JOURNAL OF COMBINATORIAL THEORY, Series A 49, 180–187 (1988)

Note

On Induced Subgraphs of the Cube

F. R. K. CHUNG

Bell Communication Research, Morristown, New Jersey 07960

ZOLTÁN FÜREDI

AT & T Bell Laboratories, Murray Hill, New Jersey 07974,
Mathematical Institute of the Hungarian Academy of Science,
1364 Budapest Pf. 127, Hungary, and
Department of Mathematics, Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139

R. L. GRAHAM

AT & T Bell Laboratories, Murray Hill, New Jersey 07974

AND

P. SEYMOUR

Bell Communication Research, Morristown, New Jersey 07960

Communicated by the Managing Editors

Received April 4, 1987

Consider the usual graph Q^n defined by the n-dimensional cube (having 2^n vertices and $n2^{n-1}$ edges). We prove that if G is an induced subgraph of Q^n with more than 2^{n-1} vertices then the maximum degree in G is at least $(\frac{1}{2}-o(1))\log n$. On the other hand, we construct an example which shows that this is not true for maximum degree larger than $\sqrt{n}+1$. © 1988 Academic Press, Inc.

1. PRELIMINARIES

Denote by Q^n the graph of the *n-dimensional cube*, i.e., the vertex set of Q^n consists of all the $(0, 1)$-vectors of length n, and two vectors $x, y \in \{0, 1\}^n$ are adjacent if they differ from each other in exactly one

0097-3165/88 $3.00

component. For a graph $G=(V, E)$ we denote the *maximum degree* by $\Delta(G)$, i.e.,

$$\Delta(G)=\max_{v\in V(G)} \deg_G(v).$$

The *average degree* $\bar{d}(G)$ is defined to be $\sum_{v\in V(G)} \deg_G(v)/|V(G)|$. We say $G\in Q^n(N)$ if G is an induced subgraph of Q^n with N vertices, i.e., $|V(G)|=N$, $V(G)\subseteq\{0,1\}^n$, and $E(G)=E(Q^n)\cap(V(G)\times V(G))$.

Q^n is a bipartite graph, so we have a $G\in Q^n(2^{n-1})$ without any edge, namely, G^n_{odd} and G^n_{even}, where $V(G^n_{\text{odd}})=\{x\in\{0,1\}^n: \sum_{i=1}^n x_i\equiv 1 \pmod 2\}$, $V(G^n_{\text{even}})=\{0,1\}^n - V(G^n_{\text{odd}})$. Our main result shows that even though the average degree of a graph $G\in Q^n(2^{n-1}+1)$ can be very small (only $2n/(2^{n-1}+1)$), these graphs must have large degree.

THEOREM 1.1. *Let G be an induced subgraph of Q^n with at least $2^{n-1}+1$ vertices. Then for some vertex v of G we have*

$$\deg_G(v)>\tfrac{1}{2}\log n-\tfrac{1}{2}\log\log n+\tfrac{1}{2}. \tag{1.1}$$

On the other hand, there exists a $G\in Q^n(2^{n-1}+1)$ with

$$\Delta(G)<\sqrt{n}+1. \tag{1.2}$$

2. Related Results and Problems from Computer Science

A (Boolean) function $f\colon\{0,1\}^n\to\{0,1\}$ is said to *depend on coordinate i* if there exists an input vector x such that $f(x)$ differs from $f(x^{(i)})$, where $x^{(i)}$ agrees with x in every coordinate except the ith. In this case x is said to be *critical* for f with respect to i. The function f is called *nondegenerate* if it depends on all n coordinates. For an input vector x, let $c(f,x)$ denote the number of coordinates i such that x is critical for f with respect to i, and let $c(f):=\max\{c(f,x): x\in\{0,1\}^n\}$. $c(f)$ is called the *critical complexity* of f. This notion is due to Cook and Dwork [3] and Reischuk [5], who showed that $\log_A c(f)$ is a lower bound to the time needed by a parallel RAM to compute the function f (where $A=\frac{1}{2}(5+\sqrt{21})=4.7\cdots$). (A parallel RAM is a collection of synchronous parallel processors sharing a global memory with no write-conflicts allowed. For precise delfinitions, see [1].) Simon [6] showed that the critical complexity of any nondegenerate Boolean function is at least

$$\alpha(n):=\tfrac{1}{2}\log n-\tfrac{1}{2}\log\log n+\tfrac{1}{2}, \tag{2.1}$$

which implies a $O(\log\log n)$ lower bound for parallel complexity. More results on this topic can be found in [7].

Call a subgraph G of Q^n *nondegenerate* if $E(G)$ contains edges from each of the n directions. Thus, the crucial point of the above problem can be reformulated as follows:

> Let U, V be a partition of $\{0, 1\}^n$ and consider the induced bipartite graph $G(U, V)$. If $G(U, V)$ is nondegenerate then $\Delta(G) > \alpha(n)$. (2.2)

This is completely analogous to our theorem (even the proof is similar). However, we need a slightly more powerful lemma (see Lemma 4.1). Reischuk (see [5]) has a simple example proving that in (2.2), $\Delta(G) = \lfloor \log n \rfloor + 2$ is possible, and it is very likely that this is the right value of

$$b(n) = \min\{\Delta(G)\colon G \text{ as is in (2.2)}\}.$$

Another interesting property of the induced bipartite graphs is proved by Ben-Or and Linial [2] (also dealing with a problem arising in theoretical computer science):

> If U, V is a partition of $\{0, 1\}^n$ then there exists a direction i such that at least $\min\{|U|, |V|\}/n$ edges go from U to V parallel to i. (2.3)

They have an upper bound of $\log n \min\{|U|, |V|\}/n$ and also this seems to be the right order of magnitude.

3. Proof of the Upper Bound

Denote the set of integers $\{1, 2, ..., n\}$ by $[n]$. Since there is a natural bijection between $\{0, 1\}^n$ and $2^{[n]}$, so we will speak about families of finite sets with the underlying set $[n]$. There exists a partition of $[n] = F_1 \cup \cdots \cup F_k$ such that $|k - \sqrt{n}| < 1$ and $|\,|F_i| - \sqrt{n}| < 1$, $1 \leqslant i \leqslant k$. Define the family $\mathbf{X}$ as follows: consider all the even sets (i.e., subsets of $[n]$ with cardinality an even number) which contain some F_i, $1 \leqslant i \leqslant k$, and all the odd sets which do not contain any F_i.

Claim 3.1. $|\mathbf{X}| = 2^{n-1} \pm 1$ according to whether $n + k$ is odd or even.

Claim 3.2. For the subgraphs induced by $\mathbf{X}$ and $2^{[n]} - \mathbf{X}$ we have $\Delta \leqslant k$.

Remark. We can generalize the above construction in the following way. Let $\mathbf{F} \subset 2^{[n]}$ be a collection of finite sets. (Later we will see that it is enough to consider Sperner families, with $\bigcup \mathbf{F} = [n]$.) Define

$$\mathbf{X}(\mathbf{F}) = \{S \subset [n]: |S| = \textit{even}, \text{ and there exists } F \in \mathbf{F} \text{ with } F \subset S\}$$
$$\cup \{S \subset [n]: |S| = \textit{odd}, F \setminus S \neq \varnothing \text{ for all } F \in \mathbf{F}\}.$$

Let $G(\mathbf{F})$ be the induced subgraph of Q^n with vertex set $\mathbf{X}(\mathbf{F})$, and $G'(\mathbf{F})$ the induced subgraph on $2^{[n]} - \mathbf{F}$. The *rank* of $\mathbf{F}$ is the largest size of its edges, i.e., $r(\mathbf{F}) = \max\{|F|: F \in \mathbf{F}\}$. Denote by $t(\mathbf{F})$ the maximum value of t such that one can find $F_1, F_2, ..., F_t \in \mathbf{F}$ and $x_i \in F_i$, $1 \leqslant i \leqslant t$, so that for $i \neq j$ we have $x_i \notin F_j$. In other words, $t(\mathbf{F})$ is the largest size of the disjointly representable subsystems of $\mathbf{F}$.

PROPOSITION 3.3. *$\Delta(G(\mathbf{F})) \leqslant \max\{r(\mathbf{F}), t(\mathbf{F})\}$, and the same holds for $\Delta(G'(\mathbf{F}))$.*

Proof. If (S, S') is an edge of Q^n, $S, S' \in \mathbf{X}(\mathbf{F})$, and S is even then $S' \subsetneq S$. Moreover if $F \subset S$, $F \in \mathbf{F}$, then $(S \setminus S') \in F$, so we have

$$\deg(S) \leqslant \left| \bigcap \{F: F \in \mathbf{F}, F \subset S\} \right| \leqslant r(\mathbf{F}). \tag{3.1}$$

On the other hand, if S is odd then $S \subset S'$ so there exists an $F \in \mathbf{F}$, $F \subset S'$, $F \not\subseteq S$. Hence if $S \subset S'_1, S'_2, ..., S'_a$ then $F_1, ..., F_a$ (where $F_i \subset S'_i$) are disjointly representable, so $a \leqslant t(\mathbf{F})$. The statement $\Delta(G'(\mathbf{F})) \leqslant \max\{r(\mathbf{F}), t(\mathbf{F})\}$ can be proved in the same way. ∎

Now use the sieve method to determine the cardinality of $\mathbf{X}(\mathbf{F})$. Let $F \subset [n]$, and $n \equiv \varepsilon \pmod 2$ ($\varepsilon = 0$ or 1). Then

$$\text{the number of even sets containing } F = \begin{cases} 2^{n-|F|-1} & \text{if } |F| < n, \\ 0 & \text{if } |F| = n, \text{ and } n \text{ is odd}, \\ 1 & \text{if } |F| = n, \text{ and } n \text{ is even}. \end{cases}$$

Similarly,

$$|\{S: F \subset S \subset [n], |S| \text{ odd}\}| = \begin{cases} 2^{n-|F|-1} & \text{if } |F| < n, \\ \varepsilon & \text{if } |F| = n. \end{cases}$$

Let $\mathbf{F} = \{F_1, ..., F_N\}$. The cardinality of the first part of $\mathbf{X}(\mathbf{F})$ is

$$\sum_{i \in [N]} (2^{n-|F_i|-1})^* - \sum\sum_{\{i,j\} \subset [N]} (2^{n-|F_i \cup F_j|-1})^* + - \cdots, \tag{3.2}$$

where $(2^A)^*$ means 2^A for $A \geqslant 0$ and $1-\varepsilon$ for $A=-1$. The cardinality of the second part of $\mathbf{X}(\mathbf{F})$ is

$$2^{n-1} - \sum_{i \in [N]} (2^{n-|F_0|-1})^{**} + \sum\sum_{\{i,j\} \subset [N]} (2^{n-|F_i \cup F_j|-1})^{**} - + \cdots, \quad (3.3)$$

where $(2^A)^{**}$ means 2^A for $A \geqslant 0$ and ε for $A=-1$. We have $(2^A)^* - (2^A)^{**} = 0$ or $1-2\varepsilon$ according to whether $A \geqslant 0$ or $A=-1$. So summing up (3.2) and (3.3) we have

$$|\mathbf{X}(\mathbf{F})| = 2^{n-1} + (1-2\varepsilon)\left[\sum_{\substack{F_i \in \mathbf{F} \\ |F_i|=n}} 1 - \sum\sum_{\substack{F_i, F_j \in \mathbf{F} \\ |F_i \cup F_j|=n}} 1 + \sum\sum\sum_{\substack{F_i, F_j, F_l \in \mathbf{F} \\ |F_i \cup F_j \cup F_l|=n}} 1 - + \cdots\right]. \quad (3.4)$$

Denote by $f(\mathbf{F})$ the bracketed expression on the right-hand side of (3.4). It is clear that if $\mathbf{F}$ is a k-partition of $[n]$ (into nonempty parts) then $f(\mathbf{F}) = (-1)^{k+1}$, which implies

$$|\mathbf{X}(\mathbf{F})| = 2^{n-1} + (-1)^{n+k+1},$$

proving Claim 3.2.

In general we are not able to calculate $f(\mathbf{F})$ explicitly since it tends to get complicated. Some properties of f are:

(i) If $F_0 = [n] \in \mathbf{F}$ then $f(\mathbf{F}) = f(\mathbf{F} - \{F_0\})$;

(ii) If $F_0 = \varnothing \in \mathbf{F}$ then $f(\mathbf{F}) = 0$;

(iii) If $\mathbf{F} = \{F_0, F_1, ..., F_N\}$, $\varnothing \neq F_0 \neq [n]$ then

$$f(\{F_0, ..., F_N\} \mid [n]) = f(\{F_1, ..., F_N\} \mid [n]) - f(\{F_1 - F_0, ..., F_n - F_0\} \mid [n] - F_0);$$

(iv) If $F_0 \neq \varnothing$ and for some $F_i \supset F_0$ then $f(\mathbf{F}) = f(\mathbf{F} - \{F_0\})$.

PROPOSITION 3.4. *Suppose* $f(\mathbf{F}) \neq 0$. *Then* $\max\{r(\mathbf{F}), t(\mathbf{F})\} \geqslant \sqrt{n}$.

Proof. Suppose that $|F| < \sqrt{n}$ holds for all $F \in \mathbf{F}$. $f(\mathbf{F}) \neq 0$ implies that $|\bigcup \mathbf{F}| = n$. Let $\{F_1, ..., F_s\}$ be a minimal subfamily of $\mathbf{F}$ with $\bigcup \mathbf{F} = [n]$. Then $\{F_1, ..., F_s\}$ is disjointly representable and $s \geqslant \sqrt{n}$. ∎

However, it may be possible that using a more complicated $\mathbf{F}$ with large $f(\mathbf{F})$ and deleting some members of $\mathbf{X}(\mathbf{F})$ (but fewer than $f(\mathbf{F})$) one can obtain a $G \in \mathbf{G}^n(2^{n-1}+1)$ with $\Delta(G) \ll \sqrt{n}$.

4. Proof of the Lower Bound

We begin with a lemma.

LEMMA 4.1. *Let G be a subgraph of the cube with average degree $\bar{d}$. Then* $|V(G)| \geqslant 2^{\bar{d}}$.

A similar lemma was used in [6], where $|V(G)| \geqslant 2^{\min \deg(v)}$ was proved. We point out that a related result of Kleitman et al. [4] immediately implies Lemma 4.1 in the case that $\bar{d}$ is an integer.

Proof. We use induction on $|V(G)|$. Split Q^n into two $(n-1)$-dimensional subcubes Q_1 and Q_2 such that $V_1 = Q_1 \cap V(G) \neq \varnothing$ and $V_2 = Q_2 \cap V(G) \neq \varnothing$. Suppose that $|V_2| \geqslant |V_1|$ and there are s edges between V_1 and V_2 in G (so that $|V_1| \geqslant s$). The restriction of G to V_i, $i = 1, 2$, is denoted by G_i. The induction hypothesis gives

$$|V_i| \log |V_i| \geqslant \sum \deg_{G_i}(v) = \sum_{v \in V_i} \deg_G(v) - s,$$

so that

$$|V_1| \log |V_1| + |V_2| \log |V_2| + 2s \geqslant \sum_{v \in V(G)} \deg_G(v). \tag{4.1}$$

However,

$$\begin{aligned}(|V_1| + |V_2|) &\log(|V_1| + |V_2|) \\ &\geqslant |V_1| \log |V_1| + |V_2| \log |V_2| + 2\,|V_1|\end{aligned}$$

if $|V_2| \geqslant |V_1|$. (Here we used the fact that the base of the logarithm is 2.) ∎

Of course, Q^n is decomposable into two $(n-1)$-dimensional subcubes Q_1^i, Q_2^i, $1 \leqslant i \leqslant n$, in natural ways according to the n directions. We prove slightly more than (1.1).

LEMMA 4.2. *Suppose $G \in Q^n(2^{n-1})$* and G contains edges from all the n directions. Then $\Delta(G) > \alpha(n)$.

This immediately implies (1.1). Indeed, let $G \in Q^n(2^{n-1} + b)$ with $\Delta(G) < (n-1)/2$. Delete b vertices from G arbitrarily. In the resulting graph G_0 every direction must occur, since otherwise $\Delta(G) \geqslant (n-1)/2$ would be forced.

Proof of Lemma 4.2. Let $X_i = \{x \in V(G): x^{(i)} \in V(G)\}$, i.e., the set of endpoints of the edges of G in direction i. Define $Y_i = \{y \notin V(G): y^{(i)} \notin V(G)\}$, $A_i = V(Q^n) - X_i - Y_i$. Then

$$|X_i| = |Y_i| > 0.$$

Let $\Delta = \Delta(G)$ and consider a pair $x, x^{(i)} \in X_i$.

Claim 4.3. x has at most $(2\Delta - 2)$ neighbours in A_i.

Proof. Let us denote the neighbours of x in A_1 by $x^{(j_1)}, ..., x^{(j_s)}$. Then $x^{(j_1)(i)}, ..., x^{(j_s)(i)}$ are neighbours of $x^{(i)}$ in A_i and either $x^{(j_u)}$ or $x^{(j_u)(i)}$ belong to $V(G)$. Thus, $s \leqslant 2(\Delta - 1)$. ∎

Claim 4.3 implies that every $x \in X_i$ has at least $(n - 2\Delta + 1)$ neighbours in Y_i. Hence

$$|E(G(X_i \cup Y_i))| \geqslant \tfrac{1}{2}|X_i| + \tfrac{1}{2}|Y_i| + (n - 2\Delta + 1)|X_i|,$$

implying

$$\bar{d}(G(X_i \cup Y_i)) \geqslant n - 2\Delta + 2.$$

Lemma 4.1 gives

$$|X_i| \geqslant 2^{n - 2\Delta + 1}. \tag{4.2}$$

Counting the degrees in $V(G)$ we have

$$\Delta \cdot 2^{n-1} \geqslant \sum_{v \in V(G)} \deg_G(v) = \sum_{i=1}^{n} |X_i| \geqslant n2^{n - 2\Delta + 1}.$$

An easy calculation now gives $\Delta \geqslant \alpha(n)$, as desired.

References

1. A. V. Aho, J. E. Hopcroft, and J. D. Ullman, "The Design and Analysis of Computer Algorithms," Addison–Wesley, Reading, MA, 1974.
2. M. Ben-Or and N. Linial, Collective coin flipping, robust voting schemes and minima of Banzhaf values, *in* "Proceedings, 26th Symposium on Foundation of Computer Science, 1985," pp. 408–416.
3. S. Cook and C. Dwork, Bounds on the time for parallel RAM's to compute simple functions, *in* "Proceedings, 14th ACM STOC, 1982," pp. 231–233.
4. D. J. Kleitman, M. M. Krieger, and B. L. Rothschild, Configurations maximizing the number of pairs of Hamming-adjacent lattice points, *Stud. Appl. Math.* **50** (1971), 115–119.

5. R. REISCHUK, "A Lower Time Bound for Parallel RAM's without Simultaneous Writes," IBM Research Report RJ3431 (40917), 3/1982.
6. H.-U. SIMON, A tight $\Omega(\log \log n)$ bound on the time for parallel RAM's to compute non-degenerate Boolean functions, *in* "Foundations of Computing Theory," Proceedings, 1983 International FCT-Conference, Borgholm, Sweden (M. Karpinski, Ed.), Lecture Notes in Computer Science, Vol. 158, Springer-Verlag, Berlin, 1983, pp. 439–444.
7. G. TURÁN, The critical complexity of graph properties, *Inform. Process. Lett.* **18** (1984), 151–153.