Microsoft Research

Microsoft

On Adaptivity Gaps of Influence Maximization under the Independent Cascade Model with Full-Adoption Feedback

Wei Chen 陈卫
Microsoft Research Asia

Binghui Peng 彭炳辉
Columbia University

ISAAC'2019, Dec. 11, 2019

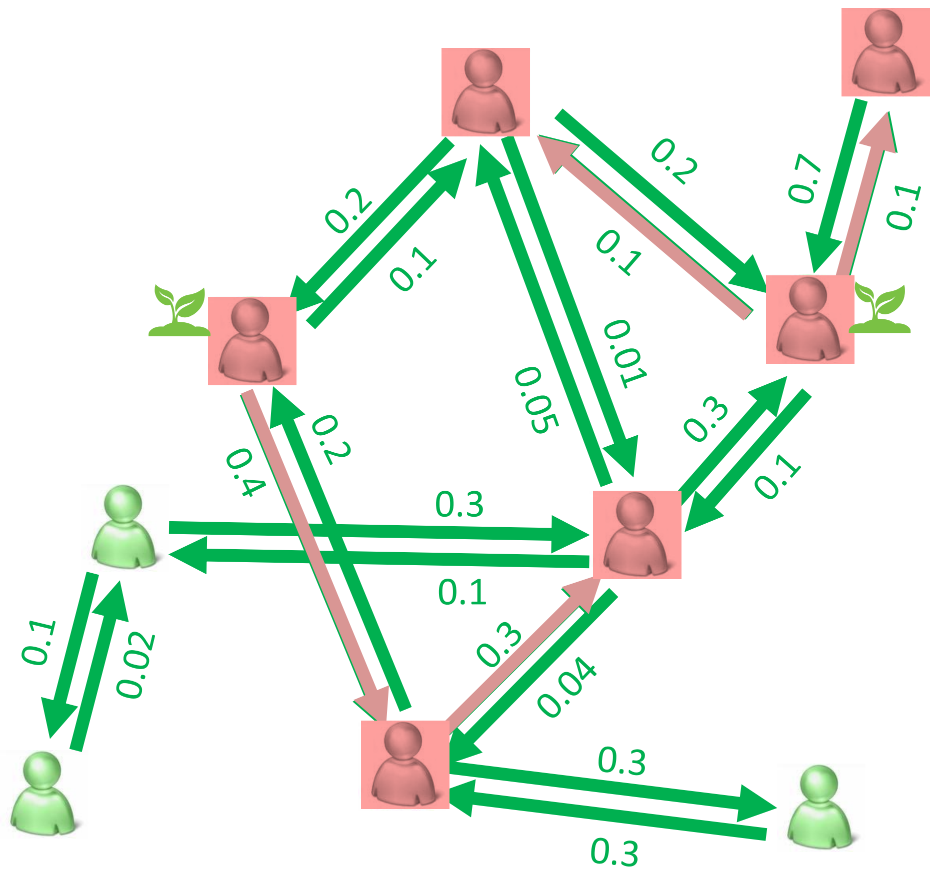
Influence Propagation Modeling and Optimizations

- How to model influence propagation in a social network?
  - Stochastic diffusion models
- How to optimize the influence propagation effect?
  - Influence maximization and its variants
- One core problem: Influence maximization
  - Find a small number of individuals in a network to generate a large influence
  - Applications in viral marketing, diffusion monitoring, rumor control, etc.

Model and Problem

Independent Cascade (IC) Model

- Social graph $G = (V, E)$, $n = |V|$
- Each edge (u, v) has an *influence probability* $p(u, v)$
- Initially seed nodes in S are activated
- At each step t, each node u activated at step $t-1$ activates its neighbor v independently with probability $p(u, v)$
- Influence spread $\sigma(S)$: expected number of activated nodes



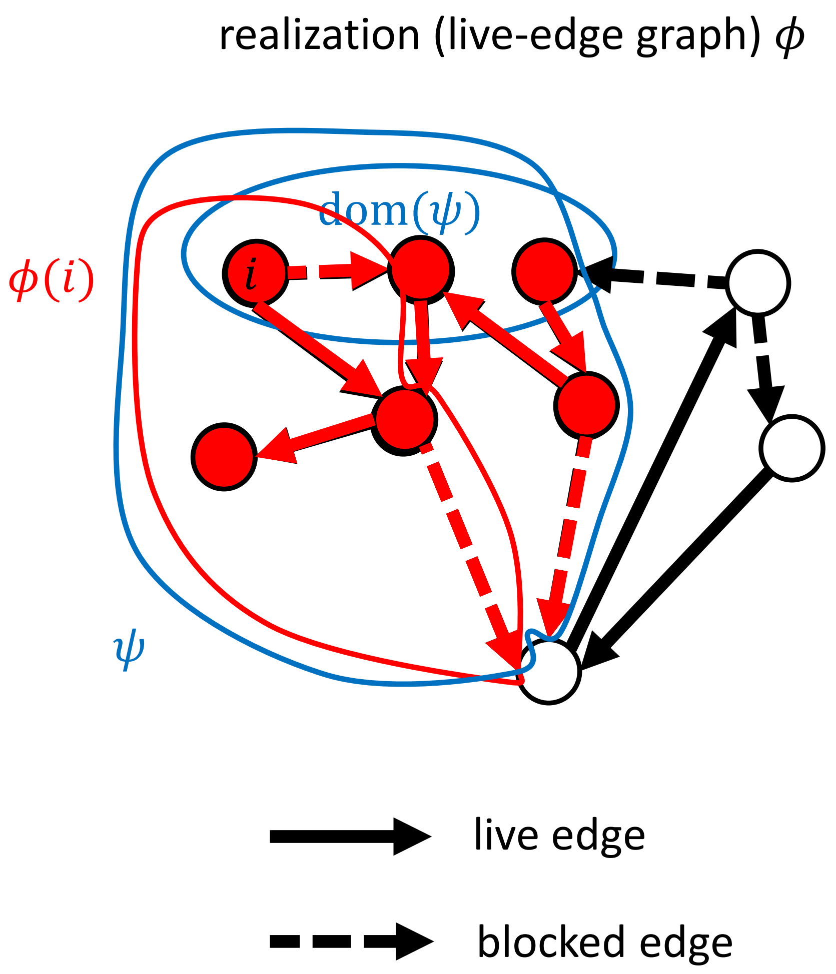
Influence Maximization

- Given a social network, a diffusion model with given parameters, and a number $\boldsymbol{k}$, find a seed set $\boldsymbol{S}$ of at most $\boldsymbol{k}$ nodes such that the influence spread of $\boldsymbol{S}$ is maximized.
- Based on *submodular function* maximization
  - Submodularity of set functions $f: 2^{\mathrm{V}} \rightarrow R$:
    - for all $S \subseteq T \subseteq V$, all $v \in V \setminus T$, $f(S \cup \{v\}) - f(S) \geq f(T \cup \{v\}) - f(T)$
  - Monotonicity: for all $S \subseteq T \subseteq V, f(S) \leq f(T)$
- Influence spread function $\boldsymbol{\sigma}(\boldsymbol{S})$ in IC model is submodular
- Greedy algorithm achieves $\mathbf{1} - \mathbf{1}/\boldsymbol{e}$ approximation

Kempe D, Kleinberg J M, and Tardos É. Maximizing the spread of influence through a social network. KDD'2003

Feedback Model: Realization and Partial Realization

- Realization $\boldsymbol{\phi}$ (Random realization $\boldsymbol{\Phi}$): all randomness in a propagation
  - In IC model, $\boldsymbol{\phi}$ (or $\boldsymbol{\Phi}$) is a (random) live-edge graph: each edge $(\boldsymbol{u}, \boldsymbol{v})$ is selected with probability $\boldsymbol{p}(\boldsymbol{u}, \boldsymbol{v})$
  - For full-adoption feedback, for each node $\boldsymbol{i}$, $\boldsymbol{\phi}(\boldsymbol{i})$ is the full cascade sequence and edge status, i.e., all reachable edges and nodes in live-edge graph $\boldsymbol{\phi}$
- Partial realization $\boldsymbol{\psi}$ (random partial realization $\boldsymbol{\Psi}$): feedback collected (partial propagation) from the currently selected seeds $\mathbf{dom}(\boldsymbol{\psi})$



Adaptive Influence Maximization

- Adaptive policy π: given any ψ, select the next node $\pi(\psi)$
  - adaptive influence spread $\sigma(\pi)$: expected number of nodes activated by π
- Adaptive influence maximization: find best policy π^* that selects at most k nodes, and maximizes adaptive influence spread $\sigma(\pi)$

Adaptivity Gap

- Supremum ratio of the adaptive optimal vs. non-adaptive optimal

$$\sup_{G,k} \frac{\mathrm{OPT}_A\,(G,k)}{\mathrm{OPT}_N(G,k)}$$

- Important to measure the effectiveness of adaptivity
- Related work (with two different approaches)
  - [ANS08, AN15]: stochastic submodular optimization on matroid, adaptivity gap: $\frac{e}{e-1}$
    - Approach: multilinear extension + Poisson process
  - [GNS16, GNS17, BSZ19]: stochastic probing, adaptivity gap: 2
    - Approach: decision tree + random-walk + fictitious hybrid policy
  - Implicitly rely on adaptive submodularity *and* feedback independence

Adaptive Submodularity

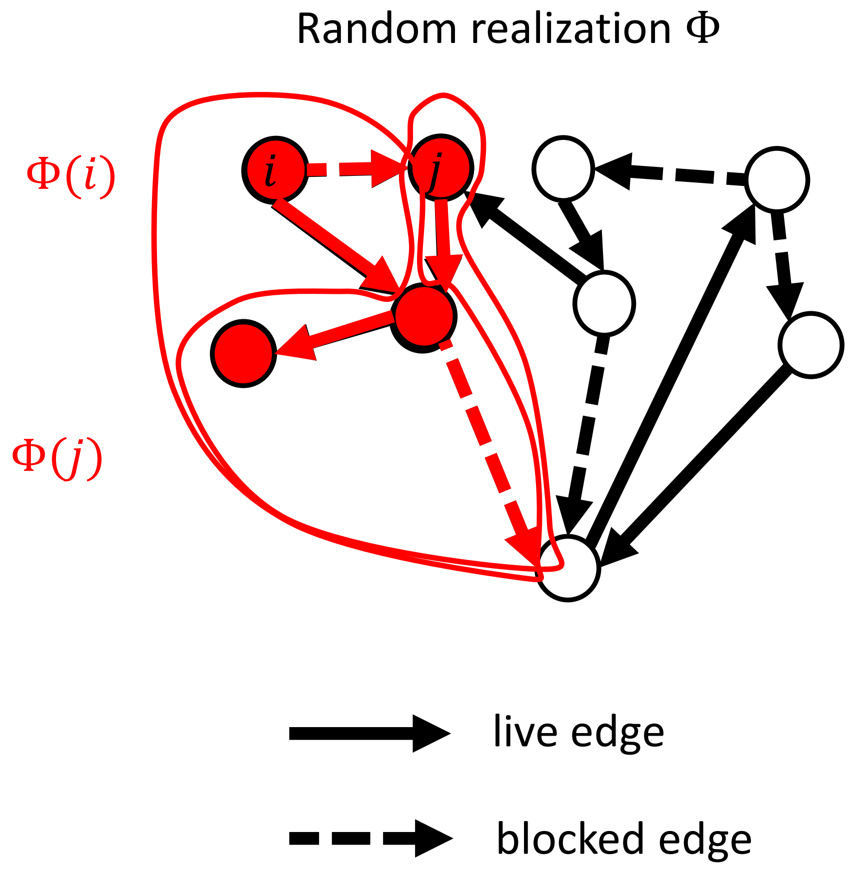
- Adaptive Submodularity [GK11]: a node $\boldsymbol{u}$'s marginal influence is higher on a smaller partial realization than on a larger partial realization
  - $\psi \subseteq \psi' \Rightarrow \Delta(u|\psi') \leq \Delta(u|\psi)$
- Adaptive Monotonicity: a node $\boldsymbol{u}$'s marginal influence on any partial realization is nonnegative
  - $\boldsymbol{\Delta(u|\psi) \geq 0}$, as long as $\boldsymbol{\psi}$ has non-zero probability to occur
- IC + Full-adoption is adaptive submodular [GK11]

Golovin D and Krause A. Adaptive submodularity: theory and applications in active learning and stochastic optimization. Journal of Artificial Intelligence Research, 2011

Adaptivity Gap in IC Model with Full-Adoption Feedback

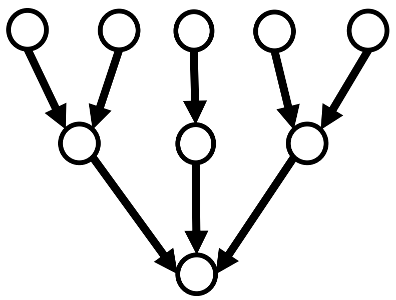
IC Model + Full-Adoption Feedback

- But not feedback-independent:
  - Propagation from two nodes may overlap
  - And thus the feedback are dependent
- Thus prior results on adaptivity gap do not apply
- Adaptivity gap on general graphs is still open

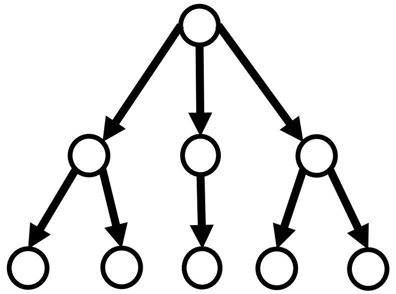


Our Results

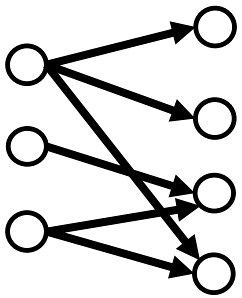
- In-arborescences: $\left[\frac{e}{e-1}, \frac{2e}{e-1}\right]$
  - In-arborescence: a tree structure with directed edges all pointing towards the root



- Out-arborescences: $\left[\frac{e}{e-1}, 2\right]$
  - Out-arborescence: a tree structure with directed edges all pointing towards the leaves



- One-directional bipartite graphs: $\frac{e}{e-1}$
  - Directed edges all pointing from one side to the other side



Analysis for In-Arborescence, Gap $\leq \frac{2e}{e-1}$

- Adapted from the approach in [AN16]
  - Multilinear extension + Poisson process + handling correlated feedback
- Multilinear extension F of influence spread function $\sigma(S)$

$$F(x_1, \ldots, x_n) = \sum_{S \subseteq V} \left[\left(\prod_{i \in S} x_i \prod_{i \notin S} (1 - x_i) \right) \sigma(S) \right]$$

- Poisson process, with parameters $(x_1, \ldots, x_n)$
  - n independent Poisson clocks $C_1, \ldots, C_n$, C_i has rate x_i
  - when C_i signals, node i is selected as a seed, and get feedback $\phi(i)$
  - $\Psi(t)$ is the partial realization by time t, process stops at $t = 1$

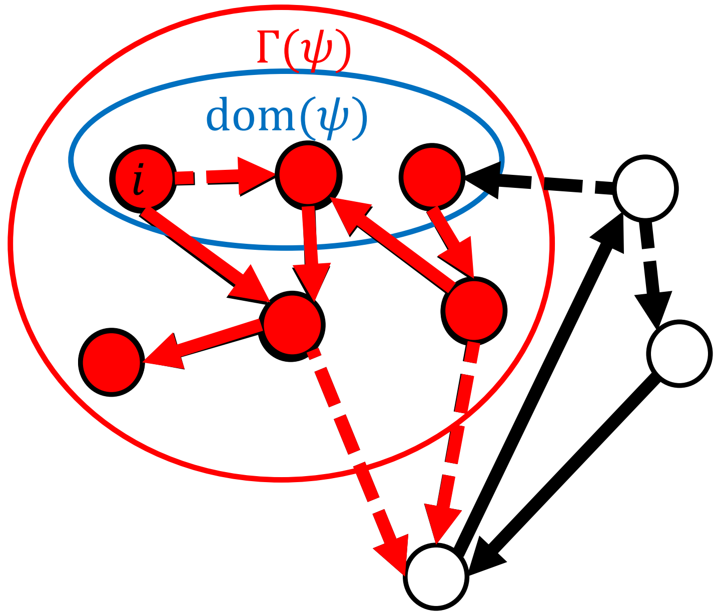
Asadpour A and Nazerzadeh H. Maximizing stochastic monotone submodular functions. Management Science, 2016

Connecting Multilinear Extension to Poisson Process

▶ **Lemma 3.2.** $\mathbb{E}\left[f\left(\Psi(1)\right)\right] = F\left(1 - e^{-x_1}, \ldots, 1 - e^{-x_n}\right) \le F(x_1, \ldots, x_n)$.

- $\boldsymbol{\Gamma(\psi)}$: set nodes activated from seeds in $\mathbf{dom}(\boldsymbol{\psi})$, based on feedback ψ
- Function $f(\psi) = |\Gamma(\psi)|$, number of activated nodes in ψ
- Proof:

$$\mathbb{E}\left[f\left(\Psi(1)\right)\right] = \sum_{S \subseteq V} \left[\left(\prod_{i \in S}\left(1 - e^{-x_i}\right)\right) \prod_{i \notin S}\left(e^{-x_i}\right) \sigma(S)\right]$$
$$= F(1 - e^{-x_1}, \ldots, 1 - e^{-x_n}) \le F(x_1, \ldots, x_n).$$



Parametrized Optimal Adaptive Influence Spread Function

$$f^+(x_1, \ldots, x_n) = \sup_{\pi} \left\{ \sigma(\pi) : \Pr_{\Phi \sim \mathcal{P}} [i \in V(\pi, \Phi)] = x_i, \forall i \in [n] \right\}.$$

- f^+: the adaptive influence spread of the best adaptive policy, among all adaptive policies π guaranteeing that node i is selected as a seed with probability x_i (for all i)

Connecting Poisson Process with Optimal Adaptive Influence Spread

- **Lemma 3.3.** *For any $t \in [0,1]$ and any fixed partial realization ψ, we have*

$$\mathbb{E}\left[\frac{df\left(\Psi(t)\right)}{dt} \mid \Psi(t)=\psi\right] \geq f^{+}\left(x_1,\ldots,x_n\right)-\sigma\left(\Gamma(\psi)\right).$$

due to feedback correlation

- Intuition: differential on process $\Psi(t)$ is related to adaptive marginal $\Delta(i|\psi)$
- Poisson process is the bridge linking non-adaptive influence spread (Lemma 3.2) with adaptive influence spread (Lemma 3.3)
- $\sigma(\Gamma(\psi))$ is due to feedback correlation.
  - If feedback were independent, it would be $|\Gamma(\psi)|$ [AN16]
  - resulting in the extra factor of 2 in adaptivity gap
- Require $\Gamma(\psi) \subseteq \Gamma(\psi') \Rightarrow \Delta(u|\psi') \leq \Delta(u|\psi)$, slightly stronger than adaptive submodularity

In-arborescence --- Shrinking Boundary

- $\partial(\psi)$: boundary of ψ, nodes separating internal nodes in $\Gamma(\psi)$ from outside nodes in $V \setminus \Gamma(\psi)$
- Boundary $\partial(\psi)$ shrinks from $\mathbf{dom}(\psi)$ in in-arborescences

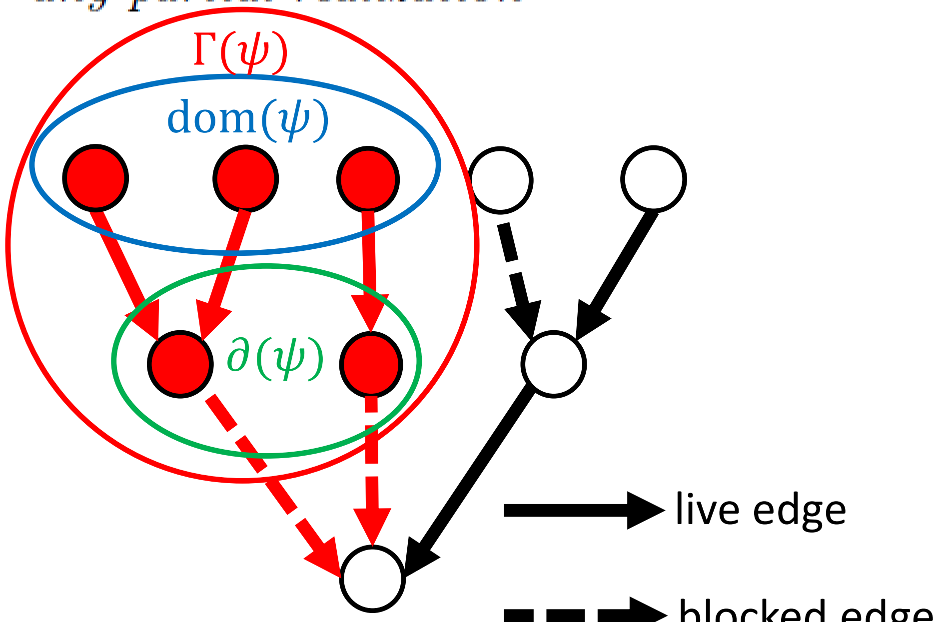
▶ **Lemma 3.7.** *When the influence graph is an in-arborescence, for any partial realization ψ, we have* $|\partial(\psi)| \leq |\mathrm{dom}(\psi)|$.

▶ **Lemma 3.8.** *For any partial realization ψ*

$$\sigma(\Gamma(\psi)) \leq |\Gamma(\psi)| + \sigma(\partial(\psi)).$$

Moreover, when the influence graph is an in-arborescence, we have

$$\sigma(\Gamma(\psi)) \leq |\Gamma(\psi)| + \mathrm{OPT}_N(G, |\mathrm{dom}(\psi)|).$$



Weak Concaveness of $\mathbf{OPT}_N(G,k)$

▶ **Lemma 3.9.** *For any fixed influence graph G, let X be a random variable taking value from $\{0, 1 \dots, n\}$, with mean value $\mathbb{E}[X] = k$. Then we have*

$$\mathbb{E}\left[\mathrm{OPT}_N(G,X)\right] \le \frac{e}{e-1}\mathrm{OPT}_N\left(G, \mathbb{E}[X]\right) = \frac{e}{e-1}\mathrm{OPT}_N(G,k). \tag{17}$$

- Non-adaptive optimal solution $\mathbf{OPT}_N(G,k)$ is weakly concave over $\boldsymbol{k}$
- Prove through greedy solution, which is concave over $\boldsymbol{k}$ by submodularity

Putting Together

- Differential inequality

$$\frac{d}{dt}\mathbb{E}\left[f\left(\Psi(t)\right)\right] \geq f^{+}(x_1,\ldots,x_n) - \frac{e}{e-1}\mathrm{OPT}_N(G,k) - \mathbb{E}\left[f(\Psi(t))\right]$$

- Solution, when $t = 1$:

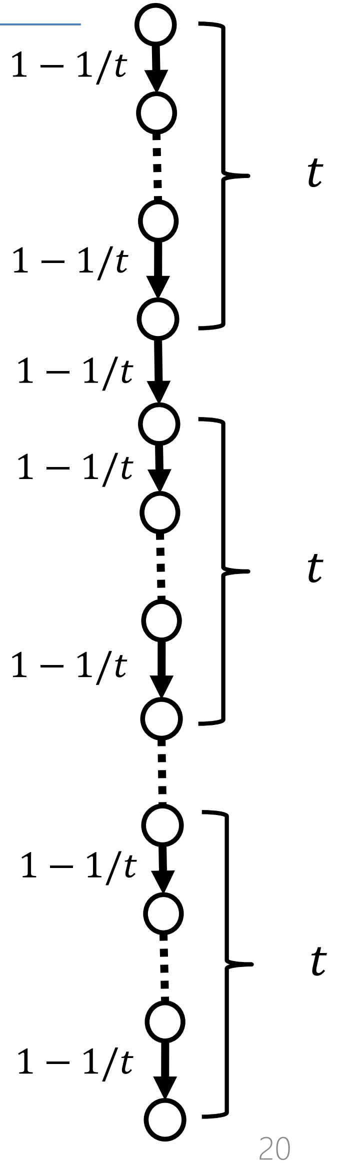
$$\mathbb{E}\left[f(\Psi(1))\right] \geq \left(1-\frac{1}{e}\right)\left[f^{+}(x_1,\ldots,x_n) - \frac{e}{e-1}\mathrm{OPT}_N(G,k)\right]$$

- By pipage rounding, and previous results:

$$\begin{aligned}\mathrm{OPT}_N(G,k) &= \sup_{x_1+\cdots+x_n=k} F(x_1,\ldots,x_n)\\ &\geq \sup_{x_1+\cdots+x_n=k} \mathbb{E}\left[f(\Psi(1))\right]\\ &\geq \sup_{x_1+\cdots+x_n=k} \left(1-\frac{1}{e}\right)\left[f^{+}(x_1,\ldots,x_n) - \frac{e}{e-1}\mathrm{OPT}_N(G,k)\right]\\ &\geq \left(1-\frac{1}{e}\right)\mathrm{OPT}_A(G,k) - \mathrm{OPT}_N(G,k).\end{aligned}$$

Other Results

- Out-arborescences, gap ≤ 2:
  - multilinear extension but no Poisson process
  - direct linking multilinear extension F with optimal adaptive solution f^+
    - $2F(x_1, \ldots, x_n) \geq f^+(x_1, \ldots, x_n)$
    - by observation: each node is only influenced by its closest ancestor seed
- In- (out-) arborescences, gap $\geq \frac{e}{e-1}$
  - directed line of length kt, edge probability $1 - 1/t$
    - best non-adaptive solution: one seed every t nodes
    - best adaptive solution: select next node not activated as a seed
- One-directional bipartite graph
  - gap $\leq \frac{e}{e-1}$: direct showing $F(x_1, \ldots, x_n) \geq \left(1 - \frac{1}{e}\right) f^+(x_1, \ldots, x_n)$
  - gap $\geq \frac{e}{e-1}$: [PC19], myopic feedback = full-adoption feedback here



Many Open Problems

- Adaptivity gap:

	IC model	LT model	Triggering model (more general)
Full-adoption	result on general graphs? tighter result for special graphs?	?	unbounded
Myopic	$\left[\frac{e}{e-1}, 4\right]$, tight result?	?	?

- Greedy adaptivity gap:

	IC model	LT model	Triggering model (more general)
Full-adoption	upper bound?	upper bound?	lower: $1 - 1/e$ upper: unbounded
Myopic	upper bound $\leq \frac{4e}{e-1}$ tight upper bound?	upper bound?	upper bound?

- Better adaptive algorithms than greedy?

Thanks!