



RESEARCH ARTICLE

Local aspects of the Sidorenko property for linear equations

Daniel Altman

Department of Mathematics, Stanford University, Stanford, USA; E-mail: daniel.h.altman@gmail.com.

Received: 14 September 2024; **Revised:** 20 July 2025; **Accepted:** 20 July 2025

2020 Mathematics Subject Classification: Primary – 11B30

Abstract

A system of linear equations in $\mathbb{F}_p^n$ is *Sidorenko* if any subset of $\mathbb{F}_p^n$ contains at least as many solutions to the system as a random set of the same density, asymptotically as $n \rightarrow \infty$. A system of linear equations is *common* if any two-colouring of $\mathbb{F}_p^n$ yields at least as many monochromatic solutions to the system of equations as a random 2-colouring, asymptotically as $n \rightarrow \infty$. Both classification problems remain wide open despite recent attention.

We show that a certain generic family of systems of two linear equations is not Sidorenko. In fact, we show that systems in this family are not locally Sidorenko, and that systems in this family which do not contain additive tuples are not weakly locally Sidorenko. This endeavour answers a conjecture and question of Kamčev–Liebenau–Morrison. Insofar as methods, we observe that the true complexity of a linear system is not maintained under Fourier inversion; our main novelty is the use of higher-order methods in the frequency space of systems which have complexity one. We also give a shorter proof of the recent result of Kamčev–Liebenau–Morrison and independently Versteegen that any linear system containing a four-term arithmetic progression is uncommon.

1. Introduction

1.1. Context

In graph theory, Sidorenko’s conjecture [Sid93] predicts that if H is a bipartite graph, then among all graphs G with n vertices and fixed average degree, the number of copies of H in G is asymptotically (as $|V(G)| \rightarrow \infty$) minimised when G is random. We will call graphs H which satisfy this property *Sidorenko*. Sidorenko’s conjecture has been resolved for certain classes of bipartite graph (e.g., trees, cycles, complete bipartite graphs [Sid91], see also [CFS10, Hat10, LS11, Sze14, KLL16, CL17, CKLL18, CL21]) though the conjecture in general remains open, and is indeed of great interest. There is a related colouring property which is also of interest: for which graphs H is it true that, among all two-colourings of the edges of K_n , the number of monochromatic copies of H is minimised by a random two-colouring (asymptotically as $n \rightarrow \infty$)? Such graphs H are called *common*. It is not difficult to see that if a graph satisfies the Sidorenko property, then it is common. Goodman showed that K_3 is common [Goo59]. Erdős [Erd62] subsequently conjectured that K_4 is common and Burr–Rosta [BR80] that every graph is common. Sidorenko [Sid89] showed that a triangle with a pendant edge is not common and Thomason [Tho89] showed that in fact K_4 is not common, together disproving both conjectures. In fact, any graph which contains K_4 is uncommon [JvT96]. It is not difficult to see that if H is *not* bipartite then it is not Sidorenko, and so we have a conjectural classification of all graphs. Even a reasonable conjectured classification of common graphs seems out of reach at the moment.

There has been recent interest in obtaining analogous conjectures and results in arithmetic contexts, initiated by Saad and Wolf [SW17]. Which systems of linear forms Ψ have the property that, among all subsets S of $\mathbb{F}_p^n$ of fixed size, the number of solutions to Ψ in S is minimised when S is chosen randomly? Such systems are called Sidorenko. Those systems with the property that random two-colourings minimise the number of monochromatic solutions among all two-colourings are called common. Again, Sidorenko implies common. In the arithmetic setting, neither property has a conjectured classification.

One situation in which a classification is known is when the system of linear forms has codimension one (i.e., its image is described by a single equation). Here, an equation is Sidorenko if [SW17] and only if [FPZ21] its coefficients may be partitioned into pairs which sum to zero, and is common and not Sidorenko if and only if it has an odd number of variables. For systems of higher codimension, one makes the trivial observation that the union of two Sidorenko systems on disjoint variable sets is Sidorenko. Kamčev, Liebenau and Morrison give a nontrivial method by which higher codimension Sidorenko systems can be built from Sidorenko equations; see [KLM23, Theorem 1.3]. There are also a number of results in the negative direction [KLM24, KLM23, SW17, Ver21]. Finally, we note that Král'–Lamaison–Pach have reported on a partial classification of common systems comprising two equations in $\mathbb{F}_2^n$ [KLP22] and have since obtained a near-complete classification of these systems, which is due to appear in forthcoming work (personal communications).

Lovász [Lov11] showed that the graph-theoretic Sidorenko conjecture is locally true: for all bipartite graphs H , there is no suitably small perturbation of the random graph G which increases the number of copies of H in G . He subsequently classified locally Sidorenko graphs: a graph H is locally Sidorenko if and only if it is a forest or has even girth [Lov12, Theorem 16.26] (see also [FW17] to this end). The study of local notions of commonness has also attracted recent interest: [CHL22, HKKV22, Lov12]. One of the purposes of this article is to introduce these local notions to the arithmetic setting. Finally, to further motivate the study of these local properties, we note that [Alt23] uses a local Sidorenko property of a particular linear system as a key ingredient in the main result there.

1.2. Results

After Fox–Pham–Zhao [FPZ21] completed the classification of codimension one systems, the next natural goal ought perhaps to be a classification of codimension two systems.

Codimension one systems are Sidorenko/common if and only if they are locally Sidorenko/common. Therefore, the classification of these properties for codimension one systems is also complete, and local notions are only interesting when the codimension is at least two.

We study two different notions of locality. Loosely speaking, the weaker notion ('weakly locally Sidorenko/common') allows the radius of perturbation to depend on the function by which we perturb, and the stronger notion ('locally Sidorenko/common') asks that the radius is uniform in the choice of function.

One of our main theorems shows that a certain generic family of systems of codimension 2 is not Sidorenko, and (nearly) classifies this family by whether they are weakly locally Sidorenko or locally Sidorenko. We direct the reader to Subsection 2.1 and in particular Definition 2.2 and Definition 2.3 for formal definitions of the weak local Sidorenko property and the local Sidorenko property. We note that Sidorenko implies locally Sidorenko, which in turn implies weakly locally Sidorenko. We note also that in fact all proofs in the literature listed above which show that a particular system is not Sidorenko in fact show the stronger property that it is not locally Sidorenko.

Let p be an odd prime and let $\Psi = (\psi_1, \dots, \psi_t)$ be a system of linear forms, each mapping $\mathbb{F}_p^D$ to $\mathbb{F}_p$, whose image is determined by a $2 \times t$ system of linear equations with coefficients

$$M := \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1t} \\ b_{21} & b_{22} & \cdots & b_{2t} \end{pmatrix}.$$

We will call Ψ *linearly generic* if every 2×2 minor of M has nonzero determinant.¹ An *additive k -tuple* is a tuple $(x_1, x_2, \dots, x_k)$ satisfying the equation $x_1 - x_2 + x_3 - x_4 + \dots + (-1)^{k+1}x_k = 0$. We say that Ψ *contains* an additive k -tuple if a vector of the form $(0, \dots, 0, 1, -1, 1, \dots, (-1)^{k+1})$ lies in the row span of M , up to reordering of columns.

Theorem 1.1. *Let Ψ be a codimension two system of linear equations in t variables which is linearly generic. Then Ψ is not locally Sidorenko (and so not Sidorenko). If t is even then Ψ is not weakly locally Sidorenko. Furthermore, for p sufficiently large in terms of t , if t is odd and Ψ is weakly locally Sidorenko, then Ψ contains an additive $(t-1)$ -tuple.*

Thus, in the situation that p is large in terms of t , we have classified all linearly generic systems by whether they are Sidorenko, locally Sidorenko, or weakly locally Sidorenko, except for those systems with t odd and which contain an additive $(t-1)$ -tuple. We note that the condition of containing a $(t-1)$ -tuple is more than a reflection of the limitation of our methods: there does exist a linearly generic system with t odd, which contains an additive $(t-1)$ -tuple, and which is weakly locally Sidorenko (see Example 4.2).

As a consequence of Theorem 1.1, we answer the following conjecture and question of Kamčev–Liebenau–Morrison.

Conjecture 1.2 [KLM23, Conjecture 5.1]. *Let Ψ be a system of 5 linear forms whose image has codimension 2. If Ψ is linearly generic then Ψ is not Sidorenko.*

Question 1.3 [KLM23, Question 5.3]. *Let $t \geq 7$ be odd. Does there exist a system of t linear forms Ψ whose image has codimension 2, which is linearly generic and which is Sidorenko?*

We note that the above conjecture and question as they appear in [KLM23] do not refer directly to the linearly generic condition, but the formulation there is equivalent. We have the following corollary of Theorem 1.1.

Corollary 1.4. *Conjecture 1.2 is true and the answer to Question 1.3 is no.*

Define the *complexity* of $\Psi = (\psi_1, \dots, \psi_t)$, where each ψ_i is a linear map from $\mathbb{F}_p^D$ to $\mathbb{F}_p$, to be the smallest positive integer s such that the functions $(\psi_i^s)_{i=1}^t$ are linearly independent over $\mathbb{F}_p$. Loosely speaking, if the complexity of a system is 1, then its analysis should be controlled by Fourier analysis, and one ought not require the higher-order theory. We direct the reader to [GW10, GW11a, GW11b], for relevant definitions, discussions and results to this end. One may easily compute that for $t \geq 5$ and p sufficiently large, systems whose images have codimension two generically have complexity one. Where it is not difficult, using higher-order constructions, to contrive systems of complexity at least 2 which are weakly locally Sidorenko but not Sidorenko, the situation is less clear for complexity one systems (we will justify this claim in Subsection 4.2). We believe therefore that the following corollary of Theorem 1.1 is also worth noting.

Corollary 1.5. *Among complexity one systems, the weak local Sidorenko property is not equivalent to the local Sidorenko property.*

The explicit example demonstrating Corollary 1.5 is Example 4.2 in Section 4.

Of independent interest is the observation and method underlying Theorem 1.1 and its corollaries. We observe that although these systems have complexity one, their ‘dual’ system (i.e., the linear system obtained in frequency space after Fourier inversion) need not have complexity one. The most involved part of Theorem 1.1 utilises higher-order constructions in the frequency space (Theorem 3.5). More broadly, the proof of Theorem 1.1 also utilises other combinatorics on the dual pattern of a linear system (Theorem 3.1). Since the uploading of this document to the arXiv, the use of combinatorics on the dual

¹We note here that in [KLM23] the same class of systems is considered, albeit in different language. A reader wishing to compare the texts may find it helpful to note that our linearly generic condition is equivalent to the condition that a system has codimension two and (in the notation of [KLM23]) $s(L) = t - 1$.

pattern of a linear system has been called a *Fourier template* by [DLZ24], and is used in both [AL25] and [DLZ24].

We also provide the first example of a system which is locally Sidorenko but not Sidorenko; see Example 4.1.

Proposition 1.6. *In the arithmetic setting, the Sidorenko property is not equivalent to the local Sidorenko property.*

This result is perhaps slightly less trivial than it sounds; for example, in the graph-theoretic setting, it remains an open question whether local commonness (appropriately defined, see [CHL22]) is equivalent to commonness. One of the reasons we have chosen to study these local properties is that all existing proofs in the literature that systems are not Sidorenko [FPZ21, KLM23, SW17, Ver21] prove the stronger statement that in fact these systems are not locally Sidorenko.

Finally, we note that the linear genericity condition in Theorem 1.1 is necessary. For example, it has been observed by Kamčev–Liebenau–Morrison [KLM23] that the system

$$\begin{aligned}x_1 + 2x_2 - x_3 - 2x_4 &= 0 \\ 3x_1 + x_2 - 3x_3 - x_4 &= 0\end{aligned}$$

is Sidorenko, and it clearly does not contain an additive tuple.

In another direction, it has been known for some time [JvT96, Theorem 12] that any graph containing K_4 is uncommon. Saad and Wolf [SW17] asked whether the same is true for four-term arithmetic progressions in the arithmetic context. This question was answered in the affirmative recently by Kamčev–Liebenau–Morrison [KLM24] and independently by Versteegen [Ver21]. Both proofs build on a ‘quadratic’ construction of Gowers [Gow20]. In Section 5, we give (together with a lemma from [KLM24]) a short proof of Saad and Wolf’s conjecture which does not rely on higher-order constructions. Our proof is also more amenable to generalisation to rank two systems in $2k$ variables, which would be a significant step in the classification of weakly locally common systems². In fact, since this document has been on the arXiv, both the author and Liebenau [AL25] and Dong, Li and Zhao [DLZ24] have used the method introduced in Section 5 below to achieve this generalisation.

We set up some notational and technical preliminaries in Section 2. We prove Theorem 1.1 in Section 3. In Section 4 we give examples to prove Corollary 1.5 and Proposition 1.6. Finally in Section 5 we give a new proof of the answer to Saad and Wolf’s question [SW17, Question 3.6].

2. Preliminaries

2.1. Definitions, basic properties

Throughout we let $\Psi = (\psi_1, \dots, \psi_t)$ be a system of $\mathbb{F}_p$ -linear forms, each mapping $\mathbb{F}_p^D$ to $\mathbb{F}_p$; we often abuse notation and let them induce linear maps from $(\mathbb{F}_p^n)^D$ to $\mathbb{F}_p^n$ in the natural way. Define the operator T_Ψ acting on functions $f : \mathbb{F}_p^n \rightarrow \mathbb{C}$ by the formula

$$T_\Psi(f) := \mathbb{E}_{x_1, \dots, x_D \in \mathbb{F}_p^n} f(\psi_1(x_1, \dots, x_D)) \cdots f(\psi_t(x_1, \dots, x_D)),$$

where throughout $\mathbb{E}$ is a normalised count, so here it is a shorthand for $\frac{1}{p^{nD}} \sum$. We will always assume that $D \leq t$ and that $\Psi^{-1}(0) = 0$ (it is clear that one may reduce to this case). At times it will be more convenient to describe a system of linear forms by the equations which define its image. Suppose that M_Ψ is an $\mathbb{F}_p$ -valued matrix such that $\text{Im } \Psi = \ker M_\Psi$. Then we have

²We do not consider weak local commonness in this document and so we do not give this concept a formal definition here. However, we hope that the enthusiastic reader may formulate a definition from that of the weak local Sidorenko property Definition 2.3 and the common property Definition 2.4.

$$T_{\Psi}(f) = \mathbb{E}_{\mathbf{x} \in \ker M_{\Psi}} f(x_1) \cdots f(x_t),$$

where $\mathbf{x} = (x_1, \dots, x_t) \in (\mathbb{F}_p^n)^t$.

Throughout this document we will deal with functional versions of the Sidorenko and common properties in which we ask the following of all functions $f : \mathbb{F}_p^n \rightarrow [0, 1]$, rather than just restricting to characteristic functions of sets.

Definition 2.1. A system of t linear forms Ψ is *Sidorenko* if, for all $n \geq 1$ and all $f : \mathbb{F}_p^n \rightarrow [0, 1]$, we have

$$T_{\Psi}(f) \geq (\mathbb{E}_{x \in \mathbb{F}_p^n} f(x))^t.$$

Note that any function can be written as the sum of a constant function and a function with zero average, whence the inequality in Definition 2.1 can be written as $T_{\Psi}(\alpha + f) \geq \alpha^t$, where the condition must hold for all $\alpha \in [0, 1]$ and all $f : \mathbb{F}_p^n \rightarrow [-\alpha, 1 - \alpha]$ with $\mathbb{E}_x f(x) = 0$. A system is locally Sidorenko if the above inequality is satisfied for all sufficiently small perturbations of the constant function.

Definition 2.2. A system of t linear forms Ψ is *locally Sidorenko* if for all $\alpha \in [0, 1]$ there exists $\varepsilon > 0$ such that for all $n \geq 1$ and all $f : \mathbb{F}_p^n \rightarrow [-\alpha, 1 - \alpha]$ with $\mathbb{E}_x f(x) = 0$, we have

$$T_{\Psi}(\alpha + \varepsilon f) \geq \alpha^t.$$

We remark that that this notion of locality considers small perturbations of the constant function with respect to the ℓ^∞ norm on the space of functions. One may also consider analogous notions of locality for different norms. We do not do so in this document, but note that, for example, considering the local Sidorenko property with respect to the ℓ^∞ norm on the Fourier side was an important part of the main result of [Alt23].

Above, the size of the neighbourhood of the constant function is fixed independently of n, f . If the size of the neighbourhood is allowed to vary with these parameters, then we will say that the linear system is weakly locally Sidorenko.

Definition 2.3. A system of t linear forms Ψ is *weakly locally Sidorenko* if for all $n \geq 1$, for all $\alpha \in [0, 1]$ and all $f : \mathbb{F}_p^n \rightarrow [-\alpha, 1 - \alpha]$ with $\mathbb{E}_x f(x) = 0$, there exists $\varepsilon_f > 0$ such that for all $\varepsilon < \varepsilon_f$ we have

$$T_{\Psi}(\alpha + \varepsilon f) \geq \alpha^t.$$

Definition 2.4. A system of t linear forms Ψ is *common* if, for all $n \geq 1$ and all $f : \mathbb{F}_p^n \rightarrow [0, 1]$, we have

$$T_{\Psi}(f) + T_{\Psi}(1 - f) \geq 2^{1-t}.$$

Local commonness and weak local commonness may be defined analogously.

Let $\Psi = (\psi_1, \dots, \psi_t)$ be a system of t linear forms. Let α be a constant and f a function on $\mathbb{F}_p^n$ with $\mathbb{E}_x f(x) = 0$. For a subset $S \subset [t]$, we denote the corresponding subsystem of Ψ by $\Psi(S)$ (i.e., the restriction of Ψ to the coordinates in S). Then, by the multilinearity of the operator T we have

$$T_{\Psi}(\alpha + f) = \sum_{S \subset [t]} \alpha^{t-|S|} T_{\Psi(S)}(f). \quad (1)$$

Note that if f has $\mathbb{E}_x f(x) = 0$ and if the image of $\Psi(S)$ contains a coordinate hyperplane then $T_{\Psi(S)}(f) = 0$. It follows from the linear genericity condition (recall: every 2×2 minor of the matrix defining its equations has nonzero determinant), then for all nonempty $S \subset [t]$ with $|S| \leq t - 2$, the subsystem $\Psi(S)$ contains a coordinate hyperplane and so yields $T_{\Psi(S)}(f) = 0$. Thus for linearly generic systems we have

$$T_{\Psi}(\alpha + f) = \alpha^t + \alpha \sum_{|S|=t-1} T_{\Psi(S)}(f) + T_{\Psi}(f). \quad (2)$$

Furthermore, for such systems, if $|S| = t - 1$ then the image of $\Psi(S)$ has codimension one (i.e., its image is described by a single equation).

2.2. Fourier inversion

The reader is directed to [Zha23, Chapter 6.1] for further background on the basics of Fourier analysis in $\mathbb{F}_p^n$.

Recall the Fourier transform defined by

$$\hat{f}(h) = \mathbb{E}_{x \in \mathbb{F}_p^n} f(x) e_p(-h \cdot x),$$

where $e_p(\cdot) := e^{\frac{2\pi i \cdot}{p}}$. Recall also the Fourier inversion formula:

$$f(x) = \sum_{h \in \mathbb{F}_p^n} \hat{f}(h) e_p(h \cdot x),$$

and Parseval's identity

$$\mathbb{E}_x |f(x)|^2 = \sum_h |\hat{f}(h)|^2.$$

For a linear system Ψ , let $M = M_{\Psi}$ have $\text{Im } \Psi = \ker M$. We will make use of the following consequence of the Fourier inversion formula:

$$T_{\Psi}(f) = \sum_{\mathbf{h} \in \text{Im } M^T} \hat{f}(h_1) \cdots \hat{f}(h_t), \quad (3)$$

where $\mathbf{h} = (h_1, \dots, h_t) \in (\mathbb{F}_p^n)^t$.

2.3. Tensoring

Definition 2.5. Let $f_1 : \mathbb{F}_p^{n_1} \rightarrow \mathbb{C}$ and $f_2 : \mathbb{F}_p^{n_2} \rightarrow \mathbb{C}$. Then the tensor product $f := f_1 \otimes f_2 : \mathbb{F}_p^{n_1+n_2} \rightarrow \mathbb{C}$ is the function defined by $f(x) = f_1(\pi_1(x))f_2(\pi_2(x))$, where π_1 is the projection onto the first n_1 coordinates of $\mathbb{F}_p^{n_1+n_2}$ and π_2 is the projection onto the last n_2 coordinates.

It is clear that $\otimes$ is an associative operation and so one may naturally define the tensor product of a finite collection of functions.

Observation 2.6. For a system of linear forms Ψ , we have

$$T_{\Psi}(f_1 \otimes f_2) = T_{\Psi}(f_1)T_{\Psi}(f_2). \quad (4)$$

Note that there is a slight abuse of notation here: we interpret T_{Ψ} as an operator on $\mathbb{F}_p^{n_1+n_2}$, $\mathbb{F}_p^{n_1}$ and $\mathbb{F}_p^{n_2}$ as it appears respectively. The equation above is easily verified by direct computation.

3. Proof of Theorem 1.1

Recall that we call a system of two equations linearly generic if the $2 \times t$ matrix

$$M := \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1t} \\ b_{21} & b_{22} & \cdots & b_{2t} \end{pmatrix}$$

which defines the equations has the property that all of its 2×2 minors have nonzero determinant.

3.1. Systems without additive tuples are not weakly locally Sidorenko

We first work towards proving the following theorem, which is one of the statements in Theorem 1.1.

Theorem 3.1. *Let Ψ be a system of two linear equations in t variables which is linearly generic. If t is even or if Ψ does not contain an additive tuple and p is sufficiently large in terms of t , then Ψ is not weakly locally Sidorenko.*

We now prove the theorem, modulo the proof of two propositions which will appear subsequently and occupy the remainder of the subsection. In both propositions, we prove something more general than what is needed for the application in this document. We do so to illustrate the scope of the constructions and because it is clear that these more general statements can be used to obtain more general results. See also Remark 3.4.

Proof of Theorem 3.1. Recall from Equation (2) that we have

$$T_{\Psi}(\alpha + \varepsilon f) = \alpha^t + \alpha \varepsilon^{t-1} \sum_{|S|=t-1} T_{\Psi(S)}(f) + \varepsilon^t T_{\Psi}(f).$$

It suffices to show that there exists f with $\mathbb{E}f = 0$ such that $\sum_{|S|=t-1} T_{\Psi(S)}(f) < 0$; then one makes ε appropriately small to conclude. If $t-1$ is odd, then this is trivial: take f for which $\sum_{|S|=t-1} T_{\Psi(S)}(f) \neq 0$ (it is not difficult to see that such an f must exist), and if $\sum_{|S|=t-1} T_{\Psi(S)}(f) > 0$ then replace f with $-f$. Henceforth we assume that t is odd.

Our strategy will be to find, for each $S \subset [t]$ with $|S| = t-1$, a function f_S with $\mathbb{E}f_S = 0$ such that $T_{\Psi(S)}(f_S) = 0$ and $T_{\Psi}(f_S) \neq 0$. Then we let $f = \bigotimes_{|S|=t-1} f_S$ and by (4) we see that

$$T_{\Psi}(\alpha + \varepsilon f) = \alpha^t + \varepsilon^t T_{\Psi}(f),$$

where $T_{\Psi}(f) = \prod_{|S|=t-1} T_{\Psi}(f_S) \neq 0$. Recall that t is odd so potentially replacing f with $-f$ depending on the sign of $T_{\Psi}(f)$, we may conclude that for all ε sufficiently small (indeed, for all ε), we have $T_{\Psi}(\alpha + \varepsilon f) < \alpha^t$, so Ψ is not weakly locally Sidorenko.

It remains to prove the existence of the functions f_S . Recall that each of the systems $\Psi(S)$ has codimension one, so is described by a single equation. This equation cannot be of the form $\sum_i (-1)^i x_i = 0$ by our assumption that Ψ does not contain an additive tuple. We prove the existence of such an f_S in Proposition 3.2 and Proposition 3.3. The first deals with the easier case in which there is a pair of coefficients whose ratio is not equal to ± 1 , and the second deals with the remaining cases which are not additive tuples. $\square$

In the following and throughout, we use Vinogradov notation $\ll$ and $\gg$ to conceal a multiplicative constant. The appearance of parameters in the subscript denotes that this constant may depend on these parameters.

Proposition 3.2. *Let Ψ be a codimension 2 linear system in t variables which is linearly generic. Let Φ be determined by a single equation $\sum_{i=1}^l a_i x_i = 0$, where each a_i is nonzero, such that there are i, j satisfying $a_i \neq \pm a_j$. Then for all $n \geq 2$, there exists $f : \mathbb{F}_p^n \rightarrow [-1, 1]$ with $\mathbb{E}_x f(x) = 0$, $T_{\Phi}(f) = 0$ and $|T_{\Psi}(f)| \gg_t 1$.*

Proof. Without loss of generality a_1, a_2 are such that $a_1 \neq \pm a_2$. We construct f from its Fourier coefficients. Let $M = M_{\Psi} = (b_{ij})$ be the $2 \times t$ matrix whose coefficients correspond to the two linear equations of Ψ . Using (3) we have that

$$T_{\Phi}(f) = \sum_{h \in \mathbb{F}_p^n} \prod_{i=1}^l \hat{f}(a_i h), \quad (5)$$

and

$$T_{\Psi}(f) = \sum_{h_1, h_2 \in \mathbb{F}_p^n} \prod_{i=1}^t \hat{f}(b_{1i}h_1 + b_{2i}h_2). \tag{6}$$

We claim that if h_1, h_2 are linearly independent, then $(u_1, \dots, u_t) := M^T \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}$ satisfies $u_i \neq 0$ for all i and $a_1 u_i \neq \pm a_2 u_j$ for all i, j . Indeed, that $u_i \neq 0$ follows from linear independence and the fact that M cannot have a zero column, and if $a_1 u_i = \pm a_2 u_j$, then by the linear independence of h_1, h_2 we have $a_1 \begin{pmatrix} b_{1i} \\ b_{2i} \end{pmatrix} = \pm a_2 \begin{pmatrix} b_{1j} \\ b_{2j} \end{pmatrix}$ which contradicts the fact that all 2×2 minors of M are full rank.

Choose any two linearly independent h_1, h_2 and define $u_1, \dots, u_t$ as above. Now define $\hat{f}(\pm u_i) = 1/(2t)$ for all i and $\hat{f}(h) = 0$ otherwise so that f is clearly real-valued and takes values in $[-1, 1]$ by Fourier inversion and the triangle inequality. Since $u_i \neq 0$ for all i we have $\hat{f}(0) = \mathbb{E}_x f(x) = 0$. Next, since $a_1 \neq \pm a_2$ and since $a_1 u_i \neq \pm a_2 u_j$ for all i, j , we have that for all h at least one of $\hat{f}(a_1 h), \hat{f}(a_2 h)$ is zero and so $T_{\Phi}(f) = 0$ from (5). Finally, $T_{\Psi}(f) \geq 2/(2t)^t$ from (6). $\square$

In the following proposition we prove a more general statement than what is needed for the proof of Theorem 3.1. Below we consider equations of arbitrary length with ± 1 coefficients. For the theorem above, we just need to consider equations $\sum_{i=1}^{t-1} a_i x_i = 0$ of length $t - 1$ (which we recall may be assumed to be even), where $a_i \in \pm 1$. In this case, since $t - 1$ is even, we have for parity reasons that $|\#\{a_i = 1\} - \#\{a_i = -1\}|$ must be even, and so the condition that $t \neq (2m + 1)|\#\{a_i = 1\} - \#\{a_i = -1\}|$ for all $m \in \mathbb{Z}$ from the statement of the proposition below is automatically satisfied.

Proposition 3.3. *Let Ψ be a codimension 2 linear system in t variables which is linearly generic. Let Φ be determined by the equation $\sum_{i=1}^{t'} a_i x_i = 0$, where each $a_i \in \{-1, 1\}$ and $\#\{a_i = 1\} \neq \#\{a_i = -1\}$. If p is sufficiently large in terms of t and $t \neq (2m + 1)|\#\{a_i = 1\} - \#\{a_i = -1\}|$ for any integer m , then there exists $f : \mathbb{F}_p \rightarrow [-1, 1]$ with $\mathbb{E}_x f(x) = 0$, with $T_{\Phi}(f) = 0$ and with $|T_{\Psi}(f)| \gg_{p,t} 1$.*

Proof. Let $V = \text{Im } M_{\Psi}^T \leq \mathbb{F}_p^t$. For $i = 1, \dots, t$, let P_i be the i th coordinate hyperplane (i.e., $\{x \in \mathbb{F}_p^t : x_i = 0\}$). Let the symmetric group on t elements $\text{Sym}(t)$ act on $\mathbb{F}_p^t$ by coordinate permutation, and let $(\mathbb{Z}/2\mathbb{Z})^t$ act on $\mathbb{F}_p^t$ by reflection in each coordinate hyperplane (so, e.g., $(1, 0, 1, 0, \dots) \in (\mathbb{Z}/2\mathbb{Z})^t$ maps $(u_1, u_2, u_3, u_4, \dots)$ to $(-u_1, u_2, -u_3, u_4, \dots)$). Let G be the group of permutations of $\mathbb{F}_p^t$ generated by $\text{Sym}(t)$ and $(\mathbb{Z}/2\mathbb{Z})^t$ and let $S = G - \{\text{id}, (1, 1, \dots, 1)\}$.

We claim that $V - (\bigcup_{i=1}^t P_i \cup \bigcup_{\sigma \in S} \sigma V)$ is nonempty. Indeed,

$$\left| V - \left(\bigcup_{i=1}^t P_i \cup \bigcup_{\sigma \in S} \sigma V \right) \right| \geq |V| - \sum_{i=1}^t |V \cap P_i| - \sum_{\sigma \in S} |V \cap \sigma V|,$$

where each $V \cap P_i$ and $V \cap \sigma V$ is a subspace of V , so has size 1, p or p^2 . By taking p sufficiently large in terms of t , it suffices to show that $V \neq V \cap P_i$ and $V \neq \sigma V$ for each $i \in [t]$ and $\sigma \in S$. That $V \neq V \cap P_i$ follows from the fact that M cannot have a zero column, and that $V \neq \sigma V$ for every $\sigma \in S$ is an easy exercise in linear algebra, using the fact that each 2×2 minor of M has full rank and that every element of G may be written in the form $\sigma \tau$ for some $\sigma \in \text{Sym}(t)$ and $\tau \in (\mathbb{Z}/2\mathbb{Z})^t$.

Let $u := (u_1, \dots, u_t) \in V - (\bigcup_{i=1}^t P_i \cup \bigcup_{\sigma \in S} \sigma V)$. We may assume that $\#\{a_i = 1\} > \#\{a_i = -1\}$ so that by Fourier inverting

$$T_{\Phi}(f) = \sum_{h \in \mathbb{F}_p^n} |\hat{f}(h)|^{2k} \hat{f}(h)^l, \tag{7}$$

where $t' = 2k + l$, $k \geq 0$, $l = \#\{a_i = 1\} - \#\{a_i = -1\} > 0$. We have by construction that each $\pm u_i$ is distinct and nonzero. Let $\frac{1}{2} \leq c_1, \dots, c_t \leq 1$ be real numbers to be chosen later and define

$\hat{f}(u_i) := c_i e^{\frac{\pi i}{2l}} / (2t)$ and $\hat{f}(-u_i) = \overline{\hat{f}(u_i)}$ for all i . Set $\hat{f}(h) = 0$ for all other h . Then $\mathbb{E}_x f(x) = \hat{f}(0) = 0$, f takes values in $[-1, 1]$ and $\Re(\hat{f}(\pm u_i)^l) = 0$ for all i . Thus we have

$$T_\Phi(f) = \Re(T_\Phi(f)) = \sum_{h \in \mathbb{F}_p^n} |\hat{f}(h)|^{2k} \Re(\hat{f}(h)^l) = 0.$$

It remains to argue that $|T_\Psi(f)| \gg_{p,t} 1$. Fourier inverting we have

$$T_\Psi(f) = \sum_{y_1, y_2 \in \mathbb{F}_p} \prod_{i=1}^t \hat{f}(b_{1i}y_1 + b_{2i}y_2).$$

Let $U = \{\pm u_1, \dots, \pm u_t\}$ and let $\mathcal{I}$ be the collection of multisubsets of $[t]$ which have t elements. For a multisubset W of U with t elements, let $i(W) \in \mathcal{I}$ be the multisubset of indices of W (e.g., if $t = 4$ we may have $i([u_1, u_1, u_3, -u_1]) = [1, 1, 3, 1]$), and let $\Sigma(W)$ be the sum of the signs in W (so $\Sigma([u_1, u_1, u_3, -u_1]) = 1 + 1 + 1 - 1 = 2$). Then

$$T_\Psi(f) = \sum_{I \in \mathcal{I}} \left(\sum_{W: i(W)=I} \# \left\{ y_1, y_2 : M^T \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = W \right\} \cdot \frac{1}{(2t)^t} e^{\frac{\Sigma(W)\pi i}{2l}} \right) \prod_{i \in I} c_i, \quad (8)$$

where we abuse notation by implicitly ‘unordering’ the vector $M^T \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ and identifying it with the corresponding multiset. Viewing (8) as a polynomial in $c_1, \dots, c_t$, we may choose $c_1, \dots, c_t$ in such a way as to force $|T_\Psi(f)| \gg_{p,t} 1$ unless (8) is the zero polynomial. Note that when $I = [t]$, we have by our construction of $u = (u_1, \dots, u_t)$ that the only multisubsets W for which $\#\{y_1, y_2 : M^T \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = W\}$ is nonzero are $W = [u_1, u_2, \dots, u_t]$ and $W = [-u_1, -u_2, \dots, -u_t]$, and in these cases $\#\{y_1, y_2 : M^T \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = W\} = 1$. Furthermore, for these W we have that $\Sigma(W) = t$ and $-t$ respectively. Therefore, the coefficient of the term $\prod_{i=1}^t c_i$ in (8) is $\frac{2}{(2t)^t} \Re \left(e^{\frac{t\pi i}{2l}} \right)$, which is nonzero since t is not of the form $(2m+1)l$ for any integer m . Thus we may choose $c_1, \dots, c_t$ so that $|T_\Psi(f)| \gg_{p,t} 1$, completing the proof. $\square$

Remark 3.4. It is not difficult to adapt the construction from the previous proposition to relax the assumption that $t \neq (2m+1)l$ to the assumption that $l \neq 1$ or t is even; one proceeds by defining $\hat{f}(u_i) := c_i e^{\frac{(2n_i+1)\pi i}{2l}}$ for an appropriate choice of integers $\{n_i\}_{i=1}^t$. In fact, it seems likely that a modest perturbation of the above argument may allow us to remove any assumption of this kind completely. We do not pursue the matter in this document as we have no immediate need for the more general statement. We hope that the above example is illustrative.

3.2. Systems with additive tuples are not locally Sidorenko

Together with Theorem 3.1, the following completes the proof of Theorem 1.1.

Theorem 3.5. *Let Ψ be a system of two linear equations which is linearly generic. Then Ψ is not locally Sidorenko.*

The main observation underpinning the upcoming construction is that the complexity of a system of linear forms is not maintained under Fourier inversion. We have not investigated the extent to which higher-order Fourier analysis may be applied in frequency space, but the upcoming construction (and natural generalisations thereof) at least demonstrates one application of the aforementioned observation.

In light of Theorem 3.1 we may assume that t is odd. We note that if $\Psi(S)$ corresponds to an additive $(t-1)$ -tuple, then by Fourier inversion, $T_{\Psi(S)}$ is the $((t-1)$ th power of the) ℓ^{t-1} norm in frequency space. In particular, if $T_{\Psi(S)}(f) = 0$ then $f = 0$, so the proof strategy from the previous subsection fails. We will have to settle for a function f for which the ratio $T_{\Psi(S)}(f)/T_\Psi(f)$ is arbitrarily large in absolute value. The existence of such a function is proven in Proposition 3.7. First, we need a lemma.

Lemma 3.6. *The following facts hold:*

1. Define the quadratic $q : \mathbb{F}_p^n \rightarrow \mathbb{F}_p$ by $q(x) = x^T Mx + h^T x$, where M is a matrix of rank r and $h \in \mathbb{F}_p^n$. Then

$$|\mathbb{E}_{x \in \mathbb{F}_p^n} e_p(q(x))| \leq p^{-r/2}.$$

2. Let Ψ be a system of linear forms, each mapping $(\mathbb{F}_p^n)^D$ to $\mathbb{F}_p^n$. Let $A \subset \mathbb{F}_p^n$ be the zero set of the quadratic defined by $x^T x$. Then

$$|T_\Psi(1_A) - p^{-k}| \leq p^{-n/2},$$

where k is the dimension of the $\mathbb{F}_p$ -span of the functions $\{\psi_i(x)^T \psi_i(x)\}_{i=1}^t$.

Sketch proof. Part 1 is a very standard Gauss sum estimate; a proof may be found in [Gre07], for example. In [GW10, Proof of Theorem 3.1] it is shown (using Part 1 as the main ingredient) that if a system of linear forms $\tilde{\Psi} := (\tilde{\psi}_i)_{i=1}^l$ has $\{\tilde{\psi}_i(x)^T \tilde{\psi}_i(x)\}_{i \in [l]}$ linearly independent, then $|T_{\tilde{\Psi}}(1_A) - p^{-l}| \leq p^{-n/2}$. Let $S \subset [t]$ be a set such that $\{\psi_i(x)^T \psi_i(x)\}_{i \in S}$ is a basis for $\{\psi_i(x)^T \psi_i(x)\}_{i=1}^t$. Then we observe the equality $\prod_{i \in S} 1_A(\psi_i(x)) = \prod_{i \in [t]} 1_A(\psi_i(x))$ for all x , so $T_\Psi(1_A) = T_{\Psi(S)}(1_A)$ and our result then follows from the second sentence of this proof.

For a $2 \times k$ matrix M with values in $\mathbb{F}_p$, let $M^{(2)}$ be the $3 \times k$ matrix defined by

$$M^{(2)} := \begin{pmatrix} \mathbf{r}_1^2 \\ \mathbf{r}_2^2 \\ \mathbf{r}_1 \mathbf{r}_2 \end{pmatrix},$$

where $\mathbf{r}_1$ and $\mathbf{r}_2$ are the first and second rows of M respectively, and where multiplication of vectors is conducted coordinatewise. We claim that if M is linearly generic and $k \geq 3$ then $M^{(2)}$ has rank 3. One may, for example, observe that if 3 values of $\mathbf{r}_1$ are nonzero, then $M^{(2)}$ contains a 3×3 Vandermonde determinant (up to rescaling columns). After similarly dealing with the case that 3 values of $\mathbf{r}_2$ are nonzero, and noting that each of $\mathbf{r}_1$ and $\mathbf{r}_2$ cannot contain more than one zero entry, we may conclude immediately in the remaining case by a cofactor expansion along the row $\mathbf{r}_1 \mathbf{r}_2$ (two of whose entries are zero).

Proposition 3.7. *Let $t \geq 5$ be odd. Let Ψ be a system of t linear forms whose image has codimension 2 and which is linearly generic. Let Φ be the system defined by an additive $(t-1)$ -tuple. Then, for all n sufficiently large in terms of p and t , there is a function $f : \mathbb{F}_p^n \rightarrow [-1, 1]$ such that $\mathbb{E}_x f(x) = 0$, $|T_\Psi(f)| \gg_{p,t} p^{-n(t-4)/2}$ and $T_\Phi(f) < p^{-n(t-3)/2}$.*

Proof. We construct f by specifying its Fourier transform $\hat{f} : \mathbb{F}_p^n \rightarrow \mathbb{C}$. Let A be the zero set of the quadratic form given by $x^T x$. Define $\hat{f}(0) = 0$ and $\hat{f}(h) = \frac{1}{p^{n/2+1}}(1_A(h) - p^{-1})$ otherwise. Firstly, $T_\Phi(f) = \sum_h |\hat{f}(h)|^{t-1} < p^n \cdot (p^{-n/2})^{t-1} = p^{-n(t-3)/2}$.

It is clear that for all $h \in \mathbb{F}_p^n$ we have $\hat{f}(h) = \hat{f}(-h)$, and so f takes values in $\mathbb{R}$. Furthermore, for all $x \neq 0$:

$$\begin{aligned} |f(x)| &= \left| \sum_h \hat{f}(h) e_p(x \cdot h) \right| \\ &= \frac{1}{p^{n/2+1}} \left| \sum_h (1_A(h) - p^{-1}) e_p(x \cdot h) - (1 - p^{-1}) \right| \\ &\leq \frac{1}{p^{n/2+1}} \left(\left| p^n \mathbb{E}_h \left(\mathbb{E}_{a \in \mathbb{F}_p} e_p(ah^T h) \right) e_p(x^T h) \right| + (1 - p^{-1}) \right) \end{aligned}$$

$$\begin{aligned} &\leq \frac{1}{p^{n/2} + 1} \left(p^n \mathbb{E}_{a \in \mathbb{F}_p} |\mathbb{E}_h e_p(a h^T h + x^T h)| + (1 - p^{-1}) \right) \\ &\leq \frac{1}{p^{n/2} + 1} \left(p^{n/2} + 1 - p^{-1} \right) \\ &< 1, \end{aligned}$$

by Fourier inverting in the first line, an expansion of 1_A in terms of quadratic phases in the third line (this follows from the orthogonality of characters), and Lemma 3.6 Part 1 in the fourth. For $x = 0$ we have

$$|f(0)| = \frac{1}{p^{n/2} + 1} \left| \sum_h (1_A(h) - p^{-1}) - 1 + p^{-1} \right| < 1,$$

by Lemma 3.6 Part 2 with the trivial set of linear forms $x \mapsto (x)$. Thus f takes values in $[-1, 1]$.

By (3) and the condition that the image of Ψ has codimension 2 we have

$$|T_\Psi(f)| = p^{2n} |T_{\Psi^\perp}(\hat{f})| \gg_{p,t} p^{2n} \cdot p^{-tn/2} |T_{\Psi^\perp}(1_A - p^{-1}) + O_t(p^{-n/2})|,$$

where $\Psi^\perp$ denotes the dual set of linear forms which may be determined as in (3). It remains to show that $|T_{\Psi^\perp}(1_A - p^{-1})| \gg_{p,t} 1$. We have

$$\begin{aligned} T_{\Psi^\perp}(1_A - p^{-1}) &= \sum_{S \subset [t]} (-p)^{-t+|S|} T_{\Psi^\perp(S)}(1_A) \\ &= \sum_{S \subset [t]} (-p)^{-t+|S|} \mathbb{E}_{x_1, x_2} \prod_{i \in S} 1_0(\psi_i^\perp(x_1, x_2)^T \psi_i^\perp(x_1, x_2)). \end{aligned}$$

Since the image of Ψ has codimension 2, we have that each $\psi_i^\perp \in \Psi^\perp$ maps $(\mathbb{F}_p^n)^2 \rightarrow \mathbb{F}_p^n$. Thus each $\psi_i^\perp(x)^T \psi_i^\perp(x)$ maps $x = (x_1, x_2) \in (\mathbb{F}_p^n)^2$ to a linear combination of $\{x_1^T x_1, x_1^T x_2, x_2^T x_2\}$, so the dimension of the span of the functions $\{\psi_i^\perp(x)^T \psi_i^\perp(x)\}_{i=1}^t$ is at most 3. By Lemma 3.6,

$$T_{\Psi^\perp}(1_A - p^{-1}) = \sum_{S \subset [t]} (-p)^{-t+|S|} p^{-\dim \text{span}\{\psi_i^\perp(x)^T \psi_i^\perp(x)\}_{i \in S}} + O_t(p^{-n/2}).$$

One observes that $\dim \text{span}\{\psi_i^\perp(x)^T \psi_i^\perp(x)\}_{i \in S} = \text{rank } M_S^{(2)}$ where M_S is the matrix obtained by restricting the columns of M_Ψ to the coordinates in S . We show that $|T_\Psi(1_A - p^{-1})| \gg_{p,t} 1$ by showing that $\sum_{S \subset [t]} (-p)^{-t+|S|} p^{-\text{rank } M_S^{(2)}}$ is nonzero. As discussed above, the fact that every 2×2 minor of M_Ψ has nonzero determinant implies that for each $S \subset [t]$ with $|S| \geq 3$, $\text{rank } M_S^{(2)} = 3$. Furthermore it is clear that if $|S| = 1, 2$ then $\text{rank } M_S^{(2)} = 1, 2$ respectively. Thus we have

$$\begin{aligned} \sum_{S \subset [t]} (-p)^{-t+|S|} p^{-\text{rank } M_S^{(2)}} &= p^{-3} \sum_{|S| \geq 3} (-p)^{-t+|S|} + \left(\binom{t}{2} - t + 1 \right) (-p)^{-t} \\ &= \frac{(1-p)^t + p^3 \left(\binom{t}{2} - t + 1 \right) - (p^2 \binom{t}{2} - pt + 1)}{(-p)^t p^3}. \end{aligned}$$

We claim that the numerator is nonzero for pairs (p, t) with p prime and $t \geq 5$. Indeed, for $p = 2$ and t odd we obtain a quadratic whose roots satisfy $t < 5$. The same is true of $p = 2$ and t even. Otherwise, for large values of (p, t) it is clear that the term $(1-p)^t$ dominates, and the claim may be checked by direct computation for small values of (p, t) (the author obtained a bona fide proof when $t \geq 11$, or $t \in \{5, 7, 9\}$ and $p \geq 11$, and checked the cases $(p, t) \in \{3, 5, 7\} \times \{5, 7, 9\}$ by hand; for the sake of brevity we do not include the details). $\square$

Remark 3.8. In the above proposition, the assumption that Ψ is linearly generic may be removed quite easily, at least if one allows p to be sufficiently large in terms of t . Indeed one writes $\sum_{S \subset [t]} (-p)^{-t+|S|} p^{-\text{rank } M_S^{(2)}}$ as a p -adic series $\sum_{i=-t}^{-1} a_i p^i$. Each $|a_i|$ may be bounded in terms of t , so if p is large enough in terms of t then this series representation is unique and so the sum is zero if and only if each a_i is zero. One may use some linear algebra to show that at least one of the a_i must be nonzero. It is certainly plausible that the more general result may be used in the further classification of Sidorenko systems in a directly analogous way to how Proposition 3.7 is used here, but we have no need for such a statement in this document and so we do not write out the details.

Proof of Theorem 3.5. Recall that we may assume that t is odd. As in the proof of Theorem 3.1, using (2), we have

$$T_{\Psi}(\alpha + \varepsilon f) = \alpha^t + \alpha \varepsilon^{t-1} \sum_{|S|=t-1} T_{\Psi(S)}(f) + \varepsilon^t T_{\Psi}(f).$$

As we have seen already, each $\Psi(S)$ has codimension one and so $T_{\Psi(S)}(f)$ is of the form $\sum_h \prod_{i=1}^{t-1} \hat{f}(a_i h)$. By Hölder's inequality then

$$|T_{\Psi(S)}(f)| \leq \sum_h |\hat{f}(h)|^{t-1} = T_{\Phi}(f),$$

where we let Φ denote an additive $(t-1)$ -tuple. Thus

$$T_{\Psi}(\alpha + \varepsilon f) \leq \alpha^t + (t-1)\alpha \varepsilon^{t-1} T_{\Phi}(f) + \varepsilon^t T_{\Psi}(f).$$

Let f be the function from Proposition 3.7, and if $T_{\Psi}(f) > 0$ then replace f with $-f$, so

$$T_{\Psi}(\alpha + \varepsilon f) \leq \alpha^t + (t-1)\alpha \varepsilon^{t-1} p^{-n(t-3)/2} - \varepsilon^t c_{p,t} p^{-n(t-4)/2}.$$

For ε, α fixed, we may choose n large enough so that $T_{\Psi}(\alpha + \varepsilon f) < \alpha^t$, completing the proof. $\square$

This also completes the proof of Theorem 1.1, which is a union of Theorem 3.1 and Theorem 3.5.

4. Examples

4.1. Locally Sidorenko but not Sidorenko

Recall that a system Ψ comprising t linear forms is locally Sidorenko if for all $\alpha \in [0, 1]$ there exists $\varepsilon > 0$ such that for all $n \geq 1$ and all $f : \mathbb{F}_p^n \rightarrow [-\alpha, 1 - \alpha]$ with $\mathbb{E}f = 0$ we have $T_{\Psi}(\alpha + \varepsilon f) \geq \alpha^t$. We begin by providing the following example of a system which is locally Sidorenko but not Sidorenko. This proves Proposition 1.6.

Example 4.1. Let Φ denote the following system of two linear equations in nine variables:

$$\begin{aligned} x_1 - x_2 + x_3 - x_4 &= 0 \\ x_5 - x_6 + x_7 - x_8 + x_9 &= 0. \end{aligned}$$

We firstly observe that Φ is not Sidorenko. Indeed, let $A = \{1\} \subset \mathbb{F}_p$. Then $\mathbb{E}_{x \in \mathbb{F}_p} 1_A(x) = 1/p$, but A contains no solutions to Φ , so $T_{\Phi}(1_A) = 0$.

We claim that Φ is locally Sidorenko. Let A4 be shorthand for the first equation in Φ and A5 for the second. We observe (see Observation 2.6) that for a function f on $\mathbb{F}_p^n$ we have $T_{\Phi}(f) = T_{A4}(f)T_{A5}(f)$. Also, since A4 has codimension one, we get $T_{A4}(\alpha + f) = \alpha^4 + T_{A4}(f)$ for functions f with $\mathbb{E}f = 0$, and similarly for A5. Thus, for $f : \mathbb{F}_p^n \rightarrow [-\alpha, 1 - \alpha]$ with $\mathbb{E}f = 0$, we see using (1) that

$$\begin{aligned} T_{\Phi}(\alpha + \varepsilon f) &= (\alpha^4 + T_{A4}(\varepsilon f))(\alpha^5 + T_{A5}(\varepsilon f)) \\ &= \alpha^9 + \varepsilon^4 \left(\alpha^5 T_{A4}(f) + \alpha^4 \varepsilon T_{A5}(f) + \varepsilon^5 T_{A4}(f) T_{A5}(f) \right). \end{aligned}$$

Fourier-inverting, we have that $T_{A4}(f) = \sum_h |\hat{f}(h)|^4$ and $T_{A5}(f) = \sum_h |\hat{f}(h)|^4 \hat{f}(h)$. In particular we have that $|T_{A5}(f)| \leq T_{A4}(f) \leq 1$. Thus,

$$T_{\Phi}(\alpha + \varepsilon f) \geq \alpha^9 + \varepsilon^4 T_{A4}(f) (\alpha^5 - \alpha^4 \varepsilon - \varepsilon^5),$$

and so setting $\varepsilon \leq \alpha/2$ we have that $T_{\Phi}(\alpha + \varepsilon f) \geq \alpha^9$, so Φ is indeed locally Sidorenko.

4.2. Weakly locally Sidorenko but not locally Sidorenko

Recall that a system Ψ comprising t linear forms is weakly locally Sidorenko if for all $n \geq 1$, for all $\alpha \in [0, 1]$ and all $f : \mathbb{F}_p^n \rightarrow [-\alpha, 1 - \alpha]$ with $\mathbb{E}f = 0$, there exists $\varepsilon_f > 0$ such that for all $\varepsilon < \varepsilon_f$ we have $T_{\Psi}(\alpha + \varepsilon f) \geq \alpha^t$.

We claimed earlier that it is easy to contrive systems of complexity at least two which are weakly locally Sidorenko, but not Sidorenko. One may, for example, begin with a system of complexity at least two whose shortest subsystems have length at least 5, and take the union of this system with an additive 4-tuple on a separate variable set. The weak local Sidorenko property follows from the fact that its smallest subsystem is an additive 4-tuple, and the (non) Sidorenko property of the entire systems essentially follows from that of the complexity two part by tensoring with an appropriate quadratic phase function. We omit the details.

In this subsection we exhibit a *complexity one*, linearly generic system which is weakly locally Sidorenko, but not locally Sidorenko, proving Corollary 1.5. The upcoming example is taken from [KLM23, Example 4.6], as is most of the analysis towards showing that it is weakly locally Sidorenko. That it is not locally Sidorenko follows from Theorem 1.1.

Example 4.2. Let $p \notin \{2, 3\}$ and let Φ be the following system of two equations in five variables:

$$\begin{aligned} x_1 - x_2 + x_3 - x_4 &= 0 \\ x_1 + 2x_2 - x_3 - 2x_5 &= 0. \end{aligned}$$

One sees that Φ is linearly generic. By Theorem 1.1, Φ is not locally Sidorenko. We claim that Φ is weakly locally Sidorenko. For f with $\mathbb{E}f = 0$ we have from (2) that

$$T_{\Phi}(\alpha + \varepsilon f) = \alpha^5 + \alpha \varepsilon^4 \sum_{|S|=4} T_{\Phi(S)}(f) + \varepsilon^5 T_{\Phi}(f),$$

so it suffices to show that $\sum_{|S|=4} T_{\Phi(S)}(f) > 0$ whenever $f \neq 0$. One sees that equations for the five subsystems S of size 4 may be obtained by performing row operations to eliminate the variable x_i , $i = 1, \dots, 5$. Therefore, applying (3) to each of these subsystems, and recalling that $T_{\Phi(S)}(f)$ is real-valued, we may obtain that

$$\begin{aligned} \sum_{|S|=4} T_{\Phi(S)}(f) &= \sum_h |\hat{f}(h)|^4 + 2|\hat{f}(h)|^2 |\hat{f}(2h)|^2 + 2\hat{f}(-h) \hat{f}(2h)^2 \hat{f}(-3h) \\ &= \sum_h \frac{1}{2} |\hat{f}(2h)|^4 + \frac{1}{2} |\hat{f}(3h)|^4 + 2|\hat{f}(h)|^2 |\hat{f}(2h)|^2 \\ &\quad + 2\Re \left(\hat{f}(-h) \hat{f}(2h)^2 \hat{f}(-3h) \right) \end{aligned}$$

$$\begin{aligned}
&\geq \sum_h |\hat{f}(-2h)\hat{f}(3h)|^2 + 2|\hat{f}(h)|^2|\hat{f}(2h)|^2 \\
&\quad + 2\Re\left(\hat{f}(-h)\hat{f}(2h)^2\hat{f}(-3h)\right) \\
&= \sum_h \left|\hat{f}(h)\hat{f}(-2h) + \overline{\hat{f}(-2h)\hat{f}(3h)}\right|^2 + |\hat{f}(h)|^2|\hat{f}(2h)|^2 \\
&\geq \sum_h |\hat{f}(h)|^2|\hat{f}(2h)|^2.
\end{aligned}$$

This is positive unless $|\hat{f}(h)|^2|\hat{f}(2h)|^2 = 0$ for all h . But in this case $\sum_{|S|=4} T_{\Phi(S)}(f) = \sum_h |\hat{f}(h)|^4$, which is positive unless f is zero. This completes the proof of the claim that f is weakly locally Sidorenko.

4.3. Locally Sidorenko, not Sidorenko, and no additive tuple

Recall that if Φ comprises only an additive $(2k)$ -tuple, the corresponding functional T_Φ is a norm on the space of functions and in particular is positive definite: $T_\Phi(f) = 0 \implies f = 0 \implies T_\Psi(f) = 0$, for any Ψ . This is the main obstruction to the proof of Theorem 3.1 going through for systems with an additive $(2k)$ -tuple. It transpires that there exists a nonlinearly generic Ψ with a subsystem Φ whose corresponding functional is not positive definite but with the property that $T_\Phi(f) = 0 \implies T_{\Psi(S)}(f) = 0$ for all subsystems $\Psi(S) \subseteq \Psi$. This motivates the upcoming example.

Example 4.3. Let Φ be the following system of two equations in eight variables:

$$\begin{aligned}
x_1 - x_2 + 2x_3 - 2x_4 + x_5 + x_6 + x_7 + x_8 &= 0 \\
x_5 - x_6 + 2x_7 - 2x_8 &= 0.
\end{aligned}$$

We claim that Φ is locally Sidorenko (and it clearly does not contain any additive tuples). Fix $\alpha \in (0, 1)$ and let $f : \mathbb{F}_p^n \rightarrow [-\alpha, 1 - \alpha]$. Firstly invoke (1) to write $T_\Phi(\alpha + \varepsilon f)$ in terms of the subsystems of Φ . These subsystems may be computed by parameterising the solution space to Φ . Alternatively, one sees that they may be read off the defining equations for Φ after performing row operations (recall that to identify subsystems which contribute nontrivially, we simply need to find those coordinate restrictions whose image do not contain a coordinate hyperplane; see remarks after (1)). Doing so, and invoking (3) to pass to frequency space, one finds that

$$\begin{aligned}
T_\Phi(\alpha + \varepsilon f) &= \alpha^8 + \alpha^4 \varepsilon^4 \sum_h |\hat{f}(h)|^2 |\hat{f}(2h)|^2 \\
&\quad + \alpha \varepsilon^7 \sum_h |\hat{f}(h)|^2 |\hat{f}(2h)|^2 \hat{f}(-2h) \hat{f}(h) \hat{f}(-3h) \\
&\quad + \alpha \varepsilon^7 \sum_h |\hat{f}(h)|^2 |\hat{f}(2h)|^2 \hat{f}(2h) \hat{f}(3h) \hat{f}(-h) \\
&\quad + \alpha \varepsilon^7 \sum_h |\hat{f}(h)|^2 |\hat{f}(2h)|^2 \hat{f}(-h) \hat{f}(-3h) \hat{f}(-4h) \\
&\quad + \alpha \varepsilon^7 \sum_h |\hat{f}(h)|^2 |\hat{f}(2h)|^2 \hat{f}(3h) \hat{f}(h) \hat{f}(4h) \\
&\quad + \varepsilon^8 \sum_{h_1, h_2} |\hat{f}(h_1)|^2 |\hat{f}(2h_1)|^2 \\
&\quad \cdot \hat{f}(h_1 + h_2) \hat{f}(h_1 - h_2) \hat{f}(h_1 + 2h_2) \hat{f}(h_1 - 2h_2).
\end{aligned}$$

The absolute value of the final sum is at most

$$\begin{aligned} \sum_{h_1} |\hat{f}(h_1)|^2 |\hat{f}(2h_1)|^2 \sum_{h_2} |\hat{f}(h_1 + h_2) \hat{f}(h_1 - h_2) \hat{f}(h_1 + 2h_2) \hat{f}(h_1 - 2h_2)| \\ \leq \sum_{h_1} |\hat{f}(h_1)|^2 |\hat{f}(2h_1)|^2 \sum_{h_2} |\hat{f}(h_1 + h_2) \hat{f}(h_1 - h_2)|, \end{aligned}$$

by the bound $\|\hat{f}\|_\infty \leq 1$. Applying Cauchy–Schwarz to the sum over h_2 and subsequently Parseval’s identity yields an upper bound of $\sum_h |f(h)|^2 |f(2h)|^2$. Thus

$$T_\Phi(\alpha + \varepsilon f) \geq \alpha^8 + \varepsilon^4 \left(\sum_h |\hat{f}(h)|^2 |\hat{f}(2h)|^2 \right) (\alpha^4 - 4\alpha\varepsilon^3 - \varepsilon^4).$$

Thus setting $\varepsilon \leq \alpha/2$, we have that $T_\Phi(\alpha + \varepsilon f) \geq \alpha^8$, and so Φ is locally Sidorenko.

We note that the previous example is not Sidorenko, so it gives another, albeit less simple, example verifying Proposition 1.6.

5. Localising to subsystems with fewest variables

In this section we observe how to ‘localise’ to subsystems with the fewest variables. As an application we outline a relatively short proof of the recent result of Kamčev–Liebenau–Morrison [KLM24] and independently Versteegen [Ver21] that any system containing a four-term arithmetic progression (or indeed any system of two independent equations in four variables) is uncommon. Our proof does not use any quadratic methods, which is also new here.

Previously we have localised by considering small perturbations of a constant function $\alpha + \varepsilon f$, whereupon, defining s to be the size of smallest subsystem of Ψ for which some $T_{\Psi(S)}(f)$ is nonzero, we have

$$T_\Psi(\alpha + \varepsilon f) \approx \alpha^t + \varepsilon^s \sum_{|S|=s} T_{\Psi(S)}(f).$$

For fixed positive integers n, k and $f : \mathbb{F}_p^n \rightarrow [-1, 1]$, define the function $f^{(k)} : \mathbb{F}_p^{n+k} \rightarrow [-1, 1]$ by $f^{(k)} := f \otimes 1_0$ where 1_0 is the characteristic function for $0 \in \mathbb{F}_p^k$ (i.e., $f^{(k)} = (f \circ \pi_1)(1_0 \circ \pi_2)$, where π_1 projects onto the first n coordinates, π_2 projects onto the final k coordinates). Then for an injective system $\Psi = (\psi_1, \dots, \psi_t)$ in D variables, we have recalling (4) that

$$T_\Psi(f^{(k)}) = T_\Psi(f) T_\Psi(1_0) = p^{-kD} T_\Psi(f).$$

Thus, expanding $T_\Psi(\alpha + f \otimes 1_0)$ multilinearly and letting k be large we have that

$$T_\Psi(\alpha + f \otimes 1_0) = \alpha^t + p^{-kd} \sum_{\dim \Psi(S)=d} T_{\Psi(S)}(f) + O_t(p^{-k(d+1)}),$$

where $\dim \Psi(S)$ denotes the dimension of the image of $\Psi(S)$ and d is the minimal such quantity for a subsystem for which $T_{\Psi(S)}(f)$ is nonzero.

Now we outline an alternative proof of the fact that any system containing a 4AP is uncommon. In fact we will prove the stronger result (as [KLM24] did) that any system containing a rank two system with four linear forms is uncommon. We note that some parts of the argument remain the same. In particular, we need the result that rank 2 systems in 4 variables can have $T_\Psi(f) < 0$. Rather than reproving it we just quote the following weak version of [KLM24, Lemma 4.1] whose proof we note does not use higher-order methods.

Lemma 5.1 [KLM24, Lemma 4.1]. *Let p be a prime. There exists an integer $n \geq 1$ such that the following is true. Let $(\Psi_i)_{i=1}^k$ be a collection of linear systems each in four variables with codimension two. Suppose that Ψ_i is linearly generic for all i . Then there exists $f : \mathbb{F}_p^n \rightarrow [-\frac{1}{2}, \frac{1}{2}]$ with $\mathbb{E}f = 0$ such that $T_{\Psi_i}(f) < 0$ for $i = 1, \dots, k$.*

Theorem 5.2 [KLM24, Theorem 1.1], generalisation of [Ver21, Theorem 1.4]. *Any linear system of distinct, nonzero forms which contains a dimension 4 subsystem of rank two is uncommon.*

Proof. Let t be the number linear forms in Ψ and let $f : \mathbb{F}_p^n \rightarrow [-\frac{1}{2}, \frac{1}{2}]$ have $\mathbb{E}f = 0$ and be otherwise unspecified for the meantime. Let $k \geq 0$ to be determined later, $\varepsilon = p^{-k}$ and $1_0 : \mathbb{F}_p^k \rightarrow \{0, 1\}$ be the characteristic function of zero. Then

$$\begin{aligned} T_{\Psi} \left(\frac{1}{2} + (\varepsilon f) \otimes 1_0 \right) + T_{\Psi} \left(\frac{1}{2} - (\varepsilon f) \otimes 1_0 \right) \\ = \frac{1}{2^{t-1}} + \sum_{\substack{\emptyset \neq S \subset [t] \\ |S| \text{ even}}} \frac{1}{2^{t-|S|-1}} p^{-k(|S| + \dim \Psi(S))} T_{\Psi(S)}(f). \end{aligned}$$

Letting k be arbitrarily large, we see it suffices to show that there is a choice of f such that $\sum_{S: \dim \Psi(S) + |S| \text{ minimal}} T_{\Psi(S)}(f) < 0$, where the minimum is taken over subsystems whose image does not contain a coordinate hyperplane (if the image contains a coordinate hyperplane then $T_{\Psi(S)}(f) = 0$ since $\mathbb{E}f = 0$). Note that this minimum is at most 6 since Ψ contains a system of four nonzero linear forms with codimension two, and for the minimum we must have that $\dim \Psi(S) < |S|$.

If the minimum is equal to 6, then the minimal systems $\Psi(S)$ must have $|S| = 4$, $\dim \Psi(S) = 2$. If this is the case then we must have that $\Psi(S)$ is linearly generic for all of these minimal systems, or else by an easy linear algebraic check they would give rise to a further subsystem S' with $|S'| + \dim \Psi(S')$ smaller. Thus we may conclude by invoking Lemma 5.1 (with an appropriate choice of n) in this case.

Otherwise any $\Psi(S)$ contributing to the minimum must have $\dim \Psi(S) = 1$, and indeed cannot have $|S| = 4$ since any system $\Psi(S)$ with $|S| = 4$, $D = 1$ gives rise to $\binom{4}{2}$ subsystems $\Psi(S_i)$ with $\dim \Psi(S_i) = 1$, $|S_i| = 2$. In the remaining case that the minimal systems have $\dim \Psi(S) = 1$, $|S| = 2$ we note that these systems are described by a single equation which cannot have its two coefficients summing to zero since this would mean that Ψ has a pair of repeated forms. Thus we may conclude by the Fox–Pham–Zhao random Fourier sampling argument to find f such that $\sum_{S: \dim \Psi(S) + |S| \text{ minimal}} T_{\Psi(S)}(f) < 0$.³ This completes the proof. $\square$

Remark 5.3. Kamčev–Liebenau–Morrison conjecture [KLM24, Conjecture 6.1] that an analogue of Lemma 5.1 holds for systems comprising an even number of linear forms whose image has codimension two. They note that their proof of [KLM24, Theorem 1.1] however does not generalise to the case of higher number of variables, even if [KLM24, Conjecture 6.1] is confirmed. We note that our proof above would generalise should [KLM24, Conjecture 6.1] be confirmed. This would go some way towards the classification of weakly locally common systems.

Remark 5.4. Since this document was released on the arXiv, the generalisation suggested in Remark 5.3 has been carried out using the methods of this chapter in both [AL25] and [DLZ24].

Acknowledgements. The author thanks Ben Green for drawing his attention to the study of the local Sidorenko property in the graph-theoretic setting (in particular [Lov11]) and for valuable feedback on an earlier version of this document. We also thank Julia Wolf for valuable feedback on an earlier version of this document and an anonymous referee for a careful reading and feedback.

Competing interests. The author has no competing interests to declare.

³In fact this is a very easy case of the Fox–Pham–Zhao argument since the equations comprise only two variables. It is likely an even simpler argument exists, but this is a known (and nice) argument, so we quote it for brevity.

References

- [AL25] D. Altman and A. Liebenau, ‘On the uncommonness of minimal rank-2 systems of linear equations’, *Bull. Lond. Math. Soc.*, 2025. <https://doi.org/10.1112/blms.70135>.
- [Alt23] D. Altman, ‘On a common-extendable, non-Sidorenko linear system’, *Comb. Theory*, **3**(3):Paper No. 5, 12, 2023.
- [BR80] S. A. Burr and V. Rosta, ‘On the Ramsey multiplicities of graphs—problems and recent results’, *J. Graph Theory* **4**(4): 347–361, 1980.
- [CFS10] D. Conlon, J. Fox, and B. Sudakov, ‘An approximate version of Sidorenko’s conjecture’, *Geom. Funct. Anal.* **20**(6): 1354–1366, 2010.
- [CHL22] E. Cs’oka, T. Hubai, and L. M. Lov’asz, ‘Locally common graphs’, *J. Graph Theory* **102**: 472–483, 2022.
- [CKLL18] D. Conlon, J. H. Kim, C. Lee, and J. Lee, ‘Some advances on Sidorenko’s conjecture’, *J. Lond. Math. Soc. (2)* **98**(3): 593–608, 2018.
- [CL17] D. Conlon and J. Lee, ‘Finite reflection groups and graph norms’, *Adv. Math.* **315**: 130–165, 2017.
- [CL21] D. Conlon and J. Lee, ‘Sidorenko’s conjecture for blow-ups’, *Discrete Anal.*, page 21472, 2021.
- [DLZ24] D. Dong, A. Li, and Y. Zhao, ‘Uncommon linear systems of two equations’, [arXiv:2404.17005](https://arxiv.org/abs/2404.17005), 2024.
- [Erd62] P. Erdős, ‘On the number of complete subgraphs contained in certain graphs’, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **7**: 459–464, 1962.
- [FPZ21] J. Fox, H. T. Pham, and Y. Zhao, ‘Common and Sidorenko linear equations’, *Q. J. Math.* **72**(4): 1223–1234, 2021.
- [FW17] J. Fox and F. Wei, ‘On the local approach to Sidorenko’s conjecture’, *Electron. Notes Discret. Math.* **61**: 459–465, 2017.
- [Goo59] A. W. Goodman, ‘On sets of acquaintances and strangers at any party’, *Amer. Math. Mon.* **66**: 778–783, 1959.
- [Gow20] W. T. Gowers, ‘A uniform set with fewer than expected arithmetic progressions of length 4’, *Acta Math. Hungar.* **161**(2): 756–767, 2020.
- [Gre07] B. Green, ‘Montréal notes on quadratic Fourier analysis’, in *Additive Combinatorics*, vol. 43 of *CRM Proc. Lecture Notes*, 69–102, American Mathematical Society, Providence, RI, 2007.
- [GW10] W. T. Gowers and J. Wolf, ‘The true complexity of a system of linear equations’, *Proc. Lond. Math. Soc. (3)* **100**(1): 155–176, 2010.
- [GW11a] W. T. Gowers and J. Wolf, ‘Linear forms and higher-degree uniformity for functions on $\mathbb{F}_p^n$ ’, *Geom. Funct. Anal.* **21**(1): 36–69, 2011.
- [GW11b] W. T. Gowers and J. Wolf, ‘Linear forms and quadratic uniformity for functions on $\mathbb{F}_p^n$ ’, *Mathematika* **57**(2): 215–237, 2011.
- [Hat10] H. Hatami, ‘Graph norms and Sidorenko’s conjecture’, *Israel J. Math.* **175**: 125–150, 2010.
- [HKKV22] R. Hancock, D. Král’, M. Krnc, and J. Volec, ‘Toward characterizing locally common graphs’, *Random Struct. Algorithms* **62**: 181–218, 2022.
- [JvT96] C. Jagger, P. Šťovíček, and A. Thomason, ‘Multiplicities of subgraphs’, *Combinatorica* **16**(1): 123–141, 1996.
- [KLL16] J. H. Kim, C. Lee, and J. Lee, ‘Two approaches to Sidorenko’s conjecture’, *Trans. Amer. Math. Soc.* **368**(7): 5057–5074, 2016.
- [KLM23] N. Kamčev, A. Liebenau, and N. Morrison, ‘Towards a characterization of Sidorenko systems’, *Q. J. Math.* **74**(3): 957–974, 2023.
- [KLM24] N. Kamčev, A. Liebenau, and N. Morrison, ‘On uncommon systems of equations’, *Israel J. Math.* **264**(1): 331–362, 2024.
- [KLP22] D. Král’, A. Lemaître, and P. P. Pach, ‘Common systems of two equations over the binary field’, in *Discrete Math. Days*, Conf. Proc., 2022, 169–173. Extended abstracts available at <https://dmd2022.unican.es/978-84-19024-03-9.pdf>.
- [Lov11] L. Lovász, ‘Subgraph densities in signed graphons and the local Simonovits–Sidorenko conjecture’, *Electron. J. Combin.* **18**(1):Paper 127, 21, 2011.
- [Lov12] L. Lovász, *Large Networks and Graph Limits*, Amer. Math. Soc. Colloq. Publ. 60, American Mathematical Society, Providence, RI, 2012.
- [LS11] J. X. Li and B. Szegedy, ‘On the logarithmic calculus and Sidorenko’s conjecture’, [arXiv:1107.1153](https://arxiv.org/abs/1107.1153), 2011.
- [Sid89] A. F. Sidorenko, ‘Cycles in graphs and functional inequalities’, *Mat. Zametki* **46**(5): 72–79, 104, 1989.
- [Sid91] A. F. Sidorenko, ‘Inequalities for functionals generated by bipartite graphs’, *Diskret. Mat.* **3**(3): 50–65, 1991.
- [Sid93] A. F. Sidorenko, ‘A correlation inequality for bipartite graphs’, *Graphs Combin.* **9**(2): 201–204, 1993.
- [SW17] A. Saad and J. Wolf, ‘Ramsey multiplicity of linear patterns in certain finite abelian groups’, *Q. J. Math.* **68**(1): 125–140, 2017.
- [Sze14] B. Szegedy, ‘An information theoretic approach to Sidorenko’s conjecture’, [arXiv:1406.6738](https://arxiv.org/abs/1406.6738), 2014.
- [Tho89] A. Thomason, ‘A disproof of a conjecture in Ramsey theory’, *J. Lond. Math. Soc. (2)* **39**(2): 246–255, 1989.
- [Ver21] L. Versteegen, ‘Linear configurations containing 4-term arithmetic progressions are uncommon’, *J. Combin. Theory Ser. A* **200**:105792, 2021.
- [Zha23] Y. Zhao, *Graph Theory and Additive Combinatorics—Exploring Structure and Randomness*, Cambridge University Press, Cambridge, 2023.