Adaptivity Gaps for Influence Maximization: Percolation and Depletion

Gabriel de Azevedo* and Ulysse Hennebelle†

Abstract

In the *influence maximization* problem, a set of at most k vertices, called *seeds*, must be selected in a network where each arc is independently present with a known probability in order to maximize the expected number of vertices that are reached from the seeds. A *static policy* selects all seeds in advance, while an *adaptive policy* selects them sequentially and draws on the feedback provided by previous cascades. We study the *adaptivity gap* between these two classes and establish three new upper bounds. First, for undirected graphs, we prove an adaptivity gap of at most $5\sqrt{k}$, where k is the seed budget. This is the first non-trivial upper bound that holds independently of the graph. Our proof relies on two connection inequalities from percolation theory and extends to matroid constraints. Second, for general directed graphs on n vertices, we obtain an $\mathcal{O}(\log n/\log\log n)$ bound, improving the previous $\lceil n^{1/3}\rceil$ guarantee. Finally, we establish a constant bound of 2 for bi-directed forests, unifying and improving existing constant bounds for various orientations of forests. Our last two results rely on a depletion approach that bounds the influence remaining after successive rounds of seed selection.

*Cornell University. Email: gm587@cornell.edu

†Cornell Tech, Cornell University. Email: uh34@cornell.edu

1 Introduction

Influence maximization is the problem of selecting a set of at most k vertices, called *seeds*, to maximize the expected spread of an activation cascade in a network. The idea of influence maximization was developed by Domingos and Richardson [DR01; RD02] who distinguished the intrinsic and the network value of a customer in viral marketing. Since then, influence maximization has seen many applications both in practice and theory such as viral marketing [HSBB11; CWW10], microfinance diffusion [BCDJ13], and outbreak detection in networks [Les+07]. In their seminal work, Kempe, Kleinberg, and Tardos [KKT03] studied this problem under two standard diffusion models: the *independent cascade model*, in which each arc is live independently with a probability given by a vector $\mathbf{p}$, and a seed activates every vertex reachable through live paths, and the *linear threshold model*, where each vertex becomes active when the total weight of its active in-neighbors reaches a randomly drawn threshold. For surveys of diffusion models, algorithms, and applications, see [LFWT18; BJP20]. In this paper, we focus on the independent cascade model.

In the same work, Kempe, Kleinberg, and Tardos [KKT03] showed that influence is a monotone submodular function of the seed set. The classical analysis of greedy maximization therefore yields a $(1 - 1/e)$-approximation to the optimal set of k seeds [NWF78]. For undirected influence networks, in the independent cascade model, Khanna and Lucier [KL14] proved that the same greedy algorithm achieves a $(1 - 1/e + c)$-approximation for a universal constant $c > 0$. This improvement shows that undirected diffusion obeys a structure that permits stronger guarantees than monotone submodularity alone.

A *static policy*, also called a *non-adaptive policy*, chooses all seeds before observing the diffusion process. In many applications, however, seeds can be *probed* one at a time, so an *adaptive policy* can use *feedback* from earlier seeds to guide later choices. Adaptive influence maximization has, for example, been used to plan health campaigns, adapting participant selection to information revealed during earlier interventions [Yad+16]. The two main feedback models were introduced by Golovin and Krause [GK10; GK11]. Under *full-adoption feedback*, the policy observes the entire cascade of each seed before choosing the next one. They showed that this problem fits the framework of adaptive submodularity and thus that adaptive greedy achieves a $(1 - 1/e)$-approximation to the optimal adaptive policy. Under *myopic feedback*, the policy observes only the states of arcs leaving the selected seed. Peng and Chen [PC19] obtained a constant-factor approximation algorithm for influence maximization under this feedback model. Our focus is on full-adoption feedback.

Approximation guarantees compare algorithms with the optimal policy for a given model, but do not determine how much feedback improves the optimum itself. First formalized by Dean, Goemans, and Vondrák [DGV04], the *adaptivity gap* measures, for a given model, how much the information given by the feedback can increase the optimal expected influence by measuring the ratio between the expected value of an optimal adaptive policy and that of an optimal static policy. In the context of influence maximization, adaptivity gaps have already been studied heuristically in practice [Jan+17], and in theory, first by Chen and Peng [CP19] who established a lower bound of $e/(e-1)$ as well as constant upper bounds for several restricted classes of directed graphs. For general directed graphs on n vertices, D'Angelo, Poddar, and Vinci [DPV23] obtained the first sublinear upper bound, $\lceil n^{1/3} \rceil$. For undirected graphs, they also proved a bound of $\sqrt{\alpha} + \mathcal{O}(1)$, where α is the sum of the degrees of vertices of degree greater than two. Before the present work, no general upper bound depending only on the number of seeds k improved on the elementary linear bound of k implied by adaptive submodularity, even for undirected graphs.

This limited understanding contrasts with the well-established bounds for the related model of stochastic monotone submodular maximization. For that setting, Asadpour and Nazerzadeh [AN16] proved a tight adaptivity gap of $e/(e-1)$ under a matroid constraint. This bound is permitted

by the independence of item outcomes. In influence maximization, cascades generated by different seeds share a realization of the graph and can therefore be dependent, so these independence-based bounds do not apply directly.

1.1 Problem definition

An instance of *influence maximization* consists of a finite set of vertices V of cardinality n and a probability vector $\mathbf{p} \in [0,1]^{V^2}$. Throughout, vectors are written in bold and their entries in plain type. Each instance $(V, \mathbf{p})$ is associated with a random arc set $A \subseteq V^2$. For $A' \subseteq V^2$,

$$\mathbb{P}(A = A') = \prod_{a \in A'} p_a \prod_{a \in V^2 \setminus A'} (1 - p_a).$$

We call an arc *live* if it belongs to A. Hence, every potential arc $a \in V^2$ is live independently with probability p_a. Unless otherwise specified, probabilities and expectations are taken over A with the distribution given above and over any policy or sampling randomness present.

We define the *cascade function* $C_V : 2^V \to 2^V$ by letting $C_V(S)$ be the set of vertices reachable from at least one vertex of $S \subseteq V$ along paths of live arcs. We call $C_V(S)$ the *cascade* of S. We say that a vertex in $C_V(S)$ is *activated* by S. The *influence* of S is $\sigma_V(S) = |C_V(S)|$, the number of vertices reached by S. When the vertex set V is clear from the context, we omit it from the subscript. When the set consists of a single seed v, we use $C(v)$ and $\sigma(v)$ as shorthand for $C(\{v\})$ and $\sigma(\{v\})$.

Given an instance $(V, \mathbf{p})$ and an integer $k \in \mathbb{N}$, we define the value of the best static policy as

$$\textsc{Static}_k(V, \mathbf{p}) = \max_{S \subseteq V,\, |S| \leq k} \mathbb{E}[\sigma(S)].$$

We refer to this number k as the *budget* and the set of vertices chosen by the policy as the set of *seeds*.

For an instance $(V, \mathbf{p})$, we define an *adaptive policy* π with budget k to be a random sequence $(v_1, \ldots, v_k)$ of vertices in V such that, for every $i \in [k]$, the choice of the vertex v_i is a function of $C(v_1), \ldots, C(v_{i-1})$ and any external randomness independent of A. We use $S(\pi) = \{v_1, \ldots, v_k\}$ for the set of seeds probed by π, $C(\pi)$ as shorthand for $C(S(\pi))$ for the set of vertices activated by π, and $\sigma(\pi) = |C(\pi)|$ for the influence of π. We denote the set of policies with budget k by Π_k. We define the value of the best adaptive policy as

$$\textsc{Adaptive}_k(V, \mathbf{p}) = \max_{\pi \in \Pi_k} \mathbb{E}[\sigma(\pi)].$$

We omit the parameters $(V, \mathbf{p})$ from $\textsc{Static}_k$ and $\textsc{Adaptive}_k$ when the instance is clear from the context. We will use that $\textsc{Adaptive}_k \leq k\textsc{Static}_1$, which stems from adaptive submodularity [GK11] and provides the elementary upper bound of k on the adaptivity gap.

The *adaptivity gap* of an instance $(V, \mathbf{p})$ with seed budget k is given by the ratio

$$\textsc{Adaptive}_k(V, \mathbf{p}) / \textsc{Static}_k(V, \mathbf{p}),$$

which is always at least 1. We formulate the problem for unweighted vertices. More generally, each vertex v may have a positive weight, and the influence of a seed set is the total weight of its cascade. All our proofs extend immediately to weighted vertices, and we use unweighted vertices for simplicity of exposition. In general, integer-weighted instances can be exactly replaced by unweighted instances by increasing the number of vertices.[1]

[1] For a vertex v with desired weight w_v, attach $w_v - 1$ new vertices to it through arcs in both directions, each live with probability 1. The resulting group activates together, and seeding any of its vertices is equivalent to seeding v.

1.2 Our results

We establish three upper bounds on the adaptivity gap for the independent cascade model with full-adoption feedback. For general undirected graphs, we obtain the first upper bound that is sublinear in the seed budget k and independent of the graph size, improving on the elementary linear bound.

Theorem 1 (Undirected graphs)**.** *For every undirected instance of influence maximization,*

$$\textsc{Adaptive}_k \leq 5\sqrt{k}\,\textsc{Static}_k.$$

We derive this bound by comparing the influence of each seed of an adaptive policy with the influence that this seed would add to a fixed set S of seeds. Since connections to S are positively correlated, a seed may also activate vertices that S reaches, and we control this overlap by splitting the seeds according to their probability of connecting to S. For seeds that rarely connect to S, extending a percolation result of Gladkov [Gla24], we show that a seed is unlikely to reach both a vertex and S unless that vertex is itself likely to connect to S. For seeds that connect to S with probability at least β, at most one of them activates each vertex, so their total overlap is at most the influence of the set they form. We bound this influence using a recently proved result conjectured by Kozma and Nitzan [KN24] in the context of percolation. Combining both bounds over all thresholds β yields the factor $\sqrt{k}$. We give the proof of Theorem 1 in Section 2. We then extend these results to any matroid constraints over the ser of seeds and obtain a bound of $5\sqrt{r}$ where r is the rank of the matroid, see Corollary 7.

For general directed graphs, we improve the upper bound from $\lceil n^{1/3} \rceil$ [DPV23], obtaining the first sublogarithmic upper bound.

Theorem 2 (Directed graphs)**.** *For every directed instance of influence maximization with n vertices,*

$$\textsc{Adaptive}_k \leq \mathcal{O}_{n\to\infty}\left(\frac{\log n}{\log\log n}\right)\textsc{Static}_k.$$

We obtain this result by studying an auxiliary batched policy that repeatedly seeds an optimal set of at most k vertices in the remaining graph until no high-influence vertex remains. We show that the probability that a vertex retains high influence decreases geometrically with the number of batches, so the policy needs only a sublogarithmic number of batches in expectation. We then compare this policy with any adaptive policy by bounding each batch's expected additional influence by $\textsc{Static}_k$ and using the absence of high-influence vertices at termination to control the additional influence obtainable from k adaptive seeds. We give the proof of Theorem 2 in Section 3.

Finally, we give an upper bound of 2 for bi-directed forests, unifying several previous bounds for specific orientations of forests and giving the first constant upper bound for undirected forests. A *bi-directed forest* is obtained from an undirected forest by replacing each edge by two independent arcs in opposite directions with possibly different live probabilities.

Theorem 3 (Bi-directed forests)**.** *For every bi-directed forest instance of influence maximization,*

$$\textsc{Adaptive}_k \leq 2\,\textsc{Static}_k.$$

The proof builds on the depletion idea used for general directed graphs in Section 3. The forest structure lets us choose a small seed set whose cascade leaves only vertices with small expected influence. We then bound the additional influence that any adaptive policy can obtain beyond this cascade, giving an adaptivity gap of at most 2. We give the proof of Theorem 3 in Section 4.

Table 1 summarizes our bounds alongside the best previously known upper bounds.

Graph Class	Previous Upper Bounds	This Work
Undirected graphs	$\lceil n^{1/3} \rceil$, $\sqrt{\alpha} + \mathcal{O}(1)$ [DPV23]	$5\sqrt{k}$
Directed graphs	$\lceil n^{1/3} \rceil$ [DPV23]	$\mathcal{O}(\log n / \log \log n)$
In-arborescences	$\frac{2e^2}{e^2-1}$ [DPV23]	2
Out-arborescences	2 [CP19]	2
Bi-directed forests	$\lceil n^{1/3} \rceil$, $\sqrt{\alpha} + \mathcal{O}(1)$ [DPV23]	2

Table 1: Adaptivity-gap upper bounds under full-adoption feedback. Here α is the sum of degrees greater than two in the underlying undirected graph.

We discuss the connections with previous work in the following subsection. We hope that the techniques developed here will also prove useful for other problems in influence maximization and, more generally, for graph diffusion problems.

1.3 Related work

Influence Maximization. Kempe, Kleinberg, and Tardos [KKT03] formulated influence maximization as a discrete optimization problem for the independent cascade and the linear threshold models and established approximation guarantees through submodularity. The connection between diffusion and submodularity extends beyond the independent cascade model: Mossel and Roch [MR07; MR10] showed that monotone submodular local activation functions yield a monotone submodular expected influence function in the general threshold model.

A substantial body of work addresses the cost of evaluating and maximizing influence. Chen, Wang, and Wang [CWW10] proved that computing influence exactly under the independent cascade model is #P-hard. Efficient approximation algorithms can achieve a ratio of $(1 - 1/e - \varepsilon)$ without exact expected influence evaluation [BBCL14; TXS14]. For undirected networks in the static model, Khanna and Lucier [KL14] used properties of bond percolation to show an improved approximation ratio for the greedy algorithm. Our analysis leverages several tools from percolation theory to study the adaptivity gap in undirected networks. For adaptive policies, adaptive greedy achieves a $(1 - 1/e)$-approximation under full-adoption feedback [GK11], while its best known approximation ratio under myopic feedback is $1 - 1/\sqrt{e}$ [DPV21].

Adaptivity Gaps. The concept of adaptivity gap was introduced by Dean, Goemans, and Vondrák [DGV04] as a quantification of the gain of adaptivity in stochastic decision problems. Adaptivity gaps have since been studied in stochastic submodular maximization and stochastic probing. For independent item outcomes and a matroid constraint, Asadpour and Nazerzadeh [AN16] proved the tight $e/(e-1)$ bound and gave a polynomial-time stochastic continuous greedy algorithm that computes a static policy within a $(1 - 1/e - \varepsilon)$ factor of the adaptive optimum. Constant adaptivity gaps for submodular objectives under prefix-closed probing constraints were established by Gupta, Nagarajan, and Singla [GNS17]. For monotone submodular objectives under these constraints, a tight bound of 2 was proved by Bradac, Singla, and Zuzic [BSZ19]. To the contrary, under XOS or subadditive valuations, the adaptivity gap is not constant in general [GNS17; LLZ25].

For influence maximization, under full-adoption feedback, Chen and Peng [CP19] proved upper bounds of $2e/(e-1)$ for in-arborescences and 2 for out-arborescences, and a tight bound of $e/(e-1)$ for one-directional bipartite graphs. They also proved a lower bound of $e/(e-1)$ even on directed

paths. D'Angelo, Poddar, and Vinci [DPV23] improved the in-arborescence bound to $2e^2/(e^2-1)$ and established an upper bound of $\lceil n^{1/3} \rceil$ for general directed graphs. They also obtained a $\sqrt{\alpha} + \mathcal{O}(1)$ bound, where α is the sum of degrees greater than two in the underlying undirected graph; opposite arcs may have different live probabilities. Their results cover several restricted classes of bi-directed forests, for which our theorem improves the upper bound to 2.

Under myopic feedback, Peng and Chen [PC19] showed that the adaptivity gap lies between $e/(e-1)$ and 4. D'Angelo, Poddar, and Vinci [DPV21] subsequently improved the upper bound to $2e/(e-1)$.

Percolation Inequalities. The conjecture of absence of percolation at criticality was a central open problem in percolation theory concerning the existence of infinite connected components in random graphs. Consider the lattice graph with vertex set $\mathbb{Z}^d$, where two vertices are adjacent if their ℓ_1 distance is 1. Each edge is present independently with probability p. There exists a critical probability p_c such that, for $p < p_c$, all connected components are almost surely finite, whereas for $p > p_c$, an infinite component exists almost surely [Gri99]. The conjecture asked whether all components remain finite almost surely at $p = p_c$. Equivalently, it asks whether $\theta(p_c) = 0$, where $\theta(p)$ denotes the probability that the origin belongs to an infinite component. The conjecture had been proved for $d = 2$ [Har60; Kes80] and for $d \geq 19$ [HS90], with the latter result subsequently extended to $d \geq 11$ [Fv17].

Kozma and Nitzan [KN24] reduced the conjecture in all dimensions to connection inequalities for finite random graphs with arbitrary independent edge probabilities. These inequalities were recently formalized in Lean and proved with the help of AI systems [Led26; Bou26; Ope26]. Subsequently, Blanc-Renaudie [Bla26] and Gouéré [Gou26] provided simpler, human-readable proofs. We use the first inequality in Kozma and Nitzan [KN24] (stated as Conjecture 1 in that work) in our proof of the bound for undirected graphs.

We also use the three-point inequality of Gladkov [Gla24, Thm. 6.2], which bounds the probability that three vertices belong to the same connected component in terms of their three pairwise connection probabilities.[2] We extend this version to a set in place of one of the vertices; see Lemma 5.

2 An $\mathcal{O}(\sqrt{k})$ Bound for Undirected Graphs

In this section, we prove the upper bound of Theorem 1 for undirected instances. An instance $(V, \mathbf{p})$ is *undirected* if $p_{u,v} = p_{v,u}$ for every $u, v \in V$. In the independent cascade model, this is equivalent to having undirected edges $\{u, v\}$ that are live independently with probability $p_{u,v}$.[3] We therefore index $\mathbf{p}$ by unordered pairs $\{u, v\} \subseteq V$. In this representation, A denotes the random set of live undirected edges.

For vertices $u, v \in V$, we write $u \leftrightarrow v$ when they are connected by a path of live edges; equivalently, $v \in C(u)$, or $u \in C(v)$, or $C(u) = C(v)$, each of these conditions implying the others. For a set $S \subseteq V$, we write $u \leftrightarrow S$ if $u \in C(S)$, equivalently if $C(u) \subseteq C(S)$. Given an adaptive policy π, the *seed probability* of each vertex $v \in V$ is $x_v = \mathbb{P}(v \in S(\pi))$. We write $\mathbf{x} = (x_v)_{v \in V}$ for the vector of seed probabilities. Our proof revolves around establishing the following lemma.

[2] The same inequality with a better constant was also claimed by Khanna and Lucier [KL14] and applied in the context of influence maximization. However, we encountered gaps in the original argument that we could not resolve.

[3] For each pair $\{u, v\}$, only the arc from the endpoint activated first to the other endpoint can produce a new activation. Since both arcs have the same live probability, we may represent the pair by a single independent undirected edge. Even after observing earlier cascades, the arcs between vertices not yet activated remain independent with their original probabilities, so this replacement also applies to adaptive seed choices. Consequently, $\textsc{Static}_k$ and $\textsc{Adaptive}_k$ are the same in both representations.

Lemma 4. *Let $(V, \mathbf{p})$ be an undirected instance of influence maximization and π be an adaptive policy with seed probabilities $\mathbf{x}$. Then, for every vertex $u \in V$ and every set $S \subseteq V$,*

$$\mathbb{P}(u \leftrightarrow S(\pi)) \leq 2 \sum_{v \in V} x_v \mathbb{P}(u \leftrightarrow v,\ u \not\leftrightarrow S) + 4\sqrt{\sum_{v \in V} x_v}\,\mathbb{P}(u \leftrightarrow S).$$

Notice that the function $f : S \mapsto \mathbb{P}(u \leftrightarrow S)$ is a monotone and submodular function, and that $\mathbb{P}(u \leftrightarrow v, u \not\leftrightarrow S) = f(S \cup \{v\}) - f(S)$. Thus, it holds that

$$f^+(\mathbf{x}) \leq \min_{S \subseteq V} \left\{ f(S) + \sum_{v \in V} x_v [f(S \cup \{v\}) - f(S)] \right\},$$

where f^+ is the concave closure of f [Von07]. When $\sum_{v \in V} x_v \leq k$, Lemma 4 can be seen as an analog of this inequality up to a factor $\mathcal{O}(\sqrt{k})$ and upper bounds the performance of any adaptive policy π, which may both correlate seeds and choose them depending on the graph realization.

Proof of Theorem 1. Let π be an adaptive policy with $\sum_{v \in V} x_v \leq k$. Let S be a set of k seeds maximizing $\mathbb{E}[\sigma(S)]$, so that $\mathbb{E}[\sigma(S)] = \textsc{Static}_k$. Fix a vertex $v \notin S$. Optimality of S and submodularity then imply that $\mathbb{E}[\sigma(S \cup \{v\}) - \sigma(S)] \leq \mathbb{E}[\sigma(S)]/k$. By summing Lemma 4 over all vertices $u \in V$, we get, for $k \geq 4$,

$$\begin{aligned}\mathbb{E}[\sigma(\pi)] &\leq 2 \sum_{v \in V} x_v \mathbb{E}[\sigma(S \cup \{v\}) - \sigma(S)] + 4\sqrt{\sum_{v \in V} x_v}\,\mathbb{E}[\sigma(S)] \\ &\leq 2\mathbb{E}[\sigma(S)] + 4\sqrt{k}\mathbb{E}[\sigma(S)] \leq 5\sqrt{k}\textsc{Static}_k.\end{aligned}$$

The second inequality stems from $\sum_v x_v \leq k$. For $k \leq 3$, the bound is elementary. □

Notably, this result requires the adaptive policy to satisfy the budget constraint only in expectation, while the static solution must satisfy it in every realization. Our techniques allow us to generalize the budget constraints to matroid constraints; see Corollary 7.

To prove Lemma 4, fix a vertex $u \in V$ and a set $S \subseteq V$. The *contribution* of a seed v is the probability that v is probed by π and activates u, meaning that u is connected to v but is not connected to the seeds probed by π before v. The sum of these contributions over $v \in V$ is $\mathbb{P}(u \leftrightarrow S(\pi))$. We define the *excess* of a seed as

$$e(v) \coloneqq (\mathbb{P}(v \text{ activates } u) - 2x_v \mathbb{P}(u \leftrightarrow v,\ u \not\leftrightarrow S))_+,$$

where $y_+ = \max(y, 0)$. The probability $\mathbb{P}(u \leftrightarrow v,\ u \not\leftrightarrow S)$ is the marginal increase in the activation probability of u when v is added to S. Thus, $e(v)$ measures how much v's contribution under π exceeds twice this marginal probability weighted by x_v. We bound this excess using the probability that v connects to both u and S. Since $\mathbb{P}(u \leftrightarrow S(\pi)) \leq 2\sum_{v \in V} x_v \mathbb{P}(u \leftrightarrow v,\ u \not\leftrightarrow S) + \sum_{v \in V} e(v)$, it suffices to show that $\sum_{v \in V} e(v) \leq 4\sqrt{\sum_{v \in V} x_v}\,\mathbb{P}(u \leftrightarrow S)$. The excess depends on the probability $\kappa(v) \coloneqq \mathbb{P}(v \leftrightarrow S)$ that v connects to S. We bound the excess individually for seeds with small $\kappa(v)$, which we call *weakly connected*, and collectively for seeds with large $\kappa(v)$, which we call *strongly connected*.

Weakly Connected Seeds. For a seed v to have positive excess, it must be able to reach both u and S. At the time of choosing v, the edges between the still inactive vertices follow their original distribution. Thus, whenever v is selected, its probability of activating u is at most $\mathbb{P}(u \leftrightarrow v)$,

and since v is selected with probability x_v, we have that $\mathbb{P}(v \text{ activates } u) \le x_v \mathbb{P}(u \leftrightarrow v)$. Since $\mathbb{P}(u \leftrightarrow v) = \mathbb{P}(u \leftrightarrow v,\, v \leftrightarrow S) + \mathbb{P}(u \leftrightarrow v,\, u \not\leftrightarrow S)$, we obtain that

$$e(v) \le x_v(\mathbb{P}(u \leftrightarrow v,\, v \leftrightarrow S) - \mathbb{P}(u \leftrightarrow v,\, u \not\leftrightarrow S))_+. \tag{1}$$

The first term can be much larger than $\kappa(v)\kappa(u)$, since the connections of u and v to S are positively correlated. The factor two in the definition of excess allows us to bound this difference by a constant multiple of $\kappa(v)\kappa(u)$. To bound the first term of the expression above, we use a slight generalization of the three-point inequality of Gladkov [Gla24].

Lemma 5 (Three-point inequality with sets)**.** *For all vertices $u, v \in V$ and every set $S \subseteq V$,*

$$\mathbb{P}(u \leftrightarrow v,\ v \leftrightarrow S)^2 \le 8\mathbb{P}(u \leftrightarrow v)\,\mathbb{P}(v \leftrightarrow S)\,\mathbb{P}(u \leftrightarrow S).$$

Proof of Lemma 5. For every undirected instance $(V, \mathbf{p})$ and vertices $u, v, z \in V$, the three-point inequality of Gladkov [Gla24, Thm. 6.2] gives

$$\mathbb{P}(u \leftrightarrow z,\, v \leftrightarrow z)^2 \le 8\mathbb{P}(u \leftrightarrow v)\mathbb{P}(u \leftrightarrow z)\mathbb{P}(v \leftrightarrow z). \tag{2}$$

We obtain the set version by applying (2) to a graph with an added vertex representing S.

Add a vertex z joined to every vertex of S by an edge of probability one, and write $\mathbb{P}^+$ for probabilities in the resulting graph. Connections to z correspond exactly to connections to S in the original instance. After adding z, the vertices u and v are connected exactly when they were already connected in the original instance or both were connected to S; hence, the connection probability increases by $\mathbb{P}(u \not\leftrightarrow v,\, u \leftrightarrow S,\, v \leftrightarrow S)$. If $\mathbb{P}(u \leftrightarrow v) = 0$, then $\mathbb{P}(u \leftrightarrow v,\, v \leftrightarrow S) = 0$, so the desired inequality holds. We may therefore assume $\mathbb{P}(u \leftrightarrow v) > 0$. Since $\mathbb{P}^+(u \leftrightarrow v) \ge \mathbb{P}(u \leftrightarrow v)$, we obtain that

$$\frac{\mathbb{P}(u \leftrightarrow v,\, v \leftrightarrow S)}{\mathbb{P}(u \leftrightarrow v)} = 1 - \frac{\mathbb{P}(u \leftrightarrow v,\, u \not\leftrightarrow S)}{\mathbb{P}(u \leftrightarrow v)} \le 1 - \frac{\mathbb{P}(u \leftrightarrow v,\, u \not\leftrightarrow S)}{\mathbb{P}^+(u \leftrightarrow v)} = \frac{\mathbb{P}^+(u \leftrightarrow z,\, v \leftrightarrow z)}{\mathbb{P}^+(u \leftrightarrow v)}.$$

Combining this with $\mathbb{P}(u \leftrightarrow v,\, v \leftrightarrow S) \le \mathbb{P}^+(u \leftrightarrow z,\, v \leftrightarrow z)$ gives

$$\frac{\mathbb{P}(u \leftrightarrow v,\, v \leftrightarrow S)^2}{\mathbb{P}(u \leftrightarrow v)} \le \frac{\mathbb{P}^+(u \leftrightarrow z,\, v \leftrightarrow z)^2}{\mathbb{P}^+(u \leftrightarrow v)} \le 8\mathbb{P}(u \leftrightarrow S)\mathbb{P}(v \leftrightarrow S),$$

where the second inequality follows from (2) applied to the augmented graph. Multiplying by $\mathbb{P}(u \leftrightarrow v)$ concludes the proof. □

Intuitively, v is unlikely to connect to both u and S unless u is itself likely to connect to S. Formally, by Lemma 5,

$$\mathbb{P}(u \leftrightarrow v,\, v \leftrightarrow S) \le \sqrt{8\mathbb{P}(u \leftrightarrow v)\kappa(v)\kappa(u)} \le \frac{\mathbb{P}(u \leftrightarrow v)}{2} + 4\kappa(v)\kappa(u),$$

where the second inequality follows from the AM–GM inequality. Rearranging and applying (1) gives

$$e(v) \le 8x_v\kappa(v)\kappa(u) \qquad \text{for all } v \in V. \tag{3}$$

Strongly Connected Seeds. For strongly connected seeds, bounding the excess of each seed separately is too loose. These individual bounds may count several seeds that connect to u, although at most one activates it in an execution of π. We therefore bound the total contribution of the

strongly connected seeds at once. Since u is activated by at most one seed, the probability that this seed belongs to a set $T \subseteq V$ is at most $\mathbb{P}(u \leftrightarrow T)$. Let $T_\beta \coloneqq \{\, v \in V : \kappa(v) \geq \beta \}$ be the set of vertices that connect to S with probability at least β. Since T_β does not depend on the realization, we obtain that

$$\sum_{v \in T_\beta} e(v) \leq \sum_{v \in T_\beta} \mathbb{P}(v \text{ activates } u) \leq \mathbb{P}(u \leftrightarrow T_\beta).$$

It remains to compare $\mathbb{P}(u \leftrightarrow T_\beta)$ with $\mathbb{P}(u \leftrightarrow S)$. To do this, we leverage the Kozma–Nitzan inequality [KN24; Gou26], formulated in the context of percolation, which we state as follows for all vertices $u, v \in V$ and every nonempty set $T \subseteq V$.

$$\mathbb{P}(u \leftrightarrow v) \geq \mathbb{P}(u \leftrightarrow T) \min_{t \in T} \mathbb{P}(t \leftrightarrow v). \tag{4}$$

If $T_\beta \neq \varnothing$, write $\mathbb{P}^\star$ for probabilities in the instance obtained by contracting S into a single vertex $s^\star$. Contraction preserves connections to S, so $\mathbb{P}^\star(w \leftrightarrow s^\star) = \mathbb{P}(w \leftrightarrow S)$ for every $w \in V \setminus S$, while it can only increase the probability of connecting u to T_β. Applying (4) to the contracted graph therefore gives

$$\mathbb{P}(u \leftrightarrow S) = \mathbb{P}^\star(u \leftrightarrow s^\star) \geq \mathbb{P}^\star(u \leftrightarrow T_\beta) \min_{t \in T_\beta} \mathbb{P}^\star(t \leftrightarrow s^\star) \geq \beta \mathbb{P}^\star(u \leftrightarrow T_\beta) \geq \beta \mathbb{P}(u \leftrightarrow T_\beta).$$

Combining both inequalities gives the following bound for strongly connected seeds:

$$\sum_{v \in V \colon \kappa(v) \geq \beta} e(v) \leq \kappa(u)/\beta \qquad \text{for all } 0 < \beta \leq 1. \tag{5}$$

Merging the Bounds. Finally, we combine the bounds into a proof of Lemma 4.

Proof of Lemma 4. It suffices to show that $\sum_{v \in V} e(v) \leq 4\sqrt{\sum_{v \in V} x_v}\, \kappa(u)$. By (3), $e(v) = 0$ whenever $\kappa(v) = 0$, and

$$\sum_{v \in V \colon \kappa(v) > 0} \frac{e(v)}{\kappa(v)} \leq 8\kappa(u) \sum_{v \in V} x_v.$$

On the other hand, applying (5) with $\beta = \kappa(v)$ gives $\sum_{t \in V \colon \kappa(t) \geq \kappa(v)} e(t) \leq \kappa(u)/\kappa(v)$ for every v with $\kappa(v) > 0$. Combining both bounds gives

$$\left(\sum_{v \in V} e(v) \right)^2 \leq 2 \sum_{v \in V \colon \kappa(v) > 0} e(v) \sum_{t \in V \colon \kappa(t) \geq \kappa(v)} e(t) \leq 2\kappa(u) \sum_{v \in V \colon \kappa(v) > 0} \frac{e(v)}{\kappa(v)} \leq 16\kappa(u)^2 \sum_{v \in V} x_v.$$

Taking square roots on both sides concludes the proof. □

Matroid Constraints. We show that Theorem 1 extends beyond cardinality constraints to matroid constraints. We first derive the following consequence of Lemma 4, which compares an adaptive policy with independent sampling according to its seed probabilities.

Proposition 6. *Let $(V, \mathbf{p})$ be an undirected instance of influence maximization and π be an adaptive policy with seed probabilities $\mathbf{x}$, and let $S_\mathbf{x}$, drawn independently of A, contain each vertex v independently with probability x_v. Then,*

$$\mathbb{P}(u \leftrightarrow S_\mathbf{x}) \geq \frac{\mathbb{P}(u \leftrightarrow S(\pi))}{2 + 4\sqrt{\sum_{v \in V} x_v}} \qquad \textit{for all } u \in V.$$

Proof. Fix a vertex $u \in V$. Order the vertices of V as $v_1, \dots, v_n$ and write $S_i = S_{\mathbf{x}} \cap \{v_1, \dots, v_i\}$, with $S_0 = \varnothing$. Adding v_i activates u exactly when v_i is sampled, $u \leftrightarrow v_i$, and $u \not\leftrightarrow S_{i-1}$. By independence of the sampling and since $S_{i-1} \subseteq S_{\mathbf{x}}$,

$$x_{v_i}\mathbb{P}(u \leftrightarrow v_i,\, u \not\leftrightarrow S_{\mathbf{x}}) \le x_{v_i}\mathbb{P}(u \leftrightarrow v_i,\, u \not\leftrightarrow S_{i-1}) = \mathbb{P}(u \leftrightarrow S_i) - \mathbb{P}(u \leftrightarrow S_{i-1}).$$

Summing over i, we obtain that $\sum_{v \in V} x_v \mathbb{P}(u \leftrightarrow v,\, u \not\leftrightarrow S_{\mathbf{x}}) \le \mathbb{P}(u \leftrightarrow S_{\mathbf{x}})$. Since $S_{\mathbf{x}}$ is independent of A, averaging Lemma 4 over $S_{\mathbf{x}}$ concludes the proof. □

A *matroid* $\mathcal{M} = (V, \mathcal{I})$ is a pair where V is a finite set of elements, called the *ground set*, and $\mathcal{I} \subseteq 2^V$ is a family of subsets whose members are called *independent sets*. The family $\mathcal{I}$ satisfies three properties: (i) $\varnothing \in \mathcal{I}$; (ii) if $I_1 \in \mathcal{I}$, then $I_2 \in \mathcal{I}$ for all $I_2 \subseteq I_1$; (iii) if $I_1, I_2 \in \mathcal{I}$ with $|I_1| > |I_2|$, then there exists $w \in I_1 \setminus I_2$ such that $I_2 \cup \{w\} \in \mathcal{I}$.

The *rank* of $\mathcal{M}$ is $r = \max_{I \in \mathcal{I}} |I|$. Let $\textsc{Static}_{\mathcal{M}}$ and $\textsc{Adaptive}_{\mathcal{M}}$ denote the optimal expected influence of static and adaptive policies, respectively, whose seed sets are independent in $\mathcal{M}$ in every execution. Swap rounding [CVZ10] rounds the seed probabilities of a feasible adaptive policy to a random independent set with expected influence at least that of $S_{\mathbf{x}}$. Since seed probabilities sum to at most the rank r, Proposition 6 gives the factor $2 + 4\sqrt{r} \le 5\sqrt{r}$ for $r \ge 4$, extending Theorem 1 to matroid constraints. For $r \le 3$, the bound is elementary.

Corollary 7 (Matroid constraints)**.** *For every undirected instance of influence maximization and every matroid $\mathcal{M}$ on the vertex set with rank r,*

$$\textsc{Adaptive}_{\mathcal{M}} \le 5\sqrt{r}\,\textsc{Static}_{\mathcal{M}}.$$

3 A Sublogarithmic Bound for Directed Graphs

In this section, we prove the upper bound of Theorem 2 for directed instances. Our proof revolves around an auxiliary policy that repeatedly seeds an optimal batch of at most k inactive vertices until no high-influence vertex remains. For a suitable high-influence threshold, we show that the expected number of batches before depletion is sublogarithmic.

Let $(V, \mathbf{p})$ be an instance. For $U \subseteq V$, we write $\mathbf{p}|_U$ and $A[U]$ for the restrictions of $\mathbf{p}$ and A to arcs with both endpoints in U. The instance restricted to U is $(U, \mathbf{p}|_U)$. For $S \subseteq V$, we interpret $C_U(S)$ as the vertices reachable from $S \cap U$ in $A[U]$; in particular, $\sigma_U(v) = 0$ for $v \notin U$. For fixed U, the arcs of $A[U]$ are independent with their original probabilities. This remains true conditional on any cascade history with inactive set U, since exploration only reveals arcs from activated vertices.

Fix a parameter $b > 1$ and let $H(U) = \{v \in U : \mathbb{E}[\sigma_U(v)] \ge b\textsc{Static}_k(V, \mathbf{p}) \,/\, k\}$ be the set of *high-influence* vertices of U. Let T be the random number of batches used by Algorithm 1. We prove the following bound.

Lemma 8. *For every $b > 1$,*

$$\mathbb{E}[T] \le \left\lceil \frac{\log n}{\log b} \right\rceil + \frac{b}{b-1}.$$

For constant b, this gives $\mathcal{O}(\log n)$ expected batches; choosing b as a function of n will give the sublogarithmic bound. For $t > T$, set $S_t = \varnothing$ and $U_t = U_{T+1}$.

Proof of Lemma 8. Deleting vertices cannot increase the expected influence of any remaining vertex, so $U_t \subseteq U_{t-1}$ implies $H(U_t) \subseteq H(U_{t-1})$. Fix $t \ge 2$ and suppose $H(U_{t-1}) \ne \varnothing$, so batch $t-1$ is executed. Furthermore, we have the uniform bound

$$\mathbb{E}[\sigma_{U_{t-1}}(S_{t-1}) \mid U_{t-1}] = \textsc{Static}_k(U_{t-1}, \mathbf{p}|_{U_{t-1}}) \le \textsc{Static}_k(V, \mathbf{p}).$$

Algorithm 1 Batched depletion policy

Input: Instance $(V, \mathbf{p})$, budget k, parameter b
$U_1 \leftarrow V$
$t \leftarrow 1$
while $H(U_t) \neq \varnothing$ **do**
  Choose $S_t \subseteq U_t$ with $|S_t| \leq k$ such that $\mathbb{E}[\sigma_{U_t}(S_t) \mid U_t] = \textsc{Static}_k(U_t, \mathbf{p}|_{U_t})$
  Seed all vertices in S_t and observe $C_{U_t}(S_t)$
  $U_{t+1} \leftarrow U_t \setminus C_{U_t}(S_t)$
  $t \leftarrow t + 1$
end while

For every fixed $v \in V$, we have $\sigma_{U_t}(v) = \sigma_{U_{t-1}}(S_{t-1} \cup \{v\}) - \sigma_{U_{t-1}}(S_{t-1})$. Thus, by submodularity of $\sigma_{U_{t-1}}$ and optimality of S_{t-1}, we have that $\mathbb{E}[\sigma_{U_t}(v) \mid U_{t-1}] \leq \textsc{Static}_k(V, \mathbf{p}) / k$. By the tower property, it follows that

$$\mathbb{E}[\mathbb{E}[\sigma_{U_t}(v) \mid U_t] \mid U_{t-1}] \leq \frac{\textsc{Static}_k}{k}.$$

Markov's inequality and $H(U_t) \subseteq H(U_{t-1})$ therefore give

$$\mathbb{P}(v \in H(U_t) \mid U_{t-1}) \leq \frac{1}{b} \mathbb{1}[v \in H(U_{t-1})],$$

where $\mathbb{1}[E]$ denotes the *indicator* of the event E. Taking expectations and iterating from $U_1 = V$ gives, for every $t \geq 1$,

$$\mathbb{P}(v \in H(U_t)) \leq \frac{1}{b^{t-1}}.$$

Since $\{T > t\} = \{H(U_{t+1}) \neq \varnothing\}$, a union bound gives, for every integer $t \geq 0$,

$$\mathbb{P}(T > t) \leq \min\left(1, \sum_{v \in V} \mathbb{P}(v \in H(U_{t+1}))\right) \leq \min\left(1, \frac{n}{b^t}\right).$$

Finally, summing the tail bound gives

$$\mathbb{E}[T] = \sum_{t \geq 0} \mathbb{P}(T > t) \leq \sum_{t \geq 0} \min\left(1, \frac{n}{b^t}\right) \leq \left\lceil \frac{\log n}{\log b} \right\rceil + \frac{b}{b-1}. \qquad \square$$

We can now prove Theorem 2.

Proof of Theorem 2. Let D be the set of vertices activated by Algorithm 1. For every $t \leq T$, conditional on U_t, the expected additional influence of batch t is $\textsc{Static}_k(U_t, \mathbf{p}|_{U_t}) \leq \textsc{Static}_k(V, \mathbf{p})$, since the remaining arcs are independent with their original probabilities. Since $\{T \geq t\} = \{H(U_t) \neq \varnothing\}$, summing over the batches gives

$$\mathbb{E}[|D|] = \sum_{t=1}^{\infty} \mathbb{E}[\mathbb{1}[T \geq t] \sigma_{U_t}(S_t)] \leq \sum_{t=1}^{\infty} \mathbb{P}(T \geq t) \textsc{Static}_k(V, \mathbf{p}) = \mathbb{E}[T] \textsc{Static}_k(V, \mathbf{p}).$$

The inequality follows by conditioning on U_t in each term. Take an adaptive policy $\pi \in \Pi_k$. Run Algorithm 1. Then, replay π in the same realization, using its original feedback $C(v_1), \ldots, C(v_{i-1})$

to choose v_i and ignoring the depletion history. We have that $C(\pi) \subseteq D \cup (C(\pi) \cap U_{T+1})$, and so

$$\begin{aligned}\mathbb{E}[\sigma(\pi)] &\leq \mathbb{E}[|D|] + \mathbb{E}\left[\mathbb{E}[|C(\pi) \cap U_{T+1}| \mid U_{T+1}]\right] \leq \mathbb{E}[|D|] + \mathbb{E}[\textsc{Adaptive}_k(U_{T+1}, \mathbf{p}|_{U_{T+1}})] \\ &\leq \mathbb{E}[|D|] + k\,\mathbb{E}[\textsc{Static}_1(U_{T+1}, \mathbf{p}|_{U_{T+1}})] \leq \mathbb{E}[T]\textsc{Static}_k + k\,\frac{b\textsc{Static}_k}{k} \\ &= (b + \mathbb{E}[T])\,\textsc{Static}_k.\end{aligned}$$

The second inequality holds because no arc from $V \setminus U_{T+1}$ to U_{T+1} is live, while, conditional on U_{T+1}, the arcs of $A[U_{T+1}]$ keep their original distribution. Thus, $C(\pi) \cap U_{T+1}$ is reached only through arcs of $A[U_{T+1}]$, and the remaining feedback of π depends only on arcs independent of $A[U_{T+1}]$, which act as external randomness. Hence, π induces an adaptive policy with budget k on $(U_{T+1}, \mathbf{p}|_{U_{T+1}})$.

Choosing $b = \sqrt{\log n}$ and applying Lemma 8 gives

$$\mathbb{E}[\sigma(\pi)] \leq \left(\sqrt{\log n} + \left\lceil \frac{2\log n}{\log\log n} \right\rceil + \frac{\sqrt{\log n}}{\sqrt{\log n} - 1}\right) \textsc{Static}_k = \mathcal{O}_{n\to\infty}\left(\frac{\log n}{\log\log n}\right) \textsc{Static}_k.$$

Taking π to be an optimal adaptive policy concludes the proof. □

4 A Constant Bound for Bi-Directed Forests

An instance $(V, \mathbf{p})$ is a *bi-directed forest* if the undirected graph on V with an edge $\{u, v\}$ for each pair of distinct vertices with $p_{u,v} > 0$ or $p_{v,u} > 0$ is a forest. Opposite arcs are live independently and may have different probabilities. This includes every orientation of a forest (such as in- or out-arborescences) and, by the equivalence in Section 2, undirected forests. We show that the adaptivity gap of every such instance is at most 2. We use a standard result on packing and covering subtrees; see Gyárfás and Lehel [GL70]. We include a short proof for completeness.

Lemma 9. *Let $\mathcal{T} = \{T_1, \ldots, T_m\}$ be a family of nonempty subtrees of a forest F, and let $t \geq 1$ be an integer. Then either $\mathcal{T}$ contains t pairwise vertex-disjoint members, or there exists a set of at most $t-1$ vertices of F intersecting every member of $\mathcal{T}$.*

Proof. We root each component of F and begin with an empty family $\mathcal{L}$ of subtrees and an empty set L of vertices. While some member of $\mathcal{T}$ is disjoint from L, we choose such a member T whose vertex v closest to the root is as deep as possible. Every member disjoint from L that intersects T contains v; otherwise, its vertex closest to the root would lie strictly below v, contradicting our choice of T. We add T to $\mathcal{L}$ and v to L.

Thus, the members of $\mathcal{L}$ are pairwise vertex-disjoint, and upon termination L intersects every member of $\mathcal{T}$. Each step adds one new element to each set, so $|\mathcal{L}| = |L|$. If $|\mathcal{L}| \geq t$, the first alternative holds. Otherwise, $|L| = |\mathcal{L}| \leq t - 1$, giving the second alternative. □

As in Section 3, we first deplete vertices with large expected residual influence and then bound the additional influence of k adaptive selections. For forests, Lemma 9 provides a fixed set L of at most $k-1$ seeds whose removal leaves an instance in which any single seed has expected influence at most $\textsc{Static}_k/k$. Seeding L contributes at most $\textsc{Static}_k$ in expectation, and so do the k further adaptive selections in total, giving the factor 2.

Proof of Theorem 3. Let $(V, \mathbf{p})$ be a bi-directed forest instance and let F be its underlying undirected forest. As in Section 3, the instance restricted to $U \subseteq V$ is $(U, \mathbf{p}|_U)$. Identify each subtree of F

with its vertex set, and let $\mathcal{T}$ be the family of nonempty subtrees T satisfying $\textsc{Static}_1(T, \mathbf{p}|_T) > \textsc{Static}_k(V, \mathbf{p})/k$. There cannot be k pairwise vertex-disjoint members of $\mathcal{T}$. Otherwise, choosing an optimal seed in each would give expected influence greater than $\textsc{Static}_k$, since their cascades (each limited to its own subtree) are disjoint.

Applying Lemma 9 to F gives a set L of at most $k-1$ vertices intersecting every member of $\mathcal{T}$. Every component of $F \setminus L$ is a subtree disjoint from L and therefore does not belong to $\mathcal{T}$. Since every directed live path avoiding L stays within one such component, $\textsc{Static}_1(V \setminus L, \mathbf{p}|_{V\setminus L}) \leq \textsc{Static}_k(V, \mathbf{p})/k$.

Fix an adaptive policy $\pi \in \Pi_k$ and compare its activations with $C(L)$ in the same realization. The vertices activated by π inside $C(L)$ contribute at most $\sigma(L)$. At each step, a directed live path from the seed selected by π to a newly activated vertex outside $C(L)$ avoids both L and all vertices activated earlier by π. Consequently, each selection contributes at most $\textsc{Static}_1(V \setminus L, \mathbf{p}|_{V\setminus L})$ expected new activations outside $C(L)$. Since $|L| \leq k-1$, summing over the k selections gives

$$\mathbb{E}[\sigma(\pi)] \leq \mathbb{E}[\sigma(L)] + k\,\textsc{Static}_1(V \setminus L, \mathbf{p}|_{V\setminus L}) \leq \textsc{Static}_{k-1} + k\frac{\textsc{Static}_k}{k} \leq 2\textsc{Static}_k. \qquad \square$$

5 Conclusion

We established new upper bounds on the adaptivity gap of influence maximization under full-adoption feedback, for undirected graphs, directed graphs, and bi-directed forests. Our proofs develop two approaches. The first uses percolation inequalities to control the dependence between cascades. The second uses an expected depletion argument to show that a few rounds of static seeding leave only vertices of small expected influence, which bounds the influence of an adaptive policy. We hope these ideas will also prove useful in other stochastic diffusion problems over graphs.

Several questions remain open. First, can we find analogs of the percolation inequalities for directed graphs? Some direct analogs of the tools used in this work fail in that setting. It would be of independent interest to further extend some of this machinery to directed graphs. Second, and most importantly, is the adaptivity gap bounded by a universal constant? The current lower bounds, for all graph classes considered in this work, are constants. This raises the question of whether better upper bounds can be achieved or whether the lower bounds can be improved. Finally, can our methods be used to improve adaptivity-gap bounds for other diffusion models, such as the linear threshold model?

6 AI Methodology

We developed a $\mathcal{O}(\log n)$ bound for directed graphs in Section 3 independently and used ChatGPT to optimize it to $\mathcal{O}(\log n/\log\log n)$. We also used ChatGPT to investigate the connection between influence maximization and recent breakthroughs on percolation inequalities and write the first draft of the proof. We substantially rewrote that proof, and the argument presented in Section 2 differs considerably from the initial proposal.

We also used ChatGPT to help evaluate the mathematical statements and proofs of results from prior work used in this paper, and to revise the writing and proofs throughout the paper, and Claude to verify references and proofread the manuscript; the authors independently verified all mathematical claims and take full responsibility for all the results.

References

[AN16] Arash Asadpour and Hamid Nazerzadeh. "Maximizing Stochastic Monotone Submodular Functions". In: *Management Science* 62.8 (Aug. 2016), pp. 2374–2391. DOI: 10.1287/mnsc.2015.2254.

[BBCL14] Christian Borgs, Michael Brautbar, Jennifer Chayes, and Brendan Lucier. "Maximizing social influence in nearly optimal time". In: *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms*. SIAM, 2014, pp. 946–957. DOI: 10.1137/1.9781611973402.70.

[BCDJ13] Abhijit Banerjee, Arun G. Chandrasekhar, Esther Duflo, and Matthew O. Jackson. "The diffusion of microfinance". In: *Science* 341.6144 (2013), p. 1236498.

[BJP20] Suman Banerjee, Mamata Jenamani, and Dilip Kumar Pratihar. "A survey on influence maximization in a social network". In: *Knowledge and Information Systems* 62.9 (2020), pp. 3417–3455. DOI: 10.1007/s10115-020-01461-4.

[Bla26] Arthur Blanc-Renaudie. *Notes on Kozma and Nitzan's Proof of $\theta(p_c) = 0$*. arXiv:2610.10502v1 [math.PR]. 2026. URL: https://arxiv.org/abs/2610.10502v1.

[Bou26] Ahmed Bou-Rabee. *Percolation after Anthropic: Lean 4 Proofs of the Kozma–Nitzan Connection Inequalities and of Site Percolation at Criticality*. GitHub repository. 2026. URL: https://github.com/nitromannitol/percolation-after-anthropic/tree/cd090736b8f5984d49926e80363910e7c6441a4f.

[BSZ19] Domagoj Bradac, Sahil Singla, and Goran Zuzic. "(Near) Optimal Adaptivity Gaps for Stochastic Multi-Value Probing". In: *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2019)*. Ed. by Dimitris Achlioptas and László A. Végh. Vol. 145. Leibniz International Proceedings in Informatics (LIPIcs). Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2019, 49:1–49:21. DOI: 10.4230/LIPIcs.APPROX-RANDOM.2019.49.

[CP19] Wei Chen and Binghui Peng. "On Adaptivity Gaps of Influence Maximization Under the Independent Cascade Model with Full-Adoption Feedback". In: *30th International Symposium on Algorithms and Computation (ISAAC 2019)*. Ed. by Pinyan Lu and Guochuan Zhang. Vol. 149. Leibniz International Proceedings in Informatics (LIPIcs). Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2019, 24:1–24:19. DOI: 10.4230/LIPIcs.ISAAC.2019.24.

[CVZ10] Chandra Chekuri, Jan Vondrák, and Rico Zenklusen. "Dependent Randomized Rounding via Exchange Properties of Combinatorial Structures". In: *Proceedings of the 51st Annual IEEE Symposium on Foundations of Computer Science*. IEEE Computer Society, 2010, pp. 575–584. DOI: 10.1109/FOCS.2010.60.

[CWW10] Wei Chen, Chi Wang, and Yajun Wang. "Scalable influence maximization for prevalent viral marketing in large-scale social networks". In: *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2010, pp. 1029–1038. DOI: 10.1145/1835804.1835934.

[DGV04] Brian C. Dean, Michel X. Goemans, and Jan Vondrák. "Approximating the Stochastic Knapsack Problem: The Benefit of Adaptivity". In: *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. IEEE Computer Society, 2004, pp. 208–217. DOI: 10.1109/FOCS.2004.15.

[DPV21] Gianlorenzo D'Angelo, Debashmita Poddar, and Cosimo Vinci. "Improved Approximation Factor for Adaptive Influence Maximization via Simple Greedy Strategies". In: *48th International Colloquium on Automata, Languages, and Programming (ICALP 2021)*. Ed. by Nikhil Bansal, Emanuela Merelli, and James Worrell. Vol. 198. Leibniz International Proceedings in Informatics (LIPIcs). Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021, 59:1–59:19. DOI: 10.4230/LIPIcs.ICALP.2021.59.

[DPV23] Gianlorenzo D'Angelo, Debashmita Poddar, and Cosimo Vinci. "Better Bounds on the Adaptivity Gap of Influence Maximization under Full-Adoption Feedback". In: *Artificial Intelligence* 318 (May 2023), p. 103895. DOI: 10.1016/j.artint.2023.103895.

[DR01] Pedro Domingos and Matt Richardson. "Mining the network value of customers". In: *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2001, pp. 57–66. DOI: 10.1145/502512.502525.

[Fv17] Robert Fitzner and Remco van der Hofstad. "Mean-Field Behavior for Nearest-Neighbor Percolation in $d > 10$". In: *Electronic Journal of Probability* 22 (2017). Paper no. 43, pp. 1–65. DOI: 10.1214/17-EJP56.

[GK10] Daniel Golovin and Andreas Krause. "Adaptive Submodularity: A New Approach to Active Learning and Stochastic Optimization". In: *Proceedings of the 23rd Annual Conference on Learning Theory (COLT)*. arXiv:1003.3967v2. 2010.

[GK11] Daniel Golovin and Andreas Krause. "Adaptive Submodularity: Theory and Applications in Active Learning and Stochastic Optimization". In: *Journal of Artificial Intelligence Research* 42 (Nov. 2011), pp. 427–486. URL: https://jair.org/index.php/jair/article/view/10731.

[GL70] A. Gyárfás and J. Lehel. "A Helly-type problem in trees". In: *Combinatorial Theory and its Applications*. Ed. by P. Erdős, A. Rényi, and V. T. Sós. Vol. 4. Colloquia Mathematica Societatis János Bolyai. Amsterdam: North-Holland, 1970, pp. 571–584. URL: https://www.renyi.hu/~gyarfas/Cikkek/02_GyarfasLehel_AHellyTypeProblemInTrees.pdf.

[Gla24] Nikita Gladkov. *Percolation Inequalities and Decision Trees*. arXiv:2408.08457v2 [math.PR]. 2024. URL: https://arxiv.org/abs/2408.08457v2.

[GNS17] 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*. SIAM, 2017, pp. 1688–1702. DOI: 10.1137/1.9781611974782.111.

[Gou26] Jean-Baptiste Gouéré. *On the Proof of the Kozma–Nitzan Conjecture*. arXiv:2610.30063v1 [math.PR]. 2026. URL: https://arxiv.org/abs/2610.30063v1.

[Gri99] Geoffrey Grimmett. *Percolation*. 2nd. Vol. 321. Grundlehren der mathematischen Wissenschaften. Berlin: Springer, 1999. DOI: 10.1007/978-3-662-03981-6.

[Har60] Theodore E. Harris. "A Lower Bound for the Critical Probability in a Certain Percolation Process". In: *Proceedings of the Cambridge Philosophical Society* 56.1 (Jan. 1960), pp. 13–20. DOI: 10.1017/S0305004100034241.

[HS90] Takashi Hara and Gordon Slade. "Mean-Field Critical Behaviour for Percolation in High Dimensions". In: *Communications in Mathematical Physics* 128.2 (Mar. 1990), pp. 333–391. DOI: 10.1007/BF02108785.

[HSBB11] Oliver Hinz, Bernd Skiera, Christian Barrot, and Jan U. Becker. "Seeding strategies for viral marketing: An empirical comparison". In: *Journal of Marketing* 75.6 (2011), pp. 55–71.

[Jan+17] Jarosław Jankowski, Piotr Bródka, Przemysław Kazienko, Bolesław K. Szymański, Radosław Michalski, and Tomasz Kajdanowicz. "Balancing speed and coverage by sequential seeding in complex networks". In: *Scientific Reports* 7.1 (2017), p. 891.

[Kes80] Harry Kesten. "The Critical Probability of Bond Percolation on the Square Lattice Equals 1/2". In: *Communications in Mathematical Physics* 74.1 (Feb. 1980), pp. 41–59. DOI: 10.1007/BF01197577.

[KKT03] David Kempe, Jon Kleinberg, and Éva Tardos. "Maximizing the spread of influence through a social network". In: *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2003, pp. 137–146. DOI: 10.1145/956750.956769.

[KL14] Sanjeev Khanna and Brendan Lucier. "Influence Maximization in Undirected Networks". In: *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms*. Full version: https://www.cis.upenn.edu/~sanjeev/papers/soda14_undirected_influence.pdf. SIAM, 2014, pp. 1482–1496. DOI: 10.1137/1.9781611973402.109.

[KN24] Gady Kozma and Shahaf Nitzan. *A Reduction of the $\theta(p_c) = 0$ Problem to a Conjectured Inequality*. arXiv:2401.12397v1 [math.PR]. 2024. URL: https://arxiv.org/abs/2401.12397v1.

[Led26] Justin Leder. *$\theta(p_c) = 0$ for Bernoulli Bond Percolation on $\mathbb{Z}^d$ in All Dimensions $d \geq 2$, in Lean 4 / Mathlib*. GitHub repository. 2026. URL: https://github.com/anthropics/formal-math/tree/795efb86f191735c5481675763537cfb4ff37e55/percolation.

[Les+07] Jure Leskovec, Andreas Krause, Carlos Guestrin, Christos Faloutsos, Jeanne VanBriesen, and Natalie Glance. "Cost-effective outbreak detection in networks". In: *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2007, pp. 420–429.

[LFWT18] Yuchen Li, Ju Fan, Yanhao Wang, and Kian-Lee Tan. "Influence maximization on social graphs: A survey". In: *IEEE Transactions on Knowledge and Data Engineering* 30.10 (2018), pp. 1852–1872. DOI: 10.1109/TKDE.2018.2807843.

[LLZ25] Jian Li, Yinchen Liu, and Yiran Zhang. "Adaptivity gaps for stochastic probing with subadditive functions". In: *2025 IEEE 66th Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE. 2025, pp. 149–185.

[MR07] Elchanan Mossel and Sébastien Roch. "On the submodularity of influence in social networks". In: *Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing*. 2007, pp. 128–134.

[MR10] Elchanan Mossel and Sébastien Roch. "Submodularity of influence in social networks: From local to global". In: *SIAM Journal on Computing* 39.6 (2010), pp. 2176–2188. DOI: 10.1137/080714452.

[NWF78] George L. Nemhauser, Laurence A. Wolsey, and Marshall L. Fisher. "An analysis of approximations for maximizing submodular set functions—I". In: *Mathematical Programming* 14.1 (1978), pp. 265–294. DOI: 10.1007/BF01588971.

[Ope26] OpenAI. *Critical Bond and Site Percolation on the Cubic Lattice*. Preprint in the GitHub repository `openai/math`. 2026. URL: https://github.com/openai/math/blob/adc7f1241b42e322a6451854ab7e4b4c146bf78a/preprints/Critical-bond-and-site-percolation-on-the-cubic-lattice-September-24-2026/paper.pdf.

[PC19] Binghui Peng and Wei Chen. "Adaptive influence maximization with myopic feedback". In: *Advances in Neural Information Processing Systems*. Ed. by H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett. Vol. 32. Curran Associates, Inc., 2019. URL: https://proceedings.neurips.cc/paper/2019/hash/f6b5f8c32c65fee991049a55dc97d1ce-Abstract.html.

[RD02] Matthew Richardson and Pedro Domingos. "Mining knowledge-sharing sites for viral marketing". In: *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2002, pp. 61–70. DOI: 10.1145/775047.775057.

[TXS14] Youze Tang, Xiaokui Xiao, and Yanchen Shi. "Influence maximization: Near-optimal time complexity meets practical efficiency". In: *Proceedings of the 2014 ACM SIGMOD International Conference on Management of Data*. ACM, 2014, pp. 75–86. DOI: 10.1145/2588555.2593670.

[Von07] Jan Vondrák. "Submodularity in combinatorial optimization". PhD thesis. Univerzita Karlova, Matematicko-fyzikální fakulta, 2007. URL: https://dspace.cuni.cz/handle/20.500.11956/13738.

[Yad+16] Amulya Yadav, Hau Chan, Albert Xin Jiang, Haifeng Xu, Eric Rice, and Milind Tambe. "Using Social Networks to Aid Homeless Shelters: Dynamic Influence Maximization under Uncertainty". In: *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*. 2016, pp. 740–748.