Fast Approximate Shortest Paths in the Congested Clique

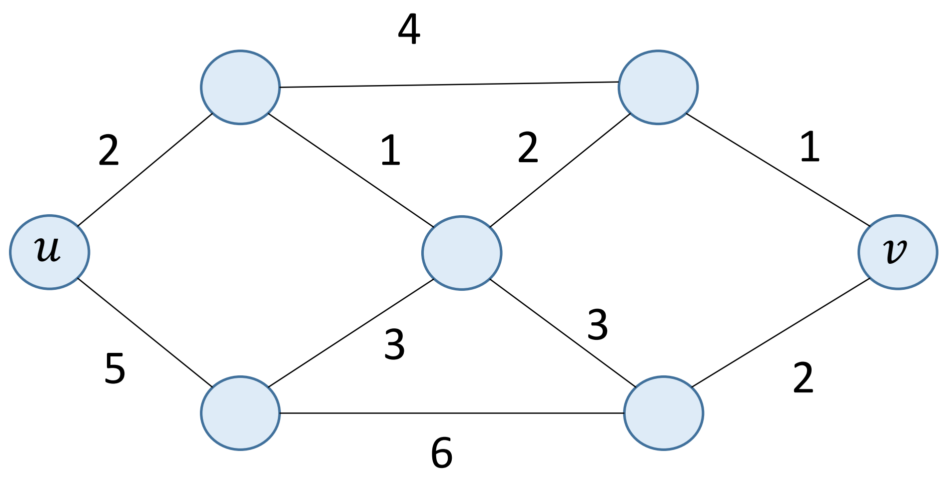
Michal Dory, Technion

Joint work with: Keren Censor-Hillel (Technion), Janne Korhonen (IST Austria), Dean Leitersdorf (Technion)

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement no. 755839

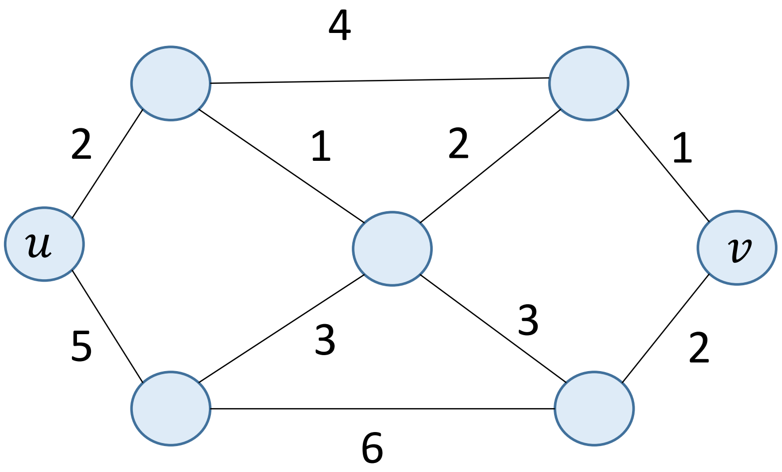
Distance Computation

- All-pairs shortest paths (APSP)
- Single-source shortest paths (SSSP)
- Multi-source shortest paths (MSSP)



The Congested Clique model

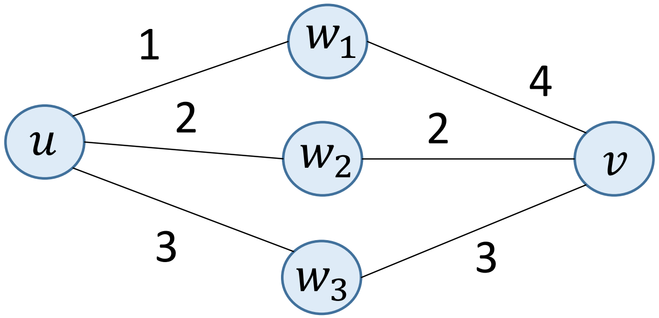
- n vertices
- Synchronous rounds
- $\Theta(\log n)$-bit messages to *all* vertices
- Input and output are *local*



Computing Distances using Matrix Multiplication

- A – weighted adjacency matrix
- Distance product:

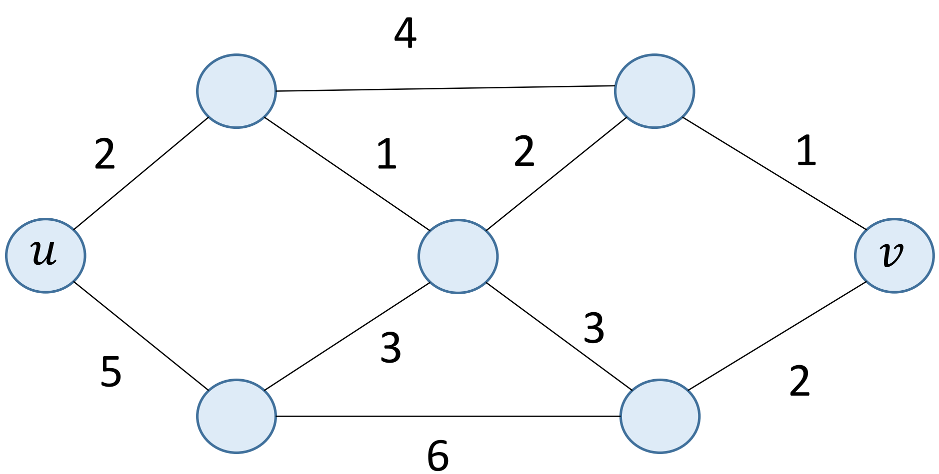
$$A^2[u,v] = \min_{w} A(u,w) + A(w,v)$$



- This is the minimum weight path between u and v of at most 2 edges

Computing Distances using Matrix Multiplication

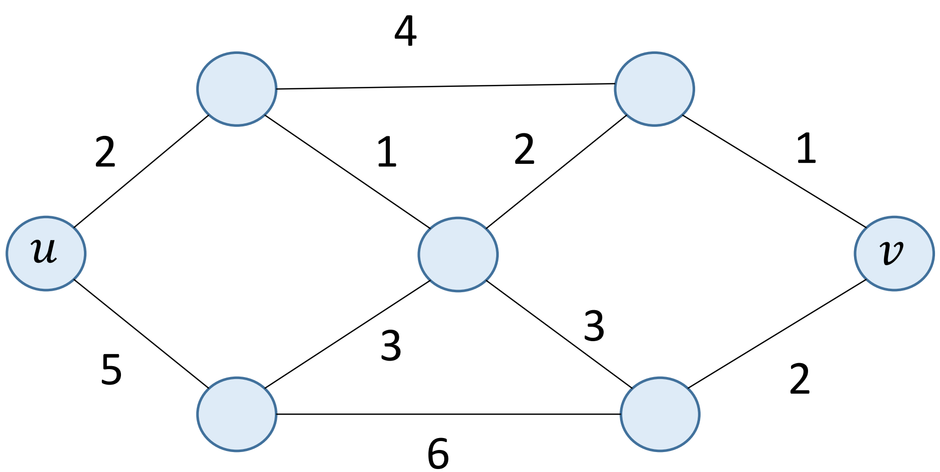
- Similarly, $A^i[u, v]$ = minimum weight path between u and v of at most i edges (hops).
- Our goal: compute A^n



Computing Distances using Matrix Multiplication

- Similarly, $A^i[u,v]$ = minimum weight path between u and v of at most i edges (hops).
- Our goal: compute A^n

$$A[u,v] = \infty$$
$$A^2[u,v] = \infty$$
$$A^3[u,v] = 7$$
$$A^4[u,v] = 6$$
$$\dots$$
$$A^n[u,v] = 6$$



Computing Distances using Matrix Multiplication

- Our goal: compute A^n
- Requires $O(\log n)$ matrix multiplications:

$$A \rightarrow A^2 \rightarrow A^4 \rightarrow \ldots \rightarrow A^n$$

- How fast can we multiply matrices?

Computing Distances using Matrix Multiplication

$O\left(n^{1-2/\omega}\right) = O(n^{0.158})$	Ring	[Censor-Hillel, Kaski, Korhonen, Lenzen, Paz, Suomela '15]
$O\left(n^{1/3}\right)$	Semiring	[Censor-Hillel, Kaski, Korhonen, Lenzen, Paz, Suomela '15]
	Rectangular, Multiple instances, more	[Le Gall '16]

Computing Distances using Matrix Multiplication

$O(n^{0.158})$	• Exact **unweighted undirected** APSP • $\left(1 + o(1)\right)$-approximation for **weighted directed** APSP	[Censor-Hillel, Kaski, Korhonen, Lenzen, Paz, Suomela '15]
$\tilde{O}\left(n^{1/3}\right)$	Exact **weighted directed** APSP	[Censor-Hillel, Kaski, Korhonen, Lenzen, Paz, Suomela '15]
$O(n^{0.2096})$	Exact APSP in **directed** graphs with **constant** weights	[Le Gall '16]

Computing Distances using Matrix Multiplication

$O(n^{0.158})$	• Exact **unweighted undirected** APSP • $(1+o(1))$-approximation for **weighted directed** APSP	[Censor-Hillel, Kaski, Korhonen, Lenzen, Paz, Suomela ‘15]
$\tilde{O}(n^{1/3})$	Exact **weighted directed** APSP	[Censor-Hillel, Kaski, Korhonen, Lenzen, Paz, Suomela ‘15]
$O(n^{0.2096})$	Exact APSP in **directed** graphs with **constant** weights	[Le Gall ‘16]

All complexities are polynomial!

What about approximations?

- We can compute a **spanner**: a sparse subgraph that approximates the distances.

$\tilde{O}(n^{1/k})$	$(2k-1)$-approximation for **weighted undirected** APSP

What about approximations?

- We can compute a **spanner**: a sparse subgraph that approximates the distances.

$\tilde{O}(n^{1/k})$	$(2k-1)$-approximation for **weighted undirected** APSP

Still polynomial for any constant k!

Computing Distances in the Congested Clique

Can we get constant approximation for APSP in sub-polynomial time?

Computing Distances in the Congested Clique

Can we get constant approximation for APSP
in sub-polynomial time?

- <u>For SSSP:</u>

$O(\epsilon^{-3} \text{polylog } n)$-round $(1+\epsilon)$-approximation
[Becker, Karrenbauer, Krinninger, Lenzen '17]

Computing Distances in the Congested Clique

Can we get constant approximation for APSP
in sub-polynomial time?

- <u>For SSSP:</u>

$O(\epsilon^{-3}\text{polylog}\, n)$-round $(1+\epsilon)$-approximation
[Becker, Karrenbauer, Krinninger, Lenzen '17]

Only for a single source!

Our Results: APSP

$O(\log^2 n/\epsilon)$	• $(2+\epsilon)$-approximation for **unweighted undirected** APSP • $(3+\epsilon)$-approximation for **weighted undirected** APSP

First polylog constant-factor approximation!

$(2-\epsilon)$-APSP implies MM	[Dor, Halperin, Zwick ‘00 Korhonen, Suomela ‘18]

Our Results: MSSP and more

$O(\log^2 n/\epsilon)$	$(1+\epsilon)$-approximation weighted undirected **MSSP** with $O(n^{1/2})$ sources
$O(\log^2 n/\epsilon)$	Near $(3/2)$-approximation for **diameter**
$\tilde{O}(n^{1/6})$	Exact weighted undirected **SSSP**

Previous results:

$\tilde{O}(n^{1/3})$	Exact weighted **SSSP**	[Censor-Hillel, Kaski, Korhonen, Lenzen, Paz, Suomela '15]
$O(\epsilon^{-3}\text{polylog}\, n)$	$(1+\epsilon)$-**SSSP**	[Becker, Karrenbauer, Krinninger, Lenzen '17]

Our Techniques

- We can multiply *sparse* matrices faster:

$O\left(1+\frac{(\rho_S\rho_T)^{1/3}}{n^{1/3}}\right)$	Semiring, Sparse	[Censor-Hillel, Leitersdorf, Turner '18]

- ρ_A= density of A, the average number of non-zero entries on a row
- Example: $O(1)$ rounds for $O(n^{3/2})$ edges.

Our Techniques

- We can multiply *sparse* matrices faster.
- How can we use this?

  - Even if A is sparse, A^2 can be dense.

  - We want to compute distances in *general* graphs.

Our Techniques

- We can multiply *sparse* matrices faster.
- How can we use this?

  - Even if A is sparse, A^2 can be dense.

  - We want to compute distances in *general* graphs.

Many *building blocks* for distance computation are actually based on computations in *sparse* graphs

Building blocks for distance computation

- k-nearest: for each vertex, compute distances to k nearest vertices

Building blocks for distance computation

- k-nearest: for each vertex, compute distances to k nearest vertices
- (S, d, k)-source detection: for each vertex, distances to k nearest sources in S, up to hop d

Building blocks for distance computation

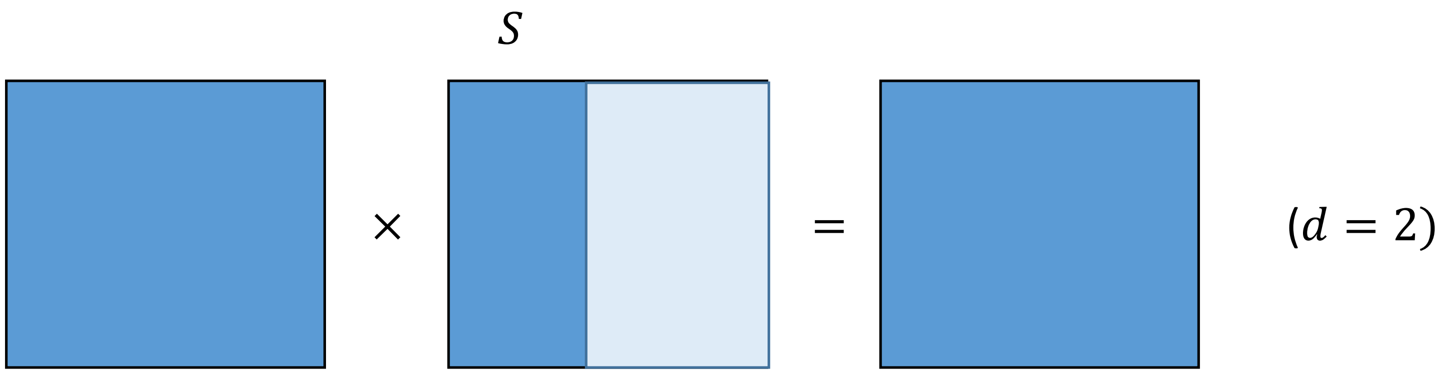
- k-nearest: for each vertex, compute distances to k nearest vertices
- (S, d, k)-source detection: for each vertex, distances to k nearest sources in S, up to hop d
- distances-through S: all distances through a set of sources S

Building blocks for distance computation

- k-nearest: for each vertex, compute distances to k nearest vertices
- (S, d, k)-source detection: for each vertex, distances to k nearest sources in S, up to hop d
- distances-through S: all distances through a set of sources S

Building blocks for distance computation

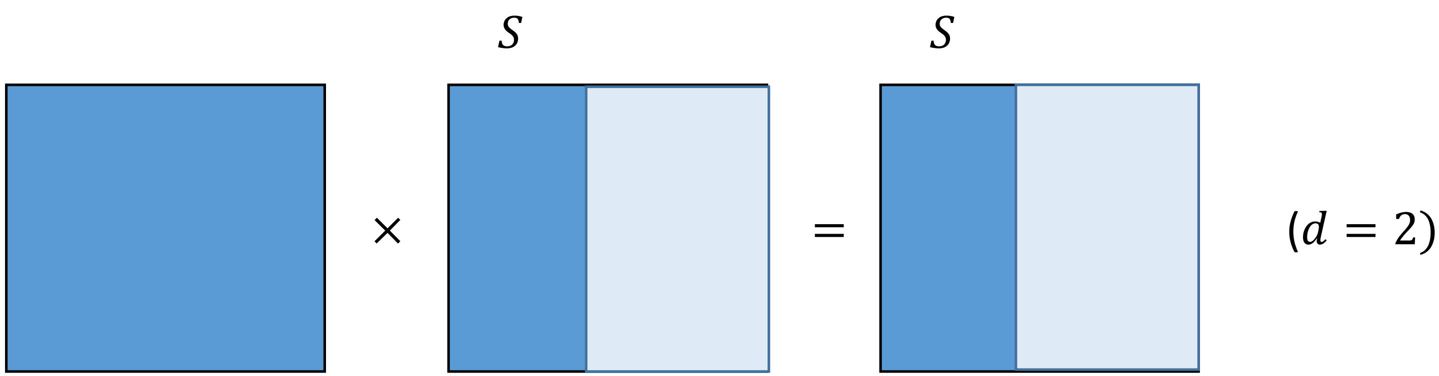
- (S, d, k)-source detection: for each vertex, compute distances to k nearest sources in S, up to hop d



Multiplication of sparse matrix by dense matrix: previous MM algorithm is still polynomial

Building blocks for distance computation

- (S, d, k)-source detection: for each vertex, compute distances to k nearest sources in S, up to hop d



Output matrix is also sparse!

Building blocks for distance computation

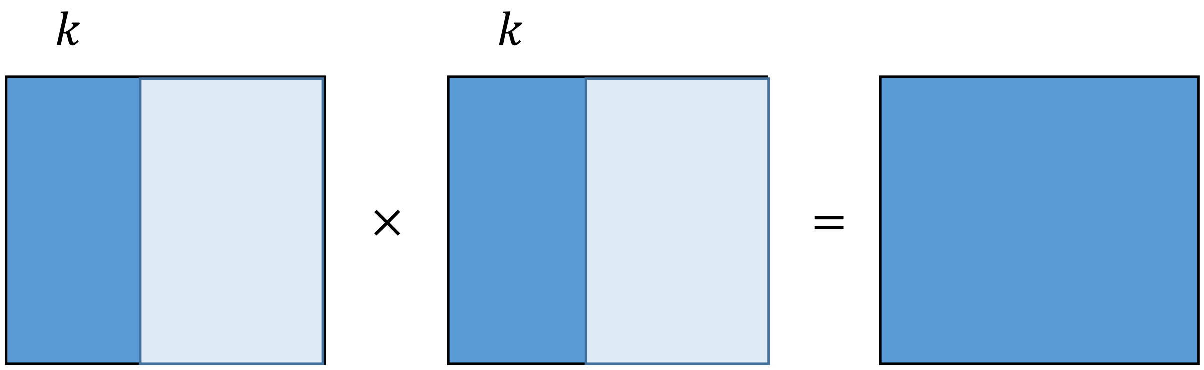
- k-nearest: for each vertex, compute distances to k nearest vertices
- (S, d, k)-source detection: for each vertex, distances to k nearest sources in S, up to hop d
- distances-through S: all distances through a set of sources S

Building blocks for distance computation

- k-nearest: for each vertex, compute distances to k nearest vertices
- (S, d, k)-source detection: for each vertex, distances to k nearest sources in S, up to hop d
- distances-through S: all distances through a set of sources S

Building blocks for distance computation

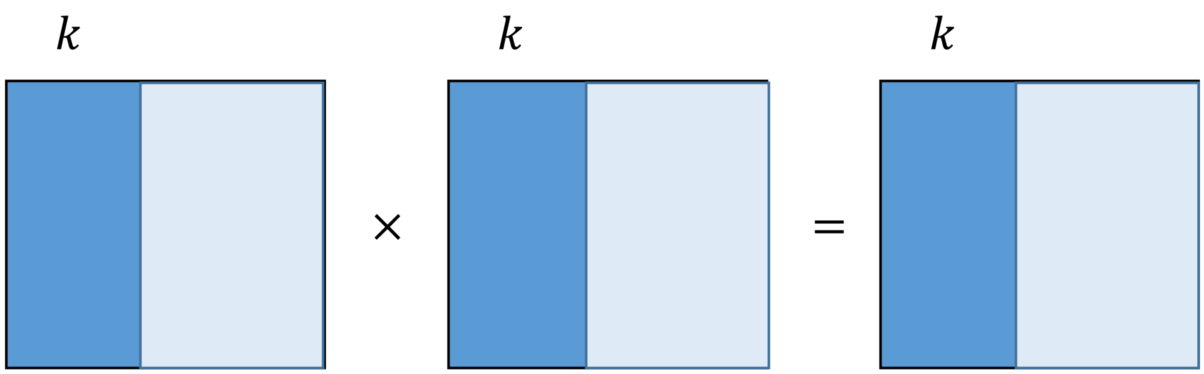
- *k-nearest*: for each vertex, compute distances to k nearest vertices



It's enough to look only at the k closest vertices to each vertex

Building blocks for distance computation

- k-nearest: for each vertex, compute distances to k nearest vertices



It's enough to look only at the k closest vertices to each vertex: also in the *output*

We don't know the identity of the k closest vertices before the computation

New Matrix Multiplication algorithm

- Previous algorithm:

$O\left(1+\frac{(\rho_S\rho_T)^{1/3}}{n^{1/3}}\right)$	Semiring, Sparse	[Censor-Hillel, Leitersdorf, Turner '18]

- Our algorithm:

$O\left(1+\frac{(\rho_S\rho_T\rho_P)^{1/3}}{n^{2/3}}\right)$	Semiring, Sparse	[Censor-Hillel, Dory, Korhonen, Leitersdorf, '19]

$$S \times T = P$$

New Matrix Multiplication algorithm

- Our algorithm:

$O\left(1+\frac{(\rho_S\rho_T\rho_P)^{1/3}}{n^{2/3}}\right)$	Semiring, Sparse	[Censor-Hillel, Dory, Korhonen, Leitersdorf, '19]

- Depends also on the sparsity of the *output* matrix
- Even if we don't know the structure of the output matrix, we can *sparsify* the output matrix *on-the-fly*, keeping only ρ_P smallest entries for each row

Application: Distance Tools

k-nearest:

$O\left(\left(\frac{k}{n^{2/3}}+1\right)\log k\right)$ rounds ⇨ $O(\log n)$ for $k = n^{2/3}$

(S, d, k)-source detection:

$O\left(\left(\frac{m^{1/3}k^{2/3}}{n}+1\right)d\right)$ rounds ($m =$ number of edges)

Work for directed weighted graphs

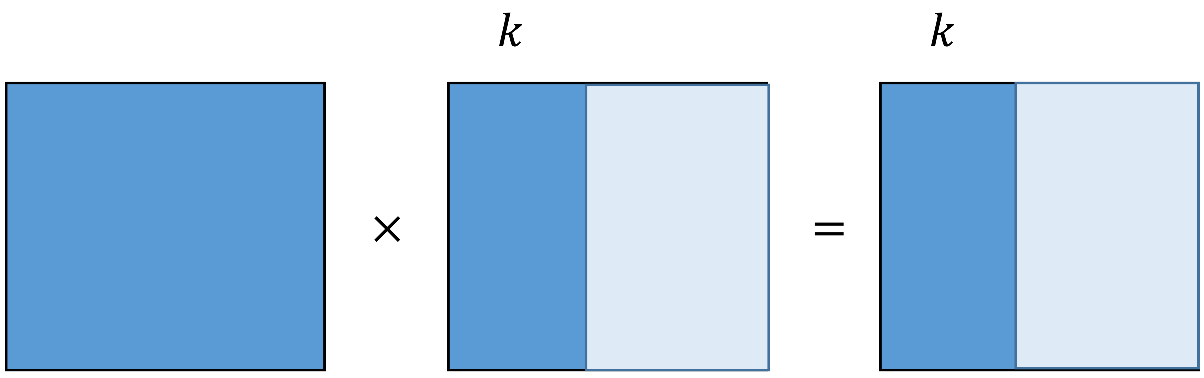
$O\left(1+\frac{(\rho_S\rho_T\rho_P)^{1/3}}{n^{2/3}}\right)$	Semiring, Sparse	[Censor-Hillel, Dory, Korhonen, Leitersdorf, '19]

Application: Distance Tools

(S, d, k)-source detection:

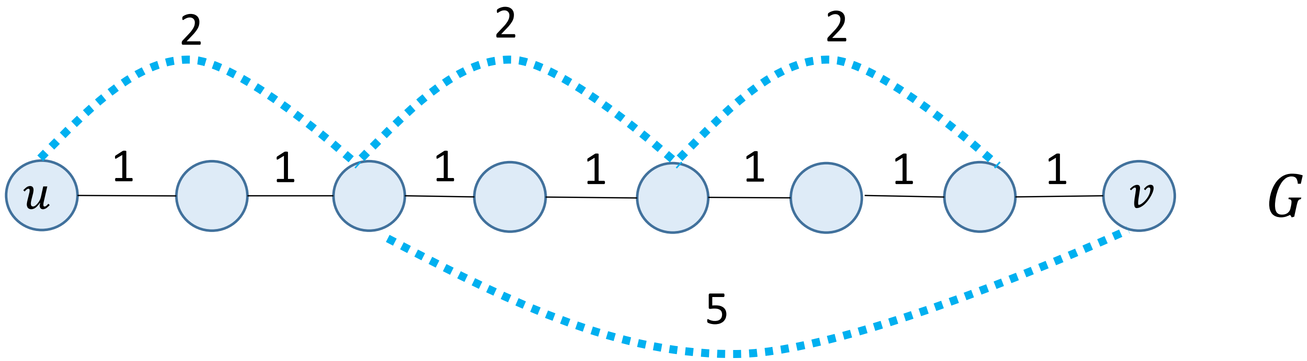
$O\left(\left(\frac{m^{1/3}k^{2/3}}{n} + 1\right) d\right)$ rounds (m = number of edges)

- To exploit the sparsity we need $\boldsymbol{d = n}$ multiplications - too expensive!



Solution: Hopsets

(β, ϵ)-hopset H:
A graph $H = (V, E')$, such that the β-hop distances in $G \cup H$ give $(1+\epsilon)$-approximation for the distances in G



Enough to look at β-hop distances!

Solution: Hopsets

(β, ϵ)-hopset H:
A graph $H = (V, E')$, such that the β-hop distances in $G \cup H$ give $(1+\epsilon)$-approximation for the distances in G

Our goal: to have small β and small running time t

What is known?

- We can get $\beta = t = O\left(\frac{\log\log n}{\epsilon}\right)^{\log\log n}$
  [Elkin, Neiman '17]

Solution: Hopsets

(β, ϵ)-hopset H:
A graph $H = (V, E')$, such that the β-hop distances in $G \cup H$ give $(1 + \epsilon)$-approximation for the distances in G

Our goal: to have small β and small running time t

What is known?

- We can get $\beta = t = O\left(\frac{\log \log n}{\epsilon}\right)^{\log \log n}$
  [Elkin, Neiman '17]

Can we get a poly-logarithmic complexity?

Solution: Hopsets

(β, ϵ)-hopset H:
A graph $H = (V, E')$, such that the β-hop distances in $G \cup H$ give $(1+\epsilon)$-approximation for the distances in G

Our goal: to have small β and small running time t

What is known?

- We can get $\beta = t = O\left(\frac{\log\log n}{\epsilon}\right)^{\log\log n}$ [Elkin, Neiman '17]

Yes! we can get: $\beta = O\left(\frac{\log n}{\epsilon}\right), t = O\left(\frac{\log^2 n}{\epsilon}\right)$

Solution: Hopsets

(β, ϵ)-hopset H:
A graph $H = (V, E')$, such that the β-hop distances in $G \cup H$ give $(1+\epsilon)$-approximation for the distances in G

Our goal: to have small β and small running time t

- We can get: $\beta = O\left(\frac{\log n}{\epsilon}\right)$, $t = O\left(\frac{\log^2 n}{\epsilon}\right)$

Idea: using our *distance tools* we can implement efficiently the hopset construction of [Elkin, Neiman '17] [Huang, Pettie '19] [Thorup, Zwick '06]

Applications: MSSP

(β, ϵ)-hopset H:
A graph $H = (V, E')$, such that the β-hop distances in $G \cup H$ give $(1+\epsilon)$-approximation for the distances in G

(S, d, k)-source detection: $O\left(\left(\frac{m^{1/3}k^{2/3}}{n} + 1\right) d\right)$ rounds

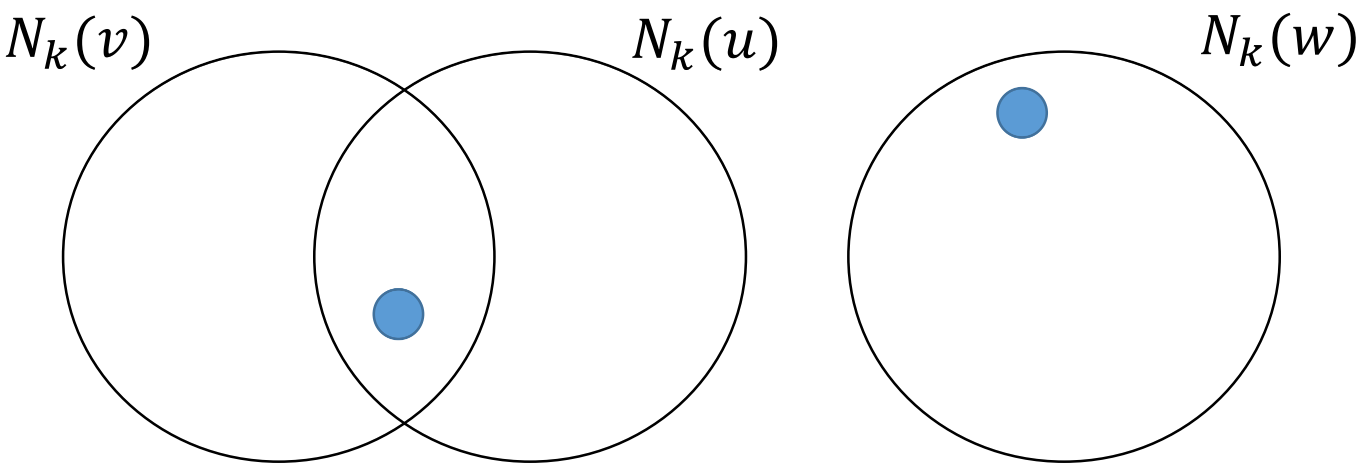
- We run our (S, d, k)- source detection in $G \cup H$ with $d = \beta$: $(1+\epsilon)$–approximation for the distances of all vertices from S

Complexity: $O\left(\left(\frac{|S|^{2/3}}{n^{1/3}} + \log n\right) \frac{\log n}{\epsilon}\right)$

➡ poly-logarithmic for $|S| = O(\sqrt{n})$

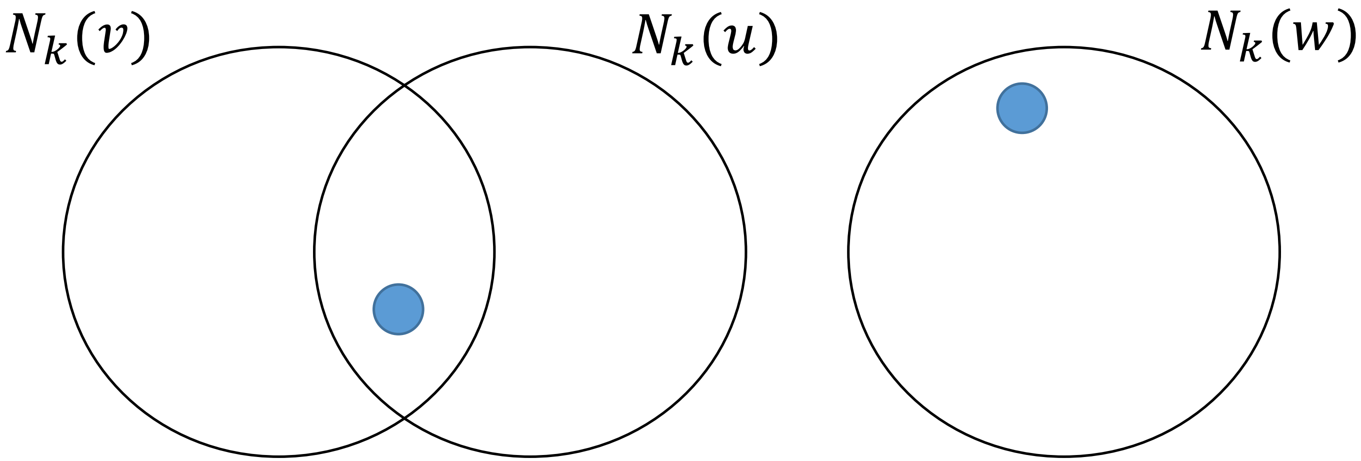
Applications: APSP

- We compute the $k = \tilde{O}(\sqrt{n})$ nearest vertices $N_k(v)$ for each v
- We compute a hitting set $\boldsymbol{A}$ of the sets $N_k(v)$ with $|A| = \tilde{O}(\sqrt{n})$



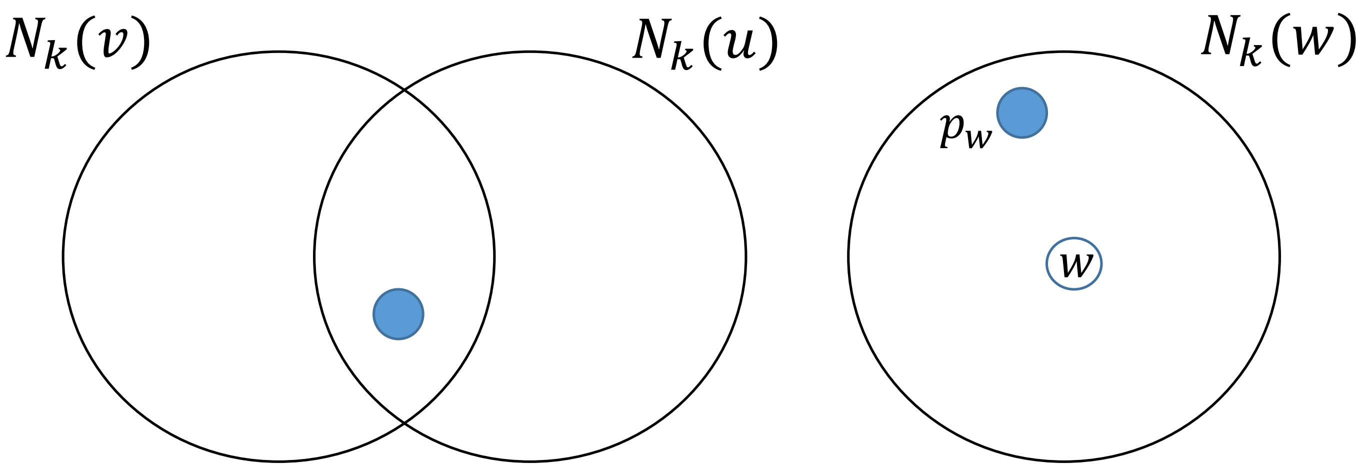
Applications: APSP

- We compute the $k = \tilde{O}(\sqrt{n})$ nearest vertices $N_k(v)$ for each v
- We compute a hitting set $\boldsymbol{A}$ of the sets $N_k(v)$ with $|A| = \tilde{O}(\sqrt{n})$: $O((\log\log n)^3)$ rounds [Parter, Yogev '18]



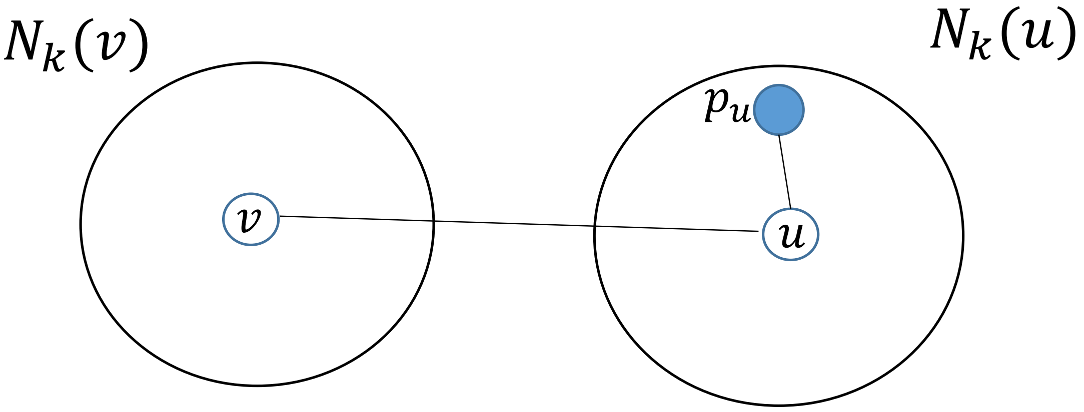
Applications: APSP

- For each v, let p_{v} = the closest vertex to v in $A \cap N_k(v)$
- We compute $(1+\epsilon)$-approximate distances $\delta(u,v)$ from all vertices to A



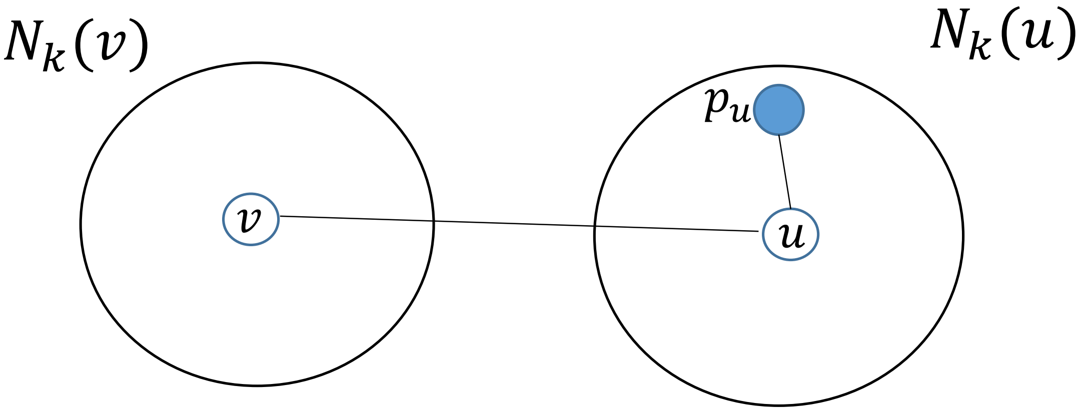
Applications: APSP

- If $v \in N_k(u)$, u knows $d(u,v)$.
- Otherwise, we estimate $\delta(u,v) = \delta(u,p_u) + \delta(p_u,v)$



Applications: APSP

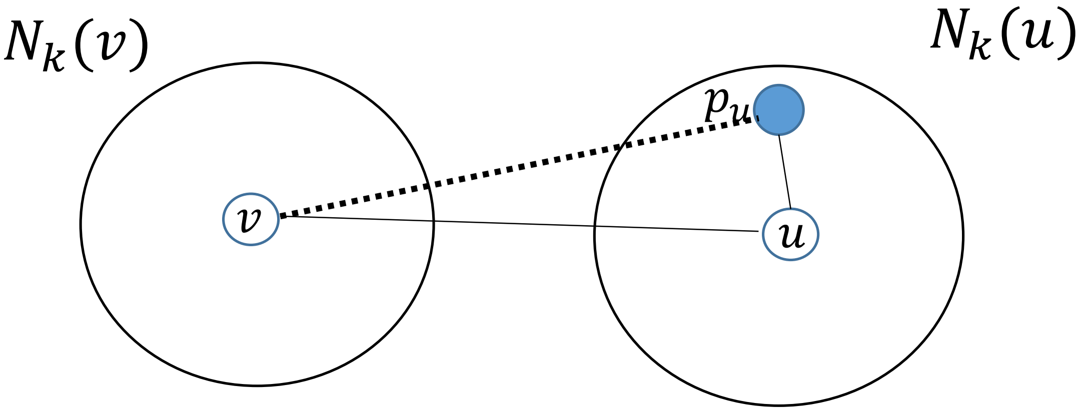
- If $v \in N_k(u)$, u knows $d(u,v)$.
- Otherwise, we estimate $\delta(u,v) = \delta(u,p_u) + \delta(p_u,v)$



- $d(u,p_u) \le d(u,v)$

Applications: APSP

- If $v \in N_k(u)$, u knows $d(u,v)$.
- Otherwise, we estimate $\delta(u,v) = \delta(u,p_u) + \delta(p_u,v)$



- $d(u,p_u) \le d(u,v)$
- $d(v,p_u) \le d(u,v) + d(u,p_u) \le 2d(u,v)$
- $\delta(u,p_u) + \delta(p_u,v)$ gives a $(\mathbf{3} + \boldsymbol{\epsilon})$-approximation

Applications: APSP

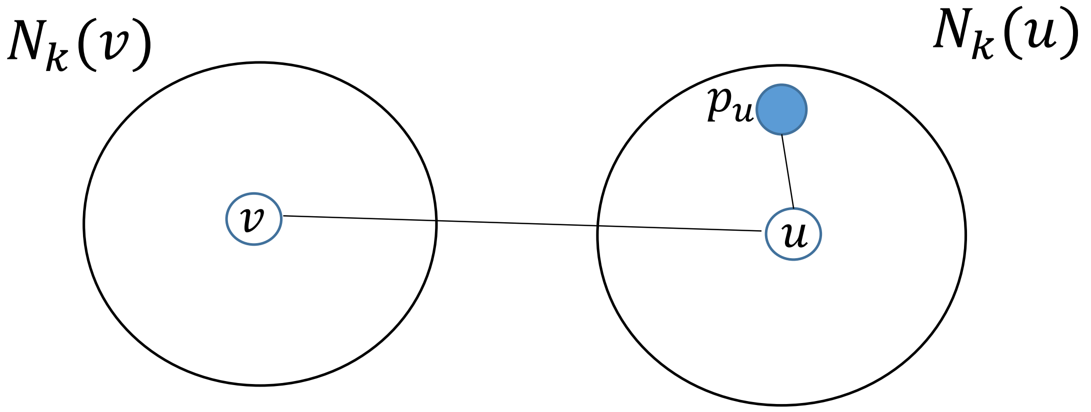
<u>Complexity:</u>

- Computing the k-nearest: $O(\log n)$ time
- Computing a hitting set: $O((\mathrm{loglog}n)^3)$ time
- Computing distances to A: $O\left(\frac{\log^2 n}{\epsilon}\right)$ time

$(3+\epsilon)$-approximation in $O\left(\frac{\log^2 n}{\epsilon}\right)$ rounds

Applications: APSP

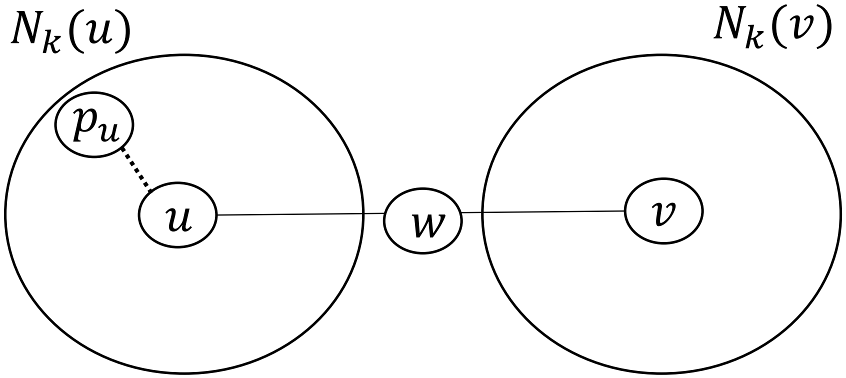
Can we improve the approximation?



If $d(u,p_u) \leq \frac{d(u,v)}{2}$ then the same analysis shows a $(2+\epsilon)$-approximation

Applications: APSP

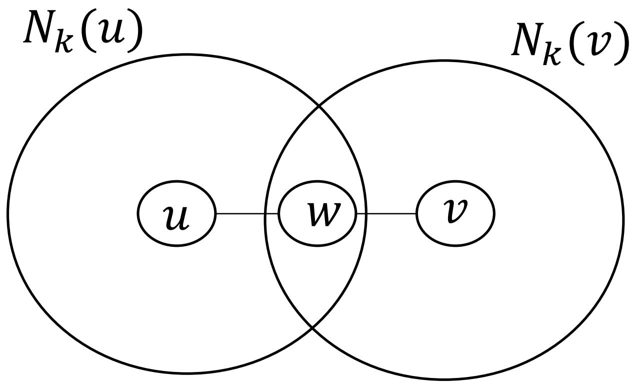
<u>Case 1:</u>



$d(u, p_u) \leq \frac{d(u,v)}{2}$ ⇒ $(2 + \epsilon)$-approximation

Applications: APSP

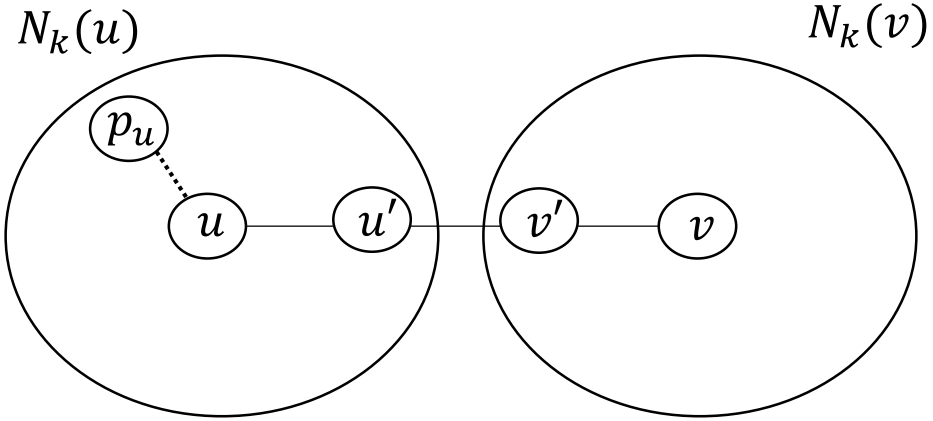
<u>Case 2:</u>



$w \in N_k(u) \cap N_k(v)$ ⇒ We can compute $d(u, v)$

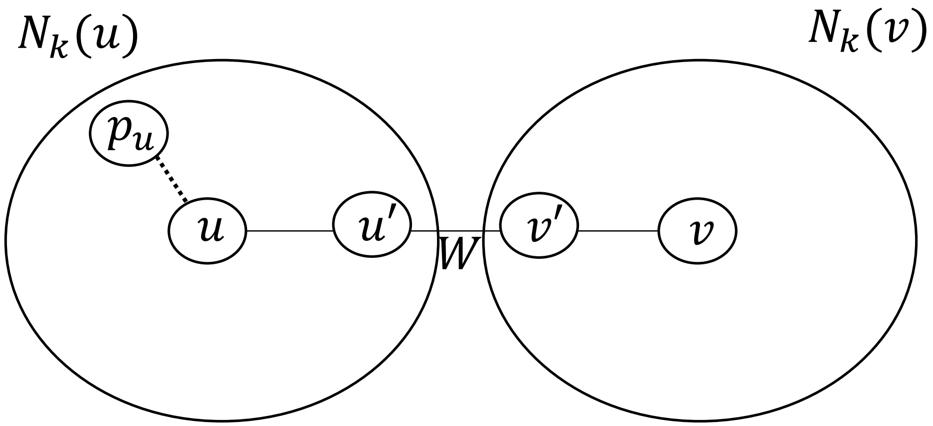
Applications: APSP

<u>Case 3:</u>



Applications: APSP

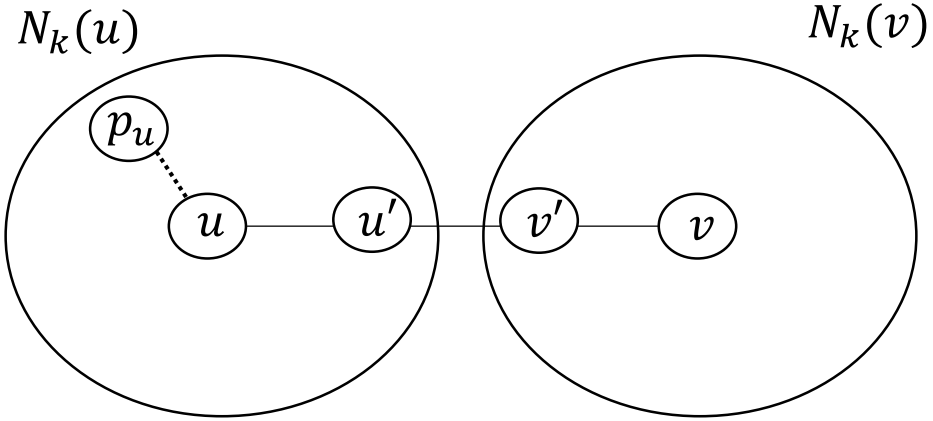
Case 3:



$d(u, p_u) \leq \frac{d(u,v)+\boldsymbol{W}}{2}$ ⇨ $\left((2+\epsilon), (1+\epsilon)\boldsymbol{W}\right)$-approximation

Applications: APSP

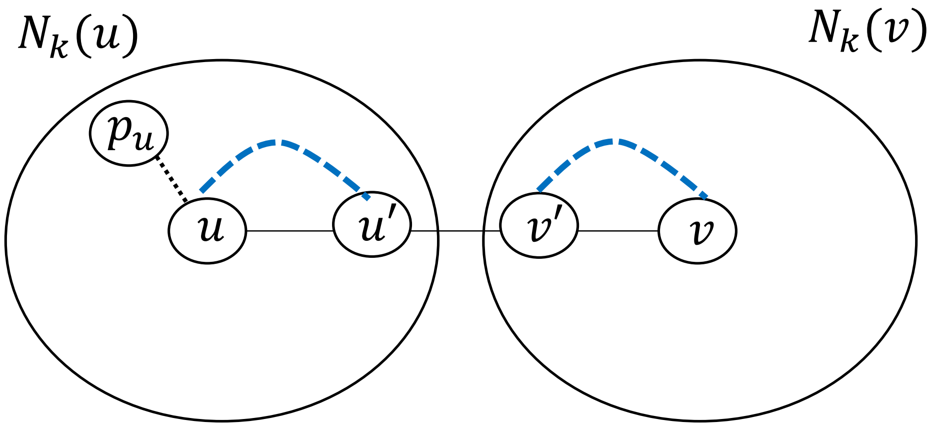
Case 3:



How do we get a $(2 + \epsilon)$-approximation?

Applications: APSP

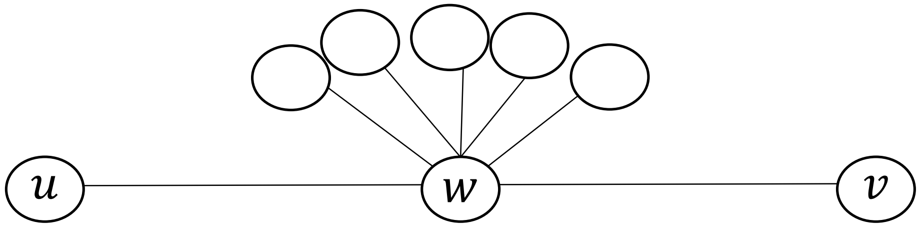
<u>Case 3:</u>



- We have a 3-hop shortest path between u and v
- However, the 3 relevant matrices are too dense

Applications: APSP

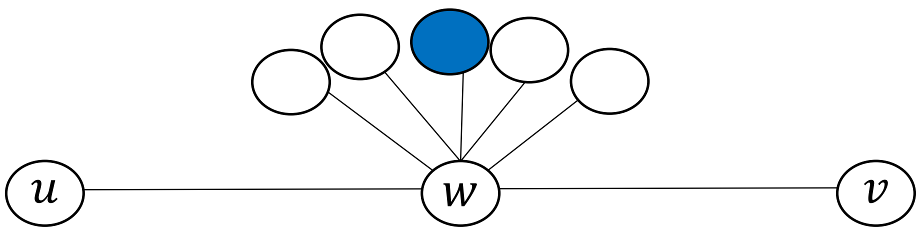
<u>Dense paths:</u>



$\deg(w) \geq \sqrt{n}$

Applications: APSP

<u>Dense paths:</u>

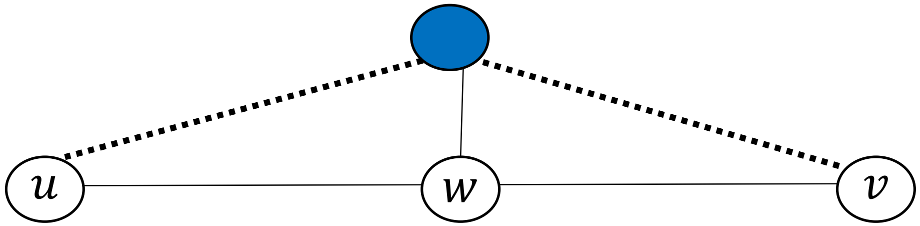


$\deg(w) \geq \sqrt{n}$

Compute a hitting set of size $\tilde{O}(\sqrt{n})$ for high-degree vertices

Applications: APSP

<u>Dense paths:</u>

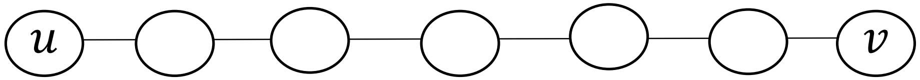


$\deg(w) \geq \sqrt{n}$

The distance through the hitting set gives a +2 approximation

Applications: APSP

Sparse paths:

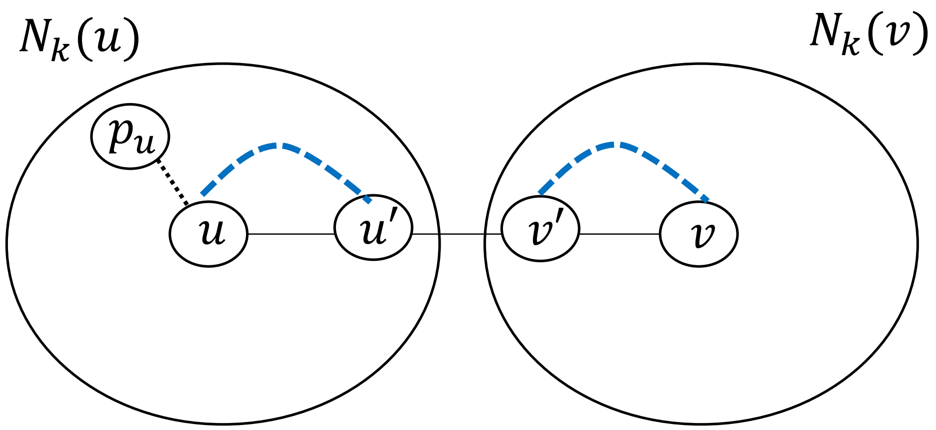


- All degrees are at most $\sqrt{n}$
- We can focus on a sparse graph with $O(n^{3/2})$ edges.

Applications: APSP

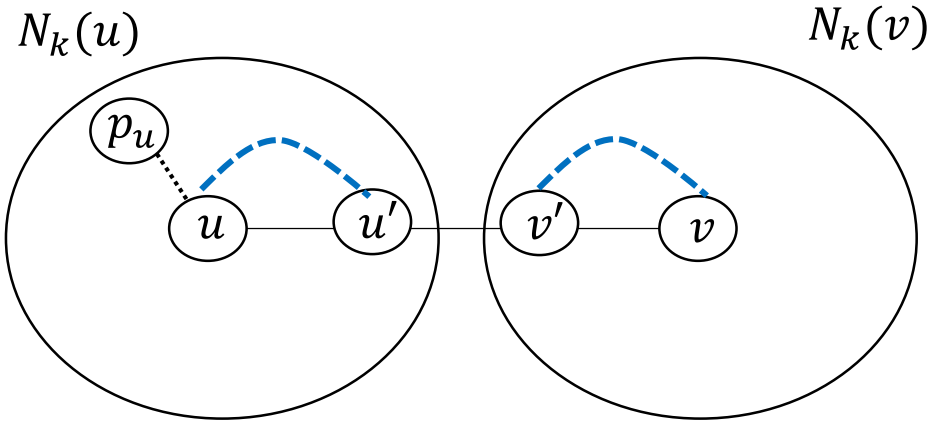
<u>Sparse paths:</u>

- We can focus on a sparse graph with $\boldsymbol{O(n^{3/2})}$ edges.
- We can compute MSSP from $\boldsymbol{\widetilde{O}(n^{3/4})}$ sources.
- Can take $k = \boldsymbol{\widetilde{O}(n^{1/4})}$.
- Now the 3 relevant matrices are sparse enough.



Applications: APSP

$(2+\epsilon)$-approximation for **unweighted** APSP in $O\left(\frac{\log^2 n}{\epsilon}\right)$ rounds



Conclusion

- We show a fast algorithm for matrix multiplication that depends on the *sparsity* and is *output*-sensitive.
- Allows to build efficient *distance tools.*
- Together with *hopsets*: polylog algorithms for MSSP, APSP.

Summary

$O(\log^2 n/\epsilon)$	• $(2+\epsilon)$-approximation for **unweighted undirected** APSP • $(3+\epsilon)$-approximation for **weighted undirected** APSP
$O(\log^2 n/\epsilon)$	$(1+\epsilon)$-approximation for **weighted** undirected **MSSP** with $O(n^{1/2})$ sources
$O(\log^2 n/\epsilon)$	Near $(3/2)$-approximation for **diameter**
$\tilde{O}(n^{1/6})$	Exact weighted undirected **SSSP**

Open Questions

- Can we get a $(2+\epsilon)$-approximation for weighted APSP?
- Can we get sub-polynomial algorithm for exact SSSP? Or directed SSSP?