Extremal Number of Trees in Hypercubes

Emily Jia* and Chiheon Kim†

October 22, 2015

Abstract

Given a tree T, we investigate bounds on the extremal number of T in the hypercube Q_d, defined as the maximum number of edges in a T-free subgraph of Q_d. We define a parameter that enables us to bound $\mathrm{ex}(Q_d, T)$ for all trees and present an analog of the Erdős-Sós conjecture in the hypercube. We calculate the extremal numbers for specific families of trees and compare them to the general bound. We demonstrate trees that achieve the lower bound and others whose extremal number is almost twice as much. From there, we provide a restriction on minimum degree that guarantees the existence of trees in subgraphs of the hypercube.

1. Introduction

The first problems in graph theory date back to 1736, when Leonhard Euler [6] determined it was impossible to walk through the city of Konigsberg and cross all seven bridges exactly once. Ever since, famous proposals such as the Traveling Salesman and Map Coloring problems have combined simple, real-life premises with graph theory research. Furthermore, the proliferation of complex physical and technological networks in the late 20th century has generated significant interest in graph theory over the past decades. Branches of graph theory include algorithmic graph theory, random graph theory, and the subject of our project: extremal graph theory.

Given a simple graph G and a family F of graphs, let the extremal number $\mathrm{ex}(G, F)$ denote the maximum number edges a subgraph of G can have without containing a graph in F. Problems that involve determining $\mathrm{ex}(G, F)$ are called *Turán-type extremal problems* and the maximal graphs are called *extremal.*

The earliest problems in extremal graph theory use a complete graph as the host graph. Forbidden graphs F include cycles, paths, and smaller complete graphs. The results are written in terms of graph parameters. A direct example is Turán's theorem [8], which bounds the extremal number for complete graphs on $r+1$ vertices in complete graphs on n vertices.

$$\mathrm{ex}(K_n, K_{r+1}) = \left(1 - \frac{1}{r}\right) \cdot \frac{n^2}{2}$$

*RSI, Department of Mathematics, Massachusetts Institute of Technology.

†Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139-4307. Email: chiheonk@math.mit.edu.

Dirac's Theorem [3] is another result central to extremal graph theory. Any graph with $n \geq 3$ vertices and minimum degree $\delta(G) \geq \frac{n}{2}$ contains a Hamiltonian cycle. An extension of Dirac's demonstrates that a graph with average degree d contains a subgraph with minimum degree of at least $\frac{d}{2}$. Thus there exists some relation between global and local properties that can affect substructures in the graph, and consequently, extremal number.

Only recently has extremal graph theory been studied in host graphs other than the complete graph K_n and the complete bipartite graph $K_{m.n}$. In 1984, Erdős [4] proposed one of the first questions in $\text{ex}(Q_d, G)$ by asking the number of edges in the maximal C_4-free subgraph of the hypercube. This remains an open problem as current bounds by Brass et al. [2] and Baber [1] still are not tight:

$$(d + \sqrt{d})2^{d-2} \leq \text{ex}(Q_d, C_4) \leq 0.6068d2^{d-1}.$$

The hypercube's potential as a network topology in parallel computing has generated new interest in exploring the extremal properties of its substructures. In 2010, Eoin Long [7] showed that there exists a path of length $2^d - 1$ in any subgraph of a hypercube with minimum degree d. Our project asks a similar question about trees in hypercubes: how can their extremal number be bounded?

There are no prior results on $\text{ex}(Q_d, T)$. However, in the complete graph K_n, bounds on on the extremal number of trees on $k+1$ vertices are well-known:

$$\frac{k-1}{2} \cdot n \leq \text{ex}(K_n, T) \leq (k-1) \cdot n.$$

The Erdős-Sós conjecture [5], an open problem since its proposal in 1962, claims that equality holds for the lower bound.

Our project defines a parameter δ_T to provide similar bounds on the extremal number of any tree T in the hypercube Q_d. We prove that there exist trees with extremal number close to each of the upper and lower bounds, thus showing a statement analogous to the Erdős-Sós conjecture does not hold in the hypercube. A related goal of the project is to demonstrate methods of calculating δ_T values for specific families of trees. We show its relation to a known parameter used in Long's paper.

The organization of the paper is as follows. In Section 2, we define additional terms used in our paper and prove the general bound on $\text{ex}(Q_d, T)$. Then, in Section 3, we present our bounds on trees whose δ_T value can be easily computed and compare their extremal number to the general bounds. In Section 4, we show methods of calculating δ_T for certain families of trees with higher diameters. Finally, in Section 5, we summarize our results, examine their relation to other work in extremal graph theory, discuss the applications of our work, and propose future directions of research.

2. Definitions and a Novel Parameter

We present unique terms used in this paper as well as a new parameter for a general bound.

2.1 Preliminary Definitions

Definition 2.1. For graphs H and F, the *extremal number* $\text{ex}(H, F)$ is the number of edges in the maximal subgraph of H not containing any copy of F.

Definition 2.2. A *hypercube* Q_d is a d-regular graph with 2^d vertices and $d2^{d-1}$ edges. We use three systems of expressing Q_d.

1. Write Q_d as two copies of Q_{d-1} that have additional edges drawn between corresponding vertices.
2. Write the vertex set as $V(Q_d) = \{0, 1\}^d$ such that each vertex is assigned a unique length d binary string of 0's and 1's. Vertices are adjacent if their string representations differ in exactly one position.
3. Write the vertex set as $V(Q_d) = \{0, 1\}^d$ and assign each vertex v a unique set $\{v\}$ containing the positions of 1's in their string representations.

The third expression leads to some terminology that allows us to group vertices with similar properties.

Definition 2.3. The *size* of v_0, denoted $|v_0|$, is the magnitude of its set $\{v_0\}$.

Definition 2.4. In a graph $X \subseteq Q_d$, define the t^{th} layer as the set of vertices in X with size t. We denote this layer as $L_t(X)$, or simply L_t.

Noting the exponential relationship between minimum degree and path or cycle length in Eoin Long's work [7], we present a known parameter used to describe subgraphs of a hypercube.

Definition 2.5. The *cubical dimension* of a graph G, expressed $\text{cd}(G)$, is the smallest d such that $G \subseteq Q_d$.

We also define the structures we examine in this paper.

Definition 2.6. A *star* S_n is a tree with n vertices of degree 1 and one vertex of degree n.

Definition 2.7. A *modified star* S'_n is a star S_n with an additional leaf attached to one of its edges.

Definition 2.8. A *twin star* $C_{j,j}$ is a tree created by connecting the central vertices of S_j and S_j with an extra edge.

Definition 2.9. A *subdivided star* S^n_k is a star with a central vertex and n disjoint paths of length k emanating from this vertex.

2.2 General Bounds

Finally, we define our own parameter δ_T that we use to bound the general case of $\text{ex}(Q_d, T)$.

Definition 2.10. Let S be a subgraph of hypercube Q_d. Let $\delta(S)$ denote the minimum degree of S. For any tree T, we define δ_T such that $T \subseteq S$ if $\delta(S) \geq \delta_T$ and $T \nsubseteq S$ otherwise. In essence, δ_T is the minimum degree condition for S that guarantees $T \subseteq S$.

Theorem 2.1. *The extremal number for any tree T in the d-dimensional hypercube Q_d is bounded*

$$\frac{\delta_T - 1}{2} 2^d \leq \text{ex}(Q_d, T) < \delta_T 2^d.$$

Proof. For the lower bound, consider the graph G containing the union of $2^{d-(\delta_T - 1)}$ disjoint $Q_{\delta_T - 1}$ in Q_d. This graph satisfies $\delta(G) < \delta_T$ so it is a construction of a T-free subgraph of Q_d with $\frac{\delta_T - 1}{2} 2^d$ edges.

By Dirac's theorem, there exists a subgraph with minimium degree δ_T in a graph with average degree $2\delta_T$. Thus, any $S \subseteq Q_d$ with $\delta_T 2^d$ edges contains T. □

Remark. We note a similarity between the general bound for trees in the complete graph and trees in the hypercube. As stated earlier, for tree T on $k + 1$ vertices

$$\frac{k-1}{2} \cdot n \leq \text{ex}(K_n, T) \leq (k-1) \cdot n.$$

This motivates us to examine an analog of the Erdős-Sós conjecture in the hypercube.

Conjecture 2.2. Equality holds for the lower bound of $\text{ex}(Q_d, T.)$

$$\text{ex}(Q_d, T) = \frac{\delta_T - 1}{2} 2^d$$

3. Bounding Extremal Number using Average Degree

We bound the extremal number for low-diameter structures such as stars, modified stars, and twin stars. A lower bound is constructed and an upper bound is provided with a counting argument. We then demonstrate that Conjecture 2.2 is false; a relation like the Erdős-Sós conjecture does not exist in the hypercube.

3.1 Stars

Every subgraph of Q_d with average degree greater than $n - 1$ contains S_n. An extremal graph X can be constructed by taking the union of $2^{d-(n-1)}$ copies of Q_{n-1}. By the recursive definition of a hypercube, this is possible for all $d \geq n-1$. This construction creates an $n-1$ regular graph so all vertices have degree less than n. Thus, $S_n \nsubseteq X$ and

$$\text{ex}(Q_d, S_n) = \frac{1}{2}(n-1)2^d.$$

3.2 Modified Stars

We consider the extremal number for modified stars find that every subgraph G of Q_d with average degree greater than $n-1$ contains the modified star S'_n. This is equivalent to the following theorem.

Theorem 3.1. *The extremal number for S'_n is*

$$\mathrm{ex}(Q_d, S'_n) = \frac{1}{2}(n-1)2^d.$$

Proof. To demonstrate the upper bound, we use contradiction. Assume that $\mathrm{ex}(Q_d, S'_n) > \frac{1}{2}(n-1)2^d$. Then the average degree of an extremal graph X is greater than $n-1$ and there exists a vertex $v \in X$ such that $\deg(v) \geq n$. The neighbors of v must all have degree 1 for the graph to be extremal. Else, $S'_n \subseteq X$, which is forbidden.

For all vertices $v' \in X$ such that $\deg(v') > n-1$, there must be a disjoint star $S_{\deg(v')}$ that contains v'. Else $S'_n \subseteq X$ Then we can partition X into two disjoint sets. Let Γ_1 denote the set of disjoint stars in X and Γ_2 denote the set of other vertices and edges. The average degree in Γ_1 must be less than 2. Furthermore, the maximum degree of any vertex in Γ_2 is $n-1$. Thus, the average degree of X cannot be greater than $n-1$, contradiction.

We construct the same extremal graph as $EX(Q_d, S_n)$. Because X does not contain S_n, X also does not contain S'_n. Thus $\frac{1}{2}(n-1)2^d \leq \mathrm{ex}(Q_d, S'_n) \leq \frac{1}{2}(n-1)2^d$ and the bounds for the extremal number are tight. □

Remark. We note some interesting characteristics of these two extremal numbers. In K_n, the extremal number for a large graph is usually greater than the extremal number for a smaller graph. However, we demonstrate that $\mathrm{ex}(Q_d, S_n) = \mathrm{ex}(Q_d, S'_n)$. Furthermore, S_n and S'_n can be embedded in Q_n but not Q_{n-1}, so $\mathrm{cd}(S_n) = \mathrm{cd}(S'_n) = n$. Thus the average degree in the extremal graph is one less than the cubical dimension of the stars.

3.3 Twin Stars

We find that every subgraph G of Q_d with average degree greater than $n-1$ contains the modified star S'_n. This is equivalent to the following theorem.

Theorem 3.2. *The extremal number for twin star $C_{k,k}$ is*

$$\mathrm{ex}(Q_d, C_{k,k}) \leq \frac{(k-1)d}{d+k-1} \cdot 2^d.$$

When $d = 2^n - (k-1)$ for integer k, equality for the upper bound holds.

Proof. We first provide a counting argument for the upper bound. Let $X \subseteq Q_d$ be a $C_{k,k}$ free graph. Partition the vertices of X into two sets, A and B, such that A is the set of all vertices $v_a \in V(X)$ satisfying $\deg(v_a) < k$, and B is the set of all vertices in $v_b \in V(X)$ satisfying $\deg(v_b) \geq k$. Because X does not contain $C_{k,k}$, no two vertices in B can be adjacent.

Denote the size of B as b. This implies that the size of A is $2^d - b$. Furthermore let y denote the number of edges with an endpoint in A and another in B, and let x denote the number of edges with both endpoints in A. We wish to maximize $|E(X)| = x + y$.

Every vertex in B has degree at least k and at most d. This implies $kb \leq y \leq db$. Every vertex in A has degree at most $k-1$, so the sum of the degrees of each vertex in A is at most $(k-1)(2^d - b)$. Each edge with both endpoints in X is counted twice in this sum, implying $2x + y \leq (k-1)(2^d - b)$.

Now we wish to maximize $x + y$ in the system

$$\begin{cases} kb \leq y \leq db \\ 2x + y \leq (k-1)(2^d - b) \end{cases}$$

for non-negative integers a, b, k, d. Consider the solution set of these inequalities for constant b in the x-y plane. Because $2x+y$ has a steeper negative slope than $x+y$, $x+y$ is maximized at the maximum value y in the solution set, and $y = db$.

Substituting this into the second inequality, we have $2x + db \leq (k-1)(2^d - b)$ and we can solve for $x = \frac{1}{2}[(k-1)2^d - (d+k-1)b]$. Then we can rewrite $|E(X)| = x + y \leq \frac{1}{2}(k-1)2^d + [d - \frac{d+k-1}{2}]b$. Because $d > k-1$, (otherwise the caterpillar cannot be embedded in Q_d) $|E(X)|$ is maximized when b is maximized. However x is nonnegative so $b \leq \frac{k-1}{d+k-1}2^d$ with an extremal graph when $x = 0, y = db, b = \frac{k-1}{d+k-1}2^d$.

For values of b satisfying $db \geq (k-1)(2^d - b)$ or $b \geq \frac{k-1}{d+k-1}2^d$, the system of inequalities does not have solutions on the line $y = db$. Then $x + y$ is still maximized at the largest value of y, which occurs at $x = 0, y = (k-1)(2^d - b)$. We substitute to get $|E(X)| = x + y = (k-1)(2^d - b)$ which is maximized when b is minimized. Thus the extremal graph occurs when $x = 0, y = db, b = \frac{k-1}{d+k-1}2^d$.

Both cases produce the same value of $|E(X)| = x + y = \frac{k-1}{d+k-1}2^d$ as the maximum number of edges in an extremal graph. Furthermore $x = 0$ implies that X is bipartite with one partition containing degree d vertices and the other containing degree $k-1$. Now we provide a construction for when the upper bound is tight.

Using Definition 2.2, write $V(Q_d) = \{0,1\}^d$. Let $d = (k-1)(2^n - 1)$. Then H denotes the n by d matrix with each of the binary representations of 1 through $2^d - 1$ appearing exactly $k-1$ times in its columns, ordered from least to greatest with leading zeroes.

Let B be the set of all column vectors $y \in \{0,1\}^d$ that satisfy $Hy = 0$ in $\mathbb{Z}_2$. Let e_j denote a column vector with a 1 in the ith entry and 0's elsewhere. We claim that for any column vector $z \in \{0,1\}^d$ satisfying $Hz \neq 0$, there exist $k-1$ distinct column vectors $e_{j+1}, ..., e_{j+k-1}$ such that $z + e_{j+1}, ..., z + e_{j+k-1} \in B$. We show this with a lemma.

Lemma 3.3. *Consider the column vectors in A and B as vertices of Q_d. Draw an edge between elements $z \in A$ and $y \in B$ if and only if $z + e_j = y$ for some $j \leq d$. Then A and B create a bipartite graph X on 2^d vertices such that every vertex in A has degree $k-1$ and every vertex in B contains vertices of degree d.*

Proof. For $z \in A$, let $Hz = w$ for $w \neq 0$. Because we work in $\mathbb{Z}_2$, w must appear $k-1$ times in the columns of H, and also $w + w = 0$. Thus, for each z there are $k-1$ distinct e_i such that $z + e_i \in B$ and each element in A has degree $k-1$.

Similarly for any $y \in B$, the product $H(y + e_i)$ equals the ith column of H. Because all columns of H are non-zero, all vectors $y + e_i$ are contained in A. Furthermore the vector y has d entries, so there are d distinct vectors $e_1, ..., e_d$. Thus each element in B has degree d. □

Because X is bipartite, the ratio of $|V(A)|$ to $|V(B)|$ equals d to $k-1$. Thus the number of edges in X is $\frac{(k-1)d}{d+k-1} \cdot 2^d$ and we have a construction that demonstrates a tight upper bound

$$\mathrm{ex}(Q_d, C_{k,k}) = \frac{(k-1)d}{d+k-1} \cdot 2^d$$

for d in the form $(k-1)(2^n - 1)$. □

Remark. This construction resembles the Hamming Code, an error connecting code used in information theory. However, our construction that uses H does not work well for $d \neq (k-1)(2^n-1)$. Because n must be an integer satisfying $n \geq \lceil \log_2(\frac{d}{k-1} - 1) \rceil$, the average degree in A decreases exponentially as n increases linearly, creating a weak lower bound.

It is also interesting to note that the average degree of the upper bound construction asymptotically approaches $2(k-1)$, even though having an average degree of $2(k+1)$ is sufficient to guarantee the existence of $C_{k,k}$. Comparing this to Theorem 2.1:

$$\frac{\delta - 1}{2} 2^d \leq \mathrm{ex}(Q_d, T) < \delta 2^d.$$

we extremal number for stars and modified stars is equal to the lower bound. Meanwhile, the extremal number provided by Theorem 3.2 is very close to the upper bound.

We now move on to finding δ_T which gives bounds on extremal number.

4. Bounding Extremal Number using Minimum Degree

We bound the extremal number for trees with cubical dimension three, subdivided stars, and depth two trees by expressing subgraphs of hypercubes in terms of their layers. We show that certain minimum degree conditions are sufficient to guarantee the existences of these trees in a subgraph $X \subset Q_d$.

4.1 Trees with Cubical Dimension 3

Earlier, we observed that $\mathrm{cd}(S_n) = \mathrm{cd}(S'_n) = n$, and Section 3.1 and 3.2 demonstrate that any subgraph $G \subseteq Q_d$ satisfying $\delta(G) \geq n$ contains S_n, S'_n. Furthermore it is simple to show that $\mathrm{cd}(C_{k,k}) = k + 1$ and any $M \subseteq Q_d$ with $\delta(G) \geq k + 1$ contains $C_{k,k}$. We investigate all trees T satisfying $\mathrm{cd}(T) = 3$ to see if they could be embedded into a graph with minimum degree 3.

Theorem 4.1. *Let M denote a subgraph of Q_d with minimum degree 3. Let T_3 denote the set of trees with cubical dimension 3. For all trees t in T_3, t is a subgraph of M.*

The main idea behind this proof is that, for every tree in T_3, we can assign a specific vertex to L_0 and use the size of its neighbors to show that the entire tree exists in any subgraph with minimum degree 3. We provide an example here. The details of this proof can be found in Appendix A.

Case 4.1. We claim that the tree in Figure 4.1 can be embedded in all $M \subseteq Q_d$ with $\delta(M) \geq 3$.

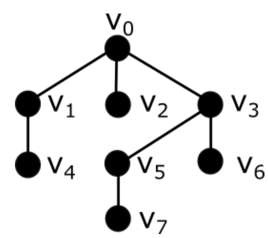

Figure 1: Labeling for Case 4.1

Proof. Without loss of generality let $|v_0| = 0$. Then neighbors v_1, v_2, v_3 must exist in M such that they all have size 1. V_1 must have at least two neighbors of size 2, so we select one not adjacent to v_3 and label it v_4. Then we know that v_3 has at least two neighbors of size 2, so we label them v_5, v_6. Finally v_5 must have at least one neighbor of size 3 in M and we label it v_7. □

Using casework for trees with cubical dimension 4 is difficult for two reasons: first, the number of such trees is much higher, and second, trees with long central paths may require constructing edges that start in L_t and end in L_{t-1}. However, noting that this method seems to work with trees that have small depth, we investigate a minimum degree bound for depth two trees.

4.2 Trees with Depth 2

In this section we calculate the minimum degree condition that guarantees the existence of certain depth two trees. We note that the set of vertices in a tree with the same depth is analogous to a layer in a hypercube. We also use a restriction argument, similar to that of Section 4.1, to find neighbors of vertices that are not adjacent to other vertices.

Theorem 4.2. *Let X denote a subgraph of Q_d with minimum degree k. Call a tree a depth two tree if one of its vertices can be assigned as a root such that all of the other vertices have depth no greater than two.*

Let T_k denote the set of depth 2 trees that have at most k edges emanating from the root and at most $\lfloor \frac{k-1}{2} \rfloor + 1$ edges emanating from non-root vertices. Then $T \subset X$.

Proof. Without loss of generality, assign the root of the depth two tree T_k to L_0. Because $\delta(X) = k$, the root must have k neighbors in $L_1(X)$. Label these neighbors as $v_1, v_2, ..., v_k \in L_1$. Each vertex has at least $k-1$ neighbors in L_2.

Consider vertex v_j for $1 \leq j \leq k$. Because j has at least $k-1$ neighbors in L_2, we can connect v_j to all of its neighbors that are not shared with $v_{j+1}, ..., v_{j+1+\lceil \frac{k-1}{2} \rceil}$, where subscripts are taken mod k, and guarantee that $\deg v_j \geq \lfloor \frac{k-1}{2} \rfloor + 1$. It suffices to show that if we repeat this restriction for all v_j, then we will have the largest tree in T_k.

We use contradiction. Assume that the resulting structure is not a tree- else, it is the largest tree in T_k. This implies that two vertices in L_1 have an edge drawn to the same vertex in L_2. Then there must exist v_g, v_h such that $v_g \notin v_{h+1}, ..., v_{h+1+\lceil \frac{k-1}{2} \rceil}$ and $v_h \notin v_{g+1}, ..., v_{g+1+\lceil \frac{k-1}{2} \rceil}$. This implies that there must be at least $2 + 2\lceil \frac{k-1}{2} \rceil$ vertices v_j, contradiction.

□

This restriction method requires significant casework to translate directly to degree 3 stars. However, considering the number and formation fo edges between layers of X may be useful for generalizing to all trees in the future. We move on to subdivided stars.

4.3 Subdivided Stars

In this section, we calculate the minimum degree condition that guarantees the existence of some S_n^k. We first present a lemma that double counts the edges between adjacent layers of a hypercube in order to relate the sizes of the layers.

Lemma 4.3. *For all subgraphs of Q_d with minimum degree k, let S denote some set of vertices $S \subseteq L_t$ and let $N(S)$ denote the neighbors of S in L_{t+1}. Then*

$$\binom{|N(S)|}{2} \geq \binom{k-t}{2} |S|.$$

Proof. Let p denote the number of pairs of vertices v_g, v_h in $N(S)$ that share a neighbor in S. Each vertex of a pair in in $N(S)$ shares at most one neighbor in S with the other, else the two vertices in the pair are not distinct. Hence the maximum number of pairs is $p \leq \binom{|N(S)|}{2}$.

However, there may not be enough edges between S and $N(S)$ so that every pair of vertices in L_{t+1} share a neighbor in L_t. If $\delta(X) = k$, any vertex in L_t has at most $\binom{t}{t-1}$ neighbors in L_{t-1}, so each vertex in L_t has at least $k-t$ neighbors in L_{t+1}. Hence there are at least $\binom{k-t}{2}$ pairs of vertices in L_{t+1} that are adjacent to each $v_t \in L_t$ and $\binom{k-1}{2}|S| \leq p$.

Thus $\binom{|N(S)|}{2} \geq p \geq \binom{k-t}{2}|S|$. □

With this lemma, we use Hall's Marriage Theorem find perfect matchings between layers of a hypercube and guarantee the existence of subdivided stars.

Theorem 4.4. *In any subgraph X in Q_d with minimum degree k, we can find a subdivided star with paths of length*

$$\left\lfloor \frac{2k-1}{2} - \sqrt{\frac{4k-3}{4}} \right\rfloor.$$

Proof. Without loss of generality, place the center of the star at L_0. We wish to calculate t such that there is a perfect matching from L_{i-1} to L_i for all $i \leq t$. By Hall's Marriage Theorem it suffices to show that for any set $S \in L_{i-1}$, and the set of its neighbors $N(S) \in L_i$, $|N(S)| \geq |S|$.

It is simple to show that $|N(S)| \geq |S|$ if $\binom{|N(S)|}{2} \geq \binom{|S|}{2}$. From Lemma 4.3 this gives the restriction $\binom{k-t}{2} \geq \frac{|S|-1}{2}$ which is equivalent to $(k-t)(k-t-1)+1 \geq |S|$. Each layer needs to contain at least k vertices that have a perfect matching, so substituting $|S| = k$ and completing the square in terms of t gives us:

$$t \leq \frac{2k-1}{2} - \sqrt{\frac{4k-3}{4}}.$$

Thus, any graph $X \subseteq Q_d$ with $\delta(X) = k$ contains a subdivided star with paths of length $\left\lfloor \frac{2k-1}{2} - \sqrt{\frac{4k-3}{4}} \right\rfloor$. □

Remark. This method creates bounds for stars in which the size of vertices increases as the distance from the central vertex increases. The length of a subdivided star's paths can be greater if the vertices in the path do not strictly increase in size.

5. Discussion and Future Work

In this paper, we provide the first results on the extremal number of trees in hypercubes. We derived a general bound by defining a new parameter δ_T and compared the extremal number of specific trees to this general bound.

$$\frac{\delta_T - 1}{2} 2^d \leq \mathrm{ex}(Q_d, T) < \delta_T 2^d$$

For stars, modified stars, and twin stars whose δ_T value is straightforward to compute, we were able to bound the extremal number more tightly. While stars and modified stars had extremal numbers equal to the lower bound, the twin stars had an extremal number close to the upper bound. We then demonstrated methods of calculating δ_T for structures with higher diameter.

Nonetheless, the problem of bounding $\mathrm{ex}(Q_d, T)$ is far from resolved. A direction for future research remains in calculating δ_T. Although we demonstrate methods of determining δ_T for specific types of trees, there is no general method for calculating δ_T.

For all trees investigated in this paper, we find that $\delta_T = \text{cd}(T)$. Along with the results from Long's [7] paper, which relates the cubical dimension of paths to the minimum degree of their extremal graph, this leads us to conjecture

Conjecture 5.1. For all trees T,

$$\delta_T = \text{cd}(T).$$

Equivalently, all cubical dimension T trees can be found in a subgraph $X \subseteq Q_d$ if $\delta(X) = \text{cd}(T)$.

As with δ_T, there are no known methods of calculating $\text{cd}(T)$ for any given T. Proving this conjecture would generate quantitative insight as how the structure a tree influences its ability to be embedded in a non-complete graph.

The applications of our project and future research will assist in constructing parallel network architecture, which is largely based on the hypercube graph. For networks storing location-sensitive information, it is desirable to restrict connectivity structures that would allow for the rapid propagation of confidential data between processors in the case of an attack. Our project provides the maximum number of links that a network can have without containing potentially dangerous structures. This would also be useful in quantum computing, in which entangled qubits are liable to collapse together. By maximizing the amount of entanglement, without becoming too entangled, our project demonstrates how to construct efficient and stable networks.

6. Acknowledgments

I would like to express my sincerest gratitude to my mentor Mr. Chiheon Kim for his guidance and support while completing this research. I would also like to thank Dr. Tanya Khovanova and Dr. John Rickert for their helpful advice regarding the project's content and this paper. Furthermore I would like to thank Professor David Jerison, Professor Ankur Moitra, and Dr. Slava Gerovitch for supervising the research process. Finally I would like to acknowledge the Center of Excellence in Education and the Massachusetts Institute of Technology for hosting the Research Science Institute, as well as my sponsors Mr. and Mrs. Sam Leung, Dr. and Mrs. Jason Koh, Dr. Zhengtian Lu and Mrs. Diyang Wu, Mr. and Mrs. Daniel Witte, Dr. and Mrs. Choong H. Koh, and Mr. Nicholas S. Gouletas for their support.

References

[1] R. Baber, "Turan densities of hypercubes," *arXiv:1201.3587*, 2012.

[2] P. Brass, H. Harborth, and H. Nienborg, "On the maximum number of edges in a c4-free subgraph of qn," *Journal of Graph Theory*, vol. 19, no. 1, pp. 17–23, 1995.

[3] G. A. Dirac, "Some theorems on abstract graphs," *Proceedings of the London Mathematical Society*, vol. 3, no. 1, pp. 69–81, 1952.

[4] P. Erdős, "On some problems in graph theory, combinatorial analysis, and combinatorial number theory," *Graph Theory and Combinatorics*, pp. 1–17, 1984.

[5] P. Erdős, A. Rényi, and V. Sós, "On a problem in the theory of graphs," *Publ. Math. Inst. Hungar. Acad. Sci*, vol. 7, pp. 215–235, 1962.

[6] L. Euler, *The Seven Bridges of Konigsberg.* Wm. Benton, 1956.

[7] E. Long, "Long paths and cycles in subgraphs of the cube," *Combinatorica*, vol. 33, no. 4, pp. 395–428, 2013.

[8] P. Turán, "On an extremal problem in graph theory," *Mat. Fiz. Lapok*, vol. 48, no. 436-452, p. 137, 1941.

7. Appendix

We complete the casework for Theorem 4.1 and demonstrate that every spanning tree in Q_3 can be found in a subgraph $M \subseteq Q_d$ satisfying that has minimum degree three. From Long's work, is is known that a path of length seven can be found in M.

Case A.1. We claim that the tree in Figure 2 can be found in M.

Proof. Assign v_0 to $L_0(M)$. Then neighbors v_1, v_2, v_3 must exist in M such that they all have size 1. V_1 must have at least two neighbors of size 2 in M, so we select one that is not adjacent to v_3 and label it v_4. Then we label a size 2 neighbor of v_3 as v_6. V_6 must have at least one size 3 neighbor, so find we v_7 such that $|v_7| = 3$. This implies that v_7 has either a size 2 or a size 4 neighbor that is not already labeled, so we assign it as v_5. □

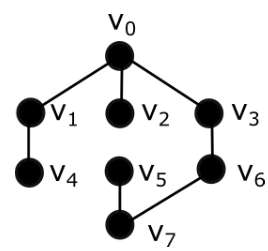

Figure 2: Labeling for Case A.1.

Case A.2. We claim that the tree in Figure 3 can be found in M.

Proof. Assign v_0 to $L_0(M)$. Then its neighbors v_1, v_2, v_3 all have size 1. Because $\deg v_1 \geq 3$, we know that v_1 must have at least two neighbors in M that have size 2, which we label v_4, v_5. Similarly v_3 must have at least two neighbors of size 2, so at least one of its neighbors is not adjacent to v_1. We label this as v_6. Since $\deg v_6 \geq 3$, v_6 must have at least one neighbor of size 3, which we label as v_7. □

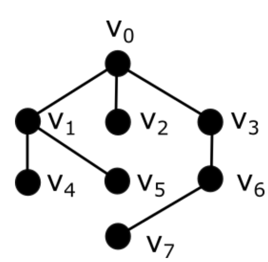

Figure 3: Labeling for Case A.2.

Case A.3. We claim that the tree in Figure 4 can be found in M.

Proof. Assign v_0 to L_0. Then neighbors v_1, v_3 must exist in M such that $|v_1| = |v_3| = 1$. Because $\deg v_3 \geq 3$, we can pick a neighbor of v_3 not adjacent to v_1 and label it v_6. Then we label two other neighbors of v_6 as v_2, v_7. Since $|v_6| = 2$ and v_6 does not neighbor v_3, we can also find two neighbors of v_1 with size 2 and label them v_4, v_5. □

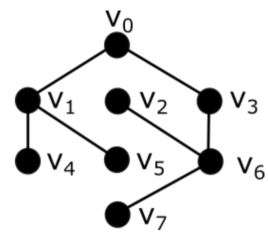

Figure 4: Labeling for Case A.3.

Case A.4. We claim that the tree in Figure 5 can be found in M.

Proof. Assign v_0 to L_0. Then we can find three of its neighbors v_1, v_2, v_3 in M that have size 1. We can find a neighbor of v_1 that is not adjacent to v_2 and label it v_4. Similarly we can find a neighbor of v_2 not adjacent to v_3 and label it v_5, and we can find a neighbor of v_3 not adjacent to v_1 and label it v_6. Then v_5 has at least one neighbor in M of size 3, and we label it v_7. □

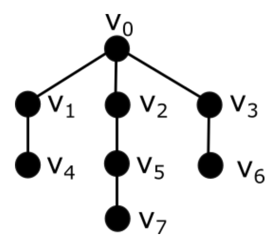

Figure 5: Labeling for Case A.4.

Along with Case 4.1 from Section 4, this casework demonstrates that all spanning trees in Q_3, and thus all trees in Q_3 exist in a subgraph $M \subseteq Q_d$ if $\delta(M) \geq 3$. Furthermore, it shows that given any vertex of M, we can find any spanning tree of Q_3 rooted at that vertex.