\ell} \Lambda_{L=0}(f)$, and hence $\Lambda_{L=0}(f) + \Lambda_{L=0}(f) < 2^{1-k-\ell}$, so that $L'$ is not Sidorenko and not common. \hfill \Box

Proof of Theorem 1.7. The equation $L'(x) = b$ with $b \neq 0$ has no solutions with all $k$ coordinates of $x = (x_1, \ldots, x_k)$ lying in a subspace of $\mathbb{F}_q^k$ not containing $b$. Thus $L'$ cannot be inhomogeneous-Sidorenko.

Let us define

$$\Lambda_{L=b}(f) := \mathbb{E}_{x = (x_1, \ldots, x_k) \in (\mathbb{F}_q^k)^{k}: L(x) = b} [f(x_1) \cdots f(x_k)].$$

An extension of (2) for inhomogeneous equations gives

$$\Lambda_{L=b}(f) = \sum_{r \in \mathbb{F}_q^k} \widehat{f}(a_1r) \cdots \widehat{f}(a_kr)r(b)$$

(6)

Similar to (3), we have

$$\Lambda_{L=b}(f) + \Lambda_{L=b}(1-f) = (Ef)^{k} + (1 - Ef)^{k} + (1 + (-1)^{k}) \sum_{r \neq 0} \widehat{f}(a_1r) \cdots \widehat{f}(a_kr)r(b).$$

(7)

Thus when $\ell = 0$ and $k$ is odd, then the above expression is at least $(Ef)^{k} + (1 - Ef)^{k} \geq 2^{1-k}$ by the convexity of $t \mapsto t^k$, so that $L$ is inhomogeneous-common.

Now we turn to the negative cases. To show that $L'$ is not inhomogeneous-common, it suffices to exhibit a function $f: \mathbb{F}_q \to [0, 1]$ such that $\Lambda_{L'=1}(f) + \Lambda_{L'=1}(1-f) < 2^{1-k-\ell}$.

When $k$ is even, set $Ef = 1/2$ and $\widehat{f}(r) = 1/(2q)$ for all $r \neq 0$, say, so that $f$ takes values in $[0, 1]$, and then the value of (7) is less than $2^{1-k}$ since $\sum_{r \neq 0} r(b) = -1$, so that $\Lambda_{L=1}(f) + \Lambda_{L=1}(1-f) < 2^{1-k}$. Adding $\ell$ free variables in this case gives $\Lambda_{L=1}(f) + \Lambda_{L=1}(1-f) = (Ef)^{\ell} \Lambda_{L=1}(f) + (1 - Ef)^{\ell} \Lambda_{L=1}(1-f) < 2^{1-k-\ell}$.

Finally, suppose $\ell > 0$ and $k$ is odd. Take $f$ as in (4) except that now we set $\beta = -1/(2q)$. Then, using (5) and (7), and using that $\sum_{r \neq 0} r(1) = -1$, we have

$$\Lambda_{L=1}(f) = \left( \frac{1}{2} + c \right)^{k} - \left(- \beta \right)^{k}$$

and

$$\Lambda_{L=1}(1-f) = \left( \frac{1}{2} - c \right)^{k} - \beta^{k}.$$
Since $\Lambda_{L'=1}(f) = (Ef)^{\ell} \Lambda_{L=1}(f)$ and $\Lambda_{L'=1}(1-f) = (1-Ef)^{\ell} \Lambda_{L=1}(1-f)$, and recall that $k$ is odd, we have
\[
\Lambda_{L'=1}(f) + \Lambda_{L'=1}(1-f) = \left(\frac{1}{2} + c\right)^{\ell} \left(\left(\frac{1}{2} + c\right)^{k} - (-\beta)^{k}\right) + \left(\frac{1}{2} - c\right)^{\ell} \left(\left(\frac{1}{2} - c\right)^{k} - \beta^{k}\right)
\]
\[
= \left(\frac{1}{2} + c\right)^{k+\ell} + \left(\frac{1}{2} - c\right)^{k+\ell} + \beta^{k} \left(\frac{1}{2} + c\right)^{\ell} - \left(\frac{1}{2} - c\right)^{\ell}
\]
\[
= 2^{1-k-\ell} + \beta^{k}2^{1-\ell}c + O(c^2),
\]
which is less $2^{1-k-\ell}$ if $c > 0$ is small enough (recall that $\beta = -1/(2q)$ and $k$ is odd).

3. Concluding Remarks

In this paper, we characterized which linear equations are common and which are Sidorenko. It is natural to try to characterize which systems of linear equations satisfy these properties, and Saad and Wolf [19] proved several results on this. To be more precise, a system of linear equations $xM = 0$ in $k$ variables is Sidorenko if the density of solutions to $xM = 0$ in a set $A$ is at least the $k$-th power of the density of $A$. Here, we think of $x$ as a matrix whose columns are $x_1, \ldots, x_k$. A system of linear equations $xM = 0$ in $k$ variables is common if the density of monochromatic solutions to $xM = 0$ in any two-coloring of $\mathbb{F}_q^n$ is at least $2^{1-k}$. We currently don’t have a guess for a characterization of these properties for general linear systems. One reason why this is of interest is that it might lead to a better understanding of the analogous properties for graphs and hypergraphs.

**Question 3.1.** Which systems of linear equations are Sidorenko? Which are common?

While the linear homogeneous equation $x_1 - 2x_2 + x_3 = 0$ giving three-term arithmetic progressions is not Sidorenko, Green [13] introduced a weakening of the Sidorenko property which this equation does satisfy. Green proved that, for each $\epsilon > 0$ there is $N(\epsilon)$ such that if $G$ is an abelian group with $|G| \geq N(\epsilon)$ and $A \subseteq G$ has density $\alpha$, then there is a nonzero $d \in G$ such that the density of three-term arithmetic progressions with common difference $d$ is at least $\alpha^3 - \epsilon$. That is, while the total density of three-term arithmetic progressions can be much less than given by the random bound, there is a nonzero $d$ for which the density of three-term arithmetic progressions with common difference $d$ is at least almost the random bound. The proof uses an arithmetic regularity lemma, and consequently gives a tower-type bound on $N(\epsilon)$. The authors later proved that such a tower-type bound is needed (see [10, 11, 12]).

It is natural to try to see whether other systems of linear equations satisfy such a popular differences property. Green and Tao [14, 15] proved that the linear system giving four-term arithmetic progressions has this property, while Ruzsa [1] proved that for any longer length arithmetic progression, the corresponding linear system does not have this property.

We say that a single linear homogeneous equation $L = 0$ is popular if, for each $\epsilon > 0$ there is $n_L(\epsilon)$ such that if $n \geq n_L(\epsilon)$ and $A \subseteq \mathbb{F}_q^n$ has density $\alpha$, then there are nonzero and distinct $d_1, \ldots, d_{k-1}$ such that the density of solutions to $L = 0$ with $x_{i+1} - x_1 = d_i$ for $i = 1, \ldots, k-1$ is at least $\alpha^k - \epsilon$. We refer to $d_1, \ldots, d_{k-1}$ as the popular differences. A linear homogeneous equation $L = 0$ is translation-invariant if and only if the sum of the coefficients is zero. If the equation $L = 0$ is popular, then it must be translation-invariant. Indeed, if $L = 0$ is not translation-invariant, then the affine subspace $S$ of codimension one (so density $\alpha = 1/q$) consisting of those elements whose first coordinate is one is such that $L = 0$ has no solution in $S$. Hence, it follows that $n_L(\epsilon)$ does not exist for $\epsilon < 1/q^k$. 

For a translation-invariant linear equation, the reader may compare the above definition of popular differences to the discussion of popular difference for three-term arithmetic progressions, where the popular difference is a single parameter \( d = x_2 - x_1 \) instead of the tuple \( (d, 2d) = (x_2 - x_1, x_3 - x_1) \). Indeed, in the general case, fixing \( d_1, \ldots, d_{k-2} \) would determine a unique choice of \( d_{k-1} \) for which there exists a solution to \( L(x) = 0 \) with \( x_{i+1} - x_i = d_i \) for \( i = 1, \ldots, k - 1 \). However, in the case of three-term arithmetic progressions over abelian groups of odd order, fixing \( x_2 - x_1 = d \neq 0 \) ensures that \( x_3 - x_1 = 2d \neq 0 \), whereas fixing nonzero and distinct \((d_1, \ldots, d_{k-2})\) in the general case does not guarantee that \( x_k - x_1 = d_{k-1} \) is nonzero or distinct from \( d_1, \ldots, d_{k-2} \). Thus, the inclusion of \( d_{k-1} \) in the definition of popular differences does not correspond to a degree of freedom, but only serves to ensure that \( x_1, \ldots, x_k \) are distinct when \( x_{i+1} - x_i = d_i \) for \( 1 \leq i \leq k - 1 \).

Note that if \( L = 0 \) is Sidorenko, then simply by averaging, it is also popular and furthermore, \( n_L(\epsilon) \) is bounded above by \( O(\log(1/\epsilon)) \). Green’s theorem shows that the equation \( x_1 - 2x_2 + x_3 = 0 \) is popular. More generally, Green’s argument in [13] can be extended to show that \( L = 0 \) is popular if and only if it is translation-invariant. Indeed, we showed above that if \( L = 0 \) is not translation-invariant, then it is not popular. If \( L = 0 \) is translation-invariant, then the arithmetic regularity lemma proof of Green’s theorem goes as follows. For each subset \( A \subseteq \mathbb{F}_q^n \) of density \( \alpha \), by a Szemerédi type lemma, there is a regular subspace \( H \) of bounded codimension, and the counting lemma and Jensen’s inequality gives that the density of solutions to \( L = 0 \) with \( x_1, \ldots, x_k \) all in the same translate of \( H \) is at least almost \( \alpha^k \). By throwing out the solutions with \( x_1, \ldots, x_k \) not all distinct (which is of smaller order for \( n \) sufficiently large) and averaging, we get that there exists nonzero and distinct \( d_1, \ldots, d_{k-1} \in H \) for which the density of solutions to \( L = 0 \) with \( x_{i+1} - x_i = d_i \) for \( 1 \leq i \leq k - 1 \) is at least \( \alpha^k - \epsilon \). This proof gives an upper bound on \( n_L(\epsilon) \) which is a tower of height \( \epsilon^{-O(1)} \). The bound given by this proof here actually depends on \( q \) as well, but the regularity proof in [13] using Bohr sets instead of subspaces gives a bound that does not depend on \( q \).

The first two authors [10] showed, in the case \( L = 0 \) corresponding to three-term arithmetic progressions, \( n_L(\epsilon) \) is in fact bounded above and below by a tower of height \( \Theta(\log(1/\epsilon)) \). We can directly adapt the proof of the upper bound in [10] to show that for all translation invariant \( L = 0 \), we have \( n_L(\epsilon) \) is bounded above by a tower of height \( \Theta(\log(1/\epsilon)) \), using a density increment argument with the mean \( k \)-th power density, defined as \( b_k(H) = \mathbb{E}[f_H(x)^k] \), where \( f_H(x) \) is the average value of \( f \) on the affine translate of \( H \) containing \( x \).

The lower bound construction in [10] heavily depends on the fact that the equation \( x_1 - 2x_2 + x_3 = 0 \) is not Sidorenko, as a crucial ingredient of our construction is a model function with relatively low density of three-term arithmetic progressions. As mentioned above, if \( L = 0 \) is Sidorenko, then \( n_L(\epsilon) \) is not of tower-type, but in fact, only logarithmic in \( \epsilon \).

The converse does not hold. Indeed, we next exhibit a linear homogeneous equation in eight variables which is not Sidorenko but \( n_L(\epsilon) \) is only logarithmic in \( \epsilon^{-1} \). For example, we may take \( L = 0 \) with \( L(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8) = -6x_1 + 3x_2 + x_3 + 7x_4 + 2x_5 - 4x_6 - 2x_7 - x_8 \). A Hilbert cube of dimension \( t \) is a sequence of \( 2^t \) numbers \( x_1 + \sum_{i=1}^t \epsilon_i d_i \) with \( \epsilon_i \in \{0, 1\} \). The Hilbert cubes of dimension \( t \) are the solution set to a system \( xM_t = 0 \) of linear equations in \( 2^t \) variables. It follows from repeated application of the Cauchy–Schwarz inequality that the system \( xM_t = 0 \) is Sidorenko (see Example 2.5 in [19]). It follows that if \( n \) is more than logarithmic in \( \epsilon^{-1} \) and \( A \subseteq \mathbb{F}_q^n \), letting \( t = 3 \) and averaging, there is a set \( \{d_1, d_2, d_3\} \) such that all eight subsets of this set have nonzero and distinct sums, and the density of Hilbert cubes
of dimension three in $A$ with this choice of $d_1, d_2, d_3$ is at least $E[1_A]^8 - \epsilon$. Let $d_4 = d_1 + d_2$, $d_5 = d_1 + d_3$, $d_6 = d_2 + d_3$, and $d_7 = d_1 + d_2 + d_3$. The Hilbert cubes $(x_1, x_1 + d_1, x_1 + d_2, x_1 + d_3, x_1 + d_1 + d_2, x_1 + d_1 + d_3, x_1 + d_2 + d_3, x_1 + d_1 + d_2 + d_3)$ of dimension three with differences $d_1, d_2, d_3$ are precisely the solutions to the example $L = 0$ in eight variables above which further satisfy $x_{i+1} - x_i = d_i$ for $i = 1, \ldots, 7$. Thus there are equations $L = 0$ like the one above which are not Sidorenko but $n_L(\epsilon)$ grows only logarithmic in $\epsilon^{-1}$.

We do not know if there is a translation-invariant linear equation $L = 0$ with at least four variables for which $n_L(\epsilon)$ has tower-type growth. It would be further interesting to characterize them. For equations $L = 0$ with three variables which are translation-invariant, we can show a tight tower-type lower bound on $n_L(\epsilon)$ by a direct generalization of the argument in [10].

**Question 3.2.** Characterize the growth rate of $n_L(\epsilon)$ for linear equations $L = 0$ with at least four variables.

Related to the previous question, it would be interesting to know for which linear equations $L = 0$ can we get a logarithmic bound on $n_L(\epsilon)$ through the above technique by finding a linear system which is Sidorenko and whose solution set is a subset of the set of solutions to $L = 0$.

**Question 3.3.** For which linear equation $L = 0$ can we find a Sidorenko linear system of equations $xM = 0$ whose solutions are also solutions of $L = 0$?

In light of the regularity lemma proof of the popular differences theorem, it is natural to consider another notion of popular, replacing differences by subspaces. Call a translation-invariant linear homogeneous equation $L = 0$ in $k$ variables *subspace popular* if for each $\epsilon > 0$ there is some $\tilde{n}_L(\epsilon)$ (and define $\tilde{n}_L(\epsilon)$ to be the smallest such integer) such that if $A \subseteq \mathbb{F}_q^n$, then there is a subspace $H$ of $\mathbb{F}_q^n$ of codimension at most $\tilde{n}_L(\epsilon)$ such that the density of solutions in $A$ to $L(x_1, \ldots, x_k) = 0$ for which $x_{i+1} - x_i \in H$ for $i = 1, \ldots, k - 1$ (this is the same as saying $x_1, \ldots, x_k$ lie in the same translate of $H$) is at least $E[1_A]^k - \epsilon$. The regularity proof shows that if $L = 0$ is translation-invariant, then $\tilde{n}_L(\epsilon)$ exists and is at most tower-type in $\epsilon^{-1}$. If $L = 0$ is translation-invariant but not Sidorenko, must $\tilde{n}_L(\epsilon)$ have tower-type growth?

**Question 3.4.** For each translation-invariant $L = 0$, what is the growth rate of $\tilde{n}_L(\epsilon)$?

**References**

[1] V. Bergelson, B. Host, and B. Kra, Multiple recurrence and nilsequences, *Invent. Math.* **160** (2005), 261–303. With an appendix by Imre Ruzsa.

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

[3] P. J. Cameron, J. Cilleruelo, and O. Serra, On monochromatic solutions of equations in groups, *Rev. Mat. Iberoam.* **23** (2007), 385–395.

[4] D. Conlon, J. Fox, and B. Sudakov, An approximate version of Sidorenko’s Conjecture, *Geom. Funct. Anal.* **20** (2010), 1354–1366.

[5] D. Conlon, J. H. Kim, C. Lee and J. Lee, Some advances on Sidorenko’s Conjecture, *J. London Math. Soc.* **98** (2018), 593–608.

[6] D. Conlon and J. Lee, Finite reflection groups and graph norms, *Adv. Math.* **315** (2017), 130–165.

[7] D. Conlon and J. Lee, Sidorenko’s conjecture for blow-ups, preprint, arXiv:1809.01259.

[8] P. Erdős, On the number of complete subgraphs contained in certain graphs, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **7** (1962), 459–464.

[9] P. Erdős and M. Simonovits, Supersaturated graphs and hypergraphs, *Combinatorica* **3**, (1983), 181–192.

[10] J. Fox and H. T. Pham, Popular progression differences in vector spaces, *Int. Math. Res. Not. IMRN*, to appear.
[11] J. Fox and H. T. Pham, Popular progression differences in vector spaces II, *Discrete Anal.*, to appear.
[12] J. Fox, H. T. Pham, and Y. Zhao, Tower-type bounds for Roth’s theorem with popular differences, preprint.
[13] B. Green, A Szemerédi-type regularity lemma in abelian groups, with applications, *Geom. Funct. Anal.* **15** (2005), 340–376.
[14] B. Green, Montréal Lecture Notes on Quadratic Fourier Analysis. In *Additive Combinatorics* (Montréal 2006, ed. Granville et al.). CRM Proceedings 43, 69?102. Providence: AMS, 2007.
[15] B. Green and T. Tao. An arithmetic regularity lemma, an associated counting lemma, and applications. In *An Irregular Mind (Szemerédi is 70)*, 261–334. Berlin-Heidelberg: Springer, 2010.
[16] H. Hatami. Graph norms and Sidorenko’s conjecture, *Israel J. Math.* **175** (2010), 125–150.
[17] J. H. Kim, C. Lee, and J. Lee, Two approaches to Sidorenko’s conjecture, *Trans. Amer. Math. Soc.* **368** (2016), 5057–5074.
[18] J. X. Li and B. Szegedy, On the logarithmic calculus and Sidorenko’s conjecture, *Combinatorica*, to appear.
[19] A. Saad and J. Wolf, Ramsey multiplicity of linear patterns in certain finite abelian groups, *Q. J. Math.* **68** (2017), 125–140.
[20] A. F. Sidorenko, A correlation inequality for bipartite graphs, *Graphs Combin.* **9** (1993), 201–204.
[21] A. F. Sidorenko, Cycles in graphs and functional inequalities, *Math. Notes* **46** (1989), 877–882.
[22] M. Simonovits, Extremal graph problems, degenerate extremal problems and super-saturated graphs, *Progress in Graph Theory (Waterloo, Ont., 1982)*, Academic Press, Toronto, ON (1984), 419–437.
[23] B. Szegedy, An information theoretic approach to Sidorenko’s conjecture, preprint, arXiv:1406.6738.
[24] A. G. Thomason, A disproof of a conjecture of Erdős in Ramsey theory, *J. London Math. Soc.* **39** (1989), 246–255.

E-mail address: jacobfox@stanford.edu

DEPARTMENT OF MATHEMATICS, STANFORD UNIVERSITY, STANFORD, CA 94305.

E-mail address: huypham@stanford.edu

DEPARTMENT OF MATHEMATICS, STANFORD UNIVERSITY, STANFORD, CA 94305.

E-mail address: yufeiz@mit.edu

DEPARTMENT OF MATHEMATICS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MA 02139.