Keeping Up with the Winner!
Targeted Advertisement to Communities in Social Networks

Shailaja Mallick[*1]**, Vishwaraj Doshi**[*2]**, Do Young Eun**[1]

[1]North Carolina State University
[2]Data Science & Advanced Analytics IQVIA Inc.
mallick.shailaja@gmail.com, vishwaraj.doshi@iqvia, dyeun@ncsu.edu

Abstract

When a new product enters a market already dominated by an existing product, will it survive along with this dominant product? Most of the existing works have shown the coexistence of two competing products spreading/being adopted on overlaid graphs with same set of users. However, when it comes to the survival of a weaker product on the same graph, it has been established that the stronger one dominates the market and wipes out the other. This paper makes a step towards narrowing this gap so that a new/weaker product can also survive along with its competitor with a positive market share. Specifically, we identify a locally optimal set of users to induce a community that is targeted with advertisement by the product launching company under a given budget constraint. To this end, we model the system as competing Susceptible-Infected-Susceptible (SIS) epidemics and employ perturbation techniques to quantify and attain a positive market share in a cost-efficient manner. Our extensive simulation results with real-world graph dataset show that with our choice of target users, a new product can establish itself with positive market share, which otherwise would be dominated and eventually wiped out of the competitive market under the same budget constraint.

1 Introduction

Motivation: In this competitive world of rapid technological advancements, one of the key strategic decisions that a company launching a new product has to make is how to market itself (Cooper and Kleinschmidt 1987; Montoya-Weiss and Calantone 1994). It can carve out a niche by differentiating itself from its competitors, either by its product design or brand image (Fuchs and Diamantopoulos 2010; Jalkala and Keränen 2014), and highlighting them through marketing techniques such as advertisements. In such situations, advertisements serve as a medium of communication and play a key role for the firm to achieve this differentiation. There has been evidence that products with no mass advertisements have gained popularity over time as the companies have focused on targeting advertisements to niche communities (Dalgic and Leeuw 1994). The effectiveness of targeting a small portion of users (communities) who can further spread and influence their neighborhood network via
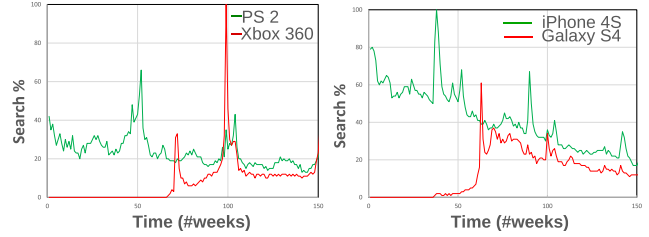
Figure 1: Real web-search percentage over a time period showing the introduction of a new product (Galaxy S4 & Xbox 360) alongside an existing product (iPhone 4S & PlayStation 2).

the word-of-mouth, rather than focusing on the whole population, has proven beneficial and also recognized by businesses in the market (Armstrong and Kotler 2003; Yang et al. 2006).

For example, Nintendo dominated the video game consoles market in the US for almost a decade until the introduction of the PlayStation (PS) (Gallagher and Park 2002). Over the next few years, PS was able to establish and eventually dominate the video game console market, forcing Nintendo to focus on hand-held consoles. Subsequently, Xbox initially entered a PS-dominated video game consoles market, and was able to sustain itself in the market (Nic Healey 2013). Similarly, in the aviation industry, Airbus entered and managed to sustain in a Boeing dominated market (Irwin and Pavcnik 2004). Figure 1 shows how before the introduction of Xbox 360, PS 2 already dominated the US market and since the emergence of Xbox, it took some time to mark its presence in the market against its competitor (similar for iPhone 4S and Galaxy S4)[1].

Note that, in the above examples, the new product always enters a monopoly, i.e., market dominated by a single competitor, and manages to survive. Hence, given these motivating examples, we focus on the following scenario: any given market is dominated by a single product and a competing product strategizes (i.e., advertising to niche communities) to survive in that market. Specifically, our mathemat-

---
[1]Real-data collected from Google Trends for the United States region which provides "insights into search pattern"(https://trends.google.com/trends/?geo=US).

ical model captures the early phase of market scenario when there is only one product in the market and a 'new' product tries to enter this monopoly.

When it comes to modeling product competition where individuals are the adopters, the interactions between them can be modeled as edges of a graph, which are linked via socially meaningful relationships like friendship or information exchange. Such social networks have become the paradigm for studying the spread of competing products and their adoption among individuals (Kalish 1985). Although here we stress on 'product' adoption and its influence among neighbors, we can draw a similar analogy within different contexts, for example, having an opinion (or preference) towards a political party/idea/information spreading in a network and how targeted advertisement plays an influential role in opinion formation/spreading mechanism. Given these motivating scenarios, we focus on settings where a new product enters a market dominated by a competitor.

Choice of Modeling Approach - bi-SIS: When two products compete with each other on a network, the eventual outcomes can be one of the following: absolute dominance, where one of the products dominates by wiping out its competitor and is adopted widely; die-out, where none of the products are adopted; or coexistence, where both of the products are adopted within the network (Prakash et al. 2012; Beutel et al. 2012; Sahneh and Scoglio 2014; Wei et al. 2016). Various mathematical models, such opinion dynamics-based models (i.e., analysis of how opinions form/spread in a population) such as De-Groot, Hegselmann-Krause, Friedkin-Johnsen, variants of epidemic models (Omic and Van Mieghem 2009; Liu et al. 2019), and others (Proskurnikov and Tempo 2017, 2018), have been proposed to capture the dynamics of competing products. Most of the works around classical opinion dynamics-based models focus on showing eventual consensus (everyone agrees to one of the opinions/rumors eventually), which has also been extended to capture the scenario of coexistence of two opinions (or clusters of opposing opinions). These works have shown coexistence by either relying on external factors in the form of stubborn agents (initially or currently forming bias towards an opinion), zealots (who never alter their initial stance) or via the inclusion of confidence bounds beyond which individuals do not interact (Gargiulo, Lottini, and Mazzoni 2008; Yang et al. 2012; Dandekar, Goel, and Lee 2013; Mobilia 2015; Amelkin, Bullo, and Singh 2017; Mukhopadhyay, Mazumdar, and Roy 2020; Anagnostopoulos et al. 2022; Nguyen et al. 2020). The addition of stubborn agents already guarantees that at least one of the opinions (the one followed by the stubborn agents) is always adopted and hence results in clusters of opinions. With the inclusion of these additional factors in the model setups, eventual consensus, or complete wipe-out is infeasible.

The propagation of information, influence maximization, idea, or opinion/rumor diffusion in a network has also been extensively studied via Linear Threshold (LT)/ Information Cascade (IC) model (Kempe, Kleinberg, and Tardos 2003; Chen, Yuan, and Zhang 2010). In these models, individuals can be either in an active (i.e., influenced) or an inac-

tive (i.e., free of influence) state. Here, the overarching assumption is that individuals either never change states or can only switch between active or inactive states randomly (Granovetter 1978; Pathak, Banerjee, and Srivastava 2010), with all nodes (users) becoming activated eventually. This is not always true in the context of products/opinions, as people tend to change their preferences/opinions/views from time to time.

The Susceptible-Infected-Susceptible (SIS) epidemic model has been used to study the spread of viruses in a network of populations (Lajmanovich and Yorke 1976; Van Mieghem, Omic, and Kooij 2008). However, the *bi-SIS* model, a variant of the SIS model (single virus), has been used to characterize the spreading dynamics of two competing products/opinions in a network (Prakash et al. 2012; Sahneh and Scoglio 2014). In such a model, via word-of-mouth or by being influenced by their neighbors, users adopt one of the two competing products (get infected) and gradually recover to become susceptible to infection again. Unlike the opinion dynamics-based model or LT/IC model, the bi-SIS model does not necessitate that the individuals have an opinion, positive or negative, ex-ante; in contrast, it allows individuals to switch states from being susceptible (inactive) to an infected (active) state and vice-versa. It has proved to be effective in capturing eventual consensus or coexistence in any general arbitrary graph, including special graphs such as complete or random graphs (Wang et al. 2003; Li, van de Bovenkamp, and Van Mieghem 2012; Prakash et al. 2012; Santos, Moura, and Xavier 2015; Yang, Yang, and Tang 2017). It can be seen that the bi-SIS model aptly captures competitiveness without the need to have any prior assumptions or dependence on a specific network structure. Thus, the bi-SIS model captures all the three possible state spaces along with providing conditions for each of the outcomes i.e., die-out, absolute dominance, and coexistence.

Although the bi-SIS model is effective in capturing all the possible outcomes, most of the existing works only focus on the stability analysis of the equilibrium points or identifying which of the three possible outcomes the system converges to (Van Mieghem, Omic, and Kooij 2008; Li, van de Bovenkamp, and Van Mieghem 2012; Prakash et al. 2012; Sahneh and Scoglio 2014; Doshi, Mallick, and Eun 2021). Furthermore, none of these works explicitly quantify the magnitude of existence; for example, the market share of either one or both of the competing products, especially in a general graph setting. Note that with respect to opinion dynamics, for example, the simple linear French-De-Groot model with stubborn agents (French Jr 1956; Proskurnikov and Tempo 2017) has shown quantifying results on a general graph. However, as mentioned earlier, unlike the bi-SIS model, the opinion dynamics-based models do not capture all the three possible state space and thus pose a limitation when it comes to the analysis of the market share of competing products.

Our Contributions: In this paper, we take a step forward and propose techniques to quantify and improve the market share of the new/weaker product when it competes with an existing dominant product in the market via bi-SIS epidemic model. Under limited budget constraints, this new company

launching its product would aim for survival, and the best strategy in this situation is to target certain users via advertisement/promotional offers to form a small support users' group (community) who, with their positive reviews, can spread about this new product. The motivation behind promoting such communities (additional edges among the users of the graph) is to increase the rate of influence so as to maximize the chance of its survival. Specifically, a resource-constrained company would want to identify an optimal set of individuals to form such communities to maximize its (positive) market share as much as possible.

When it comes to quantifying the resulting market share on a general graph, there is no known closed-form for the positive fixed point (market share) even for a single-SIS model. The situation is further compounded when the competition comes into play (bi-SIS). To tackle this challenge, we employ the perturbation approach to optimally assign the resources under a specific budget structure based on model parameters and approximate the resulting market share of this original losing/weaker product. Our proposed budget structure also accounts for a minimum budget below which the company, despite spending positive amount of budget on advertisement expecting a non-zero return on investment, still gets zero market share and thus cannot compete with its competitor. By leveraging our theoretical findings, we develop a heuristic approach which returns a locally optimal choice of users, under a total budget constraint with heterogeneous costs for different users. We also demonstrate the efficacy of our approach by comparing against different benchmarks such as different standard centrality measures, over real network topology with various cost constraints for the users in the community.

2 System Model and Problem Setup

In this section, we take an initial look at the problem of optimally targeting individual users and recruiting them to participate in community, with the goal of maximizing the long term market share of the new product. It has been shown that when two products compete on a 'same' graph[2] with same set of users, at steady state, the stronger one dominates and wipes-out the weaker one (Prakash et al. 2012). WLOG, in our setting, on a 'same' graph, *Product 1* is the *dominant* product in the market and *Product 2* is the *new/weaker* product (with zero market share) trying to compete against Product 1 for a positive market share. In our context, a community or online discussion forum provides a platform where a group of interacting users actively exchange information about common interests. Few pertinent examples can be a subreddit – a community in Reddit (Reddit 2009) or influencer pods in social media platforms. In a subreddit, members can post information which is accessible to all members of the group, irrespective of whether members follow each other or not. In the case of influencer pods, groups of influencers work together to boost each other's engagement (e.g., product's marketing) (Contributor 2024). We focus on the problem of *creating* one such community. We

---

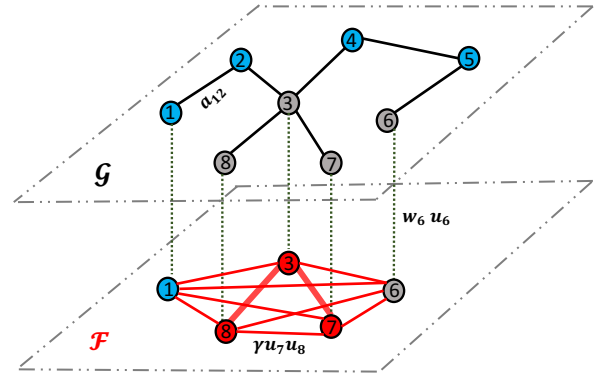[2]Same graph is analogous to one social network platform, for example, Facebook.

Figure 2: Illustration of users forming *community* (Nodes 1,3,6-8 in layer $\mathcal{F}$). The blue (red) nodes denote users adopting Product 1 (Product 2) at any given instance, and gray nodes are the users who do not own any of the products.

assume that any interactions or discussions between users in this newly formed community are solely about the new company's product, and the influence among the users of the community is uniform/homogeneous.

Suppose that any user whom the company has targeted in the form of advertisement/content-marketing chooses to engage in the community-based discussions with a certain probability. We can safely assume that the more aggressive the company is with its advertising efforts targeted towards the user, the higher is the probability of the user participating in this created community. In real life, these targeted advertisement efforts take various forms, such as product placement, early-bird discounts, or promotional offers, and cost a certain amount of resources. Later in this section, we give a more detailed discussion on the modelling aspects of the cost structure for our problem.

The System Model

Let u_i denote the probability with which any user $i \in \mathcal{N}$ actively engages in such a community. We assume that all users participating in the community are able to communicate with one another, and Product 2's influence spreads across this community with rate $\gamma\beta_2$ (where γ is a factor capturing the rate of influence relative to β_2). To clearly visualize this creation of a community, we draw it as a separate layer, as shown in Figure 2. The top layer $\mathcal{G}$ denotes the original graph where the mutual follower relationships are according to adjacency matrix $\mathbf{A}$, and layer $\mathcal{F}$ denotes the created community, in which a user $i \in \mathcal{N}$ participates with probability u_i. The spread of influence within this community, is akin to adding an edge of weight γ for any two users i and j *known* to be actively participating in the community. Specifically, the parameter γ denotes the amount the company is willing to spend on the members of the community to improve their communication efficiency (user's interaction). If the two users were already connected prior to being a part of the community, that is, $a_{ij} = 1$ (a_{12} in Figure 2), then γ is simply an additional reinforcement of their connection as a consequence of them interacting through mul-

Symbol	Definition/Description
$\mathcal{G}(\mathcal{N},\mathcal{E})$	Undirected, connected graph
$\mathcal{N} = \{1, 2, \ldots, N\}$	Set of nodes
$\mathbf{A} \triangleq [a_{ij}]$	Adjacency matrix of graph $\mathcal{G}$
β_1 (β_2)	Adoption rate of product 1 (or product 2)
δ_1 (δ_2)	Disown rate of product 1 (or product 2)
$\tau_1 = \beta_1/\delta_1$ (τ_2)	Effective adoption rate of product 1 (or product 2)
u_i	Probability with which any user I actively engages in community
w_i	Cost of recruiting user i in community
γ	Influence rate among community users (uniform)
C	Budget allocated by company launching new product (product 2)
$\lambda = \lambda(\mathbf{H}) > 0$	Largest eigenvalue of matrix $\mathbf{H}$
$\mathbf{x}^*(\mathbf{y}^*) \gg \mathbf{0}$	Positive fixed points corresponding to single-SIS
$\mathbf{S_v} \triangleq diag(\mathbf{1} - \mathbf{v})$	Diagonal matrix with elements $(1 - v_i)$

Table 1: Symbols

tiple media. The thick 'red' lines in Figure 2 highlight this reinforcement whereas the 'black' lines indicate the original interaction links. Specifically, the introduction of additional edge weight (as shown in layer $\mathcal{F}$ in Figure 2) is to induce an artificial community-like structure (clique) where the edge weight denotes how much the users are influenced from other users through their posts/shared information. The probability with which both i and j engage in the forum thus becomes $u_i u_j$ ($u_7 u_8$ in Figure 2), and the total weight of edges connecting them is, on average, given by $a_{ij} + \gamma u_i u_j$. Table 1 explains the terminology we have used in this paper. We denote vectors ($\mathbf{v} \in \mathbb{R}^N$) and matrices ($\mathbf{H} \in \mathbb{R}^{N \times N}$) as bold lowercase and uppercase letters, respectively. Note that in our model we assume that the Product 1 does not advertise to counter this 'incursion' by Product 2 and a user cannot adopt both products at the same time. We also assume that the complete network structure information is available to the company before launching the new product.

As mentioned earlier in Section 1, we use the bi-SIS ODE system to model our system dynamics, which yields the following

$$\dot{x}_i(t) = \beta_1(1 - x_i(t) - y_i(t)) \sum_{j \in \mathcal{N}} a_{ij} x_j(t) - \delta_1 x_i(t)$$

$$\dot{y}_i(t) = \beta_2(1 - x_i(t) - y_i(t)) \sum_{j \in \mathcal{N}} [a_{ij} + \gamma u_i u_j] y_j(t) \quad (1)$$

$$- \delta_2 y_i(t),$$

where $x_i(t)$ ($y_i(t)$) is the probability with which user i owns Product 1 (Product 2) at time $t \geq 0$. Also, $\beta_2(a_{ij} + \gamma u_i u_j)$ is the overall expected rate at which the user j, say who already owns Product 2, causes the susceptible (owns neither products) user i to adopt Product 2. Note that (1) is an extension of the original bi-SIS model (Prakash et al. 2012) whose system dynamics for any two competing products propagating on a graph with adjacency matrix $\mathbf{A}$ is given by

$$\dot{x}_i(t) = \beta_1\big(1 - x_i(t) - y_i(t)\big) \sum_{j \in \mathcal{N}} a_{ij} x_j(t) - \delta_1 x_i(t)$$

$$\dot{y}_i(t) = \beta_2\big(1 - x_i(t) - y_i(t)\big) \sum_{j \in \mathcal{N}} a_{ij} y_j(t) - \delta_2 y_i(t).$$

It has been established that the above system dynamics results in only the stronger product to dominate the market and wipe out the weaker one. We, in this paper, try to bridge this gap by inducing a community ($\gamma u_i u_j$ in (1)) within the original graph so that the weaker product can survive (i.e., a positive market share) in the market.

Using $\mathbf{u} = [u_i]_{i \in \mathcal{N}}$ to denote the vector of probabilities with which users participate in the community, we can write (1) in a matrix-vector form as follows:

$$\dot{\mathbf{x}}(t) = \beta_1 \text{diag}(\mathbf{1} - \mathbf{x}(t) - \mathbf{y}(t))\mathbf{A}\mathbf{x}(t) - \delta_1 \mathbf{x}(t)$$

$$\dot{\mathbf{y}}(t) = \beta_2 \text{diag}(\mathbf{1} - \mathbf{x}(t) - \mathbf{y}(t)) \big[\mathbf{A} + \gamma \mathbf{u}\mathbf{u}^T\big] \mathbf{y}(t) \quad (2)$$

$$- \delta_2 \mathbf{y}(t).$$

As can be seen in (2), the extra edge weights added to the original adjacency matrix $\mathbf{A}$, as a result of creating community in the manner as described earlier, are in the form of a rank one update given by $\mathbf{A} + \gamma \mathbf{u}\mathbf{u}^T$.[3] Later in Section 3, we use this special form of the update to aid our analysis. Our propagation model (2) yields the equilibrium equations

$$\mathbf{x} = \tau_1 \text{diag}(\mathbf{1} - \mathbf{x} - \mathbf{y})\mathbf{A}\mathbf{x} \quad (3)$$

$$\mathbf{y} = \tau_2 \text{diag}(\mathbf{1} - \mathbf{x} - \mathbf{y}) \big[\mathbf{A} + \gamma \mathbf{u}\mathbf{u}^T\big] \mathbf{y}, \quad (4)$$

where $\mathbf{x}$ and $\mathbf{y}$ are the fixed points of (2) for Product 1 and Product 2, respectively.

---

[3] $\gamma \mathbf{u}\mathbf{u}^T$ leads to a clique with homogeneous influence rate among the users. Note that heterogeneous influence rate (while ensuring more generality) introduces further complications in the analysis, which we leave as a future work.

Problem Setup

We present the problem of maximizing the market share $\bar{y} = 1/N \sum_{i \in \mathcal{N}} y_i$ (where $\mathbf{y} = [y_i]$ is the fixed point for Product 2 from (4)) under a limited budget constraint, say C, by the optimization problem given by

$$(\mathcal{M}) \qquad \underset{\gamma, \mathbf{u}}{\text{maximize}} \quad \bar{y} = \frac{1}{N} \sum_{i \in \mathcal{N}} y_i$$

$$\text{subject to} \quad \sqrt{\gamma} \sum_{i \in \mathcal{N}} w_i u_i \leq C$$

$$u_i \in [0,1] \quad \forall i \in \mathcal{N}.$$

We structure the cost constraint in this optimization problem in a heterogeneous manner. Under this setting, w_i is the cost of recruiting user i in a newly created community. This user i will be included in the forum with probability u_i such that the 'average cost' of having user i in the forum would be $w_i u_i$ ($w_6 u_6$ in Figure 2). This setting is very general and can capture various scenarios such as w_i being larger (more expensive to recruit) for more popular users/influencers as would be the case in reality. The company also allocates a portion of its budget towards γ, which captures the overall communication efficiency across the channels established for user interaction, as part of creating the discussion group[4]. We reflect this in the budget constraint in our optimization problem $\mathcal{M}$, with the presence of the $\sqrt{\gamma}$ term. This ensures that the scaling between γ and $\mathbf{u}$ is the same as $\gamma \mathbf{u}\mathbf{u}^T$ as in (4). To be more specific, for any choice of $\mathbf{u}$ up to a multiplicative constant, we can allocate C between γ and $\mathbf{u}$ without affecting the fixed point $\mathbf{y}$. Our scaling ensures that $\hat{\mathbf{u}} = c\mathbf{u}$ and $\hat{\gamma} = \gamma/c^2$ for any constant c renders the terms $\gamma \mathbf{u}\mathbf{u}^T$ unchanged. Any other scaling would force us to consider every possible value of $\mathbf{u}$ and γ *separately* and thus obfuscate the meaning of the budget constraint for our optimization problem and subsequently our analysis on the optimal choice of $(\gamma, \mathbf{u})$ on the maximal market share $\bar{y}$. This completes our problem set up, and we are now ready to address its analysis in the next section.

3 Analysis of Product 2's Market share $\bar{y}$

In this section, we begin by first analyzing the budget constraint in $\mathcal{M}$ and the minimum amount of budget needed to ensure positive market share of Product 2, $\bar{y} > 0$. We also establish a key result on the behavior of $\bar{y}$ as the entries u_i of the vector of probabilities $\mathbf{u}$ (probability with which any user $i \in \mathcal{N}$ actively engages in the community) are perturbed. We then define an optimization problem that helps in updating $\mathbf{u}$, consequently helping towards obtaining an improved market share of Product 2 under the same budget constraint. We defer to the Appendix all the technical proofs of the results presented in this section.

Minimum Budget for $\bar{y} > 0$

To begin with, we provide the following proposition regarding the budget constraint in $\mathcal{M}$.

Proposition 3.1. *Any solution $(\gamma^*, \mathbf{u}^*)$ to the optimization problem $\mathcal{M}$ always satisfies the budget constraint with equality, that is,*

$$\sqrt{\gamma^*} \sum_{i \in \mathcal{N}} w_i u_i^* = C.$$

Proposition 3.1 tells us that the optimal market share of Product 2 is always obtained by spending all the available budget. This follows from the monotonicity property[5] of the ODE system (2) in that $\bar{y}$ turns out to be an increasing function of γ and $\mathbf{u}$, albeit being nonlinear. See Appendix A for details.

For general bi-SIS models, Product 2 needs to satisfy a threshold type survival condition $\tau_2 \lambda(\mathbf{S}_{\mathbf{x}^*}[\mathbf{A} + \gamma \mathbf{u}\mathbf{u}^T]) > 1$, in order to secure a positive market share in the long run[6]. This condition captures the relationship between the initial market share of Product 1, given by $\bar{x}^* = 1/N \sum_{i \in \mathcal{N}} x_i^*$, and the budget C, which is a function of the pair $(\gamma, \mathbf{u})$. Specifically, for a given τ_2 and any $\mathbf{u}$ fixed up to a multiplicative constant, the survival condition indicates that when $\bar{x}^*$ is large, the term $\gamma \mathbf{u}\mathbf{u}^T$ also needs to be large (to ensure a large enough eigenvalue)[7] in order for Product 2 to survive in the long term, subsequently needing a large budget to do so. This implies that as the market share of Product 1 increases, the minimum budget required for Product 2 to survive also increases, which can also be seen from the heat-map as depicted[8] in Figure 3. For an arbitrarily chosen $\mathbf{u}$, fixed up to a multiplicative constant, Figure 3 illustrates how the steady-state market share of Product 2 varies with C and $\mathbf{x}^*$. The 'red' curve in Figure 3 indicates the minimum budget required to attain $\bar{y} > 0$, and any amount spent below the minimum budget leads to a zero market share. In other words, a positive budget spent by the company towards its product placement does not necessarily guarantee even a small non-zero return; the company launching a new product must account for the existing competitor's market share to decide the minimum budget to ensure its non-zero long-term market share.

Main Result

We now turn our attention towards establishing a key result on the sensitivity of $\bar{y}$, with respect to entries u_i of $\mathbf{u}$. For

---

[4]The parameters γ and w_i set in this model are part of the budget, which in a practical situation would be determined as a function of the total budget allocated by the company launching the new product.

[5]Arising from the fact that the bi-SIS system is a *monotone dynamical system* (Doshi, Mallick, and Eun 2021).

[6]Both products survive (coexistence equilibria) when $\tau_1 \lambda(\mathbf{S}_{\mathbf{y}^*}\mathbf{A}) > 1$ and $\tau_2 \lambda(\mathbf{S}_{\mathbf{x}^*}[\mathbf{A} + \gamma \mathbf{u}\mathbf{u}^T]) > 1$, and the system globally converges to $(\mathbf{x}, \mathbf{y}) \gg (\mathbf{0}, \mathbf{0})$. Here, $\mathbf{x}, \mathbf{y}$ correspond to coexistence fixed points (Sahneh and Scoglio 2014; Doshi, Mallick, and Eun 2021).

[7]This follows from the fact that the spectral radius of a nonnegative matrix is a non-decreasing function of its elements (Meyer 2000) and the terms in $\mathbf{S}_{\mathbf{x}^*} = \text{diag}(\mathbf{1} - \mathbf{x}^*)$ is decreasing in $\mathbf{x}^*$.

[8]We numerically simulate the ODE system (2) for a small graph of dolphins' population (62 nodes, 159 edges) (Lusseau et al. 2003).
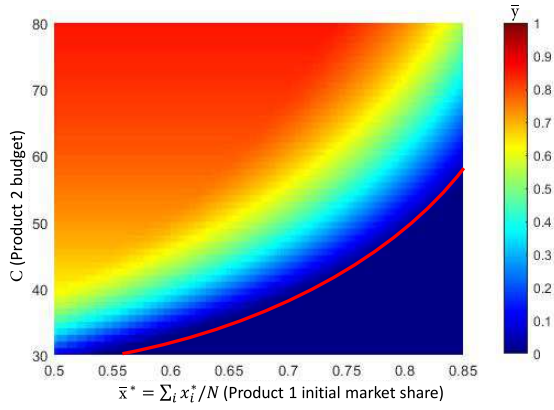
Figure 3: Heat map illustration of the market share as a function of budget C and $\bar{x}^*$.

a general graph with adjacency matrix $\mathbf{A}$, it is impossible to obtain the optimal solution to $\mathcal{M}$ in a closed form; the main issue being the absence of a closed-form expression of $\bar{y}$. Even for the single-SIS model in a general graph setting, there is no known closed-form for the positive fixed point. The situation is further compounded when competition comes into play in the bi-SIS model. Therefore, to gain better insight into $\bar{y}$ and how it would change with model parameters, we begin by applying small perturbation to u_i for each $i \in \mathcal{N}$. As previously mentioned, change in γ does not affect the fixed point in (3), (4) when the budget is kept constant, since the term $\gamma \mathbf{u}\mathbf{u}^T$ remains the same under the aforementioned scaling for γ and $\mathbf{u}$ with the same budget C. The only way to obtain an increase in $\bar{y}$ is by *heterogeneously* changing the entries u_i, and thereby obtaining an *updated* version which is no longer a multiplicative constant of the initial $\mathbf{u}$.

To proceed, we first define the terms *critical curve* and *critical parameters*, which we use throughout the rest of this section. The 'critical curve' is the set of parameters at which the aforementioned survival condition is satisfied with equality, and we call any such set of parameters 'critical' if they are on the critical curve. Let $(\gamma^c, \mathbf{u}^c)$ denote a pair of such critical parameters for which $\tau_2\lambda(\mathbf{S}_{\mathbf{x}^*}[\mathbf{A} + \gamma^c \mathbf{u}^c \mathbf{u}^{cT}]) = 1$. Lemma 3.2 provides a feasible 'initial' pair of critical parameters.

Lemma 3.2. *Let $\mathbf{v} \in \mathbb{R}^N$ be the PF eigenvector of $\mathbf{S}_{\mathbf{x}^*}\mathbf{A}$ associated with $\lambda(\mathbf{S}_{\mathbf{x}^*}\mathbf{A})$. Then, for the choice of $\mathbf{u}^c$ and γ^c such that*

$$\mathbf{u}^c \propto \mathbf{S}_{\mathbf{x}^*}^{-1}\mathbf{v}, \quad and \quad \gamma^c = \frac{1/\tau_2 - \lambda(\mathbf{S}_{\mathbf{x}^*}\mathbf{A})}{\mathbf{u}^{cT}\mathbf{S}_{\mathbf{x}^*}\mathbf{u}^c},$$

we have $\tau_2\lambda(\mathbf{S}_{\mathbf{x}^*}[\mathbf{A} + \gamma^c \mathbf{u}^c \mathbf{u}^{cT}]) = 1$.

For any given such pair, let $\boldsymbol{\nu}^c = [\nu_i^c]$ be the PF[9] eigenvector of $\mathbf{S}_{\mathbf{x}^*}[\mathbf{A} + \gamma^c \mathbf{u}^c \mathbf{u}^{cT}]$. From the threshold condition we know that $\bar{y} = 0$ when $\tau_2\lambda(\mathbf{S}_{\mathbf{x}^*}[\mathbf{A} + \gamma^c \mathbf{u}^c \mathbf{u}^{cT}]) = 1$.

---

[9]The Perron-Frobenius (PF) eigenvector is the eigenvector corresponding to the largest eigenvalue of a non-negative, irreducible matrix (Meyer 2000).

As a first step towards obtaining a positive $\bar{y}$, we apply perturbation $\alpha_i > 0$ ($\boldsymbol{\alpha} = [\alpha_i]$) to any i'th entry u_i^c of $\mathbf{u}^c$. The updated entry is then given by $u_i^c + \alpha_i$, and we have the following result.

Theorem 3.3. *The market share of Product 2 as a function of the perturbation $\alpha_i > 0$ to u_i^c, can be written as*

$$\bar{y}(\alpha_i) = \zeta\nu_i^c\alpha_i + O(\alpha_i^2), \tag{5}$$

where $\zeta > 0$ is a constant independent of i.

Theorem 3.3 tells us that the magnitude of increase in $\bar{y}$ due to increase in u_i^c (that is, α_i) is roughly proportional to ν_i^c with an error term[10]. This implies that there is a greater reward in terms of market share to invest in nodes with larger ν_i^c. The result from Theorem 3.3 helps in obtaining a direction leading to an increase in $\bar{y}$.

Suppose we have an initial pair of model parameters which satisfy the threshold condition, and thus form a critical pair $(\gamma^c, \mathbf{u}^c)$. These chosen parameters would then satisfy the budget $C = \sqrt{\gamma^c}\sum_{i \in \mathcal{N}} w_i u_i^c$. By keeping the budget constant, we intend to move in a direction that obtains the best possible increase in $\bar{y}$ from its initial value of $\bar{y} = 0$ at the critical point. We do so by applying a perturbation to u_i^c, which needs to have both positive and negative components to ensure that the budget constraint remains binding. Also, Theorem 3.3 indicates that the increase in $\bar{y}$ is roughly proportional to ν_i^c.

We leverage this result and the perturbation approach to formulate another optimization problem $\mathcal{P}(\boldsymbol{\nu}, \epsilon, \mathbf{u})$. Since Theorem 3.3 holds in a small region around the critical curve, we limit the maximum size of perturbation to $\epsilon > 0$. Specifically, for a given initial $\mathbf{u}^c$, we define $\tilde{u}_i \triangleq u_i^c + \alpha_i$, for all $i \in \mathcal{N}$ where $\alpha_i \in [-\epsilon, \epsilon]$ and the step size ϵ is a small constant that controls the maximal size of perturbation in $\mathbf{u}$. We then consider the following optimization problem:

$$(\mathcal{P}(\boldsymbol{\nu}, \epsilon, \mathbf{u})) \qquad \underset{\alpha_1, \cdots, \alpha_N}{\text{maximize}} \quad \sum_{i \in \mathcal{N}} \alpha_i\nu_i$$
$$\text{subject to} \quad \sum_{i \in \mathcal{N}} w_i\alpha_i = 0,$$
$$\alpha_i \in [-\epsilon, \epsilon].$$

The solution $\boldsymbol{\alpha}$ to $\mathcal{P}(\boldsymbol{\nu}, \epsilon, \mathbf{u})$ is based on the decreasing order of (ν_i/w_i) such that it satisfies the given constraints. This solution is essentially a relaxation of the well-known *0-1 Knap-sack* problem (Garey and Johnson 1979), which for a given initial critical parameters $(\gamma^c, \mathbf{u}^c)$ leads to a direction resulting in the best possible increase in $\bar{y}$. The perturbed $\mathbf{u}$ is the optimal number of users who can be targeted with advertisements of Product 2, and thus results in a positive market share.

---

[10]Note that Theorem 3.3 shows the sensitivity of $\bar{y}$ with respect to 'one' particular $\alpha_i > 0$ for which we expect to see the best possible increase in the market share. Any $\alpha_i < 0$ results in a zero market share $\bar{y} = 0$ since $\tau_2\lambda(\mathbf{S}_{\mathbf{x}^*}[\mathbf{A} + \gamma^c \mathbf{u}^c \mathbf{u}^{cT}]) \leq 1$.

Algorithm 1: Local Search for bi-SIS

Input: $\mathbf{A}, \tau_1, \tau_2, \mathbf{w}, \epsilon$
Output: $\mathbf{u}$
1 **Initialize:**
2     $\mathbf{v}(0) = \boldsymbol{\nu}^c$ /* $\mathbf{S_{x^*}A}$'s eigenvector */
3     $\mathbf{u}(0) = \mathbf{u}_0$ /* as in Lemma 3.2 */
4     $\gamma(0) = \gamma_0$ /* as in Lemma 3.2 */
5     $\epsilon(0) = \epsilon$
6 $\epsilon(1) \leftarrow \min\{\epsilon(0), \min\{\mathbf{u}(0)\}\}$
7 $\boldsymbol{\alpha} \leftarrow \mathcal{P}(\mathbf{v}(0), \epsilon(1), \mathbf{u}(0))$ /* Solution to optimization problem */
8 $\mathbf{u}(1) \leftarrow \mathbf{u}(0) + \boldsymbol{\alpha}$

4 Experimental Setup & Results

In this section, we present simulation results on real-world graphs under a cost constraint for a given initial $(\gamma, \mathbf{u})$ pair according to Lemma 3.2. We first briefly describe the simulation setup for our experiments, which includes the budget structure and then describe the baseline methods that we use to compare its performance against our approach.

Simulation Setup

We use the following two real-world social network based graph datasets for the simulations, which are undirected, connected graphs from SNAP (Leskovec and Krevl 2014) and Network (Rossi and Ahmed 2015) repositories.

1. Social circles- Facebook: Friends list of users using Facebook App, which contains 4,039 nodes and 88,234 edges (McAuley and Leskovec 2012).

2. Facebook pages of public figures: Nodes which are blue-verified public figures' facebook pages and edges are mutual likes among them. It contains 11,565 nodes and 67,000 edges (Rozemberczki and Sarkar 2020).

As Product 1 already exists in the network and being stronger than Product 2, the effective adoption rates are such that initially $\tau_1 \lambda(\mathbf{S_{y^*}A}) > 1$ and $\tau_2 \lambda(\mathbf{S_{x^*}A}) \le 1$ (zero market share). Here, $\tau_1 = 0.8, \tau_2 = 0.05$ for Facebook dataset with 4039 nodes and $\tau_1 = 0.4, \tau_2 = 0.03$ for Facebook dataset with $\approx 12k$ nodes. In the simulation results, we are interested in the immediate increment of $\lambda(\mathbf{S_{x^*}B})$ (where $\mathbf{B} = \mathbf{A} + \gamma \mathbf{u}\mathbf{u}^T$) from its initial value $\tau_2 \lambda(\mathbf{S_{x^*}B}) = 1$ for a given initial pair $(\gamma^c, \mathbf{u}^c)$.

We set the budget as per Proposition 3.1 and run our simulation for different values of ϵ (parameter that controls the perturbation of $\mathbf{u}$) under two scenarios: (i) when the cost of resource allocation is homogeneous across all nodes, for example, $w_i = 1$ for all $i \in \mathcal{N}$ and, (ii) when the cost of resource allocation is heterogeneous, for example, one instance of $\mathbf{w}$ sampled uniformly at random. The proposed algorithm for obtaining the locally optimal set of users is given in Algorithm 1.

Comparison with Baseline Methodologies

To evaluate the performance, we compare our proposed approach of selecting a community against two methodologies: (i) standard centrality measures such as degree-based centrality and eigenvector-based centrality, and (ii) NetShield

Algorithm 2: Centrality Measure Methodology

Input: $\mathbf{A}, \tau_1, \tau_2, \gamma^c, N, C, \mathbf{w}, \mu$
   /* μ = chosen centrality measure */
Output: $\mathbf{u}$
1 $u_i = \frac{C \mu_i w_i}{\sum_{j \in T} \mu_j w_j}$ for all $i \in N$

algorithm (Chen et al. 2015), as baseline algorithms. In this section, we describe each of the methodologies along with the reasoning for selecting these as baseline algorithms.

Degree-based centrality is a measure of node connectivity based on the number of links (degree) each node has with others which is useful in identifying popular nodes in a network who can further connect to a wider network (Newman 2018). Similar to this, eigenvector-based centrality (EVC) measures a node's influence that it has on other nodes in the network. This is done by assigning an importance score based on the fact that a node is important when it is linked to another important node. EVC identifies not only the influence of a node on a direct connection, but also the impact on the entire network (Wasserman and Faust 1994). We chose to compare against these measures to gain a holistic analysis against such well-known centrality methods which are generally used to understand the relationship structure in a social network (Freeman 1978; Friedkin 1991). Also, one of the reasons to choose these centrality measures is because one would assume that spending resources on influential users (users with higher degree or evc score) would return higher market share, which we show otherwise in our simulation results.

From the methodological perspective, Chen et. al. (Chen et al. 2015) uses a very similar epidemic modeling approach and have proposed an algorithm 'NetShield' that helps in determining the number of users/nodes who can be immunized to contain the epidemic. Within our context, we utilize this algorithm as one of the baselines to determine the number of users who can be selected to form the community. In (Chen et al. 2015), NetShield algorithm aims to minimize the influence spread by first selecting and then deleting those nodes who are more likely to further spread the influence in the graph. For baseline comparison, we utilize this algorithm to determine those highly influential users who can be recruited to form a community who can then increase the influence spread of the new product in the graph. In each of these baseline methodologies, while the total budget remains the same, that is, $C = \sum_i \sqrt{\gamma} w_i u_i$, we assign the cost w_i (either homogeneous or heterogeneous) based on the ratio of its degree and importance score for degree and eigenvector centrality measure, respectively as outlined in Algorithm 2. Similarly, the cost w_i is assigned to the users obtained from the NetShield algorithm. The obtained $\mathbf{u}$ is then similarly fed to (4) to obtain the market share of Product 2.

Simulation Results

Here, we present the comparison results for both the homogeneous and heterogeneous cost distribution for the above mentioned real-world graphs dataset. To visualize our system in the two dimensions, we represent Product 2's (Prod-

uct 1) market share on the y-axis as $AvgY = (1/N) \sum_{i \in \mathcal{N}} y_i$ ($AvgX = (1/N) \sum_{i \in \mathcal{N}} x_i$).

As mentioned in Section 3, the step size ϵ is a small constant that controls the maximal size of perturbation in $\mathbf{u}$. The obtained $\mathbf{u}$ is then fed into the bi-SIS ODE (4) along with other associated parameters which is then simulated over 1000 time steps to obtain the market share of Product 2 (Similarly for Product 1 using (3)). Figures 4 – 7 show how the market shares of Product 1 and Product 2 change based on different ϵ values when the cost is homogeneously and heterogeneously distributed. The results indicate that as the ϵ value increases, the market share of Product 2 also increases while the market share of Product 1 decreases, under both homogeneous and heterogeneous cost distribution. This demonstrates that the users' probability of willingness to participate in the community also increases as the ϵ value increases, thereby resulting in a higher market share.
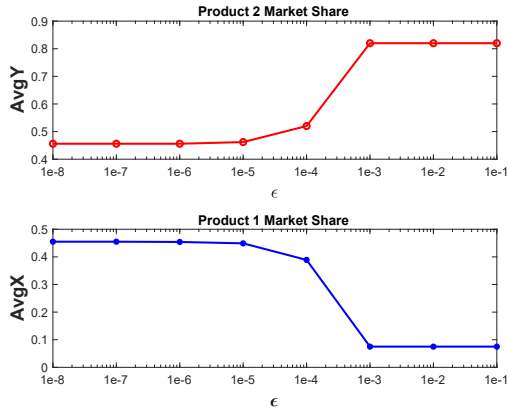

Figure 4: Facebook ($\approx 4k$ nodes): Change in $\bar{\mathbf{y}}$ (AvgY) and $\bar{\mathbf{x}}$ (AvgX) for *Homogeneously* distributed cost and different values of ϵ.
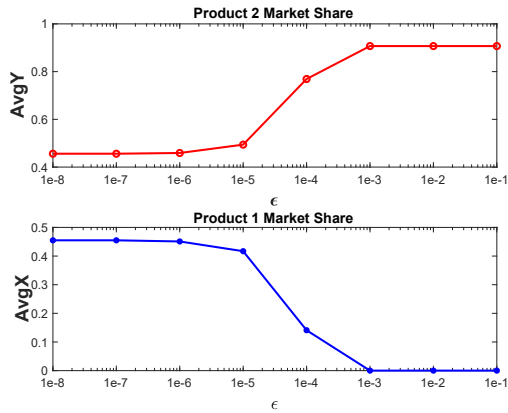

Figure 5: Facebook ($\approx 4k$ nodes): Change in $\bar{\mathbf{y}}$ (AvgY) and $\bar{\mathbf{x}}$ (AvgX) for *Heterogeneously* distributed cost and different values of ϵ.

Figure 8 shows that for the locally optimal choice of users returned by our approach for different values of ϵ, under both
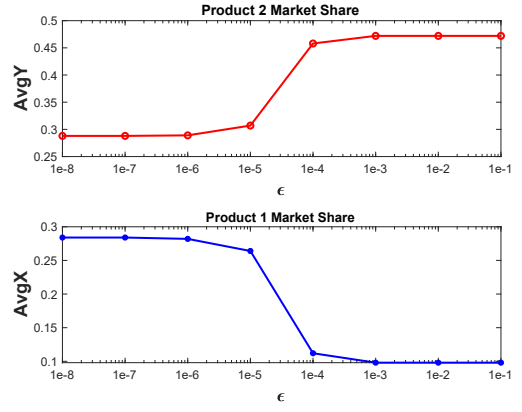

Figure 6: Facebook ($\approx 12k$ nodes): Change in $\bar{\mathbf{y}}$ (AvgY) and $\bar{\mathbf{x}}$ (AvgX) for *Homogeneously* distributed cost and different values of ϵ.
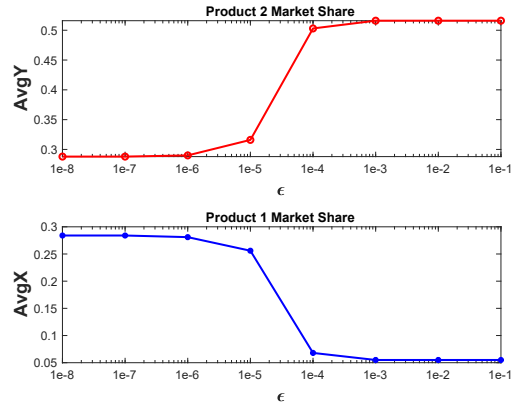

Figure 7: Facebook ($\approx 12k$ nodes): Change in $\bar{\mathbf{y}}$ (AvgY) and $\bar{\mathbf{x}}$ (AvgX) for *Heterogeneously* distributed cost and different values of ϵ.
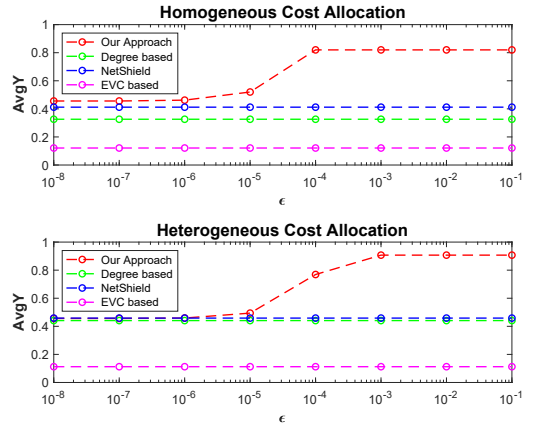

Figure 8: Facebook ($\approx 4k$ nodes): Change in $\bar{\mathbf{y}}$ (AvgY) for Homogeneously & Heterogeneously distributed cost against Degree-based, Eigenvector-based centrality measure, & NetShield Algorithm.

heterogeneous and homogeneous cost distribution, the initial weaker/new product has a higher market share than the base-

line centrality measures and NetShield algorithm, thus validating our theoretical results. Although it is usually assumed that spending resources on users with the highest number of followers (influential users) will result in a higher market share, our simulation results show otherwise, i.e., any set of users selected by our algorithm, irrespective of the number of followers, can result in a higher market share. This result has profound implications on the budget constraint, i.e., maximum amount that can be spent on recruiting users, as users with more followers are usually very expensive. Note that similar numerical results can indeed be obtained for even larger graphs than shown here, for any choice of budget, C, as long as the cost assigned to each u_i is not too large and is thus, omitted here to avoid repetition.

5   Conclusion & Future Work

We have developed an algorithm that returns an optimal set of users that form a community, that is, the best possible pair of $(\gamma, \mathbf{u})$ and these set of users when targeted with advertisement under a given budget constraint results in a positive market share of the new product, Product 2. Through perturbation approach we identify the set of users that can be influenced through advertisement so that Product 2 can enter into the network with a positive market share. We also compare our proposed approach against standard centrality measures and NetShield algorithm (used in finding nodes in SIS epidemic model) using two different datasets. The results indicate that our approach of targeting the chosen users outperforms in both the cases when the budget is assigned either homogeneously or heterogeneously to users that are part of the community. We believe that our work is the first step towards finding an optimal size of community for a new firm (launching a product) to compete against a dominant product in a market. By basing our theoretical analysis using the bi-SIS epidemic model and extensive simulation on different real-world graphs demonstrate the efficacy of our approach, including its exceeding performance over the standard centrality measures.

While this paper focuses on finding an optimal set of users to maximize the market share of a new product competing against an existing dominant product, we abstract away from analyzing multiple products' (more than two) competition. As an avenue of future research, utilizing a multi-layer graph structure and incorporating multiple products' competition would greatly contribute to a better understanding of how a new product strives to improve its market share in the presence of multiple products.

Broader Impact

Our goal is to develop a heuristic approach through which a new/weaker product can survive (with a positive market share) along with an existing dominant product. This has wider managerial implications for brand managers running new product introduction campaigns. Specifically, such campaigns involve targeting niche communities through advertisements or promotional offers. For example, brand managers utilize social media platforms such as Instagram, where they perform brand advertisements through influencers/niche groups, while operating within a limited advertisement budget. Our model captures this exact dynamics by identifying an optimal set of users under a budget constraint to form a community who can be similarly targeted.

We use real-world social network data of Facebook users and compare it against standard centrality measures. The comparison results indicate that it is not always necessary to spend all resources on more famous influencers/public figures for product promotion, but the community identified by our approach provides a much better market share. Again, this has wider implications and provides insights on how best to spend the limited advertisement budget. Finally, we assume that companies have the complete network information, which might not be always true. However, even with incomplete network information and through our approach, companies can identify niche set of users and ensure survival. Further, more information regarding the network will only result in improving the market share.

Ethical Considerations

The data (graph structure) for simulation results is accessed from publicly available repositories. There is no other ethical issue associated with our work.

Acknowledgements

References

Amelkin, V.; Bullo, F.; and Singh, A. K. 2017. Polar opinion dynamics in social networks. *IEEE Transactions on Automatic Control*, 62(11): 5650–5665.

Anagnostopoulos, A.; Becchetti, L.; Cruciani, E.; Pasquale, F.; and Rizzo, S. 2022. Biased opinion dynamics: when the devil is in the details. *Information Sciences*, 593: 49–63.

Armstrong, G.; and Kotler, P. 2003. *Marketing: an introduction*. Pearson Educación.

Beutel, A.; Prakash, B. A.; Rosenfeld, R.; and Faloutsos, C. 2012. Interacting viruses in networks: can both survive? In *ACM SIGKDD*.

Bru, R.; Canto, R.; Soto, R. L.; and Urbano, A. M. 2012. A Brauer's theorem and related results. *Central European Journal of Mathematics*, 10: 312–321.

Chen, C.; Tong, H.; Prakash, B. A.; Tsourakakis, C. E.; Eliassi-Rad, T.; Faloutsos, C.; and Chau, D. H. 2015. Node immunization on large graphs: Theory and algorithms. *IEEE Transactions on Knowledge and Data Engineering*, 28(1): 113–126.

Chen, W.; Yuan, Y.; and Zhang, L. 2010. Scalable influence maximization in social networks under the linear threshold model. In *IEEE international conference on data mining*, 88–97.

Contributor, G. 2024. What Are Influencer Pods? – Everything You Need to Know. https://grin.co/blog/influencer-pods/. Accessed: 2024-03-22.

Cooper, R. G.; and Kleinschmidt, E. J. 1987. New products: what separates winners from losers? *Journal of Product Innovation Management*, 4(3): 169–184.

Dalgic, T.; and Leeuw, M. 1994. Niche marketing revisited: concept, applications and some European cases. *European journal of marketing*, 28(4): 39–55.

Dandekar, P.; Goel, A.; and Lee, D. T. 2013. Biased assimilation, homophily, and the dynamics of polarization. *Proceedings of the National Academy of Sciences*, 110(15): 5791–5796.

Doshi, V.; Mallick, S.; and Eun, D. Y. 2021. Competing Epidemics on Graphs–Global Convergence and Coexistence. In *IEEE INFOCOM*.

FORCE11. 2020. The FAIR Data principles. https://force11.org/info/the-fair-data-principles/.

Freeman, L. C. 1978. Centrality in social networks conceptual clarification. *Social networks*, 1(3): 215–239.

French Jr, J. R. 1956. A formal theory of social power. *Psychological review*, 63(3): 181.

Friedkin, N. E. 1991. Theoretical foundations for centrality measures. *American journal of Sociology*, 96(6): 1478–1504.

Fuchs, C.; and Diamantopoulos, A. 2010. Evaluating the effectiveness of brand-positioning strategies from a consumer perspective. *European Journal of Marketing*, 44(11/12): 1763–1786.

Gallagher, S.; and Park, S. H. 2002. Innovation and competition in standard-based industries: A historical analysis of the US home video game market. *IEEE transactions on engineering management*, 49(1): 67–82.

Garey, M. R.; and Johnson, D. S. 1979. *Computers and intractability*, volume 174. Freeman San Francisco.

Gargiulo, F.; Lottini, S.; and Mazzoni, A. 2008. The saturation threshold of public opinion: are aggressive media campaigns always eective? In *Proceedings of ESSA 2008 Conference*.

Gebru, T.; Morgenstern, J.; Vecchione, B.; Vaughan, J. W.; Wallach, H.; Iii, H. D.; and Crawford, K. 2021. Datasheets for datasets. *Communications of the ACM*, 64(12): 86–92.

Granovetter, M. 1978. Threshold models of collective behavior. *American journal of sociology*, 83(6): 1420–1443.

Greenbaum, A.; Li, R.-c.; and Overton, M. L. 2020. First-order perturbation theory for eigenvalues and eigenvectors. *SIAM review*, 62(2): 463–482.

Irwin, D. A.; and Pavcnik, N. 2004. Airbus versus Boeing revisited: international competition in the aircraft market. *Journal of international economics*, 64(2): 223–245.

Jalkala, A. M.; and Keränen, J. 2014. Brand positioning strategies for industrial firms providing customer solutions. *Journal of Business & Industrial Marketing*, 29(3): 253–264.

Kalish, S. 1985. A new product adoption model with price, advertising, and uncertainty. *Management science*, 31(12): 1569–1585.

Kempe, D.; Kleinberg, J.; and Tardos, É. 2003. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD*, 137–146.

Lajmanovich, A.; and Yorke, J. A. 1976. A deterministic model for gonorrhea in a nonhomogeneous population. *Mathematical Biosciences*, 28(3-4): 221–236.

Leskovec, J.; and Krevl, A. 2014. SNAP Datasets: Stanford Large Network Dataset Collection. http://snap.stanford.edu/data. Accessed: 2022-02-20.

Li, C.; van de Bovenkamp, R.; and Van Mieghem, P. 2012. Susceptible-infected-susceptible model: A comparison of N-intertwined and heterogeneous mean-field approximations. *Physical Review E*, 86(2): 026116.

Liu, J.; Paré, P. E.; Nedić, A.; Tang, C. Y.; Beck, C. L.; and Başar, T. 2019. Analysis and control of a continuous-time bi-virus model. *IEEE Transactions on Automatic Control*, 64(12): 4891–4906.

Lusseau, D.; Schneider, K.; Boisseau, O. J.; Haase, P.; Slooten, E.; and Dawson, S. M. 2003. The bottlenose dolphin community of Doubtful Sound features a large proportion of long-lasting associations. *Behavioral Ecology and Sociobiology*, 54(4): 396–405.

McAuley, J. J.; and Leskovec, J. 2012. Learning to discover social circles in ego networks. In *NIPS*, volume 2012, 548–56.

Meyer, C. D. 2000. *Matrix Analysis and Applied Linear Algebra*, volume 71. SIAM.

Mobilia, M. 2015. Nonlinear q-voter model with inflexible zealots. *Physical Review E*, 92(1): 012803.

Montoya-Weiss, M. M.; and Calantone, R. 1994. Determinants of new product performance: A review and meta-analysis. *Journal of product innovation management*, 11(5): 397–417.

Mukhopadhyay, A.; Mazumdar, R. R.; and Roy, R. 2020. Voter and majority dynamics with biased and stubborn agents. *Journal of Statistical Physics*, 181: 1239–1265.

Newman, M. 2018. *Networks*. Oxford university press.

Nguyen, V. X.; Xiao, G.; Xu, X.-J.; Wu, Q.; and Xia, C.-Y. 2020. Dynamics of opinion formation under majority rules on complex social networks. *Scientific reports*, 10(1): 456.

Nic Healey. 2013. Battle of the ad campaign: PS4 vs Xbox One. https://www.cnet.com/tech/gaming/battle-of-the-ad-campaign-ps4-vs-xbox-one/. Accessed: 2023-01-12.

Omic, J.; and Van Mieghem, P. 2009. Epidemic spreading in networks—variance of the number of infected nodes. *Delft University of Technology, report20090707. http://www. nas. ewi. tudelft. nl/people/Piet/TUDelftReports*.

Pathak, N.; Banerjee, A.; and Srivastava, J. 2010. A generalized linear threshold model for multiple cascades. In *IEEE International Conference on Data Mining*, 965–970.

Prakash, B. A.; Beutel, A.; Rosenfeld, R.; and Faloutsos, C. 2012. Winner takes all: competing viruses or ideas on fair-play networks. In *ACM World Wide Web*.

Proskurnikov, A. V.; and Tempo, R. 2017. A tutorial on modeling and analysis of dynamic social networks. Part I. *Annual Reviews in Control*, 43: 65–79.

Proskurnikov, A. V.; and Tempo, R. 2018. A tutorial on modeling and analysis of dynamic social networks. Part II. *Annual Reviews in Control*, 45: 166–190.

Reddit. 2009. Reddit. https://www.reddit.com. Accessed: 2022-10-10.

Rossi, R. A.; and Ahmed, N. K. 2015. The Network Data Repository with Interactive Graph Analytics and Visualization. In *AAAI*.

Rozemberczki, B.; and Sarkar, R. 2020. Characteristic Functions on Graphs: Birds of a Feather, from Statistical Descriptors to Parametric Models. In *Proceedings of the 29th ACM International Conference on Information and Knowledge Management*, 1325–1334.

Sahneh, F. D.; and Scoglio, C. 2014. Competitive epidemic spreading over arbitrary multilayer networks. *Physical Review E*, 89(6): 062817.

Santos, A.; Moura, J. M.; and Xavier, J. M. 2015. Bi-virus SIS epidemics over networks: Qualitative analysis. *IEEE Transactions on Network Science and Engineering*, 2(1): 17–29.

Smith, H. L. 1995. *Monotone dynamical systems: an introduction to the theory of competitive and cooperative systems*. 41. American Mathematical Soc.

Van Mieghem, P.; Omic, J.; and Kooij, R. 2008. Virus spread in networks. *IEEE/ACM Transactions On Networking*, 17(1): 1–14.

Wang, Y.; Chakrabarti, D.; Wang, C.; and Faloutsos, C. 2003. Epidemic spreading in real networks: An eigenvalue viewpoint. In *IEEE Reliable Distributed Systems*.

Wasserman, S.; and Faust, K. 1994. Social network analysis: Methods and applications. *(No Title)*.

Wei, X.; Chen, S.; Wu, X.; Ning, D.; and Lu, J.-a. 2016. Cooperative spreading processes in multiplex networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 26(6).

Yang, H.-X.; Wang, W.-X.; Lai, Y.-C.; and Wang, B.-H. 2012. Convergence to global consensus in opinion dynamics under a nonlinear voter model. *Physics Letters A*, 376(4): 282–285.

Yang, L.-X.; Yang, X.; and Tang, Y. Y. 2017. A bi-virus competing spreading model with generic infection rates. *IEEE Transactions on Network Science and Engineering*, 5(1): 2–13.

Yang, W.-S.; Dia, J.-B.; Cheng, H.-C.; and Lin, H.-T. 2006. Mining social networks for targeted advertising. In *Proceedings of the 39th Annual Hawaii International Conference on System Sciences*, volume 6, 137a–137a.

Paper Checklist

1. For most authors...

   (a) Would answering this research question advance science without violating social contracts, such as violating privacy norms, perpetuating unfair profiling, exacerbating the socio-economic divide, or implying disrespect to societies or cultures? Yes, see the Section Broader Impact and the Ethical Considerations.

   (b) Do your main claims in the abstract and introduction accurately reflect the paper's contributions and scope? Yes

   (c) Do you clarify how the proposed methodological approach is appropriate for the claims made? Yes, see the subsection Choice of Modeling Approach and the Our Contributions.

   (d) Do you clarify what are possible artifacts in the data used, given population-specific distributions? Yes, see the Section Experimental Setup & Results.

   (e) Did you describe the limitations of your work? Yes, see the footnote 3 and the Section Conclusion & Future Work.

   (f) Did you discuss any potential negative societal impacts of your work? NA

   (g) Did you discuss any potential misuse of your work? NA

   (h) Did you describe steps taken to prevent or mitigate potential negative outcomes of the research, such as data and model documentation, data anonymization, responsible release, access control, and the reproducibility of findings? Yes, see the Section Ethical Consideration.

   (i) Have you read the ethics review guidelines and ensured that your paper conforms to them? Yes

2. Additionally, if your study involves hypotheses testing...

   (a) Did you clearly state the assumptions underlying all theoretical results? Yes, see the Section System Model and Problem Setup.

   (b) Have you provided justifications for all theoretical results? Yes, see the Section Analysis of Product 2's Market share $\bar{y}$ and the Appendix.

   (c) Did you discuss competing hypotheses or theories that might challenge or complement your theoretical results? NA

   (d) Have you considered alternative mechanisms or explanations that might account for the same outcomes observed in your study? Yes, see the subsection Comparison with Baseline Methodologies.

   (e) Did you address potential biases or limitations in your theoretical framework? Yes, see the footnote 3 and the Section Conclusion & Future Work.

   (f) Have you related your theoretical results to the existing literature in social science? NA

   (g) Did you discuss the implications of your theoretical results for policy, practice, or further research in the social science domain? Yes, see the Section Broader Impact.

3. Additionally, if you are including theoretical proofs...

   (a) Did you state the full set of assumptions of all theoretical results? Yes

   (b) Did you include complete proofs of all theoretical results? Yes, see the Appendix.

4. Additionally, if you ran machine learning experiments...

(a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? NA

(b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? NA

(c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? NA

(d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? NA

(e) Do you justify how the proposed evaluation is sufficient and appropriate to the claims made? NA

(f) Do you discuss what is "the cost" of misclassification and fault (in)tolerance? NA

5. Additionally, if you are using existing assets (e.g., code, data, models) or curating/releasing new assets, **without compromising anonymity**...

(a) If your work uses existing assets, did you cite the creators? Yes, see the subsection Simulation Setup.

(b) Did you mention the license of the assets? NA

(c) Did you include any new assets in the supplemental material or as a URL? NA

(d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? Yes, see the Section Ethical Considerations.

(e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? No, it does not contain as the data is a graph dataset with nodes represented as numbers.

(f) If you are curating or releasing new datasets, did you discuss how you intend to make your datasets FAIR (see FORCE11 (2020))? NA

(g) If you are curating or releasing new datasets, did you create a Datasheet for the Dataset (see Gebru et al. (2021))? NA

6. Additionally, if you used crowdsourcing or conducted research with human subjects, **without compromising anonymity**...

(a) Did you include the full text of instructions given to participants and screenshots? NA

(b) Did you describe any potential participant risks, with mentions of Institutional Review Board (IRB) approvals? NA

(c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? NA

(d) Did you discuss how data is stored, shared, and deidentified? NA

A   Appendix

Proof of Proposition 3.1

Proof. Let $(\gamma, \mathbf{u})$ and $(\hat{\gamma}, \hat{\mathbf{u}})$ be such that we have $\gamma u_i u_j \le \hat{\gamma}\hat{u}_i\hat{u}_j$, with strict inequality for at least one pair of i, j. Let

$\mathbf{y}, \mathbf{z}$ be the resulting fixed point for Product 2 under parameters $(\gamma, \mathbf{u})$, $(\hat{\gamma}, \hat{\mathbf{u}}))$, respectively. We show that $\bar{y} < \bar{z}$, that is, it is always better to spend more if the budget allows for it. Let $(\mathbf{x}, \mathbf{y})$ and $(\hat{\mathbf{x}}(t), \hat{\mathbf{y}}(t))$ denote the fixed points of the ODE system (2) with parameters $(\gamma, \mathbf{u})$ and $(\hat{\gamma}, \hat{\mathbf{u}})$, respectively, where the initial point $(\hat{\mathbf{x}}(0), \hat{\mathbf{y}}(0)) = (\mathbf{x}, \mathbf{y})$. Then, we have[11]

$$\left.\frac{d\hat{\mathbf{x}}(t)}{dt}\right|_{t=0} = \beta_1 \text{diag}(\mathbf{1}-\mathbf{x}-\mathbf{y}(t))\mathbf{A}\mathbf{x} - \delta_1\mathbf{x} = 0$$

$$\left.\frac{d\hat{\mathbf{y}}(t)}{dt}\right|_{t=0} = \beta_2 \text{diag}(\mathbf{1}-\mathbf{x}-\mathbf{y})\left[\mathbf{A} + \hat{\gamma}\hat{\mathbf{u}}\hat{\mathbf{u}}^T\right]\mathbf{y} - \delta_2\mathbf{y}$$

$$> \beta_2 \text{diag}(\mathbf{1}-\mathbf{x}-\mathbf{y})\left[\mathbf{A} + \gamma\mathbf{u}\mathbf{u}^T\right]\mathbf{y} - \delta_2\mathbf{y} = 0.$$

Therefore, there exists an $\epsilon > 0$ for which we have $\mathbf{x} \ge \hat{\mathbf{x}}(\epsilon)$ and $\mathbf{y} \le \hat{\mathbf{y}}(\epsilon)$. From (Doshi, Mallick, and Eun 2021), we know that the bi-SIS system is strongly monotone (Smith 1995) for $\mathbf{x}, \mathbf{y} \in (0,1)^N$. Hence, $\mathbf{y} < \hat{\mathbf{y}}(\epsilon) \ll \mathbf{y}(t + \epsilon)$ for all $t > 0$, and since $\mathbf{z} = \lim_{t\to\infty} \mathbf{y}(t+\epsilon)$, we have $\mathbf{y} \ll \mathbf{z}$. □

Proof of Lemma 3.2

Proof. Consider the matrix $\mathbf{M} = \mathbf{S}_{\mathbf{x}^*}(\mathbf{A}+\gamma\mathbf{u}\mathbf{u}^T)$. By substituting $\mathbf{u} = \mathbf{S}_{\mathbf{x}^*}^{-1}\mathbf{v}$,[12] it can be written as

$$\mathbf{M} = \mathbf{S}_{\mathbf{x}^*}(\mathbf{A} + \gamma\mathbf{S}_{\mathbf{x}^*}^{-1}\mathbf{v}\mathbf{u}^T) = \mathbf{S}_{\mathbf{x}^*}\mathbf{A} + \gamma\mathbf{v}\mathbf{u}^T,$$

which is now in the form of a rank-one perturbation to the matrix $\hat{\mathbf{M}} = \mathbf{S}_{\mathbf{x}^*}\mathbf{A}$, and is in the same form as in Proposition 1.1 in (Bru et al. 2012) with $\mathbf{v}$ as the first eigenvector of $\hat{\mathbf{M}}$, and $\gamma\mathbf{u}$ being the additional N−dimensional vector. Then, from Proposition 1.1 in (Bru et al. 2012), $\mathbf{v}$ is also the first eigenvector of $\mathbf{M} = \mathbf{S}_{\mathbf{x}^*}(\mathbf{A} + \gamma\mathbf{u}\mathbf{u}^T)$ associated with eigenvalue $\lambda(\mathbf{M}) = \lambda(\mathbf{S}_{\mathbf{x}^*}\mathbf{A}) + \gamma\mathbf{u}^T\mathbf{v} = \lambda(\mathbf{S}_{\mathbf{x}^*}\mathbf{A}) + \gamma\mathbf{u}^T\mathbf{S}_{\mathbf{x}^*}\mathbf{u}$. Finally, by rearranging the above terms, we can check that the value of γ for which $\lambda(\mathbf{M}) = 1/\tau_2$ is given by $\gamma = \frac{1/\tau_2 - \lambda(\mathbf{S}_{\mathbf{x}^*}\mathbf{A})}{\mathbf{u}^T\mathbf{S}_{\mathbf{x}^*}\mathbf{u}}$. □

Proof of Theorem 3.3

We begin by taking the derivative of (4) with respect to γ and u_r evaluated at $(\gamma^c, \mathbf{u}^c)$ which yields[13]

$$\left.\frac{\partial y_i}{\partial \gamma}\right|_{c.p.} = \tau_2(1-x_i^*)\sum_{j\in\mathcal{N}}(a_{ij} + \gamma^c u_i^c u_j^c)\left.\frac{\partial y_i}{\partial \gamma}\right|_{c.p.} \quad (6)$$

$$\left.\frac{\partial y_i}{\partial u_r}\right|_{c.p.} = \tau_2(1-x_i^*)\sum_{j\in\mathcal{N}}(a_{ij} + \gamma u_i^c u_j^c)\left.\frac{\partial y_i}{\partial u_r}\right|_{c.p.}, \quad (7)$$

---

[11] For any two vectors $\mathbf{p} = [p_i]$ and $\mathbf{q} = [q_i]$, the inequality $\mathbf{p} > \mathbf{q}$ implies $p_i \ge q_i$, with strict inequality for at least one entry.

[12] The vector $\mathbf{u}$ can actually be taken as any scalar multiple of $\mathbf{S}_{\mathbf{x}^*}^{-1}\mathbf{v}$, that is $\mathbf{u} = c\mathbf{S}_{\mathbf{x}^*}^{-1}\mathbf{v}$ for any $c \in \mathbb{R}$. We drop the 'c' notation for clarity.

[13] We define *c.p.* to refer to the point at which we take derivative of (4), i.e., $c.p. \triangleq (\gamma, \mathbf{u}) = (\gamma^c, \mathbf{u}^c)$.

where $\partial\mathbf{y}/\partial\gamma = [\partial y/\partial\gamma]_i$ and $\partial\mathbf{y}/\partial u_r = [\partial y/\partial u_r]_i$ are eigenvectors of $\mathbf{S}_{\mathbf{x}^*}[\mathbf{A} + \gamma\mathbf{u}\mathbf{u}^T]$ up to some multiplicative constants (PF theorem (Meyer 2000)) corresponding to eigenvalue τ_2. We first give the following result, to help us derive Lemma A.2, which in turn helps in determining the multiplicative constants.

Lemma A.1. *Let $\epsilon_1, \epsilon_2 > 0$ be any small constants such that $\gamma(\epsilon_1)$ and $\mathbf{u}(\epsilon_2)$ are perturbations around $\gamma, \mathbf{u}$. Then, for all $\epsilon_3 > 0$ such that $\lambda(\gamma, \mathbf{u}) = \lambda(\gamma^c, \mathbf{u}^c) + \epsilon_3$, there exists $\epsilon_4 > 0$ such that $\mathbf{y} = \epsilon_4 \mathbf{v}(\gamma, \mathbf{u})$.*

Proof. Since (6) is an eigenvalue problem of the form $\mathbf{A}\nu = \lambda\nu$, it allows for $\left.\frac{dy_i}{d\gamma}\right|_{\gamma=\gamma^c} > 0$ as $\lambda(\mathbf{S}_{\mathbf{x}^*}[\mathbf{A} + \gamma\mathbf{u}\mathbf{u}^T])|_{\gamma=\gamma^c} = 1/\tau_2$. Similarly, it can be shown that $\left.\frac{dy_i}{du_i}\right|_{u_i=u_i^c} > 0$.

From (6), we have that $1/\tau_2 = \lambda(\gamma^c, \mathbf{u}^c) > 0$ and $\frac{dy_i}{d\gamma} \triangleq [v]_i$, i.e., $\mathbf{y} = \epsilon\mathbf{v}(\gamma, \mathbf{u})$ is the eigenvector of $\mathbf{S}_{\mathbf{x}^*}[\mathbf{A} + \gamma\mathbf{u}\mathbf{u}^T]$. For $\lambda(\gamma^c, \mathbf{u}^c) > \lambda(\gamma, \mathbf{u})$, $1/\tau_2$ cannot be the eigenvalue and only possible solution is $[v]_i = 0$. Hence, for $\lambda(\gamma, \mathbf{u}) > \lambda(\gamma^c, \mathbf{u}^c)$, i.e., for any small arbitrary constant ϵ, $\lambda(\gamma, \mathbf{u}) = \lambda(\gamma^c, \mathbf{u}^c) + \epsilon$, $1/\tau_2$ is the eigenvalue with $\mathbf{y} = \epsilon\mathbf{v}(\gamma, \mathbf{u})$. $\square$

Now, the following lemma helps in determining the constants.

Lemma A.2. *Let $\lambda(\gamma, \mathbf{u}) \triangleq \lambda(\mathbf{S}_{\mathbf{x}^*}[\mathbf{A} + \gamma\mathbf{u}\mathbf{u}^T])$, and $(\gamma^c, \mathbf{u}^c)$ be any pair of critical parameters, and let $\bar{\mathbf{y}}(\gamma, \mathbf{u}) \triangleq \mathbf{1}^T\mathbf{y}(\gamma\mathbf{u})/N$ denote the market share as a function of any feasible $(\gamma, \mathbf{u})$. Then, $\bar{\mathbf{y}}(\gamma, \mathbf{u})$ can be written as,*

$$\bar{\mathbf{y}}(\gamma, \mathbf{u}) = \frac{\tau_2 \mathbf{1}^T\mathbf{v}(\gamma^c, \mathbf{u}^c)}{N\sum_{i=1}^{N}\frac{v_i(\gamma^c, u_i^c)^3}{(1-x_i^*)^2}}\left(\lambda(\gamma, \mathbf{u}) - 1/\tau_2\right)$$
$$+ O\left(\left(\lambda(\gamma, \mathbf{u}) - 1/\tau_2\right)^2\right). \tag{8}$$

Proof. From (3) and (4), we have the following expressions and then Proposition A.3 regarding orthogonality of right eigenvectors.

$$\frac{\mathbf{1}^T\mathbf{y}}{\tau_2} = (\mathbf{1} - \mathbf{x})^T[\mathbf{A} + \gamma\mathbf{u}\mathbf{u}^T]\mathbf{y} - \mathbf{y}^T[\mathbf{A} + \gamma\mathbf{u}\mathbf{u}^T]\mathbf{y} \quad (9)$$

$$\frac{\mathbf{y}^T\mathbf{S}_{\mathbf{x}}^{-1}\mathbf{y}}{\tau_2} = \mathbf{y}^T[\mathbf{A} + \gamma\mathbf{u}\mathbf{u}^T]\mathbf{y}$$
$$- \mathbf{y}^T\mathbf{S}_{\mathbf{x}}^{-1}\text{diag}(\mathbf{y})[\mathbf{A} + \gamma\mathbf{u}\mathbf{u}^T]\mathbf{y} \tag{10}$$

Proposition A.3. *For all $k \neq j$, the right eigenvectors $\mathbf{v}_k$ of $\mathbf{M} = \mathbf{S}_{\mathbf{x}^*}[A + \gamma\mathbf{u}\mathbf{u}^T]$ satisfy $\mathbf{v}_k^T\mathbf{S}_{\mathbf{x}}^{-1}\mathbf{v}_j = 0$.*

Proof. Let, $\mathbf{S}_{\mathbf{x}}^{-1/2}\mathbf{M}\mathbf{S}_{\mathbf{x}^*}^{1/2} = \mathbf{S}_{\mathbf{x}^*}^{1/2}[\mathbf{A} + \gamma\mathbf{u}\mathbf{u}^T]\mathbf{S}_{\mathbf{x}^*}^{1/2} \triangleq \hat{\mathbf{M}}$ hence, $\mathbf{M}$ and $\hat{\mathbf{M}}$ are similar matrices with same set of eigenvalues (Meyer 2000), with right eigenvector of $\hat{\mathbf{M}}$ corresponding to eigenvalue λ_k given by $\mathbf{S}_{\mathbf{x}^*}^{-1/2}\mathbf{v}_k$. $\hat{\mathbf{M}}$ is a symmetric matrix as $\mathbf{A} + \gamma\mathbf{u}\mathbf{u}^T$ is symmetric, and its eigenvectors are orthogonal to each other. Hence, for all $k \neq j$, we have $\mathbf{v}_k^T\mathbf{S}_{\mathbf{x}^*}^{-1/2}\mathbf{S}_{\mathbf{x}^*}^{1/2}\mathbf{v}_j = \mathbf{v}_k^T\mathbf{S}_{\mathbf{x}}^{-1}\mathbf{v}_j = 0$. $\square$

From $\mathbf{y}$ as a linear combination of eigenvectors, we have

$$\mathbf{y} = \sum_{k=1}^{N}c_k\mathbf{v}_k. \tag{11}$$

From the above equation, we have

$$\frac{\mathbf{1}^T\mathbf{y}}{\tau_2} = \sum_{k=1}^{N}\tau_2^{-1}c_k\mathbf{1}\mathbf{v}_k, \qquad \frac{\mathbf{y}^T\mathbf{S}_{\mathbf{x}}^{-1}\mathbf{y}}{\tau_2} = \sum_{k=1}^{N}\tau_2^{-1}c_k^2.$$
$$(12) \qquad\qquad\qquad\qquad (13)$$

Substituting (11) in (9) and in (10) gives

$$\frac{\mathbf{1}^T\mathbf{y}}{\tau_2} = \sum_{k=1}^{N}c_k\lambda_k\mathbf{1}^T\mathbf{v}_k - c_k^2\lambda_k. \tag{14}$$

$$\frac{\mathbf{y}^T\mathbf{S}_{\mathbf{x}}^{-1}\mathbf{y}}{\tau_2} = \sum_{j=1}^{N}c_j^2\lambda_j$$
$$- \sum_{j=1}^{N}\sum_{k=1}^{N}\sum_{l=1}^{N}c_jc_kc_l\lambda_l\sum_{i=1}^{N}\frac{[\mathbf{v}_j]_i[\mathbf{v}_k]_i[\mathbf{v}_k]_i}{(1-x_i)^2} \tag{15}$$

By equating (12) with (14), (13) with (15) yields

$$\sum_{k=1}^{N}(\lambda_k - \tau_2^{-1})c_k\mathbf{1}^T\mathbf{v}_k = \sum_{k=1}^{N}c_k^2\lambda_k \tag{16}$$

$$\sum_{j=1}^{N}(\lambda_j - \tau_2^{-1})c_j^2 = \sum_{j=1}^{N}\sum_{k=1}^{N}\sum_{l=1}^{N}c_jc_kc_l\lambda_l\sum_{i=1}^{N}\frac{[\mathbf{v}_j]_i[\mathbf{v}_k]_i[\mathbf{v}_k]_i}{(1-x_i)^2} \tag{17}$$

From Lemma A.1, we know $\lambda(\gamma, \mathbf{u}) = \lambda(\gamma^c, \mathbf{u}^c) + \epsilon_1$ for some $\epsilon_1 > 0$, there exists an $\epsilon_2 > 0$ such that $\mathbf{y} = \epsilon_2\mathbf{v}(\gamma^c, \mathbf{u}^c)$. Also, $\mathbf{y} = \alpha\mathbf{v}(\gamma^c, \mathbf{u}^c) + \beta\mathbf{w}$, where $\mathbf{w}$ is a vector orthogonal to $\mathbf{v}(\gamma, \mathbf{u})$, and β goes down to zero faster than α as $(\gamma, \mathbf{u}) \to (\gamma^c, \mathbf{u}^c)$, or $\lambda(\gamma, \mathbf{u}) \to \lambda(\gamma^c, \mathbf{u}^c)\tau_2^{-1}$. Let $\alpha = \alpha_0(\lambda_1 - \lambda_1^c)^p + o((\lambda_1 - \lambda_1^c)^p)$ and $\beta = \beta_0(\lambda_1 - \lambda_1^c)^q + o((\lambda_1 - \lambda_1^c)^q)$, where $q > p > 0$. From (11), $\mathbf{y}$'s component c_1 and c_k (for all $k > 1$) is associated with $\mathbf{v}$ and vectors orthogonal to $\mathbf{v}$, respectively. Then, we have $c_1 = r_1(\lambda_1 - \lambda_1^c)^p + o((\lambda_1 - \lambda_1^c)^p)$, and $c_k = r_k(\lambda_1 - \lambda_1^c)^q + o((\lambda_1 - \lambda_1^c)^q)$. Substituting these in (16) when $\lambda_1 \to \lambda_1^c$ shows that it is of the order,

$$\sum_{k=2}^{N}r_k\mathbf{1}^T\mathbf{v}_k(\lambda_k - \tau_2^{-1})(\lambda_1 - \lambda_1^c)^q + r_1\mathbf{1}^T\mathbf{v}_1(\lambda_1 - \lambda_1^c)^{p+1}$$

$+ o\left((\lambda_1 - \lambda_1^c)^{\min(q,p+1)}\right)$, and $\sum_{k=2}^{N}r_k^2\lambda_k(\lambda_1 - \lambda_1^c)^{2q} +$
$r_1^2\lambda_1(\lambda_1 - \lambda_1^c)^{2p} + o\left((\lambda_1 - \lambda_1^c)^{2p}\right)$.

Similarly (17) gives,

$$r_1^2(\lambda_1 - \lambda_1^c)^{2p+1} + \sum_{k=2}^{N}r_k^2(\lambda_k - \tau_2^{-1})(\lambda_1 - \lambda_1^c)^{2q}$$

$$+ o\left((\lambda_1 - \lambda_1^c)^{\min(2q,2p+1)}\right),$$

$$r_1^3\lambda_1(\lambda_1 - \lambda_1^c)^{3p}\sum_{i=1}^{N}\frac{v_{1i}}{(1-x_i)^2} + o((\lambda_1 - \lambda_1^c)^{3p})$$

$$+ O((\lambda_1 - \lambda_1^c)^{2p+q}).$$

From these we have the following cases: *Case A: $q \geq p + 1$*, which means that $p + q = 2p$ implying $p = 1$; *Case B: $q < p + 1$*, implying $q = 2p$; *Case C: $2q \geq 2p + 1$*, in which case $2p + 1 = 3p$ and $2q = 2p + q$, implying $p = 1$ and $q = 2$; *Case D: $2q < 2p + 1$*, which gives us $2q = 3p$ and $2p + 1 = 2p + q$. Observe that Cases A and C are the only ones complementing each other, while the others pairs are contradictory. Hence, it's true that $p = 1$ and $q = 2$. By equating the corresponding powers allows us to rewrite (11) as

$$\mathbf{y} = \frac{\mathbf{v}(\gamma, \mathbf{u})}{\lambda_1(\gamma, \mathbf{u}) \sum\limits_{i=1}^{N} \frac{v_i(\gamma, \mathbf{u})}{(1 - x_i(\gamma, \mathbf{u}))^2}} (\lambda_1(\gamma, \mathbf{u}) - \lambda_1(\gamma^c, \mathbf{u}^c))$$

$$+ O\left((\lambda_1(\gamma, \mathbf{u}) - \lambda_1(\gamma^c, \mathbf{u}^c))^2\right).$$

A Taylor series expansion with respect to $\lambda_1(\gamma, \mathbf{u})$ and centred around $\lambda_1(\gamma^c, \mathbf{u}^c) = \tau_2^{-1}$, gives us[14]

$$\mathbf{y} = \frac{\tau_2 \mathbf{v}(\gamma^c, \mathbf{u}^c)}{\sum\limits_{i=1}^{N} \frac{v_i(\gamma^c, \mathbf{u}^c)}{(1 - x_i^*)^2}} \left(\lambda_1(\gamma, \mathbf{u}) - \tau_2^{-1}\right)$$

$$+ O\left((\lambda_1(\gamma, \mathbf{u}) - \tau_2^{-1})^2\right).$$

□

Now, we use the results from Lemma A.2 towards the proof of Theorem 3.3, which is as follows. Note that $\lambda(\gamma, \mathbf{u}) - 1/\tau_2$ on the RHS in Lemma A.2 is actually based on given $(\gamma^c, \mathbf{u}^c)$ at the critical point, i.e., $\mathbf{x} = \mathbf{x}^*, \mathbf{y} = \mathbf{0}$.

Proof. From first-order approximation, we know that $f(x) \approx f(a) + f'(a)(x - a)$. Utilizing this, $\lambda(\gamma^c, \mathbf{u})$ with respect to change in u_i based on given $\lambda(\gamma^c, \mathbf{u}^c)$ at critical point, we have

$$\lambda(\gamma, \mathbf{u}) \approx \lambda(\gamma^c, \mathbf{u}^c) + \lambda'(\gamma^c, \mathbf{u}^c) \Delta \mathbf{u}. \qquad (18)$$

Utilizing Theorem 1 from (Greenbaum, Li, and Overton 2020), we know that at $u_i = u_i^c$, $\lambda'(\gamma^c, \mathbf{u}^c) = \mathbf{q}^* \mathbf{M}' \mathbf{p}$ where $\lambda'(\gamma^c, \mathbf{u}^c)$ and $\mathbf{M}'$ are, respectively, the derivatives of $\lambda(\gamma, \mathbf{u})$ and $\mathbf{M}$ at $u_i = u_i^c$. Also, $\mathbf{M} = \mathbf{S}_{\mathbf{x}^*}(\mathbf{A} + \gamma \mathbf{u}\mathbf{u}^T)$ with largest eigenvalue $\lambda(\gamma^c, \mathbf{u}^c)$ corresponding to right eigenvector $\mathbf{p} = \mathbf{v}$ and left eigenvector $\mathbf{q} = \mathbf{v}^T \mathbf{S}_{\mathbf{x}^*}^{-1}$. Substituting this in (18) and from Lemma 3.2, we have

$$\lambda(\gamma^c, \mathbf{u}) - \frac{1}{\tau_2} = \mathbf{q}^* \mathbf{M}' \mathbf{p} \Delta \mathbf{u}, \qquad (19)$$

where $\Delta \mathbf{u} = u_i - u_i^c$. The derivative of $\mathbf{M}$ with respect to u_i gives $\frac{\partial \mathbf{M}}{\partial u_i} = \gamma^c \mathbf{S}_{\mathbf{x}^*} \left[e_i \mathbf{u}^T + \mathbf{u} e_i^T\right]$.

This gives us $\mathbf{M}' = \gamma \mathbf{S}_{\mathbf{x}^*} \left[e_i \mathbf{u}^T + \mathbf{u} e_i^T\right]$ at u_i. Substituting this on the RHS of (19) gives $\mathbf{q}^* \mathbf{M}' \mathbf{p} \Delta \mathbf{u} = 2\gamma^c \mathbf{u}^T \mathbf{v}(u_i - u_i^c)v_i$, where the term $2\gamma^c \mathbf{u}^T$ always remains constant. Substituting this again in (19) gives

$$\lambda(\gamma^c, \mathbf{u}) - \frac{1}{\tau_2} = 2\gamma^c \mathbf{u}^T \mathbf{v}(u_i - u_i^c)v_i, \qquad (20)$$

which implies that the change is in the direction of $\mathbf{v} = [v_i]$ for all $i \in \mathcal{N}$. □

---

[14] $\lambda_1(\gamma, \mathbf{u})$ corresponds to the first eigenvalue from PF Theorem and is equivalent to $\lambda(\gamma, \mathbf{u})$.