On induced subgraphs of the Hamming graph

Dingding Dong

To the memory of Ron Graham, one of my mentor's mentors

Abstract

In connection with his solution of the Sensitivity Conjecture, Hao Huang (arXiv: 1907.00847, 2019) asked the following question: Given a graph $G$ with high symmetry, what can we say about the smallest maximum degree of induced subgraphs of $G$ with $\alpha(G) + 1$ vertices, where $\alpha(G)$ denotes the size of the largest independent set in $G$? We study this question for $H(n,k)$, the $n$-dimensional Hamming graph over an alphabet of size $k$. Generalizing a construction by Chung et al. (JCT-A, 1988), we prove that $H(n,k)$ has an induced subgraph with more than $\alpha(H(n,k))$ vertices and maximum degree at most $\left\lceil \frac{n}{2} \right\rceil$. Chung et al. proved this statement for $k = 2$ (the $n$-dimensional cube).

KEYWORDS

Hamming graph, induced subgraphs

1 INTRODUCTION

For a graph $G = (V, E)$, let $\alpha(G)$ denote the independence number of $G$ (the maximum size of an independent set). In this paper we study the quantity $f(G)$ defined as the smallest maximum degree of induced subgraphs of $G$ with $\alpha(G) + 1$ vertices. The Hamming graph $H(n,k)$ is a graph on the vertex set $\Sigma^n$ where $\Sigma$ is an alphabet of size $k$, such that two vertices are adjacent if and only if they differ in precisely one coordinate.

Notice that $H(n,2) = Q^n$ is the $n$-dimensional cube. The study of $f(Q^n)$ goes back to a 1988 paper by Chung, Füredi, Graham, and Seymour [1], who proved $\frac{1}{2}(\log n - \log \log n + 1) <$...
Their proof of the lower bound used the connection between \( f(Q^n) \) and the sensitivity of Boolean functions. Gotsman and Linial [2] made a significant further step in formalizing this connection. They showed that the inequality \( f(Q^n) \geq n^c \) for some constant \( c > 0 \) is equivalent to the Sensitivity Conjecture for Boolean functions, proposed by Nisan and Szegedy [5]. (See the survey by Hatami et al [3] on the Sensitivity Conjecture).

Hao Huang [4] recently proved that \( f(Q^n) \leq \lceil \sqrt{n} \rceil \). This lower bound, according to the result by Gotsman and Linial, confirmed the Sensitivity Conjecture for Boolean functions. In the last section of his paper, Huang asks the following question.

**Question 1.1** (Huang). What can we say about \( f(G) \) for graphs \( G \) with high symmetry?

We study this question for the Hamming graphs. Generalizing the proof of the inequality \( f(Q^n) \leq \lceil \sqrt{n} \rceil \) by Chung et al. [1], in this note we prove the following bound.

**Theorem 1.2.** For all \( k, n \geq 1 \), we have \( f(H(n, k)) \leq \lceil \sqrt{n} \rceil \). In fact, \( H(n, k) \) has a bipartite induced subgraph with maximum degree at most \( \lceil \sqrt{n} \rceil \) and more than \( \alpha(H(n, k)) \) vertices.

### 2 | PRELIMINARIES AND NOTATION

For a graph \( G = (V, E) \), define \( \Delta(G) \) to be the maximum degree of \( G \). For a subset \( W \subseteq V \), define \( G[W] \) to be the induced subgraph of \( G \) on vertex set \( W \). It will be convenient to take \( \mathbb{Z}_k = \mathbb{Z}/k\mathbb{Z} \) as the alphabet for the Hamming graph, so the set of vertices of \( H(n, k) \) is \( \mathbb{Z}_k^n \). We view the elements \( v \in \mathbb{Z}_k^n \) as functions from \( [n] = \{1,...,n\} \) to \( \mathbb{Z}_k \) and set \( v = (v(1), ..., v(n)) \). It is not hard to see that for all \( k, n \geq 1 \), we have \( \alpha(H(n, k)) = k^{n-1} \).

### 3 | CONSTRUCTING A FAMILY OF INDUCED SUBGRAPHS

In this section, we construct a family \( \mathcal{A} \) of \( k^2 \) induced bipartite subgraphs of \( H(n, k) \). In Section 4 we show that (a) each member of \( \mathcal{A} \) is bipartite (Proposition 4.1), (b) each member of \( \mathcal{A} \) has maximum degree at most \( \lceil \sqrt{n} \rceil \) (Proposition 4.5), and (c) at least one member of \( \mathcal{A} \) has more than \( \alpha(H(n, k)) \) vertices (Proposition 4.7). Theorem 1.2 is an immediate consequence of these statements.

Following Chung et al. [1], we note that for all \( n \in \mathbb{N} \), there exists a partition \( [n] = F_1 \cup \cdots \cup F_q \) such that \( |q - \sqrt{n}| < 1 \) and each \( |F_j| - \sqrt{n} | < 1 \). We fix such a partition for the rest of this paper and use it to define partitions of \( \mathbb{Z}_k^n \), the set of vertices.

**Definition 3.1.** We partition the vertex set of \( H(n, k) \) as

\[
\mathbb{Z}_k^n = X \cup Y,
\]

where

\[
X := \{ v \in \mathbb{Z}_k^n : F_j \subseteq v^{-1}(0) \text{ for some } j \in [q]\},
\]  

\[
\text{(1)}
\]
\[
Y := \{ v \in \mathbb{Z}_k^n : F_j \mathcal{G} v^{-1}(0) \text{ for all } j \in [q] \}.
\]  \hfill (2)

**Definition 3.2.** We further partition \( X \) and \( Y \) as

\[
X = X_0 \cup X_1 \cup \cdots \cup X_{k-1},
\]

\[
Y = Y_0 \cup Y_1 \cup \cdots \cup Y_{k-1}
\]

where

\[
X_i := X \cap \left\{ v \in \mathbb{Z}_k^n : \sum_{\ell=1}^n v(\ell) = i \right\},
\]  \hfill (3)

\[
Y_i := Y \cap \left\{ v \in \mathbb{Z}_k^n : \sum_{\ell=1}^n v(\ell) = i \right\}.
\]  \hfill (4)

We consider the following family of \( k^2 \) induced subgraphs of \( H(n,k) \).

\[
\mathcal{A} := \{ H(n,k)[X_{i_1} \cup Y_{i_2}] : i_1, i_2 \in \mathbb{Z}_k \}.
\]  \hfill (5)

\section{Analysis of the Construction}

In this section, we prove properties (a), (b), and (c) of the family \( \mathcal{A} \) indicated in the first paragraph of Section 3. We first address (a).

**Proposition 4.1.** Each of the sets \( X_i, Y_i (i \in \mathbb{Z}_k) \) is independent in \( H(n,k) \). In particular, for all \( i_1, i_2 \in \mathbb{Z}_k \), the induced subgraph \( H(n,k)[X_{i_1} \cup Y_{i_2}] \) is bipartite.

**Proof:** This follows from the fact that the coordinates of vertices in each of the sets \( X_i, Y_i \) have the same sum. \( \square \)

\subsection{Maximum degree bound}

We prove that for all \( i_1, i_2 \in \mathbb{Z}_k \), the maximum degree of \( H(n,k)[X_{i_1} \cup Y_{i_2}] \) is at most \( \lceil \sqrt{n} \rceil \). First notice that if \( v \in X \) has a neighbor in \( Y \), then \( v^{-1}(0) \) contains exactly one of \( F_1, \ldots, F_q \). By contradiction, if \( v^{-1}(0) \) contains more than one of \( F_1, \ldots, F_q \), then any neighbor of \( v \) still contains some \( F_j \) and therefore cannot be in \( Y \). This allows us to make the following definition.

**Definition 4.2.** For any \( v \in X \) with a neighbor in \( Y \), define \( j(v) \) to be the unique \( j \in [q] \) such that \( F_j \subseteq v^{-1}(0) \).

**Proposition 4.3.** Fix \( v_2 \in Y \). If \( v_1, v_1' \in X \) are neighbors of \( v_2 \) and \( j(v_1) = j(v_1') \), then \( v_1 = v_1' \). In particular, each vertex in \( Y \) has at most \( q \) neighbors in \( X \).

**Proof:** Suppose \( v_1 \in X \) is a neighbor of \( v_2 \). Since \( F_{j(v_1)} \subseteq v_1^{-1}(0) \) and \( F_{j(v_1)} \nsubseteq v_2^{-1}(0) \), there exists some \( \ell_0 \in F_{j(v_1)} \) such that \( (v_2^{-1}(0) \cap F_{j(v_1)}) \cup \{ \ell_0 \} = F_{j(v_1)} \). Hence \( v_1 \) is the only element of
\[ \{ v \in \mathbb{Z}_k^n : v \sim v_2, F_{j(v)} \subseteq v^{-1}(0) \}, \]

which we obtain from \( v_2 \) by changing the image of \( \ell_0 \) from \( v_2(\ell_0) \) to 0.

**Proposition 4.4.** If \( v_1 \in X_i \) and \( v_2 \in Y_i \) are neighbors, then there exists \( \ell_0 \in F_{j(v)} \) such that the following hold.

1. \( v_1(\ell_0) = 0; \)
2. \( v_2(\ell_0) = i_2 - i_1; \)
3. \( v_1(\ell) = v_2(\ell) \) for all \( \ell \in [n] \setminus \{\ell_0\}. \)

In particular, for all \( i_1, i_2 \in \mathbb{Z}_k \), each vertex in \( X_i \) has at most \( \max|F_j| \) neighbors in \( Y_i \).

**Proof.** Items 1 and 3 follow from the fact that \( v_1 \in X, v_2 \in Y \) and \( v_1 \sim v_2 \). To see item 2, notice that

\[ i_2 - i_1 = \sum_{\ell=1}^{n} (v_2(\ell) - v_1(\ell)) = v_2(\ell_0) - 0. \]

**Proposition 4.5.** For all \( i_1, i_2 \in \mathbb{Z}_k \),

\[ \Delta(H(n, k)[X_i \cup Y_i]) \leq \lceil \sqrt{n} \rceil. \]

**Proof.** Combining Propositions 4.1, 4.3, and 4.4, we obtain that

\[ \Delta(H(n, k)[X_i \cup Y_i]) \leq \max\left\{ \max_{j \in [q]}|F_j|, q \right\} \leq \lceil \sqrt{n} \rceil. \]  \hfill (6)

4.2  Size greater than the independence number

We now prove that there exist \( i_1, i_2 \in \mathbb{Z}_k \) such that the induced subgraph \( H(n, k)[X_i \cup Y_i] \) has size greater than \( \alpha(H(n, k)) \).

**Lemma 4.6.** \( |X| \equiv (-1)^{q+1} \mod k \). In particular, \( |X| \) is not divisible by \( k \).

**Proof.** The proof generalizes the inclusion-exclusion argument of Chung et al. \[1\]. By the inclusion-exclusion principle, we have

\[
|X| = \sum_{j \in [q]} k^{n-|F_j|} - \sum_{j_1, j_2 \in [q]} k^{n-|F_{j_1} \cup F_{j_2}|} + \cdots + (-1)^q \sum_{j_1, \ldots, j_{q-1} \in [q]} k^{n-|F_{j_1} \cup \cdots \cup F_{j_{q-1}}|} + (-1)^{q+1}.
\]

Since all terms but the last are divisible by \( k \), the statement follows.
Proposition 4.7. There exist \( i_1, i_2 \in \mathbb{Z}_k \) such that \( H(n, k)[X_i \cup Y_i] \) has size greater than \( \alpha(H(n, k)) \).

Proof. Choose \( i_1, i_2 \in \mathbb{Z}_k \) such that \( |X_{i_1}| = \max_{i \in \mathbb{Z}_k} |X_i| \) and \( |Y_{i_2}| = \max_{i \in \mathbb{Z}_k} |Y_i| \). By Lemma 4.6, it follows that

\[
|X_{i_1}| > \frac{1}{k} \sum_{i \in \mathbb{Z}_k} |X_i|.
\]

Since also

\[
|Y_{i_2}| > \frac{1}{k} \sum_{i \in \mathbb{Z}_k} |Y_i|,
\]

we have

\[
|X_{i_1} \cup Y_{i_2}| = |X_{i_1}| + |Y_{i_2}| > \frac{1}{k} \left( \sum_{i \in \mathbb{Z}_k} |X_i| + \sum_{i \in \mathbb{Z}_k} |Y_i| \right) = \frac{1}{k} \left( |X| + |Y| \right) = \frac{1}{k} |\mathbb{Z}_k^n| = k^{n-1}
\]

\[
= \alpha(H(n, k)). \quad \square
\]

Proof of Theorem 1.2. Combine Propositions 4.5 and 4.7. \quad \square

Remark 4.8. Readers familiar with Chung et al. [1] will notice that our proof closely follows the steps of that paper. The challenge for us was to find the right construction of the family \( \mathcal{A} \) of induced subgraphs that permits an extension of the analysis given in [1]. We were surprised to find that the resulting bound, \( \lceil \sqrt{n} \rceil \), does not depend on \( k \).

5 | OPEN QUESTIONS ON THE LOWER BOUND

It is natural to ask whether \( \sqrt{n} \) is the true order of magnitude of \( f(H(n, k)) \) for \( k \geq 3 \). Huang [4] proved the lower bound \( f(H(n, 2)) \geq \sqrt{n} \) using linear algebra. We were unable to generalize his argument to Hamming graphs with larger alphabets. Unfortunately we were unable even to obtain a lower bound of logarithmic order like the Chung et al. lower bound for \( k = 2 \). We were able to generalize one of their lemmas toward this lower bound; this result might be of interest in its own right.

Lemma 5.1. If \( S = (V(S), E(S)) \) is a subgraph of \( H(n, k) \) with average degree \( \bar{d} \), then \( \log |V(S)| \geq \frac{\bar{d}}{k-1} \).

Proof. The statement is clear for \( n = 1 \). Inductively, split \( \mathbb{Z}_k^n \) into \( k \) copies of \( \mathbb{Z}_k^{n-1} \):

\[
\mathbb{Z}_k^n = Z_1 \cup \cdots \cup Z_k,
\]

\[
Z_i = \left\{ v \in \mathbb{Z}_k^n : v(1) = i \right\}.
\]
For every $i \in [k]$ set $V_i = V_i(S) \cap Z_i$. By throwing away the empty $V_i$ and permuting the indices, we may assume $|V_i| \leq \cdots \leq |V_k|$ with all $V_i$ nonempty.

Let $s_1$ be the number of edges between $V_1$ and $\bigcup_{j=2}^K V_j$ in $G$, $s_2$ the number of edges between $V_2$ and $\bigcup_{j=3}^K V_j$ in $G$, $\ldots$, $s_{K-1}$ the number of edges between $V_{K-1}$ and $V_K$ in $G$. This gives $|V_i| \geq \frac{n}{K-1}$ for all $i \in [K-1]$.

The inductive hypothesis gives

$$|V_i| \log |V_i| \geq \frac{1}{k-1} \left( \sum_{v \in V_i} \deg_G[|V_i|](v) \right),$$

so that

$$\sum_{i=1}^K |V_i| \log |V_i| + \frac{2}{k-1} \left( \sum_{i=1}^{K-1} s_i \right) \geq \frac{1}{k-1} \left( \sum_{v \in V} \deg_G(v) \right). \quad (7)$$

Notice that for $0 < p \leq q$, we always have

$$(p + q) \log(p + q) \geq p \log p + q \log q + 2p. \quad (8)$$

Combining (7) and (8) gives

$$(|V_1| + \cdots + |V_K|) \log(|V_1| + \cdots + |V_K|) \geq |V_1| \log |V_1| + \cdots + |V_K| \log |V_K| + (|V_1| + \cdots + |V_K|) \log(|V_1| + \cdots + |V_K|)$$

$$+ 2|V_i| \geq \cdots \geq \sum_{i=1}^K |V_i| \log |V_i| + 2 \sum_{i=1}^{K-1} |V_i| \geq \sum_{i=1}^K |V_i| \log |V_i| + \frac{2}{k-1} \left( \sum_{i=1}^{K-1} s_i \right)$$

$$\geq \frac{1}{k-1} \left( \sum_{v \in V} \deg_G(v) \right). \quad \square$$

**ACKNOWLEDGMENTS**

This work was done while the author was an undergraduate at the University of Chicago. I wish to thank Professor László Babai for suggesting that I work on this problem and discussing it with me.

**ORCID**

Dingding Dong https://orcid.org/0000-0001-8500-2897

**REFERENCES**

1. F. R. K. Chung, Z. Füredi, R. L. Graham, and P. Seymour, *On induced subgraphs of the cube*, J. Combin. Theory. 49 (1988), no. 1, 180–187. https://doi.org/10.1016/0097-3165(88)90034-9

2. C. Gotsman and N. Linial, *The equivalence of two problems on the cube*, J. Combin. Theory. 61 (1992), no. 1, 142–146. https://doi.org/10.1016/0097-3165(92)90060-8

3. P. Hatami, R. Kulkarni, and D. Pankratov, *Variations on the Sensitivity Conjecture*, Theory Comput. Library Graduate Surveys 4 (2011), 1–27. https://doi.org/10.4086/toc.gs.2011.004
4. H. Huang, *Induced subgraphs of hypercubes and a proof of the Sensitivity Conjecture*, Ann. Math. **190** (2019), no. 3, 949–955. https://doi.org/10.4007/annals.2019.190.3.6

5. N. Nisan and M. Szegedy, *On the degree of Boolean functions as real polynomials*, Comput. Complex. **4** (1994), no. 4, 301–313. https://doi.org/10.1007/BF01263419

**How to cite this article:** Dong D. On induced subgraphs of the Hamming graph. *J Graph Theory*. 2021;96:160–166. https://doi.org/10.1002/jgt.22640