Discrete Mathematics 111 (1993) 125–129
North-Holland


# On hypergraphs having evenly distributed subhypergraphs

F.R.K. Chung
*Bell Communications Research, Inc., Morristown, New Jersey 07960, USA*

R.L. Graham
*AT&T Bell Laboratories, Murray Hill, New Jersey 07974, USA*

Received 22 July 1991

*Abstract*

Chung, F.R.K. and R.L. Graham, On hypergraphs having evenly distributed subhypergraphs, Discrete Mathematics 111 (1993) 125–129.

## 1. Introduction and preliminaries

Let $k$ be a fixed positive integer. By a *k-uniform hypergraph G*, or *k-graph* for short, we mean a pair $(V, E)$, where $V=V(G)$ is a set, called the *vertices* of $G$, *and* $E=E(G)$ is a subset of $\binom{V}{k}$, the $k$-element subsets of $V$, called the *edges* of $G$ (for a full discussion of hypergraphs, see [1]). If $V$ has cardinality $|V|=n$. we denote this by writing $G=G(n)$.

For a $k$-graph $G'=(V', E')$, we say that $G'$ is an *induced subgraph of* $G$, written as $G'<G$, if there is a mapping $\lambda: V'\to V$ such that $X\in E$ if and only if $\lambda(X)\in E'$ (where for $X\in\binom{V}{k}$, $\lambda(X):=\bigcup_{x\in X}\lambda(x)$). We denote by $\#\{G'<G\}$ the number of such (ordered) mappings.

If $\mathcal{G}=\{G(n)|n\to\infty\}$ is a family of $k$-graphs, we say that $\mathcal{G}$ *satisfies* $U(r)$ if, for each $k$-graph $G'(r)$ on $r$ vertices,

$$\#\{G'(r)<G(n)\}=(1+o(1))n^r/2^{\binom{r}{k}},\quad n\to\infty. \tag{1}$$

Thus, $\mathcal{G}$ satisfies $U(r)$ if and only if all $r$-vertex $k$-graphs occur as (ordered) induced subgraphs of $G(n)$ asymptotically equally often as $n\to\infty$.

*Correspondence to*: R.L. Graham, AT&T Bell Laboratories, 600 Mountain Ave., Room 2T-102, Murray Hill, NJ 07974, USA.

In particular, if each $G(n)$ is a *random k-graph* $G_{1/2}(n)=(V(n),E(n))$ on $n$ vertices, i.e., each $X\in\binom{V(n)}{k}$ is chosen as an edge of $G_{1/2}(n)$ independently with probability 1/2, then the corresponding family $\mathscr{G}_{1/2}$ almost certainly satisfies $U(r)$ for any fixed $r$ (i.e., satisfies $U(r)$ with probability tending to 1 as $n\to\infty$).

It is not difficult to see that if $\mathscr{G}$ satisfies $U(r)$ then $\mathscr{G}$ also satisfies $U(s)$ for any $s\leqslant r$. On the other hand, it is perhaps unexpected that it is possible to reverse this implication once $s$ is as large as $2k$. More precisely, it was shown in [2] that:

$$\text{If } \mathscr{G} \text{ satisfies } U(2k) \text{ then } \mathscr{G} \text{ satisfies } U(r) \text{ for } \textit{any} \text{ fixed } r. \tag{2}$$

Families $\mathscr{G}$ satisfying (2) have been termed *quasi-random*, since it is known that they must necessarily also satisfy a large collection of other properties all shared by families of random $k$-graphs (for details, see [2–4]).

However, it was noted in [2] that (2) is no longer valid if $U(2k)$ is replaced by $U(k+1)$. The main purpose of this note is to close this gap completely, by showing that (2) no longer holds even if we assume $\mathscr{G}$ satisfies $U(2k-1)$. More generally, for each $s$, with $k\leqslant s\leqslant 2k-1$, there are families $\mathscr{G}_s$ which satisfy $U(s)$ but not $U(s+1)$. A less direct proof for this construction appears in [3].

## 2. The main construction

If $G=(V,E)$ is a $k$-graph, we let

$$\chi=\chi_G:\binom{V}{k}\to\{0,1\} \text{ be the } \textit{edge function} \text{ for } G,$$

defined for $X\in\binom{V}{k}$ by

$$\chi(X)=\begin{cases}1 & \text{if } X\in E,\\ 0 & \text{if } X\notin E.\end{cases}$$

For $a\geqslant 0$, we define the *coboundary operator* $\delta^{(a)}$ mapping $k$-graphs on $V$ to $(k+a)$-graphs on $V$ as follows. If $G=(V,E)$ is a $k$-graph with edge function $\chi$, then $\delta^{(a)}G=(V,E^{(a)})$ is a $(k+a)$-graph with edge function $\chi^{(a)}$, given, for $Y\in\binom{V}{k+a}$, by

$$\chi^{(a)}(Y)\equiv\sum_{X\in\binom{Y}{k}}\chi(X)\ (\text{mod}\,2). \tag{3}$$

Thus, $Y$ is an edge of $\delta^{(a)}G$ if and only if $Y$ contains an *odd* number of edges $X$ of $G$ as subsets.

For $1\leqslant j\leqslant k-1$, choose a random $j$-graph $G_{1/2}^{(j)}$ on $V$ and 'lift' it to a $k$-graph $G_j:=\delta^{(k-j)}G_{1/2}^{(j)}$ on $V$ with edge function $\chi_j$. Next, form the 'symmetric difference' $k$-graph $G^*(n)=(V,E^*(n))=\nabla_{j=1}^{k-1}G_j$ with edge function $\chi^*$, defined by

$$\chi^*(X)\equiv\sum_{j=1}^{k-1}\chi_j(X)\ \ (\text{mod } 2)\quad \text{for } X\in\binom{V}{k}.$$

**Theorem 2.1.** *For almost all choices of* $G_{1/2}^{(j)}$, $\mathscr{G}^*=\{G^*(n)\mid n\to\infty\}$ *satisfies* $U(2k-1)$ *but not* $U(2k)$, *provided* $k\neq 2^t$, $t\geq 0$.

**Proof.** Consider an arbitrary fixed set $W=\{w_1,w_2,\ldots,w_{2k-1}\}$ of $2k-1$ vertices of $V$. Form a matrix $M$ with rows indexed by $X\in\binom{W}{k}$, and columns indexed by $Y_j\in\binom{W}{j}$, $1\leq j\leq k-1$, with the $(X,Y_j)$-entry $M(X,Y_j)$ defined to be 1 if $Y_j\subset X$, and 0 otherwise. We can regard each column $C(Y_j)$ of $M$ as a function mapping $\binom{W}{k}$ to $\{0,1\}$ by defining

$$C(Y_j)(X)=M(X,Y_j),\quad X\in\binom{W}{k}.$$

It is easy to see that

$$\chi_j\equiv\sum_{Y_j\in\binom{W}{j}} C(Y_j)\,(\mathrm{mod}\,2).$$

so that

$$\chi^*\equiv\sum_{j=1}^{k-1}\sum_{Y_j\in\binom{W}{j}} C(Y_j)\,(\mathrm{mod}\,2).$$

We now apply a result of Wilson [6] (see also [5]) which asserts that, when $k\neq 2^t$, $M$ has mod 2 rank equal to $\binom{2k-1}{k}$. Actually, Wilson's result implies that if we adjoin a column of all 1's to form an augmented matrix $M^+$, then for any prime $p$, $M^+$ has mod $p$ rank equal to $\binom{2k-1}{k}$. However, when $k\neq 2^t$, then some $i$, with $1\leq i\leq k-1$, has $\binom{k}{i}$ odd. Summing all the columns $C(Y_i)$, $Y_i\in\binom{V}{i}$, gives us a column of all 1's (mod 2), from which it follows that $M$ itself has mod 2 rank equal to $\binom{2k-1}{k}$.

Now, as $W$ ranges over all $(2k-1)$-element subsets of $V$, since the edges of the various corresponding $G_{1/2}^{(j)}$ are chosen independently and uniformly, then an easy argument shows that almost certainly each of the possible $\binom{2k-1}{k}$ (0, 1)-vectors occurs $(1+o(1))n^{2k-1}/2^{\binom{2k-1}{k}}$ times as $n\to\infty$. But this just means that for almost all choices of the $G_{1/2}^{(j)}$, each of the possible $k$-graphs $G(2k-1)$ on $2k-1$ vertices occurs $(1+o(1))n^{2k-1}/2^{\binom{2k-1}{k}}$ times as an induced subgraph of $G^*(n)$ as $n\to\infty$. This implies that $\mathscr{G}^*$ satisfies $U(2k-1)$, as claimed.

To show that $\mathscr{G}^*$ does not satisfy $U(2k)$, we do the following. Let $Z=\{z_1(0),z_1(1),\ldots,z_k(0),z_k(1)\}$ be an arbitrary $2k$-element subset of $V$. Consider the inclusion matrix $\bar{M}$ with rows indexed by $X\in\binom{Z}{k}$, and columns indexed by $Y_j\in\binom{Z}{j}$, $1\leq j\leq k-1$. Let us restrict our attention to the $2^k$ rows indexed by all the $k$-sets of the form $Z(\varepsilon_1,\ldots,\varepsilon_k)=\{z_1(\varepsilon_1),\ldots,z_k(\varepsilon_k)\}$, where $\varepsilon_i\in\{0,1\}$, $1\leq i\leq k$. For a fixed $Y_j\in\binom{Z}{j}$, since $j\leq k-1$, there must exist at least one index $i$ such that $z_i(0)\notin Y_j$, $z_i(1)\notin Y_j$. Thus, $Z(\varepsilon_1,\ldots,\varepsilon_{i-1},0,\varepsilon_{i+1},\ldots,\varepsilon_k)\supset Y_j$ if and only if $Z(\varepsilon_1,\ldots,\varepsilon_{i-1},1,\varepsilon_{i+1},\ldots,\varepsilon_k)\supset Y_j$. This implies that for any column of $\bar{M}$, the total number of 1's in the $2^k$ rows indexed by the $Z(\varepsilon_1,\ldots,\varepsilon_k)$ is always *even*. Hence, this also holds for the mod 2 sum of any set of columns of $\bar{M}$. Consequently, $\mathscr{G}^*$ contains no indexed subgraph on a $2k$-set $Z$ in which an *odd* number of $k$-sets $Z(\varepsilon_1,\ldots,\varepsilon_k)$ are edges. This shows that $\mathscr{G}^*$ does not satisfy $U(2k)$, and Theorem 2.1 is proved. □

In the case that $k=2^t$, an additional step is required. As before, we first construct the $k$-graphs $G^*(n)=(V_n, E)$. We then take the complement $\bar{G}^*(n)=(V_n', E')$ on a disjoint vertex set $V_n'$, and form the $k$-graph $\hat{G}(2n):=(V_n \cup V_n', \hat{E})$ by defining $\hat{E}$ to be $E \cup E'$, together with a random selection of all the $k$-sets $X$ which intersect both $E$ and $E'$. That is, each such $X$ is chosen independently with probability 1/2 to be an edge of $\hat{G}(2n)$.

**Theorem 2.2.** *For almost all choices of* $G_{1/2}^{(j)}$, $\hat{\mathscr{G}}=\{\hat{G}(2n) \mid n \to \infty\}$ *satisfies* $U(2k-1)$ *but not* $U(2k)$.

**Proof.** The case not covered by Theorem 2.1 is when $k=2^t$, which we now assume. By the previously mentioned result of Wilson, if $G(2k-1)$ is a $k$-graph on $2k-1$ vertices then

$$\#\{G(2k-1) < G^*(n)\}$$

$$= \begin{cases} 2(1+o(1))n^{2k-1}/2^{\binom{2k-1}{k}} & \text{if } G(2k-1) \text{ has an } \textit{even} \text{ number of edges,} \\ 0 & \text{if } G(2k-1) \text{ has an } \textit{odd} \text{ number of edges.} \end{cases}$$

Since $\binom{2k-1}{k}$ is odd for $k=2^t$, the situation is reversed for the complement $\bar{G}^*(n)$. This implies that $\mathscr{G}^*$ satisfies $U(2k-1)$.

To see that $\mathscr{G}^*$ does not satisfy $U(2k)$, consider the $k$-graph $H=H((2k-1)^2)$ formed from disjoint copies of $H_i(2k-1)=(W_i, E_i)$, $1 \leqslant i \leqslant 2k-1$, where each $H_i(2k-1)$ is a complete $k$-graph on $2k-1$ vertices, i.e., $|W_i|=2k-1$ and $E=\binom{W_i}{k}$. We claim:

$$\#\{H < \hat{G}(2n)\} = 0 \quad \text{for all } n. \tag{4}$$

To see this, suppose the contrary. Note that we must have $W_i \nsubseteq V_n$ for $1 \leqslant i \leqslant 2k-1$, since otherwise $H_i(2k-1) < G^*(n)$, which is impossible, because each $H_i(2k-1)$ has an odd number $\binom{2k-1}{k}$ of edges. Thus, for each $i$ there is some $w_i \in V_n'$, $1 \leqslant i \leqslant 2k-1$. However, the $k$-graph induced by the vertex set $\{w_1, \ldots, w_{2k-1}\}$ has *no* edges, which contradicts the fact that $\bar{G}^*(n)$ only has induced subgraphs on $2k-1$ vertices having an odd number of edges. This proves (4). Finally, by (2) this implies that $\hat{\mathscr{G}}$ does not satisfy $U(2k)$, and the proof is complete. □

We remark that essentially the same arguments can be applied for any $s$ with $k \leqslant s \leqslant 2k-1$, i.e., showing that families $\mathscr{G}_s$ exist satisfy $U(s)$ but not $U(s+1)$. We omit the details.

## 3. Concluding remarks

It would be interesting to know whether, in fact, the cases $k=2^t$ are inherently different, or whether this is simply an artifact of the approach we have taken. For example, we do not know, for any $k=2^t$, whether there exists a family $\mathcal{G}=\{G(n)\,|\,n\to\infty\}$ of $k$-graphs satisfying $U(2k-1)$ but for which, for some $H(2k)$,

$$\#\{H(2k)<G(n)\}=0 \quad \text{for all } n.$$

## References

[1] C. Berge, Hypergraphs (North-Holland, Amsterdam, 1989).
[2] F.R.K. Chung and R.L. Graham, Quasi-random hypergraphs, Random Structures and Algorithms 1 (1990) 105–124.
[3] F.R.K. Chung and R.L. Graham, Quasi-random set systems, J. Amer. Math. Soc. 4 (1991) 151–196.
[4] F.R.K. Chung, R.L. Graham and R.M. Wilson, Quasi-random graphs, Combinatorica 9 (1989) 345–362.
[5] P. Frankl, Intersection theorems and mod $p$ rank of inclusion matrices, J. Combin. Theory Ser. A 54 (1990) 85–94.
[6] R.M. Wilson, A diagonal form for the incidence matrices of $t$-subsets vs. $k$-subsets, European J. Combin. 11 (1990) 609–615.