

On the Efficiency of Influence-and-Exploit Strategies for Revenue Maximization under Positive Externalities^{*}

Dimitris Fotakis and Paris Siminelakis

School of Electrical and Computer Engineering,
National Technical University of Athens, 157 80 Athens, Greece
fotakis@cs.ntua.gr, psiminelakis@gmail.com

Abstract. We consider the marketing model of (Hartline, Mirrokni, Sundararajan, WWW '08) for selling a digital product in a social network under positive externalities. The seller seeks for a marketing strategy, namely an ordering in which he approaches the buyers and the prices offered to them, that maximizes her revenue. We restrict our attention to the Uniform Additive Model of externalities, and mostly focus on Influence-and-Exploit (IE) marketing strategies. We show that in undirected social networks, revenue maximization is NP-hard not only when we search for a general optimal marketing strategy, but also when we search for the best IE strategy. Rather surprisingly, we observe that allowing IE strategies to offer prices smaller than the myopic price in the exploit step leads to a significant improvement on their performance. Thus, we show that the best IE strategy approximates the maximum revenue within a factor of 0.911 for undirected and of roughly 0.553 for directed networks. Utilizing a connection between good IE strategies and large cuts in the underlying social network, we obtain polynomial-time algorithms that approximate the revenue of the best IE strategy within a factor of roughly 0.9. Hence, we significantly improve on the best known approximation ratio for the maximum revenue to 0.8229 for undirected and to 0.5011 for directed networks (from $2/3$ and $1/3$, respectively).

1 Introduction

Understanding the flow of information, influence, and epidemics through the social fabric has become increasingly important due to the high interconnectedness brought about by technological advances. The digitization of communications (e.g., cell phones, emails, text messages) and of the social interaction (e.g., Facebook, Twitter) not only has provided the researchers with a strong empirical footing upon which they can base their theories and test their predictions, but also has opened the frontier of algorithmic applications related to social networks. Particularly, there has been a shift from aggregate descriptive theories, in the spirit of *Diffusion of Innovations*, to models incorporating the structure of social networks, culminating with the algorithmic paradigm of *Influence Maximization*.

^{*} This work was supported by the project AlgoNow, co-financed by the European Union (European Social Fund - ESF) and Greek national funds, through the Operational Program "Education and Lifelong Learning", under the research funding program Thales, and by an NTUA Basic Research Grant (PEBE 2009).

Firms operating in such a reticular environment, where information about products and services diffuses rapidly between individuals, have acknowledged the importance of revisiting their approach. The availability of information about users and the mitigated effectiveness of traditional forms of marketing occasion the need for intelligent marketing strategies. Towards realizing this goal, there are three main challenges: mining individual preferences, quantifying the influence that buyers exert upon each other, and fusing these information along a marketing strategy. The ideal solution would be an algorithm that intelligently adjusts its actions (e.g., prices, individuals to approach) based on the current state of the network, and maximizes the seller's revenue.

In this work, we are interested in the latter challenge of designing efficient marketing strategies that exploit the positive influence between buyers. We focus on the setting where the utility of the product depends inherently on the scale of the product's adoption, e.g., the value of a social network depends on the fraction of the population using it on a regular basis. In fact, for many products, their value to a buyer depends on the set of her friends using them (e.g., cell phones, online gaming). In the presence of such positive externalities between the potential buyers, the seller seeks for a marketing strategy that guarantees a significant revenue through a wide adoption of the product, which leads to an increased value, and consequently, to a profitable pricing of it.

Marketing Model. More formally, we adopt the model of Hartline, Mirrokni, and Sundararajan [14], where a digital product is sold to a set of potential buyers under positive externalities. We assume an unlimited supply of the product and that there is no production cost for it. A (possibly directed) weighted social network $G(V, E, w)$ on the set V of potential buyers models how their value of the product is affected by other buyers who already own the product. Specifically, an edge $(j, i) \in E$ denotes that the event that j owns the product has a positive influence on i 's value of the product. The strength of this influence is quantified by a non-negative weight w_{ji} associated with edge (j, i) . Also, buyer i may have an intrinsic value of the product, quantified by a non-negative weight w_{ii} . The product's value to each buyer i is given by a non-decreasing function $v_i : 2^{N_i} \mapsto \mathbb{R}_+$, which depends on w_{ii} and on the set $S \subseteq N_i$ of i 's neighbors who already own the product, where $N_i = \{j \in V \setminus \{i\} : (j, i) \in E\}$. The exact values $v_i(S)$ are unknown and are treated as random variables of which only the distributions $F_{i,S}$ are known to the seller. In particular, we assume that for each buyer i and each set $S \subseteq N_i$, the seller only knows the probability distribution $F_{i,S}(x) = \mathbb{Pr}[v_i(S) < x]$ that buyer i rejects an offer of price x for the product.

Regarding the distribution of $v_i(S)$'s, the most interesting cases outlined in [14] are: (i) the *Concave Graph Model*, where the weights w_{ji} are random variables, and the values $v_i(S)$ are determined by a concave function of the total influence $M_{i,S} = \sum_{j \in S \cup \{i\}} w_{ji}$ perceived by buyer i from the set S of her neighbors owning the product, and (ii) the *Uniform Additive Model*, where the weights w_{ji} are deterministic, and the values $v_i(S)$ are uniformly distributed in $[0, M_{i,S}]$. In this work, we restrict our attention to the Uniform Additive Model, which can be regarded as an extension of the widely accepted Linear Threshold Model of social influence [15]. Though technically simpler, the Uniform Additive Model incorporates all the main features of the marketing model of [14]. An important special case of the Uniform Additive Model is the undirected (or the symmetric) case, where $w_{ij} = w_{ji}$ for all edges $\{i, j\}$ of the social network.

In this setting, the seller approaches each potential buyer once and makes an offer to him. Thus, a *marketing strategy* (π, \mathbf{x}) consists of a permutation π of the buyers and a pricing vector $\mathbf{x} = (x_1, \dots, x_n)$, where π determines the order in which the buyers are approached and \mathbf{x} the prices offered to them. Given the set S of i 's neighbors who own the product when the seller approaches her, buyer i accepts the offer with probability $1 - F_{i,S}(x_i)$, in which case she pays the price x_i , or rejects it, with probability $F_{i,S}(x_i)$, in which case she pays nothing and never receives an offer again. The seller's goal is to compute a marketing strategy (π, \mathbf{x}) that maximizes her expected revenue, namely the total amount paid by the buyers who accept the offer.

Previous Work. Using a transformation from Maximum Acyclic Subgraph, Hartline et al. [14] proved that if we have complete knowledge of the buyers' valuations, computing a revenue-maximizing ordering is NP-hard for directed social networks. Combined with the result of [12], this suggests an upper bound of 0.5 on the approximation ratio of revenue maximization for directed networks and deterministic additive valuations. On the positive side, they gave a polynomial-time dynamic programming algorithm for a fully symmetric special case, where the order of the buyers is insignificant.

An interesting contribution of [14] is a class of elegant marketing strategies called *Influence-and-Exploit* (IE). An IE strategy first offers the product for free to a selected subset of buyers, aiming to increase the value of the product to the remaining buyers (influence step). Then, in the exploit step, it approaches the remaining buyers, in a random order, and offers them the product at the so-called *myopic price*. The myopic price ignores the current buyer's influence on the subsequent buyers and maximizes the expected revenue extracted from her. In the Uniform Additive Model, each buyer accepts the myopic price with probability $1/2$. Hence, there is a notion of uniformity in the prices offered in the exploit step, in the sense that the buyers accept them with a fixed probability, and we can say that the IE strategy uses a *pricing probability* of $1/2$.

As for the revenue extracted by IE strategies compared against the maximum revenue extracted by general marketing strategies, Hartline et al. [14] proved that the best IE strategy approximates the maximum revenue within a factor of 0.25 for the Concave Graph Model, which improves to $\frac{e}{4e-2} \approx 0.306$ if the distributions $F_{i,S}$ satisfy the monotone hazard rate condition, and within a factor of 0.94 for the (polynomially solvable) fully symmetric case of the Uniform Additive Model. Combined with the recent algorithm of [16] for unconstrained submodular maximization, which can be used to approximate the revenue of the best IE strategy within a factor of 0.5, the results of [14] imply an approximation ratio of 0.125 for the maximum revenue in the Concave Graph Model, which improves to 0.153 if the distributions $F_{i,S}$ satisfy the monotone hazard rate condition. As for the Uniform Additive Model, Hartline et al. [14] proved that if each buyer is selected in the influence set randomly, with an appropriate probability, the expected revenue of IE is at least $2/3$ (resp. $1/3$) times the maximum revenue of undirected (resp. directed) networks. Since [14], the Influence-and-Exploit paradigm has been applied to a few other settings where one seeks to maximize revenue in the presence of positive externalities (see e.g. [4,5,13]).

Contribution and Techniques. Although IE strategies are simple, elegant, and promising in terms of efficiency, their performance against the maximum revenue and their approximability are not well understood. Moreover, the absence of any strong bounds

on the fraction of the maximum revenue extracted by the best IE strategy and the poor approximation ratios for the maximum revenue in the Concave Graph Model suggest looking into simpler cases of the model. This is also suggested by previous work on Influence Maximization, where focusing on simpler cases provides insights, which, in turn, can enhance our understanding of more general settings. In this work, we focus on the important case of the Uniform Additive Model, and obtain a comprehensive collection of results on the efficiency and the approximability of IE strategies. Our results also imply a significant improvement on the best known approximation ratio for revenue maximization in the Uniform Additive Model.

We first show that in the Uniform Additive Model, revenue maximization is **NP**-hard for undirected networks¹ not only when we search for a general optimal marketing strategy, but also when we search for the best IE strategy. Next, we embark on a systematic study of the algorithmic properties of IE strategies (Section 3). In [14], IE strategies are restricted, by definition, to the myopic pricing probability, which for the Uniform Additive Model is $1/2$. Rather surprisingly, we observe that we can achieve a significant improvement on the efficiency of IE strategies if we use smaller prices (equivalently, a larger pricing probability) in the exploit step. Thus, we let IE strategies use a carefully selected pricing probability $p \in [1/2, 1)$.

We prove the existence of an IE strategy with pricing probability 0.586 (resp. $2/3$) which approximates the maximum revenue, extracted by an unrestricted marketing strategy, within a factor of 0.911 for undirected (resp. 0.55289 for directed) networks. The proof assumes a revenue-maximizing pricing probability vector \mathbf{p} and constructs an IE strategy with the desired expected revenue by applying randomized rounding to \mathbf{p} . An interesting consequence is that the upper bound of 0.5 on the approximation ratio of the maximum revenue for directed networks does not apply to the Uniform Additive Model. In Section 3, we discuss the technical reasons behind this and show a pair of upper bounds on the approximation ratio achievable for directed networks. Specifically, assuming the Unique Games conjecture, we show that it is **NP**-hard to approximate the maximum revenue within a factor greater than $27/32$, if we use any marketing strategy, and greater than $3/4$, if we are restricted to IE strategies with pricing probability $2/3$.

The technical intuition behind most of our results comes from the apparent connection between good IE strategies and large cuts in the underlying social network. Following this intuition, we optimize the parameters of the random-partitioning IE strategy of [14] and slightly improve the approximation ratio to 0.686 (resp. 0.343) for undirected (resp. directed) networks. Building on the idea of generating revenue from large cuts in the network, we discuss, in Section 4, a natural generalization of IE strategies that use more than two pricing classes. We show that a simple random partitioning of the buyers in six pricing classes further improves the approximation ratio for the maximum revenue to 0.7032 for undirected networks and to 0.3516 for directed social networks.

The main hurdle in obtaining better approximation guarantees for the maximum revenue problem is the lack of any strong upper bounds on it. In Section 5, we introduce a strong Semidefinite Programming (SDP) relaxation for the problem of computing the

¹ If the seller has complete knowledge of the buyers' valuations, finding a revenue-maximizing ordering for undirected networks is polynomially solvable (Lemma 1). Therefore, the reduction of [14] does not imply that revenue maximization for undirected networks is **NP**-hard.

best IE strategy with any given pricing probability. Our approach exploits the resemblance between computing the best IE strategy and the problems of MAX-CUT and MAX-DICUT, and builds on the elegant approach of Goemans and Williamson [11] and Feige and Goemans [8]. Solving the SDP relaxation and using randomized rounding, we obtain a 0.9032 (resp. 0.9064) approximation for the best IE strategy with a pricing probability of 0.586 for undirected networks (resp. of $2/3$ for directed networks). Combining these results with the bounds on the fraction of the maximum revenue extracted by the best IE strategy, we significantly improve on the best known approximation ratio for revenue maximization to 0.8229 for undirected networks and 0.5011 for directed networks (from $2/3$ and $1/3$, respectively, in [14]). To the best of our knowledge, this is the first time an (approximate) SDP relaxation for a pricing model under positive externalities is suggested and exploited to improve the approximation ratio for the corresponding revenue (or welfare) maximization problem. Actually, we believe that our SDP-based approach may find applications to other pricing models under externalities.

Other Related Work. Our work lies in the area of pricing and revenue maximization under positive externalities, and more generally, in the area of social contagion and influence maximization (see e.g., [7,15]). Recent research has studied the impact of externalities in a variety of settings (see e.g. [14,4,1,3,6,5,13,9]). Hartline et al. [14] were the first to consider social influence in the framework of revenue maximization. Since then, relevant research has focused either on posted price strategies, where there is no price discrimination, or on game theoretic considerations, where the buyers act strategically according to their value of the product. To the best of our knowledge, our work is the first that considers the approximability of the revenue extracted by an optimal strategy and by the best IE strategy, which were the central problems in [14].

Regarding posted pricing, Arthur et al. [4] considered a model where recommendations about the product cascade through the network from early adopters, and presented an IE-based $O(1)$ -approximation algorithm for the maximum revenue. Akhlaghpour et al. [1] considered iterative posted pricing, where all interested buyers can buy the product at the same price at a given time. They studied revenue maximization under two different repricing models allowing for at most k prices. They proved that if frequent repricing is allowed, revenue maximization is NP-hard to approximate, while if the repricing rate is limited, there is an FPTAS. Anari et al. [3] considered a posted price setting with historical externalities. Given a fixed price trajectory, the buyers decide when to buy the product. In this setting, they studied existence and uniqueness of equilibria, and presented an FPTAS for special cases of revenue maximization.

In a complementary direction, Chen et al. [6] investigated the (Bayesian-)Nash equilibria when each buyer's value of the product depends on the set of buyers who own the product. They focused on two classes of equilibria, pessimistic and optimistic ones, and showed how to compute these equilibria and how to find revenue-maximizing prices. Candogan et al. [5] investigated a scenario where a monopolist sells a divisible good to buyers under positive externalities. They considered a two-stage game where the seller first sets an individual price for each buyer, and then the buyers decide on their consumption level. They proved that the optimal price for each buyer is proportional to her Bonacich centrality, and that if the buyers are partitioned into two pricing classes (which is conceptually similar to IE), the problem is reducible to MAX-CUT.

2 The Model and Preliminaries

The Influence Model. The social network is a (possibly directed) weighted network $G(V, E, w)$ on the set V of potential buyers. There is a positive weight w_{ij} associated with each edge $(i, j) \in E$ (we assume that $w_{ij} = 0$ if $(i, j) \notin E$). A social network is undirected (or symmetric) if $w_{ij} = w_{ji}$ for all $i, j \in V$, and directed otherwise. There may exist a non-negative weight w_{ii} associated with each buyer i ². Each buyer i has a value $v_i : 2^{N_i} \mapsto \mathbb{R}_+$ of the product, which depends on w_{ii} and on the set $S \subseteq N_i$ of i 's neighbors who already own the product, where $N_i = \{j \in V \setminus \{i\} : (j, i) \in E\}$. However, the exact values $v_i(S)$ are unknown to the seller, who, for each buyer i and each set $S \subseteq N_i$, only knows the probability distribution $F_{i,S}(x) = \Pr[v_i(S) < x]$ that buyer i rejects an offer of price x for the product.

In the *Uniform Additive Model* [14, Section 2.1], the values $v_i(S)$ are drawn from the uniform distribution in $[0, M_{i,S}]$, where $M_{i,S} = \sum_{j \in S \cup \{i\}} w_{ji}$ is the total influence perceived by i by the set S of her neighbors owning the product. Then, the probability that buyer i rejects an offer of price x is $F_{i,S}(x) = x/M_{i,S}$.

Myopic Pricing. The *myopic price* disregards any externalities imposed by i on her neighbors, and simply maximizes the expected revenue extracted from buyer i , given that S is the current set of i 's neighbors who own the product. For the Uniform Additive Model, the myopic price is $M_{i,S}/2$, the probability that buyer i accepts it is $1/2$, and the expected revenue extracted from her with the myopic price is $M_{i,S}/4$, which is the maximum revenue one can extract from buyer i alone.

Marketing Strategies and Revenue Maximization. We can usually extract more revenue from G by employing a marketing strategy that exploits the positive influence between the buyers. A *marketing strategy* (π, \mathbf{x}) consists of a permutation π of the buyers and a pricing vector $\mathbf{x} = (x_1, \dots, x_n)$, where π determines the order in which the buyers are approached and \mathbf{x} the prices offered to them.

We observe that for any buyer i and any probability p that i accepts an offer, there is an (essentially unique) price x_p such that i accepts an offer of x_p with probability p . For the Uniform Additive Model, $x_p = (1 - p)M_{i,S}$ and the expected revenue extracted from buyer i with such an offer is $p(1 - p)M_{i,S}$. Throughout this paper, we equivalently regard marketing strategies as consisting of a permutation π of the buyers and a vector $\mathbf{p} = (p_1, \dots, p_n)$ of pricing probabilities. We note that if $p_i = 1$, i gets the product for free, while if $p_i = 1/2$, the price offered to i is (the myopic price of) $M_{i,S}/2$. We assume that $p_i \in [1/2, 1]$, since any expected revenue in $[0, M_{i,S}/4]$ can be achieved with such pricing probabilities. The expected revenue of a marketing strategy (π, \mathbf{p}) is:

$$R(\pi, \mathbf{p}) = \sum_{i \in V} p_i(1 - p_i) \left(w_{ii} + \sum_{j: \pi_j < \pi_i} p_j w_{ji} \right) \tag{1}$$

The problem of *revenue maximization* under the Uniform Additive Model is to find a marketing strategy (π^*, \mathbf{p}^*) that extracts a maximum revenue of $R(\pi^*, \mathbf{p}^*)$ from a given social network $G(V, E, w)$.

² Wlog., we ignore w_{ii} 's for directed networks, since we can replace each w_{ii} by an edge (i', i) of weight w_{ii} from a new node i' with a single outgoing edge (i', i) and no incoming edges.

Bounds on the Maximum Revenue. Let $N = \sum_{i \in V} w_{ii}$ and $W = \sum_{i < j} w_{ij}$, if the social network G is undirected, and $W = \sum_{(i,j) \in E} w_{ij}$, if G is directed. Then an upper bound on the maximum revenue of G is $R^* = (W + N)/4$, and follows by summing up the myopic revenue over all edges of G [14, Fact 1]. A lower bound on the maximum revenue is $(W + 2N)/8$ (resp. $(W + 4N)/16$), if G is undirected (resp. directed), and follows by approaching the buyers in any order (resp. in a random order) and offering them the myopic price. Thus, myopic pricing achieves an approximation ratio of 0.5 for undirected networks and of 0.25 for directed networks.

Ordering and NP-Hardness. Revenue maximization exhibits a dual nature involving optimizing both the pricing probabilities and the sequence of offers. For directed networks, finding a good ordering π of the buyers bears a resemblance to the Maximum Acyclic Subgraph problem, where given $G(V, E, w)$, we seek for an acyclic subgraph of maximum total edge weight. In fact, any permutation π of V corresponds to an acyclic subgraph of G that includes all edges going forward in π , i.e., all edges (i, j) with $\pi_i < \pi_j$. [14, Lemma 3.2] shows that given a directed network G and a pricing probability vector \mathbf{p} , computing an optimal ordering of the buyers (for the particular \mathbf{p}) is equivalent to computing a Maximum Acyclic Subgraph of G , with each edge (i, j) having a weight of $p_i p_j (1 - p_j) w_{ij}$. Consequently, computing an ordering π that maximizes $R(\pi, \mathbf{p})$ is NP-hard and Unique-Games-hard to approximate within a factor greater than 0.5 [12]. On the other hand, we can show that in the undirected case, if the pricing probabilities are given, we can easily compute the best ordering of the buyers.

Lemma 1. *Let $G(V, E, w)$ be an undirected social network, and let \mathbf{p} be any pricing probability vector. Then, approaching the buyers in non-increasing order of their pricing probabilities maximizes the revenue extracted from G under \mathbf{p} .*

Therefore, [14, Lemma 3.2] does not imply the NP-hardness of revenue maximization for undirected networks. The following lemma employs a reduction from monotone One-in-Three 3-SAT [10, LO4], and shows that revenue maximization is NP-hard for undirected networks.

Lemma 2. *Computing a marketing strategy that extracts the maximum revenue from an undirected social network is NP-hard.*

3 Influence-and-Exploit Strategies

An *Influence-and-Exploit* (IE) strategy $\text{IE}(A, p)$ consists of a set of buyers A receiving the product for free and a pricing probability p offered to the remaining buyers in $V \setminus A$, approached in a random order. We slightly abuse the notation, and let $\text{IE}(q, p)$ denote an IE strategy where each buyer is selected in A independently with probability q . For directed networks, $\text{IE}(A, p)$ extracts an expected (wrt the random ordering of the exploit set) revenue of:

$$R_{\text{IE}}(A, p) = p(1 - p) \sum_{i \in V \setminus A} \left(w_{ii} + \sum_{j \in A} w_{ji} + \sum_{j \in V \setminus A, j \neq i} \frac{p w_{ji}}{2} \right) \quad (2)$$

Specifically, $\text{IE}(A, p)$ extracts a revenue of $p(1 - p)w_{ji}$ from each edge (j, i) with $j \in A$ and $i \in V \setminus A$, and a revenue of $p^2(1 - p)w_{ji}$ from each edge (j, i) with both $j, i \in V \setminus A$, if j is before i in the random order, which happens with probability $1/2$.

The problem of finding the best IE strategy is to compute a subset of buyers A^* and a pricing probability p^* that extract a maximum revenue of $R_{\text{IE}}(A^*, p^*)$ from a given social network $G(V, E, w)$. The following lemma employs a reduction from monotone One-in-Three 3-SAT, and shows that computing the best IE strategy is NP-hard.

Lemma 3. *The problems of computing the best IE strategy and of computing the best IE strategy with a given pricing probability p , for any fixed $p \in [1/2, 1)$, are NP-hard, even for undirected networks.*

Simple IE strategies extract a significant fraction of the maximum revenue. E.g., for undirected networks, $R_{\text{IE}}(\emptyset, 2/3) = (4W + 6N)/27$, and $\text{IE}(\emptyset, 2/3)$ achieves an approximation ratio of $\frac{16}{27}$. Moreover, $\text{IE}(X, 1/2)$ extracts the maximum revenue from any simple undirected bipartite network $G(X, Y, E)$. For directed networks, $R_{\text{IE}}(\emptyset, 2/3) = (2W + 6N)/27$, and $\text{IE}(\emptyset, 2/3)$ achieves an approximation ratio of $\frac{8}{27}$. We next show that carefully selected IE strategies extract a larger fraction of the maximum revenue.

Exploiting Large Cuts. A natural idea is to exploit the apparent connection between a large cut in the social network and a good IE strategy. For example, in the undirected case, an IE strategy $\text{IE}(q, p)$ is conceptually similar to the randomized 0.5-approximation algorithm for MAX-CUT, which puts each node in set A with probability $1/2$. However, in addition to a revenue of $p(1 - p)w_{ij}$ from each edge $\{i, j\}$ in the cut $(A, V \setminus A)$, $\text{IE}(q, p)$ extracts a revenue of $p^2(1 - p)w_{ij}$ from each edge $\{i, j\}$ between nodes in the exploit set $V \setminus A$. Thus, to optimize the performance of $\text{IE}(q, p)$, we carefully adjust the probabilities q and p so that $\text{IE}(q, p)$ balances between the two sources of revenue. The proof of Proposition 1 extends the proof of [14, Theorem 3.1].

Proposition 1. *Let $G(V, E, w)$ be an undirected (resp. directed) social network, let $\lambda = N/W$, and let $q = \max\{1 - \frac{\sqrt{2}(2+\lambda)}{4}, 0\}$, Then, $\text{IE}(q, 2 - \sqrt{2})$ approximates the maximum revenue of G within a factor of 0.686 (resp. 0.343).*

On the Efficiency of Influence-and-Exploit. IE makes a rough discretization of the pricing space, and exploits the fact that the combinatorial structure of partitioning the vertices into two sets is well understood. Nevertheless, we are left with the nontrivial task of correlating the maximum revenue with only two prices and the maximum revenue with any set of prices. We next show that the best IE strategy, which is NP-hard to compute, manages to extract a significant fraction of the maximum revenue.

Theorem 1. *For any undirected social network, there exists an IE strategy with pricing probability 0.586 whose revenue is at least 0.9111 times the maximum revenue.*

Proof. We consider an undirected social network $G(V, E, w)$, start from the revenue-maximizing pricing probability vector \mathbf{p} , and obtain an IE strategy $\text{IE}(A, \hat{p})$ by applying randomized rounding to \mathbf{p} . We show that for $\hat{p} = 0.586$, the expected (wrt the randomized rounding choices) revenue of $\text{IE}(A, \hat{p})$ is at least 0.9111 times the revenue extracted from G by the best ordering for \mathbf{p} .

By Lemma 1, the best ordering is to approach the buyers in non-increasing order of pricing probabilities. Hence, we let $p_1 \geq \dots \geq p_n$, and let π be the identity permutation. Then,

$$R(\pi, \mathbf{p}) = \sum_{i \in V} p_i(1 - p_i)w_{ii} + \sum_{i < j} p_i p_j(1 - p_j)w_{ij}$$

For the IE strategy, we assign each buyer i to the influence set A independently with probability $I(p_i) = \alpha(p_i)(p_i - 0.5)$, and to the exploit set with probability $E(p_i) = 1 - I(p_i)$, where $\alpha(x) : [0.5, 1] \mapsto [0, 2]$ is a piecewise linear function with breakpoints at $(0.5, 0.7, 0.8, 0.9, 1.0)$ and values $(0.0, 1.0, 1.33, 1.63, 2.0)$ at these points. By linearity of expectation, the expected revenue of $\text{IE}(A, \hat{p})$ is:

$$R_{\text{IE}}(A, \hat{p}) = \sum_{i \in V} \hat{p}(1 - \hat{p})E(p_i)w_{ii} + \sum_{i < j} \hat{p}(1 - \hat{p})(I(p_i)E(p_j) + E(p_i)I(p_j) + \hat{p}E(p_i)E(p_j))w_{ij}$$

Specifically, $\text{IE}(A, \hat{p})$ extracts a revenue of $\hat{p}(1 - \hat{p})w_{ii}$ from each loop $\{i, i\}$, if i is included in the exploit set. Moreover, $\text{IE}(A, \hat{p})$ extracts a revenue of $\hat{p}(1 - \hat{p})w_{ij}$ from each edge $\{i, j\}$, $i < j$, if one of i, j is included in the influence set A and the other is not, and a revenue of $\hat{p}^2(1 - \hat{p})w_{ij}$ if both i and j are included in the exploit set $V \setminus A$.

The approximation ratio of $\text{IE}(A, \hat{p})$ to the maximum revenue of G under \mathbf{p} is derived as the minimum ratio between any pair of terms in $R(\pi, \mathbf{p})$ and $R_{\text{IE}}(A, \hat{p})$ corresponding to the same loop $\{i, i\}$ or to the same edge $\{i, j\}$. Therefore, the approximation ratio of $\text{IE}(A, \hat{p})$ is no less than the minimum of:

$$\min_{0.5 \leq x \leq 1} \frac{\hat{p}(1 - \hat{p})E(x)}{x(1 - x)} \quad \text{and} \quad \min_{0.5 \leq y \leq x \leq 1} \frac{\hat{p}(1 - \hat{p})(I(x)E(y) + E(x)I(y) + \hat{p}E(x)E(y))}{xy(1 - y)}$$

Using calculus, we can show that for $\hat{p} = 0.586$, these ratios are at least 0.9111. □

For directed networks, we use the same approach, and obtain the following theorem.

Theorem 2. *For any directed social network, there is an IE strategy with pricing probability $2/3$ whose expected revenue is at least 0.55289 times the maximum revenue.*

Proof sketch. Working as in the proof of Theorem 1, we show that the approximation ratio of the IE strategy obtained by applying randomized rounding to the revenue-maximizing pricing probability vector is at least:

$$\min_{0.5 \leq x, y \leq 1} \frac{\hat{p}(1 - \hat{p})(I(x)E(y) + 0.5\hat{p}E(x)E(y))}{xy(1 - y)}$$

For $\hat{p} = 2/3$ and $\alpha(x) = 1.0$, for all x , this is simplified to $\min_{y \in [0.5, 1]} \frac{2(3-2y)}{27y(1-y)}$, which attains its minimum value of ≈ 0.55289 at $y = \frac{3-\sqrt{3}}{2}$. □

Similarly, we can show that there is an IE strategy that uses the myopic pricing probability of $1/2$ and extracts a revenue of at least 0.8857 (resp. 0.4594) times the maximum revenue for undirected (resp. directed) social networks.

On the Approximability of the Maximum Revenue for Directed Networks. The results of [14, Lemma 3.2] and [12] suggest that given a pricing probability vector \mathbf{p} , it is Unique-Games-hard to compute a vertex ordering π of a directed network G for which $R(\pi, \mathbf{p})$ is at least 0.5 times the maximum revenue of G under \mathbf{p} . An interesting consequence of Theorem 2 is that the inapproximability bound of 0.5 does not apply to revenue maximization in the Uniform Additive Model. In particular, given the prices \mathbf{p} , Theorem 2 computes, in linear time, an IE strategy with an expected revenue of at least 0.55289 times the maximum revenue of G under \mathbf{p} . This does not contradict the results of [14,12], because the pricing probabilities of the IE strategy are different from \mathbf{p} .

In the Uniform Additive Model, different acyclic (sub)graphs (equivalently, different vertex orderings) allow for a different fraction of their edge weight to be translated into revenue, while in the reduction of [14, Lemma 3.2], the weight of each edge in an acyclic subgraph is equal to its revenue. Thus, although the IE strategy of Theorem 2, with pricing probability $2/3$, gives a 0.55289-approximation to the maximum revenue of G under \mathbf{p} , its vertex ordering combined with \mathbf{p} may generate a revenue of less than 0.5 times the maximum revenue of G under \mathbf{p} . Next, we obtain a pair of inapproximability results for revenue maximization in the Uniform Additive Model.

Lemma 4. *Assuming the Unique Games conjecture, it is NP-hard to approximate within a factor greater than $27/32$ (resp. to compute an IE strategy with pricing probability $2/3$ that approximates within a factor greater than $3/4$) the maximum revenue of a directed social network in the Uniform Additive Model.*

4 Generalized Influence-and-Exploit

Building on the idea of generating revenue from large cuts between pricing classes, we obtain a class of generalized IE strategies, which employ a partition of buyers in more than two pricing classes. A generalized IE strategy consists of $K \geq 3$ classes. Each class $k, k = 1, \dots, K$, is associated with a pricing probability of $p_k = 1 - \frac{k-1}{2^{(K-1)}}$, and each buyer is assigned to the class k independently with probability q_k , where $\sum_{k=1}^K q_k = 1$, and is offered a pricing probability of p_k . The buyers are considered in non-increasing order of pricing probability, i.e., the buyers in class k are considered before the buyers in class $k + 1$, and the buyers in the same class are considered in a random order.

Let $\text{IE}(\mathbf{q}, \mathbf{p})$ be such a generalized IE strategy, where $\mathbf{q} = (q_1, \dots, q_K)$ is the assignment probability vector and $\mathbf{p} = (p_1, \dots, p_K)$ is the pricing probability vector. We can show that the approximation ratio of $\text{IE}(\mathbf{q}, \mathbf{p})$ for undirected networks is at least:

$$\min \left\{ 4 \sum_{k=1}^K q_k p_k (1 - p_k), 4 \sum_{k=1}^K q_k p_k (1 - p_k) \left(q_k p_k + 2 \sum_{\ell=1}^{k-1} q_\ell p_\ell \right) \right\}, \quad (3)$$

while for directed social networks, the approximation ratio of $\text{IE}(\mathbf{q}, \mathbf{p})$ is at least half of the quantity in (3). We can now select the assignment probability vector \mathbf{q} so that (3) is maximized. With the pricing probability vector \mathbf{p} fixed, this involves maximizing a quadratic function of \mathbf{q} over linear constraints. Thus, we obtain the following:

Theorem 3. For any undirected (directed) network G , the generalized IE strategy with $K = 6$ classes and assignment probabilities $\mathbf{q} = (0.183, 0.075, 0.075, 0.175, 0.261, 0.231)$ approximates the maximum revenue of G within a factor of 0.7032 (0.3516).

5 Influence-and-Exploit via Semidefinite Programming

The main hurdle in obtaining better approximation guarantees for the maximum revenue is the loose upper bound of $(N + W)/4$. We do not know how to obtain a stronger upper bound on the maximum revenue. However, in this section, we obtain a Semidefinite Programming (SDP) relaxation for the problem of computing the best IE strategy with any given pricing probability $p \in [1/2, 1)$. Our approach exploits the resemblance between computing the best IE strategy and the problems of MAX-CUT (for undirected networks) and MAX-DICUT (for directed networks), and builds on the approach of [11,8]. Solving the SDP relaxation and using randomized rounding, we obtain, in polynomial time, a good approximation to the best influence set for the given p . Then, employing the bounds of Theorems 1 and 2, we obtain strong approximation guarantees for the maximum revenue in both directed and undirected networks.

Directed Social Networks. The case of a directed network $G(V, E, w)$ is a bit simpler, because we can ignore loops (i, i) without loss of generality. We observe that for any given pricing probability $p \in [1/2, 1)$, the problem of computing the best IE strategy $\text{IE}(A, p)$ is equivalent to solving the following Quadratic Integer Program:

$$\begin{aligned} \max \quad & \frac{p(1-p)}{4} \sum_{(i,j) \in E} w_{ij} \left(1 + \frac{p}{2} + (1 - \frac{p}{2})y_0y_i - (1 + \frac{p}{2})y_0y_j - (1 - \frac{p}{2})y_iy_j \right) \quad (\text{Q1}) \\ \text{s.t.} \quad & y_i \in \{-1, 1\} \qquad \qquad \qquad \forall i \in V \cup \{0\} \end{aligned}$$

In (Q1), there is a variable y_i for each buyer i and an additional variable y_0 denoting the influence set A . A buyer i is assigned to A , if $y_i = y_0$, and to the exploit set, otherwise. For each edge (i, j) , $1 + y_0y_i - y_0y_j - y_iy_j$ is 4, if $y_i = y_0 = -y_j$ (i.e., if i is assigned to the influence set and j is assigned to the exploit set), and 0, otherwise. Moreover, $\frac{p}{2}(1 - y_0y_i - y_0y_j + y_iy_j)$ is $2p$, if $y_i = y_j = -y_0$ (i.e., if both i and j are assigned to the exploit set), and 0, otherwise. Therefore, the contribution of each edge (i, j) to the objective function of (Q1) is equal to the revenue extracted from (i, j) by $\text{IE}(A, p)$.

Following the approach of [11,8], we relax (Q1) to the following Semidefinite Program, where $v_i \cdot v_j$ denotes the inner product of vectors v_i and v_j :

$$\begin{aligned} \max \quad & \frac{p(1-p)}{4} \sum_{(i,j) \in E} w_{ij} \left(1 + \frac{p}{2} + (1 - \frac{p}{2})v_0 \cdot v_i - (1 + \frac{p}{2})v_0 \cdot v_j - (1 - \frac{p}{2})v_i \cdot v_j \right) \\ \text{s.t.} \quad & v_i \cdot v_j + v_0 \cdot v_i + v_0 \cdot v_j \geq -1 \qquad \qquad \qquad (\text{S1}) \\ & v_i \cdot v_j - v_0 \cdot v_i - v_0 \cdot v_j \geq -1 \\ & -v_i \cdot v_j - v_0 \cdot v_i + v_0 \cdot v_j \geq -1 \\ & -v_i \cdot v_j + v_0 \cdot v_i - v_0 \cdot v_j \geq -1 \\ & v_i \cdot v_i = 1, \quad v_i \in \mathbb{R}^{n+1} \qquad \qquad \qquad \forall i \in V \cup \{0\} \end{aligned}$$

Any feasible solution to (Q1) can be translated into a feasible solution to (S1) by setting $v_i = v_0$, if $y_i = y_0$, and $v_i = -v_0$, otherwise. An optimal solution to (S1) can be computed within any precision ε in time polynomial in n and in $\ln \frac{1}{\varepsilon}$ (see e.g. [2]).

Given a directed social network $G(V, E, w)$, a pricing probability p , and a parameter $\gamma \in [0, 1]$, the algorithm SDP-IE(p, γ) first computes an optimal solution v_0, v_1, \dots, v_n to (S1). Then, following [8], the algorithm maps each vector v_i to a rotated vector v'_i which is coplanar with v_0 and v_i , lies on the same side of v_0 as v_i , and forms an angle with v_0 equal to $f_\gamma(\theta_i) = (1 - \gamma)\theta_i + \gamma\pi(1 - \cos \theta_i)/2$, where $\pi = 3.14\dots$ and $\theta_i = \arccos(v_0 \cdot v_i)$ is the angle of v_0 and v_i . Finally, the algorithm computes a random vector r uniformly distributed on the unit $(n + 1)$ -sphere, and assigns each buyer i to the influence set A , if $\text{sgn}(v'_i \cdot r) = \text{sgn}(v_0 \cdot r)$, and to the exploit set $V \setminus A$, otherwise where $\text{sgn}(x) = 1$, if $x \geq 0$, and -1 , otherwise. We next show that:

Theorem 4. *For any directed social network G , SDP-IE($2/3, 0.722$) approximates the maximum revenue extracted from G by the best IE strategy with pricing probability $2/3$ within a factor of 0.9064.*

Proof. We let v_0, v_1, \dots, v_n be an optimal solution to (S1), let $\theta_{ij} = \arccos(v_i \cdot v_j)$ be the angle of any two vectors v_i and v_j , and let $\theta_i = \arccos(v_0 \cdot v_i)$ be the angle of v_0 and any vector v_i . Similarly, we let $\theta'_{ij} = \arccos(v'_i \cdot v'_j)$ be the angle of any two rotated vectors v'_i and v'_j , and let $\theta'_i = \arccos(v_0 \cdot v'_i)$ be the angle of v_0 and any rotated vector v'_i . Building on the proof of [11, Lemma 7.3.2], we can show that:

Lemma 5. *The IE strategy of SDP-IE(p, γ) extracts from each edge (i, j) an expected revenue of:*

$$w_{ij} p(1 - p) \frac{(1 - \frac{p}{2}) \theta'_{ij} - (1 - \frac{p}{2}) \theta'_i + (1 + \frac{p}{2}) \theta'_j}{2\pi} \tag{4}$$

Since (S1) is a relaxation of the problem of computing the best IE strategy with pricing probability p , the revenue of an optimal IE(A, p) strategy is at most:

$$\frac{p(1-p)}{4} \sum_{(i,j) \in E} w_{ij} \left(1 + \frac{p}{2} + (1 - \frac{p}{2}) \cos \theta_i - (1 + \frac{p}{2}) \cos \theta_j - (1 - \frac{p}{2}) \cos \theta_{ij} \right) \tag{5}$$

On the other hand, by Lemma 5 and linearity of expectation, the IE strategy computed by SDP-IE(p, γ) generates an expected revenue of:

$$\frac{p(1-p)}{2\pi} \sum_{(i,j) \in E} w_{ij} \left((1 - \frac{p}{2}) \theta'_{ij} - (1 - \frac{p}{2}) \theta'_i + (1 + \frac{p}{2}) \theta'_j \right) \tag{6}$$

We recall that for each i , $\theta'_i = f_\gamma(\theta_i)$. In [8, Section 4], it is shown that for each i, j ,

$$\begin{aligned} \theta'_{ij} &= \arccos\left(\cos f_\gamma(\theta_i) \cos f_\gamma(\theta_j) + \frac{\cos \theta_{ij} - \cos \theta_i \cos \theta_j}{\sin \theta_i \sin \theta_j} \sin f_\gamma(\theta_i) \sin f_\gamma(\theta_j) \right) \\ &\equiv g_\gamma(\theta_{ij}, \theta_i, \theta_j) \end{aligned}$$

The approximation ratio of SDP-IE(p, γ) is derived as the minimum ratio of any pair of terms in (6) and (5) corresponding to the same edge (i, j) . Thus, the approximation ratio of SDP-IE(p, γ) is at least:

$$\rho(p, \gamma) = \frac{2}{\pi} \min_{0 \leq \theta_{ij}, \theta_i, \theta_j \leq \pi} \frac{(1 - \frac{p}{2}) g_\gamma(\theta_{ij}, \theta_i, \theta_j) - (1 - \frac{p}{2}) f_\gamma(\theta_i) + (1 + \frac{p}{2}) f_\gamma(\theta_j)}{1 + \frac{p}{2} + (1 - \frac{p}{2}) \cos \theta_i - (1 + \frac{p}{2}) \cos \theta_j - (1 - \frac{p}{2}) \cos \theta_{ij}},$$

where $\cos \theta_{ij} = v_i \cdot v_j$, $\cos \theta_i = v_0 \cdot v_i$, and $\cos \theta_j = v_0 \cdot v_j$ must satisfy the inequality constraints of (S1). It can be shown numerically that $\rho(2/3, 0.722) \geq 0.9064$. \square

Combining Theorem 4 and Theorem 2, we conclude that:

Theorem 5. *For any directed social network G , the IE strategy of $SDP-IE(2/3, 0.722)$ approximates the maximum revenue of G within a factor of 0.5011.*

Undirected Social Networks. We apply the same approach to an undirected network $G(V, E, w)$. The important difference is that the objective function of the SDP relaxation now is:

$$\max \frac{p(1-p)}{2} \sum_{i \in V} w_{ii} (1 - v_0 \cdot v_i) + \frac{p(1-p)}{4} \sum_{i < j} w_{ij} (2 + p - p v_0 \cdot v_i - p v_0 \cdot v_j - (2 - p) v_i \cdot v_j)$$

Apart from the SDP relaxation, the algorithm is the same as that for directed networks. Working as in the proof of Theorem 4, we can prove that:

Theorem 6. *For any undirected social network G , $SDP-IE(0.586, 0.209)$ approximates the maximum revenue extracted from G by the best IE strategy with pricing probability 0.586 within a factor of 0.9032.*

Combining Theorem 6 and Theorem 1, we conclude that:

Theorem 7. *For any undirected network G , the IE strategy of $SDP-IE(0.586, 0.209)$ approximates the maximum revenue of G within a factor of 0.8229.*

Remark. By the same approach, we compute the approximation ratio of $SDP-IE(p, \gamma)$ against the best IE strategy, for any pricing probability $p \in [1/2, 1)$. Viewed as a function of p , both the best value of γ and the approximation ratio of $SDP-IE(p, \gamma)$ against the best IE strategy increase slowly with p . For example, for directed networks, the approximation ratio of $SDP-IE(0.5, 0.653)$ (resp. $SDP-IE(0.52, 0.685)$ and $SDP-IE(0.52, 0.704)$) is 0.8942 (resp. 0.8955 and 0.9005). For undirected social networks, the approximation ratio of $SDP-IE(0.5, 0.176)$ (resp. $SDP-IE(0.52, 0.183)$ and $SDP-IE(2/3, 0.425)$) is 0.899 (resp. 0.9005 and 0.907). Then, for any $p \in [1/2, 1)$, we can multiply the approximation ratio of $SDP-IE(p, \gamma)$ and the bound obtained by the approach of Theorems 1 and 2 on the fraction of the maximum revenue extracted by the best IE strategy with pricing probability p , and obtain the approximation ratio of $SDP-IE(p, \gamma)$ against the (unrestricted) optimal marketing strategy.

References

1. Akhlaghpour, H., Ghodsi, M., Haghpanah, N., Mirrokni, V.S., Mahini, H., Nikzad, A.: Optimal Iterative Pricing over Social Networks (Extended Abstract). In: Saberi, A. (ed.) WINE 2010. LNCS, vol. 6484, pp. 415–423. Springer, Heidelberg (2010)
2. Alizadeh, F.: Interior point methods in Semidefinite Programming with applications to combinatorial optimization. *SIAM J. on Optimization* 5, 13–51 (1995)
3. Anari, N., Ehsani, S., Ghodsi, M., Haghpanah, N., Immorlica, N., Mahini, H., Mirrokni, V.S.: Equilibrium Pricing with Positive Externalities (Extended Abstract). In: Saberi, A. (ed.) WINE 2010. LNCS, vol. 6484, pp. 424–431. Springer, Heidelberg (2010)
4. Arthur, D., Motwani, R., Sharma, A., Xu, Y.: Pricing Strategies for Viral Marketing on Social Networks. In: Leonardi, S. (ed.) WINE 2009. LNCS, vol. 5929, pp. 101–112. Springer, Heidelberg (2009)
5. Candogan, O., Bimpikis, K., Ozdaglar, A.: Optimal Pricing in the Presence of Local Network Effects. In: Saberi, A. (ed.) WINE 2010. LNCS, vol. 6484, pp. 118–132. Springer, Heidelberg (2010)
6. Chen, W., Lu, P., Sun, X., Tang, B., Wang, Y., Zhu, Z.A.: Optimal Pricing in Social Networks with Incomplete Information. In: Chen, N., Elkind, E., Koutsoupias, E. (eds.) WINE 2011. LNCS, vol. 7090, pp. 49–60. Springer, Heidelberg (2011)
7. Domingos, P., Richardson, M.: Mining the network value of customers. In: Proc. of the 7th ACM-SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2001), pp. 57–66. ACM (2001)
8. Feige, U., Goemans, M.X.: Approximating the value of two prover proof systems, with applications to MAX 2SAT and MAX DICUT. In: Proc. of the 3rd Israel Symposium on Theory of Computing and Systems, pp. 182–189 (1995)
9. Fotakis, D., Krysta, P., Telelis, O.: Externalities among Advertisers in Sponsored Search. In: Persiano, G. (ed.) SAGT 2011. LNCS, vol. 6982, pp. 105–116. Springer, Heidelberg (2011)
10. Garey, M.R., Johnson, D.S.: *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W.H. Freeman (1979)
11. Goemans, M.X., Williamson, D.P.: Improved approximation algorithms for Maximum Cut and Satisfiability problems using Semidefinite Programming. *J. Assoc. Comput. Mach.* 42, 1115–1145 (1995)
12. Guruswami, V., Manokaran, R., Raghavendra, P.: Beating the random ordering is hard: Inapproximability of Maximum Acyclic Subgraph. In: Proc. of the 49th IEEE Symposium on Foundations of Computer Science (FOCS 2008), pp. 573–582 (2008)
13. Haghpanah, N., Immorlica, N., Mirrokni, V.S., Munagala, K.: Optimal auctions with positive network externalities. In: Proc. of the 12th ACM Conference on Electronic Commerce (EC 2011), pp. 11–20 (2011)
14. Hartline, J., Mirrokni, V.S., Sundararajan, M.: Optimal marketing strategies over social networks. In: Proc. of the 17th International Conference on World Wide Web (WWW 2008), pp. 189–198 (2008)
15. Kempe, D., Kleinberg, J., Tardos, É.: Maximizing the spread of influence through a social network. In: Proc. of the 9th ACM-SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2003), pp. 137–146 (2003)
16. Naor, J., Feldman, M., Schwartz, R.: A tight linear time 0.5-approximation for unconstrained submodular maximization. In: Proc. of the 53rd IEEE Symposium on Foundations of Computer Science (FOCS 2012) (to appear, 2012)