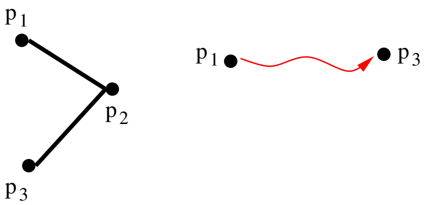
Approximation Algorithms and Hardness Results for Shortest Path Based Graph Orientations

Dima Blokh[1] Danny Segev[2] Roded Sharan[1]

1. School of Computer Science, Tel Aviv University,
2. Department of Statistics, University of Haifa

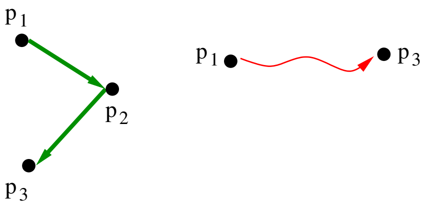
Motivation

- Protein-protein interaction networks are the skeleton for cellular signaling.
- Signaling is directed in nature
- Protein-protein interaction measurements output the presence of an interaction, but not its direction.
- Source-target experiments can be used to infer directions.



Motivation

- Protein-protein interaction networks are the skeleton for cellular signaling.
- Signaling is directed in nature
- Protein-protein interaction measurements output the presence of an interaction, but not its direction.
- Source-target experiments can be used to infer directions.



Maximum Graph Orientation (MGO) Problem Statement

- **Input:** An undirected graph G and a collection of source-target vertex pairs $P = \{(s_1, t_1), \ldots, (s_k, t_k)\}$.

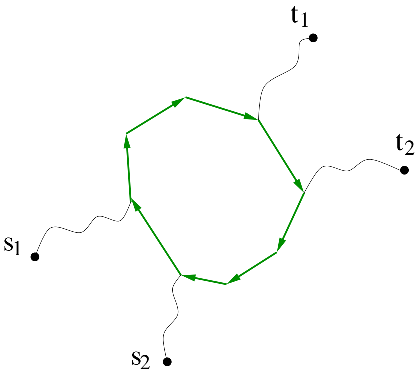
- **Objective:** Compute an orientation of G that satisfies a maximum number of pairs. (s, t) is *satisfied* by an orientation $\vec{G}$ when the latter contains a directed s-t path.

Related work

- **Medvedovsky et al.'08, Gamzu et al.'10 and Elberfeld et al.'11**:
  - MGO is NP-hard to approximate to within a factor better than 8/9.
  - $\Omega(\log\log n/\log n)$ approximation algorithm for MGO.
- **Silverbush et al.'11 and Medvedovsky et al.'08**:
  - Efficient integer programming formulations for MGO.
- **Gitter et al.'10**:
  - $O(r/2^r)$ approximation for paths whose length is bounded by a parameter r

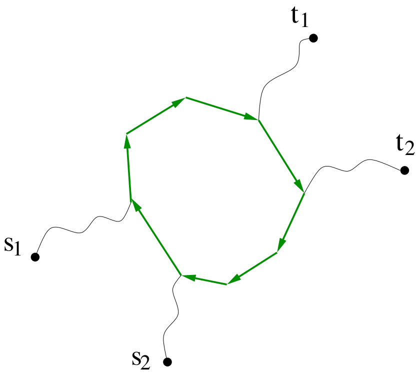
MGO Drawback

- In MGO cycles may be consistently oriented and contracted.
- End up with a tree network.



MGO Drawback

- In MGO cycles may be consistently oriented and contracted.
- End up with a tree network.



- Large fraction of the pairs are arbitrarily oriented.
- Long source-target paths.

Maximum Shortest Path Orientation (MSPO)

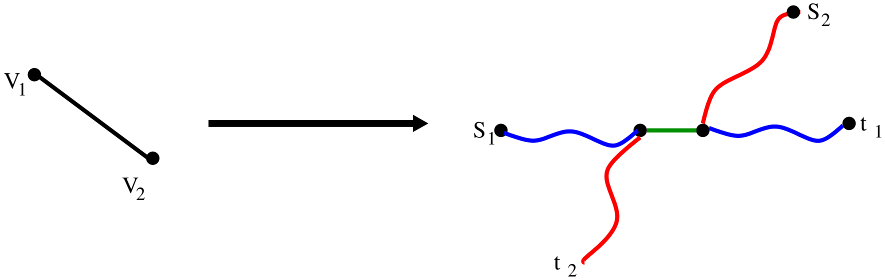
We focused on Maximum Shortest Path Orientation (MSPO) variant, in which the directed paths connecting each pair of source-target vertices are required to be shortest.

Our Contribution

- **Bad News:** NP-hard to approximate within factors of $O(1/k^{1-\epsilon})$ and $O(1/m^{1/3-\epsilon})$. This result can also be extended to the bounded length paths version studied by Gitter et al.
- **Good News:** $\Omega(1/\max\{n,k\}^{1/\sqrt{2}})$ approximation algorithm.
- **More Good News:** Significantly better approximation algorithms for almost shortest paths.

Hardness of Approximation

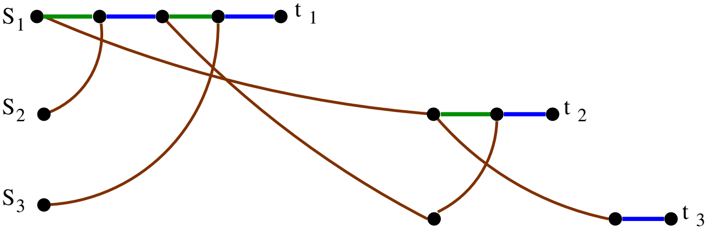
- Reduction from Independent Set.
  - Vertices in the *IS* graph $\rightarrow$ pairs.
  - Two vertices are adjacent $\leftrightarrow$ The corresponding shortest paths have a common edge in opposite directions.



The single-pair gadget

- Given k the single-pair gadget is an MSPO instance (G, P) with k pairs, $O(k^2)$ vertices and $O(k^2)$ edges, such that:
  1. For every pair in P there is some orientation that satisfies it.
  2. Any orientation of G satisfies at most one pair in P.

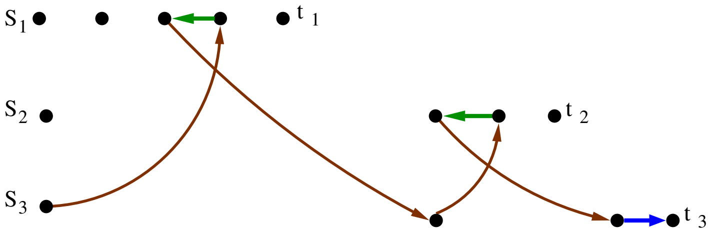
The single-pair gadget (cont.)



- The weight of blue edges is 2
- The weight of brown edges is 1
- The weight of green edges is $1/k$

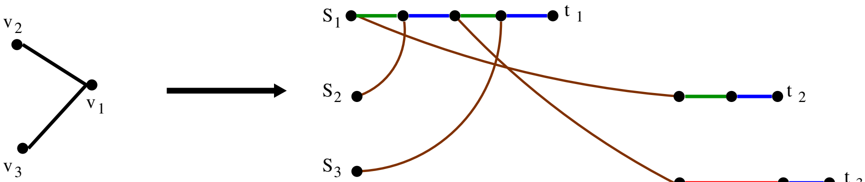
Main Claim

For every i, there is a unique shortest $s_i - t_i$ path. This path takes two brown edges and one green edge $i - 1$ times, and then traverses in left-to-right direction to reach t_i.



Reduction from Independent Set

- Given $G = (V, E)$, construct a single-pair gadget for $k = |V|$.
- $v_i \to s_i - t_i$.
- If $(v_i, v_j) \notin E$, replace the corresponding pair of brown edges by a single edge of weight 2.

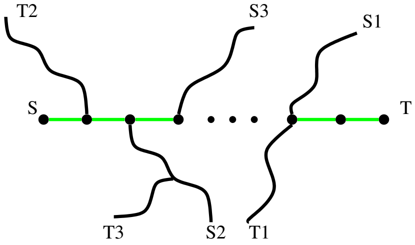


Reduction(cont.)

- $\{v_i : i \in I\}$ is an Independent Set $\leftrightarrow$ $\{(s_i, t_i) : i \in I\}$ can be satisfied.
- Resulting MSPO instance consists of n pairs and $O(n^2)$ edges.
- Bounds of $\Omega(1/k^{1-\epsilon})$ and $\Omega(1/m^{1/2-\epsilon})$ on the approximability of weighted MSPO.
- Bound of $\Omega(1/m^{1/3-\epsilon})$ for the unweighted MSPO.

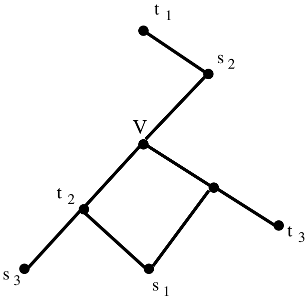
Approximation Algorithm

- While there is no vertex v which is common to β of the remaining paths:
  1. Orient the minimum length path.
  2. Discard this path as well as any path that overlaps with it.



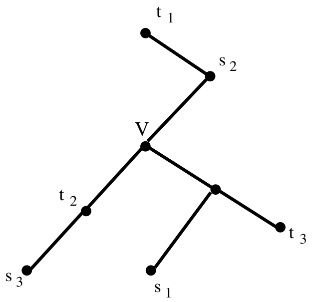
Approximation Algorithm (cont.)

- Satisfy at least 1/4 of the pairs going through v by constructing a shortest paths tree rooted in v and randomly directing every subtree from v or to v.



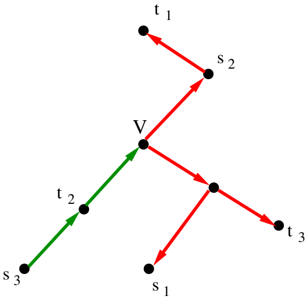
Approximation Algorithm (cont.)

- Satisfy at least 1/4 of the pairs going through v by constructing a shortest paths tree rooted in v and randomly directing every subtree from v or to v.



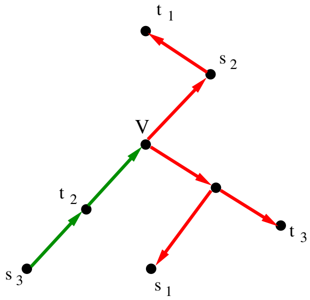
Approximation Algorithm (cont.)

- Satisfy at least 1/4 of the pairs going through v by constructing a shortest paths tree rooted in v and randomly directing every subtree from v or to v.



Approximation Algorithm (cont.)

- Satisfy at least 1/4 of the pairs going through v by constructing a shortest paths tree rooted in v and randomly directing every subtree from v or to v.



- After optimizing β the resulting approximation ratio is $\Omega(1/\max\{n,k\}^{1/\sqrt{2}})$.

Results Summary

- It is NP-hard to approximate MSPO to within factors $O(1/k^{1-\epsilon})$ and $O(1/m^{1/3-\epsilon})$.
- $\Omega(1/\max\{n,k\}^{1/\sqrt{2}})$ approximation algorithm.
- $\tilde{\Omega}(1/\sqrt{k})$ approximation for $1+\epsilon$-stretched paths.
- $\Omega(1/\log n)$ approximation for $\tilde{O}(\log n)$-stretched paths.