Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Diffusion in Random Networks: Impact of Degree Distribution

Vahideh Manshadi, Sidhant Misra, Scott Rodilitz

To cite this article:
Vahideh Manshadi, Sidhant Misra, Scott Rodilitz (2020) Diffusion in Random Networks: Impact of Degree Distribution. Operations Research

Published online in Articles in Advance 04 May 2020

. https://doi.org/10.1287/opre.2019.1945

http://pubsonline.informs.org/journal/opre

OPERATIONS RESEARCH
Articles in Advance, pp. 1–20
ISSN 0030-364X (print), ISSN 1526-5463 (online)

Methods

Diffusion in Random Networks: Impact of Degree Distribution

Vahideh Manshadi,[a] Sidhant Misra,[b] Scott Rodilitz[a]

[a] Yale School of Management, New Haven, Connecticut 06511; [b] Los Alamos National Laboratory, Santa Fe, New Mexico 87545

Contact: vahideh.manshadi@yale.edu, https://orcid.org/0000-0001-9103-7797 (VM); sidhant@lanl.gov (SM); scott.rodilitz@yale.edu, https://orcid.org/0000-0002-1343-7901 (SR)

Received: September 3, 2018
Revised: June 12, 2019
Accepted: August 21, 2019
Published Online in Articles in Advance: May 4, 2020

Subject Classification: Probability: diffusion, stochastic model applications; networks/graphs: stochastic
Area of Review: Stochastic Models

https://doi.org/10.1287/opre.2019.1945

Abstract. Motivated by viral marketing on social networks, we study the diffusion process of a new product on a network where each agent is connected to a random subset of others. The number of contacts (i.e., degree) varies across agents, and the firm knows the degree of each agent. Further, the firm can seed a fraction of the population and incurs a fixed cost per contact. Under any bounded degree distribution and for any target adoption proportion, we compute both the cost and the time it takes to reach the target in the limit of network size. Our characterization of the diffusion process in such a general setting is the first of its kind for a problem that is generally deemed intractable and solved using approximation methods such as mean-field. Our solution indicates that the degree distribution impacts the diffusion process even beyond its first and second moments. Using our limit results, we conduct comparative statics on degree distribution and uncover a trade-off between cost-efficiency and fast growth. Fixing the average degree, a minimum-variance degree distribution incurs the minimum cost to reach any adoption proportion. On the other hand, higher variance results in faster growth for low or moderately high target adoption proportions, but it incurs higher cost. This trade-off arises partly because of an endogenous effect of diffusion on the distribution of adopters' neighbors: as the diffusion progresses, adopters become more likely to be connected to other adopters. This highlights the benefit of our exact analysis compared to mean-field approximation methods that rely on perfect mixing of adopters and non-adopters (i.e., there is no change in the distribution of adopters' neighbors with the progress of diffusion). Further, we study the impact of the degree distribution on optimal seeding strategies for a given seeding budget. Somewhat surprisingly, we show that to minimize cost, it is optimal to seed low-degree agents. Even if the objective is to minimize time, for certain regimes, the optimal seeding strategy is a mixture of low- and high-degree agents.

Supplemental Material: The online appendix is available at https://doi.org/10.1287/opre.2019.1945.

Keywords: diffusion processes • seeding • network effects

1. Introduction

Word-of-mouth (WoM) is known as a powerful marketing force, as numerous empirical studies reveal that consumers' purchasing decisions are based on the advice of those in their social networks rather than on direct advertising. A recent international survey by Nielsen reports that 92% of consumers around the world count on recommendations from friends and family more than all other forms of advertising (Nielsen 2012). With technological advances in online communications that enable consumers to easily share their experience with their "friends," the effect and importance of WoM has only grown.[1] In this new era, marketers not only harness the power of WoM, but also improve its efficacy by targeting consumers based on the wealth of information available about their online activities. In particular, marketers can utilize information on connections among consumers (1) to predict the diffusion trajectory (for both time and cost), and (2) to devise effective seeding strategies to impact the trajectory. In this work, we provide a theoretical framework to study the diffusion process for a general class of network models and drive insights about the impact of heterogeneity in the degree of connections on the cost and speed of diffusion as well as on optimal seeding strategies.

To this end, we study a diffusion process of a new product that spreads through the contacts that adopters make with their neighbors. In particular, we assume that an adopter makes contact with each of her neighbors according to an independent Poisson process with rate γ. We assume that the network underlying the connections is a random network with a given degree distribution. This general class of network models has been extensively used in the study of social networks (Jackson and Yariv 2005, Jackson 2010, Shakkottai and Johari 2010, Dover et al. 2012, Fainmesser and Galeotti 2016, Ajorlou et al. 2017),

and serves as the network model when the firm's knowledge about the pattern of connections is limited to the degree of each agent, rather than having access to the identity of every neighbor of an agent.

In our setting, the marketer incurs a fixed cost c for each contact by an adopter. Further, we assume that the marketer has a budget for *seeding*. In particular, prior to the adoption process, he can directly contact a fraction $q > 0$ of agents, who become adopters. The marketer decides who to seed with the goal of minimizing the total cost or the total time to reach his target proportion of adopters. Targeting agents based on social network information is a common practice. It has been studied in operations and network economics literature in monopolistic settings (Kempe et al. 2003, Lim et al. 2015) as well as in competitive settings (Fazeli and Jadbabaie 2012, Goyal et al. 2014, Bimpikis et al. 2016), mainly under the assumption that the marketer has complete information about the network structure. However, in our setting the firm can only target based on the degree of an agent, since he does not have access to more detailed information. Seeding agents based on their degrees seems to be more practical (as it requires the firm to acquire much less information), and empirically it has been shown to be effective (Hinz et al. 2011).

1.1. Summary of Our Contributions

In the above setting,

1. We compute the cost and time to reach any adoption proportion $s \in (q, 1)$, for any general bounded degree distribution in the limit as the number of agents grows (Theorems 4.1 and 4.2). To the best of our knowledge, this is the first exact characterization of the diffusion process for such a general class of degree distributions. The other exact characterizations are for the special cases of a complete network (Massoulie and Draief 2010), which is equivalent to the Bass model (Bass 1969), and a one-dimensional grid (i.e., a cycle) (Fibich and Gibori 2010).

2. Using our exact characterization, we study the impact of degree distribution on the cost and time to reach any proportion $s \in (q, 1)$ and demonstrate a trade-off between contact cost (which is proportional to the total number of contacts) and speed. In particular, we show that lower variability in the degree distribution results in lower cost. Fixing the average degree $k \in \mathbb{N}$, the most cost-efficient network (to reach any adoption fraction s) is the k-regular network (Proposition 5.1). The impact of degree distribution on timing is more involved, as it depends on the target adoption level (i.e., s) as well as higher moments of the degree distribution (see Table 1). However, our numerical analysis suggests that unless the target level is very high (e.g., $s = 0.9$), higher variance improves the speed.

Table 1. Time Required to Reach $s = 0.5$ (Top) and $s = 0.95$ (Bottom)

		Skewness						
		–0.45	–0.30	–0.15	0	0.15	0.30	0.45
Variance	2	0.977	0.977	0.977	0.978	0.978	0.978	0.978
	4	0.956	0.956	0.956	0.956	0.957	0.957	0.957
	6	0.936	0.936	0.936	0.937	0.937	0.937	0.937
	8	0.917	0.917	0.917	0.918	0.918	0.918	0.919
	10	0.899	0.899	0.900	0.900	0.901	0.901	0.901

		Skewness						
		–0.45	–0.30	–0.15	0	0.15	0.30	0.45
Variance	2	1.004	1.004	1.003	1.003	1.002	1.002	1.001
	4	1.017	1.015	1.014	1.012	1.010	1.008	1.007
	6	1.037	1.033	1.029	1.025	1.021	1.017	1.014
	8	1.063	1.055	1.048	1.041	1.035	1.028	1.022
	10	1.093	1.082	1.071	1.060	1.050	1.040	1.031

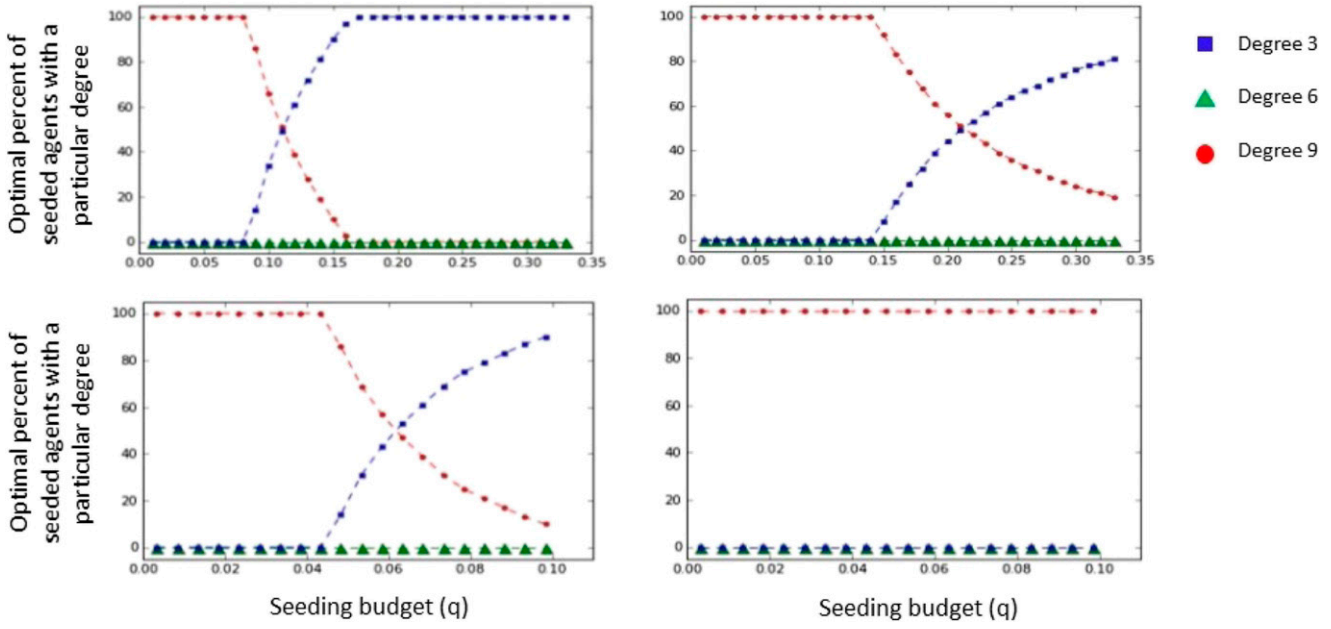
Notes. In both cases, we have $k_{\text{avg}} = 8$, $q = 0.01$, and the times are normalized by the diffusion time for the 8-regular network.

3. We also study the problem of optimal seeding that the firm faces for a network with a given degree distribution. Somewhat contrary to the general wisdom, we show that the optimal strategy does not necessarily entail seeding high-degree agents. In fact, we prove that for the objective of minimizing cost (to achieve any target level of adoption), the optimal strategy is to maximally seed low-degree agents (Proposition 5.5). For the objective of minimizing time, the optimal strategy depends on the target level s and also the seeding budget q. We show that if the target level is not too high, it is optimal to seed the highest-degree agents (Proposition 5.6). However, we present examples illustrating that for larger target levels, the optimal strategy can be to seed a mixture of high- and low-degree agents (Figure 1).

4. In the absence of seeding, diffusion has a very slow start simply because there are not enough adopters to make contacts. We also study diffusion in such a setting by assuming that the diffusion starts with a single (randomly selected) agent. We characterize the cost and time it takes to reach $\alpha \log(n)$ adopters, where n is the number of agents in the network and $\alpha > 0$ is a constant (Theorem 4.5). We call this phase of diffusion the *early adoption regime.* We provide comparative statics with respect to the degree distribution in the early adoption regime, and we show that the cost is independent of degree distribution. Further, the time to diffuse to $\alpha \log(n)$ adopters only depends on the first two moments of the degree distribution. Fixing the average degree, the time decreases as variance increases (Corollary 4.6).

We note that many other papers consider the study of diffusion processes in random networks with general degree distributions, using mean-field approximation

Figure 1. (Color online) Percentage of Seeded Agents with a Particular Degree in the Optimal Seeding Strategy (i.e., $100(q_l^*, q_m^*, q_h^*)/q$) as a Function of q

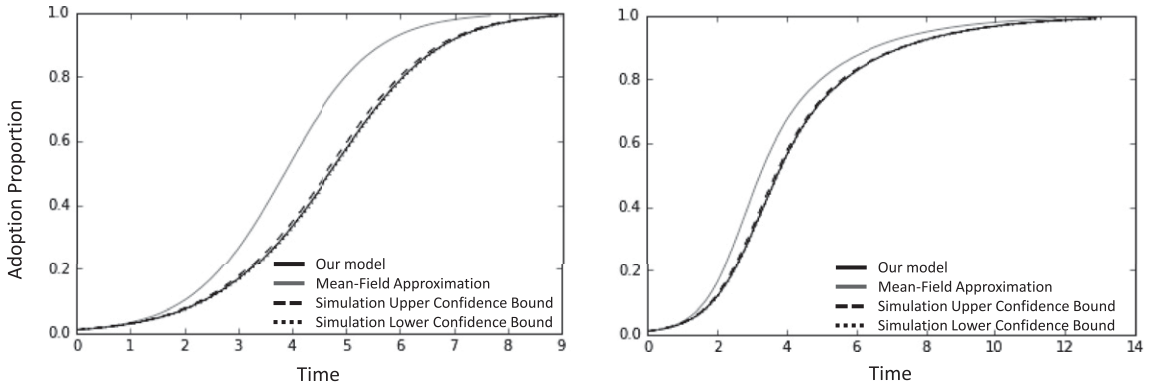


Notes. Top left: $(p_l, p_m, p_h) = (1/3, 1/3, 1/3)$, $s = 0.99$; top right: $(p_l, p_m, p_h) = (1/3, 1/3, 1/3)$, $s = 0.85$; bottom left: $(p_l, p_m, p_h) = (1/10, 1/3, 17/30)$, $s = 0.99$; bottom right: $(p_l, p_m, p_h) = (1/10, 1/3, 17/30)$, $s = 0.85$.

methods without providing error bounds (Jackson and Yariv 2005, 2007; López-Pintado 2008; Shakkottai and Johari 2010; Jackson and López-Pintado 2013). Mean-field approximation is a technique commonly used in statistical physics to approximate dynamic behavior of complex systems (Goldenfeld 2018). In our setting, a mean-field approximation assumes a perfect mixing of adopters and non-adopters regardless of the progression of the diffusion process. For example, consider an agent a with degree 5; at any adoption proportion s, a mean-field approximation assumes that, in expectation, $5 \times s$ of a's neighbors are adopters and $5 \times (1 - s)$ of them are not—that is, the neighbors of a are uniform samples from the entire population regardless of the value of s. We show, however, that such perfect mixing does not hold, as adopters become more likely to be neighbors with other adopters (see Figure 3 and the related discussion in Section 7.3). The impact of this endogenous change in the distribution of adopters' neighbors is twofold: (1) it slows down the diffusion and (2) it increases cost because it results in more "wasteful" contacts between adopters. We present examples to show that at any given time, a mean-field approximation overestimates the proportion of adopters, and the gap can be significantly large (Figure 2 and Table 5).

Another approximation is to assume that the network is a tree—that is, it does not include any cycles.

Figure 2. Time Required to Reach an Adoption Proportion of $0.01 \le s \le 0.99$ (Where $q = 0.01$ Fraction of Agents Selected Uniformly at Random Are Seeded) in Our Model, in a Mean-Field Approximation Given by (31), and in a Simulated Network with 5,000 Agents



Notes. Left: for an 8-regular network; right: for a two-degree network $\{(3, 0.5), (13, 0.5)\}$.

Dover et al. (2012) use such an approximation to uncover the degree distribution based on data from the early stages of the adoption process. This result coincides with our analysis of the early adoption regime. In fact, our analysis in the early adoption regime relies on the *locally tree-like* structure of random networks (Dembo and Montanari 2010) that ensures that the subnetwork of adopters in the early adoption regime is a tree with high probability (see Claim EC.6.1 in Online Appendix EC.6). However, as the diffusion progresses, the subnetwork of adopters will include cycles, and therefore such an approximation will again overestimate the proportion of adopters at any given time.

The above approximations were made for the sake of tractability, because in general form, the diffusion process is a high-dimensional dynamic system where we need to keep track of the status of every agent in the network. However, in our novel analysis, we overcome this challenge by exploiting a special structure of random networks. In particular, it is known that a random network with a given degree distribution can be generated by a simple process called the *configuration model* (Wormald 1999, Jackson 2010). Utilizing the configuration model and the Principle of Deferred Decisions (Mitzenmacher and Upfal 2005), we *interleave* the diffusion and graph generation processes. This idea allows us to approximate diffusion using a low-dimensional system of stochastic differential equations. Through a careful analysis of errors, we prove that in the limit the diffusion almost surely converges to the solution of our system of differential equations (see Section 6).

We now provide brief intuition on our comparative static results and the trade-off between cost and speed. To that end, let us consider a simple degree distribution with two degrees, k_h and k_l where $k_h \geq k_l$, each with proportion 0.5. We fix the average degree to be $k_{\text{avg}} = 0.5k_l + 0.5k_h$, and discuss the effects of increasing k_h (which also implies decreasing k_l). At the beginning, most of the adopters are high degree, as they have more contacts. However, as diffusion progresses, the likelihood of a "wasteful" contact between two adopters increases. The endogenous effect on the distribution of adopters' neighbors (discussed above) exacerbates such a waste. Consequently, when $k_h > k_l$, the contact cost is larger as compared to the cost in a network where $k_h = k_l$. In the early adoption regime, however, there exist no wasteful contacts with high probability (a wasteful contact implies that there exists a cycle in the subnetwork of adopters, which cannot happen if the subnetwork is a tree). This explains why the contact cost is independent of degree distribution in the early adoption regime. In terms of timing, having a larger k_h is helpful when the target fraction is small or moderately large, as higher-degree adopters make more contacts. Even though some of their contacts are wasteful, overall they expedite diffusion. However, when the target proportion is very high, another factor comes into play. Toward the end, most of the agents that have not yet adopted are low-degree ones, and the time to reach these low-degree agents increases as k_l decreases (or equivalently, as k_h increases).

Before concluding this section, we remark that our framework has applications beyond viral marketing. In particular, it is also applicable to settings where a policy maker seeks to spread information or induce adoption in a society through WoM or social interactions (for example, see Banerjee et al. 2013, Akbarpour et al. 2018, and Beaman et al. 2018). A policy maker may also not have access to the identity of every neighbor of an individual in a society. In this paper, we consider a more realistic framework—in which the policy maker is assumed to only know the degrees—and study the diffusion trajectory as well as the policy question of effective seeding.

The rest of the paper is organized as follows. In Section 2, we review the related literature. In Section 3, we introduce the model and define two metrics of interest for the diffusion process. In Sections 4 and 5, we present the main theoretical results and provide comparative statics, respectively. Section 6 outlines the main proof ideas. In Section 7, we discuss a few generalizations of our model. Further, we compare the results of our exact analysis with that of a mean-field approximation and identify significant differences due to the endogenous impact of diffusion on the distribution of adopters' neighbors. Section 8 concludes the paper. For the sake of brevity, the proofs of all statements are deferred to clearly marked appendices.

2. Related Work

Our model of diffusion in a network is a form of the classical Susceptible-Infected (SI) diffusion model, which is in turn a special case of the Susceptible-Infected-Recovered (SIR) model used in epidemiology to model the spread of contagious diseases (Kermack and McKendrick 1927, Durrett 2006) and to study how the pattern of connections impacts diffusion (Anderson and May 1984, Keeling 1999, Draief et al. 2008, Rahmandad and Sterman 2008). Variants of the SI model have also been utilized to model the spread of information or adoption of a new product in online social networks that are usually modeled as random networks (see Dover et al. 2012, Shah and Zaman 2016, Akbarpour et al. 2018, and references therein). The focus of our work is different from the above papers, as we are concerned with exactly characterizing the entire diffusion trajectory and understanding the impact of degree distribution on diffusion without resorting to approximation methods. As mentioned in Section 1, such a characterization for a

general class of random networks has mainly been an unsolved problem, with the exception of a complete network (Massoulie and Draief 2010) and a one-dimensional grid (Fibich and Gibori 2010).

We further discuss similarities and differences between our work and that of Akbarpour et al. (2018). The aforementioned paper also studies random seeding in random networks, albeit in a different setting. In their model, the diffusion starts with a set of seeds and then proceeds as follows: upon adopting at time t, an adopter v has a single chance to contact her neighbors who will independently become adopters with a certain probability at time $t+1$ (in other words, agent v will not make any further contact after t). They mainly focus on Erdős–Rényi network models (which cannot generate a network with an arbitrary degree distribution) and study the number of adopters either at the end of the process or at a fixed time $T > 0$ (which will not be the entire network for most parameters). Because of this structure, their analysis is closely related to the well-known results on percolation in Erdős–Rényi networks and the size of connected components (e.g., results stated in Van Der Hofstad 2016). As mentioned above, our setting is different, as an adopter makes contact with each neighbor (i.e., along each edge) according to a Poisson process (consequently, an agent who adopts at time t asynchronously and repeatedly makes contacts with her neighbors after time t). Thus, in our setting, every agent eventually adopts, and our goal is to characterize the time and cost of reaching any target adoption proportion. This amounts to characterizing the entire diffusion trajectory. Further, we study random networks with an arbitrary degree distribution (under two conditions: that the minimum degree is at least 3 and that the maximum degree is bounded). Because of these differences, our analysis does not follow the standard results (even though we utilize known results on configuration models) and we need to introduce and analyze the Interleaved Graph Generation and Diffusion Process (see Algorithm 1).

In theoretical and empirical studies, various *influence models* have been proposed to describe how a non-adopter adopts a product or an innovation as a result of her neighbors' influence (see Granovetter 1978, Morris 2000, Aral et al. 2013, and references therein). The influence models are then used to design effective seeding strategies, mainly under the assumption that the entire network structure is known (Kempe et al. 2003, Lim et al. 2015). Our model can also be viewed as a simple influence model, in which at any time, the probability that a non-adopter switches to an adopter is proportional to the number of her neighbors that are currently adopters. We complement the literature by studying the seeding problem under this simple influence model in a random network. In our setting, the information available about the network is limited to the degree distribution; thus, the seeding decision simplifies to deciding what proportion of each degree to seed. Hinz et al. (2011) empirically examines the effectiveness of such simple seeding strategies.[2]

Our work is related to the growing literature on designing optimal policies for a firm that faces a network of consumers that impose *local externalities* on each other. The work of Hartline et al. (2008), Candogan et al. (2012), Campbell (2013), Fainmesser and Galeotti (2016), Ajorlou et al. (2017), Cohen and Harsha (2020), and Makhdoumi et al. (2017) is all concerned with pricing policies in the presence of network externalities. Leduc et al. (2017) and Lobel et al. (2017) study the design of referral programs in a network of consumers that have heterogeneous numbers of connections. The work of Sunar et al. (2019) is concerned with jointly optimizing product development, targeting, and pricing of a durable good in a network with local externality effects. Allon and Zhang (2017) examine service differentiation polices in a network of users that have heterogeneous levels of social influence. Momot et al. (2019) and Belloni et al. (2017) study design problems in the presence of negative network effects. WoM can be viewed as a positive network externality and is in fact incorporated into several of the aforementioned papers. We contribute to this line of work by studying diffusion (as a model of WoM) in random networks with a given degree distribution and investigating the impact of degree distribution on the diffusion trajectory and the optimal seeding strategy.

Finally, our work contributes to the literature on using random graph models to study stochastic processes in large, complex networks (Golub and Jackson 2012a, b; Ding et al. 2015; Bayati et al. 2016; Anderson et al. 2017; Ashlagi et al. 2019). In analyzing our diffusion process, we use several ideas and results from the random graph theory literature (e.g. McKay 1985 and Britton et al. 2006). In particular, our analysis of the interleaved graph generation and diffusion process was inspired by Molloy and Reed (1995). Further, we use ideas from Wormald (1995) to approximate discrete stochastic processes with deterministic functions.

3. Model

We represent the social network by a graph $G_n = (V, E)$, where $|V| = n$. Each node $v \in V$ represents an agent in the network; agents v and u are neighbors if $(v, u) \in E$. At time 0, a set of agents chosen by the firm adopts a new product. We will refer to these agents as *seeds*. The new product spreads via the local contacts between neighbors. In particular, if agent v has adopted the product and if v and u are neighbors, then v contacts u according to a Poisson process with rate γ.

Contact across edges emanating from adopters are independent and identically distributed. Suppose agent $v \in V$ becomes an adopter at time t_0. At any time $t_1 > t_0$, if v contacts a neighbor u who has not adopted the product, then we assume that agent u becomes an adopter at time t_1 with probability ρ. Given the thinning property of the Poisson process, without loss of generality, we set $\rho = 1$. If v contacts a neighbor u who has adopted the product, then v will cease contacting u.

We also assume that the firm incurs a cost c each time an adopter makes a contact. We will refer to the sum of these costs as the firm's *contact cost*. Cost c aims to model the firm's opportunity cost that arises when a consumer (i.e., an adopter) participates in the viral marketing campaign and exerts effort by contacting a neighbor in the hope of turning her into an adopter. Too many unsuccessful contacts (i.e., contacting a neighbor who is already an adopter) may frustrate the current adopters and discourage them from participating in the firm's current and future viral marketing campaigns. Therefore, the firm prefers to reduce the number of wasteful contacts and achieve his target level of adoption with a smaller number of contacts. Though the firm does not pay a real cost per contact, our metric of contact cost serves as a concrete measure to model the firm's preference for reducing wasted opportunities (i.e., unsuccessful contacts).

The underlying network is assumed to be sampled uniformly at random from the set of all connected networks with a given degree distribution $\{(p_i, k_i), i \in I\}$; here, $I = \{1, 2, \ldots, |I|\}$ is the set of indices of possible degrees, and p_i is the fraction of agents with degree k_i. The average degree, minimum degree, and maximum degree are defined as follows: $k_{\text{avg}} := \sum_{i\in I} p_i k_i$, $k_{\min} := \min_{i\in I}\{k_i\}$, and $k_{\max} := \max_{i\in I}\{k_i\}$, where we use the symbol ":=" for definitions. Throughout the paper, we assume that $k_{\min} \geq 3$ and $k_{\max}$ is bounded. As we explain in Section 6, our analysis crucially depends on generating the network using the configuration model. The above technical assumptions on the degree distribution ensure that with a nonvanishing probability, the network generated by the configuration model is simple and connected (Wormald 1999).

For such a network, the firm can seed a fraction $q < 1$ of the agents, and his goal is to select the seeds in order to minimize his cost or the time it takes to reach a targeted adoption proportion.[3] As discussed in Section 1, the firm can only differentiate the agents based on their degrees, as he does not have further network information about them. Thus, the firm's decision is to determine the fraction of agents with degree k_i he wants to seed; we denote this fraction by q_i. We formally define a *feasible seeding configuration* as follows:

Definition 3.1 (Feasible Seeding Configuration)**.** We call $\{q_i, i \in I\}$ a feasible seeding configuration if (1) $0 \leq q_i \leq p_i$, for $i \in I$, and (2) $\sum_{i\in I} q_i = q$.

In our adoption process, the number of adopters can only increase over time. Further, since we are only considering connected networks, all agents will adopt the new product after a finite time. Given a number of individuals $qn < y \leq n$, we denote the contact cost required to have y adopters as $W(y)$ and the time required to have y adopters as $T(y)$. Our goal is to characterize $W(y)$ and $T(y)$ for any y, in the limit as the network size n grows to infinity. Then, we use the characterizations to understand the impact of degree distribution on the cost and the time to diffuse to any target adoption level as well as on the optimal seeding strategy.

A few remarks are in order: (1) Fixing y, both $W(y)$ and $T(y)$ are random variables. Therefore, our limit results take the form of convergence of these random variables (or a scaled version of them) almost surely—which is denoted by symbol $\overset{a.s.}{\to}$ as used in Theorems 4.1 and 4.2—or in probability—denoted by symbol $\overset{p}{\to}$ as used in Theorem 4.5. (2) $W(y)$ and $T(y)$ also depend on the degree distribution and seeding configuration. Therefore, a more precise notation would be $W(y; \{(p_i, k_i), i \in I\}, \{q_i, i \in I\})$ and $T(y; \{(p_i, k_i), i \in I\}, \{q_i, i \in I\})$. However, for ease of notation, we simply write them as $W(y)$ and $T(y)$. (3) In our characterizations of $W(y)$ and $T(y)$, presented in Section 4, the seeding configuration is assumed to be given. In other words, Theorems 4.1 and 4.2 hold for any feasible seeding configuration. In fact, the only section concerned with optimizing over the seeding configuration is Section 5.2.

4. Main Results

In this section, we establish limit results for both the contact cost and the time needed to reach y adopters. In Section 4.1, we present our limit results for the case that a fraction q of the network adopts at time 0. In Section 4.2, we establish asymptotic results for the early adoption regime which is of interest in the absence of seeds.

4.1. Diffusion with Seeds

Contact Cost. Recall that $W(y)$ represents the firm's cost to reach y adopters. In the following theorem, we prove that for $y = sn$, where $s \in (q, 1)$, the contact cost, $W(y)$, almost surely converges to a deterministic function:

Theorem 4.1 (Limit of Contact Cost)**.** *For a given degree distribution* $\{(p_i, k_i), i \in I\}$[4] *and a feasible seeding configuration* $\{q_i, i \in I\}$*, suppose for all* $n > 1$ *that* G_n *is sampled uniformly at random from the set of all connected networks with* n *agents with degree distribution* $\{(p_i, k_i), i \in I\}$*, in*

which nq_i agents of degree k_i are seeded.[5] *Then, for any $s \in (q,1)$, the following limit holds:*

$$\frac{W(sn)}{n} \overset{a.s.}{\to} w(s) := cf^{-1}(1-s), \tag{1}$$

where

$$f(x) = \sum_{i\in I}(p_i - q_i)\left(1 - \frac{2x}{k_{\text{avg}}}\right)^{k_i/2}. \tag{2}$$

In formula (1), $f^{-1}(1-s)$ represents the average number of contacts per agent required to reach adoption proportion s. In general, this inverse function cannot be written in closed form. However, the limit of contact cost, $w(s)$, can easily be calculated, since $f(\cdot)$ is a polynomial function with bounded degree. The inverse is well defined, as $f(\cdot)$ is strictly decreasing in $x \in (0, k_{\text{avg}}/2)$.

Timing. In the following theorem, we focus on the timing of diffusion, and we establish an almost sure limit for the time it takes to reach sn adopters, where $s \in (q,1)$:

Theorem 4.2 (Limit of Timing). *For a given degree distribution $\{(p_i,k_i), i \in I\}$ and a feasible seeding configuration $\{q_i, i \in I\}$, suppose for all $n > 1$ that G_n is sampled uniformly at random from the set of all connected networks with n agents with degree distribution $\{(p_i,k_i), i \in I\}$, in which nq_i agents of degree k_i are seeded. Then, for any $s \in (q,1)$, the following limit holds:*

$$T(sn) \overset{a.s.}{\to} t(s) := \int_0^{f^{-1}(1-s)} \frac{1}{\gamma g(x)}\, dx, \tag{3}$$

where $f^{-1}(\cdot)$ is the inverse of $f(x)$ defined in (2) and

$$g(x) = k_{\text{avg}} - 2x - \sum_{i\in I}(p_i - q_i)k_i\left(1 - \frac{2x}{k_{\text{avg}}}\right)^{k_i/2}. \tag{4}$$

The main idea in proving both Theorems 4.1 and 4.2 is outlined in Section 6, with the details left for Online Appendices EC.3 and EC.4, respectively. We remark that unlike in the case of contact cost, the time to reach sn adopters does not scale with n, for any $s \in (q,1)$, because the rate of contact itself scales with n. Also, diffusion processes are more commonly represented in a differential form that describes the diffusion growth rate—for example, the original Bass model (Bass 1969). In the following corollary of Theorem 4.2, we also provide the growth rate:

Corollary 4.3 (Growth Rate). *Let $S(t)$ be the number of adopters at time t. Directly from Theorem 4.2, we have that $\frac{S(t)}{n} \overset{a.s.}{\to} s(t)$, where $s(t)$ solves the differential equation*

$$\frac{ds}{dt} = \gamma \sum_{i\in I}(p_i - q_i)k_i\psi^{k_i/2}\left[1 - \sum_{i\in I}(p_i - q_i)\frac{k_i}{k_{\text{avg}}}\psi^{k_i/2-1}\right], \tag{5}$$

and ψ is the solution to the polynomial equation:

$$\sum_{i\in I}(p_i - q_i)\psi^{k_i/2} = 1 - s. \tag{6}$$

Utilizing the above corollary, in the following proposition, we characterize the adoption proportion at which peak growth occurs. Define $s_p := \arg\max_{s\in[q,1)} \frac{ds}{dt}$. We have the following:

Proposition 4.4 (Peak Growth). *There is a unique adoption proportion s_p at which point the rate of adoption is maximized:*

$$s_p = \max\left\{1 - \sum_{i\in I}(p_i - q_i)\psi_p^{k_i/2},\ \ q\right\}, \tag{7}$$

where ψ_p is the solution in the interval $(0,1)$ to the polynomial equation[6]

$$\left(\sum_{i\in I}(p_i - q_i)k_i^2\psi_p^{k_i/2}\right)\left(\sum_{i\in I}(p_i - q_i)k_i\psi_p^{k_i/2} - \frac{1}{2}k_{\text{avg}}\psi_p\right) - \left(\sum_{i\in I}(p_i - q_i)k_i\psi_p^{k_i/2}\right)^2 = 0. \tag{8}$$

4.1.1. Closed Form Solutions for Special Cases. In general, we cannot find closed-form expressions for $w(s)$ and $\frac{ds}{dt}$, as they require solving Equations (2) and (6). However, in some special cases, we can solve these polynomial equations, and therefore we have closed-form solutions. In the following, we present the solutions in two special cases: k-regular networks and two-degree networks (k_l, k_h) where $k_l = 0.5k_h$.

***k*-Regular Network.** For a k-regular network and any feasible seeding configuration, we have the following:

$$w(s) = \frac{k}{2}\left(1 - \left(\frac{1-s}{1-q}\right)^{2/k}\right), \tag{9}$$

$$\frac{ds}{dt} = \gamma k\left(1 - (1-q)^{2/k}(1-s)^{1-2/k}\right)(1-s). \tag{10}$$

We remind the reader that the Bass model takes the form of $\frac{ds}{dt} \propto s(1-s)$. Note that in (10), as we make q very small and k large, the differential equation for the growth rate becomes close to that of the Bass model.

A Two-Degree Network $(k_l = 0.5k_h, k_h)$. Here, the degree distribution is assumed to be $\{(p_l, k_l = 0.5k_h), (p_h, k_h)\}$, where p_l, and p_h are the fraction of low- and high-degree agents, respectively. Let us define the following functions.

$$\zeta(s) := \sqrt{(p_l - q_l)^2 + 4(p_h - q_h)(1-s)},$$

$$\omega(s) := \frac{\zeta(s) - (p_l - q_l)}{2(p_h - q_h)}.$$

For any feasible seeding configuration, we have the following:

$$w(s) = \left(\frac{1}{2}p_l + p_h\right)k_l\left(1-\omega(s)^{2/k_l}\right), \tag{11}$$

$$\frac{ds}{dt} = k_l\omega(s)\zeta(s)\left(1-\frac{\omega(s)^{1-2/k_l}\zeta(s)}{p_l+2p_h}\right). \tag{12}$$

4.2. Diffusion Without Seeds: Early Adoption Regime

Recall that we define the early adoption regime for a setting where the firm does not seed a fraction of agents, and the diffusion starts with a single agent selected uniformly at random. In the following theorem, we characterize the contact cost and time needed to reach $\alpha \log n$ adopters, where $\alpha > \varepsilon$ and ε is a positive constant independent of n.

Theorem 4.5 (Limits of Contact Cost and Timing in Early Adoption)**.** *Suppose for all $n > 1$ that G_n is sampled uniformly at random from the set of all connected networks with n agents with degree distribution $\{(p_i, k_i), i \in I\}$. Further, suppose that ε is a positive constant independent of n. Then, for any $\alpha > \varepsilon$, the following limits hold:*

$$\frac{W(\alpha \log n)}{\alpha \log n} \xrightarrow{p} c, \tag{13}$$

$$\frac{T(\alpha \log n)}{\log\log n} \xrightarrow{p} t^E(\alpha) := \frac{1}{\gamma\left(k_{\text{avg}} + \frac{\sigma^2}{k_{\text{avg}}} - 2\right)}, \tag{14}$$

where $\sigma^2 := \sum_{i\in I} p_i(k_i - k_{\text{avg}})^2$ is the variance of the degree distribution.

We point out that the limit of contact cost (13) implies that in the early adoption regime, on average, every additional adopter costs c. Said another way, there are $o(\log n)$ wasteful contacts in this regime.[7] Further note that the timing scales with $\log\log n$, which implies that the diffusion grows exponentially. We also remark that the right-hand side of (14) does not depend on α, implying that for $\alpha' > \alpha > 0$, $T(\alpha' \log n) - T(\alpha \log n) = o(\log\log n)$. In words, Theorem 4.5 establishes that the time needed to go from $\alpha \log n$ to $\alpha' \log n$ grows sublinearly with $\log\log n$ and is dominated by the time it takes to reach $\alpha \log n$ adopters (from one seed). The following corollary describes the impact of degree distribution on contact cost and timing in the early adoption regime.

Corollary 4.6 (Impact of Degree Distribution in Early Adoption)**.** *In the limit,*

1. *The scaled contact cost, $\frac{W(\alpha \log n)}{\alpha \log n}$, is independent of the degree distribution.*
2. *The scaled timing, $\frac{T(\alpha \log n)}{\log\log n}$, only depends on the first two moments—that is, k_{avg} and σ.*
3. *Fixing k_{avg}, the scaled timing decreases with σ.*

5. Impact of Degree Distribution

In this section, we conduct comparative statics that shed light on how the degree distribution impacts the diffusion process. In Section 5.1, we focus on the contact cost and time to diffuse to an adoption fraction $s \in (q, 1)$. Section 5.2 studies the problem of optimal seeding for a given degree distribution and seeding budget and explains the impact of degree distribution on the optimal seeding strategy. Finally, we remark that we have already provided comparative statics in the early adoption regime in Corollary 4.6. However, in Section 5.3, we provide further results on the impact of degree distribution on the speed of growth in the early adoption regime. Throughout this section, when we compare two degree distributions, we mean to compare two different networks with the specified degree distributions (as opposed to taking a network and changing its degree distribution by adding or removing edges).

5.1. Cost and Timing

We first devote our attention to how the degree distribution impacts functions $w(s)$ and $t(s)$ that are the limits of the scaled contact cost and timing defined in (1) and (3), respectively. To separate the impact of degree distribution from that of the seeding configuration, throughout this section, we assume that seeds are selected uniformly at random regardless of their degrees. The problem of optimizing over seeding configurations is studied in Section 5.2.

Contact Cost. In the following proposition, we show that if the total number of edges in the network is fixed (or equivalently k_{avg} is fixed), a network with minimum variations in the degree is the most cost-efficient one.

Proposition 5.1 (Impact of Degree Variation on Contact Cost)**.** *For a fixed average degree k_{avg}, any seeding budget $0 < q < 1$, and any adoption fraction $s \in (q, 1)$, contact cost $w(s)$ is minimized when $k_{\max} - k_{\min}$ is minimized. In particular, if $k_{\text{avg}} \in \mathbb{N}$, then the most cost-efficient network is the k_{avg}-regular network.*

Intuition for the above proposition was provided in Section 1, and the proof is presented in Online Appendix EC.7. We remark that given any k_{avg}, it is possible to create a network such that $\max_{i,j\in I}\{|k_j - k_i|\} \le 1$. If $k_{\text{avg}} \in \mathbb{N}$, we use a k_{avg}-network. Otherwise, we set $k_l = \lfloor k_{\text{avg}} \rfloor$, $k_h = \lceil k_{\text{avg}} \rceil$, $p_l = k_h - k_{\text{avg}}$, and $p_h = k_{\text{avg}} - k_l$. The same distribution also minimizes the variance, fixing the average degree to be k_{avg}.

Next, we investigate the impact of having more edges in the network on the contact cost. In general, we expect that having more edges would result in more wasteful contacts and thus increase the contact cost.

In the following proposition, we show that this intuition holds for regular networks.

Proposition 5.2 (Impact of More Edges on Contact Cost)**.** *Consider two regular networks with degrees k_1 and k_2, and let $w_1(s)$ and $w_2(s)$ be their corresponding contact cost functions, respectively. If $k_1 > k_2$, then $w_1(s) \geq w_2(s)$, for any $s \in (q, 1)$.*

Timing. The impact of degree distribution on timing is more involved as it depends on the targeted adoption level. We investigate the impact of degree distribution mainly by numerically solving for $t(s)$ given a target level s, as defined by (3). In the following, we describe two representative examples of our numerical analysis.

Let us start by explaining the setting of the timing results in Table 1. Here, we fix the average degree to be 8. We set $q = 0.01$, and we assume that seeds are selected uniformly at random regardless of their degrees. In the top table, we set $s = 0.5$, and we first compute the time to reach that target in an 8-regular network. We then normalize the timing results for other networks with that of an 8-regular one. Each row of the top table reports timing for degree distributions where the first two moments are kept fixed but the third moment (i.e., the skewness) varies between −0.45 and 0.45.[8] The setting for the bottom table is similar except that the target adoption level is $s = 0.95$.

We make the following observations for adoption level $s = 0.5$: (1) The diffusion is faster in all of the considered nonregular networks compared to the regular one. (2) Higher variance speeds up the diffusion. (3) Higher skewness, on the other hand, slows down the diffusion; however, the impact is fairly small.

Interestingly, we make the opposite observations for the high adoption level of $s = 0.95$. (1) All of the nonregular networks are slower to reach $s = 0.95$ compared to the regular network. (2) Higher variance slows down the diffusion time. (3) Skewness has a more pronounced effect, and it positively impacts the diffusion—that is, the diffusion speeds up when skewness increases.

The above observations highlight three main points: (1) There is a trade-off between cost and speed for reaching s adopters when s is not very high. (2) Unlike the early adoption regime—where the diffusion timing only depends on the first two moments (Corollary 4.6)—when the target level is high, the diffusion also depends on the third moment (i.e., the skewness). (3) The impact of variance is reversed when the target level is high.

Overall, if minimizing contact cost is of paramount importance, a regular network is ideal. However, if the goal is to reach a moderate level of adoption as quickly as possible, a regular network performs poorly. Finally, in the following proposition, we show that for regular networks, more edges results in faster diffusion. Comparing the message of the following proposition with the opposite message of Proposition 5.2 again highlights the trade-off between contact cost and speed.

Proposition 5.3 (Impact of More Edges on Timing)**.** *Consider two regular networks with degrees k_1 and k_2, and let $t_1(s)$ and $t_2(s)$ be their corresponding timing functions, respectively. If $k_1 > k_2$, then $t_1(s) \leq t_2(s)$, for any $s \in (q, 1)$.*

Further, we briefly discuss the impact of degree distribution on s_p, the adoption proportion at which peak growth occurs. In Table 2, we report s_p in the same setting as the numerical results in Table 1—that is, we fix the average degree to be 8, and we set $q = 0.01$. We make the following observations: (1) Higher variance results in having a peak growth at a smaller adoption proportion (i.e., smaller s_p). In particular, in all of the networks with variance higher than 4, peak growth happens before half of the population adopts. We remark that in the basic Bass model—which results in a symmetric adoption curve—peak growth occurs when half the population has adopted. (2) Higher skewness, on the other hand, makes peak growth occur at a higher adoption proportion; however, the effect is not that strong. Comparing the observations from the top of Table 1—which is the time required to reach $s = 0.5$—with those from Table 2, we note that both the variance and skewness have similar effects on the time to reach $s = 0.5$ and the value of s_p.

Before finishing this section, we briefly discuss the impact of degree distribution on timing in power law networks, which are commonly used to model networks whose degree distribution is heavy tailed. A power law degree distribution is characterized by a single parameter ν. The probability of having degree k is $ck^{-\nu}$, where c is a normalization factor. The most commonly used range for the exponent in the power law model is $\nu \in (2, 3]$. In Table 3, we study how the time to reach adoption levels of 0.5 and 0.95 changes

Table 2. The Impact of Variance and Skewness on s_p

		Skewness						
		−0.45	−0.30	−0.15	0	0.15	0.30	0.45
Variance	0				0.524			
	2	0.512	0.512	0.512	0.512	0.512	0.512	0.512
	4	0.499	0.499	0.499	0.500	0.500	0.500	0.500
	6	0.486	0.487	0.487	0.488	0.488	0.488	0.489
	8	0.474	0.475	0.475	0.476	0.477	0.477	0.478
	10	0.462	0.463	0.464	0.465	0.466	0.466	0.467

Notes. We have $k_{\text{avg}} = 8$, $q = 0.01$.

Table 3. Diffusion Time in Power Law Networks

		Scale-free parameter (ν)				
		2.2	2.4	2.6	2.8	3.0
Adoption Proportion	0.5	1.000	1.365	1.840	2.420	3.074
	0.95	1.000	1.160	1.352	1.571	1.808

Notes. In both cases, $q = 0.01$ and we assume random seeding, a minimum degree of 3, and a maximum degree of 1,000. The times in each row are normalized by the diffusion time of the first column—that is, $\nu = 2.2$.

as we increase ν from 2.2 to 3 for a network with the following distribution: minimum degree and maximum degree are 3 and 1,000, respectively, and the probability of having degree k is $\tilde{c}k^{-\nu}$ (with $\tilde{c}$ being the appropriate normalization factor when the degree ranges from 3 to 1,000). For the ease of comparison, in each row, we normalize the diffusion times for different values of ν by that of $\nu = 2.2$—that is, its first column. We note that as ν increases, the average degree decreases. This, in turn, results in having slower diffusion for both target adoption levels.

We note that this is in line with Proposition 5.3, which states that an increased average degree always speeds up diffusion (recall that Proposition 5.3 proves this behavior for regular networks). In addition, as ν increases, the variance also decreases. We observe that as ν increases, the diffusion slows down, and the effect is more pronounced when $s = 0.5$. This is aligned with our discussion on the comparative static results in Table 1, which suggests that increased variance speeds up diffusion for low target levels (e.g. $s = 0.5$) but slows down diffusion for high target levels (e.g. $s = 0.95$).

5.2. Optimal Seeding

We now turn our attention to the firm's seeding decision. Recall the firm's decision is to select a feasible seeding configuration, as defined in Definition 3.1, with the goal of minimizing the contact cost or time required to reach a target adoption level. We remark that any feasible seeding configuration $\{q_i, i \in I\}$ exhausts the seeding budget—that is, $\sum_{i\in I} q_i = q$. First, in the following proposition, we show that when minimizing the contact cost or timing, restricting the feasible region to include only seeding configurations that exhaust the budget is without loss of generality.

Proposition 5.4 (Monotonicity of Contact Cost and Timing in Seeding Budget)**.** *For any seeding configuration $\{q_i, i \in I\}$, any i such that $q_i < p_i$, and any adoption proportion $s \in (q, 1)$, increasing q_i reduces both the contact cost (function $w(s)$) and timing (function $t(s)$).*

Consider a seeding configuration such that $\sum_{i\in I} q_i < q$. This implies that there exists an index i such that $q_i < p_i$; now, if we increase q_i to $q_i + \epsilon$—such that $\epsilon > 0$ and we still have $q_i + \epsilon \le p_i$ and $\epsilon + \sum_{i\in I} q_i \le q$—the above proposition implies that both functions $w(s)$ and $t(s)$ will decrease. Therefore, the optimal seeding configuration will exhaust the budget. This confirms that restricting the feasible region to include only configurations that exhaust the budget is without loss of generality. In the following proposition, we characterize the optimal seeding strategy for the objective of minimizing cost.

Proposition 5.5 (Optimal Seeding Strategy for Minimizing Contact Cost)**.** *Given a fixed q and any $s \in (q, 1)$, denote the seeding configuration that minimizes $w(s)$ by $\{q_i^*, i \in I\}$. The following holds: If $q_i^* > 0$ and $k_i > k_j$, then $q_j^* = p_j$—that is, we maximally seed low-degree agents until we exhaust the budget q.*

The optimal strategy for minimizing time depends on the target level s and the seeding budget q. Intuitively, when the target fraction is small or moderately large, seeding high-degree agents is helpful as they make more contacts and speed up the adoption. However, when the target proportion is very high, seeding low-degree agents can prove useful. Once most agents have adopted the product, diffusion will be slower if the remaining agents have low degree, since they will be contacted at a low rate. If low-degree agents are not seeded at all, they will substantially slow down the "tail" of the adoption curve, which makes up a large proportion of the total time—especially when the seeding budget is high.

Mathematically speaking, the function $t(s)$ (defined in (3)) is an integral where both the integrand (i.e., $\frac{1}{\gamma g(x)}$) and the integral range (i.e., $f^{-1}(1-s)$) vary with the seeding configuration. Now, suppose we aim to compare $t(s)$ for two seeding configurations: $\{q_{1,i}, i \in I\}$ and $\{q_{2,i}, i \in I\}$, where $q_{1,i} = q_{2,i}$ for $i \in I \setminus \{l, h\}$, and $q_{2,l} = q_{1,l} - \varepsilon$ and $q_{2,h} = q_{1,h} + \varepsilon$—that is, the second seeding configuration shifts ε of the seeding budget from degree k_l to degree k_h. Let us define $t_1(\cdot), f_1(\cdot)$, and $g_1(\cdot)$ as the corresponding functions for the first seeding configuration and similarly $t_2(\cdot)$, $f_2(\cdot)$, and $g_2(\cdot)$ for the latter. As a byproduct of the proof of Proposition 5.5, we can establish certain monotonicity properties of the function $f^{-1}(1-s)$ regardless of the value of s—for example, we can show that if $k_h > k_l$, then $f_1^{-1}(1-s) < f_2^{-1}(1-s)$. However, such behavior does not hold for the integrand; functions $g_1(x)$ and $g_2(x)$ can cross. Consequently, the relation between $t_1(s)$ and $t_2(s)$ can vary with s—that is, for a certain range of s we can have $t_1(s) \ge t_2(s)$, while for another range the reverse relation holds.

In the following, we first focus on the regime where the target adoption proportion is low. In Proposition 5.6, we prove that for low target levels, maximally seeding the highest-degree agents minimizes the time to reach an adoption proportion s. This is

aligned with our intuition that when the target fraction is sufficiently low, seeding high-degree agents is effective as they make more contacts and speed up the adoption process.

Proposition 5.6 (Optimal Seeding Strategy for Minimizing Time)**.** *Suppose $q < p_i$ for all $i \in I$. Then, there is a threshold $\bar{s} \geq q$ such that if $s \leq \bar{s}$, then the function $t(s)$ is minimized when we only seed the highest-degree agents.*

The threshold $\bar{s}$ in the above proposition depends on both the degree distribution and the seeding budget q. Unfortunately, $\bar{s}$ does not admit a closed-form solution, but it is easy to compute. We defer the definition of $\bar{s}$ to the proof of the proposition presented in Online Appendix EC.11.[9] In Table 4, we compute $\bar{s}$ for different seeding budgets, q, and different three-degree distributions $\{(\frac{1}{3}, k_{\min}), (\frac{1}{3}, 8), (\frac{1}{3}, 16 - k_{\min})\}$, where we vary $k_{\min}$ from 3 to 7. Note that all these distributions have the same average degree; however, as we increase $k_{\min}$, the variance decreases. We further remark that even though we do not prove that $\bar{s} > q$, our numerical study shows that this is often the case. In fact, in every entry of Table 4, we have $\bar{s} > q$.

Solving for the optimal seeding strategy when the target level exceeds $\bar{s}$ (as defined by the above proposition) is unfortunately intractable. Even though we cannot characterize the optimal strategy in this regime, in the example of Figure 1, we rule out the optimality of simple policies such as maximally seeding the high- or low-degree agents.

In Figure 1, we focus on a three-degree distribution ($k_l = 3, k_m = 6, k_h = 9$). In the top left panel, $p_l = 1/3$, $p_m = 1/3$, and $p_h = 1/3$. We vary the seeding budget, q, from 0.01 to p_l. Further, we set the target, s, to be 0.99 and we plot the percentage of seeded agents with degrees 3, 6, and 9 in the optimal seeding that minimizes the timing (i.e. $100(q_l^*, q_m^*, q_h^*)/q$). The top right panel presents the same quantities for the same distribution but a smaller target adoption proportion, $s = 0.85$. In the bottom left plot, we have: $p_l = 1/10$, $p_m = 1/3$, and $p_h = 17/30$, while $s = 0.99$. Finally, in the bottom right plot, $p_l = 1/10$, $p_m = 1/3$, and $p_h = 17/30$, while $s = 0.85$.

Table 4. Threshold $\bar{s}$ (Defined in Proposition 5.6) for Different Values of q for Networks with Degree Distribution $\{(\frac{1}{3}, k_{\min}), (\frac{1}{3}, 8), (\frac{1}{3}, 16 - k_{\min})\}$ with $k_{\min} \in \{3, \ldots, 7\}$

		Seeding budget (q)					
		0.01	0.05	0.10	0.15	0.20	0.25
Minimum degree	3	0.324	0.325	0.326	0.327	0.324	0.320
	4	0.348	0.350	0.351	0.352	0.351	0.348
	5	0.369	0.372	0.375	0.377	0.378	0.378
	6	0.389	0.392	0.397	0.401	0.404	0.406
	7	0.405	0.410	0.416	0.423	0.428	0.434

We make the following observations based on Figure 1: (1) For both distributions, when the target is very high, there are certain seeding budgets where the optimal seeding strategy is a mixture of the high and low degrees. (2) Further, it is not optimal to allocate any amount of the seeding budget to the middle-degree agents in any of the four settings shown. We make a similar observation (on not seeding the middle-degree agents) in our exhaustive numerical study of three-degree distributions. (3) Finally, when the target is set lower (i.e., 0.85) for the degree distribution with fewer low-degree agents, it is optimal to only seed high-degree agents. Note that for any network and for any target level, Proposition 5.5 implies that the optimal strategy for achieving minimum contact cost constitutes seeding only the low-degree agents. Once again, for many parameters, we have a trade-off between cost and time. Seeding the lowest-degree agents reduces cost, but seeding the highest-degree agents often reduces time (if the target level is not very high).

5.3. Early Adoption Regime

We remark that we have already provided comparative statics for the early adoption regime in Corollary 4.6. Here, we further investigate the impact of degree distribution on the speed of growth. In particular, in the following proposition, we describe a rather counterintuitive phenomenon that if a degree k_j satisfies a certain condition, having a lower k_j may result in faster growth in the early adoption regime.

Proposition 5.7 (Impact of Lowering Degrees)**.** *Consider the degree distribution $\{(p_i, k_i), i \in I\}$ and an index $j \in I$ such that $p_j > 0$ and the degree k_j satisfies the following condition*:

$$k_j < \frac{1}{2}\left(k_{\text{avg}} + \frac{\sigma^2}{k_{\text{avg}}}\right).$$

Then, consider another network with a similar degree distribution defined as $\{(\tilde{p}_i, \tilde{k}_i), i \in I\}$, where

$$\tilde{p}_i = p_i, i \in I, \qquad \tilde{k}_i = k_i, i \in I \setminus \{j\}, \qquad \tilde{k}_j = k_j - 1.$$

In the early adoption regime, the diffusion grows faster in a network with the latter distribution—that is, for any $\alpha > 0$, $t^E(\alpha) \geq \tilde{t}^E(\alpha)$, where $t^E(\alpha)$ and $\tilde{t}^E(\alpha)$ are the scaled timing for the former and latter network respectively, as defined in (14).

We explain the intuition behind the above result by an example: consider two networks with degree distributions $\{(4, 0.5), (10, 0.5)\}$ and $\{(3, 0.5), (10, 0.5)\}$. In words, both networks have two possible degrees, a low degree and a high degree. In the former, the low degree is 4, whereas in the latter, the low degree is 3.

The main factor determining the timing in the early adoption regime is the expected degree of the adopters (as explained in detail in Section 6.3—in particular (26), which is equivalent to random incidence). In this example, the expected degree for distribution $\{(3,0.5),(10,0.5)\}$ is $\frac{0.5\times3^2+0.5\times10^2}{0.5\times3+0.5\times10}=8.385$, whereas for distribution $\{(4,0.5),(10,0.5)\}$, it is $\frac{0.5\times4^2+0.5\times10^2}{0.5\times4+0.5\times10}=8.286$. This shows that even though the total number of edges is smaller in the distribution $\{(3,0.5),(10,0.5)\}$, the expected degree of adopters (in the early adoption regime) is in fact larger. For intuition as to why the expected degree of an early adopter is larger, note that there are two effects of reducing the low degree (from 4 to 3): (1) an early adopter is more likely to be degree 10, but (2) other early adopters will be degree 3 as opposed to degree 4. In this example, the first effect more than compensates for the second. Because the expected degree of early adopters is larger, the diffusion time will also be faster.

We wish to emphasize that when we compare two degree distributions, we mean to compare two completely different networks with the specified degree distributions. Said differently, when comparing the two networks (e.g., in the above example), we do not mean that the latter network (with degree distribution $\{(3,0.5),(10,0.5)\}$) was generated by removing edges from the former network.[10]

6. Outline of Proof Ideas

In this section, we present an outline of the proof ideas. In proving our theoretical results, we crucially use the fact that a random network with a given degree distribution can be generated using a simple iterative matching process called the configuration model (Wormald 1999), which works as follows: It begins with n isolated agents; an agent with degree k_i has k_i unmatched half-edges. Then, in each iteration, we select an arbitrary half-edge and match it with another one selected uniformly at random among all the remaining half-edges. Thus, in each iteration, we create one edge. The process ends when all half-edges are matched—that is, after a total of $\frac{1}{2}nk_{\text{avg}}$ iterations.

In our analysis, we utilize the Principle of Deferred Decisions (Mitzenmacher and Upfal 2005) and interleave the above graph generation process and the diffusion one as follows: We start the process with n isolated agents. Of these n agents, a subset are seeded and become initial adopters. At any time, we define the set of *active* half-edges to consist of half-edges that belong to adopters and have not been matched so far. We note that each active half-edge "wakes up" and makes a contact according to an independent and identically distributed Poisson process with rate γ. Once an active half-edge wakes up, we realize its edge—that is, we match that half-edge with another one selected uniformly at random among all of the unmatched half-edges. If the selected half-edge belongs to a non-adopter, the diffusion progresses; otherwise, the number of adopters does not increase. This process is summarized in Algorithm 1 in Section 6.1.

Further, in Section 6.1, we show that the evolution of the discrete-time interleaved process that results from observing the system at wake-up epochs (i.e., when one half-edge is matched with another) can be modeled by a low-dimensional process (with state-space dimension of $|I|+2$). This allows us to keep track of the number of adopters of any given degree. Then, in Section 6.2, we prove that the number of adopters at each iteration can be approximated with a deterministic function. This result, in turn, helps us prove Theorem 4.1. We also analyze the time between two consecutive edge formations to characterize the timing and prove Theorem 4.2. Finally, the main idea of the proof for the early adoption regime uses the locally tree-like structure of random networks (Dembo and Montanari 2010), which is outlined in Section 6.3.

6.1. Interleaved Graph Generation and Diffusion Process

The interleaved process outlined at the beginning of Section 6 is described in detail in Algorithm 1, with the following notation: N_0 denotes the set of seeded agents; A denotes the set of active half-edges (i.e., those unmatched half-edges that belong to adopters); L represents the set of all unmatched half-edges.[11]

Algorithm 1. Interleaved Graph Generation and Diffusion

1. Initialize the set of adopters with N_0 (the set of seeded agents).
2. **For** all agents in N_0, add all of their half-edges to A, the set of active half-edges.
3. Add all half-edges to L, the set of unmatched half-edges.
4. **While** L is non-empty,
   - (a) **If** A is non-empty,
     - i. Select an arbitrary half-edge e_1 from A.
     - ii. Select a half-edge e_2 from L uniformly at random. Let v_2 denote the agent to which the half-edge e_2 belongs.
     - iii. **If** $e_2 \notin A$
       - A. Add agent v_2 to the set of adopters.
       - B. Add all the half-edges of v_2 to A.
     - iv. Remove e_1 and e_2 from both A and L.
   - (b) **Else**
     - i. Select an arbitrary non-adoptive agent u.
     - ii. Remove two of u's half-edges from L and add the rest to A.

First, we study the discrete-time process that results from observing the system at wake-up epochs.[12] We only need to keep track of the following quantities at each iteration $0 \le j \le \frac{1}{2}nk_{\text{avg}}$:

1. For $i \in I$, $N_i(j)$, the number of non-adoptive agents with degree k_i.
2. $A(j)$, the number of unmatched half-edges connected to adopters.
3. $L(j)$, the total number of unmatched half-edges.

This represents a significant improvement over the naive approach (keeping track of the status of *every* agent) since the dimensionality of our stochastic process no longer scales with n. The initial conditions of these $|I|+2$ quantities can be calculated given an initial set of adopters. Further, we can track their evolution as follows:

1. For each $i \in I$, with probability $\frac{k_i N_i(j)}{L(j)-1}$ the half-edge selected uniformly at random—that is, e_2 in Algorithm 1—belongs to a non-adoptive agents with degree k_i. In this case, we have the following:

$$N_i(j+1) = N_i(j) - 1 \tag{15a}$$
$$N_h(j+1) = N_h(j), \ \forall h \in I, \text{ and } h \neq i \tag{15b}$$
$$A(j+1) = A(j) + k_i - 2 \tag{15c}$$
$$L(j+1) = L(j) - 2. \tag{15d}$$

2. With probability $1 - \sum_{i\in I} \frac{k_i N_i(j)}{L(j)-1}$, the half-edge selected uniformly at random—that is, e_2 in Algorithm 1—belongs to an adoptive agent; thus, we have the following:

$$N_i(j+1) = N_i(j), \ \forall i \in I \tag{16a}$$
$$A(j+1) = A(j) - 2 \tag{16b}$$
$$L(j+1) = L(j) - 2\,. \tag{16c}$$

Based on the evolution above, three observations enable a thorough analysis of our process. First, note that $L(j)$ deterministically decreases by two in each iteration. As a result,

$$L(j) = nk_{\text{avg}} - 2j, \ 0 \le j \le \frac{1}{2} nk_{\text{avg}}.$$

Second, if we know $L(j)$ and each $N_i(j)$, we can solve exactly for $A(j)$:

$$A(j) = L(j) - \sum_{i\in I} k_i N_i(j), \ 0 \le j \le \frac{1}{2} nk_{\text{avg}}. \tag{17}$$

Third, if we know the state of the stochastic process at iteration j, $H(j) := \{N_i(j), i \in I, L(j), A(j)\}$, then we can specify the expected (negative) growth rate of each $N_i(j)$:

$$\mathbb{E}\big[N_i(j+1) - N_i(j) | H(j)\big] = \frac{-k_i N_i(j)}{L(j)-1} = \frac{-k_i N_i(j)/n}{k_{\text{avg}} - 2j/n - 1/n},$$
$$\forall i \in I, \ 0 \le j \le \frac{1}{2} nk_{\text{avg}} - 1.$$

6.2. Deterministic Approximations

We now adjust an approach proposed in Wormald (1999) for approximating discrete-time stochastic processes with deterministic functions. We define a deterministic approximation for the proportion of agents that are non-adoptive and have degree k_i:

$$f_i(x) = (p_i - q_i)\left(1 - \frac{2x}{k_{\text{avg}}}\right)^{k_i/2}, \tag{18}$$

where $x \in [0, \frac{1}{2}k_{\text{avg}})$ represents the average number of iterations per node. Note that $f_i(x)$ is the unique solution to

$$f_i'(x) = \frac{-k_i f_i(x)}{k_{\text{avg}} - 2x} \tag{19}$$

with initial condition $f_i(0) = p_i - q_i$. In the following lemma, we prove that $N_i(j)$ can be approximated by the deterministic function $f_i(\cdot)$ with appropriate scaling.

Lemma 6.1. *For any $i \in I$ and constant $\varepsilon_0 > 0$, at any iteration $0 \le j < (\frac{1}{2}k_{\text{avg}} - \varepsilon_0)n$, with probability $1 - o(n^{-3})$*

$$|N_i(j) - nf_i(j/n)| \le \delta_i(n) \tag{20}$$

uniformly over j with $\delta_i(n) = o(n)$.

A proof of this lemma can be found in Online Appendix EC.1. Let us define $f(x) := \sum_{i\in I} f_i(x)$. Based on Lemma 6.1, Equation (17), and the assumption that $k_{\max}$ is bounded, we have an immediate corollary.

Corollary 6.2. *At any iteration $0 \le j < (\frac{1}{2}k_{\text{avg}} - \varepsilon_0)n$, with probability $1 - o(n^{-3})$*

$$|A(j) - ng(j/n)| \le \sum_{i\in I} \delta_i(n) \le |I| \max_{i\in I}\{\delta_i(n)\} := \delta_0(n) \tag{21}$$

uniformly over j, where

$$g(x) = k_{\text{avg}} - 2x - \sum_{i\in I} (p_i - q_i) k_i \left(1 - \frac{2x}{k_{\text{avg}}}\right)^{k_i/2}, \tag{22}$$

and $\delta_0(n) = o(n)$.

We now turn our attention to the number of iterations required to reach an adoption proportion s, which we will denote J_s. Intuitively, we would expect J_s to be close to $nf^{-1}(1-s)$ because $f(j/n)$ represents an approximation of the proportion of agents that are non-adopters at iteration j. We show that this intuition is correct under certain conditions. First, we define $\mathbf{A}$ as the vector comprised of the random values of $A(j)$ for $0 \le j < (\frac{1}{2}k_{\text{avg}} - \varepsilon_0)n$. Let $\hat{\mathbf{a}}$ be a particular realization of this vector. We now define $\mathbf{S}$ as the set of all such $\hat{\mathbf{a}}$ where $ng(j/n)$ approximates $A(j)$ well. Specifically, $\mathbf{S} := \{\hat{\mathbf{a}} : |\hat{\mathbf{a}}(j) - ng(j/n)| \le \delta_0(n), 0 \le j < (\frac{1}{2}k_{\text{avg}} - \varepsilon_0)n\}$. We have the following lemma:

Lemma 6.3. *Let $j_s := nf^{-1}(1-s)$ and $\mu = \min_{i\in I} |f_i'(\frac{1}{2}k_{\text{avg}} - \varepsilon_0)|$. If $\mathbf{A} \in \mathbf{S}$, then for any $\varepsilon_0 n \le j_s \le (\frac{1}{2}k_{\text{avg}} - 2\varepsilon_0)n$, with probability $1 - o(n^{-3})$,*

$$|J_s - j_s| \le \frac{2}{\mu}\delta_0(n). \tag{23}$$

A proof of Lemma 6.3 can be found in Online Appendix EC.2. Together, these lemmas help us analyze the behavior of the adoption process for any $s \in (q, 1)$. Since the number of iterations can equivalently be thought of as the number of contacts, this essentially completes the proof of Theorem 4.1, minus some technical details ensuring the connectedness and simplicity of the graph generated through the interleaved process (presented in Online Appendix EC.3).

Up until this point, we have abstracted away from the time taken by this stochastic process and instead focused on the number of iterations. We now incorporate the time by letting e_1 in Algorithm 1 be the active half-edge whose clock wakes up first. Further, let us now define $M(j)$ as a random variable representing the time between iteration j and iteration $j+1$. Recall that each half-edge wakes up according to an independent Poisson process with homogeneous rate γ. Thus, $M(j)$ is an exponential random variable with rate $\gamma A(j)$, where $A(j)$ is the total number of unmatched half-edges belonging to adopters (i.e., the number of active half-edges) at iteration j. We then sum these exponential random variables to specify the total time until adoption proportion s is reached.

$$T(sn) = \sum_{j=0}^{J_s} M(j). \tag{24}$$

In Online Appendix EC.4, we prove Theorem 4.2 by showing that the right-hand side of Equation (24) converges almost surely to the sum of the expected values of each $M(j)$, which is given by $\frac{1}{\gamma A(j)}$ (conditioned on $A(j)$).

6.3. Proof Idea for the Early Adoption Regime

The first part of the proof of Theorem 4.5 is a direct result of showing that when there are only y adopters where $y = O(\log n)$, then the subgraph of adopters is a connected tree with high probability, where a tree is a graph with no cycles.[13] We prove this result in Online Appendix EC.6.

To prove the second part, we note that if the subgraph of adopters is a connected tree, each iteration j of the interleaved process (i.e., Algorithm 1 when we set N_0 to be a single randomly selected agent) adds a new adopter. In that case, after j iterations, we will have $j+1$ adopters. Therefore for $y = \alpha \log n$, we have the following:

$$T(y) = \sum_{j=0}^{y-1} M(j). \tag{25}$$

Next, we compute $\mathbb{E}[M(j)]$, $0 \le j \le \alpha \log n$. This requires us to count the number of active half-edges after j iterations, which we again denote by $A(j)$. If the subgraph of adopters is a connected tree, we have the following:

$$\mathbb{E}[A(0)] = k_{\text{avg}}, \quad \text{and}$$
$$\mathbb{E}[A(j)] = \mathbb{E}[A(0)] + (\mathbb{E}[\text{degree of agent } a \mid a \text{ has adopted}] - 2)j$$

In order to calculate $\mathbb{E}[\text{degree of agent } a \mid a \text{ has adopted}]$, we rely on the fact that while there are only $o(n)$ adopters, the proportion of remaining non-adopters that are degree k_i is given by $p_i - o(1)$. Thus, we have the following:

$$\mathbb{E}[\text{degree of agent } a \mid a \text{ has adopted}] = \sum_{i \in I} k_i \frac{(p_i - o(1))k_i}{k_{\text{avg}}} = k_{\text{avg}} + \frac{\sigma^2}{k_{\text{avg}}} - o(1). \tag{26}$$

Note that the above calculation is equivalent to random incidence, which appears in numerous settings. It was, for instance, applied to diffusion in networks in Dover et al. (2012). In Online Appendix EC.6, we prove that after scaling, (25) converges in probability to its expected value.

7. Model Discussion: Generalizations and Comparisons

In this section, we discuss two generalizations of our model and explain how our framework can be utilized to study these models. First, in Section 7.1, we consider an adoption process with an innovation factor—an innovator adopts the product on her own and not through contacts by her neighbors—and discuss how our solution can be generalized to such a setting. Next, in Section 7.2, we discuss a diffusion process where contact rates vary with time. In Section 7.3, we compare the outcome of our exact analysis with that of a mean-field approximation for different degree distributions, and we highlight the endogenous impact of diffusion on the distribution of adopters' neighbors. Finally, in Section 7.4, we examine how our solution—as well as that of a mean-field approximation—performs when applied to networks generated using the preferential attachment mechanism (Barabási and Albert 1999) and the friends-of-friends mechanism (Jackson and Rogers 2007). We illustrate that our solution performs well when applied to these random networks that are not generated by the configuration model.

7.1. Diffusion with Innovation

Suppose that, independent of the contact process, each agent may adopt the product on her own. Such an independent adoption can be the result of other forms of marketing forces such as mass media advertising. The original Bass model (Bass 1969) also

includes an *innovation* factor to incorporate such external influences. We generalize our model to include an innovation factor. In particular, suppose that each agent has an independent innovation Poisson clock with rate λ. When this clock wakes up, if the agent is a non-adopter, she will adopt with probability 1. In the following, we explain how we can modify our analysis to incorporate the innovation factor. In particular, we again observe the system at wake-up epochs (of either the non-adopters' innovation clocks or the active half-edges' contact clocks), and we modify the evolution given by (15a)–(16c) as follows:

1. For each $i \in I$, with probability $(\frac{\gamma A(j)}{\gamma A(j)+\lambda \sum_{i\in I} N_i(j)}) \times (\frac{k_i N_i(j)}{L(j)-1})$:

$$N_i(j+1) = N_i(j) - 1 \tag{27a}$$
$$N_h(j+1) = N_h(j), \ \forall h \in I, \text{ and } h \neq i \tag{27b}$$
$$A(j+1) = A(j) + k_i - 2 \tag{27c}$$
$$L(j+1) = L(j) - 2\,. \tag{27d}$$

2. With probability $(\frac{\gamma A(j)}{\gamma A(j)+\lambda \sum_{i\in I} N_i(j)})(1 - \sum_{i\in I} \frac{k_i N_i(j)}{L(j)-1})$:

$$N_i(j+1) = N_i(j), \ \forall i \in I \tag{28a}$$
$$A(j+1) = A(j) - 2 \tag{28b}$$
$$L(j+1) = L(j) - 2\,. \tag{28c}$$

3. For each $i \in I$, with probability $\frac{\lambda N_i(j)}{\gamma A(j)+\lambda \sum_{i\in I} N_i(j)}$:

$$N_i(j+1) = N_i(j) - 1 \tag{29a}$$
$$N_h(j+1) = N_h(j), \ \forall h \in I, \text{ and } h \neq i \tag{29b}$$
$$A(j+1) = A(j) + k_i \tag{29c}$$
$$L(j+1) = L(j)\,. \tag{29d}$$

Note that $\frac{\gamma A(j)}{\gamma A(j)+\lambda \sum_{i\in I} N_i(j)}$ is the probability that the clock that has just woken up belongs to an active half-edge. Similarly, $\frac{\lambda N_i(j)}{\gamma A(j)+\lambda \sum_{i\in I} N_i(j)}$ is the probability that the innovation clock of a non-adopter with degree k_i has just woken up.

Following our approach in Section 6.2, one can define a system of differential equations whose solutions approximate the discrete stochastic processes $A(j)$, $L(j)$, and $N_i(j)$ for all $i \in I$ with the evolution defined in (27a)–(29d). The solution to this system of nonlinear differential equations can then be used to compute the contact cost and timing (i.e., functions similar to $w(s)$ and $t(s)$).

7.2. Time-Varying Contact Rate

So far, we have assumed that the contact rate remains the same throughout the entire diffusion process. However, one can imagine that the rate of contacts can vary over time. For example, if the marketer provides incentives for the contacts—a method known as incentive-based viral marketing—he may reduce the incentives over time, which in turn, may lower the adopters' contact rate. We incorporate a time-varying contact rate by replacing the homogenous contact rate per edge γ with $\gamma(t) > 0$. In that case, our analysis remains almost identical. First note that the change in the contact rate does not impact the contact cost or the discrete-time interleaved process (defined in Section 6.1). However, it impacts our analysis of timing since we need to substitute γ with $\gamma(t)$ when computing the rate of $M(j)$ (defined in Section 6.2) and consequently (24). Instead of presenting the result for timing, we directly express the growth rate for this new model:

$$\frac{ds}{dt} = \gamma(t) \sum_{i\in I} (p_i - q_i) k_i \psi^{k_i/2} \left[1 - \sum_{i\in I} (p_i - q_i) \frac{k_i}{k_{\text{avg}}} \psi^{k_i/2-1} \right], \tag{30}$$

where ψ is given by the solution to (6). Note that this is identical to the growth rate given by (5) in Corollary 4.3, except that γ is replaced by $\gamma(t)$.

7.3. Comparison with Mean-Field Approximation

In this section, we compare the diffusion trajectory predicted by our analysis with that of a mean-field approximation and explain the difference. Recall that a mean-field approximation assumes that the distribution of neighbors of an adopter is exogenous with respect to the diffusion process. Going back to the example from Section 1, suppose agent a has degree 5. A mean-field approximation assumes that the neighbors of a are uniformly sampled from the entire population. Therefore, at any adoption proportion s, in expectation, agent a has $5 \times s$ adoptive neighbors and $5 \times (1 - s)$ non-adoptive ones. Under such an exogeneity assumption, one only needs to keep track of the fraction of adoptive agents of each degree in order to characterize the growth rate. In particular, let $s_i(t)$ be the proportion of agents with degree k_i that are adopters at time t. We adjust the analysis of Shakkottai and Johari (2010) to express the differential equations for growth in our setting as follows:

$$\frac{ds_i}{dt} = \gamma \frac{1}{k_{\text{avg}}} k_i (1 - s_i) \sum_{j\in I} p_j k_j s_j, \quad i \in I \qquad \text{and}$$
$$\frac{ds}{dt} = \gamma \frac{1}{k_{\text{avg}}} \sum_{i\in I} p_i k_i (1 - s_i) \sum_{j\in I} p_j k_j s_j. \tag{31}$$

For a given seeding configuration $\{q_i, i \in I\}$, the initial conditions for the above differential equations are $s_i(0) = \frac{q_i}{p_i}$.

In Figure 2, we compare the diffusion trajectory resulting from solving the above system of differential equations with the diffusion trajectory given in Theorem 4.2, Equation (3). We consider two networks: an 8-regular network (left) and a two-degree network $\{(3, 0.5), (13, 0.5)\}$ (right). Here, we have $q = 0.01$ and

assume that seeds are selected uniformly at random regardless of their degree. As we can see, for both networks, the solution resulting from a mean-field approximation always overestimates the proportion of adopters while our model is within the confidence interval of simulated results for 25 sample networks with 5,000 agents and the given degree distribution.

The difference between our solution and a mean-field approximation stems from an endogeneity effect that diffusion has on the distribution of the neighbors of an adopter. In particular, as the diffusion progresses, an adopter becomes more likely to be connected to other adopters. We demonstrate this effect with a numerical example. Focusing on random 3-regular networks, in the left plot of Figure 3, we plot the fraction of edges between adopters, for any proportion of adopters s, as predicted by both our analysis and a mean-field approximation. Note that the prediction of a mean-field approximation is fairly straightforward: in expectation, each adopter has $3 \times s$ adoptive neighbors. Therefore, among all edges, a fraction s^2 of them are between adopters. Computing the fraction of edges between adopters follows from our analysis in Sections 6.1 and 6.2 (it is omitted for the sake of brevity). As shown in Figure 3, for any adoption proportion s, there are many more edges between adopters than what a mean-field approximation predicts.

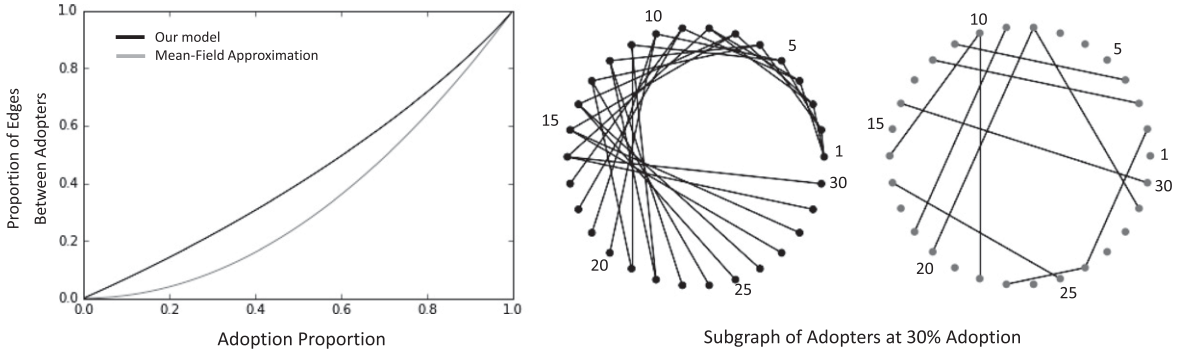
To better understand the effect that diffusion has on the distribution of the neighbors of an adopter, let us consider the following example. Suppose we have a random 3-regular network with 100 agents and diffusion starts with a single randomly selected agent a. Agent a begins contacting her neighbors at time 0. By the time that the diffusion reaches 30 agents (i.e., 30% of the population), it is very likely that agent a has already contacted all of her neighbors and therefore they are all adopters. To demonstrate this effect among adopters, in the middle panel of Figure 3, we visualize the subgraph of adopters in a sample 3-regular network that starts with a single randomly selected seeded agent. Note that in this visualization, adopters are arranged on the circle in order of adoption. In particular, the first adopter is placed at angle 0° and labeled 1. There are 32 edges between the adopters in this subgraph. We contrast this subgraph with that of a hypothetical subgraph that would be the result of a mean-field approximation. In particular, in the right panel of Figure 3, we show a subgraph resulting from sampling neighbors of adopters uniformly at random among all agents, regardless of the diffusion process. Note that adopters are again arranged on the circle in order of adoption. There are only 11 edges between the adopters in this subgraph.

We conclude this section by numerically studying the impact of degree distribution on the accuracy of a mean-field approximation. To that end, we consider a setting where $q = 0.01$ of the agents are seeded uniformly at random, regardless of their degrees, and we compute the time required to reach half of the population using a mean-field approximation as well as our analysis. In Table 5, we report the percent difference between our model and a mean-field approximation for various degree distributions. More precisely, let $t(0.5)$ and $t^{MF}(0.5)$ be the time required to reach an adoption proportion of 0.5 under our analysis and that of a mean-field approximation, respectively. We report $(t(0.5) - t^{MF}(0.5))/t(0.5)$. We observe from the results of Table 5 that a mean-field approximation underestimates the time in all of the considered networks. However, as the average degree and the variance increase, the gap between our exact solution and a mean-field approximation decreases.

7.4. Performance on Other Network Models

In this section, we examine the accuracy of our predicted diffusion curve when applied to networks

Figure 3. Illustrating the Effect of Diffusion on the Distribution of Adopters' Neighbors



Notes. Left: fraction of edges in between the first s proportion of adopters; middle and right: the subgraph of the first 30 adopters when $n = 100$. The subgraph in the middle panel is a visualization of a simulation of the diffusion process. The subgraph in the right panel is a visualization of sampling neighbors of adopters uniformly at random among all agents.

Table 5. Percent Difference Between Our Model and a Mean-Field Approximation When Computing the Time It Takes for Diffusion to Reach $s = 0.5$ When $q = 0.01$ Proportion of Randomly Selected Agents Are Seeded

		Variance				
		0	1	5	25	100
Average degree	4	38.4%	36.2%			
	8	18.3%	18.0%	17.1%	13.2%	
	15	9.6%	9.5%	9.4%	8.7%	6.6%
	25	5.7%	5.7%	5.6%	5.5%	4.9%
	50	2.8%	2.8%	2.8%	2.8%	2.7%
	100	1.4%	1.4%	1.4%	1.4%	1.4%

generated from the preferential attachment mechanism (Barabási and Albert 1999) and the friends-of-friends mechanism (Jackson and Rogers 2007). We show that the diffusion curve derived from our formulas is quite accurate and compares favorably to a mean-field approximation on both types of networks.

Many networks exhibit scale-free properties, where some agents have extremely high degree. The Barabási–Albert model can be used to generate a class of such networks, where the degrees of agents follow a power law distribution. The model begins with a small connected network and sequentially adds agents. When an agent is added, she forms an edge with m different existing agents. The probability that a new agent u forms an edge with an existing agent v is proportional to the degree of v.

Our numerical study suggests that for networks generated by the above process, our model performs quite well for $m \geq 3$. For the sake of brevity, we only present one instance of our numerical study, as shown in the left panel of Figure 4. To create the left panel, we first generate one specific network using the Barabási–Albert model with $n = 5000$, $m = 5$, and 5 initial agents. We then simulate 25 instances of diffusion on that network when randomly seeding 1% of agents. Finally, using that network's degree distribution, we use our model to predict the diffusion trajectory and compare that to a mean-field approximation. As we can see, the solution resulting from a mean-field approximation overestimates the proportion of adopters, while our model is close to (and usually within) the confidence interval of simulated diffusion trajectories.

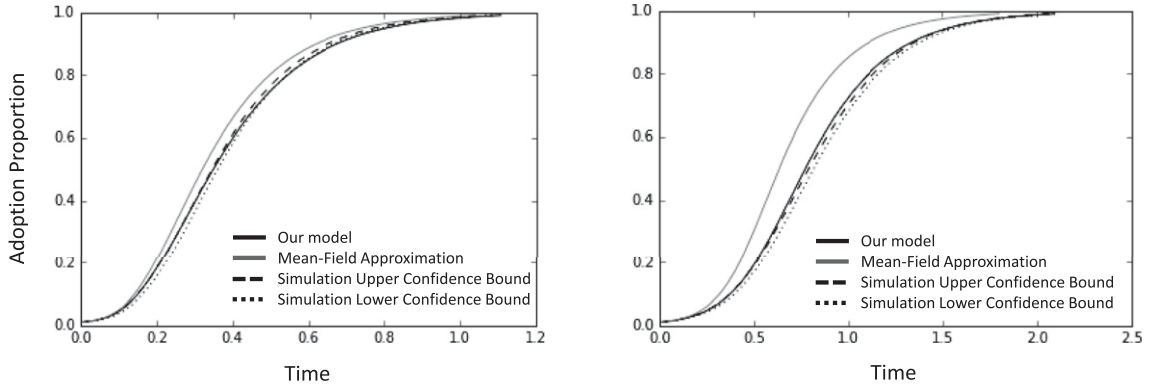
Many networks also exhibit clustering, meaning an agent's neighbors are more likely to be themselves neighbors. A network's clustering behavior is commonly measured by its clustering coefficient defined as

$$C(V) = \frac{\sum_{i,j,k \in V} e_{ij} e_{ik} e_{jk}}{\sum_{i,j,k \in V} e_{ij} e_{ik}},$$

where e_{uv} is an indicator variable that equals 1 if there is an edge between u and v with $u \neq v$. The clustering coefficients of networks generated using the configuration model or the preferential attachment mechanism (described above) are small; in fact, they asymptotically approach zero. Thus, mechanisms such as friends-of-friends have been proposed (Jackson and Rogers 2007) to generate networks with non-negligible clustering coefficients.

Here, we describe a simple version of the friends-of-friends mechanism and then examine the accuracy of our solution when applied to such networks. Similar to the Barabási–Albert model, the Jackson–Rogers model begins with a small connected network and sequentially adds agents and edges. Each arriving agent u meets m_r previously arrived agents and forms an undirected edge with each of them. This set of agents is called u's parents. Then, u meets m_n agents selected uniformly at random from the union of her parents' neighbors, excluding her parents. Agent u forms an undirected edge with each of these m_n agents.[14]

Figure 4. Time Required to Reach an Adoption Proportion of $0.01 \leq s \leq 0.99$ in Our Model, in a Mean-field Approximation Given by (31), and in a Simulated Network with 5,000 Agents



Notes. Left: for a network formed via preferential attachment with $m = 5$ (Barabási–Albert model); right: for a network formed via the friends-of-friends mechanism with $m_r = 2$ and $m_n = 1$ (Jackson–Rogers model). In both settings, a fraction $q = 0.01$ of agents are selected uniformly at random and seeded.

If we set $m_r = 2$ and $m_n = 1$, then by Theorem 2 in Jackson and Rogers (2007), the clustering coefficient of a network generated by the friends-of-friends mechanism tends to approximately 0.105.[15] Such a clustering coefficient is within the range of clustering coefficients estimated in real networks, such as subnetworks of the World Wide Web (Jackson and Rogers 2007). We would expect this clustering behavior to slow down the diffusion process, since adopters would be even more likely to contact other adopters. However, our model still performs quite well empirically. Again, for the sake of brevity, we only present one instance of our extensive numerical study, as shown in the right panel of Figure 4. The right panel is created in a similar manner to the left panel: we first generate one specific network using the friend-of-friends mechanism with $n = 5000, m_r = 2$, $m_n = 1$, and 5 initial agents. We then simulate 25 samples of diffusion on that network when randomly seeding 1% of agents. Finally, using that network's degree distribution, we use our model to predict the diffusion trajectory and compare that to a mean-field approximation. We make the following observations: the solution resulting from a mean-field approximation overestimates the diffusion trajectory as it ignores both the clustering and lack of perfect mixing (as discussed in Section 7.3); the outcome of our analysis also slightly overestimates as it ignores the clustering behavior of the underlying network. However, its predicted trajectory is much closer to the simulated one.

8. Conclusion

In large social networks, where each agent only interacts with a few others, patterns of connections between agents (also known as network structure) play an important role in the diffusion of an innovation or a new product. Random networks with given degree distributions are extensively used to model such social networks, as they capture the heterogeneity in agents' degree of connections without requiring full information about the network. However, exact characterization a diffusion process in random networks and an understanding of the impact of degree distribution on diffusion remained largely unsolved problems, perhaps because of the high dimensionality of the underlying system. For the most part, these problems have been studied using mean-field approximations without quantifying the approximation error. In this paper, we proposed a framework to rigorously study diffusion processes in random networks without resorting to approximation methods such as mean-field. The novelty of our analysis lies in exploiting an iterative process, known as the configuration model, to interleave the graph generation and the diffusion process. This allows us to significantly reduce the dimension of the underlying stochastic process.

Aside from our methodological contributions, we derived insights into the impact of degree distribution by conducting comparative statics. We demonstrated a trade-off between contact cost-efficiency and fast growth in a network. While degree distributions with minimum variance in degree are cost-efficient, the diffusion in such networks grows slower than the diffusion in networks with high variability in degree. We also study the problem of optimal seeding to minimize cost or time. Somewhat counterintuitively, our results imply that maximally seeding high-degree agents does not necessarily minimize either the contact cost or the time required to reach a target level of adopters.

Beyond the questions studied in this paper, several other operational and managerial questions arise in the presence of product diffusion (sometimes referred to as social learning or word-of-mouth effects). For example, how to manage demand when supply is constrained (Ho et al. 2002, Kumar and Swaminathan 2003, Shen et al. 2011), how to price optimally (Shen et al. 2014), and how to facilitate diffusion for the adoption of green technologies (Diaz-Rainey and Tzavara 2012, Alizamir et al. 2016). These questions have been mainly studied utilizing the Bass model, which abstracts away from heterogeneity in the agents' level of connections. We believe our framework can serve as a starting point to study the above questions in a richer setting that incorporates the network structure.

Acknowledgments

The authors would like to thank the area editor, associate editor, and two anonymous referees for their valuable comments. Further, they would like to acknowledge Jiwoong Shin, Edward Kaplan, Ramesh Johari, Michael Blair, and participants at NetEcon 2018, the 2018 INFORMS Annual Meeting, and 2019 INFORMS RMP for their helpful suggestions.

Endnotes

[1] We use the following pairs of terms interchangeably: friends and neighbors; firm and marketer; and consumer and agent.

[2] We also note that in the context of epidemiology, several papers study the dual problem of who to immunize in order to control/minimize the network epidemic (see Borgs et al. 2010, Drakopoulos et al. 2016, and references therein).

[3] We assume that q is a constant independent of network size n.

[4] Recall that we defined $I = \{1, 2, \ldots, |I|\}$ to be the set of indices of possible degrees; further, we assume that $k_{\min} \geq 3$ and $k_{\max}$ is bounded.

[5] We note that, for some n, it may not be possible to have a network with n agents, with degree distribution $\{(p_i, k_i), i \in I\}$, and with seeding configuration $\{(q_i), i \in I\}$—for example, np_i is not an integer. Thus, a more accurate statement is to consider the subsequence of networks with sizes $n_1, n_2, \ldots$ such that a network with the given

number of agents and the degree distribution exists. However, for the sake of brevity, we do not mention this in our statements.

[6] For large values of q, it is possible that there is no solution in the interval $(0,1)$. In that case, $s_p = q$.

[7] For two functions $d, l : \mathbb{N} \to \mathbb{R}$, $l(n) = o(d(n))$ if $\lim_{n\to\infty} \frac{l(n)}{d(n)} = 0$.

[8] The skewness of a degree distribution is given by $\sum_{i\in I} p_i(k_i - k_{\text{avg}})^3/\sigma^3$.

[9] Further, we have examples (not included in the paper) that show $\bar{s}$ is not a tight upper bound—that is, s can be higher than $\bar{s}$, but seeding only the highest-degree agents remains optimal.

[10] In fact, it is not possible to generate a random network with degree distribution $\{(3, 0.5), (10, 0.5)\}$ by removing edges from another random network with degree distribution $\{(4, 0.5), (10, 0.5)\}$. To see this, note that in the latter network, there exists a constant fraction of degree 4 agents who are only connected to agents with degree 10. Suppose agent v is one such agent. Removing an edge from v will result in the existence of a degree 9 agent; on the other hand, not removing an edge from v will imply that v's degree will remain 4.

[11] As a technical detail, to ensure that the process runs for $\frac{1}{2}nk_{\text{avg}}$ iterations, we must designate what to do if the set of active half-edges becomes empty (this is case (b) in Algorithm 1).

[12] Later, when we incorporate the diffusion time, we let e_1 in Algorithm 1 be the active half-edge whose clock wakes up first.

[13] For two functions $d, l : \mathbb{N} \to \mathbb{R}$, $l(n) = O(d(n))$ if $\limsup_{n\to\infty} \frac{|l(n)|}{d(n)} < \infty$.

[14] In the more general version, edges can be directed and the probability of forming an edge can be less than 1.

[15] The calculation itself is based on a mean-field analysis.

References

Ajorlou A, Jadbabaie A, Kakhbod A (2017) Dynamic pricing in social networks: The word-of-mouth effect. *Management Sci.* 64(2): 971–979.

Akbarpour M, Malladi S, Saberi A (2018) Diffusion, seeding, and the value of network information. *Proc. 2018 ACM Conf. Econ. Comput.* (ACM, New York), 641.

Alizamir S, de Véricourt F, Sun P (2016) Efficient feed-in-tariff policies for renewable energy technologies. *Oper. Res.* 64(1):52–66.

Allon G, Zhang DJ (2017) Managing service systems in the presence of social networks. Working paper, The Wharton School, Philadelphia.

Anderson R, Ashlagi I, Gamarnik D, Kanoria Y (2017) Efficient dynamic barter exchange. *Oper. Res.* 65(6):1446–1459.

Anderson RM, May RM (1984) Spatial, temporal, and genetic heterogeneity in host populations and the design of immunization programmes. *Math. Med. Biol.* 1(3):233–266.

Aral S, Muchnik L, Sundararajan A (2013) Engineering social contagions: Optimal network seeding in the presence of homophily. *Network Sci.* 1(2):125–153.

Ashlagi I, Burq M, Jaillet P, Manshadi V (2019) On matching and thickness in heterogeneous dynamic markets. *Oper. Res.* 67(4):927–949.

Banerjee A, Chandrasekhar AG, Duflo E, Jackson MO (2013) The diffusion of microfinance. *Science* 341(6144):1236498.

Barabási A-L, Albert R (1999) Emergence of scaling in random networks. *Science* 286(5439):509–512.

Bass FM (1969) A new product growth for model consumer durables. *Management Sci.* 15(5):215–227.

Bayati M, Montanari A, Saberi A (2016) Generating random networks without short cycles. Working paper, Stanford University, Stanford, CA.

Beaman L, BenYishay A, Magruder J, Mobarak AM (2018) Can network theory-based targeting increase technology adoption? NBER Working Paper 24912, National Bureau of Economic Research, Cambridge, MA.

Belloni A, Deng C, Pekeč S (2017) Mechanism and network design with private negative externalities. *Oper. Res.* 65(3):577–594.

Bimpikis K, Ozdaglar A, Yildiz E (2016) Competitive targeted advertising over networks. *Oper. Res.* 64(3):705–720.

Borgs C, Chayes J, Ganesh A, Saberi A (2010) How to distribute antidote to control epidemics. *Random Structures Algorithms* 37(2): 204–222.

Britton T, Deijfen M, Martin-Löf A (2006) Generating simple random graphs with prescribed degree distribution. *J. Statist. Phys.* 124(6): 1377–1397.

Campbell A (2013) Word-of-mouth communication and percolation in social networks. *Amer. Econom. Rev.* 103(6):2466–2498.

Candogan O, Bimpikis K, Ozdaglar AE (2012) Optimal pricing in networks with externalities. *Oper. Res.* 60(4):883–905.

Cohen MC, Harsha P (2020) Designing price incentives in a network with social interactions. *Manufacturing Service Oper. Management* 22(2):292–309.

Dembo A, Montanari A (2010) Ising models on locally tree-like graphs. *Ann. Appl. Probab.* 20(2):565–592.

Diaz-Rainey I, Tzavara D (2012) Financing the decarbonized energy system through green electricity tariffs: A diffusion model of an induced consumer environmental market. *Tech. Forecasting Soc. Change* 79(9):1693–1704.

Ding Y, Ge D, He S, Ryan CT (2015) A non-asymptotic approach to analyzing kidney exchange graphs. *Proc. 16th ACM Conf. Econom. Comput.* (ACM, New York), 257–258.

Dover Y, Goldenberg J, Shapira D (2012) Network traces on penetration: Uncovering degree distribution from adoption data. *Marketing Sci.* 31(4):689–712.

Draief M, Ganesh A, Massoulié L (2008) Thresholds for virus spread on networks. *Ann. Appl. Probab.* 18(2):359–378.

Drakopoulos K, Ozdaglar A, Tsitsiklis JN (2016) When is a network epidemic hard to eliminate? *Math. Oper. Res.* 42(1):1–14.

Durrett R (2006) *Random Graph Dynamics*. Cambridge Series in Statistical and Probabilistic Mathematics (Cambridge University Press, New York).

Fainmesser IP, Galeotti A (2016) Pricing network effects. *Rev. Econom. Stud.* 83(1):165–198.

Fazeli A, Jadbabaie A (2012) Targeted marketing and seeding products with positive externality. *Proc. 50th Annual Allerton Conf. Comm. Control Comput.* (IEEE, New York), 1111–1117.

Fibich G, Gibori R (2010) Aggregate diffusion dynamics in agent-based models with a spatial structure. *Oper. Res.* 58(5): 1450–1468.

Goldenfeld N (2018) *Lectures on Phase Transitions and the Renormalization Group* (CRC Press, Boca Raton, FL).

Golub B, Jackson MO (2012a) Does homophily predict consensus times? Testing a model of network structure via a dynamic process. *Rev. Network Econom.* 11(3), https://doi.org/10.1515/1446-9022.1367.

Golub B, Jackson MO (2012b) How homophily affects the speed of learning and best-response dynamics. *Quart. J. Econom.* 127(3): 1287–1338.

Goyal S, Heidari H, Kearns M (2014) Competitive contagion in networks. *Games Econom. Behav.* 113:58–79.

Granovetter M (1978) Threshold models of collective behavior. *Amer. J. Sociol.* 83(6):1420–1443.

Hartline J, Mirrokni V, Sundararajan M (2008) Optimal marketing strategies over social networks. *Proc. 17th Internat. Conf. World Wide Web* (ACM, New York), 189–198.

Hinz O, Skiera B, Barrot C, Becker JU (2011) Seeding strategies for viral marketing: An empirical comparison. *J. Marketing* 75(6): 55–71.

Ho T-H, Savin S, Terwiesch C (2002) Managing demand and sales dynamics in new product diffusion under supply constraint. *Management Sci.* 48(2):187–206.

Jackson MO (2010) *Social and Economic Networks* (Princeton University Press, Princeton, NJ).
Jackson MO, López-Pintado D (2013) Diffusion and contagion in networks with heterogeneous agents and homophily. *Network Sci.* 1(1):49–67.
Jackson MO, Rogers BW (2007) Meeting strangers and friends of friends: How random are social networks? *Amer. Econom. Rev.* 97(3):890–915.
Jackson MO, Yariv L (2005) Diffusion on social networks. *Economie Publique* 16(1):3–16.
Jackson MO, Yariv L (2007) Diffusion of behavior and equilibrium properties in network games. *Amer. Econom. Rev.* 97(2):92–98.
Keeling MJ (1999) The effects of local spatial structure on epidemiological invasions. *Proc. Royal Soc. London B Biol. Sci.* 266(1421): 859–867.
Kempe D, Kleinberg J, Tardos É (2003) Maximizing the spread of influence through a social network. *Proc. 9th ACM SIGKDD* (ACM, New York), 137–146.
Kermack WO, McKendrick AG (1927) A contribution to the mathematical theory of epidemics. *Proc. Royal Soc. London A Math. Phys. Engrg. Sci.* 115(772):700–721.
Kumar S, Swaminathan JM (2003) Diffusion of innovations under supply constraints. *Oper. Res.* 51(6):866–879.
Leduc MV, Jackson MO, Johari R (2017) Pricing and referrals in diffusion on networks. *Games Econom. Behav.* 104:568–594.
Lim Y, Ozdaglar A, Teytelboym A (2015) A simple model of cascades in networks. Technical report, Google, Mountain View, CA.
Lobel I, Sadler E, Varshney LR (2017) Customer referral incentives and social media. *Management Sci.* 63(10):3514–3529.
López-Pintado D (2008) Diffusion in complex social networks. *Games Econom. Behav.* 62(2):573–590.
Makhdoumi A, Malekian A, Ozdaglar AE (2017) Strategic dynamic pricing with network externalities. arXiv (submitted June 4), https://arxiv.org/abs/1706.01131.
Massoulie L, Draief M (2010) *Epidemics and Rumours in Complex Networks* (Cambridge University Press, Cambridge, UK).
McKay B (1985) Asymptotics for symmetric 0-1 matrices with prescribed row sums. *Ars Combinatoria* 19A:15–20.
Mitzenmacher M, Upfal E (2005) *Probability and Computing: Randomized Algorithms and Probabilistic Analysis* (Cambridge University Press, Cambridge, UK).
Molloy M, Reed B (1995) A critical point for random graphs with a given degree sequence. *Random Structures Algorithms* 6:161–180.
Momot R, Belavina E, Girotra K (2019) The use and value of social information in selective selling of exclusive products. *Management Sci.*, ePub ahead of print, October 15, https://doi.org/10.1287/mnsc.2019.3310.
Morris S (2000) Contagion. *Rev. Econom. Stud.* 67(1):57–78.
Nielsen (2012) Consumer trust in online, social and mobile advertising grows. Accessed November 15, 2019, http://www.nielsen.com/us/en/insights/news/2012/consumer-trust-in-online-social-and-mobile-advertising-grows.html.
Rahmandad H, Sterman J (2008) Heterogeneity and network structure in the dynamics of diffusion: Comparing agent-based and differential equation models. *Management Sci.* 54(5):998–1014.
Shah D, Zaman T (2016) Finding rumor sources on random trees. *Oper. Res.* 64(3):736–755.
Shakkottai S, Johari R (2010) Demand-aware content distribution on the internet. *IEEE/ACM Trans. Networking* 18(2):476–489.
Shen W, Duenyas I, Kapuscinski R (2011) New product diffusion decisions under supply constraints. *Management Sci.* 57(10): 1802–1810.
Shen W, Duenyas I, Kapuscinski R (2014) Optimal pricing, production, and inventory for new product diffusion under supply constraints. *Manufacturing Service Oper. Management* 16(1):28–45.
Sunar N, Birge JR, Vitavasiri S (2019) Optimal dynamic product development and launch for a network of customers. *Oper. Res.* 67(3):770–790.
Van Der Hofstad R (2016) *Random Graphs and Complex Networks*, vol. 1 (Cambridge University Press, Cambridge, UK).
Wormald NC (1995) Differential equations for random processes and random graphs. *Ann. Appl. Probab.* 5(4):1217–1235.
Wormald NC (1999) Models of random regular graphs. *Surveys in Combinatorics* (Cambridge University Press, Cambridge, UK), 239–298.

Vahideh Manshadi is an associate professor of operations at Yale School of Management with a secondary appointment in the Department of Statistics and Data Science. She was previously a postdoc at MIT Operations Research Center and received a PhD from Stanford University and BSc from Sharif University of Technology. Her research interests include sequential and dynamic decision making with applications to the design of online and matching marketplaces, and information platforms and networks.

Sidhant Misra received his PhD in electrical engineering and computer science from the Massachusetts Institute of Technology in 2014. He is currently a research scientist in the Theory Division and the Advanced Network Science Initiative at the Los Alamos National Laboratory. His research interests lie at the intersection of machine learning and optimization with applications in network structure identification, planning, and operations of energy networks.

Scott Rodilitz is a PhD student at the Yale School of Management. He is interested in the growth of matching markets, network effects, and applications in humanitarian logistics. Before coming to Yale, he received his BA in mathematics and economics from Williams College.