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

<table>
<thead>
<tr>
<th></th>
<th>insertions</th>
<th>deletions</th>
<th>arbitrary order</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>dynamic</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>insert-only</strong></td>
<td>✓</td>
<td>🚭</td>
<td>✓</td>
</tr>
<tr>
<td><strong>adjacency-list</strong></td>
<td>✓</td>
<td>🚭</td>
<td>🚭</td>
</tr>
</tbody>
</table>

*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 $\alpha$.

**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 $\alpha r$ edges.

Planar graphs have arboricity at most 3.

In dynamic stream, intermediate graphs can have high arboricity.
## Results

<table>
<thead>
<tr>
<th></th>
<th>space</th>
<th>approx factor</th>
<th>work</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>dynamic</strong></td>
<td>$\tilde{O}(\alpha n^{4/5})$</td>
<td>$(5\alpha + 9)(1 + \epsilon)$</td>
<td>CCEHMMMV16</td>
</tr>
<tr>
<td></td>
<td>$\tilde{O}(\alpha n^{4/5})$</td>
<td>$(\alpha + 2)(1 + \epsilon)$</td>
<td>MV16</td>
</tr>
<tr>
<td></td>
<td>$\tilde{O}(\alpha^{10/3} n^{2/3})$</td>
<td>$(22.5\alpha + 6)(1 + \epsilon)$</td>
<td>CJMM17*</td>
</tr>
<tr>
<td></td>
<td>$\Omega(\sqrt{n}/\alpha^{2.5})$</td>
<td>$O(\alpha)$</td>
<td>AKL17</td>
</tr>
<tr>
<td><strong>insert-only</strong></td>
<td>$\tilde{O}(\alpha n^{2/3})$</td>
<td>$(5\alpha + 9)(1 + \epsilon)$</td>
<td>EHLMO15</td>
</tr>
<tr>
<td></td>
<td>$\tilde{O}(\alpha n^{2/3})$</td>
<td>$(\alpha + 2)(1 + \epsilon)$</td>
<td>MV16</td>
</tr>
<tr>
<td></td>
<td>$O(\alpha \epsilon^{-3} \log^2 n)$</td>
<td>$(22.5\alpha + 6)(1 + \epsilon)$</td>
<td>CJMM17</td>
</tr>
<tr>
<td></td>
<td>$O(\epsilon^{-2} \log n)$</td>
<td>$(\alpha + 2)(1 + \epsilon)$</td>
<td>MV18</td>
</tr>
<tr>
<td><strong>adj</strong></td>
<td>$O(1)$</td>
<td>$\alpha + 2$</td>
<td>MV16</td>
</tr>
</tbody>
</table>

*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 $\Sigma$ that is an $(\alpha + 2)$ approximation of $\text{match}(G)$

- **Algorithm**: $(1 + \epsilon)$ approximation of $\Sigma$ in streaming (exact computation in adjacency list stream)

**Dynamic**: $\Sigma_{dyn}$

- $(1 + \epsilon)$-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**: $\Sigma_{ins}$

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

**Adjacency list**: $\Sigma_{adj}$

- 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: $\Sigma_{adj}$

$$\Sigma_{adj} = |E^L| + |V^H|(\alpha + 1) - |E^H|$$
Structural Results: Definitions: $\Sigma_{dyn}$

$$x_e = x_{uv} = \min \left( \frac{1}{d(u)}, \frac{1}{d(v)}, \frac{1}{\alpha + 1} \right)$$
Structural Results: Definitions: $\Sigma_{dyn}$

$$x_e = x_{uv} = \min \left( \frac{1}{d(u)}, \frac{1}{d(v)}, \frac{1}{\alpha + 1} \right)$$
Structural Results: Definitions: $\Sigma_{\text{dyn}}$

\[ x_e = x_{uv} = \min \left( \frac{1}{d(u)}, \frac{1}{d(v)}, \frac{1}{\alpha + 1} \right) \]
Structural Results: Definitions: $\Sigma_{dyn}$

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

\[ \Sigma_{dyn} = (\alpha + 1) \sum_e x_e \]
**Structural Results:** \( \Sigma_{\text{dyn}} \) and \( \Sigma_{\text{adj}} \)

\[
\text{match}(G) \leq |V^H| + |E^L|
\]

since a matched edge is either light or incident to a heavy vertex

\[
\leq |E^L| + |V^H| (\alpha + 1) - |E^H| = \Sigma_{\text{adj}}
\]

since \( |E^H| \leq \alpha |V^H| \)

\[
\leq (\alpha + 1) \sum_e x_e = \Sigma_{\text{dyn}}
\]

Lemma 1

\[
\leq (\alpha + 2) \text{match}(G)
\]

Lemma 2
Structural Results: $\Sigma_{dyn}$ and $\Sigma_{adj}$: Lemma 1

Lemma

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

**Light edge:**

$$x_e = \min \left( \frac{1}{d(l_1)} \geq \frac{1}{\alpha + 1}, \frac{1}{d(l_2)} \geq \frac{1}{\alpha + 1}, \frac{1}{\alpha + 1} \right)$$
Structural Results: $\Sigma_{dyn}$ and $\Sigma_{adj}$: Lemma 1

Light edge:

\[
x_e = \min \left( \frac{1}{d(l_1)} \geq \frac{1}{\alpha + 1}, \frac{1}{d(l_2)} \geq \frac{1}{\alpha + 1}, \frac{1}{\alpha + 1} \right) = \frac{1}{\alpha + 1}
\]
Structural Results: $\Sigma_{dyn}$ and $\Sigma_{adj}$: Lemma 1

Light edge:

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

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: $\Sigma_{dyn}$ and $\Sigma_{adj}$: Lemma 1

Light edge:

$$x_e = \min \left( \frac{1}{d(l_1)} \geq \frac{1}{\alpha + 1}, \frac{1}{d(l_2)} \geq \frac{1}{\alpha + 1}, \frac{1}{\alpha + 1} \right) = \frac{1}{\alpha + 1}$$

Edge with 1 light and 1 heavy endpoints:

$$x_e = \min \left( \frac{1}{d(l)} \geq \frac{1}{\alpha + 1}, \frac{1}{d(h)} < \frac{1}{\alpha + 1}, \frac{1}{\alpha + 1} \right)$$
Light edge:

\[ x_e = \min \left( \frac{1}{d(l_1)} \geq \frac{1}{\alpha + 1}, \frac{1}{d(l_2)} \geq \frac{1}{\alpha + 1}, \frac{1}{\alpha + 1} \right) = \frac{1}{\alpha + 1} \]

Edge with 1 light and 1 heavy endpoints:

\[ x_e = \min \left( \frac{1}{d(l)} \geq \frac{1}{\alpha + 1}, \frac{1}{d(h)} < \frac{1}{\alpha + 1}, \frac{1}{\alpha + 1} \right) = \frac{1}{d(h)} \]
Heavy edge:

\[ x_e = \min \left( \frac{1}{d(h_1)}, \frac{1}{d(h_2)}, \frac{1}{\alpha + 1} \right) \]
Heavy edge:

\[ x_e = \min \left( \frac{1}{d(h_1)} < \frac{1}{\alpha + 1}, \frac{1}{d(h_2)} < \frac{1}{\alpha + 1}, \frac{1}{\alpha + 1} \right) \]
**Structural Results: $\Sigma_{dyn}$ and $\Sigma_{adj}$: Lemma 1**

Heavy edge:

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

$$= \min \left( \frac{1}{d(h_1)}, \frac{1}{d(h_2)} \right)$$
Structural Results: $\Sigma_{\text{dyn}}$ and $\Sigma_{\text{adj}}$: Lemma 1

Heavy edge:

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

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

\[ = \frac{1}{d(h_1)} + \frac{1}{d(h_2)} - \max \left( \frac{1}{d(h_1)}, \frac{1}{d(h_2)} \right) \]
Heavy edge:

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

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

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

\[ > \frac{1}{d(h_1)} + \frac{1}{d(h_2)} - \frac{1}{\alpha + 1} \]
Structural Results: $\Sigma_{dyn}$ and $\Sigma_{adj}$: 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: \( \Sigma_{dy\!n} \) and \( \Sigma_{adj} \): Lemma 1

\[
\sum_{e} x_e = \sum_{e \in E^L} x_e + \sum_{e \not\in E^L, E^H} x_e + \sum_{e \in E^H} x_e \\
\geq \sum_{e \in E^L} \frac{1}{\alpha + 1} + \sum_{e \not\in E^L, E^H} \frac{1}{d(h)} + \sum_{e \in E^H} \left( \frac{1}{d(h_1)} + \frac{1}{d(h_2)} - \frac{1}{\alpha + 1} \right)
\]
 Structural Results: $\Sigma_{dyn}$ and $\Sigma_{adj}$: Lemma 1

\[ \sum_{e} x_e = \sum_{e \in E^L} x_e + \sum_{e \not\in E^L, E^H} x_e + \sum_{e \in E^H} x_e \]

\[ \geq \sum_{e \in E^L} \frac{1}{\alpha + 1} + \sum_{e \not\in E^L, E^H} \frac{1}{d(h)} + \sum_{e \in E^H} \left( \frac{1}{d(h_1)} + \frac{1}{d(h_2)} - \frac{1}{\alpha + 1} \right) \]

\[ = \sum_{e \in E^L} \frac{1}{\alpha + 1} + \sum_{h \in V^H} \sum_{e: h \in e} \frac{1}{d(h)} - \sum_{e \in E^H} \frac{1}{\alpha + 1} \]
**Structural Results:** \( \Sigma_{dyn} \) and \( \Sigma_{adj} \): Lemma 1

\[
\sum_{e} x_{e} = \sum_{e \in E^{L}} x_{e} + \sum_{e \notin E^{L}, E^{H}} x_{e} + \sum_{e \in E^{H}} x_{e} \geq \sum_{e \in E^{L}} \frac{1}{\alpha + 1} + \sum_{e \notin E^{L}, E^{H}} \frac{1}{d(h)} + \sum_{e \in E^{H}} \left( \frac{1}{d(h_{1})} + \frac{1}{d(h_{2})} - \frac{1}{\alpha + 1} \right) \]

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

\[
= \frac{|E^{L}|}{\alpha + 1} + |V^{H}| - \frac{|E^{H}|}{\alpha + 1}
\]
Structural Results: $\Sigma_{dyn}$ and $\Sigma_{adj}$: Lemma 1

\[
\sum_{e} x_e = \sum_{e \in E^L} x_e + \sum_{e \notin E^L, E^H} x_e + \sum_{e \in E^H} x_e \\
\geq \sum_{e \in E^L} \frac{1}{\alpha + 1} + \sum_{e \notin E^L, E^H} \frac{1}{d(h)} + \sum_{e \in E^H} \left( \frac{1}{d(h_1)} + \frac{1}{d(h_2)} - \frac{1}{\alpha + 1} \right) \\
= \sum_{e \in E^L} \frac{1}{\alpha + 1} + \sum_{h \in V^H} \sum_{e:h \in e} \frac{1}{d(h)} - \sum_{e \in E^H} \frac{1}{\alpha + 1} \\
= \frac{|E^L|}{\alpha + 1} + |V^H| - \frac{|E^H|}{\alpha + 1}
\]

Therefore:

$\Sigma_{adj} = |E^L| + |V^H| (\alpha + 1) - |E^H| \leq (\alpha + 1) \sum_{e} x_e = \Sigma_{dyn}$
Structural Results: $\Sigma_{dyn}$ and $\Sigma_{adj}$: Lemma 2

Lemma

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

Fun fact (from Edmond’s thm)

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

$$\sum_{e} z_e \leq (1 + \lambda) \text{match}(G)$$
Structural Results: $\Sigma_{dyn}$ and $\Sigma_{adj}$: 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:

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

$\alpha = 3$
Let $E_\alpha$ 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 $\alpha$.

$\alpha = 3$

$e \in E_\alpha$
Structural Results: Definitions: $\Sigma_{ins}$

Let $E_\alpha$ 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 $\alpha$.

$\alpha = 3$
Structural Results: Definitions: $\Sigma_{ins}$

Let $E_\alpha$ 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 $\alpha$.

\[
\alpha = 3 \\
e \notin E_\alpha \\
E_\alpha \text{ depends on stream ordering}
\]
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 \( E_{\alpha}^t \) be \( E_\alpha(G_t) \). Then

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

Let \( \Sigma_{\text{ins}} = \max_t |E_{\alpha}^t| = |E_{\alpha}^T| \).

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

\[ \text{match}(G) \leq |E_\alpha| \leq \Sigma_{\text{ins}} = |E_{\alpha}^T| \leq (\alpha+2) \text{match}(G_T) \leq (\alpha+2) \text{match}(G) \]
Structural Results: $\Sigma_{ins}$: Lemma 3

**Lemma**

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

**Upper bound**

**Fun fact (from Edmond’s thm)**

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

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

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,

$$\frac{|E_\alpha|}{\alpha + 1} = \sum_e y_e \leq \frac{\alpha + 2}{\alpha + 1} \cdot \text{match}(G)$$
Structural Results: $\Sigma_{ins}$: Lemma 3

Lemma

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

Lower bound
Structural Results: $\Sigma_{ins}$: Lemma 3

Lemma

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

Lower bound

$$V^H = \text{heavy vertices of degree } \geq \alpha + 2$$
$$E^H = \text{heavy edges with 2 heavy endpoints}$$
$$V^L = \text{light vertices}$$
$$E^L = \text{light edges}$$
Structural Results: $\Sigma_{ins}$: Lemma 3

**Lemma**

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

**Lower bound**

$$B_u = \text{last } \alpha + 1 \text{ edges on } u \text{ in the stream}$$
Structural Results: $\Sigma_{ins}$: Lemma 3

**Lemma**

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

**Lower bound**

$$B_u = \text{last } \alpha + 1 \text{ edges on } u \text{ in the stream}$$
Structural Results: $\Sigma_{ins}$: 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: $\Sigma_{ins}$: 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**: $\Sigma_{ins}$: 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: $\Sigma_{ins}$: 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: $\Sigma_{ins}$: Lemma 3

**Lemma**

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

**Lower bound**

$$\sum_{h \in V^H} |B_h| = g_1 + 2g_2 + w_2$$
Structural Results: Σ_{ins}: Lemma 3

Lemma

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: \( \Sigma_{ins} \): Lemma 3**

**Lemma**

\[
\text{match}(G) \leq |E_\alpha| \leq (\alpha + 2) \text{match}(G)
\]

**Lower bound**

![Diagram showing a lower bound expression](image)

(4)

\[
|E^L| = g_0
\]
Structural Results: $\Sigma_{ins}$: 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: $\Sigma_{ins}$: 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) $|E^L| = g_0$

\[ |E_\alpha| = g_0 + g_1 + g_2 \] (1)
Structural Results: $\Sigma_{ins}$: 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. $|E^L| = g_0$

$$|E_\alpha| = g_0 + g_1 + g_2$$  \hspace{1cm} (1)

$$= |E^L| + g_1 + g_2$$  \hspace{1cm} (4)
Structural Results: $\Sigma_{ins}$: 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. $|E^L| = g_0$

\[
|E_\alpha| = g_0 + g_1 + g_2 \\
= |E^L| + g_1 + g_2 \\
= |E^L| + (g_1 + 2g_2 + w_2) - (g_2 + w_2)
\]
Structural Results: $\Sigma_{\text{ins}}$: 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. $|E^L| = g_0$

\[
|E_\alpha| = g_0 + g_1 + g_2 = |E^L| + g_1 + g_2 = |E^L| + (g_1 + 2g_2 + w_2) - (g_2 + w_2) = |E^L| + (\alpha + 1)|V^H| - (g_2 + w_2)
\]
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) \(|E^L| = g_0\)

\[
|E_\alpha| = g_0 + g_1 + g_2 \\
= |E^L| + g_1 + g_2 \quad (1) \\
= |E^L| + (g_1 + 2g_2 + w_2) - (g_2 + w_2) \quad (4) \\
= |E^L| + (\alpha + 1)|V^H| - (g_2 + w_2) \quad (2) \\
\geq |E^L| + (\alpha + 1)|V^H| - \alpha|V^H| \quad (3)
\]
Structural Results: $\Sigma_{ins}$: 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. $|E^L| = g_0$

\[|E_\alpha| = g_0 + g_1 + g_2 \]
\[= |E^L| + g_1 + g_2 \] (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)
\[= |E^L| + |V^H| \]
Structural Results: $\Sigma_{ins}$: Lemma 3

**Lower bound**

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

\[|E_\alpha| = g_0 + g_1 + g_2\]
\[= |E^L| + g_1 + g_2\]
\[= |E^L| + (g_1 + 2g_2 + w_2) - (g_2 + w_2)\]
\[= |E^L| + (\alpha + 1)|V^H| - (g_2 + w_2)\]
\[\geq |E^L| + (\alpha + 1)|V^H| - \alpha|V^H|\]
\[= |E^L| + |V^H|\]
\[\geq \text{match}(G)\]
Structural Results: $\Sigma_{ins}$: 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) \[ |E^L| = g_0 \]

\[
|E_\alpha| = g_0 + g_1 + g_2 \\
= |E^L| + g_1 + g_2 \\
= |E^L| + (g_1 + 2g_2 + w_2) - (g_2 + w_2) \\
= |E^L| + (\alpha + 1)|V^H| - (g_2 + w_2) \\
\geq |E^L| + (\alpha + 1)|V^H| - \alpha|V^H| \\
= |E^L| + |V^H| \\
\geq \text{match}(G)
\]
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 \( p = \tilde{O}(1/n^{1/5}) \)
- Compute degrees of vertices in \( T \)
- Let \( E_T \) be edges with both endpoints in \( T \)
- Sample \( \min(|E_T|, \tilde{O}(\alpha n^{4/5})) \) 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

\[ \Sigma_{ins} = \max_t |E^t_\alpha| \]

where \( E^t_\alpha \) 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 \( \alpha \).

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 \( e \notin E^t_\alpha \), 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

\[ \Sigma_{ins} = \max_t |E_t^\alpha| \]

where \( E_t^\alpha \) 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 \( \alpha \).

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 \( |E_t^\alpha| \pm \epsilon \Sigma_{ins} \) 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 $\alpha$.
- 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!