COMMUN. MATH. SCI.
Vol. 14, No. 6, pp. 1723–1740

AN SIS-TYPE MARKETING MODEL ON RANDOM NETWORKS*

REINHARD ILLNER† AND JUNLING MA‡

Abstract. Marketing on random networks displays similarities to epidemiological models in the sense that "word-of-mouth" information passes between individuals and may "infect" susceptible buyers such that they end up buying the product. The difference to epidemics is that there are usually many competing products (rather than just one disease), and in addition to word-of-mouth transmission, products are also advertised by the producers, which can be thought of as external nodes connected to the network. In this paper we develop a model in which these various transmission pathways compete, and, in addition, where product fatigue and product switching are possible. This is a genuine and realistic extension of the model developed in [M. Li, R. Edwards, R. Illner, and J. Ma, Commun. Math. Sci., 13, 497–509, 2015], where a customer would never abandon a product after purchase. The model presented here is similar to and was inspired by SIS epidemiological models. We discuss the homogeneous limit for a fully connected graph, present some analytical properties of the models and conduct a number of numerical experiments, including an investigation of a modelling assumption we call "edge chaos". The validity of this assumption turns out to depend on the type of the underlying random network.

Key words. Random network, SIS marketing model, product fatigue, product switch.

AMS subject classifications. 92D60.

1. Introduction

The classical Bass model [1] and its variants are widely used product diffusion models, and they have been shown to agree well with empirical data (see, *e.g.* [11, 12]). These models describe the propagation of a new product with advertising and word-of-mouth marketing in a completely connected and homogeneous population. Li *et al.* [8] generalizes the classical Bass model to populations modelled as a random graph of configuration type [16], and to several competing products. The underlying assumptions in this model are that a potential buyer (a "susceptible") is subject to independent efforts by the competing companies (efforts realized by advertising) and to the independent influences transmitted by his/her neighbours or acquaintances, who will praise the virtues of the product they have already bought. Once the susceptible buyer has in fact purchased a product, he/she is assumed to stick with this product forever. These assumptions allow to generalize the ideas first brought forward by Miller and Volz [14, 20] in the context of epidemiology to marketing.

However, buyers may not stay loyal to their product, as always happens in reality. Products age, break down, become obsolete, or a customer may simply be intrigued by a competing product, abandon what he/she already owns and go shopping for a new product. Product diffusion with repeat purchases in a homogeneous population has been well studied both empirically (see, *e.g.* [6,9]) and theoretically (see, *e.g.* [2,13,17]).

This paper aims to model product diffusion with repeat purchases on social networks. This more realistic scenario has a counterpart in epidemiology known as SIS (susceptible-infected-susceptible) models, in contrast to SIR (susceptible-infected-recovered) models. In SIS models, an infectious individual becomes susceptible imme-

*Received: May 22, 2015; accepted (in revised form): January 15, 2016. Communicated by Pierre Degond.
This research is supported by NSERC Discovery grants (RI and JM).
†Department of Mathematics and Statistics, University of Victoria, Victoria, BC, Canada V8W 2Y2 (rillner@uvic.ca).
‡Department of Mathematics and Statistics, University of Victoria, Victoria, BC, Canada V8W 2Y2 (junlingm@uvic.ca).

diately after recovery; in SIR models, an infectious individual recovers to an immune state. There are two approaches to model SIS disease dynamics. A node based approach (see, *e.g.* [18]) classifies the nodes by their infection status and degrees. However, such a model over-estimates both the growth rate and the equilibrium prevalence $I(\infty)$ [10], because it ignores the fact that a node has fixed neighbours. Lindquist *et al.* [10] extends this approach by properly accounting for the neighbours available for infection. The resulting effective degree model agrees with the mean of the underlying stochastic process very well, even though it is very complex and difficult to analyse. Lindquist *et al.* [10] also show that, on a network, epidemics described by SIS models grow faster than epidemics described by SIR models with identical disease parameters. This is because an SIS type disease can transmit multiple times along an edge, while an SIR type disease can transmit at most once. Another approach is edge based, originates from the pair-approximation models (see, *e.g.* [7]). This model is precise on regular random networks (on which all nodes have the same degree), but is less precise on other networks. This approach is extended by House and Keeling [5] using the ideas of the Volz model [20], describing the dynamics of edges as classified by the infection status of the two nodes connected by the edge. Here we adapt the ideas from the House and Keeling model [5] to the marketing problem, with product fatigue ("recovery"), where the recovered customer returns to being a susceptible customer.

Our model is complicated yet uses some contentious simplifying assumptions (in particular something we call "edge chaos", a concept borrowed from the kinetic theory of gases (see Section 2). Without such assumptions, the modelling would become rather intractable, as will be explained in due time. Further, we will assume that a customer who has bought a product will not abandon this product because of aggressive advertising by the competition or by his contacts; rather, we assume that the customer will abandon the product due to general product fatigue (like a car, a computer, or a refrigerator getting so old that it needs to be replaced), and that the customer will then again be susceptible to all competing products. Generalizations to more sophisticated behaviours are certainly possible but will complicate the model further. The survey article [19] offers a comprehensive discussion of the many different factors which enter into the diffusion of products in the marketplace; our work aims to address the issues of advertising, word-of-mouth, competition and product fatigue in a unified manner. Incorporation of some of the additional effects discussed in [19] is an option for future work.

The fundamental idea, similar to the methodology used in the references on epidemiology, is to study edge dynamics on random networks. Our model will investigate changes in edge types depending on the status of neighbouring edges. This is difficult, and it is in this context that an edge chaos assumption enters into our discussion. This is a major difference to the models derived in [8]. For a fully connected graph and $N \to \infty$, ordinary differential equations like the classical Bass model should emerge in the limit, as discussed for the models in [8] in that article.

We design the model for two competing products, similarly to the scenario discussed in [8]. The setup of the model is done in Section 2, and we provide some discussion and analysis. We also present the model for the homogeneous limit (which arises from a fully connected graph as $N \to \infty$). A proof that our model is consistent with this limit is provided in Appendix A. For simplicity this is only done for the case of one product. In the subsequent section we compare predictions of the model with microscopic numerical simulations based on Gillespie's algorithm [3, 4]. The agreement between the two approaches is excellent, although high fluctuations on some types of scale-free random networks impose limitations on the applicability of the model in such cases. In Section 4 we investigate how the market shares of competing companies depends on the parameters; we present a theorem (Proposition 1) which shows how advertising efforts influence these market shares, provided that the word-of-mouth transmission rates are identical for both products. For a variety of network types we also study the dependence of market shares on network parameters; the homogeneous limit arises naturally in these experiments as a limit case.

The edge chaos assumption is a sore point in our modeling, and we perform a numerical test in which we check whether the assumption holds approximately. Not surprisingly, edge chaos is violated after an initial period; the deviation remains reasonably small for Erdős–Rényi networks, but is significant for our example of scale-free networks. In spite of this deficiency, the agreement between our model predictions and an average of microscopic simulations is very good.

2. Prerequisites, notation, and the model

We consider a population modelled as an undirected random network of configuration type. Such a network is defined as a random graph with N nodes (individuals),where N is large and a priori not known (in practice, N is the realistic potential number of customers). The network is completely described by its edge distribution $\{P_k\}_{k=0,1,2,\dots}$, where P_k is the probability that a randomly chosen node has degree k, i.e., has k neighbours.

For the marketing context we assume that each node (individual) is connected to two competing companies, denoted as A and B, which can "reach" the individual via advertising. Each company has therefore N connections into the network. The (unique) products of A and B will simply be denoted by A and B as well. For convenience we label individuals as $S=S(t)$ (susceptible, has not bought either product), or as $A=A(t)$ or $B=B(t)$ if the individual has bought A or B, respectively. Assuming that no one will buy both products, we immediately get the conservation equation

$$S(t)+A(t)+B(t)=N.$$

If we define S_k as the number of all susceptibles with degree k we have $S=\sum_{k=0}^{\infty} S_k$. If nobody has bought a product at time 0 we have the initial conditions $S_k(0)= P_kN, S(0)=N$. Similarly, defining N_k, A_k, and B_k in the obvious way, we have

$$S_k+A_k+B_k=N_k,$$

with $N_k=NP_k$. The random graph structure enters at this point. We have suppressed the dependence on t and will continue to do this in the sequel, but the initial conditions are $A_k(0)=B_k(0)=0$ for all k.

To summarize the small notational abuse implied by the above: A and B denote two competing companies, their (unique) products, and the number of individuals that hold product A or B at a given time. We will assume that nobody will hold both products at the same time.

Marketing the products A and B involves both direct advertising and word-of-mouth "transmission". In order to incorporate the structure of the network we will keep track of the status of edges. Even though the social network is assumed to be undirected, we assume that each edge consists of two directed edges of opposite direction to keep track of who markets to whom, pointing from *source* to *target*. These directed edges are denoted by pairs like SS, SA, AS, BA, AA, etc. The first letter will always denote the state of the target, the second the state of the source. M_{SS} will be the number of all directed edges with susceptible target and source; note that this number is twice the number of SS-edges, as each edge is counted twice in opposite directions. Similarly, we define M_{AS}, M_{SA}, M_{SB}, M_{BA}, etc. Note that transmission (advertising) under consideration can occur across M_{SA} but not across M_{AS} edges, (because A can transmit to S but not vice versa). However, we have $M_{SA} = M_{AS}$ because they count the same edges. Similarly, $M_{SB} = M_{BS}$ and $M_{AB} = M_{BA}$.

Further, M_S denotes the number of edges with target S. Hence

$$M_S = \sum_{k=0}^{\infty} kS_k = M_{SS} + M_{SA} + M_{SB}. \tag{2.1}$$

Similarly,

$$M_A = \sum_{k=0}^{\infty} kA_k = M_{AS} + M_{AA} + M_{AB}, \tag{2.2}$$

$$M_B = \sum_{k=0}^{\infty} kB_k = M_{BS} + M_{BA} + M_{BB}. \tag{2.3}$$

We denote by β_A (and β_B) the transmission rates from A (B) buyers to susceptible nodes along an edge. β_A and β_B are assumed constant and positive. They could in principle be time- or state-dependent. In contrast, the per capita "conversion rates" of susceptibles due to outside advertising by companies A and B will be denoted by α_A, α_B. We also take them constant, but in optimal advertising strategies they will certainly vary with time. Define further

$$p_A := \frac{M_{SA}}{M_S}, \quad p_B := \frac{M_{SB}}{M_S}. \tag{2.4}$$

These are the probabilities that an edge with target S has source in A or B, respectively.

A last feature which we implement into our model is "recovery". Here this means that the products A and B have a finite lifespan (this could be due to technical innovation (the product becomes obsolete), tiredness of the buyer, or malfunction). For simplicity, we assume that there is a constant "recovery" rate $\gamma > 0$, same for A, B, such that owners of A or B recover to join class S at this rate. As mentioned in the introduction, this is quite simplistic and may be modified in future refinements of the model.

These assumptions naturally lead to a first set of equations for the evolution of S_k,

$$\frac{d}{dt} S_k = -\beta_A p_A k S_k - \beta_B p_B k S_k - \alpha_A S_k - \alpha_B S_k + \gamma(A_k + B_k). \tag{2.5}$$

While this equation is linear in terms of the unknowns S_k, A_k, B_k, it is the appearance of the p_A, p_B, defined in terms of the unknowns M_{SA} etc., which introduces nonlinearity into the system. As $A_k + B_k = N_k - S_k$, the equation implicitly contains the edge distribution via $N_k = NP_k$.

Similarly,

$$\frac{d}{dt} A_k = \beta_A p_A k S_k + \alpha_A S_k - \gamma A_k,$$
$$\frac{d}{dt} B_k = \beta_B p_B k S_k + \alpha_B S_k - \gamma B_k.$$

As mentioned earlier, the initial conditions will be $S_k(0)=P_kN$, $A_k(0)=B_k(0)=0$, and $p_A(0)=p_B(0)=0$.

Summing over k and using $\sum A_k=A$, $\sum B_k=B$, $p_AM_S=M_{SA}$, and $p_BM_S=M_{SB}$ give

$$\frac{d}{dt}A=\beta_A M_{SA}-\gamma A+\alpha_A S, \tag{2.6}$$
$$\frac{d}{dt}B=\beta_B M_{SB}-\gamma B+\alpha_B S. \tag{2.7}$$

Similarly,

$$\frac{d}{dt}M_A=\sum_k k\frac{dA_k}{dt}=\beta_A p_A\sum_k k^2 S_k+\alpha_A M_S-\gamma M_A, \tag{2.8}$$
$$\frac{d}{dt}M_B=\sum_k k\frac{dB_k}{dt}=\beta_B p_B\sum_k k^2 S_k+\alpha_B M_S-\gamma M_B. \tag{2.9}$$

with $M_A(0)=M_B(0)=0$, $\quad M_S(0)=\sum_k kS_k(0)=N\sum_k kP_k$. The number N of nodes is considered constant; it can of course be scaled out of the equations by considering fractions of susceptibles, A-buyers, B-buyers, and different types of edges rather than numbers. The values of the S_k are needed in the remainder of the model setup, as seen in the sequel.

Deriving equations for the numbers of different types of arcs is more involved and will require some fairly strong assumptions. In particular, we need the concept of the "average excess degree" of a susceptible node after following an edge (this node could be source or target). Specifically, observe that $\sum_{k=0}^{\infty} kS_k$ is the total number of edges starting from a susceptible. Hence, if we follow an edge and arrive at a node of class S, the probability that this node has j edges is $jS_j/\sum_{k=0}^{\infty} kS_k$. (this expression reflects the fact that it is proportionally more likely to reach nodes with more edges if one follows an edge). The *average excess degree* E of such nodes is then

$$E:=\sum_{j=1}^{\infty}(j-1)\frac{jS_j}{\sum_{k=0}^{\infty} kS_k}. \tag{2.10}$$

The factor $j-1$ appears here because the edge we followed is not counted. This E clearly depends on the S_k, but for ease of notation we suppress this dependency in our formulas. E involves only the S_k because nodes of type A or B cannot be changed by neighbours; they can recover, but our assumptions are that the recovery does not depend on the neighbouring nodes. For (2.10) to be meaningful, we assume that

$$\sum_{j=1}^{\infty}(j-1)\frac{jP_j}{\sum_{k=0}^{\infty} kP_k}<\infty. \tag{2.11}$$

Also, observe that from the definition

$$p_A M_{SB}=p_B M_{SA}. \tag{2.12}$$

The dynamics of M_{SA} involves no fewer than 10 terms on the right hand side. Here is the equation.

$$\frac{d}{dt}M_{SA}=-\underbrace{\beta_A M_{SA}}_{1}-\underbrace{\gamma M_{SA}}_{2}+\underbrace{\gamma M_{AA}}_{3}+\underbrace{\gamma M_{BA}}_{4}$$

$$+\underbrace{\beta_A E p_A M_{SS}}_{5} - \underbrace{\beta_B E p_B M_{SA}}_{6} - \underbrace{\beta_A E p_A M_{SA}}_{7}$$
$$+\underbrace{\alpha_A M_{SS}}_{8} - \underbrace{\alpha_A M_{SA}}_{9} - \underbrace{\alpha_B M_{SA}}_{10}. \tag{2.13}$$

A chart illustrating the meaning of these 10 numbered gain and loss terms on the right-hand side that represent the flows of edges to and from the M_{SA} class is shown in Figure 2.1. We discuss these numbered terms one at a time. Terms 1–4 are easily

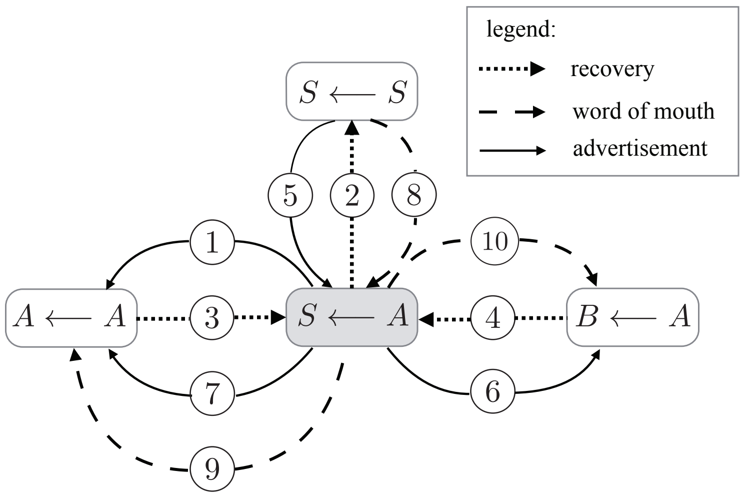

FIG. 2.1. *A flow chart for the gain and loss of the edges in the M_{SA} class (for edges $S \longleftarrow A$, pointing from the source node A to the target node S), whose dynamics is given in* (2.13)*. The circled number on each flow corresponds to the such numbered term in* (2.13)*. Flows 2, 3, and 4 represent gains and losses due to recovery of a buyer on one end of an edge; flows 8, 9, and 10 represent the gains and losses due to the conversion of a susceptible to a buyer by advertisement; the other flows represent the conversion of a susceptible to a buyer by word-of-mouth.*

understood. First, an arc of type SA can leave this class because the target end (S) is recruited (infected) by the source. Second, the source (A) might recover, and the arc leaves the class. Third and fourth, the target might have been of type A or B and recovered, thus adding to the class SA.

Terms 5–7 are a lot harder to explain and include, in fact, an "edge chaos" assumption. For example, let us focus on the gain term 5: This term is there because an arc of type M_{SS} can turn into type M_{SA} because its source can be converted into type A by another neighbour; such neighbours are available with average excess degree E (as defined above) and are of type A with probability p_A (this latter statement is a hidden independence (chaos) assumption; it is not automatically given that if we follow an SS edge from target to source, then neighbours of the source will be type A with probability p_A; making this assumption means that we neglect possible correlations between the edges.)

Similarly, terms 6 and 7 are present because the target S in an SA arc could be recruited (or infected) by edges other than the source, and such edges could be of type A or B. The terms 4–7 can be written in alternate ways by using equation (2.12). Again, hidden chaos assumptions are made here, to be avoided only at the expense of going to much more complicated models, where two-arc distribution densities would have to be introduced, and the model would lose closure.

Terms 8–10 are much simpler: Term 8 is present because the source of an arc of type SS may be recruited externally by advertising, adding the arc to class M_{SA}. Similarly, terms 9 and 10 are present because the target of an arc of type SA may be recruited by company A or B, moving the arc into class M_{AA} or M_{BA}, respectively.

The initial conditions for the quantities appearing in this and subsequent equations are

$$\begin{aligned}&M_{SA}(0)=0, M_{SB}(0)=0, M_{AA}(0)=0,\\&M_{BB}(0)=0, M_{AB}(0)=0,\end{aligned}$$

and

$$M_{SS}(0)=M_S(0)-M_{SA}(0)-M_{SB}(0)=N\sum kP_k.$$

A corresponding equation, with the same caveats, holds for M_{SB}. Because of the symmetry between A and B, this equation can be derived by exchanging A and B in (2.13).

$$\begin{aligned}\frac{d}{dt}M_{SB}=&-\beta_B M_{SB}-\gamma M_{SB}+\gamma M_{BB}+\gamma M_{AB}\\&+\beta_B Ep_B M_{SS}-\beta_A Ep_A M_{SB}-\beta_B Ep_B M_{SB}\\&+\alpha_B M_{SS}-\alpha_B M_{SB}-\alpha_A M_{SB}.\end{aligned}\qquad(2.14)$$

We further recall that $M_S=\sum kS_k$. This means, from the equations for S_k, we know the dynamics of M_S, and as $M_{SS}=M_S-M_{SA}-M_{SB}$,

$$M_{SS}=\sum_k kS_k-M_{SA}-M_{SB}.\qquad(2.15)$$

We do not need to write equations for M_{AS} and M_{BS}, because $M_{AS}=M_{SA}$, and $M_{SB}=M_{BS}$.

It remains to derive equations for M_{AA}, M_{BB}, and M_{AB} (which also gives M_{BA}). We begin with M_{AB}, as the other two can be derived from it similarly to the derivation of M_{SS}.

$$\frac{d}{dt}M_{AB}=-\underbrace{2\gamma M_{AB}}_{1}+\underbrace{\beta_A Ep_A M_{SB}}_{2}+\underbrace{\beta_B Ep_B M_{SA}}_{3}\qquad(2.16)$$

$$+\underbrace{\alpha_B M_{AS}}_{4}+\underbrace{\alpha_A M_{BS}}_{5}\qquad(2.17)$$

Term 1 here is present because either end of the edge can recover (hence the factor 2); terms 2 and 3 are present because the S-end of an SB- or a SA-edge may become recruited by another neighbour (with the previous caveats regarding independence). And last, terms 4 and 5 account for external recruitment (by advertisements) of the S-end of SA- and SB-edges. To close the equations, we need the following from (2.2) and (2.3).

$$M_{AA}=M_A-M_{AS}-M_{AB}=\sum_k kA_k-M_{SA}-M_{AB},\qquad(2.18)$$

$$M_{BB} = M_B - M_{BS} - M_{BA} = \sum_k kB_k - M_{SB} - M_{AB}. \tag{2.19}$$

In summary, the full model consists of the following equations:

$$M_S = \sum_{k=0}^{\infty} kS_k\,;\; E = \frac{\sum_{j=1}^{\infty}(j-1)jS_j}{M_S}\,; \tag{2.20a}$$

$$\frac{d}{dt}S_k = -\beta_A p_A k S_k - \beta_B p_B k S_k - \alpha_A S_k - \alpha_B S_k + \gamma(N_k - S_k)\,; \tag{2.20b}$$

$$\frac{d}{dt}A = \beta_A M_{SA} - \gamma A + \alpha_A S; \tag{2.20c}$$

$$\frac{d}{dt}B = \beta_B M_{SB} - \gamma B + \alpha_B S; \tag{2.20d}$$

$$\frac{d}{dt}M_A = \beta_A p_A \sum_k k^2 S_k + \alpha_A M_S - \gamma M_A\,; \tag{2.20e}$$

$$\frac{d}{dt}M_B = \beta_B p_B \sum_k k^2 S_k + \alpha_B M_S - \gamma M_B\,; \tag{2.20f}$$

$$\begin{aligned}\frac{d}{dt}M_{SA} = &-\beta_A M_{SA} - \gamma M_{SA} + \gamma(M_A - M_{SA})\\ &+ \beta_A E p_A M_{SS} - \beta_B E p_B M_{SA} - \beta_A E p_A M_{SA}\\ &+ \alpha_A M_{SS} - \alpha_A M_{SA} - \alpha_B M_{SA};\end{aligned} \tag{2.20g}$$

$$\begin{aligned}\frac{d}{dt}M_{SB} = &-\beta_B M_{SB} - \gamma M_{SB} + \gamma(M_B - M_{SB})\\ &+ \beta_B E p_B M_{SS} - \beta_A E p_A M_{SB} - \beta_B E p_B M_{SB}\\ &+ \alpha_B M_{SS} - \alpha_B M_{SB} - \alpha_A M_{SB}\end{aligned} \tag{2.20h}$$

$$\begin{aligned}\frac{d}{dt}M_{AB} = &-2\gamma M_{AB} + \beta_A E p_A M_{SB} + \beta_B E p_B M_{SA}\\ &+ \alpha_B M_{SA} + \alpha_A M_{SB}.\end{aligned} \tag{2.20i}$$

2.1. The homogeneous mixing limit. In the limit of a homogeneous mixing population, i.e., on a complete graph, our model becomes

$$S' = -\lambda_A \frac{AS}{N} - \lambda_B \frac{BS}{N} - \alpha_A S - \alpha_B S + \gamma A + \gamma B, \tag{2.21a}$$

$$A' = \lambda_A \frac{AS}{N} + \alpha_A S - \gamma A, \tag{2.21b}$$

$$B' = \lambda_B \frac{BS}{N} + \alpha_B S - \gamma B. \tag{2.21c}$$

Here S, A, and B are the number of potential buyers, the buyers of product A, and the buyers of product B, and $S+A+B=N$ is the population size; $\lambda_A = \beta_A(N-1)$ and $\lambda_B = \beta_B(N-1)$ is the per capita word-of-mouth transmission rate. We need to assume that λ_A and λ_B have a limit as $N \to \infty$. The edge dynamics becomes simple in this limit, because $M_{SA} = SA$. In Appendix A we show the proof for a single product model, but the same technique is easily extended to our full model (2.20).

3. Comparison with stochastic simulations

Our model aims to describe, in simplified terms, the following stochastic marketing process. On a random network, each node is labelled by its buyer status: a potential buyer (or susceptible, denoted as S), a product A buyer (denoted as A), or a product B buyer (denoted as B). The status transition in the absence of word-of-mouth contact is described by the following matrix (which rows and columns are in the order of S, A, and B).

$$T = \begin{bmatrix} -\alpha_A - \alpha_B & \alpha_A & \alpha_B \\ \gamma & -\gamma & 0 \\ \gamma & 0 & -\gamma \end{bmatrix}.$$

The first row here represents the conversion from an S node to either A or B by advertisement, the second row represents that an A-node recovers (i.e., loses the buyer status) and becomes an S node. That last row represents the recovery process of a B node. Along each edge connected to an A source node (or a B source node), contact events occur as a Poisson process with rate β_A (or β_B). Upon contact, if the target node is an S node, the target is converted to A (or B). Otherwise the contact event is ignored. The contact process along an edge stops when the source node recovers.

We use the Gillespie method [3,4] to simulate this stochastic process. A realization of the random network is generated from the configuration model as described in [15]. This model needs a given degree distribution and a given number of nodes. For each node, a random number k is drawn from the distribution and assigned as its degree. Then k stubs (half edges) are attached to the node. Two random stubs that are not from the same node or nodes that are already neighbours are randomly selected and connected to form an edge. This connection process is repeated until no such edges can be formed. Leftover stubs are discarded.

Each node is initially labelled S. On a realization of the random network, 100 simulations are conducted, and their averages are compared with the solutions of our differential equation model. The degree distribution in our model is computed from the randomly generated network.

Our experiments are for Poisson (also known as Erdős–Rényi) networks ($P_k = \frac{\lambda^k}{k!} e^{-\lambda}$) where the expectation and variance of connections are both $\lambda > 0$. These arise in the limit $N \to \infty$ if each pair of nodes has identical probability $p = \lambda/N$ of being connected. Other relevant examples are so-called scale-free networks with $P_k = Ck^{-r}$. (C is a normalizing constant). These latter networks are popular in epidemiological and social applications and arise when nodes are added to a growing network and new edges attach to an existing node with a probability proportional to its degree: more popular nodes attract more connections. Depending on the parameter r and the size of N, the expected number of edges and their variance can be huge. For example, with $r = 2$ and $N = \infty$ the expected value $\langle k \rangle$ is infinite!

Figure 3.1 shows that, on randomly generated Poisson and scale-free networks with $N = 10^4$ nodes, our model is in excellent agreement with the ensemble average of the stochastic simulations. The minimum degree of the scale-free network was chosen to be 2 to avoid isolated components.

Figure 3.2 shows that, on a scale-free network with exponent $r = 3$, with average degree $\langle k \rangle = 3.2$ and variance $\mathrm{Var}[k] = 17$, the stochastic simulations behave similarly to their mean shown in Figure 3.1. However, on a scale-free network with exponent $r = 2$, with $\langle k \rangle = 10$ and $\mathrm{Var}[k] = 3144$, the stochastic simulations show huge variances around their mean. We conclude that, on such a scale-free network, the ensemble mean is not a reliable representation of the behaviour of the stochastic marketing process. In fact, our simulations suggest that the large variance of such a network makes reliable predictions of market penetration by products almost impossible.

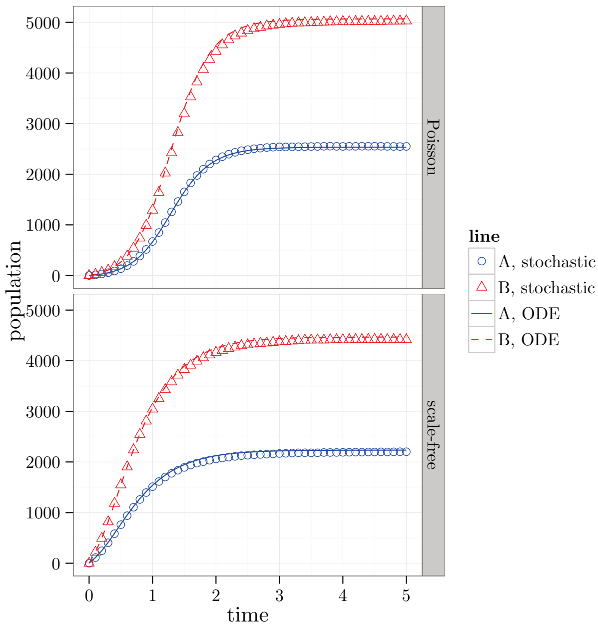

Fig. 3.1. *The comparison of the ensemble average of 100 stochastic simulations (symbols) and the solution of the ODE model (lines) on a Poisson network with average degree* $\langle k \rangle = 5$*, and a scale-free network with exponent* $r = 3$ *and minimum degree 2 (to avoid isolated pairs). Both networks have* 10^4 *nodes. The parameters are* $\beta_A = \beta_B = 1$*,* $\gamma = 1$*,* $\alpha_A = 0.01$*,* $\alpha_B = 0.02$*.*

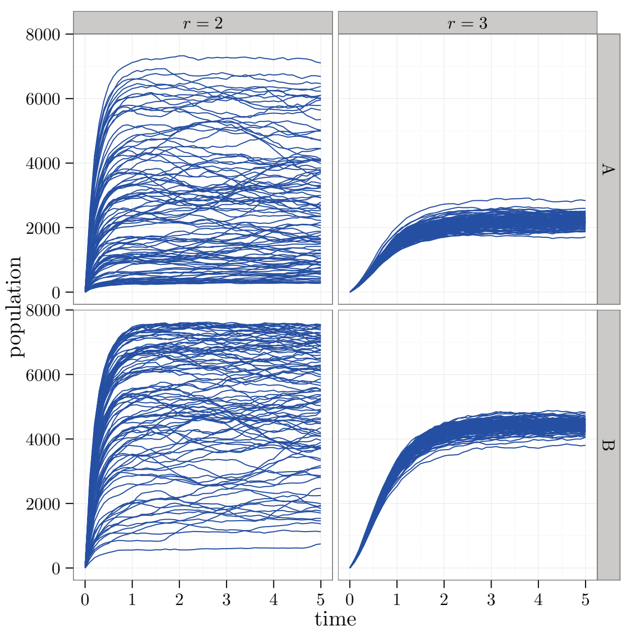

Fig. 3.2. *The comparison of 100 stochastic simulations on two scale-free networks. Both networks contain* 10^4 *nodes, and have minimum degree 2 to avoid isolated components. The exponents of the power-law degree distributions are* $r = 2$ *and* $r = 3$*, respectively. The marketing parameters are the same as in Figure 3.1.*

4. Market share

4.1. Dependence on marketing parameters. We begin with a remark on the elementary case where there is no word-of-mouth effect. In this case, the word-of-mouth network becomes irrelevant, and our model reduces to the much simpler linear system

$$\frac{d}{dt}S = -(\alpha_A+\alpha_B)S+\gamma(A+B) \tag{4.1}$$

$$\frac{d}{dt}A = \alpha_A S-\gamma A \tag{4.2}$$

$$\frac{d}{dt}B = \alpha_B S-\gamma B \tag{4.3}$$

and it is straightforward to compute the asymptotic values of S,A,B as

$$S_\infty = \frac{\gamma N}{\alpha_A+\alpha_B+\gamma}, \quad A_\infty = \frac{\alpha_A N}{\alpha_A+\alpha_B+\gamma}, \quad B_\infty = \frac{\alpha_B N}{\alpha_A+\alpha_B+\gamma}. \tag{4.4}$$

These are stable steady solutions of the simplified model. Let $q=\alpha_B/\alpha_A$, then $B(\infty)=qA(\infty)$. In fact, since $\frac{d}{dt}(qA-B)=-\gamma(qA-B)$, starting with $A(0)=B(0)=0$, $qA(t)=B(t)$ for all $t\geq 0$. We have similar results for the network Bass model [8], provided that $\beta_A=\beta_B$. Given that the recovery rate γ is the same for both products, our model (2.20) show the same behavior.

PROPOSITION 4.1. *If $\beta_A=\beta_B$ and $\alpha_B=q\alpha_A$, then the solution of the full model with initial conditions $B(0)=A(0)=0$ satisfies $B(t)=qA(t)$ and $M_{SB}(t)=qM_{SA}(t)$.*

The figures in the previous section show an example for this case ($q=2$). The proof of the proposition is a tedious but elementary calculation, which is given in Appendix B.

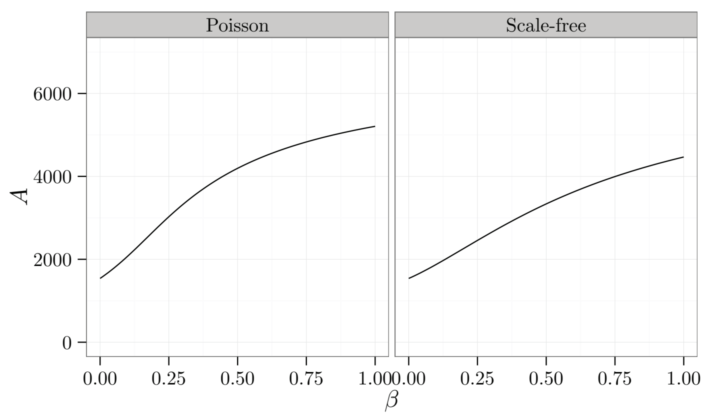

FIG. 4.1. *The dependency of the equilibrium market share of product A on β_A. Here $\beta:=\beta_A=\beta_B$, $\alpha_A=0.2$, $\alpha_B=0.1$, $\gamma=1$. The network is the same used in Figure 3.1.*

The formulas (4.4) provide a baseline for studying the effect of word-of-mouth recruitment on the equilibrium market share. For simplicity, we keep $\beta_A=\beta_B$, and so knowledge of $A(\infty)$ yields both $S(\infty)$ and $B(\infty)$ by Proposition 4.1 and

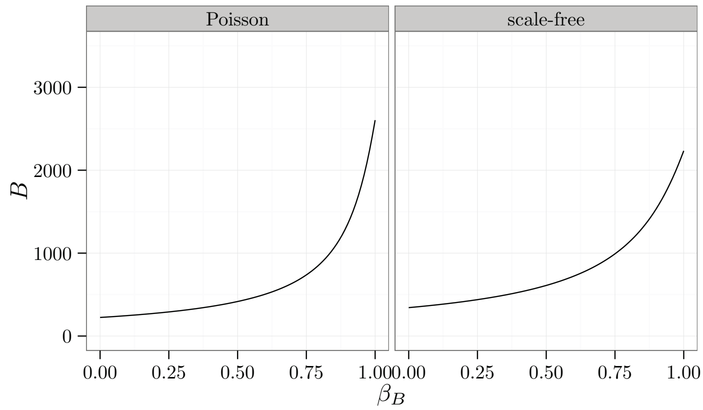

FIG. 4.2. *The dependency of the equilibrium market share of product B on β_B, with $\beta_A = 1$. The networks, α_A, α_B, and γ are the same as in Figure 4.1.*

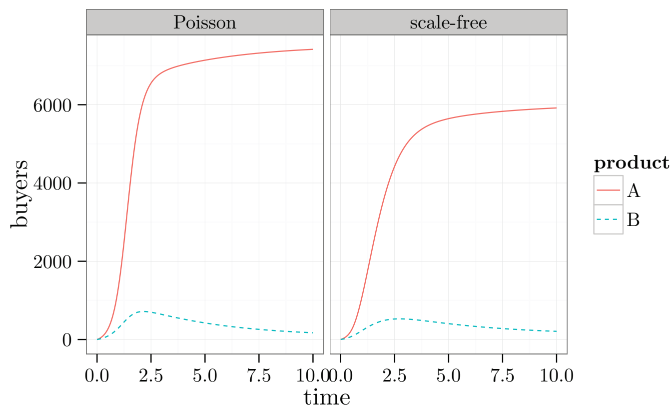

FIG. 4.3. *The time evolution of market shares of products A and B, with $\beta_A = 1$, $\beta_B = 0.7$, $\alpha_A = 0.02$, and $\alpha_B = 0.01$. The networks are the same as in Figure 3.1.*

$S(\infty) + A(\infty) + B(\infty) = N$. Figure 4.1 shows the influence of word-of-mouth recruitment rates with $\beta_A = \beta_B$. Not surprisingly, word-of-mouth recruitment is a significant factor in gaining market share.

Figure 4.2 shows the equilibrium market share of product B as a function of β_B, with a fixed β_A. We observe that the equilibrium market share of B starts with negligible values if $\beta_B = 0$, then increases slowly with β_B, and finally increases quickly as β_B grows beyond 0.5. However, because of the presence of advertisements there is no threshold phenomenon.

When β_B is small, even though the equilibrium market share of product B is small, it can still initially catch a significant market share, but then be out-competed by A. Eventually its market share declines to a small value as is illustrated in Figure 4.3. This is reminiscent of the competitions between Betamax and VHS video tape formats, and between HD DVD and Blu-ray Disc formats, which showed a similar pattern.

Figure 4.7 shows $p_S(k)$ for the model presented by Li *et al.* [8]. It is clear that

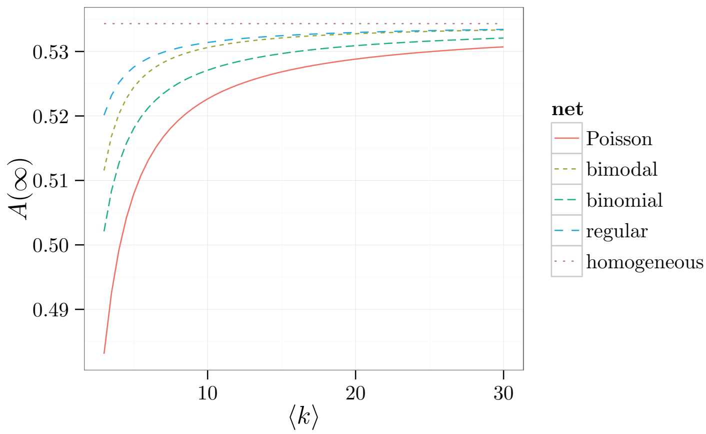

FIG. 4.4. *The final market share of Product A as a function of the average degree $\langle k \rangle$, with a fixed per node transmission rate $\beta_A \langle k \rangle = \beta_B \langle k \rangle = 5$. The variance of the degree distribution of the binomial networks are half of that of the Poisson networks with the same $\langle k \rangle$, the variances of the bimodal network are 1, and those of the regular network are 0. The parameters α_A, α_B, and γ are identical to those in Figure 3.1.*

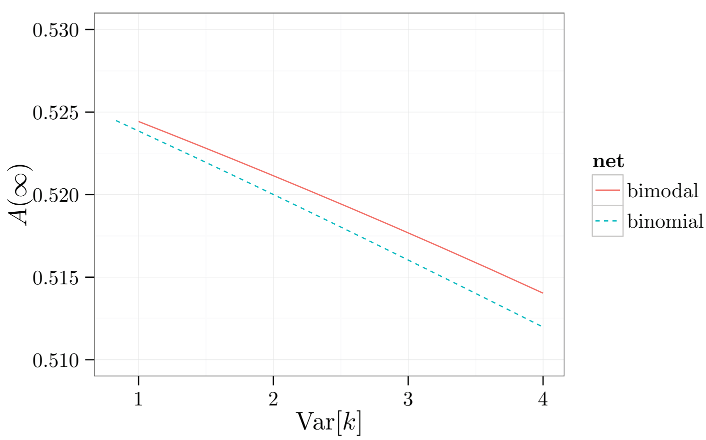

FIG. 4.5. *The final market share of Product A as a function of the variance of the degree distribution on binomial networks and bimodal networks with identical average degree. The marketing parameters are the same as in Figure 3.1.*

the edge chaos assumption is valid for all time in this model. Thus, repeat purchase will gradually introduce correlation between the degree of a susceptible node and the probability that its neighbours are susceptible.

4.2. Dependence on the network degree distribution. To illustrate the dependence of the final market share on the underlying social network, we fix the average per node transmission rate $\lambda_A = \beta_A \langle k \rangle$ and $\lambda_B = \beta_B \langle k \rangle$ and vary the degree distribution. Fixing λ_A and λ_B guarantees that, in the homogeneous mixing limit, the final market shares on networks approach those in a homogenous population. We consider the following examples: Poisson (Erdős–Rényi) random networks with average degree

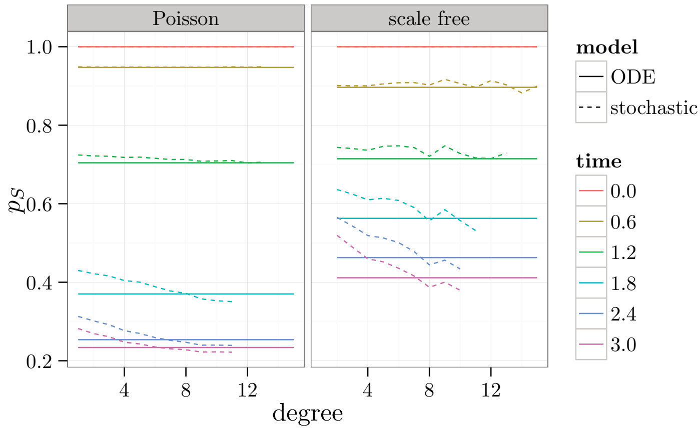

FIG. 4.6. *The time evolution of* p_S *as a function of the degree of the source node, for the simulations shown in Figure 3.1. Each curve represents* $p_S(k)$ *at a specific time.*

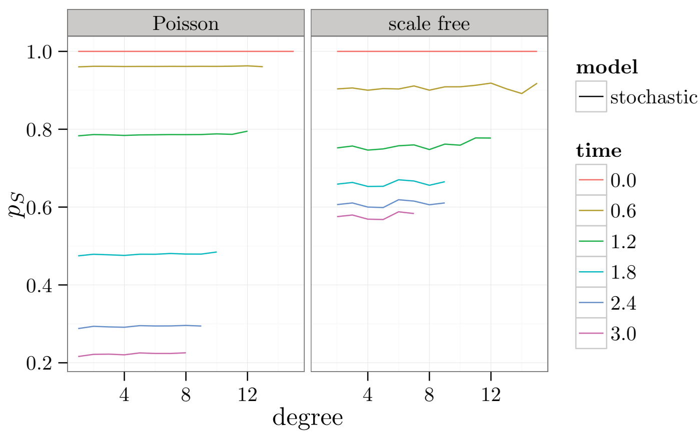

FIG. 4.7. *The time evolution of* p_S *as a function of the degree of the source node, for the Li* et al. *[8] model. Each curve represents* $p_S(k)$ *at a specific time. Marketing parameters and the networks are identical to Figure 3.1, except that* $\gamma = 0$.

$\langle k \rangle$, binomial random networks with half the variance of the Poisson networks for the same average degree, bimodal networks with degrees $\langle k \rangle \pm 1$ (and thus with variance 1), and regular networks on which all nodes have the same degree $\langle k \rangle$. The final market shares $A(\infty)$ of the solutions of our model (2.20) on these networks are compared to the final market share of the homogeneous model (2.21) in Figure 4.4 (we do not present the final market shares for product B; the figures would look similar.) It is visible that the final market shares on social networks increase and approach that in a homogeneously mixed population as the average degree $\langle k \rangle$ becomes large. This shows that the network structure slows the product diffusion. Furthermore, networks with a larger variance result in a smaller final market share. To illustrate this further, the final market share on binomial networks and bimodal networks with identical average degree are shown in Figure 4.5. It shows that the final market share indeed decreases with the variance.

In addition, the thrid and higher moments of the degree distribution also have a small influence on the final market share.

5. On the validity of the edge chaos assumption

Our model, like the earlier model presented by Li *et al.* [8], is closed by using the edge chaos assumption. In this section, we perform a check of the validity of this assumption. Since the random network is fully characterized by its degree distribution, we scrutinize whether the probability that a neighbour of a susceptible is susceptible (or a buyer) depends on the degree of the source node, i.e., whether $p_S = M_{SS}/M_S$, $p_A = M_{SA}/M_S$, and $p_B = M_{SB}/M_S$ will in microscopic simulations depend on the degree of the source node (and hence contradict the edge chaos assumption: the independence of k of these terms is a necessary, but not sufficient conditions for edge chaos). To do so, we compute the average fraction of susceptible neighbours across all susceptible nodes with a degree k, namely, $p_S(k) = M_{SS}^{(k)}/(kS_k)$ where $M_{SS}^{(k)}$ is the total number of susceptible neighbours of degree-k susceptible nodes. The time evolution of $p_S(k)$ is shown in Figure 4.6. It appears that this test is consistent with the edge chaos assumption during the growth phase of the product diffusion, i.e., $p_S(k)$ is independent of k. But as the diffusion process reaches equilibrium, $p_S(k)$ decreases with the degree k, i.e., there is an anti-correlation between the degree of a susceptible source node and the probability that its neighbour is susceptible.

6. Conclusion, remarks, and outlook

We have introduced a model for SIS-market competitions on random graphs based on node and edge dynamics. An "edge chaos" assumption is used to close the dynamic equations for the fractions of edge types. The model was tested for Poisson and scale free random networks and showed very good agreement with averaged microscopic simulations. Both analytical and numerical studies on equilibrium market shares were conducted, including some predictions of final market share as a function of advertising and word-of-mouth recruitment efforts, and as a function of network parameters. We also showed that the model is consistent with the homogeneous limit case arising in the limit $N \to \infty$ for a complete graph.

A weakness of our model is that it makes predictions about averages but not about fluctuations, and this is where the underlying type of random network may really make a difference. The numerical experiments regarding product diffusion on scale-free networks are a case in point. Further, our assumption of edge chaos is, as seen above, only an approximation, and its viability will depend on the network type.

In our model (2.20), we assume that the recovery rates for the two products are identical. In reality, they may be different. It is straightforward to incorporate different recovery rates in our model. It may also be interesting to consider the word-of-mouth transmission rates β_A and β_B fading with time, because the interest to talk about a new product may disappear. Even though β_A and β_B are assumed constants, they represent the average transmission rates in the network. A harder but interesting question is how the variances of transmission rates affect the dynamics. More profoundly, market dynamics may change the network itself, which may invalidate our edge chaos assumption.

As presented, our model assumes constant advertising rates. However, the number of potential buyers decreases with time. Thus advertisements becomes less efficient over time. With a limited advertising budget, a company may choose to advertise intensely to gain initial market share, then ramp down. Our model can likely be useful in finding optimal strategies in allocating advertisement efforts.

Appendix A. Proof of the homogeneous mixing limit. In the limit of a homogeneous mixing population, i.e., on a complete graph, the fraction of edges belonging to M_{SA}, namely $x=\frac{M_{SA}}{N(N-1)}$, should equal to $y=\frac{SA}{N(N-1)}$. To show that our model gives this limit, we first derive the dynamics of the fraction x. On a complete graph $M_S=(N-1)S$, $M_A=(N-1)A$, and $E=(N-1)$. For simplicity we only study a single product model, i.e., all the dependent variable referring to the product B, namely B, M_{SB}, M_B, M_{AB}, α_B, and β_B, are 0. Then, from (2.13)

$$\begin{aligned} x' &= \frac{M'_{SA}}{N(N-1)} \\ &= -\beta_A x - \gamma x + \gamma(\frac{A}{N} - x) + \alpha_A(\frac{S}{N} - x) - \alpha_A x \\ &\quad + \beta_A \frac{N(N-1)x}{S}[\frac{S}{N} - x] - \beta_A \frac{x^2 N(N-1)}{S} \\ &= \beta_A(N-2)x + \gamma(\frac{A}{N} - 2x) + \alpha_A(\frac{S}{N} - 2x) - 2\beta_A(N-1)\frac{x^2}{S/N}. \end{aligned}$$

The seven terms on the right correspond, in order, to the terms labelled 1,2,3,8,9,5,7 in (2.13). For examples, Term 4 there becomes $\frac{M_{AA}}{N(N-1)} = \frac{A(N-1)-MSA}{N(N-1)} = \frac{A}{N} - x$.

Now, let $\lambda_A = \lim_{N\to\infty} \beta_A N$ be constant. We convert to fractions $s := S/N, a = A/N$ then, as $N \to \infty$,

$$x' = \lambda_A x - 2\lambda_A \frac{x^2}{s} + \gamma(a - 2x) + \alpha_A(s - 2x).$$

On the other hand,

$$\begin{aligned} y' &= \frac{S'A + SA'}{N(N-1)} \\ &= \frac{-\beta_A M_{SA} A - \alpha_A SA + \gamma A(N-S) + \beta_A M_{SA} S + \alpha_A S(N-A) - \gamma SA}{N(N-1)} \\ &= \beta_A N x \frac{S-A}{N} + \alpha_A(\frac{S}{N-1} - 2y) + \gamma(\frac{A}{N-1} - 2y). \end{aligned}$$

In the limit $N \to \infty$,

$$y' = \lambda_A x - 2\lambda_A x a + \alpha_A(s - 2y) + \gamma(a - 2y)$$

Thus, using that in the limit $y = sa$,

$$\begin{aligned} (x-y)' &= -2\lambda_A x(\frac{x}{s} - a) - 2\alpha_A(x-y) - 2\gamma(x-y) \\ &= -2\frac{\lambda_A x}{s}(x-y) - 2\alpha_A(x-y) - 2\gamma(x-y). \end{aligned}$$

Given that $x(0) = y(0) = 0$, this guarantees that $x = y$ for all time t. For the fully connected model the equation (2.13) are redundant, as in this case $M_{SA} = SA$. In particular, as $N \to \infty$,

$$p_A = \frac{M_{SA}}{(N-1)S} = \frac{A}{N} \to a$$

and

$$S' = S'_{N-1} = -\lambda \frac{A}{N} S - \alpha_A S + \gamma A.$$

Similarly, we can show that the full model has (2.21) as the limit.

Appendix B. Proof of Proposition 4.1. Multiply equation (2.6) by $q = \alpha_B/\alpha_A$ and subtract from equation (2.7). With the assumption $\beta_A = \beta_B = \beta$, this gives

$$\frac{d}{dt}(B - qA) = \beta(M_{SB} - qM_{SA}) - \gamma(B - qA).$$

Thus, provided that $M_{SB} = qM_{SA}$ and $B(0) - qA(0) = 0$, this implies that $B - qA = 0$ is a solution, which is the result of the proposition.

To show that $M_{SB} - qM_{SA} = 0$, multiply equation (2.13) by q and subtract from equation (2.14). This gives

$$\begin{aligned}\frac{d}{dt}(M_{SB} - qM_{SA}) = & -\beta(M_{SB} - qM_{SA}) - \gamma(M_{SB} - qM_{SA}) \\ & + \gamma(M_B - M_{SB} - qM_A - qM_{SA}) \\ & + \beta \frac{EM_{SS}}{M_S}(M_{SB} - qM_{SA}) \\ & - \beta \frac{E}{M_S}(M_{SB} - M_{SA})(M_{SB} - qM_{SA}) \\ & + (\alpha_B + \alpha_A)(M_{SB} - qM_{SA}).\end{aligned} \tag{B.1}$$

Again, provided that $M_{SB}(0) - qM_{SA}(0) = 0$, which is true given $A(0) = B(0) = 0$, $M_{SB} - qM_{SA} = 0$ is a solution provided $M_B = qM_A$.

Similarly, multiply equation (2.8) by q and subtract from equation (2.9). This gives

$$\frac{d}{dt}(M_B - qM_A) = \beta(E+1)(M_{SB} - qM_{SA}) - \gamma(M_B - qM_A). \tag{B.2}$$

Note that $M_{SB} - qM_{SA} = 0$ and $M_B - qM_A = 0$ are a solution of the system (B.1) and (B.2). Thus, with the initial condition $A(0) = B(0) = 0$, $M_{SB} - qM_{SA} = 0$, and $M_B - qM_A = 0$ hold for all time $t \geq 0$. Thus, Proposition 4.1 holds.

REFERENCES

[1] F.M. Bass, *A simultaneous equation regression study of advertising and sales of cigarettes*, J. of Marketing Research, 6, 291–300, 1969.

[2] J.A. Dodson Jr and E. Muller, *Models of new product diffusion through advertising and word-of-mouth*, Manage. Sci., 24, 1568–1578, 1978.

[3] D.T. Gillespie, *A general method for numerically simulating the stochastic time evolution of coupled chemical reactions*, J. Comput. Phys., 22, 403–434, 1976.

[4] D.T. Gillespie, *Exact stochastic simulation of coupled chemical reactions*, J. Phys. Chem., 81, 2340–2361, 1977.

[5] T. House and M.J. Keeling, *Insights from unifying modern approximations to infections on networks*, Physical Review D Particles Fields, 84(12), 5241–5259, 2011.

[6] W.A. Kamakura and S.K. Ealasubramanian, *Long-term forecasting with innovation diffusion models: The impact of replacement purchases*, Journal of Forecasting, 6, 1–19, 1987.

[7] M.J. Keeling and K.T.D. Eames, *Networks and epidemic models*, J. R. Soc. Interface, 2, 295–307, 2005.

[8] M. Li, R. Edwards, R. Illner, and J. Ma, *Marketing new products: Bass models on random graphs*, Commun. Math. Sci., 13, 497–509, 2015.

[9] G.L. Lilien, A.G. Rao, and S. Kalish, *Bayesian estimation and control of detailing effort in a repeat purchase diffusion environment*, Management Science, 27, 493–506, 1981.

[10] J. Lindquist, J. Ma, P. van den Driessche, and F.H. Willeboordse, *Effective degree network disease models*, J. Math. Biol., 62, 143–164, 2010.

[11] V. Mahajan, E. Muller, and F.M. Bass, *New product diffusion models in marketing: A review and directions for research*, J. Marketing, 1–26, 1990.

[12] V. Mahajan, E. Muller, and Y. Wind, *New-product diffusion models*, Springer Science & Business Media, 19(2), 327–328, 2003.

[13] V. Mahajan, J. Wind, and S. Sharma, *An approach to repeat-purchase diffusion analysis*, Wharton School University of Pennsylvania, 1982.

[14] J.C. Miller, *A note on a paper by Erik Volz: SIR dynamics in random networks*, J. Math. Biol., 62349–358, 2011.

[15] M. Molloy and B. Reed, *A critical point for random graphs with a given degree sequence*, Random Struct. Algor., 6(2–3), 161–180, 1995.

[16] M.E.J. Newman, *Spread of epidemic disease on networks*, Phys. Rev. E, 66(1), 016128, 2002.

[17] J. Olson and S. Choi, *A product diffusion model incorporating repeat purchases*, Technological Forecasting and Social Change, 27, 385–397, 1985.

[18] R. Pastor-Satorras and A. Vespignani, *Epidemic dynamics in finite size scale-free networks*, Phys. Rev. E, 65, 035108(R), 2002.

[19] R. Peres, E. Muller, and V. Mahajan, *Innovation diffusion and new product growth models: a critical review and research directions*, Int. J. of Research in Marketing, 27, 91–106, 2010.

[20] E.M. Volz, *SIR dynamics in random networks with heterogeneous connectivity*, J. Math. Biol., 56, 293–310, 2008.