The Value of Community Information for Pricing Under Network Externalities

Calvin Roth[1], Ankur Mani[2], and Jiali Huang[3]

[1]Department of Industrial and Systems Engineering, University of Minnesota,
[2]Department of Industrial and Systems Engineering, University of Minnesota,
[3]Meta Inc.

April 20, 2023

Abstract

We study the value of network information for pricing decisions for a monopolist selling divisible goods with positive network externalities to consumers in a social network. In this setting we consider consumers whose utility functions are homogeneous except for there position in the network. We study optimal pricing structure when there is uncertainty about the network structure and provide general expressions for the optimal price and profit. Prior work has shown that in a large class of networks pricing using the information about the social network does not improve profit significantly for a large class of random networks. Yet, it is unclear if there is value of network information for some random networks. We study the value of network information for dense stochastic block models. For these networks each block gives information on how communities on average interact with the rest of the network instead of consumers. We find that for this class of networks there will be significant losses when pricing decisions are made with no information about the network compared to optimal profits under full network information. We also find that for asymptotically large networks making pricing decisions using only the information about community structure and community memberships of the consumers does not introduce and significant change in expected profits over optimal pricing under full network information.

1 Introduction

The popularity of social platforms such as Facebook, Twitter, Wechat and Instagram has created an ecosystem that makes it easier to harness peer effects for marketing and revenue management. On one hand, these platforms enable convenient online interactions, peer monitoring and peer influence. On the other hand, they provide rich information about network structure, peer interactions and influence. Firms do use this network information to harness peer effects and shape global outcomes in social networks such as spreading information and ideas, increasing product adoption, increasing sales of products through discriminative pricing. While price discrimination is an old practice the availability of social network information allows for fine grained price discrimination.

Yet, with all the optimism, there are caveats associated with using peer effects and social network information. On one hand, social network data is still costly, noisy, and its use raises concerns about privacy, balance of power, and equity. Yet, recent work surprisingly shows that the value of network information for pricing vanishes in a class of large social networks **??**. The work suggests that for these social networks optimal uniform price can generate optimal profit for the monopolist and network information may be ignored. This work raises valid concerns about the practice of using social network information for price discrimination. Although, the work shows that there may be some value of network information for pricing in moderate sized Erdös-Renyi networks with critical density, whether there exists a class of large networks for which there is significant value of network information for pricing is still an open question. Our work is motivated by this open question. We aim to answer the following main research questions.

- Does there exist a class of networks for which there is significant value of network information for pricing such that the vale does not vanish with network size?
- Further, if there is value of network information for a class of networks, is the full network information useful or is a smaller amount of information sufficient?

The answer to these questions is important for managers. The full network information is almost always unavailable and often the managers have to work with partial and noisy network information. The full network information also has very large dimensionality and is costly to obtain. If there is a lower dimensional information that is useful and easy to obtain then that will help the managers' problem.

Our contributions are as follows.

- We first study the monopolist's pricing problem under perfect network information and group homogeneity constraints when the set of agents are partitioned into certain groups. Under such constraints the prices must be homogeneous within each group but can differ across groups. We identify the optimal prices for each group given the network information and groups. We show that these optimal prices reduce to optimal uniform price when all agents belong to one group and optimal discriminative prices when each agent forms its own singleton group.
- We then study the monopolist's pricing problem under partial network information. Partial information introduces a distribution over the set of possible networks. We identify the structure of optimal pricing under uncertainty and the optimal expected profit. This general pricing framework could be useful for studying pricing under a variety of possible partial information sets with corresponding distributions over networks.
- We then introduce the monopolist's pricing problem when the monopolist only knows the community structure in the network and community memberships of the agents. In particular the networks are assumed to be sampled using the stochastic block model and the monopolist knows the asymmetric block matrix where each element represents the edge probabilities from one community to another. We first identify the optimal prices for the expected network and show that it is equivalent to optimal price under group homogeneity constraints if the groups are aligned with communities in the expected network. We then identify the optimal price under uncertainty conditioned upon the community information building upon our results on optimal pricing under partial network information.
- We show that for stochastic block models with density $\Omega(log(n)/n)$ the value of network information is significant and the optimal pricing under full network information generates significantly more profits

in expectation than optimal uniform pricing ignoring the network information. Thus, we are able to identify a class of networks for which there is significant value of network information for pricing answering the first main research question we stated.

- We also show that for stochastic block models with density $\Omega(log(n)/n)$ the expected profit under optimal pricing given the information about the community structure and community memberships of agents is near optimal and the difference in profit due to lack of full network information vanishes in the size of the network. This answers the second main research question. The amount of information in the block matrix grows quadratically in the number of communities instead of the number of agents and the amount of information about the community memberships grows linearly in the number of agents. The amount of information needed for perfect network information on the other hand grows quadratically in the number of agents.

The outline of the rest of the paper is as follows. In section 3 we introduce the basic model of network externalities. Following that in section 4 we consider problem of pricing under an additional family of constraints where consumers in the same group most be charged the same price as well as the concept of pricing when there is uncertainty in the network. Section 5 introduces the primary family of network we will study the Stochastic Block Model. In section 6 we state the main results of this work which is following by section 7 we discuss at a high level the ideas used in the proof with the details in the Appendix 11. In section 8 we provide an extension on how to price when there is uncertainty about the underlying structure used in the Block Stochastic Model. The validity of this results is then demonstrated through simulations in section 9. Finally, we give conclusions and suggestions for future work.

2 Literature Review

Using social network information to harness peer effects has been an active research topic in recent years. Network effects shape critical outcomes in a social network such as the spread of information, ideas and disease (see, e.g., **?**, **?**, **?**, **?**), adoption of products (see, e.g., **?**,**?**, **?**) among others.

Our work is related to the broader literature on positive network externalities **?**, **?** and in particular to marketing with positive network externalities. This literature is categorized in two streams. The first stream aims at maximizing adoption through *influence maximization* and the second stream aims at *revenue maximization* through pricing decisions.

Influence maximization initially introduced in **?** aims at identifying the best *seeds* to trigger cascades of influence that maximize the overall spread of influence and adoption in the social network. Subsequent literature has studies influence maximization in various settings, see, e.g., **?** and **?**. **?** showed that this problem is computationally hard, but approximation algorithms provide good performance with provable guarantees. Recent work **?** studied the value of network information for influence maximization and showed that for a class of random networks, random seeding strategies with a few more seeds can lead to larger cascades than optimal seed selection. Follow up work **?** has showed that this may not be the case for all networks and uncertainty structures. This follow up work is in the spirit of our work.

Our work mainly belongs to the revenue maximization stream. Revenue maximization problems aim at maximizing the revenue through pricing. There has been much research using pricing under network effects in the setting of sequential purchase decisions (see, e.g., **?**, **?**, **?**, **?**, **?**) and simultaneous purchase decisions (see, e.g., **?**, **?**, **?**, **?**). Such problems are often computationally hard and most literature focuses on the computational aspect.

Our work is closely related to the static pricing problem of selling a divisible product to a group of consumers with positive network externalities, in particular, the works of **?**, **?**,**?**, **?** and **?**, **?**. In particular, we build upon the model in **?**. These works consider a two-stage game, in which the monopolist first chooses the prices and then the consumers, embedded in a social network, make purchasing decisions simultaneously. **?** and **?** assume full network information, while **?**, **?**,**?**, and **?** assume incomplete network information. In these works, prices can be set differently based upon consumers' positions in the social network. The question of interest in **?**, **?** and **?** is to characterize and identify optimal prices and profits in the networks and the complexity of computing optimal prices. In contrast, **?**, **?** quantify the value of network information for pricing in Erdös-Renyi random networks and Power-law networks under configuration model. Their results shows that in such networks the value of network information vanishes asymptotically in the size of the network. Our work is the first to show that there exists a class of random networks (with community structure that persists asymptotically) for which the value of network information does not vanish asymptotically. Further, our work is the first to show that the in such networks, the monopolist does not need the full network information and the community structure and membership information is sufficient to derive near optimal profit, thus establishing that the community information is very valuable while rest of the information about the network may not have much value.

Moreover, in developing the asymptotic value of network information in random networks, we find connections with graph theory and random graph theory, in particular, the literature on counting the number of walks of different lengths in a network (see, e.g., **?**, **?**), on spectral graph theory (see, e.g., **?**, **?**, **?**), and on network centrality (see, e.g., **?**).

3 Basic Model

This section introduces the basic ideas of utility model of the network, and the profit function of the monopolist. Specifically, we consider optimal discriminative pricing, uniform pricing, and pricing in an intermediate range where the network is partitioned and everyone in the same partition must be charged the same price. We consider a firm selling a divisible good to consumers in a social network. Such a network with n consumers will be represented as a matrix $G \in \mathbb{R}_+^{n\times n}$ where G_{ij} is the influence the jth consumer has on the ith. We note that these entries can't be non-negative. We impose the requirement that $G_{ii} = 0, i \in [n]$ which means we can't influence ourselves. This model has been studied mathematically by **?**, **?**, **?**, **?** and **?** . This is modelling choice that reflects real world social networks where the behavior of those around us positively influences our actions. **?** showed the impact of peer behavior by looking specifically at how winning the lottery affected how the neighbors of the winners made purchases. **?** analyzed the social network of recommendations for several products including books and DVDs. We assume that the utility of the ith consumer is a function of how much they consume, x_i, how much everyone else consumes, $\mathbf{x}_{-i}$, and the price this ith consumer is charged p_i. Specifically, the utility function is given as

$$u_i(x_i, \mathbf{x}_{-i}, p_i) = ax_i - p_i x_i - x_i^2 + 4\rho x_i \sum_{j\neq i} \frac{G_{ij}}{\|G+G^T\|} x_j$$

The ax_i is an amount of positive utility that the consumer gets from purchasing gets from purchasing the good where $a > 0$ is a parameter reflecting the amount of utility they get per unit of good. $-p_i x_i$ is the utility they lose from purchasing where p_i is the price they are charged. The term $-x_i^2$ encapsulates a concept of decreasing marginal return and ensures that consumers always purchase a finite amount. Next, the term $4\rho \sum_{j\neq i} \frac{G_{ij}}{\|G+G\|} x_j$ adds another source of marginal utility that consumer i gets in proportion to the consumption their peers consumers. For instance if j is a peer of i and j purchases a lot of the good then this increases the marginal utility that consumer i would get from purchasing more of the product. The $\frac{1}{\|G+G^T\|}$ term is 1 over the largest magnitude eigenvalue of $G + G^T$ and functions as a normalization factor for the network effort introduced in **?** allows. This allows the strength of the network effect to be tuned with a single parameter $\rho \in (0, 1)$. This model of utility has been previously studied in work of **?**, **?** and **?** . In this model, the only difference in any two consumers' utility function is the price they are charged and their position in the network.

The monopolist first determines the price to charge each consumer, which can be discriminative in the sense that different people can be charged different prices. We use a vector $\mathbf{p}$ to denote the prices. Given these prices **?**, each consumer will choose a level of consumption to maximize their utility and this forms a consumption game. There is an equilibrium point of this game that is solved by a system of linear equations with a solution of

$$\mathbf{x}^*(\mathbf{p}) = \frac{1}{2}\left(I - 2\rho\frac{G}{\|G+G^T\|}\right)^{-1}(a\mathbf{1} - \mathbf{p})$$

Where $\mathbf{x}$ is a vector where the ith component is the consumption of consumer i. The expression $\left(I - 2\rho\frac{G}{\|G+G^T\|}\right)^{-1}$ is well defined and exists because $2\rho\frac{G}{\|G+G^T\|}$ has all eigenvalues all stricter stricter less than one, and therefore this matrix will have a well defined inverse **?** . The term $(I - 2\frac{\rho}{\|G+G^T\|}G)^{-1}$ can be expanded as $\sum_{k=0}^{\infty}(2\frac{\rho}{\|G+G^T\|})^k G^k$. The ijth element of the matrix for each k in the sum represents the number walks of length k from one consumer to another weighted by a discount factor of $2\frac{\rho}{\|G+G^T\|} < 1$.

Given the consumption equilibrium $\mathbf{x}^*(\mathbf{p})$, the monopolist's profit is defined as

$$\pi(\mathbf{p}) = (\mathbf{p} - c\mathbf{1})^T \mathbf{x}^*(\mathbf{p}),$$

where c is the cost to produce one unit of a good and we assume that $c < a$. The monopolist's problem is to set prices to maximize their profit. By effectively using certain network information, the monopolist could benefit from offering discounts to those who are highly influential in the network and then make this money back by charging others more.

4 Group Homogeneous Pricing

Now we consider actions of a monopolist under a group homogeneity constraint. In this constraint the set of agents are partitioned into groups and the monopolist is required to charge the same price for each member of a group. The prices may vary across groups. The monopolist will have knowledge of these group memberships. **?** studied a situation where the monopolist given two prices had to determine how to partition the network. They then showed that this problem is NP-Hard by relating it to the MAX-CUT problem. Our work differs because the monopolist is given a partition of agents which can be more than 2 groups and using this partition we derive the optimal price the monopolist should charge under these constraints. Often times the monopolist will not have control over how to partition the agents. For example, each group may represent consumers living in the same geographical region, members of the same organizations, be part of the same demographic, or share common interests. Formerly, we require that the groups are disjoint and form a partition of consumers in the network but otherwise groups are not restricted to a particular form. For instance we don't require that a group be connected in the network. We will specify the group identity of consumers by matrix $\mathcal{R} \in \mathbb{R}^{m \times n}$ where m is the number of groups and n is the network size. We define

$$\mathcal{R}_{ij} = \begin{cases} 1 & \text{If j is member of group i} \\ 0 & \text{Otherwise} \end{cases}$$

As a consequence of the groups forming a partition of the network, we have that each column of $\mathcal{R}$ sums to 1. The monopolist's problem under the context of constrained prices is to maximize the profit with the constraint that people in the same group be charged the same price. Mathematically, the monopolist's problem is

$$\begin{aligned} &max_p (\mathbf{p} - c\mathbf{1})^T \mathbf{x}^*(\mathbf{p}) \\ &s.t. \; p_i = p_j \text{ if } \; \mathcal{R}_i^T \cdot \mathcal{R}_j = 1 \end{aligned}$$

where $\mathcal{R}_i$ is the ith column of $\mathcal{R}$.

Our first result says that form of the optimal price to charge for each community, when we know the community structure $\mathcal{R}$ and the network G is the solution to a set of linear equations.

Theorem 4.1. *With group homogeneous constraints $\mathcal{R}$ the optimal price to charge each group $\boldsymbol{p}^*_{\mathcal{R}}$ is the solution to the system:*

$$\mathcal{R}\left(B + B^T\right)\mathcal{R}^T \boldsymbol{p}^*_R = \mathcal{R}\left(aB + cB^T\right)\boldsymbol{1},$$

where $B = (I - 2\frac{\rho}{\|G+G^T\|}G)^{-1}$

This theorem specifies the price to charge each group so the optimal price vector for each consumer pricing is $\mathcal{R}^T \mathbf{p}^*_R$. Due to it's frequent use we will use $B = (I - 2\frac{\rho}{\|G+G^T\|}G)^{-1}$ throughout this paper. The price the monopolist should charge explicitly is

$$\mathbf{p}^*_{\mathcal{R}} = \mathcal{R}^T \left(\mathcal{R}(B + B^T)\mathcal{R}^T\right)^{-1} \mathcal{R}(aB + cB^T)\mathbf{1} \tag{1}$$

There are two special cases of this constraint. The first is when there is only one group which equivalently is requiring that everyone is charged the same price. In this case $\mathcal{R}$ has only one row of all 1s. The corollary below shows under this case, we recover the optimal uniform price vector as studied by **?** and **?**.

Corollary 4.1. *When* $\mathcal{R} = \boldsymbol{1} \in \mathbb{R}^{1\times n}$ *the optimal price vector is*

$$\boldsymbol{p}^*_{\boldsymbol{1}} = \frac{a+c}{2}\boldsymbol{1} \tag{2}$$

.

The other special case is if everyone is in their own community which effectively means there are no constraints on price. In this case the matrix $\mathcal{R}$ is an $n \times n$ identity matrix.

We will show three separate forms of the price vector in this context and show they are equivalent. The first two are useful for the analysis needed in this paper and the third is form of the optimal price used in the work of **?** and **?** .

Corollary 4.2. *When the constrained matrix is* $\mathcal{R} = I$ *the optimal price of the monopolist* $\boldsymbol{p}^*_I$ *can by expressed in three equivalent forms:*

1. $(B + B^T)^{-1}(aB + cB^T)\boldsymbol{1}$

2. $\frac{a+c}{2}\boldsymbol{1} + \frac{a-c}{2}(B + B^T)^{-1}(B - B^T)\boldsymbol{1}$

3. $\frac{a+c}{2}\boldsymbol{1} + \frac{a-c}{2}\frac{\rho}{\|G+G^T\|}(G - G^T)\left(I - \frac{\rho}{\|G+G^T\|}(G + G^T)\right)^{-1}\boldsymbol{1}$

Example Before we move on we consider an example of this restricted pricing. We consider the following network and a possible partition of it.

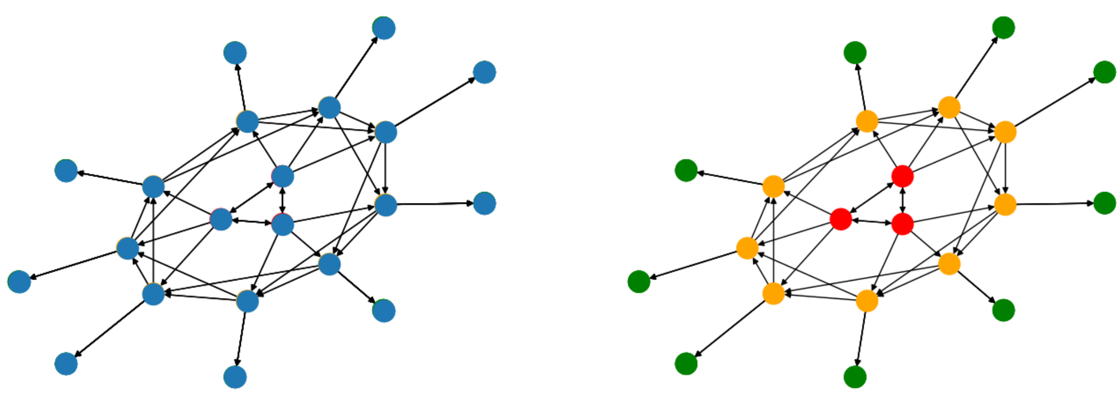


Figure 1: One the left is the original network and one the right we partition it where we require that each consumer of the same color be charged identical prices.

This partition is a natural choice because all consumers in a color are indistinguishable from one another. For this reason we might guess that even without this same color same price constraint we might charge each node in a color the same price anyways. This guess is right. Using $a = 6, c = 4, \rho = 0.9$ both the unconstrained and constrained to colors case we find that the red consumers are equally charged 10.015, each yellow is charged 3.946, and every green consumer is charged 3.328 and both get a profit of 108.77. But this example motivates the concept that consumers that have similar positions in the network can be charged the same price with low loss. In this example, the consumers are exactly the same and we don't lose any profit if we price only knowing a consumer's "color label". This observation that are partition of the network into a small number of groups relative to the size of the network isn't necessarily worse than using full information motivates exploration later in the paper of the stochastic block matrix.

4.1 Partial Information

We now consider the monopolist's pricing under partial information of the network. Combined with the previous section we will be able to look at the effects of group homogeneous pricing in the context of random graph models.

In our analysis, we will introduce two notions to establish results for a distribution of networks. The first is the notion of the true network. The monopolist instead of seeing this network has some information $\mathcal{I}$ that generates a probability distribution $\mathcal{F}$ of networks. The monopolist's problem in this regime is to maximize the profit given their information $\mathcal{I}$.

$$max_{\mathbf{p}} \frac{1}{2}(\mathbf{p} - c\mathbf{1})^T \mathbb{E}_{\mathcal{F}} \left[\left(I - \frac{2\rho}{\|G + G^T\|} G \right)^{-1} |\mathcal{I} \right] (a\mathbf{1} - \mathbf{p}) \tag{3}$$

In general computing $B_{\mathcal{F}} := \mathbb{E}_{\mathcal{F}}[(I - 2\frac{G}{\|G+G^T\|})^{-1}|\mathcal{I}]$ may be difficult for arbitrary classes of information. But if we can do this computation then the analysis for the optimal price is essentially the same as in the deterministic case but we use $B_{\mathcal{F}}$, the expected matrix.

Proposition 4.1. *If the monopolist has $\mathcal{I}$ about the state of the network the optimal price to charge is*

$$\boldsymbol{p}^* = \frac{a+c}{2}\boldsymbol{1} + \frac{a-c}{2}(B_{\mathcal{F}} + B_{\mathcal{F}}^T)^{-1}(B_{\mathcal{F}} - B_{\mathcal{F}}^T)\boldsymbol{1}$$

Where $B_{\mathcal{F}} = \mathbb{E}_{\mathcal{F}}[(I - 2\frac{\rho}{\|G+G^T\|}G)^{-1}|\mathcal{I}]$.

$B_{\mathcal{F}}$ like it's deterministic counterpart has the interpretation of being equal $\mathbb{E}_F[\sum_{k=0}^{\infty}(2\frac{\rho}{\|G+G^T\|})^k G_{\mathcal{F}}^k|\mathcal{I}]$ but now it is the discounted sum of walks between consumers over a distribution of possible networks. The profit the monopolist experiences can be lower because the consumption levels are governed by the true matrix $B := \left(I - 2\rho\frac{G}{\|G+G^T\|}\right)^{-1}$.

Proposition 4.2. *The expected profit of the monopolist given $B_{\mathcal{F}} = \mathbb{E}_{\mathcal{F}}[(I - 2\frac{\rho}{\|G+G^T\|}G)^{-1}|\mathcal{I}]$ is given by*

$$\frac{(a-c)^2}{8}\left(\boldsymbol{1} + (B_{\mathcal{F}} + B_{\mathcal{F}}^T)^{-1}(B_{\mathcal{F}} - B_{\mathcal{F}}^T)\boldsymbol{1}\right)^T B_{\mathcal{F}} \left(\boldsymbol{1} - (B_{\mathcal{F}} + B_{\mathcal{F}}^T)^{-1}(B_{\mathcal{F}} - B_{\mathcal{F}}^T)\boldsymbol{1}\right)$$

This proof follows from expanding the profit function in terms of first a particular true network to get a profit the monopolist would experience if that network were selected and then calculating the expected value over these possible profits. Details can be found in 11.2

5 Pricing in Stochastic Block Models

Our main results in this paper focus on studying the monopolist's pricing problem with networks generated from the stochastic block models. In this section, we will formally introduce the stochastic block model, the monopolist's pricing problem under the given network structure, and a framework to evaluate the value of different pricing strategies.

5.1 Stochastic Block Model

A stochastic block model is a generative network model for random networks. In this model, the agents are partitioned into m communities. The consumers in a given community act in a statistically similar way. Specifically, if i and i' are in the same community and j and j' are in the same community then $Pr(G_{ij} = 1) = Pr(G_{i'j'} = 1)$. Such a system can be described when one knows the probabilities between each block, including a block to itself, and how many consumers are in each block. We assume the number of consumers in each block to be the same, and denote the size of each community by n. Since we also assume the agents are partitioned into m communities, the total number of agents in the network is mn. To describe the relationship between two communities, we introduce a matrix $\mathcal{M} \in \mathbb{R}^{m\times m}$, where

$$\mathcal{M}_{ij} = Pr(\text{consumer } k \text{ in community } i \text{ is connected to consumer } l \text{ in community } j).$$

This matrix will be referred to as the community structure. The community structure need not be symmetric.

To generate a random block stochastic matrix, we need the number of communities, the size of each community, the probability of an edge from a consumer in one community to a consumer in another community, and labels for which consumers are a part of which community. To handle this last part, we introduce a matrix $\mathcal{R} \in \mathbb{R}^{m \times mn}$ to track what community each consumer is in. Specifically, we define

$$\mathcal{R}_{ij} = \begin{cases} 1 & \text{If } j \in \text{Community } i, \\ 0 & \text{Otherwise.} \end{cases}$$

Given the above information, we generate a block stochastic matrix $G_{\mathcal{M}}(n)$ by filling in an $mx \times mn$ adjacency matrix that is randomly filled with ones and zeros. The value $G_{\mathcal{M}}(n)_{kl}$ is one with probability M_{ij}, where k is in community i and l is in community j. Each edge is sampled independently.

For the purpose of later analysis, we also define $\bar{G}_{\mathcal{M}}(n)$ to be the edge-wise expected network. Since each element in the block stochastic matrix $G_{\mathcal{M}}(n)$ is a Bernoulli random variable, therefore the expected weight of an edge from one consumer in community i to one in community j is $\mathcal{M}_{ij}$.

5.2 The Monopolist's Pricing Problem

We are interested in how to maximize the expected profit for the monopolist, given that the monopolist knows the network is randomly generated by a community model with known community structure $\mathcal{M}$. The monopolist problem is now

$$max_{\mathbf{p}} \frac{1}{2}(\mathbf{p} - c\mathbf{1})^T \mathbb{E}[(I - \frac{2\rho}{\|G_{\mathcal{M}}(n) + G_{\mathcal{M}}(n)^T\|} G_{\mathcal{M}}(n))^{-1} |\mathcal{M}](a\mathbf{1} - \mathbf{p}).$$

This is is a specific version of the monopolist problem under general information; but now we have a more specific form of what information the monopolist has and this information shapes the distribution of networks they will encounter. Given this community information, the monopolist has no way to distinguish two consumers in the community because they are statistically identical. Therefore, there isn't a reason why one should be charged more or less than another from the monopolist's point of view. This implies that monopolist is pricing in such a way that every agent in the same community is charged the same price. Formally, the monopolist's optimal pricing is characterized in the following proposition.

Proposition 5.1. *In a network generated by block stochastic model with community structure $\mathcal{M}$ the optimal for the monopolist to charge is*

$$\boldsymbol{p}^* = \mathcal{R}^T \left(\mathcal{R}(B + B^T)\mathcal{R}^T\right)^{-1} \mathcal{R}(aB + cB^T)\boldsymbol{1}$$

Where $B = \mathbb{E}[(I - \frac{2\rho}{\|G_{\mathcal{M}}(n)+G_{\mathcal{M}}(n)^T\|} G_{\mathcal{M}}(n))^{-1}|$ Given the monopolist knows the structure$\mathcal{M}$ and $\mathcal{R}$ is the partition matrix describing the community identity of each consumer.

Intuitively, the consumers in each block are already similar, so charging agents in the same community a homogeneous price should turn out to be a small loss compared to making pricing decisions based on the full network information. In Section 6, we will prove this conjecture rigorously. To evaluate, we will need a framework to compare the value of different pricing strategies, which is introduced in the next section.

5.3 Regret and Fractional Regret

To evaluate the value of information needed for pricing, we introduce the notion of regret and fractional regret as used in **?**. The regret of a particular price vector $\mathbf{q}$ is how much profit is lost using this price compared to the optimal profit. We define the regret R as

$$R(\mathbf{q}) = \pi(\mathbf{p}^*) - \pi(\mathbf{q}),$$

where $\mathbf{p}^*$ is the optimal profit with no constraints or uncertainty. We also define the fractional regret to be the fraction of profit we miss out and it's given by

$$R_F(\mathbf{q}) = 1 - \frac{\pi(\mathbf{q})}{\pi(\mathbf{p}^*)}.$$

We will focus on how the regret and fractional regret change as a function of the network size.

6 Main Results

In this section, we present our main results about the value of network information and community information for stochastic block models. We will study the pricing problem of a monopolist who is given the information that the network is a block stochastic model generated by community structure $\mathcal{M} = p(n)\mathcal{M}_0$ where $\mathcal{M}_0$ is a matrix whose elements define the probability of edge between communities. We allow probabilities that we actually consider to change with the size of the network n. This is governed by the function $p(n)$. This means that all connection probabilities scale with n at the same rate. We are particularly interested in how much profit the monopolist can generate under different levels of information asymptotically as the network size n grows.

We start our analysis by showing that for stochastic block models with density $p(n) = \Omega\left(\frac{\log(n)}{n}\right)$, the value of network information is significant. We focus on this range of $p(n) = \Omega\left(\frac{\log(n)}{n}\right)$ because as shown in **?**, **?**, the network effect is weak in networks that are sparser and therefore network information is not valuable. Specifically, our results show that using the uniform pricing, which does not depend on any network information, leads to a constant and non-zero amount of regret asymptotically.

Theorem 6.1. *For the sequence of networks $G_{\mathcal{M}}(n)$ where $\mathcal{M} = p(n)\mathcal{M}_0$ and $p(n) = \Omega\left(\frac{\log(n)}{n}\right)$, as $n \to \infty$, the expected profit from the uniform pricing policy (i.e., charge everyone the same price $\frac{a+c}{2}$) approaches*

$$n\frac{(a-c)^2}{8}\boldsymbol{1}^T\left(I - 2\frac{\rho}{\|\mathcal{M}_0 + \mathcal{M}_0^T\|}\mathcal{M}_0\right)^{-1}\boldsymbol{1}.$$

Moreover, the fractional regret from uniform pricing converges to a positive constant β almost surely, where

$$\beta = 1 - \frac{\boldsymbol{1}^T\left(I - \frac{\rho}{\|\mathcal{M}_0+\mathcal{M}_0^T\|}(2\mathcal{M}_0)\right)^{-1}\boldsymbol{1}}{\boldsymbol{1}^T\left(I - \frac{\rho}{\|\mathcal{M}_0+\mathcal{M}_0^T\|}(\mathcal{M}_0 + \mathcal{M}_0^T)\right)^{-1}\boldsymbol{1}}. \tag{4}$$

Theorem 6.1 characterizes a class of random networks where the value of network information is substantial, and the monopolist's loss will grow linearly in the size of the network if he completely ignores the network information and charges everyone in the network the same price.

Though the above analysis emphasizes the value of network information, our next result shows that it's not necessary for the monopolist to invest in obtaining the full network information, since knowing the community information is sufficient to capture most of the profit.

Theorem 6.2. *For the sequence of networks $G_{\mathcal{M}}(n)$ where $\mathcal{M} = p(n)\mathcal{M}_0$ and $p(n) = \Omega\left(\frac{\log(n)}{n}\right)$, as $n \to \infty$, the expected profit under optimal pricing based on the networks' community structure $\mathcal{M}$ approaches*

$$n\frac{(a-c)^2}{8}\boldsymbol{1}^T\left(I - 2\frac{\rho}{\|\mathcal{M}_0 + \mathcal{M}_0^T\|}(\mathcal{M}_0 + \mathcal{M}_0^T)\right)^{-1}\boldsymbol{1}.$$

Moreover, the corresponding fractional regret converges to 0 almost surely.

Theorem 6.2 shows that for the sequence of networks generated from stochastic block model with density $p(n) = \Omega\left(\frac{\log(n)}{n}\right)$, the optimal pricing given community structure is good enough to guarantee nearly optimal profit when the network size is large enough. In other words, the additional value from learning network information other than the community structure is negligible asymptotically.

Alternatively, the results in both Theorem 6.1 and Theorem 6.2 provide a framework to evaluate the value of community information in large networks where community structure persists. If the monopolist's cost to learn the information about community structure is more than $\beta\pi(\mathbf{p}^*)$, where β is defined in equation (4), then it's not worth spending the cost to obtain the community structure information, and charging the uniform pricing (which doesn't depend on any network information) will be a better option for the monopolist.

The full proofs of Theorem 6.1 and Theorem 6.2 can be found in the Appendix.

7 Proof ideas of main results

In this section, we briefly introduce the techniques and ideas we use to prove the main results. We begin by analyzing the average graph which relies on algebraic manipulations that are possible with Kronecker product. Next, we show that the degrees of communities randomly generated stochastic block model networks will be concentrated tightly around the average degree. This leads to an expression for the spectral norm $G_{\mathcal{M}}(n) + G_{\mathcal{M}}(n)^T$ in terms of the norm of $\|\mathcal{M} + \mathcal{M}^T\|$ and the network size. Together this expression for the spectral norm, the analysis of the average graph, and the concentration of degrees allows us to prove theorems 6.1 and 6.2.

7.1 Analysis of the average network

We start by analyzing the average network of a stochastic block matrix. It is deterministic and completely known to the monopolist with the value of G_{ij} being the probability that i has an edge to j over any sampled network. The terms we get here will later be compared to the case where G is a random block stochastic matrix.

We can write $\bar{G}_{\mathcal{M}}(n) = \mathcal{M} \otimes \mathbf{1}_{n,n}$ where $\mathbf{1}_{n,n}$ is the $n \times n$ matrix of all ones and $\otimes$ is the Kronecker product. The facts we need in this case are an expression for the norm $\|\bar{G}_{\mathcal{M}}(n) + \bar{G}^T_{\mathcal{M}}(n)\|$, $\mathbf{1}^T(I - 2\frac{\rho}{\|G+G^T\|}G)^{-1}\mathbf{1}, \mathbf{1}^T(I - \frac{\rho}{\|G+G^T\|}(G + G^T))^{-1}\mathbf{1}$. The second term is needed to assess the profit from uniform price and the third is needed for the profit from the optimal pricing strategy when there are not constraints.

Lemma 7.1. $\|\bar{G}_{\mathcal{M}}(n) + \bar{G}_{\mathcal{M}}(n)^T\| = n * \|\mathcal{M} + \mathcal{M}^T\|$

Proof.

$$
\begin{aligned}
\|\bar{G}_{\mathcal{M}}(n) + \bar{G}_{\mathcal{M}}(n)^T\| &= \|(\mathcal{M} + \mathcal{M}^T) \otimes \mathbf{1}_{n,n}\| \\
&= \|\mathcal{M} + \mathcal{M}^T\| * \|\mathbf{1}_{n,n}\| \text{ Known property of Kronecker product} \\
&= \|\mathcal{M} + \mathcal{M}^T\| * n
\end{aligned}
$$

□

For $\mathbf{1}^T\left(2\overline{G}_{\mathcal{M}}(n)\right)^k\mathbf{1}$ and $\mathbf{1}^T\left(\overline{G}_{\mathcal{M}}(n) + \overline{G}^T_{\mathcal{M}}(n)\right)^k\mathbf{1}$ we will also use the observation that the ones vector of length nm can be written as $\mathbf{1}_m \otimes \mathbf{1}_n$

Lemma 7.2. $\boldsymbol{1}^T\left(2\overline{G}_{\mathcal{M}}(n)\right)^k\boldsymbol{1} = n^{k+1} * \boldsymbol{1}^T(2\mathcal{M})^k\boldsymbol{1}$

Lemma 7.3. $\boldsymbol{1}^T\left(\overline{G}_{\mathcal{M}}(n) + \overline{G}_{\mathcal{M}}(n)^T\right)^k\boldsymbol{1} = n^{k+1}\boldsymbol{1}^T(\mathcal{M} + \mathcal{M}^T)\boldsymbol{1}$

These proofs both make use of properties of the Kronecker product and detailed explanation can be found in appendix 11.3

These facts are used to show that in the expected network the uniform profit and optimal profit experienced by the monopolist both scale linearly with the network size with slope dependent on the community structure. Proofs are found in 11.4

Proposition 7.1. *Applied to the expected network of a block stochastic model network defined by community structure $\mathcal{M}$ and community size n will make a profit of* $n\frac{(a-c)^2}{8}\boldsymbol{1}^T\left(I - 2\frac{\rho}{\|\mathcal{M}+\mathcal{M}^T\|}\mathcal{M}\right)^{-1}\boldsymbol{1}$

For the average network we first compute the optimal price vector and show that it is the same price with and without the group constraints.

Proposition 7.2. *Under the constraint that each block of a block stochastic model network with community structure $\mathcal{M}$ is charged the same the optimal price $\boldsymbol{p}^*_{\mathcal{M}}$ to charge each block is*

$$\boldsymbol{p}^*_{\mathcal{M}} = \left((I - \frac{2\rho}{\|\mathcal{M}+\mathcal{M}\|}M)^{-1} + (I - \frac{2\rho}{\|\mathcal{M}+\mathcal{M}\|}M^T)^{-1}\right)^{-1}\left(a(I - \frac{2\rho}{\|\mathcal{M}+\mathcal{M}\|}M)^{-1} + c(I - \frac{2\rho}{\|\mathcal{M}+\mathcal{M}\|}M)^{-1}\right)\boldsymbol{1}_m$$

Notice that it must be the case that for this particular family of networks the unconstrained price we charge consumers must match the above price. This is because the price vector depends on a constant term and the bonanich centrality of each consumer. But for the expected network, consumers in the same block have precisely the identical positions in the network so they must have the same centrality. Therefore in this case we must have that everyone in a community is charged the same price which implies that the price they are charged is given by the above price.

For the average network the optimal profit and profit from community information match and both have a profit of

Proposition 7.3. *Under the constraint that each block of a block stochastic model network with community structure $\mathcal{M}$ size n a monopolist charging the optimal profit will recieve a profit of*

$$n\frac{(a-c)^2}{8}\boldsymbol{1}^T\left(I-\frac{\rho}{\|\mathcal{M}+\mathcal{M}^T\|}(\mathcal{M}+\mathcal{M}^T)\right)^{-1}\boldsymbol{1}$$

7.2 Sampled network and Expected network

In this section we handle the case where the network is generated randomly according to stochastic block model with the community structure $\mathcal{M}$. The goal of the monopolist is maximize its expected profit given that it knows $\mathcal{M}$. Let $\mathcal{R}\in\mathbb{R}^{m\times mn}$ be the membership matrix of the communities where m is the number of communities and n is the size of each community. From proposition 4.1, the price they should charge each community is:

$$\mathbf{p}=\mathcal{R}^T\left(\frac{a+c}{2}\mathbf{1}+\frac{a-c}{2}(B_{\mathcal{M}}+B_{\mathcal{M}})^{-1}(B_{\mathcal{M}}-B_{\mathcal{M}}^T)\mathbf{1}\right)$$

Where $B_{\mathcal{M}}=(I-2\frac{\rho}{\|\mathcal{M}+\mathcal{M}^T\|}\mathcal{M})^{-1}$. We show that as the network size becomes large the degrees of all consumers in a community concentrate around the average degree of the consumers in the community.

Proposition 7.4. *Let $c(n)$ be a function such that $\lim_{n\to\infty}c(n)=\infty$ and $c(n)\log(n)<n$ for all n. Let $\delta(n)$ be another function where $\sqrt{\frac{12}{c(n)}}<\delta(n)<1$ and $\delta(n)=\Theta\sqrt{\frac{1}{c(n)}}$. If the scaling of network is $p(n)=\frac{c(n)\log(n)}{n}$ then asymptotically almost surely $G(n)\boldsymbol{1}\in[(1-\delta(n))\overline{G}\boldsymbol{1},(1+\delta(n))\overline{G}\boldsymbol{1}]$*

The important of this theorem is what we utilize to as a core component of following proofs where this close concentration of the degrees is required. The proof follows from applying Chernoff bounds and a detailed proof can be found in appendix 11.4.

7.3 Norm of block stochastic matrix

At this point the last piece of the correspondence between the facts shown about the expected network and the randomly sampled block stochastic matrix is to evaluate the norm of $\|G+G^T\|$. The next result shows that this will asymptotically grow linearly.

Proposition 7.5. *As $\lim_{n\to\infty}\|G+G^T\|$ of a block stochastic with community structure $\mathcal{M}$ is almost surely*

$$n\|\mathcal{M}+\mathcal{M}^T\| \tag{5}$$

The norm of $G+G^T$ for large networks is bounded between largest degree of $G+G^T$ and the average degree. In the proof in the appendix 11.4 we show that these two quantities converge to the same quantity within a community.

8 Uncertainty about the Community Structure

Throughout the previous sections we assumed that the Monopolist knew for certain what community structure the network would have. In this section we introduce an extension where the monopolist instead knows that the network will be generated by one of a finite number of possible structures and the probability of each structure. We will carry out our analysis in terms of $(B_{\mathcal{M}})_i := (I - \frac{2\rho}{\|\mathcal{M}_i + \mathcal{M}_i^T\|}\mathcal{M}_i)^{-1}$ where $\mathcal{M}_i$ is each of the community structures that are possible. We assume that the ith community is picked with probability α_i.

If the monopolist knew that the ith community structure would be used then they would price as

$$\mathbf{p}^{(i)} = \frac{a+c}{2}\mathbf{1} + \frac{a-c}{2}((B_{\mathcal{M}})_i + (B_{\mathcal{M}})_i^T)^{-1}((B_{\mathcal{M}})_i - (B_{\mathcal{M}})_i^T)\mathbf{1}$$

And their expected profit would be of the form

$$\frac{n(a-c)^2}{8}\sum_i \alpha_i(\mathbf{1}+((B_{\mathcal{M}})_i+(B_{\mathcal{M}})_i^T)^{-1}((B_{\mathcal{M}})_i-(B_{\mathcal{M}})_i^T)\mathbf{1})^T(B_{\mathcal{M}})_i(\mathbf{1}-((B_{\mathcal{M}})_i+(B_{\mathcal{M}})_i^T)^{-1}((B_{\mathcal{M}})_i-(B_{\mathcal{M}})_i^T)\mathbf{1})$$

But when the monopolist is uncertain about what structure will be used heir best option is to act on the network on average by considering $(B_{\mathcal{M}}) = \sum_i \alpha_i (B_{\mathcal{M}})_i$ Where they will charge a price per community as:

$$\mathbf{p} = \frac{a+c}{2}\mathbf{1} + \frac{a-c}{2}((B_{\mathcal{M}}) + (B_{\mathcal{M}})^T)^{-1}((B_{\mathcal{M}}) - (B_{\mathcal{M}})^T)\mathbf{1}$$

Given that the ith structure is actually the one chosen the profit they earn is

$$\frac{n(a-c)^2}{8}\sum_i \alpha_i\left(\mathbf{1} + ((B_{\mathcal{M}}) + (B_{\mathcal{M}})^T)^{-1}((B_{\mathcal{M}}) - (B_{\mathcal{M}})^T)\mathbf{1}\right)^T (B_{\mathcal{M}})_i\left(\mathbf{1} - ((B_{\mathcal{M}}) + (B_{\mathcal{M}})^T)^{-1}((B_{\mathcal{M}}) - (B_{\mathcal{M}})^T)\mathbf{1}\right)$$

The natural question is how much regret does this pricing strategy induce. The answer is it grows linearly with slope determined by $\alpha_i, (B_{\mathcal{M}})_i, \mathbf{u}, and \Delta\mathbf{u}^{(i)}$.

Proposition 8.1. *The regret of using the expected structure $\mathcal{M}$ instead of the actual structures being used is*

$$\frac{n(a-c)^2}{8}\sum_i \alpha_i\left(-(\boldsymbol{1}+\boldsymbol{u})^T(B_{\mathcal{M}})_i\Delta\boldsymbol{u}^{(i)} + \Delta\boldsymbol{u}^{(i)}(B_{\mathcal{M}})_i(\boldsymbol{1} - (\boldsymbol{u}+\Delta\boldsymbol{u}^{(i)})\right)$$

Proof. Let $\mathbf{u} = \left((B_{\mathcal{M}}) + (B_{\mathcal{M}})^T)^{-1}((B_{\mathcal{M}}) - (B_{\mathcal{M}})^T\right)\mathbf{1}, \mathbf{u}^{(i)} = \left((B_{\mathcal{M}})_i + (B_{\mathcal{M}})_i^T)^{-1}((B_{\mathcal{M}})_i - (B_{\mathcal{M}})_i^T\right)\mathbf{1}$, and $\Delta\mathbf{u}^{(i)} = \mathbf{u}^{(i)} - \mathbf{u}$. With this notation, the regret $\Delta\Pi$ is

$$\begin{aligned}
\Delta\Pi &= \frac{n(a-c)^2}{8}\sum_i \alpha_i\left((\mathbf{1}+\mathbf{u}^{(i)})^T(B_{\mathcal{M}})_i(\mathbf{1}-\mathbf{u}^{(i)}) - (\mathbf{1}+\mathbf{u})^T(B_{\mathcal{M}})_i(\mathbf{1}-\mathbf{u})\right) \\
&= \frac{n(a-c)^2}{8}\sum_i \alpha_i\left((\mathbf{1}+\mathbf{u}+\Delta\mathbf{u}^{(i)})^T(B_{\mathcal{M}})_i(\mathbf{1}-(\mathbf{u}+\Delta\mathbf{u}^{(i)})) - (\mathbf{1}+\mathbf{u})^T(B_{\mathcal{M}})_i(\mathbf{1}-\mathbf{u})\right) \\
&= \frac{n(a-c)^2}{8}\sum_i \alpha_i\left(-(\mathbf{1}+\mathbf{u})^T(B_{\mathcal{M}})_i\Delta\mathbf{u}^{(i)} + \Delta\mathbf{u}^{(i)}(B_{\mathcal{M}})_i(\mathbf{1}-(\mathbf{u}+\Delta\mathbf{u}^{(i)})\right)
\end{aligned}$$

The whole term inside the summation is constant with respect to changes in the network size so the regret grows linearly.

□

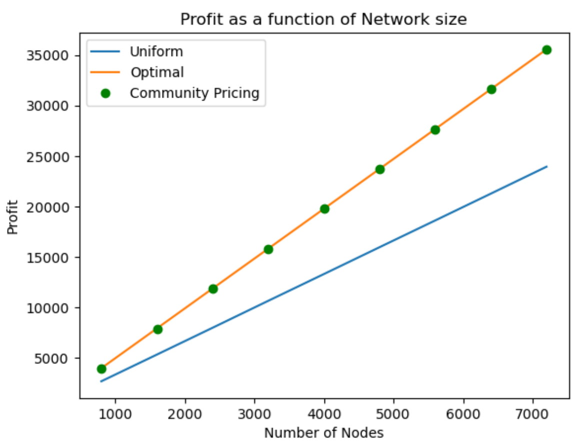
This result suggests that any uncertainty in the community structure creates regret that grows with the decrease in the quality of information for the monopolist and grows linearly in the size of the network. Thus high quality community information is important for the monopolist.

9 Experiments

9.1 Random Trials

For the first trial we set the number of communities to be 4 and generate randomly a 4x4 matrix specifying the community structure. For reproducibility we use a fixed seed and the code for this simulation was supplied as an auxiliary resource. This serves as the community structure. Keep this community structure fixed while varying the number of consumers in each community to build block stochastic matrices. For each network we calculate the profit the monopolist gets from using the uniform price, the optimal price with no constraints, and the optimal price with a group constraint that each consumer in the same block stochastic community is charged the same price. We should the constants the same throughout the experiment with $a = 6, c = 4, \rho = .9$. We report how to profit grows as a function of network size. We also report how much regret the monopolist has using the uniform price and the price from group homogeneous pricing instead of being able to use the true optimal price.

The results of these experiments match up extremely well with what theory suggests, The optimal profit, uniform profit, and profit from using only the community structure all scale linearly but the profit optimal profit and the profit from community information scale at essentially the same rate. We also see that the fractional regret for using both the uniform price and the community level price are close to constant respective of the network size. But the difference is that the uniform price experiences approximately a regret of 0.32 while the community pricing suffers almost no regret.



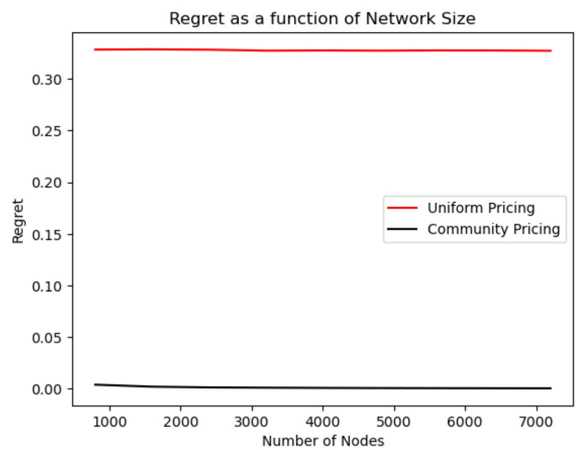


Figure 2: a) Profit as a function of the network size. Here there are 4 equally sized communities. In blue is the profit obtained from using the uniform price which is equivalent to using no information of the network. The orange line represents using full network information and the green dots represent using only community level information. b) Here we display the fractional regret $1 - \frac{\text{Profit found}}{\text{Optimal Profit}}$ from using the uniform price at price from the community structure. The uniform price is around 32% while the community pricing is nearly 0.

9.2 Gnutella data

Next, we demonstrate the validity of this approach on the Gnutella dataset **? ?**. This data set is series of snapshots of the directed network topology of the peer to peer file sharing platform Gnutella. In this network nodes represent computational units and an edge represents a communication linked existing from one person to the other. This network is directed so it is not needed that two people can mutually communicate to each other. Here we look at one snap from this data. This network consists of 6301 nodes and 20,777 vertices.

We set $\rho = 0.9, a = 6, c = 4$. The optimal profit is found to be \$6096.8 and the uniform profit \$4085.3. This is fractional regret of 33% In this network there are no predefined communities like what we had in the stochastic block model so we choose communities according to a heuristic. We partition based on the price vector we would charge in the unconstrained case with the idea the gentlest should

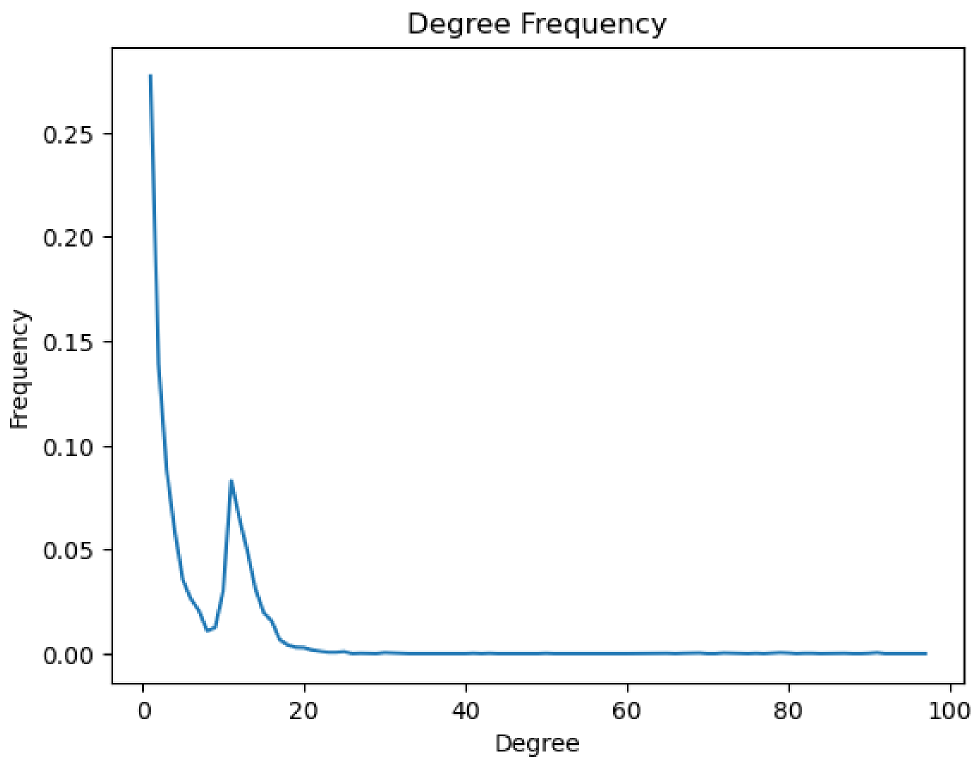


Figure 3: Degree frequency in Gnutella data. Here are using the in degree plus out degree of each node.

Case	Profit	Fractional Regret
Optimal	6096.8	0
Uniform	4085.3	33.0
Block Structure	5443	10.7

Table 1: Differences in performance between using the optimal price vector, uniform price vector, and the block pricing where the blocks have been defined by percentile ranges of the optimal price.

be type of constraints are force prices to change the least from the unconstrained case. To accomplish this we put nodes whose prices are in the same percentile range in the same community. If we are allowed to charge at most m different prices the question is what should the bounds of each percentile range be. Rather than splitting uniformly where each community gets the same number of nodes we found that it is better to split unevenly to allow the highest magnitude nodes to be split as more finely. For this experiment use 16 bins where the ith bins takes nodes whose price is in the ith range defined by $[[0,1],(1,2],(2,3],(3,4],(4,5],(5,10],(10,30],(30,50],(50,70],(70,90],(90,95],(95,96],(96,97],(97,98]$, $(98-99],(99,100]]$. For example, all the nodes optimally would charge strictly more than 50th percentile price but less than or equal to 70th percentile price are in one community. Computing the optimal profit with this constraint, we find that profit obtained is \$5443 which has a fractional regret of only 10.7%.

10 Conclusion and Future work

In this paper, we studied the pricing problem of a monopolist under partial network information. We first answered how to price under group homogeneity constraint. Then we studied pricing under general uncertainty about the network. We merged these two ideas to motivate the study stochastic block models. We showed that knowing only the community structure of the matrix resulted in prices as though we had the constraint that all the agents in a community must be charged the same price.

We proceeded to show that this level of information is valuable to the monopolist. For large block stochastic matrices they make nearly the same amount of profit as they would if they knew the full network information. But also that this information was needed to get this level of performance because under pricing with no prior information of the network the reduction in profit is significant. Such a result is valuable because it shows there are families of random networks where a limited amount of information could be useful. Since community information is often easy to obtain and is less noisy, it suggests that the monopolist should use this information while ignoring other network information. There are several possible directions of future work. First, it will be interesting to find other classes of networks that admit significant value of network information and what set of information about network is valuable for those networks. Second, it will be interesting to study how the profit changes with the scale of the community information. Third, if the community information is not available then it is often inferred from clustering or side information in the network. It will be useful to study how the fineness of clustering affect the profits. Finally, it is not clear is partitioning the network into communities is the only way or are their other ways of partitioning that could generate near-optimal profits.

11 Appendix

11.1 Proofs for group homogeneous section

Proof of Theorem 4.1

Proof. We can separate the price p into it's block parts i.e. $p = \sum_{k=1}^{b} p_k e_k$ where

$$(e_k)_i = \begin{cases} 1 & i \in \text{Block k} \\ 0 & Else \end{cases}$$

That is it selects the consumers belonging to the kth community.

$$max_{p_1,p_2,\cdots,p_b} \frac{1}{2}([\sum_{k=1}^{b} p_k e_k] - c\mathbf{1})^T (I - 2\frac{2\rho}{\|G+G^T\|}G)^{-1}(a\mathbf{1} - \sum_{k=1}^{b} p_k e_k) \tag{6}$$

Let $B = \frac{1}{2}(I - 2\frac{\rho}{\|G+G^T\|}G)^{-1}$, $q = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_b \end{bmatrix}$

This turns the objective into

$$\max_q (\mathcal{R}^T q - c\mathbf{1})^T B(a\mathbf{1} - \mathcal{R}^T q)$$

Taking the gradient with respect to q and equating to 0 we find that the stationary point of the objective is

$$\begin{aligned} a\mathcal{R}B\mathbf{1} + c\mathcal{R}B\mathbf{1} &= \mathcal{R}(B+B^T)\mathcal{R}^T \mathbf{q} \\ q &= \left(\mathcal{R}(B+B^T)\mathcal{R}^T\right)^{-1}\left(a\mathcal{R}B + c\mathcal{R}B^T\right)\mathbf{1} \end{aligned}$$

In addition, the Hessian is $-\left(\mathcal{R}(B+B^T)\mathcal{R}^T\right)$. B has already been established to be positive definite and so the hessian is negative definite and therefore finding the stationary point of this function is sufficient to find the maximum of the profit. □

Proof of Corollary 4.1

Proof. We take the restriction $\mathcal{R}$ to be a vector of size $1 \times n$ and calculate as follows:

$$\begin{aligned} \mathbf{p}_R^* &= (\mathcal{R}(B+B^T)\mathcal{R}^T)^{-1}\mathcal{R}(aB + cB^T)\mathbf{1} \\ &= (\mathbf{1}^T(B+B^T)\mathbf{1})^{-1}\mathbf{1}^T(aB + cB^T)\mathbf{1} \\ &= (2\mathbf{1}^T B\mathbf{1})^{-1}\mathbf{1}^T(a+c)B)\mathbf{1} \\ &= \frac{a+c}{2} \end{aligned}$$

□

Proof of Corollary 4.2

Proof. We start by using the optimal price is the first form by take $\mathcal{R} = I$.

$$\begin{aligned} \mathbf{p}_R^* &= (I(B+B^T)I)^{-1}I(aB + cB^T)\mathbf{1} \\ &= (B+B^T)^{-1}(aB + cB^T)\mathbf{1} \end{aligned}$$

We show the equivalence between the first and the second from by taking a difference between the first term and $\frac{a-c}{2}(B+B^T)(B-B^T)\mathbf{1}$ and showing that the remainder is $\frac{a+c}{2}\mathbf{1}$

$$(B+B^T)^{-1}(aB+cB^T)\mathbf{1} - \frac{a-c}{2}(B+B^T)^{-1}(B-B^T)\mathbf{1} \tag{7}$$
$$= \frac{a+c}{2}(B+B^T)^{-1}(B+B^T)\mathbf{1} \tag{8}$$
$$= \frac{a+c}{2}\mathbf{1} \tag{9}$$

Which is the remaining term of proposed price.

Finally, we show the equivalence between the second and third forms. Notice that it will suffice to show that $(B+B^T)^{-1}(B-B^T)\mathbf{1} = \frac{\rho}{\|G+G^T\|}(G-G^T)(I-\frac{\rho}{\|G+G^T\|}(G+G^T))^{-1}\mathbf{1}$ this is simply ignoring constants. This proof depends on a simplified form of Woodbury matrix identity which states that for invertible matrices A and C

$$(A+C)^{-1} = A^{-1} - A^{-1}(A^{-1}+C^{-1})A^{-1} \tag{Woodbury Matrix Identity}$$

For our purposes we will apply this to the term $(B+B^T)^{-1}$ to get:

$$\begin{aligned}
(B+B^T)^{-1} &= B^{-1} - B^{-1}(B^{-1}+B^{-T})^{-1}B^{-1} \\
&= B^{-1} - B^{-1}(2I - 2\frac{\rho}{\|G+G^T\|}(G+G^T))^{-1}B^{-1} \\
&= B^{-1} - \frac{1}{2}B^{-1}(I - \frac{\rho}{\|G+G^T\|}(G+G^T))^{-1}B^{-1} \\
&= (I - \frac{1}{2}B^{-1}(I - \frac{\rho}{\|G+G^T\|}(G+G^T))^{-1})B^{-1}
\end{aligned}$$

And by symmetry of $(B+B^T)^{-1} = (B^T+B)^{-1}$ we also have that

$$(B+B^T)^{-1} = (I - \frac{1}{2}B^{-T}(I - \frac{\rho}{\|G+G^T\|}(G+G^T))^{-1})B^{-T}$$

Applying this to the profit formula we derived previously and ignoring the constant terms we get:

$$\begin{aligned}
(B+B^T)^{-1}(B-B^T)\mathbf{1} &= (B+B^T)^{-1}B\mathbf{1} - (B+B^T)^{-1}B^T\mathbf{1} \\
&= (I - \frac{1}{2}B^{-1}(I - \frac{\rho}{\|G+G^T\|}(G+G^T))^{-1})B^{-1}B\mathbf{1} \\
&- (I - \frac{1}{2}B^{-T}(I - \frac{\rho}{\|G+G^T\|}(G+G^T))^{-1})B^{-T}B^T\mathbf{1} \\
&= \frac{1}{2}(B^{-T} - B^{-1})(I - \frac{\rho}{\|G+G^T\|}(G+G^T))^{-1}\mathbf{1} \\
&= \frac{1}{2}((I - 2\frac{\rho}{\|G+G^T\|}G^T) - (I - 2\frac{\rho}{\|G+G^T\|}G))(I - \frac{\rho}{\|G+G^T\|}(G+G^T))^{-1}\mathbf{1} \\
&= \frac{1}{2}(2\frac{\rho}{\|G+G^T\|}(G-G^T))(I - \frac{\rho}{\|G+G^T\|}(G+G^T))^{-1}\mathbf{1} \\
&= \frac{\rho}{\|G+G^T\|}(G-G^T)(I - \frac{\rho}{\|G+G^T\|}(G+G^T))^{-1}\mathbf{1}
\end{aligned}$$

□

11.2 Proofs of partial information theorems

Proof of Proposition 3

Proof. First we show what the profit they experience for an arbitatary network G with $B = (I - 2\frac{\rho}{\|G+G^T\|}G)^{-1}$ and $B_{\mathcal{F}} = \mathbb{E}_{\mathcal{F}}[(I - 2\frac{\rho}{\|G+G^T\|}G)^{-1}|\mathcal{I}]$.

$$\begin{aligned}
\frac{1}{2}(\mathbf{p}^* - c\mathbf{1})^T B_N(a\mathbf{1} - \mathbf{p}^*) &= \frac{1}{2}(\frac{a+c}{2}\mathbf{1} + \frac{a-c}{2}(B_{\mathcal{F}} + B_{\mathcal{F}}^T)^{-1}(B_{\mathcal{F}} - B_{\mathcal{F}}^T)\mathbf{1} - c\mathbf{1})^T B(a\mathbf{1} - \mathbf{p}^*) \\
&= \frac{1}{2}\frac{a-c}{2}(\mathbf{1} + (B_{\mathcal{F}} + B_{\mathcal{F}}^T)^{-1}(B_{\mathcal{F}} - B_{\mathcal{F}}^T)\mathbf{1} - c\mathbf{1})^T B(a\mathbf{1} - \frac{a-c}{2}(B_{\mathcal{F}} + B_{\mathcal{F}}^T)^{-1}(B_{\mathcal{F}} - B_{\mathcal{F}}^T)\mathbf{1} - c\mathbf{1}) \\
&= \frac{(a-c)^2}{8}(\mathbf{1} + (B_{\mathcal{F}} + B_{\mathcal{F}}^T)^{-1}(B_{\mathcal{F}} - B_{\mathcal{F}}^T)\mathbf{1})^T B \\
&* (\mathbf{1} - (B_{\mathcal{F}} + B_{\mathcal{F}}^T)^{-1}(B_{\mathcal{F}} - B_{\mathcal{F}}^T)\mathbf{1})
\end{aligned}$$

The expected profit is simply the some over each of these terms weight by the probability that G is the true network. Let α_i be the probability that G_i is the true network. So

$$\begin{aligned}
\mathbb{E}_{\mathcal{F}}[\pi] &= \frac{(a-c)^2)}{8}\sum_i \alpha_i(\mathbf{1} + (B_{\mathcal{F}} + B_{\mathcal{F}}^T)^{-1}(B_{\mathcal{F}} - B_{\mathcal{F}}^T)\mathbf{1})^T B_i \\
&* (\mathbf{1} - (B_{\mathcal{F}} + B_M\mathcal{F}^T)^{-1}(B_M\mathcal{F} - B_{\mathcal{F}}^T)\mathbf{1}) \\
&= \frac{(a-c)^2)}{8}(\mathbf{1} + (B_{\mathcal{F}} + B_{\mathcal{F}}^T)^{-1}(B_{\mathcal{F}} - B_{\mathcal{F}}^T)\mathbf{1})^T(\sum_i \alpha_i B_i) \\
&* (\mathbf{1} - (B_{\mathcal{F}} + B_M\mathcal{F}^T)^{-1}(B_M\mathcal{F} - B_{\mathcal{F}}^T)\mathbf{1}) \\
&= \frac{(a-c)^2)}{8}(\mathbf{1} + (B_{\mathcal{F}} + B_{\mathcal{F}}^T)^{-1}(B_{\mathcal{F}} - B_{\mathcal{F}}^T)\mathbf{1})^T B_{\mathcal{F}} \\
&* (\mathbf{1} - (B_{\mathcal{F}} + B_{\mathcal{F}}^T)^{-1}(B_{\mathcal{F}} - B_{\mathcal{F}}^T)\mathbf{1})
\end{aligned}$$

.

□

11.3 Proofs for the expected network

Proof of Lemma 7.2

Proof.

$$\mathbf{1}^T(2\overline{G}_{\mathcal{M}}(n))^k 1 = 1^T(2(\mathcal{M} \otimes \mathbf{1}_{n,n})^k)\mathbf{1} \tag{10}$$
$$= \mathbf{1}^T 2^k(\mathcal{M}^k \otimes \mathbf{1}_{n,n}^k)\mathbf{1} \tag{11}$$
$$= \mathbf{1}^T 2^k(\mathcal{M}^k \otimes n^{k-1}\mathbf{1}_{n,n})\mathbf{1} \tag{12}$$
$$= (\mathbf{1}_m \otimes \mathbf{1}_n)^T 2^k(\mathcal{M}^k \otimes n^{k-1}\mathbf{1}_{n,n})(\mathbf{1}_m \otimes \mathbf{1}_n) \tag{13}$$
$$= \mathbf{1}_m^T(2\mathcal{M})^k\mathbf{1}_m \otimes n^{k-1}\mathbf{1}_n\mathbf{1}_{n,n}\mathbf{1}_n \tag{14}$$
$$= \mathbf{1}_m^T(2\mathcal{M})^k\mathbf{1}_m * n^{k+1} \tag{15}$$

□

Proof of Lemma 7.3

Proof.

$$
\begin{aligned}
\mathbf{1}^T(\overline{G}_{\mathcal{M}}(n)+\overline{G}_{\mathcal{M}}^T(n))^k\mathbf{1} &= \mathbf{1}^T((\mathcal{M}+\mathcal{M}^T)\otimes \mathbf{1}_{n,n})^k)\mathbf{1} \\
&= \mathbf{1}^T((\mathcal{M}+\mathcal{M}^T)^k\otimes \mathbf{1}_{n,n}^k)\mathbf{1} \\
&= \mathbf{1}^T((\mathcal{M}+\mathcal{M}^T)^k\otimes n^{k-1}\mathbf{1}_{n,n})\mathbf{1} \\
&= (\mathbf{1}_m\otimes \mathbf{1}_n)^T((\mathcal{M}+\mathcal{M}^T)^k\otimes n^{k-1}\mathbf{1}_{n,n})(\mathbf{1}_m\otimes \mathbf{1}_n) \\
&= \mathbf{1}_m^T(\mathcal{M}+\mathcal{M}^T)^k\mathbf{1}_m\otimes n^{k-1}\mathbf{1}_n\mathbf{1}_{n,n}\mathbf{1}_n \\
&= \mathbf{1}_m^T(\mathcal{M}+\mathcal{M}^T)^k\mathbf{1}_m * n^{k+1}
\end{aligned}
$$

□

Proof of Proposition 7.1

Proof. The profit is

$$
\begin{aligned}
\frac{(a-c)^2}{8}\mathbf{1}^T(I-2\frac{\rho}{\|G_{\mathcal{M}}(n)+G_{\mathcal{M}}^T(n)\|}G_{\mathcal{M}}(n))^{-1}\mathbf{1} &= \frac{(a-c)^2}{8}\mathbf{1}^T(\sum_{k=0}^{\infty}(2\frac{\rho}{\|G_{\mathcal{M}}(n)+G_{\mathcal{M}}^T(n)\|})^kG^k)\mathbf{1} \\
&= \frac{(a-c)^2}{8}(\sum_{k=0}^{\infty}(2\frac{\rho}{n\|\mathcal{M}+\mathcal{M}^T\|})^k\mathbf{1}^TG\mathbf{1}) \\
&= \frac{(a-c)^2}{8}(\sum_{k=0}^{\infty}(2\frac{\rho}{n\|\mathcal{M}+\mathcal{M}^T\|})^k\mathbf{1}^T(n^{k+1}\mathcal{M})\mathbf{1}) \\
&= n\frac{(a-c)^2}{8}(\sum_{k=0}^{\infty}(2\frac{\rho}{\|\mathcal{M}+\mathcal{M}^T\|})^k\mathbf{1}^T\mathcal{M}\mathbf{1}) \\
&= n\frac{(a-c)^2}{8}\mathbf{1}^T(I-\frac{2\rho}{\|\mathcal{M}+\mathcal{M}^T\|}\mathcal{M})^{-1}\mathbf{1}
\end{aligned}
$$

□

Proof of Proposition 7.2

Proof. We expand the price vector $(\mathcal{R}(B+B^T)\mathcal{R}^T)^{-1}\mathcal{R}(aB+cB)\mathbf{1}$ in separate terms as $(\mathcal{R}(B+B^T)\mathcal{R}^T)$ and $\mathcal{R}(aB+cB)\mathbf{1}$

$$
\begin{aligned}
&\mathcal{R}(B+B^T)\mathcal{R}^T \\
&= (I_m\otimes \mathbf{1}_n)^T\left(\sum_{k=0}^{\infty}(\frac{2\rho G_{\mathcal{M}}(n)}{\|G_{\mathcal{M}}(n)+G_{\mathcal{M}}^T(n)\|})^k+\sum_{k=0}^{\infty}(\frac{2\rho G_{\mathcal{M}}^T(n)}{\|G_{\mathcal{M}}(n)+G_{\mathcal{M}}^T(n)|})^k\right)(I_m\otimes \mathbf{1}_n) \\
&= (I_m\otimes \mathbf{1}_n)^T\left(\sum_{k=0}^{\infty}(\frac{2\rho}{\|\mathcal{M}+\mathcal{M}^T\|}\mathcal{M})^k\otimes(\frac{\mathbf{1}_{n,n}}{n})^k+\sum_{k=0}^{\infty}(\frac{2\rho\mathcal{M}^T}{\|\mathcal{M}+\mathcal{M}^T\|})^k\otimes(\frac{\mathbf{1}_{n,n}}{n})^k\right)(I_m\otimes \mathbf{1}_n) \\
&= \left(\sum_{k=0}^{\infty}(\frac{2\rho}{\|\mathcal{M}+\mathcal{M}^T\|}\mathcal{M})^k\otimes(\mathbf{1}_n^T(\frac{\mathbf{1}_{n,n}}{n})^k\mathbf{1}^T)+\sum_{k=0}^{\infty}(\frac{2\rho\mathcal{M}^T}{\|\mathcal{M}+\mathcal{M}^T\|})^k\right)\otimes(\mathbf{1}_n^T\frac{\mathbf{1}_{n,n}}{n})^k\mathbf{1}_n^T) \\
&= n*\left(\sum_{k=0}^{\infty}(\frac{2\rho}{\|\mathcal{M}+\mathcal{M}^T\|}\mathcal{M})^k+\sum_{k=0}^{\infty}(\frac{2\rho}{\|\mathcal{M}+\mathcal{M}^T\|}\mathcal{M}^T)^k\right) \\
&= n(B_{\mathcal{M}}+B_{\mathcal{M}}^T)
\end{aligned}
$$

We will let $B_{\mathcal{M}} = (I - \frac{2\rho}{\|\mathcal{M}+\mathcal{M}^T\|}^{-1}$ And

$$\begin{aligned}
&\mathcal{R}(aB + cB)\mathbf{1} \\
&= (I_m \otimes \mathbf{1}_n^T) \left(a(\sum_{k=0}^{\infty}(\frac{2\rho}{\|\mathcal{M}+\mathcal{M}^T\|}\mathcal{M})^k \otimes (\frac{\mathbf{1}_{n,n}}{n})^k) + c(\sum_{k=0}^{\infty}(\frac{2\rho}{\|\mathcal{M}+\mathcal{M}^T\|}\mathcal{M}^T)^k) \otimes (\frac{\mathbf{1}_{n,n}}{n})^k \right) (\mathbf{1}_m \otimes \mathbf{1}_n) \\
&= n(aB_{\mathcal{M}} + cB_{\mathcal{M}}^T)\mathbf{1}_m
\end{aligned}$$

Taken together we have that for the average network the price to charge is

$$\begin{aligned}
&(B + B^T)^{-1}(aB + cB^T)\mathbf{1} \\
&= \frac{1}{n}(B_{\mathcal{M}} + B_{\mathcal{M}}^T)^{-1} n(aB_{\mathcal{M}} + cB_{\mathcal{M}}^T)\mathbf{1}_m \\
&= (B_{\mathcal{M}} + B_{\mathcal{M}}^T)^{-1}(aB_{\mathcal{M}} + cB_{\mathcal{M}}^T)\mathbf{1}_m
\end{aligned}$$

□

Proof of Proposition 7.3

Proof. The optimal profit has the form

$$\begin{aligned}
\frac{(a-c)^2}{8}\mathbf{1}^T(I-\frac{\rho}{\|\overline{G}_P(n)+\overline{G}_P(n)^T\|}\overline{G}_P(n)+\overline{G}_P(n)^T)^{-1}\mathbf{1} &= \frac{(a-c)^2}{8}\sum_{k=0}^{\infty}(\frac{\rho}{\|\overline{G}_P(n)+\overline{G}_P(n)^T\|})^k\mathbf{1}^T(\overline{G}_P(n)+\overline{G}_P(n)^T)^k\mathbf{1} \\
&= n\frac{(a-c)^2}{8}\sum_{k=0}^{\infty}(\frac{\rho}{n\|\mathcal{M}+\mathcal{M}^T\|})^k\mathbf{1}^T(n*(\mathcal{M}+\mathcal{M}^T))^k\mathbf{1} \\
&= n\frac{(a-c)^2}{8}\sum_{k=0}^{\infty}(\frac{\rho}{\|\mathcal{M}+\mathcal{M}^T\|})^k\mathbf{1}^T((\mathcal{M}+\mathcal{M}^T))^k\mathbf{1} \\
&= \frac{n(a-c)^2}{8}\mathbf{1}^T(I-\frac{\rho}{\|\mathcal{M}+\mathcal{M}^T\|}(\mathcal{M}+\mathcal{M}^T))\mathbf{1}
\end{aligned}$$

□

11.4 Proofs concerning random block matrix

Proof of 7.4

Proof. Let X_i be a sequence of Bernoulli random variables with $Pr(X_i = 1) = p_i$. Let also $X = \sum X_i$ and suppose that $E[X] = \mu$. For any $\delta, 0 < \delta < 1$ the Chernoff bounds state that $Pr(X \geq (1+\delta)\mu) \leq exp(-\frac{\delta^2\mu}{2+\mu}) \leq exp(-\frac{\delta^2\mu}{3})$ and $Pr(X \leq (1-\delta)\mu) \leq exp(-\frac{\delta^2\mu}{2})$

and $Pr(X \geq (1-\delta)\mu) \leq exp(\frac{\delta^2\mu}{2})$. We can break the degree of a consumer down into the degree it has to each community. We will denote by $deg_v(i)$ to mean the degree of v to the ith community. A consumer's degree to one community is the sum of n i.i.d. Bernoulli random variables so by the above calculation

$$\begin{aligned}
Pr(deg_v(i)) \geq (1+\delta)c(n)\log(n) \leq exp\left(-\frac{\delta^2 c(n)\log(n)}{3}\right) \\
Pr(deg_v(i)) \leq (1-\delta)c(n)\log(n) \leq exp\left(-\frac{\delta^2 c(n)\log(n)}{2}\right)
\end{aligned}$$

The probability that any consumer in same community of v has degree less than $(1-\delta)c(n)\log(n)$ is less than $n*exp(-\frac{\delta^2 c(n)\log(n)}{3})$ by the union bound. Likewise the probability that any consumer in same community of v has degree more than $(1+\delta)c(n)\log(n)$ is less than $nexp(-\frac{\delta^2 c(n)\log(n)}{2})$ by

By our assumptions of δ and c(n) we have that

$$\begin{aligned}
n*exp(-\frac{\delta^2 c(n)\log(n)}{3}) &\leq n*exp(-\frac{1^2 n}{3}) \\
&= 0
\end{aligned}$$

and

$$\begin{aligned}
n*exp(-\frac{\delta^2 c(n)\log(n)}{2}) &\leq n*exp(-\frac{1^2 n}{2}) \\
&= 0
\end{aligned}$$

Which imply that

$$\begin{aligned}
\sum_{n=1}^{\infty} Pr(deg_v(i) \geq (1+\delta)c(n)\log(n)) &\leq \sum_{n=1}^{\infty} n*exp(-\frac{-\delta^2 c(n)\log(n)}{3}) \\
&\leq \sum_{n=1}^{\infty} n*exp(-\frac{n}{3}) \\
&\leq 0
\end{aligned}$$

and

$$\sum_{n=1}^{\infty} Pr(deg_v(i) \geq (1+\delta)c(n)\log(n)) \leq \sum_{n=1}^{\infty} n * exp(-\frac{-\delta^2 c(n)\log(n)}{3})$$
$$\leq \sum_{n=1}^{\infty} n * exp(-\frac{n}{3})$$
$$\leq 0$$

The fact the is finite by the Borel-Cantelli lemma tells us that almost surely consumers in the community of v will have degree between $[(1-\delta)c(n)\log(n), (1+\delta)c(n)\log(n)]$.

But if the degrees of to one community is so concentrated then so is the sum of a finite fixed number of communities. □

Proof of Proposition 7.5

Proof. We start by considering the vector of degrees in G that community i has but counting edges to community j. We call this vector $\mathbf{d}_{ij}$ and the degrees of community unrestricted $\mathbf{d}_i$. Clearly, $\mathbf{d}_i = \sum_{j=1}^{m} \mathbf{d}_{ij}$. A known property of the spectral norm of graphs is that is bounded between the average degree and the maximum degree. Proposition 7 says that $G(n)\mathbf{1} \in [(1-\delta(n)\overline{G}\mathbf{1}, (1+\delta(n)\overline{G}\mathbf{1}]$.

This tells us that

$$\mathbf{d}_{ij}^{G} \in [(1-\delta(n))\mathbf{d}_{ij}^{\overline{G}}, (1-\delta(n))\mathbf{d}_{ij}^{\overline{G}}]$$
$$[(1-\delta(n))\mathbf{d}_{ij}^{\overline{G}}, (1-\delta(n))\mathbf{d}_{ij}^{\overline{G}}] = [(1-\delta(n))np(n)\mathcal{M}_i j, (1-\delta(n))np(n)]$$

Likewise the maximum degree of $\mathbf{d}_{ij}$ can by expressed by

$$\max \mathbf{d}_{ij}^{G} \in [(1-\delta(n))\max \mathbf{d}_{ij}^{\overline{G}}, (1-\delta(n))\max \mathbf{d}_{ij}^{\overline{G}}]$$
$$[(1-\delta(n))\max \mathbf{d}_{ij}^{\overline{G}}, (1-\delta(n))\max \mathbf{d}_{ij}^{\overline{G}}] = [(1-\delta(n))np(n)\mathcal{M}_i j, (1-\delta(n))np(n)$$

So the maximum and average degree for $\mathbf{d}_{ij}$ are both in $[(1-\delta(n))p(n)n\mathcal{M}_{ij}, (1+\delta(n))np(n)\mathcal{M}_{ij}]$. But if we extend this to d_i the maximum and average degrees are now in $[\sum_{j=1}^{m}(1-\delta(n))p(n)n\mathcal{M}_{ij}, \sum_{j=1}^{m}(1+\delta(n))np(n)\mathcal{M}_{ij}]$. But $\delta(n) \to 0$ with n so this reduces to $\mathbf{d}_i$ converges to $\sum_{j=1}^{m} p(n)n\mathcal{M}_{ij}$. This means the average degree is then $\frac{p(n)}{mn}\sum_{j=1}^{m} nn\mathcal{M}_{ij} = p(n)n\text{Average degree in } \mathcal{M}$. And a maximum degree of $\max p(n)n \max_{j=1}^{m} \mathcal{M}_{ij} = p(n)n\text{Maximum degree of } \mathcal{M}$

□

Proof of Theorem 6.1

Let E_i be the indicator vector for the ith community in G and e_i be a corresponding vector of length m in the probability matrix $\mathcal{M}$. That is $(e_i)_j = \begin{cases} 1 & j = i \\ 0 & Otherwise \end{cases}$

For this section we will write $G(n)\mathbf{1} = \overline{G}\mathbf{1} + \mathbf{r}$ where $\mathbf{r}$ is defined to be the difference between the sampled $G(n)\mathbf{1}$ and $\overline{G}\mathbf{1}$. Under assumptions that the network is sufficiently dense we can establish tight bounds on the distribution of degrees we can see in G. We let $A_{ik} := \mathbf{1}^T(2\rho G)^k E_i$, $A_i := \sum_{k=1}^{\infty} A_{ik}$ and $A = \sum_{i=1}^{m} A_i$. Note that A is the term we must calculate to find the uniform profit except that it is undercounting by $\mathbf{1}^T I \mathbf{1}^T = n * m$.

Proof. We have the follow bounds based on the approximation of GE_i to $\bar{G}E_i$

$$\begin{aligned}
A_j &= \sum_{k=1}^{\infty} \mathbf{1}^T (2\frac{\rho}{\|G+G^T\|}G)^k E_j \\
&= \sum_{k=1}^{\infty} \mathbf{1}^T (2\frac{\rho}{\|G+G^T\|}G)^{k-1} (2\frac{\rho}{\|G+G^T\|}GE_j) \\
&\leq \sum_{k=1}^{\infty} \mathbf{1}^T (2\frac{\rho}{\|G+G^T\|}G)^{k-1} (2\frac{\rho}{\|G+G^T\|}(1+\delta(n)))\bar{G}E_j \\
&= \sum_{k=1}^{\infty} \mathbf{1}^T (2\frac{\rho}{\|G+G^T\|}G)^{k-1} (2\frac{\rho}{\|G+G^T\|}(1+\delta(n)))(\mathcal{M}\otimes \mathbf{1}_{n,n})(e_j \otimes \mathbf{1}_n) \\
&= \sum_{k=1}^{\infty} \mathbf{1}^T (2\frac{\rho}{\|G+G^T\|}G)^{k-1} (2\frac{\rho}{\|G+G^T\|}(1+\delta(n)))(\mathcal{M}e_j \otimes n\mathbf{1}_n) \\
&= \sum_{k=1}^{\infty}\sum_{i=1}^{m} \mathbf{1}^T (2\frac{\rho}{\|G+G^T\|}G)^{k-1} (2\frac{\rho}{\|G+G^T\|}(1+\delta(n)))\mathcal{M}_{ij} n E_i
\end{aligned}$$

The last step is justified because $\mathcal{M}e_j \otimes E_i$ is a vector where the first m elements are all $\mathcal{M}_{1,j}$, the next are $\mathcal{M}_{2,j}$ and so on. Continuing from here we switch the order of summation.

$$\begin{aligned}
&= \sum_{i=1}^{m}\sum_{k=1}^{\infty} \mathbf{1}^T (2\frac{\rho}{\|G+G^T\|}G)^{k-1} \underbrace{(\frac{\rho}{\|G+G^T\|}n(1+\delta(n)))}_{\Delta^+} 2\mathcal{M}_{ij}E_i \\
&= \Delta^+ \sum_{i=1}^{m} 2\mathcal{M}_{ij} \sum_{k=1}^{\infty} \mathbf{1}^T (2\frac{\rho}{\|G+G^T\|}G)^{k-1} E_i \\
&= \Delta^+ \sum_{i=1}^{m} 2\mathcal{M}_{ij} \sum_{k=0}^{\infty} \mathbf{1}^T (2\frac{\rho}{\|G+G^T\|}G)^{k} E_i \\
&= \Delta^+ \sum_{i=1}^{m} 2\mathcal{M}_{ij} (n + \sum_{k=1}^{\infty} \mathbf{1}^T (2\frac{\rho}{\|G+G^T\|}G)^{k} E_i) \\
&= \Delta^+ \sum_{i=1}^{m} 2\mathcal{M}_{ij}(n + A_i) \\
&= \Delta^+ \sum_{i=1}^{m} (n + A_i)2\mathcal{M}_{ij}
\end{aligned}$$

By replacing Δ^+ with $\Delta^- = \frac{\rho}{\|G+G^T\|}(1-\delta(n))$ we can a get corresponding lower bound of $\Delta^- \sum_{i=1}^{m} 2(n+A_i)\mathcal{M}_{ij} \leq A_i$.

Now we consider the case when $\lim_{n\to\infty} \frac{\rho}{\|G+G^T\|}n\delta(n) \to 0$

In such a case A_i is squeezed to $\frac{\rho}{\|G+G^T\|}n\sum_{i=1}^{m} 2(n+A_i)\mathcal{M}_{ij}$ and we have that

$$\begin{aligned}
A_j &= 2\frac{\rho}{\|G+G^T\|}n\sum_i \mathcal{M}_{ij}(n+A_i) \\
&= 2\frac{\rho}{\|G+G^T\|}n^2\mathcal{M}_j + 2\frac{\rho}{\|G+G^T\|}n\sum_i \mathcal{M}_{ij}A_i
\end{aligned}$$

Where $\mathcal{M}_j = \sum_i \mathcal{M}_{ij}$

This induces a system of equations where the jth equation is

$$(1-2\rho n P_{ji})A_j - \sum_{i\neq j} 2\rho n P_{ij} = 2rn^2 P_j$$

Then we have in matrix notation

$$(I-2\rho n P^T)\widehat{A} = 2\rho n^2 \widehat{P}$$

where $\hat{A} = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{bmatrix}$ and $\hat{P} = \begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_m \end{bmatrix}$.

Which results in: $\widehat{A} = 2\rho n^2 (I-2\rho n P^T)^{-1}\widehat{P}$

Making the replacement of the spectral norm of $G+G^T$ in the term and adding back in the missing nm term which we lost by not starting the sum at 0. we get

$$\begin{aligned}
A + nm &= n\sum_{k=0}^{\infty} \mathbf{1}^T (\frac{\rho}{\|G+G^T\|} n(2\mathcal{M}))^k \mathbf{1} \\
&= n\sum_{k=0}^{\infty} \mathbf{1}^T (2\mathcal{M})^k ((1+o(1))\frac{\rho}{\|\mathcal{M}+\mathcal{M}^T\|}\frac{n}{n})^k \mathbf{1} \\
&= \lim_{n\to\infty} n\sum_{k=0}^{\infty} \mathbf{1}^T (\frac{\rho}{\|\mathcal{M}+\mathcal{M}^T\|}(2\mathcal{M}))^k \mathbf{1} \\
&= n\mathbf{1}^T (I - \frac{\rho}{\|\mathcal{M}+\mathcal{M}^T\|}(2\mathcal{M}))^{-1}\mathbf{1}
\end{aligned}$$

□

Proof of theorem 6.2

Proof. Let E_i be the indicator vector for the ith community in G and e_i be a corresponding vector of length m in the probability matrix $\mathcal{M}$. That is $(e_i)_j = \begin{cases} 1 & j = i \\ 0 & Otherwise \end{cases}$

For this section we will write $G(n)\mathbf{1} = \overline{G}\mathbf{1} + \mathbf{r}$ where $\mathbf{r}$ is defined to be the difference between the sampled $G(n)\mathbf{1}$ and $\overline{G}\mathbf{1}$. Under assumptions that the network is sufficiently dense we can establish tight bounds on the distribution of degrees we can see in G. Let Let $B_{ik} := \frac{\rho}{\|G+G^T\|} 1^T (G+G^T)^k E_i, B_i = \sum_{k=1}^{\infty} \mathbf{1}^T \frac{\rho}{\|G+G^T\|}^k (G+G^T)^k E_i$ and $B = \sum_{i=1}^{m} B_i$.

$$
\begin{aligned}
B_j &= \sum_{k=1}^{\infty} \mathbf{1}^T (\frac{\rho}{\|G+G^T\|})^k (G+G^T)^k E_j \\
&= \sum_{k=1}^{\infty} \mathbf{1}^T (\frac{\rho}{\|G+G^T\|})^{k-1} (G+G^T)^{k-1} (\frac{\rho}{\|G+G^T\|}(G+G^T)E_j) \\
&\geq \sum_{k=1}^{\infty} \mathbf{1}^T \frac{\rho}{\|G+G^T\|}^{k-1} (G+G^T)^{k-1} (\frac{\rho}{\|G+G^T\|}(1-\delta(n)))\overline{G+G^T} E_j \\
&= \sum_{k=1}^{\infty} \mathbf{1}^T (\frac{\rho}{\|G+G^T\|}(G+G^T))^{k-1} (\frac{\rho}{\|G+G^T\|}(1-\delta(n)))((\mathcal{M}+\mathcal{M}^T)\otimes 1_{n,n})(e_j \otimes 1_n) \\
&= \sum_{k=1}^{\infty} \mathbf{1}^T (\frac{\rho}{\|G+G^T\|}(G+G^T))^{k-1} (\frac{\rho}{\|G+G^T\|}(1-\delta(n)))((\mathcal{M}+\mathcal{M}^T)e_j \otimes n1_n) \\
&= \sum_{k=1}^{\infty} \sum_{i=1}^{m} \mathbf{1}^T (\frac{\rho}{\|G+G^T\|}(G+G^T))^{k-1} (\frac{\rho}{\|G+G^T\|}(1-\delta(n)))(\mathcal{M}_{ij}+\mathcal{M}_{ji})nE_i \\
&= \sum_{i=1}^{m} \sum_{k=1}^{\infty} \mathbf{1}^T (\frac{\rho}{\|G+G^T\|}(G+G^T))^{k-1} \underbrace{(\frac{\rho}{\|G+G^T\|}n(1+\delta(n)))}_{\Delta^-}(\mathcal{M}_{ij}+\mathcal{M}_{ji}E_i \\
&= \Delta^- \sum_{i=1}^{m} (\mathcal{M}_{ij}+\mathcal{M}_{ji} \sum_{k=1}^{\infty} \mathbf{1}^T (\frac{\rho}{\|G+G^T\|}(G+G^T))^{k-1} E_i \\
&= \Delta^- \sum_{i=1}^{m} (\mathcal{M}_{ij}+\mathcal{M}_{ji} \sum_{k=0}^{\infty} \mathbf{1}^T (\frac{\rho}{\|G+G^T\|}(G+G^T))^{k} E_i \\
&= \Delta^- \sum_{i=1}^{m} (\mathcal{M}_{ij}+\mathcal{M}_{ji})(n + \sum_{k=1}^{\infty} \mathbf{1}^T (\frac{\rho}{\|G+G^T\|}(G+G^T))^{k} E_i) \\
&= \Delta^- \sum_{i=1}^{m} \mathcal{M}_{ij}+\mathcal{M}_{ji})(n+B_i) \\
&= \Delta^- \sum_{i=1}^{m} (n+B_i)(\mathcal{M}_{ij}+P_{ji})
\end{aligned}
$$

When we assume that $\delta(n) \to 0$ as $n \to \infty$ we get that $(I - \frac{\rho}{\|G+G^T\|}n(\mathcal{M}+\mathcal{M}^{\mathcal{T}}))\widehat{B} = \rho n^2(\widehat{\mathcal{M}} + \widehat{\mathcal{M}^{\mathcal{T}}})$ where $\widehat{\mathcal{M}}$.

Again in the same manner we proved the last theorem we have that:

$$
\begin{aligned}
B + nm &= \sum_{j=1}^{m} B_i + nm \\
&= 2rn^2 1^T (I - \frac{\rho}{\|G+G^T\|}n(\mathcal{M}+\mathcal{M}^T))(\widehat{\mathcal{M}} + \widehat{\mathcal{M}^T}) + nm \\
&= n1^T \sum_{k=0} (2\frac{\rho}{\|G+G^T\|}n)^{k+1}(\mathcal{M}+\mathcal{M}^T)(\widehat{\mathcal{M}} + \widehat{\mathcal{M}^T}) + nm \\
&= n1^T \sum_{k=1} (2\frac{\rho}{\|G+G^T\|}n)^{k+1}(\mathcal{M}+\mathcal{M}^T)^{k+1}\mathbf{1} + nm \\
&= n1^T \sum_{k=0} (2\frac{\rho}{\|G+G^T\|}n(\mathcal{M}+\mathcal{M}^T))^{k}\mathbf{1}
\end{aligned}
$$

Making the replacement of the spectral norm of $G + G^T$ in the term and adding back in the missing nm term which we lost by not starting the sum at 0. we get

$$\begin{aligned}
B + nm &= n \sum_{k=0}^{\infty} \mathbf{1}^T \left(\frac{\rho}{\|G + G^T\|} n(\mathcal{M} + \mathcal{M}^T)\right)^k \mathbf{1} \\
&= n \sum_{k=0}^{\infty} \mathbf{1}^T (\mathcal{M} + \mathcal{M}^T)^k \left((1 + o(1)) \frac{\rho}{\|\mathcal{M} + \mathcal{M}^T\|} \frac{n}{n}\right)^k \mathbf{1} \\
&= \lim_{n \to \infty} n \sum_{k=0}^{\infty} \mathbf{1}^T \left(\frac{\rho}{\|\mathcal{M} + \mathcal{M}^T\|} (\mathcal{M} + \mathcal{M}^T)\right)^k \mathbf{1} \\
&= n \mathbf{1}^T \left(I - \frac{\rho}{\|\mathcal{M} + \mathcal{M}^T\|} (\mathcal{M} + \mathcal{M}^T)\right)^{-1} \mathbf{1}
\end{aligned}$$

□