On Polyhedral Formulations for (Subset) Feedback Vertex Set

Chandra Chekuri
Univ of Illinois, Urbana-Champaign

CMU OR Seminar, Dec 8, 2023

Based on 2 papers

- **Polyhedral Aspects of Feedback Vertex Set and Pseudoforest Deletion Set**
  (with Karthik Chandrasekaran, Samuel Fiorini, Shubhang Kulkarni, and Stefan Weltge)
  submitted to journal, 2023.

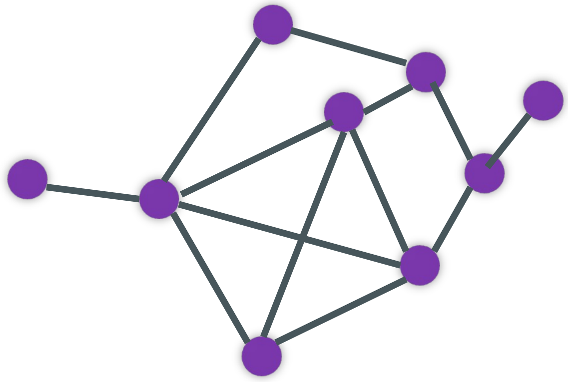
- **Constant Factor Approximation for Subset Feedback Problems via a new LP relaxation**
  (with Vivek Madan)
  SODA 2016.

Feedback Vertex Set (**FVS**)

Graph $G = (V, E)$ with vertex weights $w(v)$, $v \in V$

S is a feedback vertex set if $G - S$ has no cycles. In other words, S is a *hitting set* for all cycles in G

Goal: given G find FVS of min cardinality/weight

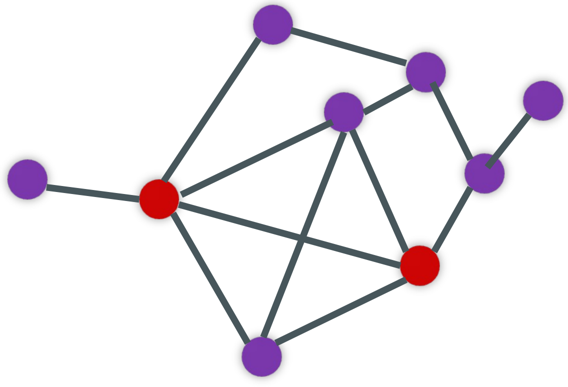


Feedback Vertex Set (**FVS**)

Graph $G = (V, E)$ with vertex weights $w(v),\ v \in V$

S is a feedback vertex set if $G - S$ has no cycles. In other words, S is a *hitting set* for all cycles in G

Goal: given G find FVS of min cardinality/weight

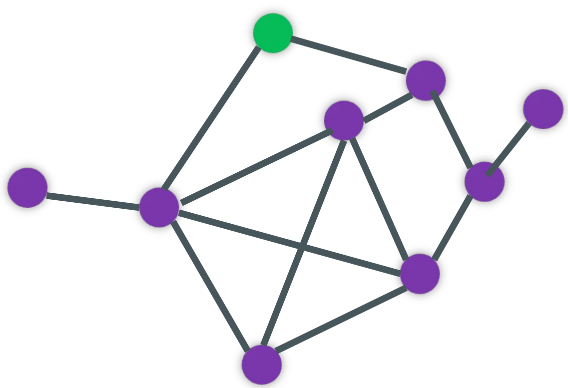


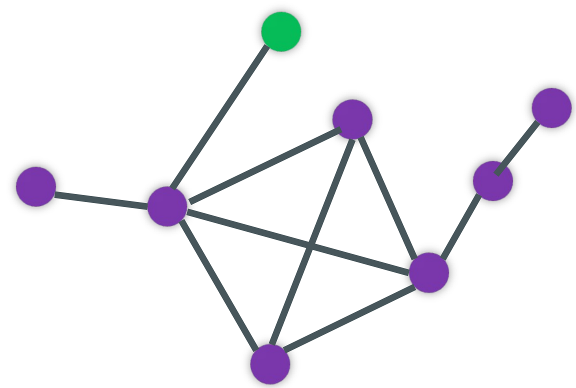
Subset Feedback Vertex Set (**SFVS**)

Graph $G = (V, E)$ with vertex weights $w(v)$, $v \in V$

$T \subseteq V$ is a set of terminals

Goal: find min cardinality/weight set S that hits all cycles that contain a terminal





Motivation

- **Applications:** Wikipedia: *FVS has wide applications in operating systems, database systems, and VLSI chip design*
  - An early application is from Bayesian inference from AI
- **Graph theory:** connection to Erdos-Posa theorem, graph minor theory, …
- **Algorithms:** approximation, fixed parameter tractability, canonical deletion problem, …

Directed graph FVS is very interesting and useful but this talk is about *undirected* graphs

Erdos-Posa Theorem

Suppose min FVS in a graph G is k. Then G contains $\Omega(k/\log k)$ vertex disjoint cycles. Moreover, this bound is tight in an infinite family of graphs.

Approximability of **FVS**

- **FVS** is NP-Hard (directed case in Karp's original list)
- α-approx. for **FVS** implies α-approx. for **Vertex Cover**
  - no $(2-\epsilon)$ approx. under UGC [Khot-Regev'08]
  - no 1.3606 approx. under $P \neq NP$ [Dinur-Safra'05]
- 2-approx. via "combinatorial" local-ratio method [Bafna-Berman-Fujito'95, Becker-Geiger'96]
- 2-approx. via prima-dual [Chudak-Goemans-Hochbaum-Williamson'98]

Approximability of **SFVS**

- 8-approx. [Even-Naor-Zosin'96] complicated and based on mix of combinatorial and LP ideas

Motivating Questions

- Is there a 2-approx. for **SFVS**? Lower bound is only 2 so far.
- Is there an *explicit poly-time solvable* LP relaxation for **FVS** that has an integrality gap of 2? [CGHW'98] formulation is not known to be solvable in poly-time.
- Is there an *explicit poly-time solvable* LP relaxation for **SFVS** with O(1) factor integrality gap?

LP Formulation for **SFVS**

[C-Madan'16]

- An explicit poly-time solvable LP formulation for **SFVS** (and hence also for **FVS**)
- Integrality gap of LP for **SFVS,** and hence also for **FVS,** is at most 13. Proof based on a primal rounding algorithm

Conjecture/Question: Is the integrality gap of CM-LP at most 2 for **SFVS**? At least for **FVS**?

Recent Results

[Chandrasekaran-C-Fiorini-Kulkarni-Weltge'23]

- A new explicit poly-time solvable formulation for **FVS** with integrality gap at most 2
- CM-LP integrality gap is at most 2 for **FVS**
- Connections to pseudo-forest deletion **(PFDS)** and Densest-Subgraph (**DSG**)
- Extreme point conjecture and evidence via **PFDS**

Rest of Talk

- Background on past LP formulations
- Pseudo-forest deletion (PFDS) and connection to FVS
- LP formulation for PFDS via density and densest subgraph
- LP formulations for FVS
- Summary and open problems

Cycle cover LP for **FVS**

$$\min \sum_u w_u x_u$$

$$\sum_{u \in C} x_u \geq 1 \quad \text{for all cycles } C$$

$$x_u \geq 0 \quad \text{for all } u \in V$$

Integrality gap is $\Theta(\log n)$ [Bar-Yehuda-Geiger-Naor-Roth]

Lower bound via expanders/high girth constant deg graphs

Dual is fractional cycle packing LP in unweighted case.

Gap is related to Erdos-Posa theorem

[CGHW'98] LP Relaxation

SD-LP (strong density LP)

$\min \sum_u w_u x_u$

$\sum_{u \in S} (d_S(u) - 1) x_u \geq |E(S)| - |S| + 1$ for all $S \subseteq V$ $s.t$ $E[S] \neq \emptyset$

$x_u \geq 0$ for all $u \in V$

Notation: $d_S(u)$ is degree of u in induced graph G[S]

[CGHW'98] LP Relaxation

$\sum_{u \in S} (d_S(u) - 1) x_u \geq |E(S)| - |S| + 1$ for all $S \; s.t \; E[S] \neq \emptyset$

Why is inequality valid? Let $\boldsymbol{x} \in \{0,1\}^V$

- Consider S = V. Say $F \subset V$ is an FVS.
- G – F has no cycle: $E[V - F] \leq |V| - |F| - 1$
- But $\sum_{u \in F} \deg(u) \geq |E| - E[V - F]$
- Rearranging gives the desired claim for $x_u = 1, u \in F$

New Formulation

We do not know efficient separation oracle for SD-LP

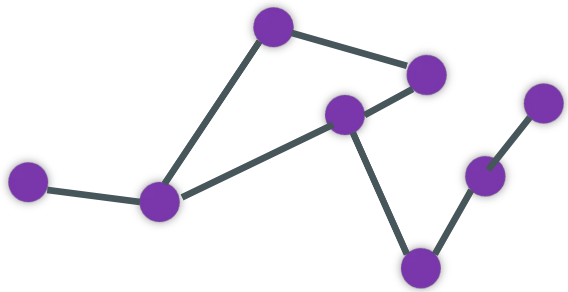
New formulation inspired by considering

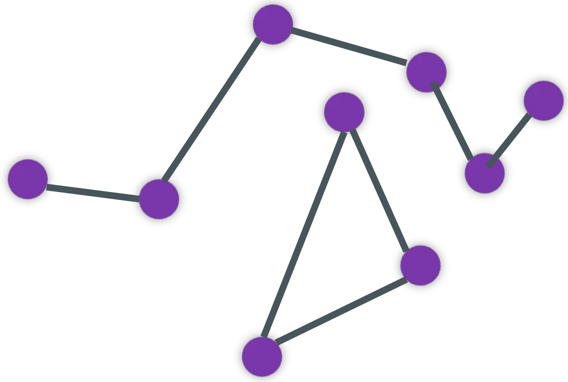
- Related problem pseudo-forest deletion (**PFDS**)
- Connecting to Densest-Subgraph (**DS**) and LP for it by [Charikar'00]

Pseudo tree/forest

pseudo-tree is a tree + at most one edge

pseudo-forest: each connected comp. is a pseudo-tree





Pseudo-forest Deletion Set (**PFDS**)

PFDS: given G remove vertices to get a pseudo-forest

[Lin-Feng-Fu-Wang'19]

PFDS admits 2-approximation via *local-ratio*

Reduction from Vertex-Cover shows hardness of 2

No LP connections

Weak Density LP for **PFDS**

WD-LP (weak density LP)

$\min \sum_u w_u x_u$

$\sum_{u \in S} (d_S(u) - 1) x_u \geq |E(S)| - |S|$ for all $S \subseteq V$

$x_u \geq 0$ for all $u \in V$

Notation: $d_S(u)$ is degree of u in induced graph G[S]

Validity reasoning is similar to that for **FVS**

Inequalities also hold for **FVS**

Weak Density LP for **PFDS**

WD-LP (weak density LP)

$\min \sum_u w_u x_u$

$\sum_{u \in S} (d_S(u) - 1) x_u \geq |E(S)| - |S|$ for all $S \subseteq V$

$x_u \geq 0$ for all $u \in V$

Notation: $d_S(u)$ is degree of u in induced graph G[S]

Validity reasoning is similar to that for **FVS**

Inequalities also hold for **FVS**

WD-LP for **PFDS**

Theorem: Worst-case integrality gap of WD-LP for **PFDS** is 3

Theorem: Suppose G is not a pseudo-forest. Then for every extreme point **x** of WD-LP for G there is some vertex u such that $x_u \geq \frac{1}{3}$.

WD-LP for **PFDS**

Theorem: Worst-case integrality gap of WD-LP for **PFDS** is 3

Theorem: Suppose G is not a pseudo-forest. Then for every extreme point **x** of WD-LP for G there is some vertex u such that $x_u \geq \frac{1}{3}$.

Remarks:

- Do not know how to solve WD-LP in poly-time
- There is a 2-approximation for PFDS via local-ratio (more later)
- What about **FVS**?

WD-LP for **FVS**

Question: Is WD-LP good for **FVS?**

Take G to be a simple cycle. Then $\boldsymbol{x} = \mathbf{0}$ is feasible so formulation is not good enough

WD + Cycle cover for **FVS**

WD+CycleCover-LP (weak density + cycle cover inequalities)

min $\sum_u w_u x_u$

$\sum_{u \in S} (d_S(u) - 1) x_u \geq |E(S)| - |S|$ for all $S \subseteq V$

$$\sum_{u \in C} x_u \geq 1 \quad \text{for all cycle } C$$

$x_u \geq 0$ for all $u \in V$

WD + Cycle cover for **FVS**

Theorem: Integrality gap of WD+Cycle cover LP is at most 2 for **FVS.**

Proof:

Follow primal-dual analysis of [CGHW'98]

Notice that weak-density constraints are "weak" only for the case when G is a cycle. Use cycle cover inequality in that case. Need to do it formally …

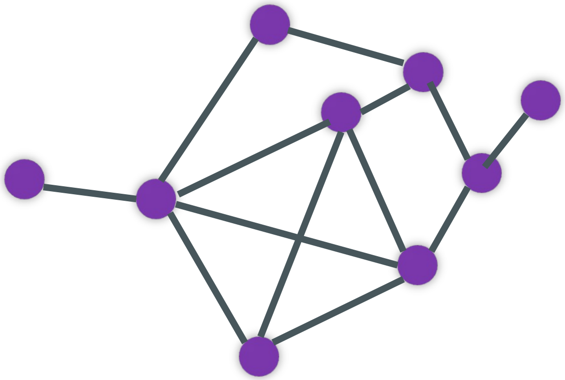
But

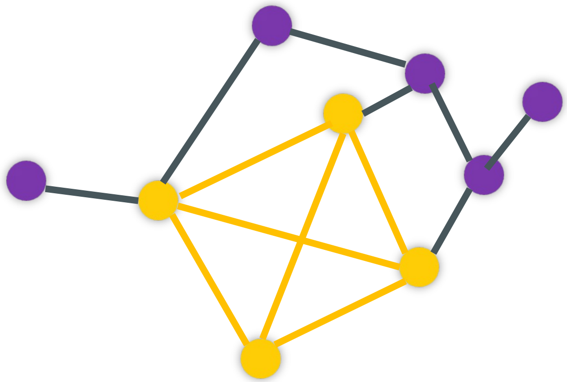
you promised an explicitly solvable LP!

Densest Subgraph

Given G=(V,E) and $S \subseteq V$, $den(S) = \frac{|E(S)|}{|S|}$

Densest subgraph (DSG): Given G=(V,E) find S to maximize den(S)





$\lambda^* = \frac{6}{4}$

LP Relaxation for **DSG**

Primal

$$\max \sum_{uv \in E} x_{uv}$$

$$\sum_v z_v = 1$$

$$x_{uv} \leq \min(z_u, z_v) \quad uv \in E$$

$$x, z \geq 0$$

Dual

$$\min D$$

$$y_{uv,u} + y_{uv,v} \geq 1 \quad uv \in E$$

$$\sum_{uv \in E} y_{uv,v} \leq D \quad v \in V$$

$$y \geq 0$$

Theorem: [Charikar'00] LP is optimal for **DSG**

Interpreting Dual

$$\min D$$

$$y_{uv,u} + y_{uv,v} = 1 \quad uv \in E$$

$$\sum_{uv \in E} y_{uv,v} \leq D \quad v \in V$$

$$y \geq 0$$

Optimal density is equal to "fractional arboricity" of G

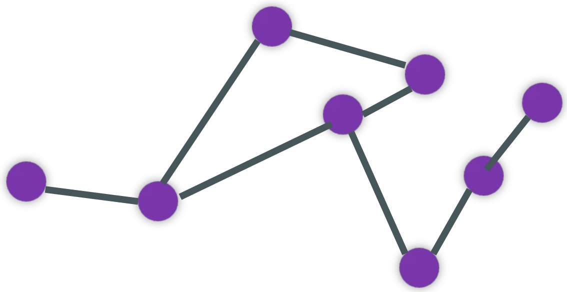
- *Orient* each edge uv *fractionally*
- Load on vertex u is the total fraction oriented *into* u
- Minimize maximum load

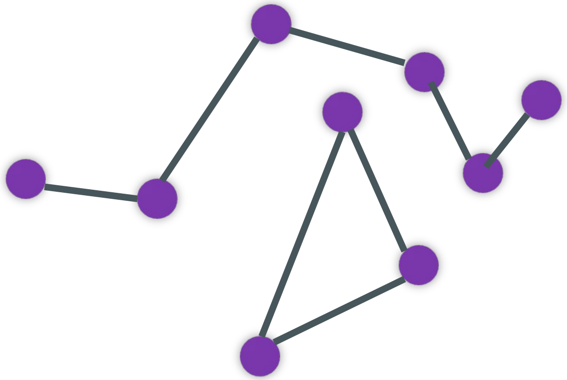
Density and PFDS

Given G=(V,E) and $S \subseteq V$, $den(S) = \frac{|E(S)|}{|S|}$

PFDS:

given G remove S such that G – S has density *at most* 1





Orientation based LP for **PFDS**

Orientation-LP

$\min \sum_u w_u x_u$

$y_{uv,v} + y_{uv,u} \geq 1 - x_u - x_v$ for each edge uv $\in E$

$\sum_{e \in \delta(u)} y_{e,u} \leq 1 - x_u$ for each $u \in V$

$x_u \geq 0$ for each $u \in V$

$y_{e,u} \geq 0$ for each $u \in V, e \in \delta(u)$

$e = uv$ deleted if u *or* v chosen. Hence amount of edge left "fractionally" is $\geq 1 - x_u - x_v$ which needs to be oriented

Want density at most 1, and amount of vertex u left is $1 - x_u$

Orientation and WD LPs

Lemma: Orientation-LP *strictly* stronger than WD-LP

Fix *any subgraph* H = (V', E') of G.

$$|E'| \leq \sum_{e=uv \in E'} (x_u + x_v + y_{e,u} + y_{e,v})$$

$$= \sum_{v \in V'} (d_{H(v)} - 1)\, x_v + \sum_{v \in V'} (x_v + \sum_{e \in \delta_{H(v)}} y_{e,v})$$

$$\leq \sum_{v \in V'} (d_{H(v)} - 1)\, x_v + |V'|$$

Explicit LP for **FVS**

Theorem: Integrality gap of WD+Cycle cover LP is at most 2 for **FVS.**

Lemma: Orientation-LP *strictly* stronger than WD-LP

Hence,

Orient-LP + Cycle cove LP has integrality gap at most 2 for **FVS.**

Can write cycle cover inequalities explicitly/compactly with distance variables

Explicit LP for **FVS**

Orientation-LP+Cycle cover inequalities

min $\sum_u w_u x_u$

$y_{uv,v} + y_{uv,u} \geq 1 - x_u - x_v$ for each edge uv $\in E$

$\sum_{e \in \delta(u)} y_{e,u} \leq 1 - x_u$ for each $u \in V$

$\sum_{u \in C} x_u \geq 1$ $for\ each\ cycle\ C$

$x_u \geq 0$ for each $u \in V$

$y_{e,u} \geq 0$ for each $u \in V, e \in \delta(u)$

Two other LPs

Theorem: CM-LP for FVS (based on labeling approach) is at least as strong as Orientation+Cycle Cover LP.

Theorem: There is an explicit LP based on orientation constraints that is at least as strong as the Strong-Density LP.

Back to **PFDS**

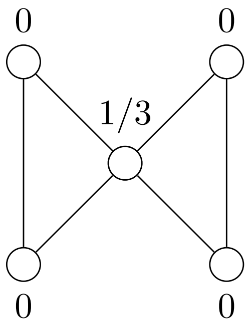
Theorem: Worst-case integrality gap of WD-LP for **PFDS** is 3

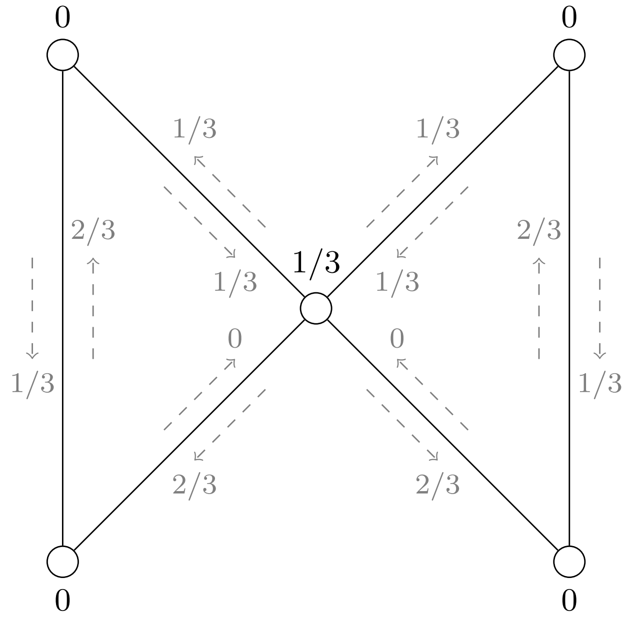
Theorem: Suppose G is not a pseudo-forest. Then for every extreme point **x** of WD-LP for G there is some vertex u such that $x_u \geq \frac{1}{3}$.

Saw that Orient-LP is at least as strong as WD-LP

Worst-case integrality gap for Orient-LP is also 3

Integrality Example for **PFDS**





Improving Integrality Gap for **PFDS**

Recall there is primal-dual 2-approx. for **PFDS** due to [Lin-Feng-Fu-Wang'19]

Can we strengthen WD-LP?

Minimal violation of pseudo-tree is 2-pseudo-tree: a connected graph with $|V|+2$ edges (like the butterfly graph)

Add a constraint that at least one node chosen from each 2-pseudo-tree

Improving Integrality Gap for **PFDS**

WD+2-pseudo-tree cover LP

$\min \sum_u w_u x_u$

$\sum_{u \in S} (d_S(u) - 1) x_u \geq |E(S)| - |S|$ for all $S \subseteq V$ $s.t$ $E[S] \neq \emptyset$

$$\sum_{u \in C} x_u \geq 1 \quad \text{for each 2pseudotree } C$$

$x_u \geq 0$ $for\ each\ u \in V$

Theorem: Integrality gap of above LP is at most 2.

Follow primal-dual analysis of [Lin-Feng-Fu-Wang'19]

Separation Oracle

$$\sum_{u \in C} x_u \geq 1 \quad \text{for each 2pseudotree } C$$

Lemma: There is a polynomial-time separation oracle for above constraint.

Guess two edges and use node-weighted Steiner tree algorithm on four terminals (exists an FPT algorithm for any fixed number of terminals)

Rounding LPs

Integrality gap of 2 for FVS/PFDS are based on *primal-dual* analysis.

Exceptions:

- Integrality gap of 3 for PDFS via *iterated rounding* wrt WD-LP or Orientation-LP
- Integrality gap of 13 for via CM-LP via *primal rounding*

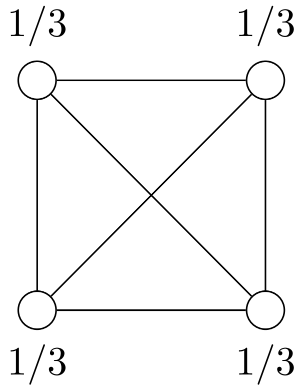
Conjecture for **FVS**

Conjecture: Let **x** be an extreme point solution for Strong-Density LP for FVS. If G is not a forest then there is some vertex u such that $x_u \geq 1/2$.

Conjecture for **FVS**

Conjecture: Let **x** be an extreme point solution for Strong-Density LP for FVS. If G is not a forest then there is some vertex u such that $x_u \geq 1/2$.

Subtlety: Weak-Density + Cycle Cover LP has integrality gap at most 2 but example shows extreme point property is false



Proof idea for **PFDS**

Theorem: Suppose G is not a pseudo-forest. Then for every extreme point **x** of WD-LP for G there is some vertex u such that $x_u \geq \frac{1}{3}$.

Proof?

Like other iterated rounding it is based on uncrossing and token counting but with a small twist

Conditional supermodularity

WD-LP (weak density LP)

min $\sum_u w_u x_u$

$\sum_{u \in S} (d_S(u) - 1) x_u \geq |E(S)| - |S|$ for all $S \subseteq V$ $s.t$ $E[S] \neq \emptyset$

$x_u \geq 0$ for each $u \in V$

Lemma: Suppose x is a fractional solution s.t $x_u < \frac{1}{2}$ for each vertex u.
Consider $f_x : 2^V \to R$ where

$$f_x(S) = |E(S)| - |S| - \sum_{u \in S} (d_S(u) - 1) x_u$$

Then f_x is supermodular.

Open Problems

- Extreme point conjecture for **FVS**. In general, explicit primal-rounding procedures for LP relaxations achieving factor of 2 approx.
- Is there a 2-approx. for **SFVS**?
  - What is the integrality gap of CM-LP for **SFVS**? Currently at most 13
  - Alternative LP relaxations for easier analysis?
- Is there a better than 2 approximation for **SFES**? Only hardness is via **Multiway-Cut**
- **Original inspiration for Fiorini:** Deletion to small treewidth (weighted case and LP/SDP formulations). See [Gupta etal, Bansal etal, Bonnet et al]

Thank You!