Quasi-random Classes of Hypergraphs

Fan R. K. Chung
Bellcore, Morristown, NJ 07960

ABSTRACT

We investigate the relations among a number of different graph properties for k-uniform-hypergraphs, which are shared by random hypergraphs. Various graph properties form equivalence classes which in turn constitute a natural hierarchy. The analogues for binary functions on k-tuples and for hypergraphs with small density are also considered. Several classes are related to communication complexity and expander graphs.

1. INTRODUCTION

The theory of "quasi-random" graphs involves a study of graph properties and their relations. Although the properties of interest are usually satisfied by random graphs, there is no direct relation to random graphs. However, these properties can be viewed as measurements of randomness.

Quasi-random graphs were first introduced in [12] by presenting a large class of graph properties that are mutually equivalent in the sense that any graph having one of these properties must necessarily have all of them. The analogous version for hypergraphs was given in [9]. While the roots of quasi-random graphs could be traced back to various problems in extremal graphs [13, 14, 19, 33, 34], there has been growing interest and applications as reflected in the recent work in [3, 18, 20, 25, 29, 30].

In this paper we establish a hierarchy by equivalence classes $\mathcal{A}_i$ for k-hypergraphs (or k-graphs for short) as follows:

$$\mathcal{A}_0 \supsetneq \mathcal{A}_1 \supsetneq \cdots \supsetneq \mathcal{A}_k$$

$\mathcal{A}_k$ is the so called "quasi-random" class that was introduced in [9] and it consists of various graph properties for a k-graph G such as: 'All k-graphs on $2k$ vertices

Random Structures and Algorithms, Vol. 1, No. 4 (1990)

appear almost equally often as induced subgraphs of G," "For any fixed $s \geq 2k$ all k-graphs on s vertices appear almost equally often as induced subgraphs of G," etc. From the opposite end, in $\mathcal{A}_0$ there is the property that the number of edges in G is approximately the same as the number of nonedges in G. In $\mathcal{A}_1$ there is the property that G is "almost regular." $\mathcal{A}_2$ contains the "jumbled graphs" property considered by Thomason and others [20, 29, 30]. It turns out that each $\mathcal{A}_i$ corresponds to an interesting invariant, the so-called i-deviation (defined in Section 2) which provides a quantitative indication as to how much the graph deviates from random graphs.

We will also discuss analogous quasi-random classes for k-tuples, or for functions from V^k to $\{-1, 1\}$. These classes are of particular interest because of their connection to communication complexity, detailed in Section 4.

In Section 5, we consider quasi-random classes involving functions or hypergraphs with small density. Some of these classes are related to expanders, which arise in many topics in extremal graphs and computational complexity.

2. NOTATION

A k-graph $G = (V, E)$ consists of a set $V = V(G)$, called the *vertices* of G, and a subset $E = E(G)$ of the set $\binom{V}{k}$ of k-element subsets of V, called the *edges* of V. We use the notation $G(n)$ to indicate that V has n elements. Throughout this paper, G denotes a k-graph unless otherwise specified.

For a k-graph $G' = (V', E')$, we let μ_G: $\binom{V}{k} \rightarrow \{-1, 1\}$ denote the edge function of G, i.e., for

$$x \in \binom{V}{k}, \quad \mu_G(x) = \begin{cases} -1 & \text{if } x \in E\,, \\ 1 & \text{otherwise}\,. \end{cases}$$

Let V^k denote the set of k-tuples $(v_1, \ldots, v_k)$, $v_i \in V$, where the v's are not necessarily distinct. Let $\Pi_G^{(i)}$: $V^{k+i} \rightarrow \{-1, 1\}$ denote the following function of G.

$$\Pi_G^{(i)}(u_1, u_2, \ldots, u_{2i}, v_{i+1}, \ldots, v_k) = \prod_{\varepsilon_1} \cdots \prod_{\varepsilon_i} \mu_G(\varepsilon_1, \ldots, \varepsilon_i, v_{i+1}, \ldots, v_k)$$

where $\varepsilon_j \in \{u_{2j-1}, u_{2j}\}$ for $j \leq i$ and $\mu_G(w_1, \ldots, w_k)$ is defined to be 0 if two of the w's are equal. $\Pi_G^{(i)}$ is a product of 2^i terms each of which is an edge function. For $i = 0$, we define $\Pi_G^0 = \mu_G$.

The *i-deviation* of G, denoted by $dev_i(G)$, is defined as follows:

$$dev_i(G) = \frac{1}{n^{k+i}} \sum_{u_1, \ldots, u_{k+i}} \Pi_G^{(i)}(u_1, \ldots, u_{k+i})$$

Thus $dev_i(G)$ assumes a value between -1 and 1. (Another interpretation is that $n^{k+i} dev_i$ is the difference of the number of "even partial (squashed) octahedrons" and the "odd partial (squashed) octahedrons" as defined in [9]; also see [11] for more details.)

For $X \subseteq V$, $G[X]$ denotes the subgraph of G induced by X, i.e., $G[X] = \left(X, E \cap \binom{X}{k}\right)$.

Let H denote an l-graph where $l < k$ and $V(H) = V(G)$. The set $E(G, H)$ of edges of G induced by H is defined to be:

$$E(G, H) = \left\{x \in E(G): \binom{x}{l} \subseteq E(H)\right\}$$

For $l = 1$, the edge set of H is just a subset of $V(G)$ and $E(G, H) = E(G[H])$. We denote $e(G) = |E(G)|$ and $e(G, H) = |E(G, H)|$.

3. STATEMENT OF THE MAIN RESULTS

We will use the following convention. Suppose we have two classes $P = P(o(1))$ and $P' = P'(o(1))$, each with occurences of the asymptotic $o(1)$ notation. By the implication "$P \Rightarrow P'$", we mean that for each $\varepsilon > 0$ there is a $\delta > 0$ (a function of ε and k but independent of n) such that if $G(n)$ satisfies $P(\delta)$ then it also satisfies $P'(\varepsilon)$, provided $n > n_0(\varepsilon)$. Two properties P and P' are said to be equivalent if $P \Rightarrow P'$ and $P' \Rightarrow P$.

In [9] it was shown that the property $dev_k(G) = o(1)$ for a hypergraph G is equivalent to a number of disparate properties, among which are:

Q: For all k-graphs G' on $2k$ vertices, the number of (labeled) occurrences of G' in G as an induced subgraph is $(1 + o(1))n^{2k}2^{-\binom{2k}{k}}$.

Let s denote a fixed integer and $s \geq 2k$.

$Q(s)$: For all k-graphs $G'(s)$ on s vertices the number of (labeled) occurrences of G' in G as an induced subgraph is $(1 + o(1))n^s 2^{-\binom{s}{k}}$.

For $k = 2$, the above properties are also equivalent to some additional properties [10, 12, 27], including the following:

Q': For each subset $S \geq V(G)$, $e(G) = \frac{1}{2}\binom{|S|}{2} + o(n^2)$.

As noted in [9], the analogous version of Q' for hypergraphs, (i.e., $e(G) = \frac{1}{2}\binom{|S|}{k} + o(n^k)$) is not equivalent to Q and therefore is not quasi-random.

In an attempt to generalize Q', Frankl and Rödl suggested the following property for 3-graphs G as a possible quasi-random property:

$$e(G, H) = \frac{1}{2}\, e(K_n^{(3)}, H) + o(n^3)$$

where $K_n^{(3)}$ denotes the complete 3-graph on n vertices.

Here we further generalize the above property FR for fixed integers i and k:

P_i: $dev_i(G) = o(1)$.
R_i: For every $(i-1)$-graph H, with $i \geq 2$,

$$e(G, H) - e(\bar{G}, H) = o(n^k)$$

where $\bar{G}$ denotes the complement of G with edge set $\left\{x \in \binom{V}{k} : x \not\in E(G)\right\}$.

Also, for $i = 0$ and 1, we define

R_0: $e(G) - e(\bar{G}) = o(n^k)$.
R_1: G is almost regular. That is,

$$\sum_{u_1, \ldots, u_{k-1}} (d^+(u_1, \ldots, u_{k-1}) - d^-(u_1, \ldots, u_{k-1}))^2 = o(n^{k+1})$$

where $d^+(u_1, \ldots, u_{k-1}) = |\{v \in V : \{u_1, \ldots, u_{k-1}, v\} \in E(G)\}|$,

$$d^-(u_1, \ldots, u_{k-1}) = |\{v \in V : \{u_1, \ldots, u_{k-1}, v\} \not\in E(G)\}|\,.$$

Theorem 1. *Properties P_i and R_i are equivalent for $i = 0, \ldots, k$. In particular for $i \geq 2$, we have*

(i) $disc_i(G) = \underset{H:(i-1)\text{-graph}}{Max} \dfrac{|e(G, H) - e(\bar{G}, H)|}{|V(G)|^k} < (dev_i(G))^{1/2^i}$

(ii) $dev_i(G) < 4^i(disc_i(G))^{1/2^i}$

We remark that disc_2 is often called discrepancy (see [14–16]) and disc_i (so-called the i-discrepancy) is a natural generalization of discrepancy. We note that the constant 4 in (ii) can probably be improved by more careful analysis. However, it would be significant if the power $1/2^i$ on the right-hand sides of the inequalities could be improved.

Theorem 2. *Let $\mathcal{A}_i$ denote the equivalence class of k-graphs for which P_i holds. Then,*

$$\mathcal{A}_0 \underset{\neq}{\supset} \mathcal{A}_1 \underset{\neq}{\supset} \mathcal{A}_2 \underset{\neq}{\supset} \cdots \underset{\neq}{\supset} \mathcal{A}_k$$

The proofs for Theorems 1 and 2 will be given in Sections 6–8. In [12] Theorems 1 and 2 are proved for the case of $k = 2$ (i.e., "usual" graphs). Frankl and Rödl proved independently $P_3 \Leftrightarrow R_3$ (private communication).

The family $\mathcal{A}_i = \mathcal{A}_i^{(k)}$ of k-graphs is said to be (k, i)-quasi-random, or sometimes i-quasi-random if there is no confusion. The term "k-quasi-random" for k-graphs is the same as "quasi-random" as in previous papers.

In a k-graph G, we define the neighborhood graph G_v of a vertex v having vertex set $G - \{v\}$ and edge set $E(G_v) = \left\{x \in \binom{V}{k-1} : \{v\} \cup x \in E(G)\right\}$. We then have the following:

Corollary 1. *G is (k, i)-quasi-random if and only if for almost all vertices v, the neighborhood graphs G_v are $(k-1, i)$-quasi-random, where $i < k$.*

For two vertices u and v in a k-graph G we define the common graph $G_{u,v}$ having vertex set $G - \{u, v\}$ and edge set $E(G_{u,v}) = \left\{x \in \binom{V}{k-1}: \mu_G(x \cup \{u\}) = \mu_G(x \cup \{v\})\right\}$.

Corollary 2. *G is a (k, i)-quasi-random if and only if for almost all pairs of vertices u and v, the common graphs $G_{u,v}$ are $(k-1, i-1)$-quasi-random.*

Corollaries 1 and 2 are immediate consequences of the definition of dev_i and Theorems 1 and 2.

4. QUASI-RANDOM CLASSES AND COMMUNICATION COMPLEXITY

Let f denote a function from V^k to $\{-1, 1\}$. All the definitions for k-graph have natural analogs for functions on k-tuples, or k-functions for support.

Let I denote a subset of size i of $\{1, \dots, k\} = [k]$ and define $\Pi_{f,I}^{(i)}$: $V^k \to \{1, -1\}$ by

$$\Pi_{f,I}(u_1, \dots, u_{k+c}) = \prod_{\varepsilon_1} \cdots \prod_{\varepsilon_k} f(\varepsilon_1, \dots, \varepsilon_k)$$

where $\varepsilon_j \in \{u_{j+m-1}, u_{j+m}\}$ if $j \in I$ and $m = |I \cap [j]|$; and $\varepsilon_j = u_{j+m}$ if $j \not\in I$. The i-deviation of f is defined to be:

$$dev_i\, f = \operatorname*{Max}_I \frac{1}{n^{k+i}} \sum_{u_1 \cdots u_{k+i}} \Pi_{f,I}\,(u_1, \dots u_{k+i})$$

where I ranges over all subsets of $[k]$ of size i. For a k-tuple $u = (u_1, \dots, u_k)$, we define u_I to be an i-tuple $(u_{a_1}, \dots, u_{a_i})$ where $a_1 < \cdots < a_i$ and $a_i \in I$. Let $\mathcal{H}_i$ denote a family of i-functions where $i < k$ and the members of $\mathcal{H}_i$ are indexed by $\binom{[k]}{i}$, denoted by h_I. The number $e(f, \mathcal{H}_i)$ is defined as follows:

$$\begin{aligned} e(f, \mathcal{H}_i) = \{x \in V^k &: \text{for every } I \subseteq [k] \text{ with } |I| = i, \text{and } h_I \in \mathcal{H}_i, h_I(u_I) \\ &= -1 \text{ and } f(x) = -1\} \end{aligned}$$

For fixed i, the properties for a k-function f can be stated as follows:

$\tilde{P}_i$: $dev_i(f) = o(1)$;
$\tilde{R}_i$: For $i \geq 2$, for every family $\mathcal{H}_{i-1}$ of $(i-1)$-functions

$$e(f, \mathcal{H}_{i-1}) - e(-f, \mathcal{H}_{i-1}) = o(n^k)\,.$$

For $i = 0$ and 1, we define

$\tilde{R}_0$: $e(f) - e(-f)) = o(n^k)$

where $$e(f) = |\{x \in V^k \colon f(x) = -1\}| .$$

$\tilde{R}_1$: f is almost regular, i.e.,

$$\operatorname*{Max}_{1 \le i \le k} \sum_{u_1, \ldots, u_{k-1}} (d_i^+(u_1, \ldots, u_{k-1}) - d_i^-(u_1, \ldots, u_{k-1}))^2 = o(n^{k+1})$$

where for $1 \le i \le k-1$,

$$d_i^+(u_1, \ldots, u_{k-1}) = |\{v \in V \colon f(u_1, \ldots, u_{i-1}, v, u_i, \ldots, u_{k-1}) = -1\}|$$

and

$$d_k^+(u_1, \ldots, u_{k-1}) = |\{v \in V \colon f(u_1, \ldots, u_{k-1}, v) = -1\}| ,$$

d_i^- is defined similarly for the f-value 1.

Theorem 3. *Property $\tilde{P}_i$ is equivalent to Property $\tilde{R}_i$ for $0 \le i \le k$. In particular, for $i \ge 2$, we have*

$$\begin{aligned} disc_i(f) &= \operatorname*{Max}_{\mathcal{H}_{i-1}} \frac{|e(f, \mathcal{H}_{i-1}) - e(-f, \mathcal{H}_{i-1})|}{|V|^k} \\ &< (dev_i(f))^{1/2^i} \end{aligned}$$

and

$$dev_i(f) < 4^i (disc_i(f))^{1/2^i} .$$

Theorem 4. *Let $\tilde{\mathcal{A}}_i$ denote the equivalence class for functions on $\{1, -1\}$ that satisfy $\tilde{Q}_i$. Then*

$$\tilde{A}_0 \underset{\neq}{\supset} \tilde{\mathcal{A}}_1 \underset{\neq}{\supset} \tilde{\mathcal{A}}_2 \underset{\neq}{\supset} \cdots \underset{\neq}{\supset} \tilde{\mathcal{A}}_k$$

Although the definitions for k-functions seem to be more complicated than for hypergraphs, the proofs are very similar and, in fact, simpler in most cases. We will omit the proofs for k-functions since they are already well-suggested by the proofs in Sections 6–8.

In [5], Babai, Nisan, and Szegedy considered the communication complexity for a k-party communication protocol f. The communication complexity for f is bounded below by $\log_2 1/\Gamma(f)$ where $\Gamma(f)$ is as follows (see [5]):

$$\begin{aligned} \Gamma(f) &= \max_S \operatorname{Prob}(x \in S \text{ and } f(x) = -1) - \operatorname{Prob}(x \in S \text{ and } f(x) = 1)) \\ &= \max_{\mathcal{H}_{k-1}} \frac{e(f, \mathcal{H}_{k-1}) - e(-f, \mathcal{H}_{k-1})}{n^k} \end{aligned}$$

where S ranges over so-called "cylinder intersections" which correspond to $\mathcal{H}_{k-1}$. Therefore property $\tilde{R}_k$ is equivalent to the property that G has small $\Gamma(f)$ and thus large communication complexity. Theorem 3 can then be used to show that numerous classes of functions have large communication complexity.

Communication complexity is used to derive lower bounds for time-space trade-off and generating pseudo-random numbers among numerous other applications in complexity theory and distributed computing [1, 4, 7, 31, 35]. The reader is referred to [23] for an extensive survey on this topic.

5. QUANTITATIVE QUASI-RANDOM CLASSES

Suppose α is a real number between 0 and 1. For a function f from V^k to $\{1, -1\}$, we define $f_\alpha(x) = 1 - \alpha$ if $f(x) = -1$ and $f_\alpha(x) = -\alpha$ if $f(x) = 1$. Although we could choose $\alpha = |\{x \in V^k\colon f(x) = -1\}|/|V|^k$, in general α is independent of f. We define $dev_i\, f_\alpha$ in an analogous way as $dev_i\, f$ (by replacing each occurrence of f by f_α in the product sum) and we have the following:

Theorem 5. *The following two properties* $P_{i,\alpha}$ and $R_{i,\alpha}$ *for a function* $f\colon V^k \to \{1, -1\}$ *are equivalent*:

$P_{i,\alpha}$: $dev_i(f_\alpha) = o(1)$.
$R_{i,\alpha}$: For $i \geq 2$, and for every family $\mathcal{H}_{i-1}$ of $(i-1)$-functions,

$$(1-\alpha)e(f, \mathcal{H}_{i-1}) - \alpha e(-f, \mathcal{H}_{i-1}) = o(n^k)$$

and

$$R_{0,\alpha}\colon\ (1-\alpha)e(f, \mathcal{H}_{i-1}) - \alpha e(-f, \mathcal{H}_{i-1}) = o(n^k)$$

$$R_{1,\alpha}\colon\ \max_i \sum_{u_1, \ldots, u_{k-1}} ((1-\alpha)d_i^+(u_1, \ldots, u_{k-1})) - \alpha d_i^-(u_1, \ldots, u_{k-1}))^2$$
$$= o(n^{k+1})\,.$$

In particular we have, for $i \geq 0$

$$\text{(i)}\ \ disc_i(f_\alpha) = \underset{\mathcal{H}_{i-1}}{Max} \frac{|(1-\alpha)e(f_\alpha, \mathcal{H}_{i-1}) - \alpha e(-f_\alpha, \mathcal{H}_{i-1})|}{|V|^k}$$
$$< (dev_i(f_\alpha))^{1/2^i}$$

and

$$\text{(ii)}\ \ dev_i(f_\alpha) < 4^i(disc_i(f_\alpha))^{1/2^i}$$

Theorem 6. *Let* $\mathcal{A}_{i,\alpha}$ *denote the equivalence class for functions for k-tuples of V to* $\{1, -1\}$ *that satisfy* $Q_{i,\alpha}$ *where* $\alpha = |\{v \in V^k\colon f(x) = -1\}|/|V|^k$. *Then*

$$\mathcal{A}_{0,\alpha} \underset{\neq}{\supset} \mathcal{A}_{1,\alpha} \underset{\neq}{\supset} \cdots \underset{\neq}{\supset} \mathcal{A}_{k,\alpha}$$

The proof of Theorem 5 is very similar to that of Theorem 1 in Section 6–8 (by replacing each occurrence of f by f_α) and will be omitted.

We note that the 2-discrepancy $disc_2$ can be modified as follows:
For a subset X of V (which can be viewed as the edge set of a 1-graph),

$$disc_2(f) = \underset{X \subseteq V}{Max} \frac{e(f, X) - \alpha|X|^k}{|X|^k}$$

where

$$e(f, X) = \left|\left\{x \in \binom{X}{k} : f(x) = -1\right\}\right| .$$

Suppose we choose α to be $\left|\left\{x \in \binom{V}{k} : f(x) = -1\right\}\right| \Big/ |V|^k$ (which can be viewed as the density of "ordered" hyper-edges). Therefore $disc_2(f)$ associates with the maximum quantity that the number of ordered-edges in a subset X can differ from the average. If we can use $dev_2(f)$ to (upper) bound $disc_2(f)$, we can then (lower) bound the number of edges leaving X from every $X \subseteq V$ and therefore assert the expanding property of the hypergraphs.

Expanding properties often arise in various problems ranging from nonblocking networks [24], sorting [2], amplification of weak random sources [28], and various problems in computational complexity. The proof for Theorem 6 consists of constructions which are obtained by using character sums, included in Section 9.

6. $P_i \Rightarrow R_i$

Suppose $dev_i(G) < \varepsilon$ and H is an $(i-1)$-graph. Recall that, for each $j \leq i$, we define

$$\Pi_G(u_1, \ldots, u_{2j}; v_{j+1}, \ldots, v_k) = \prod_{\varepsilon_i} \cdots \prod_{\varepsilon_j} \mu(\varepsilon_1, \ldots, \varepsilon_j; v_{j+1}, \ldots, v_k)$$

where $\varepsilon_{j'} \in \{u_{2j'-1}, u_{2j'}\}$ for $1 \leq j' \leq j$ and Π_G is a product of 2^j terms.

We consider

$$n^{k+i} dev_i(G) = \sum_{u_1, \ldots, u_{2i}, v_{i+1}, \ldots, v_k} \Pi_G(u_1, \ldots, u_{2i}; v_{i+1}, \ldots, v_k)$$

Eventually, we want to show the above sum can be bounded below by a subsum ranging over $u_1,, \ldots, u_{2i}, v_{i+1}, \ldots, v_k$ satisfying the property that all $(i-1)$-sets $\delta_{t_1}, \ldots, \delta_{t_i-1}$ are in $E(H)$ where

$$\delta_{t_j} = v_{t_j} \text{ if } t_j > i \text{ and } \delta_{t_j} = u_{2t_j-1} \text{ or } u_{2t_j} \text{ if } t_j \leq i . \tag{1}$$

This property will be denoted by $u_1, \ldots, u_{2i}, v_{i+1}, \ldots, v_k \in P_K^H$ where $K = \{1, \ldots, k\}$. Let T be a subset of K. We write $u_1, \ldots, u_{2i}, v_{i+1}, \ldots, v_k \in P_T^H$, if for every $(i-1)$-set, $t_1, \ldots, t_{i-1}$ of T we have $\{\delta_{t_1}, \ldots, \delta_{t_i-1}\}$ in $E(H)$ where δ_{t_j} is defined as in (1). Since

$$n^{k+i}dev_i(G) = \sum_{u_1,\ldots,u_{2i},v_{i+1},\ldots,v_k} \sum \Pi_G(u_1,\ldots,u_{2i}; v_{i+1},\ldots,v_k)$$

$$= \sum_{u_1,\ldots,u_{2i-2}} \sum_{v_{i+1}\cdots v_k} \sum_{u_{2i-1},u_{2i}} \Pi_G(u_1,\ldots,u_{2i-1}; u_{2i-2}, v_{i+1},\ldots,v_k)$$

$$\times \Pi_G(u_1,\ldots,u_{2i-2}; u_{2i}, v_{i+1},\ldots,v_k)$$

$$= \sum_{u_1,\ldots,u_{2i-2}} \sum_{v_{i+1},\ldots,v_k} \Big(\sum_{v\in V} \Pi_G(u_1,\ldots,u_{2i-2}; v, v_{i+1},\ldots,v_k)\Big)^2$$

In the sum above we can then choose a subsum which then is smaller than $n^{k+i}dev_i(G)$. Namely, we restrict ourselves to those u's and v's in P^H_{K-i}. In other words, we only select those $u_1,\ldots,u_{2i-2}, v_{i+1},\ldots,v_k$ such that all $(i-1)$-subsets $\delta_{t_1},\ldots,\delta_{t_{i-1}}$ for $\delta_{t_j}\in\{u_{2t_j-1}, u_{2t_j}\}$, if $t_j<i$, and $\delta_{t_j}=v_{t_j}$ if $t_j>i$, are contained in $E(H)$. By expanding the squared terms, the above quantity can be rewritten as follows:

$$n^{k+i}dev_i(G) > \sum_{u_1,\ldots,u_{2i},v_{i+1},\ldots,v_k\in P^H_{K-i}} \Pi\, \mu(\varepsilon_1,\ldots,\varepsilon_{i-1},\varepsilon_i; v_{i+1},\ldots,v_k)$$

It is not difficult to check that if $u_1,\ldots,u_{2i}, v_{i+1},\ldots,v_k$ satisfies P^H_{K-j} for every $1\le j\le i$, then $u_1,\ldots,u_{2i}, v_{i+1},\ldots,v_k$ satisfies P^H_K since every $(i-1)$-set of K is in $K-j$ for some j in $\{1,\ldots,i\}$. We will show the above quantity is lower bounded by a subsum of those $u_1,\ldots,u_{2i}, v_{i+1},\ldots,v_k$ satisfying P^H_{K-1} and P^H_{K-i} (denoted by $u_1,\ldots,u_{2i}, v_{i+1},\ldots,v_k\in P^H_{K-1}\cap P^H_{K-i}$). We have

$$n^{k+i}dev_i(G) \ge \sum_{u_1,\ldots,u_{2i},v_{i+1},\ldots,v_k\in P^H_{K-i}} \Pi\, \mu(\varepsilon_1,\ldots,\varepsilon_i, v_{i+1},\ldots,v_k)$$

$$= \sum_{u_3,\ldots,u_{2i},v_{i+1},\ldots,v_k} \sum_{u_1,u_2\in D_{K-i}} \Pi\, \mu(\varepsilon_1,\ldots,\varepsilon_i, v_{i+1},\ldots,v_k)$$

where $D_{K-i}=\{v\colon \{v,\delta_{t_2},\ldots,\delta_{t_{i-1}}\}$ is in $E(H)$ for every choice of $(i-2)$-set $\{t_2,\ldots,t_{i-1}\}$ of $K-i$ and δ_t's are defined as in (1)$\}$.

The above quantity can be rewritten as follows:

$$n^{k+i}dev_i(G) \ge \sum_{u_3,\ldots,u_{2i},v_{i+1},\ldots,v_k} \Big(\sum_{v\in D_{K-i}} \Pi\, \mu(v,\varepsilon_2,\ldots,\varepsilon_i, v_{i+1},\ldots,v_k)\Big)^2$$

We can then choose a subsum of those $u_3,\ldots,u_{2i}, v_{i+1},\ldots,v_k$ which satisfy P^H_{K-1}.

Therefore we have proved

$$n^{k+i}dev_i(G) \ge \sum_{u_1,\ldots,u_{2i},v_{i+1},\ldots,v_k\in P^H_{K-1}\cap P^H_{K-i}} \Pi\mu(\varepsilon_1,\ldots,\varepsilon_i, v_{i+1},\ldots,v_k)\,.$$

In a similar way, we can choose subsums which satisfy P^H_{K-j} for other j's. The proof is straightforward and will not be repeated here. So we have

$$n^{k+i}dev_i(G) \ge \sum_{u_1,\ldots,u_{2i},v_{i+1},\ldots,v_k\in P^H_{K-1}\cap P^H_{K-2}\cap\cdots\cap P^H_{K-i}} \Pi$$

$$\times\, \mu(\varepsilon_1,\ldots,\varepsilon_i, v_{i+1},\ldots,v_k)$$

$$= \sum_{u_1,\ldots,u_{2i},v_{i+1},\ldots,v_k\in P^H_K} \Pi\, \mu(\varepsilon_1,\ldots,\varepsilon_i, v_{i+1},\ldots,v_k)$$

We now apply the Cauchy Schwarz inequality:

$$n^{k+i}dev_i(G) \geq \sum_{u_1,\ldots,u_{2i-2},v_{i+1},\ldots,v_k} \left(\sum_{v \in F_{K-i}} \Pi\mu(\varepsilon_1,\ldots,\varepsilon_{i-1},v,v_{i+1},\ldots,v_k) \right)^2$$

$$\geq \frac{1}{n^{k+i-2}} \left(\sum_{u_1,\ldots,u_{2i-2},v_2,v_{i+1},\ldots,v_k \in P_K^H} \Pi\, \mu(\varepsilon_1,\ldots,\varepsilon_{i-1},v_i,\ldots,v_k) \right)^2$$

Note that we rename v by v_i and $u_1,\ldots,u_{2i-2},v_i,v_{i+1},\ldots,v_k \in P_K^H$ means all $\{\delta_{j_1},\ldots,\delta_{j_i-1}\}$ are in $E(H)$ where $\delta_{t_j}=v_{t_j}$ or $(\delta_{t_j}=u_{2t_j-1}$ or $u_{2t_j})$. Also $F_{K-i}=\{v\colon u_1,\ldots,u_{2i-2},v,v_{i+1},\ldots,v_k \in P_K^H\}$.

After we repeat the process $i-1$ times, we will arrive at the following:

$$n^{k+i}dev_i(G) \geq \frac{1}{n^{k2^i-k-i}} \left(\sum_{v_1,\ldots,v_k} \mu(v_1,\ldots,v_k) \right)^{2^i}$$

where all $(i-1)$-subsets of $\{v_1,\ldots,v_k\}$ are in $E(H)$ and therefore $\Sigma_{v_1,\ldots,v_k}\,\mu(v_1,\ldots,v_k)$ is exactly the difference of two numbers, one of which is the number of ordered k-sets x in $E(G)$ with $\binom{x}{i-1} \subset E(H)$, while the other is the number of ordered k-sets y not in $E(G)$ with $\binom{y}{i-1} \subset E(H)$. Therefore

$$n^{k+1}dev_i(G) \geq \frac{1}{n^{k2^i-k-i}}\,(k!(e(G,H)-e(\bar{G},H)))^{2^i}.$$

Since $dev_i(G)<\varepsilon$ by Q_i, we have

$$\begin{aligned}|e(G,H)-e(\bar{G},H)| &\leq (k!)^{-1/2^i}n^k(dev_i(G))^{1/2^i}\\ &\leq n^k\varepsilon^{1/2^i}.\end{aligned}$$

This completes the proof of $P_i \Rightarrow R_i$. ∎

7. $R_i \Rightarrow P_i$

Suppose that for every $(i-1)$-graph H, we have

$$|e(G,H)-e(\bar{G},H)| = \left|2e(G,H)-e\left(\binom{V}{k},H\right)\right| < 2\varepsilon^2 n^k$$

We want to show $dev_i(G)<4^i\varepsilon^{1/2^i}$. We will proceed by induction on both k and i. For $k=2$, it is mainly proved in [12] with the following clarification: It obviously holds if $16\varepsilon^{1/4}\geq 1$. Suppose $16\varepsilon^{1/4}<1$. Fact 9 in [12] states that if $|e(G,X)-e(\bar{G},X)|<2\varepsilon^2n^2$ for all subsets X of vertices, then

$$\sum_{u,v}\left|s(u,v)-\frac{n}{2}\right|<20\sqrt{3}n^3 \text{ where } s(u,v)=|\{w: \mu(u,w)=\mu(v,w)\}|\,.$$

Since $dev_2(G)=\frac{4}{n^4}\sum_{u,v}\left|s(u,v)-\frac{n}{2}\right|^2$, we have

$$\begin{aligned} dev_2(G) &< \frac{2}{n^3}\sum_{u,v}\left|s(u,v)-\frac{n}{2}\right| \\ &\le 40\sqrt{3}\varepsilon \le 16\varepsilon^{1/4}\,. \end{aligned}$$

We may assume $2\le i\le k$ and $k\ge 3$.

There are two cases:

Case 1. $i<k$.

It is straightforward to verify that

$$dev_i(G)=\frac{1}{n}\sum_{v\in V} dev_i(G_v)$$

where G_v denotes the neighborhood graph of v.

Let S denote $\{v: dev_i(G_v)>(4^i-2)\varepsilon^{1/2^i}\}$. If $|S|\le 2\varepsilon n$, we have

$$\begin{aligned} dev_i(G) &\le \frac{1}{n}\left(\sum_{v\in S} dev_i(G_v)+\sum_{v\notin S} dev_i(G_v)\right) \\ &\le \frac{|S|}{n}+(4^i-2)\varepsilon^{1/2^i}<4^i\varepsilon^{1/2^i} \end{aligned}$$

We may assume $|S|>2\varepsilon n$. For each G_v, $v\in S$, by the inductive hypothesis, there is a $(i-1)$-graph $H(v)$ so that $\left|2e(G_v,H(v))-e\left(\binom{V}{k-1},H(v)\right)\right|\ge 2\delta^2n^{k-1}$ where $\delta^2>3\varepsilon$ (since $4^i\cdot\delta^{1/2^i}<(4^i-2)\varepsilon^{1/2^i}$). There is a subset S' of S with $|S'|=\varepsilon n$ so* that either

(a): $$e(G_v,H(v))\ge\frac{1}{2}e\left(\binom{V}{k-1},H(v)\right)+2\delta^2n^{k-1}$$

for all $v\in S'$,

or (b):

$$e(G_v,H(v))\le\frac{1}{2}e\left(\binom{V}{k-1},H(v)\right)-2\delta^2n^{k-1}$$

We will only treat case (a) while the other case can be proved similarly and the proof will not be included here. Now we define an $(i-1)$-graph H with $V(H)=V$ and $E(H)=\{x\in\binom{V}{i-1}: x\in E(H(v)) \text{ for some } v \text{ in } S'\}\cup\binom{V-S'}{i-1}$.

* Strictly speaking, we should use $\lfloor\varepsilon n\rfloor$ instead of εn. However, we will usually not bother with this type of detail since it has no significant effect on the arguments or results.

We consider

$$\sum_v e(G_v, H(v)) = \sum_{x \in \binom{V}{k}} |\{u \in x \cap S' : x - u \in E(G_u, H(u))\}|$$

For each set $x \in \binom{V}{k}$ with $x \cap S' \neq \emptyset$, one of the two situations occurs:

(i) $|x \cap S'| \geq 2$.

There are at most $\varepsilon^2 n^k$ such x.

(ii) $x = u \cup y$ for $u \in S$ and $y \cap S' = \emptyset$. x is in $E(G, H)$ if and only if $x - u$ is in $E(G_u, H(u))$.

Therefore, we have

$$\begin{aligned}\sum_y e(G_v, H(v)) &\leq e(G, H) - e\left(G, \binom{V - S'}{i-1}\right) + \left|\left\{x \in \binom{V}{k} : |x \cap S'| \geq 2\right\}\right| \\ &\leq e(G, H) - e\left(G, \binom{V - S'}{i-1}\right) + \varepsilon^2 n^k \\ &< \frac{1}{2}\left(e\left(\binom{V}{k}, H\right) - e\left(\binom{V}{k}, \binom{V - S'}{i-1}\right)\right) + 3\varepsilon^2 n^k \qquad (2)\end{aligned}$$

On the other hand, (a) implies

$$\begin{aligned}\sum_v e(G_v, H_v) &\geq \sum_{v \in S'} \frac{1}{2} e\left(\binom{V}{k-1}, H(v)\right) + |S'| 2\delta^2 n^{k-1} \\ &\geq \frac{1}{2}\left(e\left(\binom{V}{k}, H\right) - e\left(\binom{V}{k}, \binom{V - S'}{i-1}\right)\right) - 3\varepsilon^2 n^k + 2\varepsilon\delta^2 n^k\end{aligned}$$

Together with (2) it implies:

$$\delta^2 < 3\varepsilon$$

which is impossible. This completes the proof for Case 1. ∎

Case 2. $i = k$.

Recall that for two vertices u and v, we define the $(k-1)$-graph $G_{u,v}$ with the same vertex set and $E(G_{v,v}) = \left\{y \in \binom{V}{k-1} : \mu_G(u, y) = \mu_G(v, y)\right\}$.

$$\begin{aligned}dev(G) &= \frac{1}{n^{2k}} \sum_{u_1, \ldots, u_{2k}} \Pi\, \mu_G(\varepsilon_1, \ldots, \varepsilon_k) \\ &= \frac{1}{n^{2k}} \sum_{u_1, u_2} \sum_{u_3 \cdots u_{2k}} \Pi \mu_{u_1, u_2}(\varepsilon_2, \ldots, \varepsilon_k) \\ &= \frac{1}{n^2} \sum_{u_1, u_2} dev(G_{u_1, u_2})\end{aligned}$$

For each $u_1 \in V$, we define $S_{u_1} = \{u_2: dev\ G_{u_1,u_2} \geq (4^k - 2)\varepsilon^{1/2^k}\}$.

$$S = \{u_1 \in V: |S_{u_1}| > 2\varepsilon n\}$$

If $|S| \leq 2\varepsilon n$, then we have

$$dev(G) \leq \frac{1}{n^2}(|S|n + (4^k - 2)\varepsilon^{1/2^k} n^2) \leq 4^k \varepsilon^{1/2^k}$$

We may assume $|S| > 2\varepsilon n$. For a fixed $u_1 \in S$, the induction assumption implies that there exists a $(k-2)$-graph $H(u_1, u_2)$ for each u_2 in S_{u_1}, satisfying

$$\left| e(G_{u_1,u_2}, H(u_1, u_2)) - \frac{1}{2} e\left(\binom{V}{k-1}, H(u_1, u_2)\right) \right| > 6\delta^2 n^{k-1}$$

where $\delta^2 = 4\varepsilon$.

We can choose a subset S'_{u_1} of S_{u_1} with $|S'_{u_1}| = \varepsilon n$ so that either

(a): $$e(G_{u_1,u_2}, H(u_1, u_2)) \geq \frac{1}{2} e\left(\binom{V}{k-1}, H(u_1, u_2)\right) + 3\delta^2 n^{k-1}$$

for all u_2 in S'_{u_1},

or (b): $$e(G_{u_1,u_2}, H(u_1, u_2)) \leq \frac{1}{2} e\left(\binom{V}{k-1}, H(u_1, u_2)\right) - 3\delta^2 n^{k-1}$$

for all u_2 in S'_{u_1}.

We will just treat (a) and omit the similar proof for (b).

We consider the following $(k-1)$-graphs H', with $V(H') = V(G)$ and

$$E(H') = \{u_2 \cup y: y \in E(H(u_1, u_2)) \text{ for } u_2 \in S'_{u_1}\} \cap \left(E(G_{u_1}) \cup \binom{V - S'_{u_1}}{k-1}\right).$$

For each set $x \in E(G, H')$, one of the three situations occurs:

(i) $|x \cap S'_{u_1}| \geq 2$.

There are at most $\varepsilon^2 n^k$ such x.

(ii) $x = u_2 \cup y$ for $u_2 \in S'_{u_1}$, $y \cap S' = \emptyset$, and so, $y \in e(G_{u_1,u_2}, H(u_1, u_2))$.

(iii) $|x \cup S'_{u_1}| = 0$. Thus $x \in E(H, H'(u))$ where $E(G'(u_1)) = E(G_{u_1}) \cap \binom{V - S'_{u_1}}{k-1}$.

Similarly, we define a $(k-1)$-graph $G''(u_1)$ with edge set $E(\bar{G}(u_1)) \cap \binom{V - S'_{u_1}}{k-1}$ and a $(k-1)$-graph H'' with edge set $E(H'') = \{u_2 \cup y: y \in E(H(u_1, u_2))$ for $u_2 \in S'_{u_1}\} \cup E(G''(u_1))$.

Therefore

$$\sum_{u_2 \in S'_{u_1}} e(G_{u_1,u_2}, H(u_1, u_2))$$

$$= \sum_{u_2 \in S'_{u_1}} |\{v \cup y: y \in E(G_{u_1,u_2}, H(u_1, u_2))\}|$$

$$\leq \left| \left\{ x \in \binom{V}{k}: |x \cap S'_{u_1}| = 1, x = u_2 \cup y, x \in E(H), y \in E(G_{u_1,u_2}, H(u_1, u_2)) \right\} \right|$$

$$+\left|\left\{x \in \binom{V}{k} : |x \cap S'_{u_1}| = 1, x = u_2 \cup y, x \in E(\bar{H}), y \in E(G_{u_1,u_2}, H(u_1, u_2))\right\}\right|$$

$$+\left|\left\{Z \in \binom{V}{k} : |Z \cap S'| \geq 2\right\}\right|$$

$$\leq e(G, H') - e(G, G'(u_1)) + e(\bar{G}, H'') - e(\bar{G}, G''(u_1))$$

$$+\left|z \in \binom{V}{k} : |z \cap S'| \geq 2\right\}\right|$$

We now apply the induction assumption to H' and $G'(u_1)$, (and H'' and $G''(u_1)$, respectively). We have

$$e(G, H') \leq \frac{1}{2}\, e\left(\binom{V}{k}, H'\right) + \varepsilon^2 n^k,\ e(G, G(u_1)) \leq \frac{1}{2}\, e\left(\binom{V}{k}, G'(u_1)\right) + \varepsilon^2 n^k\,,$$

$$e(\bar{G}, H'') \leq \frac{1}{2}\, e\left(\binom{V}{k}, H''\right) + \varepsilon^2 n^k \text{ and } e(\bar{G}, G''(u_1)) \leq \frac{1}{2}\, e\left(\binom{V}{k}, G''(u_1)\right) + \varepsilon^2 n^k\,.$$

Therefore

$$\sum_{u_2 \in S'_{u_1}} e(G_{u_1,u_2}, H(u_1, u_2))$$

$$\leq \frac{1}{2}\left(e\left(\binom{V}{k}, H'\right) - e\left(\binom{V}{k}, G'(u_1)\right) + e\left(\binom{V}{k}, H''\right) - e\left(\binom{V}{k}, G''(u_1)\right)\right) + 5\varepsilon^2 n^k \tag{3}$$

and, on the other hand, (a) implies

$$\sum_{u_2 \in S'_{u_1}} e(G_{u_1,u_2}, H(u_1, u_2))$$

$$> \frac{1}{2} \sum_{u_2 \in S'_{u_1}} e\left(\binom{V}{k-1}, H(u_1, u_2)\right) + 3\delta^2 |S_{u_1}| n^k$$

$$\geq \frac{1}{2}\left(e\left(\binom{V}{k}, H'\right) - e\left(\binom{V}{k}, G(u_1)\right) + e\left(\binom{V}{k}, H''\right) - e\left(\binom{V}{k}, G''(u_1)\right)\right)$$

$$- 5\varepsilon^2 n^{k+1} + 3\delta^2 \varepsilon n^{k+1}$$

Together with (3) this implies $\delta^2 < 4\varepsilon$ which is a contradiction. Therefore we have completed the proof of $R_i \Rightarrow P_1$ ∎

8. PROOF OF THEOREM 2

It is easy to see that $P_i \Rightarrow P_{i-1}$ for any i. Therefore we have $\mathcal{A}_i \supset \mathcal{A}_{i+1}$. To show $\mathcal{A}_i \underset{\neq}{\supset} \mathcal{A}_{i+1}$ for $i = 0, \ldots, k-1$, we will construct k-graphs G_i with the property that $G_i \in \mathcal{A}_i$ and $G_i \not\in \mathcal{A}_{i+1}$. The basic building blocks are quasi-random graphs in $\mathcal{A}_k$.

In [9], two families of quasi-random k-graphs are given. One example is the

Paley k-graph P_k with $V(P_k)=\{1,2,\ldots,p\}$ and $\mu_{P_k}(u_1,\ldots,u_k)=1$ if $u_1+\cdots+u_k$ is a quadratic residue modulo p.

For each i, we choose a quasi-random i-graph H_i on n vertices, say H_i, to be the Paley graph P_i. We define the k-graph G_i as follows:

$$V(G_i)=V(H_i)=V$$

$$E(G_i)=\left\{x\in\binom{V}{k}:\left|\binom{x}{i}\cap E(H_i)\right|\equiv 0(\bmod 2)\right\}$$

It suffices to show $G_i\in\mathscr{A}_i-\mathscr{A}_{i-1}$.

Claim 1. $G_i\in\mathscr{A}_i$.

Proof. The proof is a straightforward application of the following character sum inequality of Burgess [6] (see also Weil [32]). Let χ denote the nonprincipal character modulo p given by

$$\chi(a)=\begin{cases}1 & \text{if } a \text{ is a quadratic residue modulo } p\\ -1 & \text{otherwise}\end{cases}.$$

Then for distinct $a_1,\ldots,a_s$ in $\mathscr{Z}_p$,

$$\left|\sum_{x\in\mathscr{Z}_p}\chi(x+a_1)\cdots\chi(x+a_s)\right|\le(s-1)\sqrt{p}$$

Now we consider

$$dev_i(G_i)=\frac{1}{n^{k+i}}\sum_{u_1,\ldots,u_{2i},v_{i+1},\ldots,v_k}\Pi\,\mu_{G_i}(\varepsilon_1,\ldots,\varepsilon_i,v_{i+1},\ldots,v_k)$$

Let $[k]$ denote $\{1,\ldots,k\}$. We note that for $z_1,\ldots,z_k$ all distinct

$$\mu_{G_i}(z_1,\ldots,z_k)=\prod_{S\in\binom{[k]}{i}}\chi\Big(\sum_{j\in S}z_j\Big)$$

Therefore

$$dev_i(G_i)=\sum_{v_{i+1}\cdots v_k}\sum_{u_1,\ldots,u_{2i}}\prod_{\varepsilon_1}\cdots\prod_{\varepsilon_i}\prod_{S\in\binom{[k]}{i}}\chi\Big(\sum_{j\in S}z_j\Big)+O(n^{-1})$$

where $z_j=\varepsilon_j$ if $j\le i$ and $z_j=v_j$ if $j>i$.

We have

$$\begin{aligned}dev_i(G_i)&\le\frac{1}{n^{k+i}}\sum_{v_{i+1}\cdots v_k}O(n^{2i-1/2})+O(n^{-1})\\&=O(n^{-1/2})\end{aligned}$$

Therefore G_i satisfies Property P_i and hence is in $\mathscr{A}_i$. Claim 1 is proved. ∎

Claim 2. *G_i is not in $\mathscr{A}_{i+1}$.*

Proof. We consider the set $E(G_i, H_i)$ of edges of G_i induced by the Paley graph H_i. An edge x is in $E(G_i, H_i)$ means every i-subset of x has a sum which is a quadratic nonresidue. By definition, x contains an even number of i-sets each of which has a sum which is quadratic nonresidue. This can only happen when $\binom{k}{i} \equiv 0 (\mathrm{mod}\, 2)$. Therefore either $E(G_i, H_i)$ is empty or $E(\bar{G}_i, H_i)$ is empty. Since k and i are all fixed integers,

$$\begin{aligned} |E(G_i, H_i) - E(\bar{G}_i, H_i)| &= \left|E\left(\binom{V}{k}, H_i\right)\right| \\ &= (1 + o(1)) \frac{n^k}{2^{\binom{k}{i}}} \\ &\neq o(n^k) \end{aligned}$$

■

9. PROOF FOR THEOREM 6

In order to construct functions (or "ordered" hypergraphs) with density $\alpha < 1/2$, we need the following variation of the characters and character sum inequalities.

Let p denote a prime number. Suppose H is a subgroup of $GF(p)^*$ with index α^{-1} (i.e., $\alpha = |H|/(p-1)$). Let Φ_H denote the set of all nontrivial characters χ from $GF(p)^*$ to C^* such that $\chi/H = 1$ and $\chi(0) = 0$.

We define

$$\Phi = \alpha \cdot \sum_{\chi \in \Phi_H} \chi$$

It is not difficult to see that

$$\phi(x) = \begin{cases} 1 - \alpha & \text{if } x \in H \\ -\alpha & \text{if } x \notin H \end{cases}$$

and $|\Phi_H| =$ index of $H = \alpha^{-1} - 1$ (also see [22] and [26]).

Therefore

$$\begin{aligned} \sum_{x \in GF(p)} \phi(x + a) &= \alpha \sum_{\chi \in \Phi_H} \sum_{x \in GF(p)} \chi(x + a) \\ &\leq \alpha \sum_{\chi \in \Phi_H} \left| \sum_{x \in GF(p)} \chi(x + a) \right| \\ &\leq \alpha |\Phi_H| \sqrt{p} \\ &\leq (1 - \alpha) \sqrt{p} \end{aligned}$$

For a distinct $a_1, \ldots, a_s \in GF(p)$, we have

$$\begin{aligned} \sum_{x \in GF(p)} \phi(x + a_1) \cdots \phi(x + a_s) &= \alpha^s \sum_{\chi_i \in \Phi_H} \sum_{x \in GF(p)} \chi_1(x + a_1) \cdots \chi_s(x + a_s) \\ &\leq \alpha^s (\alpha^{-1} - 1)^s ()s - 1)\sqrt{p} = (1 - \alpha)^s (s - 1)\sqrt{p} \end{aligned}$$

since the nontrivial characters $\chi_1, \ldots, \chi_s$, we have (see [32]

$$\left| \sum_{x \in GF(p)} \chi_1(x + a_1) \cdots \chi_s(x + a_s) \right| \leq (s-1)\sqrt{p} .$$

Now, if we define $f_\alpha : (GF(p))^k \to \{(1-\alpha), -\alpha\}$ by $f_\alpha(u_1, \ldots, u_k) = \phi(u_1 + \cdots + u_k)$. Then

$$\begin{aligned} dev_i f_\alpha &= \frac{1}{p^{k+i}} \sum_{u_1, \ldots, u_{2i}} \sum_{v_{i+1} \cdots v_k} \prod_{\varepsilon_1} \prod_{\varepsilon_2} \cdots \prod_{\varepsilon_i} \phi(\varepsilon_1 + \varepsilon_2 + \cdots + \varepsilon_i + v_{i+1} + \cdots + v_k) \\ &\leq \frac{1}{p^{k+i}} \sum_{u_2 \cdots u_{2i}} \sum_{v_{i+1} \cdots v_k} \left| \sum_{u_1} \prod_{\varepsilon_1} \prod_{\varepsilon_2} \cdots \prod_{\varepsilon_i} \phi(\varepsilon_1 + \varepsilon_2 + \cdots + \varepsilon_i + v_{i+1} + \cdots + v_k) \right| \\ &\leq \frac{1}{p} (1-\alpha)^{2^i} (2^i - 1)\sqrt{p} = O(p^{-1/2}) \end{aligned}$$

Therefore f_α is in $\mathscr{A}_{i,\alpha}$.

To distinguish $\mathscr{A}_{i,\alpha}$ from $\mathscr{A}_{i-1,\alpha}$ we consider

$$f_{\alpha,i}(u_1, \ldots, u_k) = \phi(u_{i+1} + \cdots + u_k)$$

$$\begin{aligned} dev_i f_{\alpha,i} &= \operatorname*{Max}_{I \in \binom{[k]}{i}} \frac{1}{n^{k+i}} \sum_{u_1 \cdots u_{k+i}} \Pi_{f_{\alpha,i},I}(u_1, \ldots, u_{k+i}) \\ &= \frac{1}{n^{k+i}} \sum_{u_1 \cdots u_{k+i}} \Pi_{f_{\alpha,i},I_0}(u_1, \ldots, u_{k+i}) \end{aligned}$$

where $I_0 = \{1, \ldots, i\}$. Since $\Pi_{f_{\alpha,i},I_0}(u_1, \ldots, u_{k+i}) \geq \min\{\alpha^k, (1-\alpha)^k\}$, we have $dev_i f_{\alpha,i} \neq o(1)$. Therefore $f_{\alpha,i} \not\in \mathscr{A}_{i,\alpha}$.

On the other hand,

$$\begin{aligned} dev_{i-1} f_{\alpha,i} &= \operatorname*{Max}_{I' \in \binom{[k]}{i-1}} \frac{1}{n^{k+i-1}} \sum_{u_1 \cdots u_{k+i-1}} \Pi_{f_{\alpha,i}},I'(u_1, \ldots, u_{k+i}) \\ &\leq O(p^{-1/2}) \end{aligned}$$

Therefore we have $f_{\alpha,i} \in \mathscr{A}_{i-1,\alpha}$. This completes the proof for Theorem 6. ∎

We remark that we can use variations of character sum inequalities, for example, those in [8] and [21], to get better bounds for dev_i which can be useful for proving expanding properties for functions of hypergraphs with small density α.

10. FUTURE DIRECTIONS

Numerous questions can be asked for the whole spectrum of quasi-random classes. For example, given a graph property, how is it placed in the quasi-random hierarchy? A property of special interest is the Ramsey property ($G(n)$ does not contain a clique or an independent set of size $c \log n$). In [11] we further examine the subgraph property (for fixed t, $G(n)$ contains all graphs on t vertices almost equally often) and answer the questions raised in [9].

Another direction is to search for finer classifications of quasi-random graphs. For example, between $\mathscr{A}_0$ and $\mathscr{A}_1$, many more classes can be specified, such as: $\mathscr{A}_0 = \mathscr{B}_0 \supseteq \mathscr{B}_1 \supseteq \cdots \supseteq \mathscr{B}_{k-1} = \mathscr{A}_1$ where $\mathscr{B}_i$, for $i \geq 1$, is the equivalence class which includes the property:

$$\sum_{x \in \binom{V}{i}} |d^+(x) - d^-(x)| = o(n^k) \text{ where } d^+(x) = \left|\left\{ y \in \binom{V}{k-i} : x \cup y \in E(G) \right\}\right|$$

and

$$d^-(x) = \left|\left\{ y \in \binom{V}{k-i} : x \cup y \not\in E(G) \right\}\right| .$$

In Section 6, we showed that for a k-graph G and any $(i-1)$-graph H, we have

$$|e(G, H) - e(\bar{G}, H)| \leq n^k (dev_i(G))^{1/2^i}$$

Can the power $1/2^i$ of $dev_i(G)$ in the above inequality be replaced by, say, the reciprocal of a polynomial in i? If the answer is affirmative, this can be used to improve the result on pseudo-random number generation as described in [5].

For a random k-graph G, it is not hard to verify that with probability approaching 1, we have $|e(G, H) - e(\bar{G}, H)| = O(n^{(k+i)/2})$ for every i-graph H and this is best possible. Also it can be shown $dev_i(G) = O(n^{-1})$, which can be proved by similar methods as in [15].

Finally, a general direction is to extend the concept of quasi-randomness to *other* combinatiorial structures. Much more remains to be explored.

ACKNOWLEDGMENTS

The author is indebted to Prasad Tetali for pointing out that the two examples of quasi-random hypergraphs in [9] have large communication complexity which led to the connection of these two topics. Thanks are due to Ron Graham for many inspiring comments. The author is grateful to Hendrick Lenstra and Nick Katz for illuminating discussions on character sums.

REFERENCES

[1] A. V. Aho, J. D. Ullman, and M. Yannakakis, On notions of information transfer in VLSI circuits, *Proc. 15th Symposium on Theory of Computing*, 133–139 (1983).

[2] M. Ajtai, J. Komlós, and E. Szemerédi, Sorting in $c \log n$ parallel steps, *Combinatorica*, **3**, 1–9 (1983).

[3] N. Alon and F. R. K. Chung, Explicit constructions of linear-sized tolerant networks, *Discrete Math.* **72**, 15–20 (1988).

[4] L. Babai, P. Frankl, and J. Simon, Complexity classes in communication complexity theory, *Foundation of Computer Science*, 337–347 (1986).

[5] L. Babai, N. Nisan, and M. Szegedy, Multi-party protocols and logspace-hard pseudorandom sequences, *Twenty-First Symposium on the Theory of Computing*, 1–11 (1989).

[6] D. A. Burgess, On character sums and primitive roots, *Proc. London Math. Soc.*, **12**, 179–192 (1962).

[7] A. K. Chandra, M. L. Furst, and R. J. Lipton, Multiparty protocols, *Foundation of Computer Science*, 94–99 (1983).

[8] F. R, K. Chung, Diameters and eigenvalues, *J. AMS*, **2**, 187–196 (1989).

[9] F. R. K. Chung and R. L. Graham, Quasi-random hypergraphs, *Random Struct. Algorithms*, **1**, 105–124 (1990).

[10] F. R. K. Chung and R. L. Graham, Maximum cuts and quasi-random graphs, preprint.

[11] F. R. K. Chung and R. L. Graham, Quasi-random set systems, preprint.

[12] F. R. K. Chung, R. L. Graham, and R. M. Wilson, Quasi-random graphs, *Combinatorica*, **9**, 345–362 (1989).

[13] P. Erdös and A. Hajnal, On spanned subgraphs of graphs, *Beitrage zur Graphentheorie und deren Anwendungen, Kolloq. Oberhof (DDR)*, 80–96 (1977).

[14] P. Erdös and V. T. Sós, On Ramsey-Turán type theorems for hypergraphs, *Combinatorica*, **2**, 289–295 (1982).

[15] P. Erdös and J. Spencer, Imbalances in k-colorations, *Networks*, **1**, 379–385 (1972).

[16] P. Erdös and J. Spencer, *Probabilistic Methods in Combinatorics*, Akadémiai Kiadó, Budapest, 1974.

[17] P. Frankl and V. Rödl (personal communication).

[18] P. Frankl, V. Rödl, and R. M. Wilson, The number of submatrices of given type in a Hadamard matrix and related results (to appear).

[19] R. L. Graham and J. H. Spencer, A constructive solution to a tournament problem, *Can. Math. Bull.*, **14**, 45–48 (1971).

[20] J. Haviland and A. Thomason, Pseudo-random hypergraphs, *Discrete Math.* **75**, 255–278 (1989).

[21] N. M. Katz, An estimate for character sums, *J. AMS*, **2**, 197–200 (1989).

[22] R. Lidl and H. Niederreiter, *Finite Fields, Encyclopedia of Mathematics and its Applications*, Cambridge University Press, Cambridge, England, 1984.

[23] L. Lovász, Communication complexity, in *Paths, Flows and VLSI Design, Proceedings of the Bonn Conference* (1989).

[24] G. A. Margulis, Explicit constructions of concentrators, *Prob. Per. Infor.*, **9**, 71–80 (1973). English translation in *Problems of Infor. Trans.*, 325–332 (1975).

[25] V. Rödl, On the universality of graphs with uniformly distributed edges, *Discrete Math.*, **59**, 125–134 (1986).

[26] J. P. Serre, *Linear Representations of Finite Groups*, Springer-Verlag, New York, 1982.

[27] M. Simonovits and V. T. Sós, Szemerédi partition and quasi-randomness, preprint.

[28] M. Sipser, Expanders, randomness, or time versus space, *JCSS*, 379–383 (1988).

[29] A. Thomason, Random graphs, strongly regular graphs and pseudo-random graphs, in *Surveys in Combinatorics* 1987, C. Whitehead, Ed., LMS Lecture Notes Series 123, Cambridge University Press, Cambridge, England, 1987, pp. 173–196.

[30] A. Thomason, Pseudo-random graphs, in *Proceedings of Random Graphs*, Poznán, 1985, M. Karonski, Ed., *Ann. Discrete Math.*, **33**, 307–331 (1987).

[31] U. V. Vazirani, Towards a strong communication complexity theory of generating quasi-random sequences from two communicating semirandom sources, *Proc. 17th STOC*, 366–378 (1985).

[32] A. Weil, Sur les courbes algébrique et les variétés que s'en déduisent, *Actualités Sci. Ind.*, **1041** (1948).

[33] R. M. Wilson, Cyclotomy and difference families in abelian groups, *J. Number Th.*, **4**, 17–47 (1972).

[34] R. M. Wilson, Constructions and use of pairwise balanced designs, in *Combinatorics*, M. Hall, Jr. and J. H. van Lint, Eds., *Math. Centre Tracts*, **55**, 18–41, Amsterdam (1974).

[35] A. C. Yao, Some complexity questions related to distributed computing, *Proc. 11th STOC*, 209–312 (1979).

Received January 19, 1990