Streaming Algorithms for Matchings in Low Arboricity Graphs

Sofya Vorotnikova
University of Massachusetts Amherst

Joint work with Andrew McGregor

Streaming Model(s)

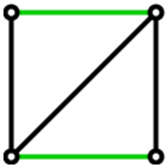
- Vertex set is fixed
- Edge updates arrive in a sequence
- One pass

	insertions	deletions	arbitrary order
dynamic	✓	✓	✓
insert-only	✓	✗	✓
adjacency-list	✓	✗	✗

edges incident to the same vertex arrive together; see every edge twice

Approximating Size of Maximum Matching

Matching is a set of edges that don't share endpoints.



In insert-only stream can easily obtain *maximal* matching, which is a 2-approximation of *maximum* matching.

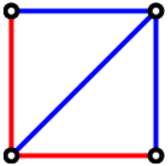
Maximum matching can be as large as $n/2$.

By approximating the **size** of the matching without finding the matching itself, we can use smaller space.

Low Arboricity Graphs

We concentrate on the class of graphs of arboricity α.

Arboricity is the minimum number of forests into which the edges of the graph can be partitioned.



Property: Every subgraph on r vertices has at most αr edges.

Planar graphs have arboricity at most 3.

In dynamic stream, intermediate graphs can have high arboricity.

Results

	space	approx factor	work
dynamic	$\tilde{O}(\alpha n^{4/5})$	$(5\alpha + 9)(1 + \epsilon)$	CCEHMMV16
	$\tilde{O}(\alpha n^{4/5})$	$(\alpha + 2)(1 + \epsilon)$	MV16
	$\tilde{O}(\alpha^{10/3} n^{2/3})$	$(22.5\alpha + 6)(1 + \epsilon)$	CJMM17*
	$\Omega(\sqrt{n}/\alpha^{2.5})$	$O(\alpha)$	AKL17
insert-only	$\tilde{O}(\alpha n^{2/3})$	$(5\alpha + 9)(1 + \epsilon)$	EHLMO15
	$\tilde{O}(\alpha n^{2/3})$	$(\alpha + 2)(1 + \epsilon)$	MV16
	$O(\alpha \epsilon^{-3} \log^2 n)$	$(22.5\alpha + 6)(1 + \epsilon)$	CJMM17
	$O(\epsilon^{-2} \log n)$	$(\alpha + 2)(1 + \epsilon)$	MV18
adj	$O(1)$	$\alpha + 2$	MV16

*Restriction: $O(\alpha n)$ deletions.

Space is specified in words. An edge or a counter = one word.

Approach

All our results have the following two parts:

- **Structural result**: define Σ that is an $(\alpha + 2)$ approximation of match(G)
- **Algorithm**: $(1 + \epsilon)$ approximation of Σ in streaming (exact computation in adjacency list stream)

Dynamic: Σ_{dyn}

- $(1 + \epsilon)$-approximation in $\tilde{O}(\alpha n^{4/5})$ space
- Also gives $\tilde{O}(\alpha n^{2/3})$ space algorithm in insert-only streams

Insert-only: Σ_{ins}

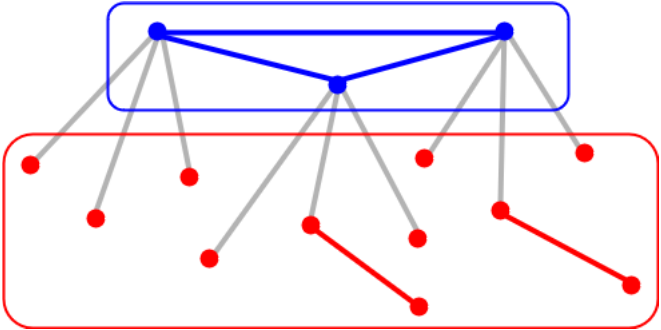
- $(1 + \epsilon)$-approximation in $O(\epsilon^{-2} \log n)$ space

Adjacency list: Σ_{adj}

- Exact computation in $O(1)$ space

Structural Results

Structural Results: Definitions



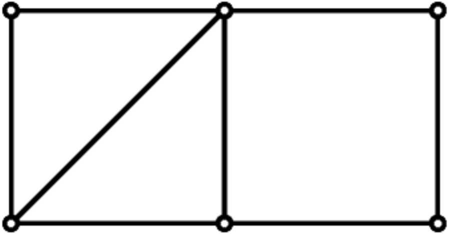
V^H = heavy vertices of degree $\geq \alpha + 2$
E^H = heavy edges with 2 heavy endpoints
V^L = light vertices
E^L = light edges

Structural Results: Definitions: Σ_{adj}

$$\Sigma_{adj} = |E^L| + |V^H|(\alpha + 1) - |E^H|$$

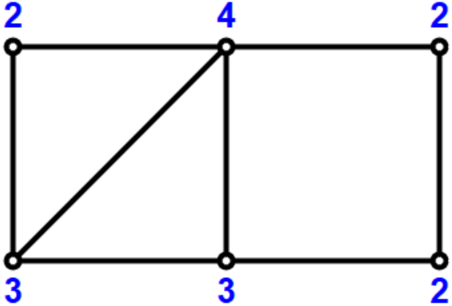
Structural Results: Definitions: Σ_{dyn}

$$x_e = x_{uv} = \min\left(\frac{1}{d(u)}, \frac{1}{d(v)}, \frac{1}{\alpha+1}\right)$$



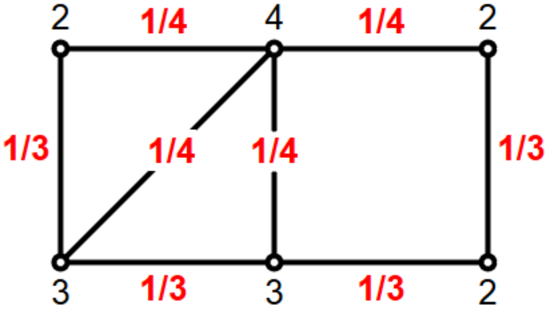
Structural Results: Definitions: Σ_{dyn}

$$x_e = x_{uv} = \min\left(\frac{1}{d(u)}, \frac{1}{d(v)}, \frac{1}{\alpha+1}\right)$$



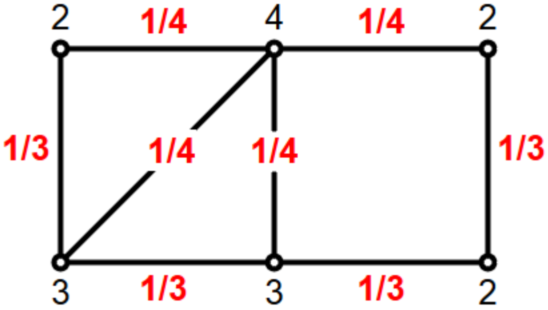
Structural Results: Definitions: Σ_{dyn}

$$x_e = x_{uv} = \min\left(\frac{1}{d(u)}, \frac{1}{d(v)}, \frac{1}{\alpha+1}\right)$$



Structural Results: Definitions: Σ_{dyn}

$$x_e = x_{uv} = \min\left(\frac{1}{d(u)}, \frac{1}{d(v)}, \frac{1}{\alpha+1}\right)$$



$$\Sigma_{dyn} = (\alpha+1)\sum_{e} x_e$$

Structural Results: Σ_{dyn} and Σ_{adj}

$$\text{match}(G) \leq |V^H| + |E^L|$$ since a matched edge is either light or incident to a heavy vertex

$$\leq |E^L| + |V^H|(\alpha+1) - |E^H| = \Sigma_{adj}$$ since $|E^H| \leq \alpha|V^H|$

$$\leq (\alpha+1)\sum_e x_e = \Sigma_{dyn}$$ Lemma 1

$$\leq (\alpha+2)\,\text{match}(G)$$ Lemma 2

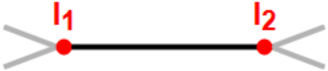
Structural Results: Σ_{dyn} and Σ_{adj}: Lemma 1

Lemma

$$\Sigma_{adj} = |E^L| + |V^H|(\alpha + 1) - |E^H| \leq (\alpha + 1) \sum_e x_e = \Sigma_{dyn}$$

Structural Results: Σ_{dyn} and Σ_{adj}: Lemma 1

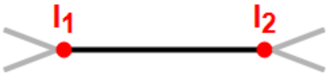
Light edge:



$$x_e = \min\left(\frac{1}{d(\ell_1)}, \frac{1}{d(\ell_2)}, \frac{1}{\alpha + 1}\right)$$

Structural Results: Σ_{dyn} and Σ_{adj}: Lemma 1

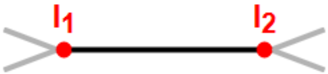
Light edge:



$$x_e = \min\left(\frac{1}{d(\ell_1)} \geq \frac{1}{\alpha+1}, \frac{1}{d(\ell_2)} \geq \frac{1}{\alpha+1}, \boldsymbol{\frac{1}{\alpha+1}}\right)$$

Structural Results: Σ_{dyn} and Σ_{adj}: Lemma 1

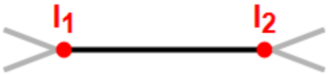
Light edge:



$$x_e = \min\left(\frac{1}{d(\ell_1)} \geq \frac{1}{\alpha+1}, \frac{1}{d(\ell_2)} \geq \frac{1}{\alpha+1}, \boldsymbol{\frac{1}{\alpha+1}}\right) = \frac{1}{\alpha+1}$$

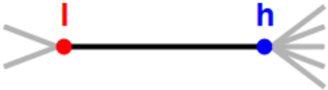
Structural Results: Σ_{dyn} and Σ_{adj}: Lemma 1

Light edge:



$$x_e = \min\left(\frac{1}{d(\ell_1)} \geq \frac{1}{\alpha+1}, \frac{1}{d(\ell_2)} \geq \frac{1}{\alpha+1}, \boldsymbol{\frac{1}{\alpha+1}}\right) = \frac{1}{\alpha+1}$$

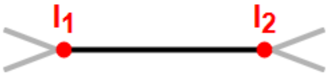
Edge with 1 light and 1 heavy endpoints:



$$x_e = \min\left(\frac{1}{d(\ell)}, \frac{1}{d(h)}, \frac{1}{\alpha+1}\right)$$

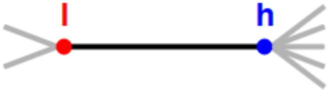
Structural Results: Σ_{dyn} and Σ_{adj}: Lemma 1

Light edge:



$$x_e = \min\left(\frac{1}{d(\ell_1)} \geq \frac{1}{\alpha+1}, \frac{1}{d(\ell_2)} \geq \frac{1}{\alpha+1}, \boldsymbol{\frac{1}{\alpha+1}}\right) = \frac{1}{\alpha+1}$$

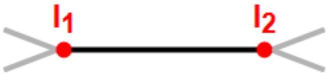
Edge with 1 light and 1 heavy endpoints:



$$x_e = \min\left(\frac{1}{d(\ell)} \geq \frac{1}{\alpha+1}, \boldsymbol{\frac{1}{d(h)}} < \frac{1}{\alpha+1}, \frac{1}{\alpha+1}\right)$$

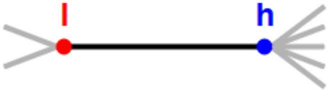
Structural Results: Σ_{dyn} and Σ_{adj}: Lemma 1

Light edge:



$$x_e = \min\left(\frac{1}{d(\ell_1)} \geq \frac{1}{\alpha+1}, \frac{1}{d(\ell_2)} \geq \frac{1}{\alpha+1}, \boldsymbol{\frac{1}{\alpha+1}}\right) = \frac{1}{\alpha+1}$$

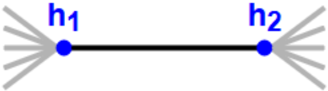
Edge with 1 light and 1 heavy endpoints:



$$x_e = \min\left(\frac{1}{d(\ell)} \geq \frac{1}{\alpha+1}, \boldsymbol{\frac{1}{d(h)}} < \frac{1}{\alpha+1}, \frac{1}{\alpha+1}\right) = \frac{1}{d(h)}$$

Structural Results: Σ_{dyn} and Σ_{adj}: Lemma 1

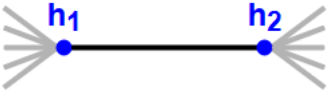
Heavy edge:



$$x_e = \min\left(\frac{1}{d(h_1)}, \frac{1}{d(h_2)}, \frac{1}{\alpha+1}\right)$$

Structural Results: Σ_{dyn} and Σ_{adj}: Lemma 1

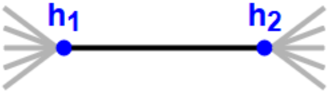
Heavy edge:



$$x_e = \min\left(\frac{\mathbf{1}}{\boldsymbol{d(h_1)}} < \frac{1}{\alpha+1}, \frac{\mathbf{1}}{\boldsymbol{d(h_2)}} < \frac{1}{\alpha+1}, \frac{1}{\alpha+1}\right)$$

Structural Results: Σ_{dyn} and Σ_{adj}: Lemma 1

Heavy edge:

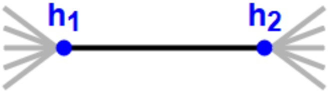


$$x_e = \min\left(\frac{\mathbf{1}}{\boldsymbol{d(h_1)}} < \frac{1}{\alpha+1}, \frac{\mathbf{1}}{\boldsymbol{d(h_2)}} < \frac{1}{\alpha+1}, \frac{1}{\alpha+1}\right)$$

$$= \min\left(\frac{1}{d(h_1)}, \frac{1}{d(h_2)}\right)$$

Structural Results: Σ_{dyn} and Σ_{adj}: Lemma 1

Heavy edge:



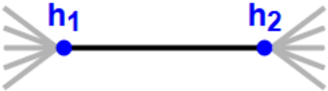
$$x_e = \min\left(\frac{\mathbf{1}}{\boldsymbol{d(h_1)}} < \frac{1}{\alpha+1}, \frac{\mathbf{1}}{\boldsymbol{d(h_2)}} < \frac{1}{\alpha+1}, \frac{1}{\alpha+1}\right)$$

$$= \min\left(\frac{1}{d(h_1)}, \frac{1}{d(h_2)}\right)$$

$$= \frac{1}{d(h_1)} + \frac{1}{d(h_2)} - \max\left(\frac{1}{d(h_1)}, \frac{1}{d(h_2)}\right)$$

Structural Results: Σ_{dyn} and Σ_{adj}: Lemma 1

Heavy edge:



$$
\begin{aligned}
x_e &= \min\left(\frac{\mathbf{1}}{\boldsymbol{d(h_1)}} < \frac{1}{\alpha+1}, \frac{\mathbf{1}}{\boldsymbol{d(h_2)}} < \frac{1}{\alpha+1}, \frac{1}{\alpha+1}\right) \\
&= \min\left(\frac{1}{d(h_1)}, \frac{1}{d(h_2)}\right) \\
&= \frac{1}{d(h_1)} + \frac{1}{d(h_2)} - \max\left(\frac{1}{d(h_1)}, \frac{1}{d(h_2)}\right) \\
&> \frac{1}{d(h_1)} + \frac{1}{d(h_2)} - \frac{1}{\alpha+1}
\end{aligned}
$$

Structural Results: Σ_{dyn} and Σ_{adj}: Lemma 1

$$\sum_{e} x_e = \sum_{e \in E^L} x_e + \sum_{e \notin E^L, E^H} x_e + \sum_{e \in E^H} x_e$$

Structural Results: Σ_{dyn} and Σ_{adj}: Lemma 1

$$\sum_{e} x_e = \sum_{e \in E^L} x_e + \sum_{e \notin E^L, E^H} x_e + \sum_{e \in E^H} x_e$$

$$\geq \sum_{e \in E^L} \frac{1}{\alpha + 1} + \sum_{e \notin E^L, E^H} \frac{1}{d(h)} + \sum_{e \in E^H} \left(\frac{1}{d(h_1)} + \frac{1}{d(h_2)} - \frac{1}{\alpha + 1} \right)$$

Structural Results: Σ_{dyn} and Σ_{adj}: Lemma 1

$$\sum_{e} x_e = \sum_{e \in E^L} x_e + \sum_{e \notin E^L, E^H} x_e + \sum_{e \in E^H} x_e$$

$$\geq \sum_{e \in E^L} \frac{1}{\alpha + 1} + \sum_{e \notin E^L, E^H} \frac{1}{d(h)} + \sum_{e \in E^H} \left(\frac{1}{d(h_1)} + \frac{1}{d(h_2)} - \frac{1}{\alpha + 1} \right)$$

$$= \sum_{e \in E^L} \frac{1}{\alpha + 1} + \sum_{h \in V^H} \underbrace{\sum_{e : h \in e} \frac{1}{d(h)}}_{1} - \sum_{e \in E^H} \frac{1}{\alpha + 1}$$

Structural Results: Σ_{dyn} and Σ_{adj}: Lemma 1

$$\sum_{e} x_e = \sum_{e \in E^L} x_e + \sum_{e \notin E^L, E^H} x_e + \sum_{e \in E^H} x_e$$

$$\geq \sum_{e \in E^L} \frac{1}{\alpha+1} + \sum_{e \notin E^L, E^H} \frac{1}{d(h)} + \sum_{e \in E^H} \left(\frac{1}{d(h_1)} + \frac{1}{d(h_2)} - \frac{1}{\alpha+1} \right)$$

$$= \sum_{e \in E^L} \frac{1}{\alpha+1} + \sum_{h \in V^H} \underbrace{\sum_{e : h \in e} \frac{1}{d(h)}}_{1} - \sum_{e \in E^H} \frac{1}{\alpha+1}$$

$$= \frac{|E^L|}{\alpha+1} + |V^H| - \frac{|E^H|}{\alpha+1}$$

Structural Results: Σ_{dyn} and Σ_{adj}: Lemma 1

$$\sum_{e} x_e = \sum_{e \in E^L} x_e + \sum_{e \notin E^L, E^H} x_e + \sum_{e \in E^H} x_e$$

$$\geq \sum_{e \in E^L} \frac{1}{\alpha+1} + \sum_{e \notin E^L, E^H} \frac{1}{d(h)} + \sum_{e \in E^H} \left(\frac{1}{d(h_1)} + \frac{1}{d(h_2)} - \frac{1}{\alpha+1} \right)$$

$$= \sum_{e \in E^L} \frac{1}{\alpha+1} + \sum_{h \in V^H} \underbrace{\sum_{e: h \in e} \frac{1}{d(h)}}_{1} - \sum_{e \in E^H} \frac{1}{\alpha+1}$$

$$= \frac{|E^L|}{\alpha+1} + |V^H| - \frac{|E^H|}{\alpha+1}$$

Therefore:

$$\Sigma_{adj} = |E^L| + |V^H|(\alpha+1) - |E^H| \leq (\alpha+1) \sum_{e} x_e = \Sigma_{dyn}$$

Structural Results: Σ_{dyn} and Σ_{adj}: Lemma 2

Lemma

$$\Sigma_{dyn} = (\alpha + 1) \sum_e x_e \leq (\alpha + 2)\,\mathsf{match}(G)$$

Fun fact (from Edmond's thm)

For any fractional matching with $z_e \leq \lambda$ for all e,

$$\sum_e z_e \leq (1 + \lambda)\,\mathsf{match}(G)$$

Structural Results: Σ_{dyn} and Σ_{adj}: Lemma 2

1. $\{x_e\}_{e \in E}$ is a fractional matching
2. $x_e \leq 1/(\alpha+1)$ for all e

From the fact:

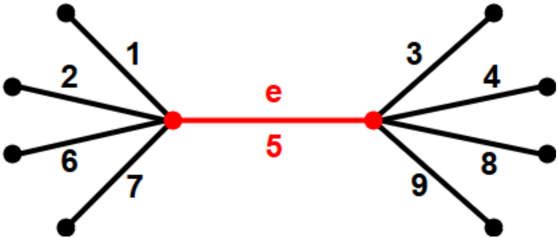
$$\sum_e x_e \leq \left(1 + \frac{1}{\alpha+1}\right) \text{match}(G) = \frac{\alpha+2}{\alpha+1} \text{match}(G)$$

Therefore:

$$\Sigma_{dyn} = (\alpha+1) \sum_e x_e \leq (\alpha+2)\, \text{match}(G)$$

Structural Results: Definitions: Σ_{ins}

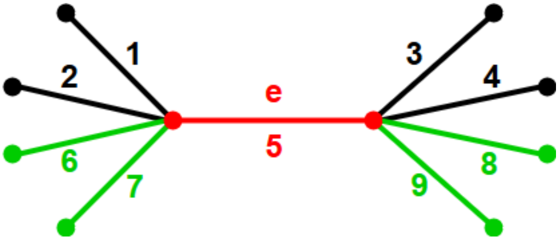
Let E_α be the set of edges uv where the number of edges incident to u or v that appear in the stream after uv are both at most α.



$\alpha = 3$

Structural Results: Definitions: Σ_{ins}

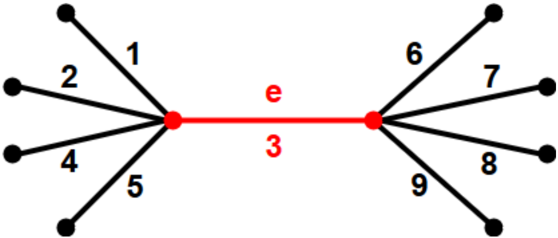
Let E_α be the set of edges uv where the number of edges incident to u or v that appear in the stream after uv are both at most α.



$\alpha = 3$
$e \in E_\alpha$

Structural Results: Definitions: Σ_{ins}

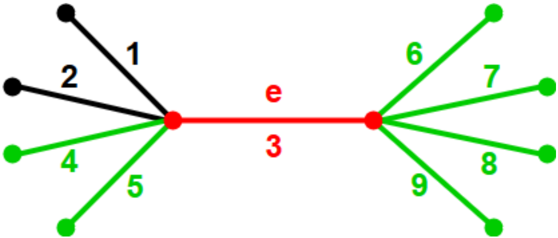
Let E_α be the set of edges *uv* where the number of edges incident to *u* or *v* that appear in the stream after *uv* are both at most α.



$\alpha = 3$

Structural Results: Definitions: Σ_{ins}

Let E_α be the set of edges *uv* where the number of edges incident to *u* or *v* that appear in the stream after *uv* are both at most α.



$\alpha = 3$
$e \notin E_\alpha$
E_α depends on stream ordering

Structural Results: Definitions: Σ_{ins}

Lemma 3

$$\text{match}(G) \leq |E_\alpha| \leq (\alpha + 2)\,\text{match}(G)$$

Let G_t be the graph defined by the first t edges in the stream.

Let E_α^t be $E_\alpha(G_t)$. Then

$$\text{match}(G_t) \leq |E_\alpha^t| \leq (\alpha + 2)\,\text{match}(G_t)$$

Let $\Sigma_{ins} = \max_t |E_\alpha^t| = |E_\alpha^T|$.

Since $\text{match}(G_t)$ is non-decreasing function of t,

$$\text{match}(G) \leq |E_\alpha| \leq \Sigma_{ins} = |E_\alpha^T| \leq (\alpha+2)\,\text{match}(G_T) \leq (\alpha+2)\,\text{match}(G)$$

Structural Results: Σ_{ins}: Lemma 3

Lemma

$$\mathrm{match}(G) \le |E_\alpha| \le (\alpha + 2)\,\mathrm{match}(G)$$

Upper bound

Fun fact (from Edmond's thm)

For any fractional matching with $z_e \le \lambda$ for all e,

$$\sum_e z_e \le (1 + \lambda)\,\mathrm{match}(G)$$

Let

$$y_e = \begin{cases} 1/(\alpha + 1) & \text{if } e \in E_\alpha \\ 0 & \text{otherwise} \end{cases}$$

$\{y_e\}_{e \in E}$ is a fractional matching with max weight $1/(\alpha + 1)$. Thus,

$$\frac{|E_\alpha|}{\alpha + 1} = \sum_e y_e \le \frac{\alpha + 2}{\alpha + 1} \cdot \mathrm{match}(G)$$

Structural Results: Σ_{ins}: Lemma 3

Lemma

$$\text{match}(G) \leq |E_\alpha| \leq (\alpha + 2)\,\text{match}(G)$$

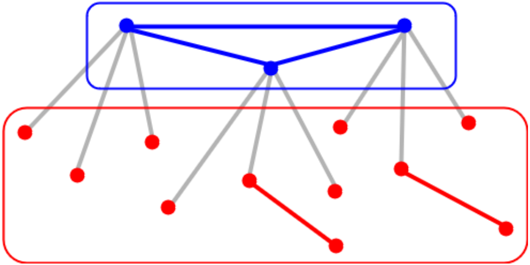
Lower bound

Structural Results: Σ_{ins}: Lemma 3

Lemma

$$\text{match}(G) \leq |E_\alpha| \leq (\alpha+2)\,\text{match}(G)$$

Lower bound



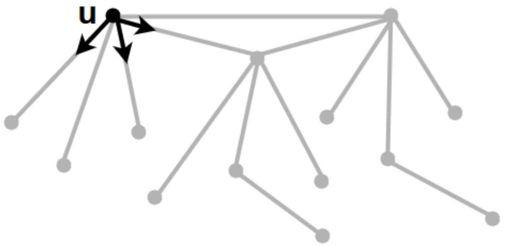
V^H = heavy vertices of degree $\geq \alpha + 2$
E^H = heavy edges with 2 heavy endpoints
V^L = light vertices
E^L = light edges

Structural Results: Σ_{ins}: Lemma 3

Lemma

$$\text{match}(G) \leq |E_\alpha| \leq (\alpha + 2)\,\text{match}(G)$$

Lower bound



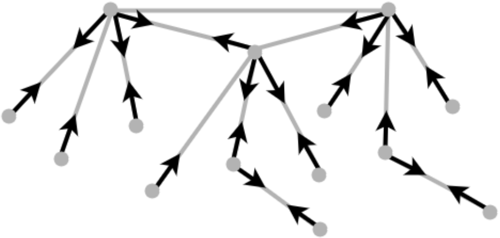
B_u = last $\alpha + 1$ edges on u in the stream

Structural Results: Σ_{ins}: Lemma 3

Lemma

$$\text{match}(G) \leq |E_\alpha| \leq (\alpha + 2)\,\text{match}(G)$$

Lower bound



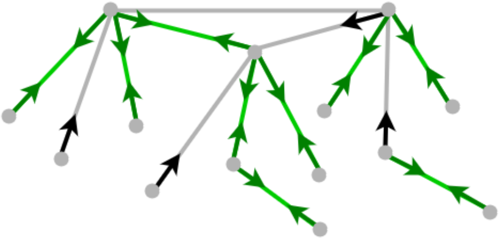
B_u = last $\alpha + 1$ edges on u in the stream

Structural Results: Σ_{ins}: Lemma 3

Lemma

$$\text{match}(G) \leq |E_\alpha| \leq (\alpha + 2)\,\text{match}(G)$$

Lower bound



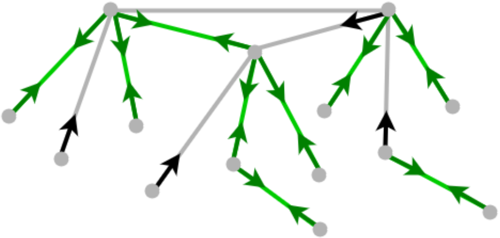
Edge uv is good if $uv \in B_u$ and $uv \in B_v$

Structural Results: Σ_{ins}: Lemma 3

Lemma

$$\text{match}(G) \leq |E_\alpha| \leq (\alpha + 2)\,\text{match}(G)$$

Lower bound



Edge uv is good if $uv \in B_u$ and $uv \in B_v$

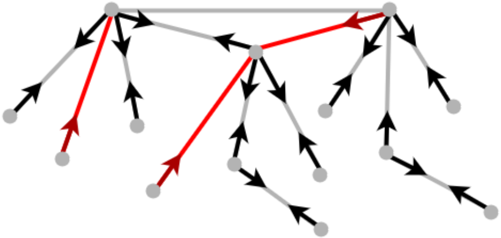
g_i is the number of good edges with i heavy endpoints

Structural Results: Σ_{ins}: Lemma 3

Lemma

$$\text{match}(G) \leq |E_\alpha| \leq (\alpha + 2)\,\text{match}(G)$$

Lower bound



Edge uv is good if $uv \in B_u$ and $uv \in B_v$
g_i is the number of good edges with i heavy endpoints

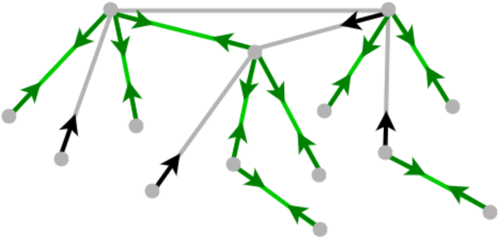
Edge uv is wasted if $uv \in B_u$ or $uv \in B_v$, but not both
w_2 is the number of wasted edges with 2 heavy endpoints

Structural Results: Σ_{ins}: Lemma 3

Lemma

$$\text{match}(G) \leq |E_\alpha| \leq (\alpha + 2)\,\text{match}(G)$$

Lower bound



(1)

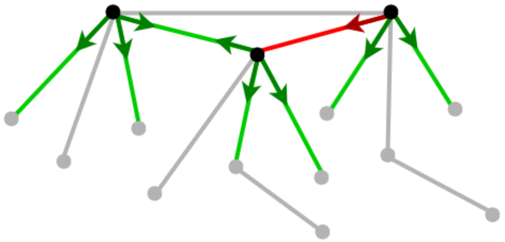
$$|E_\alpha| = g_0 + g_1 + g_2$$

Structural Results: Σ_{ins}: Lemma 3

Lemma

$$\mathrm{match}(G) \leq |E_\alpha| \leq (\alpha+2)\,\mathrm{match}(G)$$

Lower bound



(2)

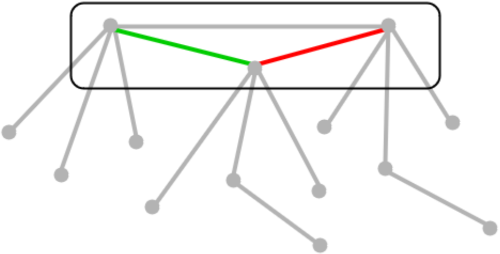
$$(\alpha+1)|V^H| = \sum_{h \in V^H} |B_h| = g_1 + 2g_2 + w_2$$

Structural Results: Σ_{ins}: Lemma 3

Lemma

$$\text{match}(G) \leq |E_\alpha| \leq (\alpha + 2)\,\text{match}(G)$$

Lower bound



(3)

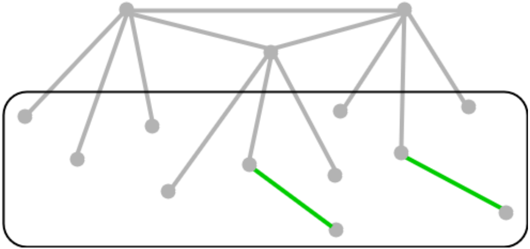
$$\alpha|V^H| \geq |E^H| \geq g_2 + w_2$$

Structural Results: Σ_{ins}: Lemma 3

Lemma

$$\text{match}(G) \leq |E_\alpha| \leq (\alpha + 2)\,\text{match}(G)$$

Lower bound



(4)

$$|E^L| = g_0$$

Structural Results: Σ_{ins}: Lemma 3

Lower bound

(1) $|E_\alpha| = g_0 + g_1 + g_2$
(2) $(\alpha + 1)|V^H| = g_1 + 2g_2 + w_2$
(3) $\alpha|V^H| \geq g_2 + w_2$
(4) $|E^L| = g_0$

Structural Results: Σ_{ins}: Lemma 3

Lower bound

(1) $|E_\alpha| = g_0 + g_1 + g_2$
(2) $(\alpha + 1)|V^H| = g_1 + 2g_2 + w_2$
(3) $\alpha|V^H| \geq g_2 + w_2$
(4) $|E^L| = g_0$

$$|E_\alpha| = g_0 + g_1 + g_2 \tag{1}$$

Structural Results: Σ_{ins}: Lemma 3

Lower bound

(1) $|E_\alpha| = g_0 + g_1 + g_2$
(2) $(\alpha + 1)|V^H| = g_1 + 2g_2 + w_2$
(3) $\alpha|V^H| \geq g_2 + w_2$
(4) $|E^L| = g_0$

$$\begin{aligned} |E_\alpha| &= g_0 + g_1 + g_2 && (1) \\ &= |E^L| + g_1 + g_2 && (4) \end{aligned}$$

Structural Results: Σ_{ins}: Lemma 3

Lower bound

(1) $|E_\alpha| = g_0 + g_1 + g_2$
(2) $(\alpha + 1)|V^H| = g_1 + 2g_2 + w_2$
(3) $\alpha|V^H| \geq g_2 + w_2$
(4) $|E^L| = g_0$

$$\begin{aligned} |E_\alpha| &= g_0 + g_1 + g_2 && (1) \\ &= |E^L| + g_1 + g_2 && (4) \\ &= |E^L| + (g_1 + 2g_2 + w_2) - (g_2 + w_2) \end{aligned}$$

Structural Results: Σ_{ins}: Lemma 3

Lower bound

(1) $|E_\alpha| = g_0 + g_1 + g_2$

(2) $(\alpha + 1)|V^H| = g_1 + 2g_2 + w_2$

(3) $\alpha|V^H| \geq g_2 + w_2$

(4) $|E^L| = g_0$

$$\begin{aligned} |E_\alpha| &= g_0 + g_1 + g_2 && (1) \\ &= |E^L| + g_1 + g_2 && (4) \\ &= |E^L| + (g_1 + 2g_2 + w_2) - (g_2 + w_2) \\ &= |E^L| + (\alpha + 1)|V^H| - (g_2 + w_2) && (2) \end{aligned}$$

Structural Results: Σ_{ins}: Lemma 3

Lower bound

(1) $|E_\alpha| = g_0 + g_1 + g_2$
(2) $(\alpha + 1)|V^H| = g_1 + 2g_2 + w_2$
(3) $\alpha|V^H| \geq g_2 + w_2$
(4) $|E^L| = g_0$

$$\begin{aligned}
|E_\alpha| &= g_0 + g_1 + g_2 && (1)\\
&= |E^L| + g_1 + g_2 && (4)\\
&= |E^L| + (g_1 + 2g_2 + w_2) - (g_2 + w_2) &&\\
&= |E^L| + (\alpha + 1)|V^H| - (g_2 + w_2) && (2)\\
&\geq |E^L| + (\alpha + 1)|V^H| - \alpha|V^H| && (3)
\end{aligned}$$

Structural Results: Σ_{ins}: Lemma 3

Lower bound

(1) $|E_\alpha| = g_0 + g_1 + g_2$
(2) $(\alpha + 1)|V^H| = g_1 + 2g_2 + w_2$
(3) $\alpha|V^H| \geq g_2 + w_2$
(4) $|E^L| = g_0$

$$\begin{aligned}
|E_\alpha| &= g_0 + g_1 + g_2 & (1)\\
&= |E^L| + g_1 + g_2 & (4)\\
&= |E^L| + (g_1 + 2g_2 + w_2) - (g_2 + w_2) & \\
&= |E^L| + (\alpha + 1)|V^H| - (g_2 + w_2) & (2)\\
&\geq |E^L| + (\alpha + 1)|V^H| - \alpha|V^H| & (3)\\
&= |E^L| + |V^H| &
\end{aligned}$$

Structural Results: Σ_{ins}: Lemma 3

Lower bound

(1) $|E_\alpha| = g_0 + g_1 + g_2$
(2) $(\alpha + 1)|V^H| = g_1 + 2g_2 + w_2$
(3) $\alpha|V^H| \geq g_2 + w_2$
(4) $|E^L| = g_0$

$$\begin{aligned}
|E_\alpha| &= g_0 + g_1 + g_2 & (1)\\
&= |E^L| + g_1 + g_2 & (4)\\
&= |E^L| + (g_1 + 2g_2 + w_2) - (g_2 + w_2) & \\
&= |E^L| + (\alpha + 1)|V^H| - (g_2 + w_2) & (2)\\
&\geq |E^L| + (\alpha + 1)|V^H| - \alpha|V^H| & (3)\\
&= |E^L| + |V^H| & \\
&\geq \text{match}(G) &
\end{aligned}$$

Structural Results: Σ_{ins}: Lemma 3

Lower bound

(1) $|E_\alpha| = g_0 + g_1 + g_2$
(2) $(\alpha + 1)|V^H| = g_1 + 2g_2 + w_2$
(3) $\alpha|V^H| \geq g_2 + w_2$
(4) $|E^L| = g_0$

$$
\begin{aligned}
|E_\alpha| &= g_0 + g_1 + g_2 && (1)\\
&= |E^L| + g_1 + g_2 && (4)\\
&= |E^L| + (g_1 + 2g_2 + w_2) - (g_2 + w_2) &&\\
&= |E^L| + (\alpha + 1)|V^H| - (g_2 + w_2) && (2)\\
&\geq |E^L| + (\alpha + 1)|V^H| - \alpha|V^H| && (3)\\
&= |E^L| + |V^H| &&\\
&\geq \text{match}(G) &&
\end{aligned}
$$

Algorithms

Algorithms: Dynamic Stream

$$\Sigma_{dyn} = (1+\alpha)\sum_e x_e = (1+\alpha)\sum_e \min\left(\frac{1}{d(u)}, \frac{1}{d(v)}, \frac{1}{\alpha+1}\right)$$

In parallel:

If matching has size $\leq n^{2/5}$,

- Use algorithm for bounded size matchings [CCEHMMV16]: $\tilde{O}(n^{4/5})$ space

If matching has size $> n^{2/5}$,

- Sample a set of vertices T with probability $p = \tilde{\Theta}(1/n^{1/5})$
- Compute degrees of vertices in T
- Let E_T be edges with both endpoints in T
- Sample $\min(|E_T|, \tilde{\Theta}(\alpha n^{4/5}))$ edges in E_T
- Use $(\alpha+1)/p \cdot \sum_{e \in E_T} x_e$ as estimate

Note: In insert-only streams, can use greedy algorithm for approximating small matching. Reduces total space to $\tilde{O}(\alpha n^{2/3})$.

Algorithms: Insert-only Stream

$$\Sigma_{ins} = \max_t |E_\alpha^t|$$

where E_α^t is the set of edges uv, s.t. the number of edges incident to u or v between arrival of uv and time t is at most α.

1. Set $p \leftarrow 1$
2. Start sampling each edge with probability p
3. If e is sampled:
   - store e
   - store counters for degrees of endpoints in the rest of the stream
   - if later we detect $e \notin E_\alpha^t$, it is deleted
4. If the number of stored edges $> 40\epsilon^{-2} \log n$
   - $p \leftarrow p/2$
   - delete every edge currently stored with probability $1/2$
5. Return $\max_t \frac{\#\text{ samples at time } t}{p \text{ at time } t}$

Algorithms: Insert-only Stream

$$\Sigma_{ins} = \max_t |E_\alpha^t|$$

where E_α^t is the set of edges uv, s.t. the number of edges incident to u or v between arrival of uv and time t is at most α.

Let k be s.t. $(20\epsilon^{-2} \log n)2^{k-1} \leq \Sigma_{ins} < (20\epsilon^{-2} \log n)2^k$.

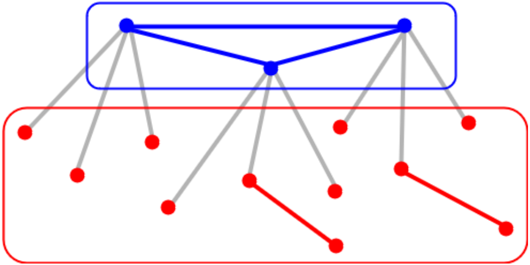
We show that whp:

1. If sampling probability is high enough ($\geq 1/2^k$), can compute $|E_\alpha^t| \pm \epsilon\Sigma_{ins}$ for all t. From Chernoff and union bounds.
2. We do not switch to probability that is too low ($< 1/2^k$), since the # edges sampled wp $1/2^k$ does not exceed $(1+\epsilon)\Sigma_{ins}/2^k < (1+\epsilon)(20\epsilon^{-2} \log n) \leq 40\epsilon^{-2} \log n$.

Algorithms: Adjacency List Stream

$$\Sigma_{adj} = |E^L| + |V^H|(\alpha + 1) - |E^H|$$

Treat adjacency stream as a degree sequence of the graph.
$|V^H|$ can be computed easily.



$$|E^L| - |E^H| = |E| - \sum_{h \in V^H} d(h)$$

which is also easy to compute.

Conclusion

Summary:

- There are quantities that provide good approximation of the size of maximum matching in graphs of arboricity α.
- Computing those quantities can be done efficiently.

Open questions:

- Better than $\alpha + 2$ approximation.
- Closing the gap between upper and lower bounds for dynamic streams.

Thank you for your attention!