

Strategic diffusion in networks through contagion[☆]

Théophile T. Azomahou^{a,c}, Daniel C. Opolot^{b,*}

^aMaastricht University, School of Business and Economics, Tongersestraat 53, 6211 LM Maastricht

^bAIFMRM, University of Cape Town, Private Bag X3, Rondebosch, 7701 Cape Town

^cUniversity of Clermont Auvergne and CERDI, 65 Boulevard F. Mitterrand, 63009 Clermont-Ferrand CEDEX 1

Abstract

We consider an evolutionary game model on finite networks where individuals revise their strategies over time through best response. Focusing on coordination games, we establish conditions under which a strategy is contagious on a given network (i.e. when it can spread contagiously starting from a small group of players to the entire population, and its associated convention is never invaded by other conventions). We define a measure of *contagion threshold* for finite networks and show that p -dominant strategies are contagious on networks with a contagion threshold equal or greater than p . We also show that in evolutionary game models where players can make mistakes, there exists a threshold level of experimentation above which contagion is not relevant because it is dominated by noisy dynamics. Above this threshold, targetting individuals in order to trigger network-wide diffusion is economically unreasonable.

Keywords: Contagion, stochastic evolution, networks, experimentation

JEL: D8, C73

1. Introduction

Contagion – a phenomenon where a product/behaviour spreads over a population from a small group of adopters through local interactions – is the reason behind targetting.¹ Indeed, if a product can spread contagiously, then it may be less costly for a firm to target a small group of consumers – *initial adopters* – who then trigger its diffusion to the entire population, than to spend resources on traditional advertising. In complex diffusion processes – where a payoff from a given action depends on the proportion of others adopting it (e.g. coordination games) – the feasibility of contagion strictly depends on the interaction structure and assumptions regarding individual behaviour.

Interesting results have been derived relying on unbounded interaction structures: for example, a risk-dominant choice is contagious when individuals follow best-response behaviour (Morris, 2000; Oyama and Takahashi, 2015), and a Pareto-dominant choice is contagious when they imitate successful choices (Alós-Ferrer and Weidenholzer, 2008). In unbounded interaction structures, contagion can be

*Corresponding author:

Email address: opodanchris@gmail.com (Daniel C. Opolot)

¹See for example Galeotti and Goyal (2009), Campbell (2013), Goyal et al. (2014), Tsakas (2017) and Beaman et al. (2018).

defined without knowledge of the exact number of initial adopters – it is sufficient to show that this number is finite. However, in bounded interaction structures, the number of initial adopters relative to the population size determines when contagion occurs (i.e. contagion occurs if this ratio is very small). Focusing on unbounded interaction structures also makes it impossible to examine contagion in situation where individuals may make mistakes, which is crucial because it determines conditions under which targetting is economically reasonable.

This paper asks the following questions: in multiple strategy coordination games and bounded interaction structures, which strategies tend to be contagious and under what conditions? In situations where individuals occasionally make mistakes and choose non-optimal strategies, when does contagion dominate experimentation?

To address these questions, we consider two related models. **Model 1:** we consider an evolutionary model where players, interacting through an arbitrary finite network, simultaneously revise their strategies over time. The two main behavioural assumptions are *best response* and *myopia* – at each period, each player chooses a strategy that maximize the average payoff given the profile of strategies chosen by her neighbours in the previous period. **Model 2:** players follow the structural and behavioural assumptions of Model 1, but in addition, occasionally make mistakes and choose strategy that are not best responses. Mistakes may result from pure experimentation, where players try-out new strategies possibly due to lack of complete information about the game; or because they are not fully rational, and hence, occasionally choose strategies that are not best responses. We use Model 1 to establish conditions under which some strategy in a multiple-strategy coordination game is contagious on a given network; and Model 2 to determine when contagion dominates experimentation, in which case targetting is economically reasonable.

A strategy is formally said to be contagious on a given network if it satisfies two conditions: (i) *spreads contagiously* – it can spread contagiously to the entire network starting from a very small (relative to the population size) group of initial adopters; (ii) *uninvadability* – the number of errors needed to leave the convention where the entire network coordinates on a contagious strategy is higher than the number of initial adopters.

We define a measure of *contagion threshold* for finite networks that determines conditions for feasibility of contagion.² Specifically, given a multiple-strategy coordination game, let a be a p -dominant strategy. That is, a is a best response whenever at least proportion p of neighbours play a (Morris et al., 1995). We define the contagion threshold of a finite network as the maximum value of p such that contagion occurs in that network starting from a neighbourhood of any player. This definition implies

²Morris (2000) defines the *contagion threshold* for unbounded networks as follows. For a 2×2 coordination game with strategies a_j and a_l , where a_l is a best response when at least proportion ϕ_{jl} of neighbours play a_l , the contagion threshold of a given network is the supremum of ϕ_{jl} such that contagion occurs in that network. However, this definition works only for unbounded and uniform network where the exact knowledge of the number of initial adopters is not necessary.

that: (i) if a coordination game contains a p -dominant strategy, for all p equal or less than the contagion threshold of a given network, then it is contagious in that network; (ii) the contagion threshold is easily computable for any finite network. And most importantly, we can compute the set of initial adopters from the smallest neighbourhood of the network.

Using Model 2, we show that there is a threshold level of experimentation below which contagion dominates experimentation. This threshold level of experimentation (more precisely, the upper bound for the threshold level) increases concavely with the population size. Above this threshold, dynamics due to experimentation dominates best-response dynamics, making contagion less relevant. Under this scenario, targetting is economically unreasonable since targetted players will eventually switch to other strategies through experimentation. Otherwise, below the threshold level of experimentation, contagion dominates experimentation and targetting is economically reasonable.

This paper is closely related to the literature on complex contagion in networks (Morris, 2000; Alós-Ferrer and Weidenholzer, 2008; Oyama and Takahashi, 2015). Morris (2000) shows that in 2×2 coordination games, a risk-dominant strategy is contagious under best reply dynamics; Alós-Ferrer and Weidenholzer (2008) shows that a Pareto-dominant choice is contagious under imitation dynamics and some assumptions regarding information; Oyama and Takahashi (2015) shows that either choice (risk-dominant and Pareto-dominant) can be contagious under best reply dynamics in the presence of a third dominated alternative. Like Oyama and Takahashi (2015), we generalize the analysis in Morris (2000) to multiple strategy coordination games. However, Oyama and Takahashi (2015), focus on two types of networks – linear and non-linear networks – so that a strategy is contagious if it can spread contagiously in either of these two networks. Our analysis is richer in that we establish conditions under which a strategy is contagious on any given network. A more fundamental difference between our analysis and Morris (2000) and Oyama and Takahashi (2015) is that we examine contagion in finite networks and provide steps for computing the smallest number of initial adopters needed to trigger contagion.³

Our analysis and results from Model 2 are related to the literature of evolutionary game dynamics in networks (Ellison, 1993, 2000; Blume, 1995; Robson and Vega-Redondo, 1996; Anderlini and Ianni, 1996; Berninghaus and Schwalbe, 1996; Young, 1998; Lee and Valentinyi, 2000; Lee et al., 2003; Alós-Ferrer and Weidenholzer, 2007; Peski, 2010). Except for Peski (2010), these papers focus on identifying *stochastically stable states* – states that are played with a positive probability at the limit of noise – by considering specific network structures, such as ring and 2-dimensional grid structures. Peski (2010) finds that p -dominant strategies are stochastically stable in networks under some mild conditions. We show that p -dominant strategies are not only stochastically stable, they are played with the highest probability even when the noise level is sufficiently high.

³A related literature on diffusion in the presence of network externalities (López-Pintado, 2008; Jackson and Yariv, 2007; Sundararajan, 2007; Galeotti et al., 2010; Galeotti and Goyal, 2009). However, these papers study binary choice diffusion processes on random networks.

The remainder of the paper is organized as follows. In Section 2, we introduce Model 1 and Model 2. Section 3 derives our main results on contagion, showing when a p -dominant strategy is contagious on a given network. In Section 4 we establish conditions under which targetting is reasonable; that is, conditions under which contagion dominates experimentation. Section 5 discusses further implications of our results on convergence rates. All lengthy proofs are relegated to the Appendix.

2. The model

2.1. Actions, network and payoffs

We study evolutionary game dynamics on arbitrary networks focusing on strict coordination games. Let $A = \{a_1, \dots, a_j, \dots, a_m\}$ 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 has a *strict coordination property* if for each $a_j \in A$, $u(a_j, a_j) > u(a_j, a_l)$ for all $a_l \neq a_j$. Hence, each strategy is a pure strategy Nash equilibrium.

Let $N = \{1, \dots, i, \dots, n\}$ denote a set of agents who interact through a social network $G(N, E)$ (or simply G where no confusion arises), where E is the set of edges connecting players in N . Let G be an $n \times n$ interaction matrix G representing all connections in $G(N, E)$, where $G_{ik} = 1$ if a link exists from i to k , and zero otherwise. We place the following restrictions of G :

- (i) G is undirected and unweighted, that is, either $G_{ik} = G_{ki} = 1$, or $G_{ik} = G_{ki} = 0$.
- (ii) 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 non-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, σ_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 σ_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 σ_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 σ , where $\sigma = (\sigma(a_1), \dots, \sigma(a_m))$. We refer to the quadruple (A, U, N, G) as a *local interaction strict coordination game*.

2.2. Behaviour and dynamics

Given payoff and interaction structures, we consider an evolutionary process where players simultaneously revise their strategies over discrete times $t = 1, 2, \dots$. This evolutionary process suitably models 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. We are particularly interested in two related evolutionary models:

- (i) **Model 1:** evolution with best response;
- (ii) **Model 2:** evolution with best response and mutations (BRM).

In Model 1, players choose strategies that are best responses to the strategy profile of their opponents. In Model 2, players not only follow best response behaviour in Model 1, but also choose strategies that are not best replies with a positive probability. Model 2 can be justified as follows: (i) players are not fully rational, and hence, may occasionally make mistakes and choose strategies that are not best responses – this argument is akin to the notion of bounded rationality in game theory; (ii) players may prefer to experiment and try-out different strategies, possibly due to lack of complete information about the game. As customary in evolutionary models, we assume that players are *myopic* – they only react to neighbours' strategies chosen one period before.

To formalize these ideas, let small bold letters e.g. $\mathbf{x}, \mathbf{y}, \mathbf{z}, \dots$ denote vectors representing profiles (configurations) of strategies. 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 strategy 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 $\sigma_i(\mathbf{x}) = (\sigma_i(a_1; \mathbf{x}), \dots, \sigma_i(a_m; \mathbf{x}))$. 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 \mathbf{x}_t is:

$$U_i(a_j, \mathbf{x}_t) = \sum_{a_l \in A} \sigma_i(a_l; \mathbf{x}_t) u(a_j, a_l) \quad (2)$$

Let $BR(\mathbf{x}_t)$ be the set of strategies that are best responses to configuration \mathbf{x}_t , that is, $BR(\mathbf{x}_t) = \{a_j : U_i(a_j, \mathbf{x}_t) \geq U_i(a_l, \mathbf{x}_t), \forall a_l \neq a_j\}$. Associated with each $BR(\mathbf{x}_t)$ is $BR_i(a_j; \mathbf{x}_t)$, which 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 $a_j \notin BR(\mathbf{x}_t)$.

For Model 1, the probability that configuration \mathbf{x} is followed by configuration \mathbf{y} , denoted by $P(\mathbf{x}, \mathbf{y})$, is given by

$$P(\mathbf{x}, \mathbf{y}) = \prod_{i=1}^n BR_i(y^i; \mathbf{x}_t = \mathbf{x}) \quad (3)$$

where the product on the right hand side of (3) follows because players revise strategies simultaneously.

For Model 2, the probability, $\mathbb{P}_i(a_j; \mathbf{x}_t)$, that player i chooses a_j (in period $t+1$) given configuration \mathbf{x}_t , is given by

$$\mathbb{P}_i(a_j; \mathbf{x}_t) = \frac{1}{m} \exp(-\beta) + (1 - \exp(-\beta)) BR_i(a_j; \mathbf{x}_t) \quad (4)$$

The first term $\frac{1}{m} \exp(-\beta)$ captures individual experimentation (mutations) on choices and is independent of the current configuration \mathbf{x}_t . Specifically, a player follows a best response behaviour with probability $(1 - \exp(-\beta))$, and with probability $\exp(-\beta)$, randomly picks any strategy (i.e. with a uniform probability $\frac{1}{m}$).

A closer examination of (4) reveals that as β increases to infinity, players become more rational and choose best responses with higher probability; and as β tends to zero, agents' choices become more random. Practically, the exact value of β may depend on the environment and may also vary across players. For the former, our analysis examines the effects of varying the level of experimentation on equilibrium behaviour. However, for simplicity, we assume that the level of experimentation is identical across players.

Analogously to (3), the probability, $P_\beta(\mathbf{x}, \mathbf{y})$, that configuration \mathbf{x} is followed by configuration \mathbf{y} in Model 2 is given by

$$P_\beta(\mathbf{x}, \mathbf{y}) = \prod_{i=1}^n \mathbb{P}_i(y^i; \mathbf{x}_t = \mathbf{x}) \quad (5)$$

The dynamics in (3) and (5) both follow a *stationary Markov chain* on the configuration space \mathbf{X} . Let P and P_β denote the respective Markov transition matrices with $P(\mathbf{x}, \mathbf{y})$ and $P_\beta(\mathbf{x}, \mathbf{y})$ as typical elements of P and P_β . The quintuples (A, U, N, G, P) and (A, U, N, G, P_β) constitute complete descriptions of Model 1 and Model 2 respectively. Given these two models, we seek to:

- (i) use Model 1 to establish conditions under which contagion occurs on a given network;
- (ii) establish conditions under which targetting is relevant. That is, conditions under which the process of contagion dominates experimentation in equilibrium.

2.3. Equilibrium behaviour

The equilibrium behaviour of Model 1 is represented by its *limit (absorbing) sets*. A subset $L \subseteq \mathbf{X}$ of states (strategy configurations) is an *absorbing set* of a Markov chain if, once entered, is never exited. If an absorbing set is a singleton then it is called an *absorbing state*; that is, any state $\mathbf{x} \in \mathbf{X}$ for which $P(\mathbf{x}, \mathbf{x}) = 1$ is an absorbing state. Subsets of states of an absorbing set that are not absorbing states form an *absorbing cycle*. For example, a pair of states \mathbf{x} and \mathbf{y} form an absorbing cycle if $P(\mathbf{x}, \mathbf{y}) = 1$ and $P(\mathbf{y}, \mathbf{x}) = 1$.

Let \mathbf{L} be a set of all absorbing sets (i.e. all absorbing states and absorbing cycles) of P . The composition of \mathbf{L} depends on the payoff and network structures. Generally, it consists of all monomorphic strategy configurations (i.e. where all players coordinate on the same strategy) corresponding to all the m strategies, and depending on the network structure, it may also contain absorbing cycle and configurations where strategies co-exist. Associated with each $L \subset \mathbf{L}$ is the basin of attraction of L , $D(L)$, which is the set of all configurations from which the Markov chain converge to L .

A strategy that is contagious under Model 1 is associated with an absorbing state where all players coordinate on that strategy. The contagion analysis thus involves examining the stability of absorbing

sets – examining the costs, measured in terms of the number of mutations, of reaching and leaving an absorbing set.

For Model 2, due to experimentation, every strategy configuration is visited multiple times in the long-run. That is, unlike Model 1 where the Markov chain eventually settles in some absorbing set, in Model 2, there is a positive probability of reaching every state at any given time. Thus, a suitable measure of equilibrium behaviour in Model 2 is the *stationary distribution*, denoted by π_β , which 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 π_β is then an m^n -vector defined as $\pi_\beta = \lim_{t \rightarrow \infty} \mathbf{q}_0 P_\beta^t$.

Thus, for Model 2, we use the notion of *long-run equilibrium*, which is a set of strategy configurations with the highest long-run probability. That is, a subset $L^* \subset \mathbf{L}$ is the long-run equilibrium of P_β if $\pi_\beta(L^*) \geq \pi_\beta(L)$ for all $L \subseteq \mathbf{L} \setminus L^*$. Our definition of long-run equilibrium is more general than the notion of *stochastic stability* often used in evolutionary game models. A subset $L^* \subset \mathbf{L}$ is stochastically stable if $\lim_{\beta \rightarrow \infty} \pi_\beta(L^*) > 0$. Thus, there must exist some value of β above which long-equilibria are stochastically stable. We show that this threshold value of β also corresponds to the value above which contagion dominates experimentation.

3. Diffusion through contagion

We aim to generalize the notion of pairwise contagion defined by Morris (2000) and Alós-Ferrer and Weidenholzer (2008) for 2×2 coordination games and unbounded networks to multiple-strategy coordination games and finite networks. For 2×2 coordination games on infinite networks, a strategy a_i is pairwise contagious relative to another strategy a_j if: (i) starting from a strategy profile where all players play a_j , a_i can spread contagiously from a finite set of initial adopters; (ii) it is *uninvadable* – once established, it should not be possible to leave it with a finite group of deviants. Uninvadability in infinite networks also implies that the number of mutations required to leave a convention where all players coordinate on strategy a_j is infinite. This concept can be extended to multiple-strategy coordination games and finite networks through the following series of conceptual definitions.

Definition 1. *Given a local interaction strict coordination game (A, U, N, G) , a sequence of strategy profiles $\{\mathbf{x}_t\}_{t=0}^\infty$ is a best response sequence if: for all $t \geq 1$, there exists at least one $i \in N$ such that $x_t^i \neq x_{t-1}^i$; (ii) if $x_t^i \neq x_{t-1}^i$, then $x_t^i \in BR(\sigma_i(\mathbf{x}_{t-1}))$; and (iii) if $\lim_{t \rightarrow \infty} x_t^i = a_i$, then for all $T \geq 0$, $a_i \in BR(\sigma_i(\mathbf{x}_t))$ for some $t \geq T$.⁴*

⁴This definition is similar to Oyama and Takahashi (2015, Definition 1) but different in that we consider simultaneous

Property (i) of Definition 1 requires that at least one player must switch a strategy at each period – this follows because we consider a dynamic process with a simultaneous revision protocol. Property (ii) requires players to switch strategies through best response. And property (iii) requires strategies that are not best responses to be eventually abandoned.

Definition 2. Let the evolutionary process (A, U, N, P) on network G start from any state $\mathbf{y} \in \mathbf{L} \setminus \mathbf{a}_l$. A strategy a_l spreads contagiously from state \mathbf{y} if there exists a small subset $N(\mathbf{y} \rightarrow \mathbf{a}_l) \subset N$ that is independent of n , such that every best response sequence $\{\mathbf{x}_t\}_{t=0}^\infty$ with $\mathbf{x}_0 = \mathbf{y}$ and $x_1^i = a_l$ for all $i \in N(\mathbf{y} \rightarrow \mathbf{a}_l)$ satisfies $\lim_{t \rightarrow \infty} x_t^i = a_l$ for each $i \in N$.

For any $\mathbf{y} \in \mathbf{L} \setminus \mathbf{a}_l$ and \mathbf{a}_l , let $n(\mathbf{y} \rightarrow \mathbf{a}_l)$ be the cardinality of $N(\mathbf{y} \rightarrow \mathbf{a}_l)$. Then for infinite networks considered in Morris (2000) and Oyama and Takahashi (2015), it is sufficient to require $n(\mathbf{y} \rightarrow \mathbf{a}_l)$ to be finite. However, for finite networks we consider in Definition 2, an equivalent requirement is for $n(\mathbf{y} \rightarrow \mathbf{a}_l)$ to be independent of the population size n so that as n , and hence, G , grows, $n(\mathbf{y} \rightarrow \mathbf{a}_l)$ stays finite and small.

Definition 3. Let $r(\mathbf{a}_l)$ be the number of mutations required to leave convention \mathbf{a}_l , where all players coordinate on strategy a_l . Convention \mathbf{a}_l , and hence, strategy a_l , is uninvadable if $r(\mathbf{a}_l) > n(\mathbf{y} \rightarrow \mathbf{a}_l)$ for all $\mathbf{y} \in \mathbf{L} \setminus \mathbf{a}_l$ and $r(\mathbf{a}_l)$ is a function of n .

The second condition for uninvadability in Definition 3, (i.e. $r(\mathbf{a}_l)$ must be an increasing function of n) ensures that as n becomes infinitely large, $\lim_{n \rightarrow \infty} r(\mathbf{a}_l) = \infty$.

Definition 4. Given a strict coordination game (A, U) , strategy a_l is contagious in network $G(N, E)$ if it spreads contagiously in $G(N, E)$ and is uninvadable.

From Definition 4, we see that strategy a_l is contagious in a finite network if, for all $\mathbf{y} \in \mathbf{L} \setminus \mathbf{a}_l$, $n(\mathbf{y} \rightarrow \mathbf{a}_l)$ is independent of the population size n , and $r(\mathbf{a}_l)$ is an increasing function of n . As the population size becomes infinitely large, we then have an equivalent definition for contagion in infinite networks where $\lim_{n \rightarrow \infty} n(\mathbf{y} \rightarrow \mathbf{a}_l) < \infty$ and $\lim_{n \rightarrow \infty} r(\mathbf{a}_l) = \infty$.

We aim to identify the type of strategies that tend to be contagious and conditions under which they are contagious on a given network. The following example highlights how the payoff structure, the number of strategies, and the network structure affect the feasibility of contagion.

3.1. A motivating example

The effect of the network structure: Consider game u_1 of Figure 1a, where strategy a_2 is a best response to any distribution that places on it a mass of more than $\frac{2}{5}$ (when a_2 has a mass of $\frac{2}{5}$, a player is indifferent between a_1 and a_2); or equivalently, a_2 is a best response whenever it is adopted by more than proportion $\frac{2}{5}$ of a player's neighbours. Thus, for a player i with degree n_i , a_2 is a best response

best response dynamics.

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

(a) Game u_1 .

(b) Game u_2 .

Figure 1: A 2×2 strict coordination game. For game u_1 , strategy a_2 is a best response to any distribution that places on it a mass of more than $\frac{2}{5}$; and for game u_2 , strategy a_2 is a best response to any distribution that places on it a mass of more than $\frac{1}{4}$.

whenever it is played by $\lceil \frac{2}{5} \times n_i \rceil$ neighbours, where $\lceil x \rceil$ is the smallest integer that is greater than or equal to x .⁵

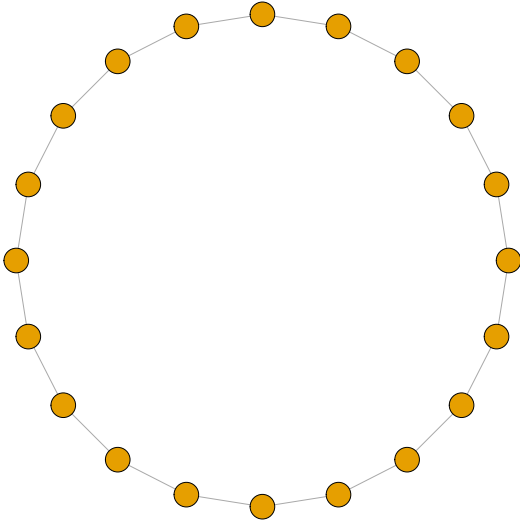
Let \mathbf{a}_1 and \mathbf{a}_2 be the respective conventions where all players coordinate on strategies a_1 and a_2 . Consider an evolutionary process P on network G_2 of Figure 2a where each player has two neighbours. If the process starts (i.e. at $t = 0$) from convention \mathbf{a}_1 , then two mutations by two consecutively players are sufficient to trigger evolution to convention \mathbf{a}_2 . This is because a_2 is a best response whenever it is played by at least $\lceil \frac{2}{5} \times 2 \rceil = 1$ neighbours. Starting from \mathbf{a}_1 , if two consecutively placed players (i.e. $N(\mathbf{a}_1 \rightarrow \mathbf{a}_2) = \{i, j\}$ so that $n(\mathbf{a}_1 \rightarrow \mathbf{a}_2) = 2$) switch to a_2 , each will have one neighbour playing a_2 and the rest playing a_1 , making a_2 a best response from $t = 1$ onward; at $t = 2$, each neighbour of $N(\mathbf{a}_1 \rightarrow \mathbf{a}_2)$ has one neighbour playing a_2 , making a_2 a best response from $t = 2$ onward; and so on until the entire network eventually switches to a_2 , and hence, a_2 spreads contagiously in this network.

Convention \mathbf{a}_2 is uninvadable in network G_2 because strategy a_1 is a best response only when it is played by $\lceil \frac{3}{5} \times 2 \rceil = 2$ neighbours, and hence, a minimum of $\frac{n}{2}$ mutations are required to leave convention \mathbf{a}_2 .

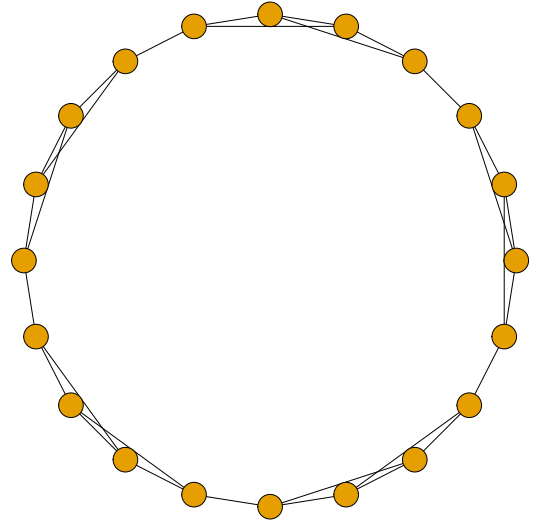
For network G_3 of Figure 2b where each player has three neighbours, a_2 is a best response only when it is played by at least $\lceil \frac{2}{5} \times 3 \rceil = 2$ neighbours. In this network, each player belongs to a *cohesive group* of four players where each player in the group has more than half (i.e. at least two) of her interactions within the group. Specifically, for each of these cohesive subgroup, denoted by N^c , each player $i \in N^c$ has at least proportion $\frac{2}{3}$ of her interactions within N^c and at most $\frac{1}{3}$ with players not in N^c . This network induces absorbing states in which a_1 and a_2 *co-exist*. That is, all states where all players in one or more N^c groups play a_2 and the rest play a_1 are absorbing.⁶ Thus, starting from convention \mathbf{a}_1 ,

⁵Note that we say a_2 is a best response whenever it is played by $\lceil \frac{2}{5} \times n_i \rceil$ neighbours because $\frac{2}{5} \times n_i$ is not an integer. If $\frac{2}{5} \times n_i$ were an integer, then a player would be indifferent between a_1 and a_2 whenever $\lceil \frac{2}{5} \times n_i \rceil = \frac{2}{5} \times n_i$ neighbours play a_2 ; under this scenario, a_2 becomes a best response only when it is played by $(\frac{2}{5} \times n_i) + 1$ neighbours.

⁶For example, a state where all players play a_1 , except for players in some N^c who play a_2 , is absorbing. This is because each $i \in N^c$ has at least two neighbours playing a_2 , making a_2 a best response; and each $j \notin N^c$ has at most one neighbour playing a_2 , making a_1 a best response.



(a) A regular cyclic network, G_2 , of degree 2.



(b) A regular cyclic network, G_3 , of degree 3.

Figure 2: Examples of regular cyclic networks

even if all four players in some N^c simultaneously mutate to a_2 , the evolutionary process will converge to absorbing state where a_1 and a_2 co-exist, and hence, a_2 does not spread contagiously in network G_3 .

However, strategy a_2 of game u_2 is contagious in network G_3 (and also network G_2). This is because for this game and network G_3 , strategy a_2 is a best response whenever it is played by $\lceil \frac{1}{4} \times 3 \rceil = 1$ neighbours. Thus starting from convention \mathbf{a}_1 one mutation to a_2 is sufficient to trigger evolution to convention \mathbf{a}_2 . Convention \mathbf{a}_2 is also uninvadable in network G_3 because strategy a_1 is a best response only when it is played by $\lceil \frac{3}{4} \times 3 \rceil = 3$ neighbours, and hence, a minimum of $\frac{3n}{4}$ mutations are required to leave convention \mathbf{a}_2 . Thus, strategy a_2 is contagious in network G_3 ; and it is easy to show that it is also contagious in network G_2 .

The intuition of this example, which we generalize below, is that the structure of network neighbourhoods may induce *intermediate absorbing states* in which strategies co-exist. In such a scenario, pairwise contagion is not feasible since the evolutionary process transitions between intermediate absorbing states before reaching a monomorphic convention.⁷ Networks with asymmetric neighbourhood structures tend to induce intermediate absorbing states while those with uniform neighbourhood structures do not. For example, network G_3 has an asymmetric neighbourhood structure consisting of cohesive subgroups with each player having no more than $\frac{1}{3}$ of her interactions with non-group members. The implication is that for any 2×2 coordination game with strategies a_j and a_l , if ϕ_{jl} is the proportion of a player's neighbours that must play a_l (and $1 - \phi_{jl}$ play a_j) above which a_l is a best response, then for all 2×2 coordination games where $\phi_{jl} > \frac{1}{3}$, an evolutionary process on network G_3 will consist of absorbing states in which a_j and a_l co-exist in different cohesive subgroups. Indeed, for game u_1 , $\phi_{12} = \frac{2}{5} > \frac{1}{3}$, and

⁷More specifically, there is no best response sequence leading from \mathbf{a}_1 to \mathbf{a}_2 , and hence, a_2 does not spread contagiously in network G .

	a_1	a_2	a_3
a_1	2 , 2	0 , 0	0 , 0
a_2	0 , 0	3 , 3	2 , 2
a_3	0 , 0	2 , 2	4.1 , 4.1

(a) Game u_3 .

	a_1	a_2	a_3
a_1	2 , 2	0 , 0	0 , 0
a_2	0 , 0	3 , 3	-9 , -9
a_3	0 , 0	-9 , -9	4.1 , 4.1

(b) Game u_4 .

Figure 3: A 3×3 strict coordination game. For game u_3 , strategy a_3 is a best response to any distribution that places on it a mass of more than $\frac{1}{3.05}$; and for game u_4 , there is no uniquely dominant strategy.

for game u_2 , $\phi_{12} = \frac{1}{4} < \frac{1}{3}$. However, for all 2×2 coordination games, no intermediate absorbing states can exist in network G_2 because $\phi_{jl} \leq \frac{1}{2}$, and each player in G_2 belongs to a cohesive subgroup of at least two players where each player within a subgroup has $\frac{1}{2}$ of her interactions with non-group members.

The effect of payoff structure and number of strategies: Consider the coordination games u_3 and u_4 in Figure 3 derived from game u_1 of Figure 1 by addition of strategy a_3 . We can think of game u_2 as the status quo (e.g. the payoff structure for two existing products in the market) and strategy a_3 as a new entrant (e.g. a new firm or an existing firm creating a new product). We then seek to establish the characteristics of strategy a_3 that ensure that it spreads contagiously and is uninvadable by the existing strategies.

Consider game u_3 . In addition to the relationship between a_1 and a_2 discussed in game u_1 above, we see that a_3 is a best response to any distribution that places on it a mass of more than $\frac{1}{3.05}$ and the rest on a_1 , and to any distribution that places on it a mass of more than $\frac{1}{5.1}$ and the rest on a_2 . Thus, strategy a_3 is a best response to any distribution that places on it a mass of more than $\frac{1}{3.05}$. Since $\frac{1}{3.05}$ is less than both $\frac{1}{2}$ and $\frac{1}{3}$, which are measures how symmetric neighbourhood structures of G_2 and G_3 respectively, there are no absorbing states where a_3 co-exists with either a_1 or a_2 . Thus, a_3 can spread contagiously from any $\mathbf{a} \in \mathbf{L} \setminus \mathbf{a}_3$ in both G_2 and G_3 .⁸ Convention \mathbf{a}_3 is uninvadable in both G_2 and G_3 since $r(\mathbf{a}_3) \geq \frac{n}{2}$ for G_2 and $r(\mathbf{a}_3) \geq \frac{3n}{4}$ for G_3 .

For game u_4 , no strategy uniquely dominates the rest. Specifically, strategy a_1 is a best response to all distributions that places on a_3 a mass between $\frac{3}{12}$ and $\frac{9}{13.1}$ and the rest on a_2 ; strategy a_3 is a best response to any distribution that places on it a mass of more than $\frac{1}{3.05}$ and the rest on a_1 ; The direct

⁸Specifically, strategy a_3 is a best response to any player whenever at least $\lceil \frac{1}{3.05} \times 2 \rceil = 1$ (and $\lceil \frac{1}{3.05} \times 3 \rceil = 1$ for network G_3) neighbour plays a_3 . For network G_2 (network G_3), starting from any strategy profile \mathbf{x} which does not contain a_3 , if two consecutively placed players (one player for G_3) mutate to a_3 at $t = 1$, then they continue to play a_3 from $t = 1$ onward; from $t = 2$ onward, all neighbours of the mutants play a_3 ; followed by neighbours of neighbours of mutants from $t = 3$ onward; and so on until the entire network converges to convention \mathbf{a}_3 .

implication is that, in both G_1 and G_3 , there are absorbing states where some two and all the three strategies co-exists.⁹ Thus, no strategy, or a group of strategies, is uniquely contagious.

The intuition of this example is that the number of strategies has no specific bearing on contagion but what matters is whether a game contains a strategy that uniquely dominates other strategies. That is, whether there exists a strategy that is a unique best response to any distribution that places on it a weight of at least $p \in (0, \frac{1}{2}]$. Such a strategy is said to be *p-dominant* (Morris et al., 1995).

We generalize the intuition of these two example to $n \times n$ coordination games and arbitrary networks that are strongly connected. Specifically, we aim to establish conditions under which *p-dominant* strategies are contagious in arbitrary networks.

3.2. The contagion of *p-dominant* strategies

Following Morris et al. (1995), a strategy is *p-dominant* if it is a best response to all distributions that place on it a mass of at least p . Formally, for a game U with strategy set A , strategy a_l is *p-dominant* if for all $a_k \in A$, and all $\sigma \in \Sigma$ with $\sigma(a_l) \geq p$,

$$\sum_{a_k \in A} \sigma(a_k) u(a_l, a_k) > \sum_{a_k \in A} \sigma(a_k) u(a_j, a_k) \quad \text{for all } a_j \neq a_l$$

In analogy to the local interaction strict coordination game, a strategy is *p-dominant* if it is a best response whenever it is played by at least proportion p , or a total of at least $\lceil pn_i \rceil$, of a player's neighbours.

Following the intuition from Section 3.1, we define a measure of how asymmetric the network neighbourhood sizes are. We take insights from Morris (2000) who defines the *contagion threshold* as a measure that captures neighbourhood structures in infinite networks. Specifically, for 2×2 coordination games with strategies a_j and a_l , where a_l is a best response when at least proportion ϕ_{jl} of neighbours play a_l , Morris (2000) defines the contagion threshold of a given network to be the supremum of ϕ_{jl} such that contagion occurs in that network. However, this definition does not directly carry-on to finite networks and multiple strategy coordination games. This is because, unlike infinite networks where contagion can be triggered from any finite group of players, for finite networks, it is necessary to compute the smallest set from which contagion can be triggered. We accordingly define the contagion threshold for finite networks as follows.

Let $S \subset N$ be a subgroup of players and $\bar{S} = N \setminus S$ its complement. For each $i \in N$, let $N_i(S) = N_i \cap S$ be the set of i neighbours belonging to subgroup S , and let $n_i(S)$ be the corresponding cardinality of $N_i(S)$; and $\alpha_i(S) = \frac{n_i(S)}{n_i}$ is the proportion of i 's neighbours in S . Write $N^q(S)$ for the players for whom at least proportion q of their interactions are with players in S , that is,

$$N^q(S) = \{i \in N : \alpha_i(S) \geq q\}$$

⁹For example, in network G_2 , a state where three adjacent players play a_2 , followed by one adjacent player playing a_1 , followed by three adjacent players playing a_3 , followed by one adjacent player playing a_1 , followed by three adjacent players playing a_2 , and so on, is an absorbing state.

For each $i \in N$, let B_{i_r} be the r th *neighbourhood* of i (i.e. the set of all players within radius r from i including i) and N_{i_r} be the r th-order neighbours of i (i.e. all players at radius r from i); we write b_{i_r} and n_{i_r} for the respective cardinalities of B_{i_r} and N_{i_r} . In analogy to the above definitions, $N_{i_r}^q(B_{i_{r-1}})$ is the set of players in N_{i_r} for whom at least proportion q of their interactions are with players in $B_{i_{r-1}}$. We then define and compute the contagion threshold of a finite network as follows.

Definition 5. *Given $G(N, E)$:*

- (i) *pick any $i \in N$ and the corresponding B_{i_1} ;*
- (ii) *for each $r \geq 2$ and $j \in N_{i_r}$, compute $\alpha_j(B_{i_{r-1}})$ and $\alpha_{i_r}^* = \min_{j \in N_{i_r}} \alpha_j(B_{i_{r-1}})$;*
- (ii) *given $i \in N$, compute $\alpha_i^* = \min_{r \geq 2} \alpha_{i_r}^*$.*

The contagion threshold of $G(N, E)$, $\eta(G)$, is given by $\eta(G) = \min_{i \in N} \alpha_i^$.*

The intuition behind our definition of contagion threshold is that for a coordination game with strategies a_j and a_l , $\eta(G)$ is the maximum ϕ_{jl} below which a_l is contagious in G .¹⁰ That is, starting from convention \mathbf{a}_j if all players within some neighbourhood of i mutate to a_l , then strategy a_l will spread contagiously in G if and only if $\phi_{jl} < \alpha_i^*$. Otherwise, if $\phi_{jl} > \alpha_i^*$, then it is possible to have absorbing states where a_j and a_l co-exist (i.e. absorbing states where players within B_{i_r} , for some $r \geq 2$, play a different strategy from players in $N \setminus B_{i_r}$), and hence, contagion is not guaranteed.

Definition 5 also provides steps for computing the smallest set of players from which contagion can be triggered. The size of this set corresponds to the number of mutations needed to trigger contagion. The steps in Definition 5 indicate that the smallest set that sufficiently triggers contagion can be computed by examining the neighbourhoods of players with the smallest $\min_{i \in N} \alpha_i^*$; that is, all $i \in \arg \max_{i \in N} \alpha_i^*$. We discuss the bounds for the smallest set below. The following proposition states the conditions for a p -dominant strategy to be contagious in a given network.

Theorem 1. *Given the diffusion process (A, U, N, P) on a strongly connected network $G(N, E)$, if $a^* \in A$ is a p -dominant strategy of U , then it is contagious in G if $p \leq \eta(G)$ and $b_2^* \leq n^{\frac{3}{5}}$, where $b_2^* = \min_{i \in N} b_{i_2}$. For all $\mathbf{a} \in \mathbf{L} \setminus \mathbf{a}^*$, $n(\mathbf{a} \rightarrow \mathbf{a}^*) \leq b_2^*$.*

Proof. See Appendix [Appendix B](#)

□

The proof of Theorem 1 follows in two steps. First, we show that if a^* is p -dominant and $p \leq \eta(G)$, then p spreads contagiously from any absorbing set after no more than b_2^* mutations. Specifically, letting r_i be the longest geodesic of i (i.e. the longest shortest distance from i to any other player), we show

¹⁰This is consistent with the definition of contagion threshold for infinite networks in [Morris \(2000\)](#). However, this definition is more suited for finite networks since it involves iterative computational steps, which cannot be executed in an infinite networks.

that starting from any $\mathbf{a} \in \mathbf{L} \setminus \mathbf{a}^*$, there exists a best response sequence $\{\mathbf{x}_t\}_{t=0}^{r^*-1}$ with $\mathbf{x}_0 = \mathbf{a}$ and $x_1^j = a^*$ for all $j \in B_2^*$ and $\lim_{t \rightarrow r^*-1} x_t^j = a^*$ for each $j \in N$, where r^* and B_2^* are the geodesic and B_{i_2} of i for whom $b_{i_2} = b_2^*$.

This follows from the definition of contagion threshold whereby, for any $i \in N$ and corresponding B_{i_1} , each $j \in N_{i_r}$ has $\alpha_j(B_{i_{r-1}}) \geq \eta(G)$ for all $r \geq 2$. Thus, for $p \leq \eta(G)$ and starting from some $\mathbf{a} \in \mathbf{L} \setminus \mathbf{a}^*$, if all players in B_{i_2} mutate to a^* at $t = 1$, then at $t = 2$, a^* is a best response for all B_{i_2} and each $j \in N_{i_3}$ since $\alpha_j(B_{i_1}) \geq \eta(G) \geq p$ (i.e. each $j \in N_{i_2}$ has at least proportion p of their interactions with players in B_{i_2} , where p dominance implies that a^* is a best response at least proportion p of neighbours play a^*); at $t = 3$, a^* is a best response for all B_{i_3} and each $j \in N_{i_4}$ since $\alpha_j(B_{i_3}) \geq \eta(G) \geq p$; and so on, until $t = r_i - 1$ when the entire network eventually plays a^* . The smallest number of mutations for a^* to spread contagiously from any absorbing set is then bounded above by b_2^* .

Note that b_2^* is an upper bound for the minimum number of mutations needed for a^* to spread contagiously on a given network. The exact number of mutations can be much less than b_2^* and can be computed by examining the dynamics within B_2^* . Specifically, the minimum number of mutations required are computed in two steps:

- (i) identify B_2^* by computing B_{i_2} for each $i \in N$ and then picking the one with the smallest cardinality;
- (ii) within B_2^* , identify the smallest set of players that should play a^* so that all players in B_2^* eventually switch to play a^* .

The second step of the proof of Theorem 1 shows that if a^* is p -dominant and $p \leq \eta(G)$, then the number of mutations needed to leave (the basin of attraction of) convention \mathbf{a}^* is greater than $n^{\frac{3}{5}}$. The intuition behind this result is that since each $j \in N_{i_r}$ has $\alpha_j(B_{i_{r-1}}) \geq \eta(G) \geq p$ for all $r \geq 2$, and that $p \leq \frac{1}{2}$, players within a given B_{i_2} will switch to a strategy different from a^* if each has at least proportion $(1 - p) > \frac{1}{2}$ of neighbours playing a strategy different from a^* . Thus, to leave convention \mathbf{a}^* , the set of mutants, $R(\mathbf{a}^*)$, must be selected in such a way that, for each $i \in N$, each $j \in B_{i_2}$ has more than half of their neighbours in $R(\mathbf{a}^*)$. This implies that the identification of $R(\mathbf{a}^*)$ is equivalent to the graph theory problem of identifying monopolies (i.e. sets of vertices of a graph containing at least half of the direct and/or indirect interactions of every player). Using well established results in graph theory (e.g. [Bermond et al. \(1996, Proposition 4\)](#)), it follows that the cardinality of $R(\mathbf{a}^*)$ is greater than $n^{\frac{3}{5}}$.

Since the number of mutations needed to leave convention \mathbf{a}^* is a function of n , it is unavoidable provided $b_2^* \leq n^{\frac{3}{5}}$. That is, provided the number of mutations needed to trigger contagion from any other absorbing set to convention \mathbf{a}^* , which is less or equal to b_2^* , is less than the number needed to leave \mathbf{a}^* .

Example: Consider the network structures in Figure 4. Network G_a of Figure 4a has contagion threshold

of $\eta(G_a) = \frac{2}{7}$ since for each i and corresponding B_{i_1} , each $j \in N_{i_r}$ for all $r \geq 2$ has $\alpha_j(B_{i_{r-1}}) \geq \frac{2}{7}$.¹¹ According to Theorem 1, a p -dominant strategy a^* spreads contagiously in networks G_a if $p \leq \frac{2}{7}$. To identify the minimum number of mutations needed to trigger contagion in this network, we examine the dynamics within each B_2^* . Network G_a consists of six B_2^* and we pick the one centred around player 1, $B_2^* = B_{1_2} = \{1, 2, 7, 8, 15, 16, 17, 18\}$.

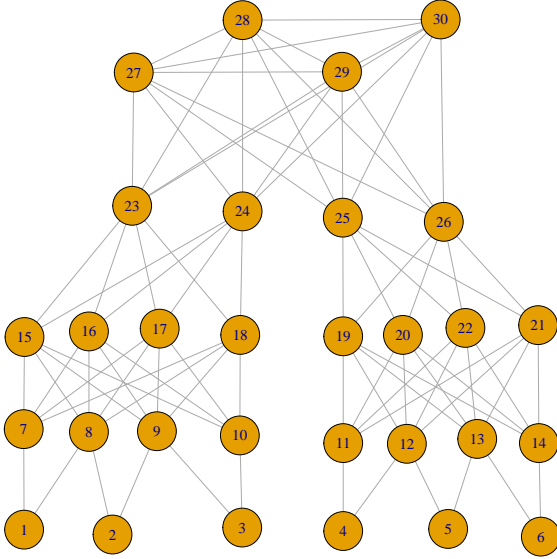
The minimum number of mutations needed to trigger contagion in G_a is then the number of mutations that ensure that all players within B_{1_2} switch to a^* . Consider the case where $p = \frac{2}{7}$ so that a^* is a best response for players $\{1, 2\}$ when $\lceil \frac{2}{7} \times 2 \rceil = 1$ neighbours play a^* ; for player 7 when $\lceil \frac{2}{7} \times 5 \rceil = 3$ neighbours play a^* ; and for players $\{8, 15, 16, 17, 18\}$ when $\lceil \frac{2}{7} \times 6 \rceil = 2$ play a^* . Hence, all players within B_{1_2} will switch to a^* if players $\{1, 2, 7, 8\}$ simultaneously switch to a^* ; any other combination of less four players within B_{1_2} may lead to an absorbing cycle containing a^* and other strategies but not convention \mathbf{a}^* .

For network G_b of Figure 4b, the contagion threshold is $\eta(G_b) = \frac{1}{6}$. Pick B_2^* centred around player 1, that is, $B_2^* = B_{1_2} = \{1, 7, 13, 14\}$. For each of these players a^* is a best response whenever at least one neighbour plays a^* . Thus, all players in B_{1_2} will switch to a^* within three steps of iteration once one player, player $\{7\}$, switches to a^* .

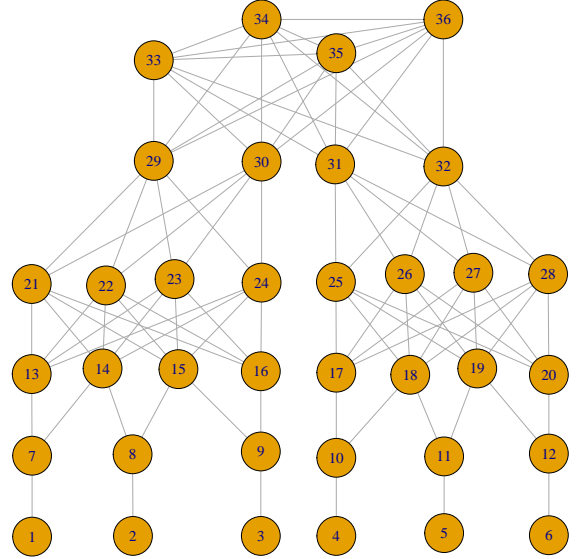
For both networks, \mathbf{a}^* is uninvadable since the cardinality of B_2^* is less $n^{\frac{3}{5}}$, which is more than 7 players for G_a and more than 8 for G_b . More specifically, for any $p \leq \frac{1}{2}$, $R(\mathbf{a}^*)$ must at the least consist of $R(\mathbf{a}^*) = \{7, 8, 9, 10, 11, 12, 13, 14, 23, 24, 25, 26\}$ for network G_a and $R(\mathbf{a}^*) = \{13, 14, 15, 16, 17, 18, 19, 20, 29, 30, 31, 32\}$ for network G_b , which is a minimum of 12 players.

The above example helps to highlight two interactive concepts in our results, which a firm/planner aiming to diffuse a product/behaviour must consider: the contagion threshold versus the number of mutations (i.e. initial adopters) needed to trigger contagion. Consider a new firm contemplating to enter a market (or an existing firm aiming to introduce a new product) where two or more products already exist. To diffuse her product through contagion, the firm would first determine the contagion threshold of the network, which in turn determines the value of p – the extent to which a new product should dominate the existing alternative products. If the contagion threshold of the network is very small, then the new product must be highly beneficial compared to all other products so that p is smaller than the contagion threshold. A firm must thus incur large initial costs to make the product more beneficial. The upside is that a very small number of initial adopters can trigger contagion of the product to the entire network, and hence, a firm need not incur large costs on targetting. If on the other hand, the contagion threshold is large, then p need not be very small, and hence, less initial costs are incurred. However, this also implies that more resources will be spent on targetting. A firm's objective

¹¹For example, if we pick player 1 and corresponding $B_{1_1} = \{7, 8\}$, we see that each $j \in N_{1_2} = \{2, 15, 16, 17, 18\}$ has $\alpha_j(B_{1_1}) \geq \frac{1}{3}$; each $j \in N_{1_3} = \{9, 10, 23, 24\}$ has $\alpha_j(B_{1_2}) \geq \frac{1}{2}$; each $j \in N_{1_4} = \{3, 27, 28, 29, 30\}$ has $\alpha_j(B_{1_3}) \geq \frac{2}{7}$; each $j \in N_{1_5} = \{25, 26\}$ has $\alpha_j(B_{1_4}) = \frac{1}{2}$; each $j \in N_{1_6} = \{19, 29, 21, 22\}$ has $\alpha_j(B_{1_5}) \geq \frac{1}{3}$; each $j \in N_{1_7} = \{11, 12, 13, 14\}$ has $\alpha_j(B_{1_6}) \geq \frac{2}{3}$; and each $j \in N_{1_8} = \{4, 5, 6, 7\}$ has $\alpha_j(B_{1_7}) = 1$.



(a) Network G_a with contagion threshold of $\eta(G_a) = \frac{2}{7}$.



(b) Network G_b with contagion threshold of $\eta(G_b) = \frac{1}{6}$.

Figure 4: Network contagion thresholds.

is thus to evaluate these two types of costs.

Overall, low-density network (i.e. networks with sparse connection) have lower contagion thresholds. This is because the smallest possible value for the contagion threshold of any network G is $\frac{1}{\bar{n}(G)}$, where \bar{n} is the maximum neighbourhood size of G . Moreover, low-density networks also have smaller B_2^* , and hence, a small number of initial adopters needed to trigger contagion. This implies that it is relatively less costly to diffuse new products and behaviour to low-density networks. One possible evidence of this implication of our results is in the observed discrepancy between the dynamics in the market for messaging and chat apps (e.g. AOL Instant Messenger, Google talk, VOIP, skype, Kik., whatsapp, snapchat, HipChat and slack), where there is a high turnover and entry rate, versus social networking apps (e.g. Friendster, Myspace and Facebook), where turnover and entry rate is relatively low. Specifically, messaging and chat apps consist of low-density 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 much 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.

Like [Oyama and Takahashi \(2015\)](#), our results extend [Morris \(2000\)](#) who examines contagion for 2×2 coordination games to multiple strategy coordination games. However, in [Oyama and Takahashi \(2015\)](#), a strategy is contagious if it is contagious in some unbounded network. As such, it is sufficient to check whether a strategy is spreads contagiously in any of two types of networks, linear and non-linear networks. Our analysis is richer in that we establish conditions under which a strategy (i.e. a p -dominant strategy) is contagious on any given network. In doing so, we also provide steps for computing the smallest number

of initial adopters needed to trigger contagion.

A more fundamental difference between our analysis and [Morris \(2000\)](#) and [Oyama and Takahashi \(2015\)](#) is that we examine contagion in finite networks. In particular, [Morris \(2000\)](#) provides bounds for contagion threshold in unbounded networks but computing the exact contagion threshold of an unbounded network is not a straightforward matter. Ideally, one would select some finite group of players in some region of the network and then iterate over the entire network following similar steps in Definition 5 above. There problem with using this method is that when the network is unbounded, there is no obvious means of knowing how many iteration are sufficient for computing the contagion threshold. One would have to assume some form of uniformity across the unbounded network, and hence, the contagion threshold can feasibly be determined for only unbounded regular networks (i.e. networks where all players have the same neighbourhood size).

4. When is it reasonable to target agents?

Since the cost of targetting specific players to aid wider diffusion of a strategy is directly related to contagion, targetting is economically reasonable only if contagion can actually occur in a given network. Contagion is a result of best response dynamics, targetting is reasonable when best response dynamics dominates experimentation. Given Model 2, we show that there exists a level of experimental below which best response dynamics, and hence, contagion, dominates experimentation. Put differently, we show that there exists a threshold value, β^* , above which the convention corresponding to the contagious strategy is the long-run equilibrium of Model 2. Below β^* , players' choices are too noisy and targetting is not economically reasonable. The following definition and notations are used to state the main results of this section.

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 5 presents two examples of \mathbf{g} -trees.

From the definition of $\Gamma(\mathbf{x})$ graphs, every $g \in \Gamma(\mathbf{x})$ spans the entire state space, except for \mathbf{x} . Each $\mathbf{y} \in \mathbf{X}$ has only one arrow emanating from it so that each $g \in \Gamma(\mathbf{x})$ has a total of $m^n - 1$ directed edges, where m^n is the cardinality of \mathbf{X} . Thus, if we let $\gamma(\mathbf{x}) = \#\Gamma(\mathbf{x})$ be the cardinality of $\Gamma(\mathbf{x})$, then $\gamma(\mathbf{x})$ is identical for any pair of configurations $\mathbf{x}, \mathbf{y} \in \mathbf{X}$; that is, $\gamma(\mathbf{x}) = \gamma(\mathbf{y}) = \gamma$.

The cardinality of $\Gamma(\mathbf{x})$, γ , is a multiple of, and not an exponential function of the cardinality of the state space, m^n . For example, when the size of the state space is 2, $\gamma = 1$; when it is 3, $\gamma = 3$; when it

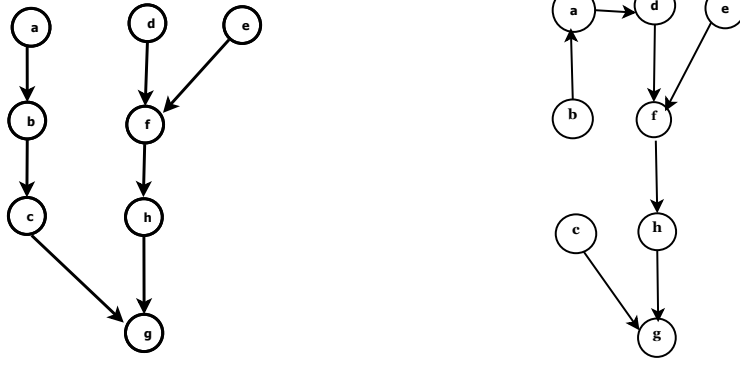


Figure 5: Examples of \mathbf{g} -trees, that is $\Gamma(\{\mathbf{g}\})$ graphs and in which the configuration space $\mathbf{X} = \{\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \mathbf{e}, \mathbf{f}, \mathbf{g}, \mathbf{h}\}$.

is 4, $\gamma = 15$,¹² and when it is 5, $\gamma = 51$. Thus, $\ln(\gamma)$ is a linear function of n .

Proposition 1. *Given a diffusion model (A, U, N, G, P_β) , if \mathbf{a}^* is contagious on network G , then there exists some $\beta^* \in (0, \ln \gamma / (n^{\frac{3}{5}} - b_2^*)]$ such that for all $\beta \geq \beta^*$, $\pi_\beta(\mathbf{a}^*) > \pi_\beta(\mathbf{a})$ for all $\mathbf{a} \in \mathbf{L} \setminus \mathbf{a}^*$.*

Proof. See Appendix [Appendix C](#)

□

The proof of Proposition 1 follows in two steps. First, we characterise the structure of stationary distributions using a graph theoretic method of [Freidlin and Wentzell \(1984\)](#). The graph theoretic method of [Freidlin and Wentzell \(1984\)](#) represents the stationary distribution of any configuration \mathbf{x} in terms of the probabilities of $\Gamma(\mathbf{x})$ graphs. Specifically, let $c(\mathbf{y}, \mathbf{z})$ be the number of mutations involve in the direct transition from \mathbf{y} to \mathbf{z} . The total cost $c(\mathbf{x}; g)$ of some $g \in \Gamma(\mathbf{x})$ is the sum of costs of all transitions in g . That is, $c(\mathbf{x}; g) = \sum_{(\mathbf{y}, \mathbf{z}) \in g} c(\mathbf{y}, \mathbf{z})$.

We in turn show that there exists some $\beta^* \in (0, \ln \gamma / (n^{\frac{3}{5}} - b_2^*)]$ such that for all $\beta \geq \beta^*$, the long-run equilibrium is the strategy configuration with the minimal cost $\Gamma(\mathbf{x})$ graph: that is, for all $\beta \geq \beta^*$, $\mathbf{x} \in \arg \min_{\mathbf{x} \in \mathbf{X}, g \in \Gamma(\mathbf{x})} c(\mathbf{x}; g)$ is the long-run equilibrium. We then show that if \mathbf{a}^* is contagious on a given network, then convention \mathbf{a}^* has the minimal cost graph.

The intuition behind Proposition 1 is that when the level of experimentation is very high (i.e. β is small and between 0 and β^*), then the evolutionary dynamics is dominated by noise and not best response. This effect is observable from the specification of Model 2 in (4), whereby, when β is less than one and small, the noise term dominates the best response term. On a larger scale, the upper bound for the threshold value of β depends on the population size. That is, β^* is bounded from above by $\ln \gamma / (n^{\frac{3}{5}} - b_2^*)$, which is a concave function of n . Specifically, $\ln \gamma / (n^{\frac{3}{5}} - b_2^*)$ is of order $(O)(n^{2/5})$ because $\ln \gamma$ is a linear function of n .

¹²That is, if the state space is $\mathbf{X} = \{\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}\}$, then the list of $\Gamma(\mathbf{d})$ graphs is: $\{\mathbf{c} \rightarrow \mathbf{b} \rightarrow \mathbf{a} \rightarrow \mathbf{d}\}, \{\mathbf{a} \rightarrow \mathbf{c} \rightarrow \mathbf{b} \rightarrow \mathbf{d}\}, \{\mathbf{a} \rightarrow \mathbf{b} \rightarrow \mathbf{c} \rightarrow \mathbf{d}\}, \{\mathbf{b} \rightarrow \mathbf{a} \rightarrow \mathbf{c} \rightarrow \mathbf{d}\}, \{\mathbf{b} \rightarrow \mathbf{c} \rightarrow \mathbf{a} \rightarrow \mathbf{d}\}, \{\mathbf{c} \rightarrow \mathbf{a} \rightarrow \mathbf{d}, \mathbf{b} \rightarrow \mathbf{d}\}, \{\mathbf{a} \rightarrow \mathbf{c} \rightarrow \mathbf{d}, \mathbf{b} \rightarrow \mathbf{d}\}, \{\mathbf{c} \rightarrow \mathbf{b} \rightarrow \mathbf{d}, \mathbf{a} \rightarrow \mathbf{d}\}, \{\mathbf{b} \rightarrow \mathbf{c} \rightarrow \mathbf{d}, \mathbf{a} \rightarrow \mathbf{d}\}, \{\mathbf{a} \rightarrow \mathbf{b} \rightarrow \mathbf{d}, \mathbf{c} \rightarrow \mathbf{d}\}, \{\mathbf{a} \rightarrow \mathbf{b} \rightarrow \mathbf{d}, \mathbf{c} \rightarrow \mathbf{d}\}, \{\mathbf{a} \rightarrow \mathbf{c}, \mathbf{b} \rightarrow \mathbf{c}, \mathbf{c} \rightarrow \mathbf{d}\}, \{\mathbf{c} \rightarrow \mathbf{a}, \mathbf{b} \rightarrow \mathbf{a}, \mathbf{a} \rightarrow \mathbf{d}\}, \{\mathbf{a} \rightarrow \mathbf{b}, \mathbf{c} \rightarrow \mathbf{b}, \mathbf{b} \rightarrow \mathbf{d}\}, \{\mathbf{c} \rightarrow \mathbf{d}, \mathbf{a} \rightarrow \mathbf{d}, \mathbf{b} \rightarrow \mathbf{d}\}$

The upper bound for β^* increases with n , although mildly, because as n increases, the size of the state space increase exponentially. This in turn implies that for a given level of experimentation, the likelihood that the evolutionary process follows paths involving only best response configurations reduces (since the number of configurations involving at least one mutation increases with the state space). Thus, as n , and hence, the state space, increases, players should experiment less if best response dynamics, and hence, contagion, is to dominates experimentation.

Proposition 1 provides conditions under which targetting is economically reasonable: that is, only when $\beta \geq \beta^*$. More broadly, Proposition 1 suggests that the predictions in evolutionary game theory that employ stochastic stability as a concept are admissible for relatively high levels of experimentation. Recall that a configuration \mathbf{x} is stochastically stable if $\lim_{\beta \rightarrow \infty} \pi_\beta(\mathbf{x}) > 0$ (Foster and Young, 1990; Kandori et al., 1993; Young, 1993; Ellison, 2000). Proposition 1 states that the level of experimentation need not be very small (i.e. β need not be asymptotically large); it is sufficient that $\beta \geq \beta^*$. And in evolutionary models with finite and a small population, β^* can be admissibly large.

5. Discussion and concluding remarks

We examined the diffusion of products and practices with coordination effects 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; myopia – the inability to remember the entire history of play in complex social interactions; and the locality of social interactions. We consider two related models that capture these structural and behavioural properties. We are particularly interested in the phenomenon of contagion through best response: the conditions under which contagion can occur in finite networks and when contagion dynamics dominates experimentation in settings where agents make mistakes.

We show that if a p -dominant strategy is present in a coordination game, then that strategy is contagious in networks with contagion threshold equal or greater than p . Our measure of contagion is easily computable. We also provide bounds and steps for computing the smallest set of players that can trigger contagion. We then examine the effects of noise on contagion, showing that there exists a threshold level of experimentation below which contagion is dominated experimentation.

Our results have broader implications for targetting individuals in the network to trigger diffusion of products and behaviour. First, we provide algorithmic steps for identifying the smallest set of players that sufficiently trigger network-wide diffusion. We find that this set can indeed be very small and independent of the population size. Second, our results imply that targetting maybe economically undesirable in settings where the extent to which players experiment is high. High levels of individual experimentation may occur in settings where individuals do not have complete information about the rewards from different choices.

Finally, we remark that our results have implications for convergence rates of evolutionary dynamics

in networks. One of the major criticism directed toward evolutionary game models is that the convergence rates to equilibrium tend to increase with the population size, so that in very large populations, equilibrium is not reached in timescales of economic relevance. In the analysis that we have relegated to Section [Appendix A](#), we show that if a contagious strategy exist, then equilibrium is reached fast.

Appendix A. 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 β 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 β is sufficiently small. Here, we examine the case where β is very large (i.e. $\beta \rightarrow \infty$) so that β'_m as defined above, is negligibly small compared to β_m . The expected waiting time is formally defined as follows.

Definition 7. 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 $\mathbb{E}[T(W) \mid \mathbf{x}_0 = \mathbf{x}]$.

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

Proposition 2. For a diffusion process (A, U, N, G, P_β) , if \mathbf{a}^* is the configuration with the maximum long-run probability, then there exists some $r^* \in (0, b_2^*)$ such that

$$\lim_{\beta \rightarrow \infty} \frac{\ln \mathbb{E}[T(\mathbf{a}^*)]}{\beta} = r_* \quad (\text{A.1})$$

Proof. See [Appendix D](#)

□

Proposition 2 shows that the expected waiting time to reach the configuration with maximum long-run probability from any other configuration takes the form

$$\mathbb{E}[T(\mathbf{a}^*)] \leq \exp[\beta r_* + f(m, n, \beta)] \quad (\text{A.2})$$

The function $f(m, n, \beta)$ increases with m and n and decreases with β . For $\beta \rightarrow \infty$, $f(m, n, \beta) \rightarrow 0$.

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 2 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 2×2 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 2 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 2 and the discussion in Section 3 provides insights into the findings in Young (2011). We showed in Section 3 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_* .

Appendix B. Proof of Theorem 1

The proof of Theorem 1 follows in two steps. First we show that if $a^* \in A$ is a p -dominant strategy of U and $p \leq \eta(G)$, then a^* spreads contagiously on G starting from a subset smaller or equal to b_2^* . Second, if $a^* \in A$ is a p -dominant strategy and $p \leq \eta(G)$, then the number of mutations needed to leave convention \mathbf{a}^* where all players play a^* is more than $n^{\frac{3}{5}}$.

Strategy a^* spreads contagiously on G : Given the diffusion process (A, U, N, P) on a strongly connected network $G(N, E)$ and a p -dominant strategy $a^* \in A$, let $p \leq \eta(G)$. Let also $A_1 = A \setminus a^*$ and $A_2 = \{a^*\}$. Since a^* is p -dominant, it is a best response whenever at least proportion p of neighbours play a^* and the rest play strategies in A_1 . This implies that if all players in a given subgroup Z play a^* , then a^* is a unique best response to any $i \in N$ for whom $\alpha_i(Z) \geq \alpha_i^* \geq p$.

Let the evolutionary process start from any state $\mathbf{a} \in \mathbf{L} \setminus \mathbf{a}^*$, where \mathbf{a} can consists of only strategies in A_1 or both a^* and strategies in A_1 . Pick any $i \in N$ and the respective B_{i_2} and let all players in B_{i_2} mutate to play a^* and the rest stick to strategies in A_1 . The evolutionary process will evolve through best response from $t = 1$ onward as follows, where we write $\bar{B}_{i_r} = N \setminus B_{i_r}$ for the complement of B_{i_r} , $i \rightarrow A_l$ to mean that i plays a strategy in A_l , $Z \rightarrow A_l$ to mean that each $j \in Z$ plays a strategy in A_l :

$t = 1$	All players in B_{i_2} play a^* ; all players in \bar{B}_{i_2} play strategies in $A_1 \cup a^*$.
$t = 2$	All players in B_{i_2} play a^* since each $j \in B_{i_2}$ has $\alpha_j(B_{i_2}) \geq \alpha_i^* \geq p$; all players in N_{i_3} play a^* since each has $\alpha_j(B_{i_2}) \geq \alpha_i^* \geq p$; all players in \bar{B}_{i_3} play strategies in $A_1 \cup a^*$.
$t = 3$	All players in B_{i_3} play a^* since each $j \in B_{i_3}$ has $\alpha_j(B_{i_3}) \geq \alpha_i^* \geq p$; all players in N_{i_4} play a^* since each has $\alpha_j(B_{i_3}) \geq \alpha_i^* \geq p$; all players in \bar{B}_{i_4} play strategies in $A_1 \cup a^*$.

$t = 4$	All players in B_{i_4} play a^* since each $j \in B_{i_4}$ has $\alpha_j(B_{i_4}) \geq \alpha_i^* \geq p$; all players in N_{i_5} play a^* since each has $\alpha_j(B_{i_4}) \geq \alpha_i^* \geq p$; all players in \bar{B}_{i_5} play strategies in $A_1 \cup a^*$.
$t = 5$	Following similar intuition, the strategy profile at $t = 5$ is: $B_{i_6} \rightarrow a^*$; and $\bar{B}_{i_6} \rightarrow A_1 \cup a^*$.
— — —	— — — — — The iterative process continues, whereby, at $t = t$ — — — — —
$t = t$	$B_{i_{t+1}} \rightarrow a^*$; and $\bar{B}_{i_{t+1}} \rightarrow A_1 \cup a^*$.
— — —	— — — And after $t = r_i - 1$ iterations, where r_i is the longest geodesic of i (i.e. the maximum shortest distance from i to any player), the strategy profile is: — — —
$t = r_i - 1$	$N \rightarrow a^*$.

We see from the above iterative process that after $t = r_i - 1$ iterations, the evolutionary process converges to convention \mathbf{a}^* . Thus, starting from any $\mathbf{a} \in \mathbf{L} \setminus \mathbf{a}^*$, there exists a best response sequence $\{\mathbf{x}_t\}_{t=0}^{r_i-1}$ with $\mathbf{x}_0 = \mathbf{a}$ and $x_1^j = a^*$ for all $j \in N(\mathbf{a} \rightarrow \mathbf{a}^*) = B_{i_2}$, where $\lim_{t \rightarrow r_i-1} x_t^j = a^*$ for each $j \in N$. This outcome follows if we pick any $i \in N$, including $i \in \arg \max_{i \in N} b_{i_2}$. Thus, the smallest set that ensures the existence of a best response sequence from any $\mathbf{a} \in \mathbf{L} \setminus \mathbf{a}^*$ to \mathbf{a}^* is B_2^* , which is the second-neighbourhood of some $i \in \arg \max_{i \in N} b_{i_2}$. Since the cardinality of B_2^* (i.e. b_2^*) is independent of the population size n , we say that strategy a^* spreads contagiously on G with $n(\mathbf{a} \rightarrow \mathbf{a}^*) \leq b_2^*$ for all $\mathbf{a} \in \mathbf{L} \setminus \mathbf{a}^*$, provided $p \leq \eta(G)$. Note that the upper bound for $n(\mathbf{a} \rightarrow \mathbf{a}^*)$ follows because it maybe possible to find a smaller set than B_2^* that trigger the contagion of a^* but only B_2^* can guarantee that a^* spreads contagiously.

Uninvadability of \mathbf{a}^* : Recall that convention \mathbf{a}^* is uninvadable if $r(\mathbf{a}^*) > n(\mathbf{a} \rightarrow \mathbf{a}_l)$ for all $\mathbf{a} \in \mathbf{L} \setminus \mathbf{a}^*$, and $r(\mathbf{a}^*)$, which is the number of mutations required to leave convention \mathbf{a}^* , is a function of n . We aim to show that if a^* spreads contagiously on a strongly connected $G(N, E)$, then $r(\mathbf{a}^*) \geq n^{\frac{3}{5}}$.

Let $R(\mathbf{a}^*) \subset N$ be the smallest set of players that should mutate to strategies in A_1 for the evolutionary process to leave the basin of attraction of \mathbf{a}^* , where $r(\mathbf{a}^*)$ is the cardinality of $R(\mathbf{a}^*)$. We see from the preceding analysis that if there exists a player i for whom all players in B_{i_2} play a^* , then a^* spreads contagiously, and hence, the evolutionary process converges to \mathbf{a}^* regardless of the strategy configuration of other players not in B_{i_2} . Thus, to leave \mathbf{a}^* , no such player must exist, and hence, each player must be at most two steps away from $R(\mathbf{a}^*)$; that is, for each i a proportion of players in B_{i_2} are in $R(\mathbf{a}^*)$. Since a^* is p -dominant, where $p \leq \frac{1}{2}$, strategies in A_1 are best responses only when more than proportion $(1 - p) > \frac{1}{2}$ of neighbours play strategies in A_1 . For strategies in A_1 to become best responses to players within B_{i_2} , more than $\frac{1}{2}$ of the interactions of players in B_{i_2} must be in $R(\mathbf{a}^*)$; and $R(\mathbf{a}^*)$ must be chosen to satisfy this condition.

The identification of the smallest $R(\mathbf{a}^*)$ is then equivalent to the problem of identifying the smallest 2-monopolies in graph theory, defined as follows. A player i in network $G(N, E)$ is said to be 2-controlled by the set $Z \subset N$ of players if at least half of the players in B_{i_2} are in Z . The set Z is called a 2-monopoly if it 2-controls every player in the network. Bermond et al. (1996, Proposition 4) show that

the minimum size of a 2-monopoly on any undirected and strongly connected network of size n is $n^{\frac{3}{5}}$, and hence, $r(\mathbf{a}^*) > n^{\frac{3}{5}}$.¹³ Note that the inequality follows because we require $R(\mathbf{a}^*)$ to contain *more than half* of interactions of B_{i_2} for all $i \in N$ while the definition of a 2-monopoly requires $R(\mathbf{a}^*)$ to contain *at least half* of the interactions of B_{i_2} .

Since $r(\mathbf{a}^*) > n^{\frac{3}{5}}$ is an increasing function of n , it follows that convention \mathbf{a}^* is uninvadable whenever $n^{\frac{3}{5}} \geq n(\mathbf{a} \rightarrow \mathbf{a}_l)$ for all $\mathbf{a} \in \mathbf{L} \setminus \mathbf{a}^*$. And since $n(\mathbf{a} \rightarrow \mathbf{a}_l) \leq b_2^*$ for all $\mathbf{a} \in \mathbf{L} \setminus \mathbf{a}^*$, convention \mathbf{a}^* is uninvadable, and hence, also contagious, whenever $b_2^* \leq n^{\frac{3}{5}}$.

Appendix C. Proof of Proposition 1

To prove Proposition 1, we first characterize the structure of stationary distributions, and in particular, for a ratio $\frac{\pi_\beta(\mathbf{x})}{\pi_\beta(\mathbf{y})}$ between any pair $\mathbf{x}, \mathbf{y} \in \mathbf{X}$. We use the following results from Freidlin and Wentzell (1984).

Lemma 2. (Freidlin and Wentzell, 1984, Lemma 3.1). *Given a diffusion process P_β , the stationary distribution $\pi_\beta(\mathbf{x})$ of some configuration $\mathbf{x} \in \mathbf{X}_n$ is given by*

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

where the total probability $P_\beta(g)$ associated with each graph g is $P_\beta(g) = \prod_{(\mathbf{z}, \mathbf{y}) \in g} P_\beta(\mathbf{z}, \mathbf{y})$ and $\Gamma(\mathbf{x})$ graphs are defined in Definition 6.

To fully characterise the structure of π_β , we first characterise the structure of transition probability $P_\beta(\mathbf{x}, \mathbf{y})$ between any pair of states $\mathbf{x}, \mathbf{y} \in \mathbf{X}$, followed by the structure of $P_\beta(g)$ for a typical $g \in \Gamma(\mathbf{x})$.

Recall that $c(\mathbf{x}, \mathbf{y})$ is the number of mutations involved in the direct transition from \mathbf{x} to \mathbf{y} . That is, the number of players who choose different strategies in state \mathbf{y} than those chosen in state \mathbf{x} , and that their choices are a result of mutations. The transition probabilities, $P_\epsilon(\mathbf{x}, \mathbf{y})$, can be directly expressed in terms of $c(\mathbf{x}, \mathbf{y})$ as

$$P_\beta(\mathbf{x}, \mathbf{y}) = \left(\frac{e^{-\beta}}{m} \right)^{c(\mathbf{x}, \mathbf{y})} \prod_{i=1}^{n-c(\mathbf{x}, \mathbf{y})} \left(\frac{m + (b_i(\mathbf{x}) - m)e^{-\beta}}{mb_i(\mathbf{x})} \right) \quad (\text{C.2})$$

where, given $\sigma_i(\mathbf{x})$ as the strategy distribution over i 's neighbourhood in configuration \mathbf{x} (see discussion in Section 2.2), $b_i(\mathbf{x})$ is the cardinality of $BR(\sigma_i(\mathbf{x}))$.

The right hand side of (C.2) follows because, first, if y^i (i.e. the strategy i plays in configuration \mathbf{y}) is not a best response to \mathbf{x} so that $BR_i(y^i; \mathbf{x}) = 0$, then from (4), the probability that i plays y^i is $\frac{e^{-\beta}}{m}$. Consequently, the probability that $c(\mathbf{x}, \mathbf{y})$ players simultaneously play strategies that are not best responses to \mathbf{x} in configuration \mathbf{y} is $\left(\frac{e^{-\beta}}{m} \right)^{c(\mathbf{x}, \mathbf{y})}$.

¹³See Peleg (2002) for a review of the literature on monopolies and local majorities in networks.

Second, if y^i is instead a best response to \mathbf{x} , so that $BR_i(y^i; \mathbf{x}) = \frac{1}{b_i(\mathbf{x})}$, then i plays y^i with probability

$$\frac{e^{-\beta}}{m} + (1 - e^{-\beta}) \frac{1}{b_i(\mathbf{x})} = \frac{m + (b_i(\mathbf{x}) - m)e^{-\beta}}{mb_i(\mathbf{x})}$$

Thus, the probability that the remaining $n - c(\mathbf{x}, \mathbf{y})$ players simultaneously play strategies that are best responses to \mathbf{x} is $\prod_{i=1}^{n-c(\mathbf{x}, \mathbf{y})} \left(\frac{m + (b_i(\mathbf{x}) - m)e^{-\beta}}{mb_i(\mathbf{x})} \right)$.

From (C.2), we see that if $b_i(\mathbf{x}) = 1$ for all $i \in N$ and for any $\mathbf{x} \in \mathbf{X}$, which is very plausible because our analysis focuses on strict coordination games (i.e. coordination games where coordination on any pure strategy is a strict Nash equilibrium), then

$$P_\beta(\mathbf{x}, \mathbf{y}) = \left(\frac{e^{-\beta}}{m} \right)^{c(\mathbf{x}, \mathbf{y})} \left(\frac{m + (1 - m)e^{-\beta}}{m} \right)^{n-c(\mathbf{x}, \mathbf{y})} \quad (\text{C.3})$$

which represents the upper bound for $P_\beta(\mathbf{x}, \mathbf{y})$. The lower bound for $P_\beta(\mathbf{x}, \mathbf{y})$ is obtained in the least likely scenario when $b_i(\mathbf{x}) = m$ for all $i \in N$ (i.e. when all strategies are best responses to \mathbf{x}) so that

$$P_\beta(\mathbf{x}, \mathbf{y}) = \left(\frac{e^{-\beta}}{m} \right)^{c(\mathbf{x}, \mathbf{y})} \left(\frac{1}{m} \right)^{n-c(\mathbf{x}, \mathbf{y})} = \left(e^{-\beta} \right)^{c(\mathbf{x}, \mathbf{y})} \left(\frac{1}{m} \right)^n \quad (\text{C.4})$$

Next, we characterise the probabilities of $\Gamma(\mathbf{x})$ graphs. From the definition of $\Gamma(\mathbf{x})$ graphs, every $g \in \Gamma(\mathbf{x})$ spans the entire state space, except for \mathbf{x} , each $\mathbf{y} \in \mathbf{X}$ has only one arrow emanating from it. Thus, each $g \in \Gamma(\mathbf{x})$ contains a total of $m^n - 1$ directed edges, where m^n is the cardinality of \mathbf{X} . Similarly, if we let $d(L_j)$ be the cardinality of $D(L_j)$ (the basin of attraction of L_j), then there are $d(L_j)$ directed edges that originate from states in $D(L_j)$.

Now, let $g(D(L))$ be a subgraph of g consisting of all $d(L_j)$ directed edges that originate from states in $D(L_j)$. Since $D(L_j)$ for all $L_j \subset \mathbf{L}$ are non-overlapping sets, we can rewrite $P_\beta(g)$ as

$$P_\beta(g) = \prod_{L_j \in \mathbf{L}} \prod_{(\mathbf{y}, \mathbf{z}) \in g(D(L_j))} P_\beta(\mathbf{y}, \mathbf{z}) \quad (\text{C.5})$$

We can further subdivide the set of edges of $g(D(L_j))$ into those that involve at least one mutations, denoted $g(D(L_j); \beta)$, and those whose dynamics are governed solely by best response, denoted $\overline{g(D(L_j); \beta)}$. That is, for each $\mathbf{x} \in D(L_j)$ and some $\mathbf{y} \neq \mathbf{x}$, a directed edge $(\mathbf{x}, \mathbf{y}) \in g(D(L_j); \beta)$ if $c(\mathbf{x}, \mathbf{y}) > 0$, and $(\mathbf{x}, \mathbf{y}) \in \overline{g(D(L_j); \beta)}$ if $c(\mathbf{x}, \mathbf{y}) = 0$. Using these definitions and notation, and substituting $c(\mathbf{x}, \mathbf{y}) = 0$ into (C.2) for all $(\mathbf{x}, \mathbf{y}) \in \overline{g(D(L_j); \beta)}$, $P_\beta(g)$ can be rewritten as

$$\begin{aligned} P_\beta(g) &= \prod_{L_j \in \mathbf{L}} \prod_{(\mathbf{y}, \mathbf{z}) \in g(D(L_j))} P_\beta(\mathbf{y}, \mathbf{z}) \\ &= \prod_{L_j \in \mathbf{L}} \prod_{(\mathbf{y}, \mathbf{z}) \in g(D(L_j))} \left(\frac{e^{-\beta}}{m} \right)^{c(\mathbf{y}, \mathbf{z})} \prod_{i=1}^{n-c(\mathbf{y}, \mathbf{z})} \left(\frac{m + (b_i(\mathbf{y}) - m)e^{-\beta}}{mb_i(\mathbf{y})} \right) \\ &= \prod_{L_j \in \mathbf{L}} \left[\prod_{(\mathbf{y}, \mathbf{z}) \in g(D(L_j); \beta)} \left(\frac{e^{-\beta}}{m} \right)^{c(\mathbf{y}, \mathbf{z})} \prod_{i=1}^{n-c(\mathbf{y}, \mathbf{z})} \left(\frac{m + (b_i(\mathbf{y}) - m)e^{-\beta}}{mb_i(\mathbf{y})} \right) \right. \\ &\quad \times \left. \prod_{(\mathbf{y}, \mathbf{z}) \in \overline{g(D(L_j); \beta)}} \left(\frac{m + (b_i(\mathbf{y}) - m)e^{-\beta}}{mb_i(\mathbf{y})} \right)^n \right] \end{aligned} \quad (\text{C.6})$$

The corresponding expressions for the upper and lower bounds of $P_\beta(g)$ using the expressions of $P_\beta(\mathbf{x}, \mathbf{y})$ in (C.3) and (C.4) are respectively

$$P_\beta(g) \leq \prod_{L_j \in \mathbf{L}} \left[\prod_{(\mathbf{y}, \mathbf{z}) \in g(D(L_j); \beta)} \left(\frac{e^{-\beta}}{m} \right)^{c(\mathbf{y}, \mathbf{z})} \left(\frac{m + (1-m)e^{-\beta}}{m} \right)^{n-c(\mathbf{y}, \mathbf{z})} \right] \quad (\text{C.7})$$

$$\times \prod_{(\mathbf{y}, \mathbf{z}) \in \overline{g(D(L_j); \beta)}} \left(\frac{m + (1-m)e^{-\beta}}{m} \right)^n \quad (\text{C.8})$$

Let $n(g; L_j; \beta)$ be the cardinality of subgraph $g(D(L_j); \beta)$ (i.e. the number of directed edges in subgraph $g(D(L_j); \beta)$) so that the cardinality of $\overline{g(D(L_j); \beta)}$ is $d(L_j) - n(g; L_j; \beta)$. Then (C.7) can be rewritten as

$$\begin{aligned} P_\beta(g) &\leq \prod_{L_j \in \mathbf{L}} \left[\left(\frac{e^{-\beta}}{m} \right)^{\sum_{(\mathbf{y}, \mathbf{z}) \in g(D(L_j); \beta)} c(\mathbf{y}, \mathbf{z})} \left(\frac{m + (1-m)e^{-\beta}}{m} \right)^{\sum_{(\mathbf{y}, \mathbf{z}) \in g(D(L_j); \beta)} (n-c(\mathbf{y}, \mathbf{z}))} \right. \\ &\quad \left. \times \left(\frac{m + (1-m)e^{-\beta}}{m} \right)^{n[d(L_j) - n(g; L_j; \beta)]} \right] \\ &= \prod_{L_j \in \mathbf{L}} \left(\frac{e^{-\beta}}{m} \right)^{\sum_{(\mathbf{y}, \mathbf{z}) \in g(D(L_j); \beta)} c(\mathbf{y}, \mathbf{z})} \left(\frac{m + (1-m)e^{-\beta}}{m} \right)^{n[d(L_j) - n(g; L_j; \beta)] + \sum_{(\mathbf{y}, \mathbf{z}) \in g(D(L_j); \beta)} (n-c(\mathbf{y}, \mathbf{z}))} \end{aligned} \quad (\text{C.9})$$

The summation $c(L_j; g) = \sum_{(\mathbf{y}, \mathbf{z}) \in g(D(L_j); \beta)} c(\mathbf{y}, \mathbf{z})$ on the numerators of the expressions of $P_\beta(g)$ in (C.9) are the total cost of leaving the basin of attraction of each L_j under graph $g \in \Gamma(\mathbf{x})$. Using this definition, we can simplify $n[d(L_j) - n(g; L_j; \beta)] + \sum_{(\mathbf{y}, \mathbf{z}) \in g(D(L_j); \beta)} (n - c(\mathbf{y}, \mathbf{z}))$ as follows:

$$\begin{aligned} n[d(L_j) - n(g; L_j; \beta)] + \sum_{(\mathbf{y}, \mathbf{z}) \in g(D(L_j); \beta)} (n - c(\mathbf{y}, \mathbf{z})) &= n[d(L_j) - n(g; L_j; \beta)] + n(g; L_j; \beta)n - c(L_j; g) \\ &= nd(L_j) - c(L_j; g) \end{aligned}$$

Equation (C.9) then simplifies to

$$\begin{aligned} P_\beta(g) &\leq \prod_{L_j \in \mathbf{L}} \left(\frac{e^{-\beta}}{m} \right)^{c(L_j; g)} \left(\frac{m + (1-m)e^{-\beta}}{m} \right)^{nd(L_j) - c(L_j; g)} \\ &= \exp \left[- \sum_{L_j \in \mathbf{L}} \left(c(L_j; g)\beta_m + (nd(L_j) - c(L_j; g))\beta'_m \right) \right] \\ &= \exp \left[- \sum_{L_j \in \mathbf{L}} \left(c(L_j; g)(\beta_m - \beta'_m) + nd(L_j)\beta'_m \right) \right] \end{aligned} \quad (\text{C.10})$$

where $\beta_m = \beta - \ln m^{-1}$ and $\beta'_m = -\ln \left[\frac{m + (1-m)e^{-\beta}}{m} \right]$. Following similar steps and using the lower bound for $P_\beta(g)$ as

$$P_\beta(g) \geq \exp \left[- \sum_{L_j \in \mathbf{L}} \left(c(L_j; g)\beta + nd(L_j) \ln m \right) \right] \quad (\text{C.11})$$

Recall the definition of the long-run equilibrium as configurations that maximize the stationary distribution. So, to compute the long-run equilibrium, we take ratios of probabilities and identify

configurations for which the ration is less than one. Specifically, configuration \mathbf{a}^* is a long-run equilibrium if $\frac{\pi_\beta(\mathbf{x})}{\pi_\beta(\mathbf{a}^*)} \leq 1$ for all $\mathbf{x} \neq \mathbf{a}^*$. From $\pi_\beta(\mathbf{x})$ in (C.1), the expression for the ratio of stationary distribution, $\frac{\pi_\beta(\mathbf{x})}{\pi_\beta(\mathbf{w})}$ of any pair of configurations $\mathbf{x}, \mathbf{w} \in \mathbf{X}$ is given by

$$\frac{\pi_\beta(\mathbf{x})}{\pi_\beta(\mathbf{w})} = \frac{\sum_{g \in \Gamma(\{\mathbf{x}\})} P_\beta(g)}{\sum_{g \in \Gamma(\{\mathbf{w}\})} P_\beta(g)} \quad (\text{C.12})$$

where the quantity $\left(\sum_{\mathbf{y} \in \mathbf{X}} \sum_{g \in \Gamma(\{\mathbf{y}\})} P_\beta(g)\right)^{-1}$ cancels out since it is identical for all configurations.

Let $\gamma(\mathbf{x}) = \#\Gamma(\mathbf{x})$ be the cardinality of $\Gamma(\mathbf{x})$. Since for any $\mathbf{x} \in \mathbf{X}$ every $g \in \Gamma(\mathbf{x})$ spans the entire state space except for one configuration (i.e. the number of edges in any $g \in \Gamma(\mathbf{x})$ is equal to $m^n - 1$), the cardinality of $\Gamma(\mathbf{x})$ is the same for any pair of configurations $\mathbf{x}, \mathbf{y} \in \mathbf{X}$; that is, $\gamma(\mathbf{x}) = \gamma(\mathbf{y}) = \gamma$. Using this notation, the following bounds hold:

$$\begin{aligned} \sum_{g \in \Gamma(\{\mathbf{x}\})} P_\beta(g) &\leq \sum_{g \in \Gamma(\{\mathbf{x}\})} \max_{g \in \Gamma(\{\mathbf{x}\})} P_\beta(g) = \gamma \max_{g \in \Gamma(\{\mathbf{x}\})} P_\beta(g) \\ \max_{g \in \Gamma(\{\mathbf{x}\})} P_\beta(g) &\leq \sum_{g \in \Gamma(\{\mathbf{x}\})} P_\beta(g) \end{aligned}$$

Thus, the ratio $\frac{\pi_\beta(\mathbf{x})}{\pi_\beta(\mathbf{w})}$ is bounded from below and above by

$$\frac{\max_{g \in \Gamma(\{\mathbf{x}\})} P_\beta(g)}{\gamma \max_{g \in \Gamma(\{\mathbf{w}\})} P_\beta(g)} \leq \frac{\pi_\beta(\mathbf{x})}{\pi_\beta(\mathbf{w})} \leq \frac{\gamma \max_{g \in \Gamma(\{\mathbf{x}\})} P_\beta(g)}{\max_{g \in \Gamma(\{\mathbf{w}\})} P_\beta(g)} \quad (\text{C.13})$$

First consider the expression for the upper bound of $P_\beta(g)$. Substituting for $P_\beta(g)$ yields the following expression.

$$\begin{aligned} \frac{\max_{g \in \Gamma(\{\mathbf{x}\})} P_\beta(g)}{\max_{g \in \Gamma(\{\mathbf{w}\})} P_\beta(g)} &= \frac{\max_{g \in \Gamma(\{\mathbf{x}\})} \exp \left[-\sum_{L_j \in \mathbf{L}} \left(c(L_j; g)(\beta_m - \beta'_m) + nd(L_j)\beta'_m \right) \right]}{\max_{g \in \Gamma(\{\mathbf{w}\})} \exp \left[-\sum_{L_j \in \mathbf{L}} \left(c(L_j; g)(\beta_m - \beta'_m) + nd(L_j)\beta'_m \right) \right]} \\ &= \exp \left[\min_{g \in \Gamma(\{\mathbf{w}\})} \sum_{L_j \in \mathbf{L}} c(L_j; g)(\beta_m - \beta'_m) - \min_{g \in \Gamma(\{\mathbf{x}\})} \sum_{L_j \in \mathbf{L}} c(L_j; g)(\beta_m - \beta'_m) \right] \\ &= \exp \left[-(\beta_m - \beta'_m) \left(\min_{g \in \Gamma(\{\mathbf{x}\})} \sum_{L_j \in \mathbf{L}} c(L_j; g) - \min_{g \in \Gamma(\{\mathbf{w}\})} \sum_{L_j \in \mathbf{L}} c(L_j; g) \right) \right] \quad (\text{C.14}) \end{aligned}$$

whereby, due to the negative in the exponents, the expressions for $P_\beta(g)$ are maximized when the costs of exiting the basins of attractions of absorbing sets are minimized. Substituting (C.14) into (C.13) yields the following lower and upper bounds for $\frac{\pi_\beta(\mathbf{x})}{\pi_\beta(\mathbf{w})}$

$$\frac{\pi_\beta(\mathbf{x})}{\pi_\beta(\mathbf{w})} = \exp \left[-(\beta_m - \beta'_m) \left(\min_{g \in \Gamma(\{\mathbf{x}\})} \sum_{L_j \in \mathbf{L}} c(L_j; g) - \min_{g \in \Gamma(\{\mathbf{w}\})} \sum_{L_j \in \mathbf{L}} c(L_j; g) \right) \pm \ln \gamma \right] \quad (\text{C.15})$$

Similarly, using the expression for the lower bound in (C.11) yields the following bounds for $\frac{\pi_\beta(\mathbf{x})}{\pi_\beta(\mathbf{w})}$

$$\frac{\pi_\beta(\mathbf{x})}{\pi_\beta(\mathbf{w})} = \exp \left[-\beta \left(\min_{g \in \Gamma(\{\mathbf{x}\})} \sum_{L_j \in \mathbf{L}} c(L_j; g) - \min_{g \in \Gamma(\{\mathbf{w}\})} \sum_{L_j \in \mathbf{L}} c(L_j; g) \right) \pm \ln \gamma \right] \quad (\text{C.16})$$

Now, notice that for all $\beta \geq 0$, $(\beta_m - \beta'_m) \geq 0$: that is,

$$\begin{aligned} (\beta_m - \beta'_m) &= \beta - \ln m^{-1} + \ln \left[\frac{m + (1-m)e^{-\beta}}{m} \right] = \beta - \ln m^{-1} + \ln [m + (1-m)e^{-\beta}] + \ln m^{-1} \\ &= \beta + \ln [m + (1-m)e^{-\beta}] \end{aligned} \quad (\text{C.17})$$

Since $m + (1-m)e^{-\beta} \geq 1$ for all $\beta \geq 0$, it follows from the right hand side of (C.17) that $(\beta_m - \beta'_m) \geq 0$ for all $\beta \geq 0$.

Thus, for all $\beta \geq 0$, we see from (C.15) that $\frac{\pi_\beta(\mathbf{x})}{\pi_\beta(\mathbf{w})} \leq 1$ whenever

$$\left(\min_{g \in \Gamma(\{\mathbf{x}\})} \sum_{L_j \in \mathbf{L}} c(L_j; g) - \min_{g \in \Gamma(\{\mathbf{w}\})} \sum_{L_j \in \mathbf{L}} c(L_j; g) \right) \geq 0$$

and that

$$-(\beta_m - \beta'_m) \left(\min_{g \in \Gamma(\{\mathbf{x}\})} \sum_{L_j \in \mathbf{L}} c(L_j; g) - \min_{g \in \Gamma(\{\mathbf{w}\})} \sum_{L_j \in \mathbf{L}} c(L_j; g) \right) + \ln \gamma \leq 0 \quad (\text{C.18})$$

Substituting for the values of β_m and β'_m in (C.18), and letting $\Phi(\mathbf{x}, \mathbf{w}) = \min_{g \in \Gamma(\{\mathbf{x}\})} \sum_{L_j \in \mathbf{L}} c(L_j; g) - \min_{g \in \Gamma(\{\mathbf{w}\})} \sum_{L_j \in \mathbf{L}} c(L_j; g)$, yields

$$\beta \geq \frac{\ln \gamma}{\Phi(\mathbf{x}, \mathbf{w})} - \ln [m + (1-m)e^{-\beta}] \quad (\text{C.19})$$

Similarly, for (C.16), $\frac{\pi_\beta(\mathbf{x})}{\pi_\beta(\mathbf{w})} \leq 1$ whenever $\left(\min_{g \in \Gamma(\{\mathbf{x}\})} \sum_{L_j \in \mathbf{L}} c(L_j; g) - \min_{g \in \Gamma(\{\mathbf{w}\})} \sum_{L_j \in \mathbf{L}} c(L_j; g) \right) \geq 0$ and

$$\beta \geq \frac{\ln \gamma}{\Phi(\mathbf{x}, \mathbf{w})} \quad (\text{C.20})$$

Note that for all $\beta \geq 0$, the function $\ln [m + (1-m)e^{-\beta}]$ in (C.19) increases from 0 to the upper $\ln m$ as $\beta \rightarrow \infty$. Thus, the minimum value of β in (C.19) is less than $\frac{\ln \gamma}{\Phi(\mathbf{x}, \mathbf{w})}$ in (C.20). That is, since $\ln [m + (1-m)e^{-\beta}]$ increases from 0 to $\ln m$, $\frac{\ln \gamma}{\Phi(\mathbf{x}, \mathbf{w})} - \ln [m + (1-m)e^{-\beta}]$ decreases from $\frac{\ln \gamma}{\Phi(\mathbf{x}, \mathbf{w})}$ to $\frac{\ln \gamma}{\Phi(\mathbf{x}, \mathbf{w})} - \ln m$, such that the equilibrium value of β in (C.19) is less than $\frac{\ln \gamma}{\Phi(\mathbf{x}, \mathbf{w})}$; that is, the 45-degree line depicting $\beta = \beta$ meets $\frac{\ln \gamma}{\Phi(\mathbf{x}, \mathbf{w})} - \ln [m + (1-m)e^{-\beta}]$ at a value less than $\frac{\ln \gamma}{\Phi(\mathbf{x}, \mathbf{w})}$.

Since lower bounds for β in (C.19) and (C.20) are obtained for the two extreme cases where $b_i(\mathbf{x}) = \frac{1}{m}$ and $b_i(\mathbf{x}) = 1$ for all i and any $\mathbf{x} \in \mathbf{X}$ respectively, and that the minimum value of β in (C.20) is greater than that in (C.19), it follows that $\frac{\pi_\beta(\mathbf{x})}{\pi_\beta(\mathbf{w})} \leq 1$ whenever $\Phi(\mathbf{x}, \mathbf{w}) \geq 0$ and

$$\beta \geq \frac{\ln \gamma}{\Phi(\mathbf{x}, \mathbf{w})} \quad (\text{C.21})$$

Thus, \mathbf{a}^* is a long-run equilibrium whenever $\Phi(\mathbf{a}^*, \mathbf{w}) \geq 0$ and $\beta \geq \frac{\ln \gamma}{\Phi(\mathbf{a}^*, \mathbf{w})}$ for all $\mathbf{x} \neq \mathbf{a}^*$. That is, if \mathbf{a}^* has the minimum cost $\Gamma(\mathbf{a}^*)$ graph, then it is a long-run equilibrium whenever $\beta \geq \frac{\ln \gamma}{\Phi(\mathbf{a}^*, \mathbf{w})}$ for all $\mathbf{x} \neq \mathbf{a}^*$.

We now show that if \mathbf{a}^* is contagious in a given network, then convention \mathbf{a}^* has the minimum cost $\Gamma(\mathbf{a}^*)$ graph. Note that configurations with the minimum cost graph belong to some absorbing state,

and hence, it suffices to focus on examining the costs of graphs for configurations within \mathbf{L} . Consider any $\mathbf{a} \in L \subset \mathbf{L}$. The cost of the minimum cost graphs, $c(\mathbf{a})$ and $c(\mathbf{a}^*)$, for the \mathbf{a}^* and \mathbf{x} respectively can be rewritten as follows.

$$c(\mathbf{a}) = \min_{g \in \Gamma(\{\mathbf{a}\})} \sum_{L_j \in \mathbf{L}} c(L_j; g) = \min_{g \in \Gamma(\{\mathbf{a}\})} \sum_{L_j \in \mathbf{L} \setminus \{\mathbf{a}^*, L\}} c(L_j; g) + \min_{g \in \Gamma(\{\mathbf{a}\})} c(\mathbf{a}^*; g) \quad (\text{C.22})$$

$$c(\mathbf{a}^*) = \min_{g \in \Gamma(\{\mathbf{a}^*\})} \sum_{L_j \in \mathbf{L}} c(L_j; g) = \min_{g \in \Gamma(\{\mathbf{a}^*\})} \sum_{L_j \in \mathbf{L} \setminus \{\mathbf{a}^*, L\}} c(L_j; g) + \min_{g \in \Gamma(\{\mathbf{a}^*\})} c(\mathbf{a}; g) \quad (\text{C.23})$$

Note that in the summations in the second equalities of (C.22) and (C.23) we exclude both \mathbf{a}^* and L . For the summation in (C.22), \mathbf{a}^* is excluded because its cost is accounted for the second equality. However, L is excluded because in the minimum cost $\Gamma(\mathbf{a})$ graph, all arrows originating from any $\mathbf{x} \in D(L)$ must point towards L , corresponding to best response transitions – by definition, the Markov chain converges to some $\mathbf{a} \in L$ from every $\mathbf{x} \in D(L)$. Thus, in a minimum cost $\Gamma(\mathbf{a})$ graph, $\min_{g \in \Gamma(\{\mathbf{a}\})} c(\mathbf{a}; g) = 0$. The same argument applies to the summation in the second equality of (C.23).

Now, notice that the cost $c(\mathbf{a})$ of any $\mathbf{a} \in L \subset \mathbf{L}$ is identical to that derived from a reduced form of $\Gamma(\mathbf{a})$, denoted by $\Gamma'(\mathbf{a})$, on a state space $\mathbf{L} \times \mathbf{L}$ (i.e. where vertices are absorbing sets), where for any $g \in \Gamma'(\mathbf{a})$ the cost of an arrow originating from some $L_j \in \mathbf{L}$ is $c(L_j; g)$. From the proof of Theorem 1 in Appendix (Appendix B), it is easy to see that for convention \mathbf{a}^* , corresponding to the contagious strategy a^* , the minimum cost graph in $\Gamma'(\mathbf{a}^*)$ involves direct transitions from every $L_j \subset L \subset \mathbf{L} \setminus \mathbf{a}^*$ to \mathbf{a}^* , and each has a cost bounded from above by $c(L_j; g) \leq b_2^*$. Thus, the minimum cost graph in $\Gamma'(\mathbf{a})$, for $\mathbf{a} \in L \subset \mathbf{L} \setminus \mathbf{a}^*$, can be constructed from the minimum cost graph in $\Gamma'(\mathbf{a}^*)$ by deleting $\mathbf{a} \rightarrow \mathbf{a}^*$ and replacing it with $\mathbf{a}^* \rightarrow \mathbf{a}$. Under this construction, $\min_{g \in \Gamma(\{\mathbf{a}\})} \sum_{L_j \in \mathbf{L} \setminus \{\mathbf{a}^*, L\}} c(L_j; g) = \min_{g \in \Gamma(\{\mathbf{a}^*\})} \sum_{L_j \in \mathbf{L} \setminus \{\mathbf{a}^*, L\}} c(L_j; g)$ such that the difference between $c(\mathbf{a})$ and $c(\mathbf{a}^*)$ is

$$c(\mathbf{a}) - c(\mathbf{a}^*) = \min_{g \in \Gamma(\{\mathbf{a}\})} c(\mathbf{a}^*; g) - \min_{g \in \Gamma(\{\mathbf{a}^*\})} c(\mathbf{a}; g) > n^{\frac{3}{5}} - b_2^* \quad (\text{C.24})$$

where the second inequality is because the minimum cost of leaving \mathbf{a}^* is $n^{\frac{3}{5}}$ mutations (see Appendix (Appendix B)). Since $n^{\frac{3}{5}} \geq b_2^*$ is a condition for strategy a^* to be contagious, it follows that $c(\mathbf{a}) - c(\mathbf{a}^*) > 0$ for all $\mathbf{a} \in \mathbf{L} \setminus \mathbf{a}^*$, and hence \mathbf{a}^* has the minimum cost \mathbf{a}^* -tree. And since $c(\mathbf{a}) - c(\mathbf{a}^*) > n^{\frac{3}{5}} - b_2^*$, we have

$$\frac{\ln \gamma}{\Phi(\mathbf{x}, \mathbf{w})} < \frac{\ln \gamma}{n^{\frac{3}{5}} - b_2^*}$$

Hence, there exists some $\beta^* \in (0, \ln \gamma / (n^{\frac{3}{5}} - b_2^*)]$ such that for all $\beta \geq \beta^*$, convention \mathbf{a}^* is the long-run equilibrium.

Appendix D. Proof of Proposition 2

The following definitions are used in the next steps of the proof. Given transition probabilities $P_\beta(\mathbf{x}, \mathbf{y})$ in (C.2), the cost function $c(\mathbf{x}, \mathbf{y})$ can be rewritten as follows

$$\lim_{\beta \rightarrow \infty} \frac{-\ln P_\beta(\mathbf{x}, \mathbf{y})}{\beta} = c(\mathbf{x}, \mathbf{y}) \quad (\text{D.1})$$

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

Definition 8. 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 .

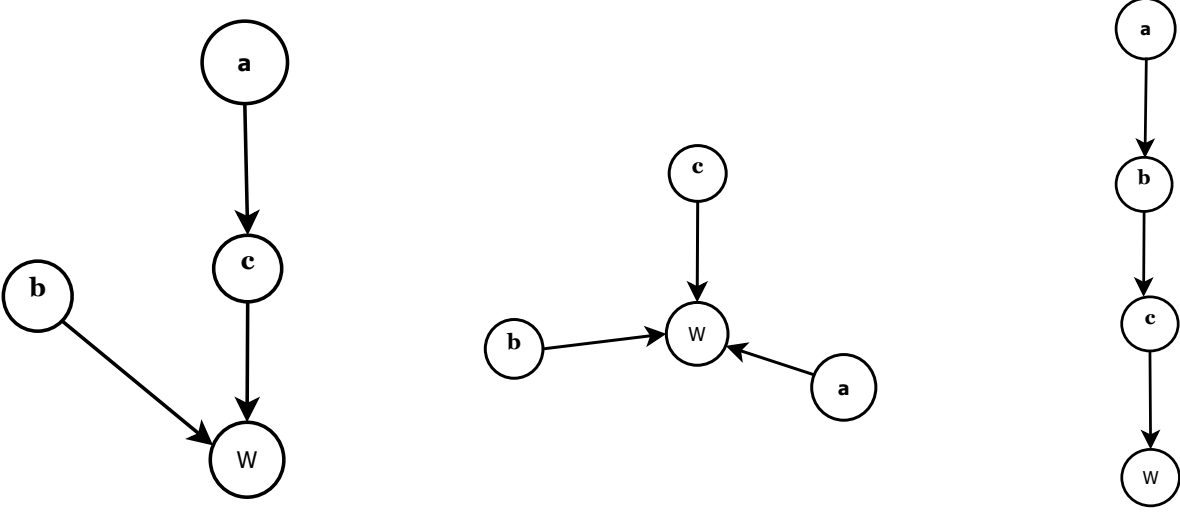


Figure D.6: Examples of $\Gamma(W)$ graphs, where $\bar{W} = \{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$.

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 5. Let $W = \{\mathbf{d}, \mathbf{e}, \mathbf{f}, \mathbf{g}, \mathbf{h}\}$, with examples of $\Gamma(W)$ graphs depicted in Figure D.6. Then examples of $\Gamma_{\mathbf{a},\mathbf{c}}(W \cup \{\mathbf{c}\})$ graphs based on $\Gamma(W)$ graphs in Figure D.6 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 3. For any $W \subset \mathbf{X}$, $W \neq \emptyset$ and $\bar{W} = \mathbf{X} \setminus W$, for any $\mathbf{x}, \mathbf{y} \in \bar{W}$

$$\lim_{\beta \rightarrow \infty} \frac{\ln \mathbb{E}[T(W) | \mathbf{x}_0 = \mathbf{x}]}{\beta} = \min_{g \in \Gamma(W)} c'(g) - \min_{\mathbf{y} \in \bar{W}} \min_{g \in \Gamma_{\mathbf{x},\mathbf{y}}(W \cup \{\mathbf{y}\})} c'(g) \quad (\text{D.2})$$

We are interested in the expected waiting time for convention \mathbf{a}^* associated with the contagious strategy \mathbf{a}^* . Thus, we can substitute $W = \{\mathbf{a}^*\}$ into (D.2). That is,

$$\lim_{\beta \rightarrow \infty} \frac{\ln \mathbb{E}[T(\mathbf{a}^*) | \mathbf{x}_0 = \mathbf{x}]}{\beta} = \min_{g \in \Gamma(\mathbf{a}^*)} c'(g) - \min_{\mathbf{y} \in \mathbf{L} \setminus \mathbf{a}^*} \min_{g \in \Gamma_{\mathbf{x},\mathbf{y}}(\mathbf{a}^* \cup \{\mathbf{y}\})} c'(g) \quad (\text{D.3})$$

Recall from the analysis in Sections Appendix B and Appendix C $c(\mathbf{a}; g) \leq b_2^*$ for all $\mathbf{a} \in \mathbf{L} \setminus \mathbf{a}^*$, and that

$$c(\mathbf{a}^*) = \min_{g \in \Gamma(\mathbf{a}^*)} c'(g) = \min_{g \in \Gamma(\mathbf{a}^*)} \sum_{L_j \in \mathbf{L}} c(L_j; g) \quad (\text{D.4})$$

Thus, there exists some $r^* \in (0, b_2^*)$ such that $\min_{g \in \Gamma(\mathbf{a}^*)} c'(g) - \min_{\mathbf{y} \in \mathbf{L} \setminus \mathbf{a}^*} \min_{g \in \Gamma_{\mathbf{x},\mathbf{y}}(\mathbf{a}^* \cup \{\mathbf{y}\})} c'(g) = r^*$.

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, 179–197.
- Alós-Ferrer, C., Weidenholzer, S., 2008. Contagion and efficiency. *Journal of Economic Theory* 143, 251–274.
- Anderlini, L., Ianni, A., 1996. Path dependence and learning from neighbors. *Games and Economic Behavior* 13, 141–177.
- Beaman, L., BenYishay, A., Magruder, J., Mobarak, A.M., 2018. Can Network Theory-based Targeting Increase Technology Adoption? Technical Report. National Bureau of Economic Research.
- Bermond, J.C., Bond, J., Peleg, D., Perennes, S., 1996. Tight bounds on the size of 2-monopolies., in: SIROCCO, pp. 170–179.
- Berninghaus, S.K., Schwalbe, U., 1996. Conventions, local interaction, and automata networks. *Journal of Evolutionary Economics* 6, 297–312.
- Blume, L.E., 1995. The statistical mechanics of best-response strategy revision. *Games and Economic Behavior* 11, 111–145.
- Campbell, A., 2013. Word-of-mouth communication and percolation in social networks. *American Economic Review* 103, 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.
- Ellison, G., 1993. Learning, local interaction, and coordination. *Econometrica* 61, 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, 17–45.
- 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, 509–532.
- Galeotti, A., Goyal, S., Jackson, M.O., Vega-Redondo, F., Yariv, L., 2010. Network games. *Review of Economic Studies* 77, 218–244.
- 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, 92–98.
- Kandori, M., Mailath, G.J., Rob, R., 1993. Learning, mutation, and long run equilibria in games. *Econometrica* 61, 29–56.
- 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, 10881–10888.
- Lee, I.H., Szeidl, A., Valentinyi, A., 2003. Contagion and State Dependent Mutations. *The B.E. Journal of Theoretical Economics* 3, 1–29.
- Lee, I.H., Valentinyi, A., 2000. Noisy contagion without mutation. *Review of Economic Studies* 67, 47–56.
- López-Pintado, D., 2008. Diffusion in complex social networks. *Games and Economic Behavior* 62, 573–590.
- Morris, S., 2000. Contagion. *Review of Economic Studies* 67, 57–78.
- Morris, S., Rob, R., Shin, H.S., 1995. p -Dominance and Belief Potential. *Econometrica* 63, 145–57.
- Oyama, D., Takahashi, S., 2015. Contagion and uninvadability in local interaction games: The bilingual game and general supermodular games. *Journal of Economic Theory* 157, 100–127.

- Peleg, D., 2002. Local majorities, coalitions and monopolies in graphs: a review. *Theoretical Computer Science* 282, 231–257.
- Peski, M., 2010. Generalized risk-dominance and asymmetric dynamics. *Journal of Economic Theory* 145, 216–248.
- Robson, A.J., Vega-Redondo, F., 1996. Efficient Equilibrium Selection in Evolutionary Games with Random Matching. *Journal of Economic Theory* 70, 65–92.
- Sundararajan, A., 2007. Local network effects and complex network structure. *The BE Journal of Theoretical Economics* 7.
- Tsakas, N., 2017. Diffusion by imitation: The importance of targeting agents. *Journal of Economic Behavior & Organization* 139, 118–151.
- Young, H.P., 1993. The evolution of conventions. *Econometrica* 61, 57–84.
- Young, H.P., 1998. *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*. Princeton University Press.
- Young, P.H., 2011. The dynamics of social innovation. *Proceedings of the National Academy of Sciences* 108, 21285–21291.