# An extremal theorem in the hypercube

David Conlon[*]

St John's College
Cambridge CB2 1TP, United Kingdom

D.Conlon@dpmms.cam.ac.uk

## Abstract

The hypercube $Q_n$ is the graph whose vertex set is $\{0,1\}^n$ and where two vertices are adjacent if they differ in exactly one coordinate. For any subgraph $H$ of the cube, let $\mathrm{ex}(Q_n, H)$ be the maximum number of edges in a subgraph of $Q_n$ which does not contain a copy of $H$. We find a wide class of subgraphs $H$, including all previously known examples, for which $\mathrm{ex}(Q_n, H) = o(e(Q_n))$. In particular, our method gives a unified approach to proving that $\mathrm{ex}(Q_n, C_{2t}) = o(e(Q_n))$ for all $t \geqslant 4$ other than 5.

## 1 Introduction

Given two graphs $G$ and $H$, $\mathrm{ex}(G, H)$ is the maximum number of edges in a subgraph of $G$ which does not contain a copy of $H$. The study of such numbers was initiated by Paul Turán [19] when he determined the maximum number of edges in a graph on $n$ vertices which does not contain a clique of size $r$, that is, $\mathrm{ex}(K_n, K_r)$. For any graph $H$, the value of $\mathrm{ex}(K_n, H)$ is known [13] to be intimately connected to the chromatic number of $H$. In particular, it is known [16] that $\mathrm{ex}(K_n, H) = o(e(K_n))$ if and only if $H$ is bipartite. In this note, we will be similarly interested in determining subgraphs $H$ of the hypercube $Q_n$ for which $\mathrm{ex}(Q_n, H) = o(e(Q_n))$.

The hypercube $Q_n$ is the graph whose vertex set is $\{0,1\}^n$ and where two vertices are adjacent if they differ in exactly one coordinate. This graph has $2^n$ vertices and, being $n$-regular, $2^{n-1}n$ edges. Erdős [11] was the first to draw attention to Turán-type problems in the cube when he asked how many edges a $C_4$-free subgraph of the cube can contain. He conjectured that the answer is $\left(\frac{1}{2} + o(1)\right) e(Q_n)$ and offered \$100 for a solution.

Improving a long-standing result of Chung [8], Thomason and Wagner [18] recently gave an upper bound of roughly $0.6226 e(Q_n)$ for $\mathrm{ex}(Q_n, C_4)$. This remains a long way

---

from the lower bound of $\frac{1}{2}(n + \sqrt{n})2^{n-1}$ (valid when $n$ is a power of 4), due to Brass, Harborth and Nienborg [6]. A related result of Bialostocki [4] implies that if the edges of the cube are 2-coloured and neither colour class contains a $C_4$, then the number of edges in each colour is at most $\frac{1}{2}(n + \sqrt{n})2^{n-1}$. Therefore, at least in some sense, the construction of Brass, Harborth and Nienborg is optimal.

Erdős [11, 12] also posed the problem of determining $\mathrm{ex}(Q_n, C_{2t})$ for all $t > 2$, suggesting that perhaps $o(e(Q_n))$ was sufficient. This is false for $t = 3$. Indeed, Chung [8] and Brouwer, Dejter and Thomassen [7] found 4-colourings of the cube without any monochromatic $C_6$ and, later, Conder [9] found a 3-colouring with the same property. This implies that $\mathrm{ex}(Q_n, C_6) \geqslant \frac{1}{3}e(Q_n)$. On the other hand, it is known [8] that $\mathrm{ex}(Q_n, C_6) \leqslant \left(\sqrt{2} - 1 + o(1)\right)e(Q_n)$.

For even $t \geqslant 4$, Chung [8] justified Erdős' intuition by showing that $\mathrm{ex}(Q_n, C_{2t}) = o(e(Q_n))$. However, until very recently, it was unknown whether a similar result holds for odd values of $t$. This problem was almost completely resolved by Füredi and Özkahya [14, 15], who showed that, for odd $t \geqslant 7$, $\mathrm{ex}(Q_n, C_{2t}) = o(e(Q_n))$. The only case that now remains unresolved is $C_{10}$.

Some progress on this problem has been made by Alon, Radoičić, Sudakov and Vondrák [2], who proved that in any $k$-colouring of the edges of $Q_n$, for $n$ sufficiently large, there are monochromatic copies of $C_{10}$. Therefore, unlike $C_4$ and $C_6$, a counterexample to the conjecture that $\mathrm{ex}(Q_n, C_{10}) = o(e(Q_n))$ cannot come from a colouring. On the other hand, Alon et al. [2] and Axenovich and Martin [3] both note that there is a 4-colouring of $Q_n$ which does not contain an induced monochromatic copy of $C_{10}$.

More generally, Alon et al. [2] gave a characterisation of all subgraphs $H$ of the cube which are Ramsey, that is, such that every $k$-edge-colouring of a sufficiently large $Q_n$ contains a monochromatic copy of $H$. It would be nice to have a similar characterisation of all graphs $H$ for which $\mathrm{ex}(Q_n, H) = o(e(Q_n))$. Unfortunately, even deciding whether this is true for $C_{10}$ seems very difficult, so a necessary and sufficient condition is probably well beyond our grasp. Nevertheless, the purpose of this note is to present a natural sufficient condition. To state the condition precisely, we will need some notation.

Recall that the vertex set of the hypercube $Q_n$ is $\{0, 1\}^n$. Any vertex, such as $[01101]$ in $Q_5$, can be written as a sequence of bits. Any edge can be written uniquely by a sequence such as $[01*11]$. The missing bit, known hereafter as the flip-bit, tells us that the edge connects the vertex where that missing bit is equal to 0 with the vertex where it is 1. In this case, $[01011]$ is connected to $[01111]$.

We say that a subgraph $H$ of the hypercube has a $k$-partite representation if there exists $l$ such that

- $H$ is a subgraph of $Q_l$;

- every edge $e = [a_1 a_2 \ldots a_l]$ in $H$ has exactly $k$ non-zero bits ($k - 1$ ones and a flip-bit);

- there exists a function $\sigma : [l] \to [k]$ such that, for each $e$, the image $\{\sigma(i_1), \ldots, \sigma(i_k)\}$ of the set of non-zero bits $\{a_{i_1}, \ldots, a_{i_k}\}$ of $e$ under $\sigma$ is $[k]$, that is, no two non-zero bits have the same image.

Let $\mathcal{H}$ be the $k$-uniform hypergraph on vertex set $[l]$ with edge set $\{\tau(e) : e \in E(H)\}$, where the function $\tau$ is defined by mapping the set of non-zero bits $\{a_{i_1}, \ldots, a_{i_k}\}$ of $e$ to the subset $\{i_1, \ldots, i_k\}$ of $[l]$. We refer to $\mathcal{H}$ as a representation of $H$. By the definition of a $k$-partite representation, $\mathcal{H}$ must, unsurprisingly, be $k$-partite. The subsets $\sigma^{-1}(1), \ldots, \sigma^{-1}(k)$ of $[l]$ will be referred to as the partite sets of the representation.

As an initial example, note that $C_8$ has a 2-partite representation. This may be seen by taking $\sigma(1) = \sigma(3) = 1$ and $\sigma(2) = \sigma(4) = 2$ in the following example.

$$
\begin{aligned}
e_1 &= [1 \quad * \quad 0 \quad 0] & \qquad e_5 &= [0 \quad 0 \quad 1 \quad *] \\
e_2 &= [* \quad 1 \quad 0 \quad 0] & e_6 &= [0 \quad 0 \quad * \quad 1] \\
e_3 &= [0 \quad 1 \quad * \quad 0] & e_7 &= [* \quad 0 \quad 0 \quad 1] \\
e_4 &= [0 \quad * \quad 1 \quad 0] & e_8 &= [1 \quad 0 \quad 0 \quad *]
\end{aligned}
$$

Similarly, for any even $t \geqslant 4$, it is easy to see that $C_{2t}$ has a 2-partite representation, namely, the cycle $C_t$ of length $t$. This observation is at the core of Chung's proof that $\mathrm{ex}(Q_n, C_{2t}) = o(e(Q_n))$.

For odd values of $t$, this is not true. However, for $t \geqslant 7$, these graphs do admit a 3-partite representation. For example, in the case of $C_{14}$, the following representation works with $\sigma(1) = \sigma(4) = 1$, $\sigma(2) = \sigma(6) = 2$ and $\sigma(3) = \sigma(5) = \sigma(7) = 3$.

$$
\begin{aligned}
e_1 &= [1 \quad 1 \quad * \quad 0 \quad 0 \quad 0 \quad 0] & \qquad e_8 &= [1 \quad * \quad 0 \quad 0 \quad 1 \quad 0 \quad 0] \\
e_2 &= [* \quad 1 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0] & e_9 &= [1 \quad 0 \quad 0 \quad 0 \quad 1 \quad * \quad 0] \\
e_3 &= [0 \quad 1 \quad 1 \quad * \quad 0 \quad 0 \quad 0] & e_{10} &= [1 \quad 0 \quad 0 \quad 0 \quad * \quad 1 \quad 0] \\
e_4 &= [0 \quad 1 \quad * \quad 1 \quad 0 \quad 0 \quad 0] & e_{11} &= [1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1 \quad *] \\
e_5 &= [0 \quad 1 \quad 0 \quad 1 \quad * \quad 0 \quad 0] & e_{12} &= [1 \quad 0 \quad 0 \quad 0 \quad 0 \quad * \quad 1] \\
e_6 &= [0 \quad 1 \quad 0 \quad * \quad 1 \quad 0 \quad 0] & e_{13} &= [1 \quad * \quad 0 \quad 0 \quad 0 \quad 0 \quad 1] \\
e_7 &= [* \quad 1 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0] & e_{14} &= [1 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0 \quad *]
\end{aligned}
$$

More generally, for all odd $t \geqslant 7$, $C_{2t}$ has a 3-partite representation which is close to a hypergraph cycle. This representation is particularly simple when $t$ is a multiple of 3, when it corresponds to the tight cycle, that is, the 3-uniform hypergraph on $t$ vertices $v_1, \ldots, v_t$ whose edge set is

$$\{v_1 v_2 v_3, \ v_2 v_3 v_4, \ v_3 v_4 v_5, \ \ldots, \ v_{t-1} v_t v_1, \ v_t v_1 v_2\}.$$

On the other hand, $C_4$, $C_6$ and $C_{10}$ do not admit $k$-partite representations for any $k$. The importance of these considerations lies in the following theorem.

**Theorem 1.1** *Let $H$ be a fixed subgraph of the cube. If, for some $k$, $H$ admits a $k$-partite representation, then*

$$ex(Q_n, H) = o(e(Q_n)).$$

We therefore have a unified proof that $\mathrm{ex}(Q_n, C_{2t}) = o(e(Q_n))$ for all $t \geqslant 4$ other than 5. However, the theorem plainly applies to a much wider class of graphs than cycles. If, for example, we add the edge $e_{15} = [1100*00]$ to the $C_{14}$ given above, we get a 3-partite

representation of a $C_{14}$ with a long diagonal, reproving another result due to Füredi and Özkahya [14].

For the sake of clarity of presentation, we will systematically omit floor and ceiling signs whenever they are not crucial. We also do not make any serious attempt to optimize absolute constants in our statements and proofs.

# 2  Subgraphs of the cube with zero Turán density

We will need a simple estimate stating that almost all vertices in the cube $Q_n$ have roughly the same number of zeroes and ones. Since there is a one-to-one correspondence between vertices of $Q_n$ and subsets of the set $[n] = \{1, 2, \ldots, n\}$, the required estimate follows from the following lemma.

**Lemma 2.1** *The number of subsets of $[n]$ containing fewer than $n/4$ or more than $3n/4$ elements is at most $(1.9)^n n$.*

**Proof:** By symmetry, it is easy to see that the number of subsets of $[n]$ with fewer than $n/4$ elements is the same as the number of subsets with more than $3n/4$ elements. Therefore, the number of subsets with fewer than $n/4$ elements or more than $3n/4$ elements is at most

$$2 \sum_{i=0}^{\lfloor n/4 \rfloor} \binom{n}{i} \leqslant n \binom{n}{n/4} \leqslant (4e)^{n/4} n,$$

where we used the estimate $j! \geqslant (j/e)^j$ with $j = n/4$. The result follows since $(4e)^{1/4} \leqslant 1.9$. $\qquad\square$

**Theorem 2.1** *Let $H$ be a subgraph of the cube with $k$-partite representation $\mathcal{H}$. Suppose $ex(K_n^{(k)}, \mathcal{H}) \leqslant \alpha n^k$. Then*

$$ex(Q_n, H) = O(\alpha^{1/k} 2^n n).$$

**Proof:** Suppose that we have a subgraph $G$ of $Q_n$ with density $\epsilon = 16k\alpha^{1/k}$. We will prove that, for $n$ sufficiently large, $G$ must contain a copy of $H$. By Lemma 2.1, the number of vertices containing fewer than $n/4$ or more than $3n/4$ ones is at most $(1.9)^n n$. Therefore, since each vertex has maximum degree $n$, the number of edges between levels $i$ and $i+1$, added over all $i$ for which $i < n/4$ or $i \geqslant 3n/4$, is at most $(1.9)^n n^2$. Since there are $2^{n-1} n$ edges in all, the density contribution of these edges is at most $2(0.95)^n n$, which, for $n$ sufficiently large, is less than $\frac{\epsilon}{2}$. Therefore, $G$ has a density of at least $\frac{\epsilon}{2}$ concentrated between levels $n/4$ and $3n/4$. In particular, there exists some $j$ with $n/4 \leqslant j < 3n/4$ such that the density of edges in $G$ between levels $j$ and $j+1$ is at least $\frac{\epsilon}{2}$.

Every edge between levels $j$ and $j+1$ may be represented by a collection of $j$ ones and a flip-bit. By the choice of $j$, there are at least $\frac{\epsilon}{2}(n-j)\binom{n}{j} \geqslant \frac{\epsilon j}{2}\binom{n}{j+1}$ edges of $G$ between levels $j$ and $j+1$. Given a subset $J$ of $[n]$ of size $j+1$, let $d(J)$ be the number of edges

for which the union of the flip-bit and the set of ones is $J$. Since $\sum d(J)$ is the number of edges between levels $j$ and $j+1$, $\mathbb{E}(d(J)) \geqslant \frac{\epsilon j}{2}$. Therefore, by convexity of the function $\binom{x}{k}$, we have

$$\sum_{J \subset [n], |J| = j+1} \binom{d(J)}{k} \geqslant \binom{n}{j+1}\binom{\mathbb{E}(d(J))}{k} \geqslant \binom{n}{j+1}\binom{\epsilon j/2}{k}.$$

For any subset $J$ of $[n]$ of size $j+1$, let $D(J) \subset J$ be the set of positions where replacing an element of $J$ with a flip-bit yields an edge. The previous equation tells us that there are at least $\binom{n}{j+1}\binom{\epsilon j/2}{k}$ pairs $(I, J)$ for which $J$ has size $j+1$, $I \subset D(J)$ and $|I| = k$. Since there are $\binom{n}{j+1-k}$ ways of choosing a subset of $n$ of size $j+1-k$, we see that there must be some set $S$ of size $j+1-k$ for which at least

$$\binom{n}{j+1}\binom{\epsilon j/2}{k} / \binom{n}{j+1-k} = \binom{n-j-1+k}{k}\binom{\epsilon j/2}{k} / \binom{j+1}{k}$$

pairs $(I, J)$ have $J \setminus I = S$. Here we used the identity $\binom{n}{j+1}\binom{j+1}{k} = \binom{n}{j+1-k}\binom{n-j-1+k}{k}$.

Fixing $S$, we see that the pair $(I, J)$ is uniquely determined by the choice of $I$. Let $\mathcal{I}$ be the $k$-uniform hypergraph whose edges are the sets $I$ taken from these pairs. Since $n \geqslant \alpha^{-1/k} = 16k/\epsilon$ and $j \geqslant n/4$, we have $j \geqslant 4k/\epsilon$. This in turn implies that $\binom{\epsilon j/2}{k} / \binom{j+1}{k} \geqslant \left(\frac{\epsilon}{4}\right)^k$. Therefore, since $j < 3n/4$, the number of edges in $\mathcal{I}$ is at least

$$\left(\frac{\epsilon}{4}\right)^k \binom{n-j-1+k}{k} \geqslant \left(\frac{\epsilon}{16}\right)^k \frac{n^k}{k!}.$$

Hence, since $\left(\frac{\epsilon}{16}\right)^k \geqslant k!\alpha$, the hypergraph $\mathcal{I}$ contains a copy of $\mathcal{H}$. Suppose that $\mathcal{H}$ is defined on vertex set $[l]$ and the mapping $g : [l] \to [n]$ describes the embedding of $\mathcal{H}$ in $\mathcal{I}$.

We define a map $f : Q_l \to Q_n$ by mapping $[a_1 \ldots a_l]$, with non-zero bits $a_{i_1}, \ldots, a_{i_r}$, to $[b_1 \ldots b_n]$, where $b_i = 1$ if and only if $i \in S \cup \{g(i_1), \ldots, g(i_r)\}$. It is straightforward to verify that this is a graph isomorphism between $Q_l$ and $f(Q_l)$. Moreover, for every edge $e = uv \in H$, the edge $f(u)f(v)$ is in $G$. To see this, suppose that the non-zero bits of $e$ are $a_{i_1}, \ldots, a_{i_k}$ and $a_{i_\ell}$ is the flip-bit. Let $J = S \cup \{g(i_1), \ldots, g(i_k)\}$. By construction, $I = \{g(i_1), \ldots, g(i_k)\} \in \mathcal{I}$ and, therefore, $I \subset D(J)$. Hence, by the definition of $D(J)$, the edge formed by replacing $b_{g(i_\ell)}$ in $f(u)$ with a flip-bit is in $G$. But this edge is just the edge between $f(u)$ and $f(v)$. We therefore have an embedding of $H$ in $G$, completing the proof. $\square$

To complete the proof of Theorem 1.1, we only need to apply the following classical result of Erdős [10] regarding the extremal number of complete $k$-partite $k$-uniform hypergraphs.

**Lemma 2.2** *Let $K_k^{(k)}(s_1, \ldots, s_k)$ be the complete $k$-partite $k$-uniform hypergraph with partite sets of size $s_1, \ldots, s_k$. Then*

$$ex(K_n^{(k)}, K_k^{(k)}(s_1, \ldots, s_k)) = O(n^{k-\delta}),$$

*where $\delta = \left(\prod_{i=1}^{k-1} s_i\right)^{-1}$.*

This yields the following, more precise, version of Theorem 1.1.

**Corollary 2.1** *Let $H$ be a subgraph of the cube which admits a $k$-partite representation, where the partite sets have sizes $s_1, \ldots, s_k$. Then*

$$ex(Q_n, H) = O(2^n n^{1 - \frac{\delta}{k}}),$$

*where $\delta = \left( \prod_{i=1}^{k-1} s_i \right)^{-1}$.*

# 3 Concluding remarks

• It still remains to decide whether $\mathrm{ex}(Q_n, C_{10}) = o(e(Q_n))$. It even remains open to decide whether there is any graph $H$ which is Ramsey with respect to the cube but which does not have zero Turán density. We have also been unable to determine whether there are graphs with zero Turán density which do not have $k$-partite representations. We conjecture that these three collections of graphs, those with $k$-partite representations, those with zero Turán density and those which are Ramsey with respect to the cube, are all distinct.

• A quantitative version of Chung's result regarding the appearance of cycles in cubes states that, for $t$ even, $\mathrm{ex}(Q_n, C_{2t}) = O(2^n n^{\frac{1}{2} + \frac{1}{t}})$. This result follows easily from Theorem 2.1 and the fact that the 2-partite representation of $C_{2t}$ is the graph $C_t$. The one additional ingredient necessary to complete the proof is the Bondy-Simonovits theorem [5], that $\mathrm{ex}(K_n, C_t) = O(n^{1 + \frac{2}{t}})$ for $t$ even.

On the other hand, for every $t$, an application of the Lovász local lemma implies that there is a subgraph $G_t$ of $Q_n$ with $\Omega(2^n n^{\frac{1}{2} + \frac{1}{2t}})$ edges which does not contain a copy of $C_{2t}$. For $t$ even, we believe that the upper bound is tight but have been unable to make any progress towards proving this.

• For odd values of $t$, the behaviour of the function $\mathrm{ex}(Q_n, C_{2t})$ is even more obscure. Füredi and Özkahya [14, 17] give an upper bound of the form $\mathrm{ex}(Q_n, C_{2t}) = O(2^n n^{1 - \epsilon_t})$ for all odd values of $t$ with $t \geqslant 7$, where $\epsilon_t > 0$ tends to $1/16$ as $t$ gets large. If the 3-partite 3-uniform graph $\mathcal{E}_t$ that represents $C_{2t}$ satisfies $\mathrm{ex}(K_n^{(3)}, \mathcal{E}_t) = O(n^{2 + \delta})$, then Theorem 2.1 would imply that $\mathrm{ex}(Q_n, C_{2t}) = O(2^n n^{\frac{2}{3} + \frac{\delta}{3}})$, improving the result of Füredi and Özkahya for $\delta$ sufficiently small. While such an improved estimate almost certainly holds for large $t$, we have not pursued this direction.

# References

[1] N. Alon, A. Krech and T. Szabó, Turán's theorem in the hypercube, *SIAM J. Discrete Math.* **21** (2007), 66–72.

[2] N. Alon, R. Radoičić, B. Sudakov and J. Vondrák, A Ramsey-type result for the hypercube, *J. Graph Theory* **53** (2006), 196–208.

[3] M. Axenovich and R. Martin, A note on short cycles in the hypercube, *Discrete Math.* **306** (2006), 2212–2218.

[4] A. Bialostocki, Some Ramsey type results regarding the graph of the $n$-cube, *Ars Combin.* **16** (1983), 39–48.

[5] J. A. Bondy and M. Simonovits, Cycles of even length in graphs, *J. Combin. Theory Ser. B* **16** (1974), 97–105.

[6] P. Brass, H. Harborth and H. Nienborg, On the maximum number of edges in a $C_4$-free subgraph of $Q_n$, *J. Graph Theory* **19** (1995), 17–23.

[7] A. E. Brouwer, I. J. Dejter and C. Thomassen, Highly symmetric subgraphs of hypercubes, *J. Algebraic Combin.* **2** (1993), 25–29.

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

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

[10] P. Erdős, On extremal problems of graphs and generalized graphs, *Israel J. Math.* **2** (1964), 183–190.

[11] P. Erdős, On some problems in graph theory, combinatorial analysis and combinatorial number theory, in: Graph Theory and Combinatorics (Cambridge, 1983), Academic Press, London, 1984, 1–17.

[12] P. Erdős, Some of my favourite unsolved problems, in: **A tribute to Paul Erdős,** Cambridge University Press, 1990, 467–478.

[13] P. Erdős and M. Simonovits, A limit theorem in graph theory, *Studia Sci. Math. Hungar.* **1** (1966), 51–57.

[14] Z. Füredi and L. Özkahya, On 14-cycle-free subgraphs of the hypercube, *Combin. Probab. Comput.* **18** (2009), 725–729.

[15] Z. Füredi and L. Özkahya, On even-cycle-free subgraphs of the hypercube, *Electronic Notes in Discrete Mathematics* **34** (2009), 515–517.

[16] T. Kővári, V. Sós and P. Turán, On a problem of K. Zarankiewicz, *Colloq. Math.* **3** (1954), 50–57.

[17] L. Özkahya, personal communication.

[18] A. Thomason and P. Wagner, Bounding the size of square-free subgraphs of the hypercube, *Discrete Math.* **309** (2009), 1730–1735.

[19] P. Turán, On an extremal problem in graph theory (in Hungarian), *Mat. Fiz. Lapok* **48** (1941), 436–452.