Adaptive Influence Maximization with Myopic Feedback

Anonymous Author(s)
Affiliation
Address
email

Abstract

We study the *adaptive influence maximization problem with myopic feedback* under the independent cascade model: one sequentially selects k nodes as seeds one by one from a social network, and each selected seed returns the immediate neighbors it activates as the feedback available for later selections, and the goal is to maximize the expected number of total activated nodes, referred as the *influence spread*. We show that the *adaptivity gap*, the ratio between the optimal adaptive influence spread and the optimal non-adaptive influence spread, is at most 4 and at least $e/(e-1)$, and the approximation ratios with respect to the optimal adaptive influence spread of both the non-adaptive greedy and adaptive greedy algorithms are at least $\frac{1}{4}(1 - \frac{1}{e})$ and at most $\frac{e^2+1}{(e+1)^2} < 1 - \frac{1}{e}$. Moreover, the approximation ratio of the non-adaptive greedy algorithm is no worse than that of the adaptive greedy algorithm, when considering all graphs. Our result confirms a long-standing open conjecture of Golovin and Krause (2011) on the constant approximation ratio of adaptive greedy with myopic feedback, and it also suggests that adaptive greedy may not bring much benefit under myopic feedback.

1 Introduction

Influence maximization is the task of given a social network and a stochastic diffusion model on the network, finding the k seed nodes with the largest expected influence spread in the model [9]. Influence maximization and its variants have applications in viral marketing, rumor control, etc. and have been extensively studied (cf. [5, 10]).

In this paper, we focus on the *adaptive influence maximization* problem, where seed nodes are sequentially selected one by one, and after each seed selection, partial or full diffusion results from the seed are returned as the feedback, which could be used for subsequent seed selections. Two main types of feedback has been proposed and studied before: (a) *full-adoption feedback*, where the entire diffusion process from the seed selected is returned as the feedback, and (b) *myopic feedback*, where only the immediate neighbors activated by the selected seed are returned as the feedback. Under the common independent cascade (IC) model where every edge in the graph has an independent probability of passing influence, Golovin and Krause [6] show that the full-adoption feedback model satisfies the key *adaptive submodularity* property, which enables a simple adaptive greedy algorithm to achieve a $(1 - 1/e)$ approximation to the adaptive optimal solution. However, the IC model with myopic feedback is not adaptive submodular, and Golovin and Krause [6] only conjecture that in this case the adaptive greedy algorithm still guarantees a constant approximation. To the best of our knowledge, this conjecture is still open before our result in this paper, which confirms that indeed adaptive greedy is a constant approximation of the adaptive optimal solution.

In particular, our paper presents two sets of related results on adaptive influence maximization with myopic feedback under the IC model. We first study the *adaptivity gap* of the problem (Section 3),

which is defined as the ratio between the adaptive optimal solution and the non-adaptive optimal solution, and is an indicator on how useful the adaptivity could be to the problem. We show that the adaptivity gap for our problem is at most 4 (Theorem 1) and at least $e/(e-1)$ (Theorem 2). The proof of the upper bound 4 is the most involved, because the problem is not adaptive submodular, and we have to create a hybrid policy that involves three independent runs of the diffusion process in order to connect between an adaptive policy and a non-adaptive policy. Next we study the approximation ratio with respect to the adaptive optimal solution for both the non-adaptive greedy and adaptive greedy algorithms (Section 4). We show that the approximation ratios of both algorithms are at least $\frac{1}{4}(1-\frac{1}{e})$ (Theorem 3), which combines the adaptivity upper bound of 4 with the results that both algorithms achieve $(1-1/e)$ approximation of the non-adaptive optimal solution (the $(1-1/e)$ approximation ratio for the adaptive greedy algorithm requires a new proof). We further show that the approximation ratios for both algorithms are at most $\frac{e^2+1}{(e+1)^2} \approx 0.606$, which is strictly less than $1-1/e \approx 0.632$, and the approximation ratio of adaptive greedy is at most that of the non-adaptive greedy when considering all graphs (Theorem 4).

In summary, our contribution is the systematic study on adaptive influence maximization with myopic feedback under the IC model. We prove both constant upper and lower bounds on the adaptivity gap in this case, and constant upper and lower bounds on the approximation ratios (with respect to the optimal adaptive solution) achieved by non-adaptive greedy and adaptive greedy algorithms. The constant approximation ratio of the adaptive greedy algorithm answers a long-standing open conjecture affirmatively. Our result on the adaptivity gap is the first one on a problem not satisfying adaptive submodularity. Our results also suggest that adaptive greedy may not bring much benefit under the myopic feedback model.

Due to the space constraint, full proof details are included in the supplementary material.

Related Work. Influence maximization as a discrete optimization task is first proposed by Kempe et al. [9], who propose the independent cascade, linear threshold and other models, study their submodularity and the greedy approximation algorithm for the influence maximization task. Since then, influence maximization and its variants have been extensively studied. We refer to recent surveys [5, 10] for the general coverage of this area.

Adaptive submodularity is formulated by Golovin and Krause [6] for general stochastic adaptive optimization problems, and they show that the adaptive greedy algorithm achieves $1-1/e$ approximation if the problem is adaptive monotone and adaptive submodular. They study the influence maximization problem under the IC model as an application, and prove that the full-adoption feedback under the IC model is adaptive submodular. However, in their arXiv version, they show that the myopic feedback version is not adaptive submodular, and they conjecture that adaptive greedy would still achieve a constant approximation in this case.

Adaptive influence maximization has been studied in [16, 17, 14, 11, 15]. Tong et al. [16] provide both adaptive greedy and efficient heuristic algorithms for adaptive influence maximization. Their theoretical analysis works for the full-adoption feedback model but has a gap when applied to myopic feedback, which is confirmed by the authors. Yuan and Tang [17] introduce the partial feedback model and develop algorithms that balance the tradeoff between delay and performance, and their partial feedback model does not coincide with the myopic feedback model. Salha et al. [11] consider a different diffusion model where edges can be reactivated at each time step, and they show that myopic feedback under this model is adaptive submodular. Sun et al. [14] study the multi-round adaptive influence maximization problem, where k seeds are selected in each round and at the end of the round the full-adoption feedback is returned. Tong [15] introduces a general feedback model and develops some heuristic algorithms for this model. A different two stage seeding process has also been studied [12, 3, 13], but the model is quite different, since their first stage of selecting a node set X is only to introduce the neighbors of X as seeding candidates for the second stage.

Adaptivity gap has been studied by two lines of research. The first line of work utilizes multilinear extension and adaptive submodularity to study adaptivity gaps for the class of stochastic submodular maximization problems and give a $e/(e-1)$ upper bound for matroid constraints [2, 1]. The second line of work [7, 8, 4] studies the stochastic probing problem and proposes the idea of random-walk non-adaptive policy on the decision tree, which partially inspires our analysis. However, their analysis also implicitly depends on adaptive submodularity. In contrast, our result on the adaptivity gap is the first on a problem that does not satisfy adaptive submodularity (see Section 3.1 for more discussions).

2

2 Model and Problem Definition

Diffusion Model. 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 $G = (V, E, p)$, where V is the set of nodes ($|V| = n$), $E \subseteq V \times V$ is the set of directed edges, and each directed edge $(u, v) \in E$ is associated with a probability $p_{uv} \in [0, 1]$. The *live edge* graph $L = (V, L(E))$ is a random subgraph of G, for any edge $(u, v) \in E$, $(u, v) \in L(E)$ with independent probability p_{uv}. If $(u, v) \in L(E)$, we say edge (u, v) is *live*, otherwise we say it is *blocked*. The dynamic diffusion in the IC model is as follows: at time $t = 0$ a live-edge graph L is sampled and nodes in a seed set $S \subseteq V$ are activated. At every discrete time $t = 1, 2, \ldots$, if a node u was activated at time $t - 1$, then all of u's out-going neighbors in L are activated at time t. The propagation continues until there are no more activated nodes at a time step. The dynamic model can be viewed equivalently as every activated node u has one chance to activate each of its out-going neighbor v with independent success probability p_{uv}. Given a seed set S, the *influence spread* of S, denoted $\sigma(S)$, is the expected number of nodes activated in the diffusion process from S, i.e. $\sigma(S) = \mathbb{E}_L[|\Gamma(S, L)|]$, where $\Gamma(S, L)$ is the set of nodes reachable from S in graph L.

Influence Maximization Problem. Under the IC model, we formalize the influence maximization (IM) problem in both non-adaptive and adaptive settings. Influence maximization in the non-adaptive setting follows the classical work of [9], and is defined below.

Definition 1 (Non-adaptive Influence Maximization). Non-adaptive influence maximization *is the problem of given a directed influence graph $G = (V, E, p)$ with IC model parameters $\{p_{uv}\}_{(u,v) \in E}$ and a budget k, finding a seed set S^* of at most k nodes such that the influence spread of S^*, $\sigma(S^*)$, is maximized, i.e. finding $S^* \in \mathrm{argmax}_{S \subseteq V, |S| \leq k} \sigma(S)$.*

We formulate influence maximization in the adaptive setting following the framework of [6]. Let O denote the set of states, which informally correspond to the feedback information in the adaptive setting. A *realization* ϕ is a function $\phi : V \to O$, such that for $u \in V$, $\phi(u)$ represents the feedback obtained when selecting u as a seed node. In this paper, we focus on the *myopic feedback* model [6], which means the feedback of a node u only contains the status of the out-going edges of u being live or blocked. Informally it means that after selecting a seed we can only see its one step propagation effect as the feedback. The realization ϕ then determines the status of every edge in G, and thus corresponds to a live-edge graph. As a comparison, the *full-adoption feedback* model [6] is such that for each seed node u, the feedback contains the status of every out-going edge of every node v that is reachable from u in a live-edge graph L. This means that after selecting a seed u, we can see the full cascade from u as the feedback. In the full-adoption feedback case, each realization ϕ also corresponds to a unique live-edge graph. Henceforth, we refer to ϕ as both a realization and a live-edge graph interchangeably. In the remainder of this section, the terminologies we introduce apply to both feedback models, unless we explicitly point out which feedback model we are discussing.

Let $\mathcal{R}$ denote the set of all realizations. We use Φ to denote a random realization, following the distribution $\mathcal{P}$ over random live-edge graphs (i.e. each edge $(u, v) \in E$ has an independent probability of p_{uv} to be live in Φ). Given a subset S and a realization ϕ, we define *influence utility function* $f : 2^V \times \mathcal{R} \to \mathbb{R}^+$ as $f(S, \phi) = |\Gamma(S, \phi)|$, where $\mathbb{R}^+$ is the set of non-negative real numbers. That is, $f(S, \phi)$ is the number of nodes reachable from S in realization (live-edge graph) ϕ. Then it is clear that influence spread $\sigma(S) = \mathbb{E}_{\Phi \sim \mathcal{P}}[f(S, \Phi)]$.

In the *adaptive influence maximization* problem, we could sequentially select nodes as seeds, and after selecting one seed node, we could obtain its feedback, and use the feedback to guide further seed selections. A *partial realization* ψ maps a subset of nodes in V, denoted $\mathrm{dom}(\psi)$ for domain of ψ, to their states. Partial realization ψ represents the feedback we could obtain after nodes in $\mathrm{dom}(\psi)$ are selected as seeds. For convenience, we also represent ψ as a relation, i.e., $\psi = \{(u, o) \in V \times O : u \in \mathrm{dom}(\psi), o = \psi(u)\}$. We say that a full realization ϕ is *consistent* with a partial realization ψ, denoted as $\phi \sim \psi$, if $\phi(u) = \psi(u)$ for every $u \in \mathrm{dom}(\psi)$.

An adaptive policy π is a mapping from partial realizations to nodes. Given a partial realization ψ, $\pi(\psi)$ represents the next seed node policy π would select when it sees the feedback represented by ψ. Under a full realization ϕ consistent with ψ, after selecting $\pi(\psi)$, the policy would obtain feedback $\phi(\pi(\psi))$, and the partial realization would grow to $\psi' = \psi \cup \{(\pi(\psi), \phi(\pi(\psi)))\}$, and policy π could pick the next seed node $\pi(\psi')$ based on partial realization ψ'. For convenience, we only consider deterministic policies in this paper, and the results we derived can be easily extend to randomized

147 policies. Let $V(\pi, \phi)$ denote the set of nodes selected by policy π under realization ϕ. For the
148 adaptive influence maximization problem, we consider the simple cardinality constraint such that
149 $|V(\pi, \phi)| \leq k$, i.e. the policy only selects at most k nodes. Let $\Pi(k)$ denote the set of such policies.

150 The objective of an adaptive policy π is its *adaptive influence spread*, which is the expected number
151 of nodes that are activated under policy π. Formally, we define the adaptive influence spread of π as
152 $\sigma(\pi) = \mathbb{E}_{\Phi \sim \mathcal{P}}[f(V(\pi, \Phi), \Phi)]$. The adaptive influence maximization problem is defined as follows.

Definition 2 (Adaptive Influence Maximization)**.** Adaptive influence maximization *is the problem*
154 *of given a directed influence graph $G = (V, E, p)$ with IC model parameters $\{p_{uv}\}_{(u,v) \in E}$ and a*
155 *budget k, finding an adaptive policy π^* that selects at most k seed nodes such that the adaptive*
156 *influence spread of π^*, $\sigma(\pi^*)$, is maximized, i.e. finding $\pi^* \in \mathrm{argmax}_{\pi \in \Pi(k)} \sigma(\pi)$.*

157 Note that for any fixed seed set S, we can create a policy π_S that always selects set S regardless of
158 the feedback, which means any non-adaptive solution is a feasible solution for adaptive influence
159 maximization. Therefore, the optimal adaptive influence spread should be at least as good as the
160 optimal non-adaptive influence spread, under the same budget constraint.

Adaptivity Gap. Since the adaptive policy is usually hard to design and analyze and the adaptive
162 interaction process may also be slow in practice, a fundamental question for adaptive stochastic
163 optimization problems is whether adaptive algorithms are really superior to non-adaptive algorithms.
164 The *adaptivity gap* measures the gap between the optimal adaptive solution and the optimal non-
165 adaptive solution. More concretely, if we use $\mathrm{OPT}_N(G, k)$ (resp. $\mathrm{OPT}_A(G, k)$) to denote the
166 influence spread of the optimal non-adaptive (resp. adaptive) solution for the IM problem in an
167 influence graph G under the IC model with seed budget k, then we have the following definition.

Definition 3 (Adaptivity Gap for IM)**.** *The adaptivity gap in the IC model is defined as the supremum*
169 *ratio of the influence spread between the optimal adaptive policy and the optimal non-adaptive policy,*
170 *over all possible influence graphs and seed set size k, i.e.,*

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

Submodularity and Adaptive Submodularity. Non-adaptive influence maximization is often
172 solved via submodular function maximization technique. A set function $f : 2^V \to \mathbb{R}$ is *submodular*
173 if for all $S \subseteq T \subseteq V$ and all $u \in V \setminus T$, $f(S \cup \{u\}) - f(S) \geq f(T \cup \{u\}) - f(T)$. Set function f
174 is monotone if for all $S \subseteq T \subseteq V$, $f(S) \leq f(T)$. Kempe et al. [9] show that the influence spread
175 function $\sigma(S)$ under the IC model is monotone and submodular, and thus a simple (non-adaptive)
176 greedy algorithm achieves a $(1 - \frac{1}{e})$ approximation of the optimal (non-adaptive) solution, assuming
177 function evaluation $\sigma(S)$ is given by an oracle.

178 Golovin and Krause [6] define *adaptive submodularity* for the adaptive stochastic optimization
179 framework. In the context of adaptive influence maximization, adaptive submodularity can be defined
180 as follows. Given a utility function f, for any partial realization ψ and a node $u \notin \mathrm{dom}(\psi)$, we define
181 the marginal gain of u given ψ as $\Delta_f(u \mid \psi) = \mathbb{E}_{\Phi \sim \mathcal{P}}[f(\mathrm{dom}(\psi) \cup \{u\}, \Phi) - f(\mathrm{dom}(\psi), \Phi) \mid \Phi \sim \psi]$,
182 i.e. the expected marginal gain on influence spread when adding u to the partial realization ψ. A
183 partial realization ψ is a *sub-realization* of another partial realization ψ' if $\psi \subseteq \psi'$ when treating
184 both as relations. We say that the utility function f is *adaptive submodular* with respect to $\mathcal{P}$ if
185 for any two fixed partial realizations ψ and ψ' such that $\psi \subseteq \psi'$, for any $u \notin \mathrm{dom}(\psi')$, we have
186 $\Delta_f(u \mid \psi) \geq \Delta_f(u \mid \psi')$, that is, the marginal influence spread of a node given more feedback is
187 at most its marginal influence spread given less feedback. We say that f is *adaptive monotone* with
188 respect to $\mathcal{P}$ if for any partial realization ψ with $\mathrm{Pr}_{\Phi \sim \mathcal{P}}(\Phi \sim \psi) > 0$, $\Delta_f(u \mid \psi) \geq 0$.

189 Golovin and Krause [6] show that the influence utility function under the IC model with full adoption
190 feedback is adaptive monotone and adaptive submodular, and thus the adaptive greedy algorithm
191 achieves $(1 - \frac{1}{e})$ approximation of the adaptive optimal solution. However, they show that the
192 influence utility function under the IC model with myopic feedback is not adaptive submodular. They
193 conjecture that the adaptive greedy policy still provides a constant approximation. In this paper, we
194 show that the adaptive greedy policy provides a $\frac{1}{4}(1 - \frac{1}{e})$ approximation, and thus finally address
195 this conjecture affirmatively.

4

3 Adaptivity Gap in Myopic Feedback Model

In this section, we analyze the adaptivity gap for influence maximization problems under the myopic feedback model and derive both upper and lower bounds.

3.1 Upper Bound on the Adaptivity Gap

Our main result is an upper bound on the adaptivity gap for myopic feedback models, which is formally stated below.

Theorem 1. *Under the IC model with myopic feedback, the adaptivity gap for the influence maximization problem is at most 4.*

Proof outline. We now outline the main ideas and the structure of the proof of Theorem 1. The main idea is to show that for each adaptive policy π, we could construct a non-adaptive randomized policy $\mathcal{W}(\pi)$, such that the adaptive influence spread $\sigma(\pi)$ is at most four times the non-adaptive influence spread of $\mathcal{W}(\pi)$, denoted $\sigma(\mathcal{W}(\pi))$. This would immediately imply Theorem 1. The non-adaptive policy $\mathcal{W}(\pi)$ is constructed by viewing adaptive policy π as a decision tree with leaves representing the final seed set selected (Definition 4), and $\mathcal{W}(\pi)$ simply samples such a seed set based on the distribution of the leaves (Definition 5). The key to connect $\sigma(\pi)$ with $\sigma(\mathcal{W}(\pi))$ is by introducing a fictitious hybrid policy $\bar{\pi}$, such that $\sigma(\pi) \leq \bar{\sigma}(\bar{\pi}) \leq 4\sigma(\mathcal{W}(\pi))$, where $\bar{\sigma}(\bar{\pi})$ is the *aggregate adaptive influence spread* (defined in Eqs. (2) and (3)). Intuitively, $\bar{\pi}$ works on three independent realizations Φ^1, Φ^2, Φ^3, such that it adaptively selects seeds just as π working on Φ^1, but each selected seed has three independent chances to activate its out-neighbors accordingly the union of Φ^1, Φ^2, Φ^3. The inequality $\sigma(\pi) \leq \bar{\sigma}(\bar{\pi})$ is immediate and the main effort is on proving $\bar{\sigma}(\bar{\pi}) \leq 4\sigma(\mathcal{W}(\pi))$.

To do so, we first introduce general notations $\sigma^t(S)$ and $\sigma^t(\pi)$ with $t = 1, 2, 3$, where $\sigma^t(S)$ is the *t-th aggregate influence spread* for a seed set S and $\sigma^t(\pi)$ is the *t-th aggregate adaptive influence spread* for an adaptive policy π, and they mean that all seed nodes have t independent chances to activate their out-neighbors. Obviously, $\bar{\sigma}(\bar{\pi}) = \sigma^3(\pi)$ and $\sigma(\mathcal{W}(\pi)) = \sigma^1(\mathcal{W}(\pi))$. We then represent $\sigma^t(S)$ and $\sigma^t(\pi)$ as a summation of k non-adaptive marginal gains $\Delta_{f^t}(u \mid \mathrm{dom}(\psi^1))$'s and adaptive marginal gains $\Delta_{f^t}(u \mid \psi^1)$'s, respectively (Definition 6 and Lemma 1), with respect to the different levels of the decision tree. Next, we establish the key connection between the adaptive marginal gain and the nonadaptive marginal gain (Lemma 3): $\Delta_{f^3}(u \mid \psi^1) \leq 2\Delta_{f^2}(u \mid \mathrm{dom}(\psi^1))$. This immediately implies that $\sigma^3(\pi) \leq 2\sigma^2(\mathcal{W}(\pi))$. Finally, we prove that the t-th aggregate non-adaptive influence spread $\sigma^t(S)$ is bounded by $t \cdot \sigma(S)$, which implies that $\sigma^2(\mathcal{W}(\pi)) \leq 2\sigma(\mathcal{W}(\pi))$. This concludes the proof.

We remark that our introduction of the hybrid policy $\bar{\pi}$ is inspired by the analysis in [4], which shows that the adaptivity gap for the *stochastic multi-value probing (SMP)* problem is at most 2. However, our analysis is more complicated than theirs and thus is novel in several aspects. First, the SMP problem is simpler than our problem, with the key difference being that SMP is adaptive submodular but our problem is not. Therefore, we cannot apply their way of inductive reasoning that implicitly relies on adaptive submodularity. Instead, we have to use our marginal gain representation and redo the bounding analysis carefully based on the (non-adaptive) submodularity of the influence utility function on live-edge graphs. Moreover, our influence utility function is also sophisticated and we have to use three independent realizations in order to apply the submodularity on live-edge graphs, which results in an adaptivity bound of 4, while their analysis only needs two independent realizations to achieve a bound of 2. We now provide the technical proof of Theorem 1. We first formally define the decision tree representation.

Definition 4 (Decision tree representation for adaptive policy). *An adaptive policy π can be seen as a decision tree $\mathcal{T}(\pi)$, where each node s of $\mathcal{T}(\pi)$ corresponds to a partial realization ψ_s, with the root being the empty partial realization, and node s' is a child of s if $\psi_{s'} = \psi_s \cup \{\pi(\psi_s), \phi(\pi(\psi_s))\}$ for some realization $\phi \sim \psi_s$. Each node s is associated with a probability p_s, which is the probability that the policy π generates partial realization ψ_s, i.e. the probability that the policy would walk on the tree from the root to node s.*

Next we define the non-adaptive randomized policy $\mathcal{W}(\pi)$, which randomly selects a leaf of $\mathcal{T}(\pi)$.

Definition 5 (Random-walk non-adaptive policy [8])**.** *For any adaptive policy π, let $\mathcal{L}(\pi)$ denote the set of leaves of $\mathcal{T}(\pi)$. Then we construct a randomized non-adaptive policy $\mathcal{W}(\pi)$ as follows: for any leaf $\ell \in \mathcal{L}(\pi)$, $\mathcal{W}(\pi)$ picks leaf ℓ with probability p_ℓ and selects $\mathrm{dom}(\psi_\ell)$ as the seed set.*

Before proceeding further with our analysis, we introduce some notations for the myopic feedback model. In the myopic feedback model, we notice that the state spaces for all nodes are mutually independent and disjoint. Thus we could decompose the realization space $\mathcal{R}$ into independent subspace, $\mathcal{R} = \times_{u \in V} O_u$, where O_u is the set of all possible states for node u. For any full realization ϕ (resp. partial realization ψ), we would use ϕ_S (resp. ψ_S) to denote the feedback for the node set $S \subseteq V$. Note that ϕ_S and ψ_S are partial realizations with domain S. Similarly, we would also use $\mathcal{P}_S$ to denote the probability space $\times_{u \in S} \mathcal{P}_u$, where $\mathcal{P}_u$ is the probability distribution over O_u (i.e. each out-going edge (u, v) of u is live with independent probability p_{uv}). With a slight abuse of notation, we further use ϕ_S (resp. ψ_S) to denote the set of live edges leaving from S under ϕ (resp. ψ). Then we could use notation $\phi_S^1 \cup \phi_S^2$ to represent the union of live-edges from ϕ^1 and ϕ^2 leaving from S, and similarly $\psi^1 \cup \phi_S^2$ with $\mathrm{dom}(\psi) = S$.

Construction for hybrid policy. For any adaptive policy π, we define a fictitious hybrid policy $\bar{\pi}$ that works on three independent random realizations Φ^1, Φ^2 and Φ^3 simultaneously, thinking about them as from three copies of the graphs G_1, G_2 and G_3. Note that $\bar{\pi}$ is not a real adaptive policy — it is only used for our analytical purpose to build connections between the adaptive policy π and the non-adaptive policy $\mathcal{W}(\pi)$. In terms of adaptive seed selection, $\bar{\pi}$ acts exactly the same as π on G_1, responding to partial realizations ψ^1 obtained so far from the full realization Φ^1 of G_1, and disregarding the realizations Φ^2 and Φ^3. However, the difference is when we define adaptive influence spread for $\bar{\pi}$, we aggregate the three partial realizations on the seed set together. More precisely, for any $t = 1, 2, 3$, we define the t-th aggregate influence utility function as $f^t : 2^V \times \mathcal{R}^t \to \mathbb{R}^+$

$$f^t\left(S, \phi^1, \cdots, \phi^t\right) := f\left(S, (\cup_{i \in [t]} \phi_S^i, \phi_{V \setminus S}^1)\right), \tag{2}$$

where $(\cup_{i \in [t]} \phi_S^i, \phi_{V \setminus S}^1)$ means a new realization ϕ' where on set S its set of out-going live-edges is the same as union of $\phi^1, \cdots \phi^t$ and on set $V \setminus S$, its set of out-going live-edges is the same as ϕ^1, and f is the original influence utility function defined in Section 2. The objective of the hybrid policy $\bar{\pi}$ is then defined as the adaptive influence spread under policy $\bar{\pi}$, i.e.,

$$\bar{\sigma}(\bar{\pi}) := \mathop{\mathbb{E}}_{\Phi^1, \Phi^2, \Phi^3 \sim \mathcal{P}}\left[f^3(V(\pi, \Phi^1), \Phi^1, \Phi^2, \Phi^3)\right]$$
$$= \mathop{\mathbb{E}}_{\Phi^1, \Phi^2, \Phi^3 \sim \mathcal{P}}\left[f\left(V(\pi, \Phi^1), (\Phi_{V(\pi, \Phi^1)}^1 \cup \Phi_{V(\pi, \Phi^1)}^2 \cup \Phi_{V(\pi, \Phi^1)}^3, \Phi_{V \setminus V(\pi, \Phi^1)}^1)\right)\right]. \tag{3}$$

In other words, the adaptive influence spread of the hybrid policy $\bar{\pi}$ is the influence spread of seed nodes $V(\pi, \Phi^1)$ selected in graph G_1 by policy π, where the live-edge graph on the seed set part $V(\pi, \Phi^1)$ is the union of live-edge graphs of G_1, G_2 and G_3, and the live-edge graph on the non-seed set part is only that of G_1. It can also be viewed as each seed node has three independent chances to activate its out-neighbors. Since the hybrid policy $\bar{\pi}$ acts the same as policy π on influence graph G_1, we can easily conclude:

Claim 1. $\bar{\sigma}(\bar{\pi}) \geq \sigma(\pi)$.

We also define t-th aggregate influence spread for a seed set S, $\sigma^t(S)$, as $\sigma^t(S) = \mathbb{E}_{\Phi^1, \cdots, \Phi^t \sim \mathcal{P}}\left[f^t(S, \Phi^1, \cdots, \Phi^t)\right]$. Then, for the random-walk non-adaptive policy $\mathcal{W}(\pi)$, we define $\sigma^t(\mathcal{W}(\pi)) = \sum_{\ell \in \mathcal{L}(\pi)} p_\ell \cdot \sigma^t(\mathrm{dom}(\psi_\ell))$, that is, the t-th aggregate influence spread of $\mathcal{W}(\pi)$ is the average t-th aggregate influence spread of seed nodes selected by $\mathcal{W}(\pi)$ according to distribution of the leaves in the decision tree $\mathcal{T}(\pi)$. Similarly, we define the t-th aggregate adaptive influence spread for an adaptive policy π as $\sigma^t(\pi) = \mathbb{E}_{\Phi^1, \cdots, \Phi^t \sim \mathcal{P}}\left[f^t(V(\pi, \Phi^1), \Phi^1, \cdots, \Phi^t)\right]$. Note that $\bar{\sigma}(\bar{\pi}) = \sigma^3(\pi)$.

Now, we could give a definition for the conditional expected marginal gain for the aggregate influence utility function f^t over live-edge graph distributions.

Definition 6. *The expected non-adaptive marginal gain of u given set S under f^t is defined as:*

$$\Delta_{f^t}(u \mid S) = \mathop{\mathbb{E}}_{\Phi^1, \cdots, \Phi^t \sim \mathcal{P}}\left[f^t\left(S \cup \{u\}, \Phi^1, \cdots, \Phi^t\right) - f^t\left(S, \Phi^1, \cdots, \Phi^t\right)\right]. \tag{4}$$

The expected adaptive marginal gain of u given partial realization ψ^1 under f^t is defined as:

$$\Delta_{f^t}(u \mid \psi^1) = \mathop{\mathbb{E}}_{\Phi^1,\cdots,\Phi^t \sim \mathcal{P}} \left[f^t\left(\text{dom}(\psi^1) \cup \{u\}, \Phi^1, \cdots, \Phi^t\right) - f^t\left(\text{dom}(\psi^1), \Phi^1, \cdots, \Phi^t\right) \mid \Phi^1 \sim \psi^1 \right].$$

(5)

The following lemma connects $\sigma^t(\pi)$ (and thus $\bar{\sigma}(\bar{\pi})$) with adaptive marginal gain $\Delta_{f^t}(u \mid \psi)$, and connects $\sigma^t(\mathcal{W}(\pi))$ with non-adaptive marginal gain $\Delta_{f^t}(u \mid S)$. Let $\mathcal{P}_i^\pi$ denote the probability distribution over nodes at depth i of the decision $\mathcal{T}(\pi)$. The proof is by applying telescoping series to convert influence spread into the sum of marginal gains.

Lemma 1. *For any adaptive policy π, and $t \geq 1$, we have*

$$\sigma^t(\pi) = \sum_{i=0}^{k-1} \mathop{\mathbb{E}}_{s \sim \mathcal{P}_i^\pi} \left[\Delta_{f^t}\left(\pi(\psi_s) \mid \psi_s\right) \right], \text{ and } \sigma^t(\mathcal{W}(\pi)) = \sum_{i=0}^{k-1} \mathop{\mathbb{E}}_{s \sim \mathcal{P}_i^\pi} \left[\Delta_{f^t}\left(\pi(\psi_s) \mid \text{dom}(\psi_s)\right) \right].$$

The next lemma bounds two intermediate adaptive marginal gains to be used for Lemma 3. The proof crucially depend on (a) the independence of realizations Φ^1, Φ^2, Φ^3, (b) the independence of feedback of different selected seed nodes, and (c) the submodularity of the influence utility function on live-edge graphs.

Lemma 2. *Let $S = \text{dom}(\psi^1)$ and $S^+ = S \cup \{u\}$ for any partial realization ψ^1 and any $u \notin \text{dom}(\psi^1)$. Then we have*

$$\mathop{\mathbb{E}}_{\Phi^1,\Phi^2,\Phi^3 \sim \mathcal{P}} \left[f\left(S^+, (\Phi_S^1 \cup \Phi_S^2 \cup \Phi_S^3, \Phi_u^1 \cup \Phi_u^2, \Phi_{V \setminus S^+}^1)\right) \right.$$
$$\left. - f\left(S, (\Phi_S^1 \cup \Phi_S^2 \cup \Phi_S^3, \Phi_{V \setminus S}^1)\right) \mid \Phi^1 \sim \psi^1 \right] \leq \Delta_{f^2}(u \mid S).$$

(6)

$$\mathop{\mathbb{E}}_{\Phi^1,\Phi^2,\Phi^3 \sim \mathcal{P}} \left[f\left(S^+, (\Phi_S^1 \cup \Phi_S^2 \cup \Phi_S^3, \Phi_u^1 \cup \Phi_u^2 \cup \Phi_u^3, \Phi_{V \setminus S^+}^1)\right) \right.$$
$$\left. - f\left(S^+, (\Phi_S^1 \cup \Phi_S^2 \cup \Phi_S^3, \Phi_u^1 \cup \Phi_u^2, \Phi_{V \setminus S^+}^1)\right) \mid \Phi^1 \sim \psi^1 \right] \leq \Delta_{f^2}(u \mid S).$$

(7)

Combining the two inequalities above, we obtain the following key lemma essential, which bounds the adaptive marginal gain $\Delta_{f^3}(u \mid \psi^1)$ with the non-adaptive marginal gain $\Delta_{f^2}(u \mid \text{dom}(\psi^1))$.

Lemma 3. *For any partial realization ψ^1 and node $u \notin \text{dom}(\psi^1)$, we have*

$$\Delta_{f^3}(u \mid \psi^1) \leq 2\Delta_{f^2}(u \mid \text{dom}(\psi^1)).$$

(8)

The next lemma gives an upper bound on the t-th aggregate (non-adaptive) influence spread $\sigma^t(S)$ using the original influence spread $\sigma(S)$. The idea of the proof is that each seed node in S has t independent chances to active its out-neighbors, but afterwards the diffusion is among nodes not in S as in the original diffusion.

Lemma 4. *For any $t \geq 1$ and any subset $S \subseteq V$, $\sigma^t(S) \leq t \cdot \sigma(S)$.*

Proof of Theorem 1. It is enough to show that for every adaptive policy π, $\sigma(\pi) \leq 4\sigma(\mathcal{W}(\pi))$. This is done by the following derivation sequence: $\sigma(\pi) \leq \bar{\sigma}(\bar{\pi}) = \sigma^3(\pi) = \sum_{i=0}^{k-1} \mathbb{E}_{s \in \mathcal{P}_i^\pi} \left[\Delta_{f^3}\left(\pi(\psi_s) \mid \psi_s\right) \right] \leq \sum_{i=0}^{k-1} \mathbb{E}_{s \in \mathcal{P}_i^\pi} \left[2\Delta_{f^2}\left(\pi(\psi_s) \mid \text{dom}(\psi_s)\right) \right] = 2\sigma^2(\mathcal{W}(\pi)) \leq 4\sigma(\mathcal{W}(\pi))$, where the first inequality is by Claim 1, the second and the third equalities are by Lemma 1, the second inequality is by Lemma 3 and the last inequality is by Lemma 4. $\square$

3.2 Lower bound

Next, we proceed to give a lower bound on the adaptivity gap for the influence maximization problem in the myopic feedback model. Our result is stated as follow:

Theorem 2. *Under the IC model with myopic feedback, the adaptivity gap for the influence maximization problem is at least $e/(e-1)$.*

Proof Sketch. We construct a bipartite graph $G = (L, R, E, p)$ with $|L| = \binom{m^3}{m^2}$ and $|R| = m^3$. For each subset $X \subset R$ with $|X| = m^2$, there is exactly one node $u \in L$ that connects to all nodes in X. We show that for any $\epsilon > 0$, there is a large enough m such that in the above graph with parameter m the adaptivity gap is at least $e/(e-1) - \epsilon$. $\square$

4 Adaptive and Non-Adaptive Greedy Algorithms

In this section, we consider two prevalent algorithms — the *greedy* algorithm and the *adaptive greedy* algorithm for the influence maximization problem. To the best of our knowledge, we provide the first approximation ratio of these algorithms with respect to the adaptive optimal solution in the IC model with myopic feedback. We formally describe the algorithms in figure 1.

Greedy Algorithm:
$S = \emptyset$
while $|S| < k$ **do**
$\quad u = \text{argmax}_{u \in V \setminus S} \Delta_f(u|S)$
$\quad S = S \cup \{u\}$
end while
return S

Adaptive Greedy Algorithm:
$S = \emptyset, \Psi = \emptyset$
while $|S| < k$ **do**
$\quad u = \text{argmax}_{u \in V \setminus S} \Delta_f(u|\Psi)$
$\quad$ Select u as seed and observe $\Phi(u)$.
$\quad S = S \cup \{u\}, \Psi = \Psi \cup \{(u, \Phi(u))\}$
end while

Figure 1: Description for *greedy* and *adaptive greedy*.

Our main result is summarized below.

Theorem 3. *Both greedy and adaptive greedy are $\frac{1}{4}(1 - \frac{1}{e})$ approximate to the optimal adaptive policy under the IC model with myopic feedback.*

Proof Sketch. The proof for the non-adaptive greedy algorithm is straightforward since the greedy algorithm provides a $(1 - \frac{1}{e})$ approximation to the non-adaptive optimal solution, which by Theorem 1 is at least $\frac{1}{4}$ of the adaptive optimal solution. For the adaptive greedy algorithm, we need to separately prove that it also provides a $(1 - \frac{1}{e})$ approximation to the non-adaptive optimal solution, and then the result is immediate similar to the non-adaptive greedy algorithm. □

Theorem 3 shows that greedy and adaptive greedy can achieve at least an approximation ratio of $\frac{1}{4}(1 - \frac{1}{e})$ with respect to the adaptive optimal solution. We further show that their approximation ratio is at most $\frac{e^2+1}{(e+1)^2} \approx 0.606$, which is strictly less than $1 - 1/e \approx 0.632$. To do so, we first present an example for non-adaptive greedy with approximation ratio at most $\frac{e^2+1}{(e+1)^2}$. Next, we show that myopic feedback does not help much to adaptive greedy, in that the approximation ratio for the non-adaptive greedy algorithm is no worse than adaptive greedy, when considering over all graphs. Combining with the first observation, we also achieve the result for the adaptive greedy algorithm.

Theorem 4. *The approximation ratio for greedy and adaptive greedy is no better than $\frac{e^2+1}{(e+1)^2} \approx 0.606$, which is strictly less than $1 - 1/e \approx 0.632$. Moreover, the approximation ratio of adaptive greedy is at most that of the non-adaptive greedy, when considering all influence graphs.*

5 Conclusion and Future Work

In this paper, we systematically study the adaptive influence maximization problem with myopic feedback under the independent cascade model, and provide constant upper and lower bounds on the adaptivity gap and the approximation ratios of the non-adaptive greedy and adaptive greedy algorithms. There are a number of future directions to continue this line of research. First, there is still a gap between the upper and lower bound results in this paper, and thus how to close this gap is the next challenge. Second, our result suggests that adaptive greedy may not bring much benefit under the myopic feedback model, so are there other adaptive algorithms that could do much better? Third, for the IC model with full-adoption feedback, because the feedback on different seed nodes may be correlated, existing adaptivity gap results in [1, 4] cannot be applied, and thus its adaptivity gap is still open even though it is adaptive submodular. One may also explore beyond the IC model, and study adaptive solutions for other models such as the linear threshold model, general threshold model etc.[9]. Finally, scalable algorithms for adaptive influence maximization is also worth to investigate.

References

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

[2] Arash Asadpour, Hamid Nazerzadeh, and Amin Saberi. Stochastic submodular maximization. In *International Workshop on Internet and Network Economics*, pages 477–489. Springer, 2008.

[3] Ashwinkumar Badanidiyuru, Christos Papadimitriou, Aviad Rubinstein, Lior Seeman, and Yaron Singer. Locally adaptive optimization: Adaptive seeding for monotone submodular functions. In *Proceedings of the twenty-seventh annual ACM-SIAM symposium on Discrete algorithms*, pages 414–429. SIAM, 2016.

[4] Domagoj Bradac, Sahil Singla, and Goran Zuzic. (near) optimal adaptivity gaps for stochastic multi-value probing. *arXiv preprint arXiv:1902.01461*, 2019.

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

[6] 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.org/abs/1003.3967) includes discussions on the myopic feedback model.

[7] Anupam Gupta, Viswanath Nagarajan, and Sahil Singla. Algorithms and adaptivity gaps for stochastic probing. In *Proceedings of the twenty-seventh annual ACM-SIAM symposium on Discrete algorithms*, pages 1731–1747. SIAM, 2016.

[8] Anupam Gupta, Viswanath Nagarajan, and Sahil Singla. Adaptivity gaps for stochastic probing: Submodular and XOS functions. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1688–1702. SIAM, 2017.

[9] David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD*, pages 137–146. ACM, 2003.

[10] 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.

[11] Guillaume Salha, Nikolaos Tziortziotis, and Michalis Vazirgiannis. Adaptive submodular influence maximization with myopic feedback. In *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pages 455–462. IEEE, 2018.

[12] Lior Seeman and Yaron Singer. Adaptive seeding in social networks. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*, pages 459–468. IEEE, 2013.

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

[14] Lichao Sun, Weiran Huang, Philip S Yu, and Wei Chen. Multi-round influence maximization. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 2249–2258. ACM, 2018.

[15] Guangmo Tong. Adaptive influence maximization under general feedback models. *arXiv preprint arXiv:1902.00192*, 2019.

[16] 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.

[17] Jing Yuan and Shao-Jie Tang. No time to observe: Adaptive influence maximization with partial feedback. In *Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI)*, 2017.

Appendix

We include the missing proofs in this appendix. For convenience, we restate the lemmas and theorems that we prove here.

A Missing Proofs of Section 3.1, Adaptivity Upper Bound

Lemma 1. *For any adaptive policy π, and $t \geq 1$, we have*

$$\sigma^t(\pi) = \sum_{i=0}^{k-1} \mathop{\mathbb{E}}_{s \sim \mathcal{P}_i^\pi} \left[\Delta_{f^t}\left(\pi(\psi_s) \mid \psi_s\right)\right], \text{ and } \sigma^t(\mathcal{W}(\pi)) = \sum_{i=0}^{k-1} \mathop{\mathbb{E}}_{s \sim \mathcal{P}_i^\pi} \left[\Delta_{f^t}\left(\pi(\psi_s) \mid \mathrm{dom}(\psi_s)\right)\right].$$

Proof. We first prove the equality on $\sigma^t(\pi)$. Let $V(\pi, \Phi)_{:i}$ (resp. $V(\pi, \Phi)_i$) denote the first i nodes (resp. the ith node) selected by policy π under realization Φ.

Then we have

$$\sum_{i=0}^{k-1} \mathop{\mathbb{E}}_{s \sim \mathcal{P}_i^\pi} \left[\Delta_{f^t}\left(\pi(\psi_s) \mid \psi_s\right)\right]$$

$$= \sum_{i=0}^{k-1} \mathop{\mathbb{E}}_{s \sim \mathcal{P}_i^\pi} \left[\mathop{\mathbb{E}}_{\Phi^1,\cdots,\Phi^t \sim \mathcal{P}} \left[f^t\left(\mathrm{dom}(\psi_s) \cup \pi(\psi_s), \Phi^1, \cdots, \Phi^t\right) - f^t\left(\mathrm{dom}(\psi_s), \Phi^1, \cdots, \Phi^t\right) \mid \Phi^1 \sim \psi_s\right]\right]$$

$$= \sum_{i=0}^{k-1} \mathop{\mathbb{E}}_{\Phi^2,\cdots,\Phi^t \sim \mathcal{P}} \left[\mathop{\mathbb{E}}_{s \sim \mathcal{P}_i^\pi} \left[\mathop{\mathbb{E}}_{\Phi^1 \sim \mathcal{P}} \left[f^t\left(\mathrm{dom}(\psi_s) \cup \pi(\psi_s), \Phi^1, \cdots, \Phi^t\right) - f^t\left(\mathrm{dom}(\psi_s), \Phi^1, \cdots, \Phi^t\right) \mid \Phi^1 \sim \psi_s\right]\right]\right]$$

$$= \sum_{i=0}^{k-1} \mathop{\mathbb{E}}_{\Phi^2,\cdots,\Phi^t \sim \mathcal{P}} \left[\mathop{\mathbb{E}}_{\Phi^1 \sim \mathcal{P}} \left[\left(f^t\left(V(\pi, \Phi^1)_{:i} \cup V(\pi, \Phi^1)_{i+1}, \Phi^1, \cdots, \Phi^t\right) - f^t\left(V(\pi, \Phi^1)_{:i}, \Phi^1, \cdots, \Phi^t\right)\right)\right]\right]$$

$$= \mathop{\mathbb{E}}_{\Phi^2,\cdots,\Phi^t \sim \mathcal{P}} \left[\mathop{\mathbb{E}}_{\Phi^1 \sim \mathcal{P}} \left[\sum_{i=0}^{k-1} \left(f^t\left(V(\pi, \Phi^1)_{:i} \cup V(\pi, \Phi^1)_{i+1}, \Phi^1, \cdots, \Phi^t\right) - f^t\left(V(\pi, \Phi^1)_{:i}, \Phi^1, \cdots, \Phi^t\right)\right)\right]\right]$$

$$= \mathop{\mathbb{E}}_{\Phi^2,\cdots,\Phi^t \sim \mathcal{P}} \left[\mathop{\mathbb{E}}_{\Phi^1 \sim \mathcal{P}} \left[f^t(V(\pi, \Phi^1), \Phi^1, \cdots, \Phi^t)\right]\right]$$

$$= \sigma^t(\pi)$$

The third equality above is by the law of total expectation, and notice that for any tree node s in $\mathcal{T}(\pi)$ and any random realization $\Phi \sim \psi_s$, we have $V(\pi, \Phi)_{:i} = \mathrm{dom}(\psi_s)$ and $V(\pi, \Phi)_{i+1} = \pi(\psi_s)$.

Next, we prove the equality on $\sigma^t(\mathcal{W}(\pi))$.

$$\sum_{i=0}^{k-1} \mathop{\mathbb{E}}_{s \sim \mathcal{P}_i^\pi} \left[\Delta_{f^t}\left(\pi(\psi_s) \mid \mathrm{dom}(\psi_s)\right)\right]$$

$$= \sum_{i=0}^{k-1} \mathop{\mathbb{E}}_{s \sim \mathcal{P}_i^\pi} \left[\mathop{\mathbb{E}}_{\Phi^1,\cdots,\Phi^t \sim \mathcal{P}} \left[f^t\left(\mathrm{dom}(\psi_s) \cup \pi(\psi_s), \Phi^1, \cdots, \Phi^t\right) - f^t\left(\mathrm{dom}(\psi_s), \Phi^1, \cdots, \Phi^t\right)\right]\right]$$

$$= \sum_{i=0}^{k-1} \mathop{\mathbb{E}}_{\Phi^1,\cdots,\Phi^t \sim \mathcal{P}} \left[\mathop{\mathbb{E}}_{s \sim \mathcal{P}_i^\pi} \left[f^t\left(\mathrm{dom}(\psi_s) \cup \pi(\psi_s), \Phi^1, \cdots, \Phi^t\right) - f^t\left(\mathrm{dom}(\psi_s), \Phi^1, \cdots, \Phi^t\right)\right]\right]$$

$$= \sum_{i=0}^{k-1} \mathop{\mathbb{E}}_{\Phi^1,\cdots,\Phi^t \sim \mathcal{P}} \left[\mathop{\mathbb{E}}_{\Phi \sim \mathcal{P}} \left[f^t\left(V(\pi, \Phi)_{:i} \cup V(\pi, \Phi)_{i+1}, \Phi^1, \cdots, \Phi^t\right) - f^t\left(V(\pi, \Phi)_{:i}, \Phi^1, \cdots, \Phi^t\right)\right]\right]$$

$$= \mathop{\mathbb{E}}_{\Phi \sim \mathcal{P}} \left[\mathop{\mathbb{E}}_{\Phi^1,\cdots,\Phi^t \sim \mathcal{P}} \left[\sum_{i=0}^{k-1} \left(f^t\left(V(\pi, \Phi)_{:i} \cup V(\pi, \Phi)_{i+1}, \Phi^1, \cdots, \Phi^t\right) - f^t\left(V(\pi, \Phi)_{:i}, \Phi^1, \cdots, \Phi^t\right)\right)\right]\right]$$

$$= \mathop{\mathbb{E}}_{\Phi \sim \mathcal{P}} \left[\mathop{\mathbb{E}}_{\Phi^1,\cdots,\Phi^t \sim \mathcal{P}} \left[f^t(V(\pi, \Phi), \Phi^1, \cdots, \Phi^t)\right]\right]$$

$$= \mathop{\mathbb{E}}_{\Phi \sim \mathcal{P}} \left[\sigma^t(V(\pi, \Phi)) \right]$$

$$= \sigma^t(\mathcal{W}(\pi)).$$

The third equality above is because the distribution of $\mathrm{dom}(\psi_s)$ and $\pi(\psi_s)$ with $s \sim \mathcal{P}_i^\pi$ is exactly the same as the distribution of $V(\pi, \Phi)_{:i}$ and $V(\pi, \Phi)_{i+1}$ with $\Phi \sim \mathcal{P}$. Note that this Φ is independent of $\Phi^1, \cdots, \Phi^t$. The last equality is because the distribution of $V(\pi, \Phi)$ with $\Phi \sim \mathcal{P}$ is exactly the distribution of the seed sets taken from the leaves of $\mathcal{T}(\pi)$, which exactly corresponds to the random-walk non-adaptive policy $\mathcal{W}(\pi)$. $\qquad\square$

Lemma 2. *Let* $S = \mathrm{dom}(\psi^1)$ *and* $S^+ = S \cup \{u\}$ *for any partial realization* ψ^1 *and any* $u \notin \mathrm{dom}(\psi^1)$. *Then we have*

$$\mathop{\mathbb{E}}_{\Phi^1, \Phi^2, \Phi^3 \sim \mathcal{P}} \left[f\left(S^+, (\Phi_S^1 \cup \Phi_S^2 \cup \Phi_S^3, \Phi_u^1 \cup \Phi_u^2, \Phi_{V \setminus S^+}^1) \right) \right.$$
$$\left. - f\left(S, (\Phi_S^1 \cup \Phi_S^2 \cup \Phi_S^3, \Phi_{V \setminus S}^1) \right) \mid \Phi^1 \sim \psi^1 \right] \le \Delta_{f^2}(u \mid S). \tag{6}$$

$$\mathop{\mathbb{E}}_{\Phi^1, \Phi^2, \Phi^3 \sim \mathcal{P}} \left[f\left(S^+, (\Phi_S^1 \cup \Phi_S^2 \cup \Phi_S^3, \Phi_u^1 \cup \Phi_u^2 \cup \Phi_u^3, \Phi_{V \setminus S^+}^1) \right) \right.$$
$$\left. - f\left(S^+, (\Phi_S^1 \cup \Phi_S^2 \cup \Phi_S^3, \Phi_u^1 \cup \Phi_u^2, \Phi_{V \setminus S^+}^1) \right) \mid \Phi^1 \sim \psi^1 \right] \le \Delta_{f^2}(u \mid S). \tag{7}$$

Proof. We first prove Inequality (6). To do so, we first expand the RHS of Eq. (6),

$$\Delta_{f^2}(u \mid S) = \mathop{\mathbb{E}}_{\Phi^2, \Phi^3 \sim \mathcal{P}} \left[f^2\left(S^+, \Phi^2, \Phi^3 \right) - f^2\left(S, \Phi^2, \Phi^3 \right) \right]$$

$$= \mathop{\mathbb{E}}_{\Phi^2, \Phi^3 \sim \mathcal{P}} \left[f\left(S^+, (\Phi_S^2 \cup \Phi_S^3, \Phi_u^2 \cup \Phi_u^3, \Phi_{V \setminus S^+}^2) \right) - f\left(S, (\Phi_S^2 \cup \Phi_S^3, \Phi_u^2, \Phi_{V \setminus S^+}^2) \right) \right]$$

$$= \mathop{\mathbb{E}}_{\Phi_S^2, \Phi_S^3 \sim \mathcal{P}_S} \left[\mathop{\mathbb{E}}_{\Phi_u^2, \Phi_u^3 \in \mathcal{P}_u} \left[\mathop{\mathbb{E}}_{\Phi_{V \setminus S^+}^2 \sim \mathcal{P}_{V \setminus S^+}} \left[f\left(S^+, (\Phi_S^2 \cup \Phi_S^3, \Phi_u^2 \cup \Phi_u^3, \Phi_{V \setminus S^+}^2) \right) - \right. \right. \right.$$
$$\left. \left. \left. f\left(S, (\Phi_S^2 \cup \Phi_S^3, \Phi_u^2, \Phi_{V \setminus S^+}^2) \right) \right] \right] \right]. \tag{9}$$

The third equality above holds because $\Phi_S^2, \Phi_S^3, \Phi_u^2, \Phi_u^3, \Phi_{V \setminus S^+}^2, \Phi_{V \setminus S^+}^3$ are mutually independent, and $\Phi_{V \setminus S^+}^3$ does not appear inside the expectation term. Next, we expand the LHS of Eq. (6),

LHS of Eq. (6)

$$= \mathop{\mathbb{E}}_{\Phi_S^1, \Phi_S^2, \Phi_S^3 \sim \mathcal{P}_S} \left[\mathop{\mathbb{E}}_{\Phi_u^1, \Phi_u^2 \in \mathcal{P}_u} \left[\mathop{\mathbb{E}}_{\Phi_{V \setminus S^+}^1 \sim \mathcal{P}_{V \setminus S^+}} \left[f\left(S^+, (\Phi_S^1 \cup \Phi_S^2 \cup \Phi_S^3, \Phi_u^1 \cup \Phi_u^2, \Phi_{V \setminus S^+}^1) \right) \right. \right. \right.$$
$$\left. \left. \left. - f\left(S, (\Phi_S^1 \cup \Phi_S^2 \cup \Phi_S^3, \Phi_u^1, \Phi_{V \setminus S^+}^1) \right) \mid \Phi^1 \sim \psi^1 \right] \right] \right]$$

$$= \mathop{\mathbb{E}}_{\Phi_S^2, \Phi_S^3 \sim \mathcal{P}_S} \left[\mathop{\mathbb{E}}_{\Phi_u^1, \Phi_u^2 \in \mathcal{P}_u} \left[\mathop{\mathbb{E}}_{\Phi_{V \setminus S^+}^1 \sim \mathcal{P}_{V \setminus S^+}} \left[f\left(S^+, (\psi^1 \cup \Phi_S^2 \cup \Phi_S^3, \Phi_u^1 \cup \Phi_u^2, \Phi_{V \setminus S^+}^1) \right) \right. \right. \right.$$
$$\left. \left. \left. - f\left(S, (\psi^1 \cup \Phi_S^2 \cup \Phi_S^3, \Phi_u^1, \Phi_{V \setminus S^+}^1) \right) \right] \right] \right].$$

$$= \mathop{\mathbb{E}}_{\Phi_S^2, \Phi_S^3 \sim \mathcal{P}_S} \left[\mathop{\mathbb{E}}_{\Phi_u^1, \Phi_u^3 \in \mathcal{P}_u} \left[\mathop{\mathbb{E}}_{\Phi_{V \setminus S^+}^2 \sim \mathcal{P}_{V \setminus S^+}} \left[f\left(S^+, (\psi^1 \cup \Phi_S^2 \cup \Phi_S^3, \Phi_u^1 \cup \Phi_u^3, \Phi_{V \setminus S^+}^2) \right) \right. \right. \right.$$
$$\left. \left. \left. - f\left(S, (\psi^1 \cup \Phi_S^2 \cup \Phi_S^3, \Phi_u^1, \Phi_{V \setminus S^+}^2) \right) \right] \right] \right].$$

$$= \mathop{\mathbb{E}}_{\Phi_S^2, \Phi_S^3 \sim \mathcal{P}_S} \left[\mathop{\mathbb{E}}_{\Phi_u^2, \Phi_u^3 \in \mathcal{P}_u} \left[\mathop{\mathbb{E}}_{\Phi_{V \setminus S^+}^2 \sim \mathcal{P}_{V \setminus S^+}} \left[f\left(S^+, (\psi^1 \cup \Phi_S^2 \cup \Phi_S^3, \Phi_u^2 \cup \Phi_u^3, \Phi_{V \setminus S^+}^2) \right) \right. \right. \right.$$

11

$$- f\left(S, (\psi^1 \cup \Phi_S^2 \cup \Phi_S^3, \Phi_u^2, \Phi_{V \setminus S^+}^2)\right)\Big]\Big]\Big]. \tag{10}$$

The first equality above holds because all these random variables are independent. The second equality above holds because $\Phi_S^1 = \psi^1$ implied by $\Phi^1 \sim \psi^1$. In the third equality, we replace $\Phi_{V \setminus S^+}^1$ with $\Phi_{V \setminus S^+}^2$ and replace Φ_u^2 with Φ_u^3, because they follow the same probability distributions and are independent to the other distributions. In the last equality, we replace Φ_u^1 with Φ_u^2.

Comparing Eq. (9) and Eq. (10), we know that it suffices to prove that for any fixed partial realizations $\phi_S^2, \phi_S^3, \phi_u^2, \phi_u^3, \phi_{V \setminus S}^2$,

$$f\left(S^+, (\psi^1 \cup \phi_S^2 \cup \phi_S^3, \phi_u^2 \cup \phi_u^3, \phi_{V \setminus S^+}^2)\right) - f\left(S, (\psi^1 \cup \phi_S^2 \cup \phi_S^3, \phi_u^2, \phi_{V \setminus S^+}^2)\right)$$

$$\leq f\left(S^+, (\phi_S^2 \cup \phi_S^3, \phi_u^2 \cup \phi_u^3, \phi_{V \setminus S^+}^2)\right) - f\left(S, (\phi_S^2 \cup \phi_S^3, \phi_u^2, \phi_{V \setminus S^+}^2)\right). \tag{11}$$

Consider any node $v \in \Gamma(S^+, (\psi^1 \cup \phi_S^2 \cup \phi_S^3, \phi_u^2 \cup \phi_u^3, \phi_{V \setminus S^+}^2)) \setminus \Gamma(S, (\psi^1 \cup \phi_S^2 \cup \phi_S^3, \phi_u^2, \phi_{V \setminus S^+}^2))$, we have the following observations: (1) under the realization $(\psi^1 \cup \phi_S^2 \cup \phi_S^3, \phi_u^2, \phi_{V \setminus S^+}^2)$ (or equivalently its live-edge graph), node v cannot be reached from nodes in S; and (2) under the realization $(\psi^1 \cup \phi_S^2 \cup \phi_S^3, \phi_u^2 \cup \phi_u^3, \phi_{V \setminus S^+}^2)$ (or equivalently its live-edge graph), node v can be reached via a path P originated from node u, and P does not contain any node in S.

Now, we are going to prove that $v \in \Gamma(S^+, (\phi_S^2 \cup \phi_S^3, \phi_u^2 \cup \phi_u^3, \phi_{V \setminus S^+}^2)) \setminus \Gamma(S, (\phi_S^2 \cup \phi_S^3, \phi_u^2, \phi_{V \setminus S^+}^2))$. Since the path P does not contain any node in S, we know that path P also exists under the realization $(\phi_S^2 \cup \phi_S^3, \phi_u^2 \cup \phi_u^3, \phi_{V \setminus S^+}^2)$, i.e., node v can be reached from node u under realization $(\phi_S^2 \cup \phi_S^3, \phi_u^2 \cup \phi_u^3, \phi_{V \setminus S^+}^2)$. Moreover, we know that the realization $((\phi_S^2 \cup \phi_S^3, \phi_u^2, \phi_{V \setminus S^+}^2)$ has less live edges than the realization $(\psi^1 \cup \phi_S^2 \cup \phi_S^3, \phi_u^2, \phi_{V \setminus S^+}^2)$, so node v can not be reached from set S under the realization $(\phi_S^2 \cup \phi_S^3, \phi_u^2, \phi_{V \setminus S^+}^2)$. As a result, we have proved

$$\Gamma\left(S^+, (\psi^1 \cup \phi_S^2 \cup \phi_S^3, \phi_u^2 \cup \phi_u^3, \phi_{V \setminus S^+}^2)\right) \setminus \Gamma\left(S, (\psi^1 \cup \phi_S^2 \cup \phi_S^3, \phi_u^2, \phi_{V \setminus S^+}^2)\right)$$

$$\subseteq \Gamma\left(S^+, (\phi_S^2 \cup \phi_S^3, \phi_u^2 \cup \phi_u^3, \phi_{V \setminus S^+}^2)\right) \setminus \Gamma\left(S, (\phi_S^2 \cup \phi_S^3, \phi_u^2, \phi_{V \setminus S^+}^2)\right). \tag{12}$$

This proves Eq. (11) and thus concludes the proof of Inequality (6). Note that the above proof on Eq. (11) resembles the proof of submodularity of influence utility function f on a live-edge graph, but Eq. (11) is a bit more complicated because it is on different live-edge graphs.

Next we prove the Inequality (7). Again, we first expand the RHS of Eq. (7).

$$\Delta_{f^2}(u \mid S) = \mathop{\mathbb{E}}_{\Phi_S^2, \Phi_S^3 \sim \mathcal{P}_S} \left[\mathop{\mathbb{E}}_{\Phi_u^2, \Phi_u^3 \in \mathcal{P}_u} \left[\mathop{\mathbb{E}}_{\Phi_{V \setminus S^+}^2 \sim \mathcal{P}_{V \setminus S^+}} \left[f\left(S^+, (\Phi_S^2 \cup \Phi_S^3, \Phi_u^2 \cup \Phi_u^3, \Phi_{V \setminus S^+}^2)\right) \right. \right. \right.$$

$$\left. \left. \left. - f\left(S, (\Phi_S^2 \cup \Phi_S^3, \Phi_u^2, \Phi_{V \setminus S^+}^2)\right) \right]\right]\right]$$

$$\geq \mathop{\mathbb{E}}_{\Phi_S^2, \Phi_S^3 \sim \mathcal{P}_S} \left[\mathop{\mathbb{E}}_{\Phi_u^2, \Phi_u^3 \in \mathcal{P}_u} \left[\mathop{\mathbb{E}}_{\Phi_{V \setminus S^+}^2 \sim \mathcal{P}_{V \setminus S^+}} \left[f\left(S^+, (\Phi_S^2 \cup \Phi_S^3, \Phi_u^2 \cup \Phi_u^3, \Phi_{V \setminus S^+}^2)\right) \right. \right. \right.$$

$$\left. \left. \left. - f\left(S^+, (\Phi_S^2 \cup \Phi_S^3, \Phi_u^2, \Phi_{V \setminus S^+}^2)\right) \right]\right]\right]. \tag{13}$$

The inequality above is by the monotonicity of $f(S, \phi)$ on S. Next, we expand the LHS of Eq. (7). LHS of Eq. (7)

$$= \mathop{\mathbb{E}}_{\Phi_S^1, \Phi_S^2, \Phi_S^3 \sim \mathcal{P}_S} \left[\mathop{\mathbb{E}}_{\Phi_u^1, \Phi_u^2, \Phi_u^3 \in \mathcal{P}_u} \left[\mathop{\mathbb{E}}_{\Phi_{V \setminus S^+}^1 \sim \mathcal{P}_{V \setminus S^+}} \left[f\left(S^+, (\Phi_S^1 \cup \Phi_S^2 \cup \Phi_S^3, \Phi_u^1 \cup \Phi_u^2 \cup \Phi_u^3, \Phi_{V \setminus S^+}^1)\right) \right. \right. \right.$$

$$\left. \left. \left. - f\left(S^+, (\Phi_S^1 \cup \Phi_S^2 \cup \Phi_S^3, \Phi_u^1 \cup \Phi_u^2, \Phi_{V \setminus S^+}^1)\right) \mid \Phi^1 \sim \psi^1 \right]\right]\right]$$

$$= \mathop{\mathbb{E}}_{\Phi_S^2, \Phi_S^3 \sim \mathcal{P}_S} \left[\mathop{\mathbb{E}}_{\Phi_u^1, \Phi_u^2, \Phi_u^3 \in \mathcal{P}_u} \left[\mathop{\mathbb{E}}_{\Phi_{V \setminus S^+}^1 \sim \mathcal{P}_{V \setminus S^+}} \left[f\left(S^+, (\psi^1 \cup \Phi_S^2 \cup \Phi_S^3, \Phi_u^1 \cup \Phi_u^2 \cup \Phi_u^3, \Phi_{V \setminus S^+}^1)\right) \right. \right. \right.$$

$$- f\left(S^+, (\psi^1 \cup \Phi_S^2 \cup \Phi_S^3, \Phi_u^1 \cup \Phi_u^2, \Phi_{V\setminus S^+}^1)\right)\Big]\Big]\Big].$$

$$= \mathop{\mathbb{E}}_{\Phi_S^2, \Phi_S^3 \sim \mathcal{P}_S} \left[\mathop{\mathbb{E}}_{\Phi_u^1, \Phi_u^2, \Phi_u^3 \in \mathcal{P}_u} \left[\mathop{\mathbb{E}}_{\Phi_{V\setminus S^+}^2 \sim \mathcal{P}_{V\setminus S^+}} \left[f\left(S^+, (\psi^1 \cup \Phi_S^2 \cup \Phi_S^3, \Phi_u^1 \cup \Phi_u^2 \cup \Phi_u^3, \Phi_{V\setminus S^+}^2)\right) \right. \right. \right.$$

$$\left. \left. \left. - f\left(S^+, (\psi^1 \cup \Phi_S^2 \cup \Phi_S^3, \Phi_u^1 \cup \Phi_u^2, \Phi_{V\setminus S^+}^2)\right)\right]\Big]\Big]. \right. \tag{14}$$

The last equality holds by replacing $\Phi_{V\setminus S^+}^1$ with $\Phi_{V\setminus S^+}^2$, because both have the same distributions and are independent from the other distributions. Similar to the previous lemma, comparing Eq. (13) and Eq. (14), it suffices to prove that for fixed partial realizations $\phi_S^2, \phi_S^3, \phi_u^1, \phi_u^2, \phi_u^3$ and $\phi_{V\setminus S^+}^2$,

$$f\left(S^+, (\psi^1 \cup \phi_S^2 \cup \phi_S^3, \phi_u^1 \cup \phi_u^2 \cup \phi_u^3, \phi_{V\setminus S^+}^2)\right) - f\left(S^+, (\psi^1 \cup \phi_S^2 \cup \phi_S^3, \phi_u^1 \cup \phi_u^2, \phi_{V\setminus S^+}^2)\right)$$

$$\leq f\left(S^+, (\phi_S^2 \cup \phi_S^3, \phi_u^2 \cup \phi_u^3, \phi_{V\setminus S^+}^2)\right) - f\left(S^+, (\phi_S^2 \cup \phi_S^3, \phi_u^2, \phi_{V\setminus S^+}^2)\right). \tag{15}$$

Consider any node $v \in \Gamma(S^+, (\psi^1 \cup \phi_S^2 \cup \phi_S^3, \phi_u^1 \cup \phi_u^2 \cup \phi_u^3, \phi_{V\setminus S^+}^2)) \setminus \Gamma(S^+, (\psi^1 \cup \phi_S^2 \cup \phi_S^3, \phi_u^1 \cup \phi_u^2, \phi_{V\setminus S^+}^2))$, we have the following observations: (1) Node v cannot be reached from any node in set S^+ under the realization $(\psi^1 \cup \phi_S^2 \cup \phi_S^3, \phi_u^1 \cup \phi_u^2, \phi_{V\setminus S^+}^2)$; and (2) node v can be reached via a simple path P originated from node u under the realization $(\psi^1 \cup \phi_S^2 \cup \phi_S^3, \phi_u^1 \cup \phi_u^2 \cup \phi_u^3, \phi_{V\setminus S^+}^2)$, and P does not contain any node in S and any edge in $\phi_u^1 \cup \phi_u^2$.

Now, we prove that $v \in \Gamma(S^+, (\phi_S^2 \cup \phi_S^3, \phi_u^2 \cup \phi_u^3, \phi_{V\setminus S^+}^2)) \setminus \Gamma(S^+, (\phi_S^2 \cup \phi_S^3, \phi_u^2, \phi_{V\setminus S^+}^2))$. Since path P does not contain any node in S and any edge in ϕ_u^1, we know that path P also exists under realization $(\phi_S^2 \cup \phi_S^3, \phi_u^2 \cup \phi_u^3, \phi_{V\setminus S^+}^2)$, i.e., node v can be reached from node u under realization $(\phi_S^2 \cup \phi_S^3, \phi_u^2 \cup \phi_u^3, \phi_{V\setminus S^+}^2)$. Moreover, we know that the realization $(\phi_S^2 \cup \phi_S^3, \phi_u^2, \phi_{V\setminus S^+}^2)$ has less live edges than the realization $(\psi^1 \cup \phi_S^2 \cup \phi_S^3, \phi_u^1 \cup \phi_u^2, \phi_{V\setminus S^+}^2)$, thus node v cannot be reached from the set S^+ under realization $(\phi_S^2 \cup \phi_S^3, \phi_u^2, \phi_{V\setminus S^+}^2)$. Thus we can conclude that $v \in \Gamma(S^+, (\phi_S^2 \cup \phi_S^3, \phi_u^2 \cup \phi_u^3, \phi_{V\setminus S^+}^2)) \setminus \Gamma(S^+, (\phi_S^2 \cup \phi_S^3, \phi_u^2, \phi_{V\setminus S^+}^2))$, this leads to Eq. (15) and concludes the proof of Inequality (7). $\square$

Lemma 3. *For any partial realization ψ^1 and node $u \notin \text{dom}(\psi^1)$, we have*

$$\Delta_{f^3}(u \mid \psi^1) \leq 2\Delta_{f^2}(u \mid \text{dom}(\psi^1)). \tag{8}$$

Proof. Again, for ease of notation, we set $S = \text{dom}(\psi^1)$ and $S^+ = \text{dom}(\psi^1) \cup \{u\}$, then we have

$$\Delta_{f^3}(u \mid \psi^1) = \mathop{\mathbb{E}}_{\Phi^1, \Phi^2, \Phi^3 \sim \mathcal{P}} \left[f\left(S^+, (\Phi_S^1 \cup \Phi_S^2 \cup \Phi_S^3, \Phi_u^1 \cup \Phi_u^2 \cup \Phi_u^3, \Phi_{V\setminus S^+}^1)\right) \right.$$

$$\left. - f\left(S, (\Phi_S^1 \cup \Phi_S^2 \cup \Phi_S^3, \Phi_{V\setminus S}^1)\right) \mid \Phi^1 \sim \psi^1 \right]$$

$$= \mathop{\mathbb{E}}_{\Phi^1, \Phi^2, \Phi^3 \sim \mathcal{P}} \left[f\left(S^+, (\Phi_S^1 \cup \Phi_S^2 \cup \Phi_S^3, \Phi_u^1 \cup \Phi_u^2 \cup \Phi_u^3, \Phi_{V\setminus S^+}^1)\right) \right.$$

$$\left. - f\left(S^+, (\Phi_S^1 \cup \Phi_S^2 \cup \Phi_S^3, \Phi_u^1 \cup \Phi_u^2, \Phi_{V\setminus S^+}^1)\right) \mid \Phi^1 \sim \psi^1 \right]$$

$$+ \mathop{\mathbb{E}}_{\Phi^1, \Phi^2, \Phi^3 \sim \mathcal{P}} \left[f\left(S^+, (\Phi_S^1 \cup \Phi_S^2 \cup \Phi_S^3, \Phi_u^1 \cup \Phi_u^2, \Phi_{V\setminus S^+}^1)\right) - f\left(S, (\Phi_S^1 \cup \Phi_S^2 \cup \Phi_S^3, \Phi_{V\setminus S}^1)\right) \mid \Phi^1 \sim \psi^1 \right]$$

$$\leq \Delta_{f^2}(u|S) + \Delta_{f^2}(u|S) = 2\Delta_{f^2}(u|\text{dom}(\psi^1)). \tag{16}$$

The inequality above is a direct consequence of Lemmas 2. $\square$

Lemma 4. *For any $t \geq 1$ and any subset $S \subseteq V$, $\sigma^t(S) \leq t \cdot \sigma(S)$.*

Proof. We have

$$\sigma^t(S) = \mathop{\mathbb{E}}_{\Phi^1, \cdots, \Phi^t \sim \mathcal{P}} \left[f^t(S, \Phi^1, \cdots, \Phi^t) \right] = \mathop{\mathbb{E}}_{\Phi^1, \cdots, \Phi^t \sim \mathcal{P}} \left[f\left(S, (\cup_{i \in [t]} \Phi_S^i, \Phi_{V\setminus S}^1)\right) \right]$$

$$= \underset{\Phi^1_{V \setminus S} \sim \mathcal{P}_{V \setminus S}}{\mathbb{E}} \left[\underset{\Phi^1_S, \cdots, \Phi^t_S \sim \mathcal{P}_S}{\mathbb{E}} \left[f\left(S, (\cup_{i \in [t]} \Phi^i_S, \Phi^1_{V \setminus S}) \right) \right] \right]. \tag{17}$$

We want to show that for any fixed $\phi^1_{V \setminus S}$,

$$\underset{\Phi^1_S, \cdots, \Phi^t_S \sim \mathcal{P}_S}{\mathbb{E}} \left[f\left(S, (\cup_{i \in [t]} \Phi^i_S, \phi^1_{V \setminus S}) \right) \right] \leq \sum_{i \in [t]} \underset{\Phi^i_S}{\mathbb{E}} \left[f\left(S, (\Phi^i_S, \phi^1_{V \setminus S}) \right) \right]. \tag{18}$$

Once Eq.(18) is shown, we can combine with Eq.(17) to obtain

$$\sigma^t(S) \leq \underset{\Phi^1_{V \setminus S} \sim \mathcal{P}}{\mathbb{E}} \left[\sum_{i \in [t]} \underset{\Phi^i_S}{\mathbb{E}} \left[f\left(S, (\Phi^i_S, \Phi^1_{V \setminus S}) \right) \right] \right]$$

$$= \sum_{i \in [t]} \underset{\Phi^1_{V \setminus S} \sim \mathcal{P}}{\mathbb{E}} \left[\underset{\Phi^i_S}{\mathbb{E}} \left[f\left(S, (\Phi^i_S, \Phi^1_{V \setminus S}) \right) \right] \right]$$

$$= \sum_{i \in [t]} \underset{\Phi^1 \sim \mathcal{P}}{\mathbb{E}} \left[f(S, \Phi^1) \right] = t \cdot \sigma(S).$$

Thus the lemma holds. Now we prove Inequality (18). To do so, we fix partial realizations $\phi^1_S, \cdots, \phi^t_S$. If node $v \in \Gamma(S, \cup_{i \in [t]} \phi^i_S, \phi^1_{V \setminus S}))$, then we conclude that under the realization $(\cup_{i \in [t]} \phi^i_S, \phi^1_{V \setminus S})$, node v can be reached via a path P originated from some node $u \in S$, and only the starting node of P is in S and all remaining nodes in P are not from S. Suppose in path P, the edge leaving node u is contained in edge set ϕ^i_u for some $i \in [t]$. Then we conclude that node $v \in \Gamma(S, (\phi^i_S, \phi^1_{V \setminus S}))$, since the path P exists under the realization $(\phi^i_S, \phi^1_{V \setminus S})$. This shows that $\Gamma(S, (\cup_{i \in [t]} \phi^i_S, \phi^1_{V \setminus S}) \subseteq \cup_{i \in [t]} \Gamma(S, (\phi^i_S, \phi^1_{V \setminus S}))$, which is sufficient to prove Inequality (18). $\qquad\square$

B Missing Proof of Section 3.2, Adaptivity Lower Bound

Theorem 2. *Under the IC model with myopic feedback, the adaptivity gap for the influence maximization problem is at least $e/(e-1)$.*

Proof. Consider the following construction for the influence graph: the influence graph $G = (L, R, E, p)$ is a bipartite graph with $|L| = \binom{m^3}{m^2}$ and $|R| = m^3$. All edges $(u, v) \in E$ are directed from the left part L to the right part R, associated with probability $1/m$. More specifically, for any subset $X \subseteq R$ with size m^2, there is a node $u_X \in L$ such that the outgoing edges of u_X are exactly (u_X, v) for every $v \in X$. Thus the out-degree of every vertex in L is m^2.

We first describe the main idea of the proof. The budget for the IM problem is m^2, i.e., we are allowed to select no more than m^2 seeds, and we would consider m to be a very large number here. Intuitively, the expected number of nodes in R that is reachable for a single node $u \in L$ is $m^2 \cdot (1/m) = m$, and the influence spread is concentrated on its expected value for large m. In an adaptive solution, we could always make the expected marginal gain for the node we select equals the expected influence spread of a single node in L, by selecting nodes in L such that none of its out-neighbors has been reached so far, unless there are too few nodes in R that are not reachable. Since $m^2 \cdot m = m^3$, the seeds we select would reach almost all but except $o(m^3)$ nodes in R, thus the influence spread of the adaptive policy is roughly m^3. While for a non-adaptive policy, it can select at most m^2 nodes from L and for each node in R, on average, it is connected with at most $m^2 \cdot m^2/m^3 = m$ seeds in L, we can easily prove that it is indeed the best allocation of seeds in L, and the expected probability for nodes in R to be reached is $1 - (1 - 1/m)^m \approx 1 - 1/e$. Moreover, since we are allowed to select no more than m^2 seeds in R and they would not reach any other node, the contribution of this part is negligible. Thus the expected influence spread for the optimal non-adaptive solution would not exceed $(1 - 1/e)m^3$ and the adaptivity gap is $e/(e-1)$ on this graph.

The following two claims would make the above intuition formal.

Claim 2. *For any $\epsilon > 0$, when m is large enough, we have $\mathrm{OPT}_A(G, m^2) \geq (1 - \epsilon)m^3$.*

Proof. For any fixed $\epsilon > 0$, we would take m such that $m \geq 48/\epsilon^2 \log m$. Consider the following adaptive policy π, which only selects nodes from the left part L. Moreover, for every node $u \in L$ selected by π, at the time of selection, none of u's out-neighbors in R has been reached yet from nodes selected by π so far (this condition can be verified by an adaptive policy with myopic feedback). When there does not exist such node or the size of the seed set already equals to the budget, π would stop. For $i \in \{1, \cdots, (1 - \epsilon/2)m^2\}$, let $\mathcal{E}_i$ denote the event that after selecting the i-th seed in L, the marginal gain of the influence spread is between $[(1 - \epsilon/2)m + 1, (1 + \epsilon/2)m + 1]$. We would give a lower bound on the conditional probability $Pr[\mathcal{E}_i \mid \mathcal{E}_1, \cdots, \mathcal{E}_{i-1}]$. Under the condition $\cup_{j=1}^{i-1} \mathcal{E}_j$, the current influence spread on the right part R is less than $(1 + \epsilon/2)m \cdot (1 - \epsilon/2)m^2 = (1 - \epsilon^2/4)m^3 < m^3 - m^2$, thus policy π would not stop by now. Thus the marginal gain is the summation of m^2 independent binomial variables with mean m. By the Chernoff bound we have

$$Pr[\mathcal{E}_i \mid \mathcal{E}_1, \cdots, \mathcal{E}_{i-1}] \geq 1 - exp(-\epsilon^2 m/12) \geq 1 - \frac{1}{m^3}. \tag{19}$$

Consequently,

$$Pr[\cup_{i=1}^t \mathcal{E}_i] = \Pi_{i=1}^t Pr[\mathcal{E}_i \mid \mathcal{E}_1, \cdots, \mathcal{E}_{i-1}] \geq (1 - \frac{1}{m^3})^{m^2} \geq 1 - \frac{1}{m^3} \cdot m^2 = 1 - \frac{1}{m}. \tag{20}$$

Thus the expected influence is greater than $(1 - \frac{1}{m}) \cdot (1 - \epsilon/2)m \cdot (1 - \epsilon/2)m^2 \geq (1 - \epsilon)m^3$. $\quad\square$

Claim 3. $\text{OPT}_N(G, m^2) \leq (1 - (1 - 1/m)^m)m^3 + 2m^2$.

Proof. Let S_L (resp. S_R) denote the seed set selected by the optimal non-adaptive policy from the left part L (resp. right part R). For any node $u_i \in R$ where $i \in [m^3]$, let x_i denote the number of u_i's in-neighbors in the seed set S_L. Since the out-degree for each node in S_L is m^2, we have $\sum_{i \in [m^3]} x_i \leq |S_L| \cdot m^2$ and the average number of in-neighbors is at most $|S_L| \cdot m^2/m^3 = |S_L|/m$. Furthermore, we can calculate the influence spread of S_L,

$$\sigma(S_L) = |S_L| + \sum_{i \in [m^3]} \Pr[u_i \text{ is reachable}]$$

$$= |S_L| + \sum_{i \in [m^3]} \left(1 - \left(1 - \frac{1}{m} \right)^{x_i} \right)$$

$$\leq |S_L| + m^3 \cdot \left(1 - \left(1 - \frac{1}{m} \right)^{|S_L|/m} \right)$$

$$\leq m^2 + m^3 \cdot \left(1 - \left(1 - \frac{1}{m} \right)^m \right). \tag{21}$$

The first inequality holds because function $g(x) = (1 - (1 - \frac{1}{m})^x)$ is concave. The last inequality holds because $|S_L| \leq m^2$. Now we have

$$\text{OPT}_N(G, m^2) = \max_{\substack{S_L \subseteq L, S_R \subseteq R, \\ |S_L| + |S_R| \leq m^2}} \sigma(S_L \cup S_R) \leq \max_{\substack{S_L \subseteq L, \\ |S_L| \leq m^2}} \sigma(S_L) + \max_{\substack{S_R \subseteq L, \\ |S_R| \leq m^2}} \sigma(S_R)$$

$$\leq m^2 + m^3 \cdot \left(1 - \left(1 - \frac{1}{m} \right)^m \right) + m^2 = \left(1 - \left(1 - \frac{1}{m} \right)^m \right) \cdot m^3 + 2m^2. \tag{22}$$

This concludes the proof. $\quad\square$

Combining Claims 2 and 3, we can conclude that for any $\epsilon > 0$, there exists large enough m such that $\text{OPT}_A(G, m^2)/\text{OPT}_N(G, m^2) \geq e/(e - 1) - \epsilon$. Letting $\epsilon \to 0$, we obtain the theorem. $\quad\square$

C   Missing Proofs in Section 4

For the proofs in this section, let $\text{Greedy}_N(G, k)$ (resp. $\text{Greedy}_A(G, k)$) denote the influence spread for the non-adaptive greedy algorithm (resp. adaptive influence spread for the adaptive greedy algorithm), on the influence graph G with a budget k.

The proof of Theorem 3 is complete once we prove the following lemma.

15

Lemma 5. *Adaptive greedy is $(1 - 1/e)$ approximate to the optimal non-adaptive policy.*

Proof. For a fixed influence graph G, let S ($|S| = k$) denote the seed set selected by the optimal non-adaptive algorithm, where s_i denotes the ith element in set S. We use $\mathcal{A}$ to denote adaptive greedy and for any $t \in \{0, 1, \cdots, k\}$, we use $U(t)$ to denote the expected adaptive influence spread of nodes selected by $\mathcal{A}$ in the first i rounds, i.e.,

$$U(t) := \mathop{\mathbb{E}}_{\Phi \sim \mathcal{P}} \left[f \left(V(\mathcal{A}, \Phi)_{:t}, \Phi \right) \right], \tag{23}$$

From the above definition, we can see that $U(0) = 0$ and $U(k) = \sigma(\mathcal{A})$. By Lemma 1, we have

$$U(t) = \sum_{i=0}^{t-1} \mathop{\mathbb{E}}_{s \sim \mathcal{P}_i^{\mathcal{A}}} \left[\Delta_f \left(\mathcal{A}(\psi_s) \mid \psi_s \right) \right]. \tag{24}$$

Now, for any $t \in \{0, 1, \cdots, k-1\}$

$$
\begin{aligned}
U(t+1) - U(t) &= \mathop{\mathbb{E}}_{s \sim \mathcal{P}_t^{\mathcal{A}}} \left[\Delta_f \left(\mathcal{A}(\psi_s) \mid \psi_s \right) \right] \\
&\geq \frac{1}{k} \sum_{i=1}^{k} \mathop{\mathbb{E}}_{s \sim \mathcal{P}_t^{\mathcal{A}}} \left[\Delta_f \left(s_i \mid \psi_s \right) \right] \\
&= \frac{1}{k} \sum_{i=1}^{k} \mathop{\mathbb{E}}_{s \sim \mathcal{P}_t^{\mathcal{A}}} \left[\mathop{\mathbb{E}}_{\Phi \sim \mathcal{P}} \left[f \left(\mathrm{dom}(\psi_s) \cup \{s_i\}, \Phi \right) - f \left(\mathrm{dom}(\psi_s), \Phi \right) | \Phi \sim \psi_s \right] \right] \\
&= \frac{1}{k} \mathop{\mathbb{E}}_{s \sim \mathcal{P}_t^{\mathcal{A}}} \left[\mathop{\mathbb{E}}_{\Phi \sim \mathcal{P}} \left[\sum_{i=1}^{k} \left(f \left(\mathrm{dom}(\psi_s) \cup \{s_i\}, \Phi \right) - f \left(\mathrm{dom}(\psi_s), \Phi \right) \right) | \Phi \sim \psi_s \right] \right] \\
&\geq \frac{1}{k} \mathop{\mathbb{E}}_{s \sim \mathcal{P}_t^{\mathcal{A}}} \left[\mathop{\mathbb{E}}_{\Phi \sim \mathcal{P}} \left[f(\mathrm{dom}(\psi_s) \cup S, \Phi) - f(\mathrm{dom}(\psi_s), \Phi) | \Phi \sim \psi_s \right] \right] \\
&\geq \frac{1}{k} \mathop{\mathbb{E}}_{s \sim \mathcal{P}_t^{\mathcal{A}}} \left[\mathop{\mathbb{E}}_{\Phi \sim \mathcal{P}} \left[f(S, \Phi) - f(\mathrm{dom}(\psi_s), \Phi) | \Phi \sim \psi_s \right] \right] \\
&= \frac{1}{k} \left(\sigma(S) - U(t) \right). \tag{25}
\end{aligned}
$$

The first inequality holds since adaptive greedy $\mathcal{A}$ chooses the node that maximizes the expected marginal gain, i.e., for any partial realization ψ, $\Delta_f(\mathcal{A}(\psi) \mid \psi) \geq \Delta_f(s_i \mid \psi)$ for any $i \in [k]$. The second inequality is because the influence utility function $f(\cdot, \Phi)$ is submodular under a fixed realization Φ. The third inequality holds because the influence utility function $f(\cdot, \Phi)$ is monotone under a fixed realization Φ. The last equality utilizes the law of total expectation.

Now via standard argument, Eq. (25) implies that

$$
\begin{aligned}
\mathrm{Greedy}_A(G, k) = U(k) &\geq \left(1 - \left(1 - \frac{1}{k} \right)^k \right) \sigma(S) = \left(1 - \left(1 - \frac{1}{k} \right)^k \right) \mathrm{OPT}_N(G, k) \\
&\geq \left(1 - \frac{1}{e} \right) \cdot \mathrm{OPT}_N(G, k). \tag{26}
\end{aligned}
$$

This concludes the proof. $\qquad\square$

We now prove Theorem 4. We first present a example showing that the non-adaptive greedy achieves at most $\frac{e^2+1}{(e+1)^2}$ approximation ratio.

Lemma 6. *Non-adaptive greedy algorithm has ratio at most $\frac{e^2+1}{(e+1)^2}$ with respect to the optimal adaptive solution, in the IC model with myopic feedback.*

Proof. Consider the following influence graph $G(V, E, p)$, where $V = V_1 \bigcup V_2 \bigcup V_3$, $|V_1| = d - 1$, $|V_2| = d$ and $|V_3| = 2d$. We would use v_j^i to denote the jth node in V_i. Nodes in V_1 and V_2 have

16

unit weight while nodes in V_3 have weight w. Note that we could achieve the weight of w by simply replacing each node with a chain of w nodes with edge probability 1, so that as long as the head of the chain is activated, the whole chain is activated. There are directed edges from V_1 to V_2 and from V_2 to V_3. More specifically, for any $j \in [d], l \in [d-1]$, there is a direct edge from the node v_l^1 to the node v_j^2, associated with probability $1/d$. The node v_j^2 is connected to node v_{2j-1}^3 and v_{2j}^3, with probability $e/(e+1)$. The budget $k = \frac{e+3}{e+1}d$. We first consider the optimal adaptive solution and we observe that the optimal adaptive strategy can reach almost all nodes in V_3.

Claim 4. *For any $\epsilon > 0$, if we set $d \geq 2\log(2/\epsilon)/\epsilon^2$, then we have $\mathrm{OPT}_A(G, k) \geq (1 - \epsilon) \cdot 2dw$*

Proof. Consider the following adaptive strategy: we first select all d nodes in V_2 and observe which nodes in V_3 have not yet been reached, this can be done with myopic feedback. We would then use the left budget to select nodes in V_3 that have not been reached. Let $X_j = \mathbb{I}\{v_j^3$ not activated by seed nodes in $V_2\}$ for $j \in [2d]$, where $\mathbb{I}\{\}$ is the indicator function. X_j's are independent Bernoulli random variables with $\mathbb{E}[X_j] = \frac{1}{e+1}$. Then by the Chernoff bound,

$$\Pr[X_1 + \cdots + X_{2d} > \frac{2}{e+1}d + \epsilon d] \leq e^{-\frac{\epsilon d \cdot \epsilon (e+1)/2}{3}} \leq e^{-d\epsilon^2/2} \leq \frac{\epsilon}{2}. \tag{27}$$

Consequently, the expected number of nodes in V_3 that have not been activated by seeds in V_2 is at most $\frac{\epsilon}{2} \cdot 2d + (1 - \frac{\epsilon}{2}) \cdot (\frac{2}{e+1}d + \epsilon d) \leq \frac{2}{e+1}d + 2\epsilon d$. But the adaptive greedy algorithm still has a budget of $\frac{2}{e+1}d$ to directly activate nodes in V_3, and thus the expected final number of non-activated nodes in V_3 is at most $2\epsilon d$. Thus we conclude the proof. $\square$

Next, we consider the greedy algorithm and have the following conclusion.

Claim 5. *The non-adaptive greedy algorithm would first select all $d-1$ nodes in V_1, and then select $\frac{2}{e+1}d + 1$ nodes in V_2. Consequently, we have that*

$$\mathrm{Greedy}_N(G, k) = (d-1) + \left[\left(\frac{2}{e+1}d + 1 \right) + \left(1 - \left(1 - \frac{1}{d} \right)^{d-1} \right) \cdot \left(\frac{e-1}{e+1}d - 1 \right) \right] \cdot (1 + \frac{2e}{e+1}w), \tag{28}$$

when $d, w \to \infty$, we know that $\frac{\mathrm{Greedy}_N(G,k)}{dw} \to \frac{2e^2+2}{(e+1)^2}$.

Proof. We first prove that greedy would first select all $d-1$ nodes in V_1. Consider that the greedy algorithm has already selected j nodes in V_1 as seeds, with $j = 0, 1, \ldots, d-1$. Let p_j denote the probability that a node in V_2 is activated in this case. We know that $p_j = 1 - (1 - \frac{1}{d})^j$. At this point, we know that the marginal gain for selecting the $(j+1)$-th node in V_1 is

$$M_1 = 1 + d \cdot \frac{1}{d}(1 - p_j) \cdot (1 + \frac{2e}{e+1}w) = 1 + (1 - p_j)(1 + \frac{2e}{e+1}w). \tag{29}$$

In contrast, the marginal gain for selecting the first node in V_2 as a seed is

$$M_2 = (1 - p_j)(1 + \frac{2e}{e+1}w), \tag{30}$$

and the marginal gain for selecting the first node in V_3 as a seed is

$$M_3 = p_j(1 - \frac{e}{e+1})w + (1 - p_j)w = \left(p_j \cdot \frac{1}{e+1} + (1 - p_j) \right) w. \tag{31}$$

Therefore $M_1 > M_2$. Comparing M_1 with M_3, we use the fact that for all $j < d$, $p_j \leq 1 - 1/e$, and thus

$$M_1 - M_3 = 1 + (1 - p_j)(1 + \frac{2e}{e+1}w) - \left(p_j \cdot \frac{1}{e+1} + (1 - p_j) \right) w$$

$$> (1 - p_j)\frac{2e}{e+1}w - \left(p_j \cdot \frac{1}{e+1} + (1 - p_j) \right) w$$

17

$$= \left(\frac{e-1}{e+1} - \frac{e}{e+1} p_j \right) w$$

$$\geq 0. \tag{32}$$

Thus we conclude that greedy would select all $(d-1)$ nodes in V_1 first. Afterwards, we compare the marginal gain of selecting a node in V_2 versus selecting a node in V_3. Notice that if we select a node in V_3, we would definitely not select a node whose in-neighbor in V_2 is already selected as a seed, because it only decreases the marginal. Therefore, the marginal gains of selecting a node in V_2 or a node in V_3 are still given us M_2 and M_3. Thus, the difference of marginal gain is

$$\begin{aligned} M_2 - M_3 &= (1-p_{d-1})(1 + \frac{2e}{e+1}w) - \left(p_{d-1} \cdot \frac{1}{e+1} + (1-p_{d-1}) \right) w \\ &> (1-p_{d-1}) \frac{2e}{e+1} w - \left(p_{d-1} \cdot \frac{1}{e+1} + (1-p_{d-1}) \right) w \\ &= \left(\frac{e-1}{e+1} - \frac{e}{e+1} p_{d-1} \right) w \\ &= \left(\frac{e-1}{e+1} - \frac{e}{e+1} \left(1 - \left(1 - \frac{1}{d} \right)^{d-1} \right) \right) w \\ &\geq 0. \end{aligned} \tag{33}$$

Thus the marginal gain for selecting nodes in V_2 is greater than nodes in V_3 and greedy would select $\frac{2}{e+1}d + 1$ nodes in V_2. All in all, the expected utility for greedy is

$$\mathrm{Greedy}_N(G,k) = (d-1) + \left[\left(\frac{2}{e+1}d + 1 \right) + \left(1 - \left(1 - \frac{1}{d} \right)^{d-1} \right) \cdot \left(\frac{e-1}{e+1}d - 1 \right) \right] \cdot (1 + \frac{2e}{e+1}w). \tag{34}$$

and when $d, w \to \infty$, we know that $\frac{\mathrm{Greedy}_N(G,k)}{dw} \to \frac{2e^2+2}{(e+1)^2}$. $\qquad \square$

Combining Claim 5 and Claim 4, we conclude that when $d, w \to \infty$,

$$\frac{\mathrm{Greedy}_N(G,k)}{\mathrm{OPT}_A(G)} \to \frac{e^2+1}{(e+1)^2} \approx 0.606. \tag{35}$$

$\qquad \square$

We then assert that the approximation ratio of adaptive greedy is no better than greedy.

Lemma 7. *The approximation ratio for the non-adaptive greedy algorithm is no worse than the adaptive greedy algorithm, over all graphs.*

Proof. Fix an influence graph $G(V, E, p)$, and any $k \in [n]$. We use c to denote the approximation ratio of greedy, i.e.,

$$c = \frac{\mathrm{Greedy}_N(G,k)}{\mathrm{OPT}_A(G,k)}.$$

We construct a family of graph $G(w)$ such that the approximation ratio for adaptive greedy is approaching to c when $w \to \infty$. The influence graph $G(w)$ consists of two parts, G_1 and G_2. The graph G_1 has same nodes as G, but it does not contain any edges, while the graph G_2 is exactly the same as G, except that the weight for each node is multiplied by a factor of w. Notice that we can always assign integral weights w to a node by connecting it to a directed chain of length $w-1$. For any node $v \in G_1$, v has exactly one outgoing edge, connecting to the corresponding node in G_2, the edge will be live with probability 1.

Now, consider adaptive greedy on $G(w)$ with the same budget. Our first observation is that adaptive greedy will never choose nodes from G_2. This is because if the corresponding node in G_1 has not been chosen, the marginal gain of choosing the node in G_1 is always larger by 1, and if it has already

18

been chosen, the marginal gain to choose the node in G_2 is 0. Consequently, the adaptive greedy algorithm would always choose nodes in G_1. However, because myopic feedback only provides one step feedback after seed selection, selecting a node in G_1 would only provide the activation of its corresponding node in G_2 as the feedback, but this is already known for sure, and thus we do not get any useful feedback under myopic feedback model on this graph. Therefore, the adaptive greedy algorithm in this case behaves exactly the same as the non-adaptive greedy algorithm on the influence graph G, and the performance for adaptive greedy is

$$\text{Greedy}_A(G(w),k) = w \cdot \text{Greedy}_N(G,k) + k \leq (w+1) \cdot \text{Greedy}_N(G,k). \tag{36}$$

Consider the optimal adaptive policy, a feasible adaptive policy is to ignore nodes in graph G_1 and perform the optimal adaptive policy on graph G_2, we have

$$\text{OPT}_A(G(w),k) \geq \text{OPT}_A(G_2(w),k) = w \cdot \text{OPT}_A(G,k). \tag{37}$$

By Eq. (36) and Eq. (37), the approximation ratio of adaptive greedy can be bounded as

$$\frac{\text{Greedy}_A(G(t),k)}{\text{OPT}_A(G(t),k)} \leq \frac{(w+1) \cdot \text{Greedy}_N(G,k))}{w \cdot \text{OPT}_A(G,k)} = \frac{w+1}{w} \cdot c \to c, \text{ when } w \to \infty. \tag{38}$$

This concludes the proof. $\qquad\square$