

# Strategic diffusion in networks through contagion<sup>☆</sup>

Théophile T. Azomahou<sup>a,c</sup>, Daniel C. Opolot<sup>b,\*</sup>

<sup>a</sup>Maastricht University, School of Business and Economics, Tongersestraat 53, 6211 LM Maastricht

<sup>b</sup>AIFMRM, University of Cape Town, Private Bag X3, Rondebosch, 7701 Cape Town

<sup>c</sup>University of Clermont Auvergne and CERDI, 65 Boulevard F. Mitterrand, 63009 Clermont-Ferrand CEDEX 1

---

## Abstract

Social networks provide a platform through which firms and governments can strategically diffuse products and desirable practices. This paper studies the process of strategic diffusion (i.e. diffusion by targeting specific individuals) in the presence of network externalities using an evolutionary game theory framework. We show that targeting is inefficient if the rate at which agents experiment on different choices is high, but below a certain threshold, targeting is economically reasonable. A firm/planner can reduce the cost of targeting by exploiting the process of contagion. A contagious choice requires a small set of initial adopters to trigger diffusion to the wider network. We show that the cost of making a product/practice contagious is lower for sparsely and uniformly connected networks than highly connected and/or less cohesive networks. We also show that the expected waiting time until a contagious choice is adopted by the entire network is independent of the population size. This implies that in large networks, even if the level of experimentation is very low, the diffusion process does not get trapped indefinitely in a suboptimal equilibrium.

*Keywords:* Strategic diffusion, stochastic evolution, networks, contagion, expected waiting time

JEL: D8, C73

---

## 1. Introduction

The diffusion of products and practices through social influence has been recognized for the past five decades.<sup>1</sup> The growth of internet and online social networks has increased interest in how firms and governments can harness the power of social networking to promote private and social goals. An active line of research seeks to examine how to target specific individuals in a network to trigger diffusion of a product or behaviour to a larger proportion of the population.<sup>2</sup> Although the literature offers many insights on the relationship between targeting strategies and individual connectivity in the network, many questions regarding when to engage in and the costs of targeting remain unanswered.

---

\*Corresponding author:

Email address: [daniel.opolot@uct.ac.za](mailto:daniel.opolot@uct.ac.za) (Daniel C. Opolot)

<sup>1</sup> See for example [Coleman et al. \(1966\)](#) on the diffusion of medical innovations, [Foster and Rosenzweig \(1995\)](#); [Conley and Udry \(2010\)](#) on diffusion of agricultural products and practices, [Reingen et al. \(1984\)](#); [Godes and Mayzlin \(2004\)](#) on diffusion of brand choices by consumers.

<sup>2</sup> See for example [Galeotti and Goyal \(2009\)](#), [Campbell \(2013\)](#), [Goyal et al. \(2014\)](#), [Chatterjee and Dutta \(2016\)](#), [Leduc et al. \(2017\)](#), [Tsakas \(2017\)](#), [Beaman et al. \(2018\)](#) and [Beaman and Dillon \(2018\)](#).

When is it economically reasonable to target specific individuals? Put differently, should a firm/planner engage in targeting or simply introduce the product/behaviour to the population (e.g. through advertising, or civic education in the case of behavioural practices) and wait for some individuals to experiment and adopt it? What are the costs associated with targeting and how can they be reduced? The intuition behind the first two questions is that some individuals, referred to as *innovators* in the diffusion literature (Bass, 1969), have a higher tendency to experiment and try-out new or different products/practices. In such situations, it may not be *economically reasonable* (i.e. resource wasteful) to conduct targeting because, (a) the agents to be targeted will experiment and switch to a desired choice anyway; (b) targeted agents may experiment and switch to a different choice. If the extent to which individuals experiment on different choices is sufficiently high, then it may be desirable to wait for innovators to trigger the diffusion process, or at most perform limited targeting.<sup>3</sup>

This paper addresses these questions through an evolutionary game theory framework. We suppose that there is a finite set of agents who interact through a social network. Agents repeatedly revise their choices/actions over time, and their choices exhibit network externalities. Network externalities exist in adoption of information technologies, fashion products, social and institutional norms, and political actions, and are characterized by inertia in that nobody adopts if no one else has adopted.

To model the tendency to experiment on choices, we consider an evolutionary process where the probability of choosing a given option is separable into components capturing social influence and individual experimentation. Including experimentation in the decision-making process is a standard assumption in evolutionary game theory and evolutionary economics. It is akin to the models of *bounded rationality*: the notion that although individuals aim to make optimal choices, their decisions are frequently subject to mistakes. In the well-studied Bass and related diffusion models, experimentation is interpreted as innovation so that the society can be divided into two categories based on the timing of adoption: *innovators* as early adopters, and *imitators* as late adopters. Here, we instead treat experimentation as a behavioural aspect of the decision process so that every agent experiments with a some probability. The social influence component of choice probabilities captures network externalities. We also assume that agents *myopically best respond* to the distribution of their neighbours' choices. Assuming myopia in the decision-making process is consistent with the notion of bounded rationality in that agents are not only prone to mistakes but are also incapable of keeping memory of the entire history of play.

There are several key differences between our model and existing models of diffusion in the literature. Firstly, in contrast to the Bass and related models of diffusion for first-time purchases (see Mahajan et al.

---

<sup>3</sup>For example, higher levels of experimentations are observable in the market for messaging and chat mobile applications (e.g. AOL Instant Messenger, Google talk, VOIP, skype, Kik., whatsapp, snapchat, HipChat and slack) where multiple new apps are introduced to the market annually and for each, there is always a considerable number of users willing to experiment and try them out. Some apps subsequently failed to penetrate the market, but many have successfully diffused to a larger proportion of the population starting from few experimenters.

(1990) and Rogers (2003) for surveys), we model a diffusion process where agents repeatedly replace the old, spoilt or no longer preferred products with new ones of the same type/brand, or with different competing products. The role that the process of experimentation plays in triggering network-wide diffusion is thus relevant in our model and less so in the models of first-time purchases. Secondly, in contrast to the mean field models of diffusion (e.g. López-Pintado (2008), Jackson and Yariv (2007), Galeotti and Goyal (2009) and Kreindler and Young (2014)) that assume infinitely large and homogeneously mixed population, we study a stochastic diffusion process on finite networks. Due to the stochastic nature of our model, the long-run probabilities are a suitable measure for characterizing the diffusion process. Moreover, unlike mean field diffusion processes that generate smooth diffusion curves, finite-population diffusion processes exhibit threshold dynamics. Our model is thus useful to firms and governments that care about the long-run survival and not just first-time purchases of their products and practices.

We proceed by first establishing conditions under which it is reasonable to engage in targeting. We do so by examining the effect of the level of experimentation on long-run probabilities. We find that long-run probabilities exhibit threshold dynamics with respect to experimentation parameter.<sup>4</sup> Specifically, there exists a threshold level of experimentation below which targeting is reasonable. Below the threshold, best response dynamics dominates experimentation so that individual decisions are influenced more by their neighbours' choices and less by randomness. Above the threshold level of experimentation, the diffusion process can enter and exit the desired state through experimentation, which makes targeting economically unreasonable.

Next, we show that a firm/planner can reduce the cost of targeting by exploiting the process of *contagion*. Contagion refers to the process where a choice spreads to a larger subregion or entire network through best response dynamics (i.e. social influence) once it has been adopted by a small number of agents—*initial adopters*. Let  $a$  be the choice a firm/planner aims to diffuse and  $\mathbf{a}$  the corresponding equilibrium configuration where  $a$  is played; that is,  $\mathbf{a}$  is a configuration that a firm/planner aims to steer the population to. Then the cost of targeting is proportional to the minimum number of experimentations (mutations) involved in transitions starting from every other equilibrium configuration and ending in  $\mathbf{a}$ . If choice  $a$  is *globally pairwise contagious* (i.e. contagious relative to every other choice), then the minimum number of mutations, and hence targeting cost, is very small. Consequently, an equilibrium state containing a globally pairwise contagious choice has the highest long-run probability.

The properties that determine the feasibility of contagion are: (i) a unique property of the network called *contagion threshold* and (ii) a property of payoffs called the *relative payoff switching-losses* between any pair of choices (i.e a loss an individual incurs from switching to a choice different from that adopted by her neighbours). Pairwise contagion of choice  $a_i$  relative to another choice  $a_j$  occurs if and only if the relative switching-loss is less than the contagion threshold. Highly connected and/or less cohesive

---

<sup>4</sup>That is, for any two configurations say  $\mathbf{x}$  and  $\mathbf{y}$  describing a profile of choices in the population, there exists an optimal level of experimentation above which a switch in the relationship between long-run probabilities of  $\mathbf{x}$  and  $\mathbf{y}$  occurs.

networks have smaller contagion thresholds; and the smaller the relative payoff-loss of  $a_l$  compared to  $a_j$ , the larger the relative utility of  $a_l$ .

A firm/planner aiming to diffuse a choice through contagion then incurs a second type of cost (different from targeting cost), here referred to as *contagion cost*. Given a network structure, the contagion cost is the cost that must be incurred in order to make the relative switching-losses sufficiently small (e.g. by increasing innovativeness and functionality of a product, reducing the price, making it user friendly, or enforcing penalties in cases where the choice is a behaviour) to satisfy the conditions for pairwise contagion. Following the discussion in the preceding paragraph, contagion cost is high for highly connected and/or cohesive network. Thus, it is relatively less costly to diffuse a choice to less connected and uniform networks through contagion, than to highly connected and/or less cohesive networks.

As an example, our findings are consistent with observable patterns of successful diffusion of products in markets for information technologies (I.T). Specifically, there is observable evidence of a high turnover and entry rate into the market for Messaging and Chat internet apps serving a similar purpose (e.g. AOL Instant Messenger, Google talk, VOIP, skype, Kik., whatsapp, snapchat, HipChat and slack) compared to the market for Social Networking apps (e.g. Friendster, Myspace and Facebook ). Once Facebook dominated the market for Social Networking apps, many new entrants (e.g. iTunes Ping, Eons, Orkut, e.t.c) tried to penetrate the market but were not as successful. Our findings suggest that the discrepancy could be due to differences in network structures. Messaging and Chat app networks are sparse and uniform (individuals interact with a few close friends and family) compared to Social Networking apps (i.e. individuals have larger numbers of friends on Facebook and the distance between any two users is small). Thus, to successfully penetrate the network (market) for messaging and chat apps, the new product needs to be only slightly better than those in the market. To penetrate a Facebook-type network however requires the new product to be far better, or simply offer different services (e.g. Youtube Pinterest, e.t.c).

In our final analysis, we show that not only does contagion reduce costs of targeting, it also drastically reduces the waiting time for the population to reach an equilibrium state containing a globally contagious choice. That is, the expected waiting time from any configuration to an equilibrium configuration containing a globally contagious choice is independent of the population size. The implication is that in large networks, even if the level of experimentation is very low, the diffusion process does not get trapped indefinitely in an equilibrium where agents choose a non-contagious (non-globally stable) option.

Our work is related to a large literature across disciplines on how firms and governments can harness the power of social networks to promote products and desirable practices. It includes [Richardson and Domingos \(2002\)](#); [Kempe et al. \(2003\)](#); [Chen et al. \(2009\)](#) in computer science, [Kirby and Marsden \(2006\)](#); [Kiss and Bichler \(2008\)](#) in marketing, and [Kelly et al. \(1991\)](#); [Valente et al. \(2003\)](#) in public health. A closely related literature in economics includes [Galeotti and Goyal \(2009\)](#), [Campbell \(2013\)](#), [Goyal et al. \(2014\)](#), [Chatterjee and Dutta \(2016\)](#), [Leduc et al. \(2017\)](#), [Tsakas \(2017\)](#), [Beaman et al. \(2018\)](#)

and [Beaman and Dillon \(2018\)](#). These papers however focus on examining how targeting strategies are influenced by the overall network structure and characteristics of individual connectivity, in particular, centrality measures such as degree, eigenvector, Bonacich and betweenness centralities. Our focus is instead on examining when targeting is reasonable, and ways in which targeting costs may be reduced.

The second strand of literature that is closely related to our paper is that on diffusion in the presence of network externalities ([Morris, 2000](#); [López-Pintado, 2008](#); [Jackson and Yariv, 2007](#); [Sundararajan, 2007](#); [Galeotti et al., 2010](#); [Galeotti and Goyal, 2009](#)). Except for [Morris \(2000\)](#), these papers study binary choice diffusion processes on random networks. In addition to the differences discussed above, our paper also considers a general diffusion process with multiple choices. Many markets, such as the I.T market discussed above, have multiple products and the interrelationship between their payoffs becomes complex as the number increases. Thus, the conclusions from a model of diffusion with only two choices do not necessarily generalise to situations with multiple choices. [Morris \(2000\)](#) studies conditions for pairwise contagion in deterministic networks, and our paper generalizes the notion of pairwise contagion and shows how firms can strategically exploit it to diffuse products.

Our paper is also related to evolutionary game theory literature that studies convergence rates.<sup>5</sup> For example [Ellison \(1993, 2000\)](#) shows that learning is fast in some families of networks, [Montanari and Saberi \(2010\)](#); [Young \(2011\)](#) show that diffusion is fast in networks made up of cohesive subgroups, and [Kreindler and Young \(2013, 2014\)](#) show that learning is fast if the level of experimentation is sufficiently large. We show that learning is fast in the presence of a contagious choice.

The remainder of the paper is organized as follows. In Section 2, we introduce a model of diffusion in the presence of network externalities. Section 3 characterizes long-run probabilities, and shows how they vary with the parameter of experimentation. Section 4 discusses the notion of contagion and shows how it can be used to diffuse a specific choice. Section 5 examines the relationship between contagion and expected waiting time, and a conclusion is offered in Section 6. All lengthy proofs are relegated to the Appendix.

## 2. The model

### 2.1. Actions, network and payoffs

We study the diffusion of products and behaviour with network effects through social networks. The simplest way to model network effects is through strict coordination games. Let  $A = \{a_1, \dots, a_j, \dots, a_m\}$

---

<sup>5</sup>There is a literature on evolutionary game dynamics in networks that includes [Ellison \(1993, 2000\)](#), [Blume \(1995\)](#), [Robson and Vega-Redondo \(1996\)](#), [Anderlini and Ianni \(1996\)](#), [Berninghaus and Schwalbe \(1996\)](#), [Lee and Valentinyi \(2000\)](#), [Lee et al. \(2003\)](#), [Alós-Ferrer and Weidenholzer \(2007\)](#) and [Alós-Ferrer and Weidenholzer \(2008\)](#). These papers focus on identifying stochastically stable states by considering specific network structures, such as ring and 2-dimensional grid structures. But as demonstrated in [Alós-Ferrer and Weidenholzer \(2007\)](#), stochastic stability is not robust to interaction structure. We consider general networks and identify the contagion threshold as the network property that matters for long-run stability.

be the choice or strategy set, and  $U$  is an  $m \times m$  symmetric payoff matrix whereby  $u(a_j, a_l)$  is a payoff to choice  $a_j$  when an opponent chooses  $a_l$ . The payoff matrix  $U$  is a *strict coordination property*, if for each  $a_j \in A$ ,  $u(a_j, a_j) > u(a_j, a_l)$  for all  $a_l \neq a_j$ ; that is, the payoff from coordination is always higher than the payoff from mis-coordination, and hence, coordinating on a given choice is pure strategy Nash equilibrium.

Let  $N = \{1, \dots, i, \dots, n\}$  denote a set of agents who interact through a social network represented by an  $n \times n$  interaction matrix  $G$  (with slight abuse of notation, we also use  $G$  to denote the underlying network). We place the following restrictions of  $G$ :

- (i)  $G$  is undirected and unweighted. Let  $G_{ik}$  be the  $i$ th row and  $k$ th column of  $G$ , denoting a link from  $i$  to  $k$ . If such a link exists, then  $G_{ik} = G_{ki} = 1$ , otherwise,  $G_{ik} = G_{ki} = 0$ .
- (i)  $G$  is strongly connected. That is, for every pair of agents  $i, k \in N$ , there exists a path of (undirected) links from  $i$  to  $k$ .

Both restrictions are simplifications that ensure that the evolutionary process (described below) converges. Directed and/or not strongly connected networks can lead to cyclic interactions and/or isolated nodes, which in turn inhibits convergence. We denote by  $N_i$  the number of  $i$ 's neighbours, that is,  $N_i = \{k \mid G_{ik} = 1\}$ ; and by  $n_i$  the cardinality of  $N_i$ , the number of  $i$ 's neighbours also commonly referred to as the *degree* of  $i$ .

Let  $\sigma_i = (\sigma_i(a_1), \dots, \sigma_i(a_m))$  be the empirical distribution of strategies in  $i$ 's neighbourhood. That is,  $\sigma_i$  is a mixed strategy representing the proportion of  $i$ 's neighbours playing each pure strategy. Then the total payoff that  $i$  receives from playing strategy  $a_j$  against  $\sigma_i$  is

$$U(a_j \mid \sigma_i) = \sum_{a_l \in A} \sigma_i(a_l) u(a_j, a_l). \quad (1)$$

Note however that the total payoff is independent of the size of a player's neighbourhood and depends only on the empirical distribution  $\sigma_i$ . Thus, where no confusion arises, we simply write  $U(a_j \mid \sigma)$  for the total payoff to any player choosing  $a_j$  against distribution  $\sigma$ , where  $\sigma = (\sigma(a_1), \dots, \sigma(a_m))$ . We refer to the double  $(U, G)$  as a *local interaction strict coordination game*.

## 2.2. Dynamics

Given payoff and interaction structures, we consider a diffusion process where agents revise their choices over time. That is, we model the diffusion of products where agents can replace the old, spoilt or no longer preferred products with new ones of the same type/brand, or with different competing products. This is opposed to the Bass diffusion model and related models for first-time purchases (Bass, 1969). But as in the Bass and related diffusion models, we assume that the probability that an agent chooses a given option is a function of social influence and individual experimentation.

Let small bold letters e.g.  $\mathbf{x}, \mathbf{y}, \mathbf{z}, \dots$  denote vectors or configurations of choices in the population. That is, each  $\mathbf{x} = (x^1, \dots, x^n)$ , where  $x^i$  is the strategy of the  $i$ th player. We write  $\mathbf{X}$  for the set of all possible configurations. The cardinality of  $\mathbf{X}$  is  $m^n$ , where  $m$  is the number of strategies.

For each player  $i$ , let  $\sigma_i(a_l; \mathbf{x})$  be the proportion of  $i$ 's neighbours playing strategy  $a_l$  in profile  $\mathbf{x}$ , and let  $\sigma_i(\mathbf{x}) = (\sigma_i(a_1; \mathbf{x}), \dots, \sigma_i(a_m; \mathbf{x}))$ . Thus, if  $\mathbf{x}_t$  is the strategy profile at time  $t$ , then the total payoff  $U_i(a_j, \mathbf{x}_t)$  to player  $i$  for adopting strategy  $a_j$  against a configuration  $\mathbf{x}$ , or equivalently, against a mixture  $\sigma_i(\mathbf{x})$ , is  $U_i(a_j, \mathbf{x}_t) = \sum_{a_l \in A} \sigma_i(a_l; \mathbf{x}_t) u(a_j, a_l)$ .

We assume that agents revise their strategies simultaneously at discrete time intervals,  $t = 1, 2, \dots$ . Following the canonical evolutionary models of best response with mutations (BRM) (Kandori et al., 1993; Young, 1993; Ellison, 2000), we assume that agents are myopic—only react to neighbours' strategies chosen one period before, and boundedly rational—they choose best responses to neighbours choices but they occasionally make mistakes and chose a different strategy at random. Specifically, the probability  $\mathbb{P}_i(a_j; \mathbf{x}_t)$  that agent  $i$  chooses  $a_j$  in the next period given the current configuration  $\mathbf{x}_t$ , is given by

$$\mathbb{P}_i(a_j; \mathbf{x}_t) = \frac{1}{m} \exp(-\beta) + (1 - \exp(-\beta)) BR_i(a_j; \mathbf{x}_t) \quad (2)$$

The first term  $\frac{1}{m} \exp(-\beta)$  captures individual experimentation (mutations) on choices and is independent of the current configuration  $\mathbf{x}_t$ . More specifically, an agent decides to experiment with a probability  $\varepsilon = \exp(-\beta)$ , common to all agents. If an agent experiments, she chooses any option with a uniform probability  $\frac{1}{m}$ .

The second term,  $BR_i(a_j; \mathbf{x}_t)$  captures social influence of other agents on  $i$ , with  $1 - \varepsilon = 1 - \exp(-\beta)$  as a measure of the strength of social influence. Let  $BR(\mathbf{x}_t)$  be the set of choices that are best responses to the configuration  $\mathbf{x}_t$ . That is

$$BR(\mathbf{x}_t) = \{a_j : U_i(a_j, \mathbf{x}_t) \geq U_i(a_l, \mathbf{x}_t), a_j, a_l \in A\}.$$

Then  $BR_i(a_j; \mathbf{x}_t)$  is the probability that  $i$  chooses  $a_j$  through best response given configuration  $\mathbf{x}_t$ . That is, if  $a_j \in BR(\mathbf{x}_t)$  and  $b$  is the cardinality of  $BR(\mathbf{x}_t)$ , then  $BR_i(a_j; \mathbf{x}_t) = \frac{1}{b}$ , and  $BR_i(a_j; \mathbf{x}_t) = 0$  if  $a_j \notin BR(\mathbf{x}_t)$ . In probabilistic terms,  $1 - \varepsilon$  is then interpreted as the probability that an agent follows a best response behaviour. Thus, as  $\beta$  increases to infinity, agents become more rational and choose best responses with higher probability; and as  $\beta$  tends to zero, agents' choices become more and more noisy.

Let  $P_\beta(\mathbf{x}, \mathbf{y}) = \mathbb{P}(\mathbf{x}_{t+1} = \mathbf{y}; \mathbf{x}_t = \mathbf{x})$  be the probability that configuration  $\mathbf{x}$  is followed by configuration  $\mathbf{y}$  after a single period of iteration, where the subscript  $\beta$  represents the parameter of experimentation defined above. Simultaneous revision of choices implies that

$$P_\beta(\mathbf{x}, \mathbf{y}) = \mathbb{P}(\mathbf{x}_{t+1} = \mathbf{y}; \mathbf{x}_t = \mathbf{x}) = \prod_{i=1}^n \mathbb{P}_i(y^i; \mathbf{x}_t = \mathbf{x}) \quad (3)$$

The dynamics in (3) follows a stationary Markov chain on the configuration space  $\mathbf{X}$ . We denote by  $P_\beta$  for the Markov transition matrix with  $P_\beta(\mathbf{x}, \mathbf{y})$ , for all pairs  $\mathbf{x}, \mathbf{y} \in \mathbf{X}$ , as its elements. Our analysis of the diffusion process defined above will then rely on the well-known properties of Markov chains. With slight abuse of meaning, we throughout the paper refer to the triple  $(U, G, P_\beta)$  as a *strict coordination game diffusion process*.



We have assumed a simple structure of choice probabilities in (2) by letting the probability of mistakes to be identical for all agents and independent of the configuration. It is possible to consider more sophisticated choice probabilities like the Logit and Probit models where the probability of mistakes depends on the state. Specifically, the probability of mistakes in the Logit and Probit models depends on payoff losses, whereby mistakes that incur larger payoff losses are rare compared to those with smaller payoff losses. A simpler model of choice probabilities considered in (2) permits a more detailed analysis of the effect of noise and network structure on long-run probabilities. Moreover, most of our results would qualitatively hold if other models of choice probabilities were to be considered. Specifically, we will show below that the long-run behaviour of the diffusion process described above exhibits threshold dynamics with respect to the probability of mistakes; such behaviour would exist in all models of mistakes although the exact threshold values may differ across the models.

### 2.3. Limit states and long-run probabilities

The long-run behaviour of a Markov chain is defined in terms of its *stationary distribution*. The stationary distribution describes the fractional amount of time the process spends in each configuration in the long-run. Or equivalently, the probability with which each configuration is visited in the long-run. Formally, let  $\mathbf{q}_0$  be an  $m^n$ -row vector representing an initial distribution of a Markov chain. For example, if the chain starts from a configuration  $\mathbf{x}$ , then  $\mathbf{q}_0$  is a vector of all zeros except a one in configuration  $\mathbf{x}$ . After  $t$  iterations, the distribution is  $\mathbf{q}_t = \mathbf{q}_0 P_\beta^t$ . The stationary distribution  $\pi_\beta$  is then an  $m^n$ -vector defined as  $\pi_\beta = \lim_{t \rightarrow \infty} \mathbf{q}_0 P_\beta^t$ .

Before examining the long-run behaviour of  $P_\beta$  for  $\beta < \infty$ , it helps to first examine the properties of the special case where  $\beta = \infty$ . Let  $P = P_\infty$ . The diffusion process with transition matrix  $P$  corresponds to the situation where agents do not experiment and only make choices based on best response evaluations. The equilibrium configurations of  $P$  are referred to as *limit* or *absorbing configurations*. Let  $\mathbf{L}$  be a set of all limit configurations of  $P$ , and write  $L$  or  $L_j, L_l, L_k, \dots$  for typical elements of  $\mathbf{L}$ ; we also write  $\mathbf{a}_l$  for a limit configuration where all agents coordinate on choice  $a_l$ . The composition of  $\mathbf{L}$  depends on the payoff and network structures. Generally, it consists of all monomorphic configurations  $\mathbf{a}_l$  for  $l = 1, \dots, m$ . Depending on the network structure,  $\mathbf{L}$  may also contain configurations where choices co-exist, and subsets of configurations within which the process cycles. We provide an example in [Appendix A](#) to illustrate how configurations with co-existing choices may arise.

For a subset of limit configurations  $\Omega \in \mathbf{L}$ , let  $D(\Omega)$  denote its basin of attraction—the set of all configurations from which the Markov chain converge to some  $L$  in  $\Omega$  through best response dynamics (i.e. in the absence of experimentation). Thus, for the process  $P$ , the set  $\mathbf{L}$  can be divided into multiple subsets  $\Omega$  with disjoint basins of attraction. Once the process enters any  $D(\Omega)$ , it gets trapped in it indefinitely, and hence, the long-run behaviour of  $P$  depends on the starting configuration.

The tendency of agents to experiment with different choices, however, ensures that in the long-run, every configuration is reached with a positive probability; in this situation,  $\pi_\beta(\mathbf{x}) > 0$  for all  $\mathbf{x} \in \mathbf{X}$ . We



focus on examining the long-run probabilities of limit configurations since the process spends most time in them even when. For the remainder of the paper, our analysis proceeds in three steps:

- (i) We first establish conditions under which targeting is economically reasonable in relation to the experimentation parameter  $\beta$ .
- (ii) We establish conditions under which a given limit state is globally stable, in the sense that it is played with the highest probability  $\pi_\beta$  in the long-run. We then use this result to examine how an economic agent can exploit the process of contagion to reduce the cost of targeting.
- (iii) Finally, we examine the speed of convergence of the strict coordination game diffusion process  $(U, G, P_\beta)$ . Specifically, we show that if a given choice satisfies a condition of contagion (to be stated below), then the waiting time to reach a state in which the said choice is adopted is independent of the population size. The implication is that even in large networks (populations), such a choice diffuses relatively fast compared to all other choices.

### 3. Experimentation and long-run probabilities: when is it reasonable to target agents?

This section characterizes long-run probabilities, focusing on how they vary with the level of experimentation. The following notations and definitions are used for the analysis that follows.

**Definition 1.** Let  $g \subset \mathbf{L} \times \mathbf{L}$  be any oriented graph defined within the configuration space  $\mathbf{L}$  of limits states. Then for a subset  $W \subset \mathbf{L}$  and its complement  $\bar{W}$ , we denote by  $\Gamma(W)$  a set of all oriented graphs satisfying two conditions: (i) no arrows start from  $W$  and exactly one arrow starts from each configuration outside of  $W$ , (ii) each  $g \in \Gamma(W)$  has no loops.

From Definition 1, if  $W$  is a singleton set, say  $W = \{L\}$ , then  $\Gamma(\{L\})$  is a set of all spanning trees of  $L$ , also known as  $L$ -trees. That is, an  $L$ -tree is a directed spanning tree consisting of the set of all limit configurations as the set of nodes, and all links are directed to  $L$  as the end point. Consider an evolutionary process on a network presented in Appendix A where  $\mathbf{L} = \{\mathbf{a}_1, \mathbf{a}, \mathbf{b}, \mathbf{a}_2\}$ . Figure 1 presents three examples of  $\Gamma(\{\mathbf{a}_2\})$  graphs (i.e.  $\mathbf{a}_2$ -trees). In total there are thirteen such  $\mathbf{a}_2$ -trees.<sup>6</sup>

Let  $W = \{L\}$ ; for each  $g \in \Gamma(\{L\})$ , and each arrow  $L_j \rightarrow L_l$  in graph  $g$ , let  $r_g(L_j, L_l)$  be the minimal number of mutations needed for a transition from  $L_j$  to  $L_l$  to occur. Specifically, given that  $L_l$  is the unique absorbing state that succeeds  $L_j$  in graph  $g$ ,  $r_g(L_j, L_l)$  is the minimum number of mutations needed to exit the basin of attraction  $D(L_j)$  of  $L_j$  to some state in the basin of attraction of  $L_l$ . Such transitions need not occur in a single step and may entail a sequence of successive mutations that in turn lead the process to evolve out of  $D(L_j)$  and eventually through best response to  $L_l$ . To reduce notational

---

<sup>6</sup>That is,  $\{\mathbf{a}_1 \rightarrow \mathbf{a} \rightarrow \mathbf{b} \rightarrow \mathbf{a}_2\}, \{\mathbf{a}_1 \rightarrow \mathbf{b} \rightarrow \mathbf{a} \rightarrow \mathbf{a}_2\}, \{\mathbf{a} \rightarrow \mathbf{a}_1 \rightarrow \mathbf{b} \rightarrow \mathbf{a}_2\}, \{\mathbf{a} \rightarrow \mathbf{b} \rightarrow \mathbf{a}_1 \rightarrow \mathbf{a}_2\}, \{\mathbf{b} \rightarrow \mathbf{a} \rightarrow \mathbf{a}_1 \rightarrow \mathbf{a}_2\}, \{\mathbf{b} \rightarrow \mathbf{a}_1 \rightarrow \mathbf{a} \rightarrow \mathbf{a}_2\}, \{\mathbf{a}_1 \rightarrow \mathbf{a} \rightarrow \mathbf{a}_2, \mathbf{b} \rightarrow \mathbf{a}_2\}, \{\mathbf{a} \rightarrow \mathbf{a}_1 \rightarrow \mathbf{a}_2, \mathbf{b} \rightarrow \mathbf{a}_2\}, \{\mathbf{a}_1 \rightarrow \mathbf{b} \rightarrow \mathbf{a}_2, \mathbf{a} \rightarrow \mathbf{a}_2\}, \{\mathbf{b} \rightarrow \mathbf{a}_1 \rightarrow \mathbf{a}_2, \mathbf{a} \rightarrow \mathbf{a}_2\}, \{\mathbf{a} \rightarrow \mathbf{b} \rightarrow \mathbf{a}_2, \mathbf{a}_1 \rightarrow \mathbf{a}_2\}, \{\mathbf{a} \rightarrow \mathbf{b} \rightarrow \mathbf{a}_2, \mathbf{a}_1 \rightarrow \mathbf{a}_2\}, \{\mathbf{a}_1 \rightarrow \mathbf{a}_2, \mathbf{a} \rightarrow \mathbf{a}_2, \mathbf{b} \rightarrow \mathbf{a}_2\}$

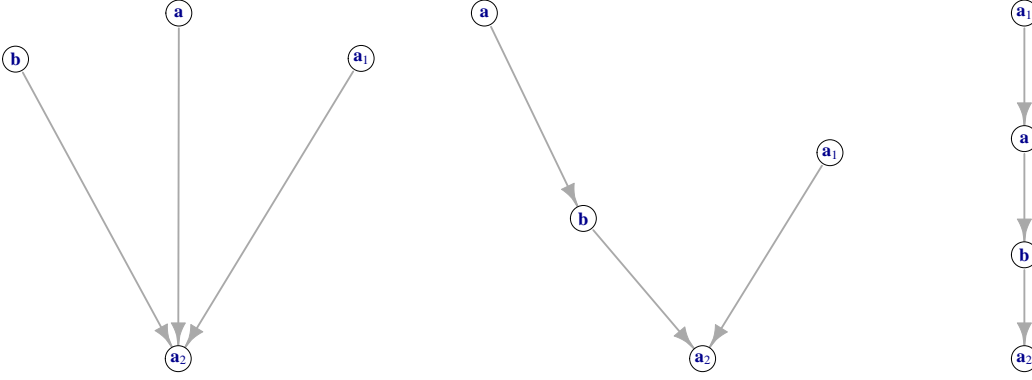


Figure 1: Examples of  $\Gamma(\{a_2\})$  graphs, where  $\mathbf{L} = \{a_1, a, b, a_2\}$ .

clutter by simply writing  $r_g(L_j)$  for  $r_g(L_j, L_l)$ , knowing that there is some  $L_l \neq L_j$  that succeeds  $L_j$  in  $g$ . This way,  $r_g(L_j)$  is simply the number of mutations needed to exit  $D(L_j)$  in graph  $g$ .

For each  $g \in \Gamma(\{L\})$ , we define  $r_g = \sum_{L_j \in \mathbf{L}} r_g(L_j)$ , where  $r_g(L) = 0$ , as the total (minimal) number of mutations among all transitions within  $g$ . We then let  $r_L = \min_{g \in \Gamma(\{L\})} r_g$  denote the total number of mutations of the *minimum cost* (i.e. minimum mutation)  $L$ -tree, and write  $r_L(L_j)$  as the minimal number of mutations required to exit  $D(L_j)$  in the minimum  $L$ -tree.<sup>7</sup> The following lemma expresses the ratio of long-run probabilities of any two limit states  $L$  and  $L'$  as a function of  $\beta$ ,  $r_L$  and  $r_{L'}$ .

**Lemma 1.** *For a strict coordination game diffusion process  $(U, G, P_\beta)$ , let  $L, L' \in \mathbf{L}$  be any pair of limit states. The ratio of long-run probabilities of  $L$  and  $L'$  is bounded by*

$$\frac{\pi_\beta(L)}{\pi_\beta(L')} = \exp [-(r_L - r_{L'}) (\beta_m - \beta'_m) \pm \kappa] \quad (4)$$

where  $\beta_m = \beta - \ln m^{-1}$ ,  $\beta'_m = -n \ln \left[ 1 - e^{-\beta} \left( 1 - \frac{1}{m} \right) \right]$ , and  $\kappa$  is a uniform constant.

*Proof.* See Appendix [Appendix B.1](#)

□

The term in the exponent of (4) is separable into  $r_L(\beta) = (\beta_m - \beta'_m)r_L$  and  $r_{L'}(\beta) = (\beta_m - \beta'_m)r_{L'}$ . For each  $L$ ,  $r_L$ , as before, is the total number of mutations in the minimum cost  $L$ -tree, and  $(\beta_m - \beta'_m)$  captures the likelihood of playing a choice(s) contained in  $L$  through experimentation. We thus refer to  $r_L(\beta)$  as the *cost* of reaching  $L$  through the minimum cost  $L$ -tree, measured in terms of the total number and the likelihood of mutations.

Lemma 1 then states that the ratio of long-run probabilities of any two limit states is an exponential function of the difference of the total costs of their minimum cost spanning trees. We denote this cost difference by  $\Phi_\beta(L, L')$ .<sup>8</sup> The following Proposition establishes the relationship between long-run probabilities and  $\beta$ .

<sup>7</sup>We refer to  $g^* \in \operatorname{argmax}_{g \in \Gamma(\{L\})} r_g$  as the *minimum cost  $L$ -tree* because, as it will become clear below, the cost of reaching  $L$  from all other limit states through  $g$  is directly proportional to  $r_g$ .

<sup>8</sup> That is,  $\Phi_\beta(L, L') = -(r_L(\beta) - r_{L'}(\beta)) = -(\beta_m - \beta'_m)(r_L - r_{L'})$ .

**Proposition 1.** For a strict coordination game diffusion process  $(U, G, P_\beta)$ , there exists a  $\beta^* > 0$  which is a solution to  $\beta_m = \beta'_m$ , such that for all  $\beta > \beta^*$ , if  $r_L > r_{L'}$ , then  $\pi_\beta(L) < \pi_\beta(L')$ .

*Proof.* See Appendix [Appendix B.2](#)

□

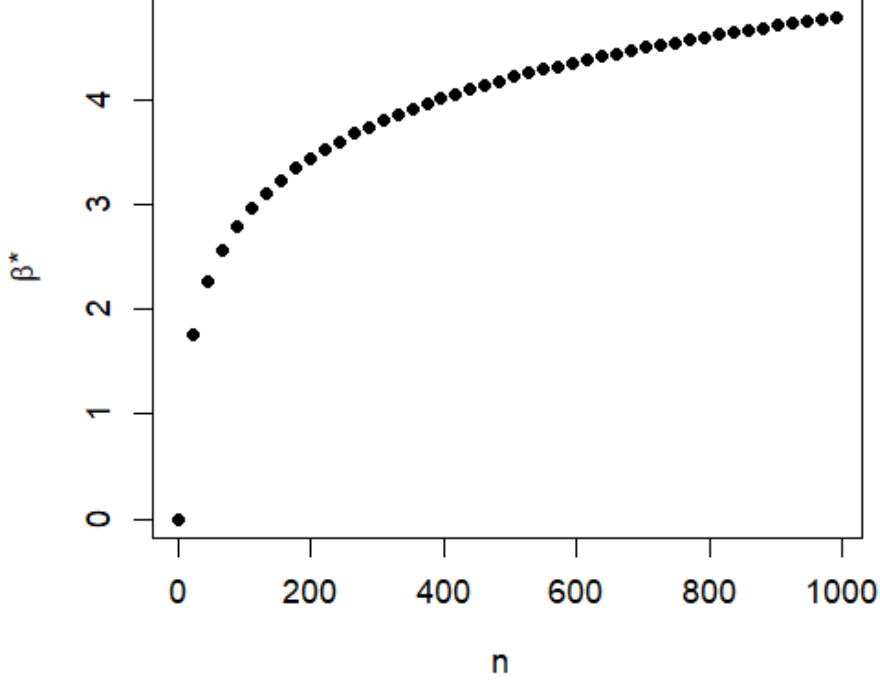


Figure 2: A plot of  $\beta^*$  against  $n$ , for  $m = 4$ .

The intuition behind Proposition 1 is that if the number of mutations involved in an  $L$ -tree is larger for an  $L'$ -tree so that  $r_L > r_{L'}$ , then reaching  $L$  from any other limit state is relatively less costly when agents' behaviour is dominated by experimentation (i.e. when  $\beta < \beta^*$ ). But when individual behaviour is dominated by best response (i.e. when  $\beta > \beta^*$ ), then reaching  $L'$  is relatively less costly since fewer mutations are involved. The threshold level of experimentation,  $\beta^*$ , increases logarithmically with the population size  $n$  (see an example of a plot of  $\beta^*$  with  $n$  for  $m = 4$  in Figure 2).

There are two main implications of Proposition 1. First, to compare the long-run probabilities of any two limit states  $L$  and  $L'$ , it is necessary to have knowledge of the level of experimentation, specifically, whether  $\beta \leq \beta^*$ , and knowledge of  $r_L$  and  $r_{L'}$  (the total minimum number of mutations necessary to reach  $L$  and  $L'$  from every other limit state), specifically, whether  $r_L \leq r_{L'}$ . Stochastic evolutionary models typically focus on examining the behaviour of long-run probabilities when the level of experimentation is very low and  $\beta \rightarrow \infty$  (Foster and Young, 1990; Kandori et al., 1993; Young, 1993; Ellison, 2000). The limit configurations, and hence choices, that retain a positive long-run probability when  $\beta \rightarrow \infty$  are said to be *stochastically stable*. In our case, this corresponds to the regime where  $\beta > \beta^*$ , and as Proposition 1 shows, the identification of stochastically stable states then reduces to searching for limit

states with the least-minimum cost  $L$ -tree. The results from the literature of stochastic evolutionary game theory regarding the type of choices that are stochastically stable thus apply to evolutionary processes with positive levels of experimentation provided  $\beta > \beta^*$ . But as we discuss in the next section, there are few papers that study stochastic evolutionary game dynamics on networks. Moreover, the few papers that do focus on specific network structures, and since stochastic stability is not robust to the interactions structure (Alós-Ferrer and Weidenholzer, 2007), such results are not generalizable to all network topologies.

The second implication of Proposition 1 is related to diffusion and targeting. The purpose of targeting is to increase the long-run probability of a given choice being adopted by the population. Since firms and governments do not have control of the extent to which agents experiment (i.e. parameter  $\beta$ ), targeting thus only makes economic sense if the intention is to reduce the number of mutations  $r_L$ . That is, if a firm/planner's objective is to target specific agents so as to steer the population to configuration  $L$ , then the resources spent on targeting will be proportional to the number of agents to be targeted, which in this case is equivalent to  $r_L$ —the minimum number of mutations needed to reach  $L$  from every other limit configuration. Each  $r_L$  can thus be viewed as the cost of targeting.

Note that  $r_L$  for each  $L$  defines the cost of targeting for long-run dynamics. It is possible to define costs of targeting for short- to medium-run dynamics. The short- to medium-run dynamics focuses on the evolution from the status quo to the desired configuration. To see the difference, consider Example Appendix A again where  $\mathbf{L} = \{\mathbf{a}_1, \mathbf{a}, \mathbf{b}, \mathbf{a}_2\}$ . The minimum cost  $\mathbf{a}_2$ -tree is  $\{\mathbf{a}_1 \rightarrow \mathbf{a} \rightarrow \mathbf{a}_2, \mathbf{b} \rightarrow \mathbf{a}_2\}$  with  $r_{\mathbf{a}_2} = 3$ .<sup>9</sup> Let the status quo be  $\mathbf{a}_1$  and assume that a firm/planner (a new entrant) is aiming to steer the population to configuration  $\mathbf{a}_2$ . For the short- to medium-run dynamics, it is sufficient to focus on the evolutionary path  $\mathbf{a}_1 \rightarrow \mathbf{a} \rightarrow \mathbf{a}_2$ . The targeting cost would thus be two agents (i.e. two mutations). Analysing short- to medium-run dynamics is equivalent to examining an evolutionary process where agents do not make mistakes. In the presence of experimentation however, focusing on the path  $\mathbf{a}_1 \rightarrow \mathbf{a} \rightarrow \mathbf{a}_2$  is not sufficient because either of agents 1, 2, 3 could experiment leading to a transition from  $\mathbf{a}_1$  to  $\mathbf{b}$ . It is therefore necessary to consider all possible limit configurations a population can end up in; and hence, the cost of targeting should be  $r_L$ .

#### 4. Diffusion through contagion

This section aims to show that one way a firm/planner can reduce the cost of targeting is by exploiting the process of contagion. Contagion is defined on a pairwise level but can be generalized to multiple choices. For a pair of choices  $a_j, a_l \in A$ , choice  $a_l$  is pairwise contagious relative to  $a_j$  ( $a_l \text{PRC} a_j$ ) if starting from a monomorphic configuration  $\mathbf{a}_j$ , choice  $a_l$  can spread through best response once it has been adopted by a *very small* set of agents called *initial adopters*. The set of initial adopters is equivalent

---

<sup>9</sup>That is, transition  $\mathbf{a}_1 \rightarrow \mathbf{a}$  requires one mutation (agent 4 to switch from  $a_1$  to  $a_2$ ), transition  $\mathbf{a} \rightarrow \mathbf{a}_2$  one mutation, and transition  $\mathbf{b} \rightarrow \mathbf{a}_2$  also one mutation.

to the number of mutations needed to exit the basin of attraction of a given limit configuration. By “very small” we mean it is much smaller than the population size. Thus, by definition, if  $a_l \text{PRC} a_j$ , then the number of initial adopters (mutations) needed to exit the basin of attraction of limit configuration in which  $a_j$  is adopted, to configuration where  $a_l$  is adopted is very small.

If there are only two choices  $a_1$  and  $a_2$ , where  $a_2$  is the new choice being introduced, then the above definition suggests that pairwise contagion of  $a_2$  relative to  $a_1$  is sufficient to reduce the number of initial adopters (i.e. mutations) needed to trigger a transition from  $\mathbf{a}_1$  to  $\mathbf{a}_2$ . But as we show below, pairwise contagion not only reduces the number of initial adopters, it also reduces the number of limit configurations (i.e. configurations where choices co-exist), and hence, reducing the cost of targeting for long-run dynamics even further. For Example [Appendix A](#), if  $a_2 \text{PRC} a_1$ , then only the monomorphic configurations  $\mathbf{a}_1$  and  $\mathbf{a}_2$  exist. To generalize the notion of pairwise contagion to situations with more than two choices, we define a *globally pairwise contagious* choice as that which is pairwise contagious relative to every other choice.

Formally, for any pair  $a_j, a_l \in A$ , let  $N(\mathbf{a}_j \rightarrow \mathbf{a}_l)$  be the set of initial adopters required to trigger contagion from  $a_j$  to  $a_l$ . Let  $n(\mathbf{a}_j \rightarrow \mathbf{a}_l)$  be the respective cardinality. Let also  $n(a_j; \mathbf{x}_t)$  be the number of agents in the population choosing option  $a_j$  when the configuration at  $t$  is  $\mathbf{x}_t$ .

**Definition 2.** For a pair of choices  $a_j, a_l \in A$ , let  $\beta > \beta^*$  and for  $n$  sufficiently large, let  $n(\mathbf{a}_j \rightarrow \mathbf{a}_l)$  be very small and independent of  $n$ , i.e.  $n(\mathbf{a}_j \rightarrow \mathbf{a}_l) \ll \frac{1}{2}n$ . A choice  $a_l$  is said to be pairwise contagious relative to  $a_j$  ( $a_l \text{PRC} a_j$ ) if  $n(a_j; \mathbf{x}_{t+1}) > n(a_j; \mathbf{x}_t)$  for all  $t \geq t'$  whenever  $n(a_l; \mathbf{x}_{t'}) = n(\mathbf{a}_j \rightarrow \mathbf{a}_l)$ .

**Definition 3.** Given a choice set  $A$ , choice  $a_l$  is said to be globally pairwise contagious if it is pair-wise contagious relative to all  $a_j \in A$ , where  $a_j \neq a_l$ .

There are two conditions in the definition of contagion above. First, contagion is defined for situations where individual behaviour is driven solely by best response evaluations; as such, we require  $\beta > \beta^*$ . Second, the set of initial adopters must be small, and its size independent of  $n$ . If the set of initial adopters is close to  $\frac{1}{2}n$  or greater than  $\frac{1}{2}n$  or scales with  $n$ , then targeting in large populations will be costly.

In the following analysis, we aim to show that a globally pairwise contagious choice, which is unique if it exists, has the least cost of targeting. Following the discussion above, this is equivalent to showing that a monomorphic limit state containing a globally pairwise contagious choice has the least cost  $L$ -tree, and hence, the largest long-run probability.

**Proposition 2.** For a strict coordination game diffusion process  $(U, G, P_\beta)$ , let  $\beta > \beta^*$ . If  $a_l$  is globally pairwise contagious then  $\pi_\beta(\mathbf{a}_l) > \pi_\beta(L_j)$  for all  $L_j \neq \mathbf{a}_l$ .

*Proof.* See [Appendix B.5](#)

□

The proof of Proposition 2 follows a series of lemmas some of which we highlight below. We focus on the lemmas establishing the conditions for pairwise contagion. The properties of the network and payoffs that jointly determine the necessary and sufficient conditions for contagion are respectively the *contagion threshold* and *relative payoff-losses (switching losses)*. We define and discuss each in detail below, starting with the former.

Let  $q \in (0, 1)$  be a critical number such that strategy  $a_2$  is a best response whenever at least proportion  $q$  of a player's neighbours choose  $a_2$ . The *contagion threshold*  $\eta(G)$  of a network  $G$  is a positive real number such that contagion occurs if and only if  $q \leq \eta(G)$ . Formally, let  $S \subset N$  be a subgroup of agents belonging to the same arbitrarily chosen neighbourhood, and  $\bar{S} = N \setminus S$  its complement. Write  $S(a_l)$  to imply that all agents in  $S$  choose  $a_l$  and write  $\bar{S}^{[p, S]}(a_j)$  for agents in  $\bar{S}$  choosing  $a_j$  for whom at least proportion  $p$  of their interactions are with agents in  $S$ .

Now, consider a sequence  $\{S_1(a_l), S_2(a_l), \dots, S_\tau(a_l), \dots, S_T(a_l)\}$  whereby  $S_1(a_l) = N(\mathbf{a}_j \rightarrow \mathbf{a}_l)$  is the set of initial adopters of  $a_l$ ,  $S_T(a_l) = N$  (i.e. corresponding to a configuration where all agents choose  $a_l$ ), and  $S_\tau(a_l) \subset S_{\tau+1}(a_l)$  for all  $\tau = 1, \dots, T$ . There exists a complementary sequence  $\{\bar{S}_1(a_l), \bar{S}_2(a_l), \dots, \bar{S}_\tau(a_l), \dots, \bar{S}_T(a_l)\}$  whereby  $\bar{S}_T(a_l) = \emptyset$  (i.e. corresponding to a configuration where all agents choose  $a_l$  and none chooses  $a_j$ ), and  $\bar{S}_\tau(a_j) \supset \bar{S}_{\tau+1}(a_j)$  for all  $\tau = 1, \dots, T$ .<sup>10</sup> Given these two complementary sequences, there exists a *maximum*  $p$ , called the *contagion threshold* of the underlying network, so that for a corresponding sequence  $\{\bar{S}_1^{[p, S]}(a_j), \bar{S}_2^{[p, S]}(a_j), \dots, \bar{S}_\tau^{[p, S]}(a_j), \dots, \bar{S}_T^{[p, S]}(a_j)\}$ ,  $a_l$  is a best response for each  $i \in \bar{S}_\tau^{[p, S]}(a_j)$ .

Every network has a contagion threshold of less or equal to  $\frac{1}{2}$  (Morris, 2000, Proposition 3). Morris (2000) provides lower and upper bounds of contagion threshold as functions of various connectivity measures.<sup>11</sup> Two of such measures are maximum neighbourhood size  $\bar{n} = \max_{i \in N} n_i$ , and network cohesion defined as follows.

Let  $Z \subset N$  and  $\bar{Z} = N \setminus Z$ . For each  $i \in Z$ , define  $n_i(Z)$  as the fraction of  $i$ 's neighbours that belong to subgroup  $Z$ . That is,

$$n_i(Z) = \frac{\#(N_i \cap Z)}{n_i}$$

where  $\#Z$  is the cardinality of set  $Z$ . Now, define  $c(Z, G) = \min_{i \in Z} n_i(Z)$  — a measure of cohesiveness of subgroup  $Z$ . The larger  $c(Z, G)$ , the more cohesive is subgroup  $Z$ . By choosing a sufficiently large subset  $Z$  with maximum cohesion  $c(Z, G)$ , the parameter of network cohesion  $c(G)$  is then defined as,

$$c(G) = 1 - \max_{\#Z \leq \frac{1}{2}n} c(Z, G) \quad (5)$$

A network  $G$  is then said to be  $c$ -cohesive if  $c(G) = c$ . The definition of  $c(G)$  in (5) captures the idea that if there exists a subgroup of agents that is highly cohesive, then the whole network is less cohesive.

**Lemma 2.** *Let  $\eta(G)$  be the contagion threshold of network  $G$ :*

<sup>10</sup>Note, it follows by this definition that  $S_{\tau+1}(a_l) \setminus S_\tau(a_l) = \bar{S}_\tau(a_j) \setminus \bar{S}_{\tau+1}(a_j)$ .

<sup>11</sup>See Appendix also Lelarge (2012) for a related characterization for random networks

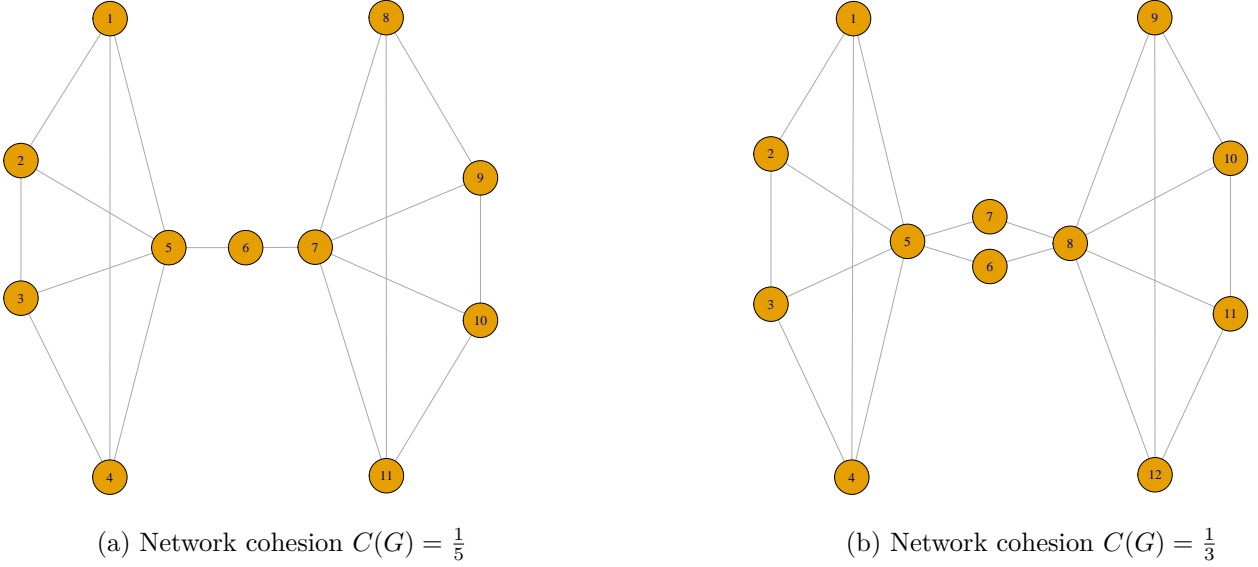


Figure 3: Networks with low cohesion

- (i) If  $\bar{n} = \max_{i \in N} n_i$  is the maximum neighbourhood size in the network, then  $\eta(G) \geq 1/\bar{n}$  (Morris, 2000, Corollary 3).
- (ii) If  $G$  is  $c$ -cohesive, then  $\eta(G) \leq c$  (Morris, 2000, Corollary 1).

Consider the networks in Figure 3. In network 3a, the subgroups  $Z_1 = \{1, 2, 3, 4, 5\}$  and  $Z_2 = \{7, 8, 9, 10, 11\}$  are both  $\frac{4}{5}$ -cohesive. Thus, the whole network is  $1 - c(Z_1, G) = 1 - c(Z_2, G) = \frac{1}{5}$ . In network 3b,  $Z_1 = \{1, 2, 3, 4, 5\}$  and  $Z_2 = \{7, 8, 9, 10, 11\}$  are both  $\frac{2}{3}$ -cohesive, and hence, the whole network is  $\frac{1}{3}$ -cohesive.

Overall, Lemma 2 shows that highly connected and/or cohesive networks have a smaller contagion threshold. Consider for example the networks in Figures 3a and 4a. Both networks have the same total number of edges, and hence, have an identical average connectivity (degree). But Figure 4a is derived from Figure 3a by redistributing the edges so that it is more uniform (less cohesive/skewed) than Figure 3a. Consequently, we expect Figure 4a, as dictated by Lemma 2, to have larger network cohesion (i.e.  $\frac{1}{3}$ ) than Figure 3a (i.e.  $\frac{1}{5}$ ).

Network cohesion sufficiently describes the upper bound for contagion threshold. In regular networks—networks where each agent has the same number of neighbours—the contagion threshold is exactly equal to network cohesion. In the grid network of Figure 4b for example, the corresponding network cohesion is  $C(G) = \frac{1}{4}$ . That is, every subgroup assuming the form in subgroup  $S$  depicted in Figure 4b is  $\frac{3}{4}$ -cohesive, making the entire network  $\frac{1}{4}$ -cohesive.

The second set of measures that determine the feasibility of contagion are *relative payoff-losses* (switching losses)  $\eta_{jl}$  and  $\phi_{jl}$ . To define  $\eta_{jl}$  and  $\phi_{jl}$ , let  $\sigma_{jl}^q = (1 - q, q, 0, \dots, 0)$  be a distribution that assigns a weight of  $1 - q$  to choice  $a_j$ ,  $q$  to choice  $a_l$  and zero to all other choices. In analogy to the local interaction coordination game,  $\sigma_{jl}^q$  is a distribution in which proportion  $q$  of an agent's neighbours choose  $a_l$  and  $1 - q$  choose  $a_j$ . The expected payoff  $U(a_k | \sigma_{jl}^q)$  to choice  $a_k$  against distribution  $\sigma_{jl}^q$  is



then

$$U(a_k | \sigma_{jl}^q) = (1 - q)u(a_k, a_j) + qu(a_k, a_l) \quad (6)$$

**Definition 4.** For a pair of choices  $a_j, a_l \in A$ , the relative payoff-losses are defined as follows:

(i)  $\eta_{jl}$  is the maximum  $q$  above which  $a_l$  is a best response to  $\sigma_{jl}^q$ ; that is, the  $q$  for which

$$U(a_l | \sigma_{jl}^q) > U(a_k | \sigma_{jl}^q) \quad \text{for all } a_k \neq a_l \quad (7)$$

(ii)  $\phi_{jl}$  is the value of  $q$  for which

$$U(a_l | \sigma_{jl}^q) > U(a_j | \sigma_{jl}^q) \quad (8)$$

The payoff-loss parameter  $\eta_{jl}$  is thus equivalent to the proportion of neighbours that must choose  $a_l$  while the complementary proportion  $1 - \eta_{jl}$  chooses  $a_j$ , for a player to be indifferent between  $a_j$  and  $a_l$ . If more than proportion  $\eta_{jl}$  choose  $a_l$ , then a player also switches from  $a_j$  to  $a_l$ . The same interpretation applies to  $\phi_{jl}$  but with the payoff matrix restricted to only two strategies  $a_j$  and  $a_l$ . The following lemma provides expressions for  $\phi_{jl}$  and  $\eta_{jl}$ , and establishes the relationship between them.

**Lemma 3.** For any pair of strategies  $a_j, a_l \in A$  in coordination game  $U$ , the respective expressions for  $\phi_{jl}$  and  $\eta_{jl}$  are

$$\phi_{jl} = \frac{u(a_j, a_j) - u(a_l, a_j)}{u(a_j, a_j) - u(a_l, a_j) + u(a_l, a_l) - u(a_j, a_l)},$$

$$\eta_{jl} = \max_{a_k \neq a_l} \left\{ \frac{u(a_k, a_j) - u(a_l, a_j)}{u(a_k, a_j) - u(a_l, a_j) + u(a_l, a_l) - u(a_k, a_l)} \right\}.$$

Similarly, for any pair of strategies  $a_j, a_l \in A$ :

(i)  $\phi_{jl} = 1 - \phi_{lj}$

(ii)  $\eta_{jl} \geq \phi_{jl}$

*Proof.* See Appendix [Appendix B.3](#)

□

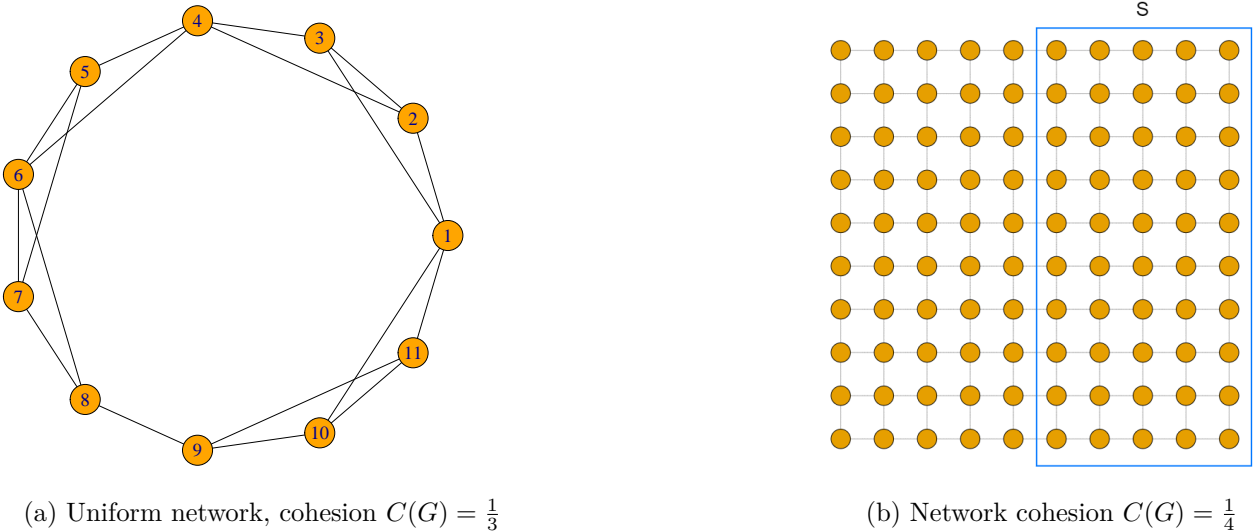


Figure 4: Uniform networks

We refer to parameters  $\eta_{jl}$  and  $\phi_{jl}$  as payoff-losses because they capture losses associated with deviating from coordination; i.e. switching away from the choice that a player's neighbours are coordinating on. First consider parameter  $\phi_{jl}$ . The quantities  $C(a_j, a_l) = u(a_j, a_j) - u(a_l, a_j)$  and  $C(a_l, a_j) = u(a_l, a_l) - u(a_j, a_l)$  are the respective losses associated with switching from  $a_j$  to  $a_l$  and from  $a_l$  to  $a_j$  given that all the player's neighbours respectively choose  $a_j$  and  $a_l$ .<sup>12</sup> The payoff parameter  $\phi_{jl}$  can thus be expressed as

$$\phi_{jl} = \frac{C(a_j, a_l)}{C(a_j, a_l) + C(a_l, a_j)} \quad (9)$$

If  $\phi_{jl} < \frac{1}{2}$ , then the losses associated with switching from  $a_j$  to  $a_l$ , while neighbour coordinate on  $a_j$ , is less than for switching from  $a_l$  to  $a_j$ , and vice versa for  $\phi_{jl} > \frac{1}{2}$ . Most importantly, the smaller  $\phi_{jl}$ , the smaller the loss of switching from  $a_j$  to  $a_l$  and the larger the loss of switching from  $a_l$  to  $a_j$ . The payoff parameter  $\eta_{jl}$  has an identical interpretation but considers all possible deviations away from  $a_j$  and  $a_l$ . When there are only two choices, deviating away from  $a_j$  leads an agent to  $a_l$ , and vice versa. But if there are more than two choices, an agent deviating away from  $a_j$  may choose any other choice other than  $a_l$ ; similarly, any agent deviating away from  $a_l$  may choose another choice different from  $a_j$ . Parameter  $\eta_{jl}$  then compares the switching loss incurred by an agent deviating away from  $a_j$  to the loss of deviating away from  $a_l$ . Thus, if  $a_k$  is the choice that maximizes the expression for  $\eta_{jl}$ , then  $\eta_{jl}$  can be re-written as<sup>13</sup>

$$\eta_{jl} = \frac{C(a_j, a_l) - C(a_j, a_k)}{C(a_j, a_l) - C(a_j, a_k) + C(a_l, a_k)} \quad (10)$$

where the numerator on the right hand side of (10) captures the losses associated with deviating from  $a_j$  to  $a_l$  while discounting for the loss of deviating to  $a_k$  from  $a_j$ . The quantity  $C(a_l, a_k)$  is the loss of deviating from  $a_l$  to  $a_k$ .

**Lemma 4.** *For an  $m \times m$  coordination game  $U$  on an arbitrary network  $G$  with contagion threshold  $\eta(G)$ , pairwise contagion of choice  $a_l$  relative to  $a_j$  occurs if and only if  $\eta_{jl} = \phi_{jl} < \eta(G) \leq \frac{1}{2}$ .*

*Proof.* See Appendix [Appendix B.4](#)

□

Lemma 4 establishes the necessary and sufficient conditions for pair-wise contagion between any pair of choices in terms of network and payoff parameters. The necessary condition is for the  $\eta_{jl}$  to be less or equal to the contagion threshold, and the sufficient condition is for  $\eta_{jl} = \phi_{jl}$ . Lemma 4 has direct

<sup>12</sup>This is because if all of an agent's neighbours choose  $a_j$ , then the payoff difference (loss) for an agent switching from  $a_j$  to  $a_l$  is  $U(a_l | \sigma_{jl}^0) - U(a_j | \sigma_{jl}^0)$ ; where, by definition of  $\sigma_{jl}^q$  above,  $\sigma_{jl}^0$  is the distribution over  $A$  where all (i.e. a unite proportion of) the neighbours choose  $a_j$ . Substituting for the expressions of expected payoffs yields  $U(a_l | \sigma_{jl}^0) - U(a_j | \sigma_{jl}^0) = u(a_j, a_j) - u(a_l, a_j)$ .

<sup>13</sup>That is, if  $a_k$  maximizes the expression for  $\eta_{jl}$ , then  $\eta_{jl} = \frac{u(a_k, a_j) - u(a_l, a_j)}{u(a_k, a_j) - u(a_l, a_j) + u(a_l, a_l) - u(a_k, a_l)}$ . The numerator can be re-written as  $u(a_k, a_j) - u(a_l, a_j) = u(a_k, a_j) - u(a_j, a_j) + u(a_j, a_j) - u(a_l, a_j) = C(a_j, a_l) - C(a_j, a_k)$ . Substituting this expression and  $u(a_l, a_l) - u(a_k, a_l) = C(a_l, a_k)$  into  $\eta_{jl}$  yields the desired result.

implications for co-existence of choices in equilibrium, and as highlighted above, contagion decreases the number of limit configurations. The following claim formalize this argument.

**Claim 1.** *If  $\eta_{jl} > \eta(G)$ , then there exists at least one limit configuration in which choices  $a_j$  and  $a_l$  co-exist. Otherwise, if  $\eta_{jl} = \phi_{jl} < \eta(G)$ , then the only limit configurations containing  $a_j$  and  $a_l$  are the respective monomorphic configurations  $\mathbf{a}_j$  and  $\mathbf{a}_l$ .*

Claim 1 follows because if  $a_l \text{PRC} a_j$ , then starting from limit configuration  $\mathbf{a}_j$ , if a group of agents switching to  $a_l$  is less than  $n(\mathbf{a}_j \rightarrow \mathbf{a}_l)$  (i.e. the number of initial adopters), the population reverts back to  $\mathbf{a}_j$  through best response. If at least  $n(\mathbf{a}_j \rightarrow \mathbf{a}_l)$  agents switch to  $a_l$ , the population converges to configuration  $\mathbf{a}_l$  through best response; and hence, there are no limit configurations where  $a_j$  and  $a_l$  co-exist.

Consider the networks in Figure 3. For the network in Figure 3a, if  $\eta_{jl} > \frac{1}{5}$ , say  $\eta_{jl} = \frac{2}{5}$ , then there exists a limit configuration in which: (i) agents  $\{1, 2, 3, 4, 5, 6\}$  all choose  $a_l$  but  $\{7, 8, 9, 10, 11\}$  choose  $a_j$ .<sup>14</sup> (ii) agents  $\{6, 7, 8, 9, 10, 11\}$  all choose  $a_l$  but  $\{1, 2, 3, 4, 5\}$  choose  $a_j$ . Similarly, if  $\eta_{jl} > \frac{1}{3}$ , then for the network in Figure 3b, there exists a limit configuration in which  $\{1, 2, 3, 4, 5, 6, 7\}$  all choose  $a_l$  but  $\{8, 9, 10, 11, 12\}$  choose  $a_j$ ; and a limit configuration in which  $\{6, 7, 8, 9, 10, 11, 12\}$  all choose  $a_l$  but  $\{1, 2, 3, 4, 5\}$  choose  $a_j$ .

If the network parameters associated with the contagion threshold are known, a firm/planner can manipulate the relative payoffs-losses to increase the long-run probability of her product/behaviour being adopted. In doing so, a firm/planner not only makes her product/behaviour globally stable, but also lowers the cost of targeting and advertising. The question that may concern a firm/planner in seeking to make the product globally contagious is the trade-off between the cost of making a choice globally contagious—*contagion cost* hereafter—versus the cost of targeting.

The contagion cost of choice  $a_l$  is the total cost that a firm/planner has to incur in order to make the quantities  $\phi_{jl}$  and  $\eta_{jl}$ , for all  $a_j \neq a_l$ , sufficiently small so as to satisfy the conditions for pairwise contagion in Lemma 4. From (9),  $\phi_{jl}$  can be decreased by increasing  $u(a_l, a_j)$  and  $u(a_l, a_l)$ . Note that a firm/planner aiming to diffuse choice  $a_l$  has control of  $u(a_l, a_l)$  and  $u(a_l, a_j)$  but not  $u(a_j, a_j)$  and  $u(a_j, a_l)$ .<sup>15</sup> To increase  $u(a_l, a_j)$  and  $u(a_l, a_l)$ , a firm/planner can increase innovativeness and functionality of a product, reduce the price, make it user friendly, or enforce penalties (in cases where  $a_l$  is a behaviour). We refer to all costs incurred in an attempt to increase  $u(a_l, a_j)$  and  $u(a_l, a_l)$  collectively

<sup>14</sup>This is because agent 7 requires at least  $\lceil \frac{2}{5} * 7 \rceil = 3$  of his neighbours to switch to  $a_l$  for him to do likewise. So, even if all agents  $\{1, 2, 3, 4, 5, 6\}$  choose  $a_l$ , agent seven, and hence agents  $\{8, 9, 10, 11\}$ , will not switch to  $a_l$  since only one out of five of his neighbour chose  $a_l$ .

<sup>15</sup>Here,  $u(a_l, a_j)$  can be viewed as private benefits (excluding the social benefits arising from coordination) of choosing  $a_l$ , and  $u(a_l, a_l)$  includes both private and social benefits of choosing  $a_l$ . Consider an example of social networking product like Facebook. Social benefits may include having a network of friends to interact with, while the private benefits may include access to other services (e.g) that do not require having a network of friends.

as contagion costs, which are different from the costs of targeting. The following corollary directly follows from Lemma 4.<sup>16</sup>

**Corollary 1.** *The contagion cost is inversely proportional to the contagion threshold.*

To quantify contagion costs, note that if  $\phi_{jl}$  is the relative loss associated with switching from  $a_j$  to  $a_l$ , then the loss associated with switching from  $a_l$  to  $a_j$  is  $\frac{1-\phi_{jl}}{\phi_{jl}}$  times more than  $\phi_{jl}$ . Note also that the larger  $\frac{1-\phi_{jl}}{\phi_{jl}}$  the more benefits (utility)  $a_l$  offers compared to  $a_j$ ; and hence, it can be interpreted as “by how much  $a_l$  is more beneficial (better) than  $a_j$ ”. Consider again the networks in Figure 3. For the network in Figure 3a, pairwise contagion of  $a_l$  relative to  $a_j$  can occur only if  $a_l$  is at least 4 times more beneficial than  $a_j$ ; and for the network in Figure 3b, only if  $a_l$  is at least 2 times more beneficial than  $a_j$ .

Corollary 1 and Lemma 2 then imply that the cost of making a choice globally contagious is higher for highly connected and/or less cohesive networks. And conversely, the cost of making a choice contagious in less connected and/or uniform networks is lower. In Section 1, we discussed the empirical evidence of diffusion in I.T markets that fits the conclusions from Corollary 1 and Lemma 2. We argued that, one potential explanation for the discrepancies between the dynamics in messaging and chat apps (e.g. AOL Instant Messenger, Google talk, VOIP, skype, Kik., whatsapp, snapchat, HipChat and slack) market versus social networking apps (e.g. Friendster, Myspace and Facebook) markets (i.e. there is a high turnover and entry rate in the messaging and chat app market than in the social networking market), results from difference in interaction structures. Specifically, messaging and chat apps consist of sparsely connected networks where individual interact with (speak, chat or message to) a few others (family and close friends). Social network apps on the other hand consist of large number of connections where some individuals may have as many as tens of thousands of friends. Thus, to invade a network of messaging and chat apps, a new product need not be much better (more beneficial) than those existing in the market. Invading a Facebook-type network however requires a new product to be far better, or simply to offer a different service. This should lead to a high turnover and entry rates in messaging and chat app markets than for social networking app markets.

In concluding this section, we highlight the following three implications of our results. First, making a choice globally contagious reduces the number of initial adopters and limit configurations, and hence, drastically reducing the cost of targeting. If the underlying game does not contain a globally contagious choice, then the outcome with the highest probability is not robust to the network structure. The lack of robustness to network structure in turn implies that if a globally contagious choice does not exist, then the identification of choices with the largest long-run probability requires going through the cumbersome process of constructing all  $\mathbf{x}$ -trees for all possible states. Our results in Proposition 2 alleviate this

---

<sup>16</sup>Corollary 1 follows because if the contagion threshold is very small, then  $\phi_{jl}$  and hence  $\eta_{jl}$  must also be very small for contagion to be feasible.

problem because it is sufficient to examine the payoff structure to identify which choice is globally stable, provided it exists.

Proposition 2 by itself is a significant contribution to the literature of evolutionary game theory. As stated, the results hold provided  $\beta > \beta^*$ , which means that globally contagious choices are stochastically stable. Some of the few papers that examine stochastic evolutionary game dynamics in networks include (Ellison, 2000), who finds that  $\frac{1}{2}$ -dominant choice are stochastically stable in a ring networks, (Berninghaus and Schwalbe, 1996; Lee et al., 2003) who find that risk-dominant choices are stochastically stable in a 2-dimension grid network and Alós-Ferrer and Weidenholzer (2007) who find that globally pairwise risk-dominant choices are stochastically stable in a ring network under some mild conditions. As the description suggests, these papers focus on either  $2 \times 2$  coordination games or specific network structures. And since stochastic stability is not robust to the network structure, their results are not generalizable. For the network in Figure 3a for example, a risk-dominant choice need not be the only stochastically stable choice. Specifically, if  $\frac{1}{5} < \phi_{12} \leq \frac{1}{2}$ , say  $\phi_{12} = \frac{3}{7}$ , so that  $a_2$  is risk-dominant but not pairwise contagious relative to  $a_1$ , then both choices can be stochastically stable.<sup>17</sup> Proposition 2 provides a robust prediction result for  $m \times m$  coordination games and for general network structures.

Second, although our analysis above, especially the results of Proposition 2, focuses on situations where a choice can spread to the entire network through contagion, contagion can also be defined for subregions of the network. For example, the sub-network in Figure 3a containing agents  $S_1 = \{1, 2, 3, 4, 5, 6\}$  has contagion threshold of  $\frac{1}{3}$ , which is less than the contagion threshold for the entire network. This implies that it is less costly to diffuse a choice through contagion to subgroup  $S_1$  than then the entire network (see Corollary 1 above). For example, in the case where there is only one product  $a_1$  in the market and a new entrant seeks to diffuse a new product  $a_2$ , then the latter needs to be only two times better than  $a_1$  for it to diffuse through contagion in subgroup  $S_1$ , but has to be four times better to diffuse to the entire network.

Third, unlike relative payoff losses, the level of experimentation that agents follow is not under a firm/planner's control. However, if  $\beta > \beta^*$ , our results in Proposition 2 and Corollary 1 hold. And even more so when the level of experimentation is very small (i.e.  $\beta$  is very large). One criticism stochastic evolutionary models receive is that in situations where  $\beta$  is very large, the convergence rate may be very slow. In particular, the expected waiting time to reach the long-run stable configuration may be an increasing function of  $n$ . The implication is that in large networks, the diffusion process takes time

---

<sup>17</sup>To see why, first notice that if  $\phi_{12} = \frac{3}{7}$ , then in addition to monomorphic limit states  $\mathbf{a}_1$  and  $\mathbf{a}_2$ , there are two additional limit states: one, denote by  $\mathbf{a}$ , where agents  $\{1, 2, 3, 4, 5, 6\}$  all choose  $a_2$  but  $\{7, 8, 9, 10, 11\}$  choose  $a_1$ ; and the other, denote by  $\mathbf{b}$ , where agents  $\{6, 7, 8, 9, 10, 11\}$  all choose  $a_2$  but  $\{1, 2, 3, 4, 5\}$  choose  $a_1$ . The minimum cost  $\mathbf{a}_2$ -tree is  $\{\mathbf{a}_1 \rightarrow \mathbf{a} \rightarrow \mathbf{a}_2, \mathbf{b} \rightarrow \mathbf{a}_2\}$ , where the cost of transition  $\mathbf{a}_1 \rightarrow \mathbf{a}$  is 2 mutations, transition  $\mathbf{a} \rightarrow \mathbf{a}_2$  is 2 mutations, and  $\mathbf{b} \rightarrow \mathbf{a}_2$  is also 2 mutations, so that the total cost of a minimum  $\mathbf{a}_2$ -tree is  $C(\mathbf{a}_2) = 6$ . The minimum cost  $\mathbf{a}_1$ -tree on the other hand is  $\{\mathbf{a}_2 \rightarrow \mathbf{a} \rightarrow \mathbf{a}_1, \mathbf{b} \rightarrow \mathbf{a}_1\}$ ; and following a similar argument  $C(\mathbf{a}_1) = 6$ ; and hence, both  $\mathbf{a}_1$  and  $\mathbf{a}_2$  are stochastically stable.

scales beyond economic relevance to reach the long-run stable configuration. We tackle this problem in the next Section, showing that if a globally contagious choice exists, then the expected waiting time to the configuration with the largest long-run probability is independent of  $n$ .

## 5. Contagion and expected waiting time

This section establishes the relationship between the expected waiting time to the configuration with the largest long-run probability, and the property of contagion. We show that if the maximum long-run probability configuration consists of agents choosing a globally pairwise contagious option, then the expected waiting time to such a configuration from every other configuration is independent of the population size when  $\beta$  is large. The direct implication of this result is that even in large networks and with low levels of experimentation, the diffusion process converges fast to the configurations with the largest long-run probabilities. The problem of slow diffusion does not arise in situations where the level of experimentation is sufficiently high. Kreindler and Young (2013) and Kreindler and Young (2014) indeed show that when  $\beta$  is sufficiently small. Here, we examine the case where  $\beta$  is very large (i.e.  $\beta \rightarrow \infty$ ) so that  $\beta'_m$  as defined above, is negligibly small compared to  $\beta_m$ . The expected waiting time is formally defined as follows.

**Definition 5.** Let  $W \subset \mathbf{X}$  be a subset of the configuration space and  $\bar{W}$  its complement. Define  $T(W) = \inf\{t \geq 0 \mid \mathbf{x}_t \in W\}$  to be the first time  $W$  is reached. The expected waiting time from some configuration  $\mathbf{x} \in \bar{W}$  to  $W$  is then defined as  $\mathbb{E}[T(W) \mid \mathbf{x}_0 = \mathbf{x}]$ .

Let  $L_*$  be the configuration with maximum long-run probability. We aim to show that diffusion is *fast* when  $L_*$  is globally pairwise contagious. That is, there exists a function  $F(\beta)$  that is independent of  $n$  so that  $\mathbb{E}[T(L_*) \mid \mathbf{x}_0 = \mathbf{x}] \leq F(\beta)$  for any initial configurations  $\mathbf{x}_0$ .

Define  $r_*(L_j) = \min_{\mathbf{x} \in \mathbf{X}} r_{\mathbf{x}}(L_j)$ , to be the minimum number of experimentations required to exit  $D(L_j)$ , the BOA of  $L_j$ , also generally referred to the *radius* of  $D(L_j)$ . Clear, for  $\beta > \beta^*$ , if  $L_*$  is the limit configuration with the maximum long-run probability, then

$$r(L_*) = \sum_{L_j \in \mathbf{L}} r_{L_*}(L_j) = \sum_{L_j \in \mathbf{L}} r_*(L_j)$$

Let  $r_* = \max_{L_j \in \mathbf{L}} r_*(L_j)$  and  $\mathbb{E}[T(L_*)] = \max_{\mathbf{x} \in \mathbf{X}} \mathbb{E}[T(L_*) \mid \mathbf{x}_0 = \mathbf{x}]$ . The following proposition relates the expected waiting time to the limit configuration with maximum long-run probability  $\mathbb{E}[T(L_*)]$ , to  $r_*$ , the maximum-minimum number of experimentations required to exit a BOA.

**Proposition 3.** For a strict coordination game diffusion process  $(U, G, P_\beta)$ , if  $L_*$  is the configuration with the maximum long-run probability, then

$$\lim_{\beta \rightarrow \infty} \frac{\ln \mathbb{E}[T(L_*)]}{\beta} = r_* \quad (11)$$

*Proof.* See Appendix [Appendix B.6](#)

□

Proposition 3 shows that the expected waiting time to reach the configuration with maximum long-run probability from any other configuration takes the form

$$\mathbb{E}[T(L_*)] \leq \exp[\beta r_* + f(m, n, \beta)] \quad (12)$$

The function  $f(m, n, \beta)$  increases with  $m$  and  $n$  and decreases with  $\beta$ . For  $\beta \rightarrow \infty$ ,  $f(m, n, \beta) \rightarrow 0$ . The quantity  $r_*$  is a maximum radius of a basin of attraction as defined above. If  $L_*$  consists of agents coordinating on a choice that is globally pairwise contagious, then for each  $L_j \in \mathbf{L}$ ,  $r_*(L_j) \ll \frac{1}{2}n$ , small and independent of  $n$ . This directly implies that  $r_* \ll \frac{1}{2}n$ , small and independent of  $n$ . Hence for  $\beta$  very large, the expected waiting time to the configuration with maximum long-run probability is independent of  $n$  if a globally pairwise contagious choice exists.

Compared to existing results on convergence rates of evolutionary processes such as Ellison (1993), Young (2011), Kreindler and Young (2013) and Kreindler and Young (2014), the result in Proposition 3 is driven more by contagion and less by noise. Kreindler and Young (2014) also find that learning is fast in networks, but they consider a  $2 \times 2$  coordination game with random sampling and with deterministic dynamics. Moreover, they define fast learning as the case in which noise is large to the extent that only one unique equilibrium exists. On a contrary, Proposition 3 shows that under appropriate conditions learning is also fast in stochastic evolutionary processes of  $m \times m$  coordination games. Young (2011) shows that learning is fast in stochastic evolutionary processes with sufficiently large noise provided the network is made up of cohesive subgroups. Proposition 3 and the discussion in Section 4 provides insights into the findings in Young (2011). We showed in Section 4 that the existence of cohesive subgroups leads to multiple limit states in which choices co-exist. The larger the number of such limit states, the smaller the number of mutations involved in the transition from any one limit state to another; and hence, the smaller  $r_*$ .

Montanari and Saberi (2010) provide bounds for expected waiting times in a stochastic evolutionary process with sequential dynamics. Just as in Proposition 3 above, they also consider the case of small noise. They however derive bounds for expected waiting times as increasing functions of the population size. This means that as  $n$  grows, the expected waiting times become infinitely large and hence learning is slow. Proposition 3 establishes conditions under which learning is fast in situations with small noise.

The results of Proposition 3 are also related to Morris (2000), Sandholm (2001) and Oyama et al. (2015). Since our definitions of global and path-wise contagion are extensions of the notion of pair-wise contagion according to Morris (2000), Proposition 3 then directly implies that learning is fast in the  $2 \times 2$  symmetric game in Morris (2000). Sandholm (2001) and Oyama et al. (2015) show that  $\frac{1}{k}$ -dominant choices are almost globally stable in a random interaction model where agents best-respond to a sample of  $k$  agents selected from the population. An interaction structure where agents randomly sample  $k$  others in the population has a contagion threshold  $\eta(G) = \frac{1}{k}$ . Lemma 4 then implies that a  $\frac{1}{k}$ -dominant choice is globally contagious in such an interaction structure. Proposition 3 thus suggests that learning is fast in evolutionary game models of Sandholm (2001) and Oyama et al. (2015).



## 6. Concluding remarks

We have studied the diffusion of products and practices that exhibit network externalities using evolutionary game theory framework. Evolutionary game theory captures many realistic aspects of individual decision processes, and most notably, the tendency to experiment or make mistakes on optimal choices, and myopia – the inability to remember the entire history of play in complex social interactions. We examined how the level of experimentation by decision makers affects diffusion, showing that it induces threshold dynamics on long-run probability of a given choice being adopted.

We demonstrated how firms and governments can harness the power of social networks to strategically diffuse products and practices. We show how the process of contagion can be exploited to reduce the cost of targeting and increase the likelihood of a product/practice being adopted.

Our analysis focused on decision processes where agents make choices through best response, and we defined contagion as spreading through best response. One direction in which our model can be extended is considering situations where agents make decisions through imitation rather than best response. There are cases in evolutionary game theory where imitation and best response lead to different long-run outcomes (see for example [Alós-Ferrer and Weidenholzer \(2008\)](#)). The notion of contagion could then be defined as spreading through imitation.

## Appendix A. An example of equilibria with co-existing choices

Consider payoffs for a bilateral interaction between the row and column players depicted in Table A.1. The underlying payoffs exhibit coordination properties, and the bilateral interaction is a coordination game. This game has three Nash equilibria: Two of which are pure strategy Nash equilibria where players coordinate on choices  $a_1$  and  $a_2$ , and the third is a mixed-strategy Nash equilibrium in which both the row and column players choose  $a_1$  with probability  $\phi_{21} = \frac{3}{5}$  and  $a_2$  with probability  $\phi_{12} = \frac{2}{5}$ . This follows from the fact that if  $\phi_{12}$  and  $\phi_{21} = 1 - \phi_{12}$  are the probabilities with which the column player chooses  $a_2$  and  $a_1$  respectively so that the row player's payoff from choosing  $a_1$  and  $a_2$  respectively are  $\phi_{12}u(a_1, a_2) + (1 - \phi_{12})u(a_1, a_1)$  and  $\phi_{12}u(a_2, a_2) + (1 - \phi_{12})u(a_2, a_1)$ , then the mixed strategy equilibrium is a solution to  $\phi_{12}u(a_1, a_2) + (1 - \phi_{12})u(a_1, a_1) = \phi_{12}u(a_2, a_2) + (1 - \phi_{12})u(a_2, a_1)$ , which yields

$$\phi_{12} = \frac{u(a_1, a_1) - u(a_2, a_1)}{u(a_1, a_1) - u(a_2, a_1) + u(a_2, a_2) - u(a_1, a_2)} = \frac{2}{5}$$

$$\phi_{21} = \frac{u(a_2, a_2) - u(a_1, a_2)}{u(a_1, a_1) - u(a_2, a_1) + u(a_2, a_2) - u(a_1, a_2)} = \frac{3}{5}$$

For any interaction structure, the monomorphic configurations  $\mathbf{a}_1 = (a_1, a_1, a_1, a_1, a_1, a_1)$  and  $\mathbf{a}_2 = (a_2, a_2, a_2, a_2, a_2, a_2)$  are limit configurations of  $P$ . Now consider the interaction structure in Figure A.5 with the interaction matrix  $G$  as defined. For this network, and payoff in Table A.1, in addition to limit configurations  $\mathbf{a}_1$  and  $\mathbf{a}_2$ , the configurations  $\mathbf{a} = (a_1, a_1, a_1, a_2, a_2, a_2)$  and  $\mathbf{b} = (a_2, a_2, a_2, a_1, a_1, a_1)$  where choices co-exist are also limit configurations of  $P$ . To see why, first note that the quantities  $\phi_{12}$

	$a_1$	$a_2$
$a_1$	4, 4	0, 2
$a_2$	2, 0	3, 3

Table A.1: A 2-player 2-choice coordination game.

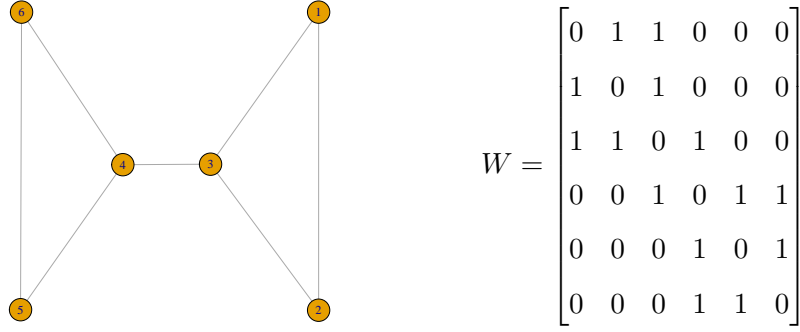


Figure A.5: An example of a network with cohesive subgroups, and the respective interaction matrix  $W$  on the right hand side.

and  $\phi_{21}$  are the minimum proportion of neighbours an agent requires to adopt  $a_2$  and  $a_1$  respectively for her to do likewise.

Similarly,  $\lceil \phi_{12} k_i \rceil$  and  $\lceil \phi_{21} k_i \rceil$ , where  $\lceil x \rceil$  is the smallest integer larger than  $x$ , are the numbers of neighbours agent  $i$  requires to choose  $a_2$  and  $a_1$  respectively for  $i$  to do likewise. For the game in Table A.1 and network in Figure A.5, it follows that agents  $\{1, 2, 5, 6\}$  who each have degree  $k = 2$ , will switch from  $a_1$  to  $a_2$  if at least  $\lceil \frac{4}{5} \rceil = 1$  of their neighbours choose  $a_2$  and from  $a_2$  to  $a_1$  if  $\lceil \frac{6}{5} \rceil = 2$  of their neighbours choose  $a_1$ . The bottleneck however occurs with agents  $\{3, 4\}$  who each require at least  $\lceil \frac{6}{5} \rceil = 2$  of their neighbours to choose  $a_2$  for them to do likewise. Thus, a state in which agents  $\{1, 2, 3\}$  play  $a_1$  and  $\{4, 5, 6\}$  play  $a_2$  is an equilibrium, and so is a state in which  $\{1, 2, 3\}$  play  $a_2$  and  $\{4, 5, 6\}$  play  $a_1$ . And hence, configurations **a** and **b** are absorbing configurations of  $P$ .

## Appendix B. Proofs

### Appendix B.1. Proof of Lemma 1

We use the following results from Freidlin and Wentzell (1984) to prove Lemma 1.

**Lemma 5.** (Freidlin and Wentzell, 1984, Lemma 3.1). *Given a diffusion process  $P_\beta$ , the stationary distribution  $\pi_\beta(\mathbf{x})$  of some configuration  $\mathbf{x} \in \mathbf{X}_n$  is given by*

$$\pi_\beta(\mathbf{x}) = \left( \sum_{g \in \Gamma(\{\mathbf{x}\})} P_\beta(g) \right) \left( \sum_{\mathbf{y} \in \mathbf{X}_n} \sum_{g \in \Gamma(\{\mathbf{y}\})} P_\beta(g) \right)^{-1} \quad (\text{B.1})$$

where the total probability  $P_\beta(g)$  associated with each graph  $g$  is  $P_\beta(g) = \prod_{(\mathbf{z}, \mathbf{y}) \in g} P_\beta(\mathbf{z}, \mathbf{y})$ .

We start by defining  $\Gamma$  on the state space  $\mathbf{X}$  as opposed to  $\mathbf{L}$  in Definition 1.

**Definition 6.** Let  $g \subset \mathbf{X} \times \mathbf{X}$  be any oriented graph defined within the configuration space  $\mathbf{X}$ . Then for a subset  $W \subset \mathbf{X}$  and its complement  $\bar{W}$ , we denote by  $\Gamma(W)$  a set of all oriented graphs satisfying two conditions: (i) no arrows start from  $W$  and exactly one arrow starts from each configuration outside of  $W$ , (ii) each  $g \in \Gamma(W)$  has no loops.

From Definition 6, if  $W$  is a singleton set, say  $W = \{\mathbf{x}\}$ , then  $\Gamma(\{\mathbf{x}\})$  is a set of all spanning trees of  $\mathbf{x}$  (i.e.  $\mathbf{x}$ -trees). Consider the case where  $\mathbf{X} = \{\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \mathbf{e}, \mathbf{f}, \mathbf{g}, \mathbf{h}\}$ ; Figure B.6 presents two examples of  $\mathbf{g}$ -trees.

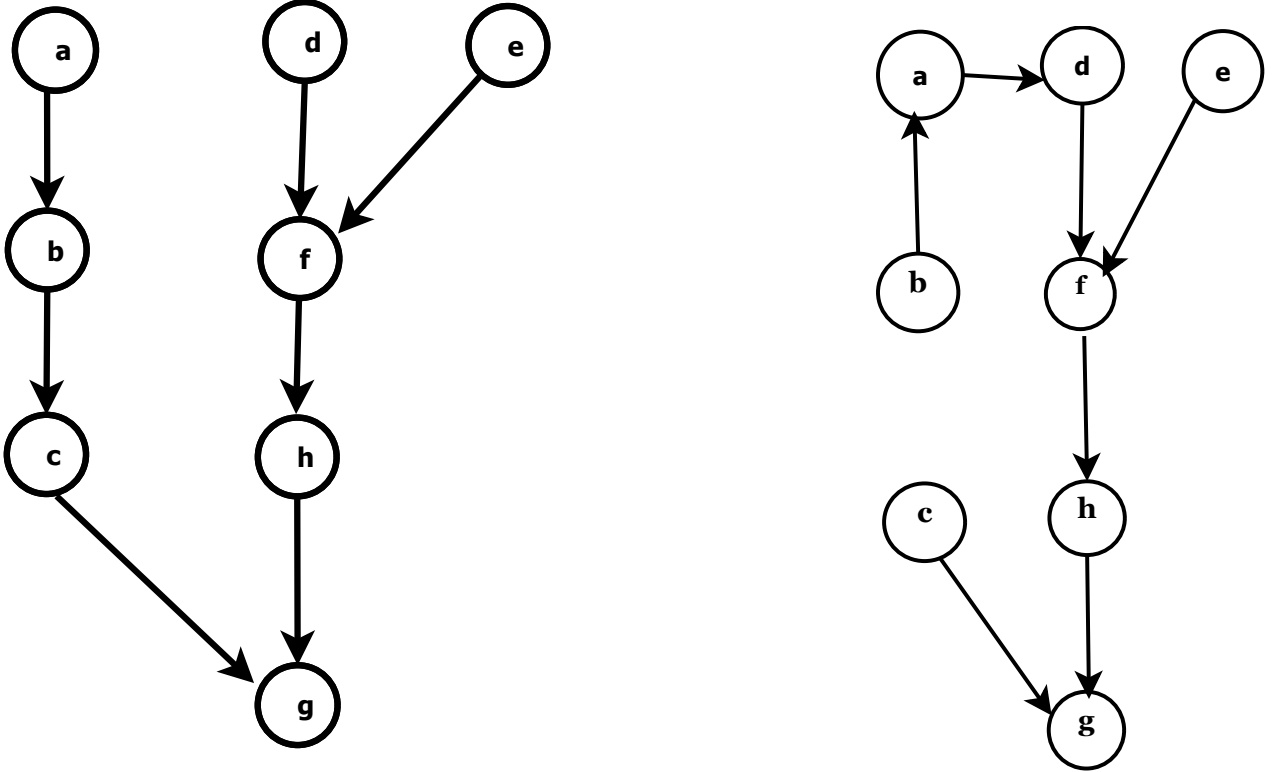


Figure B.6: Examples of  $\mathbf{g}$ -trees, that is  $\Gamma(\{\mathbf{g}\})$  graphs and in which the configuration space  $\mathbf{X} = \{\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \mathbf{e}, \mathbf{f}, \mathbf{g}, \mathbf{h}\}$ .

Now, let  $d(L_j)$  be the cardinality of  $D(L_j)$ . For each  $D(L_l)$ , let  $\partial L_l = D(L_l) \setminus L_l$ , the basin of attraction of  $L_l$  with  $L_l$  excluded. Each graph  $g \in \Gamma(\{\mathbf{x}\})$  can be partitioned into subgraphs  $g(D(L))$  and  $g(\partial L, L)$  defined as follows. For some  $L_j \in \mathbf{L}$  and corresponding subset of configurations  $D(L_j)$ ,  $g(D(L_j))$  is a set of subgraphs of  $g$  consisting of transitions between either the configurations in  $D(L_j)$  or from a configuration in  $D(L_j)$  to another not in  $D(L_j)$ , and that each of these transitions involves at least one experimentation. The subgraph  $g(\partial L_j, L_j)$  on the other hand consists of transitions of  $g$  among configurations of  $D(L_j)$  that involve best response. From the definition of  $r_g(L_j)$  in Section 3, it is equivalent to the total number of experimentations involved in the transitions of sub-graph  $g(D(L_j)) \in g$ . We let  $r_g(\partial L_j, L_j)$  be the cardinality of  $g(\partial L_l, L_j)$ .

Given the partition of  $g \in \Gamma(\{\mathbf{x}\})$  into  $g \equiv \bigcup_{L_j \in \mathbf{L}} \{g(D(L_j)) \cup g(\partial L_j, L_j)\}$ , each  $P_\beta(g)$  can be

rewritten as

$$P_\beta(g) = \prod_{L_j \in \mathbf{L}} P_\beta(g(D(L_j))) P_\beta(g(\partial L_j, L_j)) \quad (\text{B.2})$$

where  $P_\beta(g(D(L_j))) = \prod_{(\mathbf{y}, \mathbf{z}) \in g(D(L_j))} P_\beta(\mathbf{y}, \mathbf{z})$  and  $P_\beta(g(\partial L_j, L_j)) = \prod_{(\mathbf{y}, \mathbf{z}) \in g(\partial L_j, L_j)} P_\beta(\mathbf{y}, \mathbf{z})$ . The following lemma follows from above definitions.

**Lemma 6.** *Given  $g \in \Gamma(\{\mathbf{x}\})$ , let  $r_g(L_j)$  be the number of experimentations involved in the transitions of sub-graph  $g(D(L_j)) \in g$  and let  $r_g(\partial L_j, L_j)$  be the cardinality of  $g(\partial L_j, L_j)$ . Then*

$$P_\beta(g) = \exp \left[ - \sum_{L_j \in \mathbf{L}} \left( r_g(L_j) \beta_m + r_g(\partial L_j, L_j) \beta'_m \right) \right] \quad (\text{B.3})$$

where  $\beta_m = \beta - \ln m^{-1}$  and  $\beta'_m = -n \ln \left[ 1 - e^{-\beta} \left( 1 - \frac{1}{m} \right) \right]$ .

*Proof.* The first part of the proof follows from the fact that the probabilities of experimentation are identical for all agents and configurations, that is  $\frac{1}{m} \exp(-\beta)$ , it follows that

$$P_\beta(g(D(L_j))) = \left( \frac{1}{m} \exp(-\beta) \right)^{r_g(L_j)} = \exp \left( r_g(L_j) \left[ \ln m^{-1} - \beta \right] \right) \quad (\text{B.4})$$

For the probability  $P_\beta(g(\partial L_j, L_j))$ , since each transition in  $g(\partial L_j, L_j)$  is dominated by best-response dynamics, then from (3) each transition  $\mathbf{x} \rightarrow \mathbf{y}$  in  $g(\partial L_j, L_j)$  has the probability  $[(1 - e^{-\beta}) + \frac{1}{m} e^{-\beta}]^n = [1 - e^{-\beta} (1 - \frac{1}{m})]^n$ , where the power of  $n$  results from the assumption of simultaneous decision process. That is,  $y^i$  for each  $i$  is a best response to  $\mathbf{x}$  so that  $\mathbf{y}$  consists of  $n$  simultaneous best responses to  $\mathbf{x}$ . We have also assumed without loss of generality that  $b = 1$ . Let  $r_g(\partial L_j, L_j)$  be the cardinality of  $g(\partial L_j, L_j)$  under graph  $g$ , then

$$P_\beta(g(\partial L_j, L_j)) = \left( 1 - e^{-\beta} \left( 1 - \frac{1}{m} \right) \right)^{nr_g(\partial L_j, L_j)} = \exp \left( nr_g(\partial L_j, L_j) \ln \left[ 1 - e^{-\beta} \left( 1 - \frac{1}{m} \right) \right] \right) \quad (\text{B.5})$$

For notational convenience, let  $\beta_m = \beta - \ln m^{-1}$  and  $\beta'_m = -n \ln \left[ 1 - e^{-\beta} \left( 1 - \frac{1}{m} \right) \right]$ . It follows from (B.4) and (B.5) that

$$P_\beta(g) = \exp \left[ - \sum_{L_j \in \mathbf{L}} \left( r_g(L_j) \beta_m + r_g(\partial L_j, L_j) \beta'_m \right) \right] \quad (\text{B.6})$$

□

Let  $Z(\beta)$  denote the normalization factor for  $\pi_\beta$ , that is

$$Z(\beta) = \sum_{\mathbf{y} \in \mathbf{X}_n} \sum_{g \in \Gamma(\{\mathbf{y}\})} P_\beta(g)$$

The normalization factor  $Z(\beta)$  is identical for all configurations  $\mathbf{x} \in \mathbf{X}_n$ . Let  $K = \#\Gamma(\{\mathbf{x}\})$  be the cardinality of  $\Gamma(\{\mathbf{x}\})$ . Similarly,  $K$  is identical for all  $\mathbf{x} \in \mathbf{X}_n$ . Given the uniformity of  $Z(\beta)$  and  $K$ ,  $\pi_\beta(\mathbf{x})$  is then bounded by

$$Z^{-1}(\beta) \max_{g \in \Gamma(\{\mathbf{x}\})} P_\beta(g) \leq \pi_\beta(\mathbf{x}) \leq Z^{-1}(\beta) K \max_{g \in \Gamma(\{\mathbf{x}\})} P_\beta(g) \quad (\text{B.7})$$

Let  $g_{\min}(\mathbf{x}) = \operatorname{argmax}_{g \in \Gamma(\{\mathbf{x}\})} P_\beta(g)$  be the  $\mathbf{x}$ -tree with the maximum probability  $P_\beta(g)$ . Then

$$P_\beta(g_{\min}(\mathbf{x})) = \max_{g \in \Gamma(\{\mathbf{x}\})} P_\beta(g) = \exp \left[ - \sum_{L_j \in \mathbf{L}} \left( r_{\mathbf{x}}(L_j) \beta_m + r_{\mathbf{x}}(\partial L_j, L_j) \beta'_m \right) \right] \quad (\text{B.8})$$

where  $r_{\mathbf{x}}(\partial L_j, L_j)$  and  $r_{\mathbf{x}}(L_j)$  are the corresponding values of  $r_g(\partial L_j, L_j)$  and  $r_g(L_j)$  respectively when  $g = g_{\min}(\mathbf{x})$ .

It remains to show that along  $g_{\min}(\mathbf{x})$ , the quantity  $r_{\mathbf{x}}(\partial L_j, L_j) = d(L_j) - r_{\mathbf{x}}(L_j)$  for each  $L_j$ , where  $d(L_j)$  is the cardinality of  $D(L_j)$ . Note that for any  $\beta > 0$ , if  $a_j$  is a best response to  $\mathbf{x}$  then  $(1 - \exp(-\beta))BR_i(a_j; \mathbf{x}) > 0$ , which implies that the probability  $(1 - \exp(-\beta))BR_i(a_j; \mathbf{x}) + \frac{1}{m} \exp(-\beta)$  of choosing  $a_j$  given  $\mathbf{x}$  is higher than the probability  $\frac{1}{m} \exp(-\beta)$  of choosing any other  $a_l \notin BR(\mathbf{x})$  through experimentation. The direct implication is that when  $g = g_{\min}(\mathbf{x})$ , the subgraph  $g(D(L_j))$  for each  $L_j$  consists of  $r_{\mathbf{x}}(L_j)$  transitions where only one agent experiments and the rest choose a best response option, so that (see the proof of Lemma 6 above)

$$P_\beta(g(D(L_j))) = \exp \left( r_{\mathbf{x}}(L_j) \left[ \ln m^{-1} - \beta \right] \right) \quad (\text{B.9})$$

$$P_\beta(g(\partial L_j, L_j)) = \exp \left( n(d(L_j) - r_{\mathbf{x}}(L_j)) \ln \left[ 1 - e^{-\beta} \left( 1 - \frac{1}{m} \right) \right] \right) \quad (\text{B.10})$$

The second alternative for the composition of  $g(D(L_j))$  is a single direct transition where  $r_{\mathbf{x}}(L_j)$  agents simultaneously experiment to choose a non best response choice, so that  $r_{\mathbf{x}}(\partial L_j, L_j) = d(L_j) - 1$ . In this case, the total probability  $P_\beta(g(D(L_j)))$  of subgraph  $g(D(L_j))$  is identical to that in (B.9), but  $P_\beta(g(\partial L_j, L_j))$  is less than that in (B.10). When this argument is extended to all  $L_j \in \mathbf{L}$ , the conclusion is that the likelihood of reaching  $\mathbf{x}$  through the second  $\mathbf{x}$ -tree is smaller than that through  $g = g_{\min}(\mathbf{x})$ . Any other  $\mathbf{x}$ -tree different from these two alternatives will imply that the composition  $g(D(L_j))$  is so that  $r_g(L_j) > r_{\mathbf{x}}(L_j)$  and/or  $r_{\mathbf{x}}(\partial L_j, L_j) > d(L_j) - r_{\mathbf{x}}(L_j)$ , for each  $L_j \in \mathbf{L}$ . This in turn implies that  $P_\beta(g(D(L_j)))$  and  $P_\beta(g(\partial L_j, L_j))$  are greater than those in (B.9) and (B.10) respectively.

It follows from (B.7) that

$$\begin{aligned} & K^{-1} \exp \left[ - \sum_{L_j \in \mathbf{L}} \left( [r_{\mathbf{x}}(L_j) - r_{\mathbf{y}}(L_j)] \beta_m + [r_{\mathbf{x}}(\partial L_j, L_j) - r_{\mathbf{y}}(\partial L_j, L_j)] \beta'_m \right) \right] \leq \frac{\pi_\beta(\mathbf{x})}{\pi_\beta(\mathbf{y})} \\ & \leq K \exp \left[ - \sum_{L_j \in \mathbf{L}} \left( [r_{\mathbf{x}}(L_j) - r_{\mathbf{y}}(L_j)] \beta_m + [r_{\mathbf{x}}(\partial L_j, L_j) - r_{\mathbf{y}}(\partial L_j, L_j)] \beta'_m \right) \right] \end{aligned}$$

which can be rewritten as

$$\frac{\pi_\beta(\mathbf{x})}{\pi_\beta(\mathbf{y})} = \exp \left[ - \sum_{L_j \in \mathbf{L}} \left( [r_{\mathbf{x}}(L_j) - r_{\mathbf{y}}(L_j)] \beta_m + [r_{\mathbf{x}}(\partial L_j, L_j) - r_{\mathbf{y}}(\partial L_j, L_j)] \beta'_m \right) \pm \kappa \right] \quad (\text{B.11})$$

where  $\kappa = \ln(K)$ . Substituting for  $r_{\mathbf{x}}(\partial L_j, L_j) = d(L_j) - r_{\mathbf{x}}(L_j)$  in (B.11) yields

$$\frac{\pi_\beta(\mathbf{x})}{\pi_\beta(\mathbf{y})} = \exp \left[ - \sum_{L_j \in \mathbf{L}} \left( (r_{\mathbf{x}}(L_j) - r_{\mathbf{y}}(L_j)) (\beta_m - \beta'_m) \right) \pm \kappa \right] \quad (\text{B.12})$$

From (B.12), we see that it is sufficient to focus on the  $\Gamma$  graphs defined on the state space  $\mathbf{L}$  as in Definition 1, so that for any pair of limit states  $L$  and  $L'$ ,

$$\frac{\pi_\beta(L)}{\pi_\beta(L')} = \exp [-(r_L - r_{L'}) (\beta_m - \beta'_m) \pm \kappa] \quad (\text{B.13})$$

### Appendix B.2. Proof of Proposition 1

To prove Proposition 1, it helps to first examine how  $\Phi_\beta(L, L')$  varies with  $\beta$ , and in particular, the quantity  $\beta_m - \beta'_m$ . At the maximum level of experimentation, when  $\beta = 0$ ,

$$\beta_m - \beta'_m = (n - 1) \ln(m^{-1}) = -(n - 1) \ln(m) < 0 \quad (\text{B.14})$$

From (B.14), it follows that  $r_L(\beta) < 0$ , and is a decreasing function of  $n$ . When  $\beta$  increases to infinity (where agents make choices only through best response),  $\beta_m - \beta'_m$  also tends to positive infinity, so that  $r_L(\beta) = \infty > 0$ .

Now, consider the case where  $r_L > r_{L'}$  so that the number of mutations involved in the transitions of the  $L$ -tree is larger than for an  $L'$ -tree. It follows from above discussion that when  $\beta = 0$ ,  $\Phi_\beta(L, L') > 0$ , and when  $\beta = \infty$ ,  $\Phi_\beta(L, L') < 0$ . Thus, there exists a value of  $\beta = \beta^*$ , which is a solution to the equality  $\beta_m = \beta'_m$ , below which  $\Phi_\beta(L, L') > 0$  and above which  $\Phi_\beta(L, L') < 0$ .

The result then follows from Lemma 1 whereby, when  $\Phi_\beta(L, L') > 0$ ,  $\frac{\pi_\beta(L)}{\pi_\beta(L')} > 1$ , which implies that  $\pi_\beta(L) > \pi_\beta(L')$ ; and when  $\Phi_\beta(L, L') < 0$ ,  $\frac{\pi_\beta(L)}{\pi_\beta(L')} < 1$ , which implies that  $\pi_\beta(L) < \pi_\beta(L')$ .

### Appendix B.3. Proof of Lemma 3

The respective expressions for  $\eta_{jl}$  and  $\phi_{jl}$  are obtained by substituting the expressions of expected utility from (6) into (7) and (8) respectively. That is,  $(1 - q)u(a_l, a_j) + qu(a_l, a_l) > (1 - q)u(a_k, a_j) + qu(a_k, a_l)$ , for all  $a_k \neq a_l$ ; which yields

$$q > \frac{u(a_k, a_j) - u(a_l, a_j)}{u(a_k, a_j) - u(a_l, a_j) + u(a_l, a_l) - u(a_k, a_l)} \quad \text{for all } a_k \neq a_l$$

Hence,  $a_l$  is a best response to  $\sigma_{jl}^q$  whenever  $q$  is greater than

$$\eta_{jl} = \max_{a_k \neq a_l} \left\{ \frac{u(a_k, a_j) - u(a_l, a_j)}{u(a_k, a_j) - u(a_l, a_j) + u(a_l, a_l) - u(a_k, a_l)} \right\}$$

The expression for  $\phi_{jl}$  on the other hand is derived by solving for  $q$  in  $(1 - q)u(a_l, a_j) + qu(a_l, a_l) > (1 - q)u(a_j, a_j) + qu(a_j, a_l)$ , which yields

$$q > \frac{u(a_j, a_j) - u(a_l, a_j)}{u(a_j, a_j) - u(a_l, a_j) + u(a_l, a_l) - u(a_j, a_l)}$$

To prove Lemma 3 (i), note that for any pair of strategies  $a_j, a_l \in A$ ,

$$\begin{aligned} \phi_{jl} &= \frac{u(a_j, a_j) - u(a_l, a_j)}{u(a_j, a_j) - u(a_l, a_j) + u(a_l, a_l) - u(a_j, a_l)} \\ &= 1 - \frac{u(a_l, a_l) - u(a_j, a_l)}{u(a_j, a_j) - u(a_l, a_j) + u(a_l, a_l) - u(a_j, a_l)} \\ &= 1 - \phi_{lj}. \end{aligned} \quad (\text{B.15})$$

To prove Lemma 3 (ii), recall the expression for  $\eta_{jl}$  as

$$\eta_{jl} = \max_{a_k \neq a_l} \left\{ \frac{u(a_k, a_j) - u(a_l, a_j)}{u(a_k, a_j) - u(a_l, a_j) + u(a_l, a_l) - u(a_k, a_l)} \right\}$$

From the expression of  $\phi_{jl}$ , when  $a_k = a_j$ , we have  $\eta_{jl} = \phi_{jl}$ . Let  $\eta_{jl}^{-j}$  be derived from  $\eta_{jl}$  by excluding the  $j$ th row of the payoff matrix. That is,

$$\eta_{jl}^{-j} = \max_{a_k \neq a_l, a_j} \left\{ \frac{u(a_k, a_j) - u(a_l, a_j)}{u(a_k, a_j) - u(a_l, a_j) + u(a_l, a_l) - u(a_k, a_l)} \right\}$$

Then  $\eta_{jl}$  can be rewritten as

$$\eta_{jl} = \max \left\{ \phi_{jl}, \eta_{jl}^{-j} \right\}$$

And hence,  $\eta_{jl} \geq \phi_{jl}$ .

#### Appendix B.4. Proof of Lemma 4

Let  $\eta(G) = p$ . Recall the definition of contagion threshold: starting from a monomorphic absorbing state  $\mathbf{a}_j$ , define a sequence  $\{S_\tau(a_l)\}_{\tau \geq 1}$  whereby  $S_\tau(a_l) \subset S_{\tau+1}(a_l)$ , and a complementary sequence  $\{\bar{S}_\tau(a_j)\}_{\tau \geq 1}$  for which  $\bar{S}_\tau(a_j) \supset \bar{S}_{\tau+1}(a_j)$ . The contagion threshold of the underlying network is the largest  $p$ , so that for a corresponding sequence  $\{\bar{S}_\tau^{[p, S]}(a_j)\}_{\tau \geq 1}$  (each  $\bar{S}_\tau^{[p, S]}(a_l)$  is a subset of players in  $\bar{S}_\tau(a_l)$  for whom at least proportion  $p$  of their interactions are with players in  $S_\tau(a_j)$ ),  $a_l$  is a best response for each  $i \in \bar{S}_\tau^{[p, S]}(a_j)$ . Pairwise contagion of  $a_l$  relative to  $a_j$ , starting from  $\mathbf{a}_j$ , is then triggered whenever  $S(a_l) = N(\mathbf{a}_j \rightarrow \mathbf{a}_l)$ .

By definition, a contagion threshold is equivalent to the minimum proportion of neighbours any player requires to play  $a_l$ , with the complementary proportion playing  $a_j$ , for him to switch from  $a_j$  to  $a_l$ . That is, if  $p$  is the contagion threshold, then for all  $q \geq p$ ,  $a_l$  is a best response to  $\sigma_{jl}^q$  for all players. Recall the definition of  $\eta_{jl}$  as a minimum  $q$  (i.e.  $q > \eta_{jl}$ ) above which  $a_l$  is the best response to  $\sigma_{jl}^q$ . Hence, pairwise contagion of  $a_l$  relative to  $a_j$  is feasible if and only if  $\eta_{jl} < p$ .

To complete the proof, we next show that contagion occurs only if  $\eta_{jl} = \phi_{jl}$ . First, any distribution that assigns probability  $p$  to any choice  $a_k \neq a_l$  and  $1 - p$  to  $a_j$  does not support pair-wise contagion between  $a_j$  and  $a_l$ . Clearly if choice  $a_l$  is a best response to the distribution in which a proportion  $p$  of a player's neighbours play  $a_k$  and  $1 - p$  play  $a_j$ , then starting from  $\mathbf{a}_j$ , each player requires at least a fraction  $p$  of their neighbours to mutate to  $a_k$  for them to switch from  $a_j$  to  $a_l$ . Such a process is thus driven by mutations and not best response dynamics as required for contagion. Second, if  $\eta_{jl} \neq \phi_{jl}$ , then there exists a distribution in which proportion  $q$  of neighbours play  $a_l$  and  $1 - q$  play  $a_j$ , for some  $\phi_{jl} < q < \eta_{jl}$  such that another choice  $a_k \neq a_l$  is a best response. Hence in transiting from  $\mathbf{a}_j$  to  $\mathbf{a}_l$ , the process goes through some states in which choice  $a_k \neq a_l$  is played; such transitions are by definition not contagion.

Finally, Morris (2000, Proposition 3) shows that every network has a contagion threshold  $p \leq \frac{1}{2}$ . Hence, pairwise contagion of  $a_l$  relative to  $a_j$  occurs if and only if  $\eta_{jl} = \phi_{jl} < \eta(G) \leq \frac{1}{2}$ .



### Appendix B.5. Proof of Proposition 2

We aim to show that if  $a_l$  is globally pairwise contagious relative to every other choice, then the monomorphic configuration  $\mathbf{a}_l$  where all agents coordinate on  $a_l$  has the largest long-run probability whenever  $\beta > \beta^*$ ; that is,  $\pi_\beta(\mathbf{a}_l) > \pi_\beta(L)$  for all limit configurations  $L \neq \mathbf{a}_l$ . Recall that  $\pi_\beta(\mathbf{a}_l) > \pi_\beta(L)$  if the cost difference  $\Phi_\beta(L, \mathbf{a}_l) < 0$ , where

$$\Phi_\beta(L, \mathbf{a}_l) = -(\beta_m - \beta'_m) \sum_{L_j \in \mathbf{L}} ((r_L(L_j) - r_{\mathbf{a}_l}(L_j)))$$

For  $\beta > \beta^*$ , it is then sufficient to show that  $r_L > r_{\mathbf{a}_l}$  where  $r_L = \sum_{L_j \in \mathbf{L}} r_L(L_j)$  and  $r_{\mathbf{a}_l} = \sum_{L_j \in \mathbf{L}} r_{\mathbf{a}_l}(L_j)$ . Note that  $r_L$  and  $r_{\mathbf{a}_l}$  can be rewritten as

$$r_L = \sum_{L_j \neq L, \mathbf{a}_l} r_L(L_j) + r_L(\mathbf{a}_l) \quad (\text{B.16})$$

$$r_{\mathbf{a}_l} = \sum_{L_j \neq L, \mathbf{a}_l} r_{\mathbf{a}_l}(L_j) + r_{\mathbf{a}_l}(L) \quad (\text{B.17})$$

where, as defined in Section 3,  $r_L(\mathbf{a}_l)$  is the minimum number of mutations needed to exit the basin of attraction of  $\mathbf{a}_l$  to some state in the basin of attraction of the limit state that succeeds  $\mathbf{a}_l$  in the minimum cost  $L$ -tree. Specifically, if  $L'$  is the limit state succeeding  $\mathbf{a}_l$  in minimum cost  $L$ -tree, denoted by  $\bar{g}$ , then  $r_L(\mathbf{a}_l) = r_{\bar{g}}(\mathbf{a}_l, L')$ ; see the discussion in page 10 for details. A similar interpretation applies to  $r_L(\mathbf{a}_l)$ . Note that (B.16) and (B.17) follow from the expressions of  $r_L$  and  $r_{\mathbf{a}_l}$  because  $r_L(L) = 0$  and  $r_{\mathbf{a}_l}(\mathbf{a}_l) = 0$ .

From (B.16) and (B.17),  $r_L$  and  $r_{\mathbf{a}_l}$  are differentiated only by the terms  $r_L(\mathbf{a}_l)$  and  $r_{\mathbf{a}_l}(L)$ . The second terms on the right hands sides of (B.16) and (B.17) are identical. Thus, to show that  $r_L > r_{\mathbf{a}_l}$ , it is sufficient to show that  $r_L(\mathbf{a}_l) > r_{\mathbf{a}_l}(L)$ .

We first consider the case where  $L$  is a monomorphic state where all agents coordinate on some  $a_j \neq a_l$ . We then aim to show that for a pair of strategies  $a_j, a_l \in A$  and their respective monomorphic absorbing states  $\mathbf{a}_j$  and  $\mathbf{a}_l$ , if  $a_l \text{PRC} a_j$  then  $r_{\mathbf{a}_l}(\mathbf{a}_j) < r_{\mathbf{a}_j}(\mathbf{a}_l)$  for  $n$  sufficiently large. Recall that  $r_{\mathbf{a}_l}(\mathbf{a}_j)$  is the minimal number of mutations needed for the transition from  $\mathbf{a}_j$  to  $\mathbf{a}_l$  to occur. From the definition of pairwise contagion, we noted that if  $a_l \text{PRC} a_j$  then  $r_{\mathbf{a}_l}(\mathbf{a}_j) = n(\mathbf{a}_j \rightarrow \mathbf{a}_l)$ , the number of initial adopters of  $a_l$  needed to trigger contagion of  $a_l$  starting from  $\mathbf{a}_j$ . By definition,  $n(\mathbf{a}_j \rightarrow \mathbf{a}_l)$  is independent of  $n$  and is much smaller than  $\frac{n}{2}$ .

From Lemma 4, we saw that if  $a_l \text{PRC} a_j$ , then for each  $a_j \neq a_l$ ,  $\eta_{jl} = \phi_{jl} < p \leq \frac{1}{2}$ , which also implies that  $\eta_{lj} \geq \phi_{lj} > \frac{1}{2}$ . Thus, each player requires more than half of his neighbours to switch from  $a_l$  to  $a_j$  for him to do likewise. The transition from  $\mathbf{a}_l$  to  $\mathbf{a}_j$  then requires a large number of mutation that ensures the each player has more than half of his neighbours playing  $a_l$ . In this case, although the total number of mutations required depends on the network density, in general, it is a function of the network size  $n$ . For highly sparse networks (with very low density), the number of mutations required is close to  $\frac{n}{2}$ . As the network density tends to one (i.e. to a complete network), the number of mutations needed

approaches to  $\eta_{jl} \cdot (n - 1) > \frac{1}{2}$ ; this is because in a complete network, each player has  $n - 1$  neighbours. If each player requires more than  $\eta_{jl}$  of his neighbours to switch from  $a_j$  to  $a_l$  for him to do likewise, then the whole population will switch from  $a_j$  to  $a_l$  once proportion  $\eta_{jl} \cdot (n - 1)$  does so.

An example of a network with a very low density is a ring network, in which each player  $i$  has two neighbours  $i - 1$  on the left and  $i + 1$  on the right. In this network, the number of mutations required is  $\frac{n}{2}$ . This is because if  $\eta_{lj} > \frac{1}{2}$  then a player requires both of his neighbours to switch to  $a_l$  for him to do likewise. Thus, starting from  $\mathbf{a}_j$ , a mutation arrangement where for examples players  $0, 2, 4, \dots, i - 1, i + 2, \dots, n - 1$  experiment to play  $a_l$  is sufficient to switch the population from  $\mathbf{a}_j$  to  $\mathbf{a}_l$ . Pairwise contagion of  $a_l$  relative to  $a_j$  therefore implies that the cost  $r_{\mathbf{a}_j}(\mathbf{a}_l)$  is an increasing function of  $n$ ; hence, for  $n$  sufficiently large,  $r_{\mathbf{a}_l}(\mathbf{a}_j) < r_{\mathbf{a}_j}(\mathbf{a}_l)$ ; and since  $a_l$  is globally pairwise contagious,  $r_{\mathbf{a}_l}(\mathbf{a}_j) < r_{\mathbf{a}_j}(\mathbf{a}_l)$  for all  $a_j \neq a_l$ .

Next, we consider the case where  $L$  is a polymorphic limit states containing at least one choice. The first step is to note that if  $a_l$  is globally pairwise contagious (GPC), then it cannot co-exist with other choice; i.e. if  $L$  is polymorphic, then it contain two or more choices other than  $a_l$ . To see why, note from Lemma 3 and the subsequent discussion that  $L$  contains only two choice  $a_j$  and  $a_l$  if the  $\eta_{jl} \geq \phi_{jl} \geq \eta(G)$ . Since  $a_l$  is GPC,  $L$  cannot contain  $a_l$  and a second choice  $a_j \neq a_l$ . For  $L$  to contain three or more choices, then three of the choices, say  $a_j$ ,  $a_k$  and  $a_l$ , must *at least* (i.e. a necessary condition) satisfy the condition that either  $a_j \in \text{BR}(\sigma_{kl}^q)$  for  $\phi_{kl} \leq q \leq \eta_{kl}$ , or  $a_k \in \text{BR}(\sigma_{jl}^q)$  for  $\phi_{jl} \leq q \leq \eta_{jl}$ , or  $a_l \in \text{BR}(\sigma_{kj}^q)$  for  $\phi_{kj} \leq q \leq \eta_{kj}$ . The first two conditions however fail to satisfy the condition for pairwise contagion of  $a_l$  relative to  $a_j$  and  $a_k$ . Thus, only the third condition is feasible. The third condition would however imply that  $L$  constitutes a structure where agents choosing  $a_l$  are surrounded by agents choosing  $a_l$  and  $a_k$ . But since  $a_l$  is pairwise contagious relative to both  $a_j$  and  $a_k$ , it follows from the preceding discussion that  $a_l$  cannot co-exists with both  $a_j$  and  $a_k$ . Now, since  $a_l$  is GPC, and a polymorphic limit state  $L$  cannot contain  $a_l$ , it follows from the discussion in the preceding paragraphs that for  $n$  sufficiently large  $r_{\mathbf{a}_l}(L) < r_L(\mathbf{a}_l)$ . That is, if every  $a_j$  contained in  $L$  is PRC to  $a_l$ , then the minimum number of mutations needed to trigger a transition from  $L$  to  $\mathbf{a}_l$  is small and independent of  $n$ , but for the reverse transition from  $\mathbf{a}_l$  to  $L$ , it is an increasing function of  $n$ .

It follows from above discussion that if  $a_l$  is GPC, then for all  $L \neq \mathbf{a}_l$ , and  $n$  sufficiently large,  $r_{\mathbf{a}_l}(L) < r_L(\mathbf{a}_l)$ ; which in turn implies that  $r_{\mathbf{a}_l} < r_L$ , or equivalently  $\Phi_\beta(L, \mathbf{a}_l) < 0$ . Thus,  $\pi_\beta(\mathbf{a}_l) > \pi_\beta(L)$  for all  $L \neq \mathbf{a}_l$  whenever  $a_l$  is GPC.

#### Appendix B.6. Proof of Proposition 3

The following definitions are used in the next steps of the proof. Given transition probabilities  $P_\beta(\mathbf{x}, \mathbf{y})$ , define a cost function  $C(\mathbf{x}, \mathbf{y})$  as follows

$$\lim_{\beta \rightarrow \infty} \frac{-\ln P_\beta(\mathbf{x}, \mathbf{y})}{\beta} = C(\mathbf{x}, \mathbf{y}) \quad (\text{B.18})$$

The cost function  $C(\mathbf{x}, \mathbf{y})$  is equal to the number of experimentations that simultaneously occur in the transition  $\mathbf{x} \rightarrow \mathbf{y}$ . To see why, first notice that if  $a_j$  is not in the best response set of  $\mathbf{x}$  so that  $BR_i(a_j; \mathbf{x}) = 0$ , then

$$\lim_{\beta \rightarrow \infty} \frac{-\ln \mathbb{P}_i(a_j; \mathbf{x}_t)}{\beta} = \lim_{\beta \rightarrow \infty} \frac{-\ln \left[ \frac{1}{m} \exp(-\beta) + (1 - \exp(-\beta)) BR_i(a_j; \mathbf{x}) \right]}{\beta} = 1 \quad (\text{B.19})$$

So if transition  $\mathbf{x} \rightarrow \mathbf{y}$  involves  $r$  simultaneous mutations then

$$\lim_{\beta \rightarrow \infty} \frac{-\ln P_\beta(\mathbf{x}, \mathbf{y})}{\beta} = C(\mathbf{x}, \mathbf{y}) = r \quad (\text{B.20})$$

The next definition is related to  $\Gamma(W)$  graphs in Definition 6 on state space  $\mathbf{X}$ .

**Definition 7.** For any  $\mathbf{x} \in \bar{W}$  and  $\mathbf{y} \in W$  where  $\mathbf{x} \neq \mathbf{y}$ ,  $\Gamma_{\mathbf{x}, \mathbf{y}}(W)$  is a set of all  $\Gamma(W)$ -graphs which link  $\mathbf{x}$  to  $\mathbf{y}$ . For any two configurations  $\mathbf{x}, \mathbf{y} \in \bar{W}$ ,  $\Gamma_{\mathbf{x}, \mathbf{y}}(W \cup \{\mathbf{y}\})$  is the set of  $\Gamma(W)$ -graphs in which  $\mathbf{x}$  is joined to some point  $\mathbf{y}$  possibly itself and not to  $W$ , and that all other points of  $\bar{W}$  are joined to either the same point or to  $W$ .

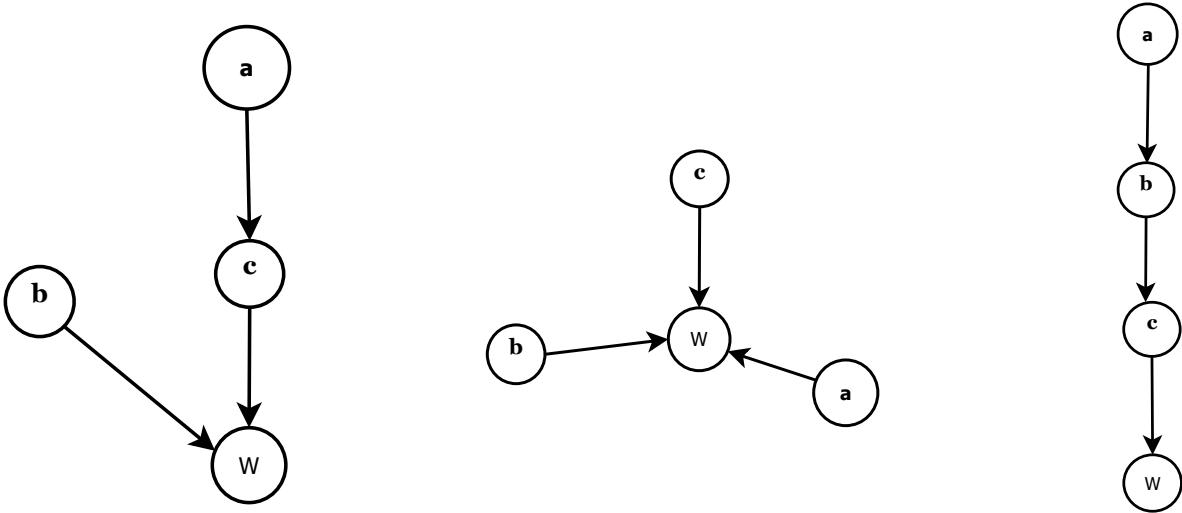


Figure B.7: Examples of  $\Gamma(W)$  graphs, where  $\bar{W} = \{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ .

Consider a configuration space  $\mathbf{X} = \{\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \mathbf{e}, \mathbf{f}, \mathbf{g}, \mathbf{h}\}$  with examples of  $\mathbf{g}$ -trees depicted in Figure B.6. Let  $W = \{\mathbf{d}, \mathbf{e}, \mathbf{f}, \mathbf{g}, \mathbf{h}\}$ , with examples of  $\Gamma(W)$  graphs depicted in Figure B.7. Then examples of  $\Gamma_{\mathbf{a}, \mathbf{c}}(W \cup \{\mathbf{c}\})$  graphs based on  $\Gamma(W)$  graphs in Figure B.7 are:  $\{\mathbf{a} \rightarrow \mathbf{c}, \mathbf{b} \rightarrow W\}$  for the graph on the left,  $\{\mathbf{c} \rightarrow W, \mathbf{b} \rightarrow W\}$  for the middle graph, and  $\{\mathbf{a} \rightarrow \mathbf{b}, \mathbf{c} \rightarrow W\}$  for the graph on the right.

Let  $C(g) = \sum_{(\mathbf{x}, \mathbf{y}) \in g} C(\mathbf{x}, \mathbf{y})$ . The following result is derived in Catoni (1999, Proposition 4.2).

**Lemma 7.** For any  $W \subset \mathbf{X}$ ,  $W \neq \emptyset$  and  $\bar{W} = \mathbf{X} \setminus W$ , for any  $\mathbf{x}, \mathbf{y} \in \bar{W}$

$$\lim_{\beta \rightarrow \infty} \frac{\ln \mathbb{E}[T(W) | \mathbf{x}_0 = \mathbf{x}]}{\beta} = \min_{g \in \Gamma(W)} C(g) - \min_{\mathbf{y} \in \bar{W}} \min_{g \in \Gamma_{\mathbf{x}, \mathbf{y}}(W \cup \{\mathbf{y}\})} C(g) \quad (\text{B.21})$$

Definition 6 directly extend to  $\Gamma$  graphs defined on state space  $\mathbf{L}$  (see Definition 1). Define  $r_*(L_j) = \min_{\mathbf{x} \in \mathbf{X}} r_{\mathbf{x}}(L_j)$ , to be the minimum number of experimentations required to exit  $D(L_j)$ , the BOA of  $L_j$ ,

also generally referred to the *radius* of  $D(L_j)$ . Clearly, provided  $\beta > \beta^*$ , if  $L_*$  is the limit configuration with the maximum long-run probability, then

$$r_{L_*} = \sum_{L_j \in \mathbf{L}} r_{L_*}(L_j) = \sum_{L_j \in \mathbf{L}} r_*(L_j)$$

Let  $W = \{L_*\}$ ; then  $\min_{g \in \Gamma(\{L_*\})} C(g) = r_{L_*}$ , and  $\min_{L_k \in \bar{W}} \min_{g \in \Gamma_{L_j, L_k}(L_* \cup \{L_k\})} C(g) = r_{L_*} - \max_{L_j \in \mathbf{L}} r_{\mathbf{x}}(L_j)$  for any  $L_l \in \bar{W}$ . Let  $r_* = \max_{L_j \in \mathbf{L}} r_*(L_j)$  and  $\mathbb{E}[T(L_*)] = \max_{\mathbf{x} \in \mathbf{X}} \mathbb{E}[T(L_*) \mid \mathbf{x}_0 = \mathbf{x}]$ . Then

$$\lim_{\beta \rightarrow \infty} \frac{\ln \mathbb{E}[T(L_*)]}{\beta} = r_* \quad (\text{B.22})$$

## Acknowledgements

We are thankful for the constructive comments from Alan Kirman, Robin Cowan, Vianney Dequiedt, Jean-Jacques Herings, Arkadi Predtetchinski, Co-Pierre Georg, François Lafond, Giorgio Triulzi and Stefania Innocenti. We also gratefully acknowledge the fruitful discussions with seminar participants at The Centre d’Economie de La Sorbonne and Paris School of Economics, University of Cergy Pontoise THEMA, Maastricht Lecture Series in Economics and Cournot seminars at Bureau d’Economie Théorique et Appliquée (BETA), University of Strasbourg. This work was supported by the FERDI (Fondation pour les Etudes et Recherches sur le Développement International) and the Agence Nationale de la Recherche of the French government through the program “Investissements d’avenir ANR-10-LABX-14-01.” The usual disclaimer applies.

## References

- Alós-Ferrer, C., Weidenholzer, S., 2007. Partial bandwagon effects and local interactions. *Games and Economic Behavior* 61 (2), 179–197.
- Alós-Ferrer, C., Weidenholzer, S., 2008. Contagion and efficiency. *Journal of Economic Theory* 143 (1), 251–274.
- Anderlini, L., Ianni, A., 1996. Path dependence and learning from neighbors. *Games and Economic Behavior* 13 (2), 141–177.
- Bass, F. M., 1969. A new product growth for model consumer durables. *Management science* 15 (5), 215–227.
- Beaman, L., BenYishay, A., Magruder, J., Mobarak, A. M., 2018. Can network theory-based targeting increase technology adoption? Tech. rep., National Bureau of Economic Research.
- Beaman, L., Dillon, A., 2018. Diffusion of agricultural information within social networks: Evidence on gender inequalities from mali. *Journal of Development Economics* 133, 147–161.

- Berninghaus, S. K., Schwalbe, U., 1996. Conventions, local interaction, and automata networks. *Journal of Evolutionary Economics* 6 (3), 297–312.
- Blume, L. E., 1995. The statistical mechanics of best-response strategy revision. *Games and Economic Behavior* 11 (2), 111–145.
- Campbell, A., 2013. Word-of-mouth communication and percolation in social networks. *American Economic Review* 103 (6), 2466–98.
- Catoni, O., 1999. Simulated annealing algorithms and markov chains with rare transitions. In: Azéma, J., Émery, M., Ledoux, M., Yor, M. (Eds.), *Séminaire de probabilités XXXIII*. Springer, Berlin Heidelberg, pp. 69–119.
- Chatterjee, K., Dutta, B., 2016. Credibility and strategic learning in networks. *International Economic Review* 57 (3), 759–786.
- Chen, W., Wang, Y., Yang, S., 2009. Efficient influence maximization in social networks. In: *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, pp. 199–208.
- Coleman, J., Katz, E., Menzel, H., of Applied Social Research, C. U. B., 1966. Medical innovation: a diffusion study. *Advanced Study in Sociology*. Bobbs-Merrill Co.  
URL <https://books.google.co.za/books?id=4e1EAAAAIAAJ>
- Conley, T. G., Udry, C. R., March 2010. Learning about a new technology: Pineapple in ghana. *American Economic Review* 100 (1), 35–69.  
URL <http://www.aeaweb.org/articles?id=10.1257/aer.100.1.35>
- Ellison, G., 1993. Learning, local interaction, and coordination. *Econometrica* 61 (5), 1047–1071.
- Ellison, G., 2000. Basins of attraction, long-run stochastic stability, and the speed of step-by-step evolution. *Review of Economic Studies* 67 (1), 17–45.
- Foster, A. D., Rosenzweig, M. R., 1995. Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of political Economy* 103 (6), 1176–1209.
- Foster, D., Young, P., 1990. Stochastic evolutionary game dynamics. *Theoret. Population Biol.*
- Freidlin, M., Wentzell, A. D., 1984. *Random perturbations of dynamical systems*. New York: Springer Verlag.
- Galeotti, A., Goyal, S., 2009. Influencing the influencers: a theory of strategic diffusion. *The RAND Journal of Economics* 40 (3), 509–532.

- Galeotti, A., Goyal, S., Jackson, M. O., Vega-Redondo, F., Yariv, L., 2010. Network games. *Review of Economic Studies* 77 (1), 218–244.
- Godes, D., Mayzlin, D., 2004. Using online conversations to study word-of-mouth communication. *Marketing science* 23 (4), 545–560.
- Goyal, S., Heidari, H., Kearns, M., 2014. Competitive contagion in networks. *Games and Economic Behavior*.
- Jackson, M. O., Yariv, L., 2007. Diffusion of behavior and equilibrium properties in network games. *The American economic review* 97 (2), 92–98.
- Kandori, M., Mailath, G. J., Rob, R., 1993. Learning, mutation, and long run equilibria in games. *Econometrica* 61 (1), 29–56.
- Kelly, J. A., St Lawrence, J. S., Diaz, Y. E., Stevenson, L. Y., Hauth, A. C., Brasfield, T. L., Kalichman, S. C., Smith, J. E., Andrew, M. E., 1991. Hiv risk behavior reduction following intervention with key opinion leaders of population: an experimental analysis. *American journal of public health* 81 (2), 168–171.
- Kempe, D., Kleinberg, J., Tardos, É., 2003. Maximizing the spread of influence through a social network. In: *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, pp. 137–146.
- Kirby, J., Marsden, P., 2006. *Connected marketing: the viral, buzz and word of mouth revolution*. Elsevier.
- Kiss, C., Bichler, M., 2008. Identification of influencers—measuring influence in customer networks. *Decision Support Systems* 46 (1), 233–253.
- Kreindler, G. E., Young, H. P., 2013. Fast convergence in evolutionary equilibrium selection. *Games and Economic Behavior* 80, 39 – 67.
- Kreindler, G. E., Young, H. P., 2014. Rapid innovation diffusion in social networks. *Proceedings of the National Academy of Sciences* 111 (Supplement 3), 10881–10888.
- Leduc, M. V., Jackson, M. O., Johari, R., 2017. Pricing and referrals in diffusion on networks. *Games and Economic Behavior* 104, 568–594.
- Lee, I. H., Szeidl, A., Valentinyi, A., 2003. Contagion and State Dependent Mutations. *The B.E. Journal of Theoretical Economics* 3 (1), 1–29.
- Lee, I. H., Valentinyi, A., 2000. Noisy contagion without mutation. *Review of Economic Studies* 67 (1), 47–56.

- Lelarge, M., 2012. Diffusion and cascading behavior in random networks. *Games and Economic Behavior* 75 (2), 752–775.
- López-Pintado, D., 2008. Diffusion in complex social networks. *Games and Economic Behavior* 62 (2), 573–590.
- Mahajan, V., Müller, E., Bass, F. M., 1990. New product diffusion models in marketing: A review and directions for research. *Journal of Marketing* 54, 1–26.
- Montanari, A., Saberi, A., 2010. The spread of innovations in social networks. *Proceedings of the National Academy of Sciences*.
- Morris, S., 2000. Contagion. *Review of Economic Studies* 67 (1), 57–78.
- Oyama, D., Sandholm, W. H., Tercieux, O., 2015. Sampling best response dynamics and deterministic equilibrium selection. *Theoretical Economics* 10 (1), 243–281.
- Reingen, P. H., Foster, B. L., Brown, J. J., Seidman, S. B., 1984. Brand congruence in interpersonal relations: A social network analysis. *Journal of Consumer Research* 11 (3), 771–783.
- Richardson, M., Domingos, P., 2002. Mining knowledge-sharing sites for viral marketing. In: *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, pp. 61–70.
- Robson, A. J., Vega-Redondo, F., 1996. Efficient Equilibrium Selection in Evolutionary Games with Random Matching. *Journal of Economic Theory* 70 (1), 65–92.
- Rogers, E. M., 2003. *Diffusion of innovations*. 5th Edition. Free Press.
- Sandholm, W. H., 2001. Almost global convergence to p-dominant equilibrium. *International Journal of Game Theory*, 107–116.
- Sundararajan, A., 2007. Local network effects and complex network structure. *The BE Journal of Theoretical Economics* 7 (1).
- Tsakas, N., 2017. Diffusion by imitation: The importance of targeting agents. *Journal of Economic Behavior & Organization* 139, 118–151.
- Valente, T. W., Hoffman, B. R., Ritt-Olson, A., Lichtman, K., Johnson, C. A., 2003. Effects of a social-network method for group assignment strategies on peer-led tobacco prevention programs in schools. *American journal of public health* 93 (11), 1837–1843.
- Young, H. P., 1993. The evolution of conventions. *Econometrica* 61 (1), 57–84.
- Young, P. H., 2011. The dynamics of social innovation. *Proceedings of the National Academy of Sciences* 108 (4), 21285–21291.