Approximation algorithm for Diameter, or lack thereof

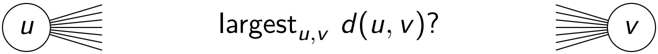
Édouard Bonnet

ENS Lyon, LIP

July 19th, 2021, LIS seminar

Diameter

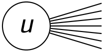
diam(G) = largest distance between a pair of vertices of G



- In weighted graphs, no better known than APSP
- In unweighted graphs, solvable in $\tilde{O}(n^{\omega})$

Diameter

diam(G) = largest distance between a pair of vertices of G

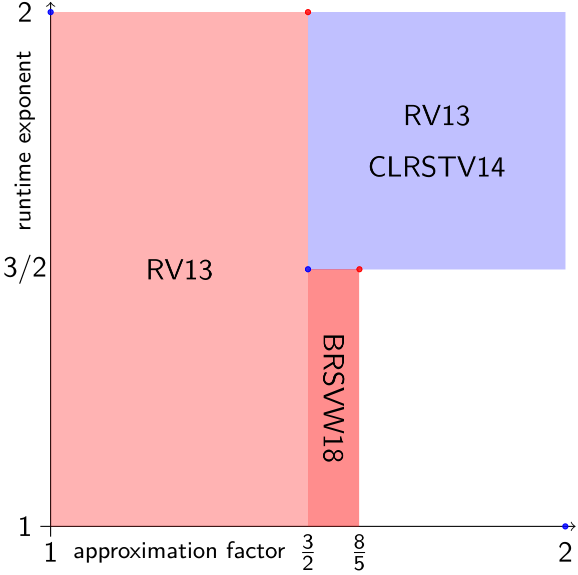


largest$_{u,v}$ $d(u, v)$?

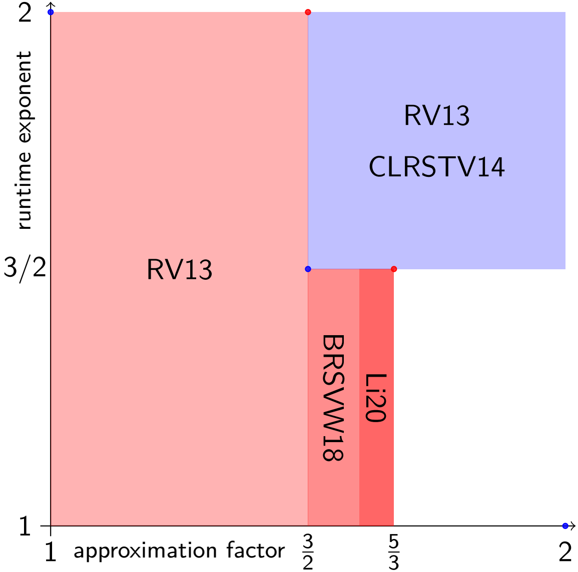
- In weighted graphs, no better known than APSP
- In unweighted graphs, solvable in $\tilde{O}(n^{\omega})$

Scope of the talk: Time vs Approximation trade-offs
Pareto front of (x, y), $\exists$ x-approximation running in time $\tilde{O}(|G|^{y})$

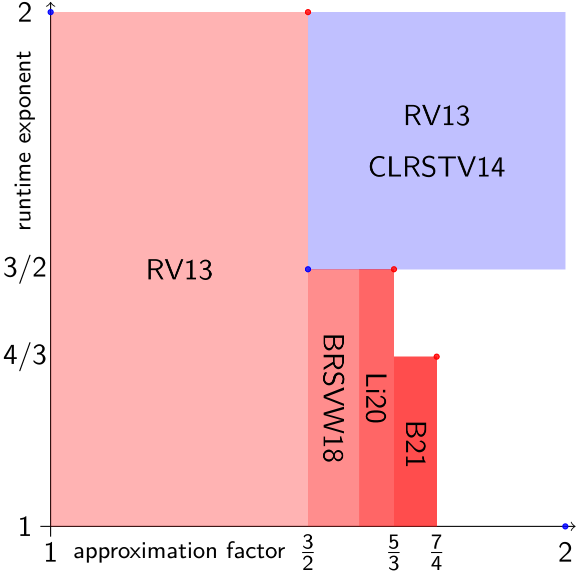
Directed (un)weighted DIAMETER



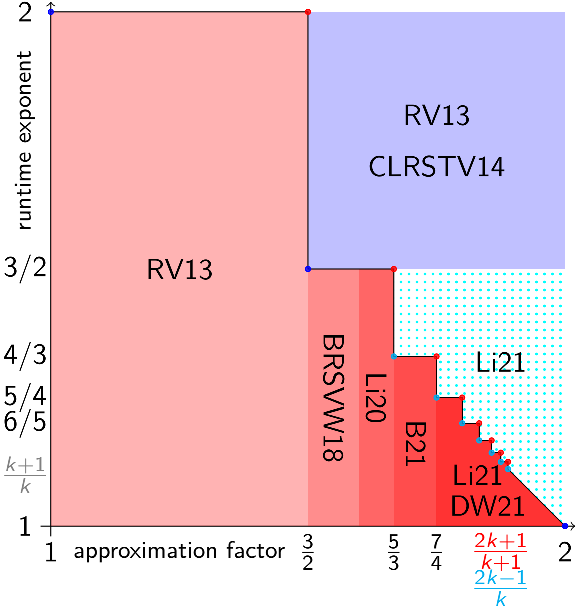
Directed (un)weighted DIAMETER



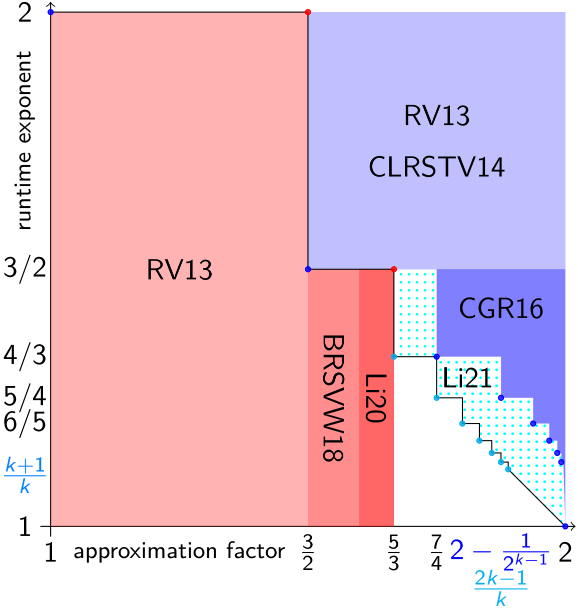
Directed (un)weighted DIAMETER



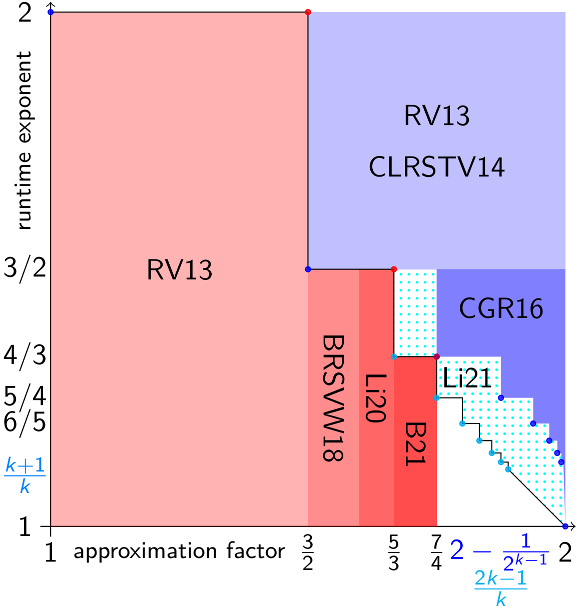
Directed (un)weighted DIAMETER



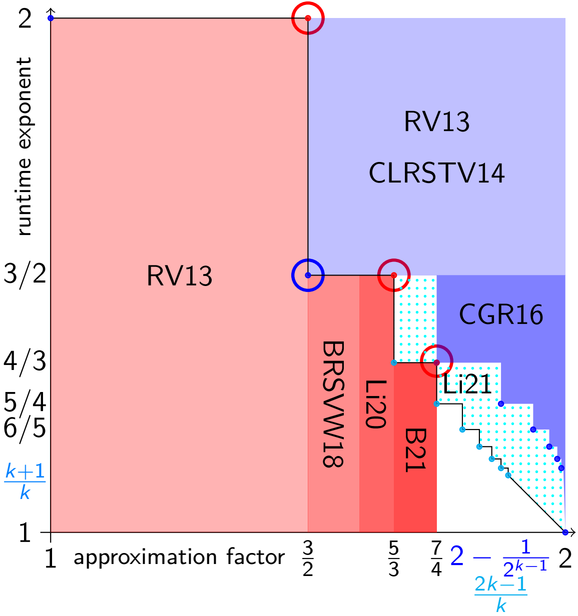
Undirected unweighted DIAMETER



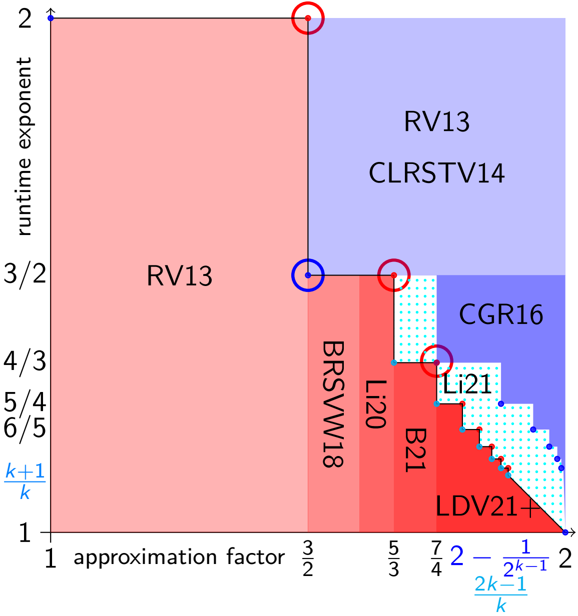
Undirected unweighted DIAMETER



Undirected unweighted DIAMETER

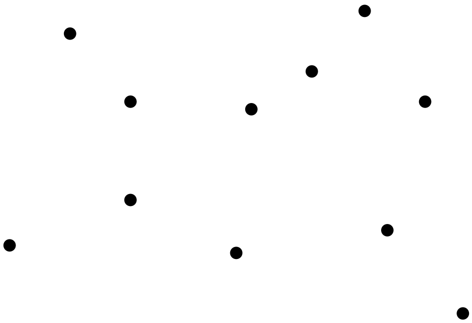


Undirected unweighted DIAMETER



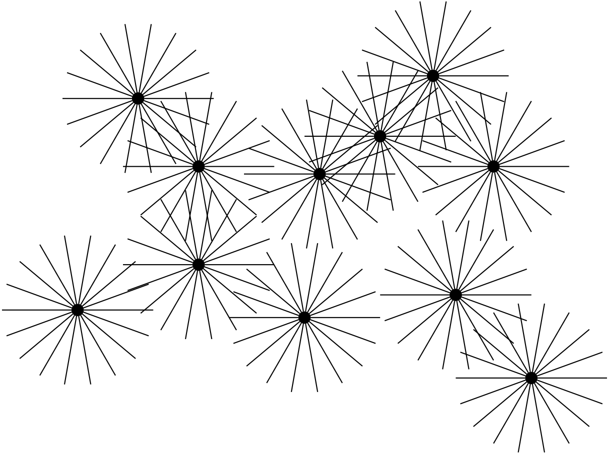
Algorithms

3/2-approximation algorithm in time $\tilde{O}(m\sqrt{n})$



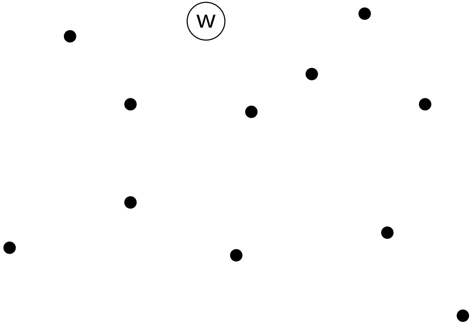
Sample $100\sqrt{n}\log n$ vertices uniformly at random $\rightarrow S$

3/2-approximation algorithm in time $\tilde{O}(m\sqrt{n})$



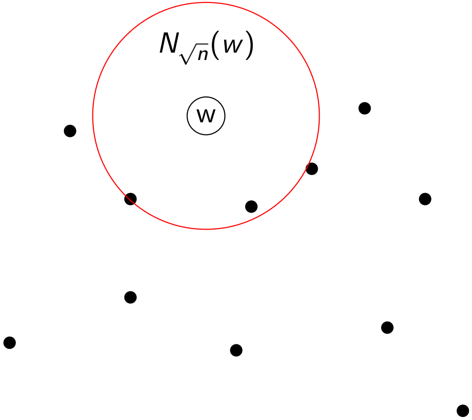
Run Dijkstra from each vertex of S

3/2-approximation algorithm in time $\tilde{O}(m\sqrt{n})$



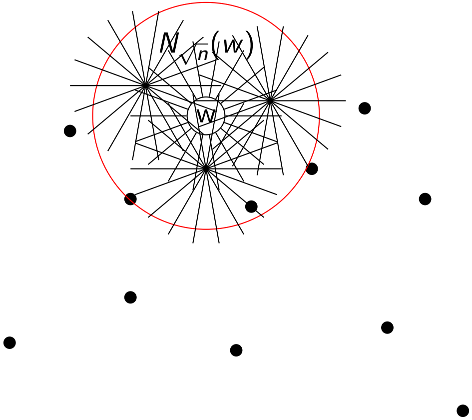
Let w be the furthest vertex to S

3/2-approximation algorithm in time $\tilde{O}(m\sqrt{n})$



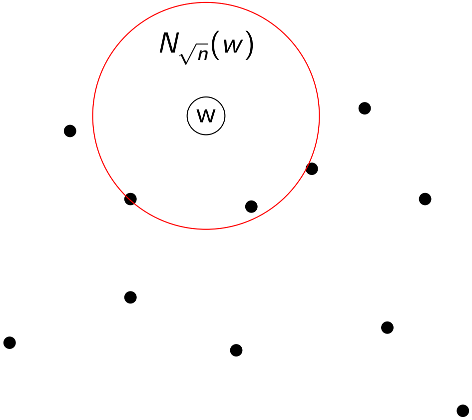
Compute $N_{\sqrt{n}}(w)$: the set of $\sqrt{n}$ closest vertices from w

3/2-approximation algorithm in time $\tilde{O}(m\sqrt{n})$



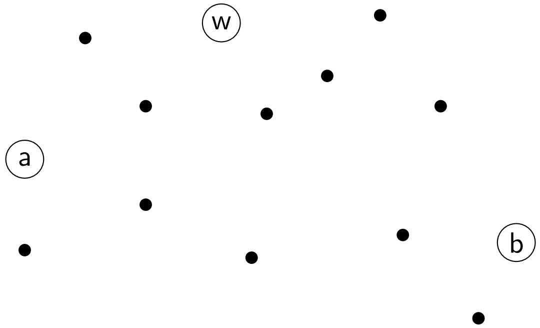
Run Dijkstra from each vertex of $N_{\sqrt{n}}(w)$

3/2-approximation algorithm in time $\tilde{O}(m\sqrt{n})$



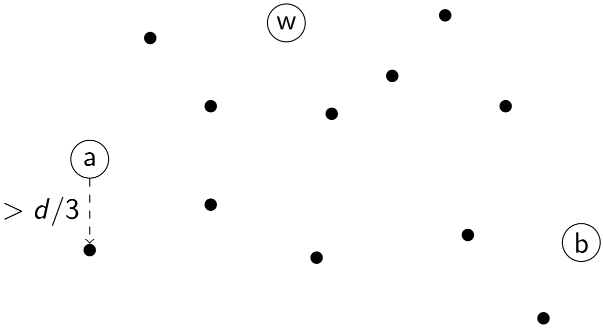
Output $\max\{\max(d(x,y), d(y,x)) \mid x \in V(G), y \in S \cup N_{\sqrt{n}}(w)\}$

Correctness of the 3/2 approximation factor



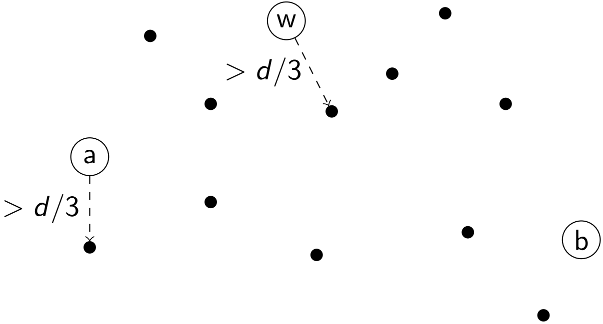
Say a and b realizes the diameter $d = \mathrm{diam}(G) = d(a, b)$

Correctness of the 3/2 approximation factor



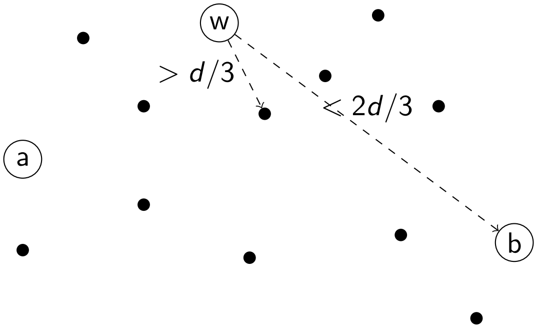
We can assume that $d(a, S) > d/3$. Why?

Correctness of the 3/2 approximation factor



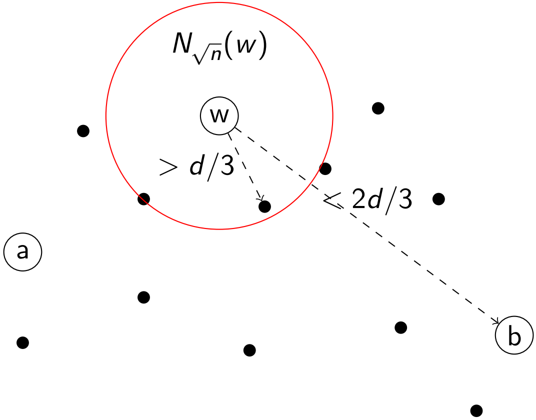
This implies that $d(w, S) > d/3$.

Correctness of the 3/2 approximation factor



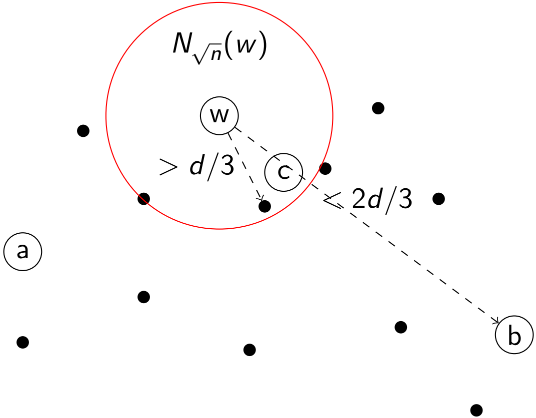
Similarly we can assume that $d(w, b) < 2d/3$.

Correctness of the 3/2 approximation factor



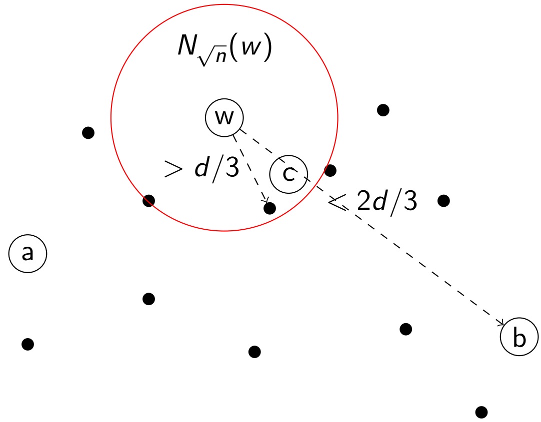
With high probability $N_{\sqrt{n}}(w)$ intersects S

Correctness of the 3/2 approximation factor



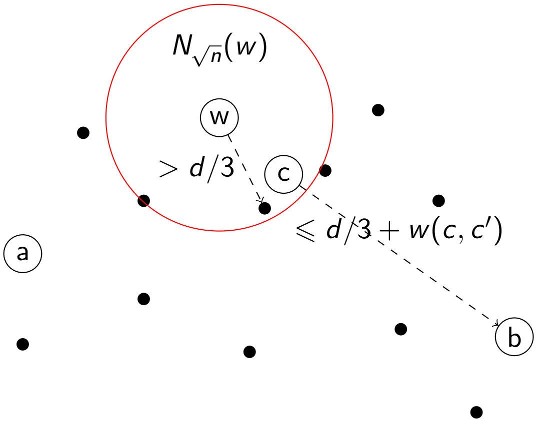
So there is $c \in N_{\sqrt{n}}(w)$ along a shortest path $w - cc' - b$...

Correctness of the 3/2 approximation factor



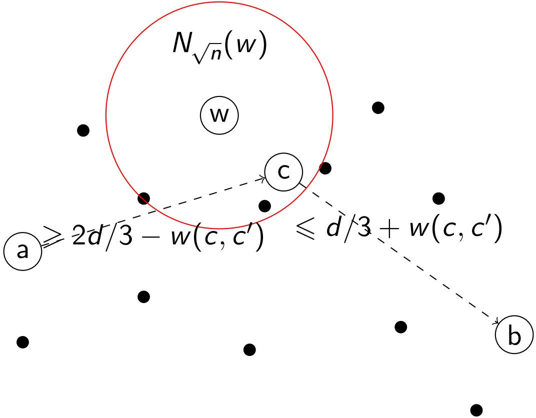
...with $d(w, c) \leqslant d/3$ and $d(w, c') > d/3$

Correctness of the 3/2 approximation factor



Thus $d(w, c) > d/3 - w(c, c')$, hence $d(c, b) \leqslant d/3 + w(c, c')$

Correctness of the 3/2 approximation factor



Finally $d(a,c) \geqslant 2d/3 - w(c,c')$

Lower bounds

SETH

$\forall k, \exists \varepsilon > 0$, no classical algorithm solves n-var k-SAT in $(2 - \varepsilon)^n$

In 1999, Impagliazzo and Paturi introduce ETH and mention a stronger version of it in their conclusion

$$\text{SETH} \Rightarrow \text{ETH} \Rightarrow \text{P} \neq \text{NP}$$

- ETH and SETH are then mainly used for NP-hard problems

SETH

$\forall k, \exists \varepsilon > 0$, no classical algorithm solves n-var k-SAT in $(2-\varepsilon)^n$

In 1999, Impagliazzo and Paturi introduce ETH and mention a stronger version of it in their conclusion

$$\text{SETH} \Rightarrow \text{ETH} \Rightarrow \text{P} \neq \text{NP}$$

- ETH and SETH are then mainly used for NP-hard problems
- In 2005, SETH is used for the first time for a problem in P

ORTHOGONAL VECTORS,

SETH

$\forall k, \exists \varepsilon > 0$, no classical algorithm solves n-var k-Sat in $(2-\varepsilon)^n$

In 1999, Impagliazzo and Paturi introduce ETH and mention a stronger version of it in their conclusion

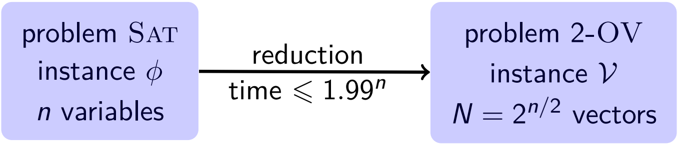
$$\text{SETH} \Rightarrow \text{ETH} \Rightarrow \text{P} \neq \text{NP}$$

- ETH and SETH are then mainly used for NP-hard problems
- In 2005, SETH is used for the first time for a problem in P
- 2014-, dozens of papers show SETH-hardness of problems in P

Orthogonal Vectors, Diameter, Fréchet Distance, Edit Distance, Longest Common Subsequence, Furthest Pair, dynamic problems, problems from Machine Learning, Model Checking, Language Theory etc.

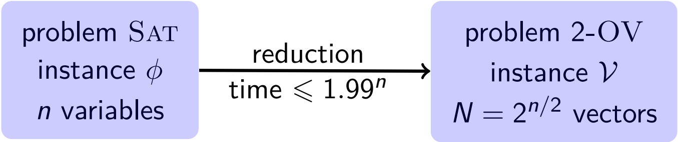
Reduction from SAT to a problem in P

2-ORTHOGONAL VECTORS (2-OV):
Are there two orthogonal vectors in a given set of N $0,1$-vectors?



Reduction from Sat to a problem in P

2-Orthogonal Vectors (2-OV):
Are there two orthogonal vectors in a given set of N $0,1$-vectors?



$\rightarrow$ Solving 2-OV in $N^{1.99}$ solves Sat $1.99^n + 2^{\frac{1.99n}{2}}$, refuting SETH

SAT $\to$ 2-ORTHOGONAL VECTORS [W05]

arbitrary equipartition of X: $x_1, x_2, \ldots, x_{\frac{n}{2}}$, $x_{\frac{n}{2}+1}, x_{\frac{n}{2}+2}, \ldots, x_n$

Find an assignment

- A of the red variables and
- B of the blue variables

such that all the clauses are satisfied by A or by B

SAT $\to$ 2-ORTHOGONAL VECTORS [W05]

arbitrary equipartition of X: $x_1, x_2, \ldots, x_{\frac{n}{2}}, x_{\frac{n}{2}+1}, x_{\frac{n}{2}+2}, \ldots, x_n$

Find an assignment

- A of the red variables and
- B of the blue variables

such that all the clauses are satisfied by A or by B

	R	B	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
A_1										
A_2										
A_3										
A_4										
B_1										
B_2										
B_3										
B_4										

SAT → 2-Orthogonal Vectors [W05]

arbitrary equipartition of X: $x_1, x_2, \ldots, x_{\frac{n}{2}}$, $x_{\frac{n}{2}+1}, x_{\frac{n}{2}+2}, \ldots, x_n$

Find an assignment

- A of the red variables and
- B of the blue variables

such that all the clauses are satisfied by A or by B

	R	B	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
A_1	1	0								
A_2										
A_3										
A_4										
B_1										
B_2										
B_3										
B_4										

A_1 assigns red variables

SAT → 2-ORTHOGONAL VECTORS [W05]

arbitrary equipartition of X: $x_1, x_2, \ldots, x_{\frac{n}{2}}$, $x_{\frac{n}{2}+1}, x_{\frac{n}{2}+2}, \ldots, x_n$

Find an assignment

- A of the red variables and
- B of the blue variables

such that all the clauses are satisfied by A or by B

	R	B	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
A_1	1	0	1							
A_2										
A_3										
A_4										
B_1										
B_2										
B_3										
B_4										

A_1 does *not* satisfy C_1

SAT → 2-ORTHOGONAL VECTORS [W05]

arbitrary equipartition of X: $x_1, x_2, \ldots, x_{\frac{n}{2}}$, $x_{\frac{n}{2}+1}, x_{\frac{n}{2}+2}, \ldots, x_n$

Find an assignment

- A of the red variables and
- B of the blue variables

such that all the clauses are satisfied by A or by B

	R	B	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
A_1	1	0	1	0						
A_2										
A_3										
A_4										
B_1										
B_2										
B_3										
B_4										

A_1 satisfies C_2

SAT → 2-ORTHOGONAL VECTORS [W05]

arbitrary equipartition of X: $x_1, x_2, \ldots, x_{\frac{n}{2}}$, $x_{\frac{n}{2}+1}, x_{\frac{n}{2}+2}, \ldots, x_n$

Find an assignment

- A of the red variables and
- B of the blue variables

such that all the clauses are satisfied by A or by B

	R	B	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
A_1	1	0	1	0	0	1	0	0	1	0
A_2										
A_3										
A_4										
B_1										
B_2										
B_3										
B_4										

first vector $(1, 0, 1, 0, 0, 1, 1, 0, 1, 0)$

Sat → 2-Orthogonal Vectors [W05]

arbitrary equipartition of X: $x_1, x_2, \ldots, x_{\frac{n}{2}}$, $x_{\frac{n}{2}+1}, x_{\frac{n}{2}+2}, \ldots, x_n$

Find an assignment

- A of the red variables and
- B of the blue variables

such that all the clauses are satisfied by A or by B

	R	B	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
A_1	1	0	1	0	0	1	0	0	1	0
A_2	1	0	0	0	0	1	1	1	0	1
A_3	1	0	0	1	0	1	0	0	1	1
A_4	1	0	0	0	1	1	0	1	1	1
B_1	0	1	1	1	0	0	1	1	1	0
B_2	0	1	0	1	0	1	0	1	0	0
B_3	0	1	1	1	1	1	0	0	0	1
B_4	0	1	0	1	0	0	1	0	0	1

SAT → 2-ORTHOGONAL VECTORS [W05]

arbitrary equipartition of X: $x_1, x_2, \ldots, x_{\frac{n}{2}}$, $x_{\frac{n}{2}+1}, x_{\frac{n}{2}+2}, \ldots, x_n$

Find an assignment

- A of the red variables and
- B of the blue variables

such that all the clauses are satisfied by A or by B

	R	B	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
A_1	1	**0**	1	**0**	**0**	1	**0**	**0**	1	**0**
A_2	1	0	0	0	0	1	1	1	0	1
A_3	1	0	0	1	0	1	0	0	1	1
A_4	1	0	0	0	1	1	0	1	1	1
B_1	0	1	1	1	0	0	1	1	1	0
B_2	0	1	0	1	0	1	0	1	0	0
B_3	0	1	1	1	1	1	0	0	0	1
B_4	**0**	1	**0**	1	**0**	**0**	1	**0**	**0**	1

Consequence for 2-Orthogonal Vectors

From a Sat-instance on n variables and m clauses, we created $N := 2^{\frac{n}{2}+1}$ vectors in dimension $d := m + 2$

Consequence for 2-Orthogonal Vectors

From a Sat-instance on n variables and m clauses, we created $N := 2^{\frac{n}{2}+1}$ vectors in dimension $d := m + 2$

An algorithm solving 2-OV in time $2^{o(d)} N^{2-\varepsilon}$ would solve Sat in $2^{o(m)} 2^{n(1-\varepsilon/2)} \rightarrow$ breaking SETH

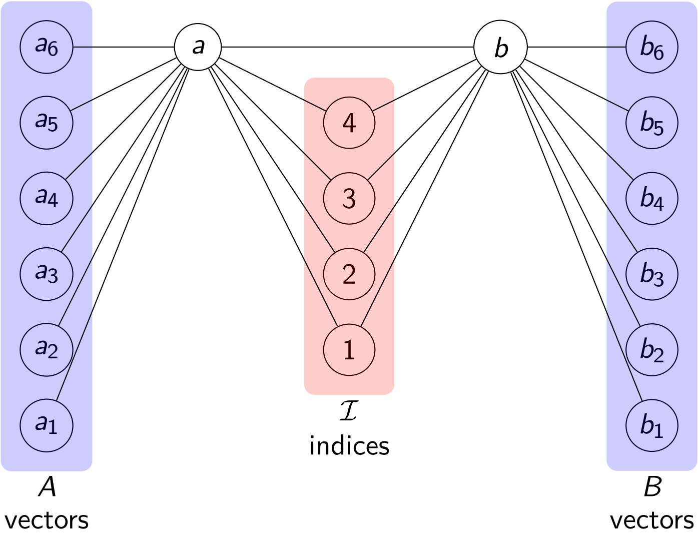
Consequence for 2-Orthogonal Vectors

From a Sat-instance on n variables and m clauses, we created $N := 2^{\frac{n}{2}+1}$ vectors in dimension $d := m + 2$

An algorithm solving 2-OV in time $2^{o(d)} N^{2-\varepsilon}$
would solve Sat in $2^{o(m)} 2^{n(1-\varepsilon/2)} \to$ breaking SETH

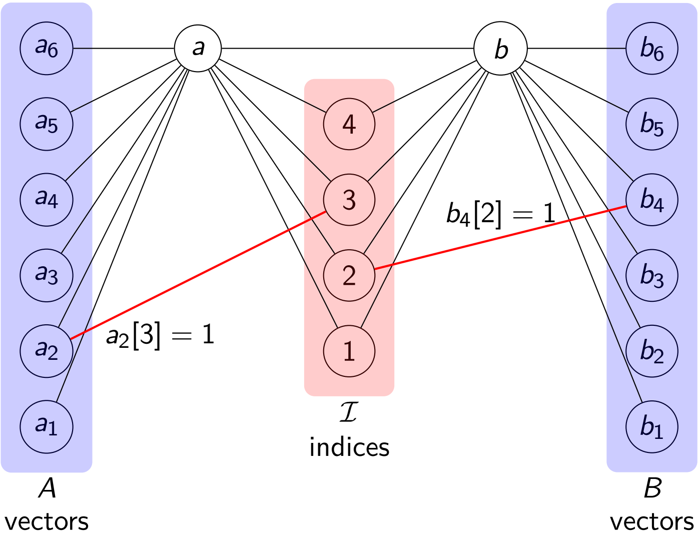
Most useful consequence here:
$N^{2-o(1)}$-time is required even if $d = \log^{O(1)} N$
Same for k-OV and $N^{k-o(1)}$-time

2-Orthogonal Vectors → Diameter [RV13]



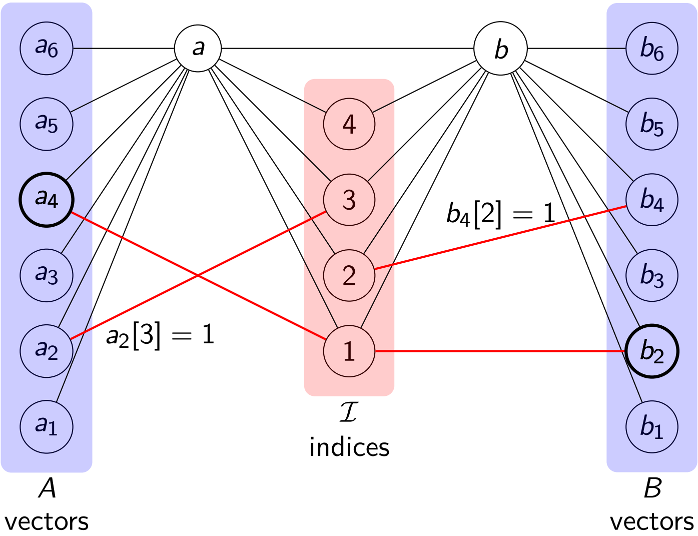
So far, all the pairs but of $A \times B$ are at distance $\leqslant 2$

2-Orthogonal Vectors → Diameter [RV13]



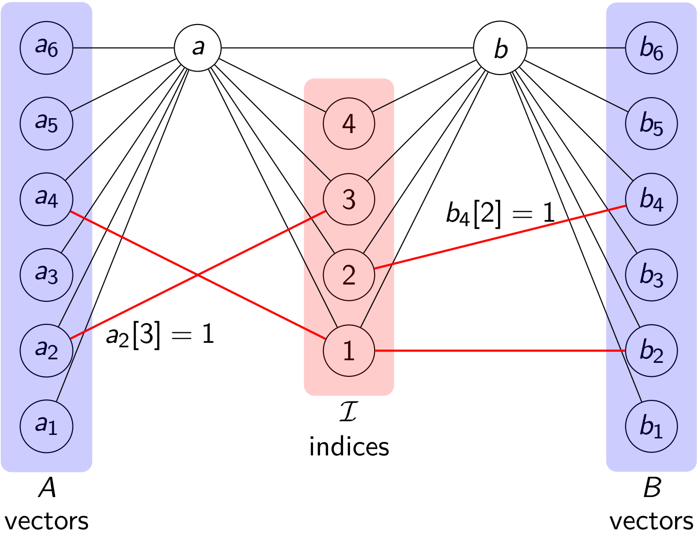
we put an edge between vector v and index i iff $v[i] = 1$

2-Orthogonal Vectors → Diameter [RV13]



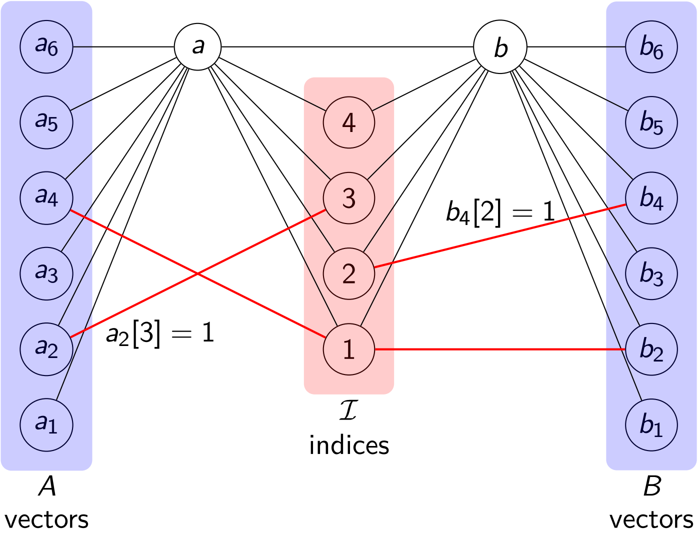
A pair (a4,b2) is at distance 2 $\Leftrightarrow \langle a_4, b_2 \rangle \neq 0$

2-Orthogonal Vectors → Diameter [RV13]



$\text{diam}(G) = 3 \Leftrightarrow \exists(a_i, b_j)$ at distance 3 $\Leftrightarrow$ orthogonal pair

2-Orthogonal Vectors → Diameter [RV13]



If no orthogonal pair, $\text{diam}(G) = 2$

3 vs 5 undirected Diameter

Theorem (Li '20)

Approximating sparse undirected unweighted DIAMETER *within factor better than* $\frac{5}{3}$ *requires time* $n^{\frac{3}{2}-o(1)}$*, unless SETH fails.*

3 vs 5 undirected Diameter

Theorem (Li '20)

Approximating sparse undirected unweighted DIAMETER *within factor better than* $\frac{5}{3}$ *requires time* $n^{\frac{3}{2}-o(1)}$*, unless SETH fails.*

Plan: hardness of 3 vs 5 DIAMETER from N-vector 3-OV to $O(N^2)$-vertex $\tilde{O}(N^2)$-edge DIAMETER-instances.

3 vs 5 undirected Diameter

Theorem (Li '20)

Approximating sparse undirected unweighted DIAMETER *within factor better than* $\frac{5}{3}$ *requires time* $n^{\frac{3}{2}-o(1)}$, *unless SETH fails.*

Plan: hardness of 3 vs 5 DIAMETER from N-vector 3-OV to $O(N^2)$-vertex $\tilde{O}(N^2)$-edge DIAMETER-instances.

2 vs 3 hardness from 2-OV [RV13]

3 vs 5 undirected Diameter

Theorem (Li '20)

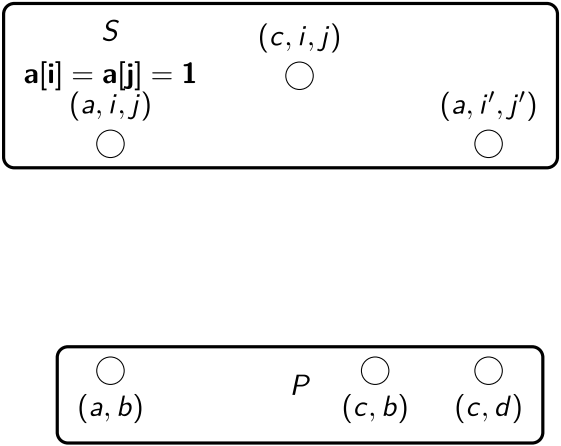
Approximating sparse undirected unweighted $\textsc{Diameter}$ *within factor better than* $\frac{5}{3}$ *requires time* $n^{\frac{3}{2}-o(1)}$*, unless SETH fails.*

Plan: hardness of 3 vs 5 $\textsc{Diameter}$ from N-vector 3-OV to $O(N^2)$-vertex $\tilde{O}(N^2)$-edge $\textsc{Diameter}$-instances.

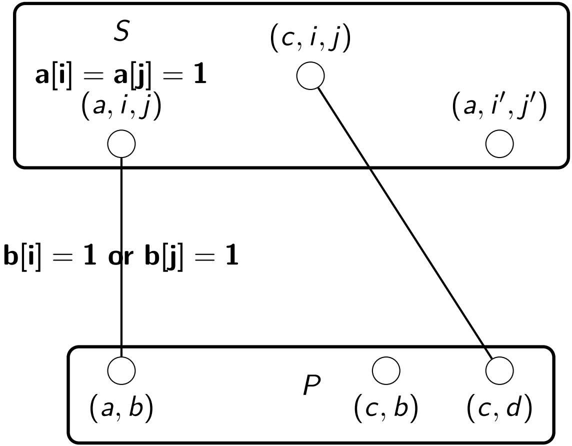
2 vs 3 hardness from 2-OV [RV13]

Vectors $a, b, c, \ldots$, Indices $i, j, k, \ldots$,
$\mathrm{ind}(a, b, c) = i$, with $a[i] = b[i] = c[i] = 1$ (exists if a, b, c not $\perp$)
index i contradicts a, b, c $\perp$

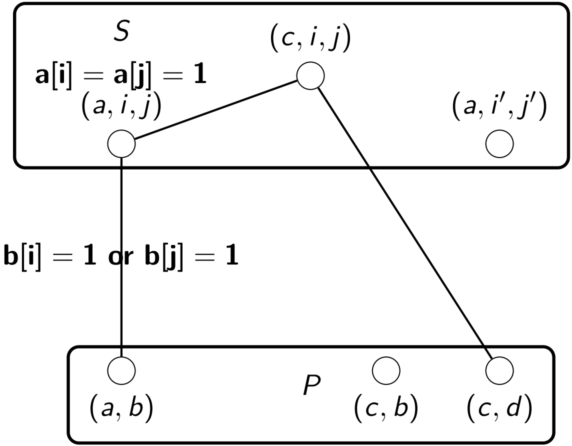
3-Orthogonal Vectors → 3 vs 5 Diameter [Li20]



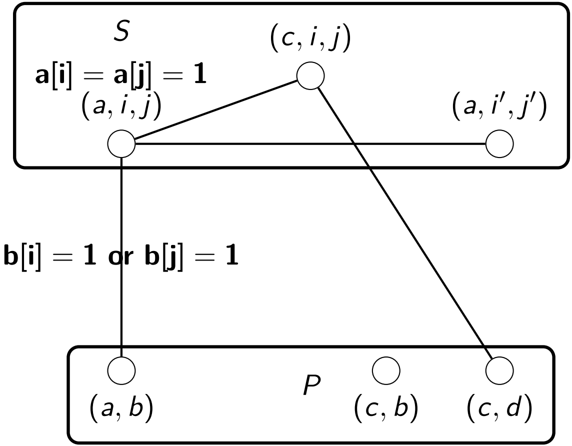
3-Orthogonal Vectors → 3 vs 5 Diameter [Li20]



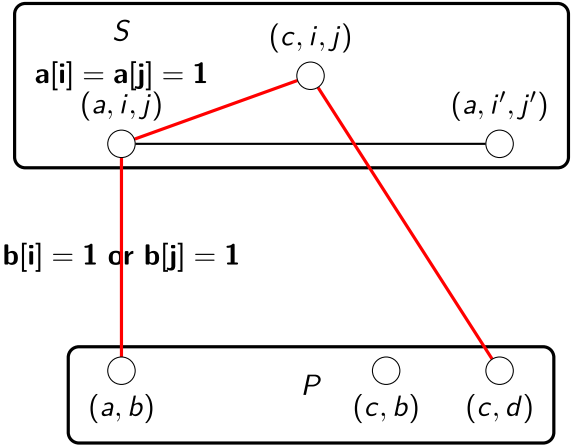
3-Orthogonal Vectors → 3 vs 5 Diameter [Li20]



3-Orthogonal Vectors → 3 vs 5 Diameter [Li20]

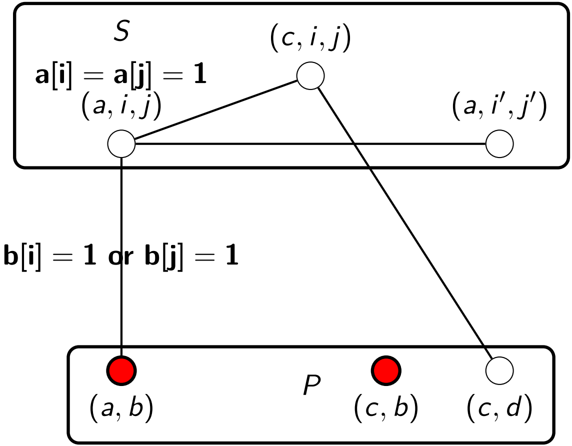


3-Orthogonal Vectors → 3 vs 5 Diameter [Li20]



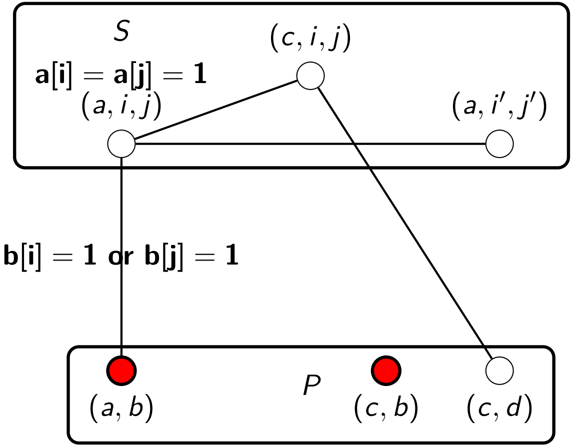
No orthogonal triple ⇒ $\text{diam}(G) = 3$,
$i = \text{ind}(a, b, c), \; j = \text{ind}(a, c, d)$

3-Orthogonal Vectors → 3 vs 5 Diameter [Li20]



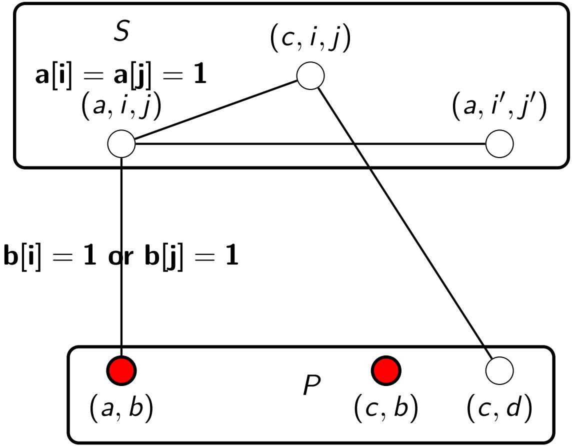
Orthogonal triple (a,b,c) $\Rightarrow d((a, b), (b, c)) = 5$,
$(a, b) - (a, i, j) - (x, i', j') - (c, i'', j'') - (c, b)$

3-Orthogonal Vectors → 3 vs 5 Diameter [Li20]



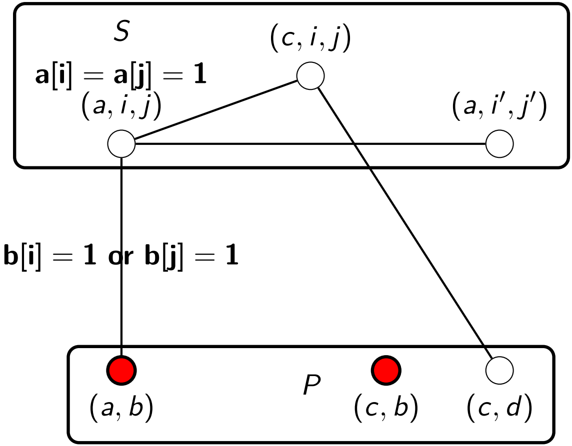
Orthogonal triple (a,b,c) $\Rightarrow d((a,b),(b,c)) = 5$,
$(a,b) - (a,i,j) - (\mathbf{a},i',j') - (c,i'',j'') - (c,b)$

3-Orthogonal Vectors → 3 vs 5 Diameter [Li20]



Orthogonal triple (a,b,c) $\Rightarrow d((a, b), (b, c)) = 5$,
$(a, b) - (a, i, j) - (a, i', j') - (c, \mathbf{i'}, \mathbf{j'}) - (c, b)$

3-Orthogonal Vectors → 3 vs 5 Diameter [Li20]



Orthogonal triple (a,b,c) $\Rightarrow d((a,b),(b,c)) = 5$,
i' or j' contradicts $a, b, c \perp$

4 vs 7 undirected Diameter

Theorem (B. '21)

Approximating sparse undirected unweighted DIAMETER *within factor better than* $\frac{7}{4}$ *requires time* $n^{\frac{4}{3}-o(1)}$*, unless SETH fails.*

4 vs 7 undirected Diameter

Theorem (B. '21)

Approximating sparse undirected unweighted DIAMETER *within factor better than* $\frac{7}{4}$ *requires time* $n^{\frac{4}{3}-o(1)}$*, unless SETH fails.*

Plan: hardness of 4 vs 7 DIAMETER from N-vector 4-OV to $O(N^3)$-vertex $\tilde{O}(N^3)$-edge DIAMETER-instances.

4 vs 7 undirected Diameter

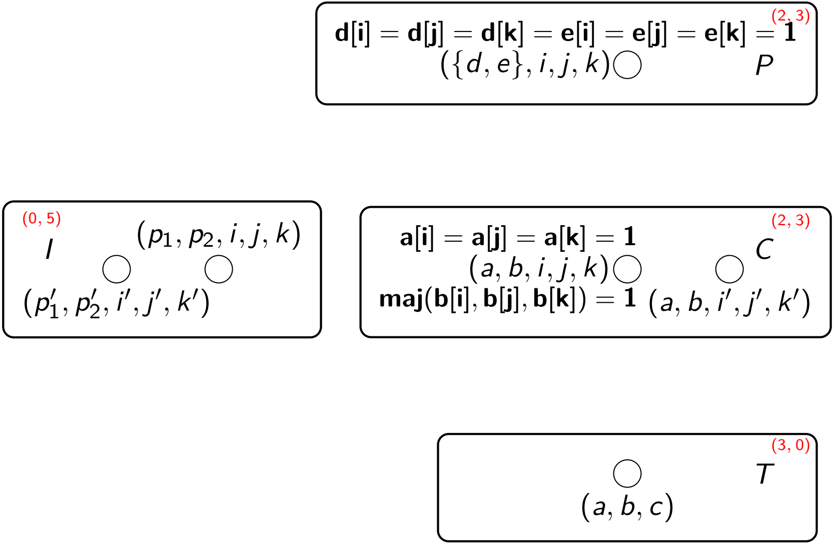
Theorem (B. '21)

Approximating sparse undirected unweighted DIAMETER *within factor better than* $\frac{7}{4}$ *requires time* $n^{\frac{4}{3}-o(1)}$*, unless SETH fails.*

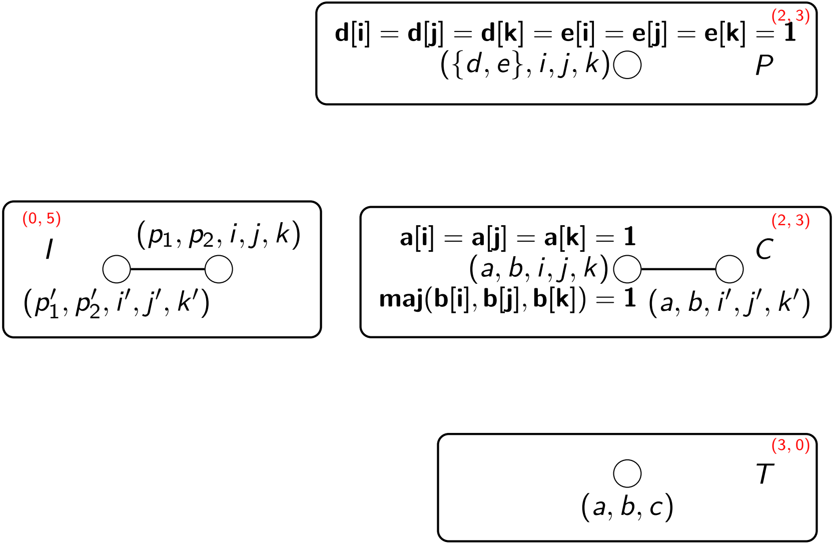
Plan: hardness of 4 vs 7 DIAMETER from N-vector 4-OV to $O(N^3)$-vertex $\tilde{O}(N^3)$-edge DIAMETER-instances.

2 vs 3 hardness from 2-OV [RV13]
3 vs 5 hardness from 3-OV [Li20]

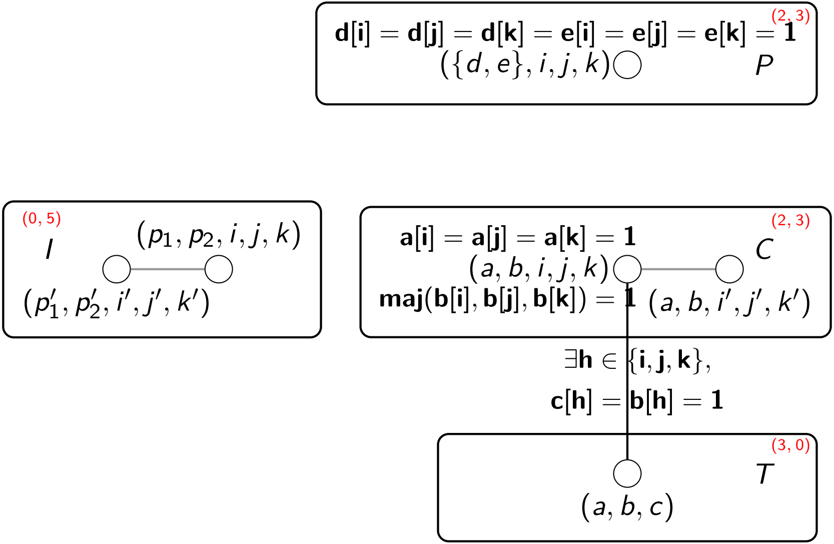
Construction *with weights*



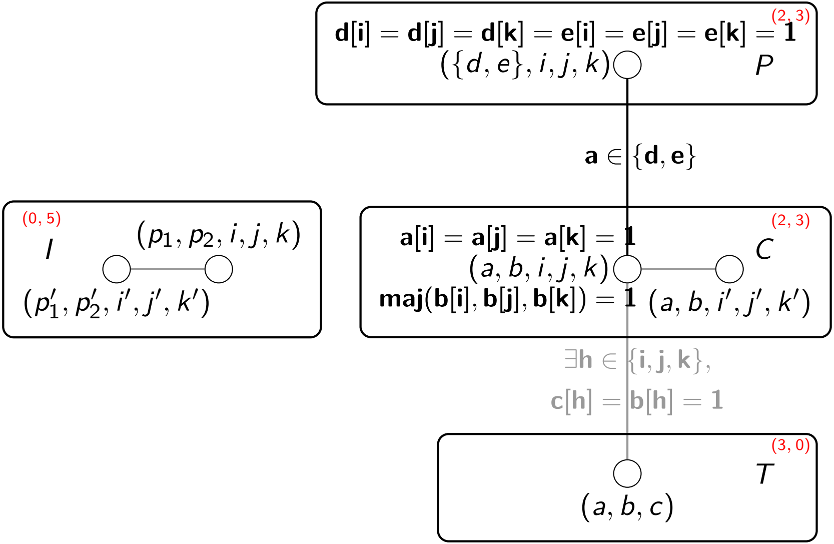
Construction *with weights*



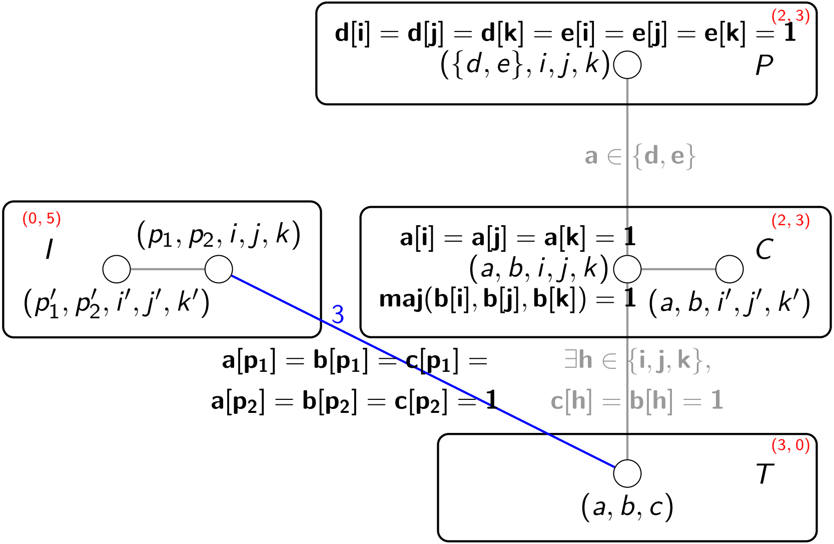
Construction *with weights*



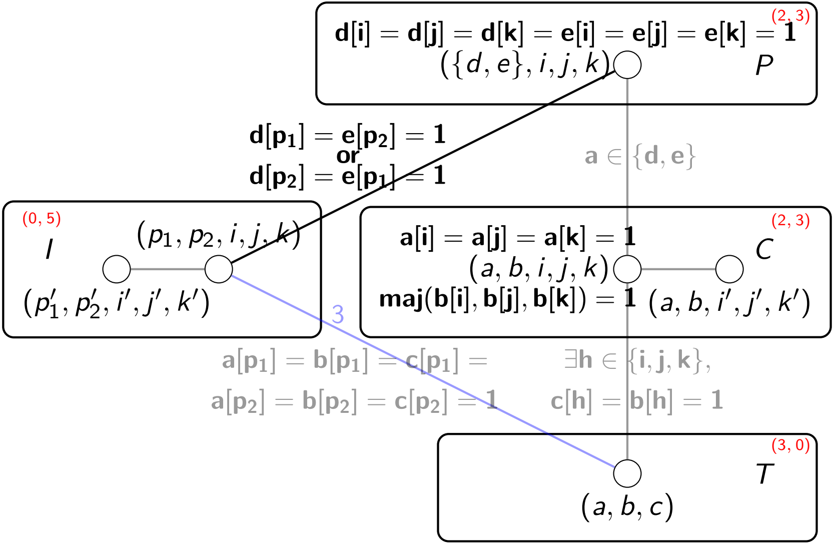
Construction *with weights*



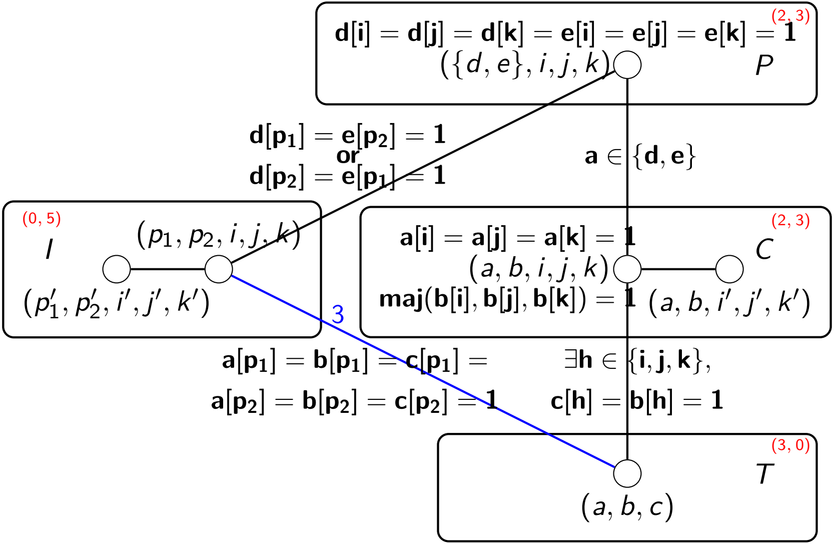
Construction *with weights*



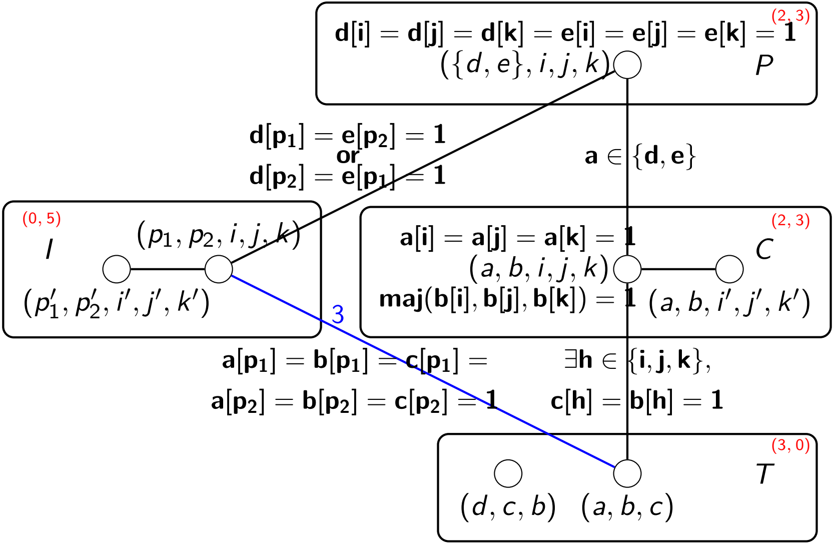
Construction *with weights*



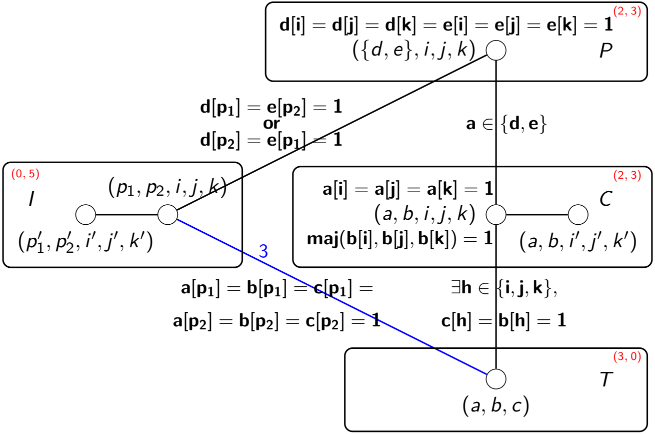
Construction *with weights*



Construction *with weights*

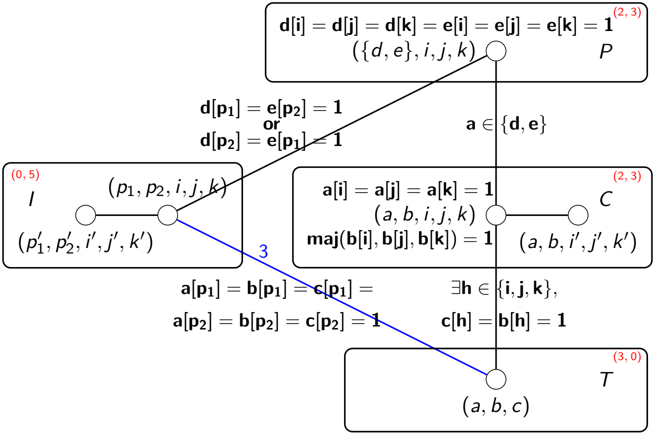


No orthogonal quadruple ⇒ diameter at most 4



Automatic paths of length at most 4, except for T-T, T-C, T-P, and C-C

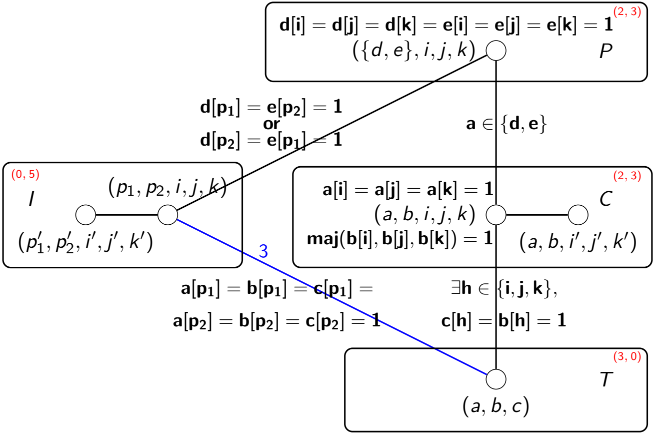
No orthogonal quadruple ⇒ diameter at most 4



T-T, T-C, C-C: (a, b, c) or $(a, b, i', j', k') - (a, b, i, j, k) - (\{a, d\}, i, j, k) - (d, e, i, j, k) - (d, e, f)$ or (d, e, i'', j'', k'')

with $i = \mathrm{ind}(a, b, c, d)$, $j = \mathrm{ind}(a, b, d, e)$, $k = \mathrm{ind}(a, d, e, f)$

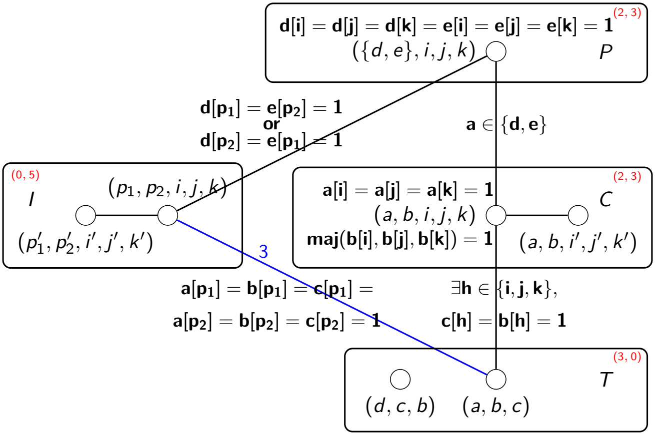
No orthogonal quadruple ⇒ diameter at most 4



T-P: $(a, b, c) - (p_1, p_2, i, j, k) - (\{d, e\}, i, j, k)$

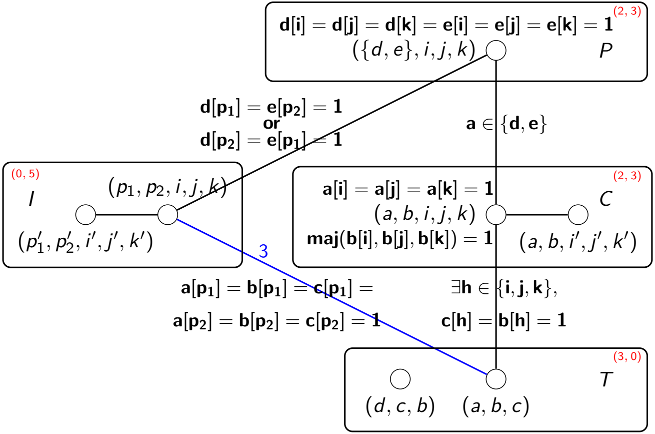
with $p_1 = \text{ind}(a, b, c, d)$, $p_2 = \text{ind}(a, b, c, e)$

a, b, c, d orthogonal $\Rightarrow d((a,b,c),(d,c,b)) \geqslant 7$



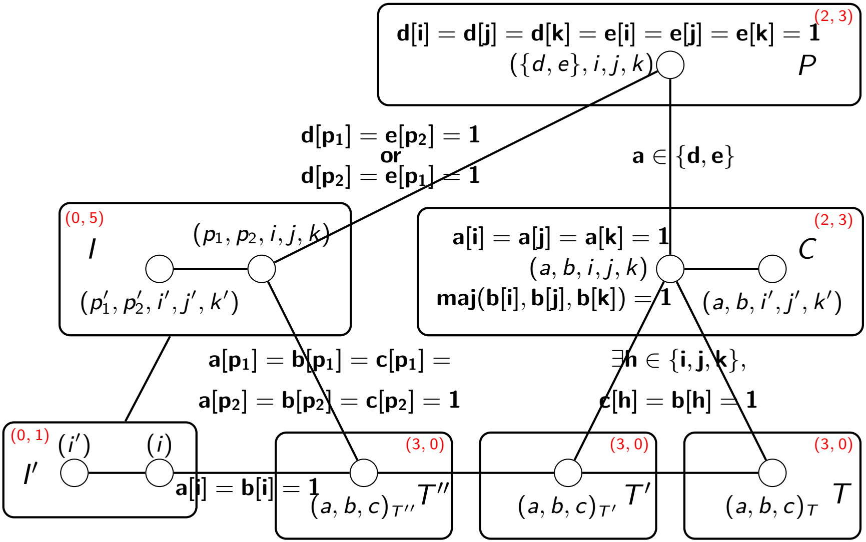
Set I cannot help for a path of length 6

a, b, c, d orthogonal $\Rightarrow d((a, b, c), (d, c, b)) \geqslant 7$

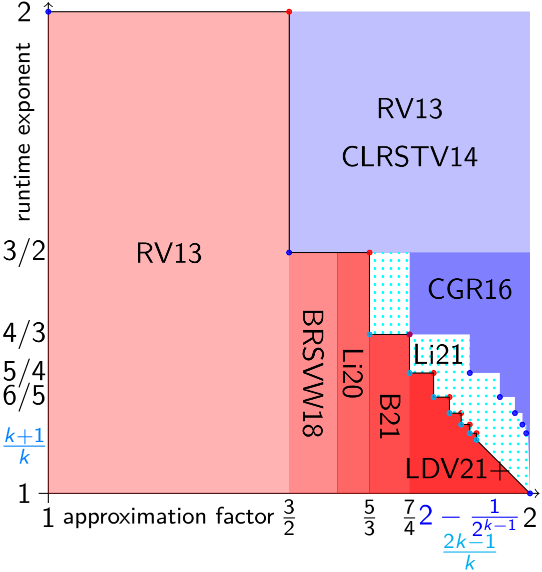


(a, b, i, j, k) and (d, c, i, j, k) have to be part of the path

Removing the weights



Undirected unweighted DIAMETER



Thank you for your attention!