Approximating Rooted Steiner Networks

Joseph Cheryian[1], Bundit Laekhanukit[2], *Guyslain Naves*[2], Adrian Vetta[2]

December 2011, Banff

[1]Waterloo University

[2]McGill University

Directed Steiner Tree problem (DST)

A network design problem:

Input:
- G a directed graph, with costs $c : E(G) \to \mathbb{N}$,
- r a vertex of G (the *root*),
- a set $T \subseteq V(G)$ of *terminals*,

Output: A subgraph G' of G such that there is one path from s to t in G', for all $t \in T$

Goal: $\min \sum_{e \in E(G')} c(e)$

McGill

Directed Rooted Connectivity problem

A network design problem:

Input:
- G a directed graph, with costs $c : E(G) \to \mathbb{N}$,
- r a vertex of G (the *root*),
- a set $T \subseteq V(G)$ of *terminals*,
- requirements $k : T \to \mathbb{N}$.

Output: A subgraph G' of G such that there are k_t disjoint paths from s to t in G', for all $t \in T$

Goal: $\min \sum_{e \in E(G')} c(e)$

Outline

1. k-DRC with $O(1)$ terminals.
2. Hardness of k-DRC (directed graph).
3. Hardness of k-URC (undirected graphs).
4. Integrality gap of k-DRC.

McGill

Directed Steiner Forest with $O(1)$ terminals

Theorem (Feldman, Ruhl (2006))

The Directed Steiner Forest with $O(1)$ terminals is polynomial-time solvable.

Proof: Guess nodes of degree > 2 and how they are linked, compute shortest paths.

Generalization to Directed Rooted Connectivity ?

McGill

Bounded connectivity requirement

Proposition

If G is an acyclic digraph and $\sum_{t \in T} k_t = O(1)$, then there is a polynomial-time algorithm.

Proof: Pebbling game (Fortune, Hopcroft, Wyllie).

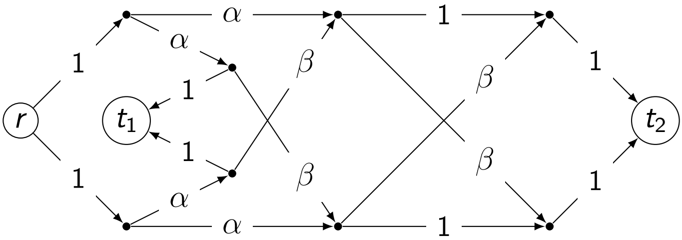
Open problem: (polynomial or NP-hard?)

$$\sum_{t \in T} k_t = O(1) \text{ but } G \text{ is not acyclic.}$$

McGill

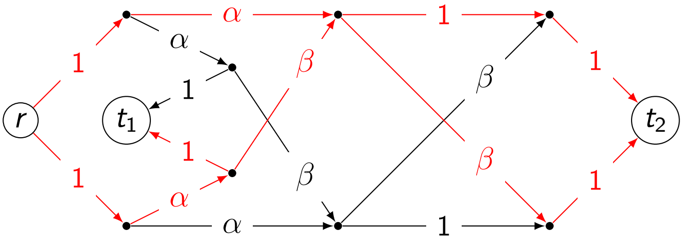
Non-integrality for requirement 3

Let $\alpha = 2\beta \geq 2$, $k_{t_1} = 1$ and $k_{t_2} = 2$.



Non-integrality for requirement 3

Let $\alpha = 2\beta \geq 2$, $k_{t_1} = 1$ and $k_{t_2} = 2$.

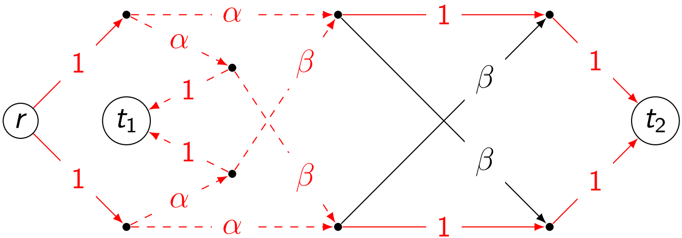


- Integral solution: $6\beta + 6$

McGill

Non-integrality for requirement 3

Let $\alpha = 2\beta \geq 2$, $k_{t_1} = 1$ and $k_{t_2} = 2$.

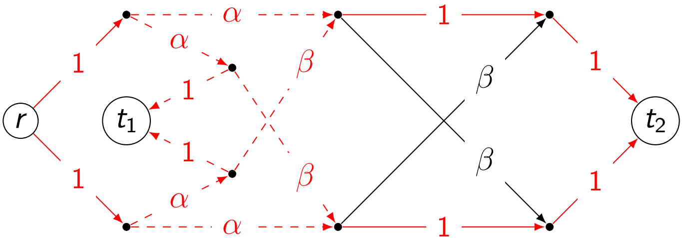


- Integral solution: $6\beta + 6$
- Fractional solution: $5\beta + 7$

McGill

Non-integrality for requirement 3

Let $\alpha = 2\beta \geq 2$, $k_{t_1} = 1$ and $k_{t_2} = 2$.

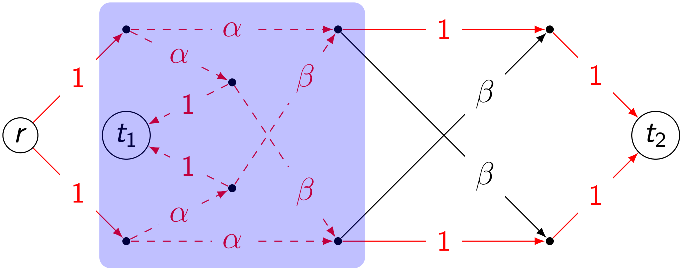


- Integral solution: $6\beta + 6$
- Fractional solution: $5\beta + 7$

Integrality gap: $\frac{6}{5}$

Non-integrality for requirement 3

Let $\alpha = 2\beta \geq 2$, $k_{t_1} = 1$ and $k_{t_2} = 2$.



- Integral solution: $6\beta + 6$
- Fractional solution: $5\beta + 7$

Integrality gap: $\frac{6}{5}$

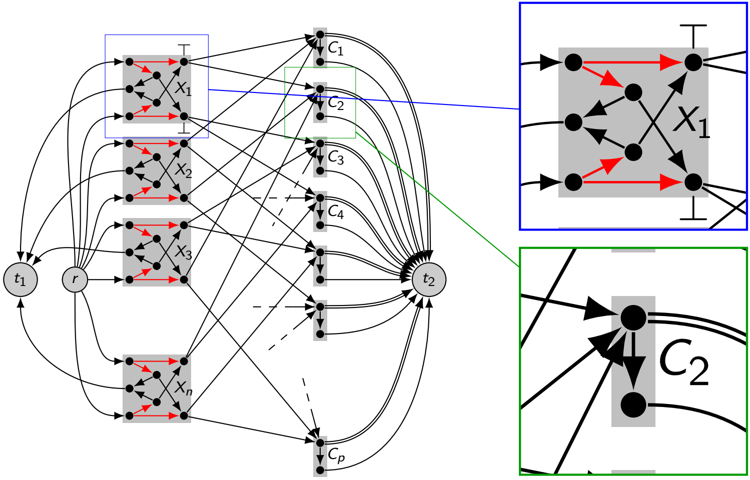
Toward an APX-hardness proof.

Theorem (Berman, Karpinski, Scott)

For every $0 < \varepsilon < 1$, *it is* NP*-hard to approximate* $\textsc{Max-3-Sat}$ *where each literal appears exactly twice, within an approximation ratio smaller than* $\frac{1016-\varepsilon}{1015}$.

McGill

Reduction for two terminals



Analysis (two terminals problem)

Using $\mathrm{OPT}_\phi \geq \frac{7q}{8}$, we get:

$$\rho \geq \frac{13n + (q - \mathrm{APP}_\phi)}{13n + (q - \mathrm{OPT}_\phi)} = 1 + \frac{\mathrm{OPT}_\phi - \mathrm{APP}_\phi}{13n + q - \mathrm{OPT}_\phi}$$
$$\geq 1 + \frac{7}{79}\frac{\mathrm{OPT}_\phi - \mathrm{APP}_\phi}{\mathrm{OPT}_\phi} = 1 + \frac{7}{79}\left(1 - \gamma^{-1}\right)$$

and finally

$$\rho \geq 1 + \frac{7}{80264} - \xi, \text{ for any } \xi > 0.$$

Analysis (two terminals problem)

Using $\mathrm{OPT}_\phi \geq \frac{7q}{8}$, we get:

$$\rho \geq \frac{13n + (q - \mathrm{APP}_\phi)}{13n + (q - \mathrm{OPT}_\phi)} = 1 + \frac{\mathrm{OPT}_\phi - \mathrm{APP}_\phi}{13n + q - \mathrm{OPT}_\phi}$$
$$\geq 1 + \frac{7}{79}\frac{\mathrm{OPT}_\phi - \mathrm{APP}_\phi}{\mathrm{OPT}_\phi} = 1 + \frac{7}{79}\left(1 - \gamma^{-1}\right)$$

and finally

$$\rho \geq 1 + \frac{7}{80264} - \xi, \text{ for any } \xi > 0.$$

Easy k-approximation when only k terminals.

McGill

Outline

1. k-DRC with $O(1)$ terminals.
2. Hardness of k-DRC (directed graph).
3. Hardness of k-URC (undirected graphs).
4. Integrality gap of k-DRC.

McGill

General directed rooted connectivity

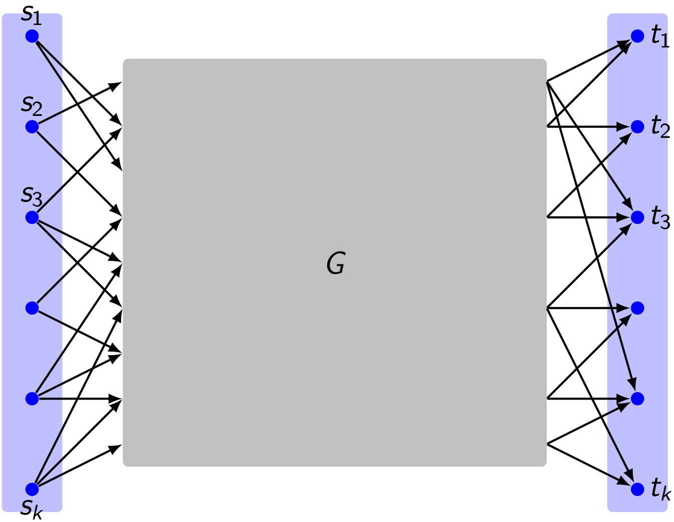
Theorem

The directed and undirected rooted k-connectivity problem are at least as hard to approximate as the label cover problem ($2^{\log^{1-\varepsilon} n}$).

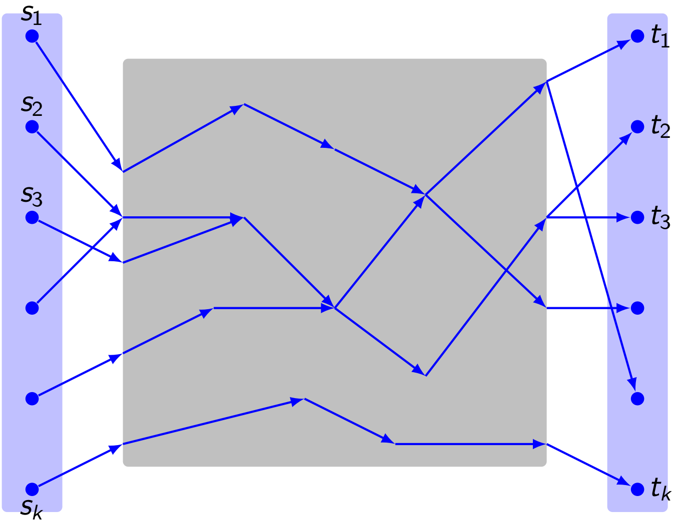
Proof: Approximation-preserving reduction from Directed Steiner Forest (Dodis, Khanna)
(pairs (s_i, t_i) to connect)

Undirected version by a reduction of Lando and Nutov.

Reduction (directed Steiner Forest)

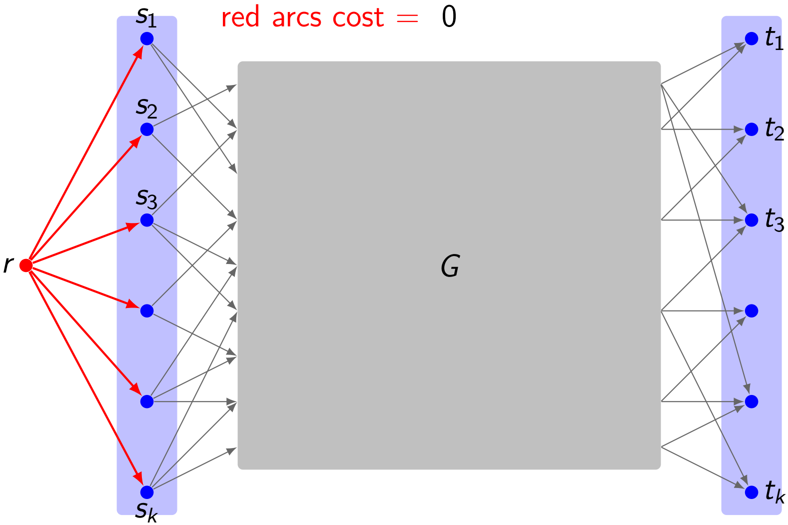


Reduction (directed Steiner Forest)

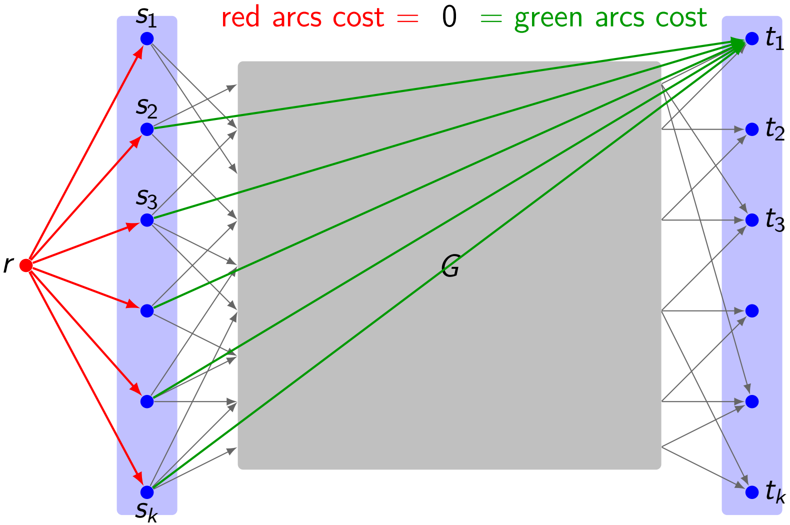


McGill

Reduction (directed Steiner Forest)

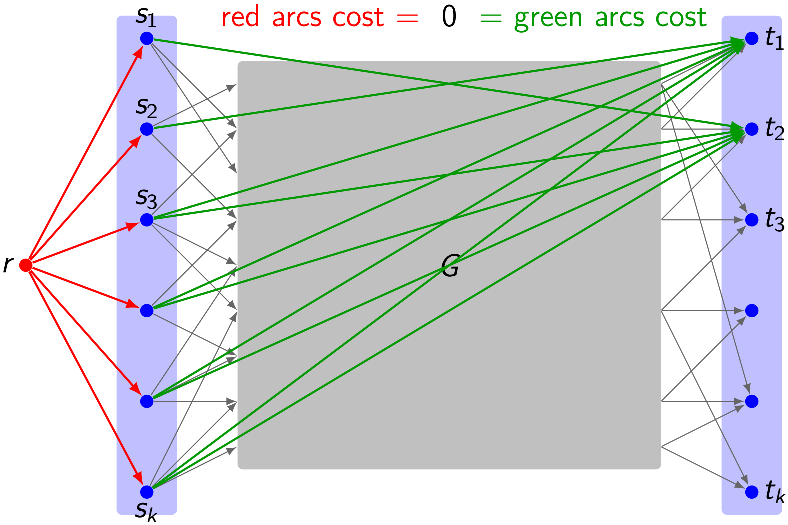


Reduction (directed Steiner Forest)



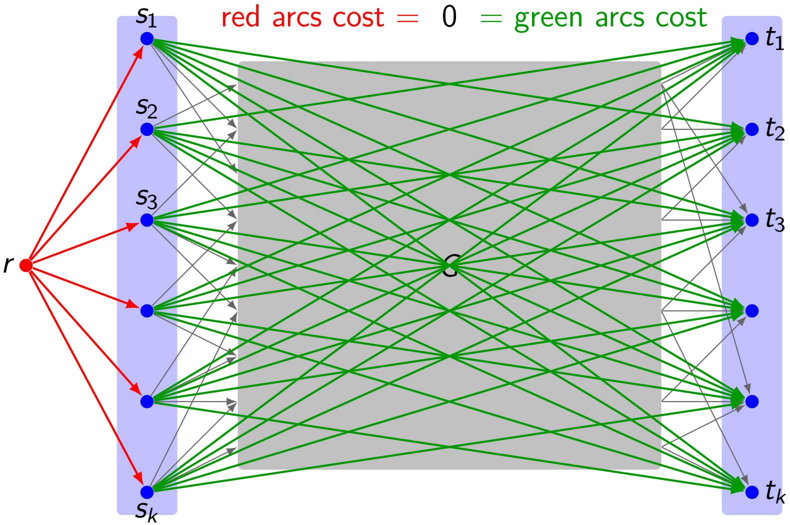
McGill

Reduction (directed Steiner Forest)



McGill

Reduction (directed Steiner Forest)



Stronger hardness result

Theorem

The directed rooted k-connectivity problem cannot be approximated to within $O(k^{\varepsilon})$, for some constant $\varepsilon > 0$, assuming that NP is not contained in $\mathrm{DTIME}(n^{polylog(n)})$.

Proof: Reduction from a label cover instance obtained from $\textsc{Max-3-Sat}(5)$ with l repetition (Chakraborty, Chuzhoy, Khanna).

McGill

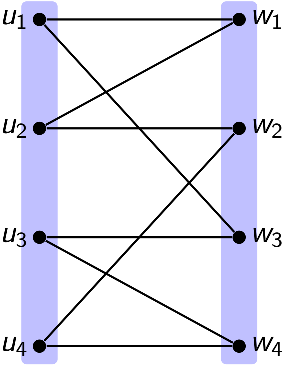
Label Cover problem

- $G = (U, W, E)$ bipartite graph,
- L set of labels,
- constraints $\Pi_e \subseteq L \times L$ for all $e \in E$,
- assign labels to every vertex to cover every edge $(\forall uw \in E, \Pi_{uw} \cap (f(u) \times f(w)) \neq \emptyset)$,
- minimize the number of labels assigned $\sum_{u \in U \cup W} |f(u)|$.

Instances obtained from $\textsc{Max-3-Sat}(5)$ with l repetition:

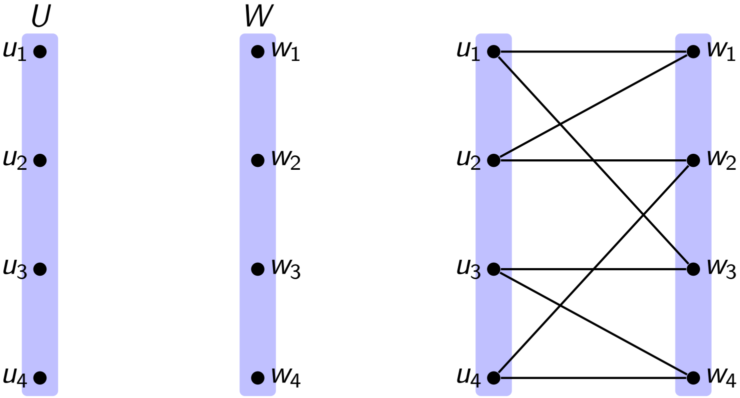
$$|U| = |W| = O(N^{O(l)}), \qquad |L| = 10^l, \qquad d = 15^l$$

Reduction from label cover



McGill

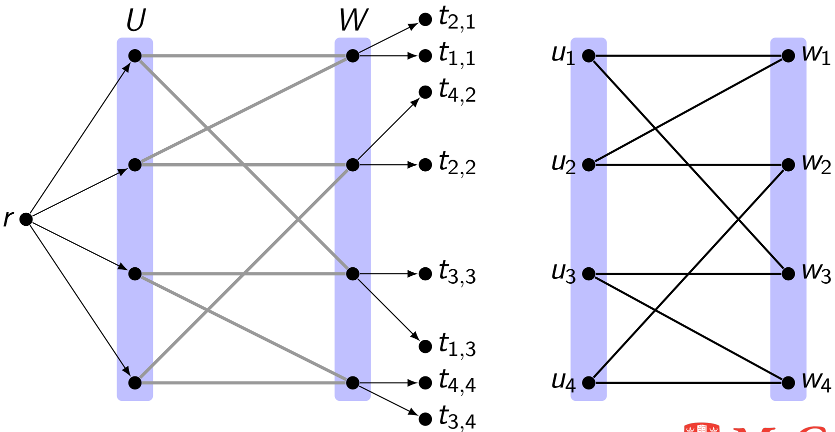
Reduction from label cover



McGill

Reduction from label cover

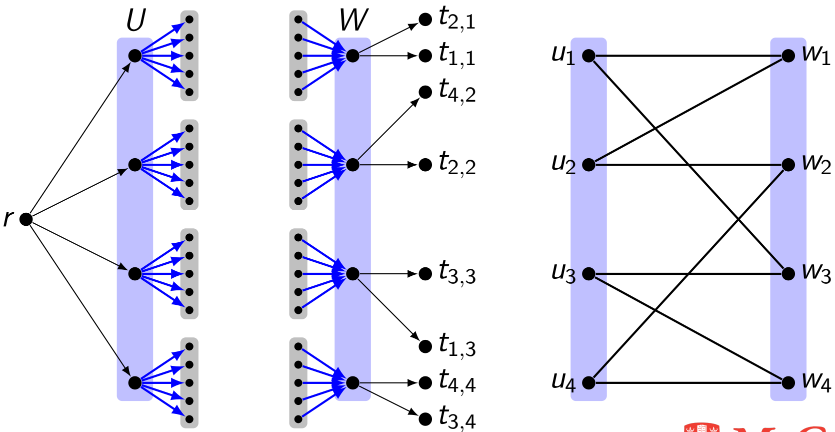
cost(⟶) = 1, cost(others) = 0



McGill

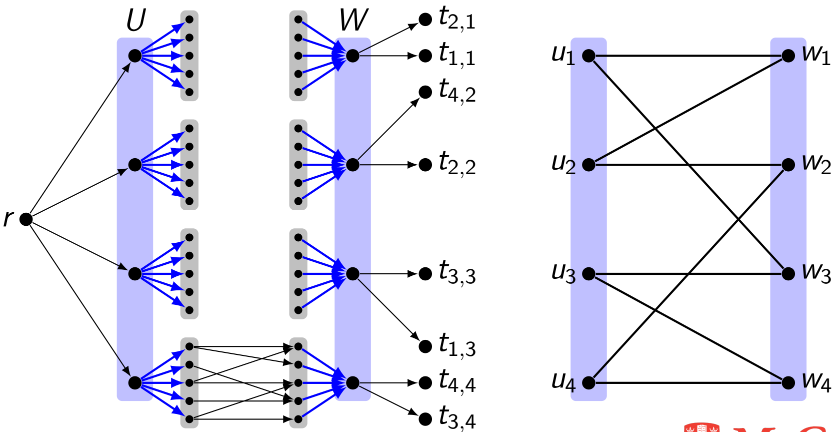
Reduction from label cost

cost(→) = 1, cost(others) = 0



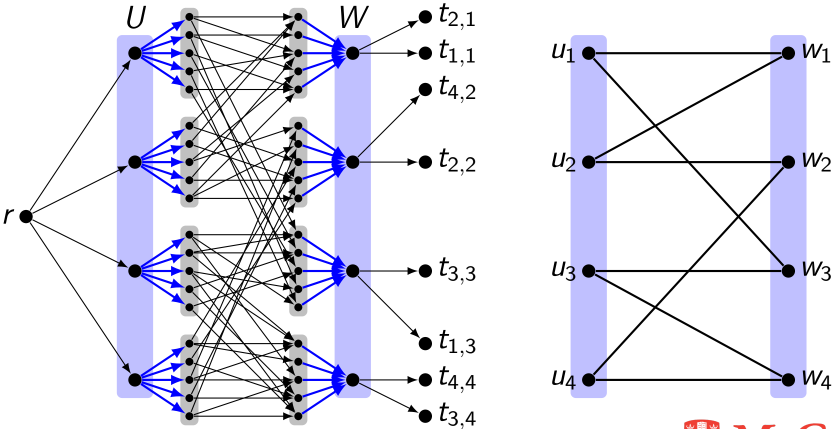
Reduction from label cover

cost(→) = 1, cost(others) = 0



Reduction from label cover

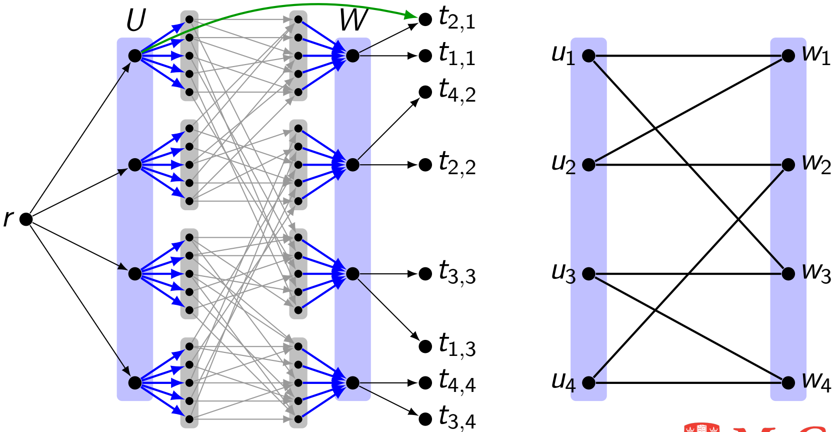
cost(→) = 1, cost(others) = 0



McGill

Reduction from label cover

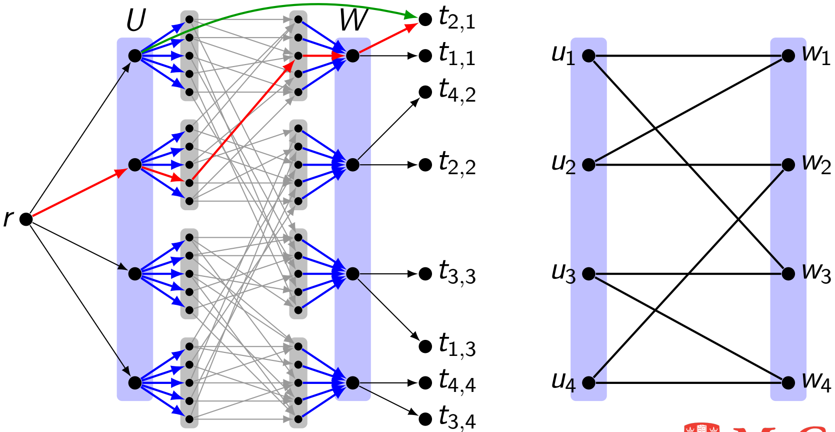
$\text{cost}(\longrightarrow) = 1$, $\text{cost}(\text{others}) = 0$



McGill

Reduction from label cover

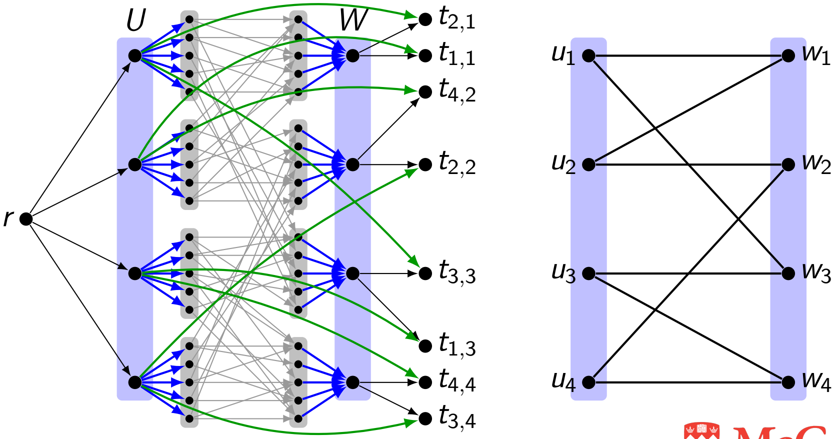
cost(→) = 1, cost(others) = 0



McGill

Reduction from label cover

cost($\longrightarrow$) $= 1$, cost(others) $= 0$



Getting the hardness ratio

Theorem (Parallel repetition theorem, Raz)

There exists a constant $\gamma > 0$ (independent of l) such that the minimum total label cover problem obtained from instances of MAX-3SAT(5) with l repetitions cannot be approximated within a factor of $2^{\gamma l}$.

In our reduction, $k = d = 15^l$, hence the k^ε-hardness!

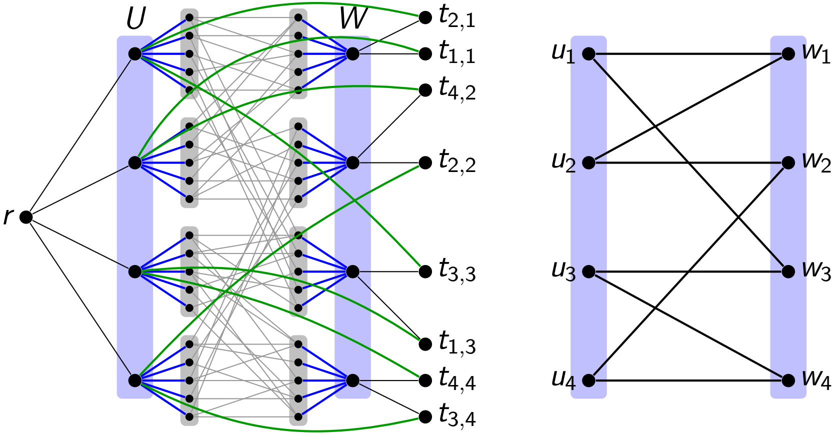
McGill

Outline

1. k-DRC with $O(1)$ terminals.
2. Hardness of k-DRC (directed graph).
3. Hardness of k-URC (undirected graphs).
4. Integrality gap of k-DRC.

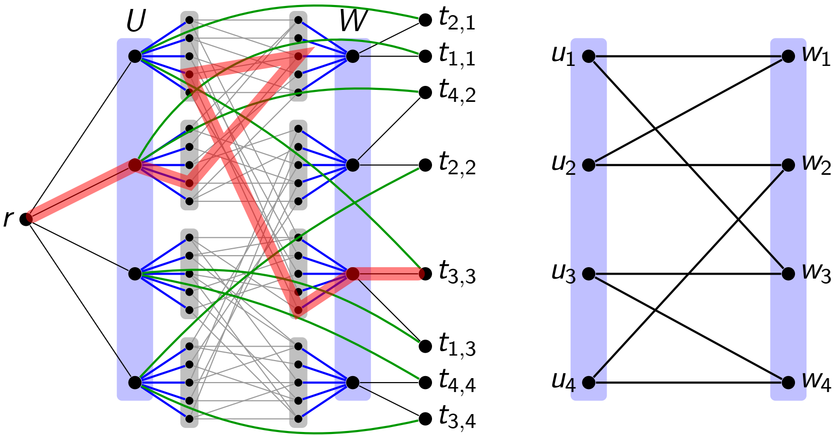
McGill

Adapting the reduction to undirected graphs



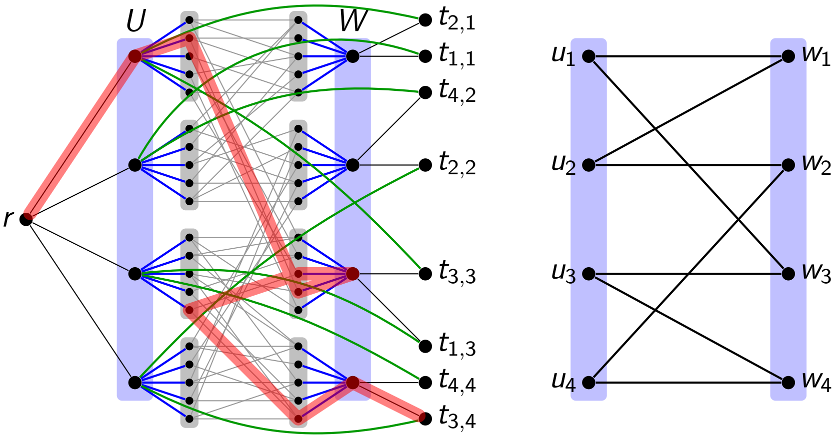
McGill

Adapting the reduction to undirected graphs

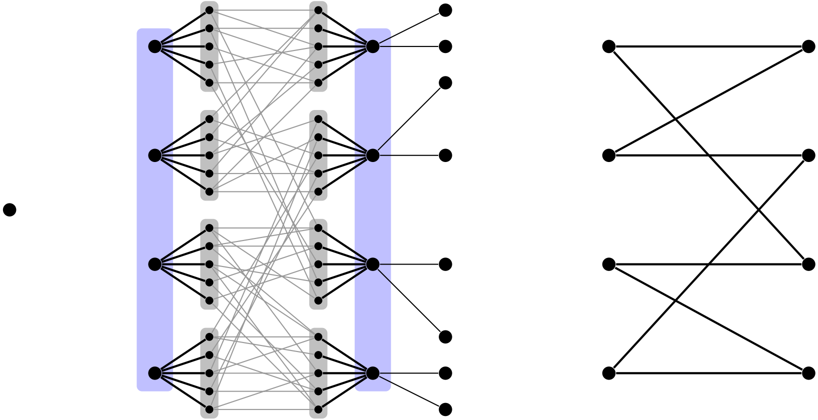


McGill

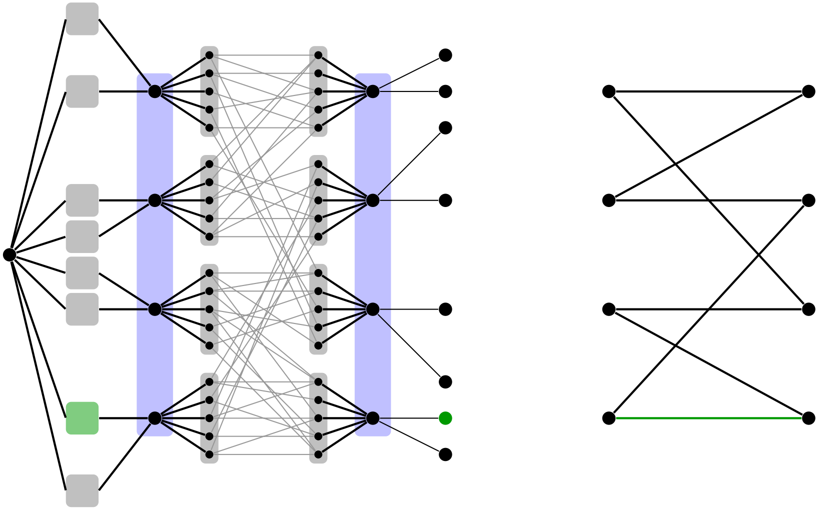
Adapting the reduction to undirected graphs



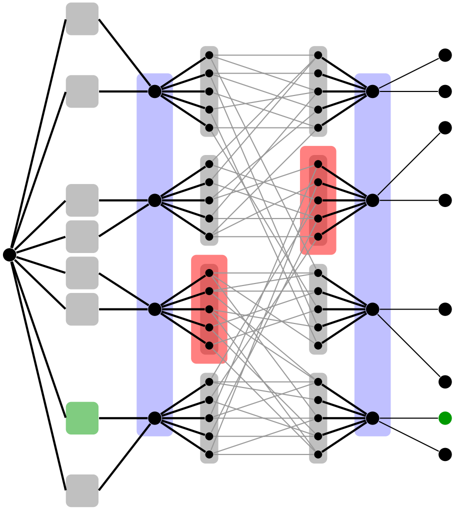
Forbidding illegal paths

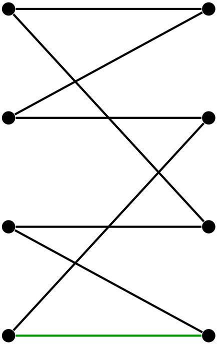


Forbidding illegal paths

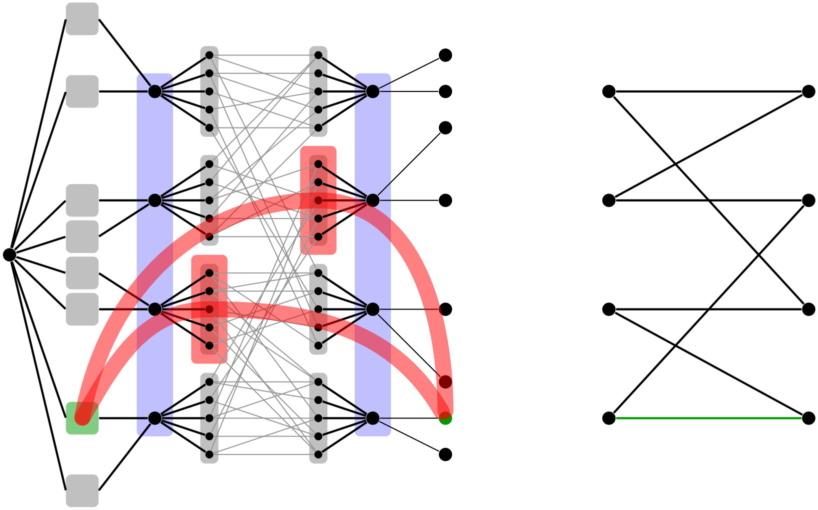


Forbidding illegal paths

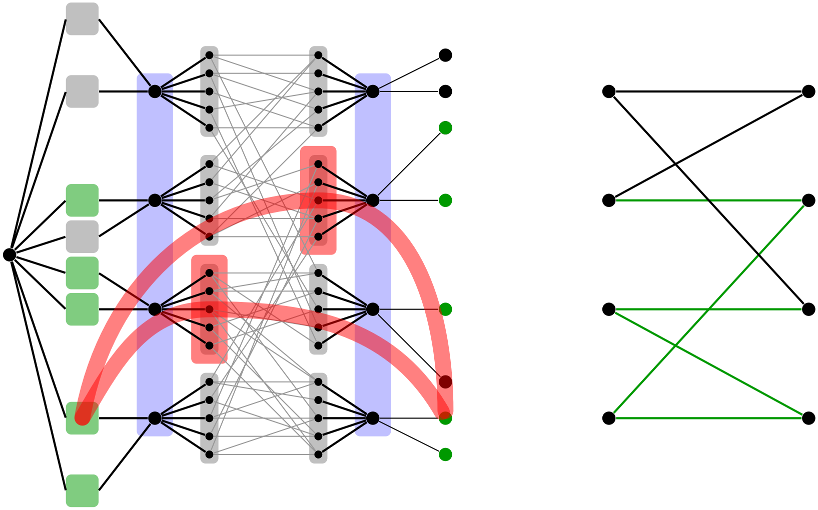




Forbidding illegal paths

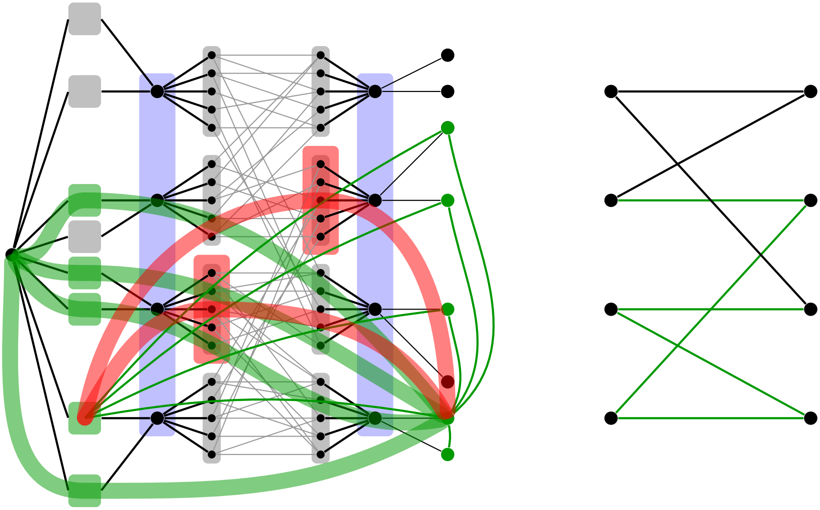


Forbidding illegal paths



McGill

Forbidding illegal paths



Very informal description

- There are illegal paths,

Very informal description

- There are illegal paths,
- add padding edges to remove illegal paths,

McGill

Very informal description

- There are illegal paths,
- add padding edges to remove illegal paths,
- this creates new illegal paths,

Very informal description

- There are illegal paths,
- add padding edges to remove illegal paths,
- this creates new illegal paths,
- add more padding edges to remove the new illegal paths,

Very informal description

- There are illegal paths,
- add padding edges to remove illegal paths,
- this creates new illegal paths,
- add more padding edges to remove the new illegal paths,
- the second padding set does not induce new illegal paths.

Very informal description

- There are illegal paths,
- add padding edges to remove illegal paths,
- this creates new illegal paths,
- add more padding edges to remove the new illegal paths,
- the second padding set does not induce new illegal paths.

We are done!

McGill

Undirected hardness

Theorem

The undirected rooted k-connectivity problem cannot be approximated to within $O(k^{\varepsilon})$, for some constant $\varepsilon > 0$, assuming that NP is not contained in $\mathrm{DTIME}(n^{polylog(n)})$.

- Improved from $\Omega(\log^{\Theta(1)} n)$,
- Best known approximation ratios are $\widetilde{O}(k)$.

Outline

1. k-DRC with $O(1)$ terminals.
2. Hardness of k-DRC (directed graph).
3. Hardness of k-URC (undirected graphs).
4. Integrality gap of k-DRC.

Integrality gap

Theorem

The natural LP relaxation of the directed rooted k-connectivity problem has an integrality ratio of $\Omega\left(\frac{k}{\log k}\right)$.

$$\min \sum_{e \in E} c_e x_e \quad \text{s.t.}$$

$$\sum_{e \in \delta^+(R)} x_e \geq k \quad (\forall R, r \in R, T \not\subseteq R)$$

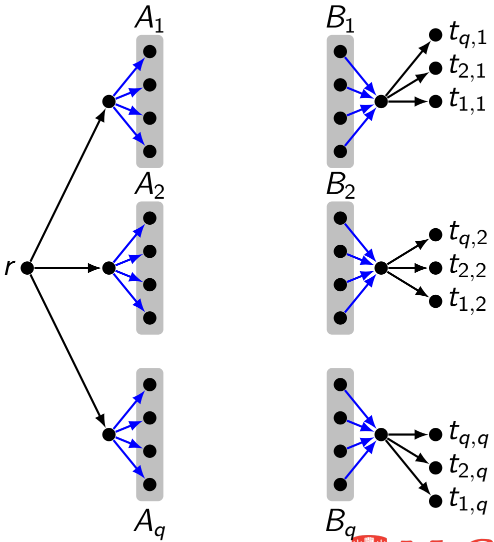
$$0 \leq x \leq 1$$

Proof: we follow a construction of Chakraborty, Chuzhoy, Khanna for SNDP integrality gap.

The construction

cost($\longrightarrow$) $= 1$
cost(others) $= 0$

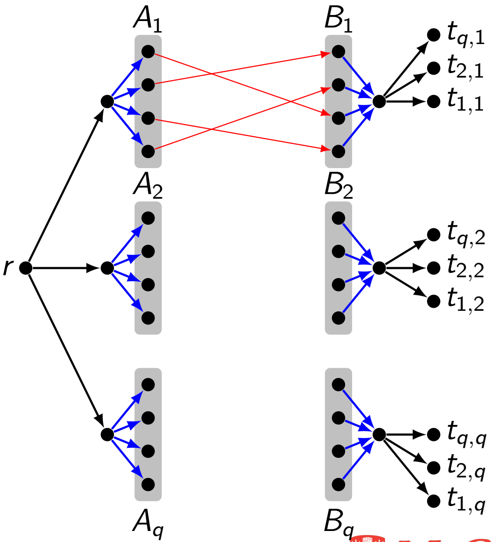
k: connectivity req.
$q = k$
$|A_i| = |B_j| = k^2$



The construction

cost($\longrightarrow$) = 1
cost(others) = 0

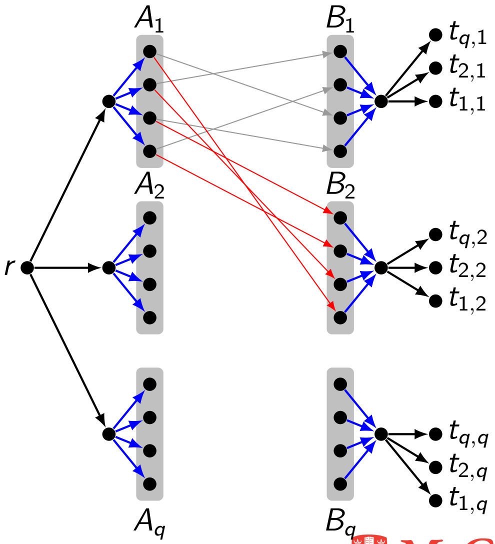
k: connectivity req.
$q = k$
$|A_i| = |B_j| = k^2$



The construction

cost($\longrightarrow$) = 1
cost(others) = 0

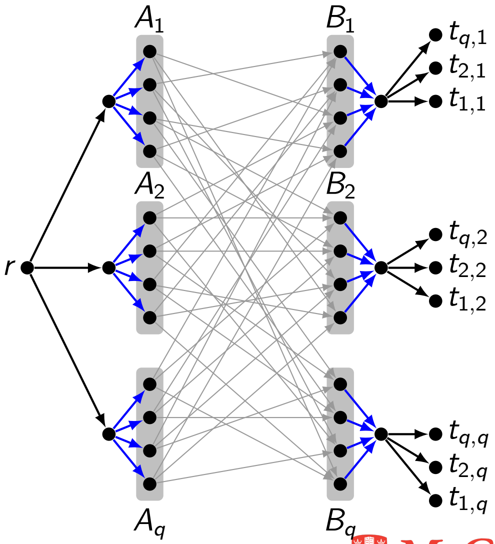
k: connectivity req.
$q = k$
$|A_i| = |B_j| = k^2$



The construction

cost($\longrightarrow$) $= 1$
cost(others) $= 0$

k: connectivity req.
$q = k$
$|A_i| = |B_j| = k^2$

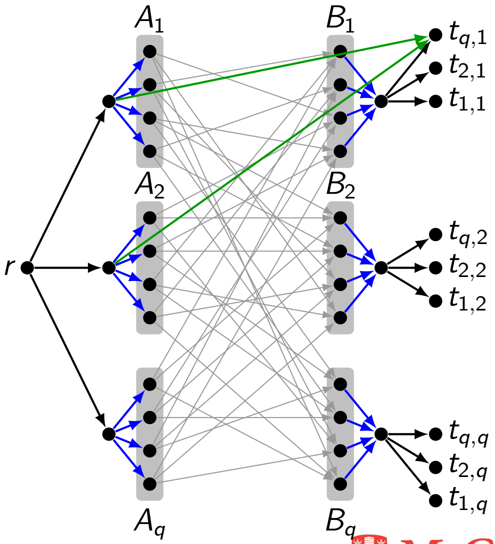


McGill

The construction

cost($\longrightarrow$) = 1
cost(others) = 0

k: connectivity req.
$q = k$
$|A_i| = |B_j| = k^2$

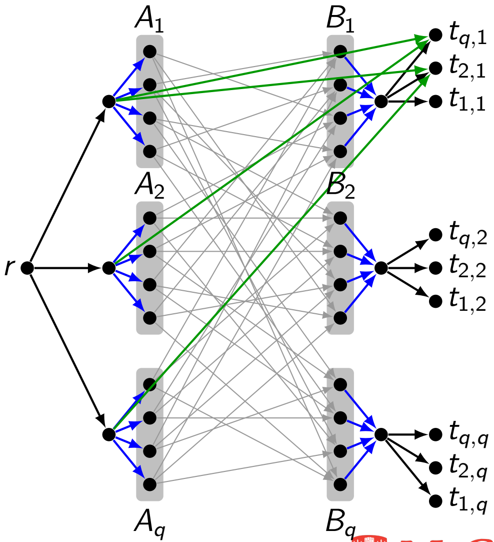


McGill

The construction

cost($\longrightarrow$) $= 1$
cost(others) $= 0$

k: connectivity req.
$q = k$
$|A_i| = |B_j| = k^2$

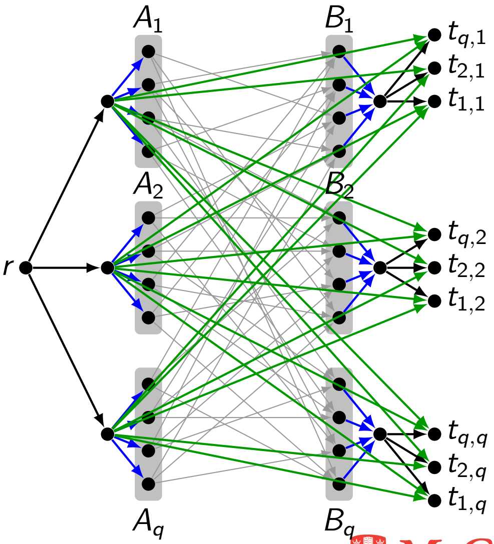


McGill

The construction

cost($\longrightarrow$) = 1
cost(others) = 0

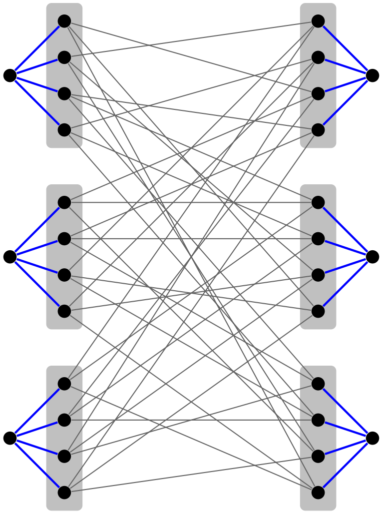
k: connectivity req.
$q = k$
$|A_i| = |B_j| = k^2$



Computing the gap

- Fractional solution:
  - $x_e = \frac{1}{k^2}$ for each $e \in E$ with $c(e) = 1$.
  - Total cost: $2q = 2k$

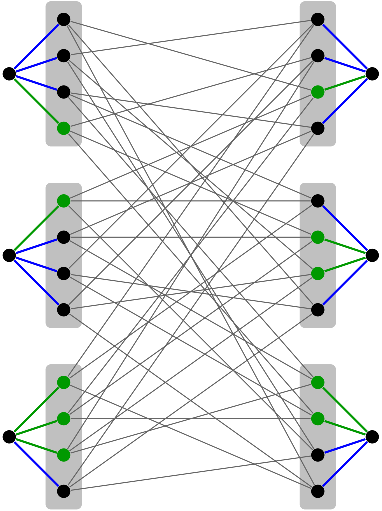
Computing the gap



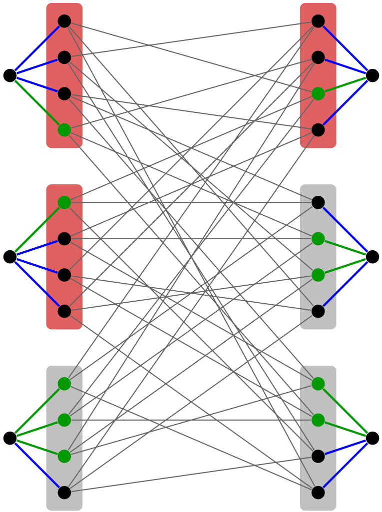
Computing the gap

- Fractional solution:
  - $x_e = \frac{1}{k^2}$ for each $e \in E$ with $c(e) = 1$.
  - Total cost: $2q = 2k$
- Integral solution:
  - Consider a subset S of arcs of cost $\leq \frac{\gamma k^2}{\log k}$,
  - prove $p_S = \Pr[S \text{ is an integral solution}]$ is very very small,
  - deduce $\sum_S p_S < 1$.
  - There is an instance without solution of cost $\leq \frac{\gamma k^2}{\log k}$.

Computing the gap



Computing the gap



Computing the gap

- Fractional solution:
  - $x_e = \frac{1}{k^2}$ for each $e \in E$ with $c(e) = 1$.
  - Total cost: $2q = 2k$
- Integral solution:
  - Consider a subset S of arcs of cost $\leq \frac{\gamma k^2}{\log k}$,
  - prove $p_S = \Pr[S \text{ is an integral solution}]$ is very very small,
  - deduce $\sum_S p_S < 1$.
  - There is an instance without solution of cost $\leq \frac{\gamma k^2}{\log k}$.

McGill

Computing the gap

- Fractional solution:
  - $x_e = \frac{1}{k^2}$ for each $e \in E$ with $c(e) = 1$.
  - Total cost: $2q = 2k$
- Integral solution:
  - Consider a subset S of arcs of cost $\leq \frac{\gamma k^2}{\log k}$,
  - prove $p_S = \Pr[S \text{ is an integral solution}]$ is very very small,
  - deduce $\sum_S p_S < 1$.
  - There is an instance without solution of cost $\leq \frac{\gamma k^2}{\log k}$.
- Integrality gap is $\Omega\left(\frac{k}{\log k}\right)$

McGill

Conclusion

- Other result:
  - Subset Connectivity problem.
- Open questions:
  - approximability when $\sum k_i = O(1)$?
  - inapproximablity when $k = O(1)$? (No better result known than DST)

McGill