Tight Approximation Algorithms for Bichromatic Graph Diameter and Related Problems

Mina Dalirrooyfard
MIT, Cambridge, MA, USA
minad@mit.edu

Virginia Vassilevska Williams
MIT, Cambridge, MA, USA
virgi@mit.edu

Nikhil Vyas
MIT, Cambridge, MA, USA
nikhilv@mit.edu

Nicole Wein
MIT, Cambridge, MA, USA
nwein@mit.edu

Abstract

Some of the most fundamental and well-studied graph parameters are the Diameter (the largest shortest paths distance) and Radius (the smallest distance for which a "center" node can reach all other nodes). The natural and important ST-variant considers two subsets S and T of the vertex set and lets the ST-diameter be the maximum distance between a node in S and a node in T, and the ST-radius be the minimum distance for a node of S to reach all nodes of T. The *bichromatic* variant is the special case in which S and T partition the vertex set.

In this paper we present a comprehensive study of the *approximability* of ST and Bichromatic Diameter, Radius, and Eccentricities, and variants, in graphs with and without directions and weights. We give the first nontrivial approximation algorithms for most of these problems, including time/accuracy trade-off upper and lower bounds. We show that nearly *all* of our obtained bounds are tight under the Strong Exponential Time Hypothesis (SETH), or the related Hitting Set Hypothesis.

For instance, for Bichromatic Diameter in undirected weighted graphs with m edges, we present an $\tilde{O}(m^{3/2})$ time [1] 5/3-approximation algorithm, and show that under SETH, neither the running time, nor the approximation factor can be significantly improved while keeping the other unchanged.

2012 ACM Subject Classification Theory of computation → Graph algorithms analysis; Theory of computation → Approximation algorithms analysis; Theory of computation → Problems, reductions and completeness; Theory of computation → Shortest paths

Keywords and phrases approximation algorithms, fine-grained complexity, diameter, radius, eccentricities

Digital Object Identifier 10.4230/LIPIcs.ICALP.2019.47

Category Track A: Algorithms, Complexity and Games

Related Version `https://arxiv.org/abs/1904.11601`

Funding *Mina Dalirrooyfard*: Supported by an Akamai Presidential Fellowship and NSF Grant CCF-6936484.
Virginia Vassilevska Williams: Supported by an NSF CAREER Award, NSF Grants CCF-1417238, CCF-1528078 and CCF-1514339, a BSF Grant BSF:2012338 and a Sloan Research Fellowship.
Nikhil Vyas: Supported by an Akamai Presidential Fellowship and NSF Grant CCF-1552651.
Nicole Wein: Supported by an NSF Graduate Fellowship and NSF Grant CCF-1514339.

Acknowledgements The authors would like to thank Arturs Backurs for discussions during the early stages of this work.

[1] $\tilde{O}$ notation hides polylogarithmic factors.

46th International Colloquium on Automata, Languages, and Programming (ICALP 2019).
Editors: Christel Baier, Ioannis Chatzigiannakis, Paola Flocchini, and Stefano Leonardi;
Article No. 47; pp. 47:1–47:15
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Introduction

A fundamental and very well studied problem in algorithms is the Diameter of a graph, where the output is the largest (shortest path) distance over all pairs of vertices. Over the years many different algorithms have been developed for the problem, both in theory (e.g. [3, 21, 24, 8, 4]) and in practice (e.g. [10, 25, 20]).

A very natural variant is the so called ST-Diameter problem [4]: given a graph and two subsets S and T of its vertex set, determine the largest distance between a vertex of S and a vertex of T. In the Subset version of ST-Diameter, we have $S = T$. Bichromatic Diameter is the version of ST-Diameter for which S and T partition the vertex set. Besides Diameter, the Radius (the smallest distance for which a "center" node can reach all other nodes) and Eccentricities (the largest distance out of every vertex) problems are also very well studied, and analogous ST, Subset, and Bichromatic versions are easy to define.

All of these parameters are simple to compute by computing all pairwise distances in the graph, i.e. by solving All-Pairs Shortest Paths (APSP). In sparse n-node graphs, where the number of edges m is $\tilde{O}(n)$, APSP still needs $\Omega(n^2)$ time, as this is the size of the output, whereas it is apriori unclear whether this much time is needed for computing the Diameter, Radius and Eccentricities or their ST and bichromatic variants, as the output is small.

A related extremely well-studied problem in computational geometry is Bichromatic Diameter on point sets (commonly known as Bichromatic Farthest Pair), where one seeks to determine the farthest pair of points in a given set of points in space (see e.g. [29, 13, 28, 2, 17]). Another related problem is the Subset version of spanners (e.g. [19, 11]), as well as the ST version of spanners (e.g. [9, 18]). Furthermore, the ST, Subset, and Bichromatic versions of many problems have been of great interest; for instance Steiner Tree, Subset TSP, and a number of problems in computational geometry such as Bichromatic Matching (e.g. [16]) and Bichromatic Line Segment Intersection (e.g. [7]).

There are several known approximation algorithms for the standard version of Diameter, most of which have been developed in the last 6 years. Trivially, running Dijkstra's algorithm from an arbitrary vertex gives a simple $\tilde{O}(m)$ time 2-approximation algorithm for directed and weighted graphs. Non-trivial algorithms achieve an improved approximation factor with an increased runtime: Building on Aingworth et al. [3], Roditty and Vassilevska W. [24] showed for instance that an "almost" 1.5 approximation for Diameter can be computed in $\tilde{O}(m\sqrt{n})$ time in m-edge n-vertex directed weighted graphs – the approximation factor is 1.5 if the Diameter is divisible by 3, and there is a slight additive error otherwise. Chechik et al. [8] gave a true 1.5 approximation at the expense of increasing the runtime to $\tilde{O}(mn^{2/3})$, and Cairo, Grossi and Rizzi [5] generalized the approach giving an $\tilde{O}(mn^{1/(k+1)})$ time, "almost" $2 - 1/2^k$ approximation algorithm for all $k \geq 1$ which works only in undirected graphs.

In STOC'18, Backurs et al. [4] gave the first non-trivial approximation algorithms for ST-Diameter: an $\tilde{O}(m^{3/2})$ time 2-approximation and an $\tilde{O}(m)$ time 3-approximation. They also showed that these algorithms cannot be improved significantly, unless the Strong Exponential Time Hypothesis (SETH) fails. Backurs et al. did not provide algorithms for ST-Eccentricities or ST-Radius, and they did not study the natural Subset and Bichromatic versions. They also only focused on undirected graphs.

We study the following natural and fundamental questions:

How well can ST-Eccentricities and ST-Radius be approximated? Are any interesting approximation algorithms possible for directed graphs for any of the ST-variants? Does the approximability of the problems change when one turns to the Subset versions in which $S = T$, or the Bichromatic versions in which S and T are required to partition the vertex set?

1.1 Our Results

We present a comprehensive study of the approximability of the ST, Subset and Bichromatic variants of the Diameter, Radius and Eccentricities problems in graphs, both with and without directions and weights. We obtain the first non-trivial approximation algorithms for most of these problems, including time/accuracy trade-off upper and lower bounds. We show that nearly *all* of our approximation algorithms are tight under SETH (or under the related Hitting Set Hypothesis for Radius). Additionally, we study a parameterized version of these problems.

Our results are summarized in Tables 1-4.

Table 1 Bichromatic undirected results. All of our parameterized algorithms and near-linear time algorithms, except for directed Subset Radius and Eccentricities, are deterministic. The rest are randomized and work with high probability[2]. Our lower bounds for Diameter and Eccentricities are under SETH and our lower bounds for Radius are under the Hitting Set (HS) Hypothesis, defined later. All of our lower bounds hold even for unweighted graphs. The trade-off lower bounds in terms of k hold for any integer $k \geq 2$. δ is any constant > 0. B and B' are parameters defined in our parameterized algorithms. The lower bound constructions for the parameterized algorithms have $|B| = \tilde{O}(1)$.
* Multiplicative approximation factor is tight, but not runtime.

	Upper Bounds			Lower Bounds	
Problem	Runtime	Approx.	Comments	Runtime	Approx.
Diameter	$O(m + n\log n)$	almost 2	unweighted, **tight**	$m^{1+o(1)}$	$2-\delta$
	$\tilde{O}(m\sqrt{n})$	almost 5/3	unweighted, nearly **tight**	$m^{\frac{k}{k-1}-o(1)}$	$2-\frac{1}{2k-1}-\delta$
	$\tilde{O}(m^{3/2})$	5/3	weighted, **tight**	"	"
	$O(m\|B\|)$	almost 3/2	unweighted, **tight***	$m^{2-o(1)}$	$3/2-\delta$
Radius	$O(m + n\log n)$	almost 2	unweighted		
	$\tilde{O}(m\sqrt{n})$	almost 5/3	unweighted, nearly **tight***	$m^{2-o(1)}$	$5/3-\delta$
	$\tilde{O}(m^{3/2})$	5/3	weighted, **tight***	"	"
	$O(m\|B\|)$	almost 3/2	unweighted, **tight***	$m^{2-o(1)}$	$3/2-\delta$
Eccentricities	$O(m + n\log n)$	3	weighted, **tight**	$m^{1+o(1)}$	$3-\delta$
	$\tilde{O}(m\sqrt{n})$	almost 2	unweighted, nearly **tight**	$m^{\frac{k}{k-1}-o(1)}$	$3-2/k-\delta$
	$\tilde{O}(m^{3/2})$	2	weighted, **tight**	"	"
	$O(m\|B\|)$	almost 5/3	unweighted, **tight***	$m^{2-o(1)}$	$5/3-\delta$

Table 2 Bichromatic directed results. See caption of Table 1.

	Upper Bounds			Lower Bounds	
Problem	Runtime	Approx.	Comments	Runtime	Approx.
Diameter	$\tilde{O}(m^{3/2})$	2	weighted, **tight***	$m^{2-o(1)}$	$2-\delta$
	$O(m\|B'\|)$	almost 3/2	unweighted, **tight***	$m^{2-o(1)}$	$3/2-\delta$
Radius	N/A	N/A	weighted, **tight**	$m^{2-o(1)}$	any finite
Eccentricities	N/A	N/A	weighted, **tight**	$m^{2-o(1)}$	any finite

All our algorithms in m-edge, n-node graphs, run in $\tilde{O}(m^{3/2})$ time or in $\tilde{O}(m\sqrt{n})$ time when a small additive error is allowed. For sparse graphs the $m^{3/2}$ runtime beats the fastest

[2] *with high probability* means with probability at least $1 - 1/n^c$ for all constants c.

Table 3 ST undirected results. See caption of Table 1.

	Upper Bounds			Lower Bounds	
Problem	Runtime	Approx.	Comments	Runtime	Approx.
Diameter[4]	$O(m + n \log n)$	3	weighted, **tight**	$m^{1+o(1)}$	$3 - \delta$
	$\tilde{O}(m\sqrt{n})$	almost 2	unweighted, nearly **tight**	$m^{\frac{k}{k-1}-o(1)}$	$3 - 2/k - \delta$
	$\tilde{O}(m^{3/2})$	2	weighted, **tight**	"	"
Radius	$O(m + n \log n)$	3	weighted		
	$\tilde{O}(m\sqrt{n})$	almost 2	unweighted, nearly **tight***	$m^{2-o(1)}$	$2 - \delta$
	$\tilde{O}(m^{3/2})$	2	weighted, **tight***	"	"
Eccentricities	$O(m + n \log n)$	3	weighted, **tight**	$m^{1+o(1)}$	$3 - \delta$ [4]
	$\tilde{O}(m\sqrt{n})$	almost 2	unweighted, nearly **tight**	$m^{\frac{k}{k-1}-o(1)}$	$3 - 2/k - \epsilon$ [4]
	$\tilde{O}(m^{3/2})$	2	weighted, **tight**	"	"

Table 4 Subset results. See caption of Table 1.

	Upper Bounds			Lower Bounds	
Problem	Runtime	Approx.	Comments	Runtime	Approx.
Diameter	$\tilde{O}(m)$	2	weighted, directed, **tight**	$m^{2-o(1)}$	$2 - \delta$
Radius	$\tilde{O}(m)$	2	weighted, undirected, **tight**	$m^{2-o(1)}$	$2 - \delta$
	$\tilde{O}(m/\delta)$	$2 + \delta$	weighted, directed, **tight** up to an additive δ	"	"
Eccentricities	$\tilde{O}(m/\delta)$	$2 + \delta$	weighted, directed, **tight** up to an additive δ	$m^{2-o(1)}$	$2 - \delta$

APSP algorithms [6, 23, 22] as they run in $\tilde{O}(mn)$ time. The $m\sqrt{n}$ time of the algorithms that allow small additive error beat the APSP algorithms for every graph sparsity.

Bichromatic Diameter and Radius

Our first contribution is an algorithm with the same running time as the 2-approximation ST-Diameter algorithm of [4], achieving a better, 5/3 approximation for Bichromatic Diameter. In other words, when S and T partition the vertex set of the graph, ST-Diameter can be approximated much better! Moreover, we show that under SETH, neither the runtime nor the approximation factor of our algorithm can be improved. The result is summarized in Theorem 1 below, and proven in Theorems 11 and 12.

▶ **Theorem 1.** *There is a randomized $\tilde{O}(m^{3/2})$ time algorithm, that given an undirected graph $G = (V, E)$ with nonnegative integer edge weights and $S \subseteq V, T = V \setminus S$, can output an estimate D' such that $3D_{ST}/5 \leq D' \leq D_{ST}$ with high probability, where D_{ST} is the ST-Diameter of G.*

Moreover, if there is an $O(m^{3/2-\varepsilon})$ time 5/3-approximation algorithm for some $\varepsilon > 0$, or if there is an $O(m^{2-\varepsilon})$ time $(5/3 - \varepsilon)$-approximation algorithm for the problem, then SETH is false.

We also obtain an $\tilde{O}(m\sqrt{n})$ time algorithm that achieves an "almost" 5/3-approximation: the guarantee for unweighted graphs is $3D_{ST}/5 - 6/5 \leq D' \leq D_{ST}$. We also obtain a near-linear time algorithm for weighted graphs that returns an estimate D' with $D_{ST}/2 - W/2 \leq D' \leq D_{ST}$ where W is the minimum weight of a $S \times T$ edge. Using our general theorem 12, we get that this result is also essentially tight, as a $(2 - \varepsilon)$-approximation for $\varepsilon > 0$ running in near-linear time would refute SETH.

To obtain our improvements for Bichromatic Diameter over the known ST-Diameter algorithms, we crucially exploit the basic fact that as S, T partition V any path that starts from a vertex $s \in S$ and ends in a vertex $t \in T$ must cross a (u, v) edge such that $u \in S, v \in T$. While this fact is clear, it not at all obvious how one might try to exploit it.

We explain our technique in more detail for the bichromatic diameter problem, and similar ideas are used for our algorithms for the other problems. Let $s^* \in S$ and $t^* \in T$ be end-points of an ST-Diameter path. Similarly to prior Diameter algorithms, our goal is to run Dijkstra's algorithm from some $s \in S$ which is close to s^*, and hence far from t^*, or from some $t \in T$ which is close to t^* and hence far from s^* (by the triangle inequality). Our 5/3-approximation algorithms are a delicate combination of two themes: (1) randomly sample nodes in S and nodes in T – similarly to prior works, the sampling works well if there are many nodes of S that are close to s^*, or if there are many nodes of T that are close to t^*. If (1) is not good enough, in theme (2) we show that we can find a node $w \in S$ close to t^* for which we can "catch" an $S \times T$ edge (s, t) on the shortest $w \to t^*$ path, such that t is close to t^*. Theme (2) is our new contribution. Because of theme (2), our algorithms are more complicated than the ST-Diameter algorithms, but run in asymptotically the same time, and achieve a better approximation guarantee. In order to better separate the ideas in our algorithms, we explain them in several steps, where Theme (1) can be seen in the first steps and Theme (2) appears towards the last steps.

Following a similar approach to our Bichromatic Diameter algorithms, we develop similar algorithms for Bichromatic Radius. First, we give a simple near-linear time almost 2-approximation algorithm, and then we adapt the 5/3-approximation for Bichromatic Diameter to also give a 5/3-approximation for Bichromatic Radius. Moreover, we show that any better approximation factor requires essentially quadratic time, under the Hitting Set (HS) Hypothesis of [1] (see also [14]).

▶ **Theorem 2.** *There is a randomized $\tilde{O}(m^{3/2})$ time algorithm, that given an undirected graph $G = (V, E)$ with nonnegative integer edge weights and $S \subseteq V, T = V \setminus S$, can output an estimate R' such that $R_{ST} \leq R' \leq 5R_{ST}/3$ with high probability, where R_{ST} is the ST-Radius of G. Moreover, if there is a $5/3 - \varepsilon$ approximation algorithm running in $O(m^{2-\delta})$ time for any $\varepsilon, \delta > 0$, then the HS Hypothesis is false.*

Similarly to the Bichromatic Diameter algorithm, if one is satisfied with a slight additive error, one can improve the runtime to $\tilde{O}(m\sqrt{n})$.

ST-Eccentricities and ST-Radius

Prior work only considered ST-Diameter but did not consider the more general ST- Eccentricities problem in which one wants to approximate for every $s \in S$, $\varepsilon_{ST}(s) := \max_{t \in T} d(s, t)$.

Here we show that one can achieve exactly the same approximation factors for ST-Eccentricities as for ST-Diameter. Since any conditional lower bound for ST-Diameter also applies for the ST-Eccentricities problem, the algorithms we obtain are conditionally optimal, similarly to the ST-Diameter algorithms in [4]. Interestingly, we show that the same conditional lower bounds apply for Bichromatic Eccentricities (see the full version [12]), and therefore our ST-Eccentricities algorithms are optimal even for the Bichromatic case.

▶ **Theorem 3.** *There is a randomized $\tilde{O}(m^{3/2})$ time algorithm, that given an undirected graph $G = (V, E)$ with nonnegative integer edge weights and $S, T \subseteq V$, can output for every $s \in S$, an estimate $\varepsilon'(s)$ such that $\varepsilon_{ST}(s)/2 \leq \varepsilon'(s) \leq \varepsilon_{ST}(s)$ with high probability. Moreover, if there is a $2 - \varepsilon$ approximation algorithm running in $O(m^{2-\delta})$ time for any $\varepsilon, \delta > 0$ or a 2-approximation algorithm running in $O(m^{3/2-\varepsilon})$ time for $\varepsilon > 0$, even for the Bichromatic case when $T = V \setminus S$, then SETH is false.*

Again, as before, one can improve the runtime to $\tilde{O}(m\sqrt{n})$ with a slight additive error, and there is a simple near-linear time 3-approximation algorithm which is tight under SETH, similar to the one in [4] for ST-Diameter. A simple argument shows that these algorithms imply algorithms with the same running time and approximation factor for ST-Radius.

Bichromatic and ST Problems in Directed Graphs

Using simple reductions we first show that there can be no $O(m^{2-\varepsilon})$ time (for $\varepsilon > 0$) algorithms that achieve any finite approximation for ST-Diameter or ST-Eccentricities (under SETH), or ST-Radius (under HS). Interestingly, the same holds for Bichromatic Eccentricities (under SETH, see the full version [12]) and Bichromatic Radius (under HS, see [12]), but not Bichromatic Diameter! Surprisingly, unlike those two problems, Bichromatic Diameter does admit a finite, in fact 2-approximation algorithm running in subquadratic time, and this algorithm is conditionally optimal:

▶ **Theorem 4.** *There is a randomized $\tilde{O}(m^{3/2})$ time algorithm, that given a directed graph $G = (V, E)$ with nonnegative integer edge weights and $S \subseteq V, T = V \setminus S$, can output an estimate D' such that $D_{ST}/2 \leq D' \leq D_{ST}$ with high probability, where D_{ST} is the ST-Diameter of G.*

Moreover, if there is an $O(m^{2-\varepsilon})$ time $2-\delta$-approximation algorithm for the problem for some $\varepsilon, \delta > 0$, then SETH is false.

The previously known techniques for approximating Diameter in directed graphs fail here. The main issue is that the prior techniques were general enough that they also gave algorithms for Eccentricities and Radius as a byproduct. In the Bichromatic case, however, there is a genuine difference between Diameter and Radius, as we noted above, and new techniques are needed. Here again it turns out that combining theme (2) with a delicate argument is sufficient to get conditionally tight algorithms under SETH.

Subset Versions

Recall that Subset Diameter, Radius, and Eccentricities are the versions of the corresponding ST problems with the constraint that $S = T$. Interestingly, Subset Diameter, Radius, and Eccentricities all exhibit the same sharp threshold behavior. For all three problems, there are near-linear time algorithms that achieve a 2 (or almost 2) approximation, as well as conditional lower bounds that show that there is no $2 - \delta$ approximation in $m^{2-o(1)}$ time.

Parameterized Algorithms

We consider the Bichromatic Diameter, Radius, and Eccentricities problems parameterized by the size of the *boundary* between the S and T sets. If S' is the set of vertices in S that have a neighbor in T, and T' is the set of vertices in T that have a neighbor in S, then the boundary B is whichever of S' or T' is smaller in size. Our lower bound constructions already have small boundary so they rule out algorithms even for graphs with small boundary. However, interestingly we obtain near-linear time algorithms for graphs with small boundary that achieve *better* multiplicative approximation factors than the optimal non-parameterized algorithms. This is not a contradiction because our parameterized algorithms have a constant additive error, while the apparently contradictory lower bounds do not tolerate additive error.

2 Preliminaries

Given a graph $G = (V, E)$ (directed or undirected, weighted or unweighted), let $d(u, v)$ denote the distance from $u \in V$ to $v \in V$. For a subset $X \subseteq V$ and $v \in V$, define $d(v, X) := \min_{x \in X} d(v, x)$. Similarly $d(X, v) := \min_{x \in X} d(x, v)$.

Unless otherwise stated, m denotes the number of edges and n the number of vertices of the underlying graph. Without loss of generality, we can assume that all undirected graphs are connected, and all directed graphs are weakly connected, so that $m \geq n - 1$.

The *Eccentricity* $\varepsilon(v)$ of a vertex $v \in V$ is $\max_{u \in V} d(v, u)$. The *Diameter* $D(G)$ of G is $\max_{v \in V} \varepsilon(v)$, and the *Radius* $R(G)$ of G is $\min_{v \in V} \varepsilon(v)$.

Given $S, T \subseteq V$, we define analogous parameters as follows. The ST-Eccentricity $\varepsilon_{ST}(v)$ of $v \in S$ is $\max_{u \in T} d(v, u)$. The ST-Diameter $D_{ST}(G)$ is $\max_{v \in S} \varepsilon_{ST}(v)$, and the ST-Radius $R_{ST}(G)$ is $\min_{v \in S} \varepsilon_{ST}(v)$.

The above parameters are called *Bichromatic* Eccentricities, Diameter, and Radius if S and T form a partition of V, i.e. $T = V \setminus S$.

The above parameters are called *Subset* Eccentricities, Diameter, and Radius if $S = T$ and are notated with subscript S instead of ST.

2.1 Preliminaries for algorithms

▶ **Lemma 5.** *Let $G = (V, E)$ be a (possibly directed and weighted graph) and let $W \subseteq V$. Let $g \geq \Omega(\ln n)$ be an integer. Let $S \subseteq W$ be a random subset of $c(|W|/g) \ln n$ vertices for some constant $c > 1$. For every $v \in V$, let $W(v)$ be the set of vertices $x \in W$ for which $d(v, x) < d(v, S)$. Then with probability at least $1 - 1/n^{c-1}$, for every $v \in V$, $|W(v)| \leq g$, and moreover, if one takes the closest g vertices of W to v, they will contain $W(v)$.*

Proof. For each $v \in V$, imagine sorting the nodes $x \in W$ according to $d(v, x)$. Define Q_v to be the first g nodes in this sorted order - those are the nodes of W closest to v (in the $v \to x$ direction).

We pick S randomly by selecting each vertex of W with probability $(c \ln n)/g$. The probability that a particular $q \in Q_v$ is not in S is $1 - (c \ln n)/g$, and the probability that no $q \in Q_v$ is in S is $(1 - (c \ln n)/g)^g \leq 1/n^c$. By a union bound, with probability at least $1 - 1/n^{c-1}$, for every $v \in V$, we have that $Q_v \cap S \neq \emptyset$.

Now, for each particular v, say that $w(v)$ is a node in $Q_v \cap S$. Since all nodes $x \in W$ with $d(v, x) < d(v, w(v))$ must be in Q_v, and since $d(v, w(v)) \geq d(v, S)$, we must have that $W(v) \subseteq Q_v$. Hence, with probability at least $1 - 1/n^{c-1}$, for every $v \in V$, $|W(v)| \leq g$ and $W(v) \subseteq Q_v$. ◀

▶ **Lemma 6.** *Let $G = (V, E)$ be a (possibly directed and weighted) graph. Let $M, W \subseteq V$ and let $S \subseteq W$ be a random subset of $c(n/g) \ln n$ vertices for some large enough constant c and some integer $g \geq 1$.*

Then, for any $D > 0$ and for any $w \in M$ with $d(w, S) > D$, if one takes the closest g vertices of W to w, they will contain all nodes of W at distance $< D$ from w, with high probability.

Proof. Let Q be the closest g vertices of W to w. By Lemma 5, with high probability Q contains all nodes of W at distance $< d(w, S)$ from w, and hence Q contains all nodes of W at distance $< D$ from w, with high probability. ◀

We sometimes sample edges instead of vertices, so analogous lemmas to Lemmas 5 and 6 hold when the sample is from a set of edges. Here is the analogue of Lemma 6. The other lemma is similar.

▶ **Lemma 7.** *Let $G = (V, E)$ be a (possibly directed and weighted graph) and let $M, W \subseteq V$. Let $E' \subseteq E$ be a random subset of $c(|E|/g) \ln n$ edges for some large enough constant c and some integer $g \geq 1$. Let Q be the endpoints of edges in E' that are in W.*

Then, for any $D > 0$, and for any w with $d(w, S) > D$, if one takes the closest g edges of E' to w wrt the distance from their W endpoints, they will contain all edges of E' whose W endpoints are at distance $< D$ from w, with high probability.

2.2 Preliminaries for lower bounds

The Strong Exponential Time Hypothesis (SETH) asserts that on a Word-RAM with $O(\log n)$ bit words, there is no $(2 - \varepsilon)^n$ time (possibly randomized) algorithm for some constant $\varepsilon > 0$ that can determine whether a given CNF-Formula with n variables and $O(n)$ clauses is satisfiable. (This version of SETH is equivalent to the original formulation by Impagliazzo, Paturi and Zane [15].) By a result of Williams [27], the following Orthogonal Vectors (OV) Problem requires $n^{2-o(1)}$poly (d) time (on a word-RAM with $O(\log n)$ bit words), unless SETH fails: given two sets $U, V \subseteq \{0, 1\}^d$ with $|U| = |V| = n$ and $d = \omega(\log n)$, determine whether there are $u \in U, v \in V$ with $u \cdot v = 0$.

Given an arbitrary instance of OV with $d = \tilde{O}(1)$ (while respecting $d = \omega(\log n)$, e.g. $d = \Theta(\log^2 n)$), consider the following graph representation, which we call the *OV-graph*: the vertex set consists of a node for every $u \in U$, for every $v \in V$ and for every coordinate $c \in [d] = C$, and there is an edge $(x \in U \cup V, c \in C)$ if and only if $x[c] = 1$. OV is then equivalent to the question of whether there exist $u \in U, v \in V$ such that $d(u, v) > 2$. In fact, it is equivalent to distinguishing whether for every $u \in U, v \in V$, $d(u, v) = 2$ (no OV-solution), or there is some $u \in U, v \in V$ such that $d(u, v) \geq 4$ (OV-solution). In other words, if we set $S = U, T = V$, the ST-Diameter of the OV-graph is 2 if and only if there is no OV-solution and at least 4 otherwise. Because the OV graph has $m = \tilde{O}(n)$, under SETH, any $(2 - \delta)$-approximation algorithm for ST-Diameter requires $m^{2-o(1)}$.

A related problem to OV is the Hitting Set (HS) problem [1, 14, 26]: given two sets $U, V \subseteq \{0, 1\}^d$ with $|U| = |V| = n$ and $d = \omega(\log n)$, determine whether there is $u \in U$ such that for all $v \in V$, $u \cdot v \neq 0$. A common hypothesis is that (on the word-RAM) HS requires $n^{2-o(1)}$ time.

If we form the OV-graph on the HS instance input, then the HS problem becomes equivalent to determining whether there is some $u \in U$ such that for all $v \in V$, $d(u, v) \leq 2$. In other words, if we set $S = U, T = V$, the ST-Radius of the OV-graph is 2 if and only if there is a HS-solution and at least 4 otherwise. Thus, under the HS hypothesis, any $(2 - \delta)$-approximation algorithm for ST-Radius requires $m^{2-o(1)}$.

Additionally for our constructions we assume that if there is a HS solution u' then for all $c \in C$, $d(u', c) \leq 3$. This is because for every coordinate index i there must be $v \in V$ with $v[i] = 1$ as otherwise we can just delete the i^{th} bit from all vectors.

Let $k \geq 2$ be an integer. Then, a generalization of the OV problem is k-OV: given k sets $U_1, \ldots, U_k \subseteq \{0, 1\}^d$, are there $u_1 \in U_1, \ldots, u_k \in U_k$ so that $\sum_{c=1}^{d} \prod_{i=1}^{k} u_i[c] = 0$? It is known that, under SETH, when $d = \omega(\log n)$, there is no $n^{k-o(1)}$ time algorithm for k-OV (in the word RAM model) [27].

Similar to the OV-graph, Backurs et al. [4] define a graph for k-OV which we will refer to as the k-OV-graph. We do not explicitly define the k-OV-graph here; instead we list its properties in the following theorem.

▶ **Theorem 8** ([4]). *Let $k \geq 2$. Given a k-OV instance consisting of sets $W_0, W_1, \ldots, W_{k-1} \subseteq \{0,1\}^d$, each of size n, we can in $O(kn^{k-1}d^{k-1})$ time construct an unweighted, undirected graph with $O(n^{k-1} + kn^{k-2}d^{k-1})$ vertices and $O(kn^{k-1}d^{k-1})$ edges that satisfies the following properties.*

1. *The graph consists of $k+1$ layers of vertices $L_0, L_1, L_2, \ldots, L_k$. The number of nodes in the sets is $|L_0| = |L_k| = n^{k-1}$ and $|L_1|, |L_2|, \ldots, |L_{k-1}| \leq n^{k-2}d^{k-1}$.*
2. *L_0 consists of all tuples $(a_0, a_1, \ldots, a_{k-2})$ where for each i, $a_i \in W_i$. Similarly, L_k consists of all tuples $(b_1, b_2, \ldots, b_{k-1})$ where for each i, $b_i \in W_i$.*
3. *If the k-OV instance has no solution, then $d(u, v) = k$ for all $u \in L_0$ and $v \in L_k$.*
4. *If the k-OV instance has a solution $a_0, a_1, \ldots, a_{k-1}$ where for each i, $a_i \in W_i$ then if $\alpha = (a_0, \ldots a_{k-2}) \in L_0$ and $\beta = (a_1, \ldots, a_{k-1}) \in L_k$, then $d(\alpha, \beta) \geq 3k - 2$.*
5. *For all i from 1 to $k-1$, for all $v \in L_i$ there exists a vertex in L_{i-1} adjacent to v and a vertex in L_{i+1} adjacent to v.*

2.3 Organization

In Section 3 we present our algorithms and conditional lower bound for undirected bichromatic diameter. We defer the rest of our algorithms and conditional lower bounds to the full version [12].

3 Algorithms and Lower Bound for Undirected Bichromatic Diameter

We begin with a simple near-linear time algorithm.

▶ **Proposition 9.** *There is an $O(m + n \log n)$ time algorithm, that given an undirected graph $G = (V, E)$ and $S \subseteq V, T = V \setminus S$, can output an estimate D' such that $D_{ST}(G)/2 - W/2 \leq D' \leq D_{ST}$, where W is the minimum weight of an edge in $S \times T$.*

Proof. Let (s, t) be a minimum weight edge of G with $s \in S$ and $t \in T$. Run Dijkstra's algorithm from s and from t. Let $D' = \max\{\max_{t' \in T} d(s, t'), \max_{s' \in S} d(s', t)\}$. Let $s^* \in S, t^* \in T$ be endpoints of an ST-Diameter path, i.e. $d(s^*, t^*) = D_{ST}$. Then, suppose that $\max_{t' \in T} d(s, t') < D_{ST}/2 - W/2$. In particular, $d(s, t^*) < D_{ST}/2 - W/2$, and hence $d(s, s^*) > D_{ST}/2 + W/2$ by the triangle inequality. Also by the triangle inequality,

$$D_{ST}/2 + W/2 < d(s, t) + d(t, s^*) \leq w(s, t) + \max_{s' \in S} d(s', t).$$

Hence, $D' > D_{ST}/2 - W/2$, where W is the minimum weight of an edge in $S \times T$. ◀

Now we turn to our 5/3-approximation algorithms. Our first theorem is for unweighted graphs. Later on, we modify the algorithm in this theorem to obtain an algorithm for weighted graphs as well, and at the same time remove the small additive error that appears in the theorem below.

▶ **Theorem 10.** *There is an $\tilde{O}(m\sqrt{n})$ time algorithm, that given an unweighted undirected graph $G = (V, E)$ and $S \subseteq V, T = V \setminus S$, can output an estimate D' such that $3D_{ST}(G)/5 \leq D' \leq D_{ST}(G)$ if $D_{ST}(G)$ is divisible by 5, and otherwise $3D_{ST}(G)/5 - 6/5 \leq D' \leq D_{ST}(G)$.*

Proof. Let $D = D_{ST}(G)$ and let us assume that D is divisible by 5. If D is not divisible by 5, the estimate we return will have a small additive error. For clarity of presentation, we omit the analysis of the case where D is not divisible by 5. However, we include such analyses in our proofs for Bichromatic Radius and ST-Eccentricities (see ARXIV) and the analysis for Diameter is analogous.

Suppose the (bichromatic) ST-Diameter endpoints are $s^* \in S$ and $t^* \in T$ and that the ST-Diameter is D. The algorithm does not know D, but we will use it in the analysis.

(Algorithm Step 1): The algorithm first samples $Z \subseteq S$ of size $c\sqrt{n}\ln n$ uniformly at random. For every $z \in Z$, run BFS, and let $D_1 = \max_{z\in Z,t\in T} d(z,t)$.

(Analysis Step 1): If for some $s' \in Z$ we have that $d(s^*,s') \le 2D/5$, then $D_1 \ge d(s',t^*) \ge D - d(s^*,s') \ge 3D/5$.

(Algorithm Step 2): Now, sample a set X from T of size $C\sqrt{n}\ln n$ uniformly at random for large enough constant C. For every $t \in X$, run BFS and find the closest node $s(t)$ of S to t. Run BFS from every $s(t)$. Let $D_2 = \max_{t\in X,t'\in T} d(s(t),t')$.

(Analysis Step 2): If s^* is at distance $\le D/5$ from some node t of X, then $d(s^*,s(t)) \le 2D/5$ (since $s(t)$ is closer to t than s^*), and so $D_2 \ge d(s(t),t^*) \ge 3D/5$.

If neither D_1, nor D_2 are good approximations, it must be that $d(s^*,X) > D/5$ and $d(s^*,Z) > 2D/5$. Consider the nodes M of S that are at distance $> 2D/5$ from Z, then the node $w \in M$ that is furthest from X among all nodes of M. If neither D_1, nor D_2 was a good approximation, $s^* \in M$ and since $d(s^*,X) > D/5$, we must have that $d(w,X) > D/5$ (and also $d(w,Z) > 2D/5$). In the next step we will look for such a w.

(Algorithm Step 3): For each $s \in S$ define D_s to be the biggest integer which satisfies $d(s,X) > D_s/5$ and $d(s,Z) > 2D_s/5$. Let $w = \arg\max D_s$ and $D' = \max D_s$.

(Analysis Step 3): By Lemma 6 we have that whp, the number of nodes of T at distance $\le D'/5$ from w and the number of nodes of S at distance $\le 2D'/5$ from w are both $\le \sqrt{n}$. Also if neither D_1, nor D_2 are good approximations, it must be that $d(s^*,X) > D/5$ and $d(s^*,Z) > 2D/5$ and hence $D' \ge D$.

(Algorithm Step 4): Run BFS from w. Take all nodes of S at distance $\le 2D'/5$ from w, call these S_w, and run BFS from them. Whp, $|S_w| \le \sqrt{n}$, so that this BFS run takes $O(m\sqrt{n})$ time. Let $D_3 := \max_{s\in S_w,t\in T} d(s,t)$.

For every $s \in S_w$, let $t(s)$ be the closest node of T to s (breaking ties arbitrarily). Run BFS from each $t(s)$. Let $D_4 := \max_{s\in S_w,s'\in S} d(s',t(s))$.

(Analysis Step 4): If $D_3 \ge 3D/5$ or $D_4 \ge 3D/5$, we are done, so let us assume that $D_3, D_4 < 3D/5$. Since $D_3 < 3D/5$, and since $D_3 \ge d(w,t^*)$, it must be that $d(w,t^*) < 3D/5$. Let P_{wt^*} be the shortest w to t^* path. Consider the node b on P_{wt^*} for which $d(w,b) = 2D/5$. If $b \in S$, then since $D' \ge D$, $b \in S_w$ and hence we ran BFS from $t(b)$. But since $d(b,t^*) = d(w,t^*) - 2D/5 < D/5$, and $d(b,t(b)) \le d(b,t^*)$ we have that $d(t(b),t^*) \le 2D/5$ and hence $D_4 \ge d(s^*,t(b)) \ge D - d(t(b),t^*) \ge 3D/5$. Thus, if $D_4 < 3D/5$, it must be that $b \in T$.

(Algorithm Step 5): Take all nodes of T at distance $\le D'/5$ from w, call these T_w and run BFS from them. Since $d(w,X) > D'/5$, whp $|T_w| \le \sqrt{n}$, so this step runs in $O(m\sqrt{n})$ time. Let $D_5 = \max_{t\in T_w,s\in S} d(t,s)$.

(Analysis Step 5): If $D_5 \ge 3D/5$, we would be done, so assume that $D_5 < 3D/5$. Let a be the node on the shortest w to t^* path P_{wt^*} with $d(w,a) = D/5$. Suppose that $a \in T$. Since $D' \ge D$, $a \in T_w$ and we ran BFS from it. However, also $d(a,t^*) = d(w,t^*) - d(w,a) < 3D/5 - D/5 = 2D/5$, and hence $D_5 \ge d(a,s^*) \ge d(t^*,s^*) - d(t^*,a) \ge D - 2D/5 = 3D/5$. Since $D_5 < 3D/5$, it must be that $a \in S$.

Now, since $a \in S$ and $b \in T$, somewhere on the a to b shortest path P_{ab}, there must be an edge (s',t') with $s' \in S, t' \in T$. Since s' is before b, $d(w,s') \le 2D/5 \le 2D'/5$, and hence $s' \in S_w$. Thus we ran BFS from $t(s')$. Since s' has an edge to $t' \in T$, $d(s',t(s')) \le d(s',t') = 1$.

Also, since $d(w, s') \geq d(w, a) = D/5$ and $d(w, t^*) \leq 3D/5 - 1$, $d(s', t^*) \leq 2D/5 - 1$. Thus,

$$\begin{aligned} D_4 \geq d(t(s'), s^*) &\geq d(s^*, t^*) - d(t(s'), t^*) \\ &\geq D - d(t(s'), s') - d(s', t^*) \\ &\geq D - 1 - 2D/5 + 1 = 3D/5. \end{aligned}$$

Hence if we set $D'' = \max\{D_1, D_2, D_3, D_4, D_5\}$, we get that $3D/5 \leq D'' \leq D$. ◀

We now modify the algorithm for unweighted graphs, both making the algorithm work for weighted graphs and removing the additive error, at the expense of increasing the runtime to $\tilde{O}(m^{3/2})$.

▶ **Theorem 11.** *There is an $\tilde{O}(m^{3/2})$ time algorithm, that given an undirected graph $G = (V, E)$ with nonnegative integer edge weights and $S \subseteq V, T = V \setminus S$, can output an estimate D' such that $3D_{ST}(G)/5 \leq D' \leq D_{ST}$.*

Proof. Suppose as before the (bichromatic) ST-Diameter endpoints are $s^* \in S$ and $t^* \in T$ and that the ST-Diameter is D.

(Algorithm Modified Step 1): The algorithm here samples $E' \subseteq E$ of size $c\sqrt{m}\ln n$ uniformly at random, for large enough c. Let Z be the endpoints of edges in E' that are in S. For every $z \in Z$, run Dijkstra's algorithm, and let $D_1 = \max_{z\in Z, t\in T} d(z, t)$.

(Analysis Step 1): If for some $s' \in Z$ we have that $d(s^*, s') \leq 2D/5$, then $D_1 \geq d(s', t^*) \geq D - d(s^*, s') \geq 3D/5$. Let us then assume that $d(s^*, Z) > 2D/5$.

(Algorithm Modified Step 2): Let X be the endpoints of edges in E' that are in T. For every $t \in X$, run Dijkstra's algorithm and find the closest node $s(t)$ of S to t. Run Dijkstra's algorithm from every $s(t)$. Let $D_2 = \max_{t\in X, t'\in T} d(s(t), t')$.

(Analysis Step 2): If s^* is at distance $\leq D/5$ from some node t of X, then $d(s^*, s(t)) \leq 2D/5$ (since $s(t)$ is closer to t than s^*), and so $D_2 \geq d(s(t), t^*) \geq 3D/5$. Let us then assume that $d(s^*, X) > D/5$.

As before, if we consider the nodes M of S that are at distance $> 2D/5$ from Z, then the node $w \in M$ that is furthest from X among all nodes of M, would have both $d(w, Z) > 2D/5$ and $d(w, X) > D/5$, as s^* is in M and satisfies $d(s^*, X) > D/5$. We will find a node w with these properties in the next step.

(Algorithm Unmodified Step 3): Perform exactly the same Step 3 as before, finding the largest integer D' such that there is some node $w \in S$ with $d(w, Z) > 2D'/5$ and $d(w, X) > D'/5$.

(Analysis Step 3): Let $w \in S$ be the node we found such that $d(w, X) > D'/5, d(w, Z) > 2D'/5$. By Lemma 7 we have that whp, the number of edges (s, g) where $s \in S, g \in V$ and $d(w, s) \leq 2D'/5$ and the number of edges (t, g') where $t \in T, g' \in V$ and $d(w, t) \leq D'/5$ is at most $\sqrt{m}$. Also, if $D_1, D_2 < 3D/5$, then $D' \geq D$, so that we also have that the number of edges (s, b) where $s \in S$ and $d(w, s) \leq 2D/5$ and the number of edges (t, b') where $t \in T$ and $d(w, t) \leq D/5$ is at most $\sqrt{m}$, whp.

(Algorithm Modified Step 4): Run Dijkstra's algorithm from w. Take all edges incident to nodes of S at dist $\leq 2D'/5$ from w. Call these edges E_S and their endpoints S_w. Run Dijkstra's algorithm from both of their end points. Whp, $|E_S| \leq \sqrt{m}$ and so $|S_w| \leq 2\sqrt{m}$, so that this Dijkstra run takes $\tilde{O}(m^{3/2})$ time. Let $D_3 := \max_{t\in S_w\cap T, s\in S} d(s, t)$.

For every $s \in S_w \cap S$, determine a closest node $t(s) \in T$ to s, and run Dijkstra's algorithm from $t(s)$ as well. This search also takes $O(m^{3/2})$ time. Let $D_4 := \max_{s\in S_w\cap S, s'\in S} d(s', t(s))$.

(Analysis Step 4): If $d(w,t^*) \geq 3D/5$, or $D_3 \geq 3D/5$ or $D_4 \geq 3D/5$, we are done, so let us assume that $d(w,t^*), D_3, D_4 < 3D/5$.
Now consider the node b on the shortest w to t^* path P_{wt^*} for which $d(w,b) \leq 2D/5$, but such that the node b' after it on P_{wt^*} has $d(w,b') > 2D/5$.
Suppose that $b \in S$. Then since $D' \geq D$, we have $d(w,b) \leq 2D'/5$ and hence $(b,b') \in E_S$. Let us consider $d(b',t^*) = d(w,t^*) - d(b',w)$. Since $d(w,t^*) < 3D/5$ and $d(b',w) > 2D/5$, $d(b',t^*) < D/5$. If $b' \in T$, then since we ran Dijkstra's algorithm from b', we got $D_3 \geq D - D/5 = 4D/5$. If $b' \in S$, then we ran Dijkstra's algorithm from $t(b')$ and $d(t(b'),t^*) \leq d(t(b'),b') + d(b',t^*) \leq 2d(b',t^*) < 2D/5$, and hence $D_4 \geq d(t(b),s^*) \geq D - 2D/5 = 3D/5$. Thus if neither $d(w,t^*)$, D_3, nor D_4 are good approximations, then $b \in T$.

(Algorithm Modified Step 5): Take all edges incident to nodes of T at dist $\leq D'/5$ from w. Call these edges E_T and their endpoints that are in T, T_w. Run Dijkstra's algorithm from all nodes in T_w.
Since $d(w,X) > D'/5$, whp $|T_w| \leq 2\sqrt{m}$, so this step runs in $O(m^{3/2})$ time. Let $D_5 = \max_{t \in T_w, s \in S} d(t,s)$.

(Analysis Step 5): If $D_5 \geq 3D/5$, we would be done, so assume that $D_5 < 3D/5$. Let a be the node on P_{wt^*} with $d(w,a) \leq D/5$ but so that the node a' after a on P_{wt^*} has $d(w,a') > D/5$. Suppose that $a' \in T$. Since $D' \geq D$, $(a,a') \in E_T$, $a' \in T_w$ and we ran Dijkstra's algorithm from a'. However, also $d(a',t^*) = d(w,t^*) - d(w,a') < 3D/5 - D/5 = 2D/5$, and hence $D_5 \geq d(a,s^*) \geq d(t^*,s^*) - d(t^*,a') \geq D - 2D/5 = 3D/5$. Since $D_5 < 3D/5$, it must be that $a' \in S$.

Now, since $a' \in S$ and $b \in T$, somewhere on the a' to b shortest path P_{ab}, there must be an edge (s',t') with $s' \in S, t' \in T$. However, since s' is before b, we have that $d(w,s') \leq d(w,b) \leq 2D/5 \leq 2D'/5$. Thus, $(s',t') \in E_S$ and we ran Dijkstra's algorithm from t'. However, $d(t',t^*) = d(w,t^*) - d(w,t') \leq d(w,t^*) - d(w,a') < 3D/5 - D/5 = 2D/5$, and hence $D_3 \geq d(t',s^*) \geq d(s^*,t^*) - d(t',t^*) > 3D/5$.

Hence if we set $D'' = \max\{d(w,t^*), D_1, D_2, D_3, D_4, D_5\}$, we get that $3D/5 \leq D'' \leq D$. ◀

Conditional Lower Bound

The following theorem implies that our algorithms for undirected Bichromatic Diameter from Theorem 11 and Proposition 9 are tight under SETH.

▶ **Theorem 12.** *Under SETH, for every $k \geq 2$, every algorithm that can distinguish between Bichromatic Diameter $2k-1$ and $4k-3$ in undirected unweighted graphs requires $m^{1+1/(k-1)-o(1)}$ time.*

In particular setting $k = 2$ and 3 in Theorem 12 implies that our $m^{3/2}$ time 5/3-approximation algorithm from Theorem 11 is tight in approximation factor and runtime, respectively. Furthermore, setting k to be arbitrarily large implies that our $\tilde{O}(m)$ time almost 2-approximation algorithm from Proposition 9 is tight under SETH.

Theorem 12 follows from the following lemma.

▶ **Lemma 13.** *Let $k \geq 2$ be any integer. Given a k-OV instance, we can in $O(kn^{k-1}d^{k-1})$ time construct an unweighted, undirected graph with $O(kn^{k-1} + kn^{k-2}d^{k-1})$ vertices and $O(kn^{k-1}d^{k-1})$ edges that satisfies the following two properties.*

1. *If the k-OV instance has no solution, then for* all *pairs of vertices $u \in S$ and $v \in T$ we have $d(u,v) \leq 2k-1$.*

2. *If the k-OV instance has a solution, then there exists a pair of vertices* $u \in S$ *and* $v \in T$ *such that* $d(u, v) \geq 4k - 3$.

Proof.

Construction of the graph. We begin with the k-OV-graph from Theorem 8. Additionally, we add $k - 1$ new layers of vertices $L_{k+1}, \ldots, L_{2k-1}$, where each new layer contains n^{k-1} vertices and is connected to the previous layer by a matching. That is, each new layer contains one vertex for every tuple $(a_1, \ldots, a_{k-1})$ where $a_i \in W_i$ for all i, and each $(a_1, \ldots, a_{k-1}) \in L_j$ is connected to its counterpart $(a_1, \ldots, a_{k-1}) \in L_{j-1}$ by an edge, for all j.

We let $S = L_0$ and we let T contain the rest of the vertices in the graph.

Correctness of the construction.

Case 1: The k-OV instance has no solution. By property 3 of Theorem 8 for all $u \in S$ and $v \in L_k$, $d(u, v) = k$. Then, since $L_k, \ldots, L_{2k-1}$ form a series of matchings, for all $u \in S$ and $v \in L_{k+1} \cup \cdots \cup L_{2k-1}$, $d(u, v) \leq 2k - 1$. Furthermore, property 5 of Theorem 8 implies that for all $u \in S$ and $v \in L_1 \cup \cdots \cup L_{k-1}$, $d(u, v) \leq 2k - 1$. Thus, we have shown that for all $u \in S$ and $v \in T$ we have $d(u, v) \leq 2k - 1$.

Case 2: The k-OV instance has a solution. Let $(a_0, a_1, \ldots, a_{k-1})$ be a solution to the k-OV instance where $a_i \in W_i$ for all i. We claim that $d((a_0, \ldots, a_{k-2}) \in S, (a_1, \ldots, a_{k-1}) \in L_{2k-1})) \geq 4k - 3$. Since $L_k, \ldots, L_{2k-1}$ form a series of matchings, every path from $(a_0, \ldots, a_{k-2}) \in S$ to $(a_1, \ldots, a_{k-1}) \in L_{2k-1}$ contains the vertex $(a_1, \ldots, a_{k-1}) \in L_k$. By property 4 of Theorem 8, $d((a_0, \ldots, a_{k-2}) \in S, (a_1, \ldots, a_{k-1}) \in L_k) \geq 3k - 2$. Thus, $d((a_0, \ldots, a_{k-2}) \in S, (a_1, \ldots, a_{k-1}) \in L_{2k-1})) \geq 4k - 3$. ◀

References

1. Amir Abboud, Virginia Vassilevska Williams, and Joshua R. Wang. Approximation and Fixed Parameter Subquadratic Algorithms for Radius and Diameter in Sparse Graphs. In *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2016, Arlington, VA, USA, January 10-12, 2016*, pages 377–391, 2016.
2. Pankaj K. Agarwal, Herbert Edelsbrunner, Otfried Schwarzkopf, and Emo Welzl. Euclidean Minimum Spanning Trees and Bichromatic Closest Pairs. *Discrete Comput. Geom.*, 6(1):407–422, December 1991.
3. D. Aingworth, C. Chekuri, P. Indyk, and R. Motwani. Fast Estimation of Diameter and Shortest Paths (Without Matrix Multiplication). *SIAM Journal on Computing*, 28(4):1167–1181, 1999.
4. Arturs Backurs, Liam Roditty, Gilad Segal, Virginia Vassilevska Williams, and Nicole Wein. Towards Tight Approximation Bounds for Graph Diameter and Eccentricities. In *Proceedings of STOC'18*, page to appear, 2018.
5. Massimo Cairo, Roberto Grossi, and Romeo Rizzi. New Bounds for Approximating Extremal Distances in Undirected Graphs. In *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2016, Arlington, VA, USA, January 10-12, 2016*, pages 363–376, 2016.
6. T. M. Chan. More Algorithms for All-Pairs Shortest Paths in Weighted Graphs. In *Proc. STOC*, pages 590–598, 2007.
7. Bernard Chazelle, Herbert Edelsbrunner, Leonidas J Guibas, and Micha Sharir. Algorithms for bichromatic line-segment problems and polyhedral terrains. *Algorithmica*, 11(2):116–132, 1994.
8. Shiri Chechik, Daniel H. Larkin, Liam Roditty, Grant Schoenebeck, Robert Endre Tarjan, and Virginia Vassilevska Williams. Better Approximation Algorithms for the Graph Diameter. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2014, Portland, Oregon, USA, January 5-7, 2014*, pages 1041–1052, 2014.
9. Don Coppersmith and Michael Elkin. Sparse sourcewise and pairwise distance preservers. *SIAM Journal on Discrete Mathematics*, 20(2):463–501, 2006.

10 Pierluigi Crescenzi, Roberto Grossi, Leonardo Lanzi, and Andrea Marino. On Computing the Diameter of Real-World Directed (Weighted) Graphs. In Ralf Klasing, editor, *Experimental Algorithms: 11th International Symposium, SEA 2012, Bordeaux, France, June 7-9, 2012. Proceedings*, pages 99–110, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

11 Marek Cygan, Fabrizio Grandoni, and Telikepalli Kavitha. On pairwise spanners. *arXiv preprint*, 2013. arXiv:1301.1999.

12 Mina Dalirrooyfard, Virginia Vassilevska Williams, Nikhil Vyas, and Nicole Wein. Tight Approximation Algorithms for Bichromatic Graph Diameter and Related Problems, 2019. arXiv:1904.11601.

13 Adrian Dumitrescu and Sumanta Guha. Extreme Distances in Multicolored Point Sets. *J. Graph Algorithms and Applications*, 8(1):27–38, 2004.

14 Jiawei Gao, Russell Impagliazzo, Antonina Kolokolova, and R. Ryan Williams. Completeness for First-Order Properties on Sparse Structures with Algorithmic Applications. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2017, Barcelona, Spain, Hotel Porta Fira, January 16-19*, pages 2162–2181, 2017.

15 R. Impagliazzo and R. Paturi. On the Complexity of k-SAT. *J. Comput. Syst. Sci.*, 62(2):367–375, 2001.

16 Piotr Indyk. A near linear time constant factor approximation for Euclidean bichromatic matching (cost). In *SODA*, volume 7, pages 39–42, 2007.

17 Naoki Katoh and Kazuo Iwano. Finding K Farthest Pairs and K Closest/Farthest Bichromatic Pairs for Points in the Plane. In *Proceedings of the Eighth Annual Symposium on Computational Geometry*, SCG '92, pages 320–329, 1992.

18 Telikepalli Kavitha. New pairwise spanners. *Theory of Computing Systems*, 61(4):1011–1036, 2017.

19 Philip N Klein. A subset spanner for planar graphs, with application to subset TSP. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 749–756. ACM, 2006.

20 Clémence Magnien, Matthieu Latapy, and Michel Habib. Fast Computation of Empirically Tight Bounds for the Diameter of Massive Graphs. *J. Exp. Algorithmics*, 13:10:1.10–10:1.9, February 2009.

21 David Peleg, Liam Roditty, and Elad Tal. Distributed Algorithms for Network Diameter and Girth. In *Automata, Languages, and Programming: 39th International Colloquium, ICALP 2012, Warwick, UK, July 9-13, 2012, Proceedings, Part II*, pages 660–672, 2012.

22 S. Pettie. A new approach to all-pairs shortest paths on real-weighted graphs. *Theor. Comput. Sci.*, 312(1):47–74, 2004.

23 Seth Pettie and Vijaya Ramachandran. A Shortest Path Algorithm for Real-Weighted Undirected Graphs. *SIAM J. Comput.*, 34(6):1398–1431, 2005.

24 Liam Roditty and Virginia Vassilevska Williams. Fast approximation algorithms for the diameter and radius of sparse graphs. In *Proceedings of the 45th annual ACM symposium on Symposium on theory of computing*, STOC '13, pages 515–524, New York, NY, USA, 2013. ACM. doi:10.1145/2488608.2488673.

25 Frank W. Takes and Walter A. Kosters. Determining the Diameter of Small World Networks. In *Proceedings of the 20th ACM International Conference on Information and Knowledge Management*, CIKM '11, pages 1191–1196, 2011.

26 Virginia Vassilevska Williams. Hardness of Easy Problems: Basing Hardness on Popular Conjectures such as the Strong Exponential Time Hypothesis (Invited Talk). In *10th International Symposium on Parameterized and Exact Computation, IPEC 2015, September 16-18, 2015, Patras, Greece*, pages 17–29, 2015.

27 R. Williams. A new algorithm for optimal 2-constraint satisfaction and its implications. *Theor. Comput. Sci.*, 348(2–3):357–365, 2005.

28 Ryan Williams. On the Difference Between Closest, Furthest, and Orthogonal Pairs: Nearly-linear vs Barely-subquadratic Complexity. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '18, pages 1207–1215, 2018.

29 A. Yao. On Constructing Minimum Spanning Trees in k-Dimensional Spaces and Related Problems. *SIAM Journal on Computing*, 11(4):721–736, 1982.