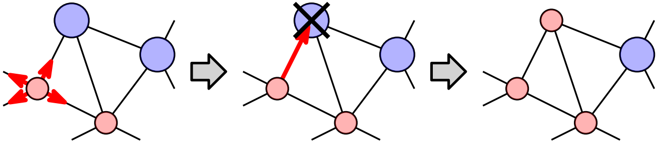
Fixation Maximization in the Positional Moran Process

Josef (Pepa) Tkadlec

Harvard University



Joint work with Aarhus U. members:
J. Brendborg, P. Karras, A. Pavlogiannis, A.U. Rasmussen

Spreading through networks

Settings:

- virus in human population,
- fake news / memes on a social network,
- interacting particle systems,
- genetic mutations in spatially structured populations, ...

Models:

1. "progressive": independent cascade, linear threshold, ...
2. "non-progressive": SIR-base, voter model, ...

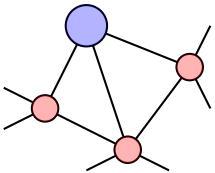
Objectives: influence maximization, epidemic control, diversity maintenance, fairness, ...

Here: Influence maximization under Moran process

Standard Moran process

[Moran '58] [Nature '05] A graph $G = (V, E)$ on n nodes.

- Nodes: agents (fitness: residents 1, mutants $1 + \delta$)

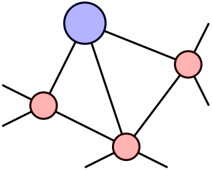


	fitness
mutant	$1 + \delta$
resident	1

Standard Moran process

[Moran '58] [Nature '05] A graph $G = (V, E)$ on n nodes.

- Nodes: agents (fitness: residents 1, mutants $1 + \delta$)
- Moran Birth-death process on a graph. Repeat:

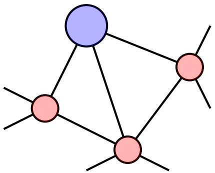


	fitness
mutant	$1 + \delta$
resident	1

Standard Moran process

[Moran '58] [Nature '05] A graph $G = (V, E)$ on n nodes.

- Nodes: agents (fitness: residents 1, mutants $1 + \delta$)
- Moran Birth-death process on a graph. Repeat:
  1. Birth: Pick a node for reproduction, proportionally to fitness

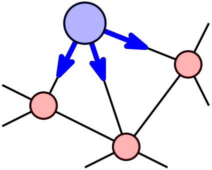


	fitness
mutant	$1 + \delta$
resident	1

Standard Moran process

[Moran '58] [Nature '05] A graph $G = (V, E)$ on n nodes.

- Nodes: agents (fitness: residents 1, mutants $1 + \delta$)
- Moran Birth-death process on a graph. Repeat:
  1. Birth: Pick a node for reproduction, proportionally to fitness

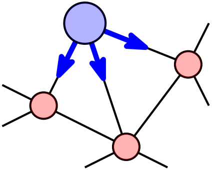


	fitness
mutant	$1 + \delta$
resident	1

Standard Moran process

[Moran '58] [Nature '05] A graph $G = (V, E)$ on n nodes.

- Nodes: agents (fitness: residents 1, mutants $1 + \delta$)
- Moran Birth-death process on a graph. Repeat:
  1. Birth: Pick a node for reproduction, proportionally to fitness
  2. Death: Pick a neighbor, randomly

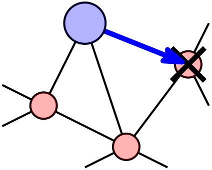


	fitness
mutant	$1 + \delta$
resident	1

Standard Moran process

[Moran '58] [Nature '05] A graph $G = (V, E)$ on n nodes.

- Nodes: agents (fitness: residents 1, mutants $1 + \delta$)
- Moran Birth-death process on a graph. Repeat:
  1. Birth: Pick a node for reproduction, proportionally to fitness
  2. Death: Pick a neighbor, randomly

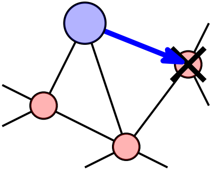


	fitness
mutant	$1 + \delta$
resident	1

Standard Moran process

[Moran '58] [Nature '05] A graph $G = (V, E)$ on n nodes.

- Nodes: agents (fitness: residents 1, mutants $1 + \delta$)
- Moran Birth-death process on a graph. Repeat:
  1. Birth: Pick a node for reproduction, proportionally to fitness
  2. Death: Pick a neighbor, randomly
  3. Replace

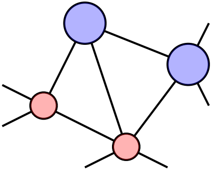


	fitness
mutant	$1 + \delta$
resident	1

Standard Moran process

[Moran '58] [Nature '05] A graph $G = (V, E)$ on n nodes.

- Nodes: agents (fitness: residents 1, mutants $1 + \delta$)
- Moran Birth-death process on a graph. Repeat:
    1. Birth: Pick a node for reproduction, proportionally to fitness
    2. Death: Pick a neighbor, randomly
    3. Replace

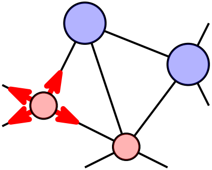


	fitness
mutant	$1 + \delta$
resident	1

Standard Moran process

[Moran '58] [Nature '05] A graph $G = (V, E)$ on n nodes.

- Nodes: agents (fitness: residents 1, mutants $1 + \delta$)
- Moran Birth-death process on a graph. Repeat:
  1. Birth: Pick a node for reproduction, proportionally to fitness
  2. Death: Pick a neighbor, randomly
  3. Replace

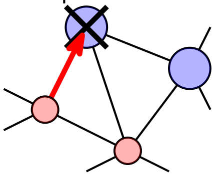


	fitness
mutant	$1 + \delta$
resident	1

Standard Moran process

[Moran '58] [Nature '05] A graph $G = (V, E)$ on n nodes.

- Nodes: agents (fitness: residents 1, mutants $1 + \delta$)
- Moran Birth-death process on a graph. Repeat:
    1. Birth: Pick a node for reproduction, proportionally to fitness
    2. Death: Pick a neighbor, randomly
    3. Replace

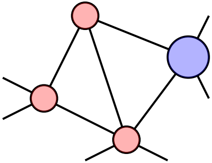


	fitness
mutant	$1 + \delta$
resident	1

Standard Moran process

[Moran '58] [Nature '05] A graph $G = (V, E)$ on n nodes.

- Nodes: agents (fitness: residents 1, mutants $1 + \delta$)
- Moran Birth-death process on a graph. Repeat:
  1. Birth: Pick a node for reproduction, proportionally to fitness
  2. Death: Pick a neighbor, randomly
  3. Replace

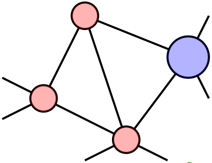


	fitness
mutant	$1 + \delta$
resident	1

Standard Moran process

[Moran '58] [Nature '05] A graph $G = (V, E)$ on n nodes.

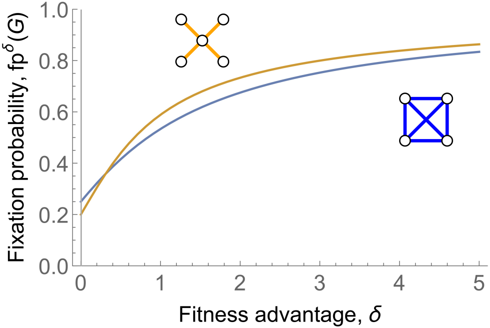
- Nodes: agents (fitness: residents 1, mutants $1 + \delta$)
- Moran Birth-death process on a graph. Repeat:
  1. Birth: Pick a node for reproduction, proportionally to fitness
  2. Death: Pick a neighbor, randomly
  3. Replace



	fitness
mutant	$1 + \delta$
resident	1

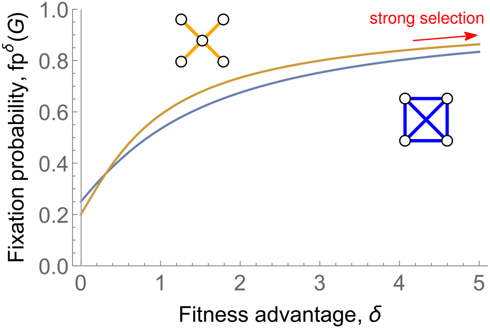
Fixation probability $\mathrm{fp}^{\delta}(G)$: Average probability that, starting from a single node, mutants spread to all sites.

Moran process: what is known?



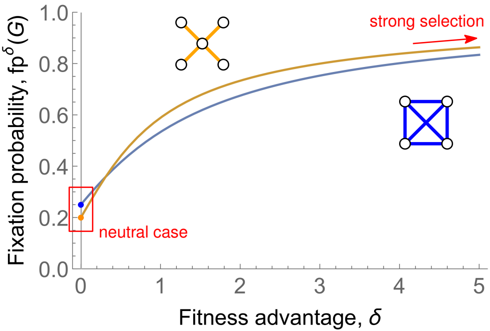
1. Strong selection: $\mathrm{fp}^{\delta\to\infty}(G) = 1$.
2. Neutral case: $\mathrm{fp}^{\delta=0}(G) = 1/n$.
3. Weak selection: $\mathrm{fp}^{\delta\to 0}(G) = 1/n + \delta \cdot c(G) + \mathcal{O}(\delta^2)$, where the "slope" $c(G)$ can be computed in P-time [JMB '21].
4. The complexity of computing $\mathrm{fp}^{\delta}(G)$ is open.
5. If G is undirected then $\mathrm{fp}^{\delta}(G)$ can be efficiently approximated (FPRAS) [SODA '12].

Moran process: what is known?



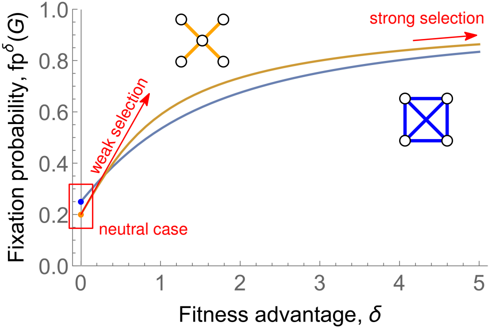
1. Strong selection: $\mathrm{fp}^{\delta\to\infty}(G) = 1$.
2. Neutral case: $\mathrm{fp}^{\delta=0}(G) = 1/n$.
3. Weak selection: $\mathrm{fp}^{\delta\to 0}(G) = 1/n + \delta \cdot c(G) + \mathcal{O}(\delta^2)$, where the "slope" $c(G)$ can be computed in P-time [JMB '21].
4. The complexity of computing $\mathrm{fp}^{\delta}(G)$ is open.
5. If G is undirected then $\mathrm{fp}^{\delta}(G)$ can be efficiently approximated (FPRAS) [SODA '12].

Moran process: what is known?



1. Strong selection: $\mathrm{fp}^{\delta\to\infty}(G) = 1$.
2. Neutral case: $\mathrm{fp}^{\delta=0}(G) = 1/n$.
3. Weak selection: $\mathrm{fp}^{\delta\to 0}(G) = 1/n + \delta \cdot c(G) + \mathcal{O}(\delta^2)$, where the "slope" $c(G)$ can be computed in P-time [JMB '21].
4. The complexity of computing $\mathrm{fp}^{\delta}(G)$ is open.
5. If G is undirected then $\mathrm{fp}^{\delta}(G)$ can be efficiently approximated (FPRAS) [SODA '12].

Moran process: what is known?

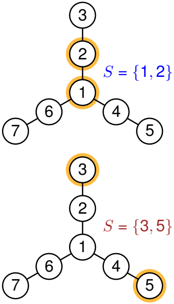


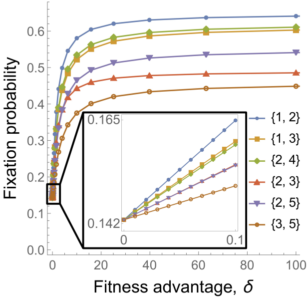
1. Strong selection: $\mathrm{fp}^{\delta\to\infty}(G) = 1$.
2. Neutral case: $\mathrm{fp}^{\delta=0}(G) = 1/n$.
3. Weak selection: $\mathrm{fp}^{\delta\to 0}(G) = 1/n + \delta \cdot c(G) + \mathcal{O}(\delta^2)$, where the "slope" $c(G)$ can be computed in P-time [JMB '21].
4. The complexity of computing $\mathrm{fp}^{\delta}(G)$ is open.
5. If G is undirected then $\mathrm{fp}^{\delta}(G)$ can be efficiently approximated (FPRAS) [SODA '12].

Positional Moran process

Idea: Mutants are better only at a subset $S \subseteq V$ of nodes.
Quantity: Fixation probability $\mathrm{fp}^{\delta}(G^S)$.

agent	fitness
	$1+\delta$
	1
	1
	1



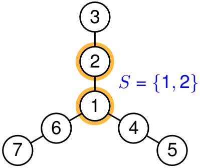


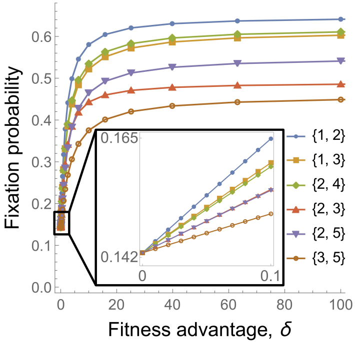
Fixation maximization $\mathtt{FM}(G, \delta, k)$: Given $k \in \mathbb{N}$, which subset $S \subseteq V$ with $|S| = k$ maximizes $\mathrm{fp}^{\delta}(G^S)$?
Specifically,

1. Strong selection: $\mathtt{FM}^{\infty}(G, k)$: max $\mathrm{fp}^{\delta \to \infty}(G^S)$.
2. Weak selection: $\mathtt{FM}^{0}(G, k)$: max "slope" $c(G, S)$ as $\delta \to 0$.

Our results

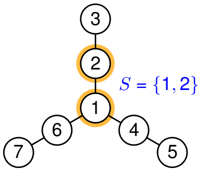
1. Strong selection ($\delta \to \infty$):
   1.1 Computing $\mathtt{FM}^\infty(G, k)$ is NP-hard, even on regular graphs.
   1.2 But it can be $(1 - 1/e)$-approximated (on any graph).
2. Weak selection ($\delta \to 0$):
   2.1 $\mathtt{FM}^0(G, k)$ can be computed in $\mathcal{O}(n^{2\omega})$ time.

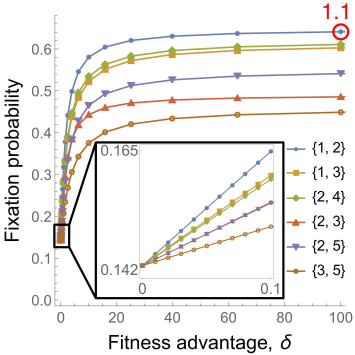




Our results

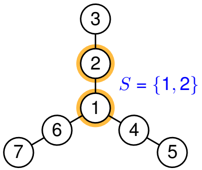
1. Strong selection ($\delta \to \infty$):
   1.1 Computing $\mathtt{FM}^{\infty}(G, k)$ is $\mathtt{NP}$-hard, even on regular graphs.
   1.2 But it can be $(1 - 1/e)$-approximated (on any graph).
2. Weak selection ($\delta \to 0$):
   2.1 $\mathtt{FM}^{0}(G, k)$ can be computed in $\mathcal{O}(n^{2\omega})$ time.

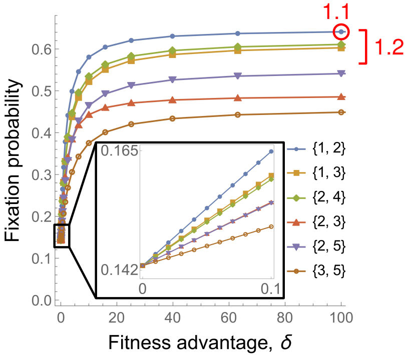




Our results

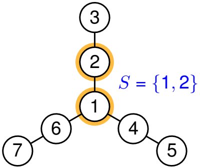
1. Strong selection ($\delta \to \infty$):
   1.1 Computing $\mathtt{FM}^{\infty}(G, k)$ is $\mathtt{NP}$-hard, even on regular graphs.
   1.2 But it can be $(1 - 1/e)$-approximated (on any graph).
2. Weak selection ($\delta \to 0$):
   2.1 $\mathtt{FM}^{0}(G, k)$ can be computed in $\mathcal{O}(n^{2\omega})$ time.

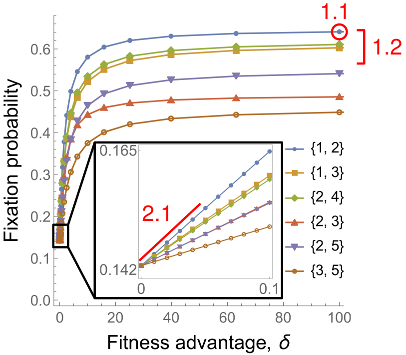




Our results

1. Strong selection ($\delta \to \infty$):
   1.1 Computing $\mathtt{FM}^{\infty}(G, k)$ is $\mathtt{NP}$-hard, even on regular graphs.
   1.2 But it can be $(1 - 1/e)$-approximated (on any graph).
2. Weak selection ($\delta \to 0$):
   2.1 $\mathtt{FM}^{0}(G, k)$ can be computed in $\mathcal{O}(n^{2\omega})$ time.

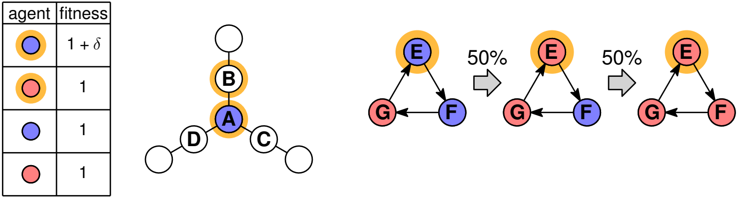




Strong selection: two key insights

(i) Observation. While no mutant is in S, the process is neutral.

(ii) Claim. Once S contains a mutant, mutants win whp.



Proof idea: Such a mutant is essentially immortal, since it forces its neighbors to be mutants almost all the time.
(But: The claim is not true if G is directed.)

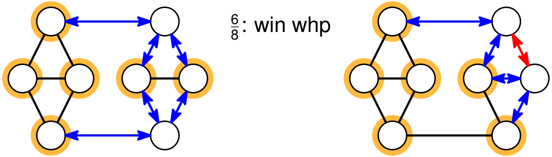
Strong selection: 1.1 hardness

Theorem

Deciding whether $\mathit{FM}^{\infty}(G, k) \geq \tau$ is NP-hard.

S is a vertex cover

not a vertex cover



$\frac{6}{8}$: win whp

$\frac{2}{8}$: each blue edge wins 50/50

$\frac{2}{8}$: red edge wins less than 50/50

Proof idea: If G is regular then $\mathrm{fp}^{\infty}(G^S) \geq \frac{n+|S|}{2n}$ iff S is a vertex cover. Indeed, if the initial mutant:

1. lands on S: then mutants win whp;
2. lands outside S: then
   - 2.1 if S is a vertex cover: mutants win with probability $1/2$;
   - 2.2 otherwise: mutants win with probability $p \lnsim 1/2$.

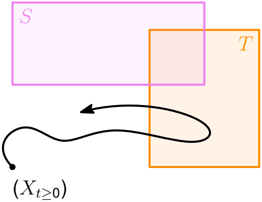
Thus, deciding whether there exists $|S| = k$ with $\mathrm{fp}^{\infty}(G^S) \geq \frac{n+k}{2n}$ is as hard as deciding whether G has a vertex cover of size k.

Strong selection: 1.2 submodularity

Lemma

The function $\mathrm{fp}^\infty(G^S)$ *is submodular, i.e.*
$\mathrm{fp}^\infty(G^S) + \mathrm{fp}^\infty(G^T) \geq \mathrm{fp}^\infty(G^{S\cup T}) + \mathrm{fp}^\infty(G^{S\cap T})$.

Corollary. Thus, greedy gives a $(1-1/e)$-approximation for $\mathtt{FM}^\infty(G,k)$ [NWF '78].



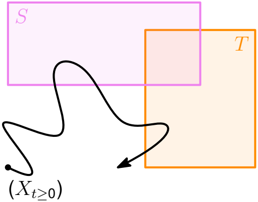
$$\mathrm{fp}^\infty(G^S) + \underline{\mathrm{fp}^\infty(G^T)} \overset{?}{\geq} \underline{\mathrm{fp}^\infty(G^{S\cup T})} + \mathrm{fp}^\infty(G^{S\cap T})$$

Proof idea. Fix a neutral trajectory $\tau = (X_0, X_1, \dots)$ and track whether it hits any of $S \setminus T$, $S \cap T$, $T \setminus S$. Small casework gives that τ contributes to LHS at least as much as it does to RHS.

Strong selection: 1.2 submodularity

Lemma
The function $\mathrm{fp}^{\infty}(G^S)$ *is submodular, i.e.*
$\mathrm{fp}^{\infty}(G^S) + \mathrm{fp}^{\infty}(G^T) \geq \mathrm{fp}^{\infty}(G^{S\cup T}) + \mathrm{fp}^{\infty}(G^{S\cap T})$.

Corollary. Thus, greedy gives a $(1 - 1/e)$-approximation for $\mathtt{FM}^{\infty}(G, k)$ [NWF '78].



$$\underline{\mathrm{fp}^{\infty}(G^S)} + \underline{\mathrm{fp}^{\infty}(G^T)} \overset{?}{\geq} \underline{\mathrm{fp}^{\infty}(G^{S\cup T})} + \mathrm{fp}^{\infty}(G^{S\cap T})$$

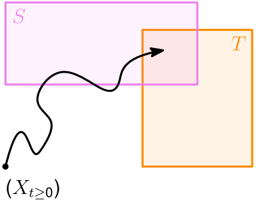
Proof idea. Fix a neutral trajectory $\tau = (X_0, X_1, \dots)$ and track whether it hits any of $S \setminus T$, $S \cap T$, $T \setminus S$. Small casework gives that τ contributes to LHS at least as much as it does to RHS.

Strong selection: 1.2 submodularity

Lemma

The function $\mathrm{fp}^{\infty}(G^S)$ *is submodular, i.e.*
$\mathrm{fp}^{\infty}(G^S) + \mathrm{fp}^{\infty}(G^T) \geq \mathrm{fp}^{\infty}(G^{S\cup T}) + \mathrm{fp}^{\infty}(G^{S\cap T})$.

Corollary. Thus, greedy gives a $(1 - 1/e)$-approximation for $\mathtt{FM}^{\infty}(G, k)$ [NWF '78].



$$\underline{\mathrm{fp}^{\infty}(G^S)} + \underline{\mathrm{fp}^{\infty}(G^T)} \overset{?}{\geq} \underline{\mathrm{fp}^{\infty}(G^{S\cup T})} + \underline{\mathrm{fp}^{\infty}(G^{S\cap T})}$$

Proof idea. Fix a neutral trajectory $\tau = (X_0, X_1, \dots)$ and track whether it hits any of $S \setminus T$, $S \cap T$, $T \setminus S$. Small casework gives that τ contributes to LHS at least as much as it does to RHS.

Weak selection: 2.1

Lemma
We have $\mathrm{fp}^{\delta\to 0}(G^S) = 1/n + \delta \cdot c(G,S) + \mathcal{O}(\delta^2)$, *where*

$$c(G,S) = \sum_{u\in S} \alpha(u),$$

for a certain function $\alpha\colon V \to \mathbb{R}$ *that can be found in* $\mathcal{O}(n^{2\omega})$ *time.*

Corollary. To maximize $\mathtt{FM}^0(G,k)$, compute all $\alpha(u_i)$ and take the top k nodes into S.

Proof idea: Building on [JMB '21]. Briefly:

1. Let $\psi_{i,j}$ be the expected amount of time for which u_i is a mutant and u_j not.
2. Then $(\psi_{i,j} \mid (u_i,u_j)\in E)$ can be found by solving a linear system.
3. The value $\alpha(u_i)$ is a certain weighted average of $\psi_{i,j}$, taking into account "how strongly u_i and u_j interact".

Summary

1. Strong selection ($\delta \to \infty$):
   1.1 Computing $\mathtt{FM}^{\infty}(G, k)$ is $\mathtt{NP}$-hard, even on regular graphs, by vertex cover.
   1.2 But it can be $(1 - 1/e)$-approximated, since $\mathrm{fp}^{\infty}(G^S)$ is submodular. (We note that for $\delta < \infty$ the function $\mathrm{fp}^{\delta}(G^S)$ is not in general submodular.)
2. Weak selection ($\delta \to 0$):
   2.1 $\mathtt{FM}^{0}(G, k)$ can be computed in $\mathcal{O}(n^{2\omega})$ time, since each $u_i \in S$ contributes certain fixed amount $\alpha(u_i)$ to the slope $c(G, S)$.

Open questions

1. Is there a better approximation factor for $\mathtt{FM}^{\infty}(G, k)$? Perhaps even a FPRAS?

2. Is there a constant-factor approximation for $\mathtt{FM}^{\delta}(G, k)$?
   - Can not be based on submodularity.

3. For a given S, what is the complexity of computing $\mathrm{fp}^{\delta}(G^{S})$?
   - Open even in special cases $S = V$ or $\delta \to \infty$ (but not both).

Thank you!