On norming systems of linear equations

Seokjoon Cho[*]   David Conlon[†]   Joonkyung Lee[‡]   Jozef Skokan[§]

Leo Versteegen[¶]

Abstract

A system of linear equations L is said to be norming if a natural functional $t_L(\cdot)$ giving a weighted count for the set of solutions to the system can be used to define a norm on the space of real-valued functions on $\mathbb{F}_q^n$ for every $n > 0$. For example, Gowers uniformity norms arise in this way. In this paper, we initiate the systematic study of norming linear systems by proving a range of necessary and sufficient conditions for a system to be norming. Some highlights include an isomorphism theorem for the functional $t_L(\cdot)$, a proof that any norming system must be variable-transitive and the classification of all norming systems of rank at most two.

1   Introduction

One of the key tools in modern additive combinatorics, introduced by Gowers [11] in his seminal work on Szemerédi's theorem, are the Gowers uniformity norms $\|\cdot\|_{U_k}$. For example, for an abelian group G and a function $f : G \to \mathbb{R}$, the U_2-norm of f is given by

$$\|f\|_{U_2}^4 = \mathbb{E}\left[f(x_1)f(x_2)f(x_3)f(x_4)\right],$$

where the expectation is taken over all solutions to the linear equation $x_1 - x_2 + x_3 - x_4 = 0$. That this functional and its generalisations are indeed norms follows from first proving a certain Hölder-type inequality using iterated applications of the Cauchy–Schwarz inequality and this in turn implies the required triangle inequality.

Our concern in this paper will be with the question of deciding which other systems of linear equations give rise to norms. Although many of our results do extend to other abelian groups, we will work throughout with groups G of the form $\mathbb{F}_q^n$ where q is a prime power and n is a parameter that we often regard as tending to infinity. In this context, an $m \times k$ *linear system on G* is a matrix $L \in \mathbb{F}_q^{m \times k}$ with linearly independent rows, though we use the term also for the corresponding system of linear equations. In line with this correspondence, by an *equation* in L, we mean a vector in the row space of L, with, for instance, the ith row $(L_{i1}, \ldots, L_{ik})$ of L corresponding to the equation $L_{i1}x_1 + \cdots + L_{ik}x_k = 0$.

Given a linear system L, we write $\mathrm{Sol}(L) = \{x \in G^k : Lx = 0\}$ for the solution set of L in G^k, i.e., those $(x_1, \ldots, x_k) \in G^k$ which satisfy $L_{i1}x_1 + \cdots + L_{ik}x_k = 0$ for all $i \in [m]$. Then, for

---

[*]Department of Mathematical Sciences, Seoul National University, South Korea. Email: `znthlz@snu.ac.kr`.

[†]Department of Mathematics, California Institute of Technology, United States. Email: `dconlon@caltech.edu`.

[‡]Department of Mathematics, Yonsei University, South Korea. Email: `joonkyunglee@yonsei.ac.kr`.

[§]Department of Mathematics, London School of Economics and Political Science, United Kingdom. Email: `j.skokan@lse.ac.uk`.

[¶]Department of Mathematics, London School Economics and Political Science, United Kingdom. Email: `lversteegen.math@gmail.com`.

real-valued functions $f_1, \ldots, f_k$ on G, we set

$$t_L(f_1, \ldots, f_k) := \mathbb{E}_{(x_1, \ldots, x_k) \in \mathrm{Sol}(L)} f_1(x_1) \cdots f_k(x_k),$$

where, here and throughout, $\mathbb{E}_{(x_1, \ldots, x_k) \in \mathrm{Sol}(L)}$ is a shorthand for $\frac{1}{|\mathrm{Sol}(L)|} \sum_{(x_1, \ldots, x_k) \in \mathrm{Sol}(L)}$.

If we now set $t_L(f) := t_L(f, \ldots, f)$ and $\|f\|_L := |t_L(f)|^{1/k}$, we can say that L is *(semi-)norming* if $\| \cdot \|_L$ defines a (semi-)norm on the space of real-valued functions on $\mathbb{F}_q^n$ for all n. As it is easy to see that $\|cf\|_L = |c|\|f\|_L$, a linear system L is semi-norming if and only if $\| \cdot \|_L$ satisfies the triangle inequality. We also say that L is *weakly (semi-)norming* if $\|f\|_{r(L)} := t_L(|f|)^{1/k}$ defines a (semi-)norm. It will turn out that only certain degenerate systems are (weakly) semi-norming rather than norming, so we will largely ignore the distinction. As such, our main question can be stated in the following more concrete terms.

Question 1.1. *Which linear systems L are norming or weakly norming?*

There is now a considerable body of work studying the analogue of Question 1.1 for graphs. In this setting, the question of classification was first raised by Lovász [17] and then studied in some depth by Hatami [12], who proved some necessary conditions, namely, that any weakly norming graph must be bipartite, balanced and biregular. He also showed that several families of graphs, including hypercubes, are weakly norming, a result which was greatly extended by Conlon and Lee [5], who showed that every finite reflection group gives rise to a family of weakly norming graphs. They also conjectured a converse saying that all weakly norming graphs should arise from finite reflection groups in this way, though this remains wide open, with the best result towards it being a result of Sidorenko [19] saying that weakly norming graphs must be edge-transitive.

Here we will prove a collection of necessary and sufficient conditions for a linear system to be norming or weakly norming, many of which run parallel to those in the graph setting, but often requiring significantly different techniques to prove. For instance, we prove an analogue of Sidorenko's edge-transitivity result saying that any weakly norming linear system is variable-transitive, in the sense that deleting any given variable yields an isomorphic subsystem. The proof of this requires an isomorphism theorem for the functional $t_L(\cdot)$, saying that if $t_L(f) = t_M(f)$ for all non-negative real-valued functions f, then the systems L and M are isomorphic. For graphs, the analogous statement, due to Lovász [16], is reasonably straightforward, while the arithmetic version requires significantly more work. On the other hand, the additional power given to us by the ability to apply Fourier methods means that classification, at least for low rank systems, comes within reach. In particular, we are able to classify all weakly norming systems of rank at most two.

Since we prove many results across several sections, we spell out the main contributions of each section below:

- In Section 2, we set out some preliminaries, first showing how the condition that L be (weakly) semi-norming can be rephrased in terms of satisfying a Hölder-type inequality. We then introduce some basic discrete Fourier analysis and rephrase this Hölder-type inequality in Fourier terms. We also define and note some basic properties of subsystems of linear systems.

- In Section 3, we establish some basic necessary conditions for a linear system to be weakly norming. For instance, we show that any weakly norming system must be translation invariant, so, in particular, each equation must have all coefficients summing to zero. We also show that any weakly norming system satisfies the arithmetic analogue of Sidorenko's conjecture, an important conjecture in graph theory (see, for example, [2, 4, 6, 21] on the graph case and [1, 9, 14] on the arithmetic case). It is also here that we say more about the distinction

2

between (weakly) norming and semi-norming systems and why we can, for the most part, ignore the distinction.

- In Section 4, we prove the aforementioned arithmetic analogue of the isomorphism theorem from graph theory and use it to prove that any weakly norming system must be variable-transitive. The main idea of the proof of the isomorphism theorem is to show that the condition that $t_L(f) = t_M(f)$ for all non-negative real-valued functions f lifts to a similar conclusion where f ranges over all complex-valued functions and then an appropriate choice of f allows us to derive the required conclusion. Once this is in place, we may follow Sidorenko's technique to establish variable-transitivity.

- In Section 5, we classify all weakly norming linear systems of rank at most two. For single equations, any such equation must be Schatten, meaning that it is an equation of the form $a_1 x_1 + \cdots + a_k x_k = 0$ with all $a_i \in \{a, -a\}$ and $|\{i \in [k] : a_i = a\}| = |\{i \in [k] : a_i = -a\}|$. For rank two systems, we need an additional operation, called subdivision, which takes a (weakly) norming system of linear equations and produces another (weakly) norming system with twice as many variables. The main result of this section then says that any rank two weakly norming linear system arises through subdividing some basic examples.

- In Section 6, we show that weakly norming linear systems are forcing. The Sidorenko property, established for weakly norming linear systems in Section 3, says that, of all functions f with a given density, the constant function is the one which minimises $t_L(f)$. The stronger forcing property, whose analogue for graphs is again well-studied, says that the constant function is also the *unique* function which minimises $t_L(f)$.

- In Section 7, we show how each (weakly) norming hypergraph gives rise to a (weakly) norming linear system. This is a key source of examples for us, though the examples are given through a natural parametrisation of the solution set of a set of linear equations rather than in terms of the linear equations themselves. For systems arising from graphs, we show that a simple technique can be used to find the system of equations itself, though we also explain why this technique does not straightforwardly extend to systems arising from hypergraphs.

- In Section 8, we prove the arithmetic analogue of an important result of Lee and Sidorenko [15] saying that any norming graph is also norming for complex-valued functions. In order for this to work, one must take the conjugate of some of the terms in the expression for $t_L(f)$ and the key difficulty in the proof, which closely follows that in [15], is in showing that there is some such assignment which works.

- Finally, in Section 9, we conclude with some brief further remarks on open problems.

2 Preliminaries

In this section, we collect some results which will be needed throughout the paper. We first discuss some reformulations of the norming property, beginning with the proposition below, an arithmetic analogue of a result of Hatami [12] on graph norms, which was itself inspired by Gowers' proof [11] that the uniformity norms are indeed norming. It says that the triangle inequality condition needed for a system to be norming is equivalent to a certain Hölder-type condition.

For the proof, we will need some notation. Given functions $f, g : \mathbb{F}_q^n \to \mathbb{C}$, their *tensor product* $f \otimes g$ is the function on $\mathbb{F}_q^n \times \mathbb{F}_q^n$ defined by

$$f \otimes g(y, z) = f(y)g(z).$$

Using the natural isomorphism between $(\mathbb{F}_q^n)^2$ and $\mathbb{F}_q^{2n}$, $f \otimes g$ can be seen as a function on $\mathbb{F}_q^{2n}$. Moreover, for $f_i, g_i : \mathbb{F}_q^n \to \mathbb{C}$,

$$t_L(f_1 \otimes g_1, f_2 \otimes g_2, \ldots, f_k \otimes g_k) = t_L(f_1, \ldots, f_k)t_L(g_1, \ldots, g_k).$$

To see this, let $(x_1, \ldots, x_k) \in (\mathbb{F}_q^{2n})^k$ and write each $x_i = (y_i, z_i)$ with $y_i, z_i \in \mathbb{F}_q^n$, so that $f \otimes g(x_i) = f(y_i)g(z_i)$. Then, since $(x_1, \ldots, x_k)$ satisfies $L(x_1, \ldots, x_k) = 0$ if and only if $L(y_1, \ldots, y_k) = L(z_1, \ldots, z_k) = 0$, we have

$$
\begin{aligned}
t_L(f_1 \otimes g_1, f_2 \otimes g_2, \ldots, f_k \otimes g_k) &= \mathbb{E}_{(x_1,\ldots,x_k)\in\mathrm{Sol}(L)} f_1 \otimes g_1(x_1) \cdots f_k \otimes g_k(x_k) \\
&= \mathbb{E}_{(y_1,\ldots,y_k)\in\mathrm{Sol}(L)} \mathbb{E}_{(z_1,\ldots,z_k)\in\mathrm{Sol}(L)} f_1(y_1)g_1(z_1) \cdots f_k(y_k)g_k(z_k) \\
&= t_L(f_1, \ldots, f_k)t_L(g_1, \ldots, g_k).
\end{aligned}
$$

For a positive integer m, the m-th tensor power $f^{\otimes m} : (\mathbb{F}_q^n)^m \to \mathbb{C}$ is the function given by

$$f^{\otimes m} = \underbrace{f \otimes f \otimes \cdots \otimes f}_{m \text{ times}}.$$

In particular, by iterating the multiplicativity of tensor products, we obtain

$$t_L(f_1^{\otimes m}, \ldots, f_k^{\otimes m}) = t_L(f_1, f_2, \ldots, f_k)^m.$$

Proposition 2.1. *A linear system L in k variables is semi-norming if and only if the inequality*

$$|t_L(f_1, \ldots, f_k)| \leq \prod_{i=1}^k \|f_i\|_L \tag{1}$$

holds for all $n \in \mathbb{N}$ and all $f_1, \ldots, f_k : \mathbb{F}_q^n \to \mathbb{R}$.

Proof. Suppose the functional $\| \cdot \|_L$ is a seminorm on the space of real-valued functions on $\mathbb{F}_q^n$ for all $n \in \mathbb{N}$ and the inequality (1) does not hold for some functions $f_1, \ldots, f_k$ on $G = \mathbb{F}_q^n$, i.e., $|t_L(f_1, f_2, \ldots, f_k)| > \prod_{i=1}^k \|f_i\|_L$. By renormalising, we may assume that $|t_L(f_1, f_2, \ldots, f_k)| = c > 1$ and $\|f_i\|_L \leq 1$ for each $i = 1, \ldots, k$. Then, for any even integer m,

$$
\begin{aligned}
\|f_1^{\otimes m} + \cdots + f_k^{\otimes m}\|_L^k = |t_L(f_1^{\otimes m} + \cdots + f_k^{\otimes m})| &= \left| \sum_{1 \leq i_1, \ldots, i_k \leq k} t_L(f_{i_1}^{\otimes m}, \ldots, f_{i_k}^{\otimes m}) \right| \\
&= \sum_{1 \leq i_1, \ldots, i_k \leq k} t_L(f_{i_1}, \ldots, f_{i_k})^m \geq t_L(f_1, \ldots, f_k)^m = c^m.
\end{aligned}
$$

On the other hand, $\|f_1^{\otimes m}\|_L + \cdots + \|f_k^{\otimes m}\|_L = \|f_1\|_L^m + \cdots + \|f_k\|_L^m \leq k$. Hence, if m is sufficiently large, we have

$$\|f_1^{\otimes m}\|_L + \cdots + \|f_k^{\otimes m}\|_L \leq k < c^{m/k} \leq \|f_1^{\otimes m} + \cdots + f_k^{\otimes m}\|_L,$$

which contradicts the triangle inequality for $\| \cdot \|_L$ on $\mathbb{F}_q^{nm}$.

Conversely, suppose that the inequality (1) holds. Then, for each f_1 and f_2,

$$\|f_1 + f_2\|_L^k = |t_L(f_1 + f_2)| \leq \sum_{1 \leq i_1,\ldots,i_k \leq 2} |t_L(f_{i_1}, \ldots, f_{i_k})|$$

$$\leq \sum_{1 \leq i_1,\ldots,i_k \leq 2} \|f_{i_1}\|_L \cdots \|f_{i_k}\|_L = (\|f_1\|_L + \|f_2\|_L)^k.$$

Thus, $\| \cdot \|_L$ satisfies the triangle inequality. $\qquad \square$

A similar argument also works in the weakly norming case, yielding the following result, whose proof we omit.

Proposition 2.2. *A linear system L is weakly semi-norming if and only if the inequality*

$$t_L(f_1, \ldots, f_k) \leq \prod_{i=1}^k \|f_i\|_L = \prod_{i=1}^k \|f_i\|_{r(L)}$$

holds for all $n \in \mathbb{N}$ and all non-negative $f_1, \ldots, f_k : \mathbb{F}_q^n \to \mathbb{R}_{\geq 0}$.

We will make considerable use of discrete Fourier analysis on the finite abelian groups $\mathbb{F}_q^n$. We give a brief refresher. Setting $G = \mathbb{F}_q^n$, we write $\widehat{G}$ for the dual group of G, i.e., the group of homomorphisms from G to the multiplicative group $\mathbb{C}$ of complex numbers, which is easily seen to be isomorphic to G itself.

Denoting the inner product of $\xi, x \in G$ by $\xi^T x$, the Fourier transform $\widehat{f} : \widehat{G} \to \mathbb{C}$ of a function $f : G \to \mathbb{C}$ is defined by

$$\widehat{f}(\xi) := \mathbb{E}_{x \in G} f(x) e(-\xi^T x),$$

where $e(y) := e^{2\pi i y / q}$, and we have the Fourier inversion formula

$$f(x) = \sum_{\xi \in \widehat{G}} \widehat{f}(\xi) e(\xi^T x).$$

Using these formulas, it is easily checked that f is real-valued if and only if $\widehat{f}(\xi)$ and $\widehat{f}(-\xi)$ are complex conjugates for every $\xi \in \widehat{G}$.

We write $f_1 * \cdots * f_k$ for the *convolution* $f_1 * \cdots * f_k(x) := \mathbb{E}_{y_1 + \cdots + y_k = x} f_1(y_1) \cdots f_k(y_k)$. In particular, if both f_1 and f_2 are always non-negative, then so is $f_1 * f_2$. The importance of convolution lies in the fact that the Fourier transform $(f_1 * \cdots * f_k)^\wedge$ of $f_1 * \cdots * f_k$ satisfies $(f_1 * \cdots * f_k)^\wedge = \widehat{f}_1 \cdots \widehat{f}_k$.

Let L be a $1 \times k$ linear system such that no entry of L is zero. For $g_j(y) := f_j(L_{1j}^{-1} y)$, we have

$$t_L(f_1, \ldots, f_k) = \mathbb{E}_{\mathrm{Sol}(L)} f_1(x_1) \cdots f_k(x_k) = g_1 * g_2 * \cdots * g_k(0) = \sum_{\xi \in \widehat{G}} \prod_{j=1}^k \widehat{g}_j(\xi) = \sum_{\xi \in \widehat{G}} \prod_{j=1}^k \widehat{f}_j(L_{1j}\xi).$$

We will also make use of the analogous Fourier inversion formula for linear systems with more than one equation.

Proposition 2.3. *Let $L \in \mathbb{F}_q^{m \times k}$ be a linear system with coefficients $(L_{ij})_{i \in [m], j \in [k]}$. Then*

$$t_L(f_1, \ldots, f_k) = \sum_{(\xi_1, \ldots, \xi_m) \in \widehat{G}^m} \prod_{j=1}^k \widehat{f}_j \left(\sum_{i=1}^m L_{ij}\xi_i \right) \qquad (2)$$

for any $f_1, \ldots, f_k : \mathbb{F}_q^n \to \mathbb{R}$.

Proof. Expanding out the right-hand side of (2), we get

$$\sum_{\xi\in\widehat{G}^m}\prod_{j=1}^k \widehat{f_j}\left(\sum_{i=1}^m L_{ij}\xi_i\right) = \sum_{\xi\in\widehat{G}^m}\prod_{j=1}^k\left(\mathbb{E}_{x_j\in G}f_j(x_j)e\left(-\sum_{i=1}^m L_{ij}\xi_i^T x_j\right)\right)$$

$$= \mathbb{E}_{x\in G^k}\sum_{\xi\in\widehat{G}^m} e\left(-\sum_{i=1}^m \xi_i^T\sum_{j=1}^k L_{ij}x_j\right)\prod_{j=1}^k f_j(x_j)$$

$$= \frac{1}{|G|^k}\sum_{x\in G^k}\prod_{i=1}^m\left(\sum_{\xi_i\in\widehat{G}} e\left(-\xi_i^T\sum_{j=1}^k L_{ij}x_j\right)\right)\prod_{j=1}^k f_j(x_j). \qquad (3)$$

For fixed $x\in G^k$, we have

$$\sum_{\xi_i\in\widehat{G}} e\left(-\xi_i^T\sum_{j=1}^k L_{ij}x_j\right) = \begin{cases} |G| & \text{if } \sum_{j=1}^k L_{ij}x_j = 0, \\ 0 & \text{otherwise.} \end{cases}$$

Inserting this into (3), we obtain

$$\sum_{\xi\in\widehat{G}^m}\prod_{j=1}^k \widehat{f_j}\left(\sum_{i=1}^m L_{ij}\xi_i\right) = \frac{|G|^m}{|G|^k}\sum_{x\in\text{Sol}(L)}\prod_{j=1}^k f_j(x_j) = t_L(f_1,\ldots,f_k),$$

as desired. $\qquad\qquad\square$

Thus, by Proposition 2.1, the linear system L is semi-norming if and only if

$$\left|\sum_{\vec{\xi}\in\widehat{G}^m}\widehat{f_1}\left(\sum_{i=1}^m L_{i1}\xi_i\right)\cdots\widehat{f_k}\left(\sum_{i=1}^m L_{ik}\xi_i\right)\right|^k \le \prod_{j=1}^k\left|\sum_{\vec{\xi}\in\widehat{G}^m}\widehat{f_j}\left(\sum_{i=1}^m L_{i1}\xi_i\right)\cdots\widehat{f_j}\left(\sum_{i=1}^m L_{ik}\xi_i\right)\right|$$

or, equivalently, writing L_j^t for the jth column of L as a row vector,

$$\left|\sum_{\vec{\xi}\in\widehat{G}^m}\widehat{f_1}\left(L_1^t\vec{\xi}\right)\cdots\widehat{f_k}\left(L_k^t\vec{\xi}\right)\right|^k \le \prod_{j=1}^k\left|\sum_{\vec{\xi}\in\widehat{G}^m}\widehat{f_j}\left(L_1^t\vec{\xi}\right)\cdots\widehat{f_j}\left(L_k^t\vec{\xi}\right)\right| \qquad (4)$$

for any $f_1,\ldots,f_k:\mathbb{F}_q^n\to\mathbb{R}$.

Example 2.4. Let L be the 2×6 system defined by the matrix

$$\begin{pmatrix} 1 & -1 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & -1 & 1 \end{pmatrix}.$$

In other words, $\text{Sol}(L)$ consists of solutions to the system of linear equations $x_1 - x_2 = x_3 - x_4 = x_5 - x_6$. Then, by (2),

$$t_L(f_1,\ldots,f_6) = \sum_{(\xi_1,\xi_2)\in\widehat{G}^2}\widehat{f_1}(\xi_1)\widehat{f_2}(-\xi_1)\widehat{f_3}(-\xi_1+\xi_2)\widehat{f_4}(\xi_1-\xi_2)\widehat{f_5}(-\xi_2)\widehat{f_6}(\xi_2).$$

In particular,

$$t_L(f) = \sum_{(\xi_1,\xi_2)\in\widehat{G}^2}|\widehat{f}(\xi_1)|^2|\widehat{f}(\xi_1-\xi_2)|^2|\widehat{f}(\xi_2)|^2.$$

We will see that this system L is weakly norming in Example 5.7.

6

We now define subsystems of linear systems, which bear some resemblance to subgraphs of graphs. Let L be an $m \times k$ system, let $i \in [k]$ and let $a^{(1)}, \ldots, a^{(r)}$ be a basis for those vectors in the row space of L for which the ith coordinate is zero. By *deleting a variable* x_i, $i \in [k]$, we mean that we pass to the $r \times (k-1)$ system L' with rows $\tilde{a}^{(1)}, \ldots, \tilde{a}^{(r)}$ where $\tilde{a}^{(j)}$ is equal to $a^{(j)}$ with the ith coordinate removed. Note that by choosing a different basis $a^{(1)}, \ldots, a^{(r)}$, we might obtain a different system L'. However, the solution space $\text{Sol}(L')$ is always the projection of $\text{Sol}(L)$ onto the set of coordinates $[k] \setminus \{i\}$, independently of the choice of basis. By *deleting an equation*, we mean removing the corresponding row vector from the basis for the row space of L, thereby reducing the dimension of the row space. A *subsystem* L' of an $m \times k$ linear system L is an $m' \times k'$ system with $k' \leq k$ and $m' \leq m$, obtained by deleting variables and equations from L. If a subsystem is obtained by deleting all variables indexed by $i \notin I$, then we say that it is *induced* on $I \subseteq [k]$.

Example 2.5. Let L be the 4×5 system defined by the matrix

$$\begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 1 & 1 & -2 & 0 & 0 \\ 0 & 0 & -2 & 1 & 1 \\ 0 & 0 & 0 & 1 & -1 \end{pmatrix}.$$

Then $\text{Sol}(L)$ is 1-dimensional, spanned by $(1, 1, 1, 1, 1)$. Deleting the variable x_3 from L yields the subsystem L' induced on $I = \{1, 2, 4, 5\}$ whose matrix can be written as

$$\begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{pmatrix},$$

where the last row is obtained by deleting the third entry from $(1, 1, 0, -1, -1)$, a vector in the row space of L. Thus, $\text{Sol}(L)$ is again 1-dimensional, spanned by $(1, 1, 1, 1)$.

If L' is an induced subsystem of L, we may write $t_{L'}(f)$ as $t_L(f_1, \ldots, f_k)$ for some appropriate $f_1, \ldots, f_k$. For graphs, this is obvious, but we have to be a little more careful in the arithmetic setting, so we include a proof.

Proposition 2.6. *Let L' be an $m' \times k'$ subsystem of an $m \times k$ linear system L, induced on $I \subseteq [k]$. Given $f : G \to \mathbb{C}$, let $f_i = f$ for $i \in I$ and $f_j = 1$ for $j \notin I$. Then $t_{L'}(f) = t_L(f_1, \ldots, f_k)$.*

Proof. We claim that if L' is induced on $I = [k-1]$, then, for any complex-valued functions $f_1, \ldots, f_{k-1}$ on G,

$$t_L(f_1, \ldots, f_{k-1}, 1) = t_{L'}(f_1, \ldots, f_{k-1}). \tag{5}$$

The proof then follows from repeatedly applying this claim to reduce the index set while possibly relabelling the variables.

Suppose first that there are no equations in $\text{Sol}(L)$ that involve x_k. Then x_k is a 'free' variable, so that $t_L(f_1, \ldots, g) = t_{L'}(f_1, \ldots, f_{k-1})\mathbb{E}[g]$ for every function g. For $g \equiv 1$, this yields (5). We may therefore assume that there is at least one equation in $\text{Sol}(L)$ that involves x_k.

Let $p_k : G^k \to G^{k-1}$ be the projection map $(x_1, \ldots, x_k) \mapsto (x_1, \ldots, x_{k-1})$. Then

$$t_L(f_1, \ldots, f_{k-1}, 1) = \frac{1}{|\text{Sol}(L)|} \sum_{(x_1, \ldots, x_k) \in \text{Sol}(L)} f_1(x_1) \cdots f_{k-1}(x_{k-1})$$

$$= \frac{1}{|\text{Sol}(L)|} \sum_{(x_1, \ldots, x_{k-1}) \in p_k(\text{Sol}(L))} f_1(x_1) \cdots f_{k-1}(x_{k-1}) |p_k^{-1}(x_1, \ldots, x_{k-1})|, \tag{6}$$

7

where $p_k^{-1}(\cdot)$ denotes the inverse image of p_k.

Since at least one equation of L involves x_k, p_k is a bijection between $\mathrm{Sol}(L)$ and $\mathrm{Sol}(L')$. Indeed, under p_k, each $(x_1, \ldots, x_k)$ projects to $(x_1, \ldots, x_{k-1})$, which satisfies all equations in L', proving that $p_k(\mathrm{Sol}(L)) \subseteq \mathrm{Sol}(L')$. Conversely, let $(x_1, \ldots, x_{k-1}) \in \mathrm{Sol}(L')$. It is enough to check that this uniquely extends to $(x_1, \ldots, x_k) \in \mathrm{Sol}(L)$. Given an equation $a_1 x_1 + \cdots + a_k x_k = 0$ in L with $a_k \neq 0$, $(x_1, \ldots, x_{k-1})$ uniquely determines

$$ x_k = -\frac{a_1}{a_k} x_1 - \cdots - \frac{a_{k-1}}{a_k} x_{k-1}. $$

Moreover, this choice of x_k is consistent with the one obtained by using any other equation $b_1 x_1 + \cdots + b_k x_k = 0$ with $b_k \neq 0$ since L' includes the equation

$$ \frac{a_1}{a_k} x_1 + \cdots + \frac{a_{k-1}}{a_k} x_{k-1} = \frac{b_1}{b_k} x_1 + \cdots + \frac{b_{k-1}}{b_k} x_{k-1}. $$

Therefore, $p_k(\mathrm{Sol}(L)) = \mathrm{Sol}(L')$ and p_k is a bijection from $\mathrm{Sol}(L)$ to $\mathrm{Sol}(L')$, which turns (6) into the identity (5). $\qquad\square$

3   Basic necessary conditions

A $k \times k$ linear system L with non-zero determinant satisfies $\mathrm{Sol}(L) = \{0\}$, which, for our purposes, may be seen as a degenerate case. As such, we shall assume $\mathrm{Sol}(L) \neq \{0\}$ in what follows, unless we are specifically speaking about degenerate systems. The following proposition asserts that all other weakly semi-norming systems are translation invariant, in the sense that if $(x_1, \ldots, x_k) \in \mathrm{Sol}(L)$ and $g \in G$, then $(x_1 + g, \ldots, x_k + g) \in \mathrm{Sol}(L)$ as well. In particular, this means that the coefficients of each equation in a weakly norming system must add to 0.

Proposition 3.1. *Every weakly semi-norming $m \times k$ system L is translation invariant. In particular, for all $i \in [k]$ and $v \in G$,*

$$ |\{(x_1, \ldots, x_k) \in \mathrm{Sol}(L) : x_i = v\}| = |\mathrm{Sol}(L)|/|G|. $$

Proof. Let $(y_1, \ldots, y_k) \in \mathbb{F}_q^k$ be a non-zero element in $\mathrm{Sol}(L)$ with $y_j = v \neq 0$. For $f_j = \mathbf{1}_{\{v\}}$ and $f_i = 1$ for all $i \neq j$, we have

$$ t_L(f_1, \ldots, f_k) = \frac{|\{(x_1, \ldots, x_k) \in \mathrm{Sol}(L) : x_j = v\}|}{|\mathrm{Sol}(L)|}, $$

which is strictly positive. By Proposition 2.2, it follows that $t_L(f_j) > 0$. That is, $\mathrm{Sol}(L)$ contains $(v, v, \ldots, v)$ and, hence, it also contains $(1, 1, \ldots, 1)$. Thus, L must be translation invariant. $\qquad\square$

By this proposition, if L is a non-degenerate weakly norming system, then each $(v, v, \ldots, v) \in \mathrm{Sol}(L)$, so we have that

$$ \|f\|_{r(L)} \geq \frac{1}{|\mathrm{Sol}(L)|} \sum_v |f(v)|^k > 0 $$

for all $f \neq 0$, implying that the system is in fact weakly norming.

Another consequence of this proposition is that $\mathbb{E}[f]^k \leq t_L(f)$ whenever L is an $m \times k$ weakly norming system and f is a non-negative function on G. Indeed, let $f_1 = f$ and $f_2 = \cdots = f_k = 1$. Then

$$ t_L(f_1, \ldots, f_k) = \mathbb{E}_{(x_1, \ldots, x_k) \in \mathrm{Sol}(L)} f(x_1) = \mathbb{E}_{x \in G} f(x) $$

8

and, therefore, Proposition 2.2 implies that $\mathbb{E}[f]^k \leq t_L(f)$. As mentioned in the introduction, linear systems which satisfy this inequality for all non-negative f are called *Sidorenko*.

Corollary 3.2. *Every weakly norming system L is Sidorenko.*

We note in passing that, since Sidorenko systems are necessarily translation invariant (see, for instance, [14]), this corollary again implies that weakly norming systems are translation invariant.

By using Proposition 2.6, we can also generalise Corollary 3.2 to say that every weakly norming system dominates all of its subsystems in the following sense (see [7] for more on the analogous property for graphs).

Corollary 3.3. *Let L' be a subsystem of an $m \times k$ weakly norming system L induced on $I \subseteq [k]$ with $|I| = k'$. Then, for every non-negative function $f : G \to \mathbb{R}_{\geq 0}$,*

$$t_L(f)^{1/k} \geq t_{L'}(f)^{1/k'}.$$

Furthermore, if L is norming, then $|t_L(f)|^{1/k} \geq |t_{L'}(f)|^{1/k'}$ for all $f : G \to \mathbb{R}$.

The *girth* of a linear system L is the minimum size of the support of a non-zero vector in the row space of L. We will now show that the girth of any Sidorenko system, and hence any weakly norming system, is even. This result was previously proved by Kamčev, Liebenau and Morrison [14], though we give a somewhat shorter proof using Fourier methods.

Proposition 3.4. *The girth of a Sidorenko system is even.*

Proof. Let L be an $m \times k$ system whose girth ℓ is odd. Suppose, without loss of generality, that one of the non-zero vectors in the row space of L with minimum support has the form $(a_1, \ldots, a_\ell, 0, \ldots, 0)$ for non-zero $a_1, \ldots, a_\ell \in \mathbb{F}_q$. Let $\alpha \in (0,1)$ and γ be a non-zero element of $\widehat{G}$ and consider the function $f_\alpha \colon G \to \mathbb{R}$ with Fourier transform

$$\widehat{f_\alpha}(\xi) = \begin{cases} 1 & \text{if } \xi = 0, \\ -\alpha & \text{if } \xi = \pm a_i \gamma, \\ 0 & \text{otherwise,} \end{cases}$$

noting that if $\alpha \leq 1/2\ell$, then f_α takes values in $[0, 2]$. By Proposition 2.3, we have

$$t_L(f_\alpha) = \sum_{\xi \in \widehat{G}^m} \prod_{j=1}^{k} \widehat{f_\alpha}\left(L_j^t \xi\right) \leq 1 + (-\alpha)^\ell + O(\alpha^{\ell+1}).$$

If α is sufficiently small, then this is less than 1, but, because $\mathbb{E}[f_\alpha] = \widehat{f_\alpha}(0) = 1$, this means that L is not Sidorenko. $\qquad\square$

Recall that if L is an $m \times k$ norming system, then $\|f\|_L$ equals $|t_L(f)|^{1/k}$. The following proposition shows that we do not need the modulus.

Proposition 3.5. *If L is a norming system, then $t_L(f) > 0$ for all $f \neq 0$. In particular, L must have an even number of variables.*

Proof. Suppose, for the sake of contradiction, that there is some non-zero f for which $t_L(f) < 0$. Let v_1 and v_2 be two distinct elements of G. By the translation invariance of L, $t_L(\mathbf{1}_{\{v_1\}})$ and $t_L(\mathbf{1}_{\{v_2\}})$ are both positive. Hence, by the continuity of t_L and the intermediate value theorem, there are $c_1, c_2 \in \mathbb{R}_{>0}$ such that $t_L(f + c_1 \mathbf{1}_{\{v_1\}}) = t_L(f + c_2 \mathbf{1}_{\{v_2\}}) = 0$. However, since L is norming, this implies that $f + c_1 \mathbf{1}_{\{v_1\}} = f + c_2 \mathbf{1}_{\{v_2\}} = 0$, which cannot be the case as $\mathbf{1}_{\{v_1\}}$ and $\mathbf{1}_{\{v_2\}}$ are clearly linearly independent. $\qquad\square$

Finally, we show that the only systems which are semi-norming but not norming are the zero matrices.

Proposition 3.6. *If a system L is semi-norming but not norming, then it must be the zero matrix.*

Proof. There must exist $i \in [k]$ such that if $x \in \mathrm{Sol}(L)$ satisfies $x_j = 0$ for all $j \neq i$, then $x_i = 0$ too. Indeed, otherwise L is the zero matrix and we are done. By the translation invariance of L, the same is true if we replace 0 by any other $v \in G$, i.e., if $x_j = v$ for all $j \neq i$, then $x_i = v$ too. Without loss of generality, we may assume that $i = 1$. Suppose now that f is a function with $\|f\|_L = 0$. By Proposition 2.1, we have that

$$|t_L(f, \mathbf{1}_{\{v\}}, \ldots, \mathbf{1}_{\{v\}})| \leq \|f\|_L/|\mathrm{Sol}(L)|^{k-1} = 0$$

for all v. On the other hand, we know that the only solution that contributes to $t_L(f, \mathbf{1}_{\{v\}}, \ldots, \mathbf{1}_{\{v\}})$ is $(v, \ldots, v)$, meaning that $|f(v)| = |t_L(f, \mathbf{1}_{\{v\}}, \ldots, \mathbf{1}_{\{v\}})||\mathrm{Sol}(L)| = 0$. Since this holds for all v, we have $f = 0$. $\qquad\square$

4 The isomorphism theorem

For graphs, it is a fundamental fact (see, for example, [18, Theorem 5.29]) that if $|\mathrm{Hom}(H_1, F)| = |\mathrm{Hom}(H_2, F)|$ for every graph F, then H_1 and H_2 must be isomorphic. In this section, we prove the arithmetic analogue of this (left-)isomorphism theorem and apply it to give some further necessary conditions for a linear system to be weakly norming.

To state the result, we should first define what isomorphism means for linear systems: two $m \times k$ linear systems L and M are *isomorphic* if M can be obtained by applying row operations and column permutations to L. That is, L and M are isomorphic if the vector space $\mathrm{Sol}(M)$ can be obtained from $\mathrm{Sol}(L)$ by simply permuting the indices of the variables.

Theorem 4.1. *Two $m \times k$ linear systems L and M are isomorphic if and only if $t_L(f) = t_M(f)$ for all positive integers n and all non-negative functions $f : \mathbb{F}_q^n \to \mathbb{R}_{\geq 0}$.*

There are two main steps in the proof. The first is encapsulated in the following lemma, which gives us much more flexibility in substituting various functions f into $t_L(f)$. We define the *symmetrised functional* $\tau_L(\cdot)$ by

$$\tau_L(f_1, \ldots, f_k) := \sum_{\pi \in S_k} t_L(f_{\pi(1)}, \ldots, f_{\pi(k)})$$

for any functions $f_1, \ldots, f_k : \mathbb{F}_q^n \to \mathbb{C}$, where S_k denotes the set of all permutations of $[k]$.

Lemma 4.2. *Let L and M be two $m \times k$ linear systems. Then the following are equivalent:*

   (i) $t_L(f) = t_M(f)$ for every non-negative function $f : \mathbb{F}_q^n \to \mathbb{R}_{\geq 0}$;

  (ii) $t_L(f) = t_M(f)$ for every real-valued function $f : \mathbb{F}_q^n \to \mathbb{R}$;

 (iii) $\tau_L(f_1, \ldots, f_k) = \tau_M(f_1, \ldots, f_k)$ for any complex-valued functions $f_1, \ldots, f_k$ on $\mathbb{F}_q^n$.

Proof. Since the implication (iii)$\Rightarrow$(i) is trivial, it will suffice to show that (i)$\Rightarrow$(ii) and (ii)$\Rightarrow$(iii). We first show that (i)$\Rightarrow$(ii). Let $\varepsilon_1, \ldots, \varepsilon_k$ be i.i.d. uniform random variables taking values in $\{\pm 1\}$ and set $\vec{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_k)$. For a real-valued function f on $\mathbb{F}_q^n$,

$$\mathbb{E}_{\vec{\varepsilon}}\left[\varepsilon_1 \varepsilon_2 \cdots \varepsilon_k t_L\left(\sum_{i=1}^{k}(|f| + \varepsilon_i f)\right)\right] = k! \, t_L(f).$$

Indeed, expanding $t_L \left(\sum_{i=1}^k (|f| + \varepsilon_i f) \right)$ yields a sum of terms of the form $t_L(h_1, h_2, \ldots, h_k)$ where each h_i is either $|f|$ or $\varepsilon_j f$ for some j. Using $\mathbb{E}[\varepsilon_j] = 0$ and $\mathbb{E}[\varepsilon_j^2] = 1$, we see that the only terms that do not vanish after averaging over $\vec{\varepsilon} \in \{\pm 1\}^k$ are those $t_L(h_1, h_2, \ldots, h_k)$ such that each $h_i = \varepsilon_j f$ for a unique j. This is exactly what the right-hand side represents. As $\sum_{i=1}^k (|f| + \varepsilon_i f) \geq 0$ for any real-valued f, we have that (i)$\Rightarrow$(ii), as required.

To show that (ii)$\Rightarrow$(iii), let $z_1, \ldots, z_k$ be i.i.d. uniform random unit complex numbers and set $\vec{z} = (z_1, \ldots, z_k)$. For complex-valued functions $f_1, \ldots, f_k$ on $\mathbb{F}_q^n$,

$$\mathbb{E}_{\vec{z}} \left[z_1 z_2 \cdots z_k t_L \left(\sum_{i=1}^k (z_i \overline{f_i} + \overline{z_i} f_i) \right) \right] = \sum_{\pi \in S_k} t_L(f_{\pi(1)}, \ldots, f_{\pi(k)}).$$

Indeed, we can use that $\mathbb{E}[z_i] = \mathbb{E}[z_i^2] = 0$ and $\mathbb{E}[z_i \overline{z_i}] = 1$ to deduce that all terms except those of the form $t_L(f_{\pi(1)}, \ldots, f_{\pi(k)})$ vanish in the expansion of $t_L \left(\sum_{i=1}^k (z_i \overline{f_i} + \overline{z_i} f_i) \right)$ after averaging over $\vec{z} \in (S^1)^k$. As $\sum_{i=1}^k (z_i \overline{f_i} + \overline{z_i} f_i)$ is always real-valued, we have that (ii)$\Rightarrow$(iii), as required. $\qquad\square$

The second step involves substituting suitable functions $\widehat{g_1}, \widehat{g_2}, \ldots, \widehat{g_k}$ into the inversion formula (2) to distinguish $t_L(\cdot)$ and $t_M(\cdot)$ if L and M are not isomorphic. We first record a technical lemma that will be needed in our computation.

Lemma 4.3. *Let L and M be $m \times k$ linear systems and let $\gamma \in \widehat{G}^m$. For each $j = 1, 2, \ldots, k$, let $f_j^{(\gamma)}$ be the function with Fourier transform $\widehat{f_j^{(\gamma)}}(\eta) := \mathbf{1}\left[\eta = L_j^t \gamma \right]$. Then*

$$t_M(f_1^{(\gamma)}, \ldots, f_k^{(\gamma)}) = \sum_{\xi \in \widehat{G}^m} \mathbf{1}\left[M^t \xi = L^t \gamma \right].$$

Proof. By (2), it follows that

$$t_M(f_1^{(\gamma)}, \ldots, f_k^{(\gamma)}) = \sum_{\xi \in \widehat{G}^m} \prod_{j=1}^k \widehat{f_j^{(\gamma)}}(M_j^t \xi) = \sum_{\xi \in \widehat{G}^m} \prod_{j=1}^k \mathbf{1}\left[M_j^t \xi = L_j^t \gamma \right],$$

which proves the desired identity. $\qquad\square$

Proposition 4.4. *Let L and M be two non-isomorphic $m \times k$ linear systems. Then there exist $n \in \mathbb{N}$ and complex-valued functions $g_1, \ldots, g_k$ on $\mathbb{F}_q^n$ such that $\tau_L(g_1, \ldots, g_k) \neq \tau_M(g_1, \ldots, g_k)$.*

Proof. Suppose $G = \widehat{G} = \mathbb{F}_q^n$ and, for each $\gamma \in \widehat{G}^m$, let $f_1^{(\gamma)}, \ldots, f_k^{(\gamma)}$ be defined as in Lemma 4.3. Then, for $\pi \in S_k$,

$$t_M(f_{\pi(1)}^{(\gamma)}, \ldots, f_{\pi(k)}^{(\gamma)}) = \sum_{\xi \in \widehat{G}^m} \mathbf{1}\left[M^t \xi = \pi(L)^t \gamma \right],$$

where $\pi(L)$ denotes the linear system obtained by permuting the k columns of L under π. For $\Gamma = (\gamma_1, \ldots, \gamma_\ell) \in (\widehat{G}^m)^\ell$, let $g_i^{(\Gamma)} = f_i^{(\gamma_1)} \otimes f_i^{(\gamma_2)} \otimes \cdots \otimes f_i^{(\gamma_\ell)}$, where ℓ will be chosen later. Then

$$t_M(g_1^{(\Gamma)}, g_2^{(\Gamma)}, \ldots, g_k^{(\Gamma)}) = t_M(f_1^{(\gamma_1)}, \ldots, f_k^{(\gamma_1)}) t_M(f_1^{(\gamma_2)}, \ldots, f_k^{(\gamma_2)}) \cdots t_M(f_1^{(\gamma_\ell)}, \ldots, f_k^{(\gamma_\ell)})$$

and, therefore,

$$\sum_{\Gamma\in(\widehat{G}^m)^\ell} t_M(g_1^{(\Gamma)}, g_2^{(\Gamma)}, \ldots, g_k^{(\Gamma)}) = \sum_{\gamma_1,\ldots,\gamma_\ell\in\widehat{G}^m} t_M(f_1^{(\gamma_1)},\ldots,f_k^{(\gamma_1)})\cdots t_M(f_1^{(\gamma_\ell)},\ldots,f_k^{(\gamma_\ell)})$$

$$= \left(\sum_{\gamma\in\widehat{G}^m} t_M(f_1^{(\gamma)},\ldots,f_k^{(\gamma)})\right)^\ell$$

$$= \left(\sum_{\gamma\in\widehat{G}^m}\sum_{\xi\in\widehat{G}^m} \mathbf{1}\big[M^t\xi = L^t\gamma\big]\right)^\ell.$$

Since both M and L have linearly independent rows, M^t and L^t are injective, so the double sum above is $|\operatorname{im}(L^t)\cap\operatorname{im}(M^t)|^n$, where $\operatorname{im}(L^t)$ and $\operatorname{im}(M^t)$ are the images of $\mathbb{F}_q^m$ under L^t and M^t, respectively. Likewise, we have

$$\sum_{\Gamma\in(\widehat{G}^m)^\ell} t_M(g_{\pi(1)}^{(\Gamma)},\ldots,g_{\pi(k)}^{(\Gamma)}) = |\operatorname{im}(\pi(L)^t)\cap\operatorname{im}(M^t)|^{n\ell},$$

$$\sum_{\Gamma\in(\widehat{G}^m)^\ell} t_L(g_{\pi(1)}^{(\Gamma)},\ldots,g_{\pi(k)}^{(\Gamma)}) = |\operatorname{im}(\pi(L)^t)\cap\operatorname{im}(L^t)|^{n\ell}.$$

If L and M are non-isomorphic, then, for all π, the rows of $\pi(L)$ and M span different spaces, so that $\operatorname{im}(\pi(L)^t)\cap\operatorname{im}(M^t)$ has dimension at most $m-1$. Thus,

$$\sum_{\Gamma\in(\widehat{G}^m)^\ell} \tau_M(g_1^{(\Gamma)},\ldots,g_k^{(\Gamma)}) = \sum_{\pi\in S_k}\sum_{\Gamma\in(\widehat{G}^m)^\ell} t_M(g_{\pi(1)}^{(\Gamma)},\ldots,g_{\pi(k)}^{(\Gamma)}) \le k! q^{n(m-1)\ell}.$$

On the other hand, $t_L(g_1^{(\Gamma)},\ldots,g_k^{(\Gamma)}) = |\operatorname{im} L^t|^{n\ell} = q^{nm\ell}$, so, as $t_L(g_{\pi(1)}^{(\Gamma)},\ldots,g_{\pi(k)}^{(\Gamma)}) \ge 0$ for all π, we have

$$\sum_{\Gamma\in(\widehat{G}^m)^\ell} \tau_L(g_1^{(\Gamma)},\ldots,g_k^{(\Gamma)}) \ge \sum_{\Gamma\in(\widehat{G}^m)^\ell} t_L(g_1^{(\Gamma)},\ldots,g_k^{(\Gamma)}) \ge q^{nm\ell}.$$

Therefore, choosing ℓ such that $q^{n\ell} > k!$ gives

$$\sum_{\Gamma\in(\widehat{G}^m)^\ell} \tau_M(g_1^{(\Gamma)},\ldots,g_k^{(\Gamma)}) < \sum_{\Gamma\in(\widehat{G}^m)^\ell} \tau_L(g_1^{(\Gamma)},\ldots,g_k^{(\Gamma)}),$$

so there must exist $\Gamma\in(\widehat{G}^m)^\ell$ such that $g_1^{(\Gamma)},\ldots,g_k^{(\Gamma)}$ are as in the claim. $\qquad\square$

Theorem 4.1 now follows immediately. Indeed, if $t_L(f) = t_M(f)$ for every non-negative f, then Lemma 4.2 implies that $\tau_L(f_1,\ldots,f_k) = \tau_M(f_1,\ldots,f_k)$ for any complex-valued functions $f_1,\ldots,f_k$, contradicting the conclusion of Proposition 4.4 if L and M are not isomorphic.

Theorem 4.1 has some interesting applications. First, we use it to prove that every weakly norming system L is *variable-transitive* in the sense that deleting any variable always leaves an isomorphic system. This is the arithmetic analogue, and closely follows the proof, of a result of Sidorenko [19, Lemma 6] saying that if H is weakly norming, then all of the graphs formed by deleting a single edge from H are isomorphic. We note that this is slightly weaker than Sidorenko's

full edge-transitivity result, which requires an additional argument that we do not currently see how to transfer to the arithmetic setting. We do however expect that such a result should hold, that is, that for any weakly norming system and any two variables x_i and x_j there is an automorphism of the system that sends x_i to x_j.

Lemma 4.5. *Let L be an $m \times k$ weakly norming system. For a non-negative function $f : G \to \mathbb{R}_{\geq 0}$, let $\alpha_i = t_L(f_1, \ldots, f_k)$, where $f_i = 1$ and $f_j = f$ for all $j \neq i$. Then $\alpha_i = \alpha_j$ for all $i, j \in [k]$.*

Proof. Since $t_L(\cdot)$ and all α_i depend continuously on f and strictly positive functions are a dense subset of the non-negative functions, we may assume that $f > 0$. Furthermore, by rearranging edge indices, it is enough to prove that $\alpha_1 = \alpha_2$. For $\varepsilon \in \mathbb{R}$ with $|\varepsilon| \leq \min_{x \in G} f(x)$,

$$t_L(f - \varepsilon, f + \varepsilon, f, f, \ldots, f) = t_L(f) + (\alpha_2 - \alpha_1)\varepsilon - t_L(1, 1, f, \ldots, f)\varepsilon^2$$

and, therefore,

$$t_L(f - \varepsilon, f + \varepsilon, f, f, \ldots, f)^k = t_L(f)^k + k(\alpha_2 - \alpha_1)t_L(f)^{k-1}\varepsilon + O(\varepsilon^2). \tag{7}$$

At the same time, we know from Proposition 2.2 that

$$t_L(f - \varepsilon, f + \varepsilon, f, f, \ldots, f) \leq \|f\|_L^{k-2}\|f - \varepsilon\|_L\|f + \varepsilon\|_L.$$

Moreover, $t_L(f \pm \varepsilon) = t_L(f) \pm \alpha\varepsilon + O(\varepsilon^2)$ for $\alpha := \alpha_1 + \cdots + \alpha_k$, so that

$$
\begin{aligned}
t_L(f - \varepsilon, f + \varepsilon, f, f, \ldots, f)^k &\leq t_L(f)^{k-2}t_L(f - \varepsilon)t_L(f + \varepsilon) \\
&= t_L(f)^{k-2}(t_L(f) - \alpha\varepsilon + O(\varepsilon^2))(t_L(f) + \alpha\varepsilon + O(\varepsilon^2)) \\
&= t_L(f)^k + O(\varepsilon^2).
\end{aligned}
$$

If $\alpha_2 \neq \alpha_1$, then (7) contradicts this inequality for ε of the same sign as $\alpha_2 - \alpha_1$ and sufficiently small in absolute value. $\qquad\square$

Given an $m \times k$ linear system L and $i \in [k]$, we let $L \setminus i$ be the induced subsystem obtained by deleting the variable x_i. Combining Lemma 4.5 with Proposition 2.6 implies that $t_{L \setminus i}(f) = t_{L \setminus j}(f)$ for all functions $f : \mathbb{F}_q^n \to \mathbb{R}_{\geq 0}$, which, by Theorem 4.1, implies the promised analogue of Sidorenko's edge-transitivity theorem.

Corollary 4.6. *If L is an $m \times k$ weakly norming linear system, then the linear subsystems $L \setminus i$ are isomorphic for all $i \in [k]$.*

As an immediate consequence of this corollary, we have the following result.

Corollary 4.7. *If a non-zero system L is weakly norming, then no column of L is zero.*

Another application of Theorem 4.1 is that it allows us to prove an analogue of [10, Theorem 1.2], which says that, aside from isolated vertices, each component of a weakly norming graph is isomorphic. Given an $m \times k$ linear system L, suppose that there exists a partition $I_1 \cup I_2 \cup \cdots \cup I_r$ of $[k]$ such that $\mathrm{Sol}(L) = \mathrm{Sol}(L_1) \oplus \cdots \oplus \mathrm{Sol}(L_r)$, where L_j is the induced subsystem on I_j. Then $t_L(f) = t_{L_1}(f) \cdots t_{L_r}(f)$ for every $f : G \to \mathbb{R}$. If there is no finer partition that decomposes $\mathrm{Sol}(L)$, then each L_i is said to be a *component* of L.

Theorem 4.8. *Let L_1 and L_2 be two components of an $m \times k$ weakly norming system L. Then L_1 and L_2 are isomorphic weakly norming systems. Furthermore, each component is norming if L is.*

We omit the proof, which uses the domination inequality in Corollary 3.3, but is otherwise exactly the same as the simplified proof of [10, Theorem 1.2] given in [7, Lemma 2.4].

5 Weakly norming systems of low rank

By Proposition 2.3, if L is a $1 \times k$ system, then

$$t_L(f_1, \ldots, f_k) = \sum_{\xi \in \widehat{G}} \widehat{f}_1(L_{11}\xi) \cdots \widehat{f}_k(L_{1k}\xi).$$

In particular, if each L_{1i} is either a or $-a$ for some non-zero $a \in \mathbb{F}_q$ and $|\{i \in [k] : L_{1i} = a\}| = |\{i \in [k] : L_{1i} = -a\}|$, then $t_L(f) = \|\widehat{f}\|_{\ell^k}^k$. As this ℓ^k-norm of the spectrum is often called the *Schatten–von Neumann norm*, we say that a vector $(a_1, \ldots, a_k)$ is *Schatten* if there exists $a \in \mathbb{F}_q$ such that each $a_i \in \{0, a, -a\}$ and $|\{i \in [k] : a_i = a\}| = |\{i \in [k] : a_i = -a\}|$. In particular, as we have seen above, if a Schatten vector $(a_1, \ldots, a_k)$ has no zero entries, then $(a_1, \ldots, a_k)$ is norming as a linear system. One of the main results of this section is that all weakly norming linear systems must have at least one Schatten vector in their row space.

To state this result, recall that the *girth* of a linear system L is the minimum size of the support of a non-zero vector in the row space of L. Furthermore, we write $\mu(L)$ for the set of non-zero vectors in the row space of L with minimum support and $s(L)$ for the number of normalised Schatten vectors in $\mu(L)$, i.e., Schatten vectors whose first non-zero coefficient is 1.

Theorem 5.1. *Let L be an $m \times k$ weakly norming system and let ℓ be the girth of L. Then either all vectors in $\mu(L)$ are Schatten or $s(L) \geq k/\ell$.*

Proof. Since L is weakly norming, L is Sidorenko by Corollary 3.2 and, therefore, ℓ is even by Proposition 3.4. Suppose that $\mu(L)$ contains a vector v that is not Schatten. Without loss of generality, we may assume that v is of the form $(a_1, \ldots, a_\ell, 0 \ldots, 0)$ for some non-zero $a_1, \ldots, a_\ell$.

Suppose first that the characteristic of $\mathbb{F}_q$ is not 2. Let $\varepsilon > 0$, let γ be a non-zero element of $\widehat{G}$ and let z be a complex number with $|z| = 1$. For each $i \in [\ell]$, take f_i to be the function whose Fourier transform is given by

$$\widehat{f}_i(\eta) = \begin{cases} 1 & \text{if } \eta = 0, \\ \varepsilon z & \text{if } \eta = (-1)^i a_i \gamma, \\ \varepsilon \overline{z} & \text{if } \eta = (-1)^{i+1} a_i \gamma, \\ 0 & \text{otherwise.} \end{cases}$$

Observe that since $\widehat{f}_i(\eta) = \overline{\widehat{f}_i(-\eta)}$ for all η, f_i is real. Furthermore, if ε is sufficiently small, f_i is positive.

By Proposition 2.3, we have

$$t_L(f_1, \ldots, f_\ell, 1, \ldots, 1) = \sum_{\xi \in \widehat{G}^m} \prod_{j=1}^{\ell} \widehat{f}_j\left(L_j^t \xi\right) \prod_{j=\ell+1}^{k} \mathbf{1}[L_j^t \xi = 0].$$

Since ℓ is the girth of L, if $\xi \neq 0$ and $L_j^t \xi = 0$ for each $j > \ell$, then we must have $L_j^t \xi \neq 0$ for all $j \leq \ell$. Thus, the only non-zero ξ that make a non-zero contribution to the sum above are those for which $L_j^t \xi \in \{\pm a_j \gamma\}$ for each $j \leq \ell$. Since v is in the row space of L, there must be $\xi \in \widehat{G}^m$ such that $(L^t \xi)_j = v_j \gamma$ for all $i \in [k]$. Furthermore, the only other vectors in the row space of L whose support is in $[\ell]$ are the multiples of v. Indeed, if there were a vector $w \in \mu(L)$ with support $[\ell]$ that is not a multiple of v, there would be a non-zero linear combination of v and w whose support is a proper subset of $[\ell]$. Thus, if $\prod_{j=1}^{\ell} \widehat{f}_j\left(L_j^t \xi\right) \neq 0$, then $L^t \xi = \pm(a_1 \gamma, \ldots, a_\ell \gamma, 0, \ldots, 0)$, so that

$$t_L(f_1, \ldots, f_\ell, 1, \ldots, 1) = 1 + 2\varepsilon^\ell.$$

14

Taking the kth power, we obtain

$$t_L(f_1, \ldots, f_\ell, 1, \ldots, 1)^k \geq 1 + 2k\varepsilon^\ell. \tag{8}$$

On the other hand, by Proposition 2.1, we have $t_L(f_1, \ldots, f_\ell, 1, \ldots, 1)^k \leq \|f_1\|_L^k \cdots \|f_\ell\|_L^k$, where each $\|f_i\|_L^k$ is equal to

$$t_L(f_i, \ldots, f_i) = \sum_{\xi \in \widehat{G}^m} \prod_{j=1}^k \widehat{f_i}\left(L_j^t \xi\right) = 1 + \sum_{\xi \in X_i} \prod_{j=1}^k \widehat{f_i}\left(L_j^t \xi\right) + O(\varepsilon^{\ell+1}) \tag{9}$$

with

$$X_i = \{\xi \in \widehat{G}^m \colon L_j^t \xi \in \{0, \pm a_i \gamma\} \text{ for all } j \in [k] \text{ and } |\{j \in [k] : L_j^t \xi \neq 0\}| = \ell\}.$$

For $\psi \in \widehat{G}^k$, let

$$\Delta_i(\psi) = |\{j \in [k] : \psi_j = (-1)^i a_i \gamma\}| - |\{j \in [k] : \psi_j = (-1)^{i+1} a_i \gamma\}|$$

and observe, by the definition of f_i, that

$$\sum_{\xi \in X_i} \prod_{j=1}^k \widehat{f_i}\left(L_j^t \xi\right) = \sum_{\xi \in X_i} \varepsilon^\ell z^{\Delta_i(L^t \xi)} = \sum_{\xi \in X_1} \varepsilon^\ell z^{\Delta_1(L^t \xi)},$$

where the last inequality holds because $X_{i_1} = a_{i_1} a_{i_2}^{-1} X_{i_2}$ for all $i_1, i_2 \in [\ell]$. Let

$$S_1 = \{\xi \in X_1 : \Delta_1(L^t \xi) = 0\}$$

and note that each normalised Schatten vector $v \in \mu(L)$ gives rise to exactly two elements of S_1. Indeed, if $u \in \mathbb{F}_q^m$ is such that $L^t u = v$, then $\pm(a_i u_1 \gamma, \ldots, a_i u_m \gamma) \in S_1$.

On the other hand, since

$$\int_{|z|=1} \sum_{\xi \in X_1 \setminus S_1} z^{\Delta_1(L^t \xi)} \mathrm{d}z = \sum_{\xi \in X_1 \setminus S_1} \int_{|z|=1} z^{\Delta_1(L^t \xi)} \mathrm{d}z = 0,$$

we may choose z such that the real part of $\sum_{\xi \in X_1 \setminus S_1} z^{\Delta_1(L^t \xi)}$ is zero. Inserting this into (9) yields

$$t_L(f_i, \ldots, f_i) = 1 + \sum_{\xi \in S_1} \prod_{j=1}^k \widehat{f_i}\left(L_j^t \xi\right) + O(\varepsilon^{\ell+1}) = 1 + 2s(L)\varepsilon^\ell + O(\varepsilon^{\ell+1}).$$

Noting that the right-hand side does not depend on i, we can combine this with (8) to obtain

$$1 + 2k\varepsilon^\ell \leq t_L(f_1, \ldots, f_\ell, 1, \ldots, 1)^k \leq \|f_1\|_L^k \cdots \|f_\ell\|_L^k \leq 1 + 2\ell s(L)\varepsilon^\ell + O(\varepsilon^{\ell+1}).$$

By taking ε sufficiently small, we see that $s(L) \geq k/\ell$, as desired.

If $\mathbb{F}_q$ has characteristic 2, the proof is essentially the same, except that we choose $z = 1$, noting that $S_1 = X_1$ trivially, as $a_i = -a_i$ for all $i \in [\ell]$. $\qquad\square$

As a corollary of Theorem 5.1, we obtain a classification of all weakly norming systems consisting of a single equation.

Theorem 5.2. *A $1 \times k$ system L is weakly norming if and only if it is Schatten with no zero entries.*

Proof. We have already seen that such a system is norming. The opposite direction follows from first applying Corollary 4.7, which tells us that all entries are non-zero, and then Theorem 5.1. $\square$

Another corollary of Theorem 5.1 is that there are at least two distinct Schatten vectors in $\mu(L)$ unless L consists of a single equation.

Corollary 5.3. *Let L be an $m \times k$ weakly norming system with $m \geq 2$. Then there are at least two independent Schatten vectors in $\mu(L)$.*

Proof. By Theorem 5.1, it is enough to show that $\ell < k$, where ℓ is the girth of L. Assume instead that $k = \ell$. Then every vector in the row space of L has support of size $k = \ell$. However, if $m \geq 2$, by using row operations, one can clearly make a vector with smaller support, contradicting that the system has girth k. $\square$

Given an $m \times k$ system L, consider the $m \times 2k$ system obtained by replacing each column v of the matrix L with two columns v and $-v$. We call this new linear system the *subdivision* of L and denote it by $\mathrm{sub}(L)$. For instance, subdividing the equation $x_1 = x_2$ that defines the L^2-norm $(\mathbb{E}_{x \in G} f(x)^2)^{1/2}$ gives the equation $x_1 - x_2 + x_3 - x_4 = 0$ that defines the Gowers U^2-norm. More generally, as we now show, subdivision preserves the (weakly) norming property (though it is perhaps worth stressing that this is not the case for graphs).

Proposition 5.4. *Let L be a weakly norming (resp. norming) system. Then the subdivision $\mathrm{sub}(L)$ of L is also weakly norming (resp. norming).*

Proof. Let L be an $m \times k$ norming system. For brevity, write $M = \mathrm{sub}(L)$. Our goal is to show that

$$\left| \sum_{\xi \in \widehat{G}^m} \widehat{f_1}\left(M_1^t \xi\right) \cdots \widehat{f_{2k}}\left(M_{2k}^t \xi\right) \right|^{2k} \leq \prod_{j=1}^{2k} \left| \sum_{\xi \in \widehat{G}^m} \widehat{f_j}\left(M_1^t \xi\right) \cdots \widehat{f_j}\left(M_{2k}^t \xi\right) \right| \tag{10}$$

for the $m \times 2k$ system M and any real-valued functions $f_1, \ldots, f_{2k}$. For each $j \in [k]$, let g_j be the function on G that satisfies $\widehat{g_j}(\xi) = \widehat{f_{2j-1}}(\xi)\overline{\widehat{f_{2j}}(\xi)}$. Then each g_j is real-valued, since $\widehat{g_j}(-\xi) = \widehat{f_{2j-1}}(-\xi)\overline{\widehat{f_{2j}}(-\xi)} = \overline{\widehat{g_j}(\xi)}$. Furthermore, for all $i, j \in [k]$,

$$\widehat{g_j}(L_i^t \xi) = \widehat{f_{2j-1}}(L_i^t \xi)\overline{\widehat{f_{2j}}(L_i^t \xi)} = \widehat{f_{2j-1}}(M_{2i-1}^t \xi)\overline{\widehat{f_{2j}}(M_{2i}^t \xi)},$$

where the second equality follows from the definition of $\mathrm{sub}(L)$.

Rewriting the left-hand side of (10) in terms of the g_j and applying (4) gives

$$\left| \sum_{\xi \in \widehat{G}^m} \widehat{f_1}\left(M_1^t \xi\right) \cdots \widehat{f_{2k}}\left(M_{2k}^t \xi\right) \right|^{2k} = \left| \sum_{\xi \in \widehat{G}^m} \widehat{g_1}(L_1^t \xi) \cdots \widehat{g_k}(L_k^t \xi) \right|^{2k}$$

$$\leq \prod_{j=1}^{k} \left| \sum_{\xi \in \widehat{G}^m} \widehat{g_j}\left(L_1^t \xi\right) \cdots \widehat{g_j}\left(L_k^t \xi\right) \right|^2. \tag{11}$$

By the Cauchy–Schwarz inequality,

$$
\left| \sum_{\xi \in \widehat{G}^m} \widehat{g_j}\left(L_1^t\xi\right) \cdots \widehat{g_j}\left(L_k^t\xi\right) \right|^2 = \left| \sum_{\xi \in \widehat{G}^m} \prod_{i=1}^{k} \widehat{f_{2j-1}}\left(L_i^t\xi\right) \widehat{f_{2j}}\left(-L_i^t\xi\right) \right|^2
$$

$$
= \left| \sum_{\xi \in \widehat{G}^m} \prod_{i=1}^{k} \widehat{f_{2j-1}}\left(M_{2i-1}^t\xi\right) \widehat{f_{2j}}\left(M_{2i}^t\xi\right) \right|^2
$$

$$
\leq \left(\sum_{\xi \in \widehat{G}^m} \prod_{i=1}^{k} \left|\widehat{f_{2j-1}}\left(M_{2i-1}^t\xi\right)\right|^2 \right) \left(\sum_{\xi \in \widehat{G}^m} \prod_{i=1}^{k} \left|\widehat{f_{2j}}\left(M_{2i}^t\xi\right)\right|^2 \right)
$$

$$
= \left(\sum_{\xi \in \widehat{G}^m} \prod_{i=1}^{k} \widehat{f_{2j-1}}\left(M_{2i-1}^t\xi\right) \widehat{f_{2j-1}}\left(M_{2i}^t\xi\right) \right) \left(\sum_{\xi \in \widehat{G}^m} \prod_{i=1}^{k} \widehat{f_{2j}}\left(M_{2i-1}^t\xi\right) \widehat{f_{2j}}\left(M_{2i}^t\xi\right) \right)
$$

$$
= \left(\sum_{\xi \in \widehat{G}^m} \prod_{i=1}^{2k} \widehat{f_{2j-1}}\left(M_i^t\xi\right) \right) \left(\sum_{\xi \in \widehat{G}^m} \prod_{i=1}^{2k} \widehat{f_{2j}}\left(M_i^t\xi\right) \right).
$$

Substituting this into (11) proves (10). For weakly norming systems L, the only difference is that we need to verify that g_j is non-negative whenever both f_{2j-1} and f_{2j} are. But $g_j = f_{2j-1} * f_{2j}^-$, where $f_{2j}^-(x) = f_{2j}(-x)$, so g_j is always non-negative. $\qquad\square$

This proposition easily generalises to the r-subdivision $\mathrm{sub}_r(L)$ of L, where each column v of L is replaced by r copies of v and $-v$, respectively. To see this, we only need to replace the use of the Cauchy–Schwarz inequality in the proof of Proposition 5.4 above by Hölder's inequality.

Proposition 5.5. *Let L be a weakly norming (resp. norming) system. Then the r-subdivision $\mathrm{sub}_r(L)$ of L is also weakly norming (resp. norming).*

In the light of this proposition, one may easily construct $m \times 2rm$ norming systems for any positive integers r and m.

Example 5.6. Let L be the $m \times m$ system given by the identity matrix. This is a degenerate case where $t_L(f) = \mathbb{E}[f]^m$, but subdividing L gives more interesting examples. For instance, $\mathrm{sub}(L)$ is the $m \times 2m$ system given by the $m \times 2m$ matrix

$$
\begin{pmatrix}
1 & -1 & & & \cdots & & \\
 & & 1 & -1 & \cdots & & \\
 & & & & \ddots & & \\
 & \cdots & & 1 & -1 & & \\
 & \cdots & & & & 1 & -1
\end{pmatrix}.
$$

As L is semi-norming, it immediately follows that $\mathrm{sub}(L)$ is as well. Concretely, since $\mathrm{Sol}(L)$ consists of the set of solutions to $x_1 = x_2$, $x_3 = x_4$, $\ldots$, $x_{2m-1} = x_{2m}$, $t_{\mathrm{sub}(L)}(f) = \mathbb{E}[f(x)^2]^m$. That is, the $\mathrm{sub}(L)$-norm is just the L^2-norm. More generally, the r-subdivision of L yields the L^{2r}-norm.

We can also construct $(m-1) \times 2rm$ (weakly) norming systems for any positive integers r and m.

Example 5.7. Let L be the $(m-1) \times m$ system given by the matrix

$$\begin{pmatrix} 1 & -1 & & \cdots & & \\ & 1 & -1 & \cdots & & \\ & & & \ddots & & \\ & & \cdots & 1 & -1 & \\ & & \cdots & & 1 & -1 \end{pmatrix}.$$

Then $\mathrm{Sol}(L)$ consists of the set of solutions to $x_1 = x_2 = \cdots = x_m$, so that $t_L(f) = \mathbb{E}[f(x)^m]$. Thus, if m is even, then L is norming and otherwise it is weakly norming. To see why it is not norming for odd m, it is enough to construct nonzero f such that $\mathbb{E}[f(x)^m] = 0$. Indeed, such a function exists, e.g., $f(x) = \mathbf{1}[x = 0] - \mathbf{1}[x = a]$ for a nonzero a. Proposition 5.5 implies that the r-subdivision of L is also weakly norming. In particular, the 2×6 weakly norming system in Example 2.4 corresponds to the case where $m = 3$ and $r = 1$.

It turns out that the two examples above include all possible $2 \times k$ weakly norming systems.

Theorem 5.8. *A $2 \times k$ system L is weakly norming if and only if it is isomorphic to one of the following three systems:*

(i) $k = 3$ and $x_1 = x_2 = x_3$.

(ii) $k = 4r$ for some $r \in \mathbb{N}$, $x_1 + \cdots + x_r = x_{r+1} + \cdots + x_{2r}$ and $x_{2r+1} + \cdots + x_{3r} = x_{3r+1} + \cdots + x_{4r}$.

(iii) $k = 6r$ for some $r \in \mathbb{N}$ and $x_1 - x_2 + \cdots - x_{2r} = x_{2r+1} - x_{2r+2} + \cdots - x_{4r} = x_{4r+1} - x_{4r+2} + \cdots - x_{6r}$.

Proof. The first one is simply the classical L^3-norm. The second is the r-subdivision of the 2×2 identity matrix L. The third one is isomorphic to the r-subdivision of the first one. Proposition 5.5 therefore proves that all three systems are weakly norming.

Let L be a $2 \times k$ weakly norming system with girth ℓ. By Corollary 5.3, the row space of L contains two distinct Schatten vectors a and b whose support has size ℓ and whose first non-zero coordinate is 1. These two vectors are linearly independent and, therefore, span the whole row space. Let I be the intersection of the support of a and b. If I is empty, then solutions to L are, up to permutation of coordinates, of the form in (ii).

Suppose now that I is non-empty. Without loss of generality, we may assume that the support of a is $[1, \ell]$ and the support of b is $[k - \ell + 1, k]$. Note that $k + |I| = 2\ell$. Recall that $L \setminus j$ is defined as the linear system in which we remove all vectors with non-zero j^{th} coefficient from the row space of L. Since $L \setminus k$ contains the equation $a_1 x_1 + \cdots + a_\ell x_\ell = 0$, we know by Corollary 4.6 that $L \setminus \ell$ also contains a Schatten vector whose support has size ℓ. Thus, the row space of L must contain a Schatten vector c with $c_\ell = 0$. We can express c as $c = \lambda a + \mu b$ for $\lambda, \mu \in \mathbb{F}_q$. Since we may assume without loss of generality that $c_1 = 1$ and, therefore, $c_i \in \{0, \pm 1\}$ for all $i \in [k]$, it follows that $\lambda = 1$ and $\mu \in \{\pm 1\}$. Furthermore, by replacing b by $-b$ if necessary, we may assume that $\mu = 1$. It follows that $c_i = 0$ precisely when $i \in I$. Thus, $\ell = k - |I|$, so, since $k + |I| = 2\ell$, we have $2k = 3\ell = 6|I|$.

Denoting $|I| = k/3$ by s, we see that because $a_1 + \cdots + a_s + b_{2s+1} + \cdots + b_{3s} = c_1 + \cdots + c_k = 0$, we have $a_1 + \cdots + a_s = -(b_{2s+1} + \cdots + b_{3s})$. We also have $a_i = -b_i$ for all $i \in [s+1, 2s]$ and, since all rows sum to zero, we know that $a_{s+1} + \cdots + a_{2s} = -(a_1 + \cdots + a_s)$. In summary, as all non-zero coefficients of a and b are 1 or -1, we may assume by reordering columns if necessary that L has

the form

$$L = \begin{pmatrix} a_1 & \cdots & a_s & -a_1 & \cdots & -a_s & 0 & \cdots & 0 \\ 0 & \cdots & 0 & a_1 & \cdots & a_s & -a_1 & \cdots & -a_s \end{pmatrix}$$

for some $a_1, \ldots, a_s \in \{\pm 1\}$. To see that L is a system of the form in (iii), it remains only to show that $\sigma = \sum_{i=1}^{s} a_i = 0$.

Assume, for the sake of contradiction, that $\sigma \neq 0$. We fix linearly independent $\gamma_1, \gamma_2 \in \mathbb{F}_q^n$ and let $\gamma_3 = \gamma_2 - \gamma_1$. We also take some $\alpha > 0$ and set $z = \exp(\pi i/\sigma)$. We define f and g to be the real-valued functions on $\mathbb{F}_q^n$ with Fourier coefficients

$$\widehat{f}(\xi) = \begin{cases} 1 & \text{if } \xi = 0, \\ \alpha z^\varepsilon & \text{if } \xi = \varepsilon\gamma_1 \text{ for } \varepsilon \in \{\pm 1\}, \\ \alpha & \text{if } \xi \in \{\pm\gamma_2, \pm\gamma_3\}, \\ 0 & \text{otherwise}, \end{cases}$$

$$\widehat{g}(\xi) = \begin{cases} 1 & \text{if } \xi = 0, \\ \alpha z^{(1-a_s\sigma)\varepsilon} & \text{if } \xi = \varepsilon\gamma_1 \text{ for } \varepsilon \in \{\pm 1\}, \\ \alpha & \text{if } \xi \in \{\pm\gamma_2, \pm\gamma_3\}, \\ 0 & \text{otherwise}. \end{cases}$$

Note that f and g take only positive real values if α is sufficiently small. Moreover, for any $\xi \in \hat{G}$, we have that

$$\widehat{g}(a_s\xi) \prod_{i=1}^{s-1} \widehat{f}(a_i\xi) = \left|\widehat{f}(\xi)\right|^s,$$

as is readily verified by checking the only non-trival case $\xi = \pm\gamma_1$. Letting $f_s = f_{2s} = f_{3s} = g$ and $f_i = f$ for all other $i \in [k]$, we therefore have, by Proposition 2.3, that

$$t_L(f_1, \ldots, f_k) = \sum_{\xi_1, \xi_2} \left(\widehat{g}(a_s\xi_1) \prod_{i=1}^{s-1} \widehat{f}(a_i\xi_1) \right) \left(\widehat{g}(a_s(\xi_2 - \xi_1)) \prod_{i=1}^{s-1} \widehat{f}(a_i(\xi_2 - \xi_1)) \right) \overline{\left(\widehat{g}(a_s\xi_2) \prod_{i=1}^{s-1} \widehat{f}(a_i\xi_2) \right)}$$

$$= \sum_{\xi_1, \xi_2} |\widehat{f}(\xi_1)|^s |\widehat{f}(\xi_2 - \xi_1)|^s |\widehat{f}(\xi_2)|^s.$$

On the other hand, denoting the real part of a complex number z by $\Re(z)$, we have that

$$t_L(f) = \sum_{\xi_1, \xi_2} \prod_{j=1}^{s} \widehat{f}(a_j\xi_1) \prod_{j=1}^{s} \widehat{f}(a_j(\xi_2 - \xi_1)) \prod_{j=1}^{s} \widehat{f}(-a_j\xi_2)$$

$$= \Re\left(\sum_{\xi_1, \xi_2} \prod_{j=1}^{s} \widehat{f}(a_j\xi_1) \prod_{j=1}^{s} \widehat{f}(a_j(\xi_2 - \xi_1)) \prod_{j=1}^{s} \widehat{f}(-a_j\xi_2) \right)$$

$$= \sum_{\xi_1, \xi_2} \Re\left(\prod_{j=1}^{s} \widehat{f}(a_j\xi_1) \prod_{j=1}^{s} \widehat{f}(a_j(\xi_2 - \xi_1)) \prod_{j=1}^{s} \widehat{f}(-a_j\xi_2) \right)$$

$$\leq \sum_{\xi_1, \xi_2} \left| \prod_{j=1}^{s} \widehat{f}(a_j\xi_1) \prod_{j=1}^{s} \widehat{f}(a_j(\xi_2 - \xi_1)) \prod_{j=1}^{s} \widehat{f}(-a_j\xi_2) \right|$$

19

$$= \sum_{\xi_1, \xi_2} |\widehat{f}(\xi_1)|^s |\widehat{f}(\xi_2 - \xi_1)|^s |\widehat{f}(\xi_2)|^s = t_L(f_1, \dots, f_k).$$

Moreover, the inequality is strict when $f_i = f$, since

$$t_L(f) = \sum_{\xi_1, \xi_2} \prod_{j=1}^{s} \widehat{f}(a_j \xi_1) \prod_{j=1}^{s} \widehat{f}(a_j(\xi_2 - \xi_1)) \prod_{j=1}^{s} \widehat{f}(-a_j \xi_2)$$

$$\le t_L(f_1, \dots, f_k) + \Re\left(\prod_{j=1}^{s} \widehat{f}(a_j \gamma_1) \widehat{f}(a_j \gamma_3) \widehat{f}(-a_j \gamma_2) \right) - |\widehat{f}(\gamma_1)|^s |\widehat{f}(\gamma_3)|^s |\widehat{f}(\gamma_2)|^s$$

$$= t_L(f_1, \dots, f_k) + \alpha^k \left(\Re(z^\sigma) - 1 \right).$$

Since $z = \exp(\pi i / \sigma)$, this yields that $t_L(f) \le t_L(f_1, \dots, f_k) - 2\alpha^k$, implying that L does not have the Hölder property. $\qquad\square$

Remark. Which of these weakly norming systems are norming? In Example 5.7, we have already seen that the first system is not norming. The second family of systems L_2 is norming, since $t_{L_2}(f)^{1/4r} = t_{L'}(f)^{1/2r}$, where L' is the single Schatten equation $x_1 + \cdots + x_r = x_{r+1} + \cdots + x_{2r}$. To verify that the third system L_3 is not norming, let f be the real function that satisfies $\widehat{f}(\xi) = \mathbf{1}[\xi = \pm a]$ for some $a \ne 0$. Then

$$t_{L_3}(f) = \sum_{\xi_1, \xi_2} |\widehat{f}(\xi_1)|^{2r} |\widehat{f}(\xi_2)|^{2r} |\widehat{f}(\xi_1 + \xi_2)|^{2r},$$

which evaluates to 0 unless $\mathbb{F}_q$ has characteristic 3. If $\mathbb{F}_q$ has characteristic 3, then $t_{L_3}(f) = 2$. However, deleting the first $2r$ variables $x_1, \dots, x_{2r}$ gives a Schatten equation L' of length $4r$, for which $t_{L'}(f) = 2$. In either case, we have a contradiction to Corollary 3.3.

One may wonder if weakly norming systems of rank three or more have a similar characterisation. Unfortunately (or fortunately), there are weakly norming systems of rank three that do not come from subdividing trivial examples.

Example 5.9. Let L be the 3×12 system given by the matrix

$$\begin{pmatrix} 1 & -1 & 1 & -1 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & -1 & 1 & -1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 1 & -1 \end{pmatrix}.$$

Then $\mathrm{Sol}(L)$ consists of all those $(x_1, x_2, \dots, x_{12})$ that can be parameterised by

$$(x_1, x_2, x_3, x_4, x_5, x_6) = (y_1 + z_{12}, z_{12} + y_2, y_2 + z_{23}, z_{23} + y_3, y_3 + z_{13}, z_{13} + y_1),$$
$$(x_5, x_6, x_7, x_8, x_9, x_{10}) = (y_3 + z_{31}, z_{31} + y_1, y_1 + z_{14}, z_{14} + y_4, y_4 + z_{34}, z_{34} + y_3),$$
$$(x_1, x_2, x_7, x_8, x_{11}, x_{12}) = (y_1 + z_{12}, z_{12} + y_2, y_1 + z_{14}, z_{14} + y_4, y_4 + z_{24}, z_{24} + y_2).$$

This system is weakly norming, since the parameterisation comes from the singleton-pair incidence graph of the set $\{1, 2, 3, 4\}$, which is a weakly norming graph. More details will be given in Section 7.

On the other hand, L is included in neither Example 5.6 nor Example 5.7. Indeed, if Example 5.6 includes L, then it must be the case that $m = 3$ and $r = 2$; however, the 2-subdivision of the 3×3 degenerate system has girth 4, whereas the girth of L is 6. Similarly, Example 5.7 does not include L for the simple reason that one would need to take $m = 4$ and $r = 3/2$.

We leave the characterisation of weakly norming systems of rank three as an open problem.

6 Forcing systems

A graph H is said to be *forcing* if the fact that $t_H(W) = p^{e(H)}$ for a graphon W with $p = t_{K_2}(W)$ implies that $W = p$ a.e. While it is easy to see that trees and non-bipartite graphs are not forcing, the forcing conjecture [20], a strengthening of Sidorenko's conjecture, says that all bipartite graphs containing a cycle should be forcing.

In analogy with the graph case, we say that an $m \times k$ linear system L is *forcing* if, for every n and every non-negative function f on $\mathbb{F}_q^n$, $t_L(f) = \mathbb{E}[f]^k$ implies that f is a constant function on $\mathbb{F}_q^n$. Taking our lead from a similar result for graphs, we will now show that every forcing system is Sidorenko.

Proposition 6.1. *Let L be an $m \times k$ linear system such that $t_L(f) < \mathbb{E}[f]^k$ for some $f : \mathbb{F}_q^n \to [0, 1]$. Then L is not forcing.*

Proof. We may assume that $n > 1$ by extending f to a larger $\mathbb{F}_q^n$ through adding an extra zero to each point in the support. Let g_1 be the indicator function of the subspace $\{(x_1, \ldots, x_n) \in \mathbb{F}_q^n : x_n = 0\}$. Then $t_L(g_1) = (q^{n-1})^{k-m}/(q^n)^{k-m} = q^{m-k}$ since L consists of m linearly independent forms and $\mathbb{E}[g_1]^k = q^{-k}$. Thus,

$$t_L(g_1) - \mathbb{E}[g_1]^k = q^{-k}(q^m - 1) > 0.$$

Now consider $P(\alpha) := t_L(\alpha g_1 + (1-\alpha)f) - \mathbb{E}[\alpha g_1 + (1-\alpha)f]^k$, which is a polynomial in α. As $P(0) < 0$ and $P(1) > 0$, there exists $\alpha_0 \in (0, 1)$ with $P(\alpha_0) = 0$. Then the function $h_1 := \alpha_0 g_1 + (1 - \alpha_0)f$ with range in $[0, 1]$ satisfies $t_L(h_1) = \mathbb{E}[h_1]^k$, so, if L is forcing, h_1 must be constant.

One can similarly obtain a function $h_2 = \beta_0 g_2 + (1 - \beta_0)f$ such that $\beta_0 \in (0, 1)$ and $t_L(h_2) = \mathbb{E}[h_2]^k$, where g_2 is the indicator function of the subspace $\{(x_1, \ldots, x_n) \in \mathbb{F}_q^n : x_{n-1} = 0\}$. It follows that h_2 is also a constant and, thus, that g_1 is a constant multiple of g_2 plus another constant; however, it is easy to check that this is not the case. $\square$

We now classify all the single equations that are forcing by showing that a single equation is forcing if and only if it is norming. Recall that, by Theorem 5.2, the norming and weakly norming properties are the same for a single equation, so all three properties agree in this case.

Theorem 6.2. *A $1 \times k$ system L is forcing if and only if it is norming.*

Proof. By Theorem 5.2, L is norming if and only if it is isomorphic to $\begin{pmatrix} 1 & \cdots & 1 & -1 & \cdots & -1 \end{pmatrix}$. Since $t_L(f) = \sum_\xi |\widehat{f}(\xi)|^{2k}$, $t_L(f) = \mathbb{E}[f]^{2k}$ if and only if $\sum_{\xi \neq 0} |\widehat{f}(\xi)|^{2k} = 0$. That is, $\widehat{f}(\xi) = 0$ for every $\xi \neq 0$ and, thus, f is a constant function.

To prove the converse, suppose L is forcing but not norming. By Proposition 6.1, L is Sidorenko and so, by [9, Theorem 1.4], isomorphic to a system of the form $\begin{pmatrix} a_1 & -a_1 & \cdots & a_r & -a_r \end{pmatrix}$ for $r = k/2$ and $a_1, \ldots, a_r \in \mathbb{F}_q$. Thus,

$$t_L(f) = \sum_\xi |\widehat{f}(a_1\xi)|^2 \cdots |\widehat{f}(a_r\xi)|^2.$$

Since L is not norming, there must be $i, j \in [r]$ such that $a_i \notin \{\pm a_j\}$. Fix a non-zero $\gamma \in \widehat{\mathbb{F}_q^n}$ and let f be the function with Fourier transform $\widehat{f}(\xi) = \frac{1}{2}\left(\mathbf{1}[\xi = \gamma] + \mathbf{1}[\xi = -\gamma]\right)$ for every $\xi \neq 0$ and $\widehat{f}(0) = 1$. Then f is a non-negative real-valued function, as $\sum_{\xi \neq 0} |\widehat{f}(\xi)| \leq 1$. Moreover, $\widehat{f}(a_i\xi)\widehat{f}(a_j\xi) = 0$ whenever $\xi \neq 0$, so $t_L(f) = |\widehat{f}(0)|^{2k} = \mathbb{E}[f]^{2k}$. But f is clearly not constant, contradicting our assumption that L is forcing. $\square$

In particular, every $1 \times k$ weakly norming system is forcing. We now prove a much stronger result, that every weakly norming system L is forcing. For weakly norming graphs, we know that they are forcing whenever they contain an even cycle because there is a 'domination' relation between the weakly norming graph and the even cycle (see, for example, [5, Section 5.2]). We may then use the simple fact that every even cycle is forcing to reach the required conclusion. Unfortunately, this proof does not transfer in an obvious way to the arithmetic setting, so we take a slightly different approach.

Theorem 6.3. *Every weakly norming system L is forcing.*

Proof. Let ℓ be the girth of L. By Theorem 5.1, the row space of L contains a Schatten vector a whose support has size ℓ. Without loss of generality, we may assume that the support of this vector is $[\ell]$. Let M be the $1 \times \ell$ system $\begin{pmatrix} a_1 & \cdots & a_\ell \end{pmatrix}$.

Suppose now that $t_L(g) = \mathbb{E}[g]^k$ for some non-negative function g on $\mathbb{F}_q^n$. Because ℓ is the girth of L, every vector in the row space of L that is only supported on $[\ell]$ must be a multiple of a. Thus, by Proposition 2.6, letting $f_i = g$ for $i \leq \ell$ and $f_i = 1$ for $i > \ell$, we have $t_M(g) = t_L(f_1, \ldots, f_k)$. By Proposition 2.2,

$$t_L(f_1, \ldots, f_k) \leq \prod_{i=1}^{k} \|f_i\|_L.$$

But $\|g\|_L = \mathbb{E}[g]$ by assumption and $\|1\|_L = 1$, which yields $t_M(g) \leq \mathbb{E}[g]^\ell$. Since M is norming, Theorem 6.2 implies that it is also forcing, so that g must be constant. $\qquad\square$

One might suspect that Theorem 6.2 generalises and the converse of Theorem 6.3 is also true. However, as we now show, there are forcing systems that are not weakly norming.

Example 6.4. Consider the 2×7 linear system L given by the matrix

$$\begin{pmatrix} 1 & 1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & -1 & -1 \end{pmatrix}.$$

Let f be a non-negative function. Then the Cauchy–Schwarz inequality gives

$$t_L(f) = \mathbb{E}_{x \in G} \left[f(x) f * f * f(x)^2 \right] \geq \frac{\mathbb{E}_{x \in G} \left[f(x) f * f * f(x) \right]^2}{\mathbb{E}_{x \in G}[f]} = \frac{\mathbb{E}_{x \in G} \left[f * f(x)^2 \right]^2}{\mathbb{E}_{x \in G}[f]}.$$

Now suppose that $t_L(f) = \mathbb{E}[f]^7$. It follows that $\sum_\xi \widehat{f}(\xi)^4 = \mathbb{E}\left[f * f(x)^2 \right] = \mathbb{E}[f]^4$, which means $\widehat{f}(\xi) = 0$ for every $\xi \neq 0$. Therefore, L is forcing.

To see that L is not weakly norming, we apply Corollary 4.6. Indeed, $L \setminus 1$ is isomorphic to the Schatten system $x_1 + x_2 = x_3 + x_4$, whereas $L \setminus 4$ is isomorphic to $x_1 + x_2 + x_3 = x_4 + x_5 + x_6$.

7 From norming hypergraphs to norming systems

In this section, we describe how (weakly) norming hypergraphs naturally give rise to (weakly) norming linear systems. Indeed, given a k-uniform hypergraph H, we may look at the homomorphism density $t_H(W)$ restricted to functions of the form $W(x_1, \ldots, x_k) := f(x_1 + \cdots + x_k)$ for a function $f : \mathbb{F}_q^n \to \mathbb{R}$. Then, in this Cayley setting, we have that

$$t_H(W) = \mathbb{E}\left[\prod_{i_1 \cdots i_k \in E(H)} W(x_{i_1}, \ldots, x_{i_k}) \right] = \mathbb{E}\left[\prod_{i_1 \cdots i_k \in E(H)} f(x_{i_1} + \cdots + x_{i_k}) \right] =: t_{L_H}(f),$$

where L_H is the system of linear equations whose solution set is defined by the parametrisation

$$y_e = x_{i_1} + \cdots + x_{i_k}, \quad e = \{i_1, \ldots, i_k\} \in E(H).$$

We call L_H the *linear H-system* and the parametrisation above the *standard parametrisation* of $\mathrm{Sol}(L_H)$.

To see that L_H is (weakly) norming whenever H is, it is enough to observe that if $f_1, \ldots, f_{e(H)} : \mathbb{F}_q^n \to \mathbb{R}$ are functions and $W_e(x_1, \ldots, x_k) := f_e(x_1 + \cdots + x_k)$ for each $e \in E(H)$, then

$$t_H(W_1, \ldots, W_{e(H)}) = \mathbb{E}\left[\prod_{e = i_1 \cdots i_k \in E(H)} W_e(x_{i_1}, \ldots, x_{i_k}) \right]$$

$$= \mathbb{E}\left[\prod_{e = i_1 \cdots i_k \in E(H)} f_e(x_{i_1} + \cdots + x_{i_k}) \right] = t_{L_H}(f_1, \ldots, f_{e(H)}).$$

Therefore, the inequality (1) holds for $L = L_H$ if and only if the corresponding inequality

$$|t_H(W_1, \ldots, W_{e(H)})| \leq \prod_{e \in E(H)} \|W_e\|_H$$

holds, where $\|W\|_H := |t_H(W)|^{1/e(H)}$. To summarise, we have the following result.

Proposition 7.1. *The linear H-system L_H is (weakly) norming whenever H is (weakly) norming.*

Though our focus in this paper is on norming systems of equations, we note that a similar result holds for Sidorenko and common systems. That is, for example, that any Sidorenko hypergraph gives rise to a Sidorenko system of equations. This gives a simple mechanism for producing many examples of Sidorenko and common systems, though it also suggests that the classification of such systems will be difficult.

We now give some examples of weakly norming systems that come from weakly norming hypergraphs.

Example 7.2. Let $H = K_{a,b}$, the complete bipartite graph on $A \cup B$ with $|A| = a$ and $|B| = b$, which is one of the simplest examples of a weakly norming graph. Then

$$t_{L_H}(f) = \mathbb{E}\left[\prod_{(a,b) \in A \times B} f(x_a + y_b) \right],$$

where the expectation is taken over uniform random vectors $(x_a)_{a \in A} \in G^A$ and $(y_b)_{b \in B} \in G^B$. The corresponding norms are the so-called *grid norms*, which played a key role in the recent paper [8]. In particular, if $a = b = 2$, then $t_{L_H}(f) = \mathbb{E}\left[f(x_1 + y_1) f(x_2 + y_1) f(x_1 + y_2) f(x_2 + y_2) \right]$, which corresponds to the fourth power of the Gowers U^2-norm. Moreover, if $a = 2$ and $b = 3$, then one can check that L_H is isomorphic to the 2×6 linear system given in Example 2.4.

Example 7.3. For H-systems, the subdivision operation defined in Section 5 can be viewed in purely graph-theoretic terms. Indeed, if H is a norming r-graph, let H^+ and H^- be two vertex-disjoint copies of H, both also vertex-disjoint from H. For an edge $e \in E(H)$, let e^+ and e^- be the edges corresponding to e in H^+ and H^-, respectively. The 2-*blow-up extension* of H, denoted by $\beta^2(H)$, is then the $(r+1)$-graph on $V(H^+) \cup V(H^-) \cup E(H)$, whose edges are those sets of the

form $e^+ \cup \{e\}$ and $e^- \cup \{e\}$. A standard application of the Cauchy–Schwarz inequality implies that $\beta^2(H)$ is also norming, a fact which first appeared implicitly in [3].

We claim that the subdivision $\mathrm{sub}(L_H)$ of L_H is equal to the $\beta^2(H)$-system. To see this, enumerate $E(H)$ and $E(\beta^2(H))$ by the integers in $[k]$ and $[2k]$ in such a way that the $(2i-1)$-th and the $2i$-th edges of $\beta^2(H)$ are obtained by adding a common vertex to two copies of the i-th edge of H, respectively. Then each vector $(z_1, z_2, \ldots, z_{2k})$ in $\mathrm{Sol}(L_{\beta^2(H)})$ can be parameterised by $z_{2j-1} = y_j + x_j$ and $z_{2j} = y'_j + x_j$, where $(y_1, \ldots, y_k)$ and $(y'_1, \ldots, y'_k)$ are vectors in $\mathrm{Sol}(L_H)$ and x_j is a free variable. Therefore, if $(y_1, \ldots, y_k)$ satisfies $a_1 y_1 + \cdots + a_k y_k = 0$ and $(y'_1, \ldots, y'_k)$ satisfies $a_1 y'_1 + \cdots + a_k y'_k = 0$, then $(z_1, z_2, \ldots, z_{2k})$ satisfies the equation $a_1(z_1 - z_2) + a_2(z_3 - z_4) + \cdots + a_k(z_{2k-1} - z_{2k}) = 0$, which is in $\mathrm{sub}(L_H)$. Conversely, if $(z_1, z_2, \ldots, z_{2k})$ satisfies a linear equation $b_1 z_1 + b_2 z_2 + \cdots + b_{2k} z_{2k} = 0$, then $b_{2i-1} = b_{2i}$ must hold, since otherwise the free variable x_i does not vanish. It therefore follows that the equation is in $\mathrm{sub}(L_H)$.

To give a concrete example, if $H = C_4$, then $\beta^2(H)$ is the line 3-graph of a cube, denoted by $M(3)$ in [3]. The H-system for $H = M(3)$, the line 3-graph of a cube, is therefore the subdivision of L_{C_4} and is defined by the single linear equation $x_1 + x_2 + x_3 + x_4 = y_1 + y_2 + y_3 + y_4$, which is just the L^8-norm of the spectrum.

For graphs H, it is reasonably straightforward to determine the corresponding system of linear equations L_H. The first step is to specify the set of all minimal linear dependencies between variables.

Theorem 7.4. *Let L_H be the linear H-system defined by a graph H and let $F \subseteq E(H)$ be an edge subset. Then $\{y_e : e \in F\}$ in the standard parametrisation of L_H is a minimal linearly dependent set if and only if F is an even cycle in H.*

Proof. It is straightforward to see that $\{y_e : e \in F\}$ is a minimal linearly dependent set of variables if F is an even cycle. Let us re-index the variables in such a way that $F = [2k]$ and $y_i = x_i + x_{i+1}$, where the addition of indices is taken modulo $2k$. Then $\sum_{j=1}^{2k} (-1)^j y_j = 0$ and every proper subset of $\{y_1, \ldots, y_{2k}\}$ is linearly independent.

Conversely, suppose that $\{y_e : e \in F\}$ is a minimal linearly dependent set of variables. Then F must be the edge set of a subgraph, which, abusing notation slightly, we also call F, with minimum degree at least two. In particular, F contains an even cycle with edge set $F' \subseteq F$; however, $\{y_e : e \in F'\}$ is a linearly dependent set of variables, which means $F = F'$. $\qquad\square$

The *cycle matroid* of a graph H is the matroid defined on the edge set $E(H)$ of a graph H whose independent sets are those edge subsets of H which do not contain a cycle. In other words, every minimal dependent set, i.e., every *circuit*, is the edge set of a cycle, which for (weakly) norming graphs H, which are necessarily bipartite, exactly corresponds to the minimal linear dependence condition in Theorem 7.4. Hence, for (weakly) norming graphs H, the map $e \mapsto y_e$ is a matroid isomorphism from the cycle matroid of H to $\{y_e : e \in E(H)\}$. This observation now allows us to compute the rank of the linear system L_H.

Proposition 7.5. *Let L_H be the linear H-system defined by a (weakly) norming graph H. Then L_H has rank $|E(H)| - |V(H)| + \kappa(H)$, where $\kappa(H)$ denotes the number of connected components in H.*

Proof. It is a standard fact that the rank of the cycle matroid of a graph H is $|V(H)| - \kappa(H)$, the number of edges in a spanning forest of H. Therefore, the dimension of $\mathrm{Sol}(L_H)$, which is the rank of $\{y_e : e \in E(H)\}$, is also $|V(H)| - \kappa(H)$, so the rank of the matrix L_H with columns indexed by $E(H)$ is $|E(H)| - |V(H)| + \kappa(H)$. $\qquad\square$

24

Together, Proposition 7.5 and Theorem 7.4 allow us to explicitly write out the $m \times k$ linear system L_H corresponding to a bipartite graph H in matrix form in two steps. First, compute the rank m of L_H using Proposition 7.5. Second, assign alternating signs to the edges in each cycle of H. Once we find m such cycles whose corresponding ± 1 row vectors are linearly independent, we are done.

Example 7.6. Let H be the 1-subdivision of the complete graph K_t. That is, we replace each edge of K_t with a path of length two. By the results of [5], this is a connected weakly norming graph for every $t \geq 3$, though it was shown in [15] that it is not norming for any $t \geq 4$. As $|E(H)| = t(t-1)$ and $|V(H)| = t + t(t-1)/2 = t(t+1)/2$, L_H has rank $t(t-3)/2 + 1$. In particular, if $t = 4$, this gives the 3×12 system given in Example 5.9.

For another example, let $H = K_{a,b}$ as given in Example 7.2. Then we have an $m \times k$ system L_H with $m = (a-1)(b-1)$ and $k = ab$. In particular, we can write the 4×9 matrix L_H for the case $a = b = 3$ as

$$
\begin{pmatrix}
1 & -1 & & 1 & -1 & & & & \\
& & & -1 & 1 & -1 & 1 & & \\
-1 & & & & 1 & & & 1 & -1 \\
& & -1 & & & & 1 & & 1 & -1
\end{pmatrix}.
$$

Thus, the grid norms generalise Example 2.4 in a different way to Example 5.7.

Unfortunately, the simple characterisation of minimal dependent sets given in Theorem 7.4 does not generalise in a straightforward manner to hypergraphs. A generalisation, if it exists, must at least depend on the characteristic of the underlying field $\mathbb{F}_q$. To illustrate the dependence on the characteristic, suppose first that $\mathbb{F}_q$ is a field of characteristic 2. Then the set of variables $\{y_e : e \in F\}$ for an edge subset $F \subseteq E(H)$ in the standard parametrisation of the linear H-system L_H is linearly dependent if and only if F is an even-degree subgraph. That is, the number of edges of F incident to each vertex must be even. For example, let $\{x_i + y_j + z_k : i, j, k \in \{0,1\}\}$ be the standard parametrisation of the linear H-system L_H, where H is the 'octahedron' 3-graph. This is the standard 3-graph that defines the Gowers octahedral norm, while L_H defines the Gowers U^3-norm. Then $\{x_i + y_j + z_k : i + j + k = 0 \bmod 2\}$ is a minimal linearly dependent set. However, this is not the case over other characteristics. Indeed, if $\{x_i + y_j + z_k : i + j + k = 0 \bmod 2\}$ is a linearly dependent set of variables, then $\sum_{i+j+k \text{ even}} c_{ijk}(x_i + y_j + z_k) = 0$ for some $c_{ijk} \in \{\pm 1\}$; however, one may easily check that this is impossible for characteristics other than 2.

Before moving on, let us show that every linear H-system L defined by a norming graph H gives a norm which is topologically equivalent to the Gowers U^2-norm.

Theorem 7.7. *Let H be a norming graph and let L be the linear H-system. Then, for every $\varepsilon > 0$ and any $f : G \to \mathbb{R}$ with $|f| \leq 1$, the following hold:*

(i) $\|f\|_L \leq \varepsilon$ implies $\|f\|_{U^2} \leq \varepsilon^{2/e(H)}$ and

(ii) $\|f\|_{U^2} \leq \varepsilon$ implies $\|f\|_L \leq \varepsilon$.

Proof. By Theorem 5.1, every norming system contains a Schatten equation of length 2ℓ for some $\ell \in [2, e(H)/2]$. Thus, $\|f\|_L \leq \varepsilon$ implies $\sum_\xi |\widehat{f}(\xi)|^{2\ell} \leq \varepsilon$ by Corollary 3.3. In particular, we must have $|\widehat{f}(\xi)|^2 \leq \varepsilon^{1/\ell}$ for all ξ. Since the U^2-norm arises from the equation $x_1 - x_2 + x_3 - x_4 = 0$, we have by Proposition 2.3 and Plancherel's theorem that

$$
\|f\|_{U^2} = \sum_{\xi \in \widehat{G}} |\widehat{f}(\xi)|^4 \leq \left(\max_{\psi \in \widehat{G}} |\widehat{f}(\psi)|^2 \right) \sum_{\xi \in \widehat{G}} |\widehat{f}(\xi)|^2 = \left(\max_{\psi \in \widehat{G}} |\widehat{f}(\psi)|^2 \right) \mathbb{E}_{x \in G} |f(x)|^2 \leq \varepsilon^{1/\ell},
$$

yielding the required conclusion.

Conversely, let H be a graph on vertex set $[d]$. Since a norming graph must have at least one edge, we may assume that $\{1, 2\}$ is an edge of H. By the definition of $L = L_H$, we then have

$$\|f\|_L^{e(H)} = \mathbb{E}_{x_1,\ldots,x_d \in G} \left[\prod_{ij \in E(H)} f(x_i + x_j) \right]$$

$$= \left(\mathbb{E}_{x_2,\ldots,x_d \in G} \left[\prod_{\substack{ij \in E(H) \\ 1 \notin ij}} f(x_i + x_j) \right] \right) \left(\mathbb{E}_{x_1 \in G} \left[\prod_{j \in N(1)} f(x_1 + x_j) \right] \right).$$

By the Cauchy–Schwarz inequality, this is at most

$$\left(\mathbb{E}_{x_2,\ldots,x_d \in G} \left[\left(\prod_{\substack{ij \in E(H) \\ 1 \notin ij}} f(x_i + x_j) \right)^2 \right] \right)^{1/2} \left(\mathbb{E}_{x_2,\ldots,x_d \in G} \mathbb{E}_{x_1, x_1' \in G} f(x_1 + x_j) f(x_1' + x_j) \right)^{1/2}.$$

Using the fact that $|f| \leq 1$ and applying Cauchy–Schwarz once more, we see that this is at most

$$\left(\mathbb{E}_{x_3,\ldots,x_d \in G} \mathbb{E}_{x_1, x_1', x_2, x_2' \in G} f(x_1 + x_2) f(x_1' + x_2) f(x_1 + x_2') f(x_1' + x_2') \right)^{1/4}.$$

Since $(x_1 + x_2, x_1' + x_2, x_1 + x_2', x_1' + x_2')$ parametrises the solutions of $y_1 - y_2 - y_3 + y_4 = 0$, this is precisely the U^2 norm of f. $\qquad \square$

Example 7.8. Theorem 7.7 does not generalise to hypergraphs, in that, for $r \geq 3$, different r-graphs may define inequivalent systems. Example 7.3 shows that the line 3-graph of a cube gives rise to the L^8-norm of the spectrum. This is equivalent to the U^2-norm, whereas the 'octahedron' 3-graph gives rise to the U^3-norm. On the other hand, Hatami, Hatami and Lovett [13] proved that $\| \cdot \|_L$ is equivalent to the U^k-norm for some k whenever L is a norming system, so norming hypergraphs H always give rise to norms $\| \cdot \|_{L_H}$ that are equivalent to some U^k-norm.

8   Complex-valued functions

Gowers norms were originally defined for complex-valued functions [11]. For instance, for any function $f : G \to \mathbb{C}$,

$$\|f\|_{U_2}^4 := \mathbb{E} \left[f(x_1) \overline{f(x_2) f(x_3)} f(x_4) \right],$$

where the expectation is taken over all solutions to the equation $x_1 - x_2 - x_3 + x_4 = 0$ and, unlike the real-valued case, some of the f's have been conjugated. The question we will be concerned with here is whether, for a given (real-)norming system, it is possible to assign conjugates in a similar fashion so that the system defines a norm on the space of complex-valued functions. We stress that here we are only concerned with norming systems, since weakly norming systems have an absolute value under the expectation which makes the distinction moot.

The analogous question for graphs was already answered in the affirmative by Lee and Sidorenko [15], who proved that, given a (real-)norming graph, there exists a conjugation assignment under

which the graph density defines a norm for complex-valued functions. That is, the (real-)norming property is equivalent to the complex-norming property. In this section, following their technique, we prove that the same holds for linear systems.

Let L be an $m \times k$ linear system and α a 0/1-valued function defined on $[k]$. The *complex L-density $t_{L,\alpha}$ with respect to α* is defined by

$$t_{L,\alpha}(f) := \mathbb{E}_{(x_1,\ldots,x_k)\in \mathrm{Sol}(L)} \left[\prod_{i=1}^{k} f(x_i)^{\alpha(i)} \overline{f(x_i)}^{1-\alpha(i)} \right].$$

That is, we use $\overline{f(x_i)}$ if $\alpha(i) = 0$ and $f(x_i)$ otherwise for each $i \in [k]$. We say that the pair (L, α) is *complex-norming* if $|t_{L,\alpha}(f)|^{1/k}$ defines a norm on the space of complex-valued functions on $\mathbb{F}_q^n$ for every n. With this notation, the main result of this section is as follows.

Theorem 8.1. *Let L be an $m \times k$ norming system. Then there exists $\alpha : [k] \to \{0,1\}$ such that (L, α) is complex-norming.*

To prove Theorem 8.1, we need some notation. Let $\mathcal{F}$ be the set of all complex-valued functions defined on $\mathbb{F}_q^{\mathbb{N}}$ with finite support. That is, each $f \in \mathcal{F}$ can be seen as a complex-valued function on $\mathbb{F}_q^n$ for a large enough n. Following [15], we define a *decoration functional* on $\mathcal{F}^k$ to be a function from $\mathcal{F}^k$ to $\mathbb{C}$ that satisfies the following conditions:

(i) $|\tau(cf_1,\ldots,cf_k)| = |c|^k |\tau(f_1,\ldots,f_k)|$ for each $c \in \mathbb{C}$;

(ii) $\tau(f_1,\ldots,f_{i-1},g+ah,f_{i+1},\ldots,f_k)$
$= \tau(f_1,\ldots,f_{i-1},g,f_{i+1},\ldots,f_k) + a\tau(f_1,\ldots,f_{i-1},h,f_{i+1},\ldots,f_k)$
for any $f_1,\ldots,f_k,g,h \in \mathcal{F}$ and any *real* number a;

(iii) $\tau(\overline{f}_1,\ldots,\overline{f}_k) = \overline{\tau(f_1,\ldots,f_k)}$;

(iv) $\tau(f_1 \otimes g_1,\ldots,f_k \otimes g_k) = \tau(f_1,\ldots,f_k)\tau(g_1,\ldots,g_k)$.

For brevity, we will sometimes write $\tau(f) = \tau(f,\ldots,f)$. Following [15], we set $s_L(f_1,\ldots,f_k) := \max_\alpha |t_{L,\alpha}(f_1,\ldots,f_k)|$, where the maximum is taken over all $\alpha : [k] \to \{0,1\}$. Then $s_L(\cdot)$ only satisfies a weaker 'subadditivity' property rather than the linearity condition (ii), namely,

(ii') $|\tau(f_1,\ldots,f_{i-1},g+h,f_{i+1},\ldots,f_k)|$
$\leq |\tau(f_1,\ldots,f_{i-1},g,f_{i+1},\ldots,f_k)| + |\tau(f_1,\ldots,f_{i-1},h,f_{i+1},\ldots,f_k)|.$

If τ satisfies (i) and (ii'), then we say that τ is a *weak decoration functional*. In particular, $s_L(\cdot)$ is a weak decoration functional. By adapting the proof of Proposition 2.1, we have the following lemma.

Lemma 8.2. *Let τ be a weak decoration functional on $\mathcal{F}^k$. Then $|\tau(\cdot)|^{1/k}$ is a seminorm on $\mathcal{F}$ if, for any $f_1,\ldots,f_k \in \mathcal{F}$,*

$$|\tau(f_1,\ldots,f_k)|^k \leq \prod_{i=1}^{k} |\tau(f_i)|.$$

The converse also holds if τ is a decoration functional.

This in turn allows us to verify the following property of $s_L(\cdot)$. Since the proof is verbatim the same as that of [15, Lemma 5.5], we omit it.

Corollary 8.3. *Let L be an $m \times k$ norming system. Then $s_L(\cdot)^{1/k}$ is a norm on $\mathcal{F}$.*

We will also need the following easy fact [15, Lemma 2.2] saying that $|\tau(\cdot)|^{1/k}$ is a seminorm if and only if $|\tau(\cdot)|$ is convex.

Lemma 8.4. *Let τ be a weak decoration functional on $\mathcal{F}^k$. Then $|\tau(\cdot)|^{1/k}$ is a seminorm on $\mathcal{F}$ if and only if $|\tau(\cdot)|$ is convex, i.e., $|\tau(\frac{1}{2}(f+g))| \leq \frac{1}{2}(|\tau(f)| + |\tau(g)|)$ for all $f, g \in \mathcal{F}$.*

Thus, combining Corollary 8.3 with this lemma shows that $s_L(\cdot)$ is a convex functional.

Corollary 8.5. *Let L be an $m \times k$ norming system. Then $s_L(\cdot)$ is convex on $\mathcal{F}$.*

The final ingredient we need is the following lemma [15, Lemma 2.3], which says that the convexity of $s_L(\cdot)$ implies the convexity of *some* $|t_{L,\alpha}(\cdot)|$ under appropriate technical conditions. To state the conditions, we need some more definitions. A *density functional* is a functional of the form $\tau(f) = \tau(f, f, \ldots, f)$ where τ is a decoration functional (and, crucially, not a weak decoration functional). A set $\mathcal{C}$ in a vector space $\mathcal{F}$ is called *algebraically open at* $f \in \mathcal{C}$ if, for any $g \in \mathcal{F}$, there exists $\varepsilon > 0$ such that $f + xg \in \mathcal{C}$ for each $x \in (-\varepsilon, \varepsilon)$. We simply say that $\mathcal{C}$ is *algebraically open* if it is algebraically open at every $f \in \mathcal{C}$.

Lemma 8.6. *Let $\tau_1, \tau_2, \ldots, \tau_r$ be non-negative real-valued density functionals on $\mathcal{F}$ that satisfy:*

(a) $\tau(f) := \max_j \tau_j(f)$ is convex on $\mathcal{F}$;

(b) for each $j \in [r]$, there is $f \in \mathcal{F}$ such that $\max_{\ell \neq j} \tau_\ell(f) < \tau_j(f)$;

(c) for any non-empty algebraically open subset $\mathcal{C} \subseteq \mathcal{F}$, there is $f \in \mathcal{C}$ such that $\tau_j(f) > 0$ for all $j \in [r]$.

Then there exists $j \in \{1, 2, \ldots, r\}$ such that $\tau_j(\cdot)$ is convex.

With these results in hand, we are now ready to prove Theorem 8.1, which states that every norming system is also complex norming under an appropriate conjugation assignment.

Proof of Theorem 8.1. Let us first give a brief sketch. By Corollary 8.5, we already know that $s_L(\cdot)$ is convex, which verifies condition (a) of Lemma 8.6 for $\tau = s_L$ and $\tau_1, \ldots, \tau_r$ any collection of functionals of the form $|t_{L,\alpha}|$. Therefore, if (b) and (c) hold, there exists $|t_{L,\alpha}(\cdot)|$ that is convex. Finally, Lemma 8.4 and the 'converse' part of Lemma 8.2 show that (L, α) is norming, with the latter excluding the possibility that $|t_{L,\alpha}(\cdot)|^{1/k}$ defines a seminorm that is not a norm.

Thus, it remains to verify the technical conditions (b) and (c) in Lemma 8.6. Let $\mathcal{A}$ be a minimal collection of 0/1-valued functions on $[k]$ such that $s_L(f) = \max_{\alpha \in \mathcal{A}} |t_{L,\alpha}(f)|$. Then (b) immediately follows from minimality once we take $\tau_1, \ldots, \tau_r$ to be the collection of $|t_{L,\alpha}|$ with $\alpha \in \mathcal{A}$.

It therefore suffices to verify (c). For 0/1-valued functions $\alpha_1, \ldots, \alpha_r$ on $[k]$, consider the augmented function $\alpha : [rk] \to \{0, 1\}$ that maps each $j \in [rk]$ to $\alpha_q(i)$ if $j = (q-1)k + i$ for $1 \leq i \leq k$. Let $L(r)$ be the $rm \times rk$ system obtained by taking the disjoint union of r copies of the $m \times k$ system L, i.e.,

$$L(r) = \begin{pmatrix} L & & \cdots & & \\ & L & \cdots & & \\ & & \ddots & & \\ & & \cdots & L & \\ & & \cdots & & L \end{pmatrix}.$$

Let g be the constant function 1 and let $f \in \mathcal{C}$ for an algebraically open set $\mathcal{C} \subseteq \mathcal{F}$. Then there exists $\varepsilon > 0$ such that $\mathcal{C}$ contains all functions of the form $f + xg$ with $x \in (-\varepsilon, \varepsilon)$. The function $P(x) := t_{L(r),\alpha}(f + xg)$ is a polynomial of degree rk with complex coefficients, since the coefficient $t_{L(r),\alpha}(f + xg)$ of x^{rk} is $t_{L(r),\alpha}(g) = 1 \neq 0$. Thus, $P(x)$ has only a finite number of zeros on the interval $(-\varepsilon, \varepsilon)$ and, therefore, there exists some $c \in (-\varepsilon, \varepsilon)$ and $h = f + cg \in \mathcal{G}$ such that $\prod_{i=1}^{r} t_{L,\alpha_i}(h) = t_{L(r),\alpha}(h) = P(c) \neq 0$. □

9 Concluding remarks

Some surprisingly basic questions about the structure of weakly norming systems remain open. For instance, while we have shown that every weakly norming system must include a Schatten vector in its row space, it remains open as to whether every weakly norming system has a basis consisting only of Schatten vectors. Moreover, if this is true, is it the case that there is a basis for the row space consisting of Schatten vectors all of which have the same support size?

In Section 7, we saw that each (weakly) norming hypergraph gives rise to a (weakly) norming linear system. But is it the case that every weakly norming linear system arises in this way? If so, then the conjecture (see, for instance, [5, 15]) that every weakly norming hypergraph is associated to a finite reflection group in a specific way would also extend to weakly norming linear systems. Since Fourier techniques are available to us in the latter context, it may be that such statements are then easier to prove.

References

[1] Daniel Altman. On a common-extendable, non-Sidorenko linear system. *Comb. Theory*, 3(3):Paper No. 5, 12, 2023.

[2] David Conlon, Jacob Fox, and Benny Sudakov. An approximate version of Sidorenko's conjecture. *Geom. Funct. Anal.*, 20(6):1354–1366, 2010. `doi:10.1007/s00039-010-0097-0`.

[3] David Conlon, Hiêp Hàn, Yury Person, and Mathias Schacht. Weak quasi-randomness for uniform hypergraphs. *Random Structures Algorithms*, 40(1):1–38, 2012. `doi:10.1002/rsa.20389`.

[4] David Conlon, Jeong Han Kim, Choongbum Lee, and Joonkyung Lee. Some advances on Sidorenko's conjecture. *J. London Math. Soc.*, 98(3):593–608, 2018. `doi:10.1112/jlms.12142`.

[5] David Conlon and Joonkyung Lee. Finite reflection groups and graph norms. *Adv. Math.*, 315:130–165, 2017. `doi:10.1016/j.aim.2017.05.009`.

[6] David Conlon and Joonkyung Lee. Sidorenko's conjecture for blow-ups. *Discrete Anal.*, 2021:Paper No. 2, 13 pp., 2021. `doi:10.19086/da`.

[7] David Conlon and Joonkyung Lee. Domination inequalities and dominating graphs. *Math. Proc. Cam. Phil. Soc.*, pages 1–18, 2024. `doi:10.1017/S0305004124000185`.

[8] Yuval Filmus, Hamed Hatami, Kaave Hosseini, and Esty Kelman. Sparse graph counting and Kelley–Meka bounds for binary systems. Preprint available at arXiv:2311.12248.

[9] Jacob Fox, Huy Tuan Pham, and Yufei Zhao. Common and Sidorenko linear equations. *Q. J. Math.*, 72(4):1223–1234, 2021. `doi:10.1093/qmath/haaa068`.

[10] Frederik Garbe, Jan Hladký, and Joonkyung Lee. Two remarks on graph norms. *Discrete Comput. Geom.*, 67(3):919–929, 2022. `doi:10.1007/s00454-021-00280-w`.

[11] William Timothy Gowers. A new proof of Szemerédi's theorem. *Geom. Funct. Anal.*, 11(3):465–588, 2001. `doi:10.1007/s00039-001-0332-9`.

[12] Hamed Hatami. Graph norms and Sidorenko's conjecture. *Israel J. Math.*, 175:125–150, 2010. `doi:10.1007/s11856-010-0005-1`.

[13] Hamed Hatami. Private communication. 2024.

[14] Nina Kamčev, Anita Liebenau, and Natasha Morrison. Towards a characterization of Sidorenko systems. *Q. J. Math.*, 74(3):957–974, 2023. URL: `https://doi.org/10.1093/qmath/haad013`, `doi:10.1093/qmath/haad013`.

[15] Joonkyung Lee and Alexander Sidorenko. On graph norms for complex-valued functions. *J. Lond. Math. Soc.*, 106(2):1501–1538, 2022. `doi:https://doi.org/10.1112/jlms.12604`.

[16] László Lovász. Operations with structures. *Acta Math. Acad. Sci. Hungar.*, 18:321–328, 1967. URL: `https://doi.org/10.1007/BF02280291`, `doi:10.1007/BF02280291`.

[17] László Lovász. Graph homomorphisms: Open problems. Unpublished manuscript, 2008.

[18] László Lovász. *Large networks and graph limits*, volume 60 of *Amer. Math. Soc. Colloq. Publ.* American Mathematical Society, 2012. `doi:10.1090/coll/060`.

[19] Alexander Sidorenko. Weakly norming graphs are edge-transitive. *Combinatorica*, 40(4):601–604, 2020. `doi:10.1007/s00493-020-4468-3`.

[20] Jozef Skokan and Lubos Thoma. Bipartite subgraphs and quasi-randomness. *Graphs Combin.*, 20(2):255–262, 2004. `doi:10.1007/s00373-004-0556-1`.

[21] Balázs Szegedy. An information theoretic approach to Sidorenko's conjecture. Preprint available at arXiv:1406.6738.