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

Wei Chen
Microsoft Research, Beijing, China
weic@microsoft.com

Binghui Peng[1]
Columbia University, New York, United States
bp2601@columbia.edu

Abstract

In this paper, we study the *adaptivity gap* of the influence maximization problem under the independent cascade model when *full-adoption* feedback is available. Our main results are to derive upper bounds on several families of well-studied influence graphs, including in-arborescences, out-arborescences and bipartite graphs. Especially, we prove that the adaptivity gap for the in-arborescences is between $[\frac{e}{e-1}, \frac{2e}{e-1}]$, and for the out-arborescences the gap is between $[\frac{e}{e-1}, 2]$. These are the first constant upper bounds in the full-adoption feedback model. Our analysis provides several novel ideas to tackle the correlated feedback appearing in adaptive stochastic optimization, which may be of independent interest.

2012 ACM Subject Classification Theory of computation → Social networks

Keywords and phrases Adaptive influence maximization, adaptivity gap, full-adoption feedback

Digital Object Identifier 10.4230/LIPIcs.ISAAC.2019.24

1 Introduction

Following the celebrated work of Kempe et al. [19], the *influence maximization* (IM) problem has been extensively studied over past decades. Influence maximization is the problem of selecting at most k *seed nodes* that maximize the influence spread on a given social network and diffusion model. Influence maximization has many real world applications such as viral markets, rumor controls, etc. In the past years, the IM problem has been studied in different context such as outbreak detection [20], topic-aware influence propagation [5], competitive and complementary influence maximization [23] etc., and both theoretically and practically efficient algorithms have been developed [13, 12, 7, 32, 31]. See the recent survey [11, 21] for more detailed reference.

Meanwhile, stimulated by the real life demand, researchers in recent years begin to consider this classical problem in the adaptive setting. In the *adaptive influence maximization* problem, instead of selecting the full seed set all at once, we are allowed to select seeds one after another, making future decisions based on the propagation feedback gathered from the previous seeds selected. Two feedback models are typically considered [15]: *myopic feedback*, where only the one-step propagation from the selected seed to its immediate out-neighbors are included in the feedback, and *full-adoption feedback*, where the entire cascade from the seed is included in the feedback. This adaptive decision process can potentially bring huge benefits but it also brings technical challenges, since adaptive policies are usually hard to

[1] Work is mostly done while Binghui was at Tsinghua University and visiting Microsoft Research Asia as an intern.

30th International Symposium on Algorithms and Computation (ISAAC 2019).
Editors: Pinyan Lu and Guochuan Zhang; Article No. 24; pp. 24:1–24:19
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

design and analyze, and the adaptive decision process can be slow in practice. Thus, a crucial task in this area is to decide whether and how much adaptive policy is really superior over the non-adaptive policy. The *adaptivity gap* quantifies to what extent adaptive policy outperforms a non-adaptive one and it is defined as the supremum ratio between the influence spread of the optimal adaptive policy and that of the optimal non-adaptive policy. The above question has been answered recently when only myopic feedback are available [25, 14] and constant upper bounds on the adaptivity gap have been derived.

In this paper, we consider the influence maximization problem in the independent cascade (IC) model with *full-adoption feedback*. Even though the full-adoption feedback under the IC model satisfies an important property called *adaptive submodularity*, the analysis of its adaptivity gap is more challenging because the feedback obtained from different seed nodes are no longer independent – feedback from one seed contains multiple-step cascade results, and thus it may overlap with the feedback from other seed nodes. Therefore, results from existing studies on the adaptivity gap of general classes of stochastic adaptive optimization problems [3, 16, 17, 8] cannot be applied, since they all rely on the independent feedback assumption.

In this study, we are able to derive nontrivial constant upper bounds on several families of graphs, including in-arborescences, out-arborescence and bipartite graphs, which have been the targets of many studies in influence maximization (see Section 1.1 for more details). Formally, we have (i) when the influence graph is an in-arborescence, the adaptivity gap is between $[\frac{e}{e-1}, \frac{2e}{e-1}]$ (Section 3 and Section 6), (ii) when the influence graph is an out-arborescence, the adaptivity gap is between $[\frac{e}{e-1}, 2]$ (Section 4 and Section 6) and (iii) the adaptivity gap for the bipartite influence graph is $\frac{e}{e-1}$ (Section 5). Our upper bounds on arborescences are the first constant upper bounds in the full-adoption feedback model with multiple steps of cascades and thus dependent feedback, and our upper bound on bipartite graphs improves the results in [14, 18].

The main technical contributions in this paper are on the adaptivity gaps for arborescences, in which the feedback information can be correlated and all previous methods failed. We adopt two different proof strategies to overcome the difficulty of dependent feedback. For in-arborescences, we follow the framework in [3] and construct a Poisson process to relate the influence spread of the optimal adaptive policy and the *multilinear extension*. The analyses are non-trivial due to the correlated feedback. We need to delicately decompose the marginal gain of the Poisson process and give upper bounds on each terms. The key observation we have for in-arborescences is that the *boundary* of the active nodes shrinks during the diffusion process. For out-arborescences, we again relate the influence spread of the multilinear extension to the optimal policy, but using a completely different proof strategy. The key observation for out-arborescences is that the predecessors of each node form a directed line thus proving a stronger results on this line is sufficient. We derive a family of constraints on the optimal adaptive policy and telescope the marginal gains of the multilinear extension, combining these two could yield our results.

1.1 Related Work

A number of studies [6, 36, 35, 24, 22] have focused on the influence maximization problem on arborescences and interesting theoretical results have been found with this special structural assumption. Bharathi et al. [6] initiate the study on arborescences and derive a polynomial-time approximation scheme (PTAS) for bidirected trees. Wang et al. [36] convert diffusion in the IC model in the local region into the diffusion in in-arborescences to design efficient heuristic algorithms for influence maximization. For in-arborescences, Wang et al. [35] give a polynomial time algorithm in the linear threshold (LT) model and Lu et al. [24] prove NP hardness results under the independent cascade model.

The influence maximization problem on one-directional bipartite graphs has been studied by [2, 29, 18], and it has applications in advertisement selections. Especially, Hatano et al. [18] consider the problem in the adaptive setting and derive adaptive algorithms with theoretical guarantees.

Initiated by the pioneering work of [15], a recent line of work [33, 37, 26, 30, 14, 25] focus on the adaptive influence maximization problem and develop both theoretical results and practical methods. Golovn and Krause [15] propose the novel concept of adaptive submodularity and apply it to the adaptive influence maximization problem. They prove that with full-adoption feedback in the IC model, the influence spread function satisfies the adaptive submodularity, thus a simple adaptive greedy algorithm could achieve the $(1-1/e)$ approximation ratio. Fujii et al. [14] generalize the notion and propose weakly adaptive submodularity. They consider the adaptivity gap on both LT and IC models, when the influence graph is one-directional bipartite, which means that the diffusion is always done in one step and there is no difference between myopic feedback and full-adoption feedback. While they prove a tight upper bound of 2 for the LT model, their bound for IC model depends on the structure of the graph and can be far worse than $1-1/e$. In contrast, in this paper we provide the tight bound of $1-1/e$ with a simple analysis in this case. Recently, Peng and Chen [25] consider the myopic feedback model and prove an upper bound of 4 for the adaptivity gap in the IC model. Singer and his collaborators have done a series of studies on adaptive seeding and studied the adaptivity gap in their setting [27, 28, 4], but their model is a two-step adaptive model with the first step purely for referring to the seed candidates, and thus their model is very different from adaptive influence maximization of this paper and other related work above.

From the theoretical side, there are two lines of works [3, 1, 16, 17, 8] on the adaptivity gap that are most relevant to ours. Asadpour et al. [3] study the stochastic submodular optimization problem. They use multilinear extensions to transform an adaptive strategy to a non-adaptive strategy and give a tight upper bound of $\frac{e}{e-1}$. Their method inspires our work but theirs cannot be directly applied to our settings, since the feedback information is not independent in the full-adoption feedback model. We defer further discussion about this key difference to Section 3. Another line of work [16, 17, 8] focuses on the stochastic probing problem. They transform any adaptive policy to a random walk non-adaptive policy and Bradac et al. [8] finally prove a tight upper bound of 2 for prefix constraints.

2 Preliminaries

In this paper, we focus on the well known *independent cascade* (IC) model as the diffusion model. In the IC model, the social network is described by a directed influence graph $G = (V, E, p)$ ($|V| = n$), and there is a probability p_{uv} associated with each edge $(u, v) \in E$. The *live-edge* graph $L = (V, L(E))$ is a random subgraph of the influence graph G, where each edge $(u, v) \in E$ appears in $L(E)$ independently with probability p_{uv}. If the edge appears in $L(E)$, we say it is *live*, otherwise we say it is *blocked*. We use $\mathcal{L}$ to denote all possible live-edge graphs and $\mathcal{P}$ to denote the probability distribution over $\mathcal{L}$. The diffusion process can be described by the following discrete time process. At time $t = 0$, a seed set $S \subseteq V$ is activated and a live-edge graph L is sampled from the probability distribution $\mathcal{P}$ (i.e., each edge (u, v) is live with probability p_{uv}). At time $t = 1, 2, \ldots$, a node $u \in V$ is active if (i) u is active at time $t-1$ or (ii) one of u's in-neighbor in the live-edge graph L is active at time $t-1$. The diffusion process ends at a step when there are no new nodes activated at this step. We use $\Gamma(S, L)$ to denote the set of active nodes at the end of diffusion, or equivalently, the set of nodes reachable from set S under live-edge graph L. We define the *influence reach function* $f : \{0,1\}^V \times \mathcal{L} \to \mathbb{R}^+$ as $f(S,L) := |\Gamma(S,L)|$. Then the influence spread of a set S, denoted as $\sigma(S)$, is defined as the expected number of active nodes at the end of the diffusion process, i.e., $\sigma(S) := \mathbb{E}_{L\sim\mathcal{P}}[f(S,L)]$.

We formally state the (non-adaptive) influence maximization problem as follows.

▶ **Definition 2.1** (Non-adaptive influence maximization). *The non-adaptive influence maximization (IM) problem is the problem of given an influence graph $G = (V,E,p)$ and a budget k, finding a seed set $S^\star$ of size at most k that maximizes the influence spread, i.e., finding $S^\star = \mathrm{argmax}_{S\subseteq V,|S|\le k}\sigma(S)$.*

In the adaptive setting, instead of committing the entire seed set all at once, we are allowed to select the seed node one by one. After we select a seed, we can get some feedback about the diffusion state from the node. Formally, a *realization* ϕ is a function $\phi : V \to O$, mapping a node u to its state, i.e., the feedback we obtain when we select the node u as a seed. The realization ϕ determines the status of all edges in the influence graph and it is one-to-one correspondence to a live-edge graph. Henceforth, in the rest of the paper, we would use ϕ to refer to both the realization and the live-edge graph interchangeably. The feedback information depends on the feedback model and in this paper, we consider the *full-adoption* feedback model. In the full-adoption feedback model, after we select a node u, the feedback $\phi(u)$ of u we see is the status of all out-going edges of nodes v that are reachable from u in the live-edge graph corresponding to ϕ. In other words, we get to see the full cascade starting from the node u. At each step of the adaptive seeding process, our observation so far is represented by a *partial realization* $\psi \subseteq V \times O$, which is a collection of nodes and states, $(u,\phi(u))$, we have observed so far. We use $\mathrm{dom}(\psi)$ to denote the set $\{u : u \in V, \exists(u,o) \in \psi\}$, that is, all nodes we have selected so far. For two partial realizations ψ and ψ', we say ψ is a sub-realization of ψ' if $\psi \subseteq \psi'$ when treating ψ and ψ' as subsets of $V \times O$.

An adaptive policy π is a mapping from partial realizations to nodes. Given a partial realization ψ, we use $\pi(\psi)$ to represent the next seed selected by π. After selecting node $\pi(\psi)$, our observation (partial realization) grows as $\psi' = \psi \cup (\pi(\psi), \phi(\pi(\psi)))$ and the policy π would pick the next node based on the new partial realization ψ'. Given a realization ϕ, we use $V(\pi,\phi)$ to denote the seed set selected by the policy π. The *adaptive influence spread* of the policy π is defined as the expected number of active nodes under the policy π, i.e., $\sigma(\pi) := \mathbb{E}_{\Phi\sim\mathcal{P}}[f(V(\pi,\Phi),\Phi)]$. We define $\Pi(k)$ as the set of policies π, such that for any possible realization ϕ, $|V(\pi,\phi)| \le k$. For convenience, we treat the adaptive policy π as deterministic, but our results apply to randomized policies too. The adaptive influence maximization problem is formally stated as follow.

▶ **Definition 2.2** (Adaptive influence maximization). *The adaptive influence maximization (AIM) problem is the problem of given an influence graph $G = (V,E,p)$ and a budget k, finding a feasible policy $\pi \in \Pi(k)$ that maximizes the adaptive influence spread, i.e., finding $\pi^\star = \mathrm{argmax}_{\pi\in\Pi(k)}\sigma(\pi)$.*

In this paper, we study the *adaptivity gap* of the influence maximization problem under the full-adoption feedback model. The adaptivity gap measures the supremacy of the optimal adaptive policy over the optimal non-adaptive policy. We use $\mathrm{OPT}_N(G,k)$ (resp. $\mathrm{OPT}_A(G,k)$) to denote the influence spread of the optimal non-adaptive (resp. adaptive) policy for the IM problem on the influence graph G with a budget k.

▶ **Definition 2.3** (Adaptivity gap). *The adaptivity gap for the IM problem in the IC model with full-adoption feedback is defined as*

$$\sup_{G,k} \frac{\mathrm{OPT}_A(G,k)}{\mathrm{OPT}_N(G,k)}. \tag{1}$$

We prove constant upper bounds on the adaptivity gap for several classes of graphs, including the *in-arborescence* and the *out-arborescence*.

▶ **Definition 2.4** (In-arborescence). *We say an influence graph $G = (V, E, p)$ is an in-arborescence when the underlying graph is a directed tree with a root u, such that for any node $v \in V$, the unique path between nodes u and v is directed from v to u. In other words, the information propagates from leaves to the root.*

▶ **Definition 2.5** (Out-arborescence). *We say an influence graph $G = (V, E, p)$ is an out-arborescence when the underlying graph is a directed tree with a root u, such that for any node $v \in V$, the unique path between nodes u and v is directed from u to v. In other words, the information propagates from the root to leaves.*

A set function $f : V \to \mathbb{R}^+$ is said to be submodular if for any set $A \subseteq B \subseteq V$ and any element $u \in V \backslash B$, $\Delta_f(u|A) := f(A \cup \{u\}) - f(A) \geq f(B \cup \{u\}) - f(B) = \Delta_f(u|B)$. We call $\Delta_f(u|A)$ the marginal gain for adding element u to the set A. Moreover, the function f is said to be monotone if $f(B) \geq f(A)$. Under the IC model, the influence spread function $\sigma(\cdot)$ is proved to be *submodular* and *monotone* [19], thus given value oracles for $\sigma(\cdot)$, the greedy algorithm is $1 - 1/e$ approximate to the optimal non-adaptive solution.

In the adaptive submodular optimization scenario, a similar notion to the submodularity is called the *adaptive submodularity*. For a function f and a partial realization ψ, the conditional marginal gain for an element $u \notin \mathrm{dom}(\psi)$ is defined as $\Delta_{f,\mathcal{P}}(u|\psi) := \mathbb{E}_{\Phi\sim\mathcal{P}}\left[f\left(\mathrm{dom}(\psi) \cup \{u\}\right) - f\left(\mathrm{dom}(\psi)\right) | \Phi \sim \psi\right]$, where we write $\Phi \sim \psi$ to say that the realization Φ is *consistent* with the partial realization ψ, i.e., $\Phi(u) = \psi(u)$ for every $u \in \mathrm{dom}(\psi)$. A function f is said to be *adaptive submodular* with respect to $\mathcal{P}$ if for any partial realizations $\psi \subseteq \psi'$ and any element $u \in V \backslash \mathrm{dom}(\psi')$, $\Delta_{f,\mathcal{P}}(u|\psi) \geq \Delta_{f,\mathcal{P}}(u|\psi')$. Moreover, the function f is *adaptive monotone* with respect to $\mathcal{P}$ if $\Delta_{f,\mathcal{P}}(u|\psi) \geq 0$ for any feasible partial realization ψ with $\Pr_{\Phi\sim\mathcal{P}}\{\Phi \sim \psi\} > 0$. Golovin and Krause [15] show the following important result, which will be used in our analysis.

▶ **Proposition 2.6** ([15]). *The influence reach function f is adaptive submodular and adaptive monotone with respect to the live-edge graph distribution $\mathcal{P}$ under the independent cascade model with full-adoption feedback.*

The following two definitions are very important to our later analysis.

▶ **Definition 2.7** (Multilinear extension). *The multilinear extension $F : [0,1]^V \to \mathbb{R}^+$ of the influence spread function σ is defined as*

$$F(x_1, \ldots, x_n) = \sum_{S\subseteq V} \left[\left(\prod_{i\in S} x_i \prod_{i\notin S} (1 - x_i) \right) \sigma(S) \right]. \tag{2}$$

We remark that the multilinear extension $F(\cdot)$ is monotone and DR-submodular [19], when the original function $\sigma(\cdot)$ is monotone and submodular. A vector function f is DR-submodular if for any two vectors $(x_1, \ldots, x_n) \leq (y_1, \ldots, y_n)$ (coordinate-wise), for any $\delta > 0$, any $j \in [n]$, $f(x_1, \ldots, x_j + \delta, \ldots, x_n) - f(x_1, \ldots, x_n) \geq f(y_1, \ldots, y_j + \delta, \ldots, y_n) - f(y_1, \ldots, y_n)$. For any configuration $(x_1, \ldots, x_n)$, we use $f^+(x_1, \ldots, x_n)$ to denote the adaptive influence spread of the optimal adaptive strategy consistent with this configuration. Formally,

▶ **Definition 2.8** (Adaptive influence spread function based on an optimal adaptive policy)**.** *We define* $f^+ : [0,1]^{|V|} \to \mathbb{R}^+$ *as:*

$$f^+(x_1, \ldots, x_n) = \sup_{\pi} \left\{ \sigma(\pi) : \Pr_{\Phi \sim \mathcal{P}} [i \in V(\pi, \Phi)] = x_i, \forall i \in [n] \right\}. \tag{3}$$

3 Adaptivity Gap for In-arborescence

In this section, we give an upper bound on the adaptivity gap when the influence graph is an in-arborescence, as stated in the following theorem.

▶ **Theorem 3.1.** *When the underlying influence graph is an in-arborescence, the adaptivity gap for the IM problem in the IC model with full-adoption feedback is at most* $\frac{2e}{e-1}$*.*

Our approach follows the general framework of [3], i.e., we use the multilinear extension to transform an adaptive policy to a non-adaptive policy, and construct a Poisson process to connect the influence spread of the non-adaptive policy to the adaptive policy. Once we have done this, combining with the rounding procedure in [9, 10], we can derive an upper bound on the adaptive policy. However, remembering that the main difficulty of our problem comes from the correlation of the feedback, directly applying the analyses in [3] does not work. Our methods have several key differences comparing to [3]. To be more specific, we can no longer directly relate the dynamic marginal gain of the Poisson process to the influence spread of the adaptive policy. Instead, we need to delicately decompose the marginal gain into two parts (see Lemma 3.3 and Lemma 3.9). The first part can be related to the optimal adaptive strategy by telescoping the summation and applying a coupling argument, while the second part can be related to a (randomized) non-adaptive policy. However, this non-adaptive policy is not guaranteed to be bounded by $\mathrm{OPT}_N(G, k)$ (the optimal non-adaptive policy of budget k), because the size of the seed set is random and can potentially be very large. We utilize the *"weak concavity"* of the optimal solution to show that it is enough to consider the expected size of the (random) seed set, and then we give an upper bound on this expected size for an in-arborescence. This upper bound relies on a crucial property of the in-arborescence, i.e., the *boundary* (see Definition 3.7) size always shrinks during the diffusion process of the information (see Lemma 3.8). Putting things together, we get a differential inequation that relates the dynamic marginal gain of the Poision process to both an optimal adaptive policy and the optimal non-adaptive policy. Solving the differential inequation yields a lower bound on the multilinear extension and it gives an upper bound on the adaptivity gap. We remark that one noticeable difference of our bound on the adaptivity gap is that it does not hold for the matroid constraint (which holds in [3]), even though the multilinear extension was originally designated to handle matroid constraints.

Following the work [3, 34], for any configuration $(x_1, \ldots, x_n)$, we consider the following Poisson process, which indirectly relates the multilinear extension $F(x_1, \ldots, x_n)$ to the optimal adaptive solution $f^+(x_1, \ldots, x_n)$.

Poisson Process. There are n independent Poisson clocks $C_1, \ldots, C_n$, the clock C_i ($i \in [n]$) sends signals with rate x_i. Whenever a clock C_i sends out a signal, we select node i as a seed and gather feedback $\phi(i)$ according to the underlying realization ϕ. We use $\Psi(t)$ to denote the partial realization at time t and we start with $\Psi(0)$ as $\emptyset$. We note that $\Psi(t)$ is a random partial realization that contains (a) random time points $t_i \le t$ at which clock C_i sends a signal, for $i \in [n]$; and (b) for each $t_i \le t$, the feedback $\phi(i)$ of seed node i based on the live-edge graph corresponding to realization ϕ. The Poisson process ends at $t = 1$. Note that the Poisson process is parameterized by $(x_1, x_2, \ldots, x_n)$, but we ignore these parameters in the notation $\Psi(t)$.

Given a partial realization ψ, we slightly abuse the notation and use $\Gamma(\psi)$ to denote all the nodes that seed nodes $\mathrm{dom}(\psi)$ activate in the diffusion. Since we are considering the full-adoption feedback, all nodes activated from $\mathrm{dom}(\psi)$ are included in the partial realization ψ, and thus $\Gamma(\psi)$ is a fixed node set when ψ is fixed. We define $f(\psi) = |\Gamma(\psi)|$. The following lemma states that at the end of the Poisson process, i.e., when $t = 1$, the expected influence spread $\mathbb{E}[f(\Psi(1))]$ is no greater than the influence spread of $F(x_1, \ldots, x_n)$, so it links the Poisson process to the non-adaptive optimal solution.

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

Proof. Notice that in the Poisson process, the selection of seeds are actually independent of the realization of the influence graph. Moreover, seed nodes are selected independently. At the end of the process (when $t = 1$), the node i is selected as a seed with probability $1 - e^{-x_i}$. Thus we have

$$\begin{aligned}\mathbb{E}\left[f\left(\Psi(1)\right)\right] &= \sum_{S \subseteq V}\left[\left(\prod_{i\in S}(1 - e^{-x_i})\right)\prod_{i\notin S}\left(e^{-x_i}\right)\sigma(S)\right]\\ &= F(1 - e^{-x_1}, \ldots, 1 - e^{-x_n}) \le F(x_1, \ldots, x_n).\end{aligned} \tag{4}$$

The inequality holds due to the monotonicity of the multilinear extension $F(\cdot)$ and the fact that $1 - e^{-x} \le x$. ◀

In the following lemma, we give a lower bound on the dynamic marginal gain of the influence spread in the Poisson process, which links the process to the optimal adaptive solution f^+.

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

$$\mathbb{E}\left[\frac{df\left(\Psi(t)\right)}{dt} \mid \Psi(t) = \psi\right] \ge f^+\left(x_1, \ldots, x_n\right) - \sigma\left(\Gamma(\psi)\right). \tag{5}$$

Before proving the above lemma, we need to first establish a useful and generic result concerning the marginal influence spread given a partial realization (Lemma 3.6), which is proved in turn from two basic and intuitive results (Lemmas 3.4 and 3.5). The result of Lemma 3.6 also provides an alternative and perhaps more intuitive proof of the adaptive submodularity of the independent cascade model under the full-adoption feedback, and thus it may be of independent interest.

Given an influence graph $G = (V, E, p)$, we use $\sigma^G(S)$ to denote the expected influence spread of the seed set S on influence graph G. For a node set $A \subseteq V$, we use $G \setminus A$ to denote a reduced influence graph from G in which all nodes in A and its incident edges are removed, and the remaining edges has the same influence probabilities as in G.

▶ **Lemma 3.4.** *Under the independent cascade model with full-adoption feedback, for any influence graph $G = (V, E, p)$, partial realization ψ, and any node $i \notin \Gamma(\psi)$, we have*

$$\Delta_f(i|\psi) = \sigma^{G\setminus\Gamma(\psi)}(\{i\}), \tag{6}$$

i.e., the marginal influence spread of node i given a partial realization equals to its influence on the reduced graph where nodes in $\Gamma(\psi)$ are removed.

Proof. We use E_1 to denote all edges that are included in ψ, i.e., all edges that have already revealed their status. These are the outgoing edges from nodes activated in ψ, i.e. outgoing edges from $\Gamma(\psi)$. We use E_2 to denote edges in graph $G\backslash\Gamma(\psi)$ and set $E_3 = E\backslash(E_1 \cup E_2)$. Then E_2 are the edges in G that are not incident to nodes in $\Gamma(\psi)$ and E_3 are incoming edges of nodes in $\Gamma(\psi)$ that come from nodes in $V \setminus \Gamma(\psi)$. We use Φ_i to denote a random realization of edges in E_i $(i = 1, 2, 3)$, and we know that they are mutually independent in the IC model. Let $S = \mathrm{dom}(\psi)$. Now, we have

$$\begin{aligned}\Delta_f(i|\psi) &= \mathop{\mathbb{E}}_{\Phi}[f\left(S \cup \{i\}, \Phi\right) - f(S, \Phi)|\Phi \sim \psi] \\ &= \mathop{\mathbb{E}}_{\Phi_1,\Phi_2,\Phi_3}[f\left(S \cup \{i\}, (\Phi_1, \Phi_2, \Phi_3)\right) - f\left(S, (\Phi_1, \Phi_2, \Phi_3)\right)|(\Phi_1, \Phi_2, \Phi_3) \sim \psi] \\ &= \mathop{\mathbb{E}}_{\Phi_2}\left[\mathop{\mathbb{E}}_{\Phi_3}\left[f(S \cup \{i\}, (\psi, \Phi_2, \Phi_3)) - f(S, (\psi, \Phi_2, \Phi_3))\right]\right]. \end{aligned} \tag{7}$$

The second equality comes from the fact that Φ_1, Φ_2 and Φ_3 are independent and the third equality comes from the fact that $\Phi_1 = \psi$, due to the full-adoption feedback model. We also have

$$\sigma^{G\backslash\Gamma(\psi)}(\{i\}) = \mathop{\mathbb{E}}_{\Phi_2}\left[f(\{i\}, \Phi_2)\right]. \tag{8}$$

To prove that Eq.(7) is the same as Eq.(8), it is suffice to show that for any fixed ϕ_2, ϕ_3 that are realizations of edges in E_2 and E_3 respectively, we have

$$\Gamma(\{i\}, \phi_2) = \Gamma(S \cup \{i\}, (\psi, \phi_2, \phi_3))\backslash\Gamma(S, (\psi, \phi_2, \phi_3)). \tag{9}$$

Recall that we equate a realization with a live-edge graph, and thus in the above notation, ϕ_2 is considered as a live-edge graph in graph $G \setminus \Gamma(\psi)$ and $\Gamma(\{i\}, \phi_2)$ means the set of nodes reachable from i in the live-edge graph of ϕ_2. We first prove that $\Gamma(\{i\}, \phi_2) \subseteq \Gamma(S \cup \{i\}, (\psi, \phi_2, \phi_3))\backslash\Gamma(S, (\psi, \phi_2, \phi_3))$. For any node $u \in \Gamma(\{i\}, \phi_2)$, we know that it can be reached from node i via a path in graph $G\backslash\Gamma(\psi)$, the path only contains edges in ϕ_2, which also means u is in the graph $G \setminus \Gamma(\psi)$. Due to the existence of the path, we know that $u \in \Gamma(S \cup \{i\}, (\psi, \phi_2, \phi_3))$. We claim that $u \notin \Gamma(S, (\psi, \phi_2, \phi_3))$. This is because we are dealing with the full-adoption feedback model, and thus $\Gamma(S, (\psi, \phi_2, \phi_3))$ is the nodes reachable from $S = \mathrm{dom}(\psi)$, which means $\Gamma(S, (\psi, \phi_2, \phi_3)) = \Gamma(\psi)$, and thus $u \in \Gamma(S, (\psi, \phi_2, \phi_3))$ would conflict with the fact u is in $G \setminus \Gamma(\psi)$. Therefore, $u \in \Gamma(S \cup \{i\}, (\psi, \phi_2, \phi_3))\backslash\Gamma(S, (\psi, \phi_2, \phi_3))$.

On the other side, for any node $u \in \Gamma(S \cup \{i\}, (\psi, \phi_2, \phi_3))\backslash\Gamma(S, (\psi, \phi_2, \phi_3))$, the node u can be reached from the node i via a path in the live-edge graph (ψ, ϕ_2, ϕ_3), but u cannot be reached in the live-edge graph (ψ, ϕ_2, ϕ_3) from any node v that can be activated by $\mathrm{dom}(\psi)$ in ψ, because otherwise $u \in \Gamma(S, (\psi, \phi_2, \phi_3))$. This implies that the path from i to u must be in the live-edge graph ϕ_2 in graph $G \setminus \Gamma(\psi)$. That is, $u \in \Gamma(\{i\}, \phi_2)$. Therefore, Eq.(9) holds, and this completes the proof. ◀

▶ **Lemma 3.5.** *Given an influence graph $G = (V, E, p)$. For any two node sets $A, B \subseteq V$ with $A \subseteq B$, for any $u \in V \setminus B$, we have $\sigma^{G\backslash B}(\{i\}) \le \sigma^{G\backslash A}(\{i\})$.*

Proof (Sketch). The proof is intuitively straightforward, since removing nodes could only hurt the influence spread. A more rigorous proof could be carried out by arguing for any random realization Φ_1 for edges in $G \setminus B$ and any random realization Φ_2 for edges in $G \setminus A$ but not in $G \setminus B$, (a) Φ_1 and Φ_2 are independent in the IC model; (b) $\sigma^{G\backslash B}(\{i\}) = \mathbb{E}_{\Phi_1}[|\Gamma(\{i\}, \Phi_1)|]$; (c) $\sigma^{G\backslash A}(\{i\}) = \mathbb{E}_{\Phi_1,\Phi_2}[|\Gamma(\{i\}, \Phi_1 \cup \Phi_2)|]$; and (d) $\Gamma(\{i\}, \Phi_1) \subseteq \Gamma(\{i\}, \Phi_1 \cup \Phi_2)$. ◀

▶ **Lemma 3.6.** *Under the independent cascade model with full-adoption feedback, for any two partial realizations ψ_1 and ψ_2, if $\Gamma(\psi_1) \subseteq \Gamma(\psi_2)$, then for any node $i \notin \Gamma(\psi_1)$, $\Delta_f(i|\psi_2) \leq \Delta(i|\psi_1)$.*

Proof. First, note that if $i \in \Gamma(\psi_2) \setminus \Gamma(\psi_1)$, then $\Delta(i|\psi_2) = 0$, so $\Delta_f(i|\psi_2) \leq \Delta(i|\psi_1)$. We now consider that $i \notin \Gamma(\psi_2)$. By Lemma 3.4, $\Delta_f(i|\psi_1) = \sigma^{G\setminus\Gamma(\psi_1)}(\{i\})$, and $\Delta_f(i|\psi_2) = \sigma^{G\setminus\Gamma(\psi_2)}(\{i\})$. Since $\Gamma(\psi_1) \subseteq \Gamma(\psi_2)$, by Lemma 3.5, we have $\sigma^{G\setminus\Gamma(\psi_2)}(\{i\}) \leq \sigma^{G\setminus\Gamma(\psi_1)}(\{i\})$. Thus, the lemma holds. ◀

We are now ready to prove Lemma 3.3.

Proof Lemma 3.3. First, we consider the left-hand side of Eq. (5). For any $t \in [0, 1], i \in [n]$ and small enough amount of time dt, the clock C_i sends out signals with probability $x_i dt$ during the time interval $[t, t + dt]$. Since signals are sent out independently, the probability that more than one clock send out signals simultaneously in time interval $[t, t+dt]$ is of order $O((dt)^2)$, which can be considered negligible comparing to dt. Thus we have

$$\mathbb{E}\left[f\left(\Psi(t+dt)\right) - f\left(\Psi(t)\right)|\Psi(t) = \psi\right] = \sum_{i \notin \mathrm{dom}(\psi)} x_i dt \cdot \Delta_f(i|\psi), \tag{10}$$

Rewriting the above equation, we derive that

$$\mathbb{E}\left[\frac{df\left(\Psi(t)\right)}{dt}|\Psi(t) = \psi\right] = \sum_{i \notin \mathrm{dom}(\psi)} x_i \Delta_f(i|\psi) = \sum_{i \notin \Gamma(\psi)} x_i \Delta_f(i|\psi). \tag{11}$$

The second equality holds because $\Delta_f(i|\psi) = 0$ for any node $i \in \Gamma(\psi)$ in the full-adoption feedback model.

Next, we consider the right-hand side of Eq. (5). We write $\boldsymbol{x} = (x_1, \ldots, x_n)$ and use the indicator vector $\boldsymbol{I}_S \in \{0, 1\}^n$ to denote an n-dimensional 0-1 vector such that the coordinate i is 1 if and only if $i \in S$. By the monotonicity of the function $f^+(\cdot)$, we have

$$f^+\left(x_1, \ldots, x_n\right) \leq f^+\left(\boldsymbol{x} \vee \boldsymbol{I}_{\Gamma(\psi)}\right). \tag{12}$$

Consider the optimal adaptive policy π^+ of $f^+(\boldsymbol{x} \vee \boldsymbol{I}_{\Gamma(\psi)})$ as defined in Definition 2.8. We can assume π^+ selects nodes in $\Gamma(\psi)$ at the beginning since they will eventually appear in the seed set regardless of the realization of the live-edge graph. For $i \notin \Gamma(\psi)$, π^+ would select node i as a seed with probability x_i, according to Definition 2.8. Let Ψ_i denote the partial realization just before π^+ selects node i, given that π^+ first selects nodes in $\Gamma(\psi)$ and sees their feedback. Then we have $\Gamma(\psi) \subseteq \Gamma(\Psi_i)$. Conditioned on Ψ_i, the selection of i provides a marginal gain of $\Delta_f(i|\Psi_i)$ for the influence spread. When we take its expectation over Ψ_i and then multiply it with x_i, we obtain the overall marginal gain of selecting i as a seed in policy π^+. When summing over all $i \notin \Gamma(\psi)$, together with the non-adaptive influence spread of seed nodes in $\Gamma(\psi)$, we thus obtain:

$$f^+\left(\boldsymbol{x} \vee \boldsymbol{I}_{\Gamma(\psi)}\right) = \sum_{i \notin \Gamma(\psi)} x_i \cdot \mathop{\mathbb{E}}_{\Psi_i}[\Delta_f(i|\Psi_i)] + \sigma\left(\Gamma(\psi)\right). \tag{13}$$

Combining Eq. (11), (12), (13), it suffices to prove

$$\Delta_f(i|\Psi_i) \leq \Delta_f(i|\psi) \tag{14}$$

for any $i \notin \Gamma(\psi)$ and any partial realization Ψ_i such that $\Gamma(\psi) \subseteq \Gamma(\Psi_i)$, but this is exactly what is shown in Lemma 3.6. This completes our proof. ◀

A couple of remarks are now in order concerning Lemma 3.3 and its proof. First, Inequality (14) looks very much like the adaptive submodularity condition, and we could directly apply adaptive submodularity should we have $\psi \subseteq \Psi_i$. However, by our construction, we only have $\Gamma(\psi)$ as the initial seed set for Ψ_i, and the partial realization in Ψ_i from the seeds in $\mathrm{dom}(\psi)$ may not be exactly the same as ψ, so we do not have $\psi \subseteq \Psi_i$. Instead, we only have $\Gamma(\psi) \subseteq \Gamma(\Psi_i)$. Therefore, we provide a separate proof (Lemmas 3.4, 3.5, and 3.6) to show that the above condition is enough to prove Inequality (14). Even though our proof does not directly apply the adaptive submodularity result of the IC model with the full-adoption feedback, the essence of proof related to Inequality (14) is still due to the diminishing return behavior of the model in the adaptive setting.

Second, the overall proof structure of Lemma 3.3 roughly follows the proof structure in [3]. However, because in the full-adoption feedback model the feedback from two different nodes may be correlated, we cannot exactly follow the proof structure in [3]. In particular, we have to incorporate $\Gamma(\psi)$ directly into the seeds of an adaptive strategy π^+ to avoid such correlation to interfere with our analysis. This results in the term $\sigma(\Gamma(\psi))$ in Inequality (5) of Lemma 3.3 instead of the term $|\Gamma(\psi)|$ that would be derived should we exactly follow [3], and eventually the term $\sigma(\Gamma(\psi))$ leads to the extra factor of 2 in the adaptivity gap given in Theorem 3.1, which does not appear in [3].

Next, we introduce the concept of the *boundary* of a partial realization. We need to use the property of the boundary in in-arborescences (Lemma 3.8) to properly bound the extra term $\sigma(\Gamma(\psi))$ (Lemma 3.9).

▶ **Definition 3.7** (Boundary of a partial realization)**.** *In the full-adoption feedback model, for any partial realization ψ, we use $\partial(\psi)$ to denote the boundary of the partial realization, i.e., the set of nodes $\partial(\psi) \subseteq \Gamma(\psi)$ with minimum cardinality such that there is no directed edges in the original graph G from $\Gamma(\psi)\backslash\partial(\psi)$ to $V\backslash\Gamma(\psi)$. We remark that when there are more than one such sets, we take an arbitrary one.*

The main property we rely on the structure of an in-arborescence is that the boundary of any partial realization can be bounded by the number of seeds that have been selected. Formally, we have

▶ **Lemma 3.8.** *When the influence graph is an in-arborescence, for any partial realization ψ, we have $|\partial(\psi)| \leq |\mathrm{dom}(\psi)|$.*

Proof. Consider any partial realization ψ and any node $v \in \mathrm{dom}(\psi)$. Take the unique directed path from node v to the root u, let $\bar{v}$ denote the node on the path which is (i) contained in $\Gamma(\psi)$ and (ii) closest to the root u. Then we set $S = \{\bar{v} : v \in \mathrm{dom}(\psi)\}$. Clearly there is no directed edge from $\Gamma(\psi)\backslash S$ to $V\backslash\Gamma(\psi)$ and we have $|\partial(\psi)| \leq |S| \leq |\mathrm{dom}(\psi)|$. ◀

Now, we give an upper bound on the term $\sigma(\Gamma(\psi))$.

▶ **Lemma 3.9.** *For any partial realization ψ*

$$\sigma(\Gamma(\psi)) \leq |\Gamma(\psi)| + \sigma(\partial(\psi)). \tag{15}$$

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

$$\sigma(\Gamma(\psi)) \leq |\Gamma(\psi)| + \mathrm{OPT}_N(G, |\mathrm{dom}(\psi)|). \tag{16}$$

Proof. Fix any realization $\phi \sim \psi$, and then consider any node v in $\Gamma(\Gamma(\psi), \phi)\backslash\Gamma(\partial(\psi), \phi)$. There must exist a directed path P from $\Gamma(\psi)\backslash\partial(\psi)$ to v, and the path P does not contain any nodes in $\partial(\psi)$. According to the definition of the boundary set $\partial(\psi)$, there is no directed path from $\Gamma(\psi)\backslash\partial(\psi)$ to $V\backslash\Gamma(\psi)$, unless it goes through a node in $\partial(\psi)$. Thus we conclude that $v \in \Gamma(\psi)\backslash\partial(\psi)$ and this gives proof for Eq. (15). With Lemma 3.8, we have $\sigma(\partial(\psi)) \le \mathrm{OPT}_N(G, |\mathrm{dom}(\psi)|)$. Therefore, Inequality (16) holds. ◀

For any fixed influence graph G, we can view $\mathrm{OPT}_N(G, k)$ as a function of the budget k, we prove that $\mathrm{OPT}_N(G, k)$ is *"weakly concave"* for k, as stated in the following lemma.

▶ **Lemma 3.10.** *For any fixed influence graph G, let X be a random variable taking value from $\{0, 1 \ldots, n\}$, with mean value $\mathbb{E}[X] = k$. Then we have*

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

Proof. Let $\mathrm{Greedy}_N(G, k)$ denote the non-adaptive greedy solution that selects k seed nodes. For $X \in \{0, 1, \ldots, n\}$, the greedy solution is $1 - 1/e$ approximate to the optimal solution, i.e.,

$$\mathrm{OPT}_N(G, X) \le \frac{e}{e-1}\mathrm{Greedy}_N(G, X). \tag{18}$$

We note that the greedy solution $\mathrm{Greedy}_N(G, X)$ is concave in X, due to the submodularity of the influence spread function. Then taking expectation over both sides of Eq. (18), by Jensen's inequality, we have

$$\begin{aligned}\mathbb{E}\left[\mathrm{OPT}_N(G, X)\right] &\le \frac{e}{e-1}\,\mathbb{E}[\mathrm{Greedy}_N(G, X)] \le \frac{e}{e-1}\mathrm{Greedy}_N\left(G, \mathbb{E}[X]\right) \\ &\le \frac{e}{e-1}\mathrm{OPT}_N\left(G, \mathbb{E}[X]\right) = \frac{e}{e-1}\mathrm{OPT}_N(G, k).\end{aligned} \tag{19}$$

This concludes the proof. ◀

Putting things together, we are able to prove Theorem 3.1 as follows.

Proof of Theorem 3.1. When the influence graph is an in-arborescence, for any configuration $(x_1, \ldots, x_n)$ satisfying $\sum_i x_i = k$, for any $t \in [0, 1]$ and any fixed partial realization ψ, we have

$$\begin{aligned}&\mathbb{E}\left[\frac{df\left(\Psi(t)\right)}{dt}|\Psi(t) = \psi\right] \\ &\ge f^+\left(x_1, \ldots, x_n\right) - \sigma\left(\Gamma(\psi)\right) && \text{by Lemma 3.3} \\ &\ge f^+(x_1, \ldots, x_n) - |\Gamma(\psi)| - \mathrm{OPT}_N(G, |\mathrm{dom}(\psi)|) && \text{by Lemma 3.9} \\ &= f^+(x_1, \ldots, x_n) - f(\Psi(t)) - \mathrm{OPT}_N(G, |\mathrm{dom}(\Psi(t))|).\end{aligned} \tag{20}$$

Taking expectation over $\Psi(t)$, we have for any $t \in [0, 1]$,

$$\begin{aligned}\frac{d}{dt}\mathbb{E}\left[f\left(\Psi(t)\right)\right] = \mathbb{E}\left[\frac{df\left(\Psi(t)\right)}{dt}\right] &= \mathop{\mathbb{E}}_{\Psi(t)}\mathbb{E}\left[\frac{df\left(\Psi(t)\right)}{dt} \mid \Psi(t)\right] \\ &\ge f^+(x_1, \ldots, x_n) - \mathbb{E}\left[f(\Psi(t))\right] - \mathbb{E}\left[\mathrm{OPT}_N(G, |\mathrm{dom}(\Psi(t))|)\right] \\ &\ge f^+(x_1, \ldots, x_n) - \frac{e}{e-1}\mathrm{OPT}_N(G, k) - \mathbb{E}\left[f(\Psi(t))\right].\end{aligned} \tag{21}$$

The first equality above is by the linearity of expectation. The second equality above is by the law of total expectation. The first inequality is by Eq.(20), and the second inequality holds due to Lemma 3.10 and the fact that

$$\mathbb{E}\left[|\mathrm{dom}(\Psi(t))|\right] = \sum_i (1 - e^{-tx_i}) \le \sum_i tx_i \le k$$

for any $t \le 1$. Solving the above differential inequality in Eq.(21), we obtain

$$\mathbb{E}\left[f(\Psi(t))\right] \ge (1-e^{-t})\left[f^+(x_1,\dots,x_n) - \frac{e}{e-1}\mathrm{OPT}_N(G,k)\right]. \tag{22}$$

In particular, when $t=1$, we have

$$\mathbb{E}\left[f(\Psi(1))\right] \ge \left(1-\frac{1}{e}\right)\left[f^+(x_1,\dots,x_n) - \frac{e}{e-1}\mathrm{OPT}_N(G,k)\right]. \tag{23}$$

Finally, we have

$$\begin{aligned}
\mathrm{OPT}_N(G,k) &= \sup_{x_1+\cdots+x_n=k} F(x_1,\dots,x_n) \\
&\ge \sup_{x_1+\cdots+x_n=k} \mathbb{E}\left[f(\Psi(1))\right] \\
&\ge \sup_{x_1+\cdots+x_n=k} \left(1-\frac{1}{e}\right)\left[f^+(x_1,\dots,x_n) - \frac{e}{e-1}\mathrm{OPT}_N(G,k)\right] \\
&\ge \left(1-\frac{1}{e}\right)\mathrm{OPT}_A(G,k) - \mathrm{OPT}_N(G,k).
\end{aligned} \tag{24}$$

The first equality above comes from the pipage rounding procedure in [9]. The first inequality above is by Lemma 3.2. The second inequality is by Eq. (23). The last equality is by the definition of f^+ (Definition 2.8). Thus we conclude that the adaptivity gap is at most $\frac{2e}{e-1}$ in the case of an in-arborescence. ◀

4 Adaptivity Gap for Out-arborescence

In this section, we give an upper bound on the adaptivity gap when the influence graph is an out-arborescence. Formally,

▶ **Theorem 4.1.** *When the influence graph is an out-arborescence, the adaptivity gap for the IM problem in the IC model with full-adoption feedback is at most 2.*

We first introduce some notations. For any node $u \in V$ and any seed set $S \subseteq V$, we define $\sigma_u(S) := \Pr_\Phi\left[u \in \Gamma(S,\Phi)\right]$, i.e., the probability that the node u is activated when S is the seed set. Similarly, for any adaptive policy π, we define $\sigma_u(\pi) := \Pr_\Phi\left[u \in \Gamma(V(\pi,\Phi),\Phi)\right]$, i.e., the probability that the node u is activated under policy π. We will extend the definition for the multilinear extension (Definition 2.7) and the definition for f^+ (Definition 2.8) correspondingly. To be more specific, we define

$$F_u(x_1,\dots,x_n) = \sum_{S\subseteq[n]}\left[\left(\prod_{i\in S} x_i \prod_{i\notin S}(1-x_i)\right)\sigma_u(S)\right], \tag{25}$$

and

$$f_u^+(x_1,\dots,x_n) = \sup_\pi\left\{\sigma_u(\pi) : \Pr_{\Phi\sim\mathcal{P}}\left[i \in V(\pi,\Phi)\right] = x_i,\ \forall i\in[n]\right\}. \tag{26}$$

In order to show Theorem 4.1, we again transform an adaptive policy to a non-adaptive policy and compare their influence spread. Here, we utilize a new approach based on the structure of out-arborescences and prove a stronger result. That is, we would prove that the probability for any node u to become active in the multilinear extension (policy) is at least half of the optimal adaptive policy (Lemma 4.2). This requires us to give a fine-grained bound on the optimal adaptive policy (Lemma 4.3) and the multilinear extension (Lemma 4.4).

▶ **Lemma 4.2.** *When the influence graph is an out-arborescence, for any node $u \in V$ and any configuration $(x_1, \dots, x_n)$, we have*

$$f_u^+(x_1, \dots, x_n) \le 2F_u(x_1, \dots, x_n). \tag{27}$$

Once we have the result of Lemma 4.2, we can prove Theorem 4.1 as follows.

Proof of Theorem 4.1. Given Lemma 4.2, we have

$$\begin{aligned}\mathrm{OPT}_N(G,k) &= \sup_{x_1+\cdots+x_n=k} F(x_1,\dots,x_n) = \sup_{x_1+\cdots+x_n=k} \sum_u F_u(x_1,\dots,x_n)\\ &\ge \frac{1}{2}\cdot \sup_{x_1+\cdots+x_n=k} \sum_u f_u^+(x_1,\dots,x_n) \ge \frac{1}{2}\cdot \sup_{x_1+\cdots+x_n=k} f^+(x_1,\dots,x_n) \ge \frac{1}{2}\cdot \mathrm{OPT}_A(G,k).\end{aligned}$$ ◀

We now show how to prove Lemma 4.2. We note that node u's predecessors (nodes that can reach node u in the original graph) form a directed line when the influence graph is an out-arborescence. We slightly abuse the notation and use node i to indicate the $(i-1)^{th}$ predecessor of node u, notice that node u itself is represented as node 1. We ignore all other nodes since they do not affect either sides of Eq. (27). We use p_i to denote the probability that the node i can reach node 1. The following lemma gives an upper bound on the optimal adaptive strategy.

▶ **Lemma 4.3.** *For any i, $f_1^+(x_1, \dots, x_n) \le \sum_{j=1}^{i} x_j p_j + p_{i+1}$.*

Proof. Let π be any adaptive strategy satisfying $\Pr_{\Phi\sim\mathcal{P}}[i \in V(\pi,\Phi)] = x_i, i \in [n]$. Let $\mathcal{E}_i$ denote the event that node 1 becomes active right after π chooses node i. Furthermore, we use $\mathcal{E}_{i:}$ to denote the event that node 1 becomes active right after π chooses a node from $\{i, i+1, \dots, n\}$. We notice that events $\mathcal{E}_1, \dots, \mathcal{E}_n$ are disjoint and we have

$$f_1^+(x_1, \dots, x_n) = \sum_{j=1}^{n} \Pr[\mathcal{E}_j] = \sum_{j=1}^{i} \Pr[\mathcal{E}_j] + \Pr[\mathcal{E}_{i+1:}]\ \forall i. \tag{28}$$

It is easy to see that

$$\Pr[\mathcal{E}_{i+1:}] \le p_{i+1}, \tag{29}$$

since the event $\mathcal{E}_{i+1:}$ can only happen when the node $i+1$ can reach node 1. Moreover, let $\mathcal{F}_i$ denote the event that the policy π selects the node i before any nodes in $\{1, \dots, i\}$ are active. Then we have for any $j \in [n]$,

$$\Pr[\mathcal{E}_j] = \Pr_\Phi[\mathcal{E}_j|\mathcal{F}_j] \cdot \Pr_\Phi[\mathcal{F}_j] \le \Pr_\Phi[\mathcal{E}_j|\mathcal{F}_j] \cdot \Pr_\Phi[j \in V(\pi,\Phi)] = p_j \cdot x_j. \tag{30}$$

The first equality holds since the event $\mathcal{E}_j$ can only happen when π selects the node j before any nodes $\{1, \dots, j\}$ are active. Combining Eq. (28) (29) (30), we complete the proof. ◀

We measure the marginal contribution of node i in the next lemma. Intuitively, we can see that $F_1(0, \dots, 0, x_i, \dots, x_n) - F_1(0, \dots, 0, x_{i+1}, \dots, x_n)$ measures the marginal contribution of i in activating node 1, when node i moves from no probability of being selected as a seed to probability of x_i being selected as the seed, under the situation that no nodes in $\{1, \dots, i-1\}$ can be seeds while node $j > i$ has probability x_j of being selected as a seed. Then this marginal contribution only happens when all three conditions hold: (a) possible seeds in $\{i+1, \dots, n\}$ cannot activate i, which has probability $1 - F_i(0, \dots, 0, x_{i+1}, \dots, x_n)$; (b) node i is activated as a seed, which has probability x_i, and (c) node i passes influence and activates node 1, which has probability p_i.

▶ **Lemma 4.4.** *For any i, we have*

$$F_1(0,\ldots,0,x_i,\ldots,x_n)-F_1(0,\ldots,0,x_{i+1},\ldots,x_n)=x_ip_i\left(1-F_i(0,\ldots,0,x_{i+1},\ldots,x_n)\right).$$

Proof. Since the node 1's predecessors form a directed line, for any i we have

$$\begin{aligned}
&F_1(0,\ldots,0,x_i,\ldots,x_n)-F_1(0,\ldots,0,x_{i+1},\ldots,x_n)\\
=&p_i\cdot\left(F_i(0,\ldots,0,x_i,\ldots,x_n)-F_i(0,\ldots,0,x_{i+1},\ldots,x_n)\right)\\
=&p_i\cdot\left(1-(1-x_i)\left(1-F_i(0,\ldots,0,x_{i+1},\ldots,x_n)\right)-F_i(0,\ldots,0,x_{i+1},\ldots,x_n)\right)\\
=&p_ix_i\left(1-F_i(0,\ldots,0,x_{i+1},\ldots,x_n)\right).
\end{aligned}$$

The first two equalities hold because the realization and the selection of nodes are independent. ◀

We are now back to prove Lemma 4.2.

Proof of Lemma 4.2. We use j to denote the minimum index that satisfies $F_j(0,\ldots,x_{j+1},\ldots,x_n)>\frac{1}{2}$. If such index does not exist, we simply set $j=n+1$. Then, we have

$$\begin{aligned}
&F_1(x_1,\ldots,x_n)\\
=&\sum_{i=1}^{j-1}\left(F_1(0,\ldots,0,x_i,\ldots,x_n)-F_1(0,\ldots,0,x_{i+1},\ldots,x_n)\right)+F_1(0,\ldots,0,x_j,\ldots,x_n)\\
=&\sum_{i=1}^{j-1}x_ip_i\left(1-F_i(0,\ldots,0,x_{i+1},\ldots,x_n)\right)+F_1(0,\ldots,0,x_j,\ldots,x_n)\\
=&\sum_{i=1}^{j-1}x_ip_i\left(1-F_i(0,\ldots,0,x_{i+1},\ldots,x_n)\right)+p_j\cdot F_j(0,\ldots,0,x_j,\ldots,x_n)\\
\geq&\frac{1}{2}\sum_{i=1}^{j-1}x_ip_i+\frac{1}{2}p_j\\
\geq&\frac{1}{2}f_1^+(x_1,\ldots,x_n).
\end{aligned}\tag{31}$$

The second equality comes from Lemma 4.4 and the last inequality comes from Lemma 4.3. ◀

5 Adaptivity Gap for One-Directional Bipartite Graphs

In this section, we give an upper bound on the adaptivity gap of the influence maximization problem in the IC model with full-adoption feedback under one-directional bipartite graphs $G(L,R,E,p)$, where L and R are the two set of nodes on the left side and right side respectively, and $E\subseteq L\times R$ are a set of edges only pointing from a left-side node to a right-side node, and p maps each edge to a probability. Our upper bound is tight as it matches the lower bound derived in [25] and it also improves the results developed in [14, 18]. The proof strategy adopted for bipartite graphs is a relatively easy application of our approaches in previous sections, which relates the multilinear extension and the optimal strategy (See Appendix A).

▶ **Theorem 5.1.** *When the influence graph is a one-directional bipartite graph $G(L,R,E,p)$, the adaptivity gap on the influence maximization problem in the IC model with full-adoption feedback is $\frac{e}{e-1}$.*

6 Lower Bounds on the Adaptivity Gap

In this section, we give an example showing that the adaptivity gap is no less than $e/(e-1)$ in the full-adoption feedback model, even when the influence graph is a directed line, a special case of both the in-arborescence and the out-arborescence.

▶ **Theorem 6.1.** *The adaptivity gap for the IM problem in the IC model with full-adoption feedback is at least $e/(e-1)$, even when the influence graph is a directed line.*

Proof. Consider the following influence graph $G(V, E, p)$: the graph is a directed line with vertices $v_{11}, \dots, v_{1t}, v_{21}, \dots, v_{2t}, \dots, v_{k1}, \dots, v_{kt}$, and each edge is live with probability $1-1/t$. Moreover, we have a budget k. Combining the following two claims, we can conclude that the adaptivity gap is greater than or equal to $e/(e-1)$. The proofs of the claims are in Appendix B.

▷ Claim 6.2. For any $\epsilon > 0$, if $k \geq 8/\epsilon^3$, we have $\mathbb{E}[\mathrm{OPT}_A(G, k)] \geq (1-\epsilon)kt$.

▷ Claim 6.3. The optimal non-adaptive strategy is to select $v_{11}, \dots, v_{k1}$ as seeds. Thus, we have $\mathbb{E}[\mathrm{OPT}_N(G, k)] = (1-(1-1/t)^t)kt$. ◀

Discussion on Existing Approaches. There are two types of strategies for proving upper bounds on adaptivity gaps. One common strategy is to convert any adaptive strategy to the multilinear extension as in [27, 3] and our paper. The other is to convert the adaptive strategy to the *random walk non-adaptive strategy* [16, 17, 8]. Here we claim that using the instance constructed in Theorem 6.1, we can show that these two strategies can not yield better-than-2 upper bounds on the adaptivity gap. We defer the detailed discussions to Appendix B.

7 Conclusion

In this paper, we consider several families of influence graphs and give the first constant upper bounds on adaptivity gaps for them under the full-adoption feedback model. Our methods tackle the correlations on the feedback and hopefully can be applied to other adaptive stochastic optimization problems. For future directions, there are still gaps between our lower and upper bounds for both in-arborescences and out-arborescences, so it would be interesting to close the gap. Another open question is to settle down the adaptivity gap for general influence graphs under the IC model with the full-adoption feedback. The adaptive gap for the linear threshold model and other diffusion models are also open.

References

1 Marek Adamczyk, Maxim Sviridenko, and Justin Ward. Submodular stochastic probing on matroids. *Mathematics of Operations Research*, 41(3):1022–1038, 2016.

2 Noga Alon, Iftah Gamzu, and Moshe Tennenholtz. Optimizing budget allocation among channels and influencers. In *WWW*, pages 381–388. ACM, 2012.

3 Arash Asadpour and Hamid Nazerzadeh. Maximizing stochastic monotone submodular functions. *Management Science*, 62(8):2374–2391, 2015.

4 Ashwinkumar Badanidiyuru, Christos Papadimitriou, Aviad Rubinstein, Lior Seeman, and Yaron Singer. Locally adaptive optimization: Adaptive seeding for monotone submodular functions. In *SODA*. SIAM, 2016.

5 Nicola Barbieri, Francesco Bonchi, and Giuseppe Manco. Topic-aware social influence propagation models. In *ICDM'12*, 2012.

6 Shishir Bharathi, David Kempe, and Mahyar Salek. Competitive influence maximization in social networks. In *WINE*, pages 306–311. Springer, 2007.

7 Christian Borgs, Michael Brautbar, Jennifer Chayes, and Brendan Lucier. Maximizing Social Influence in Nearly Optimal Time. In *SODA'14*, pages 946–957. ACM-SIAM, 2014.

8 Domagoj Bradac, Sahil Singla, and Goran Zuzic. (Near) Optimal Adaptivity Gaps for Stochastic Multi-Value Probing. *arXiv preprint*, 2019. `arXiv:1902.01461`.

9 Gruia Calinescu, Chandra Chekuri, Martin Pál, and Jan Vondrák. Maximizing a monotone submodular function subject to a matroid constraint. *SIAM Journal on Computing*, 40(6):1740–1766, 2011.

10 Chandra Chekuri, Jan Vondrak, and Rico Zenklusen. Dependent randomized rounding via exchange properties of combinatorial structures. In *FOCS*, pages 575–584. IEEE, 2010.

11 Wei Chen, Laks VS Lakshmanan, and Carlos Castillo. *Information and Influence Propagation in Social Networks.* Morgan & Claypool Publishers, 2013.

12 Wei Chen, Chi Wang, and Yajun Wang. Scalable influence maximization for prevalent viral marketing in large-scale social networks. In *KDD'10*, 2010.

13 Wei Chen, Yajun Wang, and Siyu Yang. Efficient influence maximization in social networks. In *Proceedings of the 15th ACM SIGKDD*. ACM, 2009.

14 Kaito Fujii and Shinsaku Sakaue. Beyond Adaptive Submodularity: Approximation Guarantees of Greedy Policy with Adaptive Submodularity Ratio. In *ICML*, pages 2042–2051, 2019.

15 Daniel Golovin and Andreas Krause. Adaptive Submodularity:Theory and Applications in Active Learning and Stochastic Optimization. *Journal of Artificial Intelligence Research*, 42:427–486, 2011. arXiv version (`arXiv:1003.3967`) includes discussions on the myopic feedback model.

16 Anupam Gupta, Viswanath Nagarajan, and Sahil Singla. Algorithms and adaptivity gaps for stochastic probing. In *SODA*. SIAM, 2016.

17 Anupam Gupta, Viswanath Nagarajan, and Sahil Singla. Adaptivity gaps for stochastic probing: Submodular and XOS functions. In *SODA*. SIAM, 2017.

18 Daisuke Hatano, Takuro Fukunaga, and Ken-Ichi Kawarabayashi. Adaptive Budget Allocation for Maximizing Influence of Advertisements. In *IJCAI*, pages 3600–3608, 2016.

19 David Kempe, Jon M. Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. *Theory of Computing*, 11(4):105–147, 2015. Conference version appeared in KDD'2003.

20 Jure Leskovec, Andreas Krause, Carlos Guestrin, Christos Faloutsos, Jeanne M Vanbriesen, and Natalie Glance. Cost-effective outbreak detection in networks. In *ACM Knowledge Discovery and Data Mining*, pages 420–429, 2007.

21 Yuchen Li, Ju Fan, Yanhao Wang, and Kian-Lee Tan. Influence Maximization on Social Graphs: A Survey. *IEEE Trans. Knowl. Data Eng.*, 30(10):1852–1872, 2018.

22 Yishi Lin, Wei Chen, and John CS Lui. Boosting information spread: An algorithmic approach. In *ICDE*, pages 883–894. IEEE, 2017.

23 Wei Lu, Wei Chen, and Laks VS Lakshmanan. From competition to complementarity: comparative influence diffusion and maximization. *Proceedings of the VLDB Endowment*, 9(2):60–71, 2015.

24 Zaixin Lu, Zhao Zhang, and Weili Wu. Solution of Bharathi–Kempe–Salek conjecture for influence maximization on arborescence. *Journal of Combinatorial Optimization*, 33(2):803–808, 2017.

25 Binghui Peng and Wei Chen. Adaptive Influence Maximization with Myopic Feedback. In *NeurIPS*, 2019.

26 Guillaume Salha, Nikolaos Tziortziotis, and Michalis Vazirgiannis. Adaptive Submodular Influence Maximization with Myopic Feedback. In *ASONAM*, pages 455–462. IEEE, 2018.

27 Lior Seeman and Yaron Singer. Adaptive seeding in social networks. In *FOCS*, pages 459–468. IEEE, 2013.

28 Yaron Singer. Influence maximization through adaptive seeding. *ACM SIGecom Exchanges*, 15(1):32–59, 2016.

29 Tasuku Soma, Naonori Kakimura, Kazuhiro Inaba, and Ken-ichi Kawarabayashi. Optimal budget allocation: Theoretical guarantee and efficient algorithm. In *ICML*, 2014.

30 Lichao Sun, Weiran Huang, Philip S Yu, and Wei Chen. Multi-round influence maximization. In *KDD*, pages 2249–2258. ACM, 2018.

31 Youze Tang, Yanchen Shi, and Xiaokui Xiao. Influence maximization in near-linear time: A martingale approach. In *SIGMOD'15*, pages 1539–1554. ACM, 2015.

32 Youze Tang, Xiaokui Xiao, and Yanchen Shi. Influence maximization: near-optimal time complexity meets practical efficiency. In *SIGMOD'14*, 2014.

33 Guangmo Tong, Weili Wu, Shaojie Tang, and Ding-Zhu Du. Adaptive influence maximization in dynamic social networks. *IEEE/ACM Transactions on Networking (TON)*, 25(1):112–125, 2017.

34 Jan Vondrák. Submodularity in combinatorial optimization, 2007.

35 Ailian Wang, Weili Wu, and Lei Cui. On Bharathi–Kempe–Salek conjecture for influence maximization on arborescence. *Journal of Combinatorial Optimization*, 31(4):1678–1684, 2016.

36 Chi Wang, Wei Chen, and Yajun Wang. Scalable influence maximization for independent cascade model in large-scale social networks. *DMKD*, 2012.

37 Jing Yuan and Shaojie Tang. No Time to Observe: Adaptive Influence Maximization with Partial Feedback. In *IJCAI*, 2017.

A Missing Proof from Section 5

▶ **Theorem 5.1.** *When the influence graph is a one-directional bipartite graph $G(L, R, E, p)$, the adaptivity gap on the influence maximization problem in the IC model with full-adoption feedback is $\frac{e}{e-1}$.*

Proof. For each node u, it suffices to prove that for any configuration $(x_1, \ldots, x_n)$,

$$F_u(x_1, \ldots, x_n) \geq \left(1 - \frac{1}{e}\right) f_u^+(x_1, \ldots, x_n), \tag{32}$$

where F_u and f_u^+ are the same as defined in the proof of Theorem 4.1. We use p_i to denote the probability that node i can reach node u, then we have

$$F_u(x_1, \ldots, x_n) = 1 - \prod_{i=1}^{n}(1 - p_i x_i). \tag{33}$$

On the other side, let $\mathcal{E}_i$ denote the event that node u becomes active right after the optimal policy π^+ chooses node i. We know that $\Pr[\mathcal{E}_i] \leq x_i \cdot p_i$ and thus we can conclude that

$$f_u^+(x_1, \ldots, x_n) = \sum_{i=1}^{n} \Pr\left[\mathcal{E}_i\right] \leq \sum_{i=1}^{n} x_i p_i. \tag{34}$$

Combining Eq. (33) (34) and the fact that

$$1 - \prod_{i=1}^{n}(1 - y_i) \geq \left(1 - \frac{1}{e}\right) \min\{1, \sum_{i=1}^{n} y_n\} \tag{35}$$

holds for all $y_i \in [0, 1]$, we can prove Eq. (32) and conclude the proof. ◀

B Missing Proofs and Further Discussions from Section 6

▷ Claim 6.2. For any $\epsilon > 0$, if $k \geq 8/\epsilon^3$, we have $\mathbb{E}[\mathrm{OPT}_A(G,k)] \geq (1-\epsilon)kt$.

Proof. Consider the following adaptive policy π: π always selects the inactive node that is closest to the origin of the directed line, until it reaches the budget. Let X_i ($i \in [k]$) denote the number of nodes that can be reached from the i^{th} seed and let $X = X_1 + \cdots + X_k$. It is easy to see that $\mathbb{E}[\mathrm{OPT}_A(G,k)] \geq \sigma(\pi) = \mathbb{E}[X]$. Let $Y_i \sim GE(1-1/t)$, i.e., Y_i is a geometric random variable parametrized with $1-1/t$. $Y_1, \ldots, Y_k$ are independent and we know that $\mathbb{E}[Y_i] = t$ and $\mathrm{Var}[Y_i] = t^2 - t$. Our key observation is that $\mathbb{E}[X] = \mathbb{E}[\min\{Y_1 + \cdots + Y_k, kt\}]$. By Chebshev bounds, we have

$$\Pr\left[Y_1 + \cdots + Y_k < (1-\epsilon/2)kt\right] \leq \frac{4k(t^2-t)}{\epsilon^2 k^2 t^2} \leq \frac{4}{\epsilon^2 k} \leq \epsilon/2. \tag{36}$$

Thus we know that

$$\begin{aligned}\mathbb{E}\left[\min\{Y_1 + \cdots + Y_k, kt\}\right] &\geq \Pr\left[Y_1 + \cdots + Y_k \geq (1-\epsilon/2)kt\right] \cdot (1-\epsilon/2)\,kt \\ &\geq (1-\epsilon/2) \cdot (1-\epsilon/2)\,kt \geq (1-\epsilon)\,kt.\end{aligned} \tag{37}$$

This concludes the proof. ◁

▷ Claim 6.3. The optimal non-adaptive strategy is to select $v_{11}, \ldots, v_{k1}$ as seeds. Thus, we have $\mathbb{E}[\mathrm{OPT}_N(G,k)] = (1-(1-1/t)^t)kt$.

Proof. In the non-adaptive setting, for any node u and seed set S, we define the distance between the node u and the set S as the distance between u and the closest predecessor of u in S. We know that the probability that the node u is active only depends on the distance between u and S. Let N_i ($i \geq 0$) denote the set of nodes that has distance i with S. Then we know that (i) nodes in N_i are active with probability $(1-1/t)^i$, (ii) $N_0, N_1, \ldots N_{kt-1}$ are disjoint and $|N_i| \leq k$. Now we have that $\sigma(S) = \sum_{i=0}^{kt-1}(1-1/t)^i \cdot |N_i| \leq \sum_{i=0}^{t-1}(1-1/t)^i \cdot k$. Thus, we can conclude that the optimal non-adaptive solution is to select $v_{11}, \ldots, v_{k1}$ as seeds and $\mathbb{E}[\mathrm{OPT}_N(G,k)] = \sum_{i=0}^{t-1}(1-1/t)^i \cdot k = (1-(1-1/t)^t)kt$. ◁

Discussion on Existing Approaches. In this paragraph, we give a hard instance showing that existing approaches cannot yield better-than-2 upper bounds on the adaptivity gap. The hard instance is exactly the directed line constructed in Theorem 6.1, i.e., a directed line of length kt and each edge is live with probability $1-1/t$. We use node i to denote the $(i-1)^{th}$ successor of the origin of the directed line, notice that the origin itself is denoted as node 1.

Multilinear Extension. One common strategy is to use the multilinear extension as in [27, 3]. In [3], they consider the *stochastic submodular optimization* problem and prove that $f^+(x_1, \ldots, x_n) \leq \frac{e}{e-1}F(x_1, \ldots, x_n)$ holds for any configuration $(x_1, \ldots, x_n)$. We show that the ratio of $f^+(x_1, \ldots, x_n)/F(x_1, \ldots, x_n)$ can approach 2 in our example. To be more specific, consider the configuration $(1, 1/t, \ldots, 1/t)$, we claim that $f^+(1, 1/t, \ldots, 1/t) = kt$. Consider the adaptive policy π that always selects the inactive node that is closest to the origin of the directed line. The policy π will select the first node with probability 1 and other nodes with probability $1/t$, since it will seed a node if and only if its incoming edge is blocked, this can happen with probability $1/t$. On the other side, we have $F(1, 1/t, \ldots, 1/t) \leq F(1-1/t, 0, \ldots, 0) + F(1/t, \ldots, 1/t) \leq (1-1/t)t + F(1/t, \ldots, 1/t) \leq t + \frac{1}{2}kt + k$. The first inequality holds because of the DR-submodularity of the multilinear extension and the third one holds because every node u in the line is active with probability

$$\begin{aligned}&\sum_{i=0}^{\infty}\Pr\left[u \text{ is activated by its } i^{th} \text{ predecessor}\right]\\&\le \sum_{i=0}^{\infty}\frac{1}{t}\cdot\left(1-\frac{1}{t}\right)^i\cdot\left(1-\frac{1}{t}\right)^i=\frac{1}{t}\cdot\frac{1}{1-(1-1/t)^2}=\frac{t}{2t-1}.\end{aligned} \tag{38}$$

We conclude that when $t,k\to\infty$, $f^+(1,1/t,\ldots,1/t)/F(1,1/t,\ldots,1/t)\to 2$.

Random Walk Non-adaptive Strategy. In [16, 17, 8], the authors consider the *adaptive stochastic probing* problem and they convert an adaptive policy to a non-adaptive policy by sampling a random leaf of the decision tree of the adaptive policy. Using our hard instance in the previous paragraph, we can show that this approach (i.e., random walk non-adaptive strategy) can give an upper bound of at most 2. To be more specific, we again consider the adaptive strategy π and its corresponding non-adaptive strategy $\mathcal{W}(\pi)$, where $\mathcal{W}(\pi)$ picks a random leaf of the decision tree of the policy π. We are going to show that $f^+(1,1/t,\ldots,1/t)/\sigma(\mathcal{W}(\pi))$ approaches 2 asymptotically and it is sufficient to show that $\sigma(\mathcal{W}(\pi))\le t+k+\frac{1}{2}kt$. We imagine that node 1 appears in $\mathcal{W}(\pi)$ with probability $1/t$ instead of 1, this is for ease of analysis and it will decrease the influence spread for at most $(1-1/t)\cdot t$ due to the submodularity of the influence spread function. For any node u, u is activated by its i^{th} predecessor (if it has one) when (i) the random seed set $\mathcal{W}(\pi)$ does not contain nodes between u and its i^{th} predecessor (this happens with probability $\left(1-\frac{1}{t}\right)^i$), (ii) its i^{th} predecessor is included in the seed set (this happens with probability $\frac{1}{t}$) and (iii) node u can be reached from its i^{th} predecessor (this happens with probability $\left(1-\frac{1}{t}\right)^i$). Moreover, we know that the above three events are independent in the non-adaptive setting, thus the probability that node u is activated by the i^{th} predecessor is $\frac{1}{t}\cdot\left(1-\frac{1}{t}\right)^i\cdot\left(1-\frac{1}{t}\right)^i$ and the probability that it is active is no more than $\frac{t}{2t-1}$. This concludes our argument.