Saturation Number of Trees in the Hypercube

Kavish Gandhi

Research Science Institute

3 Main Results

Theorem (General Lower Bound)

Given a graph G with minimum degree δ,

$$\mathrm{sat}(Q_n, G) \geq (\delta - 1 + o(1)) \cdot 2^{n-2}.$$

Theorem (Upper Bound on Stars)

Given the star S_k,

$$\mathrm{sat}(Q_n, S_k) \leq (k - 2 + o_k(1))2^{n-1}.$$

Theorem (Upper Bound on Generalized Stars)

Let m and k be positive integers, and let $m' = \lfloor \log_2(m-1) \rfloor$ and $j = \lceil \log_2(m - 2^{m'}) \rceil$. Then

$$\mathrm{sat}(Q_n, GS_{k,m}) \leq (k + 1 + m' + \frac{j}{2^{m'}}) \cdot 2^{n-2}.$$

Theorem (Upper Bound on Paths)

Let $k \geq 4$ and $i = \lfloor \log_2(k-1) \rfloor$. Then,

$$\mathrm{sat}(Q_n, P_k) \leq \begin{cases} (i+1)\frac{k-1}{2^{i+1}}2^{n-1} & \text{if } k \text{ is odd} \\ \left(\frac{i(k-2)}{2^{i+1}} + 1\right) 2^{n-1} & \text{if } k \text{ is even.} \end{cases}$$

Theorem (Upper Bound on Caterpillars)

Given $S_{k_1 \times k_2 \times \cdots \times k_m}$ with $\min\{k_1, k_2, \ldots, k_n\} = k_i = r$ for $\lfloor \log_2 m \rfloor < i < m - \lfloor \log_2 m \rfloor$, then

$$\mathrm{sat}(Q_n, S_{k_1 \times k_2 \times \cdots \times k_m}) \leq r \cdot 2^{n-1}$$

Theorem (Upper Bound on General Trees)

Let $\mathcal{T}$ the class of trees for which the removal of some edge splits the graph into two trees with smaller cubical dimension than the original tree. Then, given this cubical dimension is k for $T \in \mathcal{T}$,

$$\mathrm{sat}(Q_n, T) \leq k \cdot 2^{n-1}.$$

I also found tight bounds for symmetric trees around a central node with certain dimensional properties.

Key points about these results:

- We have found a plethora of decently tight bounds for the saturation number
- They are all $O(2^n)$ with either a linear or constant multiplier. We conjecture this to be true for all trees.
- To find these bounds, we developed two novel methods: *disjoint subcubes* and *amplified Hamming codes*. These will be discussed further in the next section.

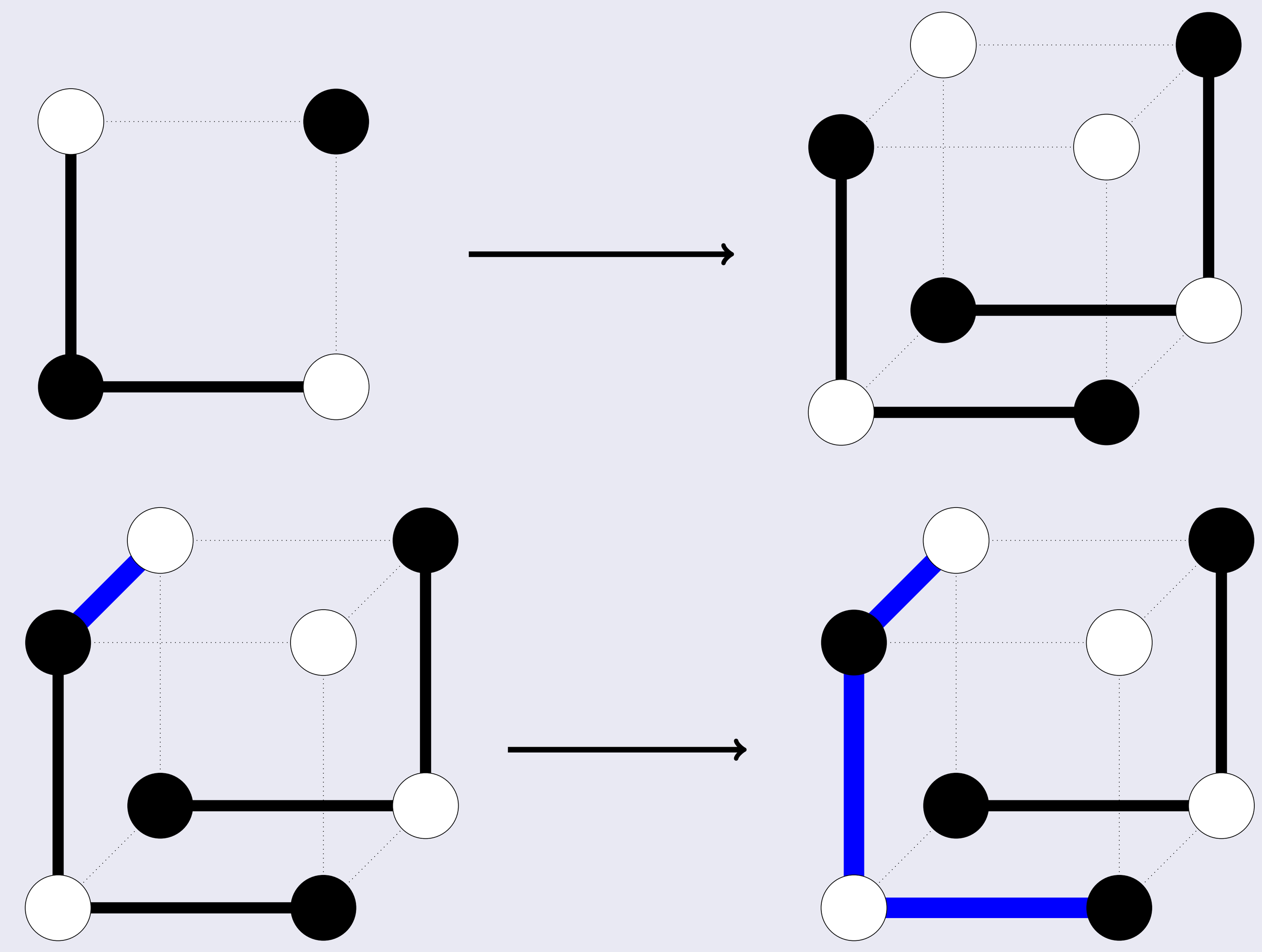
4 Method 1: Disjoint Subcubes

The main idea of this method is the following:

1. We split the tree into two components, each of which can fit into a subcube Q_k that the original tree cannot.
2. By copying this subgraph 2^{n-k} times, we create a saturated subgraph of Q_n.

By modifying this construction to only include a portion of the edges in each Q_k, we were able to find:

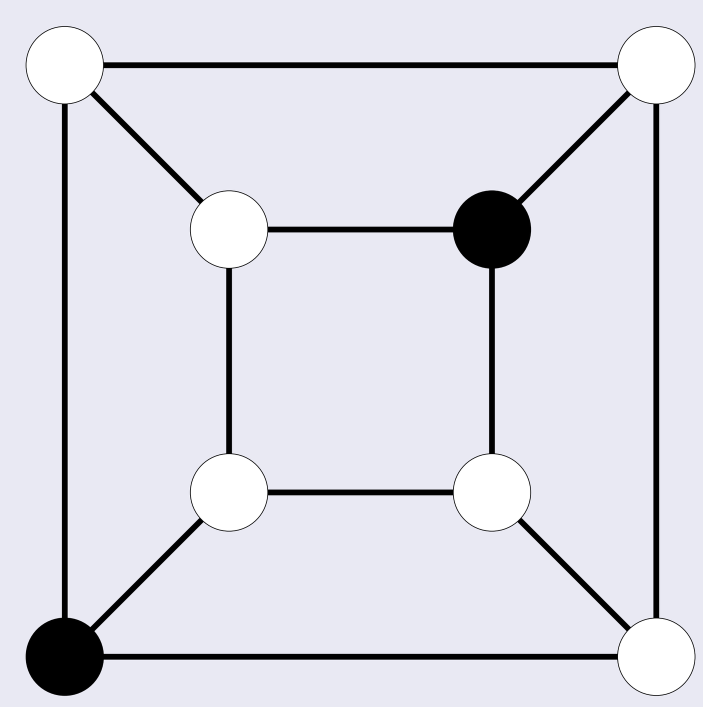
- A general upper bound for trees
- A tight upper bound for generalized stars and paths



5 Method 2: Amplified Hamming Codes

A Hamming code C is a special set of vertices with the following properties:

- $|C| = \frac{2^n}{n+1}$.
- The distance between all pairs of vertices is at least 3.
- C dominates Q_n. In other words, every vertex in $V(Q_n) \setminus C$ is adjacent to exactly one vertex in C.



A Hamming code is a dominating set of points, or Q_0's, in Q_{2^i-1}. It is also possible to have a dominating set of Q_k's in Q_{2^i-1+k} by amplifying the Hamming code. From these dominating sets, we derive our second method:

1. Fix a dominating set S of Q_k's in Q_n.
2. Around and within each of these dominating sets, create a component of the original tree.
3. Because the dominating sets are exactly 3 apart, this subgraph, with some manipulation, will be saturated.

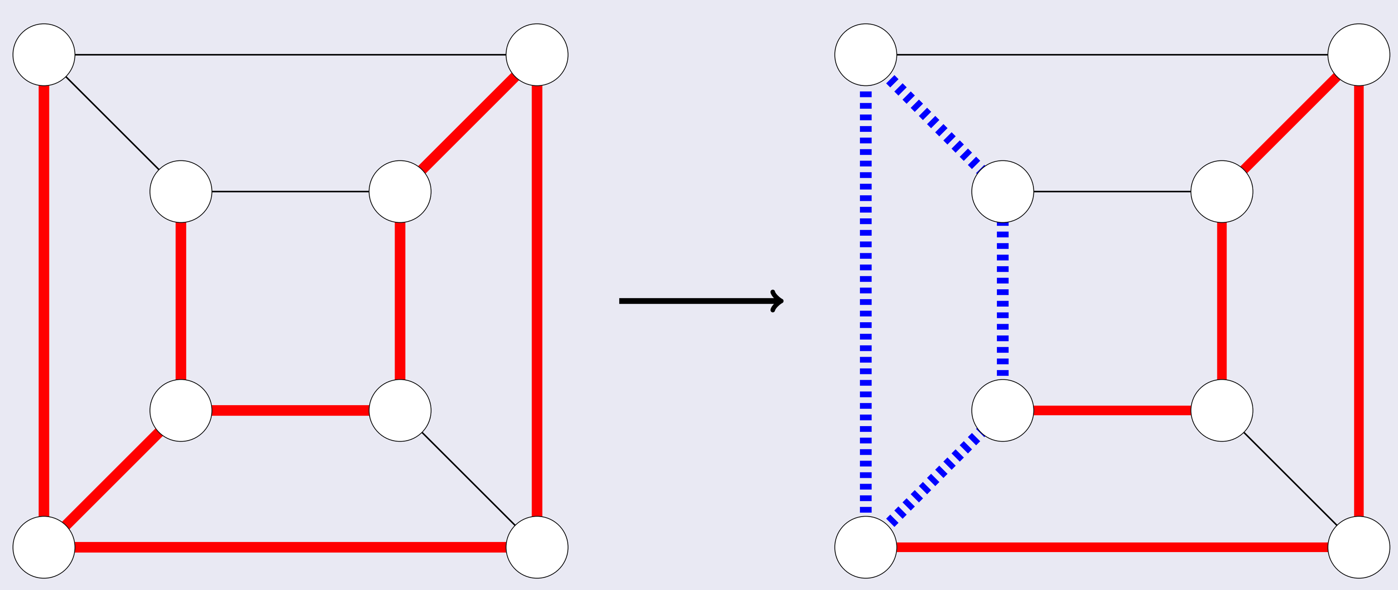
1 Introduction and Key Concepts

I studied a problem in extremal graph theory concerning the **saturation number** of graphs in the **hypercube**.

Definition

The **saturation number** $\mathrm{sat}(H, G)$ is the minimum number of edges in a subgraph H' of H that satisfies the following two properties:

- H' does not contain G
- Given any $e \in H \backslash H'$, $H' + e$ contains G.



Definition

The n-dimensional **hypercube** Q_n is the regular graph with vertex set $\{0, 1\}^n$ and edge set consisting of all pairs of vertices differing in exactly one coordinate.

2 The Problem

$\mathrm{sat}(Q_n, G)$ has not been very well-studied in the past; the only major result concerns $\mathrm{sat}(Q_n, Q_m)$.

Theorem (Morrison, Noel, Scott)

Given $m < n$,

$$(m-1) \cdot 2^n \leq \mathrm{sat}(Q_n, Q_m) < 72m^2 2^n.$$

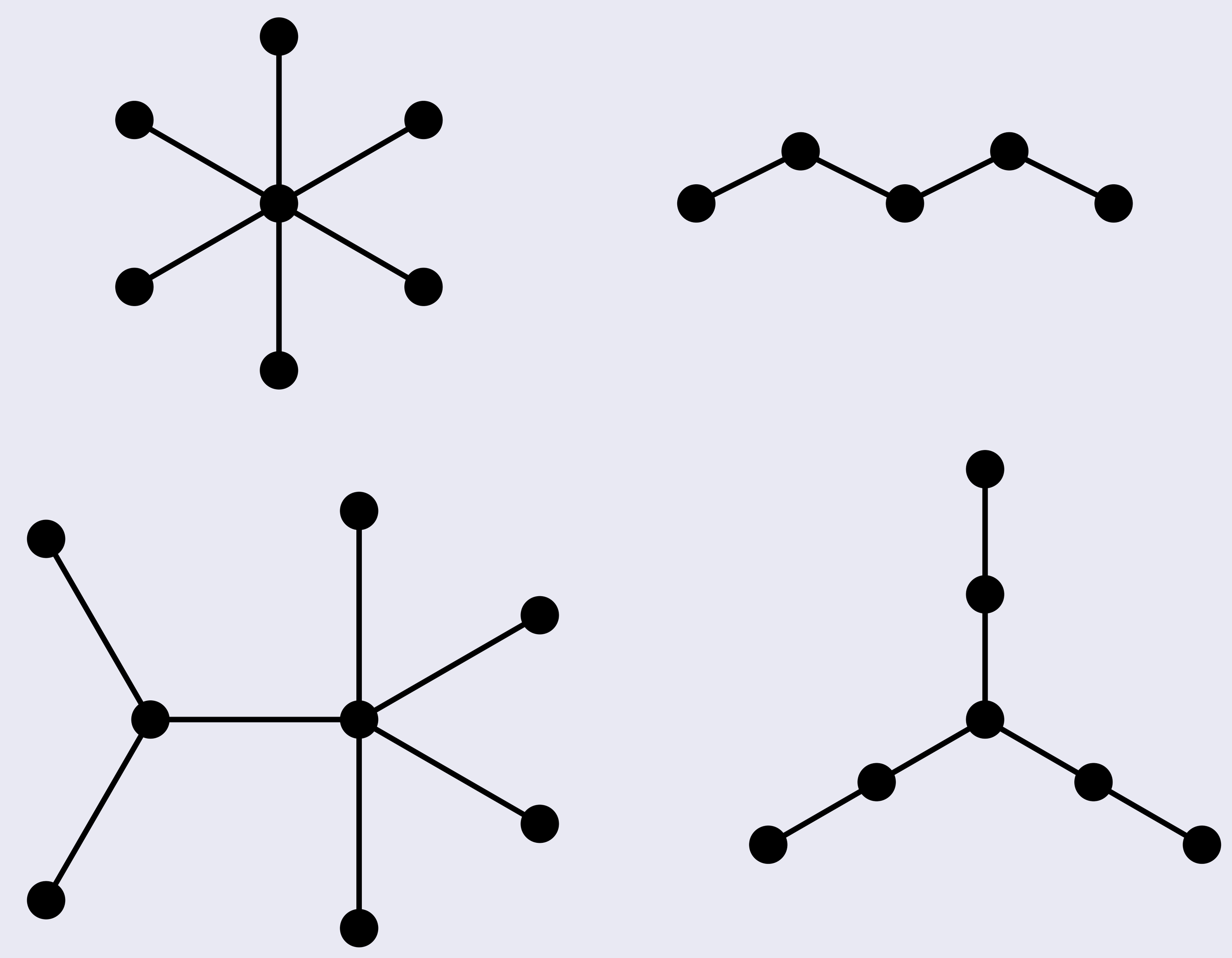
However, previously, **no results** have been found that bound the saturation number of trees (acyclic graphs).

Question

What is $\mathrm{sat}(Q_n, T)$ for a general tree T?

I answer this question the following types of trees:

- Paths of length k, denoted P_k
- Stars with k edges, denoted S_k
- Generalized stars, or trees consisting of a central node and k emanating paths of length m, denoted $GS_{k,m}$
- Caterpillars, or trees consisting of consecutive stars connected by their central nodes, denoted $S_{k_1 \times k_2 \times \ldots \times k_n}$



6 Significance of this Work

In a larger mathematical context, the major contributions of this work are:

1. The derivation of numerous new and decently tight bounds on the saturation number of trees.
2. Two novel methods for deriving such bounds based on disjoint subcubes and the Hamming code, which can hopefully be extended in the future to more subgraphs.

Given that this problem is relatively unstudied, I hope that my work serves as just the beginning in the study of this field. I wonder if the following conjecture is true:

Conjecture

Given a graph G, $sat(Q_n, G) = O(2^n)$.

7 Practical Applications

Saturation is particularly interesting in and relevant to theoretical computer science. In particular, in a parallel computing network, let

1. The processors be vertices of the hypercube.
2. The communication links be edges of the hypercube

Then, the saturation number is equivalent to the maximum number of links that must fail before some desired configuration of processors and links no longer exists. Thus, my work could potentially be useful in:

- Ensuring fault tolerance of parallel networks
- Design and implementation of efficient parallel architecture

8 Future Directions

There are many possible future directions for this research. In particular, I would like to:

- Find upper and lower bounds for all trees.
- Find exact values for the saturation number of some trees; the first such target would be stars and paths. I have made progress on this for S_2 (P_2) and S_3.
- Find bounds on $\mathrm{sat}(Q_n, C_{2k})$ for cycles of even length.
- Perhaps develop more methods to find tight bounds on $\mathrm{sat}(Q_n, G)$ for graphs G satisfying a certain property.
- Investigate bounds on the semi-saturation and weak saturation numbers of trees, both of which have slightly less strict requirements than saturation.

9 Acknowledgements

I would like to thank:

- Chiheon Kim, my mentor
- Dr. Tanya Khovanova
- Professor Pavel Etingof
- Dr. Slava Gerovitch
- The Research Science Institute
- The MIT-PRIMES Program
- Joint Mathematical Meetings