A Note on Hardness of Diameter Approximation

Karl Bringmann[*] Sebastian Krinninger[†]

Abstract

We revisit the hardness of approximating the diameter of a network. In the CONGEST model, $\tilde{\Omega}(n)$ rounds are necessary to compute the diameter [Frischknecht et al. SODA'12]. Abboud et al. [DISC 2016] extended this result to sparse graphs and, at a more fine-grained level, showed that, for any integer $1 \leq \ell \leq \operatorname{polylog}(n)$, distinguishing between networks of diameter $4\ell + 2$ and $6\ell + 1$ requires $\tilde{\Omega}(n)$ rounds. We slightly tighten this result by showing that even distinguishing between diameter $2\ell + 1$ and $3\ell + 1$ requires $\tilde{\Omega}(n)$ rounds. The reduction of Abboud et al. is inspired by recent conditional lower bounds in the RAM model, where the orthogonal vectors problem plays a pivotal role. In our new lower bound, we make the connection to orthogonal vectors explicit, leading to a conceptually more streamlined exposition. This is suited for teaching both the lower bound in the CONGEST model and the conditional lower bound in the RAM model.

1 Introduction

In distributed computing, the diameter of a network is arguably the single most important quantity one wishes to compute. In the CONGEST model [Pel00], where in each round every vertex can send to each of its neighbors a message of size $O(\log n)$, it is known that $\tilde{\Omega}(n)$ rounds are necessary to compute the diameter [FHW12] even in sparse graphs [ACK16], where n is the number of vertices. With this negative result in mind, it is natural that the focus has shifted towards *approximating* the diameter. In this note, we revisit hardness of computing a diameter approximation in the CONGEST model from a *fine-grained* perspective.

The current fastest approximation algorithm [Hol+14], which is inspired by a corresponding RAM model algorithm [RW13], takes $O(\sqrt{n \log n} + D)$ rounds and computes a $\frac{3}{2}$-approximation of the diameter, i.e., an estimate $\hat{D}$ such that $\lfloor \frac{2}{3} D \rfloor \leq \hat{D} \leq D$, where D is the true diameter of the network. In terms of lower bounds, Abboud, Censor-Hillel, and Khoury [ACK16] showed that $\tilde{\Omega}(n)$ rounds are necessary to compute a $(\frac{3}{2} - \varepsilon)$-approximation of the diameter for any constant $0 < \varepsilon < \frac{1}{2}$. At a more fine-grained level, they show that, for any integer $1 \leq \ell \leq \operatorname{polylog}(n)$, at least $\tilde{\Omega}(n)$ rounds are necessary to decide whether the network has diameter $4\ell + 2$ or $6\ell + 1$, thus ruling out any "relaxed" notions of $(\frac{3}{2} - \varepsilon)$-approximation that additionally allow small additive error. We tighten this result by showing that, for any integer $\ell \geq 1$, at least $\tilde{\Omega}(n)$ rounds

[*]Max Planck Institute for Informatics, Saarland Informatics Campus, Germany

[†]University of Vienna, Faculty of Computer Science, Austria. Work partially done while at Max Planck Institute for Informatics.

are necessary to distinguish between diameter $2\ell + 1$ and $3\ell + 1$, and more generally between diameter $2\ell + q$ and $3\ell + q$ for any $\ell, q \geq 1$.

The reduction of Abboud et al. [ACK16] is inspired by recent work on conditional lower bounds in the RAM model, where the *orthogonal vectors problem* plays a pivotal role. In particular, the Orthogonal Vectors Hypothesis (OVH) is a weaker "polynomial-time analogue" of the Strong Exponential Time Hypothesis (SETH) [IPZ01] (it is well-known that SETH implies OVH [Wil05]). In our new lower bound, we make the connection to orthogonal vectors explicit: we consider a communication complexity version of orthogonal vectors that we show to be hard *unconditionally* by a reduction from set disjointness and then devise a reduction from orthogonal vectors to diameter approximation.

Additionally, our approach has implications in the RAM model. Prior to our work, the situation there was as follows. In their seminal paper [RW13], Roditty and Vassilevska Williams showed that, for any constants $\varepsilon > 0$ and $\delta > 0$ there is no algorithm that computes a $(\frac{3}{2} - \varepsilon)$-approximation of the diameter and runs in time $O(m^{2-\delta})$, unless the Strong Exponential Time Hypothesis (SETH) fails. In particular, they show that no algorithm can decide whether a given graph has diameter 2 or 3 in time $O(m^{2-\delta})$, unless the Strong Exponential Time Hypothesis (SETH) fails. The hardness of 2 vs. 3 is already implied by the weaker Orthogonal Vectors Hypothesis (OVH), which in turn is implied by SETH [Wil05] and was popularized after the paper of Roditty and Vassilevska Williams appeared. It has then been shown by Chechik et al. [Che⁺14] that, for any integer $1 \leq \ell \leq n^{o(1)}$, there is no algorithm that distinguishes between diameter $3(\ell + 1)$ and $4(\ell + 1)$ with running time $O(m^{2-\delta})$ for some constant $\delta > 0$, unless SETH fails. Finally, Cairo, Grossi, and Rizzi [CGR16] showed that, for any integer $1 \leq \ell \leq n^{o(1)}$, there is no algorithm that distinguishes between diameter 2ℓ and 3ℓ with running time $O(m^{2-\delta})$ for some constant $\delta > 0$, unless SETH fails. Our reduction reconstructs the result of Cairo et al. under the weaker hardness assumption OVH, yielding again a more streamlined chain of reductions.

2 Reduction from Set Disjointness to Orthogonal Vectors

Set disjointness is a problem in communication complexity between two players, called Alice and Bob, in which Alice is given an n-dimensional bit vector x and Bob is given an n-dimensional bit vector y and the goal for Alice and Bob is to find out whether there is some index k at which both vectors contain a 1, i.e., such that $x[k] = y[k] = 1$ (meaning the sets represented by x and y are not disjoint). The relevant measure in communication complexity is the number of bits exchanged by Alice and Bob in any protocol that Alice and Bob follow to determine the solution. A classic result [KN97, Raz92] states that any such protocol requires Alice and Bob to exchange $\Omega(n)$ bits to solve set disjointness.

In the orthogonal vectors problem, Alice is given a set of bit vectors $L = \{l_1, \ldots, l_n\}$ and Bob is given a set of bit vectors $R = \{r_1, \ldots, r_n\}$, and the goal for them is to find out if there is a pair of orthogonal vectors $l_i \in L$ and $r_j \in R$ (i.e., such that, for every $1 \leq k \leq d$, $l_i[k] = 0$ or $r_j[k] = 0$). We give a reduction from set disjointness to orthogonal vectors.

Theorem 2.1. *Any b-bit protocol for the orthogonal vectors problem in which Alice and Bob each hold n vectors of dimension $d = 2\lceil \log n \rceil + 3$, gives a b-bit protocol for the set disjointness problem* *where Alice and Bob each hold an n-dimensional bit vector.*

Proof. We show that, without any communication, Alice and Bob can transform a set disjointness instance $\langle x, y\rangle$ with n-dimensional bit vectors into an orthogonal vectors instance $\langle L, R\rangle$ such that x and y are not disjoint if and only if $\langle L, R\rangle$ contains an orthogonal pair. For every integer $1 \le i \le n$, let s_i denote the binary representation of i with $\lceil \log n \rceil$ bits. For every bit b, let $\bar{b}$ be the result of 'flipping' bit b, i.e., $\bar{1} = 0$, and $\bar{0} = 1$. Similarly, for a bit vector b, let $\bar{b}$ be the result of flipping each bit of b. For every $1 \le i \le n$, let l_i to be the vector obtained from concatenating $x[i]$, $\bar{x}[i]$, $\bar{x}[i]$, s_i, and $\bar{s}_i$. For every $1 \le j \le n$, let r_i to be the vector obtained from concatenating $\bar{y}[i]$, $y[i]$, $\bar{y}[i]$, $\bar{s}_i$, and s_i.

We now claim that the vectors x and y are not disjoint if and only if $\langle L, R\rangle$ contains an orthogonal pair. If the vectors x and y are not disjoint, then there is some k such that $x[k] = y[k] = 1$. Clearly, s_i and $\bar{s}_i$ are orthogonal and, as the vectors $(x[i], \bar{x}[i], \bar{x}[i])$ and $(\bar{y}[i], y[i], \bar{y}[i])$ are equal to $(1, 0, 0)$ and $(0, 1, 0)$, respectively, they are also orthogonal. It follows that l_i and r_i are orthogonal.

Now assume that $\langle L, R\rangle$ contains an orthogonal pair $l_i \in L$ and $r_j \in R$. We first show that $i = j$. Suppose for the sake of contradiction that $i \ne j$. Then the binary representations s_i and s_j differ in at least one bit, say $s_i[k] \ne s_j[k]$. If $s_i[k] = 0$ and $s_j[k] = 1$, then $\bar{s}_i$ and s_j are not orthogonal and thus l_i and r_j are not orthogonal, contradicting the assumption. If $s_i[k] = 1$ and $s_j[k] = 0$, then s_i and $\bar{s}_j$ are not orthogonal and thus l_i and r_j are not orthogonal, contradicting the assumption. It follows that $i = j$ and thus the vectors $(x[i], \bar{x}[i], \bar{x}[i])$ and $(\bar{y}[i], y[i], \bar{y}[i])$ are orthogonal. Orthogonality of $x[i]$ and $\bar{y}[i]$ rules out $x[i] = 1$ and $y[i] = 0$, orthogonality of $\bar{x}[i]$ and $y[i]$ rules out $x[i] = 0$ and $y[i] = 1$, and orthogonality of $\bar{x}[i]$ and $\bar{y}[i]$ rules out $x[i] = 0$ and $y[i] = 0$. It follows that $x[i] = y[i] = 1$, making x and y not disjoint. □

The hardness of set disjointness now directly transfers to orthogonal vectors.

Corollary 2.2. *Any protocol solving the orthogonal vectors problem with n vectors of dimension* $d = 2\lceil \log n \rceil + 3$, *requires Alice and Bob to exchange* $\Omega(n)$ *bits.*

3 Reduction from Orthogonal Vectors to Diameter

We now establish hardness of distinguishing between networks of diameter $2\ell + q$ and $3\ell + q$, where $\ell \ge 1$ and in the CONGEST model $q \ge 1$, whereas in the RAM model $q \ge 0$. To unify the cases of odd and even ℓ, we introduce an additional parameter $p \in \{0, 1\}$ and change the task to distinguishing between networks of diameter $4\ell' - 2p + q$ and $6\ell' - 3p + q$ for integers $\ell' \ge 1$, $q \ge 0$, and $p \in \{0, 1\}$. This covers the original question: if ℓ is even, then set $\ell' := \ell/2$ and $p := 0$ and if ℓ is odd, then set $\ell' := \lceil \ell/2 \rceil$ and $p := 1$.

3.1 Construction and Implications

Given an orthogonal vectors instance $\langle L := \{l_1, \ldots, l_n\}, R := \{r_1, \ldots, r_n\}\rangle$ of d-dimensional vectors and parameters $\ell \ge 1$, $q \ge 0$, and $p \in \{0, 1\}$, we define an unweighted undirected graph $G := G_{L,R,\ell,p,q}$ as follows. The graph G contains the following *exterior* vertices: $u^L_1, \ldots, u^L_n$, $u^R_1, \ldots, u^R_n$, $v^L_1, \ldots, v^L_n$, $v^R_1, \ldots, v^R_n$, $w^L_1, \ldots, w^L_d$, $w^R_1, \ldots, w^R_d$, x^L, x^R, y^L, and y^R. These exterior vertices are connected by paths as follows, where each path introduces a separate set of interior vertices:

- For every $1 \le i \le n$, add paths $\pi(u_i^L, v_i^L)$ and $\pi(u_i^R, v_i^R)$, each of length $\ell - p$.
- For every $1 \le i \le n$, add paths $\pi(v_i^L, x^L)$ and $\pi(v_i^R, x^R)$, each of length ℓ.
- For every $1 \le i \le n$ and every $1 \le k \le d$ such that $l_i[k] = 1$ add a path $\pi(v_i^L, w_k^L)$ of length ℓ.
- For every $1 \le i \le n$ and every $1 \le k \le d$ such that $r_i[k] = 1$ add a path $\pi(v_i^R, w_k^R)$ of length ℓ.
- For every $1 \le k \le d$, add paths $\pi(y^L, w_k^L)$ and $\pi(y^R, w_k^R)$, each of length ℓ.
- For every $1 \le k \le d$, add a path $\pi(w_k^L, w_k^R)$ of length q. That is, if $q = 0$ then identify w_i^L and w_i^R.
- Add a path $\pi(x^L, y^L)$ and a path $\pi(x^R, y^R)$, each of length $\ell - p$.
- Add a path $\pi(y^L, y^R)$ of length $p + q$. That is, if $p + q = 0$, then identify y^L and y^R.

The graph G is visualized in Figure 1. Observe that G has $O(nd\ell + dq)$ vertices and $O(nd\ell + dq)$ edges. We show that our construction has the following formal guarantees.

Theorem 3.1. *Let $\langle L, R\rangle$ be an orthogonal vectors instance of two sets of d-dimensional vectors of size n each and let $\ell \ge 1$, $p \in \{0, 1\}$, and $q \ge 0$ be integer parameters. Then the unweighted, undirected graph $G := G_{L,R,\ell,p,q}$ has $O(nd\ell + dq)$ vertices and edges and its diameter D has the following property: if $\langle L, R\rangle$ contains an orthogonal pair, then $D = 6\ell - 3p + q$, and if $\langle L, R\rangle$ contains no orthogonal pair, then $D = 4\ell - 2p + q$.*

Before we give a proof of this statement, we motivate by discussing its immediate consequences in the CONGEST model and the RAM model. For the CONGEST model, observe that G has a small cut of size $d + 1$ between its left hand side and its right hand side. A standard simulation argument, where communication between Alice and Bob is limited to messages sent along the small cut, yields our main result.

Corollary 3.2. *In the CONGEST model, any algorithm distinguishing between graphs of diameter $2\ell + q$ and graphs of diameter $3\ell + q$ when $\ell \ge 1$ and $q \ge 1$ requires $\Omega(n/((\ell + q)\log^3 n))$ rounds.*

Proof. Let $\langle L, R\rangle$ be an orthogonal vectors instance with n vectors of dimension $d = 2\lceil \log n\rceil + 3$ and let $\mathcal{A}$ be an algorithm distinguishing between graphs of diameter $2\ell + 1$ and graphs of diameter $3\ell + 1$. If ℓ is even, then set $\ell' := \ell/2$ and $p := 0$ and if ℓ is odd, then set $\ell' := \lceil \ell/2\rceil$ and $p := 1$. Then by Theorem 3.1 the graph $G := G_{L,R,\ell',p,q}$ has diameter $3\ell + q$ if $\langle L, R\rangle$ contains an orthogonal pair and $2\ell + q$ otherwise. Observe that G has $n' = O(n(\ell + q)\log n)$ edges and since $q \ge 1$ it can be partitioned into two node sets A and B such that

- $G[A]$, the subgraph of G induced by A, is fully determined by L, ℓ', p, and q.

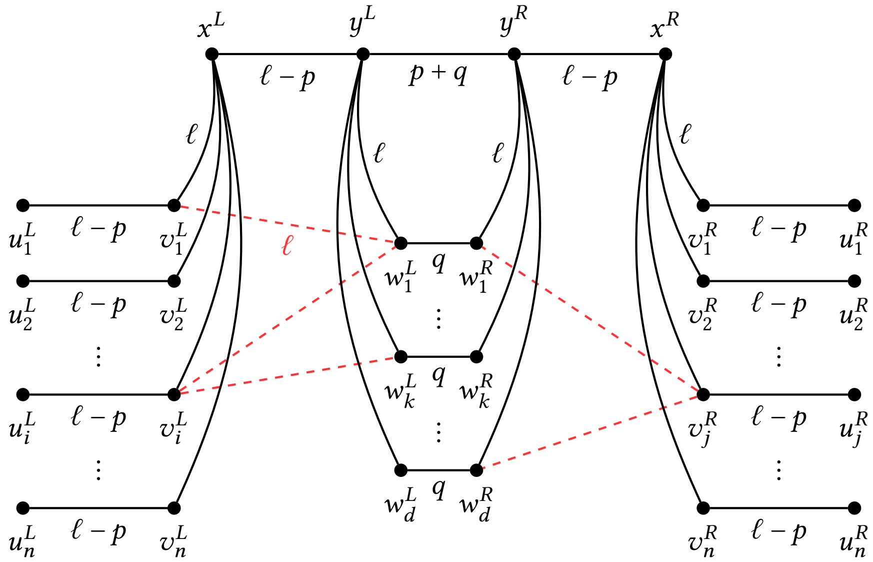

Figure 1: Visualization of the graph $G := G_{L,R,\ell,p,q}$ used in our reduction from orthogonal vectors to diameter distinction. The red, dashed edges encode the orthogonal vectors instance: For every $1 \leq i \leq n$ and every $1 \leq k \leq d$, the path $\pi(v_i^L, w_k^L)$ is contained in G if and only if $l_i[k] = 1$. Similarly, for every $1 \leq j \leq n$ and every $1 \leq k \leq d$, the path $\pi(v_i^R, w_k^R)$ is contained in G if and only if $r_j[k] = 1$.

- $G[B]$, the subgraph of G induced by B, is fully determined by R, ℓ', p, and q.
- The number of edges between A and B in G is $d + 1 = O(\log n)$.

Thus, Alice and Bob can simulate running $\mathcal{A}$ on the graph G as follows: Alice constructs the graph $G[A]$ and simulates the states of all vertices in A as well as the messages sent between them and Bob constructs the graph $G[B]$ and simulates the states of all vertices in B as well as the messages sent between them. Every time a message is sent from a node in A to a node in B, Alice communicates the $O(\log n)$-size message to Bob and every time a message is sent from a node in B to a node in A, Bob communicates the $O(\log n)$-size message to Alice. As A and B are separated by $O(\log n)$ edges, in each simulated round of $\mathcal{A}$ at most $O(\log^2 n)$ bits can be sent from Alice to Bob and vice versa. As Alice and Bob need to exchange $\Omega(n)$ bits to determine the result to the orthogonal vectors problem by Corollary 2.2, the algorithm $\mathcal{A}$ requires $\Omega(n/\log^2 n) = \Omega(n'/((\ell + q)\log^3 n'))$ rounds. □

In the RAM model, the Orthogonal Vectors Hypothesis (OVH) states that there is no algorithm that decides whether a given orthogonal vectors instance contains an orthogonal pair in time $O(n^{2-\delta}\,\mathrm{poly}(d))$ for some constant $\delta > 0$. Under this hardness assumption, our reduction has the following straightforward implication.

Corollary 3.3. *In the RAM model, under OVH, there is no algorithm distinguishing between graphs of diameter $2\ell + q$ and graphs of diameter $3\ell + q$, where $\ell \geq 1$ and $q \geq 0$, in time $O(m^{2-\delta}/(\ell + q)^{2-\delta})$ for any constant $\delta > 0$.*

Proof. Let $\langle L, R\rangle$ be an orthogonal vectors instance with n vectors of dimension d and let $\mathcal{A}$ be an algorithm distinguishing between graphs of diameter $2\ell + q$ and graphs of diameter $3\ell + q$ running in time $O(m^{2-\delta}/(\ell+q)^{2-\delta})$. If ℓ is even, then set $\ell' := \ell/2$ and $p := 0$ and if ℓ is odd, then set $\ell' := \lceil \ell/2 \rceil$ and $p := 1$. Then by Theorem 3.1 the graph $G := G_{L,R,\ell',p,q}$ has diameter $3\ell + q$ if $\langle L, R\rangle$ contains an orthogonal pair and $2\ell + q$ otherwise. Observe that G has $m = O(nd\ell + dq)$ edges and thus $\mathcal{A}$ will take time $O(n^{2-\delta}d^{2-\delta})$ on G. This yields an algorithm solving any orthogonal vectors instance in time $O(n^{2-\delta}\operatorname{poly}(d))$, contradicting OVH. □

3.2 Proof of Theorem 3.1

Before we give the proof of Theorem 3.1, we introduce the following useful terminology: For every $1 \le i \le n$, P_i^L is defined as the set of all vertices that lie on one of the the following paths: $\pi(u_i^L, v_i^L)$, $\pi(v_i^L, y^L)$ (*excluding* y^L), or $\pi(v_i^L, w_k^L)$ (*excluding* w_k^L) for some $1 \le k \le d$ such that $l_i[k] = 1$. Similarly, for every $1 \le i \le n$, P_i^R is defined as the set of all vertices that lie on one of the the following paths: $\pi(u_i^R, v_i^R)$, $\pi(v_i^R, y^R)$ (*excluding* y^R), or $\pi(v_i^R, w_k^R)$ (*excluding* w_k^R) for some $1 \le k \le d$ such that $r_i[k] = 1$. We set $V^L := \bigcup_{1 \le i \le n} P_i^R$ (left vertices), $V^R := \bigcup_{1 \le i \le n} P_i^R$ (right vertices), and $V^M := V \setminus (V^L \cup V^R)$. Note that V^M consists on all vertices that lie on $\pi(y^L, y^R)$, $\pi(x^L, y^L)$, $\pi(x^R, y^R)$, $\pi(y^L, w_k^L)$ for some $1 \le k \le d$, $\pi(y^R, w_k^R)$ for some $1 \le k \le d$, or $\pi(w_k^L, w_k^R)$ for some $1 \le k \le d$.

We now state some universal upper and lower bounds on distances in the graph G that hold regardless of whether the orthogonal vectors instance contains an orthogonal pair. Their correctness can readily be verified and we also give rigorous proofs in the appendix.

Lemma 3.4. *For every orthogonal vectors instance,* $\operatorname{dist}_G(s, v_i^L) \le \ell - p$ *for every* $1 \le i \le n$ *and every* $s \in P_i^L$ *and* $\operatorname{dist}_G(v_j^R, t) \le \ell - p$ *for every* $1 \le j \le n$ *and every* $t \in P_j^R$.

Lemma 3.5. *For every orthogonal vectors instance and every pair of vertices* $s, t \in V^M$, $\operatorname{dist}_G(s,t) \le 4\ell - 2p + q$ *and more specifically* $\operatorname{dist}_G(x^L, v) \le 2\ell - p + q$ *and* $\operatorname{dist}_G(v, x^R) \le 2\ell - p + q$ *for every vertex* $v \in V^M$.

Lemma 3.6. *For every orthogonal vectors instance and every pair of vertices s and t, if* $s \in V^L \cup V^M$ *and* $t \in V^L \cup V^M$ *or* $s \in V^R \cup V^M$ *and* $t \in V^R \cup V^M$, *then* $\operatorname{dist}_G(s,t) \le 4\ell - 2p + q$.

Lemma 3.7. *For every orthogonal vectors instance, the following holds in G:*

- $\operatorname{dist}_G(u_i^L, y^L) \ge 3\ell - 2p$ *and* $\operatorname{dist}_G(u_i^R, y^R) \ge 3\ell - 2p$ *for every* $1 \le i \le n$,
- $\operatorname{dist}_G(u_i^L, v_{i'}^L) \ge 3\ell - p$ *and* $\operatorname{dist}_G(u_i^R, v_{i'}^R) \ge 3\ell - p$ *for every* $1 \le i \le n$ *and every* $1 \le i' \le n$ *such that* $i' \ne i$,
- $\operatorname{dist}_G(u_i^L, w_k^L) \ge 2\ell - p$ *for every* $1 \le i \le n$ *and every* $1 \le k \le n$,
- $\operatorname{dist}_G(u_i^R, w_k^R) \ge 2\ell - p$ *for every* $1 \le i \le n$ *and every* $1 \le k \le n$

We finally give the proof of Theorem 3.1. We split up the two cases (containing an orthogonal pair or not) into two pieces, whose proofs follow a similar pattern.

Proposition 3.8. *If the orthogonal vectors instance $\langle L, R\rangle$ contains no orthogonal pair, then $D = 4\ell - 2p + q$.*

Proof. We first show that $D \le 4\ell - 2p + q$, i.e., $\mathrm{dist}_G(s,t) \le 4\ell - 2p + q$ for every pair of vertices $s, t \in V$. Note that by Lemma 3.6 we only have to show that $\mathrm{dist}_G(s,t) \le 4\ell - 2p + q$ whenever $s \in P_i^L$ for some $1 \le i \le n$ and $t \in P_j^R$ for some $1 \le j \le n$. By Lemma 3.4 we have $\mathrm{dist}_G(s, v_i^L) \le \ell - p$ and $\mathrm{dist}_G(v_j^R, t) \le \ell - p$ for such $s \in P_i^L$ and $t \in P_j^R$. Since the orthogonal vectors instance contains no orthogonal pair there is a $1 \le k \le d$ such that both $l_i[k] = r_j[k] = 1$. Thus, our graph G contains both paths $\pi(v_i^L, w_k^L)$ and $\pi(w_k^R, v_j^R)$, each of length ℓ. By the triangle inequality we therefore have

$$\begin{aligned}\mathrm{dist}_G(s,t) \le \underbrace{\mathrm{dist}_G(s, v_i^L)}_{\le \ell - p \text{ (Lemma 3.4)}} &+ \underbrace{\mathrm{dist}_G(v_i^L, w_k^L)}_{\le \ell} + \underbrace{\mathrm{dist}_G(w_k^L, w_k^R)}_{\le q} \\ &+ \underbrace{\mathrm{dist}_G(w_k^R, v_j^R)}_{\le \ell} + \underbrace{\mathrm{dist}_G(v_j^R, t)}_{\le \ell - p \text{ (Lemma 3.4)}} \le 4\ell - 2p + q\,.\end{aligned}$$

It remains to show that $D \ge 4\ell - 2p + q$. We will argue that $\mathrm{dist}_G(u_1^L, u_1^R) \ge 4\ell - 2p + q$. Since the set of vertices $\{y^R, w_1^R, \dots, w_d^R\}$ separates the left part of the graph from the right part of the graph, every path from u_i^L to u_j^R must contain at least one of the following paths entirely: $\pi(y^L, y^R), \pi(w_1^L, w_1^R), \dots, \pi(w_d^L, w_d^R)$. If the shortest path from u_i^L to u_j^R contains the path $\pi(y^L, y^R)$ entirely, then, since $\mathrm{dist}_G(u_i^L, y^L) = 3\ell - 2p$ and $\mathrm{dist}_G(u_j^R, y^R) = 3\ell - 2p$ by Lemma 3.7,

$$\mathrm{dist}_G(u_1^L, u_1^R) = \underbrace{\mathrm{dist}_G(u_1^L, y^L)}_{\ge 3\ell - 2p \text{ by Lemma 3.7}} + \underbrace{|\pi(y^L, y^R)|}_{= p + q} + \underbrace{\mathrm{dist}_G(y^R, u_1^R)}_{\ge 3\ell - 2p \text{ by Lemma 3.7}} \ge 6\ell - 3p + q\,.$$

If the shortest path from u_i^L to u_j^R contains the path $\pi(w_k^L, w_k^R)$ for some $1 \le k \le d$ entirely, then the argument is as follows: By Lemma 3.7 we have $\mathrm{dist}_G(u_1^L, w_k^L) \ge 2\ell - p$ and $\mathrm{dist}_G(w_k^R, u_1^R) \ge 2\ell - p$. We therefore get

$$\mathrm{dist}_G(u_1^L, u_1^R) = \underbrace{\mathrm{dist}_G(u_1^L, w_k^L)}_{\ge 2\ell - p \text{ by Lemma 3.7}} + \underbrace{|\pi(w_k^L, w_k^R)|}_{= q} + \underbrace{\mathrm{dist}_G(w_k^R, u_1^R)}_{\ge 2\ell - p \text{ by Lemma 3.7}} \ge 4\ell - 2p + q\,.$$

□

Proposition 3.9. *If the orthogonal vectors instance $\langle L, R\rangle$ contains an orthogonal pair, then $D = 6\ell - 3p + q$.*

Proof. We first show that $D \ge 6\ell - 3p + q$. Let $l_i \in L$ and $r_j \in R$ denote the orthogonal pair. We will argue that $\mathrm{dist}_G(u_i^L, u_j^R) \ge 6\ell - 3p + q$.

Since the set of vertices $\{y^R, w_1^R, \dots, w_d^R\}$ separates the left part of the graph from the right part of the graph, every path from u_i^L to u_j^R must contain at least one of the following paths entirely: $\pi(y^L, y^R), \pi(w_1^L, w_1^R), \dots, \pi(w_d^L, w_d^R)$. If the shortest path from u_i^L to u_j^R contains the path $\pi(y^L, y^R)$ entirely, then, since $\mathrm{dist}_G(u_i^L, y^L) = 3\ell - 2p$ and $\mathrm{dist}_G(u_j^R, y^R) = 3\ell - 2p$ by Lemma 3.7,

$$\mathrm{dist}_G(u_i^L, u_j^R) = \underbrace{\mathrm{dist}_G(u_i^L, y^L)}_{\geq 3\ell - 2p \text{ by Lemma 3.7}} + \underbrace{|\pi(y^L, y^R)|}_{=p+q} + \underbrace{\mathrm{dist}_G(y^R, u_j^R)}_{\geq 3\ell - 2p \text{ by Lemma 3.7}} \geq 6\ell - 3p + q\,.$$

If the shortest path from u_i^L to u_j^R contains the path $\pi(w_k^L, w_k^R)$ for some $1 \leq k \leq d$ entirely, then the argument is as follows. Since l and r are an orthogonal pair, we have $r_j[k] = 0$ or $l_i[k] = 0$. By symmetry, assume $r_j[k] = 0$, which implies that the path $\pi(w_k^R, v_j^R)$ is not contained in G. Since the shortest path is simple, it is either the case that after the vertex w_k the shortest path contains (a) the subpath $\pi(w_k^R, y^R)$ or (b) the subpath $\pi(w_k^R, v_{j'}^R)$ for some $1 \leq j' \leq n$ such that $j' \neq j$. By Lemma 3.7 we have $\mathrm{dist}_G(u_j^R, w_k^L) \geq 2\ell - p$. In case (a) we additionally use $\mathrm{dist}_G(u_j^R, y^R) = 3\ell - 2p$ from Lemma 3.7 and thus get

$$\begin{aligned}\mathrm{dist}_G(u_i^L, u_j^R) &= \underbrace{\mathrm{dist}_G(u_i^L, w_k^L)}_{\geq 2\ell - p \text{ by Lemma 3.7}} + \underbrace{|\pi(w_k^L, w_k^R)|}_{=q} + \underbrace{|\pi(w_k^R, y^R)|}_{=\ell} + \underbrace{\mathrm{dist}_G(y^R, u_j^R)}_{\geq 3\ell - 2p \text{ by Lemma 3.7}} \\ &\geq 6\ell - 2p + q \geq 6\ell - 3p + q\,.\end{aligned}$$

In case (b) we additionally use $\mathrm{dist}_G(u_j^R, v_{j'}^R) = 3\ell - p$ from Lemma 3.7 and thus get

$$\begin{aligned}\mathrm{dist}_G(u_i^L, u_j^R) &= \underbrace{\mathrm{dist}_G(u_i^L, w_k^L)}_{\geq 2\ell - p \text{ by Lemma 3.7}} + \underbrace{|\pi(w_k^L, w_k^R)|}_{=q} + \underbrace{|\pi(w_k^R, v_{j'})|}_{=\ell} + \underbrace{\mathrm{dist}_G(v_{j'}, u_j^R)}_{\geq 3\ell - p \text{ by Lemma 3.7}} \\ &\geq 6\ell - p + q \geq 6\ell - 3p + q\,.\end{aligned}$$

It remains to show that $D \leq 6\ell - 3p + q$. By Lemma 3.6 and since $4\ell - 2p + q \leq 6\ell - 3p + q$, we only have to show that $\mathrm{dist}_G(s, t) \leq 6\ell - 3p + q$ when $s \in P_i^L$ for some $1 \leq i \leq n$ and $t \in P_j^R$ for some $1 \leq j \leq n$. By Lemma 3.4 we have $\mathrm{dist}_G(s, v_i^L) \leq \ell - p$ and $\mathrm{dist}_G(v_j^R, t) \leq \ell - p$ for such $s \in P_i^L$ and $t \in P_j^R$. By the triangle inequality we therefore have

$$\begin{aligned}\mathrm{dist}_G(s, t) \leq{} & \underbrace{\mathrm{dist}_G(s, v_i^L)}_{\leq \ell - p \text{ (Lemma 3.4)}} + \underbrace{\mathrm{dist}_G(v_i^L, x^L)}_{\leq \ell} + \underbrace{\mathrm{dist}_G(x^L, y^L)}_{\leq \ell - p} + \underbrace{\mathrm{dist}_G(y^L, y^R)}_{\leq p+q} \\ & + \underbrace{\mathrm{dist}_G(y^R, x^R)}_{\leq \ell - p} + \underbrace{\mathrm{dist}_G(x^R, v_j^R)}_{\leq \ell} + \underbrace{\mathrm{dist}_G(v_j^R, t)}_{\leq \ell - p \text{ (Lemma 3.4)}} \leq 6\ell - 3p + q\,.\end{aligned}$$

□

References

[ACK16] Amir Abboud, Keren Censor-Hillel, and Seri Khoury. "Near-Linear Lower Bounds for Distributed Distance Computations, Even in Sparse Networks". In: *International Symposium on Distributed Computing (DISC)*. 2016, pp. 29–42 (cit. on pp. 1, 2).

[CGR16] Massimo Cairo, Roberto Grossi, and Romeo Rizzi. "New Bounds for Approximating Extremal Distances in Undirected Graphs". In: *Symposium on Discrete Algorithms (SODA)*. 2016, pp. 363–376 (cit. on p. 2).

[Che+14] Shiri Chechik, Daniel H. Larkin, Liam Roditty, Grant Schoenebeck, Robert Endre Tarjan, and Virginia Vassilevska Williams. "Better Approximation Algorithms for the Graph Diameter". In: *Symposium on Discrete Algorithms (SODA)*. 2014, pp. 1041–1052 (cit. on p. 2).

[FHW12] Silvio Frischknecht, Stephan Holzer, and Roger Wattenhofer. "Networks cannot compute their diameter in sublinear time". In: *Symposium on Discrete Algorithms (SODA)*. 2012, pp. 1150–1162 (cit. on p. 1).

[Hol+14] Stephan Holzer, David Peleg, Liam Roditty, and Roger Wattenhofer. "Brief Announcement: Distributed 3/2-Approximation of the Diameter". In: *International Symposium on Distributed Computing (DISC)*. 2014, pp. 562–564 (cit. on p. 1).

[IPZ01] Russell Impagliazzo, Ramamohan Paturi, and Francis Zane. "Which Problems Have Strongly Exponential Complexity?" In: *Journal of Computer and System Sciences* 63.4 (2001). Announced at FOCS'98, pp. 512–530 (cit. on p. 2).

[KN97] Eyal Kushilevitz and Noam Nisan. *Communication complexity*. New York, NY, USA: Cambridge University Press, 1997 (cit. on p. 2).

[Pel00] David Peleg. *Distributed Computing: A Locality-Sensitive Approach*. Philadelphia, PA, USA: SIAM, 2000 (cit. on p. 1).

[Raz92] Alexander A. Razborov. "On the Distributional Complexity of Disjointness". In: *Theoretical Computer Science* 106.2 (1992), pp. 385–390 (cit. on p. 2).

[RW13] Liam Roditty and Virginia Vassilevska Williams. "Fast Approximation Algorithms for the Diameter and Radius of Sparse Graphs". In: *Symposium on Theory of Computing (STOC)*. 2013, pp. 515–524 (cit. on pp. 1, 2).

[Wil05] Ryan Williams. "A new algorithm for optimal 2-constraint satisfaction and its implications". In: *Theoretical Computer Science* 348.2-3 (2005). Announced at ICALP'04, pp. 357–365 (cit. on p. 2).

Appendix

In this appendix, we provide rigorous proofs of Lemmas 3.4 to 3.7.

Proof of Lemma 3.4

We only proof the first part of the claim. The second part then follows from symmetric arguments. There are three possibilities for a vertex s to be contained in P_i^L:

1. s lies on the path $\pi(u_i^L, v_i^L)$ (which has length $\ell - p$)
2. s lies on the path $\pi(v_i^L, x^L)$ (which has length ℓ) and $s \neq x^L$

3. s lies on the path $\pi(v_i^L, w_k^L)$ (which has length ℓ) for some $1 \le k \le d$ such that $l_i[k] = 1$ and $s \neq w_k^L$

As $p \le 1$, we have $\ell - 1 \le \ell - p$ and thus in each of the three cases we have $\mathrm{dist}_G(s, v_i^L) \le \ell - p$.

Proof of Lemma 3.5

Consider all simple paths from x^L to x^R in $G[V^M]$: These are $\pi(x^L, y^L), \pi(y^L, y^R), \pi(y^R, x^R)$ as well as $\pi(x^L, y^L), \pi(y^L, w_k^L), \pi(w_k^L, w_k^R), \pi(w_k^R, y^R), \pi(y^R, x^R)$ for every $1 \le k \le d$. These paths have length $2(\ell - p) + p + q \le 4\ell - 2p + q$ and $2(\ell - p) + 2\ell + q = 4\ell - 2p + q$, respectively, in both cases we obtain length $\le 4\ell - 2p + q$. Moreover, each node in V^M is contained in (at least) one of these paths. From these paths, pick P_s, P_t containing s, t. Following P_s and then following the reversed P_t yields a cyclic walk from x^L to itself containing s and t. This walk has length at most $2(4\ell - 2p + q)$, and thus the induced walk from s to t or the one from t to s has length at most $4\ell - 2p + q$. This yields $\mathrm{dist}_G(s, t) \le 4\ell - 2p + q$.

For the stronger bound $\mathrm{dist}_G(x^L, v) \le 2\ell - p + q$ for every vertex $v \in V^M$, consider the following paths:

- $\pi(x^L, y^L), \pi(y^L, w_k^L), \pi(w_k^L, w_k^R)$ for any $1 \le k \le d$ has length $2\ell - p + q$
- $\pi(x^L, y^L), \pi(y^L, y^R), \pi(y^R, w_k^R)$ minus the last vertex, for any $1 \le k \le d$, has length $(\ell - p) + (p + q) + \ell - 1 \le (\ell - p) + (p + q) + (\ell - p) = 2\ell - p + q$.
- $\pi(x^L, y^L), \pi(y^L, y^R), \pi(y^R, x^R)$ has length $2(\ell - p) + (p + q) = 2\ell - p + q$

Since these paths cover all vertices in V^M, each vertex in V^M has distance at most $2\ell - p + q$ to x^L. An analogous argument gives $\mathrm{dist}_G(v, x^R) \le 2\ell - p + q$ for every vertex $v \in V^M$.

Proof of Lemma 3.6

By Lemma 3.5 we have $\mathrm{dist}_G(s, t) \le 4\ell - 2p + q$ for $s, t \in V^M$. Next, consider the case $s \in V^L$ and $t \in V^L$, say $s \in P_i^L$ for some $1 \le i \le n$ and $t \in P_j^L$ for some $1 \le j \le n$. Then we have

$$\begin{aligned}\mathrm{dist}_G(s,t) &\le \underbrace{\mathrm{dist}_G(s, v_i^L)}_{\le \ell - p \text{ by Lemma 3.4}} + \underbrace{\mathrm{dist}_G(v_i^L, x^L)}_{\le |\pi(x^L, v_i^L)| \le \ell} + \underbrace{\mathrm{dist}_G(x^L, v_j^L)}_{\le |\pi(x^L, v_i^L)| \le \ell} + \underbrace{\mathrm{dist}_G(v_j^L, t)}_{\le \ell - p \text{ by Lemma 3.4}} \\ &\le 4\ell - 2p\,.\end{aligned}$$

Finally, consider the case $s \in V^L$ and $t \in V^M$, say $s \in P_i^L$ for some $1 \le i \le n$. By Lemma 3.5 we have $\mathrm{dist}_G(x^L, t) \le 2\ell - p + q$ and thus get

$$\mathrm{dist}_G(s,t) \le \underbrace{\mathrm{dist}_G(s, v_i^L)}_{\le \ell - p \text{ by Lemma 3.4}} + \underbrace{\mathrm{dist}_G(v_i^L, x^L)}_{\le |\pi(x^L, v_i^L)| \le \ell} + \underbrace{\mathrm{dist}_G(x^L, t)}_{\le 2\ell - p + q \text{ by Lemma 3.5}} \le 4\ell - 2p + q\,.$$

The remaining cases require symmetric arguments in which the roles of L and R are exchanged.

Proof of Lemma 3.7

Let $1 \leq i \leq n$. Observe that all simple paths of length at most $3\ell - p$ starting at vertex u_i^L and ending at an exterior vertex must be of the following form (where some of these paths are of length at most $3\ell - p$ only if $p = 0$, and some only if $p = 0$ and $q = 0$):

1. $\pi(u_i^L, v_i^L)$ (of length $\ell - p$),
2. $\pi(u_i^L, v_i^L), \pi(v_i^L, x^L)$ (of length $2\ell - p$),
3. $\pi(u_i^L, v_i^L), \pi(v_i^L, x^L), \pi(x^L, y^L)$ (of length $3\ell - 2p$),
4. $\pi(u_i^L, v_i^L), \pi(v_i^L, x^L), \pi(x^L, y^L), \pi(y^L, y^R)$ (of length $3\ell - p + q$),
5. $\pi(u_i^L, v_i^L), \pi(v_i^L, x^L), \pi(x^L, v_{i'}^L)$ for some $1 \leq i' \leq n$ (of length $3\ell - p$),
6. $\pi(u_i^L, v_i^L), \pi(v_i^L, w_k^L)$ for some $1 \leq k \leq d$ such that $l_i[k] = 1$ (of length $2\ell - p$),
7. $\pi(u_i^L, v_i^L), \pi(v_i^L, w_k^L), \pi(w_k^L, y^L)$ for some $1 \leq k \leq d$ such that $l_i[k] = 1$ (of length $3\ell - p$),
8. $\pi(u_i^L, v_i^L), \pi(v_i^L, w_k^L), \pi(w_k^L, y^L), \pi(y^L, y^R)$ for some $1 \leq k \leq d$ such that $l_i[k] = 1$ (of length $3\ell + q$),
9. $\pi(u_i^L, v_i^L), \pi(v_i^L, w_k^L), \pi(w_k^L, v_{i'}^L)$ for some $1 \leq k \leq d$ and some $1 \leq i' \leq n$ with $i' \neq i$ such that $l_i[k] = 1$ and $l_{i'}[k] = 1$ (of length $3\ell - p$),
10. $\pi(u_i^L, v_i^L), \pi(v_i^L, w_k^L), \pi(w_k^L, w_k^R)$ for some $1 \leq k \leq d$ (of length $2\ell - p + q$),
11. $\pi(u_i^L, v_i^L), \pi(v_i^L, w_k^L), \pi(w_k^L, w_k^R), \pi(w_k^R, y^R)$ for some $1 \leq k \leq d$ (of length $3\ell - p + q$),
12. $\pi(u_i^L, v_i^L), \pi(v_i^L, w_k^L), \pi(w_k^L, w_k^R), \pi(w_k^R, y^R), \pi(y^R, y^L)$ for some $1 \leq k \leq d$ (of length $3\ell + 2q$), or
13. $\pi(u_i^L, v_i^L), \pi(v_i^L, w_k^L), \pi(w_k^L, w_k^R), \pi(w_k^R, v_{j'}^R)$ for some $1 \leq k \leq d$ and some $1 \leq j' \leq n$ such that $l_i[k] = 1$ and $r_{j'}[k] = 1$ (of length $3\ell - p + q$).

It follows that $\mathrm{dist}_G(u_i^L, v_{i'}^L) = 3\ell - p$ (as all paths of length at most $3\ell - 2p$ ending at $v_{i'}^L$ have length $3\ell - p$), $\mathrm{dist}_G(u_i^L, y^L) = 3\ell - 2p$ (as the shortest path of length at most $3\ell - 2p$ ending at y^L has length $3\ell - 2p$), and $\mathrm{dist}_G(u_i^L, w_k^L) \geq 2\ell - p$ for every $1 \leq k \leq d$ (as the only possible path of length at most $3\ell - 2p$ ending at w_k^L has length $2\ell - p$). A symmetric argument gives $\mathrm{dist}_G(u_i^R, v_{i'}^R) \geq 3\ell - p$, $\mathrm{dist}_G(u_i^R, y^R) \geq 3\ell - 2p$, and $\mathrm{dist}_G(u_i^R, w_k^R) \geq 2\ell - p$ for every $1 \leq k \leq d$.