

ABSTRACT

Title of dissertation: **DIFFUSION DYNAMICS IN
INTERCONNECTED COMMUNITIES**

Xiaoya Wei, Doctor of Philosophy, 2015

Dissertation directed by: **Professor Eyad H. Abed**
Dept. of Electrical & Computer Engineering

In this dissertation, multi-community-based Susceptible-Infected-Recovered (SIR) and Susceptible-Infected-Susceptible (SIS) models of infection/innovation diffusion are introduced for heterogeneous social networks in which agents are viewed as belonging to one of a finite number of communities. Agents are assumed to have well-mixed interactions within and between communities. The communities are connected through a backbone graph which defines an overall network structure for the models. The models are used to determine conditions for outbreak of an initial infection. The role of the strengths of the connections between communities in the development of an outbreak as well as long-term behavior of the diffusion is also studied. Percolation theory is brought to bear on these questions as an independent approach separate from the main dynamic multi-community modeling approach of the dissertation. Results obtained using both approaches are compared and found to be in agreement in the limit of infinitely large populations in all communities.

Based on the proposed models, three classes of marketing problems are formulated and studied: referral marketing, seeding marketing and dynamic marketing. It

is found that referral marketing can be optimized relatively easily because the associated optimization problem can be formulated as a convex optimization. Also, both seeding marketing and dynamic marketing are shown to enjoy a useful property, namely “continuous monotone submodularity.” Based on this property, a greedy heuristic is proposed which yields solutions with approximation ratio no less than $1-1/e$. Also, dynamic marketing for SIS models is reformulated into an equivalent convex optimization to obtain an optimal solution. Both cost minimization and trade-off of cost and profit are analyzed.

Next, the proposed modeling framework is applied to study competition of multiple companies in marketing of similar products. Marketing of two classes of such products are considered, namely marketing of durable consumer goods (DCG) and fast-moving consumer goods (FMCG). It is shown that an epsilon-equilibrium exists in the DCG marketing game and a pure Nash equilibrium exists in the FMCG marketing game. The Price of Anarchy (PoA) in both marketing games is found to be bounded by 2. Also, it is shown that any two Nash equilibria for the FMCG marketing game agree almost everywhere, and a distributed algorithm converging to the Nash equilibrium is designed for the FMCG marketing game.

Finally, a preliminary investigation is carried out to explore possible concepts of network centrality for diffusions. In a diffusion process, the centrality of a node should reflect the influence that the node has on the network over time. Among the preliminary observations in this work, it is found that when an infection does not break out, diffusion centrality is closely related to Katz centrality; when an infection does break out, diffusion centrality is closely related to eigenvector centrality.

DIFFUSION DYNAMICS IN INTERCONNECTED
COMMUNITIES

by

Xiaoya Wei

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2015

Advisory Committee:

Professor Eyad H. Abed, Chair/Advisor

Professor Nuno C. Martins

Professor Richard J. La

Professor K. J. Ray Liu

Professor Balakumar Balachandran, Dean's Representative

© Copyright by
Xiaoya Wei
2015

Dedication

For my mom.

Acknowledgments

First of all, I would like to express my sincere appreciation to my advisor Professor Eyad H. Abed for his supervision, patience and support. Without his guidance this dissertation would have not been possible. I'll miss our numerous meetings that broadened my vision, shaped my ideas and refined my research. He showed me not only how to do good research but also how to be a good man. I will always cherish my experience as one of his graduate students.

I would also like thank Professor André L. Tits. He is not only a great teacher but also a nice friend. I learned a lot from his lectures which are inspiring, insightful and technically rigorous. He provided me with a lot of valuable advice when I just started my studies in University of Maryland. I also had a lot of fun chatting with him in coffee hours.

My sincere thanks go to my parents for their support and dedication through my life. They nurtured me by setting role models when I grew up, giving me comforts when I felt weak and encouraging me to be brave and independent all the time. I view these as the most valuable treasure that I could have. There is no way that I can fully express my gratitude towards them.

I would also like to thank all my friends that I have ever met. It is their friendship that made my life colorful. I also want to thank Dr. Shan Liu for his guidance in my undergraduate stage. Finally, I am grateful for partial support of my research under a Multidisciplinary University Research Initiative (MURI) grant from the Air Force Office of Scientific Research.

Table of Contents

List of Figures	viii
List of Abbreviations	ix
1 Introduction	1
1.1 Modeling of Diffusion Dynamics	4
1.2 Optimal Social Network Marketing	9
1.3 Social Network Marketing Game	12
1.4 Diffusion Centrality	16
1.5 Organization of the Dissertation	18
2 Modeling and Analysis	20
2.1 Modeling Diffusion Among Communities	20
2.2 Accuracy of the Community-Based Approximation	28
2.2.1 Theoretical Analysis	28
2.2.2 Numerical Simulation	31
2.3 Diffusion Outbreak	33
2.3.1 Condition for Outbreak	33
2.3.2 Effect of Inter-Community Connections	37
2.3.3 Weak Interconnections: A Special Case	38
2.4 Final State of the Diffusion	40
2.4.1 Final State	40
2.4.2 Outbreak Impact Within Communities	41
2.4.3 Bounds on the Final State	43
2.5 A View from Percolation Theory	45
2.5.1 Transmissibility and Percolation	47
2.5.2 Solution to the Percolation Problem	48
2.5.3 Condition for Infection Outbreak	51
2.5.4 Final State of the Diffusion	55

3	Optimal Social Network Marketing	57
3.1	Problem Formulation	58
3.1.1	Referral Marketing	58
3.1.2	Seeding Marketing	60
3.1.3	Dynamic Marketing	62
3.2	Continuous Monotone Submodularity	63
3.3	Greedy Heuristic	66
3.3.1	General Greedy Algorithm	67
3.3.2	Application to Social Network Marketing	68
3.4	Two Variants of Dynamic Marketing	71
3.4.1	Minimum Cost Marketing	72
3.4.2	Trade-Off Between Profit and Cost	73
3.5	The Convexity of $\phi_T^{(3)}(\cdot)$ (for SIS model (2.6))	76
3.6	Numerical Experiments	77
3.6.1	Seeding Marketing	77
3.6.2	Dynamic Marketing	78
4	Social Network Marketing Games	82
4.1	DCG Marketing Game	83
4.1.1	Continuous Monotone Submodularity	87
4.1.2	Competitiveness	87
4.2	FMCG Marketing Game	88
4.2.1	Competitiveness	91
4.2.2	Continuous Monotone Submodularity	91
4.2.3	Concavity-Convexity	91
4.3	Existence of Nash Equilibrium	93
4.4	Price of Anarchy	94
4.5	Almost Uniqueness of the Nash Equilibrium	95
4.6	Distributive Algorithm for Seeking a Nash Equilibrium	96
5	Diffusion Centrality	99
5.1	Background and Related Work	99
5.1.1	Degree Centrality	100
5.1.2	Eigenvector Centrality	100
5.1.3	Katz Centrality	101
5.2	Sensitivity Centrality	102
5.3	Vulnerability Centrality	105
6	Conclusions and Suggestions for Future Work	108
6.1	Main Contributions of the Dissertation	108
6.2	Suggestions for Future Work	112
A	Derivation of the Agent-Based SIR Model	115

B	Proofs of Several Theorems	120
B.1	Proof for Theorem 3.1	120
B.2	Proof for Lemma 3.1	120
B.3	Proof for Lemma 3.2	121
B.4	Proof for Theorem 3.2	122
	B.4.1 Continuous Monotone Submodularity of $\phi^{(1)}(\cdot)$	122
	B.4.2 Continuous Monotone Submodularity of $-\phi_T^{(2)}$	127
	B.4.3 Continuous Monotone Submodularity of $-\phi_T^{(3)}$	132
B.5	Proof for Theorem 3.3	135
B.6	Proof for Theorem 3.3	136
B.7	Proof for Theorem 3.4	138
B.8	Proof for Theorem 3.5	141
B.9	Proof for Theorem 3.6	142
B.10	Proof for Theorem 4.1	144
B.11	Proof for Theorem 4.2	145
B.12	Proof for Theorem 4.3	146
B.13	Proof for Theorem 4.4	146
B.14	Proof for Theorem 4.5	148
B.15	Proof for Theorem 4.6	149
	B.15.1 The DCG Marketing Game	149
	B.15.2 The FMCG Marketing Game	152
B.16	Proof for Theorem 4.7	153
B.17	Proof for Theorem 4.8	160
B.18	Proof for Theorem 4.9	165
	Bibliography	172

List of Figures

2.1	Illustrating a sample community-based social network structure	23
2.2	Numerical simulation of infection's diffusion	32
3.1	Numerical optimal and greedy solution: case 1	80
3.2	Numerical optimal and greedy solution: case 2	80
3.3	Numerical optimal and greedy solution: case 3	81
A.1	Markov chain state transition diagram for a network of size 2	116

List of Abbreviations

$s_j(t)$	Percentage of susceptible population in community j at time t
$i_j(t)$	Percentage of infected population in community j at time t
$r_j(t)$	Percentage of recovered population in community j at time t
$u_j(t)$	Strength of external influence in community j at time t
n	Number of communities
λ_{jk}	Infection rate from community k to community j
γ_j	Recovery rate of community j
$\phi^{(1)}(\cdot)$	Object function in seeding marketing
$\phi^{(2)}(\cdot)$	Object function in dynamic marketing for SIR model
$\phi^{(3)}(\cdot)$	Object function in dynamic marketing for SIS model
m	Number of players in marketing game
$\mathbb{U}(\cdot)$	Combination of all players' strategy in marketing game
$\zeta_k(\mathbb{U}(\cdot))$	Utility for player k in marketing game
$\eta(\mathbb{U}(\cdot))$	Social welfare in marketing game, i.e. $\eta(\mathbb{U}(\cdot)) = \sum_{k=1}^m \zeta_k(\mathbb{U}(\cdot))$
$\mathbb{U}_{-k}(\cdot)$	Strategy combination of the k -th player's competitors
\mathbf{I}_n	Identity $n \times n$ matrix
SIR	Susceptible-Infected-Recovered
SIS	Susceptible-Infected-Susceptible
DCG	Durable Consumer Goods
FMCG	Fast-Moving Consumer Goods
PoA	Price of Anarchy

Chapter 1: Introduction

Social networks are a powerful means for the diffusion of ideas, information and products. With the advent of electronic media, the speed and scale of social network interactions has grown immensely. However, social networks are as old as civilization, since interactions also occur through word-of-mouth communications and other interactions. Through interactions with their circle of acquaintances, individuals' choices may be influenced toward deciding to adopt an idea or product; of course, it can also happen that the influence factors are not sufficiently strong or convincing to result in significant adoption as the idea propagates through a population. Decisions in a social network about whether or not to adopt an idea or product are primarily influenced by two types of sources, namely individual-to-individual interactions and external influences [1].

External influences can come from news media, changes in legal regulations, etc. A typical example of an external influence is commercial advertising. For example, people may choose to purchase a product after seeing it advertised on television or on the Internet. Word-of-mouth effects and inter-individual interactions occur between those who have connections with one another. Through successive interactions among individuals, an idea or innovation can diffuse through the population,

reaching and possibly being accepted by an growing set of individuals over time.

The concept of viral marketing [2–4] has attracted great interest from both academia and business [5]. Viral marketing refers to advertising strategies that capitalize on information diffusion in social networks. Here, not only does a company promote its product via advertising, but it seeks to achieve broad reach of the advertising message through subsequent information sharing among individuals through social network interactions. If an early set of individuals who are exposed to the idea or product are favorably impressed and relay their opinion to their acquaintances, the acquaintances might also have a favorable opinion and in turn share the information with their circle of acquaintances, and so on, resulting in a viral advertising campaign. When successful, this is a very efficient and effective approach to information diffusion or product advertising. While a true viral marketing campaign is difficult if not impossible to design, the viral marketing concept will lead us to address the question of how to design an advertising strategy so as to maximize social network communication effects and optimize the effectiveness of a campaign promoting an idea or product. In this thesis, sometimes we refer to any social network marketing campaign simply as viral marketing.

It is often the case that multiple company entities simultaneously promote competing products or ideas. This leads to competition in social network marketing. For such a competition among advertisers, one can ask how efficient (or inefficient) the dynamics are as a function of the selfishness of the involved companies or entities. Mathematically, the question of existence of a Nash equilibrium for such a situation arises. If a Nash equilibrium is known to exist, it is interesting to ask whether or

not the participating companies can coordinate their actions so as to reach the Nash equilibrium. One can also ask if companies can cooperate to maximize their profits. All these questions are critical for the analysis of competition and gaming in viral marketing.

Another interesting question in viral marketing is how to determine an individual's economic value as a recommender of products within a social network. Here, the interest is in the eventual benefit of an individual recommending a product to his connections, who will in turn advocate the product to their connections at least to some extent, and so on. This can be addressed by determining a reasonable measure of each node's "centrality" with regard to the diffusion process. Although there has been significant work on defining centrality measures that are exclusively a function of a network's topology [6–11], these studies are generally static in nature and it's worthwhile to investigate new concepts of centrality in the context of diffusion dynamics.

In the following sections we briefly review existing work on modeling of diffusion dynamics in social networks, optimal viral marketing, viral marketing competition, and concepts of "diffusion centrality." We follow this discussion with a summary of the contributions of this dissertation in these areas. We note that this dissertation addresses a subject (social network dynamics) which has received a great deal of interest using a large variety of approaches and from researchers with diverse scientific backgrounds. The past work that is mentioned below and in the references is chosen from this vast literature based on its relevance to the work presented in the dissertation.

1.1 Modeling of Diffusion Dynamics

Since diffusion is common in various types of networks, researchers have proposed models for diffusion of information using a variety of perspectives. This includes using analogies between information diffusion and other kinds of diffusion processes such as the spread of infectious diseases in large populations.

In sociology, there have been rich studies on the diffusion of innovations in social networks, mostly focused on studying the influence of “word of mouth” communications in spreading innovations. Two models, the threshold model [12] of Granovetter and the cascade model [3] of Goldenberg *et al.* are considered fundamental in this area [13]. The famous cascade model [14] of Kempe *et al.* further generalizes these two models.

In the economics community, researchers study the propagation of information from a game theory perspective, in which agents are postulated to receive a payoff if they adopt an innovation or learn the new information being spread. Morris [15] and Young [16] consider a model in which, at each time step, all agents decide to adopt the new innovation or not based on maximization of their utilities. Ghaderi *et al.* [17] consider a scenario containing “stubborn agents” who consistently insist on their original opinions. Haller *et al.* [18] and Bala *et al.* [19] explore situations in which an agent can form a connection with another to earn a higher payoff.

The natural analogy between the spread of an infectious disease and the spread of information is utilized by numerous researchers to study innovation diffusion. These studies derive benefit from past work in mathematical epidemiology. Two

important models from epidemiology are the susceptible-infected-recovered (SIR) model and the susceptible-infected-susceptible (SIS) model [20–23]. The SIR model assumes that agents are initially susceptible, and after exposure to disease they may become infected with some probability, and can eventually recover. Newman [24] uses percolation theory to study the steady state of the SIR model for a random network. Pastor-Satorras and Vespignani [25, 26] use mean-field approximation to derive a dynamical SIS model for scale-free networks. Wang *et al.* [27], Ganesh *et al.* [28] and Mieghem *et al.* [29] propose a dynamical SIS model for any network represented by a fixed graph. Meanwhile, Youssef *et al.* [30] explore the SIR dynamic model in a similar scenario.

Although the diffusion of innovations has been studied from various perspectives, the nature of the propagation mechanism is similar in the various approaches. Shakarian *et al.* [31] employ the generalized annotated program (GAP) as a general framework to express different types of existing diffusion models on graphs [32–34].

In this dissertation, we focus on two improvements to commonly used models of innovation diffusion. First, most diffusion models directly employ the actual social network topology where each agent is viewed as a vertex in a graph. A practical issue is that these networks are usually of extremely large size, which implies that the associated models are computationally very demanding or even intractable. Also, the connections between all agents are difficult to determine and costly to estimate. As noted by Newman [35], there is considerable difficulty in studying network dynamics for medium and large scale networks. Also, since human behavior appears random on a local scale [36–38], it is likely inefficient to exploit the

whole detailed topology that explicitly represents all agents in a large population. Therefore it is important to make further simplifications on network structure to reduce the computational burden in diffusion modeling for social networks. Thus, it is important to develop models that allow for a reduced computational burden compared with detailed multi-agent models, while maintaining essential elements of the network structure and dynamics. We address this critical need by developing a multi-community social network model that is much simpler than detailed multi-agent models, while respecting the basic network structure of a social network and maintaining the dynamical contagion effects in the model. The second direction for improvement that we address is that most models do not incorporate external dynamic influences on innovation diffusion. In addition to the “word-of-mouth” effects, the diffusion of an innovation is also shaped by influences from external sources, such as news media and actions and statements of public figures. A typical example of such external influences is commercial advertising, which functions to influence consumers’ choices.

We derive a dynamic model for innovation diffusion with external influences using the analogy between information propagation and the spreading of an infectious disease. First, we generalize the SIR model in [30] and the SIS model in [29] of infectious disease propagation by utilizing a continuous-time Markov chain setting to model the innovation diffusion process, which is based on the assumption that the arrival of an innovation to an agent and the recovery process (i.e., loss of interest in adopting the innovation) of an agent who has previously adopted the innovation are independent Poisson processes. We also model the effect of external influences

by assuming that such an influence increases the rates of the Poisson processes by which agents decide to adopt an innovation.

Next, we introduce the concept of “community” [39,40] into our modeling and view a large social network as a finite set of interacting communities, where these communities are connected to each other through a backbone network. Agents belonging to the same community are assumed to have fully-mixed interactions with each other; this assumption entails that each community consists of many individuals, and the individuals are mixed uniformly within the community. Also, we assume that agents from any two connected communities interact with each other uniformly [41]. Through this construction, we simplify our model to only consider innovation diffusion within and between communities, allowing each community to be represented as a vertex in the simplified social network model. Therefore the complexity of the social network using the new “networked communities” modeling framework is significantly reduced, rendering the new model much more tractable computationally than a full multi-agent model that explicitly includes the interactions and actions of all individuals in the population. Also, we will show that our community-based diffusion model is an extension of the well-known mean-field model [42], and that it produces results that are in agreement with full agent-based models [27, 29, 30] under certain conditions.

As noted above, one type of information innovation process involves advertising of commercial products through a social network. In this work, we at times make use of the following observation about adoption of commercial products depending on the type of product being advertised. Commercial products are commonly viewed

as falling into one of the following two broad categories: durable consumer goods (DCG) and fast-moving consumer goods (FMCG). The similarity between disease propagation and innovation diffusion enables us to interpret the SIR and SIS models in the context of product diffusion. If we view the adoption of a product as an infectious disease, then those who have not adopted the products are susceptible and those who have adopted the product and recommend it to others are infected. After having been using the DCG for some time, agents may get tired of the product and quit the diffusion by neither making recommendations to others nor making another purchase, which fits into the SIR framework. Meanwhile, when agents are tired of the FMCG, they simply stop using the product or making recommendations, but they are still likely to adopt the product (or a competing product) again later. Such a scenario for FMCG fits with the assumptions inherent in the SIS framework.

To maintain consistency with research in epidemiology and to readily observe correspondences between innovation diffusion and infectious disease propagation, we will use the terms “susceptible,” “infected” and “recovered” to refer to agents’ corresponding states in an innovation diffusion process. Also, we at times refer to models of DCG diffusion as SIR models and those for FMCG diffusion as SIS models. Note that the DCG and FMCG terminologies are used here for information diffusion in general, and that the analogy with consumer products is made for convenience and to increase intuition on the processes being studied.

1.2 Optimal Social Network Marketing

There has been extensive research aimed at maximizing the effectiveness of information diffusion through word-of-mouth communication. A famous example involves viral marketing [2–4]. A typical question formulated in viral marketing is if we can convince a given number of agents to adopt a product initially, how can we select these agents so as to maximize the word-of-mouth effects as well as inter-individual interactions and thus the global expected profits (or other similar benefit) resulting from the campaign [43].

To answer this question, Domingos *et al.* [44, 45] use data mining techniques and proposed a heuristic algorithm. Kempe *et al.* [14, 46] show that in a variety of famous models, the spreading maximization problem is NP-hard. They also determine a greedy algorithm with guaranteed approximation ratio to within $1 - 1/e$ of the true optimum. Shakarian *et al.* [31] further generalizes this algorithm into the framework of GAP with the same approximation ratio.

In reality, advertising on a social network is a continuous process occurring over time. In previous research on advertising in social networks, it is common to assume that an advertising campaign consists simply of initially targeting a few “influential” agents in a network. These research results do not use models that can incorporate continuous advertising that acts throughout the diffusion of the message. For example, in [14] Kempe *et al.* study spreading maximization primarily based on two types of models: the linear threshold model [15, 47] and the independent cascade model [3, 4]. Both of these two models cannot incorporate continuous advertising

during diffusion processes.

There is also a body of research that formulates social network marketing in terms of an optimal control problem [48,49]. Economics researchers have introduced various dynamic models for diffusion of advertising campaigns, such as the advertising capital model [50], the random walk model [51], the product quality model [52] and the market growth model [53]. Dynamic advertising is investigated based on these dynamic models. Khouzani *et al.* [54] also explore optimal controls based on epidemic models. However, all of these efforts employ homogeneous model, which means that the connection topology of a social network is not reflected in these works. Also, the reliance on classical optimal control theory leads to only qualitative descriptions of the optimal advertising strategies.

In this dissertation we consider social network marketing based on the community-based SIR and SIS models as introduced in Section 1.1. We formulate and study three types of marketing problems: referral marketing, seeding marketing and dynamic marketing. In referral marketing, we study the maximization of product diffusion via increasing a product's spreading rates at a given cost in the form of referral bonus [55]. Seeding marketing is similar to the spreading maximization problem in [14,31,46], where we aim to maximize a diffusion by convincing a set of agents to adopt a product initially. Dynamic marketing corresponds to the scenario where a company advertises its product continuously during the product's diffusion. This problem is mathematically formulated as an optimal control problem.

We will show that the referral marketing problem can be converted into a convex optimization problem, which is straightforward to solve. However, the seeding

marketing and dynamic marketing problems are nonlinear optimization and nonlinear optimal control problems respectively, which are extremely difficult, if not impossible, to solve analytically. Therefore we propose heuristics allowing us to approximately solve these problems numerically.

In discrete mathematics, submodularity is an important property of set functions with deep theoretical consequences and many applications [56]. The maximization of submodular functions is widely studied in various areas [57, 58]. In [14], Kempe *et al.* formulate the spreading maximization problem as maximization of a submodular function and solve it by using a greedy heuristic. It should be noted that the solution space in [14] is finite, which suits the definition of submodularity. However, the solution space in both the seeding and dynamic marketing problems is at least a continuum, which is beyond the scope of traditional studies using submodularity.

We therefore generalize the traditional notion of submodularity into “continuous monotone submodularity” and show that both the seeding marketing and dynamic marketing problems can be converted into maximization of continuous monotone submodular functions. Also, we propose a new greedy heuristic to maximize continuous monotone submodular functions. With these preparations, the seeding marketing and dynamic marketing problems are solved in a unified algorithmic framework. Through rigorous analysis, we show that our greedy heuristic has an approximation ratio no less than $1 - 1/e$ for both problems. Moreover, we show that the greedy heuristic can be converted into a closed-loop form for the SIR model, which improves resilience to external noises and model uncertainties.

Our continuous monotone submodularity concept can be readily described in the following way: if a mapping from investment to the corresponding profit in economics is continuous monotone submodular, then with the investment increased, the profit is increases but marginal profit decreases. Intuitively, such a property should generally hold for various readouts in product diffusion and marketing. It should be noted that whenever this property applies, the greedy heuristic achieves an approximation ratio no less than $1 - 1/e$.

We also formulate other variants of dynamic marketing problems, including minimization of cost to achieve a minimum desired profit, and studying trade-offs between profits and cost. The greedy heuristic is modified for the minimum cost problem while a bi-directional local search algorithm is designed for the trade-off problem. Analysis is conducted on corresponding approximation ratios for both problems.

Taking advantage of recent progress in monotone control systems [59], we show that the marketing problems based on the SIS model are intrinsically convex. This convexity implies that a locally optimal marketing policy found via using any numerical optimal control toolbox is automatically globally optimal.

1.3 Social Network Marketing Game

Imagine several companies simultaneously promoting similar and competing products (such as Pepsi and Coca-Cola) on a social network. Using game theory terminology, we call such a competition in marketing of multiple competing products

a social network marketing game. In a marketing game, each company is viewed as a player and chooses an advertising strategy to maximize its own product’s sales or other objectives that are formulated in terms of a utility function.

To analyze a marketing game, the first step is to model diffusions of multiple competitive products on a social network. Existing models of competitive innovation diffusion generally fall into two categories. The first category, which we call cascade competition models, include the linear threshold competition model and the independent cascade competition model [60–63]. Another category is the dynamic competition model [64–67], which generally uses ordinary differential equations (ODEs) to model competitive contagions’ diffusions.

Research on viral marketing games based on cascade competition models is classified into two categories. In the first category, it is assumed that all but one of the players already have determined their strategies and the problem is to figure out the best response of the remaining player. Bharathi *et al.* [61] and Carnes *et al.* [68] study how to maximize one’s profits, while on the contrary Budak *et al.* [69] and He *et al.* [70] consider how to minimize the competitors’ profits. In the second category, one assumes that all players choose their strategies simultaneously. Dubey *et al.* [71] and Tzoumas *et al.* [63] focus on existence of a pure Nash equilibrium for marketing games; Tsai *et al.* [72] propose a heuristic to compute a mixed Nash equilibrium; Goyal *et al.* [62] and He *et al.* [70] study price of anarchy (PoA) of such marketing games based on cascade models.

The formulation of viral marketing games based on cascade competition models allows each company to target a set of “influential” agents to use their products

initially and thus promote their product via word-of-mouth effects and other inter-individual interactions. However, in practice companies can exert external influence on product diffusion during the whole process (e.g., via online advertising) but such continuous advertising cannot be incorporated into cascade competition models. On the contrary, dynamic competition models take continuous advertising into account as an input signal to a dynamical system. Within the scope of our knowledge, all the existing research [73] on viral marketing games based on dynamic competition models [64–66] primarily employ the traditional differential game theory [74] and provide certain qualitative descriptions of a Nash equilibrium.

Although the dynamical competition model is more flexible than the cascade competition model, it is also more complicated. Existing studies on social network marketing games based on dynamic models primarily suffer from two critical problems. Firstly, the dynamic competition models that all existing research focuses on are aggregate, which means that agents’ connections in a social network are neglected. Secondly, all the research efforts to-date use classical differential game theory [74] and transform the specification of the Nash equilibrium into a two-point boundary value problem, which is extremely difficult to analytically solve in general. For this reason most of these efforts merely provide certain qualitative descriptions of the Nash equilibrium and few result in a closed-form solution.

As noted above (and in Section 1.1), commodities are usually classified into two categories: durable consumer goods (DCG) and fast-moving consumer goods (FMCG). To model diffusion dynamics of multiple competitive DCGs, we propose a DCG competition model by generalizing our community-based SIR model.

In FMCG diffusion, after a customer tires of a product, he may either switch to another competing product or choose not to use any product in the category. However, he is still likely to purchase the product again later. This process fits the assumptions of the voter model [75–77]. In this dissertation, we generalize the voter model in [76] to formulate FMCG competition models.

It should be noted that both the DCG competition model and the FMCG competition model in this dissertation are developed for social networks, and the first problem of the current marketing game studies based on dynamic competition models as mentioned above is overcome in this work. In this dissertation, we call social network marketing games based on the DCG and FMCG competition models as the DCG marketing game and the FMCG marketing game, respectively.

Next, in addition to using the differential game theory [74] to analyze existence of Nash equilibrium in viral marketing games, we are also inspired by work of Vetta [78] and Rosen [79] to perform further analysis. The inspiration from Vetta [78] sheds light on the analysis of PoA of marketing games, which yields a result that the PoA is bounded by 2 in both the DCG and the FMCG marketing games. Through generalizing the result of Rosen [79] in combination with the recent progress in monotone control systems [59, 80], we also show that FMCG marketing game enjoys almost uniqueness of the Nash equilibrium (i.e., any two Nash equilibria agree almost everywhere), and we propose a “best-response” mechanism for each company to find the Nash equilibrium in a distributed iterative algorithm.

1.4 Diffusion Centrality

An increasingly important problem arising in studies of social networks is to identify the “centrality” of network vertices. Centrality of a vertex is a measure of the vertex’s importance relative to other vertices in the network, in some well-defined sense. Notions of network centrality can be useful in identifying important vertices in various practical settings [81], such as targeted marketing [44, 82], epidemiology modeling [83] and network robustness [84], among others. Numerous centrality metrics have been proposed. A few examples of such metrics are degree centrality [6], eigenvector centrality [7], Katz centrality [8], PageRank [9], closeness centrality [10] and betweenness centrality [11].

Most of the current research on network centrality only takes static network structure into account. However, social networks function as a medium for the diffusion of information over time (such as the spread of opinions, rumors or even infectious diseases [14]). Thus, it is worthwhile to measure the importance of each vertex within the context of a specific diffusion process, which cannot be achieved using the network graphical structure alone. Instead, it is necessary to study the notion of centrality in the context of a diffusion process. In this thesis, we call this the problem of measuring *diffusion centrality*.

The idea to measure an agent’s centrality associated with a diffusion process in a network [85] has attracted increasing interest. Researchers from various areas are gradually realizing the importance of the diffusion centrality. For example, Banerjee *et al.* [86] examine how “central” agents are in diffusions of micro-finance,

and Grindrod *et al.* [87] investigate this concept in monitoring broadcast activities in a network.

Current research on measuring centrality in a diffusion primarily divides into two categories: one body of work extends the existing centrality metrics, while the other aims at development of new metrics. Ide *et al.* [88] compare the performance of various centrality metrics using simulation of spreads of an infection. Kim *et al.* [81] make such a comparison based on an empirical dataset from the Hagggle project [89]. Carreras *et al.* [90] focus on eigenvector centrality in a highly partitioned mobile network. Guimarães *et al.* [91] attempt to modify traditional centrality metrics to adapt to a diffusion process.

Lerman *et al.* [92] propose a definition of centrality metric for dynamic networks and apply it in rankings of a citation network. Kim [93] designs a temporal centrality notion and investigates its performance in the Hagggle dataset [89]. Wang *et al.* [94] design an algorithm to numerically determine agents' dynamic centrality via data mining. There are also several contributions focusing on potentials application of such dynamic centrality metrics. For instance, Mochalova *et al.* [95] consider application of dynamic centrality in influence maximization.

In this dissertation we propose a diffusion centrality notion based on the community-based SIR and SIS models introduced in Section 1.1. Since in the network associated with the community-based model, each vertex represents a community instead of an individual, the centrality we study in this dissertation measures the importance of each community in a diffusion process.

We investigate the notion of diffusion centrality in two scenarios: when the

diffusion involves an outbreak and when it does not, where the definition of outbreak will be explained later in this dissertation. In the case when an outbreak does not occur, we study the influence of one specific community over another one by extending sensitivity analysis of steady-state equations of the SIR model. We also interpret such influences (matrix) by generalizing the idea of counting walks on networks. The diffusion centrality of each community is then defined as weighted sum of its influences over other communities, which turns out to be related to Katz centrality [8].

When an outbreak occurs, we also study diffusion centrality from another viewpoint: we investigate how vulnerable each community is. More specifically, centrality can be defined as the solution to steady-state equations of both the SIR and SIS models. This definition makes sense because the more seriously a community is influenced, the closer it is to the center of a diffusion process. We show that when the outbreak is weak, i.e., a relatively small percentage of the population is influenced in the diffusion, such a centrality metric is connected with eigenvector centrality [7].

1.5 Organization of the Dissertation

The remainder of this dissertation is organized as follows: In Chapter 2, we propose community-based SIR and SIS models and study properties of diffusions based on these models. We also employ an approach based on percolation theory to study diffusion for community-based networks, and obtain conclusions that agree with those obtained using the SIR and SIS models.

In Chapter 3 we formulate three types of marketing problems based on the proposed SIR and SIS models. A greedy heuristic algorithm is designed to solve these problems in a unified framework, and approximation ratios are obtained for each of the problems. We also show that some of the marketing problems are convex in special scenarios, in which case it is feasible to find a globally optimal solution.

We study and model diffusions of multiple competitive contagions in Chapter 4. Based on the proposed models of diffusion of competitive contagions, we use game theory to investigate competition of multiple companies in marketing their commercial goods. This can also be viewed more generally in terms of competition in promoting competing ideas or opinions. Also, an algorithm is designed for each player to distributively seek a Nash equilibrium.

In Chapter 5 we initiate a study of dynamic centrality of communities in our community-based social network diffusion model. Concepts of diffusion centrality are proposed for several scenarios, and their connections with classical network centrality metrics are discussed. In Chapter 6 we give concluding remarks and suggest topics for future research.

Chapter 2: Modeling and Analysis

This chapter deals with the modeling and analysis of diffusion dynamics in social networks viewed as a finite set of communities connected by a backbone network. Starting from modeling a diffusion in a social network as a random walk and then invoking the structure of a network of communities, we propose vector SIR and SIS models for diffusion dynamics. In Section 2.1 we derive community-based dynamically models for both SIR and SIS infection diffusion. We analyze the accuracy of our model in Section 2.2. In Section 2.3 we study conditions for infection outbreak. In Section 2.4 we study the final states reached by a diffusion. Finally, in Section 2.5 we use percolation theory to study diffusion in the community-based network and compare the obtained results with those achieved using the vector dynamic system models.

2.1 Modeling Diffusion Among Communities

In [30] Youssef *et al.* consider SIR epidemic diffusion in a network consisting of N agents connected through a directed graph. In the model below, coefficient $a_{jk} = 1$ when there is a link from agent k to agent j , and $a_{jk} = 0$ otherwise. With λ and γ denoting the infection rate and the recovery rate, respectively, the agent-based

SIR epidemic model of Youssef is

$$\begin{aligned}
\frac{d s_j(t)}{d t} &= -s_j(t)\lambda \sum_{k=1}^N a_{jk}i_k(t) \\
\frac{d i_j(t)}{d t} &= s_j(t)\lambda \sum_{k=1}^N a_{jk}i_k(t) - \gamma i_j(t) \\
\frac{d r_j(t)}{d t} &= \gamma i_j(t)
\end{aligned} \tag{2.1}$$

where $s_j(t)$, $i_j(t)$ and $r_j(t)$ denote the probability that agent j is susceptible, infected or recovered at time t , respectively.

Before deriving our community-based SIR model, we make two generalizations to the agent-based SIR model of Youssef (2.1). First, in our configuration infection and recovery rates can differ for different agents. We denote the infection rate from agent k to agent j by λ_{jk} and the recovery rate of agent j by γ_j . Also, we introduce a term $u_j(t)$, which represents the strength of an external influence on agent j at time t . In this setting, if agent j is susceptible at time t , then the infection rate for this agent is increased by $u_j(t)$.

Denote the state of agent i at time t by $X_j(t)$; the state $X_j(t)$ takes values in $\{\text{sus}, \text{inf}, \text{rec}\}$. Assume that for any pair of agents, their states are independent of each other, i.e. $X_j(t) \perp\!\!\!\perp X_k(t)$ for any $1 \leq j, k \leq N$ and time t . Under this assumption, our generalized agent-based SIR model is

$$\begin{aligned}
\frac{d s_j(t)}{d t} &= -s_j(t) \left(\sum_{k=1}^N a_{jk}\lambda_{jk}i_k(t) + u_j(t) \right) \\
\frac{d i_j(t)}{d t} &= s_j(t) \left(\sum_{k=1}^N a_{jk}\lambda_{jk}i_k(t) + u_j(t) \right) - \gamma_j i_j(t) \\
\frac{d r_j(t)}{d t} &= \gamma_j i_j(t)
\end{aligned} \tag{2.2}$$

where detailed derivation of (2.2) are provided in Appendix A.

As mentioned previously, in the agent-based SIR model (2.2), each agent corresponds to a vertex (node) in the network. For a network of large size, the agent-based SIR model is greatly complicated and may be intractable in practice. In reality, because of factors such as geographic location or educational or language backgrounds, a social network can be viewed as a finite set of communities. The division of agents into communities should be performed according to attributes that are significant to the particular type of information or product whose diffusion is being considered. If we then abstract each community as an aggregated vertex in the social network, an innovation can correspondingly spread either within a community or between pairs of communities. In this section we further elaborate on this approach of viewing a large network of agents in terms of interconnected communities, and simplify the agent-based SIR model (2.2).

We assume that the whole population is divided into n communities, and that each agent belongs to one and only one community. The i -th community includes N_i agents. The primary feature of each community is that all the agents in the community have equal probability to interact with each other. Also, any two agents from a given pair of connected communities also have equal probability to be connected. These are well-mixedness assumptions that are reasonable when each community is very large and homogeneous, and when connections between individuals in different connected communities are also uniform.

In our configuration, the agents in the j -th community have a probability of p_{jj} to have an acquaintance with others in the same community. Similarly, they have probability of p_{jk} to know individuals from the k -th community. Assume that

p_{jj} is strictly positive, which means that agents of the same community are always likely to be acquainted with one another. At the same time, a positive p_{jk} describes how tightly the j -th and the k -th communities are connected, while $p_{jk} = 0$ means the two communities are not directly connected. Using this mathematical structure, we assume well-mixedness within each community within a network of communities, and uniform connectedness of agents among any two directly interconnected communities.

Figure 2.1 depicts an example of a community structure in a social network generated by $p_{ii} = 0.4$ for $i = 1 \dots 4$, $p_{12} = p_{41} = 0.05$, $p_{21} = p_{13} = p_{31} = p_{14} = 0.1$, $p_{23} = 0.15$ and with all the other probabilities set to zero. The arrowhead indicates a directed connection. For example the agent with index 1 in community 2 is connected with the agent with index 1 in community 1. A bi-directional line means the two agents are connected to each other and each influences the other.

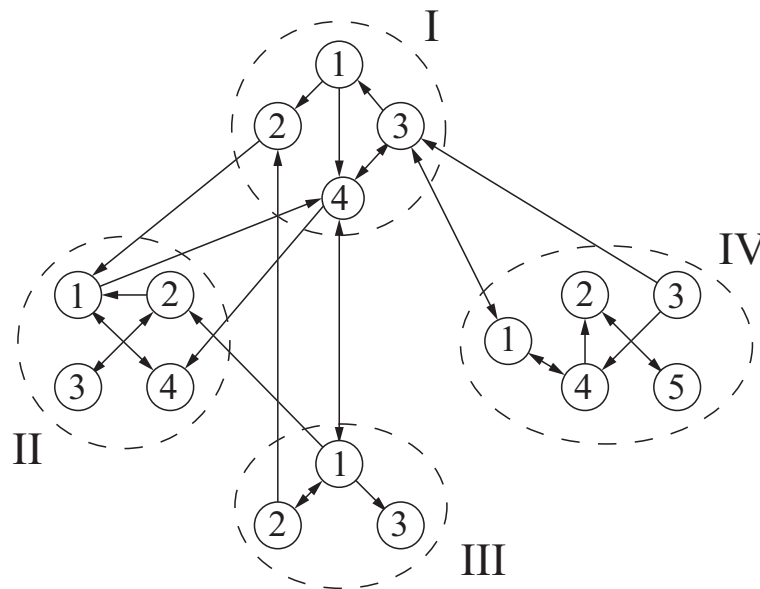


Figure 2.1: Illustrating a sample community-based social network structure

Let $X_j^v(t)$ denote the states of agent v in community j , which takes a value in $\{\text{sus}, \text{inf}, \text{rec}\}$. Denote the connection from the agent w in community k to the agent j in community v as a_{jk}^{vw} . If they are connected, $a_{jk}^{vw} = 1$ and otherwise $a_{jk}^{vw} = 0$. We easily have $\Pr [a_{jk}^{vw} = 1] = p_{jk}$. Let $s_j^v = \Pr [X_j^v(t) = \text{sus}]$, $i_j^v = \Pr [X_j^v(t) = \text{inf}]$ and $r_j^v = \Pr [X_j^v(t) = \text{rec}]$.

Let $\lambda'_{jk} \geq 0$ and $\gamma'_j > 0$ be the infection rate from an agent in community k to another agent in community j and the recovery rate of agents in community j , respectively. Assume that the advertising strengths to each agent in the same community are identical. This assumption makes sense because of the well-mixedness within each community. Let $u(t) = \begin{bmatrix} u_1(t) & u_2(t) & \dots & u_n(t) \end{bmatrix}$ be the advertising strength vector, where $u_j(t)$ is the strength of the external influence for all agents in community j at time t .

Applying our agent-based SIR model (2.2) to the social network with the community structure as specified above, we have the following model:

$$\begin{aligned} \frac{d s_j^v(t)}{d t} &= -s_j^v(t) \left(\sum_{k=1}^N \sum_{w=1}^{N_k} \lambda'_{jk} a_{jk}^{vw} i_k^w(t) + u_j(t) \right) \\ \frac{d i_j^v(t)}{d t} &= s_j^v(t) \left(\sum_{k=1}^N \sum_{w=1}^{N_k} \lambda'_{jk} a_{jk}^{vw} i_k^w(t) + u_j(t) \right) - \gamma'_j i_j^v(t) \\ \frac{d r_j^v(t)}{d t} &= \gamma'_j i_j^v(t) \end{aligned} \quad (2.3)$$

We are especially interested in the expected fraction of susceptible, infected and recovered populations in each community, i.e., $s_j(t) = \mathbf{E} \left[\frac{\sum_{v=1}^{N_j} s_j^v(t)}{N_j} \right]$, $i_j(t) = \mathbf{E} \left[\frac{\sum_{v=1}^{N_j} i_j^v(t)}{N_j} \right]$ and $r_j(t) = \mathbf{E} \left[\frac{\sum_{v=1}^{N_j} r_j^v(t)}{N_j} \right]$, where $\mathbf{E}[\cdot]$ denotes mathematical expectation. To study dynamics of the quantities $s_j(t)$, $i_j(t)$ and $r_j(t)$, the following assumption is essential.

Assumption 2.1 States of any pair of agents are independent of each other and jointly independent of whether or not there is a direct connection between them, i.e., $X_j^v(t) \perp\!\!\!\perp a_{jk}^{vw}$, $a_{jk}^{vw} \perp\!\!\!\perp X_k^w(t)$ as well as $X_k^w(t) \perp\!\!\!\perp X_j^v(t)$ for any $1 \leq j, k \leq n$, $1 \leq v \leq N_j$ and $1 \leq w \leq N_k$.

Taking the sum of each equation in (2.3) over all v for every fixed j , and replacing a_{jk}^{uv} with $\mathbf{E}[a_{jk}^{uv}] = p_{jk}$, we have

$$\begin{aligned}
\frac{ds_j(t)}{dt} &= \mathbf{E} \left[-\frac{1}{N_j} \sum_{v=1}^{N_j} s_j^v(t) \left(\sum_{k=1}^N \sum_{w=1}^{N_k} \lambda'_{jk} a_{jk}^{vw} i_k^w(t) + u_j(t) \right) \right] \\
&= - \sum_{v=1}^{N_j} \frac{1}{N_j} \left(\sum_{k=1}^N \sum_{w=1}^{N_k} \mathbf{E} [\lambda'_{jk} a_{jk}^{vw} s_j^v(t) i_k^w(t)] + \mathbf{E} [s_j^v(t) u_j(t)] \right) \\
&= - \sum_{v=1}^{N_j} \mathbf{E} \left[\frac{1}{N_j} s_j^v(t) \right] \sum_{k=1}^N \sum_{w=1}^{N_k} \lambda'_{jk} \mathbf{E} [a_{jk}^{vw} i_k^w(t)] - s_j(t) u_j(t) \tag{2.4} \\
&= - s_j(t) \left(\sum_{k=1}^N \lambda'_{jk} p_{jk} \sum_{w=1}^{N_k} N_k \mathbf{E} \left[\frac{1}{N_k} i_k^w(t) \right] + u_j(t) \right) \\
&= - s_j(t) \left(\sum_{k=1}^N N_k \lambda'_{jk} p_{jk} i_k(t) + u_j(t) \right)
\end{aligned}$$

Similarly, for $i_j(t)$ and $r_j(t)$,

$$\begin{aligned}
\frac{di_j(t)}{dt} &= \mathbf{E} \left[\frac{1}{N_j} \sum_{v=1}^{N_j} s_j^v(t) \left(\sum_{k=1}^N \sum_{w=1}^{N_k} \lambda'_{jk} a_{jk}^{vw} i_k^w(t) + u_j(t) \right) \right] - \mathbf{E} \left[\frac{1}{N_j} \sum_{v=1}^{N_j} \gamma'_i i_j^v(t) \right] \\
&= s_j(t) \left(\sum_{k=1}^N N_k \lambda'_{jk} p_{jk} i_k(t) + u_j(t) \right) - \gamma'_i i_j(t) \\
\frac{dr_j(t)}{dt} &= \mathbf{E} \left[\frac{1}{N_j} \sum_{v=1}^{N_j} \gamma'_i i_j^v(t) \right] = \gamma'_i i_j(t)
\end{aligned}$$

It should be noted that the equation $\mathbf{E} [s_i^v(t) a_{ij}^{vw} i_j^w(t)] = \mathbf{E} [s_i^v(t)] = \mathbf{E} [a_{ij}^{vw}] = \mathbf{E} [i_j^w(t)]$ is in the derivations above; this equation is a consequence of Assumption

2.1. This assumption especially makes sense for populations of large scale because if an agent has many connections, then the state of one of its acquaintances and whether or not he has a connection with another specific agent are not determining factors for his state.

Finally, define matrices Λ and Γ by $\Lambda := [\lambda_{jk}]_{N \times N}$ where each entry $\lambda_{jk} = N_k \lambda'_{jk} p_{jk}$, and $\Gamma := \text{diag}[\gamma_j]_{N \times N}$ which is diagonal with diagonal entries $\gamma_j = \gamma'_j$, where $\text{diag}[\cdot]$. The SIR model on the social network with networked community structure is then:

$$\frac{dS(t)}{dt} = -\text{diag}[S(t)] (\Lambda I(t) + U(t)) \quad (2.5a)$$

$$\frac{dI(t)}{dt} = \text{diag}[S(t)] (\Lambda I(t) + U(t)) - \Gamma I(t) \quad (2.5b)$$

$$\frac{dR(t)}{dt} = \Gamma I(t) \quad (2.5c)$$

where $S(t) := \begin{bmatrix} s_1(t) & s_2(t) & \dots & s_N(t) \end{bmatrix}^T$, $I(t) = \begin{bmatrix} i_1(t) & i_2(t) & \dots & i_N(t) \end{bmatrix}^T$,
 $R(t) := \begin{bmatrix} r_1(t) & r_2(t) & \dots & r_N(t) \end{bmatrix}^T$ and $U(t) := \begin{bmatrix} u_1(t) & u_2(t) & \dots & u_N(t) \end{bmatrix}^T$.

Using all the notations specified above and following similar derivation specified above, we obtain the community-based SIS model as follows:

$$\frac{dS(t)}{dt} = -\text{diag}[S(t)] (\Lambda I(t) + U(t)) + \Gamma I(t) \quad (2.6a)$$

$$\frac{dI(t)}{dt} = \text{diag}[S(t)] (\Lambda I(t) + U(t)) - \Gamma I(t) \quad (2.6b)$$

Unless specified otherwise, we will refer to the community-based SIR model (2.5) and the community-based SIS model (2.6) as SIR model and SIS model, respectively, in the rest of this chapter.

The SIR model (2.5) and the SIS model (2.6) are community-based diffusion

models proposed in this chapter. Compared with the agent-based models in [27–29], the model we have derived using the community-based topology is of much smaller size and therefore more convenient for simulation and analysis. Also, since human behavior appears random on a local scale [36–38], it is likely inefficient to work with the whole detailed topology explicitly representing all individuals in a large population. Instead, the community-based model considers individuals’ connections at a higher level, which is more compact and amenable for study.

The community-based model introduced above can be extended to a more general case in which the agents within each community are not fully mixed but rather carry specific degree distributions governing their number of connections within their community and to other communities. To deal with such a scenario, we can simply view individuals with the same number of degrees as belonging a “smaller” community, and then follow the same derivation as above. This approach would be similar to the mean-field approximation proposed by Vespignani *et al.* [25, 26].

It is interesting to note that Vespignani’s mean-field approximation model for large enough populations [25] is a special case of our model if we take agents who have the same number of connections into the same community and follow the approach above. Also, if we view each individual as an independent “community” then our model is reduced to Youssef’s agent-based SIR model [30]. Thus, our model generalizes these existing models while allowing a reduction in computational burden.

2.2 Accuracy of the Community-Based Approximation

Recalling that Assumption 2.1 was essential in the derivation of the SIR model (2.5), we next investigate the accuracy of this model through both theoretical analysis and numerical simulation.

2.2.1 Theoretical Analysis

The only approximation made in the derivation of the model (2.5) based on Assumption 2.1 is $\mathbf{E} [s_j^v(t) a_{jk}^{vw} i_k^w(t)] = \mathbf{E} [s_j^v(t)] \mathbf{E} [a_{jk}^{vw}] \mathbf{E} [i_k^w(t)]$. The following equation always holds:

$$\mathbf{E} [s_j^v(t) a_{jk}^{vw} i_k^w(t)] = \mathbf{E} [s_j^v(t) i_k^w(t) | a_{jk}^{vw} = 1] \mathbf{E} [a_{jk}^{vw}]$$

When $a_{jk}^{vw} = 1$, i.e. there is a connection between the two agents, we have

$$\mathbf{E} [s_j^v(t) i_k^w(t) | a_{jk}^{vw} = 1] \leq \mathbf{E} [s_j^v(t) i_k^w(t)] \leq \mathbf{E} [s_j^v(t)] \mathbf{E} [i_k^w(t)]$$

because for a given a pair of agents, the connection from one to another cannot positively influence the probability that the former agent is infected and the latter agent is susceptible. Meanwhile, the infection of one agent cannot positively influence the probability of being susceptible for another agent.

Therefore, we have $\mathbf{E} [s_j^v(t) a_{jk}^{vw} i_k^w(t)] \leq \mathbf{E} [s_j^v(t)] \mathbf{E} [a_{jk}^{vw}] \mathbf{E} [i_k^w(t)]$. Looking at (2.4), we can conclude that $s_j(t)$ is a lower bound for $\mathbf{E} \left[\frac{\sum_{v=1}^{N_j} s_j^v(t)}{N_j} \right]$. Correspondingly, $i_j(t)$ is an upper bound for $\mathbf{E} \left[\frac{\sum_{v=1}^{N_j} i_j^v(t)}{N_j} \right]$.

According to (2.4), the term $\sum_{k=1}^N \sum_{v=1}^{N_k} \lambda'_{jk} a_{jk}^{jk} i_k^w(t)$ is a component of transition

rate of $s_j(t)$. Next we are going to focus on the deviation that arises from Assumption 2.1 by studying the variance of $\sum_{k=1}^N \sum_{v=1}^{N_k} \lambda'_{jk} a_{jk}^{jk} i_k^w(t)$ as follows: Since the input term $u_j(t)$ is not directly associated with Assumption 2.1, we simply ignore the term $u_j(t)$ by setting $u_j(t) \equiv 0$.

$$\begin{aligned}
& \text{Var} \left[\sum_{k=1}^N \sum_{w=1}^{N_k} \lambda'_{jk} a_{jk}^{vw} i_k^w(t) \right] \\
&= \mathbf{E} \left[\left(\sum_{k=1}^N \sum_{w=1}^{N_k} \lambda'_{jk} a_{jk}^{vw} i_k^w(t) \right)^2 \right] - \mathbf{E} \left[\sum_{k=1}^N \sum_{w=1}^{N_k} \lambda'_{jk} a_{jk}^{vw} i_k^w(t) \right]^2 \\
&= \sum_{\substack{1 \leq k \leq N \\ 1 \leq k' \leq N}} \sum_{\substack{1 \leq w \leq N_k \\ 1 \leq w' \leq N_{k'}}} p_{jk} p_{jk'} \lambda_{jk} \lambda_{jk'} \left(\mathbf{E} \left[i_{k'}^{w'}(t) i_k^w(t) \mid a_{jk}^{vw} = 1, a_{jk'}^{vw'} = 1 \right] \right. \\
&\quad \left. - \mathbf{E} \left[i_k^w(t) \mid a_{jk}^{vw} = 1 \right] \mathbf{E} \left[i_{k'}^{w'}(t) \mid a_{jk'}^{vw'} = 1 \right] \right) \\
&= \sum_{k=1}^N \sum_{w=1}^{N_k} \text{Var} \left[\lambda'_{jk} a_{jk}^{vw} i_k^w(t) \right] + \sum_{\substack{1 \leq k \leq N \\ 1 \leq k' \leq N}} \sum_{\substack{1 \leq w \leq N_k \\ 1 \leq w' \leq N_{k'}}}^{(k,w) \neq (k',w')} p_{jk} p_{jk'} \lambda_{jk} \lambda_{jk'} \\
&\quad \cdot \left(\mathbf{E} \left[i_{k'}^{w'}(t) i_k^w(t) \mid a_{jk}^{vw} = 1, a_{jk'}^{vw'} = 1 \right] - \mathbf{E} \left[i_k^w(t) \mid a_{jk}^{vw} = 1 \right] \mathbf{E} \left[i_{k'}^{w'}(t) \mid a_{jk'}^{vw'} = 1 \right] \right) \tag{2.7}
\end{aligned}$$

Notice that if all $i_j^w(t)$'s are independent, then $\text{Var} \left[\sum_{j=1}^N \sum_{w=1}^{N_j} \lambda'_{ij} a_{ij}^{vw} i_j^w(t) \right]$ is actually the first term at the R.H.S. of (2.7). However, infection of one agent cannot have negative impacts on the probability of infection at another agent, therefore we have

$$\mathbf{E} \left[i_{k'}^{w'}(t) i_k^w(t) \mid a_{jk}^{vw} = 1, a_{jk'}^{vw'} = 1 \right] \geq \mathbf{E} \left[i_k^w(t) \mid a_{jk}^{vw} = 1 \right] \mathbf{E} \left[i_{k'}^{w'}(t) \mid a_{jk'}^{vw'} = 1 \right]$$

and hereby the second term at the R.H.S. of (2.7) is always positive. This observation leads us to understand that positive correlations among the $i_j^w(t)$'s actually exacerbate the variance of $\sum_{j=1}^N \sum_{w=1}^{N_j} \lambda'_{ij} a_{ij}^{vw} i_j^w(t)$ compared with the case that all $i_j^w(t)$'s are independent.

Also, it should be noted that $\text{Var} \left[\sum_{k=1}^N \sum_{w=1}^{N_k} \lambda'_{jk} a_{jk}^{vw} i_k^w(t) \right]$ actually scales with the number of connections that a specific agent has. The number of connections that each agent has, i.e., $p_{ij} N_j$ for each $1 \leq j \leq n$, is bounded whatever the social network is like, therefore this variance is still moderate even if the whole social network is of extremely large size.

When each agent can become infected with a small probability, i.e., $i_j^w(t) < \epsilon$ for some small ϵ , then the second term at the R.H.S. of (2.7) scales as $O(\epsilon^2)$. Therefore the variance we are investigating in this scenario is very close to the first term, which corresponds to the case that all $i_j^w(t)$'s are independent.

In another interesting case where each agent is infected with a high probability, notice that

$$\begin{aligned} \mathbf{E} \left[\left(\sum_{k=1}^N \sum_{w=1}^{N_k} \lambda'_{jk} a_{jk}^{vw} i_k^w(t) \right)^2 \right] &= \mathbf{E} \left[\sum_{k=1}^N \sum_{k'=1}^N \sum_{w=1}^{N_k} \sum_{w'=1}^{N_{k'}} \lambda'_{jk} \lambda'_{jk'} a_{jk}^{vw} a_{jk'}^{vw'} i_k^w(t) i_{k'}^{w'}(t) \right] \\ &\leq \mathbf{E} \left[\sum_{k=1}^N \sum_{w=1}^{N_k} \lambda'_{jk} a_{jk}^{vw} i_k^w(t) \right] \sum_{k'=1}^N \sum_{w'=1}^{N_{k'}} \lambda'_{jk'} p_{jk'} \end{aligned}$$

then coefficient of variation

$$\frac{\sqrt{\text{Var} \left[\sum_{k=1}^N \sum_{w=1}^{N_k} \lambda'_{jk} a_{jk}^{vw} i_k^w(t) \right]}}{\mathbf{E} \left[\sum_{k=1}^N \sum_{w=1}^{N_k} \lambda'_{jk} a_{jk}^{vw} i_k^w(t) \right]} \leq \sqrt{\frac{\sum_{k=1}^N \sum_{w=1}^{N_j} \lambda'_{jk} p_{jk}}{\mathbf{E} \left[\sum_{k=1}^N \sum_{w=1}^{N_k} \lambda'_{jk} a_{jk}^{vw} i_k^w(t) \right]} - 1}$$

Since agents are infected with high probability, then $\mathbf{E} \left[\sum_{k=1}^N \sum_{w=1}^{N_k} \lambda'_{jk} a_{jk}^{vw} i_k^w(t) \right]$ is very close to $\sum_{k=1}^N \sum_{w=1}^{N_j} \lambda'_{jk} p_{jk}$. Therefore the coefficient of variation is small in this case, which is again in favor of Assumption 2.1.

2.2.2 Numerical Simulation

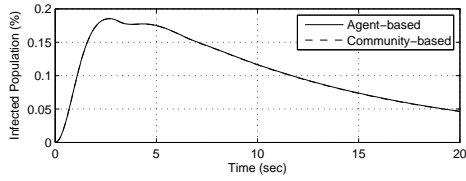
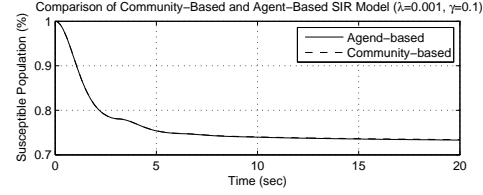
In addition to the theoretical analysis above, we also run a numerical simulation on a social network consisting of four communities. In each community there are 100 agents and we are given the following connectivity matrix for the communities:

$$P = \begin{bmatrix} 0.08 & 0.04 & 0.02 & 0.02 \\ 0.01 & 0.08 & 0.05 & 0.04 \\ 0.03 & 0.03 & 0.08 & 0.01 \\ 0.02 & 0.01 & 0.01 & 0.08 \end{bmatrix}$$

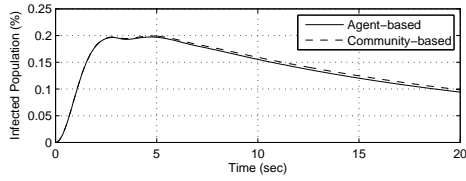
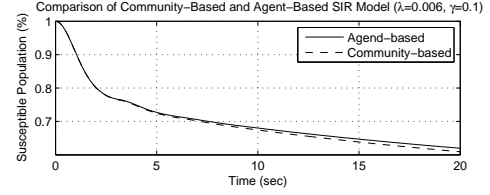
Also, we assume the parameters are uniform, i.e. $\lambda'_{ij} = \lambda$ and $\gamma'_k = \gamma$ for any $1 \leq i, j, k \leq 4$. We run the simulation for four sets of parameters: $\lambda = 0.001$, $\gamma = 0.1$; $\lambda = 0.006$, $\gamma = 0.1$; $\lambda = 0.03$, $\gamma = 0.1$ and $\lambda = 0.5$, $\gamma = 0.1$, which represent the scenarios with different spreading strength and the recovery rate. The input term $u_4(t) = |\sin t|$, while all the other input terms are identically zero. For each set of parameters, we run the simulation 10 times and average the normalized percentage of susceptible and infected population. We also make a comparison between the results both from the agent-based model (2.2) and the community-based model (2.5). The simulation results are shown in Fig. 2.2.

The simulation result in Fig. 2.2 verifies our theoretical analysis in Section 2.2.1. First of all, looking at all the four figures in Fig. 2.2, we can observe that the quantity $s_j(t)$ in (2.5) is a lower bound of that in (2.2). Similarly, the quantity $i_j(t)$ in (2.5) is an upper bound of in (2.2).

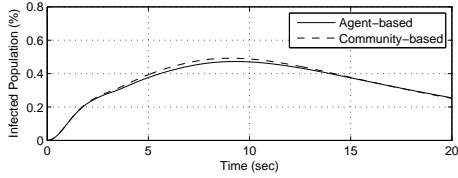
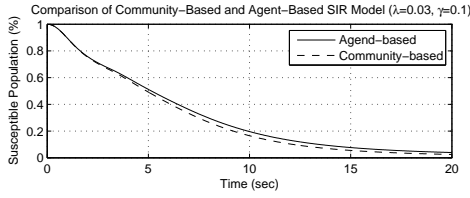
Also, the observable deviation of the community-based model (2.5) in Fig.



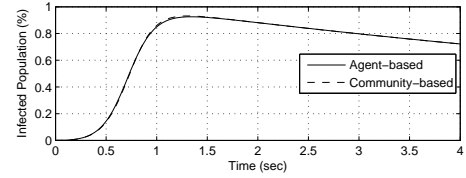
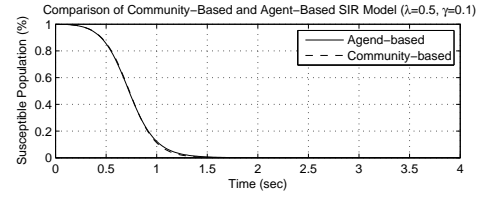
(a) $\lambda = 0.001, \gamma = 0.1$



(b) $\lambda = 0.006, \gamma = 0.1$



(c) $\lambda = 0.03, \gamma = 0.1$



(d) $\lambda = 0.5, \gamma = 0.1$

Figure 2.2: The comparison between the agent-based SIR dynamical model (A.6) and the community-based SIR dynamical model (2.5) with corresponding parameters.

2.2(b) and Fig. 2.2(c) verify that states of agents within the same community are actually not independent. Because each community is of size 100, if such independence holds true then there should be no observable deviation. Therefore the observed deviation verifies that indeed $i_j^w(t)$'s are not all independent.

Another observation is that in Fig. 2.2(a), agents are infected with a very low probability, while in Fig. 2.2(d) they are infected with a high probability. These are the two special cases that we have discussed in Section 2.2.1. As we have shown, compared with normal cases as in Fig. 2.2(b) and Fig. 2.2(c), the deviation of the model (2.5) in these two special cases are smaller than in the normal cases.

2.3 Diffusion Outbreak

Outbreak has been a topic of great interest in various studies of diffusion in networks. We will rigorously study outbreak from a mathematical perspective as well as the effect of inter-community connection strengths on outbreak of a diffusion.

2.3.1 Condition for Outbreak

Intuitively, outbreak means that even if a small fraction of the population is initially infected, the infection is still likely to reach a non-negligible portion of population without the help of an external influence, i.e., as time tends to infinity, there will be a significant percentage of agents affected by the diffusion in the case that all the input terms vanish. To make this mathematically rigorous, we will use the following definition for outbreak:

Definition 2.1 *In both the SIR model (2.5) and the SIS model (2.6), an outbreak occurs if $\exists \epsilon > 0$, such that $\forall \delta > 0$, $\exists I_0 \in [0, 1]^n \cap \{\|I_0\|_2 < \delta\}$, $I(0) = I_0$ such that $\limsup_{t \rightarrow \infty} \|1_n - S(t)\|_2 > \epsilon$. Otherwise the diffusion dies out. It should be noted that $S(t_0) = 1_n - I_0$ and $U(t) \equiv 0$. Here $1_n \in \mathbb{R}^n$ is a vector with all entries equal to 1.*

Notice that Definition 2.1 is quite similar to definitions of Lyapunov instability and stability, which actually gives us inspiration about how to determine the outbreak condition. Definition 2.1 also implies that outbreak is a property based on the network topology, instead of being dependent on initial conditions.

Theorem 2.1 *For the SIR model (2.5) and the SIS model (2.6), if the matrix $\Lambda - \Gamma \not\leq 0$ the infection breaks out. On the other hand, if $\Lambda - \Gamma \prec 0$, then the infection dies out.*

Theorem 2.1 is easy to understand intuitively. The matrix Λ represents infection rate while Γ is recovery rate. Then $\Lambda - \Gamma \prec 0$ means the infection rate is totally dominated by the recovery rate, in which case the diffusion should die out. Otherwise, if infection is not overwhelmed by recovery, i.e., $\Lambda - \Gamma \not\leq 0$, then outbreak occurs. A proof of Theorem 2.1 follows:

Proof: We primarily focus on proving Theorem 2.1 for the SIR model (2.5). A corresponding proof for the SIS model (2.6) is very similar and skipped here.

For the first part of the theorem, consider a scalar function

$$V(I(t)) = I^T(t) (\Lambda + \Lambda^T - 2\Gamma) I(t).$$

Calculating the derivative of $V(I(t))$ with respect to t , we obtain

$$\begin{aligned}
\frac{dV(I(t))}{dt} &= I^T(t) [(\Lambda + \Lambda^T - 2\Gamma) (\text{diag}[S(t)] \Lambda - \Gamma) \\
&\quad + (\text{diag}[S(t)] \Lambda - \Gamma)^T (\Lambda + \Lambda^T - 2\Gamma)] I(t) \\
&= I^T(t) (\Lambda + \Lambda^T - 2\Gamma)^2 I(t) - I^T(t) [(\Lambda + \Lambda^T - 2\Gamma) (\mathbf{I}_n - \text{diag}[S(t)]) \Lambda \\
&\quad + \Lambda^T (\mathbf{I}_n - \text{diag}[S(t)])^T (\Lambda + \Lambda^T - 2\Gamma)] I(t)
\end{aligned}$$

where \mathbf{I}_n is the identity matrix of size n .

The eigenvector associated with the greatest eigenvalue of $\Lambda + \Lambda^T - 2\Gamma$ has all its entries nonnegative and the largest eigenvalue of this matrix is positive since $\Lambda - \Gamma \not\leq 0$. This is because the matrix $\Lambda + \Lambda^T - 2\Gamma + 2\gamma_{\max} \mathbf{I}_n$, where $\gamma_{\max} = \max_{1 \leq i \leq n} \{\gamma_i\}$, is a nonnegative matrix, and by the Perron-Frobenius Theorem [96] the eigenvector associated with the greatest eigenvalue of $\Lambda + \Lambda^T - 2\Gamma + 2\gamma_{\max} \mathbf{I}_n$ is nonnegative while $\Lambda + \Lambda^T - 2\Gamma + 2\gamma_{\max} \mathbf{I}_n$ and $\Lambda + \Lambda^T - 2\Gamma$ share the same eigenvectors. Therefore $\exists x \in [0, 1]^n$ such that $x^T (\Lambda + \Lambda^T - 2\Gamma) x > 0$.

We proceed using a contradiction argument. Suppose the infection dies out, then for a small enough $\epsilon < \frac{\sigma_{\min}(\Lambda + \Lambda^T - 2\Gamma)}{2\|\Lambda\|_2}$ we can find a choice of $I(0) = I_0 \in [0, 1]^n$, such that $\|I(t)\|_2 < \epsilon$ for any $t \geq 0$, where $\sigma_{\min}(\cdot)$ denotes the smallest nonzero singular value. With such an ϵ , we will have $\frac{dV(I(t))}{dt} > 0$ if $V(I(t)) > 0$. Also since $\Lambda + \Lambda^T - 2\Gamma$ has positive eigenvalues, we have $V(0) = 0$ and $V(I_0) > 0$ for some I_0 with arbitrarily small $\|I_0\|_2$. Therefore, there $\exists \epsilon' > 0$ so that $\forall \delta > 0$, we can find a $I(0) = I_0 \in [0, 1]^n \cap \{\|I_0\|_2 < \delta\}$ so that $\|I(t')\|_2 > \epsilon'$ for some $t' \geq 0$, then $\lim_{t \rightarrow \infty} \|R(t)\|_2 > \|I(t')\|_2 > \epsilon'$, i.e., the infection does not die out. This is a contradiction.

Next we consider the second part of this theorem. Since $\forall t \geq 0, S(t) < 1_n$, from equation (2.5b) we have

$$\frac{dI(t)}{dt} = \text{diag}[S(t)] \Lambda I(t) - \Gamma I(t) < (\Lambda - \Gamma)I(t)$$

Then by the Comparison Lemma [97], we have $I(t) < \exp((\Lambda - \Gamma)t) I_0$, where $I(0) = I_0$. Because $\Lambda - \Gamma \prec 0$ by equation (2.5c), we find that

$$\begin{aligned} \lim_{t \rightarrow \infty} \|R(t)\|_2 &= \left\| \int_0^\infty \Gamma I(t) dt \right\|_2 \leq \int_0^\infty \|\Gamma\|_2 \|I(t)\|_2 dt \\ &\leq \int_0^\infty \|\Gamma\|_2 \|I_0\|_2 \|\exp((\Lambda - \Gamma)t)\|_2 dt \\ &\leq \int_0^\infty \|\Gamma\|_2 \|I_0\|_2 \exp\left(\frac{-1}{\|(\Lambda - \Gamma)^{-1}\|_2} t\right) dt \\ &= \|\Gamma\|_2 \|I_0\|_2 \|(\Lambda - \Gamma)^{-1}\|_2 \end{aligned}$$

Therefore, $\lim_{t \rightarrow \infty} \|R(t)\|_2$ can be made arbitrarily small with a choice of small enough $\|I_0\|_2$. The infection dies out according to Definition 2.1. □

It should be noted that in the SIR model, since $S(t)$ is monotone decreasing, $\|1_n - S(t)\|_2$ is convergent as $t \rightarrow \infty$. Therefore the statement $\limsup_{t \rightarrow \infty} \|1_n - S(t)\|_2 > \epsilon$ in Definition 2.1 can be replaced with $\lim_{t \rightarrow \infty} \|1_n - S(t)\|_2 > \epsilon$ for the SIR model and Theorem 2.1 still holds true.

If we view each individual as a separate community and consider the uniform parameter case, i.e., $\Lambda = \lambda A$ and $\Gamma = \gamma \mathbf{I}_n$, where $A \in \{0, 1\}^{n \times n}$ is the adjacency matrix and $\mathbf{I}_n \in \mathbb{R}^{n \times n}$ is an identity matrix, then Theorem 2.1 is reduced the epidemic threshold in Section 5.1 of [30]. Therefore, Theorem 2.1 generalizes this existing result on diffusion outbreak threshold.

2.3.2 Effect of Inter-Community Connections

Theorem 2.1 specifies a condition for infection outbreak. Next, we will employ Theorem 2.1 to gain insights on the contribution of inter-community connections to propagations infections.

In Theorem 2.1, the condition $\Lambda - \Gamma \not\leq 0$ means that the greatest eigenvalue of $\Lambda - \Gamma$ is positive. To investigate the effects of inter-community connections, we focus on the largest eigenvalue of $\Lambda - \Gamma$. Let $\Lambda - \Gamma = (\Lambda_D - \Gamma) + \Lambda_{LU}$, where Λ_D denotes the diagonal matrix with diagonal entries of Λ and $\Lambda_{LU} = \Lambda - \Lambda_D$. Denote the largest eigenvalue by $\lambda_{\max}(\cdot)$. The following theorem addresses the effects of the connections.

Theorem 2.2 *The inequality $\lambda_{\max}(\Lambda_D - \Gamma) < \lambda_{\max}((\Lambda_D - \Gamma) + \Lambda_{LU})$ holds true, where $\lambda_{\max}(\cdot)$ denotes the largest eigenvalue. In other words, the inter-community connections increase the chance of an infection outbreak.*

Proof: First, consider $\Lambda_D - \Gamma + \gamma_{\max} \mathbf{I}_n$, where $\gamma_{\max} = \max_{1 \leq i \leq n} \{\gamma_i\}$. Since $\Lambda_D - \Gamma + \gamma_{\max} \mathbf{I}_n$ and Λ_{LU} are both nonnegative, it follows that $|\Lambda_D - \Gamma + \gamma_{\max} \mathbf{I}_n| < |\Lambda_D - \Gamma + \Lambda_{LU} + \gamma_{\max} \mathbf{I}_n|$. By the Wielandt's Theorem [98], $\rho(\Lambda_D - \Gamma + \gamma_{\max} \mathbf{I}_n) < \rho(\Lambda_D - \Gamma + \gamma_{\max} \mathbf{I}_n + \Lambda_{LU})$, in which $\rho(\cdot)$ is spectral radius. By the Peron-Frobenius theorem [96], the largest eigenvalues of both of these matrices are positive, and we

have

$$\begin{aligned}
& \lambda_{\max}(\Lambda_D - \Gamma) - \lambda_{\max}((\Lambda_D - \Gamma) + \Lambda_{LU}) \\
&= [\lambda_{\max}(\Lambda_D - \Gamma + \gamma_{\max}\mathbf{I}_n) - \gamma_{\max}] - [\lambda_{\max}((\Lambda_D - \Gamma) + \gamma_{\max}\mathbf{I}_n + \Lambda_{LU}) - \gamma_{\max}] \\
&= \rho(\Lambda_D - \Gamma + \gamma_{\max}\mathbf{I}_n) - \rho(\Lambda_D - \Gamma + \gamma_{\max}\mathbf{I}_n + \Lambda_{LU}) < 0
\end{aligned}$$

□

Theorem 2.2 corroborates an intuitive fact: strong connectivity among communities enhances the propagation of the infection. Next we investigate the case of weak interconnections to get further insight into effects of such interconnections.

2.3.3 Weak Interconnections: A Special Case

To further investigate effects of inter-community connections on outbreaks, we consider the following simple case. Let $\Lambda = \Lambda_D + \Lambda_{LU}$, where $\Lambda_D = \lambda\mathbf{I}_n$ and $\Lambda_{LU} = \lambda'A$. Here $A = [a_{ij}]_{n \times n} \in \{0, 1\}^{N \times N}$ is the adjacency matrix, i.e., if community j is connected to community i then $a_{ij} = 1$ and otherwise $a_{ij} = 0$. Also we assume that all agents have the same recovery rate, i.e., $\Gamma = \gamma\mathbf{I}_n$.

For the weak interconnections case, we assume that $\lambda \gg \lambda'$. This means that the infection rate is much higher within communities than among them (hence the reference to “weak-interconnections”). Let $\tau = \frac{\lambda}{\gamma}$, $\tau' = \frac{\lambda'}{\gamma}$. If $\tau > 1$, then according to Theorem 2.1, the infection will break out even if $\tau' = 0$, i.e., even if there is no network connection. Thus we assume that $\tau < 1$, which means that the infection will die out without the help of inter-community connections, and we study the minimum value of τ' to trigger infection outbreak.

First we recall the following useful lemma.

Lemma 2.1 *If A is adjacency matrix for a graph, then the largest eigenvalue of A satisfies*

$$\bar{d} < \lambda_{\max}(A) < d_{\max} \quad (2.8)$$

where \bar{d} and d_{\max} are the average and maximum degree of the graph.

Proof: See [99]. □

With Lemma 2.1, we easily determine the following condition for infection outbreak in the weak interconnections case.

Theorem 2.3 *For the case of weak interconnections, if $\tau' > \frac{1-\tau}{d_{\max}}$, the infection breaks out.*

Proof: According to Theorem 2.1, infection outbreak occurs if

$$\lambda_{\max}(\lambda \mathbf{I}_n + \lambda' A - \gamma \mathbf{I}_n) = \gamma \lambda_{\max}[\tau' A - (1-\tau)\mathbf{I}_n] > 0$$

This is equivalent to the condition $\tau' > \frac{1-\tau}{\lambda_{\max}(A)}$. Using the upper bound on $\lambda_{\max}(A)$ from Lemma 2.1, we have

$$\tau' > \frac{1-\tau}{\lambda_{\max}(A)} > \frac{1-\tau}{d_{\max}} \quad (2.9)$$

□

For example, if the graph associated with A has chain structure, i.e., all the communities are arranged in a line, then we have $\bar{d} = 2 - \frac{2}{n}$ and $d_{\max} = 2$ when $n \geq 3$. For sufficiently large n , \bar{d} is close to d_{\max} and thus Lemma 2.1 provides an accurate estimate for $\lambda_{\max}(A)$. Theorem 2.3 specifies the condition $\tau' > \frac{1-\tau}{d_{\max}}$ and in some special cases Theorem 2.3 is accurate for determining the minimum interconnection strength τ' to result in infection outbreak.

2.4 Final State of the Diffusion

After determining the condition of infection outbreaks, it is natural to ask what will it be like at the end of a diffusion. What's the percentage of population involved in the diffusion in each community? Will all communities be influenced by the diffusion? We will answer these questions in this section.

2.4.1 Final State

First, we establish a connection between the initial state of a diffusion and the corresponding final state. It should be noted that in the SIS model (2.6), the final state is irrelevant with the initial state, hence we focus on the SIR model (2.5) here. The following theorem provides such a connection.

Theorem 2.4 *Consider the SIR model (2.5) with initial condition $S(0) = 1_n - I_0$, $I(0) = I_0$ and $R(0) = 0$ where $I_0 \in [0, 1]^n$, and suppose $\lim_{t \rightarrow \infty} S(t) = \tilde{S}$, $\lim_{t \rightarrow \infty} I(t) = \tilde{I}$ and $\lim_{t \rightarrow \infty} R(t) = \tilde{R}$. Then the following equations hold:*

$$\tilde{I} = 0 \tag{2.10a}$$

$$\tilde{S} = 1_n - \tilde{R} \tag{2.10b}$$

$$\tilde{S} = \text{diag}[1_n - I_0] \exp(-\Lambda \Gamma^{-1} \tilde{R}) \tag{2.10c}$$

where we define that $\exp\left([x_1, x_2, \dots, x_n]^T\right) := [\exp(x_1), \exp(x_2), \dots, \exp(x_n)]^T$ for $[x_1, x_2, \dots, x_n]^T \in \mathbb{R}^n$ and $\text{diag}[\cdot]$ is the diagonal matrix with each entry of \cdot on its diagonal line.

Proof: As time tends to infinity, all the infected individuals finally get recovered. Therefore equations (2.10a) and (2.10b) are trivial. With use of equation (2.5a), we have

$$\text{diag} [S(t)]^{-1} d S(t) = -\Lambda I(t) dt \quad (2.11)$$

We use the notation $\log \left([x_1, x_2, \dots, x_n]^T \right) := [\log(x_1), \log(x_2), \dots, \log(x_n)]^T$

for vectors

$[x_1, x_2, \dots, x_n]^T \in \mathbb{R}^n$. Using equation (2.5c), equation (2.11) can be rewritten as

$$d \log (S(t)) = -\Lambda \Gamma^{-1} d R(t) \quad (2.12)$$

Integrating (2.12) from $t = 0$ to ∞ gives

$$\log \left(\text{diag} [1_n - I_0]^{-1} \tilde{S} \right) = -\Lambda \Gamma^{-1} \tilde{R}$$

which is equivalent to (2.10c). □

2.4.2 Outbreak Impact Within Communities

If an infection breaks out, it is natural to ask whether all the communities are affected. We are interested in whether or not it is possible that an outbreak occurs for the overall social network, but some communities are affected by the outbreak while other communities have no internal outbreak. To rigorously study this issue we first provide a definition on whether or not a community is affected during an infection outbreak:

Definition 2.2 *The j -th community ($1 \leq j \leq n$) is affected by infection outbreak if there exists an $\epsilon_j > 0$ such that under any initial condition which leads to an*

infection outbreak, we have $\lim_{t \rightarrow \infty} \|1 - s_j(t)\|_2 > \epsilon_i$.

The following theorem concerns the impact of an outbreak within the communities.

Theorem 2.5 *In both the SIR model (2.5) and the SIS model (2.6), if Λ is irreducible and $\Lambda - \Gamma \not\leq 0$, then there exists $\epsilon_j > 0$ for each $1 \leq j \leq n$ such that with any initial condition that $S(0) = 1_n - I_0$, $I(0) = I_0$ (and $R(0) =$ for the SIR model (2.5)), we have $\lim_{t \rightarrow \infty} |1 - s_j(t)| > \epsilon_j$.*

Proof: The assertion can be proved similarly for the SIR model (2.5) and the SIS model (2.6), so we present a proof only for the SIR model (2.5). Let $\tilde{r}_j = \lim_{t \rightarrow \infty} r_j(t)$ for $1 \leq j \leq n$. We will first prove that $\tilde{r}_j > 0$ for all $1 \leq j \leq n$.

Since the infection breaks out, $\exists 1 \leq k_1 \leq n$ s.t. $\tilde{r}_{k_1} > 0$. If $\exists k_m$ and $\tilde{r}_{k_m} = 0$, since Λ is irreducible, therefore $\exists k_2, \dots, k_{m-1}$ such that $k_1, k_2, \dots, k_m \in \{1, 2, \dots, n\}$ are distinct and for any $2 \leq l \leq m$, we have $\lambda_{k_{l-1}, k_l} > 0$. Using the k_l -th row of equation (2.10c), easily if $r_{k_l \infty} = 0$ then $r_{k_{l-1} \infty} = 0$. Letting l vary from m down to 2, we have $r_{k_1 \infty} = 0$, which is a contradiction. Therefore $r_{j \infty} > 0$ for any $1 \leq j \leq n$.

Since I_0 can be taken sufficiently small, equation (2.10c) can be approximated as $1_n - \tilde{R} = \exp(-\Lambda \Gamma^{-1} \tilde{R})$. Since the function $f(\tilde{R}) = \exp(-\Lambda \Gamma^{-1} \tilde{R}) + \tilde{R} - 1$ is analytic, there is a finite number of solutions to the equation $1_n - \tilde{R} = \exp(-\Lambda \Gamma^{-1} \tilde{R})$ on $[0, 1]^n$. For any $1 \leq j \leq n$, $\tilde{r}_j > 0$ and there are finite solutions, then $\exists \epsilon_j = \frac{1}{2} \min \{\tilde{r}_j \mid \tilde{r}_j \text{ satisfies (2.10)}\}$. With such a choice of ϵ_j , according to Definition 2.2, all the communities are affected in the course of the infection outbreak.

□

Theorem 2.5 illustrates that if network topology is strongly connected, then all the communities are affected during any infection outbreak. On the other hand, if the topology is not strongly connected, i.e. there exist two communities between which there is no connecting path, such global effects cannot be guaranteed.

2.4.3 Bounds on the Final State

Since there does not exist a closed-form expression for a diffusion's final state for either the SIR model (2.5) and the SIS model (2.6), hence it is helpful to estimate the final state by coming up with an upper bound and lower bound. In Section IV.B of [29], Van Mieghem *et al.* analyze bounds of the steady state in their agent-based SIS model and their results can be easily extended to our community-based SIS model (2.6), therefore we primarily focus on the SIR model (2.5) here. The following theorem gives an upper bound of final states in the SIR model (2.5).

Theorem 2.6 *Consider a sequence $\{v_n\}$ with $v_1 = \exp(-\Lambda\Gamma^{-1}1_n)$, and for any $k \geq 1$, $v_{k+1} = \exp(-\Lambda\Gamma^{-1}(1_n - v_k))$. Note that $\{v_k\} \subset [0, 1]^n$, then for any $k \geq 1$, we have $\tilde{R} < 1 - v_k$, where $\tilde{R} \in \mathbb{R}^n$ represents the percentage of finally recovered population in each community.*

Proof: Construct a function series $\{f_k(\cdot)\}$ by the following procedure: for any $k \geq 0$ where $f_k : \mathbb{R}^n \mapsto \mathbb{R}^n$. First, let $f_0(x) = 1 - x$, then for any $k \geq 0$, let $f_{k+1}(x) = \exp(-\Lambda\Gamma^{-1}(1_n - f_k(x)))$.

If \tilde{R} is a solution to equation (2.10), it's easy to verify $f_0(\tilde{R}) = f_1(\tilde{R}) = \dots = f_k(\tilde{R}) = \dots$. We can also find that $v_k = f_k(1_n)$. Therefore for any $\forall k \geq 1$, $1 - \tilde{R} = f_0(\tilde{R}) = f_k(\tilde{R}) > f_k(1_n) = v_k$, and therefore $\tilde{R} < 1_n - v_k$. \square

Notice that because $\Lambda\Gamma^{-1}$ is a nonzero matrix, therefore $v_k < v_{k+1}$ for any $k \geq 1$, i.e. $\{v_k\}$ is monotonically increasing. Thus $\{1_n - v_k\}$ is a sequence of increasingly tighter upper bounds for \tilde{R} . The following theorem concerns lower bounding \tilde{R} .

Theorem 2.7 *For nonzero \tilde{R} satisfying equation (2.10), we have the inequality*

$$\Lambda\Gamma^{-1}\tilde{R} > 2 \operatorname{diag} [\Lambda\Gamma^{-1}1_n]^{-1} (\Lambda\Gamma^{-1} - \mathbf{I}_n) 1_n$$

Proof: Since $\Lambda\Gamma^{-1}$ is nonnegative, according to Theorem 2.5, \tilde{R} is positive, and $-\Lambda\Gamma^{-1}\tilde{R}$ is negative. Since the second order term in Taylor expansion of exponential of a negative variable is greater than the exponential itself, we have

$$\begin{aligned} 1_n &= \tilde{R} + \exp \left[-\Lambda\Gamma^{-1}\tilde{R} \right] \\ &< \tilde{R} + 1_n - \Lambda\Gamma^{-1}\tilde{R} + \frac{1}{2} \operatorname{diag} [\Lambda\Gamma^{-1}\tilde{R}] \Lambda\Gamma^{-1}\tilde{R} \end{aligned} \quad (2.13)$$

Simplifying inequality (2.13), we have

$$\frac{1}{2} \operatorname{diag} [\Lambda\Gamma^{-1}1_n] \Lambda\Gamma^{-1}\tilde{R} > (\Lambda\Gamma^{-1} - \mathbf{I}_n) 1_n \quad (2.14)$$

Inequality (2.14) is equivalent to the inequality in this theorem. \square

With Theorem 2.7, we find a bound for $\tilde{r}_{\max} := \max_{1 \leq i \leq n} \{\tilde{r}_i\}$ with the following corollary.

Corollary 2.1 *The quantity \tilde{r}_{\max} satisfies*

$$\tilde{r}_{\max} > 2 \max \left\{ \text{diag} [\Lambda \Gamma^{-1} \mathbf{1}_n]^{-2} (\Lambda \Gamma^{-1} - \mathbf{I}_n) \mathbf{1}_n \right\},$$

where $\max\{\cdot\}$ indicates the maximum entry of a vector.

Proof: Since $\Lambda \Gamma^{-1}$ and \tilde{R} are nonnegative, then $\Lambda \Gamma^{-1} \tilde{R} < \tilde{r}_{\max} \Lambda \Gamma^{-1} \mathbf{1}_n$. Combining this inequality with Theorem 2.7, the result follows. \square

Theorem 2.7 actually specifies a polyhedron as the bound for \tilde{R} . Corollary 2.1 gives a way to simplify Theorem 2.7 to obtain further insight on the lower bound of \tilde{r}_{\max} .

2.5 A View from Percolation Theory

It should be noted that our community-based model (2.5) is an approximation of the agent-based model (2.3). As noted by Youssef *et al.* in [30], even the agent-based model is also an approximation of actual diffusion dynamics in social networks. On the grounds of this observation, one may question the validity of our analysis in Section 2.3 and 2.4, which is conducted based on the approximated community-based model (2.5) and (2.6). Here, we are going to analyze diffusions in the community-based network as specified in Section 2.1 via the percolation theory [100] in order to validate our previous analysis.

Percolation theory is a favorite tool that physicists use to analyze complex networks. A representative problem in percolation theory is the following: if we pour water on the upper surface of an object made of a porous material, will the

water reach the bottom side of the object? The material can be mathematically modeled as a lattice of $m \times m \times m$ small cubes, with every two neighboring cubes either open to each other (i.e., water can flow from one to another) with probability p or closed to each other with probability $1 - p$, and these connection probabilities are all independent. In percolation theory, we are interested in finding the relationship between the probability that water flows through the material and p . It happens that under general circumstances (for very large lattices), there is a critical value p_c of p such that for $p > p_c$, water flows through with probability 1 and for smaller values of p there is 0 probability of water flowing through to the bottom.

More generally, in a large network as p increases from 0 through the critical threshold p_c , large clusters and long-range connectivity in the network appear. More general percolation problems may involve multiple such probabilities p_1, p_2 , etc. For example in our problem studied in this section, we have a probability matrix, instead of a scalar.

Newman [24] appears to have been the first to make a connection between percolation theory and epidemic spread in networks. Leicht *et al.* [101] extends percolation theory to an interacting network, which is similar to our community-based network that was presented in Section 2.1. Allard *et al.* [102] relate percolation to epidemic spreading in networks with community structure.

To maintain the readability of this chapter, we will briefly explain the connection between percolation theory and the infection diffusion of Section 2.5.1, and we will summarize the solution proposed by Leicht *et al.* [101] and Allard *et al.* [102] in Section 2.5.2. In Sections 2.5.3 and 2.5.4, we compare the results from our work

using dynamical models and from percolation theory. Also we investigate the conditions under which these results using alternate approaches agree. This comparison strengthens our analysis in Sections 2.3 and 2.4. Moreover, it provides further insights about the underlying assumptions in our work.

2.5.1 Transmissibility and Percolation

We consider a pair of connected agents, one of which (j) is infected while the other (i) is susceptible. Suppose the infection rate from j to i is λ_{ij} and the infected agent will remain infected for a duration τ_j . Then the probability that the disease will be transmitted from agent j to agent i , denoted T'_{ij} , is

$$T'_{ij} = 1 - \lim_{\Delta t \rightarrow 0} (1 - \lambda_{ij} \Delta t)^{\tau_j / \Delta t} = 1 - \exp(-\lambda_{ij} \tau_j) \quad (2.15)$$

In the SIR model, the infection rate is exactly λ_{ij} in (2.15) while the time τ_j is a random variable with an exponential distribution with parameter γ_j . Therefore the average transmission probability is

$$T_{ij} = \int_0^{\infty} \exp(-\gamma_j \tau_j) (1 - \exp(-\lambda_{ij} \tau_j)) d\tau_j = \frac{\lambda_{ij}}{\lambda_{ij} + \gamma_j} \quad (2.16)$$

Here we call T_{ij} the transmissibility of the infection. Notice that in our interconnected community model, the infection transmission from any agent in community j to another in community i is only dependent on the community indices i and j . Therefore we use T_{ij} to indicate the transmissibility from community j to i .

In this section we extend the fact [24] that we can simply use the expectation

(mean value) of the probability of infection spread between communities. Infection will propagate as if all transmission probabilities from community j to i were equal to T_{ij} . *The truth of this result in the single community case is demonstrated in [24], and we apply it in a general interconnected community context.*

Imagine that an infection outbreak starts from a single infected agent and the infection spreads with probability T_{ij} from community j to i . We retain the edge from community j to community i with probability T_{ij} and discard it with probability $1 - T_{ij}$. The remaining edges thus form an “infected network.” Then the final size of an outbreak is exactly the size of the cluster of vertices that is connected from the initial infected vertex in the “infected network.” Therefore the epidemic model here is equivalent to the percolation model on an interacting network, which is studied by Leicht *et al.* [101].

2.5.2 Solution to the Percolation Problem

We consider a system with $n \geq 2$ interacting networks. Any network i with $1 \leq i \leq n$ enjoys a multi-degree distribution $\{p_{k_1, k_2, \dots, k_n}^i\}$, where $p_{k_1, k_2, \dots, k_n}^i$ is the fraction of all vertices in network i that have k_1 connections in network 1, k_2 connections in network 2, etc. This degree distribution can be represented using a generating function

$$G_i(X) = \sum_{k_1, k_2, \dots, k_n=0}^{\infty} p_{k_1, k_2, \dots, k_n}^i x_1^{k_1} x_2^{k_2} \dots x_n^{k_n}$$

with $X = [x_1, x_2, \dots, x_n]^T$.

We need to emphasize that our system of networks is considered in the limit

of large size (i.e., there are infinitely many agents in each network) and are totally random in all aspects but obey the joint multi-degree distribution $\{p_{k_1, k_2, \dots, k_N}^i\}$. All the results given in this section are averaged over the ensemble of all possible network systems, in the limit of large network size.

A further observation that will prove crucial is the following: Consider a randomly selected edge from a vertex in network j to another vertex in network i (an $i - j$ edge). If we follow an $i - j$ edge to its end (a vertex), then that vertex is more likely to have more connections to network j than a randomly chosen vertex in network i , since the high-connection vertices have more edges connected to them than the low-connection ones. More specifically, let $q_{k_1 \dots k_j \dots, k_n}^{ij}$ denote the probability of following a randomly chosen $i - j$ edge to a vertex with k_1 connections in network 1, k_2 connections in network 2, etc., then $q_{k_1 \dots k_j \dots, k_n}^{ij} \propto (k_j + 1) p_{k_1 \dots k_j + 1 \dots, k_n}^i$ and the generating function for $\{q_{k_1 \dots k_j \dots, k_n}^{ij}\}$ is

$$\begin{aligned}
G_{ij}(X) &= \sum_{k_1, k_2, \dots, k_n=0}^{\infty} q_{k_1, k_2, \dots, k_n}^i x_1^{k_1} x_2^{k_2} \dots x_n^{k_n} \\
&= \sum_{k_1, k_2, \dots, k_n=0}^{\infty} \frac{(k_j + 1) p_{k_1 \dots k_j + 1 \dots, k_n}^i}{\sum_{l_1, l_2, \dots, l_n=0}^{\infty} (l_j + 1) p_{l_1 \dots l_j + 1 \dots, l_n}^i} x_1^{k_1} x_2^{k_2} \dots x_n^{k_n} \\
&= \left(\sum_{l_1, l_2, \dots, l_n=0}^{\infty} l_j p_{l_1, l_2, \dots, l_n}^i \right)^{-1} \frac{\partial}{\partial x_j} \sum_{k_1, k_2, \dots, k_n=0}^{\infty} p_{k_1, k_2, \dots, k_n}^i x_1^{k_1} x_2^{k_2} \dots x_n^{k_n} \\
&= \frac{\partial G_i(X) / \partial x_j}{\partial G_i(X) / \partial x_j |_{X=1_n}}
\end{aligned}$$

In order to solve the percolation problem, we also need generating functions $G_i(X|T)$ and $G_{ij}(X|T)$ for degree distributions on the ‘‘infected networks’’ which is described in Section 2.5.1, corresponding to $G_i(X)$ and $G_{ij}(X)$ on the original

networks. These are simple to derive. $G_i(X|T)$ is calculated as follows:

$$\begin{aligned}
& G_i(X|T) \\
&= \sum_{k_1, k_2, \dots, k_n=0}^{\infty} \sum_{l_1, l_2, \dots, l_n=0}^{l_1=k_1, l_2=k_2, \dots, l_n=k_n} p_{k_1, k_2, \dots, k_n}^i \prod_{m=1}^n \binom{k_m}{l_m} T_{mi}^{l_m} (1 - T_{mi})^{k_m - l_m} x_m^{l_m} \\
&= \sum_{k_1, k_2, \dots, k_n=0}^{\infty} p_{k_1, k_2, \dots, k_n}^i \sum_{l_1, l_2, \dots, l_n=0}^{l_1=k_1, l_2=k_2, \dots, l_n=k_n} \prod_{m=1}^n \binom{k_m}{l_m} (x_m T_{mi})^{l_m} (1 - T_{mi})^{k_m - l_m} \quad (2.17) \\
&= \sum_{k_1, k_2, \dots, k_n=0}^{\infty} p_{k_1, k_2, \dots, k_n}^i \prod_{m=1}^n (1 - T_{mi} + x_m T_{mi})^{k_m} \\
&= G_i(1_n + \text{diag}[T_{1i}, T_{2i}, \dots, T_{ni}](X - 1_n))
\end{aligned}$$

Through similar calculations in (2.17), we have

$$G_{ij}(X|T) = G_{ij}(1_n + \text{diag}[T_{1i}, T_{2i}, \dots, T_{ni}](X - 1_n))$$

Now we are ready to calculate the size distribution of small components in the network ensemble. A component is any closed connected cluster of vertices. Here, ‘‘small’’ means the intensive component (i.e., the component that does not scale with the network size). Define $H_{ij}(X|T)$ as the generating function for the size distribution of the component reached by following an $i - j$ edge.

Small components are typically finite and therefore the chance that any two small components that are connected to the same vertex have an edge between them goes as the inverse of the size of the whole network, and therefore is 0 in the limit of a large system of networks [24]. In other words the small component should be completely tree-like in structure.

Further, the cluster reached by an $i - j$ edge may be either 1) a single vertex without any other outgoing edges connected to it or 2) a single vertex connected by any number $m \geq 1$ of edges other than the one we reach it by, each leading to

another cluster whose size distribution is generated by $H(X|T)$ as well. With all the previous analysis, $H_{ij}(X|T)$ can be solved through the following self-consistent equation:

$$H_{ij}(X) = x_i G_{ij}(H_{1i}(X), H_{2i}(X), \dots, H_{ni}(X)) \quad (2.18)$$

Then the size of clusters connected from a randomly chosen vertex has its distribution generated by

$$H_i(X) = x_i G_i(H_{1i}(X), H_{2i}(X), \dots, H_{ni}(X)) \quad (2.19)$$

2.5.3 Condition for Infection Outbreak

As noted in Section 2.5.1, percolation theory can be applied in studying infection propagation. Using percolation theory to determine an infection outbreak condition over the interacting networks as specified in Section 2.5.1, we have the following theorem.

Theorem 2.8 *Define an $n \times n$ block matrix F , whose blocks $[F_{ij}]$ are also $n \times n$ matrices with*

$$[F_{ij}]_{pq} = \delta_{iq} \cdot \frac{\partial}{\partial x_j} G_{ip}(X|T) \Big|_{X=1_n}$$

Denote by \mathbf{I}_{n^2} the identity matrix. If $\mathbf{I}_{n^2} - F \succ 0$, the infection does not break out. If $\mathbf{I}_{n^2} - F \not\succeq 0$, the infection breaks out. (Intuitively $\lambda_{\max}(F)$ is a measure of the strength of the infection propagation.)

Proof: First, for any randomly chosen agent from community i , we compute the average number $\langle s_i \rangle_j$ of agents from community j that are connected to it:

$$\langle s_i \rangle_j = \left. \frac{\partial}{\partial x_j} H_i(X|T) \right|_{X=1_n} = \delta_{ij} + \sum_{l=1}^n \frac{\partial}{\partial x_l} G_i(X|T) \frac{\partial}{\partial x_j} H_{li}(X|T) \Big|_{X=1_n} \quad (2.20)$$

From (2.20), to compute each $\langle s_i \rangle_j$, we need to find $\left. \frac{\partial}{\partial x_l} H_{ij}(X|T) \right|_{X=1_n}$ for each i, j and l . In order to compute these quantities, we first calculate every partial derivative of each side of the equation

$$H_{ij}(X|T) = x_i G_{ij} [H_{1i}(X|T), H_{2i}(X|T), \dots, H_{ni}(X|T)],$$

after which we obtain the following linear equations:

$$\left. \frac{\partial}{\partial x_l} H_{ij}(X|T) \right|_{X=1_n} = \delta_{il} + \sum_{k=1}^n \frac{\partial}{\partial x_k} G_{ij}(X|T) \frac{\partial}{\partial x_l} H_{ki}(X|T) \Big|_{X=1_n} \quad (2.21)$$

$\forall 1 \leq i, j, l \leq n$.

To write (2.21) more compactly, define $h^l = [h_{11}^l \dots h_{1n}^l, \dots, h_{n1}^l \dots h_{nn}^l] \in \mathbb{R}^{n^2}$, with each entry $h_{pq}^l = \left. \frac{\partial}{\partial x_l} H_{pq}(X|T) \right|_{X=1_n}$. Also for convenience we define $\Delta^l = [\delta_{1l}^{(1)} \dots \delta_{1l}^{(n)} \dots \delta_{nl}^{(1)} \dots \delta_{nl}^{(n)}]$ where $\delta_{pq}^{(k)} = 1$ when $p = q$ and $\delta_{pq}^{(k)} = 0$ when $p \neq q$.

With these preparations, equation (2.21) can be written in the form

$$h^l = \Delta^l + F \cdot h^l \quad (2.22)$$

with $1 \leq l \leq n$.

If $\mathbf{I}_{n^2} - F \succ 0$, since F is nonnegative, then F is a convergent matrix, then by (2.22) combined with [103], we have $h^l = (I - F)^{-1} \Delta^l = (I + F + F^2 + \dots) \Delta^l$, and therefore h^l is finite and nonnegative. In this case, all the components across

the network are of a finite size. That is to say, the infection does not break out in this case.

However, as $\lambda_{\max}(F)$ approaches 1 from below, $h^l = (I + F + F^2 + \dots) \Delta^l$ will also increase and the component size will be correspondingly larger. At the point when $\lambda_{\max}(F) = 1$, the component size h starts to diverge. The condition $\lambda_{\max}(F) = 1$ is thus the point of onset of infection outbreak. Therefore, when $\mathbf{I}_{n^2} - F \not\leq 0$ the infection breaks out.

□

We need to emphasize that the interacting networks structure as specified in Section 2.5.1 is more general than the community-based network of Section 2.1. Hence Theorem 2.8 deals with a more general case than Theorem 2.1. However, the next corollary shows that if we apply Theorem 2.8 to the community-based network of Section 2.1, we will get the same result as Theorem 2.1 under certain approximations.

Corollary 2.2 *Applying Theorem 2.8 to the community-based network specified in Section 2.1 with infection rate λ'_{ij} significantly smaller than the recovery rate γ'_j (i.e. $\frac{1}{\lambda'_{ij} + \gamma'_j} \approx \frac{1}{\gamma'_j}$) for all $1 \leq i, j \leq n$, Theorem 2.8 reduces to Theorem 2.1.*

Proof: Because the communities that constitute the network described in Section 2.1 are all random networks, therefore all the degree distributions should be Poisson

[104], i.e., $G_i(X) = \prod_{k=1}^n \exp(-N_i p_{ki} (1 - x_k))$. It is easy to show that

$$G_{ij}(X) = \prod_{k=1}^N \exp(-N_k p_{ik} (1 - x_k)) = G_i(X)$$

. Therefore, for all $1 \leq i, j, k \leq n$,

$$\left. \frac{\partial}{\partial x_k} G_{ij}(X|T) \right|_{X=1_n} = T_{ki} \left. \frac{\partial}{\partial x_k} G_{ij}(X) \right|_{X=1_n} = \frac{\lambda'_{ik} p_{ik} N_k}{\lambda'_{ik} + \gamma'_k}$$

Since $\lambda'_{ij} \ll \gamma'_i$ for all $1 \leq i, j \leq n$, we have

$$\left. \frac{\partial}{\partial x_k} G_{ij}(X|T) \right|_{X=1_n} = \frac{\lambda'_{ik} p_{ik} N_k}{\lambda'_{ik} + \gamma'_k} \approx \frac{\lambda'_{ik} p_{ik} N_k}{\gamma'_k}$$

Moreover, noting that $H_{ij}(X|T) = H_i(X|T)$ for all $1 \leq i, j \leq n$, equation (2.21) can be simplified to

$$\begin{aligned} \left. \frac{\partial}{\partial x_j} H_i(X|T) \right|_{X=1_n} &= \delta_{ij} + \sum_{k=1}^N \left. \frac{\partial}{\partial x_k} G_i(X|T) \frac{\partial}{\partial x_j} H_k(X|T) \right|_{X=1_n} \\ &\approx \delta_{ij} + \frac{\lambda'_{ik} p_{ik} N_k}{\gamma'_k} \left. \frac{\partial}{\partial x_j} H_k(X|T) \right|_{X=1_n} \\ &= \delta_{ij} + \frac{\lambda_{ik}}{\gamma_k} \cdot \left. \frac{\partial}{\partial x_j} H_k(X|T) \right|_{X=1_n} \end{aligned} \quad (2.23)$$

for $\forall 1 \leq i, j \leq n$.

For any fixed j , equation (2.23) is a set of n linear equations. To write it in a compact form, let $h^j = [h_1^j, h_2^j, \dots, h_n^j]^T$ with $h_i^j = \left. \frac{\partial}{\partial x_j} H_i(X|T) \right|_{X=1_n}$ for all $1 \leq i \leq n$. Let $F = [f_{ij}]_{n \times n} \in \mathbb{R}^{n \times n}$, with $f_{ij} = \frac{\lambda_{ij}}{\gamma_i}$. Let $\Delta^j = [\delta_{1j}, \delta_{2j}, \dots, \delta_{nj}]^T$ where $\delta_{jj} = 1$ and $\delta_{jk} = 0$ for all $1 \leq k \leq n, k \neq j$. Then equation (2.23) is

$$h^j = \Delta^j + F \cdot h^j$$

with $1 \leq j \leq n$.

Noting that $F = \Lambda \Gamma^{-1}$, invoking Theorem 2.8 now reveals the same result as Theorem 2.1.

□

2.5.4 Final State of the Diffusion

In addition to determining the infection outbreak condition, percolation theory can also be applied to calculate final state of a diffusion. The following theorem shows how we can employ percolation theory to calculate the final state. We also compare it with Theorem 2.4, which is derived from our dynamical model.

Theorem 2.9 *For the community-based network specified in Section 2.1 with infection rate λ'_{ij} significantly smaller than the recovery rate γ'_j (i.e. $\frac{1}{\lambda'_{ij} + \gamma'_j} \approx \frac{1}{\gamma'_j}$) for all $1 \leq i, j \leq n$, we have that $\tilde{S} = \exp(-\Lambda\Gamma^{-1}\tilde{R})$ holds true where $\lim_{t \rightarrow \infty} S(t) = \tilde{S}$ and $\lim_{t \rightarrow \infty} R(t) = \tilde{R}$.*

Proof: Denote $\tilde{S} = [\tilde{s}_1, \tilde{s}_2, \dots, \tilde{s}_n]^T$ and $\tilde{R} = [\tilde{r}_1, \tilde{r}_2, \dots, \tilde{r}_n]^T$. With the analysis in Section 2.5.2, the following equation holds true:

$$H_i(X|T) = x_i G_{ij} [H_{1i}(X|T), H_{2i}(X|T), \dots, H_{ni}(X|T)] \quad (2.24)$$

If an infection does not break out, $H_i(1_n|T) = 1$. However, as Newman argues in [24], if an infection breaks out, equation (2.24) does not hold because a giant component is formed, which is extensive and contains loops. Instead, (2.24) holds true if we redefine $H_i(X|T)$ and $H_{ij}(X|T)$ as corresponding generating functions for degree distributions of the parts of the network that are not in the giant component. These parts do not constitute the entire network but instead only a portion of it.

Therefore, if the infection breaks out, we would have $H_i(1_n|T) = \tilde{s}_i$. Equation (2.24) is converted as follows

$$\tilde{s}_i = G_i(\tilde{s}_1, \tilde{s}_2, \dots, \tilde{s}_n|T) \quad (2.25)$$

Following similar analysis as in the proof for Corollary 2.2, we have

$$G_i(X|T) = \prod_{k=1}^n \exp(-N_k p_{ik} T_{ki} (1 - x_k))$$

Combined this with (2.25) and (2.16), we have

$$\begin{aligned} \tilde{s}_i &= \exp\left(\sum_{k=1}^n N_k p_{ik} \frac{\lambda'_{ik}}{\lambda'_{ik} + \gamma'_k} (1 - \tilde{s}_k)\right) \\ &\approx \exp\left(\sum_{k=1}^n \frac{N_k p_{ik} \lambda'_{ik}}{\gamma'_k} \tilde{r}_k\right) \\ &= \exp\left(\sum_{k=1}^N \frac{\lambda_{ik}}{\gamma_k} \tilde{r}_k\right) \end{aligned} \tag{2.26}$$

for $i = 1 \dots n$. Actually (2.26) is equivalent to $\tilde{S} = \exp(\Lambda \Gamma^{-1} \tilde{R})$.

□

Both Corollary 2.2 and Theorem 2.9 show connections between our community-based dynamical model and the percolation model. Through the comparison, we may have the insights that our dynamical model is valid in the case that infection rates are significantly smaller than recovery rates. Actually this condition also agrees with our assumption that any two agents' states are independent of whether or not they are directly connected. In addition, the fact that our dynamic model gives conclusions in agreement with the percolation model also adds credence to our community-based dynamical model.

Chapter 3: Optimal Social Network Marketing

This chapter deals with optimal marketing on social networks. To keep in accordance with current research [14, 44, 46, 105, 106], we formulate the maximization of diffusion through altering the network structure or convincing an initial set of agents to adopt the innovation, both for the referral marketing and seeding marketing problems. We also investigate dynamic marketing, which is formulated as an optimal control problem both for the SIR model (2.5) and the SIS model (2.6).

Then, we extend to the continuous setting a property known as submodularity, giving a concept of “continuous monotone submodularity,” and use this concept to propose a greedy heuristic to optimize a continuous monotone submodular function. The approximation ratio of the heuristic is shown to be bounded by $1 - 1/e$. By showing that the social network marketing problems that we formulate all obey this property, we apply this greedy heuristic to solve each of the marketing problems of interest in this thesis. Based on the characteristics of the SIR model (2.5), we are able to modify the greedy heuristic for SIR models into a closed-loop form.

Also, we investigate other scenarios with different objectives. More specifically, we study the problem of minimizing the cost with the profits guaranteed to be greater than a given threshold. We also conduct trade-off analysis on profits vs.

cost. Finally, we show that the formulated viral marketing problem based on the SIS model (2.6) is convex, implying that any locally optimal solution is automatically globally optimal. At the end of this chapter a numerical experiment is given for validation.

3.1 Problem Formulation

Referral marketing, seeding marketing and dynamic marketing have been introduced in Section 1.2. Next, we mathematically formulate these problems specifically for the SIR and SIS models specified in Section 2.1.

3.1.1 Referral Marketing

Referral marketing, in which the current customers are offered an in-kind or cash reward for referring their acquaintances [55], is a popular marketing strategy to stimulate contagion of an innovation through word-of-mouth effects or electronic communications. With a referral bonus as stimulus, the agents are more likely to recommend a product to others, which can be modeled as an increase in the infection rate between individuals. For simplicity of analysis, we assume that the increase in the infection rate is proportional to the amount of the corresponding investment, while the recovery rate is fixed and unaffected by the investment.

For referral marketing, we propose two natural problems: the *minimum cost problem*, which aims to specify the minimum cost to facilitate infection outbreak, and the *profit maximization problem*, which is to determine the maximum possible

profit with a fixed budget. With the SIR model (2.5), these two problems are respectively formulated as (3.1) via Theorem 2.1 and (3.2) via Theorem 2.4.

$$\begin{aligned}
\min \quad & \text{tr}(W^T X) & \max \quad & Z^T \cdot R \\
\text{s.t.} \quad & \Lambda' - \Gamma \not\leq 0 & \text{s.t.} \quad & R + \exp(-\Lambda' \Gamma^{-1} R) = 1_n \\
& \Lambda' = \Lambda + X & & \text{tr}(W^T \cdot X) \leq B \\
& X \geq 0 & & X \geq 0, \Lambda' = \Lambda + X
\end{aligned} \tag{3.1} \tag{3.2}$$

where $B > 0$ is the total budget, the vector $Z \in \mathbb{R}^n$ refers to the weight of each community and the matrix $X \in \mathbb{R}^{n \times n}$ with x_{ij} representing the increase in λ_{ij} . For the SIS model (2.6), the minimum cost problem is of the same form as (3.1) while the maximum profit problem is trivial and hereby not investigated here. By Theorem 3.1, the profit maximization problem (3.2) can be formulated into a convex form.

Theorem 3.1 *The solution of (3.2) remains unchanged with the constraint $R + \exp(-\Lambda' \Gamma^{-1} R) = 1_n$ replaced with $R + \exp(-\Lambda' \Gamma^{-1} R) \leq 1_n$. This replacement makes (3.2) convex.*

Proof: See Appendix B.1. □

The following lemma connects the existence of a nonzero solution to the equation $R + \exp(-\Lambda \Gamma^{-1} R) = 1_n$ with the occurrence of infection outbreak.

Lemma 3.1 *If there is a nonzero solution to $R + \exp(-\Lambda \Gamma^{-1} R) = 1_{n \times 1}$ for every $R \in [0, 1]^n$, then we have $\Lambda - \Gamma \not\leq 0$.*

Proof: See Appendix B.2. □

According to Lemma 3.1, we can take advantage of (3.2) to determine whether a given budget is sufficient to produce infection outbreak. Then by conducting binary search with solving (3.2) to determine whether a budget B is enough to trigger the outbreak repeatedly, we can efficiently find an approximate solution to problem (3.1).

3.1.2 Seeding Marketing

The seeding marketing strategy [107] is another useful strategy in social network marketing, where the aim is to maximize inter-agent communication effects by convincing a set of influential agents to adopt and advertise a product to their acquaintances. We mathematically formulate the *seeding marketing problem*, which aims to maximize the product's diffusion under a given budget constraint. This problem is similar to the spreading maximization problem in [14], where the problem is studied based on the linear threshold model [15, 47] and the independent cascade model [3, 4].

As found in Section 2.4, in the SIS model (2.6) the final state (i.e., after infinite time) is independent of the initial state. Since here we consider seeding marketing effects asymptotically as time approaches infinity, we consider only seeding marketing for the SIR model (2.5). Using Theorem 2.4, the seeding marketing problem can

be formulated as follows:

$$\begin{aligned}
& \max && Z^T \cdot R \\
& \text{s.t.} && R + \text{diag}[1 - I_0] \exp(-\Lambda \Gamma^{-1} R) = 1_n \\
& && W^T I_0 \leq B, I_0 \geq 0
\end{aligned} \tag{3.3}$$

Here, $Z \in \mathbb{R}^n$ denote the weights of the communities in terms of value to the advertiser, $W \in \mathbb{R}^n$ denotes the costs to infect the agents in each community and $B > 0$ is the budget (a scalar). The problem can be re-cast more compactly as follows. Transform I_0 into $X_0 = \text{diag}[W] I_0$, which is actually the investment in each community. Unless specified otherwise, we will use D to represent the feasible set for the seeding marketing problem, i.e. $D = \{X_0 \in \mathbb{R}^n \mid 1^T \cdot X_0 \leq B, X_0 \geq 0\}$. By Lemma 3.2, with any $X_0 = \text{diag}[W] I_0 \in D$, the final state R that satisfies the equation $R + \text{diag}[1 - I_0] \exp(-\Lambda \Gamma^{-1} R) = 1_n$ is unique.

Lemma 3.2 *For every nonzero $X_0 \geq 0$ with its j -th component less than or equal to w_j , then $\exists! R \in [0, 1]^n$ so that the following equation holds.*

$$R + \text{diag}[1_n - \text{diag}[W]^{-1} X_0] \exp(-\Lambda \Gamma^{-1} R) = 1_n$$

Proof: See Appendix B.3. □

Define the mapping $\phi^{(1)} : D \rightarrow \mathbb{R}^n$, which maps the investment X_0 into the solution to $R + \text{diag}[1 - I_0] \exp(-\Lambda \Gamma^{-1} R) = 1_n$. Then the seeding marketing problem (3.2) is written as

$$\begin{aligned}
& \max && Z^T \phi^{(1)}(X_0) \\
& \text{s.t.} && X_0 \in D
\end{aligned} \tag{3.4}$$

The problem (3.4) is nonlinear, and we have not found an equivalent convex formulation. Solving this problem accurately is expected to be computationally burdensome. However, the techniques that both Kempe *et al.* [14] and Shakarian *et al.* [31] employ in their research inspires us to propose a heuristic algorithm which will be given later in this chapter.

3.1.3 Dynamic Marketing

Dynamic marketing is more flexible than referral or seeding marketing. Unfortunately, there is limited existing work in this area. Through advertising, a company exerts external dynamic influences over time to maximize profits as well as the achieved extent of diffusion. The *dynamic marketing problem* aims to maximize such profits with a given budget. Taking advantage of the SIR model (2.5) and the SIS model (2.6), dynamic marketing can be formulated in terms of a class of optimal control problems.

For convenience, the configurations here are set up as follows. Let the initial condition to be $S(0) = 1_n$, $I(0) = R(0) = 0$ for the SIR model (2.5) or $S(0) = 1_n$ and $I(0) = 0$ for the SIS model (2.6), which implies that the product is adopted in nowhere initially. With a given input term $U(\cdot)$, the value of $S(T)$ is determined in either the SIR model (2.5) and the SIS model (2.6), where T is the final time.

Using the cost matrix $W \in \mathbb{R}^n$, we transform the input $U(\cdot)$ into $V(\cdot) = \text{diag}[W]U(\cdot)$, which is interpreted as the investment in each community. Any investment term $V(\cdot)$ also determines the value of $S(T)$. Unless specified other-

wise, we use \mathcal{D} to represent the feasible field for dynamic marketing, i.e. $\mathcal{D} = \{V(\cdot) \mid V(t) \geq 0, \mathbf{1}_n^T \cdot V(t) \leq B(t), \text{ for } 0 \leq t \leq T\}$ where $B(t)$ refers to the budget at time t . Then we define the mapping $\phi_T^{(2)} : \mathcal{D} \rightarrow \mathbb{R}^n$, where $\phi_T^{(2)}(V(\cdot))$ is the value $S(T)$ of the SIR model (2.5) with the input $U(\cdot) = \text{diag}[W]^{-1} V(\cdot)$. Similarly, we define $\phi_T^{(3)} : \mathcal{D} \rightarrow \mathbb{R}^n$ for the SIS model (2.6).

Then the dynamic marketing problem is formulated as follows:

$$\begin{aligned} \max \quad & Z^T \left(\mathbf{1}_n - \phi_T^{(i)}(V(\cdot)) \right) \\ \text{s.t.} \quad & V(\cdot) \in \mathcal{D}, \quad i = 2, 3 \end{aligned} \tag{3.5}$$

Here $Z \in \mathbb{R}^n$ represents the weights of the communities in the overall payoff, $i = 2$ corresponds to marketing for a system following the SIR model (2.5); similarly, $i = 3$ is used to denote marketing for a system following the SIS model (2.6). We also study other variants of the dynamic marketing problem via putting other objectives into (3.5). These variants will be discussed later in this chapter.

3.2 Continuous Monotone Submodularity

In discrete mathematics, submodularity is an important property of set functions with deep theoretical consequences and many applications [56]. A set function $f : 2^{\mathcal{M}} \rightarrow \mathbb{R}$ is submodular if for every $X, Y \subseteq \mathcal{M}$ with $X \subseteq Y$ and every $x \in \mathcal{M} \setminus Y$ we have that $f(X \cup \{x\}) - f(X) \geq f(Y \cup \{x\}) - f(Y)$. Also, such a function f is monotone if for every $X, Y \subseteq \mathcal{M}$ with $X \subseteq Y$ we have $f(X) \leq f(Y)$.

A famous theorem guarantees that, when using the greedy heuristic to maximize a submodular function, the approximation ratio is bounded by $1 - 1/e$ [56].

This theorem is used by Kempe *et al.* in [14,46] to obtain an approximate solution to the spreading maximization problem using the linear threshold model [15,47] and the independent cascade model [3,4]. However, notice that both the domain D for seeding marketing (3.3) and \mathcal{D} for dynamic marketing (3.5) are more complicated than the discrete domain \mathcal{M} in the submodularity definition. Thus, the first step in proceeding using the submodularity concept and finding associated greedy heuristics is to generalize the notion of submodularity.

We set the domain as the lattice $(\mathcal{M}, \leq, +, \cdot)$ which are closed w.r.t. two operations, addition $+$ and scalar multiplication \cdot . Denote 0 as the minimum in \mathcal{M} . The set \mathcal{M} is required to possess the following properties.

1. $x + 0 = x$.
2. $\forall x, y \in \mathcal{M}, x \leq x + y$.

Two examples of such a domain \mathcal{M} are $[0, \infty)^n$ and $\{V(\cdot) \mid V(t) \in [0, \infty)^n, 0 \leq t \leq T\}$.

These coincide, respectively, with the feasible set D in seeding marketing (3.3) and the feasible set \mathcal{D} in dynamic marketing (3.5). Then the definition of continuous monotone submodularity of a function $f : \mathcal{M} \rightarrow \mathbb{R}^n$ is formulated in the following four steps, culminating in Definition 3.4.

Definition 3.1 *A function $f : \mathcal{M} \rightarrow \mathbb{R}$ with \mathcal{M} denoting the domain is monotone if for $\forall x, y \in \mathcal{M}$ and $x \leq y$, we have $f(x) \leq f(y)$.*

Definition 3.2 *A function $f : \mathcal{M} \rightarrow \mathbb{R}$ with \mathcal{M} denoting the domain is continuous submodular if for $\forall \Delta x, x, y \in \mathcal{M}$ and $x \leq y$ we have $\left. \frac{\partial}{\partial \alpha} f(x + \alpha \Delta x) \right|_{\alpha=0} \geq \left. \frac{\partial}{\partial \alpha} f(y + \alpha \Delta x) \right|_{\alpha=0}$.*

$$\alpha \Delta x) \Big|_{\alpha=0}.$$

Definition 3.3 *A function $f : \mathcal{M} \rightarrow \mathbb{R}$ with \mathcal{M} denoting the domain is continuous monotone submodular if it is monotone and continuous submodular.*

Definition 3.4 *A vector function $f : \mathcal{M} \rightarrow \mathbb{R}^n$ with \mathcal{M} denoting the domain is continuous monotone submodular if each of its components is continuous monotone submodular.*

The continuous monotone submodularity is easily interpreted from an economics viewpoint. If we view $f(x)$ as an profit function while the argument x is the investment, then that $f(\cdot)$ is continuous monotone submodular means that with the increasing investment, the utilities also increase but the marginal utility decreases.

The following theorem connects continuous monotone submodularity in Definition 3.4 with seeding marketing (3.4) and dynamic marketing (3.5).

Theorem 3.2 *The function $\phi^{(1)}(\cdot)$ in the seeding marketing problem (3.4), and the functions $-\phi_T^{(2)}(\cdot)$ and $-\phi_T^{(3)}(\cdot)$ in dynamic marketing are all continuous monotone submodular.*

Proof: See Appendix B.4. □

Notice that Theorem 3.2 implies that both the seeding marketing problem (3.4) and the dynamic marketing problem (3.5) enjoy the continuous monotone submodularity property, which is a generalization of classical submodularity. This inspires us to generalize the traditional greedy heuristic to solve these two problems in a unified framework. We address this issue in the following sections.

3.3 Greedy Heuristic

In discrete mathematics, the maximization of submodular function is widely studied in various areas [14, 57, 58]. The simplest example is the maximization problem with cardinality constraints as follows.

$$\max_{S \subseteq \mathcal{D}} f(S), \text{ s.t. } |S| \leq m \quad (3.6)$$

for some $m \geq 0$ where the function $f : 2^{\mathcal{M}} \rightarrow \mathbb{R}$ is submodular. As Kempe *et al.* points out in [14] that this problem is NP-hard and thus to design an efficient algorithm finding the global optimal solution with polynomial complexity is impossible.

An intuitive approach to solve (3.6) is the greedy heuristic [56], which starts with the empty set $S_0 = \emptyset$, and in iteration i , adds the element maximizing that $f(S_{i-1} \cup \{e\}) - f(S_{i-1})$, i.e. $S_i = S_{i-1} \cup \arg \max_e \{f(S_{i-1} \cup \{e\}) - f(S_{i-1})\}$. Nemhauser *et al.* [108] shows that such greedy heuristic yields a $1 - 1/e$ approximation ratio.

Our notion of continuous monotone submodularity given in Section 3.2 is an extension of the traditional notion of submodularity. This inspires us to generalize the greedy heuristic so as to maximize the continuous monotone submodular functions $\phi^{(1)}(\cdot)$, $-\phi_T^{(2)}(\cdot)$ and $-\phi_T^{(3)}(\cdot)$ in both the seeding marketing problem (3.4) and the dynamic marketing problem (3.5). Section 3.3.1 generalizes the greedy heuristic to maximize the continuous monotone submodular function and Section 3.3.2 applies it in the seeding marketing (3.4) and dynamic marketing (3.5) with modifications.

3.3.1 General Greedy Algorithm

For convenience, we define the notion of “magnitude” as a mapping $|\cdot| : \mathcal{M} \rightarrow [0, +\infty)$ where the magnitude $|x|$ of $x \in \mathcal{M}$ intuitively represents how large x is. To be a magnitude, such a mapping $|\cdot|$ is required to satisfy the following properties:

1. $|0| = 0$.
2. $\forall x, y \in \mathcal{M}$, if $x \leq y$ then $|x| \leq |y|$.
3. $\forall x, y \in \mathcal{M}$ and $\forall \alpha, \beta \geq 0$, $|\alpha x + \beta y| = \alpha|x| + \beta|y|$.

For the feasible set D in seeding marketing (3.4) and \mathcal{D} in dynamic marketing (3.5), an intuitive example of the magnitude $|\cdot|$ is the 1-norm. With this magnitude function, consider the maximization of continuous monotone submodular function as follows:

$$\begin{aligned} \max_{x \in \mathcal{M}} \quad & f(x) \\ \text{s.t.} \quad & x \in \Omega \\ & |x| \leq k \end{aligned} \tag{3.7}$$

where $f : \mathcal{M} \rightarrow \mathbb{R}$ such that $f(0) = 0$, $\Omega \subseteq \mathcal{M}$ is a convex subset of \mathcal{M} , representing the general constraints on x , and $k \in \mathbb{R}^+$ is a given positive constant.

Inspired by the greedy heuristic in discrete mathematics, we design the following algorithm (Algorithm 1) to maximize a continuous monotone submodular function, which we call a “continuous greedy algorithm,” with the assumption that

$$\arg \max_{|\Delta x|=1} \left. \frac{\partial}{\partial \alpha} f(x + \alpha \Delta x) \right|_{\alpha=0} \text{ is always nonempty } \forall x \in \mathcal{M}.$$

Algorithm 1 General Continuous Greedy Algorithm

Input: function $f : \mathcal{M} \rightarrow \mathbb{R}$, general constraints Ω and upper bound l .

set $x = 0, s = 0$

while $|x| < l$ **do**

$$x \leftarrow x + \Delta x \cdot ds, \Delta x \in \arg \max_{|\Delta x|=1} \left\{ \left. \frac{\partial}{\partial \alpha} f(x + \alpha \Delta x) \right|_{\alpha=0} \text{ s.t. } \Delta x \in \text{TC}(x, \Omega) \right\}$$

$s \leftarrow s + ds$

end while

Output: x

Let $\text{TC}(x, \Omega)$ denote the tangent cone to Ω at x . It should be noted that since $|\cdot|$ satisfies Property 3 above, we have $\frac{d}{ds} |x(s)| = \left| \frac{d}{ds} x(s) \right| = 1$, and therefore $|x(l)| = l$. Let $x^*(k)$ denote the solution to the problem (3.7) while $x(l)$ is the solution generated by Algorithm 1 to (3.7) where the constraint $|x| < k$ is replaced by $|x| < l$. The next theorem connects $x^*(k)$ with $x(l)$.

Theorem 3.3 *For all positive k and l we have*

$$f(x(l)) \geq (1 - \exp(-l/k)) f(x^*(k))$$

In particular, when $k = l$, $f(x(l)) \geq \frac{e-1}{e} f(x^(k))$.*

Proof: See Appendix B.5. □

3.3.2 Application to Social Network Marketing

Recall the seeding marketing (3.4) and dynamic marketing (3.5) problems:

$$\begin{aligned}
\max \quad & Z^T \phi^{(1)}(X_0) & \max \quad & Z^T \left(1_n - \phi_T^{(i)}(V(\cdot))\right) \\
\text{s.t.} \quad & X_0 \in D & \text{s.t.} \quad & V(\cdot) \in \mathcal{D}, \quad i = 2, 3
\end{aligned} \tag{3.4} \tag{3.5}$$

According to Theorem 3.2, the functions $Z^T \phi^{(1)}(X_0)$ and $-Z^T \phi_T^{(i)}(V(\cdot))$ are both continuous monotone submodular, therefore both problems fit into the framework of continuous monotone submodularity.

Algorithm 1 requires the existence of $\arg \max_{|\Delta x|=1} \frac{\partial}{\partial \alpha} f(x + \alpha \Delta x) \Big|_{\alpha=0}$, which is an assumption and holds for seeding marketing (3.4) because D is compact but is not necessarily valid for dynamic marketing (3.5) since $V(\cdot)$ lives in a Banach space. This may cause trouble in theory but is not a big issue for practical purposes. With the discretization of $V(\cdot)$ by dividing $[0, T]$ into a number of smaller intervals, the feasible domain \mathcal{D} becomes compact and the existence of $\arg \max_{|\Delta x|=1} \frac{\partial}{\partial \alpha} f(x + \alpha \Delta x) \Big|_{\alpha=0}$ is guaranteed.

The details of applying Algorithm 1 to seeding marketing (3.4) and dynamic marketing (3.5) are straightforward and are not detailed here. The main step in solving dynamic marketing (3.5) is to discretize $V(\cdot)$ by dividing the $[0, T]$ into small intervals. Also, dynamic marketing based on the SIR model enjoys an interesting property which is helpful to further improve the corresponding greedy algorithm; the following lemma highlights this property.

Lemma 3.3 *With any $V(\cdot) \geq 0$, $\Delta V(\cdot) \geq 0$ and the time delay $\tau > 0$, then $-\frac{\partial}{\partial \alpha} \phi_T^{(2)}(V(t) + \alpha \Delta V(t)) \Big|_{\alpha=0} \geq -\frac{\partial}{\partial \alpha} \phi_T^{(2)}(V(t) + \alpha \Delta V(t - \tau)) \Big|_{\alpha=0}$. Here $\Delta V(t) = 0$ for $t \notin [0, T - \tau]$.*

Proof: See Appendix B.6. □

Lemma 3.3 highlights an interesting property of the SIR model (2.5): if part of the external influence over the diffusion process is delayed, then the corresponding inter-agent communication effects are decreased.

When solving the dynamic marketing based on the SIR model (2.5) for practical purposes, we discretize $V(\cdot)$ via partitioning the time interval $[0, T]$ into several steps. The greedy heuristic serves to iteratively find the pair of the community and time-step which has the greatest contribution towards the profits increase and add the corresponding investment by a small unit.

According to Lemma 3.3, an increase of investment in the first time-step always leads greater profit lift than in other time-steps afterwards. Therefore, we can modify the greedy heuristic to iteratively find the community which has the greatest contribution and increase it at the first time-step until the investment in the first time-step reaches its bound. Then we move on to the second time-step and repeat the same procedure. With such modification, if there's still budget to increase the investment in one time-step, there's no need to check the time-steps later than it. In such a way the algorithm efficiency is greatly improved.

More importantly, we can introduce feedback into the greedy heuristics. In essence, at each iteration the greedy heuristics is simulating the SIR model (2.5) to predict in which component of the discretized input $V(\cdot)$ should we increase the investment. Due to the model uncertainties and the external disturbances, a feedforward simulation has the risk to be divergent from the reality. By Theorem 3.3, we can calculate the investment in the first time-step, exert the input, wait until

the end of the first time-step, estimate the states, use such estimates to update our simulation to improve the prediction accuracy and calculate the investment for the second time-step, so on so forth.

It should be noted that as long as the continuous monotone submodularity holds, the Algorithm 1 achieves the approximation ratio of $1 - 1/e$. Intuitively, the continuous monotone submodularity widely exists in various aspects concerning the product diffusion and marketing, therefore the Algorithm 1 may be potentially applied in other problems besides the seeding marketing (3.4) and dynamic marketing (3.5) problems formulated in this chapter.

Taking advantage of the recent progress in monotone control systems [59], we show that the dynamic marketing problems based on the SIS model are intrinsically convex. Such convexity means that a locally optimal marketing policy that is found out via using any numerical optimal control toolboxes is automatically globally optimal. We will address this issue later in this chapter.

3.4 Two Variants of Dynamic Marketing

Up to now, dynamic marketing primarily focuses on maximizing the profits with a fixed budget. This section deals with another two scenarios: to minimize the cost with a given profit guaranteed and the trade-off between the profit versus cost.

3.4.1 Minimum Cost Marketing

Unlike the dynamic marketing problem (3.5), the minimum cost marketing problem aims to find out the minimum possible investment to achieve a certain effect of the advertisement diffusion. We formulate the minimum cost marketing problem as (3.8). Similar to how we deal with the dynamic marketing problem (3.5) in Section 3.3, we generalize (3.8) into the general minimum cost problem (3.9) and solve it by modifying the termination condition of the Algorithm 1.

$$\begin{aligned}
 \min \quad & \int_0^T 1_n^T \cdot V(t) \, dt & \min_{x \in \mathcal{M}} \quad & |x| \\
 \text{s.t.} \quad & \phi^{(i)}(V(\cdot)) \leq \theta, \quad i = 2, 3 & \text{s.t.} \quad & f(x) \geq \theta \\
 & & & x \in \Omega
 \end{aligned} \tag{3.8} \tag{3.9}$$

The function $f : \mathcal{M} \rightarrow \mathbb{R}$ is continuous monotone submodular with $f(0) = 0$ and $\theta > 0$ is a given threshold. The main idea for our approach to solving (3.9) is similar to Algorithm 1, with the only difference being the termination condition, i.e., we keep increasing x until the value of $f(x)$ reaches the threshold θ . This procedure is formulated as in algorithmic form as follows:

The following theorem specifies the approximation ratio achieved by Algorithm 2.

Theorem 3.4 For $0 \leq s \leq \eta$, we define $\xi(s) = \min_{|\Delta x|=1} \left\{ \frac{1}{\left. \frac{\partial}{\partial \alpha} f(x(s) + \alpha \Delta x) \right|_{\alpha=0}} \right\}$. Also, with a fixed $x \in \mathcal{M}$ we define $\nu(\cdot, \cdot) : [0, \eta] \times \mathcal{M} \rightarrow \mathbb{R}$ as

$$\nu(s, x) = \frac{\partial}{\partial \alpha} f \left(x(s) + \alpha \frac{x \vee x(s) - x(s)}{|x \vee x(s) - x(s)|} \right) \Bigg|_{\alpha=0} \cdot |x \vee x(s) - x(s)|$$

Let $k_1 = \frac{\xi(\eta)}{\xi(0)}$ and $k_2 = \max_{x \leq x^*} \frac{\nu(0, x)}{\nu(\eta, x)}$ where x^* is the solution to (3.9), then the

Algorithm 2 General Greedy Algorithm for Minimum Cost Problem

Input: function $f : \mathcal{M} \rightarrow \mathbb{R}$, general constraints Ω and threshold θ .

set $x = 0, s = 0$

while $f(x) < \theta$ **do**

$$x \leftarrow x + \Delta x \cdot ds, \Delta x \in \arg \max_{|\Delta x|=1} \left\{ \left. \frac{\partial}{\partial \alpha} f(x + \alpha \Delta x) \right|_{\alpha=0} \text{ s.t. } \Delta x \in \text{TC}(x, \Omega) \right\}$$

$s \leftarrow s + ds$

end while

Output: $|x|$

value of the corresponding solution obtained by the Algorithm 2, which is denoted as $\eta \in \mathbb{R}$, and the optimal value $\zeta = |x^*|$ of (3.9) satisfies that

$$\eta \leq [1 + \log \min(k_1, k_2)] \zeta$$

Proof: See Appendix B.7. □

Theorem 3.4 yields a lower bound of the approximation ratio achieved using Algorithm 2. Intuitively, if the first order derivative $\dot{f}(x)$ does not decrease too dramatically as x increases, the approximation ratio $\frac{\zeta}{\eta} \geq [1 + \log \min\{k_1, k_2\}]^{-1}$ is relatively close to 1. Application of Algorithm 2 to minimum cost marketing (3.8) is straightforward, and the details are not included here.

3.4.2 Trade-Off Between Profit and Cost

In addition to maximizing the profit and minimizing the cost, another problem of interest is the trade-off between profit and cost. For the SIR model (2.5) and the

SIS model (2.6), this trade-off problem can be formulated as follows:

$$\begin{aligned} \max \quad & Z^T \left(1_n - \phi_T^{(i)}(V(\cdot)) \right) - \int_0^T 1_n^T \cdot V(t) \, dt & \max_{x \in \mathcal{M}} \quad & f(x) \\ \text{s.t.} \quad & V(\cdot) \in \mathcal{D}, i = 2, 3 & \text{s.t.} \quad & x \in \Omega \end{aligned} \quad (3.10) \quad (3.11)$$

The objective function in the trade-off problem (3.10) is still continuous submodular but no longer monotone. Without loss of generality, we formulate the maximization of the general continuous non-monotone submodular function as (3.11). The function $f : \mathcal{M} \rightarrow \mathbb{R}$ is continuous submodular and $f(0) = 0$. Also, $\Omega \subseteq \mathcal{M}$ is a convex subset of \mathcal{M} , representing the general constraints on x . In addition, we require the Ω satisfy the following properties:

1. For $\forall x, y \in \Omega$, we have $x \vee y \in \Omega$ and $x \wedge y \in \Omega$.
2. For $\forall x, y \in \Omega$, if $x \leq y$, then $\exists z \in \Omega$ such that $x + z = y$; denote z as $z = y - x$.

An example of the feasible set \mathcal{D} in the trade-off problem (3.10) satisfying these properties is

$$\mathcal{D} = \{V(\cdot) \mid 0 \leq V(t) \leq V_{\max}(\cdot), \text{ for } 0 \leq t \leq T\}$$

where $V_{\max}(\cdot)$ is the upper bound on the investment $V(\cdot)$. Assuming the existence of such an upper bound $V_{\max}(\cdot)$ is helpful for theoretically addressing the problem; moreover we believe that it is reasonable from a practical point of view.

In [109], Feige *et al.* study the maximization of traditional non-monotone submodular set function. Inspired by their work, we propose the bi-directional local search algorithm to maximize the continuous non-monotone submodular function $f(x)$ in (3.11). At each iteration, unlike the greedy algorithm, the local search algo-

rithm seeks a descent that increases $f(x)$ by a certain ratio instead of maximizing the increase. Also, the greedy algorithm iteratively increases x but in the bi-directional local search algorithm, it is also allowed to decrease x when necessary. This is also why we call this algorithm “bi-directional.” Denote the maximum in Ω as ω and ϵ as the parameter which controls the accuracy of the algorithm, then the local search algorithm is as follows.

Algorithm 3 Bi-Directional General Local Search Algorithm

Input: function $f : \mathcal{M} \rightarrow \mathbb{R}$, general constraints Ω and the parameter $\epsilon > 0$.

```

set  $x = x_0$ ,  $s = 0$  such that  $f(x_0) > 0$ 
while  $\exists \Delta x \in \text{TC}(x, \Omega)$ ,  $|\Delta x| = 1$  s.t.  $\left. \frac{\partial}{\partial \alpha} f(x + \alpha \Delta x) \right|_{\alpha=0} \geq \frac{\epsilon}{|\omega|} f(x)$  or  $\left. \frac{\partial}{\partial \alpha} f(x - \alpha \Delta x) \right|_{\alpha=0} \geq \frac{\epsilon}{|\omega|} f(x)$  do
    if  $\exists \Delta x \in \text{TC}(x, \Omega)$ ,  $|\Delta x| = 1$  s.t.  $\left. \frac{\partial}{\partial \alpha} f(x + \alpha \Delta x) \right|_{\alpha=0} \geq \frac{\epsilon}{|\omega|} f(x)$  then
         $x \leftarrow x + \Delta x \cdot d s$ 
    else
         $x \leftarrow x - \Delta x \cdot d s$ 
    end if
     $s \leftarrow s + d s$ 
end while

```

Output: x or $\omega - x$ if $f(\omega - x) > f(x)$

The following theorem deals with the approximation ratio of the Algorithm 3.

Theorem 3.5 *Let \tilde{x} denote the solution obtained by using the Algorithm 3 to solve the maximization of non-monotone continuous submodular function (3.11) and x^**

is the optimal solution to (3.11). With ω denoting the maximum in Ω , if $f(\omega) \geq 0$, then either $f(\tilde{x}) \geq \frac{1}{3}(1 - \epsilon)f(x^*)$ or $f(\omega - \tilde{x}) \geq \frac{1}{3}(1 - \epsilon)f(x^*)$ holds true.

Proof: See Appendix B.8. □

Theorem 3.5 shows that the approximation ratio of the Algorithm 3 is $\frac{1}{3}$. It should also be noted that, the value of $f(x)$ increases exponentially until termination in the Algorithm 3, therefore this algorithm is highly efficient in the running time.

3.5 The Convexity of $\phi_T^{(3)}(\cdot)$ (for SIS model (2.6))

In addition to the continuous monotone submodularity of $\phi_T^{(3)}(\cdot)$ as specified by Theorem 3.2, with the recent progress in monotone control system [80], we can show that the $\phi_T^{(3)}(\cdot)$, which connects the investment $V(\cdot)$ and the final percentage of the susceptible agents $S(T)$ in the SIS model (2.6), is convex with the following theorem.

Theorem 3.6 *Each component of the function $\phi_T^{(3)}(\cdot) : \mathcal{D} \rightarrow \mathbb{R}^n$ as specified in Section 3.1.3 is convex.*

Proof: See Appendix B.9. □

By Theorem 3.6, the dynamic marketing (3.5), the minimum investment (3.8) and the trade-off (3.10) problems based on the SIS model (2.6) are all essentially convex. This important observation implies that any locally optimal solution to these problems generated by an numerical optimization toolbox is automatically globally optimal. Therefore, it is realistic to solve all the problems formulated

based on the SIS model (2.6) in this chapter efficiently.

3.6 Numerical Experiments

This section specifies two numerical experiments on seeding marketing (3.4) and dynamic marketing (3.5) respectively. We implement the Algorithm 1 to solve each of the problem and analyze the accuracy numerically.

3.6.1 Seeding Marketing

We consider a network consisted of four communities as for the seeding marketing (3.4). The weight vector $Z = W = [1 \ 1 \ 1 \ 1]^T$ and the recovery rate $\Gamma = \text{diag}[0.1 \ 0.1 \ 0.1 \ 0.1]$. The budget $B = 0.4$ while the spreading strength matrix is

$$\Lambda = \lambda_0 \begin{bmatrix} 0.08 & 0.04 & 0.02 & 0.02 \\ 0.01 & 0.08 & 0.05 & 0.04 \\ 0.03 & 0.03 & 0.08 & 0.01 \\ 0.02 & 0.01 & 0.01 & 0.08 \end{bmatrix}$$

where we will investigate the cases with $\lambda_0 = 0.1$, $\lambda_0 = 0.8$ and $\lambda_0 = 3$ respectively, which represents the scenarios where the outbreak does not occur, the outbreak occurs with relatively weak spreading strength and the outbreak occurs with relatively strong spreading strength.

In order to evaluate the performance of our greedy algorithm, we also need to find the globally optimal solution numerically. To achieve this, we randomly select a starting point and employ the MATLAB Optimization Toolbox to find an optimal

solution. Then we repeat this process 1000 times for different starting points and select the best solution among them. We treat the solution obtained in this way as our globally optimal solution to the seeding marketing problem (3.4). Then, in each scenario, the greedy solution and the optimal solution are given in Table 3.1.

Table 3.1: Comparison between the Greedy Solution and Globally Optimal Solution

λ_0	I_0	I_0^*	$Z^T \phi^{(1)}(\text{diag}[W] I_0)$	$Z^T \phi^{(1)}(\text{diag}[W] I_0^*)$
0.1	[0, 0.105, 0.169, 0.126]	[0, 0.109, 0.166, 0.125]	0.4640	0.4641
0.8	[0, 0, 0.111, 0.289]	[0, 0, 0.112, 0.288]	2.1386	2.1387
3	[0, 0, 0, 0.4]	[0, 0, 0, 0.4]	3.9578	3.9578

where I_0 and I_0^* are the greedy solution and the optimal solution respectively. Although Theorem 3.3 only gives a lower bound for the approximation ratio as 63.2%, from Table 3.1 we can see that in the seeding marketing problem (3.4) with the parameters listed above, the accuracy of the greedy heuristic is very high. In fact our continuous greedy algorithm yields a solution almost identical to the globally optimal solution.

3.6.2 Dynamic Marketing

We apply the Algorithm 1 to the dynamic marketing (3.5) for the SIS model (2.6) on a network consisted of four communities. Let the final time $T = 20$, the

weight vector $Z = W = [1 \ 1 \ 1 \ 1]^T$ and the recovery rate $\Gamma = \text{diag} [0.1 \ 0.1 \ 0.1 \ 0.1]$.

The spreading strength matrix is

$$\Lambda = \lambda_0 \begin{bmatrix} 0.08 & 0.04 & 0.02 & 0.02 \\ 0.01 & 0.08 & 0.05 & 0.04 \\ 0.03 & 0.03 & 0.08 & 0.01 \\ 0.02 & 0.01 & 0.01 & 0.08 \end{bmatrix}$$

with $\lambda_0 = 0.3$, $B(t) \equiv 0.04$; $\lambda_0 = 0.3$, $B(t) \equiv 0.4$ and $\lambda_0 = 1$, $B(t) \equiv 0.04$, which respectively correspond to the scenarios that the outbreak does not occur with low budget, the outbreak does not occur with high budget and the outbreak occurs. We discretize the time horizon into 20 intervals of equal length for practical implementation of Algorithm 1.

Since $\phi_T^{(3)}(\cdot)$ is convex, we can minimize it by using any numerical toolbox to find a locally optimal solution. Such a locally optimal solution is automatically globally optimal due to convexity. Figures 3.1, 3.2 and 3.3 depict the optimal control solution and the solution generated by the greedy algorithm.

The performance is summarized in Table 3.2, where the greedy and optimal results refer to the final profits achieved by the optimal solution and greedy solution respectively. The approximation ratio equals to the greedy result over the optimal result.

The experiments in this section show excellent practical performance of our greedy algorithm. Although Theorem 3.3 asserts that the approximation ratio can be as low as 63.2%, its practical accuracy is actually much better in the experiments that we have undertaken.

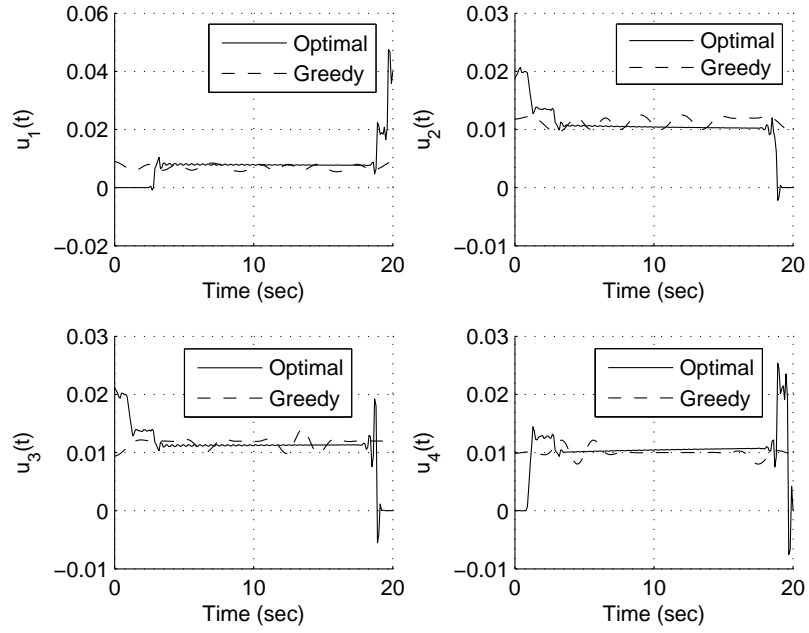


Figure 3.1: Case 1: $\lambda_0 = 0.3$ and $B(t) \equiv 0.04$

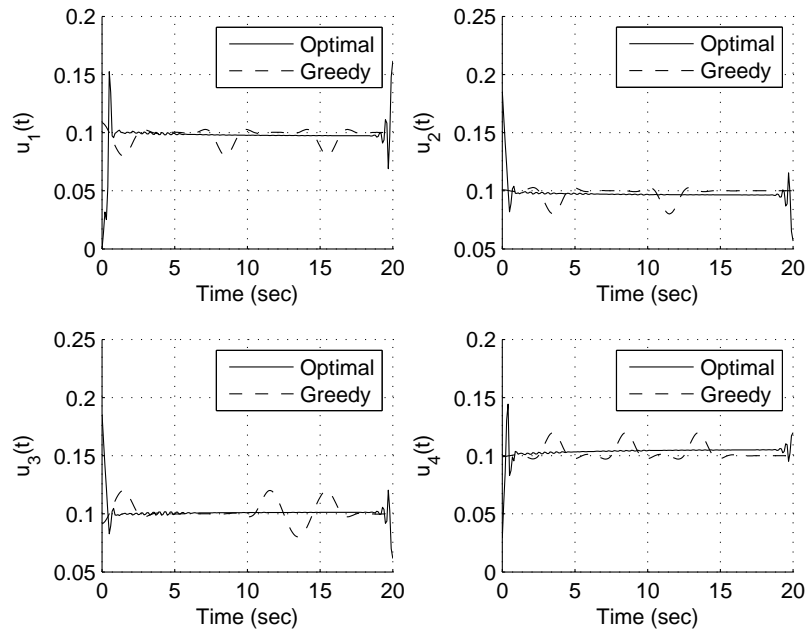


Figure 3.2: Case 2: $\lambda_0 = 0.3$ and $B(t) \equiv 0.4$

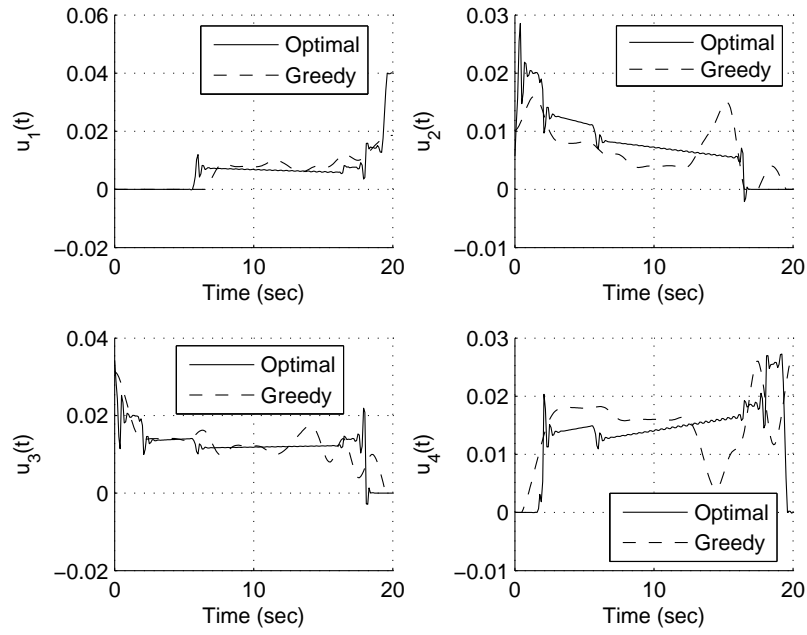


Figure 3.3: Case 3: $\lambda_0 = 1$ and $B(t) \equiv 0.04$

Table 3.2: Comparing Greedy Solution and Optimal Solution in a Dynamic Marketing Problem

	Greedy Result	Optimal Result	Approximation Ratio (%)
Case 1	0.4363	0.4369	99.86
Case 2	2.1838	2.1840	99.99
Case 3	0.9741	0.9753	99.88

Chapter 4: Social Network Marketing Games

In marketing, usually multiple companies are simultaneously promoting their products by maximizing communications between individuals, which is a diffusion of multiple competitive contagions. Since our SIR model (2.5) or SIS model (2.6) consider diffusions of one contagion, it is necessary to derive new models for diffusions of competitive products.

In DCG diffusions, after purchasing a product, an agent spontaneously recommends the product to others for a duration of time. Afterwards, he gets tired and neither make recommendations nor purchase another product in short time, which means he quits the diffusion. We generalize the SIR model (2.5) to model the DCG diffusion, where by being infected with product i or recovered from i we mean agents spontaneously recommend products or stop making the recommendations.

The main difference for FMCG diffusions is that after an agent gets tired, he may either switches to another product or choose not to use any of them. However, he is still likely to purchase this product later again. To model the FMCG diffusion, we generalize current research works on voter models [75–77].

In this chapter we formulate social network marketing games based on the DCG model and the FMCG model respectively, which is referred to as DCG mar-

keting game and FMCG marketing game. We show existence of an ϵ -equilibrium with arbitrarily small ϵ in DCG marketing games and a pure Nash equilibrium in FMCG marketing games. The Price of Anarchy (PoA) is shown to be bounded by 2. In FMCG marketing game, we show that any two Nash equilibria equal to each other almost everywhere and design a “best-response” mechanism to seek the Nash equilibrium distributively.

4.1 DCG Marketing Game

In reality, because of reasons such as geographical locations or education backgrounds [41], a social network can be divided into a set of communities and agents’ interactions within each community are well-mixed. Therefore each community can be abstracted into a vertex in the social network while a backbone graph captures the connections among the communities.

In this chapter we consider diffusions over a population consisting of n well-mixed communities connected by a backbone network, where each individual (or agent) belongs to one and only one community. The i -th community includes N_i agents. A primary feature of communities is that all agents in the same community have equal probability to be connected with each other (well mixedness). Also, any two agents from a given pair of communities have equal probability to be connected (uniform likelihood of connection between agents of two connected communities). In the remainder of this chapter, we denote by p_{jj} the probability that agents in the j -th community have an acquaintance with any other selected agent in the same

community. Similarly, they have probability of p_{ij} to know those from the i -th community.

Assume that there are m players (companies) promoting their products simultaneously. As mentioned in Section 1.3, each agent may either be susceptible (has not purchased any product), infected with the i -th product (active in recommending the i -th product after purchasing it), or recovered from the i -th product (tired of the i -th product and no longer recommending it). Because people usually purchase a DCG no more than once in a short time, there are no agents simultaneously infected with or recovered from two or more products, and thus in total there are $2m + 1$ possible states for each agent.

An agent infected with a product “spreads” this product to his susceptible neighbors at some infection rate, while he becomes recovered from the product with some recovery rate. We assume that all the DCGs in competition are substitute goods with similar price and function, and therefore on average each community does not have any personal preference for each product. In other words, both infection rates and recovery rates are intrinsic properties of each specific community and irrelevant to different products. In this chapter, we denote by λ'_{ij} the infection rate from agents in community j to agents in community i , and by γ'_i the recovery rate for agents in community i .

Also, each player can increase chances that agents in certain communities adopt their products via advertising in the corresponding communities. To resolve the conflict that a susceptible agent may be exposed to multiple products simultaneously, we assume that he will adopt the product that first “infects” (convinces)

him.

Let $S(t) = \begin{bmatrix} s^1(t) & s^2(t) & \dots & s^n(t) \end{bmatrix}^T$, $I_k(t) = \begin{bmatrix} i_k^1(t) & i_k^2(t) & \dots & i_k^n(t) \end{bmatrix}^T$ and $R_k(t) = \begin{bmatrix} r_k^1(t) & r_k^2(t) & \dots & r_k^n(t) \end{bmatrix}^T$ for $k = 1, \dots, m$, where $s^j(t)$, $i_k^j(t)$ and $r_k^j(t)$ represent the percentage of agents in the j -th community who are susceptible, infected with the k -th product, or recovered from the k -th product at time t , respectively. We also use $U_k(t) = \begin{bmatrix} u_k^1(t) & u_k^2(t) & \dots & u_k^n(t) \end{bmatrix}^T$ where $u_k^j(t)$ is the advertising strength of player k in community j at time t . Similar to the derivation of the community-based SIR model in Section 2.1, the DCG competition model is formulated as follows:

$$\frac{dS(t)}{dt} = -\text{diag}[S(t)] \sum_{j=1}^m (\Lambda I_j(t) + C_j U_j(t)) \quad (4.1a)$$

$$\frac{dI_k(t)}{dt} = \text{diag}[S(t)] (\Lambda I_k(t) + C_k U_k(t)) - \Gamma I_k(t) \quad (4.1b)$$

$$\frac{dR_k(t)}{dt} = \Gamma I_k(t) \quad (4.1c)$$

where $\Lambda = [\lambda_{ij}]_{n \times n}$ is the infection rate matrix and $\Gamma = \text{diag}[\gamma_i]_{n \times n}$ is the recovery rate matrix, with $\lambda_{ij} = N_j \lambda'_{ij} p_{ij}$ and $\gamma_i = \gamma'_i$. Also, the diagonal matrix $C_k = \text{diag}[c_k^j]_{n \times n}$ is the cost matrix, where c_k^j denotes the strength of advertising that player k can exert on community j per unit of cost.

Initially we assume that all agents in the population are susceptible (have not made any purchase yet), i.e. $S(0) = 1_n$ while $I_k(0) = R_k(0) = 0$ and the competitive DCGs spread under influence of the companies' advertising. With any strategy profile $\mathbb{U}(\cdot) = [U_1(\cdot), U_2(\cdot), \dots, U_m(\cdot)]$, the solution of (4.1) is uniquely determined, and we define that $\phi_j^{(1)}(t, \mathbb{U}) = I_j(t) + R_j(t)$ with input $\mathbb{U}(\cdot)$. Also, we use weight vector $W \in \mathbb{R}^n$ to represent the relative importance of each of the communities. We

denote the corresponding payoff function for player k as $\zeta_k^{(1)}(\mathbb{U}) = W^T \cdot \phi_k^{(1)}(T, \mathbb{U})$, which measures revenue of the k -th player at time T . Here T denotes the duration of the DCG marketing campaign.

Also, we use $B_k(t) \in \mathbb{R}$ to denote the budget of player k at time t , i.e. $1_n^T \cdot U_k(t) \leq B(t)$, where each $B_k(t)$ is continuous for $t \in [0, T]$. Then the feasible strategy set for player k is denoted as

$$\mathcal{D}_k = \{U_k(\cdot) \mid U_k(t) \geq 0, 1_n^T \cdot U_k(t) \leq B_k(t)\}$$

. Thus, $\mathbb{U}(\cdot)$ takes values in the set

$$\mathcal{D} := \mathcal{D}_1 \times \mathcal{D}_2 \times \cdots \times \mathcal{D}_m$$

In the DCG marketing game, each player k picks a strategy in \mathcal{D}_k to maximize his utility function $\zeta_k^{(1)}(\mathbb{U}(\cdot))$.

To study Price of Anarchy (PoA), we define social welfare $\eta^{(1)}(\mathbb{U}(\cdot))$, which is the total profit of all the players, as $\eta^{(1)}(\mathbb{U}(\cdot)) = \sum_{k=1}^m \zeta_k^{(1)}(\mathbb{U}(\cdot))$. For $\mathbb{U}(\cdot) = [U_1(\cdot), U_2(\cdot), \dots, U_m(\cdot)]$, we denote combinations of the k -th player's competitors' strategies as

$$\mathbb{U}_{-k}(\cdot) = [U_1(\cdot), \dots, U_{k-1}(\cdot), U_{k+1}(\cdot), \dots, U_m(\cdot)]$$

Also, we denote the strategy profile in which player k changes his strategy from $U_k(\cdot)$ into $U'_k(\cdot)$ as $\mathbb{U}(\cdot) \oplus U'_k(\cdot) = [U_1(\cdot), \dots, U'_k(\cdot), \dots, U_m(\cdot)]$. All these notations are commonly used in the existing literature on game theory, such as [110]. Unless specified otherwise, the model (4.1) and the game defined as above are referred to as the DCG competition model and the DCG marketing game in the remainder of this chapter.

We have now completed the formulation of DCG marketing games. In the rest of this section we will determine two properties of DCG marketing games that will be used later in the chapter.

4.1.1 Continuous Monotone Submodularity

In Section 3.2, we introduced the definition of continuous monotone submodularity. The first property that we observe for DCG marketing games is that the social welfare function $\eta^{(1)}(\mathbb{U}(\cdot))$ is continuous monotone submodular w.r.t. $\mathbb{U}(\cdot)$. Formally, we have the following theorem.

Theorem 4.1 *Social welfare, $\eta^{(1)}(\mathbb{U}(\cdot))$, is continuous monotone submodular w.r.t. $\mathbb{U}(\cdot)$. To wit, for any $\mathbb{U}(\cdot) \leq \mathbb{U}'(\cdot)$, $\Delta\mathbb{U}(\cdot) \in \mathcal{D}$, we have $\left. \frac{\partial}{\partial \alpha} \eta^{(1)}(\mathbb{U}(\cdot) + \alpha \Delta\mathbb{U}(\cdot)) \right|_{\alpha=0} \geq 0$ and $\left. \frac{\partial}{\partial \alpha} \eta^{(1)}(\mathbb{U}(\cdot) + \alpha \Delta\mathbb{U}(\cdot)) \right|_{\alpha=0} \geq \left. \frac{\partial}{\partial \alpha} \eta^{(1)}(\mathbb{U}'(\cdot) + \alpha \Delta\mathbb{U}(\cdot)) \right|_{\alpha=0}$.*

Proof: See Appendix B.10. □

4.1.2 Competitiveness

Next, we study a property called competitiveness for DCG marketing games. Intuitively, a game being competitive means that any player will not benefit if other players increase their investment. Formally, competitiveness is defined as follows:

Definition 4.1 *The marketing game is competitive if any player k increases his investment from $U_k(\cdot)$ to $U'_k(\cdot)$ s.t. $U_k(\cdot) \leq U'_k(\cdot)$, then all the other players' payoffs do not increase, i.e., $\zeta_i(t, \mathbb{U}(\cdot)) \leq \zeta_i(t, \mathbb{U}(\cdot) \oplus U'_k(\cdot))$ with $i \neq k$.*

Intuitively in DCG marketing games, no player can benefit from others' increasing their investment. The following theorem asserts this mathematically.

Theorem 4.2 *The DCG marketing game is competitive, i.e., for any $U_k(\cdot), U'_k(\cdot) \in \mathcal{D}_k$ such that $U_k(\cdot) \leq U'_k(\cdot)$, we have $\zeta_i^{(1)}(\mathbb{U}(\cdot)) \leq \zeta_i^{(1)}(\mathbb{U}(\cdot) \oplus U'_k(\cdot))$ for $k = 1, \dots, m$ and $i \neq k$.*

Proof: See Appendix B.11. □

We have shown two properties of the DCG marketing game: continuous monotone submodularity and competitiveness. Later in this chapter we will see that these are useful in the analysis of PoA of the DCG marketing game.

4.2 FMCG Marketing Game

The primary difference between FMCG and DCG is that customers repeatedly purchase a product in FMCG, and therefore whenever a customer tires of a product he is currently using, he will immediately switch to adopt a different product or not use any. However, he is still likely to adopt this product again later. This process is similar to dynamics of opinion diffusions and can serve as generalized voter models for the diffusion in competitive FMCGs.

The voter model, which first was proposed to formulate the dynamics of interacting particle systems [75, 111], was later applied to studying formation and diffusion of opinions across social networks [76, 77, 112–114]. Similar to the presentation in Section 4.1, here we consider the case where m players are simultaneously promoting their products. The adoption of each product (or none of the products) is

treated as a different opinion and thus there are in total $m + 1$ opinions propagating in the population.

We study diffusions of competitive FMCGs in a social network as specified in Section 4.1, where the population consists of n communities. At every time instant, an agent using one of the products (or not using any of them) attempts to persuade his neighbors to adopt his product (or not use any product). Also, this agent has the potential to switch to another product under the influence of his connected neighbors. As in Section 4.1, we assume that the agents in each community on average have no personal preference for any product(s) and use λ'_{ij} to denote the spreading rate for the influence of an agent from community j on agents in community i .

Also, each player (company) can increase the chances that agents in a community adopt his product by advertising in that community. We use $U_k(t) = \begin{bmatrix} u_k^1(t) & u_k^2(t) & \dots & u_k^n(t) \end{bmatrix}^T$ to denote the k -th player's strategy where $u_k^j(t)$ is his advertising investment in community j at time t . Also, we define $X_k(t) = \begin{bmatrix} x_k^1(t) & x_k^2(t) & \dots & x_k^n(t) \end{bmatrix}^T$ where $x_k^j(t)$ is the fraction of agents in community j who are using the k -th product at time t . We take $k = 0$ to apply to the agents choosing not to use any product at time t .

Generalizing the voter model in [76] and following a development similar to that for the community-based SIR model in Section 2.1, the FMCG competition

model is formulated as

$$\begin{aligned} \frac{d X_k(t)}{d t} = & \text{diag} [1_n - X_k(t)] (\Lambda X_k(t) + C_k U_k(t)) \\ & - \text{diag} [X_k(t)] \left(\sum_{i \neq k} \Lambda X_i(t) + C_i U_i(t) \right) \end{aligned} \quad (4.2)$$

where the notation Λ and C_k 's are as in Section 4.1. The dynamics of FMCG diffusion is composed of two terms. The first term is associated with the agents using the i -th product trying to influence others to use the same product, while the second term arises from the influence agents receive from their neighbors.

As in Section 4.1, we use \mathcal{D}_k and \mathcal{D} to denote the strategy set for player k and the strategy space, where the strategy profile $\mathbb{U}(\cdot) = [U_1(\cdot), U_2(\cdot), \dots, U_m(\cdot)]$. We assume that initially the whole population has not yet adopted any product, i.e., $X_0(0) = 1_n$. With a given strategy profile $\mathbb{U}(\cdot)$, the solution to the FMCG competition model (4.2) is uniquely determined, and is denoted as $\phi_k^{(2)}(t, \mathbb{U}(\cdot)) = X_k(t)$. The utility function for player k is $\zeta_k^{(2)}(\mathbb{U}(\cdot)) = W^T \cdot \phi_k^{(2)}(T, \mathbb{U}(\cdot))$ where T is the duration of the FMCG marketing campaign.

In the FMCG marketing game, each player k picks his strategy in \mathcal{D}_k to maximize his utility function $\zeta_k^{(2)}(\mathbb{U}(\cdot))$. In order to analyze the PoA, we also define social welfare as $\eta^{(2)}(\mathbb{U}(\cdot)) = \sum_{k=1}^m \zeta_k^{(2)}(\mathbb{U}(\cdot))$. Unless specified otherwise, the model (4.2) and the game defined as above are referred to as the FMCG competition model and the FMCG marketing game in the rest of this chapter. Next, we will give three important properties of the FMCG marketing game.

4.2.1 Competitiveness

Similar to the DCG marketing game, the FMCG marketing game is competitive. The next theorem asserts competitiveness of the FMCG marketing game.

Theorem 4.3 *The FMCG game is competitive, i.e., for any $U_k(\cdot), U'_k(\cdot) \in \mathcal{D}_k$ such that $U_k(\cdot) \leq U'_k(\cdot)$ we have $\zeta_k^{(2)}(\mathbb{U}(\cdot)) \leq \zeta_k^{(2)}(\mathbb{U}(\cdot) \oplus U'_k(\cdot))$ and $\zeta_i^{(2)}(\mathbb{U}(\cdot)) \geq \zeta_i^{(2)}(\mathbb{U}(\cdot) \oplus U'_k(\cdot))$ for any $k = 1, \dots, m$ and $i \neq k$.*

Proof: See Appendix B.12. □

4.2.2 Continuous Monotone Submodularity

The definition of continuous monotone submodularity is given in Section 4.1.1. As was the case for the DCG marketing game, the social welfare function of the FMCG marketing game is continuous monotone submodular. We record this assertion as a theorem.

Theorem 4.4 *The social welfare function $\eta^{(2)}(\cdot)$ for FMCG marketing game is continuous monotone submodular.*

Proof: See Appendix B.13. □

4.2.3 Concavity-Convexity

In the foregoing, we have shown that both the DCG and FMCG games enjoy the properties of continuous monotone submodularity and competitiveness. Notice

that the FMCG competition model (4.2) has simpler dynamics than the DCG competition model (4.1). Indeed, we will show that the FMCG marketing game features enjoys a stronger property, called “concavity-convexity.” Its mathematical definition is given as follows:

Definition 4.2 *A marketing game has the concavity-convexity property if the payoff of each player is concave w.r.t. his own strategy and convex w.r.t. his competitors’ strategies. More specifically, for a strategy profile $\mathbb{U}(\cdot)$ and two different strategies of the k -th player, $U_k(\cdot)$ and $U'_k(\cdot)$, we have*

$$\zeta_k(\alpha\mathbb{U}(\cdot) + (1 - \alpha)\mathbb{U}(\cdot) \oplus U'_k(\cdot)) \geq \alpha\zeta_k(\mathbb{U}(\cdot)) + (1 - \alpha)\zeta_k(\mathbb{U}(\cdot) \oplus U'_k(\cdot))$$

$$\zeta_i(\alpha\mathbb{U}(\cdot) + (1 - \alpha)\mathbb{U}(\cdot) \oplus U'_k(\cdot)) \leq \alpha\zeta_i(\mathbb{U}(\cdot)) + (1 - \alpha)\zeta_i(\mathbb{U}(\cdot) \oplus U'_k(\cdot))$$

where $\alpha \in [0, 1]$ and $i \neq k$. Function $\zeta_k(\cdot)$ is the payoff function of the k -th player.

Then the following theorem asserts that the FMCG marketing game has the aforementioned property.

Theorem 4.5 *The FMCG marketing game has the concavity-convexity property of Definition 4.2.*

Proof: See Appendix B.14. □

We will find that the three properties of FMCG marketing games shown in this section are very important later. It should be noted that the FMCG marketing game has stronger properties than the DCG marketing game, and thus we expect to have stronger conclusions for FMCG games in the analysis below.

4.3 Existence of Nash Equilibrium

For readability of the discussion on Nash equilibrium, we first define the concept of ϵ -equilibrium in the context of marketing games. Intuitively, if all players choose their strategies according to an ϵ -equilibrium, then a single player's switching his strategy will not bring him an additional benefit of more than ϵ .

Definition 4.3 *A given strategy profile $\mathbb{U}(\cdot)$ is an ϵ -equilibrium, where $\epsilon \geq 0$, if were the k -th player to switch his strategy from $U_k(\cdot) \in \mathcal{D}_k$ to $U'_k(\cdot) \in \mathcal{D}_k$, then $\zeta_k(t, \mathbb{U}(\cdot)) \geq \zeta_k(t, \mathbb{U}(\cdot) \oplus U'_k(\cdot)) - \epsilon$, where $\zeta_k(\cdot)$ is the utility function. The case in which $\epsilon = 0$ is special; in this situation, the corresponding strategy profile $\mathbb{U}(\cdot)$ is called a Nash equilibrium.*

It is well-known that a mixed-strategy Nash equilibrium always exists in games whose strategy space is finite dimensional. However, both DCG and FMCG marketing games are differential games, whose strategy spaces are of infinite dimension. In general a Nash equilibrium is not guaranteed to exist in such games. Our conclusions on the existence of Nash equilibrium in the DCG and FMCG marketing games are summarized in the following theorem.

Theorem 4.6 *For any $\epsilon > 0$ there exists an ϵ -equilibrium for the DCG marketing game as formulated in Section 4.1. Also, there exists a Nash equilibrium for the FMCG marketing game in Section 4.2.*

Proof: See Appendix B.15. □

As mentioned in Section 4.2, FMCG marketing games enjoy stronger properties than DCG marketing games. This observation is supported in the analysis of the Nash equilibrium's existence in this section: the former has a Nash equilibrium while the latter is only shown to have an ϵ -equilibrium.

4.4 Price of Anarchy

In game theory, Price of Anarchy (PoA) is a measure of efficiency degradation in a competition due to players' selfishness. It is a general concept because it applies in a variety of competitions, and the notion of efficiency can be defined in a variety of ways. In this chapter, we study the PoA of DCG and FMCG marketing games, where PoA is defined as follows:

Definition 4.4 *Assume that there exists a strategy profile $\mathbb{U}^*(\cdot)$ that maximizes social welfare function $\eta^{(\cdot)}(\cdot)$. Also, let $\mathbb{U}(\cdot)$ be a(n) (ϵ -)equilibrium for this marketing game. Then the Price of Anarchy (PoA) is defined as*

$$PoA = \frac{\eta^{(\cdot)}(\mathbb{U}^*(\cdot))}{\eta^{(\cdot)}(\mathbb{U}(\cdot))}$$

where $\eta^{(\cdot)}(\cdot)$ refers to the social welfare for both types of the marketing games.

Although the existence of $\mathbb{U}^*(\cdot)$ is an assumption in Definition 4.4, according to [115] or Theorem 5.1.1 in [116], such a $\mathbb{U}^*(\cdot)$ is guaranteed to exist in both the DCG and FMCG marketing games. The following theorem gives an upper bound of the PoA in marketing games.

Theorem 4.7 *For a given DCG marketing game, $\exists \delta > 0$ such that for any $0 < \epsilon < \delta$, if $\mathbb{U}(\cdot)$ is an ϵ -equilibrium, then $\eta^{(1)}(\mathbb{U}^*(\cdot)) \leq 2\eta^{(1)}(\mathbb{U}(\cdot))$. In an FMCG marketing game with $\mathbb{U}(\cdot)$ being the Nash equilibrium, $\eta^{(2)}(\mathbb{U}^*(\cdot)) \leq 2\eta^{(2)}(\mathbb{U}(\cdot))$ holds.*

Proof: See Appendix B.16. □

Theorem 4.7 shows that the PoA is bounded by 2 for both DCG and FMCG marketing games. This is reassuring, in that players' selfishness does not severely damage efficiency of the marketing game.

4.5 Almost Uniqueness of the Nash Equilibrium

Compared with existence, uniqueness of Nash equilibrium is more rare and also more valuable. In a real competition, if there are multiple Nash equilibria and players are not well coordinated, they may play according to different equilibria and results of the competition thus become unpredictable. However, if the Nash equilibrium is unique, we may be confident that all players will act based on the same equilibrium, making the game much more tractable.

For any strategy profile $\mathbb{U}(\cdot)$, if we change its value on a zero-measure subset, the payoff for each player remains unchanged. Therefore Nash equilibrium for marketing games cannot be strictly unique, nor even countable. To resolve this dilemma, we relax uniqueness of Nash equilibrium into "almost uniqueness" with the following definition.

Definition 4.5 *A marketing game has an almost unique Nash equilibrium if for any Nash equilibria $\mathbb{U}^1(\cdot)$, $\mathbb{U}^2(\cdot)$, we have $\mathbb{U}^1(\cdot) = \mathbb{U}^2(\cdot)$ almost everywhere on $[0, T]$.*

As mentioned in Section 4.2, FMCG marketing games have stronger properties than DCG games. In this section we focus on analyzing FMCG marketing games. We have the following theorem result.

Theorem 4.8 *If both $\mathbb{U}^1(\cdot)$ and $\mathbb{U}^2(\cdot)$ are Nash equilibria for an FMCG marketing game, then $\|\mathbb{U}^1(\cdot) - \mathbb{U}^2(\cdot)\|_2 = 0$.*

Proof: See Appendix B.17. □

Here $\|\mathbb{U}^1(\cdot) - \mathbb{U}^2(\cdot)\|_2 = 0$ is equivalent to requiring $\mathbb{U}^1(\cdot) = \mathbb{U}^2(\cdot)$ almost everywhere. However, it is actually extremely difficult to find one of such “almost unique” equilibria. In classical differential game theory [74], it is typical to convert specifications of a Nash equilibrium in differential games into a corresponding two-point boundary value problem, which can be quite intractable theoretically. In the next section, we will circumvent this dilemma and propose a distributive mechanism for searching for a Nash equilibrium.

4.6 Distributive Algorithm for Seeking a Nash Equilibrium

In the case where each player can repeatedly play the FMCG marketing game, here is a question we may ask: is there a way that the players can iteratively update their strategies such that the utility for each of them converges to the utility at the Nash equilibrium? Fortunately, the answer is in the affirmative. To analyze this problem, it is useful to first consider the following infinite dimensional dynamical

system:

$$\frac{dU_k(t, \cdot)}{dt} = f_k(\mathbb{U}(t, \cdot), \mu_k(\cdot)) = \nabla_k \zeta_k^{(2)}(\mathbb{U}(t, \cdot)) + \begin{bmatrix} \mu_{k1}(\cdot) \\ \mu_{k2}(\cdot) \\ \vdots \\ \mu_{kn}(\cdot) \end{bmatrix} - \begin{bmatrix} \mu_{k0}(\cdot) \\ \mu_{k0}(\cdot) \\ \vdots \\ \mu_{k0}(\cdot) \end{bmatrix} \quad (4.3)$$

The Fréchet differentiability of $\zeta_k^{(2)}(\cdot)$ by Theorem 3.2.6 in [116] ensures that the first term on the right side of (4.3) is well defined. Notice that the solution $\mathbb{U}(t, s)$ of (4.3) involves two dimensions of time. Time t refers to the time scale along which (4.3) evolves. For each t , $\mathbb{U}(t, \cdot)$ represents a strategy profile, which is a function on $[0, T]$, and thus $\mathbb{U}(t, s)$ is the value of the strategy profile $\mathbb{U}(t, \cdot)$ at time $s \in [0, T]$. Before we explain the auxiliary functional vector $\mu_k(\cdot) = [\mu_{ki}(\cdot)]_{n+1}$ for each k , we need to define an index set as follows:

$$\mathcal{J}_k(\mathbb{U}(t, s)) = \left\{ i \mid u_{ki}(t, s) \leq 0, i = 1 \dots n \right\} \cup \left\{ 0 \mid \sum_{i=1}^n u_{ki}(t, s) \geq b_k(s) \right\}$$

The index set $\mathcal{J}_k(\mathbb{U}(t, s))$ represents the indices for the constraints that are violated by the k -th player's strategy $U_k(t, \cdot)$ at time s . Then $\mu_k(s)$ is determined by the following equation:

$$\mu_k(s) = \underset{\substack{\mu_{ki}(s) \geq 0, (k,i) \in \mathcal{J}_k(\mathbb{U}(t,s)) \\ \mu_{ki}(s) = 0, (k,i) \notin \mathcal{J}_k(\mathbb{U}(t,s))}}{\arg \min}} \|f_k(\mathbb{U}(t, \cdot), \mu_k(\cdot))(s)\|_2 \quad (4.4)$$

Now, (4.3) can be interpreted intuitively: the term $\nabla_k \zeta_k^{(2)}(\mathbb{U}(t, \cdot))$ functions to maximize the utility $\zeta_k^{(2)}(\cdot)$ by driving $U_k(t, \cdot)$ in the gradient descent; the quantities $\mu_k(\cdot)$ are similar to Lagrange multipliers [117], and their inclusion aims to ensure the strategy $U_k(t, \cdot)$ stays in the feasible set \mathcal{D}_k . We call (4.3) the “best-response” mechanism.

In the FMCG marketing game, the almost uniqueness of the Nash equilibrium ensures the uniqueness of the payoff for each player k at all Nash equilibria given that the differences between the Nash equilibria on a zero-measure set make no difference to the diffusion process. Let $\zeta_k^{(2)}$ denote the utility for player k at the Nash equilibrium. The following theorem asserts that with the best-response mechanism, the strategy $U_k(t, \cdot)$ remains in \mathcal{D}_k and the corresponding utility converges to $\zeta_k^{(2)}$.

Theorem 4.9 *Let $\mathbb{U}(t, \cdot)$ evolve according to (4.3) starting from any feasible strategy profile $\mathbb{U}(0, \cdot) \in \mathcal{D}$. Then for each $1 \leq k \leq m$, $U_k(t, \cdot) \in \mathcal{D}_k$ for any time t . Also, $\lim_{t \rightarrow \infty} \zeta_k^{(2)}(\mathbb{U}(t, \cdot)) = \zeta_k^{(2)}$.*

Proof: See Appendix B.18. □

It should be noted that the best-response mechanism (4.3) is by nature a distributive algorithm, assuming that each player is aware of the opponents' strategies. More specifically, each player can update his strategy independently and Theorem 4.9 ensures that his payoff will converge to the utility at the Nash equilibrium.

Also, for practical purposes, in implementation of the algorithm, we need to discretize $\mathbb{U}(t, s)$ w.r.t. both t and s . Following such a discretization the strategy space \mathcal{D} becomes finite dimensional and analysis of both the DCG and the FMCG marketing games becomes simpler. For example, the existence of the Nash equilibrium becomes trivial and the analysis of the best-response mechanism (4.3) also simplifies. However, we emphasize the novelty of the case in which \mathcal{D} is a Banach space, since this extends existing differential game theory [74] and is therefore of theoretical value.

Chapter 5: Diffusion Centrality

In this chapter we study diffusion dynamics from a different perspective: investigating network centrality for a diffusion process. Centrality is a measure of importance of a vertex or a group of vertices in a network. It is of interest to find centrality notions such that the higher the centrality, the more important the vertex (or group of vertices) is for the network. Of course this depends on the quality that is of interest. In Section 5.1 we briefly recall several traditional network centrality metrics. In Section 5.2 we introduce a notion of sensitivity centrality for the scenario when a diffusion does not break out, and we relate this notion to Katz centrality. In Section 5.3 we introduce a notion of vulnerability centrality for the case in which a diffusion outbreak occurs, and show its connection with eigenvector centrality for the situation in which the outbreak affects only a small portion of the whole population.

5.1 Background and Related Work

In this section we recall several traditional centrality metrics, all based on the static topology of an unweighted network. Let $A = [a_{ij}]_{n \times n}$ be the adjacency matrix associated with the network, where $a_{ij} = 1$ if there is an edge connecting

vertex j to vertex i and $a_{ij} = 0$ otherwise. The traditional notions of degree centrality, eigenvector centrality and Katz centrality are summarized in the following subsections.

5.1.1 Degree Centrality

Degree centrality is perhaps the simplest centrality metric for vertices of a network. The degree centrality of a vertex equals the number of edges connected to the vertex. In a directed network, vertices have both an in-degree and an out-degree, and therefore we can consider both in-degree centrality and out-degree centrality as follows:

$$c_{in-d}(i) = \sum_{j=1}^n a_{ij} \quad c_{out-d}(i) = \sum_{j=1}^n a_{ji} \quad (5.1)$$

where $c_{in-d}(i)$ and $c_{out-d}(i)$ are in-degree and out-degree centralities for vertex i .

5.1.2 Eigenvector Centrality

If we view degree centrality of a vertex as a measure of its direct influence on other vertices, then eigenvector centrality incorporates both direct and indirect influence. With eigenvector centrality, the importance of each vertex is taken to scale with the sum of importance of its neighbors. More specifically, if $c_e(i)$ is the importance of vertex i , then we have the following equation:

$$c_e(i) = \kappa \sum_{j=1}^n a_{ij} c_e(j) \quad (5.2)$$

Here κ is a positive constant. By the Perron-Frobenius theorem, an adjacency matrix for a strongly-connected network has one and only one eigenvector with all

entries positive, and this eigenvector is associated with the largest eigenvalue $\lambda_1(A)$. Therefore by inserting $\kappa = \lambda_1(A)$ into (5.2), we find that the vector of eigenvector centralities is the eigenvector of the adjacency matrix A corresponding to the largest eigenvalue.

5.1.3 Katz Centrality

It is known that only vertices that are in a strongly connected component of two or more vertices, or the out-component of such a component, can have nonzero eigenvector centrality [118]. To overcome this issue, we define Katz centrality as

$$c_k(i) = \alpha \sum_{j=1}^n a_{ij} c_k(j) + 1 \quad (5.3)$$

where $c_k(i)$ is the Katz centrality for vertex i and α a positive constant. With $C_k = \begin{bmatrix} c_k(1) & c_k(2) & \dots & c_k(n) \end{bmatrix}$, equation (5.3) can be transformed into the following form:

$$C_k = (\mathbf{I}_n - \alpha A)^{-1} \cdot \mathbf{1}_n \quad (5.4)$$

For (5.4) to be well defined, the constant α should be taken between 0 and $\lambda_1(A)^{-1}$. A possible extension of the Katz centrality is obtained by allowing the additive term in (5.3) to not be the same for each vertex. In this way, we can generalize Katz centrality as follows:

$$c_k(i) = \alpha \sum_{j=1}^n a_{ij} c_k(j) + \beta_k \quad (5.5)$$

This generalized Katz centrality can also be written more compactly as $C_k = (\mathbf{I}_n - \alpha A)^{-1} \cdot \beta$ where $\beta = \begin{bmatrix} \beta_1 & \beta_2 & \dots & \beta_n \end{bmatrix}$. This generalization will be very helpful when discussing the diffusion centrality later in this chapter. Most researchers

set the value of α close to $\lambda_1(A)^{-1}$, in which case Katz centrality turns out to be numerically close to ordinary eigenvector centrality, but assigns small nonzero values to vertices that are not in strongly connected components or their out-components.

5.2 Sensitivity Centrality

One notion of diffusion centrality can be defined by increasing the percentage of infected agents in a specific community and investigating how this will change the fraction of individuals who are finally affected by the diffusion. This idea rests on a sensitivity analysis of the diffusion’s final state, and for this reason we refer to it as sensitivity centrality. In the scenario of viral marketing, a critical step is to initially target a few “influential” agents, and sensitivity centrality is helpful to determine how influential each agent is. As specified in Section 3.1.2, in the SIS model (2.6) the final state does not depend on the initial state. Therefore in this section we relegate our attention and analysis to the SIR model (2.5).

For the SIR model (2.5), let the initial state be $S(0) = 1_n - I_0 - R_0$, $I(0) = I_0$ and $R(0) = R_0$, then following the derivation in the proof of Theorem 2.4, we have that the final states $I(\infty) = 0$, $S(\infty) = 1_n - R(\infty)$ and $R := R(\infty)$ satisfy the following equation:

$$R + \text{diag}[1_n - I_0 - R_0] \exp(-\Gamma \Lambda^{-1}(R - R_0)) = 1_n \quad (5.6)$$

Following the proof for Lemma 3.2 in Section 3.1.2, we can show that with given $I_0, R_0 \in \mathbb{R}^n$ such that $I_0, R_0 \geq 0$ and $I_0 + R_0 < 1_n$, there exists a unique $R \in \mathbb{R}^n$ such that equation (5.6) holds. Therefore we can define a mapping $\phi : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$

such that $R = \phi(I_0, R_0)$ is the solution to (5.6).

Before proceeding to define sensitivity centrality, we first investigate the influence of one specific community over another. Intuitively, with a given initial condition, if we increase the fraction of agents that are initially infected in one community, then the influence of this community over other communities can be viewed in terms of the corresponding increase in the finally affected population. Let $t_{jk}(I_0, R_0)$ denote the influence of community j over community k with initial condition $I(0) = I_0$ and $R(0) = R_0$. Then we have

$$t_{jk}(I_0, R_0) = \left. \frac{\partial \phi_j(I, R)}{\partial i_k} \right|_{I=I_0, R=R_0} \quad (5.7)$$

where $\phi_j(\cdot)$ and i_k are the j -th entry of $\phi(\cdot)$ and k -th entry of I . Using the definition of influence $t_{jk}(I_0, R_0)$, the sensitivity centrality of each community is the weighted sum of its influence over other communities, i.e.

$$c_s(j, I_0, R_0) = \sum_{k=1}^n w_k t_{kj}(I_0, R_0) \quad (5.8)$$

Here the positive constant w_k is the weight associated with community k . The most important difference between sensitivity centrality and other static centrality metrics is that sensitivity centrality is associated with a diffusion process and hence dependent on the initial conditions of the diffusion. Formally, we define $c_s(j, I_0, R_0)$ as the sensitivity centrality of community j under the initial condition $I(0) = I_0$ and $R(0) = R_0$.

For convenience, we use the matrix form $T(I_0, R_0) = [t_{jk}(I_0, R_0)]_{n \times n}$. To investigate $t_{jk}(I_0, R_0)$, we calculate partial derivatives on both sides of (5.6) w.r.t.

each entry of I_0 , yielding the following equation:

$$T(I_0, R_0) - \text{diag}[1 - I_0 - R_0] D(R_0) \Lambda \Gamma^{-1} T(I_0, R_0) = D(R_0) \quad (5.9)$$

Here $D(R_0) = \text{diag}[\exp(-\Lambda \Gamma^{-1}(R - R_0))]$. To better interpret the influence matrix $T(I_0, R_0)$ in (5.9), we first investigate the case where a diffusion does not break out and $I_0 = R_0 = 0$, i.e., none of agents are involved in the diffusion initially. Then (5.9) is reduced to

$$T_0 - \Lambda \Gamma^{-1} T_0 = \mathbf{I}_n \quad (5.10)$$

where $T_0 = T(I_0, R_0)$. From (5.10) we know that $T_0 = (\mathbf{I}_n - \Lambda \Gamma^{-1})^{-1}$. Since the diffusion does not break out, we have $\lambda_1(\Lambda \Gamma^{-1}) < 1$ and therefore $T_0 = \sum_{i=0}^{\infty} (\Lambda \Gamma^{-1})^i$. Combining with the equation (5.8), we can conclude that in this case sensitivity centrality is connected with Katz centrality.

To better understand the connection between sensitivity centrality and Katz centrality, we consider T_0 in the following way. If information contagion spreads from one community to another via k intermediate communities, then we refer to it as a k -step spread. Correspondingly, we define k -step transmissibility as the probability that a k -step spread occurs from an agent in community j to another agent in community i , which is denoted as $t_{ij}^{(k)}$. According to Section 2.5.1, we have $t_{ij}^{(0)} = \frac{\lambda_{ij}}{\gamma_j}$, which can be written in compactly as $\left[t_{ij}^{(0)} \right]_{n \times n} = \Lambda \Gamma^{-1}$. Similar to counting walks of a specific length in a network [119], it is straightforward to show that the k -step transmissibility $\left[t_{ij}^{(k)} \right]_{n \times n} = (\Lambda \Gamma^{-1})^k$. Therefore for initial conditions $I_0 = R_0 = 0$, the influence defined in (5.7) is the sum of the transmissibilities in all steps from 0 to ∞ .

The foregoing connects sensitivity centrality to Katz centrality in the case where an outbreak does not occur and $I_0 = R_0 = 0$. As long as $\lambda_1(\Lambda\Gamma^{-1}) < 1$, i.e., if the diffusion does not break out, for any initial condition I_0 and R_0 the sensitivity centrality in (5.8) with influence matrix $T(I_0, R_0)$ determined by (5.9) takes the form of Katz centrality.

It should be noted that in most cases, at least in the scenario of social network marketing, an outbreak does not occur spontaneously. As long as a diffusion does not break out, the sensitivity centrality defined in this section is still in the form of Katz centrality and can be interpreted in the same way as in this Section.

5.3 Vulnerability Centrality

As mentioned in Section 5.2, sensitivity centrality can be readily interpreted in the case where a diffusion does not break out. However, when there is an outbreak, according to Definition 2.1 and Theorem 2.1, if spreading strengths dominates recovery rates, then an arbitrarily small fraction of initially infected agents will lead to a significant portion of the whole population being affected by the diffusion. In this case the idea to define diffusion centrality by generalizing the sensitivity analysis of final state equations does not work well.

Instead, we want to measure the importance of each community in a diffusion in terms of the size of the agent population that is influenced by the diffusion process. This idea arises from an idea that if a community is closer to the center in a diffusion process, the more information/contagion will propagate through this

community, and thus the higher the fraction of agents that will finally be influenced. In other words, we want to use the final state of a diffusion in each community as an index of its centrality in the diffusion process. We call this notion *vulnerability centrality*. We will investigate this notion based on both the SIR model (2.5) and the SIS model (2.6) in the case that a negligible fraction of agents are initially infected while all others are susceptible.

According to Theorem 2.1, when spreading strengths dominate recovery rates, i.e., $\Lambda - \Gamma \not\leq 0$ or $\lambda_1(\Lambda - \Gamma) > 0$, an outbreak occurs. Also, the larger $\lambda_1(\Lambda - \Gamma)$ is, the diffusion breaks out to a greater extent. When $\lambda_1(\Lambda - \Gamma)$ is positive but close to 0, we call the corresponding outbreak a “weak outbreak.” Here, we investigate vulnerability centrality in the scenario of a weak outbreak.

For SIR model (2.5), since it is assumed that only a negligible fraction of agents are initially infected, we just insert $I_0 = R_0 = 0$ into (5.6), obtaining

$$R + \exp(-\Lambda\Gamma^{-1}R) = 1_n \quad (5.11)$$

Use the Taylor expansion of $\exp(-\Lambda\Gamma^{-1}R) = \sum_{j=0}^{\infty} \frac{1}{j!} (-\Lambda\Gamma^{-1}R)^j$, where the j -th power of a vector is the j -th power of its entries. Then (5.11) can be written as

$$R = \Lambda\Gamma^{-1}R - \frac{1}{2!} (\Lambda\Gamma^{-1}R)^2 + \frac{1}{3!} (\Lambda\Gamma^{-1}R)^3 - \dots \quad (5.12)$$

In case of a weak outbreak, we can assume that R is small and thus neglect higher order terms in (5.12). Therefore vulnerability centrality in the weak outbreak case is expressed as follows:

$$R = \Lambda\Gamma^{-1}R \quad (5.13)$$

This equation indicates that in the weak outbreak case, vulnerability centrality based on SIR model (2.5) can be approximated as eigenvector centrality for matrix $\Lambda\Gamma^{-1}$.

For SIS model (2.6), we denote $I_\infty \in \mathbb{R}^n$ as the percentage of agents finally becoming infected in each community, then using equation (2.6b), we have

$$\text{diag}[1_n - I_\infty] \Lambda I_\infty - \Gamma I_\infty = 0 \quad (5.14)$$

This can be transformed into

$$\begin{aligned} \Lambda I_\infty &= \text{diag}[1_n - I_\infty]^{-1} \Gamma I_\infty \\ &= \Gamma (I_\infty + I_\infty^2 + \dots) \end{aligned} \quad (5.15)$$

Similar to the derivation based on SIR model (2.5), we neglect higher order terms in I_∞ in (5.15) and approximate the vulnerability centrality of SIS model (2.6) in the weak outbreak case as

$$I_\infty = \Gamma^{-1} \Lambda I_\infty \quad (5.16)$$

This equation implies that the vulnerability centrality is aligned with the primary eigenvector of $\Gamma^{-1}\Lambda$. Therefore in case of weak outbreak, vulnerability centrality for both the SIR and SIS models is connected with eigenvector centrality. Moreover, if the recovery rates are identical for all communities, then $\Gamma^{-1}\Lambda = \Lambda\Gamma^{-1}$ and therefore vulnerability centrality is the same for both models.

Chapter 6: Conclusions and Suggestions for Future Work

In this chapter we briefly summarize the work in this dissertation and discuss some issues for future work.

6.1 Main Contributions of the Dissertation

In this dissertation we have studied a variety of problems in the area of diffusion dynamics for large social networks, with a modeling theme that attempts to reduce computational burden by viewing a large network in terms of a network of connected homogeneous communities.

In Chapter 2 we proposed a community-based dynamical model for SIR and SIS infection diffusion. We have shown that our model generalize the well known mean-field approach and agent-based approach to modeling of SIR and SIS infection diffusion. Also, we investigated the accuracy of our model both through mathematical analysis and numerical simulations.

Based on our community-based dynamical model, we introduced a rigorous definition of infection outbreak and determined conditions for an outbreak to occur. We studied contributions of inter-community connections to a contagion based on the outbreak condition. Also, we investigated the final states of an infection diffusion

as well as the impact of a diffusion on the individual communities.

We also employed a percolation theory approach to study SIR diffusion in the community-based network. We observed that the predictions of analysis based on the dynamical model agree with those based on percolation theory. This was useful not only as an interesting exercise in its own right, but also as a check of reasonableness of the dynamical model and the ensuing predictions.

In Chapter 3 we studied referral marketing, seeding marketing and dynamic marketing problems based on the community-based SIR and SIS models. The referral marketing problem can be easily reformulated in a convex form. However, seeding marketing is a nonlinear optimization and dynamic marketing involves a nonlinear optimal control problem, which are both extremely difficult, if not impossible, to solve analytically. The difficulties in solving nonlinear optimization and optimal control problems drove us to seek alliterative, heuristic algorithms. In Chapter 3, we defined the continuous monotone submodularity property and showed that both seeding marketing and dynamic marketing enjoy this property. Based on this property, we proposed a greedy algorithm to solve seeding marketing and dynamic marketing problems, and found that the approximate solutions satisfy an approximation ratio bound of $1 - \frac{1}{e}$.

Also, we formulated two other variants of the dynamic marketing problem, including determining the minimum possible cost to achieve a given level of profits, and conducting a trade-off between profits and cost. The greedy heuristic was modified to solve the minimum cost problem and a local bi-directional search algorithm was designed for the trade-off problem both with analysis conducted on

corresponding approximation ratio, respectively. The greedy heuristic was also converted into closed-loop form for the SIR model, which is more resilient to external noise and model uncertainties. Also, all the marketing problems based on the SIS model were shown to be convex, and could therefore be solved numerically for the globally optimal solution.

The analysis of approximation ratios for all the heuristic algorithms designed in Chapter 3 was based on continuous submodularity. From the perspective of economics, continuous submodularity simply means the decreasing marginal profit with increasing investment. We believe that this property is satisfied in various problems related to marketing in social networks, and therefore the heuristic algorithms proposed in Chapter 3 are potentially applicable in other scenarios of marketing strategy design.

In Chapter 4 we studied the competition of multiple companies that are simultaneously promoting their products, all of which are interchangeable, substitute goods. As a first step, we proposed community-based dynamic models for diffusions of competitive DCGs as well as FMCGs. Also, we formulated DCG and FMCG marketing games based on these two models. Then we showed the existence of an ϵ -equilibrium for the DCG marketing game and a Nash equilibrium for the FMCG marketing game. We also showed that the Price of Anarchy (PoA) for both games is bounded by 2. We then showed that any two Nash equilibria of an FMCG marketing game agree almost everywhere, and designed a distributive mechanism to find Nash equilibria.

Our contributions in Chapter 4 can be viewed as addressing two basic issues.

Firstly, we proposed dynamic models for diffusion of competitive contagions on a social network. This enabled us to study continuous advertising over a specific network while existing literature either only allows each company to target a set of agents initially or fails to take the heterogeneous connections of the social network into account. Also, instead of invoking classical differential game theory, in Chapter 4 we analyzed the PoA of social network marketing games and proposed a mechanism to seek a Nash equilibrium based on its specific properties. The classical differential game theory typically converts configurations of a Nash equilibrium into a two-point boundary value problem, which is extremely difficult to solve analytically. Instead we designed a best-response mechanism (4.3), which circumvents this dilemma and finds a Nash equilibrium in an iterative fashion.

In Chapter 5 we proposed concepts of diffusion centrality, which attempts to rank the importance of each vertex in a diffusion process. The community-based SIR and SIS models formulated in Section 2.1 were employed for modeling the diffusion process. We investigated the importance of each community based on connections between initial states and final states in a diffusion process in case there is no diffusion outbreak. By undertaking a sensitivity analysis of the equation for the final state, we showed that diffusion centrality defined in this way is connected with Katz centrality. We also investigated diffusion centrality in case when an outbreak occurs. Through approximating the solution of the final state equation, we observed that this diffusion centrality notion is connected with eigenvector centrality when only a small portion of agents are affected in the diffusion process.

6.2 Suggestions for Future Work

The work presented in this dissertation gives a modeling framework that is conceptually simple and computationally tractable. However there are many issues requiring further study. Among the most basic is determining and implementing model parameter identification schemes, and tuning the parameters and models so that the models are predictive for real world social networks. Finding the best community detection algorithm for our setting and connecting them with our modeling framework can also be of great value.

In order to determine the contribution of a specific pair of communities and time-step to the objective in social network marketing problems, what we typically do is increase the corresponding input infinitesimally, simulate the diffusion and observe the resulting increase in profit. This process consumes major computational resources. Algorithm 1 is very expensive given it involves multiple such simulations at each iteration. Minoux [120] proposed an accelerated greedy algorithm maximizing discrete submodular set functions by using "lazy evaluations," which often leads to performance improvements of several orders of magnitude [121, 122]. The idea of Minoux is based on a fundamental observation that at each iteration it is not necessary to check all the remaining elements via simulation. Instead, some of the simulations can be skipped.

To accelerate Algorithm 1 for practical performance, we can further generalize the "lazy evaluation" idea from discrete submodular set functions to the continuous monotone submodular functions. With this generalization, our improved greedy

heuristics will be more practical in the application of social network marketing.

In addition, the work on dynamic marketing in Chapter 3 is based on the assumption that Λ and Γ are known. Also, it only applies in deterministic cases, which means there are either no external disturbances nor time-dependent variations in Λ or Γ . For practical purposes it is necessary to generalize our work to make it robust to such external perturbations.

In [123], Streeter *et al.* proposes an online algorithm for maximizing discrete submodular set functions [124]. His online algorithm was successfully applied in sensor selection [125] and webpage ranking [126, 127]. Guillory *et al.* [128, 129] suggests its potential application in viral marketing with hidden information. In the future, we can explore generalizing Streeter’s algorithm to continuous monotone submodular functions and thus solve dynamic marketing problems with uncertain or varying Λ or Γ .

Golovin *et al.* [122] propose an adaptive algorithm for stochastic optimization of discrete submodular set functions with external disturbances. We can examine whether it is possible to extend the work in [122] to continuous monotone submodular functions. This could allow solving dynamic marketing problems with external or even adversarial noises.

In Chapter 4, our research primarily focused on analyzing marketing games as well as a distributive algorithm for seeking a Nash equilibrium. However, since not every player will necessarily choose their strategy according to an equilibrium, it also makes sense to design an online marketing strategy for each individual player in the presence of competitors. By treating competitors’ marketing efforts as external

disturbances, the online adaptive algorithm as specified previously may be helpful for dynamic marketing against the actions of competitors.

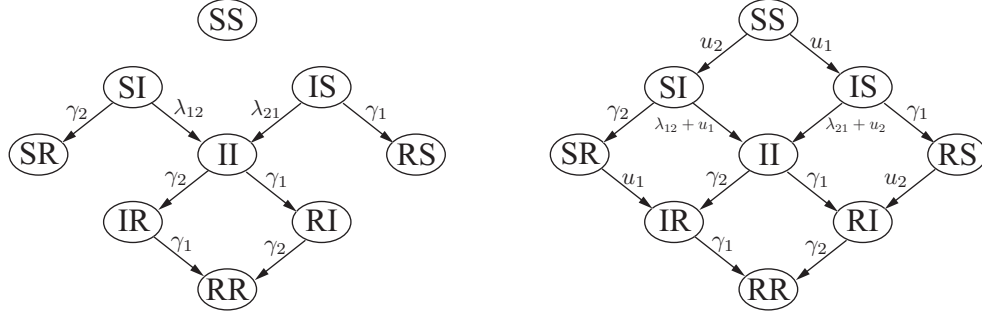
In Chapter 4 we considered social network marketing games where players independently compete with each other. However, in reality some players may cooperate with each other so as to gain higher profits, and this possibility is not taken into account in Chapter 4. How to appropriately formulate the players' cooperation and how to fairly distribute the profits becomes a problem worthy of further study.

Last but not least, it would also be interesting to apply our SIR and SIS models to empirical data. Because of the lack of a suitable dataset, the first step for this numerical work should be using a web crawler to collect the occurrence of several keywords within a specific time span from a fixed set of users in an online social network (such as Twitter). Next, we can detect community structures in the network formed by these users and fit our model to the collected data. This experimentation will allow us to assess how our models work in practice and explore possibilities for developing refinements to the models.

Appendix A: Derivation of the Agent-Based SIR Model

Consider all combinations of agents' states in the network. Since each agent can only be susceptible, infected or recovered, there are in total 3^N such combinations, which could be denoted as $x_1, \dots, x_{3^N} \in \{\text{sus}, \text{inf}, \text{rec}\}^N$. Here “sus” stands for susceptible, “inf” for infected and “rec” for recovered. Let $\mathcal{X} = \{x_1, \dots, x_{3^N}\}$ denote the set of all possible combinations and $y_i(t)$ as the probability that the combination of states at time t is x_i .

With all these preparations, a diffusion process is connected with a continuous-time Markov Chain (CTMC). Fig. A.1(a) shows an example of a Markov state transition diagram in a graph with $N = 2$ agents. For the general case, the corresponding transition rate matrix $Q = [q_{ij}]_{3^N \times 3^N}$ [30] is as follows:



(a) State Transition Diagram without Input (b) State Transition Diagram with Input

Figure A.1: Markov chain state transition diagram for a network of size 2

$$q_{ij} = \begin{cases} \sum_{l \neq k} a_{kl} \lambda_{kl} 1_{\{x_j^l = \text{inf}\}} & \text{if } x_j^k = \text{sus}, x_i^k = \text{inf} \\ & \text{and } x_j^l = x_i^l \text{ for } l = 1, 2, \dots, N, l \neq k \\ \gamma_k & \text{if } x_j^k = \text{inf}, x_i^k = \text{rec} \\ & \text{and } x_j^l = x_i^l \text{ for } l = 1, 2, \dots, N, l \neq k \\ - \sum_{k \neq j} q_{kj} & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \quad (\text{A.1})$$

Here x_i^k represents the k -th agent's state in combination x_i . In addition, $1_{\{x_j^l = \text{inf}\}} = 1$ if $x_j^l = \text{inf}$ and $1_{\{x_j^l = \text{inf}\}} = 0$ otherwise.

Then time dependence of the probability vector $y(t) = [y_1(t) \ y_2(t) \ \dots \ y_{3^N}(t)]^T$ satisfies the following differential equation:

$$\frac{d}{dt} y(t) = Q y(t) \quad (\text{A.2})$$

To serve our purpose to incorporate external influences on diffusion processes,

we would make further modifications on the previous model by Youseff [30]. We assume that recovering processes are spontaneous and cannot be altered. However, people can influence infecting processes. In addition to being influenced with infected neighbors, we can also put them under the influence of external source while the strength of such source can be artificially adjusted. In advertisement promotion scenarios, it means we can show advertisements in people's daily life in addition to the "word-of-mouth" effects as well inter-individual interactions to increase the chance of people knowing our products.

Let $u(t) = \begin{bmatrix} u_1(t) & u_2(t) & \dots & u_N(t) \end{bmatrix}$ stand for the external influence, where $u_i(t)$ is the strength of such influence for agent i at time t . With the influence $u(t)$, the Markov state transition diagram becomes Fig. A.1(b) and then transition rate matrix $Q'(u) = [q'_{ij}(u)]_{3^N \times 3^N}$ would become as follows:

$$q'_{ij}(u) = \begin{cases} \sum_{l \neq k} a_{kl} \lambda_{kl} 1_{\{x_j^l = \text{inf}\}} + u_k & \text{if } x_j^k = \text{sus}, x_i^k = \text{inf} \\ & \text{and } x_j^l = x_i^l \text{ for } l = 1, 2, \dots, N, l \neq k \\ \gamma_k & \text{if } x_j^k = \text{inf}, x_i^k = \text{rec} \\ & \text{and } x_j^l = x_i^l \text{ for } l = 1, 2, \dots, N, l \neq k \\ - \sum_{k \neq j} q_{kj} & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \quad (\text{A.3})$$

For convenience, we define $B_k = [b_{ij}^k]_{3^N \times 3^N}$ with $k = 1, 2, \dots, N$ as follows:

$$b_{ij}^k = \begin{cases} 1 & \text{if } x_j^k = \text{sus}, x_i^k = \text{inf} \\ & \text{and } x_j^l = x_i^l \text{ for } l = 1, 2, \dots, N, l \neq k \\ 0 & \text{otherwise} \end{cases}$$

With definition of B_k 's, we can write $Q'(u) = Q + \sum_{k=1}^N u_k(t)B_k$. Thus time dependence of the probability state vector $y(t)$ would obey the following differential equation:

$$\frac{d}{dt}y(t) = (Q + \sum_{k=1}^N u_k(t)B_k)y(t) \quad (\text{A.4})$$

As for a network consisted of N agents, the state of each agent j at any time t , denoted as $X_j(t)$, in the network can be susceptible, infected or recovered. Let $s_j(t) = \Pr[X_j(t) = \text{sus}]$, $i_j(t) = \Pr[X_j(t) = \text{inf}]$ and $r_j(t) = \Pr[X_j(t) = \text{rec}]$ denoting the probability that agent j is susceptible, infected or recovered. Applying Markov theory specifically to study such probabilities [29,30], we have the following equation:

$$\frac{d}{dt} \begin{bmatrix} s_j(t) \\ i_j(t) \\ r_j(t) \end{bmatrix} = \begin{bmatrix} - \left(\sum_{k=1}^N a_{jk} \lambda_{jk} 1_{\{X_k = \text{inf}\}} + u_j(t) \right) & 0 & 0 \\ \sum_{k=1}^N a_{jk} \lambda_{jk} 1_{\{X_k = \text{inf}\}} + u_j(t) & -\gamma_i & 0 \\ 0 & \gamma_i & 0 \end{bmatrix} \begin{bmatrix} s_j(t) \\ i_j(t) \\ r_j(t) \end{bmatrix} \quad (\text{A.5})$$

Here the indicator function $1_{\{X_k = \text{inf}\}}$ equals to 1 when $X_k(t) = \text{inf}$ and equals to 0 otherwise.

It should be noted that the equation (A.5) is no longer Markovian since transition rates from being susceptible to infected for each agent is also a random variable,

which depends on the rest of the network. With the assumption that $X_i(t) \perp\!\!\!\perp X_j(t)$ for any $1 \leq i, j \leq N$ and time t , we would like to use the mathematical expectation of the transition rate to replace the actual rate.

Notice that $\mathbf{E} [1_{\{X_k(t)=\text{inf}\}}] = \Pr [X_k(t) = \text{inf}] = i_k(t)$ while a_{jk} 's, λ_{kj} 's and γ_i 's are all given constants. Therefore we have

$$\mathbf{E} \left[\sum_{k=1}^N a_{jk} \lambda_{jk} 1_{\{X_k=\text{inf}\}} + u_j(t) \right] = \sum_{k=1}^N a_{jk} \lambda_{jk} i_k(t) + u_j(t)$$

Then the equation (A.5) becomes

$$\begin{aligned} \frac{d s_j(t)}{d t} &= -s_j(t) \left(\sum_{k=1}^N a_{jk} \lambda_{jk} i_k(t) + u_j(t) \right) \\ \frac{d i_j(t)}{d t} &= s_j(t) \left(\sum_{k=1}^N a_{jk} \lambda_{jk} i_k(t) + u_j(t) \right) - \gamma_j i_j(t) \\ \frac{d r_j(t)}{d t} &= \gamma_j i_j(t) \end{aligned} \tag{A.6}$$

The equation (A.6) is the agent-based SIR dynamical model we derive from Markovian model (A.4). Notice that we replace transition rates of each state with the corresponding mathematical expectation to derive agent-based SIR model (A.6), therefore (A.6) is actually an approximation of Markovian model (A.4).

Appendix B: Proofs of Several Theorems

B.1 Proof for Theorem 3.1

We prove it by contradiction here. If R is the solution to (3.2) and the equality on k -th row of this constraint does not hold, i.e. $r_k + \exp\left([- \Lambda' \Gamma^{-1} R]_k\right) < 1$, where $[\cdot]_k$ is the k -th component of the vector. Then if we increase r_k by a little bit, $r_k + \exp\left([- \Lambda' \Gamma^{-1} R]_k\right) < 1$ still holds. Meanwhile, for the other rows of this inequality, i.e. $r_j + \exp\left([- \Lambda' \Gamma^{-1} R]_j\right) \leq 1$ with $j \neq k$. Since r_j 's are not changed while r_k is increased by a little bit, noted that all the components of $\Lambda' \Gamma^{-1}$ are nonnegative, the term $\exp\left([- \Lambda' \Gamma^{-1} R]_j\right)$ is at least not increased. Thus the constraint $r_j + \exp\left([- \Lambda' \Gamma^{-1} R]_j\right) \leq 1$ still holds but the objective function $Z^T \cdot R$ is increased. This is a contradiction.

B.2 Proof for Lemma 3.1

Without loss of generality, let's assume that Λ is irreducible. This is because if Λ is reducible, then the associated network can be divided into several connected components, and we can apply Theorem 3.1 on each component respectively.

If the nonzero R satisfies that $R \in [0, 1]^n$, then first we will show that $R <$

$\Lambda\Gamma^{-1}R$ in the component-wise sense by contradiction. Otherwise we have $r_k \geq [\Lambda\Gamma^{-1}R]_k$ for some $1 \leq k \leq n$, where $[\cdot]_k$ is the k -th component of the vector. On the other hand, $r_k + \exp(-[\Lambda\Gamma^{-1}R]_k) = 1$. Since $r_k \geq [\Lambda\Gamma^{-1}R]_k$, we have $r_k + \exp(-r_k) \leq 1$. For $r_k \geq 0$, the only possibility is that $r_k = 0$. However, according to Theorem 2.5 in Section 2.4.2, if Λ is irreducible, $r_k = 0$ means that $R = 0$, which is a contradiction.

Now since $R < \Lambda\Gamma^{-1}R$, with $R' = \Gamma^{-1}R$ we have $\Lambda R' > \Gamma R'$, i.e. $(\Lambda - \Gamma)R' > 0$. Since R' is nonzero, we have $R'^T (\Lambda - \Gamma) R' > 0$. Therefore we may conclude that $\Lambda - \Gamma \not\leq 0$. Proved.

B.3 Proof for Lemma 3.2

Suppose that for some $X_0, \exists R_1, R_2 \in [0, 1]^n$ where $R_1 \neq R_2$ are both the solutions. Define $h(\alpha)$ as

$$h(\alpha) = (\alpha R_1 + (1 - \alpha) R_2) + \text{diag} [1_n - \text{diag} [W]^{-1} X_0] \exp(-\Lambda\Gamma^{-1}(\alpha R_1 + (1 - \alpha) R_2))$$

Obviously that $h(\cdot)$ is strictly convex w.r.t. α . Select $k \in \arg \min_{\substack{1 \leq k \leq n \\ r_1^k < r_2^k}} \left\{ \frac{r_2^k}{r_2^k - r_1^k} \right\}$ and $\alpha_0 = \frac{r_2^k}{r_2^k - r_1^k}$ where r_p^q represents the q -th component of R_p , $p = 1, 2$ and $1 \leq q \leq n$. Note that $h(\cdot)$ is convex, $h(0) = h(1) = 1_n$. Since $\alpha_0 > 1$, therefore $h(\alpha_0) > 1_n$. However, since $\alpha_0 R_1 + (1 - \alpha_0) R_2 \geq 0$ with its k -th entry being 0, therefore $h_k(\alpha_0) = \exp\left(-\sum_{i=1}^n \lambda_{ki} \gamma_i^{-1} (\alpha_0 r_1^i + (1 - \alpha_0) r_2^i)\right) \leq 1$, where $h_k(\cdot)$ is the k -th component of $h(\cdot)$. This is a contradiction. Proved.

B.4 Proof for Theorem 3.2

We are going to show the continuous monotone submodularity of the $\phi^{(1)}(\cdot)$, $-\phi_T^{(2)}(\cdot)$ and $-\phi_T^{(3)}(\cdot)$ respectively.

B.4.1 Continuous Monotone Submodularity of $\phi^{(1)}(\cdot)$

First, we need to show the monotonicity of $\phi^{(1)}(\cdot)$ and the following lemma is helpful.

Lemma B.1 *For any feasible $X_0 \in D$, if $R' \in [0, 1]^n$ satisfies that*

$$R' + \text{diag} [1_n - \text{diag}[W]^{-1} X_0] \exp(\Lambda \Gamma^{-1} R') \leq 1_n$$

then $R = \phi^{(1)}(X_0) \geq R'$.

Proof: We show by contradiction. Define index set $\mathcal{Z} = \{j \mid r_j < r'_j, 1 \leq j \leq n\}$ where r_j is the j -th component of R , r'_j is the j -th component of R' and x_j is the j -th component of X_0 . If $\mathcal{Z} \neq \emptyset$, then similar to the proof of Lemma 3.2, let $k \in \arg \min_{k \in \mathcal{Z}} \left\{ \frac{r'_k}{r'_k - r_k} \right\}$ and $\alpha_0 = \frac{r'_k}{r'_k - r_k}$. It's easy to check that $\alpha_0 > 1$. Also, let define $h_k(\alpha)$ as follows:

$$h_k(\alpha) = \alpha r_k + (1 - \alpha) r'_k + (1 - w_k^{-1} x_k) \exp \left(- \sum_{i=1}^n \lambda_{ki} \gamma_i^{-1} (\alpha r_i + (1 - \alpha) r'_i) \right)$$

It is easy to verify that $h_k(\alpha)$ is strictly convex w.r.t. α . Also, because $h_k(0) \leq 1$, $h_k(1) = 1$ and $\alpha_0 > 1$, we can conclude that $h_k(\alpha_0) > 1$. However, because $\alpha_0 r_k + (1 - \alpha_0) r'_k = 0$, then $h_k(\alpha_0) = (1 - w_k^{-1} x_k) \exp \left(- \sum_{i=1}^n \lambda_{ki} \gamma_i^{-1} (\alpha_0 r_i + (1 - \alpha_0) r'_i) \right) < 1$. This is a contradiction.

□

With Theorem B.1, we can show the following corollary about $\phi^{(1)}(\cdot)$ being monotone:

Corollary B.1 *For any feasible $X_0, X'_0 \in D$, if $X_0 \leq X'_0$ then $\phi^{(1)}(X_0) \leq \phi^{(1)}(X'_0)$.*

Proof: Let $R = \phi^{(1)}(X_0)$, then easily we can show

$$R + \text{diag} [1_n - \text{diag}[W]^{-1} X'_0] \exp (\Lambda \Gamma^{-1} R) \leq 1_n$$

Using Theorem B.1, we have $R' = \phi^{(1)}(X'_0) \geq \phi^{(1)}(X_0) = R$. Proved. □

The next lemma is very helpful to show the continuous submodularity of $\phi^{(1)}(\cdot)$.

Lemma B.2 *For any feasible $X_0, X'_0 \in D$ such that $X_0 \leq X'_0$, with $\forall \beta \in [0, 1]$, we have*

$$\phi^{(1)} (\beta X_0 + (1 - \beta) X'_0) \geq \beta \phi^{(1)} (X_0) + (1 - \beta) \phi^{(1)} (X'_0)$$

Proof: Let $R = \phi^{(1)}(X_0)$ and $R' = \phi^{(1)}(X'_0)$. By Theorem B.1 we know that $R \leq R'$. Then, let's look at the following inequalities:

$$\begin{aligned} & (\beta R + (1 - \beta) R') + \text{diag} [1_n - \text{diag}[W]^{-1} (\beta X_0 + (1 - \beta) X'_0)] \\ & \cdot \exp (-\Lambda \Gamma^{-1} (\beta R + (1 - \beta) R')) \\ & \leq (\beta R + (1 - \beta) R') + \text{diag} [1_n - \text{diag}[W]^{-1} (\beta X_0 + (1 - \beta) X'_0)] \\ & \cdot (\beta \exp (-\Lambda \Gamma^{-1} R) + (1 - \beta) \exp (-\Lambda \Gamma^{-1} R')) \\ & \leq (\beta R + (1 - \beta) R') + \beta \text{diag} [1_n - \text{diag}[W]^{-1} X_0] \exp (-\Lambda \Gamma^{-1} R) \\ & \quad + (1 - \beta) \text{diag} [1_n - \text{diag}[W]^{-1} X'_0] \exp (-\Lambda \Gamma^{-1} R') \end{aligned}$$

$$\begin{aligned}
&= \beta (R + \text{diag} [1_n - \text{diag}[W]^{-1}X_0] \exp (-\Lambda\Gamma^{-1}R)) \\
&\quad (1 - \beta) (R' + \text{diag} [1_n - \text{diag}[W]^{-1}X'_0] \exp (-\Lambda\Gamma^{-1}R')) \\
&= 1_n
\end{aligned}$$

The first inequality holds because each component of $\exp (\Lambda\Gamma^{-1}R)$ is convex w.r.t. R . The second equality holds in that $\exp (-\Lambda\Gamma^{-1}R) \geq \exp (-\Lambda\Gamma^{-1}R')$ since $R \leq R'$. Using Lemma B.1, we can see that $\phi^{(1)} (\beta X_0 + (1 - \beta) X'_0) \geq \beta \phi^{(1)} (X_0) + (1 - \beta) \phi^{(1)} (X'_0)$. Proved. \square

The property of $\phi^{(1)}(\cdot)$ in Lemma B.2 is quite similar to convexity. However, we need to notice that Lemma B.2 requires that $X_0 \leq X'_0$ while the convexity is not limited by it. Then we can come to the following lemma, which is very close the definition of continuous submodularity.

Lemma B.3 *For any feasible $X_0 \leq X'_0$ and any ΔX such that $X_0 + \Delta X, X'_0 + \Delta X \in D$, we have that $\phi^{(1)} (X_0 + \Delta X) - \phi^{(1)} (X_0) \geq \phi^{(1)} (X'_0 + \Delta X) - \phi^{(1)} (X'_0)$.*

Proof: For convenience let's denote $f(I, R)$ as following

$$f(X, R) = R + \text{diag} [1_n - \text{diag}[W]^{-1}X] \exp (-\Lambda\Gamma^{-1}R)$$

Then, with fixed I , the first-order derivative of $f(I, R)$ with respect to R is

$$\frac{\partial}{\partial R} f(X, R) = \mathbf{I}_n - \text{diag} [1_n - \text{diag}[W]^{-1}X] \text{diag} [\exp (-\Lambda\Gamma^{-1}R)] \Lambda\Gamma^{-1} \quad (\text{B.1})$$

where \mathbf{I}_n represents the identity matrix. Using Theorem B.1 in combination with (B.1), we can see that

$$\left. \frac{\partial}{\partial R} f(X, R) \right|_{X=X_0, R=\phi^{(1)}(X_0)+\Delta R} \leq \left. \frac{\partial}{\partial R} f(X, R) \right|_{X=X'_0, R=\phi^{(1)}(X'_0)+\Delta R} \quad (\text{B.2})$$

for any ΔR such that $R_1 + \Delta R, R_2 \Delta R \in D$.

Let $R_1 = \phi^{(1)}(X_0)$, $R'_1 = \phi^{(1)}(X_0 + \Delta X)$, $R_2 = \phi^{(1)}(X'_0)$ and $R'_2 = \phi^{(1)}(X'_0 + \Delta X)$. Then using (B.2) we have

$$\begin{aligned}
& f(X_0 + \Delta X, R_1 + (R'_2 - R_2)) - f(X_0 + \Delta X, R_1) \\
&= \int_{R_1}^{R_1 + (R'_2 - R_2)} \frac{\partial}{\partial R} f(X_0 + \Delta X, R) \, dR \\
&\leq \int_{R_2}^{R_2 + (R'_2 - R_2)} \frac{\partial}{\partial R} f(X'_0 + \Delta X, R) \, dR \\
&= f(X'_0 + \Delta X, R_2 + (R'_2 - R_2)) - f(X'_0 + \Delta X, R_2) \\
&= f(X'_0, R_2) - f(X'_0 + \Delta X, R_2)
\end{aligned}$$

On the other hand,

$$\begin{aligned}
& f(X'_0, R_2) - f(X'_0 + \Delta X, R_2) \\
&= \text{diag} [\text{diag}[W]^{-1} \Delta X] \exp(-\Lambda \Gamma^{-1} R_2) \\
&\leq \text{diag} [\text{diag}[W]^{-1} \Delta X] \exp(-\Lambda \Gamma^{-1} R_1) \\
&= f(X_0, R_1) - f(X_0 + \Delta X, R_1) \\
&= f(X_0 + \Delta X, R'_1) - f(X_0 + \Delta X, R_1)
\end{aligned}$$

Therefore we have $f(X_0 + \Delta X, R_1 + (R'_2 - R_2)) \leq f(X_0 + \Delta X, R'_1) = 1_n$.

Using Theorem B.1, we have $R_1 + (R'_2 - R_2) \leq R'_1$, namely $R'_1 - R_1 \geq R'_2 - R_2$.

Proved. \square

Now with all these preparations, we can come to the following theorem showing that $\phi^{(1)}(\cdot)$ is continuous monotone submodular.

Theorem B.1 For any X_0 and any $\Delta X \geq 0$ such that $X_0, X_0 + \Delta X \in D$, we have

$\frac{\partial}{\partial \alpha} \phi^{(1)}(X_0 + \alpha \Delta X) \Big|_{\alpha=0} \geq 0$. Also, with any $X'_0 \geq X_0$ such that $X'_0, X'_0 + \Delta X \in D$, we have $\frac{\partial}{\partial \alpha} \phi^{(1)}(X_0 + \alpha \Delta X) \Big|_{\alpha=0} \geq \frac{\partial}{\partial \alpha} \phi^{(1)}(X'_0 + \alpha \Delta X) \Big|_{\alpha=0}$.

Proof: First of all, for $\forall 0 < \alpha_1 < \alpha_2$, by using Theorem B.2, we have

$$\frac{\alpha_1}{\alpha_2} \phi^{(1)}(X_0 + \alpha_2 \Delta X) + \frac{\alpha_2 - \alpha_1}{\alpha_2} \phi^{(1)}(X_0) \leq \phi^{(1)}(X_0 + \alpha_1 \Delta X)$$

Therefore $\frac{\phi^{(1)}(X_0 + \alpha_1 \Delta X) - \phi^{(1)}(X_0)}{\alpha_1} \geq \frac{\phi^{(1)}(X_0 + \alpha_2 \Delta X) - \phi^{(1)}(X_0)}{\alpha_2}$. This means that we have $\frac{\phi^{(1)}(X_0 + \alpha \Delta X) - \phi^{(1)}(X_0)}{\alpha}$ is increasing as α decreases. Therefore $\frac{\phi^{(1)}(X_0 + \alpha \Delta X) - \phi^{(1)}(X_0)}{\alpha}$ is convergent (including being convergent to $+\infty$) as $\alpha \rightarrow 0^+$ and thus $\frac{\partial}{\partial \alpha} \phi^{(1)}(X_0 + \alpha \Delta X) \Big|_{\alpha=0}$ exists. The fact that $\phi^{(1)}(\cdot)$ is monotone guarantees that $\frac{\partial}{\partial \alpha} \phi^{(1)}(X_0 + \alpha \Delta X) \Big|_{\alpha=0} \geq 0$.

With Theorem B.3, we can show that for $\alpha > 0$,

$$\frac{\phi^{(1)}(X_0 + \alpha \Delta X) - \phi^{(1)}(X_0)}{\alpha} \geq \frac{\phi^{(1)}(X'_0 + \alpha \Delta X) - \phi^{(1)}(X'_0)}{\alpha}$$

Therefore we have that $\frac{\partial}{\partial \alpha} \phi^{(1)}(X_0 + \alpha \Delta X) \Big|_{\alpha=0} \geq \frac{\partial}{\partial \alpha} \phi^{(1)}(X'_0 + \alpha \Delta X) \Big|_{\alpha=0}$.

Proved. □

It should be noted that the derivative $\frac{\partial}{\partial \alpha} \phi^{(1)}(X_0 + \alpha \Delta X) \Big|_{\alpha=0}$ is potential to be infinite. However, since each component of $\frac{\partial}{\partial \alpha} \phi^{(1)}(X_0 + \alpha \Delta X) \Big|_{\alpha=0}$ gets smaller as X_0 gets greater, mostly $\frac{\partial}{\partial \alpha} \phi^{(1)}(X_0 + \alpha \Delta X) \Big|_{\alpha=0}$ takes a regular real value.

By now, we have showed the continuous monotone submodularity of $\phi^{(1)}(\cdot)$.

B.4.2 Continuous Monotone Submodularity of $-\phi_T^{(2)}$

Before we proceed into details, we will restate the following lemma from [59], which will be helpful later.

Lemma B.4 *Let $\Phi_A(t, s)$ be the fundamental matrix to the system*

$$\dot{x}(t) = A(t)x(t)$$

where $A(t)$ is locally bounded and Metzler, i.e. $A_{ij} \geq 0$ for $i \neq j$ then we have

$$\Phi_A(t, s) \geq 0, \quad t \geq s.$$

Proof: For any compact interval I one can find a constant c such that $B(t) := A(t) + c \cdot \mathbf{I}_n \geq 0$ for $t \in I$. We have $\Phi_A(t, s) = \exp(-c(t-s)) \Phi_B(t, s)$ where

$$\frac{d}{dt} \Phi_B(t, s) = B(t) \Phi_B(t, s)$$

$$\Phi_B(s, s) = \mathbf{I}_n$$

from which follows that $\Phi_B(t, s) \geq 0$ for $t \geq s$ and $[s, t] \subseteq I$. Proved. \square

Firstly we discuss about the monotonicity of $\phi_T^{(2)}(\cdot)$ and the following lemma illustrates this.

Lemma B.5 *For any $V(\cdot) \geq 0$ and continuous $\Delta V(\cdot) \geq 0$, we have the inequality*

$$\left. \frac{\partial}{\partial \alpha} \phi_T^{(2)}(V(\cdot) + \alpha \Delta V(\cdot)) \right|_{\alpha=0} \leq 0 \text{ and thus } \phi_T^{(2)}(V(\cdot) + \alpha \Delta V(\cdot)) \leq \phi_T^{(2)}(V(\cdot)).$$

Proof: In the SIR model (2.5) one of the three equations is redundant since $S(t) + I(t) + R(t) = 1_n$ is constant. Therefore we consider the two equations of them as

follows:

$$\begin{aligned}\frac{d S_\alpha(t)}{d t} &= -\operatorname{diag}\left[S_\alpha(t)\right]\left(\Lambda\left(1_n - S_\alpha(t) - R_\alpha(t)\right) + \operatorname{diag}[W]^{-1}\left(V(t) + \alpha \Delta V(t)\right)\right) \\ \frac{d R_\alpha(t)}{d t} &= \Gamma\left(1_n - S_\alpha(t) - R_\alpha(t)\right)\end{aligned}\tag{B.3}$$

where $S_\alpha(t)$ and $R_\alpha(t)$ means that $S(t)$ and $R(t)$ are related to α .

Fix T , $V(\cdot)$ and $\Delta V(\cdot)$. For convenience let's define $\Delta_S(\alpha, t) := S_\alpha(t) - S_0(t)$ and $\Delta_R(\alpha, t) := R_\alpha(t) - R_0(t)$. Calculate the partial derivative w.r.t. α on both sides of the equation (B.3), we have

$$\begin{aligned}\frac{d}{d t} \frac{\partial}{\partial \alpha} \Delta_S(\alpha, t) &= A_{11}(t) \frac{\partial}{\partial \alpha} \Delta_S(\alpha, t) + A_{12}(t) \frac{\partial}{\partial \alpha} \Delta_R(\alpha, t) + B_1(t) \Delta V(t) \\ \frac{d}{d t} \frac{\partial}{\partial \alpha} \Delta_R(\alpha, t) &= A_{21}(t) \frac{\partial}{\partial \alpha} \Delta_S(\alpha, t) + A_{22}(t) \frac{\partial}{\partial \alpha} \Delta_R(\alpha, t) + B_2(t) \Delta V(t)\end{aligned}\tag{B.4}$$

where $A_{11}(t) = -\operatorname{diag}\left[\Lambda\left(1_n - S_\alpha(t) - R_\alpha(t)\right) + \operatorname{diag}[W]^{-1}\left(V(t) + \alpha \Delta V(t)\right)\right] + \operatorname{diag}\left[S_\alpha(t)\right] \Lambda$, $A_{12}(t) = \operatorname{diag}\left[S_\alpha(t)\right] \Lambda$, $A_{21}(t) = A_{22}(t) = -\Gamma$. For the matrix B , $B_1(t) = -\operatorname{diag}\left[S_\alpha(t)\right] \operatorname{diag}[W]^{-1}$ and $B_2(t) = 0$.

Without loss of generality, we assume that the nonnegative $V(0) \neq 0$. Looking at the equation (B.4), since $\Delta_S(\alpha, 0) = \Delta_R(\alpha, 0) = 0$ and $B_1(0) < 0$, which means that the term $B_1(t) \Delta V(t) < 0$ dominates the other terms around the neighborhood of $t = 0^+$, therefore $\frac{\partial}{\partial \alpha} \Delta_S(\alpha, t) < 0$ around $t = 0^+$ for any $\alpha \geq 0$. If for some $\alpha > 0$, $\frac{\partial}{\partial \alpha} \Delta_S(\alpha, t)$ is not negative on $t \in [0, T]$, we can find a $t_0 \in [0, T]$ so that at least one component of $\frac{\partial}{\partial \alpha} \Delta_S(\alpha, t_0)$, denoted as $\frac{\partial}{\partial \alpha} \Delta_S^k(\alpha, t_0)$, equals to zero and for $\forall 0 < t < t_0$, $\frac{\partial}{\partial \alpha} \Delta_S(\alpha, t) < 0$.

Let's look back at (B.3), for the first equation, we can divide the j -th row of it by the corresponding $s_j(t)$. Notice that $\frac{d s_j(t)}{s_j(t)} = d \log\left(s_j(t)\right)$, and $\Gamma^{-1} \frac{d R_\alpha(t)}{d t} =$

$I_\alpha(t)$, we have

$$d \log (S_\alpha(t)) = -\Lambda \Gamma^{-1} d R_\alpha(t) - \text{diag} [W]^{-1} (V(t) + \alpha \Delta V(t)) dt$$

Integrate both sides the this equation from $t = 0$ to $t = t_0$, we have

$$\log(S(t_0)) = -\Lambda \Gamma^{-1} R(t_0) - \int_0^{t_0} \text{diag} [W]^{-1} (V(t) + \alpha \Delta V(t)) dt$$

Calculate the partial derivative w.r.t. α of both sides of the equation, we have

$$\text{diag} [S_\alpha(t_0)]^{-1} \frac{\partial}{\partial \alpha} \Delta_S(\alpha, t_0) = -\Lambda \Gamma^{-1} \frac{\partial}{\partial \alpha} \Delta_R(\alpha, t_0) - \int_0^{t_0} \text{diag} [W]^{-1} \Delta V(t) dt$$

Since $\frac{\partial}{\partial \alpha} \Delta_S^k(\alpha, t_0) = 0$, $S_\alpha(t_0) > 0$ and $\int_0^{t_0} \Delta V(t) dt > 0$, it is easy to show that $\left[\Lambda \Gamma^{-1} \frac{\partial}{\partial \alpha} \Delta_R(\alpha, t_0) \right]_k < 0$, where $[\cdot]_k$ is the k -th component of the given vector.

However, if we look at the second equation in (B.4), we have that

$$\left[\frac{\partial}{\partial \alpha} \Lambda \Gamma^{-1} \Delta_R(\alpha, t_0) \right]_k = - \int_0^{t_0} \left[\Lambda \Gamma^{-1} \exp(-\Gamma(t_0 - t)) \Gamma \frac{\partial}{\partial \alpha} \Delta_S(\alpha, t) \right]_k dt$$

As we assume, for $t \in [0, t_0]$, $\frac{\partial}{\partial \alpha} \Delta_S(\alpha, t) \leq 0$. Since $\exp(-\Gamma(t_0 - t)) \Gamma > 0$, therefore $\exp(-\Gamma(t_0 - t)) \Gamma \frac{\partial}{\partial \alpha} \Delta_S(\alpha, t) \leq 0$ and thus $\left[\frac{\partial}{\partial \alpha} \Lambda \Gamma^{-1} \Delta_R(\alpha, t_0) \right]_k \geq 0$.

This is a contradiction.

Therefore $\frac{\partial}{\partial \alpha} \Delta_S(\alpha, t) < 0$ for $t \in [0, T]$, $\alpha \geq 0$. Thus $\phi_T^{(2)}(V(t) + \alpha \Delta V(t)) \leq \phi_T^{(2)}(V(t))$. \square

Theorem B.5 shows the monotonicity of $\phi_T^{(2)}(\cdot)$. Before we proceed to the continuous submodularity part, we need first to show the following lemma, which would be helpful later.

Lemma B.6 *The matrices $A^{(1)}(t)$ and $A^{(2)}(t)$ are continuous time-variant and Metzler such that $A^{(1)}(t) \geq A^{(2)}(t)$ for any $t \geq 0$. Also, $B^{(1)}(t)$ and $B^{(2)}(t)$ are continuous time-variant matrices such that $B^{(1)}(t) \geq B^{(2)}(t) \geq 0$ for any $t \geq 0$. Then for the*

dynamical system $\dot{x}(t) = A^{(1)}(t)x(t) + B^{(1)}(t)u(t)$ and $\dot{y}(t) = A^{(2)}(t)y(t) + B^{(2)}(t)u(t)$ with equal initial condition $x(0) = y(0) > 0$ and nonnegative input $u(t)$, we have $x(t) \geq y(t)$.

Proof: Firstly, $\dot{x}(0) - \dot{y}(0) = (A^{(1)}(0) - A^{(2)}(0))x(0) + (B^{(1)}(0) - B^{(2)}(0))u(0)$. Because $A^{(1)}(t) - A^{(2)}(t) \geq 0$, $B^{(1)}(t) - B^{(2)}(t) \geq 0$, $x(0) > 0$ and $u(t) \geq 0$, we have $\dot{x}(0) - \dot{y}(0) \geq 0$. Therefore in a neighborhood around $t = 0^+$, we have $x(t) \geq y(t)$.

However, if $x(t) \geq y(t)$ does not hold for any $t \geq 0$, in that case there exist some $t_0 \geq 0$ such that $x(t) \geq y(t)$ for $t \in [0, t_0]$ and at time $t = t_0$, $\exists 1 \leq k \leq n$, we have $x_k(t_0) = y_k(t_0)$, $\dot{x}_k(t_0) < \dot{y}_k(t_0)$.

It should be noted that according to Theorem B.4, we have $x(t) \geq 0$ and $y(t) \geq 0$ for $\forall t \geq 0$. Therefore we can see that

$$\begin{aligned} \dot{x}_k(t_0) - \dot{y}_k(t_0) &= \sum_{i=1}^n \left[\left(a_{ki}^{(1)}(t_0)x_i(t_0) - a_{ki}^{(2)}(t_0)y_i(t_0) \right) + \left(b_{ki}^{(1)}(t_0) - b_{ki}^{(2)}(t_0) \right) u_i(t_0) \right] \\ &\geq \sum_{i=1}^n \left(a_{ki}^{(1)}(t_0)x_i(t_0) - a_{ki}^{(2)}(t_0)y_i(t_0) \right) \\ &\geq \sum_{i=1}^n \left(a_{ki}^{(1)}(t_0) - a_{ki}^{(2)}(t_0) \right) y_i(t_0) \geq 0 \end{aligned}$$

This is a contradiction. Therefore $x(t) \geq y(t)$ for $\forall t \geq 0$. \square

Next, we can show the continuous submodularity of $-\phi_T^{(2)}(\cdot)$ with the following lemma:

Lemma B.7 For any $0 \leq V(t) \leq V'(t)$ and $\Delta V(t) \geq 0$, we have $\frac{\partial}{\partial \alpha} \phi_T^{(2)}(V(t) + \alpha \Delta V(t)) \Big|_{\alpha=0} \leq \frac{\partial}{\partial \alpha} \phi_T^{(2)}(V'(t) + \alpha \Delta V(t)) \Big|_{\alpha=0}$.

Proof: In the proof for Theorem B.5, with the input $V(t) + \alpha\Delta V(t)$ we have that

$$\frac{d}{dt} \frac{\partial}{\partial \alpha} \Delta_R(\alpha, t) = -\Gamma \left(\frac{\partial}{\partial \alpha} \Delta_S(\alpha, t) + \frac{\partial}{\partial \alpha} \Delta_R(\alpha, t) \right) \quad (\text{B.5})$$

and

$$\frac{\partial}{\partial \alpha} \Delta_S(\alpha, t) = -\text{diag}[S_\alpha(t)] \left(\Lambda \Gamma^{-1} \frac{\partial}{\partial \alpha} \Delta_R(\alpha, t) + \int_0^t \text{diag}[W]^{-1} \Delta V(\tau) d\tau \right) \quad (\text{B.6})$$

Insert (B.6) into (B.5), we have

$$\begin{aligned} \frac{d}{dt} \frac{\partial}{\partial \alpha} \Delta_R(\alpha, t) = & \Gamma \left((\text{diag}[S_\alpha(t)] \Lambda \Gamma^{-1} - \mathbf{I}_n) \frac{\partial}{\partial \alpha} \Delta_R(\alpha, t) \right. \\ & \left. + \text{diag}[S_\alpha(t)] \int_0^t \text{diag}[W]^{-1} \Delta V(\tau) d\tau \right) \end{aligned}$$

We denote $S'_\alpha(t)$ and $R'_\alpha(t)$ as the solution to the following equations:

$$\begin{aligned} \frac{d S'_\alpha(t)}{dt} &= -\text{diag}[S'_\alpha(t)] (\Lambda I'_\alpha(t) + \text{diag}[W]^{-1} (V'(t) + \alpha \Delta V(t))) \\ \frac{d R'_\alpha(t)}{dt} &= \Gamma I'_\alpha(t) \end{aligned}$$

and $\Delta_{S'}(\alpha, t) = S'_\alpha(t) - S'_\alpha(0)$, $\Delta_{R'}(\alpha, t) = R'_\alpha(t) - R'_\alpha(0)$. Similarly, we have

$$\begin{aligned} \frac{d}{dt} \frac{\partial}{\partial \alpha} \Delta_{R'}(\alpha, t) = & \Gamma \left((\text{diag}[S'_\alpha(t)] \Lambda \Gamma^{-1} - \mathbf{I}_n) \frac{\partial}{\partial \alpha} \Delta_{R'}(\alpha, t) \right. \\ & \left. + \text{diag}[S'_\alpha(t)] \int_0^t \text{diag}[W]^{-1} \Delta V(\tau) d\tau \right) \end{aligned}$$

Notice that both $(\text{diag}[S_\alpha(t)] \Lambda \Gamma^{-1} - \mathbf{I}_n)$ and $(\text{diag}[S'_\alpha(t)] \Lambda \Gamma^{-1} - \mathbf{I}_n)$ are continuous Metzler matrices. By Lemma B.5, we have

$$\text{diag}[S_\alpha(t)] \Lambda \Gamma^{-1} - \mathbf{I}_n \geq \text{diag}[S'_\alpha(t)] \Lambda \Gamma^{-1} - \mathbf{I}_n$$

because $\text{diag}[S_\alpha(t)] \geq \text{diag}[S'_\alpha(t)]$. Using Theorem B.6, we have $\frac{\partial}{\partial \alpha} \Delta_{R'}(\alpha, t) \leq \frac{\partial}{\partial \alpha} \Delta_R(\alpha, t)$. We need to notice that

$$\frac{\partial}{\partial \alpha} \Delta_{S'}(\alpha, t) = -\text{diag}[S'_\alpha(t)] \left(\Lambda \Gamma^{-1} \frac{\partial}{\partial \alpha} \Delta_{R'}(\alpha, t) + \int_0^t \text{diag}[W]^{-1} \Delta V(\tau) d\tau \right) \quad (\text{B.7})$$

Comparing the equation (B.6) and (B.7), since $\text{diag}[S_\alpha(t)] \geq \text{diag}[S'_\alpha(t)] \geq 0$ and $\frac{\partial}{\partial \alpha} \Delta_R(\alpha, t) \geq \frac{\partial}{\partial \alpha} \Delta_{R'}(\alpha, t) \geq 0$, we have $\frac{\partial}{\partial \alpha} \Delta_S(\alpha, T) \leq \frac{\partial}{\partial \alpha} \Delta_{S'}(\alpha, T)$. Proved. \square

Up to now, we have managed to show that the mapping $\phi_T(\cdot)$ which is determined by the SIR model (2.5) is also continuous monotone submodular.

B.4.3 Continuous Monotone Submodularity of $-\phi_T^{(3)}$

Firstly we show the monotonicity of the $-\phi_T^{(3)}(\cdot)$ by the following lemma:

Lemma B.8 *For any $V(\cdot) \geq 0$, $\Delta V(\cdot) \geq 0$, we have $\phi_T^{(3)}(V(t) + \alpha \Delta V(t)) \leq \phi_T^{(3)}(V(t))$ and $\left. \frac{\partial}{\partial \alpha} \phi_T^{(3)}(V(t) + \alpha \Delta V(t)) \right|_{\alpha=0} \leq 0$.*

Proof: For convenience, we define $f(\cdot, \cdot)$ as

$$f(S(t), V(t)) = -\text{diag}[S(t)] (\Lambda(1_n - S(t)) + \text{diag}[W]^{-1}V(t)) + \Gamma(1_n - S(t))$$

and thus $\dot{S}(t) = f(S(t), V(t))$. Denote $f_i(\cdot, \cdot)$ as the i -th component of $f(\cdot, \cdot)$. It's easy to check that

$$\begin{aligned} \frac{\partial}{\partial s_j} f_i(S, V) &= \lambda_{ij} s_i \geq 0 & \forall 1 \leq i, j \leq n \text{ and } i \neq j \\ \frac{\partial}{\partial v_j} f_i(S, V) &= -w_i^{-1} s_i 1_{\{i=j\}} \leq 0 & \forall 1 \leq i, j \leq n \end{aligned}$$

which means that $\dot{S}(t) = f(S(t), V(t))$ is a monotone control system [80, 130].

Also $f(\cdot, \cdot)$ is continuously differentiable, according to the Theorem 1 [80], we have

$$\phi_T^{(3)}(V(\cdot) + \alpha \Delta V(\cdot)) \leq \phi_T^{(3)}(V(\cdot)) \text{ and } \left. \frac{\partial}{\partial \alpha} \phi_T^{(3)}(V(\cdot) + \alpha \Delta V(\cdot)) \right|_{\alpha=0} \leq 0. \quad \square$$

Next, we proceed to show the continuous submodularity of $\phi_T^{(3)}(\cdot)$ with the following lemma. The proof of the continuous submodularity is a little bit more technical than of monotonicity.

Lemma B.9 *For any $0 \leq V(\cdot) \leq V'(\cdot)$ and $\Delta V(\cdot) \geq 0$, we have $\frac{\partial}{\partial \alpha} \phi_T^{(3)}(V(\cdot) + \alpha \Delta V(\cdot)) \Big|_{\alpha=0} \leq \frac{\partial}{\partial \alpha} \phi_T^{(3)}(V'(\cdot) + \alpha \Delta V(\cdot)) \Big|_{\alpha=0}$.*

Proof: Instead of investigating $\frac{\partial}{\partial \alpha} \phi_T^{(3)}(V(\cdot) + \alpha \Delta V(\cdot)) \Big|_{\alpha=0}$, we will consider

$$\frac{\partial}{\partial \alpha} \phi_T^{(3)}(V(\cdot) - \alpha \Delta V(\cdot)) \Big|_{\alpha=0} = - \frac{\partial}{\partial \alpha} \phi_T^{(3)}(V(\cdot) + \alpha \Delta V(\cdot)) \Big|_{\alpha=0}$$

and to show that $\frac{\partial}{\partial \alpha} \phi_T^{(3)}(V(\cdot) - \alpha \Delta V(\cdot)) \Big|_{\alpha=0} \geq \frac{\partial}{\partial \alpha} \phi_T^{(3)}(V'(\cdot) - \alpha \Delta V(\cdot)) \Big|_{\alpha=0}$.

With $V(t) - \alpha \Delta V(t)$ as input, the SIS model (2.6) becomes

$$\frac{d S_\alpha(t)}{dt} = -\text{diag}[S_\alpha(t)] (\Lambda(1_n - S_\alpha(t)) + \text{diag}[W]^{-1} (V(t) - \alpha \Delta V(t))) + \Gamma(1 - S_\alpha(t)) \quad (\text{B.8})$$

Denote $\Delta_S(\alpha, t) = S_\alpha(t) - S_0(t)$ and calculate the partial derivative w.r.t. α for both sides of (B.8), we have

$$\frac{d}{dt} \frac{\partial}{\partial \alpha} \Delta_S(\alpha, t) = A(t) \frac{\partial}{\partial \alpha} \Delta_S(\alpha, t) + B(t) \Delta V(t) \quad (\text{B.9})$$

the $A(t) = -\text{diag}[\Lambda(1_n - S_\alpha(t)) + \text{diag}[W]^{-1} (V(t) - \alpha \Delta V(t))] - \Gamma + \text{diag}[S_\alpha(t)] \Lambda$ and $B(t) = \text{diag}[S_\alpha(t)] \text{diag}[W]^{-1}$.

Similarly, with $S'_\alpha(t)$ as the solution to (2.6) with the input $V'(t) - \alpha \Delta V(t)$, the variation term $\Delta_{S'}(\alpha, t) = S'_\alpha(t) - S'_0(t)$ satisfies the following equation:

$$\frac{d}{dt} \frac{\partial}{\partial \alpha} \Delta_{S'}(\alpha, t) = A'(t) \frac{\partial}{\partial \alpha} \Delta_{S'}(\alpha, t) + B'(t) \Delta V(t) \quad (\text{B.10})$$

$A'(t) = -\text{diag}[\Lambda(1_n - S'_\alpha(t)) + \text{diag}[W]^{-1}(V'(t) - \alpha\Delta V(t))] - \Gamma + \text{diag}[S'_\alpha(t)]\Lambda$
and $B'(t) = \text{diag}[S'_\alpha(t)]\text{diag}[W]^{-1}$.

Here, both $A(t)$ and $A'(t)$ are Metzler matrices. Since $S_\alpha(t) \geq S'_\alpha(t)$ and $V(t) \leq V'(t)$, we have $A(t) \geq A'(t)$ and $B(t) \geq B'(t) \geq 0$ for any $t \geq 0$. Comparing equation (B.9) and (B.10), according to Lemma B.6, since $\Delta_S(\alpha, 0) = \Delta_{S'}(\alpha, 0) = 0$ and $\Delta V(t) \geq 0$, we have $\frac{\partial}{\partial \alpha} \Delta_{S'}(\alpha, t) \leq \frac{\partial}{\partial \alpha} \Delta_S(\alpha, t)$ for $\forall \alpha, t \geq 0$. Therefore $\frac{\partial}{\partial \alpha} \phi_T^{(3)}(V(t) - \alpha\Delta V(t)) \Big|_{\alpha=0} \geq \frac{\partial}{\partial \alpha} \phi_T^{(3)}(V'(t) - \alpha\Delta V(t)) \Big|_{\alpha=0}$, i.e.

$$\frac{\partial}{\partial \alpha} \phi_T^{(3)}(V(t) + \alpha\Delta V(t)) \Big|_{\alpha=0} \leq \frac{\partial}{\partial \alpha} \phi_T^{(3)}(V'(t) + \alpha\Delta V(t)) \Big|_{\alpha=0}$$

The lemma is proved. □

Up to now we have shown that $-\phi_T^{(3)}(\cdot)$ is continuous monotone submodular.

B.5 Proof for Theorem 3.3

Firstly, let's fix l and k . Then we have the following sequence of inequalities for all $\tau < l$.

$$f(x^*(k)) \leq f(x^*(k) \vee x(\tau)) \quad (\text{B.11a})$$

$$= f(\tau) + \int_0^{|x^*(k) \vee x(\tau)| - |x(\tau)|} \frac{\partial}{\partial \alpha} f \left(x(\tau) + \alpha \frac{x^*(k) \vee x(\tau) - x(\tau)}{|x^*(k) \vee x(\tau)| - |x(\tau)|} \right) d\alpha \quad (\text{B.11b})$$

$$\leq f(\tau) + |x^*(k) \vee x(\tau) - x(\tau)| \quad (\text{B.11c})$$

$$\cdot \frac{\partial}{\partial \alpha} f \left(x(\tau) + \alpha \frac{x^*(k) \vee x(\tau) - x(\tau)}{|x^*(k) \vee x(\tau)| - |x(\tau)|} \right) \Big|_{\alpha=0} \quad (\text{B.11d})$$

$$\leq f(\tau) + |x^*(k) \vee x(\tau) - x(\tau)| \cdot \max_{|\Delta x|=1} \frac{\partial}{\partial \alpha} f(x(\tau) + \alpha \Delta x) \Big|_{\alpha=0} \quad (\text{B.11e})$$

$$= f(\tau) + |x^*(k) \vee x(\tau) - x(\tau)| \cdot \frac{d}{ds} f(x(s)) \Big|_{s=\tau} \quad (\text{B.11f})$$

$$\leq f(\tau) + k \frac{d}{ds} f(x(s)) \Big|_{s=\tau} \quad (\text{B.11g})$$

Here, the equation (B.11a) follows from the monotonicity of $f(\cdot)$. The inequality (B.11d) arises from the submodularity of $f(\cdot)$. The equation (B.11f) holds because $x(s)$ is obtained by the continuous greedy algorithm 1 and inequality (B.11g) follows from the fact that $x^*(k) + x(l) \geq x^*(k) \vee x(l)$. Hence

$$f(x^*(k)) - f(x(\tau)) \leq k \cdot \frac{d}{ds} f(x(s)) \Big|_{s=\tau} \quad (\text{B.12})$$

Now let's define $\delta(k, \tau) = f(x^*(k)) - f(x(\tau))$. Thus the equation (B.12) becomes

$$\delta(k, \tau) \leq -k \cdot \frac{d}{d\tau} \delta(k, \tau) \quad (\text{B.13})$$

Since $\delta(k, 0) = f(x^*(k))$, therefore solving (B.13) we have

$$\delta(k, l) \leq \exp(-l/k)\delta(k, 0) = \exp(-l/k)f(x^*(k))$$

Therefore

$$f(x(l)) \geq (1 - \exp(-l/k)) f(x^*(k))$$

In particular, for $k = l$, $f(x(k)) \geq \frac{e-1}{e}f(x^*(k))$. Proved.

B.6 Proof for Theorem 3.3

First, let's define $\Delta V'(t) = \Delta V(t) - \Delta V(t - \tau)$. It is easy to verify that $\int_0^{t_0} \Delta V'(t) dt \geq 0$ for any $t_0 \in [0, T]$. Without loss of generality, we assume that on a neighborhood around time $t = 0^+$, we have $\Delta V'(t) > 0$.

With the input $V(t) + \alpha\Delta V'(t)$, the SIR model (2.5) becomes

$$\begin{aligned} \frac{d S_\alpha(t)}{dt} &= -\text{diag}[S_\alpha(t)] (\Lambda (1_n - S_\alpha(t) - R_\alpha(t)) + \text{diag}[W]^{-1} (V(t) + \alpha\Delta V'(t))) \\ \frac{d R_\alpha(t)}{dt} &= \Gamma (1_n - S_\alpha(t) - R_\alpha(t)) \end{aligned} \tag{B.14}$$

For convenience, let's use the notations that $\Delta_S(\alpha, t) = S_\alpha(t) - S_0(t)$ and $\Delta_R(\alpha, t) = R_\alpha(t) - R_0(t)$. Similar to the proof of Theorem B.5, we calculate the partial derivative w.r.t. α on both sides of the equation (B.14), we have

$$\begin{aligned} \frac{d}{dt} \frac{\partial}{\partial \alpha} \Delta_S(\alpha, t) &= A_{11}(t) \frac{\partial}{\partial \alpha} \Delta_S(\alpha, t) + A_{12}(t) \frac{\partial}{\partial \alpha} \Delta_R(\alpha, t) + B_1(t) \Delta V'(t) \\ \frac{d}{dt} \frac{\partial}{\partial \alpha} \Delta_R(\alpha, t) &= A_{21}(t) \frac{\partial}{\partial \alpha} \Delta_S(\alpha, t) + A_{22}(t) \frac{\partial}{\partial \alpha} \Delta_R(\alpha, t) + B_2(t) \Delta V'(t) \end{aligned} \tag{B.15}$$

where $A_{11}(t) = -\text{diag}[\Lambda (1_n - S_\alpha(t) - R_\alpha(t)) + \text{diag}[W]^{-1} (V(t) + \alpha\Delta V'(t))] + \text{diag}[S_\alpha(t)]\Lambda$, $A_{12}(t) = \text{diag}[S_\alpha(t)]\Lambda$, $A_{21}(t) = A_{22}(t) = -\Gamma$,. For matrix B , $B_1(t) = -\text{diag}[S_\alpha(t)] \text{diag}[W]^{-1}$ and $B_2(t) = 0$.

Looking at the equation (B.15), since $\Delta_S(\alpha, 0) = \Delta_R(\alpha, 0) = 0$ and $B_1(0) < 0$, which means that the term $B_1(t)\Delta V(t) < 0$ dominates the other terms around the neighborhood of $t = 0^+$, therefore $\frac{\partial}{\partial \alpha}\Delta_S(\alpha, t) < 0$ around $t = 0^+$ for any $\alpha \geq 0$. If for some $\alpha > 0$, $\frac{\partial}{\partial \alpha}\Delta_S(\alpha, t)$ is not negative on $t \in [0, T]$, we can find a $t_0 \in [0, T]$ so that a component of $\frac{\partial}{\partial \alpha}\Delta_S(\alpha, t_0)$, denoted as $\frac{\partial}{\partial \alpha}\Delta_S^k(\alpha, t_0)$, equals to zero and for any $0 < t < t_0$, we have $\frac{\partial}{\partial \alpha}\Delta_S(\alpha, t) < 0$.

Let's look back at (B.14), for the first equation, we can divide the j -th row of it by the corresponding $s_j(t)$. Notice that $\frac{d s_j(t)}{s_j(t)} = d \log(s_j(t))$, and $\Gamma^{-1} \frac{d R(t)}{d t} = I(t)$, we have

$$d \log(S(t)) = -\Lambda \Gamma^{-1} d R(t) - \text{diag}[W]^{-1} (V(t) + \alpha \Delta V'(t)) d t$$

Integrate both sides the this equation from $t = 0$ to $t = t_0$, we have

$$\log(S(t_0)) = -\Lambda \Gamma^{-1} R(t_0) - \int_0^{t_0} \text{diag}[W]^{-1} (V(t) + \alpha \Delta V'(t)) d t$$

Calculate the partial derivative w.r.t. α of both sides of the equation, we have

$$\text{diag}[S_\alpha(t_0)]^{-1} \frac{\partial}{\partial \alpha} \Delta_S(\alpha, t_0) = -\Lambda \Gamma^{-1} \frac{\partial}{\partial \alpha} \Delta_R(\alpha, t_0) - \int_0^{t_0} \text{diag}[W]^{-1} \Delta V'(t) d t$$

Since $\frac{\partial}{\partial \alpha} \Delta_S^k(\alpha, t_0) = 0$, $S_\alpha(t_0) > 0$ and $\int_0^{t_0} \Delta V'(t) d t \geq 0$, we have that $\left[\Lambda \Gamma^{-1} \frac{\partial}{\partial \alpha} \Delta_R(\alpha, t_0) \right]_k \leq 0$.

However, if we look at the second equation in (B.15), we have that

$$\left[\Lambda \Gamma^{-1} \frac{\partial}{\partial \alpha} \Delta_R(\alpha, t_0) \right]_k = - \int_0^{t_0} \left[\Lambda \Gamma^{-1} \exp(-\Gamma(t_0 - t)) \Gamma \frac{\partial}{\partial \alpha} \Delta_S(\alpha, t) \right]_k d t$$

As we assume, for $t \in [0, t_0)$, $\frac{\partial}{\partial \alpha} \Delta_S(\alpha, t) < 0$. Since $\exp(-\Gamma(t_0 - t)) \Gamma > 0$, therefore $\exp(-\Gamma(t_0 - t)) \Gamma \frac{\partial}{\partial \alpha} \Delta_S(\alpha, t) < 0$ and thus $\left[\Lambda \Gamma^{-1} \frac{\partial}{\partial \alpha} \Delta_R(\alpha, t_0) \right]_k > 0$.

This is a contradiction.

Therefore $\frac{\partial}{\partial \alpha} \Delta_S(\alpha, t) < 0$ for any $t \in [0, T]$ and $\alpha \geq 0$. Thus easily with $\alpha = 1$ and $t = T$, we have $\frac{\partial}{\partial \alpha} \phi_T^{(2)}(V(t) + \alpha \Delta V(t)) \Big|_{\alpha=0} \leq \frac{\partial}{\partial \alpha} \phi_T^{(2)}(V(t) + \alpha \Delta V(t - \tau)) \Big|_{\alpha=0}$.

B.7 Proof for Theorem 3.4

Before we start to prove Theorem 3.4, we first reformulate the problem (3.9) using the the following lemma:

Lemma B.10 *The general minimum cost problem (3.9) is equivalent to the following problem:*

$$\begin{aligned} \min_{x \in \mathcal{D}} \quad & |x| \\ \text{s.t.} \quad & \frac{\partial}{\partial \alpha} f \left(y + \alpha \frac{x \vee y - y}{|x \vee y - y|} \right) \Big|_{\alpha=0} \cdot |x \vee y - y| \geq \theta - f(y), \quad \forall y \in \Omega \\ & x \in \Omega \end{aligned} \quad (\text{B.16})$$

Proof: The objective functions of (3.9) and (B.16) are the same. Therefore we only need to compare the constraints of them. First, for the feasible x in (3.9), i.e. $f(x) \geq \theta$, we have that

$$\begin{aligned} & \frac{\partial}{\partial \alpha} f \left(y + \alpha \frac{x \vee y - y}{|x \vee y - y|} \right) \Big|_{\alpha=0} \cdot |x \vee y - y| \\ & \geq \int_0^{|x \vee y - y|} \frac{\partial}{\partial \alpha} f \left(y + \alpha \frac{x \vee y - y}{|x \vee y - y|} \right) d\alpha \\ & = f(x \vee y) - f(y) \\ & \geq f(x) - f(y) \\ & \geq \theta - f(y) \end{aligned}$$

therefore such x is also feasible in (B.16).

At the same time, if we let $x = y$ in the inequality of (B.16), then instantly we have $f(x) \geq \theta$, which means the x feasible in (B.16) is also feasible in (3.9). Therefore Theorem B.10 is proved. \square

Denote the solution to the minimum cost problem (3.9) obtained by the Algorithm 2 as $x(\eta)$ where we have $|x(\eta)| = \eta$. Also, denote the optimal solution to (3.9) as x^* and assume that $|x^*| = \zeta$. Then the following optimization problem is easily a relaxation from (B.16).

$$\begin{aligned}
& \min_{x \in \mathcal{D}} |x| \\
& \text{s.t.} \quad \left. \frac{\partial}{\partial \alpha} f \left(x(s) + \alpha \frac{x \vee x(s) - x(s)}{|x \vee x(s) - x(s)|} \right) \right|_{\alpha=0} \cdot |x \vee x(s) - x(s)| \\
& \quad \theta - f(x(s)), \quad \forall s \in [0, \eta] \\
& \quad x \in \Omega
\end{aligned} \tag{B.17}$$

whose optimal solution is denoted as x' with $|x'| = \zeta'$. Easily we have $\zeta' \leq \zeta$. The following lemma is also helpful in showing the Theorem (3.4).

Lemma B.11 *Let $u(\cdot), x(\cdot) \in C^1[0, T]$, with $\dot{u}(t) \geq 0$ and $\dot{x}(t) \leq 0$. If we have $u(0) > 0$ and $x(T) > 0$, then as for $S = u(0)x(0) + \int_0^T \frac{d u(t)}{d t} \cdot x(t) dt = u(T)x(T) - \int_0^T \frac{d x(t)}{d t} \cdot u(t) dt$, we have*

$$S \leq \left(\max_{t \in [0, T]} u(t)x(t) \right) \left[1 + \log \min \left(\frac{x(0)}{x(T)}, \frac{u(T)}{u(0)} \right) \right]$$

Proof: First of all, since $u(t)x(t)$ is differentiable on $[0, T]$, therefore $\max_{t \in [0, T]} u(t)x(t)$ always exists. Then, because $x(t) \leq \frac{\max_{t \in [0, T]} u(t)x(t)}{u(t)}$, we have

$$\begin{aligned}
S & \leq \left(\max_{t \in [0, T]} u(t)x(t) \right) \left[1 + \int_0^T \frac{d u(t)/u(t)}{d t} dt \right] \\
& = \left(\max_{t \in [0, T]} u(t)x(t) \right) \left[1 + \log \frac{u(T)}{u(0)} \right]
\end{aligned}$$

Similarly, taking $u(t) \leq \frac{\max_{t \in [0, T]} u(t)x(t)}{x(t)}$, we have

$$S \leq \left(\max_{t \in [0, T]} u(t)x(t) \right) \left[1 + \log \frac{x(0)}{x(T)} \right]$$

Combining these two inequalities, this lemma is proved. \square

For $s \in [0, \eta]$, we define $\xi(s) = \min_{|\Delta x|=1} \left\{ \frac{1}{\frac{\partial}{\partial \alpha} f(x(s) + \alpha \Delta x)|_{\alpha=0}} \right\}$. Also, we define $\nu(\cdot, \cdot) : [0, \eta] \times \mathcal{M} \rightarrow \mathbb{R}$ as

$$\nu(s, x) = \frac{\partial}{\partial \alpha} f \left(x(s) + \alpha \frac{x \vee x(s) - x(s)}{|x \vee x(s) - x(s)|} \right) \Big|_{\alpha=0} \cdot |x \vee x(s) - x(s)|$$

Because $f(\cdot)$ is continuous monotone submodular and $|x \vee x(s) - x(s)|$ is non-increasing w.r.t. s due to the monotonicity of $x(s)$, we have that $\xi(s)$ is nondecreasing and $\nu(s, x)$ is nonincreasing w.r.t. s . Using Theorem B.11, we have that

$$\begin{aligned} & \int_0^\eta \frac{d\xi(s)}{ds} \nu(s, x) ds + \xi(0)\nu(0, x) \\ & \leq \left(\max_{s \in [0, \eta]} \xi(s)\nu(s, x) \right) \left[1 + \log \min \left(\frac{\xi(\eta)}{\xi(0)}, \frac{\nu(0, x)}{\nu(\eta, x)} \right) \right] \end{aligned}$$

Also, we need to notice that

$$\max_{s \in [0, \eta]} \xi(s)\nu(s, x) \leq \max_{s \in [0, \eta]} |x \vee x(s) - x(s)| \leq |x|$$

With $\mu(s)$ for $s \in [0, T]$ as the dual variable, the Lagrangian of the problem (B.16) is

$$L(x, \mu(\cdot)) = |x| + \int_0^\eta \mu(s) (\theta - f(x(s)) - \nu(s, x)) ds$$

With $k_1 = \frac{\xi(\eta)}{\xi(0)}$ and $k_2 = \max_{x \leq x^*} \frac{\nu(0, x)}{\nu(\eta, x)}$, if we let $\mu^*(s) = \frac{\frac{d\xi(s)}{ds} + \xi(0)\delta(s)}{1 + \log \min \{k_1, k_2\}}$

where $\delta(s)$ is the impulse function, then to achieve the minimum of the Lagrangian $L(x, \mu^*(s))$ w.r.t. x , we need that $x = 0$. By duality theory, because $\mu^*(s)$ is dual feasible, we know that $L(0, \mu^*(s)) \leq \zeta' \leq \zeta$.

On the other hand,

$$\begin{aligned}
\eta &= \int_0^\eta \xi(s) \frac{d f(x(s))}{d s} d s \\
&= \xi(\eta) f(x(\eta)) - \xi(0) f(x(0)) - \int_0^\eta f(x(s)) \frac{d \xi(x(s))}{d s} d s \\
&= \int_0^\eta \left(f(x(\eta)) \frac{d \xi(x(s))}{d s} - f(x(s)) \frac{d \xi(x(s))}{d s} \right) d s + f(x(\eta)) \xi(0) \\
&= \int_0^\eta \frac{d \xi(x(s))}{d s} (\theta - f(x(s))) d s + \theta \xi(0)
\end{aligned}$$

Therefore we have

$$\eta = [1 + \log \min \{k_1, k_2\}] L(0, \mu^*(s)) \leq [1 + \log \min \{k_1, k_2\}] \zeta$$

B.8 Proof for Theorem 3.5

First, for any $x \in \Omega$ such that $x \leq \tilde{x}$, we have

$$\begin{aligned}
f(x) - f(\tilde{x}) &= \int_0^{|\tilde{x}-x|} \frac{\partial}{\partial \alpha} f \left(\tilde{x} - \alpha \frac{\tilde{x} - x}{|\tilde{x} - x|} \right) d \alpha \\
&\leq |\tilde{x} - x| \cdot \frac{\epsilon}{|\omega|} f(\tilde{x}) \\
&\leq \epsilon \cdot f(\tilde{x})
\end{aligned}$$

therefore $(1 + \epsilon)f(\tilde{x}) \geq f(x)$. Similarly, for any $x \in \Omega$ such that $x \geq \tilde{x}$, we also have

$$(1 + \epsilon)f(\tilde{x}) \geq f(x).$$

With the previous analysis, it easily follows that $(1 + \epsilon)f(\tilde{x}) \geq f(\tilde{x} \vee x^*)$ and

$(1 + \epsilon)f(\tilde{x}) \geq f(\tilde{x} \wedge x^*)$. Therefore we have

$$2(1 + \epsilon)f(\tilde{x}) + f(\omega - \tilde{x}) \geq f(\tilde{x} \wedge x^*) + f(\tilde{x} \vee x^*) + f(\omega - \tilde{x}) \quad (\text{B.18a})$$

$$\geq f(\tilde{x} \wedge x^*) + f(x^* - \tilde{x} \wedge x^*) + f(\omega) \quad (\text{B.18b})$$

$$\geq f(\tilde{x} \wedge x^*) + f(x^* - \tilde{x} \wedge x^*) \quad (\text{B.18c})$$

$$\geq f(x^*) + f(0) = f(x^*) \quad (\text{B.18d})$$

where (B.18a) holds because of our previous analysis, (B.18b) and (B.18d) follows from the continuous submodularity of $f(\cdot)$ and (B.18c) is due to that $f(\omega) \geq 0$. Therefore easily we have either $f(\tilde{x}) \geq \frac{1}{3}(1 - \epsilon)f(x^*)$ or $f(\omega - \tilde{x}) \geq \frac{1}{3}(1 - \epsilon)f(x^*)$.

B.9 Proof for Theorem 3.6

The following lemma is modified from [59], which is very helpful showing the convexity of the $\phi_T^{(3)}(\cdot)$.

Lemma B.12 *If $f(\cdot, \cdot) : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n \in C^1$ is convex. Also, $\frac{\partial}{\partial x_j} f_i(x, u) \geq 0$ and $\frac{\partial}{\partial u_k} f_i(x, u) \leq 0$ for all i, j and k with $i \neq j$, where $f_i(\cdot, \cdot)$ denotes the i -th component of $f(\cdot, \cdot)$. Then for the dynamical system $\dot{x}(t) = f(x(t), u(t))$ with the fixed initial condition $x(0) = x_0$, then each component of the solution $x(t) = \phi(t, u(\cdot))$ is convex w.r.t. $u(\cdot)$.*

Proof: First of all, let $g(x, u, v) = f(x, u) + v$. It's easy to verify that for any i and j we have $\frac{\partial}{\partial v_j} f_i(x, u, v) \geq 0$. For dynamical system $\dot{x}(t) = g(x(t), u(t), v(t))$ and $\dot{x}'(t) = g(x'(t), u'(t), v'(t))$, define the convex cone $K = \mathbb{R}_-^n \times \mathbb{R}_+^n$. By the

Theorem 1 in [80], we know that if $x(0) = x'(0)$ and $(u(t), v(t)) \preceq_K (u'(t), v'(t))$, then $x(t) \leq x'(t)$ for $\forall t \geq 0$.

Let $x_1(t) = \phi(t, u_1(\cdot))$, $x_2(t) = \phi(t, u_2(\cdot))$, $x_\lambda(t) = \phi(t, \lambda u_1(\cdot) + (1 - \lambda)u_2(\cdot))$ with $0 \leq \lambda \leq 1$. Then we have

$$\dot{x}_\lambda(t) = f(x_\lambda(t), \lambda u_1(t) + (1 - \lambda)u_2(t))$$

Also, it will be easy to verify that

$$\lambda \dot{x}_1(t) + (1 - \lambda)\dot{x}_2(t) = f(\lambda x_1(t) + (1 - \lambda)x_2(t), \lambda u_1(t) + (1 - \lambda)u_2(t)) + v(t)$$

where $v(t) = \lambda f(x_1(t), u_1(t)) + (1 - \lambda)f(x_2(t), u_2(t)) - f(\lambda x_1(t) + (1 - \lambda)x_2(t), \lambda u_1(t) + (1 - \lambda)u_2(t))$. Since f is convex, therefore $v(t) \geq 0$ for any $t \geq 0$. Then as we discuss before, because $(\lambda u_1 + (1 - \lambda)u_2, 0) \preceq_K (\lambda u_1 + (1 - \lambda)u_2, v)$ we have $x_\lambda(t) \leq \lambda x_1(t) + (1 - \lambda)x_2(t)$ for any $t \geq 0$. \square

Theorem B.12 provides a powerful tool to show the convexity of the dynamical system. Next, we are going to reformulate the SIS model (2.6) so as to use Theorem B.12 to show its convexity. The SIS epidemic dynamical model can be formulated as follows:

$$\frac{d}{dt}S(t) = -\text{diag}[S(t)](\Lambda(1_n - S(t)) + \text{diag}[W]^{-1}V(t)) + \Gamma(1 - S(t)) \quad (\text{B.19})$$

Divide the j -th component of both sides of (B.19) by $s_j(t)$, we have

$$\frac{d}{dt}\log(S(t)) = -(\Lambda(1_n - S(t)) + \text{diag}[W]^{-1}V(t)) + \Gamma(\text{diag}[S(t)]^{-1} - 1_n) \quad (\text{B.20})$$

With the variable transform $Z(t) = \log(S(t))$, then (B.20) becomes

$$\frac{d}{dt}Z(t) = -(\Lambda(1_n - \exp(Z(t))) + \text{diag}[W]^{-1}V(t)) + \Gamma(\exp(-Z(t)) - 1_n) \quad (\text{B.21})$$

Using Theorem B.12, we know that for any $V_1(\cdot), V_2(\cdot) \in \mathcal{D}$ and $0 \leq \lambda \leq 1$ with $V_\lambda(\cdot) = \lambda V_1(\cdot) + (1 - \lambda)V_2(\cdot)$, we have that

$$\phi_T^{(3)}(V_\lambda(\cdot)) = \exp\left(\log(\phi_T^{(3)}(V_\lambda(\cdot)))\right) \quad (\text{B.22a})$$

$$\leq \exp\left(\lambda \log(\phi_T^{(3)}(V_1(\cdot))) + (1 - \lambda) \log(\phi_T^{(3)}(V_2(\cdot)))\right) \quad (\text{B.22b})$$

$$\leq \lambda \exp\left(\log(\phi_T^{(3)}(V_1(\cdot)))\right) + (1 - \lambda) \exp\left(\log(\phi_T^{(3)}(V_2(\cdot)))\right) \quad (\text{B.22c})$$

$$= \lambda \phi_T^{(3)}(V_1(\cdot)) + (1 - \lambda) \phi_T^{(3)}(V_2(\cdot)) \quad (\text{B.22d})$$

where the inequality (B.22b) holds because of the convexity of (B.21) and inequality (B.22c) holds because the exponential function is convex. The theorem is thus proved.

B.10 Proof for Theorem 4.1

With the transform that $I(t) = \sum_{k=1}^m I_k(t)$, $R(t) = \sum_{k=1}^m R_k(t)$ and $U(t) = \sum_{k=1}^m C_k U_k(t)$, the DCG competition model (4.1) is converted as follows:

$$\begin{aligned} \frac{dS(t)}{dt} &= -\text{diag}[S(t)](\Lambda I(t) + U(t)) \\ \frac{dI(t)}{dt} &= \text{diag}[S(t)](\Lambda I(t) + U(t)) - \Gamma I(t) \\ \frac{dR(t)}{dt} &= \Gamma I(t) \end{aligned} \quad (\text{B.23})$$

which is actually the community-based SIR model (2.5) in Section 2.1. Then by invoking Theorem 3.2 in Section 3.2, this theorem becomes an instant corollary.

B.11 Proof for Theorem 4.2

It should be noted that $S(t) = 1_n - \text{diag}[W]^{-1} \eta^{(1)}(t, \mathbb{U}(\cdot))$. Without loss of generality, for any $j \neq k$, let's typically investigate the following two equations from (4.1), which is shown as follows:

$$\begin{aligned} \frac{dI_j(t)}{dt} &= \text{diag}[1_n - \text{diag}[W]^{-1} \eta^{(1)}(t, \mathbb{U}(\cdot))] (\Lambda I_j(t) + C_j U_j(t)) - \Gamma I_j(t) \\ \frac{dR_j(t)}{dt} &= \Gamma I_j(t) \end{aligned} \quad (\text{B.24})$$

For convenience let's write the equation (B.24) in the compact form as $\dot{I}_j = f_1(I_j, R_j, U_j, \eta^{(1)})$ and $\dot{R}_j = f_2(I_j, R_j, U_j, \eta^{(1)})$, where we implicitly views $\eta^{(1)}(\cdot)$ as a kind of input. Denote the p -th entry of $f_{(\cdot)}(\cdot)$ as $f_{(\cdot)}^p(\cdot)$ while the q -th entry of $I_j(t)$, $R_j(t)$, $U_j(t)$ and $\eta^{(1)}(\cdot)$ as $i_j^q(t)$, $r_j^q(t)$, $u_j^q(t)$ and $\eta_q^{(1)}(\cdot)$ respectively, it will be easy to verify that for any $p \neq q$, we have

$$\frac{\partial f_1^p}{\partial i_j^q} = \left(1 - \frac{\eta^{(1)}(t, \mathbb{U}(\cdot))}{w_p}\right) \lambda_{pq} \geq 0 \quad \frac{\partial f_2^p}{\partial i_j^q} = 0 \quad \frac{\partial f_1^p}{\partial r_j^q} = 0 \quad \frac{\partial f_2^p}{\partial i_j^q} = 0$$

Also, for any p and q , we have

$$\begin{aligned} \frac{\partial f_1^p}{\partial \eta_q^{(1)}} &= -\frac{\sum_{l=1}^n \lambda_{pl} i_l}{w_p} \cdot \mathbf{1}_{\{p=q\}} \leq 0 & \frac{\partial f_2^p}{\partial \eta_q^{(1)}} &= 0 \\ \frac{\partial f_1^p}{\partial u_j^q} &= \left(1 - \frac{\eta^{(1)}(t, \mathbb{U}(\cdot))}{w_p}\right) c_j^q \cdot \mathbf{1}_{\{p=q\}} \geq 0 & \frac{\partial f_2^p}{\partial u_j^q} &= 0 \end{aligned}$$

By Theorem 1 in [80], we have that (B.24) is a monotone control system. In particular, if $\eta_1^{(1)}(\cdot)$ is increased by any means, we have that $I_j(t) + R_j(t)$ is decreased for any time $t > 0$.

Therefore if the k -th player has his strategy increased from $U_k(\cdot)$ to $U_k'(\cdot)$, then according to Theorem 4.1, we have that the social welfare $\eta^{(1)}(\mathbb{U}(\cdot))$ is increased

to $\eta^{(1)}(\mathbb{U}(\cdot) \oplus U'_k(\cdot))$. According to the analysis above, we got that $I_j(t) + R_j(t)$ is decreased for any time $t > 0$, which proves the competitiveness of the DCG marketing game.

B.12 Proof for Theorem 4.3

For simplicity of expression, we would first write the voter competitive model in the compact form as $\dot{X}_k(t) = f_k(X_k(t), \mathbb{U}(\cdot))$. Denote the i -th entry of $X_k(\cdot)$, $f(\cdot)$ and $U_k(\cdot)$ as $x_k^i(\cdot)$, $f_k^i(\cdot)$ and u_k^i respectively,. It is easy to verify the following inequalities:

$$\frac{\partial f_k^i}{\partial x_k^j} = \lambda_{ij} > 0$$

where $i \neq j$ and

$$\frac{\partial f_k^i}{\partial x_l^j} = 0 \quad \frac{\partial f_k^i}{\partial u_k^j} = (1 - x_k^i) c_k^j \cdot \mathbf{1}_{\{i=j\}} \geq 0 \quad \frac{\partial f_k^i}{\partial u_l^j} = -x_k^i c_l^j \cdot \mathbf{1}_{\{i=j\}} \leq 0$$

with any $1 \leq i, j \leq m$ and $k \neq l$. Invoking the Theorem 1 in [80], the statements in this theorem is proved.

B.13 Proof for Theorem 4.4

For the monotonicity part, consider the dynamics of $X_0(\cdot)$ in the FMCG competition model (4.2). Notice that $U_0(t) \equiv 0$, then according to Theorem 4.3, the competitiveness of FMCG diffusion ensures that any increase in $U_k(\cdot)$ will not cause an increase in the value of $\zeta_0^{(2)}(\cdot)$ for $1 \leq k \leq m$. Since $\eta^{(2)}(\cdot) = \mathbf{1}_n^T \cdot W^T - \zeta_0^{(2)}(\cdot)$, then $\eta(\mathbb{U}(\cdot))$ is monotone w.r.t. $\mathbb{U}(\cdot)$.

Then let's sum up the fraction of agents adopting one of the m products and denote it as $X(\cdot) = \sum_{k=1}^m X_k(\cdot)$. Also, we sum up the investment from all the players as $U(\cdot) = \sum_{k=1}^m U_k(\cdot)$. Then the voter competitive model (4.2) is transformed into (B.25).

$$\frac{d}{dt}X(t) = (\Lambda - \text{diag}[\Lambda 1_n])X(t) + \text{diag}[1_n - X(t)]U(t) \quad (\text{B.25})$$

Then, replace $U(\cdot)$ with $U(\cdot) + \alpha\Delta U(\cdot)$ and (B.25) becomes (B.26).

$$\frac{d}{dt}X_\alpha(t) = (\Lambda - \text{diag}[\Lambda \cdot 1_n])X_\alpha(t) + \text{diag}[1_n - X_\alpha(t)](U(t) + \alpha\Delta U(t)) \quad (\text{B.26})$$

Let's denote $\Delta_X(\alpha, t) = X_\alpha(t) - X(t)$ and calculate the partial derivative w.r.t. α for both sides of (B.26), then we have

$$\frac{d}{dt}\frac{\partial}{\partial\alpha}\Delta_X(\alpha, t) = A(t)\frac{\partial}{\partial\alpha}\Delta_X(\alpha, t) + B(t)\Delta U(t) \quad (\text{B.27})$$

where $A(t) = \Lambda - \text{diag}[\Lambda \cdot 1_n - (U(t) + \alpha\Delta U(t))]$ and $B(t) = \text{diag}[1_n - X_\alpha(t)]$.

Similarly, with $X'_\alpha(t)$ as the solution to (B.25) where $U(\cdot)$ is replaced with $U'(\cdot) + \alpha\Delta U(\cdot)$, then the variation term $\Delta_{X'}(\alpha, t) = X'_\alpha(t) - X'_0(t)$ would satisfy the following equation:

$$\frac{d}{dt}\frac{\partial}{\partial\alpha}\Delta_{X'}(\alpha, t) = A'(t)\frac{\partial}{\partial\alpha}\Delta_{X'}(\alpha, t) + B'(t)\Delta U(t) \quad (\text{B.28})$$

where $A(t) = \Lambda - \text{diag}[\Lambda \cdot 1_n - (U'(t) + \alpha\Delta U(t))]$ and $B(t) = \text{diag}[1_n - X'_\alpha(t)]$.

If $U(\cdot) \leq U'(\cdot)$, then we've already shown that $X_\alpha(t) \leq X'_\alpha(t)$. Notice that both $A(t)$ and $A'(t)$ are Metzler matrices. Also, $A(t) \geq A'(t)$ and $B(t) \geq B'(t)$ for any $t \geq 0$. Then according to Lemma B.6, we have $\frac{\partial}{\partial\alpha}\Delta_{X'}(\alpha, t) \leq \frac{\partial}{\partial\alpha}\Delta_X(\alpha, t)$ for any α and $t \geq 0$, which follows that

$$\frac{\partial}{\partial\alpha}\eta^{(2)}(U(\cdot) + \alpha\Delta U(\cdot)) \geq \frac{\partial}{\partial\alpha}\eta^{(2)}(U'(\cdot) + \alpha\Delta U(\cdot))$$

with $\mathbb{U}'(\cdot) \geq \mathbb{U}(\cdot) \geq 0$ and $\Delta\mathbb{U}(\cdot) \geq 0$. This theorem is proved.

B.14 Proof for Theorem 4.5

For the FMCG competition model (4.2), divide the j -th entry of both sides by $x_i^j(t)$, we have

$$\frac{d \log X_k(t)}{d t} = \text{diag} [X_k(t)]^{-1} \Lambda X_k(t) + \text{diag} [X_k(t)]^{-1} C_k U_k(t) - \sum_{i=1}^m C_i U_i(t) - \Lambda \cdot \mathbf{1}_n \quad (\text{B.29})$$

Replacing $\log X_i(t)$ with $Z_i(t)$, the equation (B.29) becomes

$$\begin{aligned} \frac{d Z_k(t)}{d t} &= \text{diag} [\exp (-Z_k(t))] \Lambda \exp (Z_k(t)) + \text{diag} [\exp (-Z_k(t))] C_k U_k(t) \\ &\quad - \sum_{i=1}^m C_i U_i(t) - \Lambda \cdot \mathbf{1}_n \end{aligned} \quad (\text{B.30})$$

For the equation (B.30), notice that with $U_k(\cdot)$ fixed, the time derivative $\dot{Z}_k(t)$ is convex w.r.t. $Z_k(t)$ and $\mathbb{U}_{-k}(t)$. It is also easy to verify that the dynamical system in (B.30) is monotone w.r.t. $Z_k(\cdot)$ and $\mathbb{U}_{-k}(\cdot)$. Therefore according to Theorem B.12, $\log \left(\phi_k^{(2)}(\mathbb{U}(\cdot)) \right)$ is convex w.r.t. $\mathbb{U}_{-k}(\cdot)$.

Then for given strategy profile $\mathbb{U}(\cdot)$ and $\mathbb{U}'(\cdot)$ with $\mathbb{U}_\alpha = \alpha\mathbb{U}(\cdot) + (1 - \alpha)\mathbb{U}'(\cdot)$, we have

$$\phi_k^{(2)}(t, \mathbb{U}_\alpha) = \exp \left(\log \left(\phi_k^{(2)}(t, \mathbb{U}_\alpha) \right) \right) \quad (\text{B.31a})$$

$$\leq \exp \left(\alpha \log \left(\phi_k^{(2)}(t, \mathbb{U}(\cdot)) \right) + (1 - \alpha) \log \left(\phi_k^{(2)}(t, \mathbb{U}'(\cdot)) \right) \right) \quad (\text{B.31b})$$

$$\leq \alpha \exp \left(\log \left(\phi_k^{(2)}(t, \mathbb{U}(\cdot)) \right) \right) + (1 - \alpha) \exp \left(\log \left(\phi_k^{(2)}(t, \mathbb{U}'(\cdot)) \right) \right) \quad (\text{B.31c})$$

$$= \alpha \phi_k^{(2)}(t, \mathbb{U}(\cdot)) + (1 - \alpha) \phi_k^{(2)}(t, \mathbb{U}'(\cdot)) \quad (\text{B.31d})$$

for $\alpha \in [0, 1]$. Here the inequality (B.31b) holds because of the convexity of (B.30) while the inequality (B.31c) holds because $\exp(\cdot)$ is also convex. Therefore $\phi_k^{(2)}(t, \mathbb{U}(\cdot))$ is convex w.r.t. $\mathbb{U}_{-k}(\cdot)$, and so is the payoff function $\zeta_k^{(2)}(t, \mathbb{U}(\cdot))$.

To show that $\zeta_k^{(2)}(\cdot)$ is concave w.r.t. $U_k(\cdot)$ needs a little trick. With the transform $Y_i(t) = 1_n - X_i(t)$, the voter competitive model (4.2) becomes as follows:

$$\begin{aligned} \frac{dY_i(t)}{dt} &= (\Lambda - \text{diag}[\Lambda \cdot 1_n]) Y_i(t) \\ &\quad - \text{diag}[Y_i(t)] C_i U_i(t) + \text{diag}[1_n - Y_i(t)] \sum_{k \neq i} C_k U_k(t) \end{aligned}$$

Following the same derivations as the first half of this proof, we can show that $1_n - \phi_k^{(2)}(t, \mathbb{U}(\cdot))$ is convex w.r.t. $U_k(\cdot)$, i.e. $\phi_k^{(2)}(t, \mathbb{U}(\cdot))$ is concave w.r.t. $U_k(\cdot)$ and so is the payoff function $\zeta_k^{(2)}(\mathbb{U}(\cdot))$. The theorem is proved.

B.15 Proof for Theorem 4.6

The proof is divided into two parts for the DCG marketing game and the FMCG marketing game respectively.

B.15.1 The DCG Marketing Game

In 1976 Kononenko [131] studies the existence of ϵ -equilibrium for general differential games. His conclusion is summarized in Theorem B.13.

Lemma B.13 *For a general differential game by m players with the dynamics*

$$\dot{X}(t) = f(t, X(t), U_1(t), \dots, U_m(t))$$

with X_T denoting the terminal state. In this game, the player i plays with the control $U_i(\cdot)$ which takes value in the set \mathcal{U}^i . He aims at maximizing his payoff $g_i(X_T)$. We assume the following conditions hold:

1. The sets \mathcal{U}^k with $k = 1, \dots, m$ are compact subsets of some finite dimensional spaces.
2. $f : [0, T] \times \mathbb{R}^n \times \mathcal{U}^1 \times \dots \times \mathcal{U}^m \rightarrow \mathbb{R}^n$ is continuous and bounded, and globally Lipschitz continuous w.r.t. X .
3. The maps $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$ are Lipschitz continuous and bounded for $i = 1, \dots, m$.

Let us set

$$\mathcal{U}^{-i} := \mathcal{U}^1 \times \dots \times \mathcal{U}^{i-1} \times \mathcal{U}^{i+1} \times \dots \times \mathcal{U}^m$$

Then we also assume that the following famous Isaacs condition holds.

$$\inf_{U_{-i} \in \mathcal{U}^{-i}} \sup_{U_i \in \mathcal{U}^i} \langle f(t, x, U_{-i} \oplus U_i), p \rangle = \sup_{U_i \in \mathcal{U}^i} \inf_{U_{-i} \in \mathcal{U}^{-i}} \langle f(t, x, U_{-i} \oplus U_i), p \rangle$$

with $(t, x, p) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^n$. The $\langle \cdot, \cdot \rangle$ is the inner product. Then with any initial condition, there exists an ϵ -equilibrium for the differential game for $\forall \epsilon > 0$.

Proof: See [131] or Chapter 4 in [132]. □

Then we write the DCG competition model (4.1) in a compact form as follows:

$$\dot{Y}(t) = f(t, Y(t), U_1(t), \dots, U_m(t)) \tag{B.32}$$

where $Y(t) = [S^T(t), I_1^T(t), \dots, I_m^T(t), R_1^T(t), \dots, R_m^T(t)]^T$ and $f : [0, T] \times \mathbb{R}^{(2m+1)n} \times \mathbb{R}^{mn} \rightarrow \mathbb{R}^{2m+1}$ is the dynamics of the DCG competition model (4.1). Comparing

with the conditions in Theorem B.13, the primary difference in our context is that the feasible field \mathcal{U}^k where $U_k(t)$ takes value may vary at different time because the budget $B_k(t)$ for each player is assumed to be time-variable. To overcome this issue, we define $\psi : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$ as $\psi(U, B) = U \cdot \min \left\{ \frac{B}{1^T \cdot U}, 1 \right\}$.

Then, we define $f'(\cdot)$ as

$$f'(t, Y(t), U_1(t), \dots, U_m(t)) = f(t, Y(t), \psi(U_1(t), B_1(t)), \dots, \psi(U_m(t), B_m(t)))$$

It is easy to verify that for each $U_k(t)$ such that $1^T \cdot U_k(t) \leq B_k(t)$ where $k = 1, \dots, m$, we have $f(t, Y(t), U_1(t), \dots, U_m(t)) = f'(t, Y(t), U_1(t), \dots, U_m(t))$. Therefore the DCG competition model (4.1) is recast into

$$\dot{Y}(t) = f'(t, Y(t), U_1(t), \dots, U_m(t))$$

Also, we set $\mathcal{U}^k = \left\{ U \in \mathbb{R}^n \mid U \geq 0, 1^T \cdot U \leq \max_{0 \leq t \leq T} B_k(t) \right\}$. Since $B_k(\cdot)$ is continuous, thus each \mathcal{U}^k is well defined.

It is easy to verify that each set \mathcal{U}^k is compact and of finite dimensions. The dynamics $f'(\cdot)$ is continuous and bounded, which is also globally Lipschitz w.r.t. $Y(t)$. (The $f'(\cdot)$ being continuous w.r.t. t is because $B_k(t)$ is continuous.) The payoff functions are linear w.r.t. the terminal state and thus Lipschitz and bounded. Also, since $f'(\cdot)$ is linear w.r.t. each $U_k(t)$, the Isaacs condition naturally holds.

Then according to Theorem B.13, the DCG marketing part of the theorem holds.

B.15.2 The FMCG Marketing Game

For convenience, let's define the mapping of *best response* with the following definition:

Definition B.1 *In the marketing game, the best response $\text{Br}_k(\cdot)$ for the player k with a given strategy profile $\mathbb{U}(\cdot)$ is*

$$\text{Br}_k(\mathbb{U}(\cdot)) = \arg \max_{U'_k(\cdot) \in \mathcal{D}_k} (\zeta^{(\cdot)}(\mathbb{U}(\cdot) \oplus U'_k(\cdot)))$$

In 2014 Seierstad [133] studies the existence of Nash equilibrium for a special class of nonlinear differential games. His main result, the Theorem 1 in [133], is restated in Theorem B.14.

Lemma B.14 *For a general differential game by m players with the dynamics*

$$\dot{X}(t) = f(t, X(t), U_1(t), \dots, U_m(t))$$

with X_T denoting the terminal state, each player k plays with the strategy $U_k(\cdot)$ which takes value in some compact set \mathcal{U}^k and he aims at maximizing his payoff $g_k(X_T)$. Then if his best response $\text{Br}_k(\mathbb{U}(\cdot))$ is nonempty and convex with any strategy profile $\mathbb{U}(\cdot)$, then the Nash equilibrium exists for this game.

Proof: See [133]. □

By Theorem 4.5, $\zeta_k^{(2)}(\mathbb{U}(\cdot))$ is concave w.r.t. $U_k(\cdot)$, therefore for any $U_k^1(\cdot)$ and $U_k^2(\cdot)$ that maximizes the utility function $\zeta_k^{(2)}(\mathbb{U}(\cdot) \oplus U_k(\cdot))$, it follows that $\alpha U_k^1(\cdot) + (1 - \alpha) U_k^2(\cdot)$ is also the maximal with $\alpha \in [0, 1]$. Therefore the $\text{Br}_k(\mathbb{U}(\cdot))$

is convex. According to Theorem B.14, the existence of Nash equilibrium for the FMCG marketing game is guaranteed.

B.16 Proof for Theorem 4.7

Unless specified otherwise, we will analyze both the DCG and the FMCG marketing game in the unified framework, where $\eta^{(\cdot)}(\cdot)$ and $\zeta_k^{(\cdot)}(\cdot)$ refer to the social welfare and the payoff for the player k in both games.

For simplicity of expression, we use notion $\mathbb{U}_k^*(\cdot) = [U_1^*(\cdot), \dots, U_k^*(\cdot), 0, \dots, 0]$ and $\mathbb{U}_k(\cdot) = [U_1(\cdot), \dots, U_k(\cdot), 0, \dots, 0]$. Also, we define the operator \oplus by requiring that

$$\begin{aligned}\mathbb{U}^*(\cdot) \oplus \emptyset_k &= [U_1^*(\cdot), \dots, U_{k-1}^*(\cdot), 0, U_{k+1}^*(\cdot), \dots, U_m^*(\cdot)] \\ \mathbb{U}(\cdot) \oplus \emptyset_k &= [U_1(\cdot), \dots, U_{k-1}(\cdot), 0, U_{k+1}(\cdot), \dots, U_m(\cdot)]\end{aligned}$$

First of all, we will use the following lemma to show a very important property of the social welfare function $\eta^{(\cdot)}(\cdot)$ based on its continuous monotone submodularity.

Lemma B.15 *For the social welfare function $\eta^{(\cdot)}(\cdot)$, with any strategy profile $\mathbb{U}(\cdot)$ and $\mathbb{U}'(\cdot)$ s.t. $\mathbb{U}(\cdot) \leq \mathbb{U}'(\cdot)$ as well as any variation $\Delta\mathbb{U}(\cdot) \geq 0$ we have*

$$\eta^{(\cdot)}(\mathbb{U}'(\cdot) + \Delta\mathbb{U}(\cdot)) - \eta^{(\cdot)}(\mathbb{U}'(\cdot)) \leq \eta^{(\cdot)}(\mathbb{U}(\cdot) + \Delta\mathbb{U}(\cdot)) - \eta^{(\cdot)}(\mathbb{U}(\cdot))$$

Proof: According to the Theorem 3.2.6. in [116], $\eta^{(\cdot)}(\cdot)$ is Fréchet differentiable, therefore

$$\begin{aligned}\eta^{(\cdot)}(\mathbb{U}'(\cdot) + \Delta\mathbb{U}(\cdot)) - \eta^{(\cdot)}(\mathbb{U}'(\cdot)) &= \int_0^1 \frac{\partial}{\partial\alpha} \eta^{(\cdot)}(\mathbb{U}'(\cdot) + \alpha\Delta\mathbb{U}(\cdot)) \, d\alpha \\ &\leq \int_0^1 \frac{\partial}{\partial\alpha} \eta^{(\cdot)}(\mathbb{U}(\cdot) + \alpha\Delta\mathbb{U}(\cdot)) \, d\alpha \\ &= \eta^{(\cdot)}(\mathbb{U}(\cdot) + \Delta\mathbb{U}(\cdot)) - \eta^{(\cdot)}(\mathbb{U}(\cdot))\end{aligned}$$

Here the inequality arises from that $\eta^{(\cdot)}(\cdot)$ is continuous monotone submodular. \square

The following corollary is instantly available from Theorem B.15.

Corollary B.2 *For any strategy profile $\mathbb{U}(\cdot)$ and $\mathbb{U}'(\cdot)$, we have*

$$\eta^{(\cdot)}(\mathbb{U}'(\cdot) + \mathbb{U}_k(\cdot)) \leq \sum_{k=1}^m [\eta^{(\cdot)}(\mathbb{U}'(\cdot) + \mathbb{U}_k(\cdot)) - \eta^{(\cdot)}(\mathbb{U}'(\cdot) + \mathbb{U}_{k-1}(\cdot))]$$

Proof: Notice that $\eta^{(\cdot)}(\mathbb{U}'(\cdot) + \mathbb{U}(\cdot)) - \eta^{(\cdot)}(\mathbb{U}(\cdot)) \leq \eta^{(\cdot)}(\mathbb{U}'(\cdot))$, this corollary is proved. \square

Next, we'd proceed to the following lemma, which is crucial to the analysis of the PoA.

Lemma B.16 *For any strategy profile $\mathbb{U}(\cdot) \in \mathcal{D}$, we have*

$$\begin{aligned}\eta^{(\cdot)}(\mathbb{U}^*(\cdot)) &\leq \eta^{(\cdot)}(\mathbb{U}(\cdot)) + \sum_{k=1}^m [\eta^{(\cdot)}(\mathbb{U}(\cdot) \oplus U_k^*(\cdot)) - \eta^{(\cdot)}(\mathbb{U}(\cdot) \oplus \emptyset_k)] \\ &\quad - (\eta^{(\cdot)}(\mathbb{U}^*(\cdot) + \mathbb{U}(\cdot)) - \eta^{(\cdot)}(\mathbb{U}^*(\cdot)))\end{aligned}$$

Proof: First of all, according to Theorem B.15 and Theorem B.2, we have

$$\begin{aligned}
\eta^{(\cdot)}(\mathbb{U}^*(\cdot) + \mathbb{U}(\cdot)) &\leq \eta^{(\cdot)}(\mathbb{U}(\cdot)) + \sum_{k=1}^m [\eta^{(\cdot)}(\mathbb{U}(\cdot) + \mathbb{U}_k^*(\cdot)) - \eta^{(\cdot)}(\mathbb{U}(\cdot) + \mathbb{U}_{k-1}^*(\cdot))] \\
&\leq \eta^{(\cdot)}(\mathbb{U}(\cdot)) + \sum_{k=1}^m [\eta^{(\cdot)}(\mathbb{U}(\cdot) \oplus U_k^*(\cdot)) - \eta^{(\cdot)}(\mathbb{U}(\cdot) \oplus \emptyset_k)]
\end{aligned} \tag{B.33}$$

On the other side, we have

$$\eta^{(\cdot)}(\mathbb{U}^*(\cdot) + \mathbb{U}(\cdot)) = \eta^{(\cdot)}(\mathbb{U}^*(\cdot)) + [\eta^{(\cdot)}(\mathbb{U}^*(\cdot) + \mathbb{U}(\cdot)) - \eta^{(\cdot)}(\mathbb{U}^*(\cdot))] \tag{B.34}$$

Combining (B.33) and (B.34), we have

$$\begin{aligned}
\eta^{(\cdot)}(\mathbb{U}^*(\cdot)) &\leq \eta^{(\cdot)}(\mathbb{U}(\cdot)) + \sum_{k=1}^m [\eta^{(\cdot)}(\mathbb{U}(\cdot) \oplus U_k^*(\cdot)) - \eta^{(\cdot)}(\mathbb{U}(\cdot) \oplus \emptyset_k)] \\
&\quad - (\eta^{(\cdot)}(\mathbb{U}^*(\cdot) + \mathbb{U}(\cdot)) - \eta^{(\cdot)}(\mathbb{U}^*(\cdot)))
\end{aligned}$$

The lemma is hereby proved. \square

In Theorem B.16, $\mathbb{U}(\cdot)$ is any feasible strategy profile. Now let us consider the case that $\mathbb{U}(\cdot)$ is an ϵ -equilibrium in the following lemma:

Lemma B.17 *Let $\mathbb{U}(\cdot)$ be an ϵ -equilibrium for the marketing game with $\eta^{(\cdot)}(\cdot)$ as the social welfare function and $\zeta_k^{(\cdot)}(\cdot)$ as the payoff for the player k . The strategy profile $\mathbb{U}^*(\cdot)$ maximizes the $\eta^{(\cdot)}(\mathbb{U}^*(\cdot))$. Then we have*

$$\eta^{(\cdot)}(\mathbb{U}^*(\cdot)) \leq 2\eta^{(\cdot)}(\mathbb{U}(\cdot)) - [\eta^{(\cdot)}(\mathbb{U}^*(\cdot) + \mathbb{U}(\cdot)) - \eta^{(\cdot)}(\mathbb{U}^*(\cdot)) - m \cdot \epsilon]$$

Proof: First of all, we have the following inequalities:

$$\sum_{k=1}^m [\eta^{(\cdot)}(\mathbb{U}(\cdot) \oplus U_k^*(\cdot)) - \eta^{(\cdot)}(\mathbb{U}(\cdot) \oplus \emptyset_k)] \quad (\text{B.35a})$$

$$\leq \sum_{k=1}^m \max_{U'_k(\cdot) \in \mathcal{D}_k} [\eta^{(\cdot)}(\mathbb{U}(\cdot) \oplus U'_k(\cdot)) - \eta^{(\cdot)}(\mathbb{U}(\cdot) \oplus \emptyset_k(\cdot))] \quad (\text{B.35b})$$

$$\leq \sum_{k=1}^m \max_{U'_k(\cdot) \in \mathcal{D}_k} [\zeta_k^{(\cdot)}(\mathbb{U}(\cdot) \oplus U'_k(\cdot))] \quad (\text{B.35c})$$

$$\leq \sum_{k=1}^m [\zeta_k^{(\cdot)}(\mathbb{U}(\cdot) \oplus U_k(\cdot)) + \epsilon] \quad (\text{B.35d})$$

$$= \eta^{(\cdot)}(\mathbb{U}(\cdot)) + m \cdot \epsilon \quad (\text{B.35e})$$

where (B.35b) arises from the max operation, (B.35c) is due to the competitiveness of the marketing game, (B.35d) holds because $\mathbb{U}(\cdot)$ is the ϵ -equilibrium and (B.35e) is out of the definition of $\eta^{(\cdot)}(\cdot)$ the fact that $\zeta_k^{(\cdot)}(\mathbb{U}(\cdot) \oplus \emptyset_k) = 0$.

Then, using Theorem B.16, we have

$$\begin{aligned} \eta^{(\cdot)}(\mathbb{U}^*(\cdot)) &\leq \eta^{(\cdot)}(\mathbb{U}(\cdot)) + \sum_{k=1}^m [\eta^{(\cdot)}(\mathbb{U}(\cdot) \oplus U_k^*(\cdot)) - \eta^{(\cdot)}(\mathbb{U}(\cdot) \oplus \emptyset_k)] \\ &\quad - (\eta^{(\cdot)}(\mathbb{U}^*(\cdot) + \mathbb{U}(\cdot)) - \eta^{(\cdot)}(\mathbb{U}^*(\cdot))) \\ &\leq 2\eta^{(\cdot)}(\mathbb{U}(\cdot)) - (\eta^{(\cdot)}(\mathbb{U}^*(\cdot) + \mathbb{U}(\cdot)) - \eta^{(\cdot)}(\mathbb{U}^*(\cdot)) - m \cdot \epsilon) \end{aligned}$$

The theorem is proved. □

Theorem B.17 is sufficient to show that the PoA is bounded by 2 in the FMCG marketing game because the ϵ can be set at 0. However, for the DCG marketing game, we still need to show that $\eta^{(\cdot)}(\mathbb{U}^*(\cdot) + \mathbb{U}(\cdot)) - \eta^{(\cdot)}(\mathbb{U}^*(\cdot))$ is greater than some constant which is irrelevant to the choice of $\mathbb{U}(\cdot)$. The following lemma deals with this.

Lemma B.18 Denote $\mathbb{U}^*(\cdot) \in \mathcal{D}$ as the maximizer for $\eta^{(1)}(\cdot)$ and $\mathbb{U}(\cdot) \in \mathcal{D}$ as an ϵ -equilibrium for the DCG marketing game. There exists a constant $\delta > 0$ such that $\eta^{(1)}(\mathbb{U}^*(\cdot) \oplus \mathbb{U}(\cdot)) - \eta^{(1)}(\mathbb{U}(\cdot)) \geq \delta$. Here the value of δ is irrelevant to the choice of ϵ or $\mathbb{U}(\cdot)$.

Proof: As for the DCG competition model (4.1), let $I(t) = \sum_{k=1}^m I_k(t)$ and $R(t) = \sum_{k=1}^m R_k(t)$, then (4.1) can be written as (B.36).

$$\begin{aligned} \frac{dS(t)}{dt} &= -\text{diag}[S(t)] \left(\text{diag}[V(t)] \Lambda (1_n - S(t) - R(t)) + \sum_{k=1}^m C_k U_k(t) \right) \\ \frac{dR(t)}{dt} &= \Gamma (1_n - S(t) - R(t)) \end{aligned} \quad (\text{B.36})$$

where $V(t) \equiv 1_n$. Because $S(t) + I(t) + R(t) = 1_n$, we consider the dynamics of $S(t)$ and $R(t)$ in (B.36) and write it in the compact form follows:

$$\begin{aligned} \dot{S}(t) &= f_S(S(t), R(t), U(t), V(t)) \\ \dot{R}(t) &= f_R(S(t), R(t), U(t), V(t)) \end{aligned} \quad (\text{B.37})$$

At each time t , the value of $S(t)$ and $R(t)$ are determined by the input $U(\cdot)$ and $V(\cdot)$ as well as the initial condition $S(0) = S_0$ and $R(0) = R_0$. Therefore, we use the notion that $S(t) = \phi_S(t, S_0, R_0, U(\cdot), V(\cdot))$ and $R(t) = \phi_R(t, S_0, R_0, U(\cdot), V(\cdot))$. The social welfare function as specified in Section 4.1 is $\eta^{(1)}(\mathbb{U}(\cdot)) = W^T \cdot (1_n - \phi_S(T, 1_n, 0, \mathbb{U}(\cdot), 1_n))$.

Let's define a new strategy profile $\mathbb{U}'(\cdot)$ as

$$\mathbb{U}'(t) = \begin{cases} \mathbb{U}^*(t) & 0 \leq t \leq T \\ \mathbb{U}(t - T) & T < t \leq 2T \end{cases}$$

Then using Lemma 3.3 in Section 3.3.2, we have that

$$\begin{aligned} & \phi_S(T, 1_n, 0, \mathbb{U}^*(\cdot), 1_n) - \phi_S(T, 1_n, 0, \mathbb{U}(\cdot) + \mathbb{U}^*(\cdot), 1_n) \\ & \geq \phi_S(T, 1_n, 0, \mathbb{U}^*(\cdot), 1_n) - \phi_S(2T, 1_n, 0, \mathbb{U}'(\cdot), 1_n) \end{aligned} \quad (\text{B.38})$$

Denote $f_S^k(\cdot)$ as the k -th entry of $f_S(\cdot)$, $f_R^k(\cdot)$ as the k -th entry of $f_R(\cdot)$, u_k^j as the j -th entry of U_k , v^j as the j -th entry of V , s^k as the k -th entry of S and r^k as the k -th entry of R . Then it's easy to verify that

$$\begin{aligned} \frac{\partial f_S^i}{\partial s^j} &= v_i s_i \lambda_{ij} \geq 0 & \frac{\partial f_S^i}{\partial r^j} &= v_i s_i \lambda_{ij} \geq 0 \\ \frac{\partial f_R^i}{\partial s^j} &= 0 & \frac{\partial f_R^i}{\partial r^j} &= 0 \end{aligned}$$

for any $i \neq j$. Also, we can verify that

$$\begin{aligned} \frac{\partial f_S^i}{\partial v^j} &= -s_i \sum_{k=1}^n \lambda_{jk} (1 - s^k - r^k) \cdot \mathbf{1}_{\{i=j\}} \leq 0 & \frac{\partial f_S^i}{\partial u_k^j} &= -s_i c_k^j \cdot \mathbf{1}_{i=j} \leq 0 \\ \frac{\partial f_R^i}{\partial v^j} &= 0 & \frac{\partial f_R^i}{\partial u_k^j} &= 0 \end{aligned}$$

for any i, j and k . According to [80], equation (B.37) is a monotone control system, with $\phi_0 = \phi_S(T, 1_n, 0, \mathbb{U}^*(\cdot), 1_n)$, we have

$$\begin{aligned} & \phi_S(T, 1_n, 0, \mathbb{U}^*(\cdot), 1_n) - \phi_S(2T, 1_n, 0, \mathbb{U}'(\cdot), 1_n) \\ & \geq \phi_0 - \phi_S(T, \phi_0, 1_n - \phi_0, \mathbb{U}(\cdot), 1_n) \\ & \geq \phi_0 - \phi_S(T, \phi_0, 1_n - \phi_0, \mathbb{U}(\cdot), 0) \end{aligned}$$

Looking at equation (B.36), easily we may claim that

$$\phi_0 \geq \exp \left[- \int_0^T \left(\Lambda 1_n + \sum_{j=1}^m C_j U_j^*(t) dt \right) \right]$$

For each k , since the budget $B_k(t)$ is continuous for $t \in [0, T]$, therefore $B_k(t)$ is also bounded. Therefore there exists some constant vector ξ such that $\phi_0 \geq \xi$. Similarly, we have

$$\phi_S(T, \phi_0, 1_n - \phi_0, \mathbb{U}(\cdot), 0) = \text{diag}[\phi_0] \cdot \exp \left[- \int_0^T \sum_{j=1}^m C_j U_j(t) \right]$$

Since $\mathbb{U}(\cdot)$ is the ϵ -equilibrium, each player strive to maximize his own payoff and each $\zeta_k^{(1)}(\mathbb{U}(\cdot))$ is monotone w.r.t. $U_k(\cdot)$, we have that $1_n^T \cdot U_k(t) = B_k(t)$ for each k and time $t \geq 0$. Hereby there exists some constant vector $\xi' > 0$ such that $\exp \left[- \int_0^T \sum_{j=1}^m C_j U_j(t) \right] \leq 1_n - \xi'$. Then we have

$$\begin{aligned} & \eta^{(1)}(\mathbb{U}^*(\cdot) + \mathbb{U}(\cdot)) - \eta^{(1)}(\mathbb{U}^*(\cdot)) \\ &= W^T \cdot [\phi_S(T, 1_n, 0, \mathbb{U}^*(\cdot), 1_n) - \phi_S(T, 1_n, 0, \mathbb{U}^*(\cdot) + \mathbb{U}(\cdot), 1_n)] \\ &\geq W^T \cdot [\phi_S(T, 1_n, 0, \mathbb{U}^*(\cdot), 1_n) - \phi_S(2T, 1_n, 0, \mathbb{U}^*(\cdot), 1_n)] \\ &\geq W^T \cdot [\phi_0 - \phi_S(T, \phi_0, 1_n - \phi_0, \mathbb{U}(\cdot), 0)] \\ &\geq W^T \cdot \text{diag}[\xi] \cdot \xi' \end{aligned}$$

By now this lemma is proved. □

Theorem [B.18](#) indicates that $\eta^{(1)}(\mathbb{U}^*(\cdot) + \mathbb{U}(\cdot)) - \eta^{(1)}(\mathbb{U}^*(\cdot))$ features with a positive lower bound which is not dependent on the specific choice of the $\mathbb{U}(\cdot)$. Therefore, we may claim that with $\mathbb{U}(\cdot)$ being the ϵ -equilibrium for the DCG marketing game with sufficiently small ϵ , then the PoA is no greater than 2. Up to now the who theorem has been proved.

B.17 Proof for Theorem 4.8

Before starting the specific analysis, we first define some notations for the convenience of analysis later on. According to the Theorem 3.2.6 in [116], the utility function $\zeta_k^{(2)}(\cdot)$ is twice continuously Fréchet differentiable. Then for $\zeta_k^{(2)}(\mathbb{U}(\cdot))$, we write its first order derivative as

$$\frac{\partial \zeta_k^{(2)}(\cdot)}{\partial \mathbb{U}(\cdot)}(\mathbb{U}(\cdot)) \mathbb{U}(\cdot) = \sum_{i=1}^m \frac{\partial \zeta_k^{(2)}(\cdot)}{\partial U_i(\cdot)}(\mathbb{U}(\cdot)) U_i(\cdot)$$

as well as the second order derivative as

$$\left(\frac{\partial^2 \zeta_k^{(2)}(\cdot)}{\partial \mathbb{U}(\cdot)^2}(\mathbb{U}(\cdot)) \mathbb{U}^1(\cdot) \right) \mathbb{U}^2(\cdot) = \sum_{i=1}^m \sum_{j=1}^m \left(\frac{\partial^2 \zeta_k^{(2)}(\cdot)}{\partial U_i(\cdot) \partial U_j(\cdot)}(\mathbb{U}(\cdot)) U_j^1(\cdot) \right) U_i^2(\cdot) \quad (\text{B.39})$$

Since $\zeta_k^{(2)}(\cdot)$ is Fréchet differentiable, according to the Rietz-Fréchet theorem (Theorem 4.12 in [134]), we can write the gradient of $\zeta_k(\cdot)$ as follows:

$$\nabla \zeta_k^{(2)}(\cdot) = \left[\nabla_1 \zeta_k^{(2)}(\cdot)^T \quad \nabla_2 \zeta_k^{(2)}(\cdot)^T \quad \dots \quad \nabla_m \zeta_k^{(2)}(\cdot)^T \right]^T$$

where $\nabla_i \zeta_k^{(2)}(\mathbb{U}(\cdot)) = \frac{\partial \zeta_k^{(2)}(\mathbb{U}(\cdot))}{\partial U_i(\cdot)}$. Using $\langle \cdot, \cdot \rangle$ as the notion for inner product we have that $\left\langle \nabla_i \zeta_k^{(2)}(\mathbb{U}(\cdot)), U_i(\cdot) \right\rangle = \frac{\partial \zeta_k^{(2)}(\cdot)}{\partial U_i(\cdot)}(\mathbb{U}(\cdot)) U_i(\cdot)$. We also extends the denotation of the inner product as follows:

$$\left\langle \nabla \zeta_k^{(2)}(\mathbb{U}(\cdot)), \mathbb{U}'(\cdot) \right\rangle = \sum_{i=1}^m \left\langle \nabla_i \zeta_k^{(2)}(\mathbb{U}(\cdot)), U_i'(\cdot) \right\rangle$$

Also, for simplicity of notations, let's define matrix $\mathbb{Q}(\mathbb{U}(\cdot)) = [Q_{ij}(\mathbb{U}(\cdot))]_{m \times m}$ with the entry $Q_{ij}(\mathbb{U}(\cdot))$ to be the 2nd-order partial derivative of a twice continuously Fréchet differentiable function at $\mathbb{U}(\cdot)$. Then for any $\mathbb{U}^1(\cdot), \mathbb{U}^2(\cdot) \in \mathcal{D}$, we

define the notation $\langle \cdot, \cdot, \cdot \rangle$ as follows:

$$\langle \mathbb{U}^1(\cdot), \mathbb{Q}(\mathbb{U}(\cdot)), \mathbb{U}^2(\cdot) \rangle = \sum_{i=1}^m \sum_{j=1}^m (Q_{ij}(\mathbb{U}(\cdot)) U_j^2(\cdot)) U_i^1(\cdot)$$

We would call such a matrix $\mathbb{Q}(\mathbb{U}(\cdot))$ to be positive definite if

$$\langle \mathbb{U}(\cdot), \mathbb{Q}(\mathbb{U}(\cdot)), \mathbb{U}(\cdot) \rangle > 0$$

for any feasible strategy profile $\mathbb{U}(\cdot)$ such that $\|\mathbb{U}(\cdot)\|_2 > 0$. Here the 2-norm of $\mathbb{U}(\cdot)$ is defined as follows:

$$\|\mathbb{U}(\cdot)\|_2 = \sum_{k=1}^m \sum_{i=1}^n \|u_k^i(\cdot)\|_2$$

$u_k^i(\cdot)$ is the i -th entry of $U_k(\cdot)$. As an example, (B.39) can be expressed as

$$\left\langle \mathbb{U}^1(\cdot), \frac{\partial^2 \zeta_k^{(2)}(\cdot)}{\partial \mathbb{U}(\cdot)^2}(\mathbb{U}(\cdot)), \mathbb{U}^2(\cdot) \right\rangle = \sum_{i=1}^m \sum_{j=1}^m \left(\frac{\partial^2 \zeta_k^{(2)}(\cdot)}{\partial U_i(\cdot) \partial U_j(\cdot)}(\mathbb{U}(\cdot)) U_j^2(\cdot) \right) U_i^1(\cdot)$$

where $\frac{\partial^2 \zeta_k^{(2)}(\cdot)}{\partial \mathbb{U}(\cdot)^2}$ is the matrix with $\frac{\partial^2 \zeta_k^{(2)}(\cdot)}{\partial U_i(\cdot) \partial U_j(\cdot)}$ being on its i -th row and j -th column. Also, we define the quasi gradient $g(\mathbb{U}(\cdot))$ of the social welfare function $\eta^{(2)}(\mathbb{U}(\cdot)) = \sum_{k=1}^m \zeta_k^{(2)}(\mathbb{U}(\cdot))$ as follows:

$$g(\mathbb{U}(\cdot)) = \begin{bmatrix} \nabla_1 \zeta_1^{(2)}(\mathbb{U}(\cdot)) \\ \nabla_2 \zeta_2^{(2)}(\mathbb{U}(\cdot)) \\ \dots \\ \nabla_m \zeta_m^{(2)}(\mathbb{U}(\cdot)) \end{bmatrix} \quad (\text{B.40})$$

A very important property, the quasi strict concavity of the social welfare function $\eta^{(2)}(\mathbb{U}(\cdot))$ is given by the following definition:

Definition B.2 *The social welfare function $\eta^{(2)}(\mathbb{U}(\cdot))$ is called quasi strictly concave if*

$$\langle g(\mathbb{U}^1(\cdot)), (\mathbb{U}^2(\cdot) - \mathbb{U}^1(\cdot)) \rangle + \langle g(\mathbb{U}^2(\cdot)), (\mathbb{U}^1(\cdot) - \mathbb{U}^2(\cdot)) \rangle > 0$$

with any strategy profiles $\mathbb{U}^1(\cdot), \mathbb{U}^2(\cdot) \in \mathcal{D}$ such that $\|\mathbb{U}^1(\cdot) - \mathbb{U}^2(\cdot)\|_2 > 0$.

The quasi strict concavity is essential to showing the almost uniqueness of Nash equilibrium for the FMCG marketing game, which the readers will see later. Before that, we may need first to show that the quasi gradient $g(\mathbb{U}(\cdot))$ of $\eta^{(2)}(\mathbb{U}(\cdot)) = \sum_{k=1}^m \zeta_k^{(2)}(\mathbb{U}(\cdot))$ is quasi strictly concave, which is given by Theorem B.19.

Lemma B.19 *As for the FMCG marketing game as specified in Section 4.2, the quasi gradient $g(\mathbb{U}(\cdot))$ defined by (B.40) is quasi strictly concave, i.e.*

$$\langle g(\mathbb{U}^1(\cdot)), (\mathbb{U}^2(\cdot) - \mathbb{U}^1(\cdot)) \rangle + \langle g(\mathbb{U}^2(\cdot)), (\mathbb{U}^1(\cdot) - \mathbb{U}^2(\cdot)) \rangle > 0$$

for any feasible strategy profiles $\mathbb{U}^1(\cdot), \mathbb{U}^2(\cdot) \in \mathcal{D}$ such that $\|\mathbb{U}^1(\cdot) - \mathbb{U}^2(\cdot)\|_2 > 0$.

Proof: First, let's investigate matrix $\mathbb{G}(\mathbb{U}(\cdot)) = [G_{ij}(\mathbb{U}(\cdot))]_{m \times m}$ with $G_{ij}(\mathbb{U}(\cdot)) = \frac{\partial^2 \zeta_i^{(2)}(\mathbb{U}(\cdot))}{\partial U_i(\cdot) \partial U_j(\cdot)}$. It should be noted that such $\mathbb{G}(\mathbb{U}(\cdot))$ is in fact the Fréchet derivative of the quasi gradient $g(\mathbb{U}(\cdot))$, i.e. $\mathbb{G}(\mathbb{U}(\cdot)) = \frac{\partial g(\mathbb{U}(\cdot))}{\partial \mathbb{U}(\cdot)}$.

To investigate $\mathbb{G}(\mathbb{U}(\cdot))$, we define auxiliary matrices $\mathbb{P}(\mathbb{U}(\cdot)) = [P_{ij}(\mathbb{U}(\cdot))]_{m \times m}$, $\mathbb{Q}^k(\mathbb{U}(\cdot)) = [Q_{ij}^k(\mathbb{U}(\cdot))]_{m \times m}$ and $\mathbb{H}(\mathbb{U}(\cdot)) = [H_{ij}(\mathbb{U}(\cdot))]_{m \times m}$ with the corresponding

component defined as follows:

$$\begin{aligned}
P_{ij}(\mathbb{U}(\cdot)) &= \frac{\partial^2 \zeta_i^{(2)}(\mathbb{U}(\cdot))}{\partial U_i(\cdot) \partial U_j(\cdot)} \cdot \mathbf{1}_{\{i=j\}} \\
H_{ij}(\mathbb{U}(\cdot)) &= \sum_{k=1}^m \frac{\partial^2 \zeta_k^{(2)}(\mathbb{U}(\cdot))}{\partial U_i(\cdot) \partial U_j(\cdot)} \\
Q_{ij}^k(\mathbb{U}(\cdot)) &= \frac{\partial^2 \zeta_k^{(2)}(\mathbb{U}(\cdot))}{\partial U_i(\cdot) \partial U_j(\cdot)} \cdot \mathbf{1}_{\{i \neq k, j \neq k\}}
\end{aligned}$$

It is easy to verify that for any i and j we have

$$G_{ij}(\mathbb{U}(\cdot)) + G_{ji}(\mathbb{U}(\cdot)) = P_{ij}(\mathbb{U}(\cdot)) - \sum_{k=1}^m Q_{ij}^k(\mathbb{U}(\cdot)) + H_{ij}(\mathbb{U}(\cdot))$$

Also, by Theorem 4.5, for each k the utility function $\zeta_k^{(2)}(\mathbb{U}(\cdot))$ is concave w.r.t. $U_k(\cdot)$ and convex w.r.t. $U_{-k}(\cdot)$, therefore the matrices $\mathbb{P}(\mathbb{U}(\cdot))$ and $-\mathbb{Q}^k(\mathbb{U}(\cdot))$ are both negative positive.

In addition, according to Theorem 4.5 we know that the social welfare $\eta^{(2)}(\cdot) = \sum_{k=1}^m \zeta_k^{(2)}(\cdot)$ is concave w.r.t. $\mathbb{U}(\cdot)$, therefore the matrix $\mathbb{H}(\mathbb{U}(\cdot))$, which can be viewed as the generalized ‘‘Hessian’’ of $\eta^{(2)}(\mathbb{U}(\cdot))$, is also negative positive. Therefore, we may conclude that $G(\mathbb{U}(\cdot)) + G^T(\mathbb{U}(\cdot))$ is negative positive for any strategy profile $\mathbb{U}(\cdot) \in \mathcal{D}$.

Next, let’s take any two feasible strategy sets $\mathbb{U}^1(\cdot), \mathbb{U}^2(\cdot) \in \mathcal{D}$. For any $0 \leq \alpha \leq 1$, let’s denote $\mathbb{U}^\alpha(\cdot) = \alpha \mathbb{U}^1(\cdot) + (1 - \alpha) \mathbb{U}^2(\cdot)$. Since $g(\mathbb{U}(\cdot))$ is Fréchet differentiable, we have

$$\begin{aligned}
g(\mathbb{U}^2(\cdot)) - g(\mathbb{U}^1(\cdot)) &= \int_0^1 \frac{\partial g(\mathbb{U}^\alpha(\cdot))}{\partial \mathbb{U}(\cdot)} (\mathbb{U}^2(\cdot) - \mathbb{U}^1(\cdot)) d\alpha \\
&= \int_0^1 \mathbb{G}(\mathbb{U}^\alpha(\cdot)) (\mathbb{U}^2(\cdot) - \mathbb{U}^1(\cdot)) d\alpha
\end{aligned} \tag{B.41}$$

Then, calculate the inner product of (B.41) with $\mathbb{U}^1(\cdot) - \mathbb{U}^2(\cdot)$, we have

$$\begin{aligned}
& \langle g(\mathbb{U}^1(\cdot)), (\mathbb{U}^2(\cdot) - \mathbb{U}^1(\cdot)) \rangle + \langle g(\mathbb{U}^2(\cdot)), (\mathbb{U}^1(\cdot) - \mathbb{U}^2(\cdot)) \rangle \\
&= - \langle g(\mathbb{U}^2(\cdot)) - g(\mathbb{U}^1(\cdot)), (\mathbb{U}^2(\cdot) - \mathbb{U}^1(\cdot)) \rangle \\
&= - \int_0^1 \langle \mathbb{U}^2(\cdot) - \mathbb{U}^1(\cdot), \mathbb{G}(\mathbb{U}^\alpha(\cdot)), \mathbb{U}^2(\cdot) - \mathbb{U}^1(\cdot) \rangle d\alpha \\
&= - \frac{1}{2} \int_0^1 \langle \mathbb{U}^2(\cdot) - \mathbb{U}^1(\cdot), \mathbb{G}(\mathbb{U}^\alpha(\cdot)) + \mathbb{G}^T(\mathbb{U}^\alpha(\cdot)), \mathbb{U}^2(\cdot) - \mathbb{U}^1(\cdot) \rangle d\alpha \\
&> 0
\end{aligned}$$

where the inequality holds because $\mathbb{G}(\mathbb{U}(\cdot)) + \mathbb{G}^T(\mathbb{U}(\cdot))$ is negative positive as is shown previously. This lemma is hereby proved. \square

With all these preparations, we are now ready to show this theorem by contradiction. Assume there exists two distinct Nash equilibriums $\mathbb{U}^1(\cdot), \mathbb{U}^2(\cdot) \in \mathcal{D}$ for the FMCG marketing game and $\|\mathbb{U}^1(\cdot) - \mathbb{U}^2(\cdot)\|_2 > 0$. In the strategy profile $\mathbb{U}^1(\cdot)$, by the definition of Nash equilibrium, each strategy $U_k^1(\cdot)$ is actually the solution to the following optimization problem:

$$\begin{aligned}
& \max \quad \zeta_k^{(2)}(\mathbb{U}_{-k}^1 \oplus U_k^1(\cdot)) \\
& \text{s.t.} \quad \sum_{i=1}^n u_{ki}^1(t) \leq B_k(t) \\
& \quad \quad u_{ki}^1(t) \geq 0 \\
& \quad \quad i = 1, \dots, n, \quad 0 \leq t \leq T
\end{aligned}$$

where $u_{ki}^1(\cdot)$ is the i -th entry of $U_k^1(\cdot)$ and \mathbb{U}_{-k}^1 is already given. Let the generalized Lagrange multiplier $\mu_{k0}^1(\cdot) \geq 0$ be associated with the constraint $\sum_{i=1}^n u_{ki}^1(t) \leq B_k(t)$ and $\mu_{ki}^1(\cdot) \geq 0$ associated with the constraint with $u_{ki}^1(\cdot) \geq 0$ for each i . Invoking the generalized Karush-Kuhn-Tucker (KKT) conditions (see Chapter 2 in [117]), we

have (B.42).

$$\nabla_k \zeta_k^{(2)}(\mathbb{U}^1(\cdot)) + \begin{bmatrix} \mu_{k1}^1(\cdot) \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \mu_{k2}^1(\cdot) \\ \vdots \\ 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \mu_{kn}^1(\cdot) \end{bmatrix} - \begin{bmatrix} \mu_{k0}^1(\cdot) \\ \mu_{k0}^1(\cdot) \\ \vdots \\ \mu_{k0}^1(\cdot) \end{bmatrix} = 0 \quad (\text{B.42})$$

Similarly, for the strategy set $\mathbb{U}^2(\cdot)$, we also have

$$\nabla_k \zeta_k^{(2)}(\mathbb{U}^2(\cdot)) + \begin{bmatrix} \mu_{k1}^2(\cdot) \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \mu_{k2}^2(\cdot) \\ \vdots \\ 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \mu_{kn}^2(\cdot) \end{bmatrix} - \begin{bmatrix} \mu_{k0}^2(\cdot) \\ \mu_{k0}^2(\cdot) \\ \vdots \\ \mu_{k0}^2(\cdot) \end{bmatrix} = 0 \quad (\text{B.43})$$

Calculate the inner product of (B.42) with $\mathbb{U}^2(\cdot) - \mathbb{U}^1(\cdot)$ as well as the inner product of (B.43) with $\mathbb{U}^1(\cdot) - \mathbb{U}^2(\cdot)$ and add them together, we have the follows:

$$\langle g(\mathbb{U}^1(\cdot)), (\mathbb{U}^2(\cdot) - \mathbb{U}^1(\cdot)) \rangle + \langle g(\mathbb{U}^2(\cdot)), (\mathbb{U}^1(\cdot) - \mathbb{U}^2(\cdot)) \rangle = 0$$

which is in contradiction with Theorem B.19. The theorem is proved.

B.18 Proof for Theorem 4.9

First, we will prove that the starting from any strategy profile $\mathbb{U}(0, \cdot) \in \mathcal{D}$, as $\mathbb{U}(t, \cdot)$ evolves according to (4.3), the strategy profile $\mathbb{U}(t, \cdot)$ will stay feasible, i.e. $\mathbb{U}(t, \cdot) \in \mathcal{D}$ for any $t \geq 0$ by contradiction. First let's define \hat{t} such that $\hat{t} = \inf \{t | \mathbb{U}(t, \cdot) \notin \mathcal{D}\}$. Easily there exists some $\hat{t}' > \hat{t}$ such that at the interval $t \in [0, \hat{t}']$, the right side of (4.3) remains bounded and therefore $\mathbb{U}(t, \cdot)$ is continuous for $t \in [0, \hat{t}']$.

Since \mathcal{D} is naturally a closed set, therefore at time \hat{t} , the strategy profile $\mathbb{U}(\hat{t}, \cdot) \in \mathcal{D}$ while after \hat{t} it immediately gets out of \mathcal{D} . There are only two kinds of possibilities for this. First, for some $1 \leq i \leq n$, $1 \leq k \leq m$ and $s \in [0, T]$ such that $u_{ki}(\hat{t}, s) = 0$ and $\left. \frac{d}{dt} u_{ki}(t, s) \right|_{t=\hat{t}} < 0$. Secondly, for some $1 \leq k \leq m$ and $s \in [0, T]$ we have $\sum_{i=1}^n u_{ki}(\hat{t}, s) = B_k(s)$ and $\left. \frac{d}{dt} \sum_{i=1}^n u_{ki}(t, s) \right|_{t=\hat{t}} > 0$. Notice that

$$\begin{aligned} & \|f_k(\mathbb{U}(\hat{t}, \cdot), \mu_k(\cdot))(s)\|_2 \\ &= \left\| \nabla_k \zeta_k^{(2)}(\mathbb{U}(\hat{t}, \cdot))(s) \right\|_2 \\ & \quad + \sum_{i=1}^n \left[2(\mu_{ki}(s) - \mu_{k0}(s))^T \cdot \nabla_{ki} \zeta_k^{(2)}(\mathbb{U}(t, \cdot))(s) + (\mu_{ki}(s) - \mu_{k0}(s))^2 \right] \end{aligned}$$

where $\nabla_{ki} \zeta_k^{(2)}(\mathbb{U}(\cdot))(s)$ refers to the i -th entry of $\nabla_k \zeta_k^{(2)}(\mathbb{U}(\cdot))(s)$. For the first kind of possibility, consider the first-order derivative

$$\begin{aligned} & \frac{\partial}{\partial \mu_{ki}(s)} \|f_k(\mathbb{U}(\hat{t}, \cdot), \mu(\cdot))(s)\|_2 \\ &= 2 \left(\nabla_{ki} \zeta_k^{(2)}(\mathbb{U}(\hat{t}, \cdot))(s) + (\mu_{ki}(s) - \mu_{k0}(s)) \right) \\ &= 2 \left. \frac{d}{dt} u_{ki}(t, s) \right|_{t=\hat{t}} < 0 \end{aligned}$$

where $u_{ki}(\cdot)$ is the i -th entry of $U_k(\cdot)$. However, since $u_{ki}(\hat{t}, s) = 0$, according to (4.4) we have that $\mu_{ki}(s)$ is nonnegative and minimize $\|f_k(\mathbb{U}(\hat{t}, \cdot), \mu(\cdot))(s)\|_2$, which means that we can further decrease $\|f_k(\mathbb{U}(t, \cdot), \mu(\cdot))(s)\|_2$ by increasing $\mu_{ki}(s)$, which is in contradiction to (4.4). Therefore the first possibility does not hold.

Similarly, as for the second possibility we have

$$\begin{aligned}
& \frac{\partial}{\partial \mu_{k0}(s)} \|f_k(\mathbb{U}(\hat{t}, \cdot), \mu(\cdot))(s)\|_2 \\
&= -2 \sum_{i=1}^n \left[\nabla_{k_i \zeta_k^{(2)}}(\mathbb{U}(\hat{t}, \cdot))(s) + (\mu_{ki}(s) - \mu_{k0}(s)) \right] \\
&= -2 \left. \frac{d}{dt} \sum_{i=1}^n u_{ki}(t, s) \right|_{t=\hat{t}} < 0
\end{aligned}$$

Following the similar analysis for the first possibility, the second possibility does not hold either. Therefore, we may conclude that starting from any feasible strategy profile, the solution by the best-response mechanism (4.3) is continuous and remains feasible in \mathcal{D} for each $t \geq 0$. The first part of the theorem is proved. Next, we move on to show the second part of the theorem. As the first step, we will use the following lemma to show that $\|f_k(t)\|_2$ is convergent to 0 for each k as each player updates their strategies according to the best-response mechanism (4.3).

Lemma B.20 *Starting from any initial feasible strategy set $\mathbb{U}(0, \cdot)$, for any $1 \leq k \leq m$, we have the magnitude of the dynamics $\|f_k(t)\|_2 \rightarrow 0$ as the time $t \rightarrow \infty$.*

Proof: Let's consider $f(\cdot) = \left[f_1^T(\cdot) \quad f_2^T(\cdot) \quad \dots \quad f_m^T(\cdot) \right]^T$, the time derivative of $f(\cdot)$ is given as follows:

$$\frac{d}{dt} f(\mathbb{U}(t, \cdot), \mu(\cdot)) = \langle \mathbb{G}(\mathbb{U}(t, \cdot)), f(\mathbb{U}(t, \cdot), \mu(\cdot)) \rangle + \begin{bmatrix} \dot{\mu}_1(\cdot) \\ \dot{\mu}_2(\cdot) \\ \vdots \\ \dot{\mu}_n(\cdot) \end{bmatrix} - \begin{bmatrix} \dot{\mu}_{10}(\cdot) \cdot 1_n \\ \dot{\mu}_{20}(\cdot) \cdot 1_n \\ \vdots \\ \dot{\mu}_{n0}(\cdot) \cdot 1_n \end{bmatrix}$$

where $\mathbb{G}(\cdot)$ is as defined in the proof for Theorem B.19. For convenience, we denote that $\mu_k(\cdot) = \left[\mu_{k1}(\cdot) \quad \mu_{k2}(\cdot) \quad \dots \quad \mu_{kn}(\cdot) \right]^T$. According to the definition of $\mu(\cdot)$'s in

(4.4), if $\mu_{ki}(s)$ or $\dot{\mu}_{ki}(s)$ does not equal to 0, then we must have $u_{ki}(s) = 0$ for $i = 1, \dots, n$ or $\sum_{i=1}^n u_{ki}(s) = B_k(s)$ for $i = 0$. However, since $\mu_{ki}(s)$'s minimize $f_k(\mathbb{U}(t, \cdot), \mu(\cdot))(s)$, according to the Hilbert Projection Theorem in combination with (4.3), we have that

$$\left(\begin{array}{c} \left[\begin{array}{c} \dot{\mu}_{k1}(\cdot) \\ \dot{\mu}_{k2}(\cdot) \\ \vdots \\ \dot{\mu}_{kn}(\cdot) \end{array} \right] - \left[\begin{array}{c} \dot{\mu}_{k0}(\cdot) \\ \dot{\mu}_{k0}(\cdot) \\ \vdots \\ \dot{\mu}_{k0}(\cdot) \end{array} \right] \end{array} \right)^T \cdot f_k(\mathbb{U}(t, \cdot), \mu(\cdot)) = 0$$

for $1 \leq k \leq m$. Therefore the time derivative of $\|f(\cdot)\|_2$ can be expressed as follows:

$$\frac{d}{dt} \|f(\mathbb{U}(t, \cdot), \mu(\cdot))\|_2 = \langle f(\mathbb{U}(t, \cdot), \mu(\cdot)), \mathbb{G}(\mathbb{U}(t, \cdot)), f(\mathbb{U}(t, \cdot), \mu(\cdot)) \rangle$$

According to the proof for Theorem B.19, we have

$$\begin{aligned} & \frac{d}{dt} \|f(\mathbb{U}(t, \cdot), \mu(\cdot))\|_2 \\ &= \langle f(\mathbb{U}(t, \cdot), \mu(\cdot)), \mathbb{G}(\mathbb{U}(t, \cdot)), f(\mathbb{U}(t, \cdot), \mu(\cdot)) \rangle \\ &\leq \sum_{k=1}^m \left\langle f_k(\mathbb{U}(t, \cdot), \mu(\cdot)), \nabla_k^2 \zeta_k^{(2)}(\mathbb{U}(t, \cdot)), f_k(\mathbb{U}(t, \cdot), \mu(\cdot)) \right\rangle \end{aligned}$$

According to the proof for Theorem 4.5, $\zeta_k^{(2)}(\mathbb{U}(\cdot))$ is the exponential of some other functional which is concave w.r.t. $U_k(\cdot)$. Therefore there exists some $\epsilon > 0$ such that the following equation holds:

$$\begin{aligned} & \left\langle f_k(\mathbb{U}(t, \cdot), \mu(\cdot)), \nabla_k^2 \zeta_k^{(2)}(\mathbb{U}(t, \cdot)), f_k(\mathbb{U}(t, \cdot), \mu(\cdot)) \right\rangle \\ &\leq -\epsilon \cdot \left\langle \nabla_k \zeta_k^{(2)}(\mathbb{U}(t, \cdot), \mu(\cdot)), f_k(\mathbb{U}(t, \cdot), \mu(\cdot)) \right\rangle^2 \\ &= -\epsilon \cdot \left(\left\langle f_k(\mathbb{U}(t, \cdot), \mu(\cdot)), f_k(\mathbb{U}(t, \cdot), \mu(\cdot)) \right\rangle \right) \end{aligned}$$

$$\begin{aligned}
& - \left\langle \mu_k(\cdot) - \mu_{k0}(\cdot) \cdot \mathbf{1}_n, f_k(\mathbb{U}(t, \cdot), \mu(\cdot)) \right\rangle^2 \\
& = -\epsilon \cdot \langle f_k(\mathbb{U}(t, \cdot), \mu(\cdot)), f_k(\mathbb{U}(t, \cdot), \mu(\cdot)) \rangle^2 \\
& = -\epsilon \cdot \|f_k(\mathbb{U}(t, \cdot), \mu(\cdot))\|_2^4
\end{aligned}$$

Since $\|f(\mathbb{U}(t, \cdot), \mu(\cdot))\|_2 = \sum_{k=1}^m \|f_k(\mathbb{U}(t, \cdot), \mu(\cdot))\|_2$, we have

$$\frac{d}{dt} \|f(\mathbb{U}(t, \cdot), \mu(\cdot))\|_2 \leq -\epsilon \cdot \|f(\mathbb{U}(t, \cdot), \mu(\cdot))\|_2^4$$

Therefore $\|f(\mathbb{U}(t, \cdot), \mu(\cdot))\|_2$ is convergent to 0 as time $t \rightarrow \infty$, so as each $\|f_k(\mathbb{U}(t, \cdot), \mu(\cdot))\|_2$ does. This lemma is proved. \square

Now denote $\mathbb{U}^0(\cdot)$ as a Nash equilibrium for the FMCG marketing game, then invoking the KKT condition [117], we have

$$\nabla_k \zeta_k^{(2)}(\mathbb{U}^0(\cdot)) + \begin{bmatrix} \mu_{k1}^0(\cdot) \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \mu_{k2}^0(\cdot) \\ \vdots \\ 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \mu_{kn}^0(\cdot) \end{bmatrix} - \begin{bmatrix} \mu_{k0}^0(\cdot) \\ \mu_{k0}^0(\cdot) \\ \vdots \\ \mu_{k0}^0(\cdot) \end{bmatrix} = 0 \quad (\text{B.44})$$

for each $1 \leq k \leq m$, where the quantities $\mu_{ki}^0(\cdot)$'s are the corresponding Lagrange multipliers.

For $f_k(\mathbb{U}(t, \cdot), \mu(\cdot))$, let's plus $\langle U_k^0(\cdot) - U_k(t, \cdot), f_k(\mathbb{U}(t, \cdot), \mu(\cdot)) \rangle$ with the in-

ner product of $U_k(t, \cdot) - U_k^0(\cdot)$ and (B.44) sum them over k , then we have

$$\begin{aligned}
& \langle g(\mathbb{U}^0(\cdot)), (\mathbb{U}(t, \cdot) - \mathbb{U}^0(\cdot)) \rangle + \langle g(\mathbb{U}(t, \cdot)), (\mathbb{U}^0(\cdot) - \mathbb{U}(t, \cdot)) \rangle \\
& + \sum_{k=1}^m \left[\sum_{i=1}^n \langle \mu_{ki}^0(\cdot), u_{ki}(t, \cdot) \rangle + \langle \mu_{k0}^0(\cdot), B_k(\cdot) - \mathbf{1}_n^T \cdot U_k(t, \cdot) \rangle \right. \\
& \left. + \sum_{i=1}^n \langle \mu_{ki}(\cdot), u_{ki}^0(\cdot) \rangle + \langle \mu_{k0}(\cdot), B_k(\cdot) - \mathbf{1}_n^T \cdot U_k^0(\cdot) \rangle \right] \tag{B.45} \\
& = \langle \mathbb{U}^0(\cdot) - \mathbb{U}(t, \cdot), f(\mathbb{U}(t, \cdot), \mu(\cdot)) \rangle \\
& \leq \|\mathbb{U}^0(\cdot) - \mathbb{U}(t, \cdot)\|_2 \cdot \|f(\mathbb{U}(t, \cdot), \mu(\cdot))\|_2
\end{aligned}$$

Since for any time $t \geq 0$, we have $\mathbb{U}(t, \cdot) \in \mathcal{D}$, where the feasible set \mathcal{D} is bounded, therefore the term $\|\mathbb{U}^0(\cdot) - \mathbb{U}(t, \cdot)\|_2$ is bounded. According to Theorem B.20, we know that $\|f(\mathbb{U}(t, \cdot), \mu(\cdot))\|_2$ is convergent to 0 and thus the R.H.S. of (B.45) goes to 0 as $t \rightarrow \infty$.

At the same time, because $\mathbb{U}^0(\cdot), \mathbb{U}(t, \cdot) \in \mathcal{D}$ and $g(\cdot)$ is quasi strictly concave as as specified by Theorem B.19, each term at the L.H.S. of (B.45) is actually nonnegative. Therefore, we can get the follows:

$$\lim_{t \rightarrow \infty} \langle g(\mathbb{U}^0(\cdot)), (\mathbb{U}(t, \cdot) - \mathbb{U}^0(\cdot)) \rangle + \langle g(\mathbb{U}(t, \cdot)), (\mathbb{U}^0(\cdot) - \mathbb{U}(t, \cdot)) \rangle = 0$$

On the other hand, we have that

$$\begin{aligned}
& \langle g(\mathbb{U}^0(\cdot)), (\mathbb{U}(t, \cdot) - \mathbb{U}^0(\cdot)) \rangle + \langle g(\mathbb{U}(t, \cdot)), (\mathbb{U}^0(\cdot) - \mathbb{U}(t, \cdot)) \rangle \\
& = - \langle g(\mathbb{U}(t, \cdot)) - g(\mathbb{U}^0(\cdot)), \mathbb{U}(t, \cdot) - \mathbb{U}^0(\cdot) \rangle \\
& = - \int_0^1 \langle \mathbb{U}(t, \cdot) - \mathbb{U}^0(\cdot), \mathbb{G}(\alpha \mathbb{U}(t, \cdot) + (1 - \alpha) \mathbb{U}^0(\cdot)), \mathbb{U}(t, \cdot) - \mathbb{U}^0(\cdot) \rangle d\alpha
\end{aligned}$$

By the proof for Theorem [B.19](#) and [B.20](#), we have

$$\begin{aligned}
& - \int_0^1 \langle \mathbb{U}(t, \cdot) - \mathbb{U}^0(\cdot), \mathbb{G}(\alpha \mathbb{U}(t, \cdot) + (1 - \alpha) \mathbb{U}^0(\cdot)), \mathbb{U}(t, \cdot) - \mathbb{U}^0(\cdot) \rangle d\alpha \\
& \geq - \sum_{k=1}^m \int_0^1 \langle \mathbb{U}(t, \cdot) - \mathbb{U}^0(\cdot), \nabla^2 \zeta_k^{(2)}(\alpha \mathbb{U}(t, \cdot) + (1 - \alpha) \mathbb{U}^0(\cdot)), \mathbb{U}(t, \cdot) - \mathbb{U}^0(\cdot) \rangle d\alpha
\end{aligned} \tag{B.46a}$$

$$\geq \epsilon \cdot \sum_{k=1}^m \int_0^1 \langle \nabla \zeta_k^{(2)}(\alpha \mathbb{U}(t, \cdot) + (1 - \alpha) \mathbb{U}^0(\cdot)), \mathbb{U}(t, \cdot) - \mathbb{U}^0(\cdot) \rangle^2 d\alpha \tag{B.46b}$$

$$\geq \epsilon \cdot \sum_{k=1}^m \left[\int_0^1 \langle \nabla \zeta_k^{(2)}(\alpha \mathbb{U}(t, \cdot) + (1 - \alpha) \mathbb{U}^0(\cdot)), \mathbb{U}(t, \cdot) - \mathbb{U}^0(\cdot) \rangle d\alpha \right]^2 \tag{B.46c}$$

$$= \epsilon \cdot \sum_{k=1}^m \left[\zeta_k^{(2)}(\mathbb{U}(t, \cdot)) - \zeta_k^{(2)}(\mathbb{U}^0(\cdot)) \right]^2 \tag{B.46d}$$

where [\(B.46a\)](#) holds based on the proof for Theorem [B.19](#) and [\(B.46b\)](#) is based on the proof for Theorem [B.20](#). The third inequality [\(B.46c\)](#) holds as a result of Cauchy-Schwartz inequality. Here $\epsilon > 0$ is a constant. Up to now, we have shown that

$$\lim_{t \rightarrow \infty} \sum_{k=1}^m \left[\zeta_k^{(2)}(\mathbb{U}(t, \cdot)) - \zeta_k^{(2)}(\mathbb{U}^0(\cdot)) \right]^2 = 0$$

The theorem is hereby proved.

Bibliography

- [1] J. A. Chevalier and D. Mayzlin. The effect of word of mouth on sales: Online book reviews. *JOURNAL OF MARKETING RESEARCH*, 43(3):345–354, 2006.
- [2] J. J. Brown and P. H. Reingen. Social ties and word-of-mouth referral behavior. *JOURNAL OF CONSUMER RESEARCH*, 14(3):350–362, 1987.
- [3] J. Goldenberg, B. Libai, and E. Muller. Talk of the network: A complex systems look at the underlying process of word-of-mouth. *MARKETING LETTERS*, 12(3):211–223, 2001.
- [4] J. Goldenberg, B. Libai, and E. Muller. Using complex systems analysis to advance marketing theory development: Modeling heterogeneity effects on new product growth through stochastic cellular automata. *ACADEMY OF MARKETING SCIENCE REVIEW*, 9(3):1–18, 2001.
- [5] D. Scott. *The New Rules of Marketing and PR: How to Use News Releases, Blogs, Podcasting, Viral Marketing and Online Media to Reach Buyers Directly*. New Rules of Marketing & PR: How to Use Social Media, Blogs. Wiley, 2007.
- [6] E. Costenbader and T. W. Valente. The stability of centrality measures when networks are sampled. *SOCIAL NETWORKS*, 25(4):283 – 307, 2003.
- [7] P. Bonacich. Some unique properties of eigenvector centrality. *SOCIAL NETWORKS*, 29(4):555 – 564, 2007.
- [8] L. Katz. A new status index derived from sociometric analysis. *PSYCHOMETRIKA*, 18(1):39–43, 1953.
- [9] A. Langville and C. Meyer. *Google’s PageRank and Beyond: The Science of Search Engine Rankings*. Princeton University Press, 2011.

- [10] G. Sabidussi. The centrality index of a graph. *PSYCHOMETRIKA*, 31(4):581–603, 1966.
- [11] L. C. Freeman. A set of measures of centrality based on betweenness. *SOCIOMETRY*, 40(1):35–41, 1977.
- [12] M. Granovetter. Threshold models of collective behavior. *AMERICAN JOURNAL OF SOCIOLOGY*, 83(6):1420–1443, 1978.
- [13] D. Gruhl, R. Guha, D. Liben-Nowell, and A. Tomkins. Information diffusion through blogspace. In *PROCEEDINGS OF THE 13TH INTERNATIONAL CONFERENCE ON WORLD WIDE WEB, WWW '04*, pages 491–501, New York, NY, USA, 2004. ACM.
- [14] D. Kempe, J. Kleinberg, and E. Tardos. Maximizing the spread of influence through a social network. In *PROCEEDINGS OF THE NINTH ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING, KDD '03*, pages 137–146, New York, NY, USA, 2003. ACM.
- [15] S. Morris. Contagion. *THE REVIEW OF ECONOMIC STUDIES*, 67(1):57–78, 2000.
- [16] H. P. Young. The diffusion of innovations in social networks. Economics Working Paper Archive 437, The Johns Hopkins University, Department of Economics, 2000.
- [17] J. Ghaderi and R. Srikant. Opinion dynamics in social networks: A local interaction game with stubborn agents. In *AMERICAN CONTROL CONFERENCE (ACC), 2013*, pages 1982–1987, 2013.
- [18] H. Haller and S. Sarangi. Nash networks with heterogeneous links. *MATHEMATICAL SOCIAL SCIENCES*, 50(2):181–201, 2005.
- [19] V. Bala and S. Goyal. A strategic analysis of network reliability. In B. Dutta and M. Jackson, editors, *STUDIES IN ECONOMIC DESIGN*, pages 313–336–. Springer Berlin Heidelberg, 2003.
- [20] W. O. Kermack and A. G. McKendrick. A contribution to the mathematical theory of epidemics. *PROCEEDINGS OF THE ROYAL SOCIETY OF LONDON. SERIES A*, 115(772):700–721, 1927.
- [21] N. Bailey. *Mathematical Theory of Infectious Diseases*. Mathematics in Medicine Series. Oxford University Press, Incorporated, 1987.
- [22] R. Anderson and R. May. *Infectious Diseases of Humans: Dynamics and Control*. Oxford Science Publications. OUP Oxford, 1992.

- [23] H. Hethcote. The mathematics of infectious diseases. *SIAM REVIEW*, 42(4):599–653, 2000.
- [24] M. Newman. Spread of epidemic disease on networks. *PHYSICAL REVIEW E*, 66(1, Part 2), 2002.
- [25] R. Pastor-Satorras and A. Vespignani. Epidemic dynamics and endemic states in complex networks. *PHYSICAL REVIEW E*, 63(6, Part 2), 2001.
- [26] M. Barthélemy, A. Barrat, R. Pastor-Satorras, and A. Vespignani. Dynamical patterns of epidemic outbreaks in complex heterogeneous networks. *JOURNAL OF THEORETICAL BIOLOGY*, 235(2):275–288, 2005.
- [27] Y. Wang, D. Chakrabarti, C. Wang, and C. Faloutsos. Epidemic spreading in real networks: an eigenvalue viewpoint. In *RELIABLE DISTRIBUTED SYSTEMS, 2003. PROCEEDINGS. 22ND INTERNATIONAL SYMPOSIUM ON*, pages 25–34, 2003.
- [28] A. Ganesh, L. Massoulié, and D. Towsley. The effect of network topology on the spread of epidemics. In *INFOCOM 2005. 24TH ANNUAL JOINT CONFERENCE OF THE IEEE COMPUTER AND COMMUNICATIONS SOCIETIES. PROCEEDINGS IEEE*, volume 2, pages 1455–1466 vol. 2, 2005.
- [29] P. van Mieghem, J. Omic, and R. Kooij. Virus spread in networks. *IEEE-ACM TRANSACTIONS ON NETWORKING*, 17(1):1–14, 2009.
- [30] M. Youssef and C. Scoglio. An individual-based approach to sir epidemics in contact networks. *JOURNAL OF THEORETICAL BIOLOGY*, 283(1):136–144, 2011.
- [31] P. Shakarian, M. Broecheler, V. S. Subrahmanian, and C. Molinaro. Using generalized annotated programs to solve social network diffusion optimization problems. *ACM TRANSACTIONS ON COMPUTATIONAL LOGIC*, 14(2):10:1–10:40, 2013.
- [32] F. Coelho, O. Cruz, and C. Codeco. Epigrass: a tool to study disease spread in complex networks. *SOURCE CODE FOR BIOLOGY AND MEDICINE*, 3(1):3, 2008.
- [33] D. Centola. The spread of behavior in an online social network experiment. *SCIENCE*, 329(5996):1194–1197, 2010.
- [34] S. Aral, L. Muchnik, and A. Sundararajan. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES*, 106(51):21544–21549, 2009.
- [35] M. Newman. Communities, modules and large-scale structure in networks. *NATURE PHYSICS*, 8(1):25–31, 2012.

- [36] A. Vazquez, J. Oliveira, Z. Dezso, K. Goh, I. Kondor, and A. Barabasi. Modeling bursts and heavy tails in human dynamics. *PHYSICAL REVIEW E*, 73(3, 2), 2006.
- [37] C. Song, Z. Qu, N. Blumm, and A.-L. Barabasi. Limits of predictability in human mobility. *SCIENCE*, 327(5968):1018–1021, 2010.
- [38] C. Wang and B. A. Huberman. How random are online social interactions? *SCIENTIFIC REPORTS*, 2, 2012.
- [39] M. Girvan and M. Newman. Community structure in social and biological networks. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, 99(12):7821–7826, 2002.
- [40] M. Newman. Modularity and community structure in networks. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, 103(23):8577–8582, 2006.
- [41] M. Newman and J. Park. Why social networks are different from other types of networks. *PHYSICAL REVIEW E*, 68(3, Part 2), 2003.
- [42] A. Vespignani. Modelling dynamical processes in complex socio-technical systems. *NATURE PHYSICS*, 8(1):32–39, 2012.
- [43] D. M. Chickering and D. Heckerman. A decision theoretic approach to targeted advertising. In *PROCEEDINGS OF THE SIXTEENTH CONFERENCE ON UNCERTAINTY IN ARTIFICIAL INTELLIGENCE*, UAI'00, pages 82–88, San Francisco, CA, USA, 2000. Morgan Kaufmann Publishers Inc.
- [44] P. Domingos and M. Richardson. Mining the network value of customers. In *PROCEEDINGS OF THE SEVENTH ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING*, KDD '01, pages 57–66, New York, NY, USA, 2001. ACM.
- [45] M. Richardson and P. Domingos. Mining knowledge-sharing sites for viral marketing. In *PROCEEDINGS OF THE EIGHTH ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING*, KDD '02, pages 61–70, New York, NY, USA, 2002. ACM.
- [46] D. Kempe, J. Kleinberg, and E. Tardos. Influential nodes in a diffusion model for social networks. In *AUTOMATA, LANGUAGES AND PROGRAMMING*, volume 3580 of *Lecture Notes in Computer Science*, pages 1127–1138. Springer Berlin Heidelberg, 2005.
- [47] E. Berger. Dynamic monopolies of constant size. *JOURNAL OF COMBINATORIAL THEORY, SERIES B*, 83(2):191–200, 2001.
- [48] S. P. Sethi. Dynamic optimal control models in advertising: A survey. *SIAM REVIEW*, 19(4):685–725, 1977.

- [49] G. Feichtinger, R. F. Hartl, and S. P. Sethi. Dynamic optimal control models in advertising: Recent developments. *MANAGEMENT SCIENCE*, 40(2):195–226, 1994.
- [50] R. C. Rao. A note on optimal and near optimal price and advertising strategies. *MANAGEMENT SCIENCE*, 31(3):376–377, 1985.
- [51] R. Rishel. A partially observed advertising model. *OPTIMAL CONTROL THEORY AND ECONOMICS ANALYSIS*, 2:253–262, 1985.
- [52] K. Conrad. *Quality, advertising and the formation of goodwill under dynamic conditions*. Beiträge zur angewandten Wirtschaftsforschung. Inst. für Volkswirtschaftslehre u. Statistik d. Univ., 1983.
- [53] J. T. Teng and G. L. Thompson. Optimal strategies for general price-quality decision models of new products with learning production costs. *EUROPEAN JOURNAL OF OPERATIONAL RESEARCH*, 93(3):476–489, 1996.
- [54] M. Khouzani, S. Sarkar, and E. Altman. Optimal control of epidemic evolution. In *INFOCOM, 2011 PROCEEDINGS IEEE*, pages 1683–1691, 2011.
- [55] L. J. Kornish and Q. Li. Optimal referral bonuses with asymmetric information: Firm-offered and interpersonal incentives. *MARKETING SCIENCE*, 29(1):108–121, 2009.
- [56] A. Krause and D. Golovin. Submodular function maximization. In *TRACTABILITY: PRACTICAL APPROACHES TO HARD PROBLEMS*, chapter 3, pages 71–104. Cambridge University Press, 2014.
- [57] U. Feige. A threshold of $\ln n$ for approximating set cover. *J. ACM*, 45(4):634–652, 1998.
- [58] A. Krause and C. Guestrin. Near-optimal value of information in graphical models. In *CONFERENCE ON UNCERTAINTY IN ARTIFICIAL INTELLIGENCE (UAI)*, 2005.
- [59] A. Ranzter and B. Bernhardsson. Control of convex monotone systems. In *53RD IEEE CONFERENCE ON DECISION AND CONTROL*, Los Angeles, CA, USA, 2014.
- [60] N. Alon, M. Feldman, A. D. Procaccia, and M. Tennenholtz. A note on competitive diffusion through social networks. *INF. PROCESS. LETT.*, 110(6):221–225, 2010.
- [61] S. Bharathi, D. Kempe, and M. Salek. Competitive influence maximization in social networks. In *PROCEEDINGS OF THE 3RD INTERNATIONAL CONFERENCE ON INTERNET AND NETWORK ECONOMICS*, WINE’07, pages 306–311, Berlin, Heidelberg, 2007. Springer-Verlag.

- [62] S. Goyal and M. Kearns. Competitive contagion in networks. In *PROCEEDINGS OF THE FORTY-FOURTH ANNUAL ACM SYMPOSIUM ON THEORY OF COMPUTING*, STOC '12, pages 759–774, New York, NY, USA, 2012. ACM.
- [63] V. Tzoumas, C. Amanatidis, and E. Markakis. A game-theoretic analysis of a competitive diffusion process over social networks. In *PROCEEDINGS OF THE 8TH INTERNATIONAL CONFERENCE ON INTERNET AND NETWORK ECONOMICS*, WINE'12, pages 1–14, Berlin, Heidelberg, 2012. Springer-Verlag.
- [64] G. E. Fruchter. The many-player advertising game. *MANAGEMENT SCIENCE*, 45(11):1609–1611, 1999.
- [65] J.-P. Dubé and P. Manchanda. Differences in dynamic brand competition across markets: An empirical analysis. *MARKETING SCIENCE*, 24(1):81–95, 2005.
- [66] D. Nguyen and L. Shi. Competitive advertising strategies and market-size dynamics: A research note on theory and evidence. *MANAGEMENT SCIENCE*, 52(6):965–973, 2006.
- [67] G. Erickson. *Dynamic Models of Advertising Competition*. International Series in Quantitative Marketing. Springer US, 2003.
- [68] T. Carnes, C. Nagarajan, S. M. Wild, and A. van Zuylen. Maximizing influence in a competitive social network: A follower’s perspective. In *PROCEEDINGS OF THE NINTH INTERNATIONAL CONFERENCE ON ELECTRONIC COMMERCE*, ICEC '07, pages 351–360, New York, NY, USA, 2007. ACM.
- [69] C. Budak, D. Agrawal, and A. El Abbadi. Limiting the spread of misinformation in social networks. In *PROCEEDINGS OF THE 20TH INTERNATIONAL CONFERENCE ON WORLD WIDE WEB*, WWW '11, pages 665–674, New York, NY, USA, 2011. ACM.
- [70] X. He, G. Song, W. Chen, and Q. Jiang. Influence blocking maximization in social networks under the competitive linear threshold model. In *SIAM INTERNATIONAL CONFERENCE ON DATA MINING*, pages 463–474. SIAM, Omnipress, 2012.
- [71] P. Dubey, R. Garg, and B. De Meyer. Competing for customers in a social network (r). Department of Economics Working Papers 13-01, Stony Brook University, Department of Economics, 2013.
- [72] J. Tsai, Y. Qian, Y. Vorobeychik, C. Kiekintveld, and M. Tambe. Bayesian security games for controlling contagion. In *SOCIAL COMPUTING (SOCIAL-COM)*, 2013 INTERNATIONAL CONFERENCE ON, pages 33–38, 2013.

- [73] S. Johnson, A. N.rrgensen and G. Zaccour. *Differential Games in Marketing*. International Series in Quantitative Marketing. Springer US, 2004.
- [74] R. Isaacs. *Differential Games: A Mathematical Theory with Applications to Warfare and Pursuit, Control and Optimization*. Dover books on mathematics. Dover Publications, 1999.
- [75] T. Liggett. *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*. Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen. Springer, 1999.
- [76] V. Sood and S. Redner. Voter model on heterogeneous graphs. *PHYSICAL REVIEW LETTERS*, 94:178701, 2005.
- [77] P. Holme and M. E. J. Newman. Nonequilibrium phase transition in the coevolution of networks and opinions. *PHYSICAL REVIEW E.*, 74:056108, 2006.
- [78] A. Vetta. Nash equilibria in competitive societies, with applications to facility location, traffic routing and auctions. In *FOUNDATIONS OF COMPUTER SCIENCE, 2002. PROCEEDINGS. THE 43RD ANNUAL IEEE SYMPOSIUM ON*, pages 416–425, 2002.
- [79] J. B. Rosen. Existence and uniqueness of equilibrium points for concave n-person games. *ECONOMETRICA*, 33(3):520–534, 1965.
- [80] D. Angeli and E. Sontag. Monotone control systems. *AUTOMATIC CONTROL, IEEE TRANSACTIONS ON*, 48(10):1684–1698, 2003.
- [81] H. Kim, J. Tang, R. Anderson, and C. Mascolo. Centrality prediction in dynamic human contact networks. *COMPUT. NETW.*, 56(3):983–996, 2012.
- [82] G. Adomavicius and A. Tuzhilin. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *KNOWLEDGE AND DATA ENGINEERING, IEEE TRANSACTIONS ON*, 17(6):734–749, 2005.
- [83] J. Kleinberg. Computing: The wireless epidemic. *NATURE*, 449(7160):287–288, 2007.
- [84] R. Albert, H. Jeong, and A. Barabasi. Error and attack tolerance of complex networks. *NATURE*, 406(6794):378–382, 2000.
- [85] C. Kang, C. Molinaro, S. Kraus, Y. Shavitt, and V. Subrahmanian. Diffusion centrality in social networks. In *ADVANCES IN SOCIAL NETWORKS ANALYSIS AND MINING (ASONAM), 2012 IEEE/ACM INTERNATIONAL CONFERENCE ON*, pages 558–564, 2012.

- [86] A. Banerjee, A. G. Chandrasekhar, E. Duflo, and M. O. Jackson. The diffusion of microfinance. *SCIENCE*, 341(6144):–, 2013.
- [87] P. Grindrod and D. J. Higham. A dynamical systems view of network centrality. *PROCEEDINGS OF THE ROYAL SOCIETY A: MATHEMATICAL, PHYSICAL AND ENGINEERING SCIENCE*, 470(2165):–, May 8, 2014.
- [88] K. Ide, R. Zamami, and A. Namatame. Diffusion centrality in interconnected networks. *PROCEDIA COMPUTER SCIENCE*, 24(0):227–238, 2013.
- [89] M. May. Hagggle project (<http://www.hagggleproject.org/>), 2010.
- [90] I. Carreras, D. Miorandi, G. Canright, and K. Eng-Monsen. Eigenvector centrality in highly partitioned mobile networks: Principles and applications. In F. Dressler and I. Carreras, editors, *STUDIES IN COMPUTATIONAL INTELLIGENCE*, volume 69, pages 123–145. Springer Berlin Heidelberg, 2007.
- [91] A. a. Guimarães, A. B. Vieira, A. P. C. Silva, and A. Ziviani. Fast centrality-driven diffusion in dynamic networks. In *PROCEEDINGS OF THE 22ND INTERNATIONAL CONFERENCE ON WORLD WIDE WEB COMPANION*, WWW '13 Companion, pages 821–828, Republic and Canton of Geneva, Switzerland, 2013. International World Wide Web Conferences Steering Committee.
- [92] K. Lerman, R. Ghosh, and J. H. Kang. Centrality metric for dynamic networks. In *PROCEEDINGS OF THE EIGHTH WORKSHOP ON MINING AND LEARNING WITH GRAPHS*, MLG '10, pages 70–77, New York, NY, USA, 2010. ACM.
- [93] H. Kim and R. Anderson. Temporal node centrality in complex networks. *PHYS. REV. E*, 85:026107, 2012.
- [94] Y. Wang, G. Cong, G. Song, and K. Xie. Community-based greedy algorithm for mining top-k influential nodes in mobile social networks. In *PROCEEDINGS OF THE 16TH ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING*, KDD '10, pages 1039–1048, New York, NY, USA, 2010. ACM.
- [95] A. Mochalova and A. Nanopoulos. On the role of centrality in information diffusion in social networks. *ECIS Completed Research, Paper 101*, 2013.
- [96] C. Meyer. *Matrix Analysis and Applied Linear Algebra*, chapter 8, page 673. Matrix Analysis and Applied Linear Algebra. Society for Industrial and Applied Mathematics, 2000.
- [97] H. Khalil. *Nonlinear Systems*, chapter 3.4, pages 102–103. Prentice Hall PTR, 2002.

- [98] A. Jeffrey and D. Zwillinger. *Table of Integrals, Series, and Products*, chapter 15.819, page 1078. Elsevier Science, 2000.
- [99] A. Brouwer and W. Haemers. *Spectra of Graphs*, chapter 3.1, page 33. Universitext (En ligne). Springer New York, 2012.
- [100] D. Stauffer and A. Aharony. *Introduction To Percolation Theory*. Taylor & Francis, 1994.
- [101] E. Leicht and R. M. D’Souza. Percolation on interacting networks. *Arxiv preprint arXiv:0907.0894*, 2009.
- [102] A. Allard, P.-A. Noel, L. J. Dube, and B. Pourbohloul. Heterogeneous bond percolation on multitype networks with an application to epidemic dynamics. *PHYSICAL REVIEW E*, 79(3, 2), 2009.
- [103] C. Meyer. *Matrix Analysis and Applied Linear Algebra*, chapter 7, page 630. Matrix Analysis and Applied Linear Algebra. Society for Industrial and Applied Mathematics, 2000.
- [104] M. Newman, A. Barabasi, and D. Watts. *The Structure and Dynamics of Networks*., chapter 4.1, pages 232–233. Princeton Studies in Complexity. Princeton University Press, 2006.
- [105] M. Kimura, K. Saito, and H. Motoda. Blocking links to minimize contamination spread in a social network. *ACM TRANSACTIONS ON KNOWLEDGE DISCOVERY FROM DATA*, 3(2):9:1–9:23, 2009.
- [106] X. Zhai, L. Zheng, J. Wang, and C. W. Tan. Optimization algorithms for epidemic evolution in broadcast networks. In *IEEE WIRELESS COMMUNICATIONS AND NETWORKING CONFERENCE (WCNC)*, pages 1540–1545, 2013.
- [107] O. Hinz, B. Skiera, C. Barrot, and J. U. Becker. Seeding strategies for viral marketing: An empirical comparison. *JOURNAL OF MARKETING*, 75(6):55–71, 2011.
- [108] G. L. Nemhauser and L. A. Wolsey. Best algorithms for approximating the maximum of a submodular set function. *MATHEMATICS OF OPERATIONS RESEARCH*, 3(3):177–188, 1978.
- [109] U. Feige, V. S. Mirrokni, and J. Vondr. Maximizing non-monotone submodular functions. *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*, 0:461–471, 2007.
- [110] N. Nisan, T. Roughgarden, E. Tardos, and V. Vazirani. *Algorithmic Game Theory*, chapter 15.6, pages 402–405. Cambridge University Press, 2007.

- [111] T. M. Liggett. Stochastic models of interacting systems. *THE ANNALS OF PROBABILITY*, 25(1):1–29, 1997.
- [112] D. Aldous. Interacting particle systems as stochastic social dynamics. *BERNOULLI*, 19(4):1122–1149, 2013.
- [113] E. Yildiz, A. Ozdaglar, D. Acemoglu, A. Saberi, and A. Scaglione. Binary opinion dynamics with stubborn agents. *ACM TRANSACTIONS ON ECONOMICS AND COMPUTATION*, 1(4):19:1–19:30, 2013.
- [114] M. Kimura, K. Saito, K. Ohara, and H. Motoda. Opinion formation by voter model with temporal decay dynamics. In *PROCEEDINGS OF THE 2012 EUROPEAN CONFERENCE ON MACHINE LEARNING AND KNOWLEDGE DISCOVERY IN DATABASES - VOLUME PART II*, ECML PKDD’12, pages 565–580, Berlin, Heidelberg, 2012. Springer-Verlag.
- [115] A. Filippov. On certain questions in the theory of optimal control. *JOURNAL OF THE SOCIETY FOR INDUSTRIAL & APPLIED MATHEMATICS, SERIES A: CONTROL*, 1(1):76–84, 1962.
- [116] A. Bressan and B. Piccoli. *Introduction to the Mathematical Theory of Control*. AIMS on applied math. American Institute of Mathematical Sciences, 2007.
- [117] M. Hinze, R. Pinnau, and M. Ulbrich. *Optimization with PDE Constraints*. Mathematical Modelling: Theory and Applications. Springer, 2008.
- [118] M. Newman. *Networks: An Introduction*, chapter 7.2, page 171. Oxford University Press, Inc., New York, NY, USA, 2010.
- [119] S. Wasserman and K. Faust. *SOCIAL NETWORK ANALYSIS: METHODS AND APPLICATIONS*, chapter 4.9.7, pages 159–161. Structural Analysis in the Social Sciences. Cambridge University Press, 1994.
- [120] M. Minoux. Accelerated greedy algorithms for maximizing submodular set functions. In J. Stoer, editor, *LECTURE NOTES IN CONTROL AND INFORMATION SCIENCES*, volume 7, pages 234–243–. Springer Berlin Heidelberg, 1978.
- [121] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. VanBriesen, and N. Glance. Cost-effective outbreak detection in networks. In *PROCEEDINGS OF THE 13TH ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING*, KDD ’07, pages 420–429, New York, NY, USA, 2007. ACM.
- [122] D. Golovin and A. Krause. Adaptive submodularity: Theory and applications in active learning and stochastic optimization. *J. ARTIF. INT. RES.*, 42(1):427–486, 2011.

- [123] M. Streeter and D. Golovin. An online algorithm for maximizing submodular functions. In D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou, editors, *ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS 21*, pages 1577–1584. Curran Associates, Inc., 2009.
- [124] K. Munagala, S. Babu, R. Motwani, and J. Widom. The pipelined set cover problem. In *PROCEEDINGS OF THE 10TH INTERNATIONAL CONFERENCE ON DATABASE THEORY, ICDT'05*, pages 83–98, Berlin, Heidelberg, 2005. Springer-Verlag.
- [125] D. Golovin, M. Faulkner, and A. Krause. Online distributed sensor selection. In *PROCEEDINGS OF THE 9TH ACM/IEEE INTERNATIONAL CONFERENCE ON INFORMATION PROCESSING IN SENSOR NETWORKS, IPSN '10*, pages 220–231, New York, NY, USA, 2010. ACM.
- [126] F. Radlinski, R. Kleinberg, and T. Joachims. Learning diverse rankings with multi-armed bandits. In *PROCEEDINGS OF THE 25TH INTERNATIONAL CONFERENCE ON MACHINE LEARNING, ICML '08*, pages 784–791, New York, NY, USA, 2008. ACM.
- [127] M. Streeter, D. Golovin, and A. Krause. Online learning of assignments. In Y. Bengio, D. Schuurmans, J. Lafferty, C. Williams, and A. Culotta, editors, *ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS 22*, pages 1794–1802. Curran Associates, Inc., 2009.
- [128] A. Guillory and J. Bilmes. Interactive submodular set cover. In J. Fürnkranz and T. Joachims, editors, *PROCEEDINGS OF THE 27TH INTERNATIONAL CONFERENCE ON MACHINE LEARNING (ICML-10)*, pages 415–422, Haifa, Israel, 2010. Omnipress.
- [129] A. Guillory and J. A. Bilmes. Simultaneous learning and covering with adversarial noise. In *PROCEEDINGS OF THE 28TH INTERNATIONAL CONFERENCE ON MACHINE LEARNING (ICML-11)*, pages 369–376, 2011.
- [130] D. Angeli and E. Sontag. Interconnections of monotone systems with steady-state characteristics. In M. Queiroz, M. Malisoff, and P. Wolenski, editors, *OPTIMAL CONTROL, STABILIZATION AND NONSMOOTH ANALYSIS*, volume 301 of *Lecture Notes in Control and Information Science*, pages 135–154. Springer Berlin Heidelberg, 2004.
- [131] A. Kononenko. Equilibrium positional strategies in non-antagonistic differential games. *DOKLADY AKADEMII NAUK SSSR*, 231(2):285–288, 1976.
- [132] C. Pierre. *Introduction to differential games*. Universite de Brest, Place du Marchal De Lattre De Tassigny 75775 Paris Cedex 16 - France, 2010.
- [133] A. Seierstad. Existence of open loop nash equilibria in certain types of non-linear differential games. *OPTIMIZATION LETTERS*, 8(1):247–258, 2014.

- [134] W. Rudin. *Real and Complex Analysis*. Mathematics series. McGraw-Hill, 1987.