

# Common and Sidorenko equations in Abelian groups\*

LEO VERSTEEGEN

A linear configuration is said to be common in a finite Abelian group  $G$  if for every 2-coloring of  $G$  the number of monochromatic instances of the configuration is at least as large as for a randomly chosen coloring. Saad and Wolf conjectured that if a configuration is defined as the solution set of a single homogeneous equation in an even number of variables over  $G$ , then it is common in  $\mathbb{F}_p^n$  if and only if the equation's coefficients can be partitioned into pairs that sum to zero mod  $p$ . This was proven by Fox, Pham and Zhao for sufficiently large  $n$ . We generalize their result to all sufficiently large Abelian groups  $G$  for which the equation's coefficients are coprime to  $|G|$ .

AMS 2000 SUBJECT CLASSIFICATIONS: Primary 11B75; secondary 05D10.

KEYWORDS AND PHRASES: Linear configuration, Sidorenko, common, Abelian group.

## 1. Introduction

Let  $G$  be a finite Abelian group. For a pair of integers  $k \leq d$ , we consider matrices  $L \in \mathbb{Z}^{k \times d}$  that have full rank in  $G$ , in the sense that the maps  $G^d \rightarrow G^k$  the matrices induce are surjective. The (*linear*) *configuration given by  $L$* , denoted by  $\mathcal{C}(L)$ , is the kernel of this map in  $G^d$ . We refer to the elements of  $\mathcal{C}(L)$  as the *instances* of the configuration. For a function  $f: G \rightarrow \mathbb{R}$  we define the *arithmetic multiplicity* as<sup>1</sup>

$$t_L(f) = \mathbb{E}_{v \in \mathcal{C}(L)} \prod_{i=1}^d f(v_i) = \frac{1}{|\mathcal{C}(L)|} \sum_{v \in \mathcal{C}(L)} \prod_{i=1}^d f(v_i).$$

---

arXiv: [2109.04445](https://arxiv.org/abs/2109.04445)

\*The author is grateful to be funded by Trinity College of the University of Cambridge through the Trinity External Researcher Studentship.

<sup>1</sup>Here and throughout the paper we write for a set  $X$  and a function  $f: X \rightarrow \mathbb{C}$  the *average of  $f$  over  $X$*  as  $\mathbb{E}(f) = \mathbb{E}_{x \in X} f(x) = \frac{1}{|X|} \sum_{x \in X} f(x)$ .

For a set  $A \subset G$  we denote by  $1_A$  the indicator function of  $A$  and write  $t_L(A)$  to mean  $t_L(1_A)$ . In analogy to graph-theoretical concepts, Saad and Wolf [4] introduced the notions of a configuration being *Sidorenko* or *common* (see the references in the introduction of [4]). We define commonness in a slightly different way here, which we believe is better suited to general finite Abelian groups instead of the specific classes of groups, which were the primary focus of Saad and Wolf.

**Definition 1.1.** We call  $L$  *fully Sidorenko* in  $G$  if for all  $f: G \rightarrow [0, 1]$

$$t_L(f) \geq \mathbb{E}(f)^d.$$

Furthermore,  $L$  is simply called *Sidorenko* in  $G$  if for all  $A \subset G$

$$t_L(A) \geq \left( \frac{|A|}{|G|} \right)^d.$$

We say that  $L$  is *fully common* in  $G$  if for all  $f: G \rightarrow [0, 1]$

$$t_L(f) + t_L(1 - f) \geq \left( \frac{1}{2} \right)^{k-1}.$$

Finally,  $L$  is said to be *common* in  $G$  if for all  $A \subset G$

$$t_L(A) + t_L(A^C) \geq \left( \frac{1}{2} \right)^{k-1}.$$

Clearly, if a configuration is fully common or fully Sidorenko, then it is also common or Sidorenko, respectively. Conversely, if the group  $G$  is sufficiently large and  $L$  is non-degenerate, a standard argument allows us to convert functions  $f: G \rightarrow [0, 1]$  into sets  $A \subset G$  such that  $t_C(f) \approx t_C(1_A)$ .

As a first step towards a general characterization of Sidorenko and common configurations, Saad and Wolf laid out in [4] some initial results, conjectures and questions about necessary and sufficient conditions for a configuration to be Sidorenko or common. Among these was the following.

**Conjecture 1.2.** *Let  $d \geq 2$  even. A linear configuration defined by a single equation in  $d$  variables, i.e., as the kernel of a matrix  $L \in \mathbb{Z}^{1 \times d}$ , is common in  $\mathbb{F}_p^n$  if and only if the set  $[d] := \{1, \dots, d\}$  can be partitioned into pairs such that for each pair  $\{i, j\}$  we have  $L_{1i} \equiv -L_{1j} \pmod{p}$ .*

Saad and Wolf remarked that one direction is easy: from the assumption about the coefficients of  $L$ , it follows by an application of the Cauchy-Schwarz inequality that the configuration is in fact Sidorenko. In [2], Fox, Pham and Zhao went on to show the reverse direction, thus proving the conjecture. In this paper, we generalize these results from  $\mathbb{F}_p^n$  to all finite Abelian groups  $G$  in which the coefficients of  $L$  are coprime to  $|G|$ . First we need to make precise how the condition  $L_{1i} \equiv -L_{1j} \pmod p$  translates to general finite Abelian groups.

**Definition 1.3** (Canceling partitions). A pair of integers  $(m, n)$  is *canceling in  $G$*  if their sum acts trivially on the group, i.e.,  $(m+n)x = 0$  for all  $x \in G$ . For a matrix  $L \in \mathbb{Z}^{k \times d}$ , a *canceling partition in  $G$*  is a partition of  $[d]$  into (distinct) pairs such that for each pair  $\{i, j\}$  and each  $h \in [k]$  the pair  $(L_{hi}, L_{hj})$  is canceling in  $G$ .

The following result settles which single equations are fully common and which are fully Sidorenko under a straightforward condition on the non-degeneracy of the equations.

**Theorem 1.4.** *Let  $G$  be a finite Abelian group and let  $L \in \mathbb{Z}^{1 \times d}$  be such that all coefficients in  $L$  are coprime to the order of  $G$ .*

- (i) *If  $d$  is even and a canceling partition exists for  $L$ , then  $L$  is fully Sidorenko<sup>2</sup>.*
- (ii) *If  $d$  is odd, then  $L$  is fully common in  $G$ .*
- (iii) *If  $d$  is odd, then  $L$  is not fully Sidorenko in  $G$ .*
- (iv) *If  $d$  is even but no canceling partition exists for  $L$ , then  $L$  is not fully common (and hence not fully Sidorenko) in  $G$ .*

Part (i) was observed for  $G = \mathbb{F}_p^n$  by Saad and Wolf and their proof carries over directly. The commonness of a single equation in an odd number of variables, part (ii) of the above theorem, was already proved in [1] by Cameron, Cilleruelo and Serra, even for non-Abelian groups, but we include a proof for completeness. The remaining two statements were proved in [2] for  $G = \mathbb{F}_p^n$ . Our proof resembles that in [2] in so far as we also construct a suitable function  $f$  via its Fourier coefficients. However, in [2], all coefficients are sampled essentially independently at random while we choose all coefficients depending on a single one-dimensional random variable. The advantage of our construction is that the lower bound obtained on the difference  $2^{1-d} - t_L(f) - t_L(1-f)$  does not depend on  $p$  (or, more generally

---

<sup>2</sup>It is easy to see that this holds also for  $k \geq 1$ .

speaking, the smallest size of a non-trivial subgroup). We will expand on this in our concluding remarks.

As a corollary to Theorem 1.4 we obtain the following result in terms of sets rather than functions.

**Corollary 1.5.** *For all  $d \in \mathbb{N}$ , there exists a constant  $C$  such that the following is true. Let  $G$  be a group with  $|G| > C$ , and let  $L \in \mathbb{Z}^{1 \times d}$  be such that all coefficients in  $L$  are coprime to the order of  $G$ . If  $d$  is odd, then  $L$  is not Sidorenko in  $G$ . If  $d$  is even but  $L$  has no canceling partition in  $G$ , then  $L$  is not common in  $G$ .*

Conceivably, the restriction that the coefficients should be coprime to the order of the group could be weakened in some situations. For the groups  $\mathbb{Z}_p$  or  $\mathbb{F}_p^n$ , this would mean that we allow (some) coefficients of the equation to be zero. Doing so will not change whether the equation is Sidorenko, but it might make it uncommon (see [2]). For other finite Abelian groups, the problem is even more delicate, but any formulation of a result seems to require a large number of case distinctions regarding the group decomposition, rendering it of limited interest.

## 2. The proof of Theorem 1.4

Our proof of Theorem 1.4 relies heavily on the Fourier transform. For the convenience of the reader, we recall briefly how it is defined in a discrete setting and state without proof its most important properties.

For a finite Abelian group  $G$ , consider the set  $\widehat{G}$  of *characters*, i.e., group homomorphisms  $\gamma: G \rightarrow \{z \in \mathbb{C} : |z| = 1\}$ . Equipped with pointwise multiplication as a binary operation between characters,  $\widehat{G}$  becomes a group which is isomorphic to  $G$ . For later use, we fix an arbitrary isomorphism  $\mathfrak{D}: G \rightarrow \widehat{G}$ . One reason characters are so useful to us is that they detect the additive identity of a group by the following well-known identity.

$$(2.1) \quad \sum_{\gamma \in \widehat{G}} \gamma(x) = \begin{cases} |G| & \text{if } x = 0, \\ 0 & \text{otherwise.} \end{cases}$$

The other benefit of characters is that we may use them to analyze functions via their Fourier coefficients. Recall that for a function  $f: G \rightarrow \mathbb{C}$  the *Fourier transform of  $f$*  is the function

$$\widehat{f}: \widehat{G} \rightarrow \mathbb{C} \quad \gamma \mapsto \widehat{f}(\gamma) = \mathbb{E}_{x \in G} \gamma(x) f(x).$$

The operator that maps a function to its Fourier transform is linear and a scaled isometry in the sense that for all  $f_1, f_2 \in \text{Map}(G, \mathbb{C})$  it holds that

$$\sum_{\gamma \in \widehat{G}} \widehat{f_1}(\gamma) \overline{\widehat{f_2}(\gamma)} = \mathbb{E}_{x \in G} f_1(x) \overline{f_2(x)}.$$

Furthermore, we have for all  $f$  the inversion formula

$$f(x) = \sum_{\gamma \in \widehat{G}} \widehat{f}(\gamma) \gamma(-x).$$

Real-valued functions  $f$  do not generally have a real-valued Fourier transform. If and only if  $f$  is real, then  $\widehat{\widehat{f}(\gamma^{-1})} = \widehat{f}(\gamma)$  for every  $\gamma \in \widehat{G}$ .

The identity element of  $\widehat{G}$  is the character that always takes the value 1 and it plays a special role with respect to the Fourier transform. Namely, for every function  $f: G \rightarrow \mathbb{C}$  we have that  $\widehat{f}(1) = \mathbb{E}(f)$ .

The Fourier transform allows us to express the number of solutions of an equation in a concise manner.

**Lemma 2.1.** *Let  $L \in \mathbb{Z}^{1 \times d}$  have full rank in  $G$ . Then we have for all  $f: G \rightarrow [0, 1]$*

$$(2.2) \quad t_L(f) = \sum_{\gamma \in \widehat{G}} \prod_{i=1}^d \widehat{f}(\gamma^{L_{1i}}) = \sum_{x \in G} \prod_{i=1}^d \widehat{f}(\mathfrak{D}(L_{1i}x)).$$

*Proof.* Equation (2.1) allows us to express the fact that a vector  $v \in G^d$  is a solution of  $Lv = 0$  via

$$\sum_{\gamma \in \widehat{G}} \gamma(Lv) = \begin{cases} |G| & \text{if } Lv = 0, \\ 0 & \text{otherwise.} \end{cases}$$

Using that  $|\{v \in G^d : Lv = 0\}| = |G|^{d-1}$ , this means we can rewrite the arithmetic multiplicity as

$$\mathbb{E}_{\substack{v \in G^d \\ Lv=0}} \prod_{i=1}^d f(v_i) = \mathbb{E}_{v \in G^d} \left( \prod_{i=1}^d f(v_i) \right) \left( \sum_{\gamma \in \widehat{G}} \gamma(Lv) \right) = \sum_{\gamma \in \widehat{G}} \mathbb{E}_{v \in G^d} \prod_{i=1}^d f(v_i) \gamma^{L_{1i}}(v_i).$$

If we split the expectation into the different coordinates, we see that by definition of the Fourier transform this is the same as

$$\sum_{\gamma \in \widehat{G}} \prod_{i=1}^d \widehat{f}(\gamma^{L_{1i}}),$$

as desired.  $\square$

We are now ready to prove Theorem 1.4, which we restate here for the convenience of the reader.

**Theorem 1.4.** *Let  $G$  be a finite Abelian group and let  $L \in \mathbb{Z}^{1 \times d}$  be such that all coefficients in  $L$  are coprime to the order of  $G$ .*

- (i) *If  $d$  is even and a canceling partition exists for  $L$ , then  $L$  is fully Sidorenko<sup>3</sup>.*
- (ii) *If  $d$  is odd, then  $L$  is fully common in  $G$ .*
- (iii) *If  $d$  is odd, then  $L$  is not fully Sidorenko in  $G$ .*
- (iv) *If  $d$  is even but no canceling partition exists for  $L$ , then  $L$  is not fully common (and hence not fully Sidorenko) in  $G$ .*

*Proof.* It is easy to see that systems consisting of only one equation have full rank if at least one coefficient of  $L$  is coprime to  $|G|$ . This is the case here, so  $L$  has full rank and we may apply Lemma 2.1 to see

$$(2.3) \quad t_L(f) = (\mathbb{E}(f))^d + \sum_{v \in G \setminus \{0\}} \prod_{i=1}^d \widehat{f}(\mathfrak{D}(L_{1i}x)).$$

Throughout the proof, we shall refer to the sum over the non-zero values of  $v$  on the right-hand side as the *deviation* for  $f$ . If a canceling partition exists, then the deviation is always positive because  $\widehat{f}(\mathfrak{D}(L_{1i}x)) \widehat{f}(\mathfrak{D}(-L_{1i}x)) = |\widehat{f}(\mathfrak{D}(L_{1i}x))|^2$ , which proves (i). On the other hand, proving that  $L$  is not fully Sidorenko is equivalent to finding  $f$  such that the deviation is negative. Before constructing such a function  $f$ , we will ascertain what is needed to prove or disprove commonness.

By linearity of the Fourier transform, for any  $f: G \rightarrow \mathbb{C}$  we have  $\widehat{1-f} = \widehat{1} - \widehat{f}$ . It is easy to see by (2.1) that for  $x \neq 0$ , we have  $\widehat{1}(\mathfrak{D}(x)) = 0$ . Hence  $\widehat{(1-f)}(\mathfrak{D}(x)) = -\widehat{f}(\mathfrak{D}(x))$ . Also, since all  $L_{1i}$  are coprime to the order of

---

<sup>3</sup>It is easy to see that this holds also for  $k \geq 1$ .

the group, we can argue that  $L_{1i}x$  can only be zero if  $x$  is zero. Hence,  $(\widehat{1-f})(\mathfrak{D}(L_{1i}x)) = -\widehat{f}(\mathfrak{D}(L_{1i}x))$  for  $x \neq 0$ . This gives

$$(2.4) \quad t_L(f) + t_L(1-f) = (\mathbb{E}(f))^d + (1 - \mathbb{E}(f))^d + \sum_{x \neq 0} \left( \prod_{i=1}^d \widehat{f}(\mathfrak{D}(L_{1i}x)) + \prod_{i=1}^d (-\widehat{f}(\mathfrak{D}(L_{1i}x))) \right).$$

Clearly, if  $d$  is odd, then the two products in the sum on the right cancel. The system is thereby shown to be common as the term  $(\mathbb{E}(f))^d + (1 - \mathbb{E}(f))^d$  is minimized by  $\mathbb{E}(f) = \frac{1}{2}$ . This proves (ii). If  $d$  is even, (2.4) becomes

$$(2.5) \quad t_L(f) + t_L(1-f) = (\mathbb{E}(f))^d + (1 - \mathbb{E}(f))^d + 2 \sum_{x \neq 0} \prod_{i=1}^d \widehat{f}(\mathfrak{D}(L_{1i}x)).$$

The sum on the right is of course again the deviation from (2.3). Thus, to prove both our remaining claims it suffices to construct, for a given  $L$  that lacks a canceling partition (which is always the case when  $d$  is odd), a function  $f: G \rightarrow [0, 1]$  with  $\mathbb{E}(f) = \frac{1}{2}$  such that the deviation is negative. We will construct  $f$  as the Fourier-inverse of an appropriate function.

To this end, pick an element  $a \in G$  of maximal order. Observe that the order of any other element in  $G$  divides that of  $a$ . Note that the order of  $a$  cannot be 2, because if all elements had order two then every partition of the coefficients of  $L$  into pairs would be cancelling. The order of  $a$  is also coprime to every coefficient of  $L$ . Therefore, no element of  $U = \{L_{11}a, \dots, L_{1d}a\}$  is its own inverse, in particular  $0 \notin U$ . Let now  $S = \{s_1, \dots, s_r\} \subset U$  be such that for each  $b \in U$  exactly one of  $b$  and  $-b$  is in  $S$ . For each  $\varphi \in \mathbb{R}$ , define a map

$$g_\varphi: G \rightarrow \mathbb{C} \quad y \mapsto \begin{cases} \frac{1}{4r} e\left(\frac{\varphi}{(2d)^j}\right) & \text{if } y = s_j, \\ \frac{1}{4r} e\left(-\frac{\varphi}{(2d)^j}\right) & \text{if } y = -s_j, \\ \frac{1}{2} & \text{if } y = 0, \\ 0 & \text{otherwise,} \end{cases}$$

where we have written  $e(\cdot)$  for  $\exp(2\pi i \cdot)$ . This map is well defined by definition of  $S$  and has support  $U \cup \{0\}$ . We now define  $f_\varphi$  to be the Fourier inverse of  $g_\varphi$ . To be precise, we define

$$f_\varphi: G \rightarrow \mathbb{C} \quad x \mapsto \sum_{\gamma \in \widehat{G}} g_\varphi(\mathfrak{D}^{-1}(\gamma)) \gamma(-x).$$

Because  $g_\varphi(\mathfrak{D}^{-1}(\gamma^{-1})) = \overline{g_\varphi(\mathfrak{D}^{-1}(\gamma))}$  for all  $\gamma \in \widehat{G}$ ,  $f_\varphi$  takes only real values. Moreover, the image of  $f_\varphi$  is contained in  $[0, 1]$ . Indeed, for  $x \in G$ ,

$$\left| f_\varphi(x) - \frac{1}{2} \right| = \left| \sum_{\gamma \neq 1} g_\varphi(\mathfrak{D}^{-1}(\gamma)) \gamma(-x) \right| \leq \sum_{i=1}^r |g_\varphi(s_i)| + |g_\varphi(-s_i)| = \frac{1}{2}.$$

Furthermore, the average of  $f_\varphi$  is  $g_\varphi(0) = \frac{1}{2}$  and  $\widehat{f} = g_\varphi \circ \mathfrak{D}^{-1}$ . We now wish to find  $\varphi$  such that the deviation for  $f_\varphi$  is negative.

Observe first that if a summand  $\prod_{i=1}^d \widehat{f}_\varphi(\mathfrak{D}(L_{1_i}v))$  is non-zero, it will be  $(\frac{1}{4r})^d$  times a complex number on the unit circle. Let  $X \subset G \setminus \{0\}$  be the set of  $v$  for which this is the case. Note that  $X$  is non-empty, as  $g_\varphi$  is designed so that at least  $a \in X$ . Now, fix  $x \in X$ . We have  $L_{11}x = \pm L_{1i}a$  for some  $i \in [d]$ . Therefore, if  $mx = 0$  for some  $m \in \mathbb{Z}$ , we have  $mL_{1i}a = 0$ . But since  $L_{1i}$  is coprime to the order of  $a$ , we get that  $ma = 0$ . Hence  $x$  is also of maximal order in  $G$  and the order of every element in  $G$  divides that of  $x$ .

Each factor  $\widehat{f}(L_{1_i}x)$  is non-zero, so that for each  $i \in [k]$  there are  $j_i \in [r]$  and  $\sigma_i \in \{\pm 1\}$  such that  $L_{1_i}x = \sigma_i \cdot s_{j_i}$ . The summand  $\prod_{i=1}^d \widehat{f}_\varphi(\mathfrak{D}(L_{1_i}x))$  can therefore be written as

$$\prod_{i=1}^d g_\varphi(L_{1_i}x) = (4r)^{-d} e \left( \varphi \sum_{i=1}^d \sigma_i \frac{1}{(2d)^{j_i}} \right) = (4r)^{-d} e \left( \varphi \sum_{j=1}^r \sum_{\substack{i \in [d] \\ j_i=j}} \sigma_i \frac{1}{(2d)^j} \right).$$

We claim that there is an index  $j$  such that

$$|\{i \in [d] : j_i = j, \sigma_i = 1\}| \neq |\{i \in [d] : j_i = j, \sigma_i = -1\}|,$$

for otherwise we could partition  $[d]$  into pairs such that for every pair  $\{i_1, i_2\}$  among them  $j_{i_1} = j_{i_2}$  and  $\sigma_{i_1} = -\sigma_{i_2}$ , implying that  $(L_{1_{i_1}} + L_{1_{i_2}})x = 0$ . But because the order of every element in  $G$  divides the order of  $x$ , we would have  $(L_{1_{i_1}} + L_{1_{i_2}})y = 0$  for all  $y \in G$ , i.e., there would be a partition of  $[d]$  into canceling pairs in  $G$ , which we have assumed not to be the case.

Choose the minimal index  $j$  with this property and denote it by  $j_*$ . We have

$$\left| \sum_{j=1}^r \sum_{\substack{i \in [d] \\ j_i=j}} \sigma_i \frac{1}{(2d)^j} \right| \geq \left| \sum_{\substack{i \in [d] \\ j_i=j_*}} \sigma_i \frac{1}{(2d)^{j_*}} \right| - \sum_{j=j_*+1}^r \sum_{\substack{i \in [d] \\ j_i=j}} \frac{1}{(2d)^j}$$



$$\geq \frac{1}{(2d)^{j^*}} - \frac{1}{2^{j^*+1}d^{j^*}},$$

which is greater than 0. Let

$$c_x := \sum_{i=1}^d \sigma_i \frac{1}{(2d)^{j_i}} = \sum_{j=1}^r \sum_{\substack{i \in [d] \\ j_i=j}} \sigma_i \frac{1}{(2d)^j}.$$

Since for each  $j \in [d]$  the summand on the right-hand side is an integer multiple of  $(2d)^{-r}$ ,  $c_x$  will be a multiple of  $(2d)^{-r}$  also. Because  $c_x \neq 0$ ,  $c_x$  must have modulus at least  $(2d)^{-r}$ . On the other hand, the triangle inequality gives  $|c_x| \leq \frac{1}{2}$ .

Consider now the function  $\psi$  that maps a phase  $\varphi$  to  $(4r)^d$  times the deviation of  $f_\varphi$ , i.e.,

$$\psi: [0, (2d)^r] \rightarrow \mathbb{C} \quad \varphi \mapsto (4r)^d \sum_{x \neq 0} \prod_{i=1}^d g_\varphi(L_{1^i}x) = \sum_{x \in X} e(c_x \varphi).$$

Note that  $|\psi|$  is bounded from above by  $|X|$ . If  $\varphi$  is sufficiently close to the boundaries of the interval  $[0, (2d)^r]$ ,  $\psi(\varphi)$  will have positive real part. To be precise, if  $\varphi \leq \frac{1}{4}$  or  $\varphi \geq (2d)^r - \frac{1}{4}$ , then

$$\operatorname{Re}(\psi(\varphi)) = \sum_{x \in X} \cos(2\pi c_x \varphi) \geq \frac{|X|}{\sqrt{2}}.$$

On the other hand, the average of  $\psi$  is 0 in the sense that

$$\int_0^{(2d)^r} \psi(\varphi) \, d\varphi = \frac{1}{2\pi i} \sum_{x \in X} \frac{1}{c_x} [e(c_x \varphi)]_0^{(2d)^r} = 0.$$

Now let

$$A = \left\{ \varphi \in [0, (2d)^r] : \operatorname{Re}(\psi(\varphi)) \leq -\frac{|X|}{2\sqrt{2}(2d)^r} \right\},$$

$$B = \left\{ \varphi \in [0, (2d)^r] : -\frac{|X|}{2\sqrt{2}(2d)^r} < \operatorname{Re}(\psi(\varphi)) \leq 0 \right\}.$$

Denoting by  $\lambda$  the Lebesgue measure on  $[0, (2d)^r]$  we obtain

$$\begin{aligned} \operatorname{Re} \left( \int_0^{(2d)^r} \psi(\varphi) \, d\varphi \right) &\geq \frac{1}{2} \cdot \frac{|X|}{\sqrt{2}} - \lambda(A) \cdot |X| - \lambda(B) \cdot \frac{|X|}{2\sqrt{2}(2d)^r} \\ &> |X| \left( \frac{1}{2\sqrt{2}} - \lambda(A) - \frac{(2d)^r}{2\sqrt{2}(2d)^r} \right) \\ &= -\lambda(A) \cdot |X|. \end{aligned}$$

Because we had already established that the integral on the left-hand side is 0,  $A$  must have positive measure, meaning that  $A$  is not empty. But by the definitions of  $\psi$  and  $A$ ,  $f_\varphi$  has negative deviation for any  $\varphi \in A$ , completing the proof.  $\square$

Note that the proof in fact yields an explicit upper bound on the minimal deviation in (2.3) over all  $f: G \rightarrow [0, 1]$  with  $\mathbb{E}(f) = 1/2$ . Namely, using that  $r \leq d$ , we have for at least one value of  $\varphi$

$$(2.6) \quad \sum_{x \neq 0} \prod_{i=1}^d \widehat{f_\varphi}(\mathfrak{D}(L_{1^i}x)) \leq -\frac{1}{(4r)^d} \cdot \frac{1}{2\sqrt{2}(2d)^r} \leq -\frac{1}{2^{3d+1}\sqrt{2}d^{2d}}.$$

### 3. Deriving Corollary 1.5 from Theorem 1.4

There are different ways to deduce Corollary 1.5 from Theorem 1.4, but all of them will require that the set

$$\mathcal{C}^\#(L) = \{v \in \mathcal{C}(L) \subset G^d : \exists i, j \in [d] : i \neq j \wedge v_i = v_j\}$$

of *non-injective instances* is not too large. The following lemma, which is inspired by Lemma 2.1 in [3], makes the dependence on the size of  $\mathcal{C}^\#(L)$  very explicit.

**Lemma 3.1** (Conversion of maps to sets). *Let  $G$  be an Abelian group,  $L \in \mathbb{Z}^{k \times d}$  a system of equations and  $f: G \rightarrow [0, 1]$ .*

(i) *There exists a set  $A \subset G$  with  $|A| \geq \mathbb{E}(f) \cdot |G| - 1$  and*

$$t_L(A) \leq t_L(f) + \frac{|\mathcal{C}^\#(L) \cap A^d|}{|\mathcal{C}(L)|}.$$

(ii) *There exists a set  $A \subset G$  with*

$$t_L(A) + t_L(A^C) \leq t_L(f) + t_L(1 - f) + \frac{|\mathcal{C}^\#(L) \cap A^d| + |\mathcal{C}^\#(L) \cap (A^C)^d|}{|\mathcal{C}(L)|}.$$

*Proof.* We prove only (i) since (ii) can be shown by an argument that is almost identical. Let  $\mathcal{C}^{\text{inj}} = \mathcal{C}(L) \setminus \mathcal{C}^\#(L)$ . The map

$$\psi: \text{Map}(G, [0, 1]) \rightarrow \mathbb{R} \quad g \mapsto \sum_{v \in \mathcal{C}^{\text{inj}}} \prod_{i=1}^k g(v_i)$$

attains its minimum on the compact and non-empty set  $\{g \in [0, 1]^G : \mathbb{E}(g) = \mathbb{E}(f)\}$ . Of all the maps  $g$  achieving this minimum, we denote by  $g_0$  one for which also the set  $R = \{x \in G : g(x) \notin \{0, 1\}\}$  is smallest. We show that  $R$  can contain at most one element. Assume to the contrary that there are two different elements  $a, b \in R$ . For  $\eta \in \mathbb{R}$ , consider

$$g_\eta: G \rightarrow \mathbb{R} \quad x \mapsto \begin{cases} g_0(a) + \eta & \text{if } x = a, \\ g_0(b) - \eta & \text{if } x = b, \\ g_0(x) & \text{otherwise.} \end{cases}$$

Clearly,  $\mathbb{E}(g_\eta) = \mathbb{E}(f)$  and because  $a, b \in R$  we have  $\text{im}(g_\eta) \subset [0, 1]$  for small enough  $\eta$ . It is easy to see that  $\eta \mapsto \psi(g_\eta)$  is a quadratic polynomial with constant term  $\psi(g_0)$  and a non-positive quadratic coefficient. By minimality of  $\psi(g_0)$ , this function must then be constant. But then we may choose  $\eta$  so that at least one of  $a$  and  $b$  takes a value in  $\{0, 1\}$  under  $g_\eta$ . This contradicts the minimality of the set  $R$ .

Now let  $A = g_0^{-1}(1)$ . Then  $1_A(x) = g_0(x)$  for  $x \notin R$  and because  $|R| \leq 1$ , we have that  $|A| > \mathbb{E}(f) \cdot |G| - 1$ . On the other hand, the monotonicity of  $\psi$  gives  $\psi(1_A) \leq \psi(g_0)$ , and  $\psi(g_0)$  in turn is less than  $\psi(f)$  by the minimality of  $g_0$  under  $\psi$  among the functions with average  $\mathbb{E}(f)$ . All in all, we get

$$t_L(A) = \frac{1}{|\mathcal{C}(L)|} \left( \psi(1_A) + \sum_{v \in \mathcal{C}^\#(L)} \prod_{i=1}^k 1_A(v_i) \right) \leq t_L(f) + \frac{|\mathcal{C}^\# \cap A^d|}{|\mathcal{C}(L)|}.$$

□

Unfortunately, there exist configurations in which all instances are non-injective. Having said that, the following lemma provides a useful bound on

the size of  $\mathcal{C}^\#(L)$  for all matrices  $L$  consisting of a single row under a rather weak assumption about the coefficients of the equation.

**Lemma 3.2.** *Let  $G$  be any finite Abelian group,  $d \geq 3$  and  $L \in \mathbb{Z}^{1 \times d}$ . If at least three coefficients of  $L$  are coprime to  $|G|$ , then*

$$|\mathcal{C}^\#(L)| \leq \binom{d}{2} \frac{|\mathcal{C}(L)|}{|G|}.$$

*Proof.* The matrix  $L$  describes a surjective linear map. Indeed, if  $L_{1i}$  is coprime to  $|G|$  for some  $i \in [d]$ , then fixing all but the  $i$ th coordinate of  $v \mapsto Lv$  defines an automorphism of  $G$ . This means  $|\mathcal{C}(L)| = |\ker L| = |G|^{d-1}$ . It is now enough to show that for any pair  $\{i, j\} \in [d]^{(2)}$  the number of solutions where  $v_i = v_j$  can be bounded by  $|G|^{d-2}$ . Indeed, for each such pair there is, by assumption, a third index  $m$  such that  $L_{1m}$  is coprime to  $|G|$ . Then, if we choose all coordinates except  $v_j$  and  $v_m$  freely from  $G$  and set  $v_j = v_i$ , there is exactly one  $v_m \in G$  such that  $Lv = 0$ .  $\square$

Since in Theorem 1.4 we assume that all coefficients are coprime to the order of the group, we may apply Lemma 3.2 to deduce Corollary 1.5. We recall its statement here.

**Corollary 1.5.** *For all  $d \in \mathbb{N}$ , there exists a constant  $C$  such that the following is true. Let  $G$  be a group with  $|G| > C$ , and let  $L \in \mathbb{Z}^{1 \times d}$  be such that all coefficients in  $L$  are coprime to the order of  $G$ . If  $d$  is odd, then  $L$  is not Sidorenko in  $G$ . If  $d$  is even but  $L$  has no canceling partition in  $G$ , then  $L$  is not common in  $G$ .*

*Proof.* We denote by  $\Delta$  the modulus of the deviation achieved in (2.6), i.e.,

$$\Delta = \frac{1}{2^{3d+1} \sqrt{2} d^{2d}}.$$

Choose  $C = d^2/\Delta$  and let  $G$  and  $L$  be as in the statement of Corollary 1.5. The proof of Theorem 1.4 yields a function  $f$  with  $\mathbb{E}(f) = 1/2$ ,  $t_L(f) \leq 1/2^d - \Delta$  and  $t_L(1-f) \leq 1/2^d - \Delta$ . Now apply Lemma 3.1 to obtain sets  $A_1, A_2$  such that

$$\begin{aligned} |A_1| &\geq \mathbb{E}(f)|G| - 1 \\ t_L(A_1) &\leq t_L(f) + \frac{|\mathcal{C}^\#(L)|}{|\mathcal{C}(L)|} \end{aligned}$$

$$t_L(A_2) + t_L(A_2^C) \leq t_L(f) + t_L(1-f) + \frac{2|\mathcal{C}^\#(L)|}{|\mathcal{C}(L)|}.$$

Recall Lemma 3.2, which stated that matrices  $L \in \mathbb{Z}^{1 \times d}$  with at least 3 coefficients that are coprime to  $|G|$  satisfy

$$|\mathcal{C}^\#(L)| \leq \binom{d}{2} \frac{|\mathcal{C}(L)|}{|G|}.$$

Since a matrix without a canceling partition must have at least three coefficients and the coefficients are coprime to  $|G|$  by assumption, the lemma applies. We have

$$t_L(A_2) + t_L(A_2^C) \leq \left(\frac{1}{2}\right)^{d-1} - 2\Delta + \frac{d^2}{C} \leq \left(\frac{1}{2}\right)^{d-1} - \Delta,$$

showing that  $L$  is not common. In order to show that for odd  $d$ ,  $L$  is not Sidorenko in  $G$ , we need to deal with the fact that  $|A_1|/|G|$  could be slightly less than  $1/2$ . We apply the mean value theorem to the monomial  $x^d$  to obtain the lower bound

$$\left(\frac{|A_1|}{|G|}\right)^d \geq \left(\frac{1}{2} - \frac{1}{|G|}\right)^d \geq \left(\frac{1}{2}\right)^d - \frac{d}{|G|} \geq \left(\frac{1}{2}\right)^d - \frac{d}{C},$$

whence

$$t_L(A_1) \leq \left(\frac{1}{2}\right)^d - \Delta + \frac{d^2}{2C} \leq \left(\frac{|A_1|}{|G|}\right)^d - \Delta + \frac{d^2 + 2d}{2C}.$$

Since the latter is strictly less than  $(|A_1|/|G|)^d$ , we have shown that  $L$  is not Sidorenko.  $\square$

#### 4. Concluding remarks

We conclude the article with a brief comparison between our proof and that given in [2] for the special case  $G = \mathbb{F}_p^n$ , which can easily be reduced to the case  $G = \mathbb{F}_p$ . There, the witness function  $f$  is also constructed via an inverse Fourier transform. But instead of taking just a few Fourier coefficients to be non-zero, the authors sample for  $n = 1$  every  $\hat{f}(\gamma)$  identically and (essentially) independently for  $\gamma \neq 1$  and precompose with a projection to the first coordinate for larger  $n$ . To ensure that the resulting function  $f$

takes values in  $[0, 1]$ , the values  $\widehat{f}(\gamma)$  have to be scaled appropriately; just as we had to scale by  $(4r)^{-1}$ . The key difference is that their scaling factor depends on  $p$ , which leads to a much smaller deviation for large  $p$ . Although we still arrive at a valid proof of Theorem 1.4, there are two problems with a deviation that approaches zero as the order of the group grows.

Firstly, the conversion of  $f$  into a set we require to deduce Corollary 1.5 involves an error term that might move us back above the threshold  $\mathbb{E}(f)^d$  if the deviation is too small. If the deviation does not decrease too quickly, a careful conversion might still be possible, but the bound from the proof in [2] in combination with Lemma 3.1 would be insufficient.

Secondly, and perhaps more importantly, a vanishing deviation would not be in the spirit of Conjecture 1.2. We have mentioned earlier that our definition of commonness differs slightly from that of Saad and Wolf [4]. By their definition, an equation  $L \in \mathbb{Z}^{1 \times d}$  is common in  $\mathbb{Z}_p$  if

$$\liminf_{p \rightarrow \infty} \min_{A \subset \mathbb{Z}_p} t_L(A) + t_L(A^C) \geq \left(\frac{1}{2}\right)^{d-1}.$$

It is therefore a slightly weaker property to be common in  $\mathbb{Z}_p$  in the sense of [4] than to be common in  $\mathbb{Z}_p$  for all sufficiently large  $p$  in our sense. Conversely, it is a stronger property to be uncommon in the sense of Saad and Wolf. By giving a lower bound on the deviation independent of  $p$  when  $d$  is even and  $L$  has no canceling partition, we have proved this stronger property.

## Acknowledgments

The author wishes to express his deep gratitude to his master thesis supervisor Mathias Schacht for introducing him to the area of additive combinatorics, to Emil Powierski for many hours of enlightening discussion on the topic, to his PhD supervisor Julia Wolf for her generous help in improving this paper and to the anonymous referee for his comments.

## References

- [1] P. Cameron, J. Cilleruelo, and O. Serra. “On monochromatic solutions of equations in groups”. *Rev. Mat. Iberoam.* 23(1) (2007), pp. 385–395. [MR2351139](#)
- [2] J. Fox, H. T. Pham, and Y. Zhao. “Common And Sidorenko Linear Equations”. *Q. J. Math.* (2021). DOI: [10.1093/qmath/haaa068](#). [MR4350147](#)

- [3] C. Reiher. “Counting odd cycles in locally dense graphs”. *J. Combin. Theory, Ser. B* 105 (2014), pp. 1–5. [MR3171777](#)
- [4] A. Saad and J. Wolf. “Ramsey multiplicity of linear patterns in certain finite abelian groups”. *Q. J. Math.* 68 (2016), pp. 125–140. [MR3658286](#)

LEO VERSTEEGEN

CENTRE FOR MATHEMATICAL SCIENCES, WILBERFORCE ROAD

CAMBRIDGE CB3 0WB

UNITED KINGDOM

*E-mail address:* [lvv23@dpms.cam.ac.uk](mailto:lvv23@dpms.cam.ac.uk)

RECEIVED SEPTEMBER 20, 2021