Inapproximability of Treewidth, Graph-layout and One-shot Pebbling

Per Austrin, Toniann Pitassi, Yu Wu
University of Toronto

October 20, 2011

Problems we consider

- One-shot Pebbling game :
  Pebbling game on a DAG where each node can only be pebbled once.
- Graph layout problems:
  Minimum Linear Arrangement, Interval Graph Completion, etc.
- Width parameters: treewidth, pathwidth.

Our Contribution:

- Treat those problems in a unified way.
- Prove that assuming the Small Set Expansion (SSE) Conjecture, the above problems are hard to approximate within any constant factor.

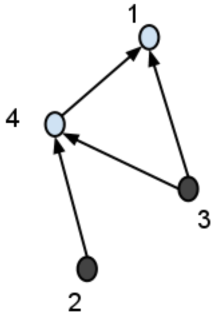
Problems we consider

- One-shot Pebbling game :
  Pebbling game on a DAG where each node can only be pebbled once.
- Graph layout problems:
  Minimum Linear Arrangement, Interval Graph Completion, etc.
- Width parameters: treewidth, pathwidth.

Our Contribution:

- Treat those problems in a unified way.
- Prove that assuming the Small Set Expansion (SSE) Conjecture, the above problems are hard to approximate within any constant factor.

One-shot Pebbling Game

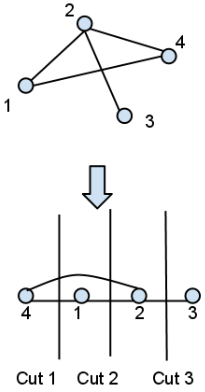


Given a DAG, pebble the sink node according to the following rules, while minimize # pebbles used.

- A pebbled can be placed on any source nodes.
- A pebble can be placed on a vertex v if all of the immediate predecessors of v are pebbled.
- A pebble can be removed from a vertex.

Additional rule: each vertex can be pebbled only once.

Graph Layout Problems

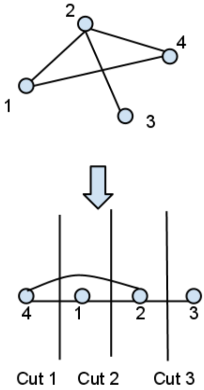


Given a graph $G = (V, E)$ $V = \{1, .., n\}$,
find a permutation π on V.
Define $\mathit{length}(u, v) = |\pi(u) - \pi(v)|$.

Minimum Linear Arrangement:

$$\min \sum_{(u,v)\in E} \mathit{length}(u, v)$$

Graph Layout Problems



Minimum Cut Linear Arrangement:

$$Cut_i(\pi) = \{e \in E | \pi(u) \leq i < \pi(v), e = (u, v)\}$$

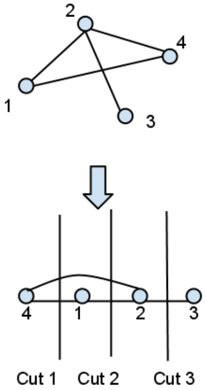
Want:

$$\min_{\pi} \max_{i \in [n]} |Cut_i(\pi)|$$

8 variations:

- Undirected / directed acyclic
- Counting edges or vertices at each cut
- Aggregation by sum or max

Graph Layout Problems



Minimum Cut Linear Arrangement:

$$Cut_i(\pi) = \{e \in E | \pi(u) \leq i < \pi(v), e = (u, v)\}$$

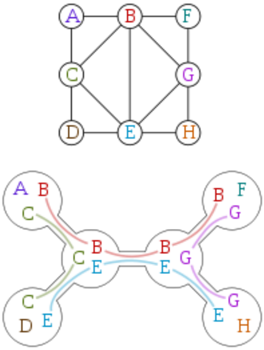
Want:

$$\min_{\pi} \max_{i \in [n]} |Cut_i(\pi)|$$

8 variations:

- Undirected / directed acyclic
- Counting edges or vertices at each cut
- Aggregation by sum or max

Width-parameters: treewidth, pathwidth



Cited from Wikipedia

$(T, \mathcal{V})$ is called a tree decomposition of G if:

(T1) $V = \cup_{t \in T} V_t$;

(T2) $\forall e \in E$, $\exists t \in T$, s.t. both endpoints of e lie in V_t;

(T3) for every $v \in V$, $\{t \in T \mid v \in V_t\}$ is a subtree of T.

$$Treewidth(G) = \min_T \max |V_t| - 1.$$

A complete list of problems covered

Problem			Also known as / Equivalent with
undir.	edge	sum	Minimum/Optimal Linear Arrangement
undir.	edge	max	MCLA, CutWidth
undir.	vertex	sum	Interval Graph Completion, SumCut
undir.	vertex	max	Pathwidth, One-shot Black-White Pebbling, Vertex Separation
DAG	edge	sum	Minimum Storage-Time Sequencing, Directed MLA/OLA
DAG	edge	max	
DAG	vertex	sum	
DAG	vertex	max	One-shot Black Pebbling,Register Sufficiency

Complexity Status

- All of the problems are **NP**-Complete (decision version).
- Approximation Algorithms known (optimization version) best known approximation ratio $poly(\log n)$.
- Very few hardness of approximation results.

Known approximation algorithms

- $O(\sqrt{\log n})$: Treewidth [Feige,Hajiaghayi and Lee 05']
- $O(\sqrt{\log n} \log \log n)$: Minimum Linear Arrangement, Interval Graph Completion, Directed MLA. [Charikar,Hajiaghayi,Karloff and Rao 10']
- $O(\log n \sqrt{\log n})$: MCLA, Vertex Separation (Pathwidth) [Feige,Hajiaghayi and Lee 05'], Register Sufficiency.

Known hardness result

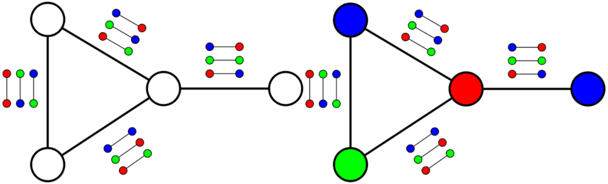
- no constant approximation for MLA assuming UGC with an additional condition (which is equivalent to SSE) [Devanur, Khot, and Saket 06', Raghavendra,Steurer and Tulsiani 10'].
- No PTAS for MLA, unless **NP**$\subseteq$ randomized subexponential time. [Ambuhl, Mastrolilli and Svensson 07']
- **NP**-hard to find a tree-decomposion with width $\mathit{Treewidth}(G) + n^{\epsilon}$ for some constant $0 < \epsilon < 1$. [Bodlaender, Gilbert, Hafsteinsson and Kloks 95']

Unique Games Conjecture

Definition

$Y = (V, E, \Pi, [R])$ consists of a graph $G = (V, E)$, a set of labels $[R] = \{1, ..., R\}$ and a set of permutations $\pi_{v \leftarrow w} : [R] \to [R]$ for each edge $e = (w, v) \in E$.

An assignment $F : V \to [R]$ of labels to vertices is said to satisfy an edge $e = (w, v)$, if $\pi_{v \leftarrow w}(F(w)) = F(v)$.



Unique Games Conjecture

Problem (Unique Games $(R, 1-\epsilon, \eta)$).
$Y = (V, E, \Pi = \{\pi_{v \leftarrow w} : [R] \to [R] | e = (w, v) \in E\}, [R])$,
distinguish between the following:

Yes There exists a labeling assignment satisfies a fraction of $(1-\epsilon)$ edges.

No No labeling assignment satisfies a fraction of η edges.

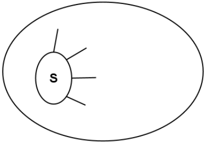
Conjecture (Unique Games Conjecture (Khot 02))

For all constants $\epsilon, \eta > 0$, there exists large enough constant R such that Unique Games $(R, 1-\epsilon, \eta)$ is **NP***-hard.*

Some facts about UGC

- Implies that approximating Vertex Cover within a factor of $2 - \epsilon$ is hard. [Khot and Regev 08]
- Implies that SDP approximation algorithms for several problems, such as MAX-CUT, are optimal. [Khot, Kindler, Mossel and O'Donnell 07]
- Proving / Disproving UGC is a hot topic in complexity.

Small Set Expansion (SSE) Conjecture



$G = (V, E)$: undirected, d-regular graph
$\Phi_G(S)$: the (normalized) edge expansion of S, for $S \subseteq V$

$$\Phi_G(S) = \frac{|E(S, \bar{S})|}{d|S|}$$

The SSE Problem asks if G has a small set S which does not expand or whether all small sets are highly expanding.
[Raghavendra and Steurer 10']

The Small Set Expansion Conjecture

- $G = (V, E)$, $\mathrm{SSE}(\eta, \delta)$ is the problem of distinguishing:
  - Yes There is an $S \subseteq V$ with $|S| = \delta|V|$ and $\Phi_G(S) \leq \eta$.
  - No For every $S \subseteq V$ with $|S| = \delta|V|$ it holds that $\Phi_G(S) \geq 1 - \eta$.
- Small Set Expansion Conjecture
  For every $\eta > 0$, there is a $\delta > 0$ such that $\mathrm{SSE}(\eta, \delta)$ is **NP**-hard.

Some facts about SSE

- SSE-hardness implies UGC-hardness.
- It is equivalent to UGC with an additional condition on the expansion of the graph.[Raghavendra, Steurer and Tulsiani 10']
- Can be solved in subexponential time (also true for Unique Games). [Arora, Barak and Steurer 11'], [Barak, Raghavendra and Steurer 11']

Strong version of the SSE conjecture

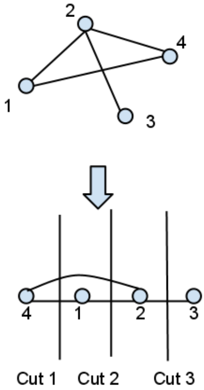
SSE Conjecture (strong form)

For every integer $q > 0$ and $\epsilon, \gamma > 0$, it is NP-hard to distinguish between the following two cases for a given regular graph $G = (V, E)$

Yes There is a partition of V into q equi-sized sets $S_1, \ldots, S_q$ such that $\Phi_G(S_i) \leq 2\epsilon$ for every $1 \leq i \leq q$.

No For $S \subseteq V$, $|V|/10 \leq |S| \leq 9|V|/10$, we have $\Phi_G(S) \geq c\sqrt{\epsilon}$

Inapproximability of MCLA



Minimum Cut Linear Arrangement:

$$Cut_i(\pi) = \{e \in E | \pi(u) \leq i < \pi(v), e = (u, v)\}$$

Want:

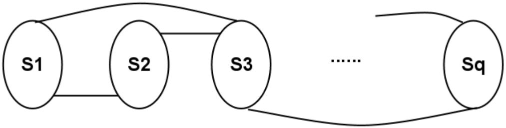
$$\min_{\pi} \max_{i \in [n]} |Cut_i(\pi)|$$

Theorem

MCLA (Layout[undirected, edge, max]) is hard to approximate within any constant factor.

MCLA is hard to approximate (1)

YES case: let $q = 1/\epsilon$

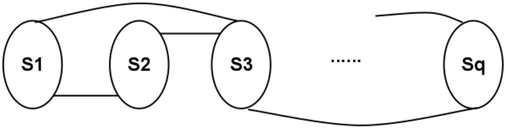


FACT: $|E(S, \bar{S})| = \Phi_G(S) d|S|$

$$
\begin{aligned}
|\#\text{edges inside each set}| &\leq d|S_i| = dn/q = 2\epsilon|E| \\
|\#\text{edges that expand}| &\leq n|E(S_i, \bar{S}_i)| = n \cdot \Phi_G(S_i) d|S_i| \leq 2\epsilon|E| \\
MCLA(G) &\leq 4\epsilon|E|
\end{aligned}
$$

MCLA is hard to approximate (1)

YES case: let $q = 1/\epsilon$

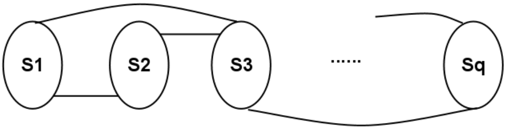


FACT: $|E(S, \bar{S})| = \Phi_G(S)d|S|$

$$\begin{aligned} |\#\text{edges inside each set}| &\leq d|S_i| = dn/q = 2\epsilon|E| \\ |\#\text{edges that expand}| &\leq n|E(S_i, \bar{S}_i)| = n \cdot \Phi_G(S_i)d|S_i| \leq 2\epsilon|E| \\ MCLA(G) &\leq 4\epsilon|E| \end{aligned}$$

MCLA is hard to approximate (2)

No case: let $S=$ the first half vertices according to a given permutation.



$$
\begin{aligned}
|\#\text{edges between the first half and the second half}| &\geq c\sqrt{\epsilon}|E| \\
MCLA(G) &\geq c\sqrt{\epsilon}|E|
\end{aligned}
$$

Hardness result of layout problems

- What we showed: For a given graph G, MCLA(G) (Layout[undirected, edge, max]) is hard to approximate within any constant factor.
- Can also get the same hardness for MLA (Layout[undirected, edge, sum]) (also proved by [Devanur, Khot, and Saket 06', Raghavendra,Steurer and Tulsiani 10']).
- For the rest of the layout problems, we need:
  1. Reduction from undirected version to directed version.
  2. Reduction from edge version to vertex version.

Hardness result of layout problems

- What we showed: For a given graph G, MCLA(G) (Layout[undirected, edge, max]) is hard to approximate within any constant factor.
- Can also get the same hardness for MLA (Layout[undirected, edge, sum]) (also proved by [Devanur, Khot, and Saket 06', Raghavendra,Steurer and Tulsiani 10']).
- For the rest of the layout problems, we need:
  1. Reduction from undirected version to directed version.
  2. Reduction from edge version to vertex version.

Reduction from undirected to directed version

Given $G = (V, E)$, produce a directed graph $D = (V', E')$.

$$
\begin{aligned}
V' &= V \cup E \\
E' &= \{(e, v) \mid e \in E, v \in V, v \in e\}.
\end{aligned}
$$

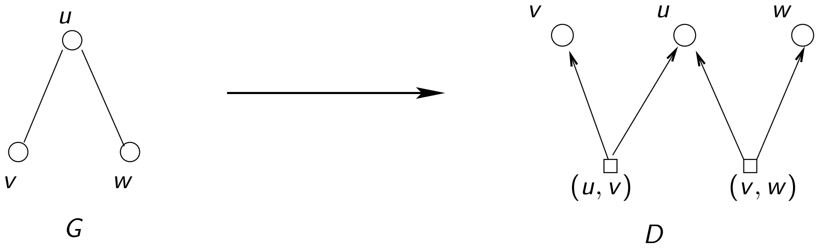
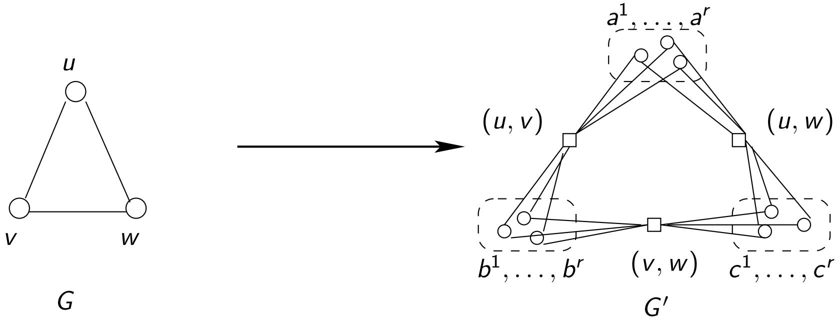


Figure: The reduction from G to D.

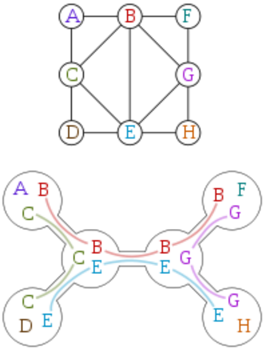
Reduction from edge to vertex version

Given $G = (V, E)$, produce $G' = (V', E')$, where

$$V' = \{v^i \mid v \in V, i \in [r]\}$$
$$E' = \{(e, v^i) \mid e \in E, v \in V, v \in e, i \in [r]\}.$$



Inapproximability of Treewidth



Cited from Wikipedia

(T1) $V = \cup_{t \in T} V_t$;

(T2) $\forall e \in E, \exists t \in T$, s.t. both endpoints of e lie in V_t;

(T3) for every $v \in V$, $\{t \in T \mid v \in V_t\}$ is a subtree of T.

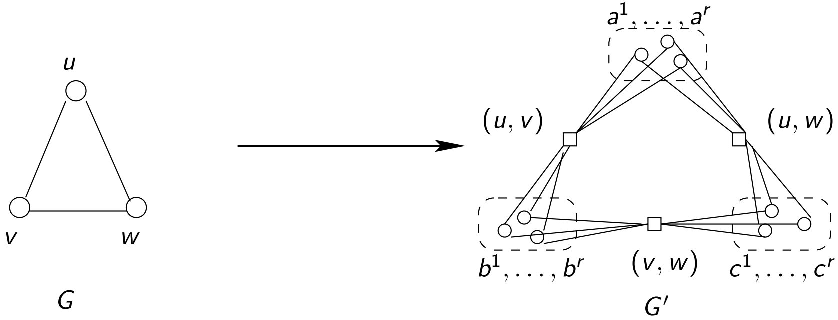
Theorem

It's SSE-hard to approximate Treewidth(G) within any constant factor.

Inapproximability for Treewidth

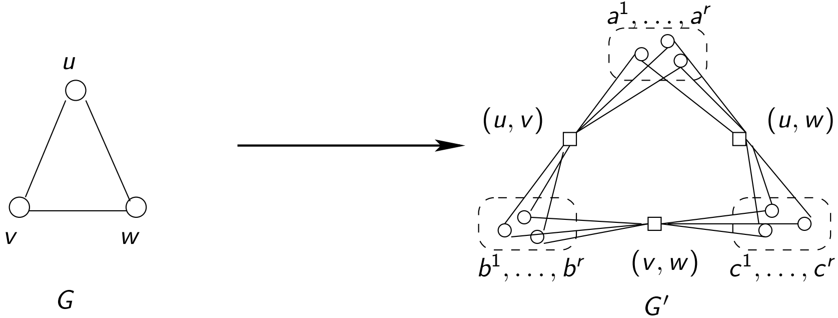
Proof: Use the following facts.

- $Treewidth(G) \leq Pathwidth(G)$.
- $Pathwidth(G) = Layout[undirected, vertex, max]$. [Kinnersley 92]



Inapproximability of Treewidth

In the Yes case, we use the fact that
$Treewidth(G') \leq Pathwidth(G') \leq O(\epsilon)|E|$.
In the No case, we show that $Treewidth(G') \geq c\sqrt{\epsilon}|E|$.



Conclusion and Discussion

Main Results

- SSE-hard to approximate graph layout problems within any constant factor.
- SSE-hard to approximate treewidth within any constant factor.

Open Problems

- Weaker Assumptions such as UGC?
- Would hardness of our problems imply hardness for the SSE problem?
- Approximability of the original pebbling problem?

Thank you all!