A note on short cycles in a hypercube

Maria Axenovich* Ryan Martin†

November 3, 2004

Abstract

How many edges can a quadrilateral-free subgraph of a hypercube have? This question was raised by Paul Erdős about 27 years ago. His conjecture that such a subgraph asymptotically has at most half the edges of a hypercube is still unresolved. Let $f(n, C_l)$ be the largest number of edges in a subgraph of a hypercube Q_n containing no cycle of length l. It is known that $f(n, C_l) = o(|E(Q_n)|)$, when $l = 4k$, $k \geq 2$ and that $f(n, C_6) \geq (\frac{1}{3})|E(Q_n)|$. Until now, it was an open question whether $f(n, C_l) = o(|E(Q_n)|)$ for $l = 4k + 2$, $k \geq 2$. Here, we give a general upper bound for $f(n, C_l)$ when $l = 4k + 2$ and provide a coloring of $E(Q_n)$ by 4 colors containing no induced monochromatic C_{10}.

1 Introduction

Let Q_n be a hypercube of dimension n. We treat its vertices as binary sequences of length n or as subsets of a set $[n] = \{1, 2, \ldots, n\}$, whichever is more convenient. The edges of Q_n correspond to pairs of sets with symmetric difference of size 1 or, equivalently, to pairs of sequences with Hamming distance 1. We denote the set of subsets of $[n]$ of size k by $\binom{[n]}{k}$. We say that a subset of edges forms an i^{th} edge layer of a hypercube if these edges join vertices in $\binom{[n]}{i}$ and $\binom{[n]}{i+1}$. It is a classical question to find a dense subgraph of a hypercube without cycles of a certain fixed length. Moreover, it is of interest to consider the Ramsey properties of cycles in Q_n. We say that a cycle C_l has a *Ramsey property* in Q_n, if for a fixed number of colors k, there is an N_0 such that if $n > N_0$ and the edges of Q_n are colored in k colors then there is always a monochromatic C_l in a such coloring.

It is easy to see that there is a coloring of $E(Q_n)$ in two colors with no monochromatic C_4. Indeed, color all edges in each edge-layer in the same color, using two alternating colors on successive layers. Each color class is a quadrilateral-free subgraph of Q_n, with the largest one having at least $\frac{1}{2}|E(Q_n)|$ edges. There is a coloring by Conder [4] using three colors on $E(Q_n)$ and containing no monochromatic C_6, in particular providing a hexagon-free subgraph of Q_n with at least $|E(Q_n)|/3$ edges. This shows that both C_4 and C_6 do not have Ramsey property in a hypercube.

*Department of Mathematics, Iowa State University, Ames, IA 50011, axenovic@math.iastate.edu
†Department of Mathematics, Iowa State University, Ames, IA 50011, rymartin@iastate.edu

In general, if $f(n, C_l)$ is the largest number of edges in a subgraph of Q_n with no cycle of length l then the following facts are known: Fan Chung [3] proved that $f(n, C_4) \leq .623|E(Q_n)|$. The conjecture of Erdős that $f(n, C_4) = (1/2 + o(1))|E(Q_n)|$ is still open. Chung, [3] also gave the following upper bound for cycles of length 0 modulo 4: $f(n, C_{4k}) \leq cn^{(1/2k)-(1/2))}|E(Q_n)|$, $k \geq 2$.

Thus, C_{4k} has Ramsey property for $k \geq 2$. In [3], Chung also raised the question about Ramsey properties of cycles C_{4k+2}, $k \geq 2$, in particular about C_{10}. In this paper, we show that C_{10}, as an induced subgraph of Q_n, does not have the Ramsey property.

Theorem 1.1 *There is a coloring of $E(Q_n)$ using 4 colors such that there is no induced monochromatic C_4, C_6, or C_{10}.*

We prove this theorem and describe some properties of the corresponding coloring which are of independent interest in Section 2. For completeness, we provide a general upper bound on the maximum number of edges in a subgraph of a hypercube containing no cycle of length C_{4k+2} in Section 3. For the general graph-theoretic definitions we refer the reader to [8].

2 A coloring with no induced monochromatic C_{10}

For a binary sequence x, let $w(x)$ be the weight (or number of 1's) in x. If the vertices corresponding to the binary sequences x and y form an edge in Q_n (i.e., they differ in exactly one position) let the common substring that precedes the change be called the *prefix* and be denoted $p = p(xy)$. Let the common substring that follows the change be called the *suffix* and be denoted $s = s(xy)$. We shall sometimes say that $p(xy)$, $(s(xy))$ is the *common prefix*, (*common suffix*) on the edge xy.

Description of a coloring c

For an edge xy, with $x \in \binom{[n]}{k}$, $y \in \binom{[n]}{k+1}$ and having common prefix $p = p(xy)$, let $c(xy) = (c_1, c_2)$, where

$$c_1 \equiv k \pmod 2, \qquad c_2 \equiv w(p) \pmod 2.$$

2.1 Proof of Theorem 1.1

We shall prove that the coloring c does not produce induced monochromatic cycles of length 4, 6, or 10. Since the sets of colors used on consecutive edge-layers of Q_n are disjoint, the monochromatic cycles can occur only within edge-layers. There is no C_4 within an edge-layer, so there is no monochromatic C_4.

Let C be a monochromatic cycle in Q_n under coloring c. Then $C = x_1, y_1, \ldots, x_m, y_m, x_1$, where $x_i \in \binom{[n]}{k}$, $y_i \in \binom{[n]}{k+1}$, for $i = 1, \ldots, m$ and for some $k \in \{1, \ldots, n-1\}$. Note also that $y_i = x_i \cup x_{i+1}$, for $i = 1, \ldots, m$ where addition is taken modulo m. In particular, we have that $\text{dist}(x_i, x_{i+1}) = 2$ and $\text{dist}(y_i, y_{i+1}) = 2$, for $i = 1, \ldots, m$.

In the following lemma, we settle the case with C_6 mostly to illustrate the properties of the coloring.

Lemma 2.1 *There is no monochromatic C_6 in Q_n under coloring c.*

Proof. Let $C = x_1y_1x_2\dots x_3y_3x_1$ be a monochromatic C_6. Then x_i's can be written as $a_01a_10a_20a_3$, $a_00a_11a_20a_3$, $a_00a_10a_21a_3$, for some binary words $a_0,\dots,a_3$. The corresponding y_i's are $a_01a_11a_20a_3$, $a_00a_11a_21a_3$, $a_01a_10a_21a_3$. Therefore, the set of common prefixes on the edges of C contains a_00a_1 and a_01a_1. These two prefixes have different weights modulo 2, thus C is not monochromatic under coloring c. ■

For a word $\mathbf{u} = u_1u_2\cdots u_n$, let the reverse of $\mathbf{u}$ be $\overline{\mathbf{u}} = u_nu_{n-1}\cdots u_1$.

Observation 1. Let $x, x' \in \binom{[n]}{k}$ and $y, y' \in \binom{[n]}{k+1}$. Let xy, $x'y'$ be edges of Q_n such that x and y have a common prefix p and a common suffix s and x' and y' have a common prefix p' and a common suffix s'. Then $w(p) \equiv w(p') \pmod 2$ iff $w(s) \equiv w(s') \pmod 2$. This is also equivalent to saying that if q is the common prefix of $\overline{x}$ and $\overline{y}$ and q' is the common prefix of $\overline{x'}$ and $\overline{y'}$ then $w(p) \equiv w(p') \pmod 2$ iff $w(q) \equiv w(q') \pmod 2$.

Main idea of the proof. Let $C = x_1, y_1, \dots, x_5, y_5, x_1$ be an induced C_{10} in Q_n with $x_i \in \binom{[n]}{k}$ and $y_i \in \binom{[n]}{k+1}$, $i = 1,\dots,5$. Let $I(C) = \{i_1,\dots,i_m\}$, $i_1 < i_2 < \cdots < i_m$, be the set of positions where x_i and x_{i+1} differ for some $i \in [5]$. First, we shall show that $m = 5$. Second we define a 5×5 matrix $\mathbf{A} = \mathbf{A}(C)$ where the $(i,j)^{\text{th}}$ entry is denoted $a(i,j)$ and $a(i,j) = 1$ if $x_i = 1$ in position i_j and $a(i,j) = 0$ otherwise. We shall show that none of the 5×5 binary matrices are possible as $\mathbf{A}(C)$ for an induced cycle C monochromatic under coloring c.

Note. For the rest of the proof we assume that C is monochromatic under coloring c and that it is an induced cycle in Q_n. We take the addition modulo 5 unless otherwise specified.

We say that a matrix $\mathbf{A}$ has **bad prefixes** if there are two pairs of consecutive rows $\mathbf{r}_i, \mathbf{r}_{i'}$ and $\mathbf{r}_j, \mathbf{r}_{j'}$ in $\mathbf{A}$, $i' = i+1 \pmod 5$, $j' = j+1 \pmod 5$, such that either

(1) $\mathbf{r}_i$ starts with $\mathbf{a}00$, $\mathbf{r}_{i'}$ starts with $\mathbf{a}01$; $\mathbf{r}_j$ starts with $\mathbf{a}01$, $\mathbf{r}_{j'}$ starts with $\mathbf{a}10$ or

(2) $\mathbf{r}_i$ starts with $\mathbf{a}00$, $\mathbf{r}_{i'}$ starts with $\mathbf{a}01$; $\mathbf{r}_j$ starts with $\mathbf{a}11$, $\mathbf{r}_{j'}$ starts with $\mathbf{a}10$,

for some binary word $\mathbf{a}$.

For a matrix $\mathbf{A}$ with rows $\mathbf{v}_1,\dots,\mathbf{v}_5$, let $\overline{\mathbf{A}}$ be a matrix with rows $\overline{\mathbf{v}_1},\dots,\overline{\mathbf{v}_5}$ in the same order.

Lemma 2.2 *Let $\mathbf{A} = \mathbf{A}(C)$, then neither $\mathbf{A}$ nor $\overline{\mathbf{A}}$ have bad prefixes.*

Proof. Assume that the matrix $\mathbf{A}$ has bad prefixes as in part (1) of the definition. Assume without loss of generality, that $i = 1, i' = 2$, $j = 3$, and $j' = 4$. Then $x_1 = a_10a_20b$, $x_2 = a_10a_21a_3$, and $y_1 = a_10a_21a_3$, for some binary words a_1, a_2, a_3 and the common prefix of x_2 and y_1 is a_10a_2. We also have that $x_3 = a_10a_21a_3$, $x_4 = a_11a_20a_3$, $y_3 = a_11a_21a_3$

and the common prefix of x_4 and y_3 is $a_1 1 a_2$. Thus, $c(x_4 y_3) \neq c(x_2 y_1)$ since the weights of corresponding prefixes are different modulo 2, a contradiction to the assumption that cycle C is monochromatic. The other case and row labellings are very similar. When $\overline{\mathbf{A}}$ has bad prefixes, we use Observation 1 to arrive at the same conclusion. ■

Lemma 2.3 *There are three consecutive x_i's, without loss of generality, x_1, x_2, x_3, such that $x_1 < x_2 < x_3$ in lexicographic order. The number of positions where some two of x_1, x_2, x_3 differ is four.*

Proof. Since there is an odd number of x_i's, there are three consecutive elements, without loss of generality x_1, x_2, x_3, such that $x_1 < x_2 < x_3$. Since $d(x_1, x_2) = d(x_2, x_3) = 2$, it is clear that the number of positions where (x_1 and x_2) or (x_2 and x_3) differ is either three of four. Assume that it is three. We must have only one 1 in these positions for each x_1, x_2, x_3 otherwise $x_1 \cup x_2 = x_2 \cup x_3$, implying that $y_1 = y_2$, a contradiction. Then it follows that $x_1 = a_1 1 a_2 0 a_3 0 a_4$, $x_2 = a_1 0 a_2 1 a_3 0 a_4$, and $x_3 = a_1 0 a_2 0 a_3 1 a_4$, for some binary words $a_1, \ldots, a_4$. Thus $y_1 = a_1 1 a_2 1 a_3 0 a_4$, $y_2 = a_1 0 a_2 1 a_3 1 a_4$. The common prefix between x_1 and y_1 is $a_1 1 a_2$, the common prefix between x_3 and y_2 is $a_1 0 a_2$. These prefixes have different weights modulo two, a contradiction. ■

For a binary sequence x of length n, we denote by $x[I]$ its restriction to positions from I. For example, if $x = 0010010110$ and $I = \{2, 3, 5, 6, 10\}$ then $x[I] = 01010$.

Lemma 2.4 $|I(C)| = 5$. *Moreover, 1's occur in consecutive positions (modulo 5) in each column of $\mathbf{A}(C)$.*

Proof. Let $|I(C)| = \{i_1, \ldots, i_m\}$. From Lemma 2.3, we know that $m \geq 4$.
Case 1. Let $m = 4$. It follows from Lemma 2.3 that each row of A has two 1's. Therefore, $y_i[I]$ has weight three for each $i = 1, \ldots, 5$. But there are only four binary words of length four and weight three, a contradiction to the fact that we have five distinct y_i's.
Case 2. Assume that $m \geq 6$. We shall introduce a (multi-)graph H on vertices $v_{i_1}, \ldots, v_{i_m}$. Let $v_{i_s} v_{i_t} \in E(H)$ if some pair of consecutive x_i's differ exactly in positions i_s and i_t. Note that since C is a cycle, a graph H must have even nonzero degrees, thus $|E(H)| \geq |V(H)| = m$. Since $|E(H)| = 5$, $m \leq 5$, a contradiction.

Since the number of edges in H is 5 and the number of vertices is 5 we have that each vertex has degree 2. Since a vertex of H corresponds to a column of $\mathbf{A}(C)$, and the degree of a vertex in H corresponds to a number of times 1 changes to 0 and 0 changes to 1 on consecutive elements in a corresponding column of $\mathbf{A}(C)$ (in cyclic order), we must have 1's occur consecutively in each column of $\mathbf{A}(C)$.

■

Lemma 2.5 *The matrix $\mathbf{A} = \mathbf{A}(C)$ has exactly two 1's in each row.*

Proof. Since all x_i's have the same weight, each row of $\mathbf{A}$ contains the same number of 1's.

Assume that each row of $\mathbf{A}$ contains exactly one 1. There are three consecutive rows of $\mathbf{A}$ which form binary words in increasing lexicographic order. It is easy to see that $\mathbf{A}$ has *bad* prefixes because of these rows.

Assume that there are three 1's and two 0's in each row of A. Then $y_i[I]$ has weight 4, for $i = 1, \ldots, 5$. There are exactly 5 binary words of length 5 and weight 4. Thus each column of $\mathbf{A}$ must have two consecutive 0's and three consecutive 1's (in cyclic order). Then each $x_i = a_0 * a_1 * a_2 * a_3 * a_4 * a_5$, where a_j's are some binary words and $* \in \{0, 1\}$ are located in positions from I.

If there is a row with two consecutive 0's in the first and second column then we immediately get *bad* prefixes since the first two columns must be as follows, up to cyclic rotation of the rows:

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

Thus we can assume that the first two columns of $\mathbf{A}$ are as follows, up to cyclic rotation of the rows:

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

Now, we see that $a(3, 3) = 0$, otherwise $a(3, 4) = a(3, 5) = 0$ and we arrive at *bad* prefixes in $\overline{\mathbf{A}}$, similarly to the previous argument. Then we have two possible cases for the first three columns of $\mathbf{A}$:

$$\begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

Note that in the first case we have prefixes $a_0 1 a_1 1 a_2$ and $a_0 1 a_1 0 a_2$, which have different weights modulo 2, and in the second case we have prefixes $a_0 1 a_1 1 a_2$ and $a_0 0 a_1 1 a_2$, which have different weight modulo 2, each case is a contradiction to the fact that C is monochromatic.

Finally, the rows of $\mathbf{A}$ cannot contain more than three 1's each since in this case $y_i = y_j$ for $i, j \in [5]$. ■

Note that Lemma 2.5 implies that the total number of ones in $\mathbf{A}$ is 10.

Lemma 2.6 *If there are indices i_1, i_2, i_3 and j_1, j_2 such that $\mathbf{A}(C)$ has entries $a(i_1, j_1) = a(i_1, j_2) = a(i_2, j_1) = a(i_2, j_2) = a(i_3, j_1) = a(i_3, j_2) = 0$, then C has a chord in Q_n.*

Proof. Consider $x_{i_1}, x_{i_2}, x_{i_3}$. Note that $x_{i_1} \cup x_{i_2} = x_{i_2} \cup x_{i_3} = x_{i_1} \cup x_{i_3}$. At least two of these three vertices must be successive vertices of C in a layer $\binom{[n]}{k}$, say without loss of generality that $x_{i_1} = x_1$ and $x_{i_2} = x_2$. But then $x_{i_3} \neq x_3$, $x_{i_3} \neq x_5$, otherwise we shall have $y_1 = y_2$ or $y_1 = y_5$. Thus $x_{i_3} = x_4$ and $y_1 x_4$ forms a chord of C in Q_n. ∎

Now, we are ready to complete the proof of the Theorem 1.1. We see from Lemma 2.6 that **A** has at most one column with at most one 1. Since the total number of 1's in **A** is 10, we have that each column of **A** has at at most three 1's. Moreover there is at most one column with three 1's. Let **A** have exactly one column with exactly one 1. Then, there must be a column with exactly three 1's. Consider possible submatrices formed by these two columns (up to cyclically rotating the rows):

$$B_1 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 0 & * \\ 0 & * \end{pmatrix}, \quad B_3 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad * \in \{0,1\}.$$

In case of B_1, we have that, for some new column i, $a(2,i) = a(3,i) = 1$ then $a(1,i) = 1$. Thus there are two columns with exactly three 1's, a contradiction. In case of B_2 we must have $a(2,i) = 1 = a(1,i) = a(3,i)$ for a new column i, therefore $x_1 = x_3$, a contradiction. B_3 is possible only if other three columns up to permutation are as follows.

$$\begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix} \tag{1}$$

(This is easily checked using only the fact that each row has two 1's and that Hamming distance between consecutive rows is 2.)

If we apply Lemma 2.2 to all pairs of columns in (1), we see that the only possible pairs for the first two and the last two columns are $\{1,2\}, \{1,5\}, \{2,5\}$. Thus it is impossible to choose two acceptable last columns and two acceptable first columns at the same time. This concludes the argument that **A** cannot have columns with exactly one 1.

Therefore **A** has exactly two 1's in each column occuring consecutively. Consider the first two columns. There are only two possibilities: when there is a row i such that $a(i,1) = a(i,2) = 1$ and when there is no such row. In both situations, we immediately see that matrix **A** has bad prefixes.

Thus no 5×5 matrix of 0's and 1's can be equal to $\mathbf{A}(C)$ for some monochromatic induced 10-cycle C. This concludes the proof of the theorem 1.1. ∎

Remark. The above proof basically reduces the analysis of the coloring c on $E(Q_n)$ to the analysis of the coloring c on $E(Q_5)$. We believe that a similar result should hold for chordless C_{4k+2} in Q_n, $k > 2$.

3 The general results for even cycles in a hypercube

The following result of Fan Chung [3] states that dense subgraphs of a hypercube contain all "not too long" cycles of lengths divisible by 4.

Theorem 3.1 *For each t, there is a constant $c = c(t)$ such that if G is a subgraph of Q_n with at least $cn^{-1/4}|E(Q_n)|$ edges then G contains all cycles of lengths $4k$, $2 \le k \le t$.*

On the other hand, we can conclude that a dense subgraph of a hypercube contains a cycle of length $4k+2$ for some k which follows from the following strengthening of a classical theorem of Bondy and Simonovits [2] by Verstraëte [7].

Theorem 3.2 *[7] Let $q \ge 2$ be a natural number and G a bipartite graph of average degree at least $4q$ and girth g. Then there exist cycles of $(g/2-1)q$ consecutive even length in G.*

Unfortunately, these results still do not guarantee the existence of a cycle of length $4k+2$ for some small fixed k in dense subgraphs of a hypercube. For completeness, we include a general bound on the maximum number of edges in a C_{4k+2}-free subgraph of a hypercube.

Theorem 3.3 $f(n, C_{4k+2}) \le (1+o(1))\frac{1}{\sqrt{2}}n2^{n-1}$, $k \ge 1$.

Proof. Let $k \ge 1$ be an integer and let G be a subgraph of Q_n containing no cycles of length $4k+2$. We follow Fan Chung's notation to carry out the counting argument over the Q_2's of different types. We shall introduce χ_2, χ_3 and χ_4 as follows: For $i = 3, 4$, let χ_i be the fraction of the number of Q_2's in Q_n sharing exactly i edges with G, i.e., the number of Q_2's sharing exactly 3 edges with G is $\chi_3\binom{n}{2}2^{n-2}$. Let χ_2 be the fraction of the number of Q_2's in Q_n sharing with G exactly two edges, such that these two edges are adjacent.

Let d_v be the degree of a vertex v in G. For each $v \in V(Q_n)$, we introduce a graph $H_v = (V, E)$ such that $V = \{u \in V(Q_n) : uv \in E(Q_n)\}$ and $E = \{\{u, w\} :$ there is a $2-$path from u to v in G different from $uvw\}$.

If $uw, u'w' \in E(H_v)$ (u might coincide with u'), then there is a unique path uxw, $x \ne v$ and a unique path $u'x'w'$, $x' \ne v$ in G, moreover $x \ne x'$. We note first that H_v does not have C_{2k+1}. To that end, consider $u_1, \ldots, u_{2k+1}, u_1$, a cycle in H_v. The previous observation implies that there is a cycle $u_1, x_1, u_2, x_2, \ldots, u_{2k+1}, x_{2k+1}, u_1$ in G, a contradiction. If n is sufficiently large, then any graph on n vertices with no copy of C_{2k+1} has at most $n^2/4$ edges, see for example [1]. Hence, H_v has at most $n^2/4$ edges, for n large enough.

Then we have

$$\sum_{v \in V(Q_n)} |E(H_v)| \le 2^n n^2/4. \tag{2}$$

Next, we consider the sum of $|E(H_v)|$ over all $v \in V(Q_n)$. Note that each Q_2 sharing 4 edges with G contributes 4 to this sum, each Q_2 in Q_n that shares 3 edges with G contributes 2 to to this sum, and each Q_2 in Q_n that shares 2 adjacent edges contributes 1.

$$\sum_{v \in V(Q_n)} |E(H_v)| = \binom{n}{2} 2^{n-2}(4\chi_4 + 2\chi_3 + \chi_2). \tag{3}$$

We can associate pairs of edges in G incident to the same vertex arbitrarily to copies of Q_2 in Q_n in which they reside, implying

$$\sum_{v\in V(Q_n)} \binom{d_v}{2} \le \binom{n}{2} 2^{n-2}(4\chi_4 + 2\chi_3 + \chi_2). \tag{4}$$

Finally, using the Cauchy-Schwarz inequality, we have

$$\begin{aligned} \sum_{v\in V(Q_n)} \binom{d_v}{2} &= \sum_{v\in V(Q_n)} (d_v^2/2 - d_v/2) \\ &= \left(\sum_{v\in V(Q_n)} d_v^2/2\right) - |E(G)| \\ &\ge 2^{-n-1}\left(\sum_{v\in V(Q_n)} d_v\right)^2 - |E(G)| \\ &= 2^{-n-1}(2|E(G)|)^2 - |E(G)|. \end{aligned} \tag{5}$$

Combining (2), (3), (4) and (5), we have

$$2^{-n-1}(2|E(G)|)^2 - |E(G)| \le 2^n n^2/4.$$

If $|E(G)| = an2^n$, then $a \le \frac{1}{2n} + \frac{1}{2}\sqrt{2 + 1/n^2} \le 1/\sqrt{2} + \epsilon$, where $\epsilon \to 0$ as $n \to \infty$. ■

Remark. The constant $1/\sqrt{2}$ in the above theorem can be improved with more refined arguments which we do not include in this note. In particular, one can obtain a better constant by a more accurate analysis of $|E(H_v)|$, as was observed by Pikhurko [6].

References

[1] B. Bollobás, *Extremal graph theory*, Reprint of the 1978 original. Dover Publications, Inc., Mineola, NY, 2004. xx+488 pp.

[2] J. Bondy, M. Simonovits, *Cycles of even lengths in graphs*, J. Combinatorial Theory B **16**, (1974), 97–105.

[3] F. R. K. Chung, *Subgraphs of a hypercube containing no small even cycles.* J. Graph Theory 16 (1992), no. 3, 273–286.

[4] M. Conder, *Hexagon-free subgraphs of hypercubes,* J. Graph Theory **17** (1993), 477–479.

[5] K. Johnson, R. Entringer, *Largest induced subgraphs of the n-cube that contain no* 4-*cycles*, J. Combin. Theory Ser. **B** 46 (1989), no. 3, 346–355.

[6] O. Pikhurko, personal communication.

[7] J. Verstraëte, *On Arithmetic Progressions of Cycle Lengths in Graphs*, submitted.

[8] D. West, *Introduction to graph theory*, Second edition, Prentice Hall (2001), xx+588 pp.