Differentially Private and Incentive Compatible Recommendation System for the Adoption of Network Goods

KEVIN HE, Harvard University
XIAOSHENG MU, Harvard University

We study the problem of designing a recommendation system for network goods under the constraint of differential privacy. Agents living on a graph face the introduction of a new good and undergo two stages of adoption. The first stage consists of private, random adoptions. In the second stage, remaining non-adopters decide whether to adopt with the help of a recommendation system $\mathcal{A}$. The good has network complementarity, making it socially desirable for $\mathcal{A}$ to reveal the adoption status of neighboring agents. The designer's problem, however, is to find the socially optimal $\mathcal{A}$ that preserves privacy. We derive feasibility conditions for this problem and characterize the optimal solution.

Categories and Subject Descriptors: K.4.4 [**Computers and Society**]: Electronic Commerce

Additional Key Words and Phrases: differential privacy; network game; recommender system

1. INTRODUCTION

Recommendation systems, in their content-based and collaborative forms, have been well-studied in the past two decades (see [Adomavicius and Tuzhilin 2005] for an overview). More recently, a new kind of social recommendation has become popular on several websites. Platforms that include a social networking component have begun to employ user friendship on the site as a tool towards better recommendations. The following examples illustrate the variety of products that this form of recommendation has been applied to:

(1) Netflix launched Netflix Social in 2013. This feature looks at a user's Netflix friends and pushes films and TV shows that her friends have marked as favorites [Johnson 2013].
(2) Last.fm's music recommendation takes into account the musical taste of a user's Last.fm friends [Konstas et al. 2009]. The site will recommend songs that the user's friends have listened to.
(3) Facebook's "people you may know" feature suggests new "products" (new Facebook friends) to a user based on data from her existing Facebook friends (their friends lists) [Facebook 2008].

This novel social approach is distinct from collaborative filtering, where the website compares a user against its whole database of users, producing a recommendation based on information from those users who are most similar to her. The new approach uses voluntary friendship in an online social network as the main basis for recommendation.

Author's addresses: K. He and X. Mu, Economics Department, Harvard University; email: {he02,xiaoshengmu}@fas.harvard.edu.

EC'14, June 8–12, 2014, Stanford, CA, USA. ACM 978-1-4503-2565-3/14/06 ...$15.00.

http://dx.doi.org/10.1145/2600057.2602841

The present work aims to examine the interplay between privacy concerns and incentive compatibility in such social recommendation systems. On one hand, privacy is an especially pressing issue here due to the sparse nature of these social recommendation systems. The website generates its product suggestions for a user not from its entire database of users, but primarily from the small subset of users who are her neighbors in a social graph. In doing so, the website may reveal some sensitive information. For instance, some users may regard their Facebook friends lists as private information. However, as the security company Quotium points out, the "people you may know" suggestions represent outputs produced from from precisely this private information [Abezgauz 2013]. On the other hand, users are not obliged to follow the recommendations and one expects that they will have some prior belief about the quality of various products. If the recommendations are too noisy, users will ignore them. Thus we would want our system to be "incentive compatible", in the sense that users who receive a recommendation will voluntarily act upon them.

In markets where the recommender can only produce non-compulsory recommendations under the constraint of privacy, we are interested in question such as: Is it possible to design a social recommendation system that preserves a prescribed level of privacy, while still producing useful outputs? If so, what does the socially optimal, incentive compatible, privacy preserving system look like?

1.1. Our Contributions

We view our contribution to the literature as threefold: (1) we propose a model that captures the interplay between privacy and incentive compatibility, in the setting of network good adoption; (2) for each desired level of privacy, we derive necessary and sufficient conditions for the existence of a recommendation system that satisfies both incentive compatibility constraint and privacy; (3) when at least one such system exists, we characterize the "optimal" system in terms of social welfare.

Since the relationship between privacy and incentive compatibility in the setting of friendship-based social recommendation systems is not a well-studied problem, the first task is to come up with the right framework. We take inspiration from the economics literature and study a model where agents living on a network face the introduction of a new network good. Each agent contemplates whether to pay a cost $c > 0$ to adopt the good, whose value to the agent depends on how many of her social network neighbors are also adopters. This model is appropriate for two reasons. First, network goods best illustrate the conflict between social gains of information disclosure and the constraints imposed by privacy concerns. Disregarding privacy, the first-best strategy of the website would be to recommend adoption of the network good to precisely those agents whose friends have adopted the good. However, this strategy reveals to every agent the adoption status of her friends, which clearly violates privacy. Second, adoption of a new good is usually a voluntary matter, so that the incentive compatibility of the recommendation system becomes important.

Next, we find that for each desired level ϵ of privacy, there exists a social recommendation system that satisfies both incentive compatibility and privacy if and only if c is sufficiently low (call such systems *feasible*). Specifically, since incentive compatibility and privacy are both concepts defined in terms of individual agents, we may first ask whether there exists a feasible system at each agent. The result is that there is a bound $\bar{c}(d, \epsilon, p)$ so that there is a feasible system for an agent with d friends in the network each with p probability of being adopters if and only if $c < \bar{c}(d, \epsilon, p)$. Taking the minimum of this bound across all agents in the network gives the analogous network-based result.

Finally, when the set of feasible systems is non-empty, the socially optimal feasible system is the one that uses a "cutoff" $\bar{k}(d, \epsilon, p, c)$. For an agent with exactly $\bar{k}$ adopter

friends, the system recommends with probability $\frac{\exp(\epsilon)}{\exp(\epsilon)+1}$. For an agent with fewer than $\bar{k}$ adopter friends, the system recommends adoption with a probability that decreases exponentially at rate ϵ as the number of adopter friends decreases. For an agent with more than $\bar{k}$ adopter friends, the probability that the system does not recommend adoption decreases at rate ϵ as the number of adopter friends increases.

2. BACKGROUND AND RELATED WORK

2.1. Differential Privacy

In the past decade, two incidents illustrated the possibility of privacy breach for even anonymized data. The first incident involved Netflix. As part of a contest, the company released anonymized viewing and rating history of some of its customers. Narayanan and Shmatikov managed to match the Netflix data to user data on IMDB, uncovering the identity of many user in the anonymized dataset [Narayanan and Shmatikov 2008]. The second incident involved Massachusetts Group Insurance Commission, who released anonymized patient information. Sweeney succeeded in discovering the identity of many anonymized patients and thus gaining access to their medical history [Dwork 2008]. One de-anonymized patient was in fact the governor of Massachusetts.

These incidences prompted the security community to seek a higher standard of data privacy than anonymity. The community settled on the notion of differential privacy, originally proposed by Dwork [Dwork 2006]. Differential privacy roughly says a single-entry change in the underlying database should not produce a clearly distinguishable effect on the output of the algorithm. Formally, write $\mathbb{D}$ for the set of possible states of the database. Then the $\mathbb{S}-$valued randomized algorithm $\mathcal{A} : \mathbb{D} \to \mathbb{S}$ is said to be ϵ-differentially private if for every $S \subseteq \mathbb{S}$ and $D, D' \in \mathbb{D}$ where D' differs from D in only a single row,

$$\mathbf{Pr}[\mathcal{A}(D) \in S] \leq \exp(\epsilon) \cdot \mathbf{Pr}[\mathcal{A}(D') \in S].$$

In some sense one may view the condition of ϵ-differential privacy as a "continuity" condition for $\mathcal{A}$. A small change in the input should produce only a correspondingly small change in the output.

2.2. Differential Privacy and Mechanism Design

Recently many papers have explored the intersection between differential privacy and mechanism design. These works fall roughly into two strands: those that treat differential privacy as a tool for mechanism design, and those that treat it as a desideratum of mechanism design.

The first strand began with McSherry and Talwar who pointed out that a differentially private mechanism has truth-telling as an approximately dominant strategy for both individuals and coalitions [McSherry and Talwar 2007]. They also proposed the "exponential mechanism", a differentially private mechanism for generic mechanism design problems. Several later papers have built on this result. Nissim et al. showed a suitable lottery between the exponential mechanism and a "commitment mechanism" implements an ex-post Nash equilibrium that is approximately optimal with respect to the objective function [Nissim et al. 2012]. Huang et al. proved that if the objective is to maximize social welfare, then a suitably instantiated exponential algorithm satisfies differential privacy, truth-telling, and approximate optimality [Huang and Kannan 2012].

More recently, Kearns et al. considered the problem of using recommendation as an equilibrium coordination device for games of incomplete information [Kearns et al. 2014]. By contrast, in the present paper agents will have myopic payoffs and thus

not suffer from coordination problems. The role of the mechanism is not so much to coordinate as to inform agents of the state of the world subject to privacy constraint.

An example of work in the second strand of literature is Hsu et al. on private allocations [Hsu et al. 2013]. They propose an efficient algorithm for solving the allocation problem of assigning n objects to m individuals, with a relaxed version of differential privacy as an additional constraint. Our work also uses differential privacy as a constraint rather than a tool. However, our interest lies in explicitly characterizing the solution to our problem, not in finding an algorithm for computing it. We also use a relaxed version of differential privacy. Namely, we ask that the recommendation given to each player i is differentially private from the types of all $j \neq i$. This is the "mirror image" of the "joint differential privacy" that Hsu et al. use, which requires the distribution of recommendations given to all $j \neq i$ be differentially private from i's type.

2.3. Social Recommendation

The literature on the use of social network data to enhance recommendations probably started with Konstas et al.'s work in 2009 [Konstas et al. 2009]. There, the authors collected data from Last.fm and showed that incorporating friendship information improves recommendations. Following their work, other groups have applied social networking data to recommendations of music [Bu et al. 2010], points of interest [Ye et al. 2010], and enterprise-level social media content [Guy et al. 2010].

Several recent papers have studied the constraints that differential privacy imposes on social recommendation. The two most relevant to this work are Machanavajjhala et al. [Machanavajjhala et al. 2011] and Zhu et al. [Zhu et al. 2013]. Machanavajjhala's group emphasized that the trade-off between accuracy of result and privacy of output is a general phenomenon and derived general lower bounds for achievable privacy parameter ϵ for algorithms that guarantee constant accuracy. Compared to their work, we aim to study a more specific problem of economic interest, where we can explicitly characterize the optimal algorithm. Zhu's group proposed a form of social collaborative filtering that protects privacy. For us, there is only a single item under consideration– the network good to be adopted. Thus the problem is not so much the collaborative filtering one of predicting user rating of a new item based on her ratings of past items, as it is disclosing the adoption status with respect to a single good in a privacy-sensitive way.

2.4. Innovation Diffusion

This work relates to the economic literature on innovation diffusion in a network setting. The typical model opens with the appearance of a new product. Agents located on a network decide whether or not to adopt the product, where their payoff from adoption depends on the adoption status of their neighbors [Ellison 1993; Jackson and Yariv 2007; Young and Kreindler 2012]. These models study the consequences of repeated local coordination games, as well as the limiting distribution of adoption in the network. In contrast to these papers with decentralized agents, our work asks what welfare improvements can be made if a central information authority can intervene in the adoption process.

3. MODEL

Let $G = (V, E)$ denote a graph where $V = \{v_1, ..., v_n\}$ represents the agents. The graph models a social network that faces the introduction of a new good with network externality. The edges represent a symmetric relationship such as friendship, along which the externality effect of the network good takes form. Agents are imperfectly informed about the value of the good, since the adoption status of other agents in the network

are treated as private information. A designer wishes to make recommendations to agents as to maximize social welfare, while preserving data privacy.

We consider the following two stage model.

STAGE I: INITIAL ADOPTION. Each node independently adopts the good with probability p. The adoption status of node v_i is $\theta_i \in \{0,1\}$, which is private to v_i. This means v_i knows the realization of θ_i, but only the distribution of θ_j for $j \neq i$. Summarize STAGE I adoption state of the network as $\theta = (\theta_1, ..., \theta_n) \in \Theta$, with $\Theta := \{0,1\}^n$.

STAGE I models an adoption process before the intervention of the recommendation system. When a new product appears on the market, agents may be heterogeneous with respect to their willingness to try out this product. In our model, this decentralized process of adoption results in p fraction of the agents being early adopters, who are perhaps fans of related products. Moreover, STAGE I gives rise to a common prior under which agents can calculate the expected payoff from adopting the network good after receiving a recommendation.

STAGE II: RECOMMENDATION. The designer sends a recommendation to a subset of STAGE I non-adopters. Then, each non-adopter decides whether to adopt. If non-adopter v_i decides to adopt in STAGE II, he receives payoff $\phi\left(\frac{\sum_{w \in N(v_i)} \theta_w}{d(v_i)}\right) - c$, where $c > 0$ represents cost of adoption and ϕ is a continuous, increasing function normalized to have $\phi(0) = 0$ and $\phi(1) = 1$. If he remains a non-adopter, his payoff is 0. Since v_i does not know the value of θ_j for $j \neq i$, v_i only observes his payoff ex-post.

We note that non-adopter agents in our model are myopic: they consider only the short run payoff of their actions and do not take into account future changes in adoption status of the network. In the alternative, game-theoretic model, the decision of each STAGE I non-adopter affects not only her payoff, but also the payoffs of her neighbors. That is, a STAGE II adopter v_i benefits not only from those of her neighbors who adopted as of the end of STAGE I, but also from new neighboring adopters in STAGE II. The downside to this alternative model is that it has too many equilibria. In particular, there is an equilibrium where all STAGE I non-adopters adopt in Stage II, in anticipation of maximal network externality at the end of STAGE II. Assuming $p < c$, there is another equilibrium where no one adopts in STAGE II, correctly expecting minimal network externality. The multitude of equilibria makes such a model unattractive.

We interpret $\phi(k/d)$ as the (gross) benefit of adoption when k out of d of one's friends have adopted the good. The assumption that ϕ is increasing captures the network complementarity of the good. Examples of ϕ include:

— $\phi(x) = x$. This is perhaps the simplest model where the gross benefit of adoption for v_i is exactly the fraction of v_i's neighbors who are STAGE I adopters.
— $\phi(x) = x^\alpha$, for parameter $\alpha > 0$. For $0 < \alpha < 1$, ϕ is strictly concave and models a network good with decreasing marginal benefit of having one more friend adopt the good. For $\alpha > 1$, ϕ is strictly convex and there is increasing marginal benefit.
— $\phi(x) = \frac{\exp(\beta(1+\alpha)x)}{\exp(\beta(1+\alpha)x)+\exp(\beta(1-x))}$, for parameters $\alpha, \beta > 0$. This is the logit response function, a commonly used model in the discrete choice literature [McFadden 1976] and stochastic learning literature [Blume 1993, 1995].

Knowing ϕ, the designer announces a stochastic function $\mathcal{A} : \Theta \to \Theta$ before STAGE I, mapping each STAGE I adoption status $\theta \in \Theta$ to a profile of recommendations. Specifically, for all i with $\theta_i = 0$, $\mathcal{A}^i(\theta) = 1$ means $\mathcal{A}$ sends a recommendation to i, while $\mathcal{A}^i(\theta) = 0$ means $\mathcal{A}$ does not.

Upon receiving or not receiving a recommendation, each node v_i with $\theta_i = 0$ computes the conditional expected payoff from adopting the good:

$$\pi_i(\mathcal{A}^i) := \mathbb{E}_\theta \left[\phi \left(\frac{\sum_{w \in N(v_i)} \theta_w}{d(v_i)} \right) - c \;\middle|\; \mathcal{A}^i \right].$$

and node v_i decides to adopt if and only if $\pi_i \geq 0$.

The problem of the designer is a constrained maximization problem.

(OBJECTIVE) Social welfare (**SW**) computes the expected social gain from $\mathcal{A}$.

$$\mathbb{E}_\theta \left[\mathbb{E}_{\mathcal{A}} \left[\sum_{v_i \in V} 1_{\{\theta_i = 0\}} \cdot 1_{\{\pi_i \geq 0\}} \cdot \left(\phi \left(\frac{\sum_{w \in N(v_i)} \theta_w}{d(v_i)} \right) - c \right) \right] \right]. \tag{SW}$$

To interpret, for every realization of θ, we restrict attention to those nodes v_i who have not adopted the network good in STAGE I. For every realization of $\mathcal{A}$, each node with $\theta_i = 0$ either receives a recommendation from the designer or receives no recommendation. These nodes then compute the conditional expected payoff from adopting the network good based on this signal. Nodes with $\pi_i \geq 0$ decide to adopt and we compute the sum of their payoffs. Since both θ and $\mathcal{A}$ are stochastic, we take expectations over these two random variables to arrive at a deterministic objective function.

Note, however, $\mathcal{A}$ is a function of θ. Choosing $\mathcal{A}$ to maximize **SW** is equivalent to choosing $\mathcal{A}(\theta)$ to maximize the inner expectation of **SW** at every θ.

(CONSTRAINT 1) Incentive compatibility (**IC**) says each agent must find it beneficial to do as recommended. For every i with $\theta_i = 0$,

$$\begin{cases} \mathbb{E}_{\theta,\mathcal{A}} \left(\phi \left(\frac{\sum_{w \in N(v_i)} \theta_w}{d(v_i)} \right) - c \;\middle|\; \mathcal{A}^i(\theta) = 1 \right) & \geq 0 \\ \mathbb{E}_{\theta,\mathcal{A}} \left(\phi \left(\frac{\sum_{w \in N(v_i)} \theta_w}{d(v_i)} \right) - c \;\middle|\; \mathcal{A}^i(\theta) = 0 \right) & \leq 0 \end{cases} \tag{IC}$$

IC says a node expects to receive at least a payoff of 0 from adoption conditional on not being a STAGE I adopter and receiving a recommendation at the start of STAGE II. Also, a node expects to receive no more than 0 payoff from adoption conditional on not being a STAGE I adopter and *not* receiving a recommendation at the start of STAGE II. (Some authors, such as Myerson, also call this the obedience constraint [Myerson 1988].)

There is a Bayesian interpretation to **IC**. Each non-adopter v_i has a prior belief about the distribution of θ, which are the payoff relevant states. The realization of $\mathcal{A}^i(\theta)$ is the data that v_i receives. Since v_i knows $\mathcal{A}$, she knows the likelihood of receiving this data given different θ. So, she forms a posterior belief about θ based on the data and computes her expected payoff from adoption according to this posterior. **IC** says following the recommendation is the right thing to do based on this posterior.

(CONSTRAINT 2) ϵ-differential privacy (**DP**) requires that $\mathcal{A}$ is not too sensitive to small changes in θ. For any θ, θ' differing in only one component and for $\chi \in \{0, 1\}$,

$$\Pr\left[\mathcal{A}^i(\theta) = \chi\right] \leq \exp(\epsilon) \cdot \Pr\left[\mathcal{A}^i(\theta') = \chi\right]. \tag{DP}$$

This is a relaxation of the usual ϵ-differential privacy condition in the context of network, where Θ plays the role of the "database" in standard settings with ϵ-differential privacy. We only ask that the recommendation given to individual i is not too sensitive to the types of all $j \neq i$. This relaxation is motivated by a no-collusion assumption. If every player only observes their own recommendation and there is no exchange of in-

formation, then there will be no breach in privacy as long as our version of differential privacy is satisfied.

A stochastic function $\mathcal{A}$ is called **feasible** if it satisfies both **IC** and **DP**.

The **constrained optimization problem** of the designer is to pick a feasible $\mathcal{A}$ to maximize **SW.**

The key feature of the model is the tension between the **IC** and **DP** constraints. **DP** requires the recommendations to not be too informative of the adoption status of the neighbors. However, **IC** requires that the recommendation be sufficiently informative. In the two extreme cases, making uniformly random recommendations to every node certainly obeys **DP**, but violates **IC** if c is sufficiently high. Making socially optimal recommendations (i.e. recommend a node to adopt if and only if its gain from adoption is positive) satisfies **IC**, but may violate **DP** for sufficiently small ϵ. These two competing constraints reflect the classic tension between preserving data privacy and producing useful outputs in differential privacy problems.

4. A TOY EXAMPLE

Consider a very simple network consisting of only two vertices connected with an edge. Set **DP** parameter $\epsilon = \ln(2)$, probability of STAGE I adoption $p = 0.2$, and cost of Stage II adoption $c = 0.25$. Suppose $\phi(x) = x$. We use this "toy example" to illustrate the geometry of the problem.

It is easy to see that a recommendation system in this case is completely characterized by two probabilities $l_0, l_1 \in [0, 1]$, where l_0 is the probability that the designer recommends adoption to an agent whose neighbor did not adopt in STAGE I, l_1 is the probability of recommendation when her neighbor did adopt. We can thus think of the set of all possible recommendation systems as $[0, 1] \times [0, 1]$.

In Figure 1 we plot the subsets of recommendation systems that satisfy **DP** and **IC**. It is convenient to consider the "two halves" of the **DP** constraint separately, namely

$$\Pr\left[\mathcal{A}^i(\theta) = 1\right] \leq \exp(\epsilon) \cdot \Pr\left[\mathcal{A}^i(\theta') = 1\right] \qquad \textbf{(DP1)}$$

and

$$\Pr\left[\mathcal{A}^i(\theta) = 0\right] \leq \exp(\epsilon) \cdot \Pr\left[\mathcal{A}^i(\theta') = 0\right] \qquad \textbf{(DP2)}$$

As one may expect, the subset of recommendation systems that satisfies each half of DP lies along a "fat diagonal", shown as **DP1** an **DP2** in the figure. In **DP1**, we need l_0 to be not too different from l_1 as to avoid disclosing the adoption status of an agent's neighbor. In **DP2**, a similar constraint is imposed on the probabilities of non-recommendation, $(1 - l_0)$ and $(1 - l_1)$. The set of recommendation systems that satisfy **DP** is the diamond-shaped intersection between **DP1** and **DP2**. On the other hand, **IC** captures the upper left quadrant of the graph. In order for the agent to adopt the good when recommended to do so, the recommendation must be sufficiently informative about the neighbor's adoption status. This is exactly the upper left corner of the plot, which contains points with low l_0 and high l_1.

One can visualize the tension between **DP** and **IC**. **DP** pushes the recommendation system to lie along the diagonal. **IC** wants the system to lie in the upper left corner. In particular, if ϵ decreases, then the width of the "**DP** diamond" shrinks. Meanwhile, if c increases or p decreases, then the "**IC** triangle" shifts further to the upper left. Thus with sufficiently small ϵ, sufficiently large c, and sufficiently small p, there will be no intersection between the "**DP** diamond" and the "**IC** triangle" except the trivial solution of $(0, 0)$, making the problem essentially infeasible.

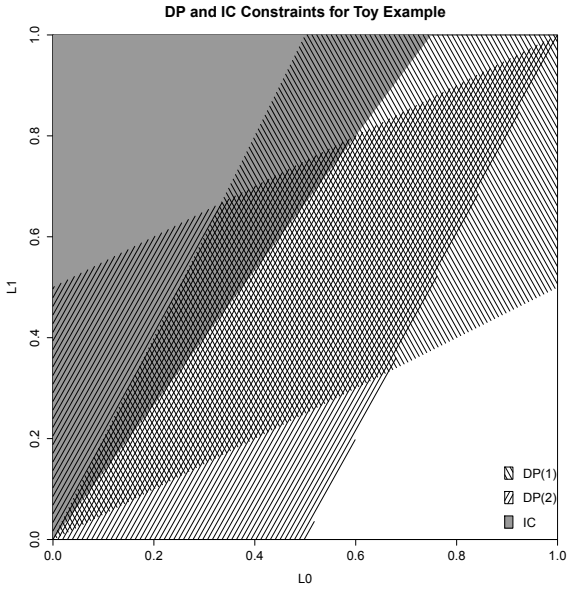

Fig. 1. Recommendation systems that satisfy **DP** and **IC** in a toy example with two agents, $\epsilon = \ln(2)$, $p = 0.2$, $c = 0.25$. Here l_0 refers to the probability of recommendation when an agent's neighbor is a STAGE I non-adopter, whereas l_1 refers to this probability when the neighbor has adopted in STAGE I. The gray region refers to (l_0, l_1) pairs that satisfy **IC**. The shaded black region refers to (l_0, l_1) that satisfy (the two halves of) **DP**. Their intersection contains all the feasible recommendation systems.

In the case where the two regions have some intersection other than $(0, 0)$, the problem is to maximize **SW** over this intersection. In Figure 2, we show a heat map of **SW** as a function of l_0 and l_1, with **DP** and **IC** drawn in as reference lines.

Note that the global maxima of **SW** occur at two corners: $l_0 = 0, l_1 = 1$ and $l_0 = 1, l_1 = 0$. In both cases, the designer's recommendation acts as a signal that perfectly reveals the adoption status of the neighbor. In the first system, an agent adopts the good if and only if recommended to do so. In the second system, an agent adopts if and only if *not* recommended to do so. Note that the second situation seems pathological and is eliminated by the **IC** constraint. In fact, none of the global maximum can be achieved due to the **DP** constraint binding the feasible region along the central diamond. In this example, the point that maximizes **SW** in the feasible region is $(\frac{1}{3}, \frac{2}{3})$.

5. THEORETICAL RESULTS

We will assume throughout this section that the cost of adoption is not too small relative to its benefits. Specifically, we make the following assumption:

$$c \geq \sum_{k=0}^{d} \phi\left(d/k\right) \cdot \binom{d}{k} \cdot p^k (1-p)^{d-k} \qquad \text{(Non-Trivial Cost)}$$

Note that this condition simplifies to $c \geq p$ when $\phi(x) = x$. Without this assumption, there is no role for recommendation – agents will find it beneficial to adopt even in its absence. In that case, the system that always sends a recommendation to every agent is feasible.

In the presence of Non-Trivial Cost, the mechanism that never recommends adoption to any agent is evidently feasible. Does there always exist another feasible $\mathcal{A}$? The

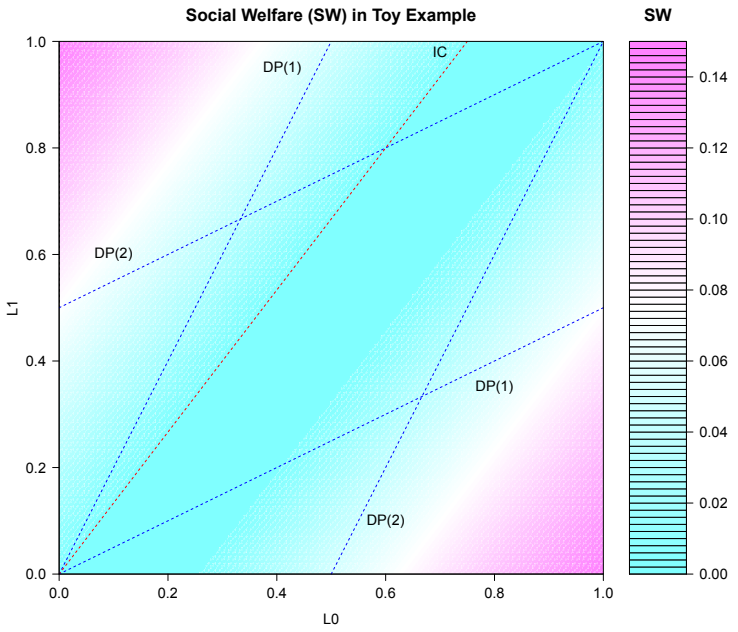

Fig. 2. Heat map of **SW** as a function of (l_0, l_1) in the toy example. The boundaries of **DP** and **IC** are shown for reference.

answer is no. The following proposition provides a complete characterization of the set of costs c that admit at least one feasible, non-trivial $\mathcal{A}$, holding privacy requirement ϵ and initial adoption parameter p fixed.

PROPOSITION 5.1. *Fix a node v that did not adopt the good in the first stage, and let $d = d(v)$. There exists a stochastic recommendation function A satisfying **IC** and **DP** for v if and only if the cost of adoption $c \leq \bar{c}(d, \epsilon, p)$, where we define*

$$\bar{c}(d, \epsilon, p) = \frac{1}{(1 - p + pe^{\epsilon})^d} \sum_{k=0}^{d} \phi(\frac{k}{d}) \binom{d}{k} (pe^{\epsilon})^k (1-p)^{d-k}. \tag{1}$$

PROOF. We first lay out some notations that will be useful later. Define $l^v(\Theta) = Pr[A^v(\Theta) = 1]$ be the likelihood of recommendation given profile Θ, in which $\theta_v = 0$. Denote by $\Theta_{N(v)}$ the projection of Θ onto the subspace spanned by $N(v) := \{v_1, v_2 \ldots, v_d\}$, the set of neighbors of v. It is often times easier to work with $l^v(\Theta_{N(v)})$, which is the conditional probability of recommendation to v, given $\theta_v = 0$ and the partial profile $\Theta_{N(v)}$. Formally,

$$A^v(\Theta_{N(v)}) = \sum_{\theta_u : u \in V - \{v\} - N(v)} \left(A^v(\theta_v, \Theta_{N(v)}, \{\theta_u\}) \prod_{u \in V - \{v\} - N(v)} Pr[\theta_u] \right). \tag{2}$$

DP1 (See Section 4) imposes the following ratio inequality

$$l^v(\Theta) \leq \exp(\epsilon) \cdot l^v(\hat{\Theta}), \text{ whenever } \Theta, \hat{\Theta} \text{ differ in at most one coordinate.} \tag{3}$$

From (2), the same must also be true for the conditional probabilities $l^v(\Theta_{N(v)})$.

Next, for each k ($0 \leq k \leq d$), let l_k be the conditional probability that the social planner makes a recommendation to v, given $\theta_v = 0$ and that the profile $\Theta_{N(v)}$ sums to k. Because each coordinate of $\Theta_{N(v)}$ is independent and identically distributed, we have

$$l_k = \frac{1}{\binom{d}{k}} \sum_{\theta_{v_1}+\cdots+\theta_{v_d}=k} l^v(\Theta_{N(v)}). \tag{4}$$

We claim that $l_{k+1} \leq \exp(\epsilon) \cdot l_k$. To see this, fix k. Create two sets S^k, S^{k+1} consisting of profiles $\Theta_{N(v)}$ that have sum k and $k+1$ respectively. For every $\Theta_{N(v)} \in S^k$ and $\hat{\Theta}_{N(v)} \in S^{k+1}$ that differ in at most one coordinate, we can apply (3). Adding up all such inequalities, we obtain

$$(k+1) \sum_{\theta_{v_1}+\cdots+\theta_{v_d}=k+1} l^v(\Theta_{N(v)}) \leq \exp(\epsilon) \cdot (d-k) \sum_{\hat{\theta}_{v_1}+\cdots+\hat{\theta}_{v_d}=k+1} l^v(\hat{\Theta}_{N(v)}),$$

or by (4)

$$l_{k+1} \leq \exp(\epsilon) \cdot l_k. \tag{5}$$

Similar analysis via **DP2** gives:

$$(1 - l_{k+1}) \geq \exp(-\epsilon) \cdot (1 - l_k). \tag{6}$$

Finally define

$$p_k^d = \binom{d}{k} p^k (1-p)^{d-k}. \tag{7}$$

to be the prior probability (before stage II) that the profile $\Theta_{N(v)}$ has sum k.

Using p_k^d, l_k and applying Bayes' rule, we calculate that the agent's expected utility conditional on recommendation is

$$\mathbb{E}[\phi(\frac{k}{d})|A^v = 1] = \frac{\sum_{k=0}^d \phi(\frac{k}{d}) \cdot p_k^d \cdot l_k}{\sum_{k=0}^d p_k^d \cdot l_k}. \tag{8}$$

The **IC** constraint can be rewritten compactly as (8) $\geq c$.

With these preliminaries we can now prove the proposition:
(Sufficiency): Consider any recommendation function given by (assuming $\theta_v = 0$):

$$l^v(\Theta) = l^v(\Theta_{N(v)}) = \lambda \cdot \exp(\epsilon(\theta_{v_1} + \cdots + \theta_{v_d})); \tag{9}$$

where $\lambda > 0$ is a constant independent of the realization of Θ. For λ sufficiently small, (9) is a well defined probability, so A is well-defined.

Under this construction **DP1** is trivially satisfied and **DP2** is satisfied for all sufficiently small λ. Moreover, (5) holds with equality, which implies that l_k is proportional to $\exp(\epsilon k)$. (7) thus evaluates to $\bar{c}(d, \epsilon, p)$, which means agents the incentive to adopt the good when recommended to do so, i.e.

$$\sum_{k=0}^d (\phi(\frac{k}{d}) - c) \cdot p_k^d \cdot l_k \geq 0.$$

Note that under "Non-Trivial Cost",

$$\sum_{k=0}^{d}(\phi(\frac{k}{d})-c)\cdot p_k^d \leq 0.$$

Therefore,

$$\sum_{k=0}^{d}(\phi(\frac{k}{d})-c)\cdot p_k^d \cdot (1-l_k) \leq 0.$$

So agents have no incentive to adopt when not recommended to do so.

(Necessity): We show that under (5), (6) and (7), the maximum of (8) is $\bar{c}(d,\epsilon,p)$. Equivalently,

$$\sum_{k=0}^{d}\left(\phi(\frac{k}{d})-\bar{c}\right)p_k^d \cdot l_k \leq 0.$$

This will follow from Proposition 5.4, which characterizes the recommendation probabilities l_k that maximize v's expected utility for arbitrary costs. □

An especially elegant version of the above bound occurs when $\phi(x)=x$. In that case:

COROLLARY 5.2. $\bar{c}(d,\epsilon,p)=\frac{\exp(\epsilon)\cdot p}{1-p+\exp(\epsilon)\cdot p}$.

PROOF. First observe that $\sum_{k=0}^{d} k\binom{d}{k}a^k b^{d-k} = a\cdot d\cdot(a+b)^{d-1}$. This comes from first expanding

$$(a+b)^d \;=\; \sum_{k=0}^{d}\binom{d}{k}a^k b^{d-k}.$$

Applying $\frac{d}{da}$ to both sides and then multiplying both sides by a to obtain:

$$a\cdot d\cdot(a+b)^{d-1} = \sum_{k=1}^{d} k\binom{d}{k}a^k b^{d-k}.$$

Hence the observation.

From equation (1), $\bar{c}(d,\epsilon,p)=\frac{1}{(1-p+p\cdot\exp(\epsilon))^d}\sum_{k=0}^{d}\phi(\frac{k}{d})\binom{d}{k}(p\cdot\exp(\epsilon))^k(1-p)^{d-k}$. In the case where $\phi(x)=x$, this simplifies to:

$$\frac{1}{(1-p+p\cdot\exp(\epsilon))^d}\cdot\frac{1}{d}\sum_{k=0}^{d} k\binom{d}{k}(p\cdot\exp(\epsilon))^k(1-p)^{d-k}.$$

Applying the observation above yields:

$$\frac{1}{(1-p+p\cdot\exp(\epsilon))^d}\cdot\frac{1}{d}\cdot(1-p+p\cdot\exp(\epsilon))^{d-1}\cdot(p\cdot\exp(\epsilon))\cdot d.$$

which simplifies readily into the desired expression. □

This corollary says when $\phi(x) = x$, $\bar{c}$ is independent of d, making the bound "topology-free". In other words, our necessary and sufficient conditions for the existence of a feasible algorithm holds without making any assumptions about the structure of the network G.

The simplicity of this result depends crucially on the linear functional form of ϕ. However, it turns out that we also know a lot about $\bar{c}$ when we only assume decreasing marginal return–ϕ concave.

PROPOSITION 5.3. *$\bar{c}(d, \epsilon, p)$ as defined in (1) is increasing in ϵ and p, with $\lim_{d\to\infty} \bar{c}(d, \epsilon, p) = \phi(\frac{p\cdot\exp(\epsilon)}{1-p+p\cdot\exp(\epsilon)})$. Moreover if ϕ is concave, then $\bar{c}(d, \epsilon, p)$ increases in d.*

PROOF. For convenience, we write $\lambda := \frac{p\cdot\exp(\epsilon)}{1-p+p\cdot\exp(\epsilon)}$. (1) can be rewritten as:

$$\bar{c}(d, \epsilon, p) = \bar{c}(d, \lambda) = \sum_{k=0}^{d} \phi(\frac{k}{d}) \binom{d}{k} \lambda^k (1-\lambda)^{d-k}. \tag{10}$$

Its derivative with respect to λ is $d \cdot \sum_{k=0}^{d-1} \left(\phi(\frac{k+1}{d}) - \phi(\frac{k}{d})\right) \binom{d-1}{k} \lambda^k (1-\lambda)^{d-1-k} \geq 0$. Because λ increases in ϵ and p, so does $\bar{c}$. Alternatively, a higher ϵ means a more relaxed **DP** constraint, while a higher p makes **IC** easier to satisfy. Either effect leads to a higher $\bar{c}$.

To prove that $\bar{c}(d, \lambda)$ converges to $\phi(\lambda)$, we first show

$$\sum_{k=0}^{d(\lambda-\delta)} \binom{d}{k} \lambda^k (1-\lambda)^{d-k} \leq \delta \text{ for all large } d, \forall \delta > 0. \tag{11}$$

Note that the ratio between consecutive summand between cons in (11) is $\frac{k+1}{d-k} \cdot \frac{1-\lambda}{\lambda}$, which is less than $1-\delta$ for $k \leq d(\lambda-\delta) + O(1)$. The inequality thus follows from a simple calculation of geometric sums.

Now using (10), (11), we obtain that

$$\begin{aligned}
\bar{c}(d, \lambda) &\geq \sum_{k > d(\lambda-\delta)} \phi(\frac{k}{d}) \binom{d}{k} \lambda^k (1-\lambda)^{d-k} \\
&\geq \phi(\lambda - \delta) \sum_{k > d(\lambda-\delta)} \binom{d}{k} \lambda^k (1-\lambda)^{d-k} \\
&\geq (1-\delta) \cdot \phi(\lambda - \delta)
\end{aligned}$$

for all large d. Similarly we have

$$\begin{aligned}
\bar{c}(d, \lambda) &= \sum_{k \leq d(\lambda+\delta)} \phi(\frac{k}{d}) \binom{d}{k} \lambda^k (1-\lambda)^{d-k} + \sum_{k > d(\lambda+\delta)} \phi(\frac{k}{d}) \binom{d}{k} \lambda^k (1-\lambda)^{d-k} \\
&\leq \phi(\lambda + \delta) + \delta.
\end{aligned}$$

Letting δ go to zero and using the continuity of ϕ, we conclude $\lim_{d\to\infty} \bar{c}(d, \lambda) = \phi(\lambda)$. This concludes the proof of the proposition and shows that in fact, the limit does not depend on the shape of ϕ.

Now fixing λ, we show that $\bar{c}$ increases in d for concave ϕ. From (10) we have

$$\begin{aligned}
\bar{c}(d,\lambda) &= \sum_{k=0}^{d} \phi(\frac{k}{d}) \binom{d}{k} \left(\lambda^{k+1}(1-\lambda)^{d-k} + \lambda^k(1-\lambda)^{d-k+1}\right) \\
&= \sum_{k=0}^{d+1} \lambda^k(1-\lambda)^{d-k+1} \left(\phi(\frac{k}{d})\binom{d}{k} + \phi(\frac{k-1}{d})\binom{d}{k-1}\right) \\
&= \sum_{k=0}^{d+1} \lambda^k(1-\lambda)^{d-k+1} \binom{d+1}{k} \left(\frac{d-k+1}{d+1} \cdot \phi(\frac{k}{d}) + \frac{k}{d+1} \cdot \phi(\frac{k-1}{d})\right) \\
&\leq \sum_{k=0}^{d+1} \lambda^k(1-\lambda)^{d-k+1} \binom{d+1}{k} \phi(\frac{k}{d+1}) \\
&= \bar{c}(d+1,\lambda).
\end{aligned}$$

In the second to last step we applied Jensen's inequality, thanks to the concavity of ϕ. □

The previous proposition roughly says that when agents exhibit decreasing marginal return, those who have a large number of friends are more likely to adopt the recommendation. To get the intuition, consider the case where the agent only cares about whether there is a friend who has adopted the product, but does not care at all about how many have adopted. We note that this is an extreme form of decreasing marginal return. To satisfy **IC** in this context, the designer has to minimize l_0, the probability that he recommends when no friend has adopted, under the privacy constraints. It is clear that a larger d makes this task easier, because the privacy constraint is less restrictive on l_0. Also if ϕ is convex, then a similar proof shows $\bar{c}$ decreases in d. These results combined give another proof of Corollary 5.2.

The next proposition confirms that the optimal recommendation system uses a "cutoff strategy" and gives an explicit formula for the cutoff.

PROPOSITION 5.4. *Under the same assumptions as in Proposition 5.1, node v's interim (knowing $\theta_v = 0$ but nothing else) expected utility given that he will adopt the social planner's recommendation is*

$$\pi_v = \sum_{k=0}^{d} \left(\phi(\frac{k}{d}) - c\right) p_k^d \cdot l_k. \tag{12}$$

Consider the optimization problem $\max \pi_v$ subject to the constraints: $0 \leq l_k \leq 1$, $l_{k+1} \leq exp(\epsilon) \cdot l_k$, and $exp(-\epsilon) \cdot (1 - l_k) \leq 1 - l_{k+1}$.

The optimization problem has the following cutoff type solution:
i) For $c \leq \bar{c}(d,\epsilon,p)$, there exists integer $\bar{k}$ between 0 and d such that π_v is maximized when

$$l_k = \begin{cases} \frac{exp(\epsilon)}{exp(\epsilon)+1} \cdot exp(\epsilon(k-\bar{k})), \text{ if } k \leq \bar{k}; \\ 1 - \frac{1}{exp(\epsilon)+1} \cdot exp(\epsilon(\bar{k}-k)), \text{ if } k > \bar{k}. \end{cases}$$

$\bar{k}(d,\epsilon,p,c)$ is such that:

$$\sum_{k=0}^{d} (\phi(\frac{k}{d}) - c) p_k^d \cdot exp(-|k-\bar{k}|\epsilon) \geq 0 > \sum_{k=0}^{d} (\phi(\frac{k}{d}) - c) p_k^d \cdot exp(-|k-\bar{k}+1|\epsilon). \tag{13}$$

ii) For $c > \bar{c}(d,\epsilon,p)$, π_v can never be positive. Thus $l_k = 0, \forall k$ is optimal.

As in the proof of Proposition 5.1, thanks to Non-Trivial Cost we need not consider the "other side" of **IC** in optimizing (12). If Non-Trivial Cost is violated, then the optimal system is either the one characterized by this Proposition or the system that always recommends to every agent.

PROOF. We assume that the maximum value is strictly positive and derive properties of the maximizing solution $\{l_k\}$. Because ϕ is strictly increasing, there exists a unique smallest integer k^* such that $\phi(\frac{k^*}{d}) \geq c$, and the multiplicative factor $\phi(\frac{k}{d}) - c$ is non-negative precisely when $k \geq k^*$. In order to maximize π_v, l_k has be to be as large as possible for $k > k^*$ and as small as possible for $k < k^*$.

We show that the sequence $\{l_k\}$ is weakly increasing. Suppose not, then $l_{k+1} < l_k$ for some k. If $k < k^*$, then we can make $l_0, l_1, \ldots, l_k$ all slightly smaller to increase $\pi(v)$ while maintaining the constraints. If $k \geq k^*$, then similarly we can make $l_{k+1}, \ldots, l_n$ all slightly larger without violating the constraints. The resulting $\pi(v)$ is again higher. Either way we get a contradiction.

We label the constraints of the optimization problem as:

$$l_{k+1} \leq exp(\epsilon) \cdot l_k \tag{DP1'}$$

$$1 - l_{k+1} \geq exp(-\epsilon) \cdot (1 - l_k). \tag{DP2'}$$

We note that if $l_{k+1} \leq \frac{exp(\epsilon)}{exp(\epsilon)+1}$ then **DP1'** implies **DP2'**. The implication goes the other way if this inequality is reversed.

Because we would like to make l_k as large as possible for $k > k^*$ and as small as possible for $k < k^*$, at each k one of the DP constraints must bind. Suppose first that $l_0 > \frac{exp(\epsilon)}{exp(\epsilon)+1}$. Then **DP2'** binds for each k, and we can write l_k as a linear function of l_0. $\pi(v)$ is also linear, so it is maximized when $l_0 = 1$ or $l_0 = \frac{exp(\epsilon)}{exp(\epsilon)+1}$. The former cannot happen because we assumed that full recommendation does not satisfy **IC**. Hence it is W.L.O.G. to assume $l_0 \leq \frac{exp(\epsilon)}{exp(\epsilon)+1}$.

Now define $\bar{k} \geq 0$ to be the largest k such that $l_k \leq \frac{exp(\epsilon)}{exp(\epsilon)+1}$. It follows from previous observations that **DP1'** binds if $k < \bar{k}$ and **DP2'** binds otherwise. So

$$l_k = \begin{cases} exp(\epsilon(k - \bar{k})) \cdot l_{\bar{k}}, \text{ if } k \leq \bar{k}; \\ 1 - exp(\epsilon(\bar{k} - k)) \cdot (1 - l_{\bar{k}}), \text{ if } k > \bar{k}. \end{cases}$$

This implies $\pi(v)$ is linear in $l_{\bar{k}}$. By the definition of $\bar{k}$, $l_{\bar{k}}$ can take value on the interval $[\frac{1}{exp(\epsilon)+1}, \frac{exp(\epsilon)}{exp(\epsilon)+1}]$. Therefore $\pi(v)$ is maximized when $l_{\bar{k}} = \frac{exp(\epsilon)}{exp(\epsilon)+1}$ or $\frac{1}{exp(\epsilon)+1}$. In the former case the characterization of optimal recommendation follows directly. In the latter case we note that $\bar{k} < n$, for otherwise **DP2'** never binds and we can multiply the sequence $\{l_k\}$ by a constant $t > 1$ to achieve a higher profit without violating the constraints. Moreover $l_{\bar{k}+1} = \frac{exp(\epsilon)}{exp(\epsilon)+1}$, and the characterization follows by taking $\bar{k}+1$ as $\bar{k}$.

Hence we have shown that the constrained optimization problem is maximized by a cutoff-type solution whenever the maximum value is positive. It remains to find the optimal $\bar{k}$. The characterization (13) is immediately obtained by considering the alternatives $\bar{k} \pm 1$ and using the fact that $\pi(v)$ could not be higher. Also note that

$$f(\bar{k}) := \sum_{k=0}^{d} (\phi(\frac{k}{d}) - c) p_k^d \cdot exp(-|k - \bar{k}|\epsilon) \tag{14}$$

satisfies

$$f(\bar{k}) = exp(-\epsilon) f(\bar{k} - 1) + (exp(\epsilon) - exp(-\epsilon)) \cdot \sum_{k=\bar{k}}^{d} (\phi(\frac{k}{d}) - c) p_k^d \cdot exp(-|k - \bar{k} + 1|\epsilon).$$

Suppose $f(\bar{k}-1) \geq 0$. If $\bar{k} - 1 \geq k^*$, then $\sum_{k=\bar{k}}^{d} (\phi(\frac{k}{d}) - c) p_k^d \cdot exp(-|k - \bar{k} + 1|\epsilon)$ is trivially positive. Otherwise

$$\sum_{k=\bar{k}}^{d} (\phi(\frac{k}{d}) - c) p_k^d \cdot exp(-|k - \bar{k} + 1|\epsilon) = f(\bar{k} - 1) - \sum_{k=\bar{0}}^{\bar{k}-1} (\phi(\frac{k}{d}) - c) p_k^d \cdot exp(-|k - \bar{k} + 1|\epsilon)$$

is again positive. It follows that whenever $f(\bar{k} - 1) \geq 0$, $f(\bar{k}) > 0$, i.e. $f(\cdot)$ is first negative then positive. From this we conclude that the $\bar{k}$ determined from (13) is unique, and the associated sequence $\{l_k\}$ maximizes $\pi(v)$.

Finally, a necessary condition for the existence of $\bar{k}$ is that $f(d) \geq 0$, or

$$\sum_{k=0}^{d} (\phi(\frac{k}{d}) - c) p_k^d \cdot exp(\epsilon k) \geq 0.$$

This implies $c \leq \bar{c}(d, \epsilon, p)$, completing the proof of this Proposition as well as the necessity part of Proposition 5.1. □

Although Proposition 5.4 is stated in terms of l_k, it is straightforward to recover a set of primitive recommendation probabilities $l^v(\Theta)$ that maximize expected utility subject to **IC** and **DP** : just let $l^v(\Theta) = l_k$ for $k = \theta_{v_1} + \cdots + \theta_{v_d}$.

The intuition for this result is that in order to make an effective recommendation while preserving privacy, the designer should recommend aggressively when there is clear benefit to the agent, i.e. when $k > \bar{k}$. Above this threshold, designer recommends as frequently as possible without violating **DP2**. Below this threshold the designer fully exploits the limits of **DP1** to minimize its impact on the agent's incentives.

PROPOSITION 5.5. *$\bar{k}(d, \epsilon, p, c)$ as defined in (13) is decreasing in p and increasing in c. Moreover if ϕ is twice differentiable with $\phi' > 0$ and $|\phi''| < M < \infty$ on (0,1), then for fixed ϵ, p and $c < \phi(\frac{p \cdot \exp(\epsilon)}{1 - p + p \cdot \exp(\epsilon)})$, we have $\bar{k}(d) = \phi^{-1}(c) \cdot d + O(1)$.*

PROOF. Recall that $\bar{k}$ was characterized in (13). Using the definition in (14), we see that $\bar{k}$ is the smallest m such that $f(m) \geq 0$, or

$$\sum_{k=0}^{d} (\phi(\frac{k}{d}) - c) \binom{d}{k} (\frac{p}{1-p})^{k-k^*} \cdot exp(-|k - m|\epsilon) \geq 0.$$

Increasing p or decreasing c makes the LHS bigger, so $\bar{k}$ is decreasing in p and increasing in c.

Now we turn to consider the asymptotic behavior of $\bar{k}$. From the assumption that $c < \phi(\lambda)$ and Proposition 5.3, we know that $c < \bar{c}(d, \epsilon, p)$ eventually. Therefore $\bar{k}(d, \epsilon, p, c)$ is

finite for all large d. Furthermore, simple algebra allows us to write $f(m)$ as a positive constant times

$$\sum_{k=0}^{m}(\phi(\frac{k}{d})-c)\binom{d}{k}\left(\frac{p\cdot exp(\epsilon)}{1-p}\right)^{k-k^*}+\sum_{k=m+1}^{n}(\phi(\frac{k}{d})-c)\binom{d}{k}\left(\frac{p}{(1-p)\cdot exp(\epsilon)}\right)^{k-k^*}. \quad (15)$$

Call the first summand in (15) $S(m)$ and the second $T(m)$. We claim that $m = k^* + \Omega(1)$ ensures $S(m) \geq 0$ (and $T(m) \geq 0$ trivially). Note that $S(m) \geq 0$ can be rewritten as:

$$\sum_{k=0}^{k^*-1}(c-\phi(\frac{k}{d}))\frac{\binom{d}{k}}{\binom{d}{k^*}}\left(\frac{p\cdot exp(\epsilon)}{1-p}\right)^{k-k^*}\leq\sum_{k=k^*}^{m}(\phi(\frac{k}{d})-c)\frac{\binom{d}{k}}{\binom{d}{k^*}}\left(\frac{p\cdot exp(\epsilon)}{1-p}\right)^{k-k^*}. \quad (16)$$

When $k \leq k^*$, the ratio $\frac{\binom{d}{k}}{\binom{d}{k^*}} = \frac{k^*\cdots(k+1)}{(d-k^*+1)\cdots(d-k+1)} \leq (\frac{k^*}{d-k^*+1})^{k^*-k}$. Since $k^* = \phi^{-1}(c)\cdot d + O(1)$ and $c < \phi(\lambda)$, we find that for some $\delta > 0$,

$$\frac{\binom{d}{k}}{\binom{d}{k^*}}\left(\frac{p\cdot exp(\epsilon)}{1-p}\right)^{k-k^*}\leq(1-\delta)^{k^*-k},\forall k\leq k^*,$$

and similarly

$$\frac{\binom{d}{k}}{\binom{d}{k^*}}\left(\frac{p\cdot exp(\epsilon)}{1-p}\right)^{k-k^*}\geq(1+\delta)^{k-k^*},\forall k=k^*+O(1).$$

It follows that the LHS of (16) is at most $\sum_{k=0}^{k^*-1}(\phi(\frac{k^*}{d})-\phi(\frac{k}{d}))(1-\delta)^{k^*-k} \leq \frac{||\phi'||}{d}\sum_{k=0}^{k^*-1}(k^*-k)(1-\delta)^{k^*-k}$ by the mean value theorem. Since $|\phi''| \leq M$, ϕ' must be bounded. Furthermore the quasi-geometric sum converges absolutely. Thus the LHS of (16) is $O(\frac{1}{d})$.

On the other hand, a crude lower bound for the RHS is $(m-k^*)\left(\phi(\frac{k^*+1}{d})-\phi(\frac{k^*}{d})\right) \geq \frac{m-k^*}{2d}\phi'(\phi^{-1}(c))$. The inequality follows again from the mean value theorem, or more rigorously from a second order Taylor expansion of $\phi(\cdot)$ at $\phi^{-1}(c)$. As $\phi' > 0$ we conclude that $m-k^* = \Omega(1)$ is sufficient for (16) to hold. This proves $\bar{k}(d,\epsilon,p,c) \leq k^*+O(1)$.

To establish $\bar{k}(d,\epsilon,p,c) \geq k^* - O(1)$, it suffices to show that $m = k^* - \Omega(1)$ implies $T(m) \leq 0$ (and $S(m) <$ trivially). From the Non-Trivial Cost assumption and using similar techniques as in **Proposition 5.3**, we have that $c \geq \phi(p)$. From this we can approximate $T(m)$ by a geometric sum and prove it is non-negative. The calculations are similar to what we have done but tedious, so we omit them here. □

ACKNOWLEDGMENTS

We thank Yaron Singer and Bo Waggoner for helpful discussions. We thank 4 anonymous reviewers for very insightful comments.

REFERENCES

ABEZGAUZ, I. 2013. Facebook vulnerability discloses friends lists defined as private. http://www.quotium.com/research/advisories/Facebook_Vulnerability_Discloses_Private_Friends_list.php.

ADOMAVICIUS, G. AND TUZHILIN, A. 2005. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering 17,* 6, 734–749. 03933.

BLUME, L. E. 1993. The statistical mechanics of strategic interaction. *Games and Economic Behavior 5,* 3, 387–424.

BLUME, L. E. 1995. The statistical mechanics of best-response strategy revision. *Games and Economic Behavior 11,* 2, 111–145.

BU, J., TAN, S., CHEN, C., WANG, C., WU, H., ZHANG, L., AND HE, X. 2010. Music recommendation by unified hypergraph: combining social media information and music content. In *Proceedings of the international conference on Multimedia*. ACM, 391–400.

DWORK, C. 2006. Differential privacy. In *Automata, languages and programming*. Springer, 1–12.

DWORK, C. 2008. Differential privacy: A survey of results. In *Theory and Applications of Models of Computation*, M. Agrawal, D. Du, Z. Duan, and A. Li, Eds. Number 4978 in Lecture Notes in Computer Science. Springer Berlin Heidelberg, 1–19. 00335.

ELLISON, G. 1993. Learning, local interaction, and coordination. *Econometrica 61,* 5, 1047–1071. 00980.

FACEBOOK. 2008. People you may know, https://www.facebook.com/help/501283333222485. https://www.facebook.com/help/501283333222485/.

GUY, I., ZWERDLING, N., RONEN, I., CARMEL, D., AND UZIEL, E. 2010. Social media recommendation based on people and tags. In *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval*. ACM, 194–201.

HSU, J., HUANG, Z., ROTH, A., ROUGHGARDEN, T., AND WU, Z. S. 2013. Private matchings and allocations. *arXiv preprint arXiv:1311.2828*.

HUANG, Z. AND KANNAN, S. 2012. The exponential mechanism for social welfare: Private, truthful, and nearly optimal. In *Foundations of Computer Science (FOCS), 2012 IEEE 53rd Annual Symposium on*. IEEE, 140–149.

JACKSON, M. O. AND YARIV, L. 2007. Diffusion of behavior and equilibrium properties in network games. *The American Economic Review 97,* 2, 92–98. 00125.

JOHNSON, C. 2013. Introducing netflix social, http://blog.netflix.com/2013/03/introducing-netflix-social.html. http://blog.netflix.com/2013/03/introducing-netflix-social.html.

KEARNS, M., PAI, M., ROTH, A., AND ULLMAN, J. 2014. Mechanism design in large games: Incentives and privacy. In *Proceedings of the 5th conference on Innovations in theoretical computer science*. ACM, 403–410.

KONSTAS, I., STATHOPOULOS, V., AND JOSE, J. M. 2009. On social networks and collaborative recommendation. In *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*. ACM, 195–202.

MACHANAVAJJHALA, A., KOROLOVA, A., AND SARMA, A. D. 2011. Personalized social recommendations: accurate or private. *Proceedings of the VLDB Endowment 4,* 7, 440–450.

MCFADDEN, D. L. 1976. Quantal choice analaysis: A survey. In *Annals of Economic and Social Measurement, Volume 5, number 4*. NBER, 363–390.

MCSHERRY, F. AND TALWAR, K. 2007. Mechanism design via differential privacy. In *Foundations of Computer Science, 2007. FOCS'07. 48th Annual IEEE Symposium on*. IEEE, 94–103.

MYERSON, R. B. 1988. Incentive constraints and optimal communication systems. In *Proceedings of the 2nd Conference on Theoretical Aspects of Reasoning about Knowledge*. Morgan Kaufmann Publishers Inc., 179–193.

NARAYANAN, A. AND SHMATIKOV, V. 2008. Robust de-anonymization of large sparse datasets. In *IEEE Symposium on Security and Privacy, 2008. SP 2008*. 111–125.

00453.
NISSIM, K., SMORODINSKY, R., AND TENNENHOLTZ, M. 2012. Approximately optimal mechanism design via differential privacy. In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*. ACM, 203–213.
YE, M., YIN, P., AND LEE, W.-C. 2010. Location recommendation for location-based social networks. In *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*. ACM, 458–461.
YOUNG, H. P. AND KREINDLER, G. E. 2012. Rapid innovation diffusion in social networks. Economics Series Working Paper 626, University of Oxford, Department of Economics. 00000.
ZHU, T., LI, G., REN, Y., ZHOU, W., AND XIONG, P. 2013. Differential privacy for neighborhood-based collaborative filtering. In *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. ACM, 752–759.