The Steiner distance problem for large vertex subsets in the hypercube

Éva Czabarka¹,², Josiah Reiswig∗³ and László Székely¹,²

¹University of South Carolina, Columbia, SC, USA
czabarka@math.sc.edu, szekely@math.sc.edu
²University of Johannesburg, South Africa
³Anderson University, SC, USA
jreiswig@andersonuniversity.edu

Abstract
We find the asymptotic behavior of the Steiner $k$-diameter of the $n$-cube if $k$ is large. Our main contribution is the lower bound, which utilizes the probabilistic method.

1 Introduction

For a connected graph $G$ of order at least 2 and $S \subseteq V(G)$, the Steiner distance $d(S)$ among the vertices of $S$ is the minimum size among all connected subgraphs whose vertex sets contain $S$. Necessarily, such a minimum subgraph must be a tree and such a tree is called a Steiner tree.

∗Corresponding author.

AMS (MOS) Subject Classifications: Primary 05C12; secondary 05C05, 05C35, 05C69

Key words and phrases: hypercube, Steiner distance, domination

Received: 19 July 2019
Accepted: 22 June 2021
The Steiner distance was introduced by G. Chartrand, O.R. Oellermann, S. Tian and H.B. Zou [2], and it has turned into a well-studied parameter of graphs. Tao Jiang, Zevi Miller, and Dan Pritikin [6] studied how large the Steiner distance of \( k \) vertices can be in the \( n \)-dimensional hypercube \( Q_n \) as \( n \to \infty \), while Zevi Miller and Dan Pritikin [5] gave near tight bounds for the Steiner distance of a layer, i.e., vertices with the same number of 1’s, in the \( n \)-dimensional hypercube \( Q_n \) as \( n \to \infty \). For a given \( 2 \leq k \leq n \), the Steiner \( k \)-diameter of the \( n \)-cube, \( \text{sdiam}_k(Q_n) \), is the maximum Steiner distance among all \( k \) subsets of \( V(Q_n) \).

In this note we give natural upper bounds for the Steiner distance of a large vertex set in the hypercube. It turns out that even the second order term in this estimate is close to tight. With these bounds, we determine \( \text{sdiam}_k(Q_n) \) asymptotically for large \( k \).

2 Upper bound

For the upper bound, we utilize connected dominating sets of \( Q_n \). A set \( S \subset V(Q_n) \) is a dominating set of \( Q_n \) if every vertex of \( Q_n \) is either an element of \( S \) or has a neighbor in \( S \). The minimum size of all dominating sets is called the domination number of \( Q_n \) and is denoted \( \gamma(Q_n) \). The connected domination number, denoted by \( \gamma_c(Q_n) \), is minimum size of all connected dominating sets.

In 1988, Kabatyanskii and Panchenko [4] showed

\[
\lim_{n \to \infty} \frac{\gamma(Q_n)}{2^n/n} = 1.
\]

In an upcoming paper, Griggs [3] utilizes this result to show that

\[
\lim_{n \to \infty} \frac{\gamma_c(Q_n)}{2^n/n} = 1.
\]

We use this last result to develop an upper bound for the Steiner diameter of subsets of \( V(Q_n) \).

**Lemma 1.** Suppose that \( S \subset V(Q_n) \). Then,

\[
d(S) \leq |S| + \frac{2^n}{n}(1 + o(1)).
\]
Proof. Begin with a spanning tree of a minimum connected dominating set of $Q_n$. Add edges as needed to connect each element of $S$ to this tree. The resulting subgraph spans $S$ and contains at most $|S| + \gamma_c(Q_n) - 1$ edges. Using [3], we then have that $d(S) \leq |S| + \frac{2^n}{n}(1 + o(1))$. \hfill $\square$

3 Lower bound

To bound the Steiner distance of large vertex subsets of $Q_n$ from below, we partition the vertices of the hypercube into two sets. Identifying each vertex of $Q_n$ into a binary string of length $n$, we let vertices with an even number of 1’s make up the set of even vertices and denote this set by $E_n$. Similarly, we let the vertices with an odd number of 1’s make up the set of odd vertices and denote this set by $O_n$. We refer to changing the value of the $i$th entry of a binary string $v = v_0 \cdots v_i \cdots v_n$ as “flipping” the $i$th entry of $v$. Given an entry $v_i$, we let $\bar{v}_i = 1 - v_i$. That is, $\bar{v}_i$ is the flipped value of $v_i$. For the proof of Theorem 2, we use probabilistic methods similar to those found in [1].

Theorem 2. If $S \subset E_n$ with $|S| \geq 2$, then

$$d(S) \geq |S| + \frac{|S|^2}{n2^n} - \frac{(n + 1)}{2}.$$ 

Proof. Suppose that $S \subset E_n$. Let $S'$ be a subset of the odd vertices which is the image of $S$ under some automorphism of $Q_n$. That is, $S \subset E_n$, $S' \subset O_n$, and $S' = \gamma(S)$ for some $\gamma \in \text{Aut}(Q_n)$. To show that such a subset $S'$ exists, consider the set of all vertices in $S$ with the first entry flipped. Since $S'$ is the image of $S$ under the automorphism $\gamma$, we have that $d(S) = d(S')$. Now suppose that $\lambda_1$ and $\lambda_2$ are automorphisms of $Q_n$ which preserve the parity of their inputs. Then, $\lambda_1(S) \subset E_n$ and $\lambda_2(S') \subset O_n$. Since $\lambda_1$ and $\lambda_2$ are automorphisms, we have that $d(\lambda_1(S)) = d(\lambda_2(S')) = d(S)$.

We now bound $d(\lambda_1(S) \cup \lambda_2(S'))$ above and below in terms of $|S|$ and $d(S)$. For the lower bound, note that $\lambda_1(S)$ and $\lambda_2(S')$ are disjoint. Hence, we have the naive bound

$$2|S| - 1 \leq d(\lambda_1(S) \cup \lambda_2(S')).$$ (1)

For the upper bound, suppose that $\lambda_1(T)$ and $\lambda_2(T')$ are Steiner trees of $\lambda_1(S)$ and $\lambda_2(S')$ in $Q_n$, respectively. Denote the respective edge sets
of the Steiner trees by \( E(\lambda_1(T)) \) and \( E(\lambda_2(T')) \). Using no more than \( n \) edges (Since \( \text{diam}(Q_n) = n \)), we may connect \( \lambda_1(T) \) and \( \lambda_2(T') \) to form a subgraph of \( Q_n \) which contains \( \lambda_1(S) \cup \lambda_2(S') \). Hence,
\[
d(\lambda_1(S) \cup \lambda_2(S')) \leq |E(\lambda_1(T)) \cup E(\lambda_2(T'))| + n. \tag{2}
\]

Linking inequalities (1) and (2) together and applying the principle of inclusion and exclusion, we have
\[
2|S| - 1 \leq |E(\lambda_1(T)) \cup E(\lambda_2(T'))| + n
= |E(\lambda_1(T))| + |E(\lambda_2(T'))| - |E(\lambda_1(T)) \cap E(\lambda_2(T'))| + n
= 2d(S) - |E(\lambda_1(T)) \cap E(\lambda_2(T'))| + n,
\]
which implies that
\[
2d(S) - |E(\lambda_1(T)) \cap E(\lambda_2(T'))| \geq 2|S| - (n + 1). \tag{3}
\]

Let \( \Gamma = \langle \alpha, \beta_{i,j} : 1 \leq 0 < j \leq n - 1 \rangle \) be the subgroup of the group of automorphisms of \( Q_n \) generated by the automorphisms
\[
\alpha : v_0v_1 \cdots v_{n-1} \mapsto v_1 \cdots v_{n-1}v_0
\]
\[
\beta_{i,j} : v_0v_1 \cdots v_i \cdots v_j \cdots v_{n-1} \mapsto v_0v_1 \cdots \hat{v}_i \cdots \hat{v}_j \cdots v_{n-1}.
\]

In words, \( \alpha \) shifts each entry of its input to the left by 1 (modulo \( n \)), while \( \beta_{i,j} \) flips only the values of the \( i \)th and \( j \)th entries of its input. Note that each element of \( \Gamma \) preserves the parity of its input. We now verify the following claim:

**Claim:** For any two edges \( e_1, e_2 \in E(Q_n) \), there exists a unique element of \( \lambda \in \Gamma \) such that \( \lambda(e_1) = e_2 \).

**Proof.** Suppose that \( e_1 = ab \) and \( e_2 = uv \) where \( a \) and \( u \) are even vertices while \( b \) and \( v \) are odd vertices. Without loss of generality, we may assume that \( a = 0 \), the vertex of all zeros. This implies that the string \( b \) contains a single 1. We shall first prove existence of an automorphism \( \lambda \in \Gamma \) mapping \( e_1 \) to \( e_2 \).

Since \( u \in \mathcal{E}_n \), using a composition of automorphisms of the form \( \beta_{i,j} \) we may map \( uv \) to \( 0\hat{v} \), where \( \hat{v} \) has a single 1. Then, using some power of the automorphism \( \alpha \), we may then map the edge \( 0\hat{v} \) to the edge \( 0b = e_1 \). Let \( \lambda \) be the composition of these automorphisms in \( \Gamma \).

To show that this automorphism is unique, we show that \( |\Gamma| = n2^{n-1} \). Since \( \alpha \circ \beta_{ij} = \beta_{i-1,j-1} \circ \alpha \) (where the indexes are taken modulo \( n \)), any
\( \lambda \in \Gamma \) can be described as first applying an appropriate power of \( \alpha \) and then flipping an even number of digits. As we have \( n \) choices for the power of \( \alpha \) and \( 2^{n-1} \) choices for the subset of digits we flip, we conclude that \( |\Gamma| = n2^{n-1} \).

Since \( Q_n \) has \( n2^{n-1} \) edges, any \( \lambda \in \Gamma \) maps the edge \( 0b \) to an edge in such a way that \( 0 \) is mapped the edge’s vertex in \( E_n \), and all edges of \( Q_n \) will be the image of \( 0b \) under some \( \lambda \in \Gamma \), the claim follows.

We now consider the experiment of selecting elements \( \lambda_1, \lambda_2 \in \Gamma \) independently with uniform probability, and applying them to \( T \) and \( T' \), respectively. Consider the random variable \( X = |E(\lambda_1(T)) \cap E(\lambda_2(T'))| \). For the expected value of \( X \), \( \mathbb{E}(X) \), we have that

\[
\max_{\lambda_1, \lambda_2} \{|E(\lambda_1(T)) \cap E(\lambda_2(T'))|\} \geq \mathbb{E}(X).
\]

Using our claim, we observe that

\[
\mathbb{E}(X) = \sum_{f \in E(Q_n)} P[(f \in E(\lambda_1(T))) \text{ and } (f \in E(\lambda_2(T')))]
\]

\[
= \sum_{f \in E(Q_n)} \frac{|E(\lambda_1(T))|}{n2^{n-1}} \cdot \frac{|E(\lambda_2(T'))|}{n2^{n-1}}
\]

\[
= \frac{|E(\lambda_1(T))|^2}{n2^{n-1}}
\]

\[
= \frac{d(S)^2}{n2^{n-1}},
\]

which implies

\[
\max_{\lambda_1, \lambda_2} \{|E(\lambda_1(T)) \cap E(\lambda_2(T'))|\} \geq \frac{d(S)^2}{n2^{n-1}}.
\]

Using \( \lambda_1 \) and \( \lambda_2 \) which achieve this maximum and applying inequality (3), we see that

\[
2d(S) - \frac{d(S)^2}{n2^{n-1}} \geq 2|S| - (n + 1).
\]

Using the above inequality, we now bound \( d(S) \) from below. Since \( |S| \geq 2 \) and \( S \subset E_n \), we have that \( n \geq 2 \). Hence, \( d(S) = |S| + x \) for some \( x \geq 0 \).
So,

\[ 2(|S| + x) - \frac{(|S| + x)^2}{n^{2n-1}} \geq 2|S| - (n + 1) \]

\[ 2|S| + 2x - \frac{|S|^2 + 2|S|x + x^2}{n^{2n-1}} \geq 2|S| - (n + 1) \]

\[ 2x - \frac{2|S|x}{n^{2n-1}} + 2|S| - \frac{|S|^2 + x^2}{n^{2n-1}} \geq 2|S| - (n + 1) \]

\[ 2x \left(1 - \frac{|S|}{n^{2n-1}}\right) \geq \frac{|S|^2 + x^2}{n^{2n-1}} - (n + 1) \]

\[ x \geq \frac{|S|^2}{n^{2n}} - \frac{(n + 1)}{2}, \]

and the result is proven. \( \square \)

**Remark:** In the above theorem, we assumed \(|S| \geq 2\). If \(|S| = 1\), we have that \(d(S) = 0\).

With these results in hand, we can determine the asymptotic growth of \(sdiam_k(Q_n)\) for large \(k\). In particular, we can determine the first and second order terms if \(k = \Omega(2^n)\), while we can determine the first order term if \(2^n/n = o(k)\).

**Corollary 3.** If \(k = k(n)\) is regarded as a function \(n\), then

1. if \(k = \Omega(2^n)\), then \(sdiam_k(Q_n) = k + \Theta(2^n/n)\), and

2. if \(2^n/n = o(k)\), then \(\lim_{n \to \infty} \frac{sdiam_k(Q_n)}{k} = 1\).

**Proof.** If \(k \leq 2^{n-1}\), let \(S \subseteq V(Q_n)\) be a subset of the even vertices of size \(k\). If \(k > 2^{n-1}\), let \(S\) contain all even vertices and choose the remaining odd vertices randomly. Applying the bounds determined in Lemma 1 and Theorem 2, we see that

\[ k + \frac{k^2}{n^{2n}} - \frac{n + 1}{2} \leq d(S) \leq sdiam_k(Q_n) \leq k + \frac{2^n}{n}(1 + o(1)). \]

If \(k = \Omega(2^n)\), then \(sdiam_k(Q_n)\) is bounded above and below by \(k + \Theta(2^n/n)\). More precisely, we have that \(k + c_1(2^n/n) \leq sdiam_k(Q_n) \leq k + c_2(2^n/n)\) for some positive constants \(c_1\) and \(c_2\). On the other hand, if only \(2^n/n = o(k)\), we have \(sdiam_k(Q_n) = k(1 + o(1))\), giving \(\lim_{n \to \infty} \frac{sdiam_k(Q_n)}{k} = 1\). \( \square \)
Acknowledgement

The last two authors were supported in part by the National Science Foundation contract DMS-1600811.

References

[1] N. Alon and J. H. Spencer, The Probabilistic Method (Second Edition), John Wiley and Sons, 2000.

[2] G. Chartrand, O.R. Oellermann, S. Tian and H.B. Zou, Steiner distance in graphs, Časopis Pest. Mat., 114 (1989), 399–410.

[3] J.R. Griggs, Spanning trees and domination in hypercubes, to appear in Integers, https://arxiv.org/abs/1905.13292

[4] G.A. Kabatyanskii and V.I. Panchenko, Unit sphere packings and coverings of the Hamming space, Problems of Inform. Transm., 24:4 (1988), 261–272.

[5] Z. Miller and D. Pritikin, Applying a result of Frank and Rödl to the construction of Steiner trees in the hypercube, Discrete Math., 131 (1994), 183–194.

[6] T. Jiang, Z. Miller, and D. Pritikin, Near optimal bounds for Steiner trees in the hypercube, SIAM J. Comp., 40(5) (2011), 1340–1360.