

# On the relationship between $p$ -dominance and stochastic stability in network games

Daniel C. Opolot

*School of Economics, University of Cape Town*

---

## Abstract

This paper studies the long-run (stochastic) stability of an evolutionary model with persistent randomness in network games. We focus on evolutionary dynamics in symmetric coordination games where players choose the best responses to their opponents' strategies. Recent studies find that the network structure determines uniqueness of stochastically stable outcomes in such models. We identify the *contagion threshold* (i.e. a number below which some strategies can spread contagiously across a network) as one of the aggregate network measures that determines the uniqueness of stochastically stable outcomes. Specifically, we show that strategies that form a  *$p$ -best response set* (i.e. strategies that are best responses to any distribution that places on them a mass of at least  $p$ ) are uniquely stochastically stable in networks with a contagion threshold equal to  $p$ . This result holds for all strongly connected networks with diameter greater or equal to 5.

**Keywords:** evolutionary dynamics, stochastic stability, networks,  $p$ -best response set, contagion threshold

**JEL codes:** C72, C73, D83, D85.

---

## 1. Introduction

The concept of Nash equilibrium plays a central role in the field of game theory and economics. However, many games, such as coordination games,<sup>1</sup> have multiple strict Nash equilibria. This raises the question of which outcomes should be regarded as more reasonable than others. To address this issue, Foster and Young (1990), Kandori et al. (1993) and Young (1993) proposed

---

\*Postal address: University of Cape Town, Private Bag X3, Rondebosch, 7701 Cape Town

Email address: opodanchris@gmail.com

<sup>1</sup>Coordination games represent a class of games where players benefit most by choosing the same strategy. They are generally applied to model technology choice, social conventions, decisions on industry standards, and political action.

examining the process by which conventions (Nash equilibria of coordination games) become established using evolutionary models with persistent randomness. In this framework, randomness ensures that all outcomes are likely, but as the amount of noise vanishes, some outcomes become more likely than others. Outcomes that retain a positive probability of being played in the long-run as the amount of noise vanishes are said to be *stochastically stable*.

A notable observation in the aforementioned papers is that stochastic stability as a selection mechanism not only rules out unstable mixed equilibria, but also selects a unique outcome in coordination games with two strategies. This outcome being the risk-dominant equilibrium, which is a strategy that is a unique best response to any distribution that places on it a mass of at least one half. However, recent studies show that the predictions of stochastic stability strictly depend on the interaction structure (Alós-Ferrer and Weidenholzer, 2007; Peski et al., 2010). Even in  $2 \times 2$  coordination games, it is easy to construct an interaction structure where both strategies are stochastically stable.

This paper seeks to identify properties of the interaction structure that determine uniqueness of stochastically stable outcomes. We consider an evolutionary model of best response with mutations (BRM) in symmetric coordination games – a framework introduced by Young (1993) and Kandori et al. (1993) – where players interact locally through a social network. In this model, players repeatedly choose strategies that are best responses to the distribution of strategies in their *neighbourhoods* (i.e. a fixed subset of the population whose strategies a player observes). Additionally, players occasionally experiment (mutate) and choose strategies that are not best responses.

Our analysis builds on the properties of contagion dynamics in social networks, and in particular, on the notion of *contagion threshold* for unbounded networks introduced by Morris (2000).<sup>2</sup> We define the contagion threshold of a finite network as some number  $\alpha$  such that for all players at any distance from some “reference neighbourhood”, at least proportion  $\alpha$  of their neighbours are closer to the reference neighbourhood. Every *strongly connected* network (i.e. a network in which every two players are connected through some path) has a unique contagion threshold bounded from above by  $\frac{1}{2}$  and from below by  $\frac{1}{\Delta}$ , where  $\Delta$  is the size of the largest neighbourhood.

The contagion threshold of a network naturally relates to the notion of *p-best response/p-dominance*, which is a generalization of risk-dominance to multiple strategies. A set of strategies

---

<sup>2</sup>Specifically, in a  $2 \times 2$  coordination game, let  $p \in (0, 1)$  be a critical number such that strategy  $a_2$  is a best response whenever at least proportion  $p$  of a player’s neighbours choose  $a_2$ . Morris (2000) defines the contagion threshold of a network as a number  $q$  such that strategy  $a_2$  spreads contagiously to the entire network starting from a finite small group whenever  $p \leq q$ .

is a *p*-best response set if the best responses to distributions that place on them a mass of at least *p* are themselves within that set (Tercieux, 2006). Equivalently, if proportion *p* of a player’s neighbours play strategies in a *p*-best response set, then that player’s best responses are themselves within that set. If a *p*-best response set contains only one strategy, then that strategy is a *p*-dominant strategy (Morris et al., 1995). In a coordination game with only two strategies, a  $\frac{1}{2}$ -dominant strategy is also a risk-dominant strategy.

The connection between contagion and *p*-best response arises from the fact that if *p* is less or equal to the contagion threshold of a network, then strategies in a *p*-best response set can spread contagiously from any *first-neighbourhood* to the entire or large proportion of that network. A first-neighbourhood of any player *i* is the set of all direct neighbours of *i*, with *i* included. Relying on this connection, we show that: *a*) the upper bound for the cost, measured in terms of the number of mutations, of reaching any state containing only strategies in a *p*-best response set is the size of the smallest first-neighbourhood; *b*) the cost of leaving the basin of attraction of an absorbing set of states containing only strategies in a *p*-best response set is greater than the size of the first-neighbourhood of any player. Invoking Ellison (2000, Theorems 2),<sup>3</sup> we then show that an absorbing set of states containing only strategies in a *p*-best response set is uniquely stochastically stable.

More specifically, our results demonstrate that if a network is strongly connected, with a diameter<sup>4</sup> greater or equal to 5, then a *p*-best response set is uniquely stochastically stable whenever *p* is less or equal to the contagion threshold. The diameter of a network captures its level of *centralization* and *density*. Centralized networks exhibit a *core-periphery* structure where a subset of densely interconnected players form the core of the network, and the rest who are adjacent to the core but not to themselves form the periphery. Highly centralized networks have a short diameter (e.g. a star network with diameter of 2). Networks with a low density – the total number of connections relative to the number of players – have a large diameter. This, together with the fact that the contagion threshold has a lower bound of  $\frac{1}{\Delta}$ , imply that our results have strong predictability in networks that are not highly centralized, and are sufficiently large and/or localized.

This paper contributes to the literature on long-run stability of evolutionary dynamics in networks.<sup>5</sup> Examples include Ellison (1993, 2000) and Weidenholzer (2012), who show that

---

<sup>3</sup>Ellison (2000, Theorems 2) states that if the cost of leaving a set of states is larger than the cost of reaching it, then that set is uniquely stochastically stable.

<sup>4</sup>The diameter of the network is the maximum shortest distance between any two players.

<sup>5</sup>There are many papers that also study alternative classes of noisy best response choice rules to the BRM

$\frac{1}{2}$ -dominant strategies are uniquely stochastically stable in regular ring-networks where each player has  $2k$  neighbours; [Blume \(1995\)](#), [Young \(1998\)](#), [Lee and Valentinyi \(2000\)](#) and [Lee et al. \(2003\)](#) show that risk-dominant strategies (for  $2 \times 2$  symmetric coordination games) are uniquely stochastically stable in 2-dimensional grid lattices; [Anderlini and Ianni \(1996\)](#) show that any or both strategies can be selected in an evolutionary model where mutations are triggered by change; and [Alós-Ferrer and Weidenholzer \(2007\)](#) show that a globally pairwise risk-dominant strategy of a  $3 \times 3$  symmetric coordination game is stochastically stable in a cyclic network where each player has two neighbours. In contrast to our paper, all these papers focus on specific (mostly regular) networks.

Our results are more closely related to [Alós-Ferrer and Weidenholzer \(2008\)](#) and [Peski et al. \(2010\)](#). Although [Alós-Ferrer and Weidenholzer \(2008\)](#) study evolutionary dynamics where players imitate strategies that yield the highest payoff (i.e. “imitate-the-best”), they present an example where best response dynamics leads to the selection of a  $\frac{1}{\Delta}$ -dominant strategy.<sup>6</sup> Our results generalize this example from [Alós-Ferrer and Weidenholzer \(2008\)](#) because the lower bound for the contagion threshold of any network is  $\frac{1}{\Delta}$ .

[Peski et al. \(2010\)](#) identifies a network parameter that determines conditions under which a  $p$ -dominant strategy is uniquely stochastically stable. For a given network  $G$ , let  $\delta_0(G)$  be the size of the smallest odd neighbourhood. [Peski et al. \(2010, Theorem 2 & Corollary 1\)](#) shows that in a BRM model, a  $p$ -dominant strategy is uniquely stochastically stable in network  $G$  if  $p$  is less or equal to  $\frac{1}{2} \left( 1 - \frac{1}{\delta_0(G)} \right)$ . Thus, [Peski et al. \(2010\)](#) identifies  $\delta_0(G)$  as one of the network parameters that determines uniqueness of stochastically stable outcomes.<sup>7</sup>

There are then two main differences between our results and [Peski et al. \(2010\)](#). First, our results have *stronger predictability* in sparsely connected networks (i.e. where the size of

---

model considered in this paper ([Maruta, 2002](#); [Myatt and Wallace, 2003](#); [Sandholm, 2010](#); [Staudigl, 2012](#); [Staudigl and Weidenholzer, 2014](#); [Newton and Sawa, 2015](#); [Klaus and Newton, 2016](#); [Sandholm and Staudigl, 2016](#); [Sawa and Wu, 2018](#)). These papers however focus on global interactions, which makes comparison to our findings non-trivial.

<sup>6</sup>However, the main result of [Alós-Ferrer and Weidenholzer \(2008\)](#) is that imitation dynamics selects a payoff dominant strategy in  $2 \times 2$  coordination games rather than a risk-dominant that is selected by best response dynamics. Related papers in this regard include [Alós-Ferrer and Weidenholzer \(2006\)](#) and [Chen et al. \(2013\)](#).

<sup>7</sup>[Sandholm \(2001\)](#) and [Oyama et al. \(2015\)](#) also show that  $\frac{1}{k}$ -dominant strategies are stable in the long-run. However, they consider a deterministic evolutionary model where players choose strategies that are best responses to the empirical distribution of  $k$  others sampled randomly from an infinitely large population. Their model thus calls for a different solution concept – *almost global convergence* – which is not immediately comparable to stochastic stability.

the largest neighbourhood is small so that the lower bound for the contagion threshold,  $\frac{1}{\Delta}$ , is large) while [Peski et al. \(2010, Theorem 2 & Corollary 1\)](#) has stronger predictability in highly connected networks. By strong predictability we mean  $p$  is as close to  $\frac{1}{2}$  as possible for a given network. If an unweighted network has at least one player with only one neighbour so that  $\delta_0(G) = 1$ , then [Peski et al. \(2010, Theorem 2 & Corollary 1\)](#) completely loses predictability. However, since the lower bound for the contagion threshold is  $\frac{1}{\Delta} > 0$ , Theorem 1 of this paper predicts that some  $p$ -dominant strategy, for  $p \geq \frac{1}{\Delta}$  is always stochastically stable in any strongly connected network. Second, we derive conditions under which a  $p$ -best response set, and not just a  $p$ -dominant strategy, is uniquely stochastically stable. Not every coordination game has a  $p$ -dominant strategy but every coordination game has a  $p$ -best response set with  $p$  less or equal to one half.

The remainder of the paper is organized as follows. Section 2 outlines a model of evolutionary dynamics with best response and mutations. Section 3 defines the notion of contagion and contagion threshold, and states the main results. Section 4 discusses the implications of our results and how they relate to the literature, in particular, to [Peski et al. \(2010\)](#). Concluding remarks are offered in Section 5 and proofs are contained in the Appendix.

## 2. An evolutionary model of best response with mutations

### 2.1. The coordination game

We consider  $m$ -strategy 2-player symmetric coordination games with a set of pure strategies denoted by  $A = \{a_1, \dots, a_j, \dots, a_m\}$ . Let  $U$  be an  $m \times m$  payoff matrix so that  $u(a_j, a_l)$  is the payoff to a player playing strategy  $a_j$  against an opponent playing strategy  $a_l$ . A coordination game is characterized by all pure strategies being strict Nash equilibria. That is,  $u(a_j, a_j) > u(a_j, a_l)$  for all  $a_j, a_l \in A$  and  $a_l \neq a_j$ . We focus on symmetric coordination games where  $u(a_j, a_l)$ , for all  $a_j, a_l \in A$ , are identical for both players (i.e. the row and column players). We refer to the double  $(A, U)$  as a *symmetric strict coordination game*.

Let  $\Sigma$  be the set of all mixed strategies over  $A$ , so that for any  $a_j \in A$  and  $\sigma \in \Sigma$ ,  $\sigma(a_j)$  is the weight or probability  $\sigma$  assigns to  $a_j$ . The corresponding payoff to a player playing strategy  $a_j$  against a mixture  $\sigma$  is  $U(a_j \mid \sigma) = \sum_{a_k \in A} \sigma(a_k)u(a_j, a_k)$ . The set of pure strategy best responses to  $\sigma$  is defined as

$$BR(\sigma) = \{a_j \in A \mid U(a_j \mid \sigma) \geq U(a_l \mid \sigma) \forall a_l \in A\}.$$

## 2.2. The social interaction structure

We consider a finite set  $N = \{1, \dots, i, \dots, n\}$  of players interacting through a *social network*. The social network is represented by a graph  $G(N, E)$ , with  $N$  representing the set of players and  $E$  the set of edges linking different pairs of players. In  $G(N, E)$ , each player is identified by her *neighbourhood* – a fixed subset  $N_i \subset N$  – with whom  $i$  directly interacts with. We write  $n_i$  for the cardinality of  $N_i$ .

We use the following definitions throughout the paper. Given  $G(N, E)$ , a path from player  $i$  to  $j$  is a set of players  $i = i_1, i_2, \dots, i_r = j$  with links  $i_1 \rightarrow i_2, i_2 \rightarrow i_3, \dots, i_{r-1} \rightarrow j$  without cycles. We write  $d_{ij}$  for the length of the shortest path from  $i$  to  $j$ . Letting  $d_i = \min_{j \in N} d_{ij}$ , the diameter of network  $G(N, E)$  is defined as  $d(G) = \max_{i \in N} d_i$ . We place the following restrictions on the network structure.

- (i)  $G(N, E)$  is *unweighted* and *undirected*: a link between any pair of agents  $i$  and  $j$  has a weight of one, and the existence of a link from  $i$  to  $j$  implies existence of a reverse link from  $j$  to  $i$ .
- (ii)  $G(N, E)$  is *strongly connected*: there exists a path of links connecting any pair of players  $i, j \in N$  and  $i \neq j$ .

Assumptions (i) and (ii) help simplify our analysis and avoids complexities that arise from heterogeneous weights and disconnected network structures. Networks that are not strongly connected may consist of disconnected subgroups, which raises complications regarding conditions under which an evolutionary process converges. We avoid such complications in order to focus on the main question regarding the structural properties of the network that determine uniqueness of stochastically stable outcomes. We further discuss the implications of each of these assumptions on our results in Section 4.

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

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

Note however that the total payoff is independent of a player's identity (the size of a player's neighbourhood) and depends only on the empirical distribution  $\sigma_i$ . If any two players have an identical empirical distribution  $\sigma$ , then their total payoffs are also identical. So, where no confusion arises, we simply write  $U(a_j \mid \sigma)$  for the total payoff to any player playing strategy  $a_j$  against distribution  $\sigma$ . We then refer to the quadruple  $(A, U, N, G)$  as a *symmetric strict coordination game on network  $G$* .

### 2.3. Unperturbed learning process

We consider a *myopic best response* dynamic process where players simultaneously play a symmetric strict coordination game against their neighbours at discrete times,  $t = 1, 2, \dots$ . We assume that  $G(N, E)$  is exogenously given and fixed over time, so that each player repeatedly plays against the same set of neighbours. At each period, a player selects a pure strategy from a set of strategies that maximize (1).

Let  $\mathbf{x} = (x^1, \dots, x^n)$  denote a profile/configuration of strategies, where  $x^i$  is the strategy played by the  $i$ th player. Each strategy profile is a state of an evolutionary process, and we denote the set of all states by  $\mathbf{X}$ . For each  $\mathbf{x}$ , let  $\sigma_i(a_l; \mathbf{x})$  be the proportion of  $i$ 's neighbours playing strategy  $a_l$  in profile  $\mathbf{x}$ , and let  $\sigma_i(\mathbf{x}) = (\sigma_i(a_1; \mathbf{x}), \dots, \sigma_i(a_m; \mathbf{x}))$  be the empirical distribution over strategies in  $i$ 's neighbourhood under strategy profile  $\mathbf{x}$ . In analogy to (1),  $U(a_j | \sigma_i(\mathbf{x})) = \sum_{a_l \in A} \sigma_i(a_l; \mathbf{x}) u(a_j, a_l)$  is the payoff to  $i$  for adopting strategy  $a_j$  against distribution  $\sigma_i(\mathbf{x})$ .

Let  $\mathbf{x}(t)$  be the strategy profile at  $t$ , and  $x^i(t)$  the respective  $i$ th strategy in profile  $\mathbf{x}(t)$ . By myopic best response we mean, at  $t + 1$ , each player chooses a strategy that is a best response to the strategy profile at  $t$ . That is,  $x^i(t + 1) \in BR(\sigma_i(\mathbf{x}(t)))$ , and the probability that player  $i$  plays strategy  $a_j$  at  $t + 1$  is  $P(x^i(t + 1) = a_j | \sigma_i(\mathbf{x}(t))) = \frac{1}{|BR(\sigma_i(\mathbf{x}(t)))|}$  if  $a_j \in BR(\sigma_i(\mathbf{x}(t)))$ , and zero otherwise, where  $|S|$  is the cardinality of set  $S$ .

The assumption of myopia is standard in the literature of evolutionary game theory and it is used to model agents with bounded memory and bounded rationality. That is, it assumes that agents are incapable of keeping track of the entire history of play and performing complex evaluations associated with forward-looking decision making.

The dynamic framework described above leads to a finite Markov chain on the state space  $\mathbf{X}$ . Let  $P$  denote the transition probability matrix so that  $P(\mathbf{x}, \mathbf{y})$  is the probability of transiting from state  $\mathbf{x}$  to  $\mathbf{y}$  in a single period. A subset  $W \subseteq \mathbf{X}$  of states 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  $P(\mathbf{y}, \mathbf{x}) = 1$ . We denote by  $\mathbf{A}$  a set of all absorbing sets. We refer to a tuple  $(A, U, N, G, P)$  as an *evolutionary process of best response* on network  $G$ .

### 2.4. Perturbed learning process and stochastic stability

Following the literature, a *perturbed process* is derived from the unperturbed process described above by adding rare mutations. We particularly follow the evolutionary framework



of best response with mutations (BRM) by [Young \(1993\)](#) and [Ellison \(2000\)](#), whereby, with a small fixed probability  $\varepsilon > 0$ , a player chooses a strategy at random, and with probability  $(1 - \varepsilon)$ , plays a strategy in the best response set. That is, the probability that  $i$  plays strategy  $a_j$  at  $t + 1$  is  $P_\varepsilon(x^i(t + 1) = a_j | \sigma_i(\mathbf{x}(t))) = \frac{\varepsilon}{m} + \frac{1-\varepsilon}{|BR(\sigma_i(\mathbf{x}(t)))|}$  if  $a_j \in BR(\sigma_i(\mathbf{x}(t)))$ , and  $P_\varepsilon(x^i(t + 1) = a_j | \sigma_i(\mathbf{x}(t))) = \frac{\varepsilon}{m}$  if  $a_j \notin BR(\sigma_i(\mathbf{x}(t)))$ .

We write  $P_\varepsilon(\mathbf{x}, \mathbf{y})$  for the probability that profile  $\mathbf{x}$  is followed by  $\mathbf{y}$ , and  $P_\varepsilon$  for the transition probability matrix of the associated Markov chain. Specifically, let  $c(\mathbf{x}, \mathbf{y})$  be the number of players for whom  $y^i$ , the  $i$ th strategy in  $\mathbf{y}$ , is not a best response to  $\mathbf{x}$  (i.e. the number of mutations involved in a direct transition from  $\mathbf{x}$  to  $\mathbf{y}$ ). Then  $P_\varepsilon(\mathbf{x}, \mathbf{y})$  can be expressed as follows, where we use the notation  $b_i(\mathbf{x}) = |BR(\sigma_i(\mathbf{x}))|$ .<sup>8</sup>

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

Mutations ensure that every state is reached with a positive probability from every other state. The perturbed process is thus ergodic in that it has only one absorbing set (the entire state space), which also implies that it has a unique *stationary distribution*  $\pi_\varepsilon$  – a probability distribution over the state space that describes the long-run average time spent in each state. For each  $\mathbf{x} \in \mathbf{X}$ ,  $\pi_\varepsilon(\mathbf{x})$  is the proportion of time the perturbed process spends in state  $\mathbf{x}$ .

As the amount of noise tends to zero, the stationary distribution of the perturbed process tends to a limit distribution  $\pi^* = \lim_{\varepsilon \rightarrow 0} \pi_\varepsilon$  that places most weight on a subset of absorbing states. The set  $\mathbf{A}^* \subseteq \mathbf{A}$  of absorbing states for which  $\lim_{\varepsilon \rightarrow 0} \pi_\varepsilon(\mathbf{A}^*) > 0$  is said to be *stochastically stable* ([Foster and Young, 1990](#); [Young, 1993](#)). If  $\mathbf{A}^* = \{\mathbf{x}^*\}$ , then  $\mathbf{x}^*$  is a stochastically stable state/strategy profile. Accordingly, a set of strategies  $A^* \subseteq A$  played in  $\mathbf{A}^*$  are said to be stochastically stable strategies.

To identify a stochastically stable set, we employ a method developed by [Ellison \(2000\)](#), which involves computing *modified coradii* and *radii of basins of attraction* of sets of states. The basin of attraction  $D(W)$  of a subset  $W \subseteq \mathbf{A}$ , is the set of initial states from which the unperturbed Markov chain  $(A, U, N, G, P)$  converges to  $W$  with probability one. That is,

$$D(W) = \left\{ \mathbf{y} \in \mathbf{X} \mid P\left(\exists t' \text{ such that } \mathbf{x}(t) \in W \forall t > t' \mid \mathbf{x}(0) = \mathbf{y}\right) = 1 \right\}$$

The radius of a basin of attraction of  $W$ ,  $R(W)$ , is the minimum number of mutations necessary to exit  $D(W)$  when an evolutionary process starts from  $W$ . Formally, define a *path*

---

<sup>8</sup>This follows from the expressions of  $P_\varepsilon(x^i(t + 1) = a_j | \sigma_i(\mathbf{x}(t)))$ , whereby, the probability that  $c(\mathbf{x}, \mathbf{y})$  players mutate to play non-best response strategies is  $(\frac{\varepsilon}{m})^{c(\mathbf{x}, \mathbf{y})}$ . And the probability that the remaining  $n - c(\mathbf{x}, \mathbf{y})$  players simultaneously play their best responses is  $\prod_{i=1}^{n-c(\mathbf{x}, \mathbf{y})} \left(\frac{\varepsilon}{m} + (1 - \varepsilon)\frac{1}{b_i(\mathbf{x})}\right) = \prod_{i=1}^{n-c(\mathbf{x}, \mathbf{y})} \left(\frac{m + (b_i(\mathbf{x}) - m)\varepsilon}{mb_i(\mathbf{x})}\right)$



from a subset  $W$  to  $Z$  as a finite sequence of distinct states  $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T)$  with  $\mathbf{x}_1 \in W$ ,  $\mathbf{x}_T \in Z$  and  $\mathbf{x}_\tau \notin Z$  for  $2 \leq \tau \leq T-1$ . Let  $S(W, Z)$  be the set of all paths from  $W$  to  $Z$ . The cost  $c(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T)$  of the path  $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T)$  is defined as

$$c(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T) = \sum_{\tau=1}^{T-1} c(\mathbf{x}_\tau, \mathbf{x}_{\tau+1})$$

where  $c(\mathbf{x}_\tau, \mathbf{x}_{\tau+1})$  is defined in (2). Let  $C(W, Z) = \min_{(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T) \in S(W, Z)} c(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T)$ . Then the radius  $R(W)$  of  $D(W)$  is defined as:

$$R(W) = \min_{(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T) \in S(W, \mathbf{X} - D(W))} c(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T) = C(W, \mathbf{X} - D(W))$$

Let  $Z_1, Z_2, \dots, Z_q$ , with  $Z_q \in W$ , be a sequence of absorbing sets through which path  $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T)$  from  $\mathbf{x}$  to  $W$  traverses consecutively. Here,  $\mathbf{x}_1$  can be an element of  $Z_1$  and  $Z_l \notin W$  for  $l < q$ . The modified cost  $c^*(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T)$  of path  $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T)$  is defined as:

$$c^*(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T) = c(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T) - \sum_{l=2}^{q-1} R(Z_l) \quad (3)$$

The minimum modified cost  $C^*(\mathbf{x}_1, W)$  of the least cost path from  $\mathbf{x}_1$  to subset  $W$  is defined as

$$C^*(\mathbf{x}, W) = \min_{(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T) \in S(\mathbf{x}, W)} c^*(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T) \quad (4)$$

The modified coradius  $CR^*(W)$  of  $W$  is then defined as  $CR^*(W) = \max_{\mathbf{x} \notin W} C^*(\mathbf{x}, W)$ . Following Ellison (2000, Theorems 2), if  $R(W) > CR^*(W)$ , then  $W$  is uniquely stochastically stable.

### 3. $p$ -best response and stochastic stability

We aim to establish the relationship between  $p$ -best response/ $p$ -dominance and stochastic stability of an evolutionary process  $(A, U, N, G, P_\varepsilon)$ . Specifically, we derive conditions under which  $p$ -best response sets are uniquely stochastically stable. A  $p$ -best response set is a set of strategies that are best responses to distributions that place on them a mass of at least  $p$ . For any nonempty subset  $A' \subseteq A$  and any  $\sigma \in \Sigma$ , let  $\sigma_{A'} = \sum_{a_j \in A'} \sigma(a_j)$  be the total mass that distribution  $\sigma$  places on  $A'$ .

**Definition 1.** A nonempty subset of strategies  $A' \subseteq A$  is a  $p$ -best response set of a symmetric strict coordination game  $(A, U)$  if for all  $\sigma \in \Sigma$  with  $\sigma_{A'} \geq p$ ,  $BR(\sigma) \subseteq A'$ .

Based on Definition 1, we can divide the strategy set  $A$  into two disjoint subsets,  $A^p$ , the  $p$ -best response set of  $(A, U)$ , and  $A^{1-p} = A \setminus A^p$ . If, for a given  $p \in (0, \frac{1}{2}]$ ,  $A^p = \{a\}$ , then  $a$  is a  $p$ -dominant strategy. In analogy to local interactions, if at least proportion  $p$  of a player's neighbours play strategies in  $A^p$ , then all of a player's best responses are themselves in  $A^p$ . Every

coordination game has a  $p$ -best response set for  $p \in (0, \frac{1}{2}]$  (Oyama et al., 2015, Proposition 2). The notion of  $p$ -best response exhibits monotonicity in that for  $p < q$ , a  $p$ -best response set of a game is also a  $q$ -best response set of that game.

Given  $A^p \subseteq A$ , let  $\mathbf{A}^p \subseteq \mathbf{A}$  be absorbing sets of states containing only strategies in  $A^p$ . Strategies in  $A^p$  are said to be uniquely stochastically stable if set  $\mathbf{A}^p$  is uniquely stochastically stable. We derive conditions under which  $\mathbf{A}^p$  is uniquely stochastically stable (i.e. under which  $R(\mathbf{A}^p) > CR^*(\mathbf{A}^p)$ ) by exploiting properties of the process of *contagion* in networks. A set of strategies, say  $A^p$ , is contagious in a given network if they can spread through best response to the entire or large proportion of the network starting from a small group. We rely on the implications of this process to place an upper bound on  $CR^*(\mathbf{A}^p)$  and a lower bound on  $R(\mathbf{A}^p)$ .

Morris (2000) defines the *contagion threshold* as a property of the network that ensures that a given strategy (in a  $2 \times 2$  coordination game) can spread contagiously in an *unbounded network* starting from any *finite group* of players. Analogously, we define the contagion threshold for a finite network as some number  $\alpha$  such that for all players at any distance from some *reference neighbourhood*, at least proportion  $\alpha$  of their neighbours are closer to the reference neighbourhood. As it becomes clear below, this definition ensures that contagion can be triggered from “the neighbourhood of any player”, which contrasts with “any finite group of players in an unbounded network”.

Formally, for each  $i \in N$ , let  $N_i(S) = N_i \cap S$  be the set of  $i$ ’s neighbours belonging to subgroup  $S$ , and let  $n_i(S)$  be the corresponding cardinality of  $N_i(S)$ . Let  $\alpha_i(S) = \frac{n_i(S)}{n_i}$  be the proportion of  $i$ ’s neighbours in  $S$ . For each  $i \in N$ , let  $B_{i_r}$  be the  $r$ th *neighbourhood* of  $i$  (i.e. the set of all players within distance  $r$  from  $i$ , with  $i$  included) and  $N_{i_r}$  be the  $r$ th-order neighbours of  $i$  (i.e. all players at distance  $r$  from  $i$ ); we write  $b_{i_r}$  and  $n_{i_r}$  for the respective cardinalities of  $B_{i_r}$  and  $N_{i_r}$ . We then define the contagion threshold of a finite network as follows.

**Definition 2.** Given  $G(N, E)$ :

- (i) pick any  $i \in N$  and the corresponding  $B_{i_1}$ ;
- (ii) for  $r = 2, 3, \dots, d_i$  and  $j \in N_{i_r}$ , compute  $\alpha_j(B_{i_{r-1}})$ ,  $\alpha_{i_r} = \min_{j \in N_{i_r}} \alpha_j(B_{i_{r-1}})$  and  $\alpha_i = \min_{r \geq 2} \alpha_{i_r}$ ;
- (iii) compute  $\bar{\alpha}_{i_2} = \max \{ \alpha_j(B_{i_1}) : \forall j \in N_{i_2} \text{ and } \alpha_j(B_{i_1}) < 1 \}$ .

The contagion threshold  $\underline{\eta}(G)$  is defined as  $\underline{\eta}(G) = \min_{i \in N} \alpha_i$ ; and the maximum contagion threshold at distance  $r = 2$  from any player  $i \in N$  is defined as  $\bar{\eta}_2(G) = \max_{i \in N} \bar{\alpha}_{i_2}$ .

We see from Definition 2 that the “reference neighbourhood” from which strategies in  $A^p$  can spread contagiously is  $B_{i_1}$  for any  $i \in N$ . Specifically, when  $p \leq \underline{\eta}(G)$ , it is possible for strategies

in  $A^p$  to spread through best response from some  $B_{i_1}$  to the entire or large proportion of the network. Note, however, that this condition does not immediately rule out the possibility of the contagious spread of strategies in  $A^{1-p}$ . We derive additional conditions that ensure that only set  $A^p$  is contagious when  $p \leq \underline{\eta}(G)$ , and hence, conditions under which  $R(\mathbf{A}^p)$  is greater than  $CR^*(\mathbf{A}^p)$ .

The algorithmic steps for computing the contagion threshold of a network directly follow from Definition 2. We provide matrix-based algorithmic steps which can be implemented in polynomial time (see Appendix [Appendix A](#)). The following theorem states the conditions under which a  $p$ -best response set is uniquely stochastically stable in a given network.

**Theorem 1.** *Given an evolutionary process of best response with mutations,  $(A, U, N, P_\epsilon)$ , on an undirected, unweighted and strongly connected network  $G(N, E)$ , a  $p$ -best response set  $A^p \subseteq A$  is uniquely stochastically stable if:*

- (i)  $p \leq \underline{\eta}(G)$  and  $d(G) \geq 7$ , or
- (ii)  $p \leq \min\{\underline{\eta}(G), 1 - \bar{\eta}_2(G)\}$  and  $5 \leq d(G) \leq 6$ .

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

□

The first part of the proof of Theorem 1 establishes an upper bound for the modified coradius of  $\mathbf{A}^p$ . Given an undirected, unweighted and strongly connected network  $G(N, E)$  with  $p \leq \underline{\eta}(G)$  and  $d(G) \geq 5$ , we first show that starting from any configuration  $\mathbf{x} \in \mathbf{X} \setminus \mathbf{A}^p$ , if all players in  $B_{i_1}$ , for any  $i \in N$ , mutate to strategies in  $A^p$ , then these strategies will spread contagiously to the entire or large proportion of the network. Specifically,  $\underline{b}_1 = \min_{i \in N} b_{i_1}$  mutations to strategies in  $A^p$  ensure that the evolutionary process converges to either an absorbing state/cycle containing only strategies in  $A^p$  or to an absorbing cycle containing both strategies in  $A^p$  and  $A^{1-p}$ .

If, after  $\underline{b}_1$  mutations to strategies in  $A^p$  starting from any  $\mathbf{x} \in \mathbf{X} \setminus \mathbf{A}^p$ , the evolutionary process converges to an absorbing state/cycle in  $\mathbf{A}^p$ , then the cost  $C(\mathbf{x}, \mathbf{A}^p) \leq \underline{b}_1$ . If it instead converges to some absorbing cycle  $W$  containing both strategies in  $A^p$  and  $A^{1-p}$ , then  $C(\mathbf{x}, W) \leq \underline{b}_1$ . We then show that the radius of the basin of attraction of  $W$  is equal to the number of mutations needed to trigger evolution from  $W$  to some configuration in  $\mathbf{A}^p$ ; that is,  $R(W) = C(W, \mathbf{A}^p)$ . This is true for any  $W \subseteq \mathcal{C}(\mathbf{A})$ , where  $\mathcal{C}(\mathbf{A})$  is the set of all absorbing cycles containing both strategies in  $A^p$  and  $A^{1-p}$ .

Recall the definition of the modified coradius of  $\mathbf{A}^p$  as  $CR^*(\mathbf{A}^p) = \max_{\mathbf{x} \in \mathbf{X} \setminus \mathbf{A}^p} C^*(\mathbf{x}, \mathbf{A}^p)$ , where  $C^*(\mathbf{x}, \mathbf{A}^p) = C(\mathbf{x}, \mathbf{A}^p) \leq \underline{b}_1$  in a scenario where the evolutionary process converges to

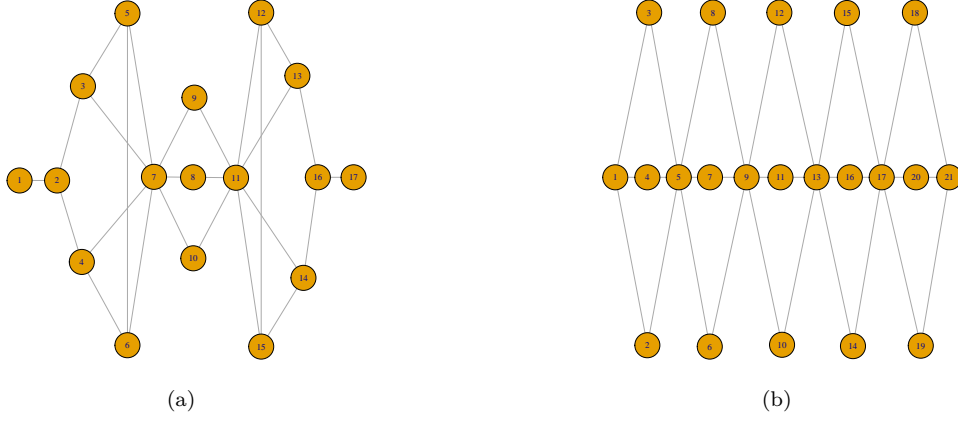


Figure 1: (a) An example of a network with contagion threshold of  $\frac{2}{7}$  and  $d(G) = 8$ . (b) A network with contagion threshold of  $\frac{1}{2}$  and  $d(G) = 10$ .

some configuration in  $\mathbf{A}^p$  after  $\underline{b}_1$  mutations. However, if it converges to some configuration in  $W \subseteq \mathcal{C}(\mathbf{A})$ , then

$$\begin{aligned}
C^*(\mathbf{x}, \mathbf{A}^p) &= \min_{(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T) \in S(\mathbf{x}, \mathbf{A}^p)} c^*(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T) \\
&= \min_{(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{T'}) \in S(\mathbf{x}, W)} c(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{T'}) \\
&\quad + \min_{(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T) \in S(W, \mathbf{A}^p)} c(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T) - R(W)
\end{aligned} \tag{5}$$

Since  $C(W, \mathbf{A}^p) = \min_{(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T) \in S(W, \mathbf{A}^p)} c(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T) = R(W)$ , we see that  $C^*(\mathbf{x}, \mathbf{A}^p) = \min_{(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{T'}) \in S(\mathbf{x}, W)} c(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{T'}) = C(\mathbf{x}, W) \leq \underline{b}_1$ . Thus, the modified coradius of  $\mathbf{A}^p$  is bounded from above by  $CR^*(\mathbf{A}^p) \leq \underline{b}_1$ .

**Example:** To further elaborate this point, consider the two networks depicted in Figures 1a and 1b. The contagion threshold for the network in Figure 1a (computed using the matrix-based algorithm in Appendix A) is  $\underline{\eta}(G) = \frac{2}{7}$  and the diameter is  $d(G) = 8$ . The players with the smallest first-neighbourhood are 1 and 17, with  $b_{1_1} = b_{17_1} = \underline{b}_1 = 2$ . Let  $p \leq \frac{2}{7}$ . Assume that the process  $(A, U, N, G, P)$  starts from some configuration  $\mathbf{x} \in \mathbf{X} \setminus \mathbf{A}^p$  where all players play strategies in  $A^{1-p}$ . Let players in  $B_{1_1} = \{1, 2\}$  both mutate to strategies in  $A^p$  at  $t = 1$ ; then  $(A, U, N, G, P)$  will evolve as follows:

$t = 1 \mid$  All players in  $B_{1_1} = \{1, 2\}$  play strategies in  $A^p$ ; the rest play strategies in  $A^{1-p}$ .

$t = 2$	Player 1 plays a strategy in $A^p$ since $\alpha_1(B_{1_1}) > p$ ; a player in $N_{1_1} = \{2\}$ plays a strategy in $A^p$ because $\alpha_2(B_{1_1}) = \frac{1}{3} > p$ (the last inequality does not generally hold for all networks); players in $N_{1_2} = \{3, 4\}$ also all play strategies in $A^p$ because each has proportion $\frac{1}{3} > \frac{2}{7} \geq p$ of neighbours play strategies in $A^p$ at $t = 1$ . Note that for each $j \in N_{1_2}$ , since $p \leq \underline{\eta}(G) \leq \alpha_1 \leq \alpha_j(B_{1_1})$ , we have $(1 - p) \geq (1 - \underline{\eta}(G)) \geq (1 - \alpha_1) \geq (1 - \alpha_j(B_{1_1})) \geq \alpha_j(N_{1_3})$ . This implies that strategies in $A^{-p}$ are not best responses for all $j \in N_{1_2}$ since they are best responses only when played by more than proportion $1 - p$ of neighbours. All other players play strategies in $A^{1-p}$ .
$t = 3$	All players in $B_{1_3} = \{1, 2, 3, 4, 5, 6, 7\}$ play strategies in $A^p$ since each has at least proportion $\frac{2}{7} \geq p$ of neighbours play strategies in $A^p$ at $t = 2$ ; all other players play strategies in $A^{1-p}$ .
---	-----
$t = 8$	All players in $B_{1_8} = N = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17\}$ play strategies in $A^p$ .

Following similar steps, for the network in Figure 1b, if all players in  $B_{4_1} = \{1, 4, 5\}$  mutate to strategies in  $A^p$  at  $t = 1$ ; then for  $p \leq \frac{1}{2}$ , the process  $(A, U, N, G, P)$  evolves as follows:

$t = 2$	All players in $4 \cup N_{4_2} = 4 \cup \{2, 3, 6, 7, 8\}$ play strategies in $A^p$ since each has at least proportion $\frac{1}{2} > p$ of neighbours in $B_{4_1}$ ; players in $N_{4_1} = \{1, 5\}$ switch back to strategies in $A^{1-p}$ because $\alpha_j(B_{4_1}) = \frac{1}{3} < \frac{1}{2} \geq p$ ; all other players play strategies in $A^{1-p}$ .
$t = 3$	All players in $4 \cup N_{4_2} = 4 \cup \{2, 3, 6, 7, 8\}$ switch back to strategies in $A^{1-p}$ ; players in $N_{4_1}$ switch to strategies in $A^p$ ; a player in $N_{4_3} = \{9\}$ plays a strategies in $A^p$ ; and all other players play strategies in $A^{1-p}$ .
$t = 4$	All players in $4 \cup N_{4_2} = 4 \cup \{2, , 3, 6, 7, 8\}$ play strategies in $A^p$ ; players in $N_{4_1} = \{1, 5\}$ switch back to strategies in $A^{1-p}$ ; player $N_{4_3} = \{9\}$ switches back to a strategy in $A^{1-p}$ ; players in $N_{4_4} = 4 \cup \{10, 11, 12\}$ all play strategies in $A^p$ ; and all other players play strategies in $A^{1-p}$ .
---	-----
$t = 9$	All players in $4 \cup N_{4_2} = 4 \cup \{2, 3, 6, 7, 8\}$ play strategies in $A^{1-p}$ ; players in $N_{4_1} = \{1, 5\}$ play strategies in $A^p$ ; player $N_{4_3} = \{9\}$ plays a strategy in $A^p$ ; players in $N_{4_4} = \{10, 11, 12\}$ play strategies in $A^{1-p}$ ; player $N_{4_5} = \{13\}$ plays a strategies in $A^p$ ; players in $N_{4_6} = \{14, 15, 16\}$ play strategies in $A^{1-p}$ ; player $N_{4_7} = \{17\}$ plays a strategy in $A^p$ ; players in $N_{4_8} = \{18, 19, 20\}$ play strategies in $A^{1-p}$ ; and player $N_{4_9} = \{21\}$ plays a strategy in $A^p$ .

Thus, for the network in Figure 1a,  $b_1$  mutations sufficiently trigger evolution from configuration  $\mathbf{x} \in \mathbf{X} \setminus \mathbf{A}^p$  containing only strategies in  $A^{1-p}$  to some configuration in  $\mathbf{A}^p$ . The modified coradius of  $\mathbf{A}^p$  in such a network is then equal to the cost  $C(\mathbf{x}, \mathbf{A}^p) \leq b_1$ .<sup>9</sup> However, for the network in Figure 1b, the evolutionary process converges to an absorbing cycle  $W \subseteq \mathcal{C}(\mathbf{A})$  where players in  $N_{i_\tau}$  and  $N_{i_{\tau+1}}$  alternate between strategies in  $A^p$  and  $A^{1-p}$ . Mutations by all players in any  $N_{i_\tau}$  is sufficient to trigger an exit from the basin of attraction of  $W$  to some configuration in  $\mathbf{A}^p$ . For example, if all players in  $N_{4_2}$  mutate to strategies in  $A^p$  at  $t = 9$ , then all players in  $N_{4_1} \cup N_{4_2} \cup N_{4_3}$  will play strategies in  $A^p$ . The evolutionary process then evolves towards  $\mathbf{A}^p$  from  $t = 10$  onward so that  $C(W, \mathbf{A}^p) \leq |N_{4_2}|$ . Note that mutations by all players in  $N_{4_2}$  to strategies in  $A^{1-p}$  at  $t = 10$  is not sufficient to trigger an exit from the basin of attraction of  $W$ , and hence,  $R(W) = C(W, \mathbf{A}^p)$ . This equality implies that  $CR^*(\mathbf{A}^p) = C(\mathbf{x}, W) \leq b_1$ .

The second part of the proof of Theorem 1 establishes a lower bound for the radius of  $D(\mathbf{A}^p)$  (i.e the basin of attraction of  $\mathbf{A}^p$ ). The number of mutations needed to trigger an exit from  $D(\mathbf{A}^p)$  depends on the diameter of the network. Consider the network in Figure 2 with  $p \leq \underline{\eta}(G) = \frac{2}{5}$  and  $d(G) = 3$ . Starting from any  $\mathbf{x} \in \mathbf{A}^p$ , let players 4 and 5 simultaneously mutate to strategies in  $A^{1-p}$  at  $t = 1$ . Then at  $t = 2$ , all players in  $\{2, 3, 6, 7, 8\}$  switch to strategies in  $A^{1-p}$  because each has at least proportion  $\frac{2}{3} > \frac{3}{5} = (1 - p)$  of neighbours play strategies in  $A^{1-p}$  at  $t = 1$ . Still at  $t = 2$ , all players in  $\{1, 4, 5\}$  play strategies in  $A^p$ . From  $t = 3$  onward, these two sets of players (i.e.  $\{2, 3, 6, 7, 8\}$  and  $\{1, 4, 5\}$ ) alternate between strategies in  $A^p$  and  $A^{1-p}$ . Thus, two mutations are sufficient to trigger an exit from  $D(\mathbf{A}^p)$  to an absorbing cycle containing both strategies in  $A^p$  and  $A^{1-p}$ .

In this example, players 4 and 5 can trigger an exit from  $D(\mathbf{A}^p)$  because all other players are within two steps from them. However, if, for any  $i \in N$ ,  $d_i$ , the maximum shortest distance from  $i$  to every other player, is sufficiently large, then mutations to strategies in  $A^{1-p}$  by players in  $B_{i_1}$  would not be sufficient to trigger an exit from  $D(\mathbf{A}^p)$  (i.e. the evolutionary process would revert to some state in  $\mathbf{A}^p$ ). The intuition is as follows.

Consider a network with  $d_i \geq 4$  for all  $i \in N$ , which also implies that  $d(G) \geq 7$  (see the proof in Appendix B Claim 1). Let  $p \leq \underline{\eta}(G)$ . For any  $i \in N$  in such a network, there exists some  $j \in N_{i_4}$  for whom  $B_{j_2} \cap B_{i_1} = \emptyset$  (i.e.  $B_{j_2}$  and  $B_{i_1}$  are disjoint sets). Now, starting from some configuration  $\mathbf{x} \in \mathbf{A}^p$ , let all players within  $B_{i_1}$  mutate to strategies in  $A^{1-p}$  at  $t = 1$ . Since  $B_{j_2}$  is a disjoint set from  $B_{i_1}$ , all players in  $B_{j_2}$  play strategies in  $A^p$  at  $t = 1$ .

---

<sup>9</sup>Note that since this is true for any  $\mathbf{x}$  containing only strategies in  $A^{1-p}$ , it must be true for all other configurations containing both strategies in  $A^p$  and  $A^{1-p}$ .





#### 4. Discussion and relation to the literature

We now discuss the implications of each of the conditions in Theorem 1. The first condition requires the network to be undirected, unweighted and strongly connected. Our results immediately extend to directed networks that are unweighted and strongly connected. This is because the requirements we make in the proof of Theorem 1 regarding the diameter of an undirected network, in particular, Claim 1, also hold for directed networks. For example, for a directed, unweighted and strongly connected network  $G(N, E)$ , if  $d(G) \geq 5$ , then  $d_i \geq 3$  for all  $i \in N$ . Thus, the steps undertaken to derive an upper bound for the modified coradius of  $\mathbf{A}^p$  and the lower bound for  $D(\mathbf{A}^p)$  hold for directed, unweighted and strongly connected networks as well.

Theorem 1 can also be extended to weighted and strongly connected networks but only after accounting for link weights in the definitions of network diameter (and other distance measures) and contagion threshold. For example, one would have to redefine the shortest path from player  $i$  to  $j$  as the path whose links have the smallest sum of weights, and the shortest distance  $d_{ij}$  being the sum of all weights along such a path.

Unlike directedness and weightedness, requiring a network to be strongly connected is necessary for our results to hold. When a network is not strongly connected, there is no guarantee that strategies in  $A^p$  can spread contagiously from a neighbourhood  $B_{i_1}$  of every  $i \in N$ . Similarly, there is no guarantee that when  $d(G) \geq 5$ ,  $d_i \geq 3$  for all  $i \in N$ . The proof, and hence validity of our results, relies on these two conditions. Note, however, that although a strongly connected network is a necessary condition for our results to hold, this condition has strong implications only for directed networks. This is because if an undirected network is not strongly connected, then it can be divided into disjoint strongly connected components that can be analysed separately.

The second condition is a restriction on the diameter,  $d(G) \geq 5$ . This condition is sufficient but not necessary for Theorem 1 to hold. It is easy to construct examples of network with  $d(G) \leq 2$  where strategies in  $A^p$  are uniquely stochastically stable, provided all other conditions of Theorem 1 hold. The condition  $d(G) \geq 5$  rules out highly centralized networks (e.g. a star network for which all strategies are stochastically stable) and densely connected networks (e.g. a complete network where each player interacts with every other player). Overall, this restriction is not limiting since real-world networks (e.g. social media networks, collaboration networks in academia, etc.) are large and sufficiently localized that their diameters are larger than 5.

The last condition relates  $p$  to the contagion thresholds  $\underline{\eta}(G)$  and  $\bar{\eta}_2(G)$ . The inequalities in Theorem 1 (i.e.  $p \leq \underline{\eta}(G)$  and  $p \leq \min\{\underline{\eta}(G), 1 - \bar{\eta}_2(G)\}$ ) are sufficient but not necessary since one can construct an example where  $p > \underline{\eta}(G)$  but strategies in  $A^p$  are uniquely stochastically

stable. However, generalizing such examples is very challenging due to the complexity of evolutionary dynamics in networks. Our definition of the contagion threshold for finite networks retains properties of an analogous definition by Morris (2000) but for unbounded networks.<sup>11</sup> In particular, as in Morris (2000), our definition ensures that contagion can be triggered from any region of the network (i.e. neighbourhood of any  $i \in N$ ). However, computing the exact value of the contagion threshold for an unbounded network is an impossible task for non-regular networks. For this reason, the analysis in Morris (2000) is limited to regular networks and to deriving lower and upper bounds for the contagion threshold.

We can deduce from Definition 2 that when  $d(G) \geq 7$ , the lower bound for the contagion threshold is  $\underline{\eta}(G) \geq \frac{1}{\Delta(G)}$ , where  $\Delta(G) = \max_{i \in N} n_i$ , is the size of the largest neighbourhood in  $G$ , and the upper bound is  $\frac{1}{2}$ . To see why, notice that even if  $\alpha_h(B_{i_{\tau-1}}) \geq \frac{1}{2}$  and all  $h \in N_{i_\tau}$  for all  $\tau = 2, 3, \dots, d_i$ , there exists some  $j \in N_{i_r}$  for  $r \geq 4$  and  $r \geq \tau + 2$  for whom  $\underline{\eta}(G) \leq \alpha_j \leq \alpha_h(B_{j_{r-\tau-1}}) \leq (1 - \alpha_h(B_{i_{\tau-1}})) \leq \frac{1}{2}$  (this is elaborated in the proof of Claim 2 in Appendix Appendix B). Similarly, for networks with  $5 \leq d(G) \leq 6$ , we have  $\underline{\eta}(G) \geq \frac{1}{\Delta(G)}$  and  $\min\{\underline{\eta}(G), 1 - \bar{\eta}_2(G)\} \leq \frac{1}{2}$ . Thus, our results imply that  $p$ -best response sets, for  $\frac{1}{\Delta(G)} \leq p \leq \frac{1}{2}$ , are uniquely stochastically stable in any undirected, unweighted and strongly connected network whose diameter and contagion threshold satisfy the conditions in Theorem 1.

Our approach and results are related to Peski et al. (2010) but differ in two main respects. Peski et al. (2010, Theorem 2 & Corollary 1) states that under Kandori et al. (1993) model, which is equivalent to the BRM model studied here, a  $(\frac{1}{2} - \delta^*)$ -dominant strategy is uniquely stochastically stable under local interactions, where  $\delta^*$  is defined below. That is, let strategy  $a \in A$  be  $(\frac{1}{2} - \delta^*)$ -dominant, and for any pair  $i, j \in N$ , let  $g_{ij} = 1$  if there is a link from  $i$  to  $j$ , and zero otherwise. Then configuration  $\mathbf{a}$ , where all players coordinate on strategy  $a$ , is uniquely stochastically stable on any network  $G$  with  $\delta^* = \max_{i \in N} \delta_i$  and

$$\delta_i = \inf \left\{ \delta : \forall S \subseteq N \setminus i \forall j^* \in S \text{ if } \frac{\sum_{j \in S} g_{ij}}{\sum_{j \neq i} g_{ij}} > \frac{1}{2} + \delta, \text{ then } \frac{\sum_{j \in S \setminus j^*} g_{ij}}{\sum_{j \neq i} g_{ij}} \geq \frac{1}{2} - \delta \right\} \quad (6)$$

If we label  $N_i$ , the first-order neighbours of player  $i \in N$ , as  $N_i = \{1, 2, \dots, n_i\}$ , then the set  $S$  satisfying (6) is such that when  $n_i$  is even,  $S = \{1, 2, \dots, \frac{n_i}{2}\}$ , and when  $n_i$  is odd,

---

<sup>11</sup>Specifically, for  $2 \times 2$  coordination games with strategies  $a_j$  and  $a_l$ , where  $a_l$  is a best response when at least proportion  $\phi_{jl}$  of neighbours play  $a_l$ , Morris (2000) defines the contagion threshold of a given network to be the supremum of  $\phi_{jl}$  such that contagion occurs in that network.

$S = \{1, 2, \dots, \lceil \frac{n_i}{2} \rceil\}$ . This implies that:

$$\delta_i = \begin{cases} 0 & \text{if } n_i \text{ is even} \\ \frac{1}{2n_i} & \text{if } n_i \text{ is odd} \end{cases} \quad (7)$$

where the equality for  $n_i$  odd follows because when  $S = \{1, 2, \dots, \lceil \frac{n_i}{2} \rceil\}$ , we have

$$\frac{\sum_{j \in S} g_{ij}}{\sum_{j \neq i} g_{ij}} = \frac{\lceil \frac{n_i}{2} \rceil}{n_i} = \frac{1}{2} + \frac{2\lceil \frac{n_i}{2} \rceil - n_i}{2n_i} = \frac{1}{2} + \frac{1}{2n_i}$$

If we let  $\delta_0(G)$  be the size of the smallest odd neighbourhood in  $G$ <sup>12</sup> so that  $\delta^* = \frac{1}{2\delta_0(G)}$ , then the results in [Peski et al. \(2010, Theorem 2 & Corollary 1\)](#) essentially state that a  $p$ -dominant strategy is uniquely stochastically stable in network  $G$  if  $p \leq \frac{1}{2} \left(1 - \frac{1}{\delta_0(G)}\right)$ .<sup>13</sup> Thus, these results depend solely on the size of the smallest odd neighbourhood. Stated this way, the contrast with our results becomes apparent.

First, the results in [Peski et al. \(2010\)](#) have stronger predictability for evolutionary dynamics in densely connected networks (i.e. where  $\delta_0(G)$  is large) – by strong predictability we mean the upper bound for  $p$  at which a  $p$ -dominant strategy is uniquely stochastically stable in a given network is tight (i.e.  $p$  is as close to  $\frac{1}{2}$  as it can be for a given network). Our results instead have stronger predictability in sparsely connected networks. For example, when the smallest odd neighbourhood size is 3, [Peski et al. \(2010, Theorem 2 & Corollary 1\)](#) predicts that a  $\frac{1}{3}$ -dominant strategy is uniquely stochastically stable. However, in the example of the network in Figure 1b above with  $\delta_0(G) = 3$ , Theorem 1 predicts that a  $\frac{1}{2}$ -dominant strategy is uniquely stochastically stable. [Peski et al. \(2010, Theorem 2 & Corollary 1\)](#) fails to predict unique stochastically stable outcomes in any network containing at least one player with  $n_i = 1$ . This is because for such networks,  $\delta_0(G) = 1$ , and hence,  $p = 0$ . An example of such a network is depicted in Figure 1a above, where Theorem 1 instead predicts that a  $\frac{2}{7}$ -dominant strategy is uniquely stochastically stable.

---

<sup>12</sup>That is, if  $\mathbb{O}$  is the set of all odd integers, then  $\delta_0(G) = \min_{i \in N; n_i \in \mathbb{O}} n_i$

<sup>13</sup>The proof of these results relies on the intuition that when  $p \leq \frac{1}{2} \left(1 - \frac{1}{\delta_0(G)}\right)$ , then for every  $i \in N$ ,  $\lceil (1-p)n_i \rceil > \lceil pn_i \rceil$ . That is, since  $(1-p) \geq 1 - \frac{1}{2} \left(1 - \frac{1}{\delta_0(G)}\right) = \frac{1}{2} \left(1 + \frac{1}{\delta_0(G)}\right)$ , we have  $\lceil (1-p)n_i \rceil \geq \lceil \frac{1}{2} \left(n_i + \frac{n_i}{\delta_0(G)}\right) \rceil \geq \lceil \frac{1}{2}(n_i + 1) \rceil$ , where the last inequality follows because  $\frac{n_i}{\delta_0(G)} \geq 1$ . Similarly,  $\lceil pn_i \rceil \leq \lceil \frac{1}{2} \left(n_i - \frac{n_i}{\delta_0(G)}\right) \rceil \leq \lceil \frac{1}{2}(n_i - 1) \rceil$ . Since  $\lceil \frac{1}{2}(n_i + 1) \rceil > \lceil \frac{1}{2}(n_i - 1) \rceil$ , it follows that  $\lceil (1-p)n_i \rceil > \lceil pn_i \rceil$ . This inequality means that, if  $a \in A$  is a  $p$ -dominant strategy, then the number of neighbours that must play strategies in  $A \setminus a$  for the best response to be within  $A \setminus a$  is strictly larger than the number that must play  $a$  for  $a$  to be a best response. The implication is that the cost of reaching  $\mathbf{a}$  from any other configuration will be strictly smaller than the cost of reaching any other absorbing set from all other configurations, which makes  $\mathbf{a}$  uniquely stochastically stable.

Second, Theorem 1 establishes conditions under which  $p$ -best response sets are stochastically stable and not just  $p$ -dominant strategies. This generalization is important because games with a  $p$ -dominant strategy are an exception rather than the rule. In the absence of a unique  $p$ -dominant strategy, the next best option may be to isolate strategies that form a  $p$ -best response set.

Our final remark concerns the implications of our results for the rate of convergence of stochastic evolutionary dynamics. As raised in Ellison (1993) and Oyama et al. (2015), one of the limitations of stochastic evolutionary models is that the number of mutations required to exit the basin of attraction of a non-stochastically stable absorbing set may be an increasing function of the population size. This leads to concerns regarding applicability of these models' predictions to large populations. However, under the conditions of Theorem 1, the number of mutations needed to exit the basin of attraction of any non-stochastically stable set is bounded from above by the size of the smallest neighbourhood and independent of the population size. Thus, the expected waiting time of evolution from any other state to set  $\mathbf{A}^p$  is independent of the population size, and hence, convergence is fast.

## 5. Concluding remarks

Evolutionary models with persistent randomness provide a framework for equilibrium selection in games with multiple equilibria. These models employ stochastic stability as a solution concept. However, recent developments have highlighted the lack of robustness of stochastic stability to the interaction structure, to the extent that it fails to distinguish between equilibria in  $2 \times 2$  coordination games in some networks.

One approach that can lead to relatively robust predictions is to identify network parameters and aggregate measures that determine the conditions under which stochastically stable outcomes are unique. We adopt this approach and identify the contagion threshold as one of such aggregate network measures. Network structures can then be categorized based on their contagion thresholds and the predictions of stochastic stability will be identical for networks with the same contagion threshold. We show that, with mild restrictions on the network diameter and connectedness, strategies that form a  $p$ -best response set are uniquely stochastically stable in networks with contagion threshold of  $p$ .

The contagion threshold is just one among a potential list of network parameters and aggregate measures that can robustly predict unique stochastically stable outcomes. More research is thus needed to further explore the role of different network properties.

## Appendix A. A matrix-based algorithm for computing the contagion threshold

The following steps provide a matrix-based algorithm for computing the contagion threshold of a network. Let  $G$  be an adjacency matrix of an undirected, unweighted and strongly connected network  $G(N, E)$ . That is,  $G$  is a zero-one matrix with elements  $g_{ij} = 1$  if there is a link from  $i$  to  $j$ , and zero otherwise. Given  $G(N, E)$ , let  $W(\tau)$  be a matrix with each element  $w_{ij}(\tau) = \alpha_j(B_{i_{\tau-1}})$  if  $j \in N_{i_{\tau-1}}$  and zero otherwise. If  $d(G)$  is the diameter of  $G(N, E)$ , then the sequence of  $W(\tau)$  matrices, for  $\tau = 2, 3, \dots, d(G)$ , can be computed through the following steps.

For any arbitrary matrix  $M$ , let  $M^{[1]}$  be a zero-one matrix derived from  $M$  by replacing all elements of  $M$  that are greater than one with one. Similarly, let  $M^{[0]}$  be a zero-one matrix derived from  $M$  by replacing all negative elements of  $M$  with zero. Let  $\text{RepZero}[M|Z]$  be an operator that takes two equal-sized square matrices  $M$  and  $Z$ , and replaces the element  $m_{ij}$  of  $M$  with zero if  $z_{ij} = 0$ , otherwise  $m_{ij} = m_{ij}$  if  $z_{ij} \neq 0$ . For a sequence of matrices  $M(0), M(1), M(2), \dots, M(r)$ , let  $S(M(r)) = \sum_{\tau=0}^r M(\tau)$ . We also write  $I$  for an identity matrix and  $D(G)$ , or simply  $D$  where no confusion arises, for the diagonal matrix with the diagonal elements being the sizes of neighbourhoods of nodes in network  $G(N, E)$ . More precisely, if  $\mathbf{c}$  is a column vector containing sums of rows of  $G$  (i.e.  $c_i = \sum_{j \in N} g_{ij}$  for each  $i \in N$ ), then  $D(G)$  is a square matrix whose diagonal is  $\mathbf{c}$  and zero elsewhere. A sequence  $\{W(\tau)\}$  for  $\tau = 2, 3, \dots, d(G)$  is computed through the following iterative steps.

$\tau = 1$	$M(0) = I; H(1) = M(1) = G.$
$\tau = 2$	$H(2) = M(1) \times G; Z(2) = H^{[1]}(2) - S(M(1)); M(2) = Z^{[0]}(2); \text{ and } W(2) = \text{RepZero}[H(2) M(2)] \times D^{-1}.$
$\tau = 3$	$H(3) = M(2) \times G; Z(3) = H^{[1]}(3) - S(M(2)); M(3) = Z^{[0]}(3); \text{ and } W(3) = \text{RepZero}[H(3) M(3)] \times D^{-1}.$
$\tau = 4$	$H(4) = M(3) \times G; Z(4) = H^{[1]}(4) - S(M(3)); M(4) = Z^{[0]}(4); \text{ and } W(4) = \text{RepZero}[H(4) M(4)] \times D^{-1}.$
----	-----
$\tau = \tau$	$H(\tau) = M(\tau - 1) \times G; Z(\tau) = H^{[1]}(\tau) - S(M(\tau - 1)); M(\tau) = Z^{[0]}(\tau); \text{ and } W(\tau) = \text{RepZero}[H(\tau) M(\tau)] \times D^{-1}.$
----	----- The iterative process continues until $\tau = d(G)$ -----

Given  $W(\tau)$  for all  $\tau = 2, 3, \dots, d(G)$ , we compute  $\alpha_i$  by taking the minimum  $w_{ij}(\tau)$  over all  $j \neq i$  and  $\tau = 2, 3, \dots, d(G)$ . That is,  $\alpha_i = \min_{j \neq i, \tau} w_{ij}(\tau)$ . The contagion threshold is  $\underline{\eta}(G) = \min_{i \in N} \alpha_i$ .

## Appendix B. Proof of Theorem 1

The proof of the Theorem involves deriving the upper bound for the modified coradius of  $\mathbf{A}^p$  and the lower bound for the radius of the basin of attraction of  $\mathbf{A}^p$ . We use the following definitions and notations for the rest of the proof. Recall that a path from player  $i$  to  $j$  is a set of players  $i = i_1, i_2, \dots, i_r = j$  with links  $i_1 \rightarrow i_2, i_2 \rightarrow i_3, \dots, i_{r-1} \rightarrow j$  without cycles. Let  $\mathcal{P}_{ij}^\tau$  be a set of all shortest paths from  $i$  to  $j$  of length  $\tau$ , and  $p_{ij}^\tau$  be a typical path in  $\mathcal{P}_{ij}^\tau$ . Similarly, let  $\mathcal{P}_i^\tau$  be a set of all shortest paths from  $i$  to all players at distance  $\tau$ , and  $p_i^\tau$  a typical path in  $\mathcal{P}_i^\tau$ . Recall also that  $d_i$  is the length of the shortest path from  $i \in N$  to any other player; we define  $\underline{d}(G) = \min_{i \in N} d_i$ . The following claim holds.

**Claim 1.** *For an undirected, unweighted and strongly connected network  $G(N, E)$  with diameter  $d(G) = \max_{i \in N} d_i$ , we have  $\underline{d}(G) \geq \frac{d(G)}{2}$  if  $d(G)$  is even, and  $\underline{d}(G) \geq \frac{d(G)+1}{2}$  if  $d(G)$  is odd.*

*Proof.* Given an undirected, unweighted and strongly connected network  $G(N, E)$  with diameter  $d(G) \geq r$ , pick any two players  $i, j \in N$  where  $j \in N_{i_r}$ . Pick any path  $p_{ij}^r \in \mathcal{P}_{ij}^r$  from  $i$  to  $j$  and label players along this path as  $i = i_1, i_2, i_3, \dots, i_{r+1} = j$ . If  $r$  is even, then:

$$d_{ij} = d_{i_1j} = r; d_{i_2j} = r - 1; d_{i_3j} = r - 2; \dots; d_{i_{\frac{r}{2}}j} = \frac{r}{2} - 1; d_{i_{\frac{r}{2}+2}j} = d_{i_{\frac{r}{2}+2}i} = \frac{r}{2}; d_{i_{\frac{r}{2}+4}i} = \frac{r}{2} - 1; \dots; d_{i_{r+1}i} = d_{ji} = r.$$

Similarly, if  $r$  is odd, then:

$$d_{ij} = d_{i_1j} = r; d_{i_2j} = r - 1; d_{i_3j} = r - 2; \dots; d_{i_{\frac{r+1}{2}}j} = \frac{r+1}{2}; d_{i_{\frac{r+1}{2}+2}i} = \frac{r+1}{2}; d_{i_{\frac{r+1}{2}+4}i} = \frac{r+3}{2}; \dots; d_{i_{r+1}i} = r - 1; d_{i_{r+1}i} = d_{ji} = r.$$

Thus, for any pair of players  $i, j \in N$  with  $j \in N_{i_r}$  and any path  $p_{ij}^r \in \mathcal{P}_{ij}^r$ ,  $d_h \geq \frac{r}{2}$  for all  $h$  along  $p_{ij}^r$  when  $r$  is even, and  $d_h \geq \frac{r+1}{2}$  for all  $h$  along  $p_{ij}^r$  when  $r$  is odd. Similarly, for every player,  $h$ , along the diameter of the network,  $d_h \geq \frac{d(G)}{2}$  when  $d(G)$  is even, and  $d_h \geq \frac{d(G)+1}{2}$  when  $d(G)$  is odd. And since the network is strongly connected, there exists a path connecting every other player  $k$  that does not lie along the diameter to some  $h$  along the diameter. For all such players,  $d_k \geq d_{kh} + \frac{d(G)}{2}$  if  $d(G)$  is even, and  $d_k \geq d_{kh} + \frac{d(G)+1}{2}$  if  $d(G)$  is odd.  $\square$

Claim 1 implies that when  $d(G) = 5$  and  $d(G) = 6$ , then  $\underline{d}(G) \geq 3$ . Similarly, when  $d(G) \geq 7$ , then  $\underline{d}(G) \geq 4$ .

### Upper bound for modified coradius of $\mathbf{A}^p$

The upper bound for the modified coradius of  $\mathbf{A}^p$  is formally stated in the following lemma.

**Lemma 2.** *Given  $(A, U, N, G, P)$ , if  $G$  is undirected, unweighted and strongly connected, with  $p \leq \underline{\eta}(G)$  and  $d_i \geq 3$  for all  $i \in N$ , then  $CR^*(\mathbf{A}^p) \leq \underline{b}_1$ , where  $\underline{b}_1 = \min_{i \in N} b_{i_1}$ .*

*Proof.* The proof follows in two steps. First, we show that starting from some  $\mathbf{x} \in \mathbf{X} \setminus \mathbf{A}^p$ , if  $\underline{b}_1$  players mutate to strategies in  $A^p$ , then the process  $(A, U, N, G, P)$  will either converge to some absorbing state/cycle consisting of only strategies in  $A^p$ , or to some absorbing cycle consisting of both strategies in  $A^p$  and  $A^{1-p}$ . In the former scenario, we then have  $CR^*(\mathbf{A}^p) = CR(\mathbf{A}^p) = \max_{\mathbf{x} \in \mathbf{X} \setminus \mathbf{A}^p} C(\mathbf{x}, \mathbf{A}^p) \leq \underline{b}_1$ . Second, we show that if  $(A, U, N, G, P)$  converges to an absorbing cycle  $W$ , then  $R(W) = C(W, \mathbf{A}^p)$  so that the coradius  $CR^*(\mathbf{A}^p) = \max_{\mathbf{x} \in \mathbf{X} \setminus \mathbf{A}^p} C(\mathbf{x}, W) \leq \underline{b}_1$ .

Starting from any  $\mathbf{x} \in \mathbf{X} \setminus \mathbf{A}^p$ , pick any player  $i \in N$ ; we pick player  $\underline{i} \in N$  for whom  $b_{\underline{i}_1} = \min_{i \in N} b_{i_1} = \underline{b}_1$ . Let all players in  $B_{\underline{i}_1}$  mutate to strategies in  $A^p$  at period  $t = 1$ . Then from  $t = 2$  onward, the process  $(A, U, N, G, P)$  evolves as follows, where we write  $i \rightarrow A'$  and  $N' \rightarrow A'$  to mean  $i$  and respectively each player in  $N'$  plays a strategy in  $A'$ ; and  $i \rightarrow A' \otimes A''$  and  $N' \rightarrow A' \otimes A''$  to mean  $i$  and respectively each player in  $N'$  either plays a strategy in  $A'$  or  $A''$ :

$t = 1$	$B_{\underline{i}_1} \rightarrow A^p$ ; and $N_{\underline{i}_2} \cup N_{\underline{i}_3} \cup \dots \cup N_{\underline{i}_{d_{\underline{i}}}} \rightarrow A^p \otimes A^{1-p}$ , depending on the exact configuration of $\mathbf{x}$ .
$t = 2$	$\underline{i} \cup N_{\underline{i}_2} \rightarrow A^p$ because at $t = 1$ , $N_{\underline{i}_1} \rightarrow A^p$ , and for each $h \in N_{\underline{i}_2}$ , $\alpha_h(N_{\underline{i}_1}) \geq p$ , which makes strategies in $A^p$ best responses. Note that no $h \in N_{\underline{i}_2}$ with $N_{h_1} \cap N_{\underline{i}_3} \neq \emptyset$ will play a strategy in $A^{1-p}$ even if all players in $N_{\underline{i}_2}$ and $N_{\underline{i}_3}$ play strategies in $A^{1-p}$ at $t = 1$ . This is because for such players, $p \leq \underline{\eta}(G) \leq \alpha_{\underline{i}} \leq \alpha_h(B_{\underline{i}_1})$ implies that $(1 - p) \geq (1 - \underline{\eta}(G)) \geq (1 - \alpha_{\underline{i}}) \geq (1 - \alpha_h(B_{\underline{i}_1})) = \alpha_h(N_{\underline{i}_2} \cup N_{\underline{i}_3})$ . Since $(1 - p) \geq \alpha_h(N_{\underline{i}_2} \cup N_{\underline{i}_3})$ , strategies in $A^{1-p}$ are not best responses to any such $h$ because they are best responses only when they are played by more than proportion $1 - p$ of neighbours. We also have $N_{\underline{i}_1} \rightarrow A^p \otimes A^{1-p}$ ; moreover, for each $h \in N_{\underline{i}_1}$ with $N_{h_1} \cap N_{\underline{i}_2} = \emptyset$ , $h \rightarrow A^p$ . And $N_{\underline{i}_3} \cup N_{\underline{i}_4} \cup \dots \cup N_{\underline{i}_{d_{\underline{i}}}} \rightarrow A^p \otimes A^{1-p}$ , depending on the exact configuration of $\mathbf{x}$ .
$t = 3$	$\underline{i} \cup N_{\underline{i}_2} \rightarrow A^p \otimes A^{1-p}$ . For each $h \in N_{\underline{i}_1}$ with $N_{h_1} \cap N_{\underline{i}_2} = \emptyset$ , $h \rightarrow A^p \otimes A^{1-p}$ ; for each $h \in N_{\underline{i}_1}$ with $N_{h_1} \cap N_{\underline{i}_2} \neq \emptyset$ but $N_{h_2} \cap N_{\underline{i}_3} = \emptyset$ , $h \rightarrow A^p \otimes A^{1-p}$ ; for each $h \in N_{\underline{i}_1}$ with $N_{h_2} \cap N_{\underline{i}_3} \neq \emptyset$ , $h \rightarrow A^p$ because $\alpha_h(N_{\underline{i}_2}) \geq p$ . To see why, notice that for each $h \in N_{\underline{i}_1}$ with $N_{h_2} \cap N_{\underline{i}_3} \neq \emptyset$ , there exists some $j \in N_{\underline{i}_3}$ with $\underline{\eta}(G) \leq \alpha_j \leq \alpha_h(B_{j_1}) \leq \alpha_h(N_{\underline{i}_2})$ . The last inequality follows because not all players in $N_{h_1} \cap N_{\underline{i}_2}$ (i.e. first-order neighbours of $h$ that constitute the proportion $\alpha_h(N_{\underline{i}_2})$ ) have their first-order neighbours in $B_{j_1}$ (see a more detailed related proof in Claim 2). Since $p \leq \underline{\eta}(G)$ , we have $\alpha_h(N_{\underline{i}_2}) \geq p$ . We also have $N_{\underline{i}_3} \rightarrow A^p$ because each $h \in N_{\underline{i}_3}$ has proportion $\alpha_h(B_{\underline{i}_2}) \geq p$ of neighbours play strategies in $A^p$ at $t = 2$ . And $N_{\underline{i}_4} \cup N_{\underline{i}_5} \cup \dots \cup N_{\underline{i}_{d_{\underline{i}}}} \rightarrow A^p \otimes A^{1-p}$ , depending on the exact configuration of $\mathbf{x}$ .



$t = 4$	$i \cup N_{i_1} \rightarrow A^p \otimes A^{1-p}.$ For each $h \in N_{i_2}$ with $N_{h_1} \cap N_{i_3} = \emptyset$ , $h \rightarrow A^p \otimes A^{1-p}$ ; for each $h \in N_{i_2}$ with $N_{h_1} \cap N_{i_3} \neq \emptyset$ , $h \rightarrow A^p$ because both the neighbours of $h$ that belong to $N_{i_1}$ and $N_{i_3}$ play strategies in $A^p$ at $t = 3$ , and the proportion of these neighbours is greater than $\alpha_h(N_{i_1}) \geq p$ . $N_{i_3} \rightarrow A^p \otimes A^{1-p}.$ $N_{i_4} \rightarrow A^p$ because $\alpha_h(N_{i_3}) \geq p$ for each $h \in N_{i_4}$ . And $N_{i_5} \cup N_{i_6} \cup \dots \cup N_{i_{d_i}} \rightarrow A^p \otimes A^{1-p}$ , depending on the exact configuration of $\mathbf{x}$ .
$t = 5$	$i \cup N_{i_2} \rightarrow A^p \otimes A^{1-p}.$ For each $h \in N_{i_1}$ with $N_{h_1} \cap N_{i_2} = \emptyset$ , $h \rightarrow A^p \otimes A^{1-p}$ ; for each $h \in N_{i_1}$ with $N_{h_1} \cap N_{i_2} \neq \emptyset$ but $N_{h_2} \cap N_{i_3} = \emptyset$ , $h \rightarrow A^p \otimes A^{1-p}$ ; for each $h \in N_{i_1}$ with $N_{h_2} \cap N_{i_3} \neq \emptyset$ ; $h \rightarrow A^p$ . $N_{i_3} \rightarrow A^p$ because for each $h \in N_{i_3}$ , all players in $N_{h_1} \cap N_{i_2}$ play strategies in $A^p$ at $t = 4$ , and since $\alpha_h(B_{i_2}) \geq p$ , a strategy in $A^p$ is a best response for each $h$ . $N_{i_4} \rightarrow A^p \otimes A^{1-p}.$ $N_{i_5} \rightarrow A^p.$ And $N_{i_6} \cup N_{i_7} \cup \dots \cup N_{i_{d_i}} \rightarrow A^p \otimes A^{1-p}$ , depending on the exact configuration of $\mathbf{x}$ .
$t = 6$	$i \cup N_{i_1} \rightarrow A^p \otimes A^{1-p}.$ For each $h \in N_{i_2}$ with $N_{h_1} \cap N_{i_3} = \emptyset$ , $h \rightarrow A^p \otimes A^{1-p}$ ; for each $h \in N_{i_2}$ with $N_{h_1} \cap N_{i_3} \neq \emptyset$ , $h \rightarrow A^p$ . $N_{i_3} \rightarrow A^p \otimes A^{1-p}.$ $N_{i_4} \rightarrow A^p.$ $N_{i_5} \rightarrow A^p \otimes A^{1-p}.$ $N_{i_6} \rightarrow A^p.$ And $N_{i_7} \cup N_{i_8} \cup \dots \cup N_{i_{d_i}} \rightarrow A^p \otimes A^{1-p}$ , depending on the exact configuration of $\mathbf{x}$ .
— —	— — — — — And so on — — — — —
—	
$t =$	Assume without loss of generality that $d_i$ is an odd number; then $i \cup N_{i_2} \rightarrow A^p \otimes A^{1-p}.$
$d_i$	For each $h \in N_{i_1}$ with $N_{h_1} \cap N_{i_2} = \emptyset$ , $h \rightarrow A^p \otimes A^{1-p}$ ; for each $h \in N_{i_1}$ with $N_{h_1} \cap N_{i_2} \neq \emptyset$ but $N_{h_2} \cap N_{i_3} = \emptyset$ , $h \rightarrow A^p \otimes A^{1-p}$ ; for each $h \in N_{i_1}$ with $N_{h_2} \cap N_{i_3} \neq \emptyset$ ; $h \rightarrow A^p$ . $N_{i_3} \rightarrow A^p$ $N_{i_4} \rightarrow A^p \otimes A^{1-p}.$ $N_{i_5} \rightarrow A^p.$ $N_{i_6} \rightarrow A^p \otimes A^{1-p}.$ $N_{i_7} \rightarrow A^p.$ — — — — — And $N_{i_{d_i}} \rightarrow A^p.$

$$\begin{array}{l|l}
t & i \cup N_{i_1} \rightarrow A^p \otimes A^{1-p}. \\
d_{\underline{i}}+1 & \text{For each } h \in N_{i_2} \text{ with } N_{h_1} \cap N_{i_3} = \emptyset, h \rightarrow A^p \otimes A^{1-p}; \text{ for each } h \in N_{i_2} \text{ with } N_{h_1} \cap N_{i_3} \neq \emptyset, \\
& h \rightarrow A^p. \\
& N_{i_3} \rightarrow A^p \otimes A^{1-p}. \\
& N_{i_4} \rightarrow A^p. \\
& N_{i_5} \rightarrow A^p \otimes A^{1-p}. \\
& N_{i_6} \rightarrow A^p. \\
& \text{-----} \\
& N_{i_{d_{\underline{i}}-1}} \rightarrow A^p. \\
& \text{And } N_{i_{d_{\underline{i}}}} \rightarrow A^p \otimes A^{1-p}. \\
- & \text{----- And so on -----} \\
- &
\end{array}$$

Thus, from  $t = d_{\underline{i}}$  iterations onward, if  $d_{\underline{i}} \geq 3$ , then the iterative process will converge to either an absorbing state/cycle containing only strategies in  $A^p$ , or to an absorbing cycle containing both strategies in  $A^p$  and  $A^{1-p}$ . In the former, this would imply that  $CR^*(\mathbf{A}^p) = CR(\mathbf{A}^p) \leq b_1$ . For the latter, let  $\mathcal{C}(\mathbf{A})$  be the set of all such absorbing cycles. The structure of a typical state  $\mathbf{x} \in W \subseteq \mathcal{C}(\mathbf{A})$  is as follows.

For  $r \geq 3$ , let  $N_{i_1}(\mathcal{P}_i^r)$  be a set of players in  $N_{i_1}$  that lie along the paths in  $\mathcal{P}_i^r$ ; that is,  $N_{i_1}(\mathcal{P}_i^r)$  is a set of all  $h \in N_{i_1}$  for whom  $N_{h_{r-1}} \cap N_{i_r} \neq \emptyset$ . Similarly,  $N_{i_2}(\mathcal{P}_i^r)$  is a set of all  $h \in N_{i_2}$  for whom  $N_{h_{r-2}} \cap N_{i_r} \neq \emptyset$ ;  $N_{i_3}(\mathcal{P}_i^r)$  is a set of all  $h \in N_{i_3}$  for whom  $N_{h_{r-3}} \cap N_{i_r} \neq \emptyset$ ;  $\dots$ ;  $N_{i_{r-1}}(\mathcal{P}_i^r)$  is a set of all  $h \in N_{i_{r-1}}$  for whom  $N_{h_1} \cap N_{i_r} \neq \emptyset$ . Then a typical  $\mathbf{x} \in W \subseteq \mathcal{C}(\mathbf{A})$  contains an alternating sequence  $N_{i_1}(\mathcal{P}_i^r) \rightarrow A^p$ ,  $N_{i_2}(\mathcal{P}_i^r) \rightarrow A^p \otimes A^{1-p}$ ,  $N_{i_3}(\mathcal{P}_i^r) \rightarrow A^p$ ,  $\dots$ ,  $N_{i_{d_{\underline{i}}}} \rightarrow A^p$  for all  $r = 3, 4, \dots, d_{\underline{i}}$  and  $d_{\underline{i}}$  is odd; or  $N_{i_1}(\mathcal{P}_i^r) \rightarrow A^p \otimes A^{1-p}$ ,  $N_{i_2}(\mathcal{P}_i^r) \rightarrow A^p$ ,  $N_{i_3}(\mathcal{P}_i^r) \rightarrow A^p \otimes A^{1-p}$ ,  $\dots$ ,  $N_{i_{d_{\underline{i}}}} \rightarrow A^p \otimes A^{1-p}$  for all  $r = 3, 4, \dots, d_{\underline{i}}$  and  $d_{\underline{i}}$  is even. In both cases, all other players not in  $N_{i_1}(\mathcal{P}_i^r) \cup N_{i_2}(\mathcal{P}_i^r) \cup N_{i_3}(\mathcal{P}_i^r) \cup \dots \cup N_{i_{d_{\underline{i}}}}$  will play any of the strategies in  $A^p \cup A^{1-p}$ .

Now, let  $(A, U, N, G, P)$  converge to some  $W \subseteq \mathcal{C}(\mathbf{A})$ . We claim that if  $\underline{d}(G) \geq 3$  (i.e.  $d_i \geq 3$  for all  $i \in N$ ), then for any  $\mathbf{x} \in W \subseteq \mathcal{C}(\mathbf{A})$ ,

$$R(W) = C(\mathbf{x}, \mathbf{A}^p) \leq \underline{n}_{i_1}(\mathcal{P}_{\underline{i}}) \quad (\text{B.1})$$

where  $\underline{n}_{i_1}(\mathcal{P}_{\underline{i}}) = \min_{r; r \geq 3} |N_{i_1}(\mathcal{P}_{\underline{i}}^r)|$  and  $|Z|$  is the cardinality of set  $Z$ .

To prove the second inequality in (B.1), consider any two configurations  $\mathbf{x}, \mathbf{z} \in W \subseteq \mathcal{C}(\mathbf{A})$ , whereby,  $\mathbf{z}$  is a subsequent state to  $\mathbf{x}$  under  $(A, U, N, G, P)$ . That is,  $\mathbf{x}$  consists of an alternating sequence  $N_{i_1}(\mathcal{P}_{\underline{i}}^r) \rightarrow A^p \otimes A^{1-p}$ ,  $N_{i_2}(\mathcal{P}_{\underline{i}}^r) \rightarrow A^p$ ,  $N_{i_3}(\mathcal{P}_{\underline{i}}^r) \rightarrow A^p \otimes A^{1-p}$ ,  $\dots$ ,  $N_{i_{d_{\underline{i}}}} \rightarrow A^p \otimes A^{1-p}$  for all  $r = 3, 4, \dots, d_{\underline{i}}$ ; and  $\mathbf{z}$  consists of  $N_{i_1}(\mathcal{P}_{\underline{i}}^r) \rightarrow A^p$ ,  $N_{i_2}(\mathcal{P}_{\underline{i}}^r) \rightarrow A^p \otimes A^{1-p}$ ,  $N_{i_3}(\mathcal{P}_{\underline{i}}^r) \rightarrow A^p$ ,

$\dots, N_{i_{d_i}} \rightarrow A^p$  for all  $r = 3, 4, \dots, d_i$ . Starting from  $\mathbf{x}$ , let all players in any  $N_{i_r}(\mathcal{P}_i^r)$  mutate to strategies in  $A^p$  at  $t = 1$ . If players in  $N_{i_2}(\mathcal{P}_i^r)$  mutate to strategies in  $A^p$ , then process  $(A, U, N, G, P)$  evolves as follows:

$$\begin{array}{l|l}
t = 1 & \begin{array}{l} N_{i_1}(\mathcal{P}_i^r) \cup N_{i_2}(\mathcal{P}_i^r) \cup N_{i_3}(\mathcal{P}_i^r) \rightarrow A^p \text{ for all } r = 3, 4, \dots, d_i; \\ N_{i_4}(\mathcal{P}_i^r) \rightarrow A^p \otimes A^{1-p}; N_{i_5}(\mathcal{P}_i^r) \rightarrow A^p; \dots; N_{i_{d_i}} \rightarrow A^p \text{ for all } r = 3, 4, \dots, d_i; \\ \text{All other players play strategies in either } A^p \text{ or } A^{1-p}. \end{array} \\
t = 2 & \begin{array}{l} N_{i_1} \cup N_{i_2} \cup N_{i_3} \cup N_{i_4}(\mathcal{P}_i^r) \rightarrow A^p \text{ for all } r = 3, 4, \dots, d_i \text{ because each } h \in B_{i_3} \cup N_{i_4}(\mathcal{P}_i^r) \text{ has} \\ \text{more than } \alpha_i \geq \underline{\eta}(G) \geq p \text{ of their neighbours in } N_{i_1}(\mathcal{P}_i^r) \cup N_{i_2}(\mathcal{P}_i^r) \cup N_{i_3}(\mathcal{P}_i^r); \\ N_{i_5}(\mathcal{P}_i^r) \rightarrow A^p \otimes A^{1-p}, N_{i_6}(\mathcal{P}_i^r) \rightarrow A^p, \dots, N_{i_{d_i}} \rightarrow A^p \otimes A^{1-p} \text{ for all } r = 3, 4, \dots, d_i; \\ \text{All other players play strategies in either } A^p \text{ or } A^{1-p}. \end{array} \\
t = 3 & \begin{array}{l} i \cup N_{i_1} \cup N_{i_2} \cup N_{i_3} \cup N_{i_4} \cup N_{i_5}(\mathcal{P}_i^r) \rightarrow A^p \text{ for all } r = 3, 4, \dots, d_i; \\ N_{i_6}(\mathcal{P}_i^r) \rightarrow A^p \otimes A^{1-p}; N_{i_7}(\mathcal{P}_i^r) \rightarrow A^p; \dots; N_{i_{d_i}} \rightarrow A^p \text{ for all } r = 3, 4, \dots, d_i; \\ \text{All other players play strategies in either } A^p \text{ or } A^{1-p}. \end{array} \\
-- & \begin{array}{l} --- \text{ And so on, until the entire network eventually plays strategies in } A^p --- \end{array}
\end{array}$$

If on the other hand players in some  $N_{i_\tau}(\mathcal{P}_i^r)$ , for  $\tau \geq 3$ , mutate to strategies in  $A^p$ , then from  $t = 1$  onward, process  $(A, U, N, G, P)$  evolves as follows:

$$\begin{array}{l|l}
t = 1 & \begin{array}{l} N_{i_{\tau-1}}(\mathcal{P}_i^r) \cup N_{i_\tau}(\mathcal{P}_i^r) \cup N_{i_{\tau+1}}(\mathcal{P}_i^r) \rightarrow A^p \text{ for all } r = 3, 4, \dots, d_i; \\ N_{i_1}(\mathcal{P}_i^r) \rightarrow A^p; N_{i_2}(\mathcal{P}_i^r) \rightarrow A^p \otimes A^{1-p}; \dots; N_{i_{\tau-2}}(\mathcal{P}_i^r) \rightarrow A^p \otimes A^{1-p} \text{ for all } r = 3, 4, \dots, d_i; \\ N_{i_{\tau+2}}(\mathcal{P}_i^r) \rightarrow A^p \otimes A^{1-p}; N_{i_{\tau+3}}(\mathcal{P}_i^r) \rightarrow A^p; \dots; N_{i_{d_i}} \rightarrow A^p \text{ for all } r = 3, 4, \dots, d_i; \\ \text{All other players play strategies in either } A^p \text{ or } A^{1-p}. \end{array}
\end{array}$$

$t = 2$   $N_{i_{\tau-2}}(\mathcal{P}_{\underline{i}}^r) \cup N_{i_{\tau-1}}(\mathcal{P}_{\underline{i}}^r) \cup N_{i_{\tau}} \cup N_{i_{\tau+1}} \cup N_{i_{\tau+2}}(\mathcal{P}_{\underline{i}}^r) \rightarrow A^p$  for all  $r = 3, 4, \dots, d_{\underline{i}}$ . This is true for players in  $N_{i_{\tau}} \cup N_{i_{\tau+1}} \cup N_{i_{\tau+2}}(\mathcal{P}_{\underline{i}}^r)$  because each has more than  $\alpha_{\underline{i}} \geq \underline{\eta}(G) \geq p$  of their neighbours in  $N_{i_{\tau-1}}(\mathcal{P}_{\underline{i}}^r) \cup N_{i_{\tau}}(\mathcal{P}_{\underline{i}}^r) \cup N_{i_{\tau+1}}(\mathcal{P}_{\underline{i}}^r)$ . Players in  $N_{i_{\tau-1}}(\mathcal{P}_{\underline{i}}^r)$  do not switch back to strategies in  $A^{1-p}$  at  $t = 2$  even if all players in  $N_{i_{\tau-2}}$  play strategies in  $A^{1-p}$  at  $t = 1$  because for any  $h \in N_{i_{\tau-1}}(\mathcal{P}_{\underline{i}}^r)$ ,  $N_{h_2} \cap N_{i_{\tau+1}} \neq \emptyset$ , which implies that there exists some  $j \in N_{h_2} \cap N_{i_{\tau+1}}$  for whom  $p \leq \underline{\eta}(G) \leq \alpha_j \leq \alpha_h(N_{j_1})$  for all  $h \in N_{i_{\tau-1}}(\mathcal{P}_{\underline{i}}^r)$ . Thus, each  $h \in N_{i_{\tau-1}}(\mathcal{P}_{\underline{i}}^r)$  has more than proportion  $p$  of her neighbours play strategies in  $A^p$  at  $t = 1$ . Moreover,  $p \leq \alpha_j \leq \alpha_h(N_{j_1})$  implies that  $(1-p) \geq (1-\alpha_j) \geq (1-\alpha_h(N_{j_1})) = \alpha_h(N_{i_{\tau-2}} \cup N_{i_{\tau-1}})$ . This in turn implies that strategies in  $A^{1-p}$  are not best responses for players in  $N_{i_{\tau-1}}(\mathcal{P}_{\underline{i}}^r)$  since they are best responses only when they are played by more than proportion  $1-p$  of neighbours. For the same reasons, players in  $N_{i_{\tau-2}}(\mathcal{P}_{\underline{i}}^r)$  will switch to strategies in  $A^p$  at  $t = 2$ . That is, for each  $h \in N_{i_{\tau-2}}(\mathcal{P}_{\underline{i}}^r)$ , there exists some  $j \in N_{h_2} \cap N_{i_{\tau}}$  for whom  $p \leq \underline{\eta}(G) \leq \alpha_j \leq \alpha_h(N_{j_1})$ , where  $p \leq \alpha_h(N_{j_1})$  implies that a strategy in  $A^p$  is a best response to  $h \in N_{i_{\tau-2}}(\mathcal{P}_{\underline{i}}^r)$ ;  
 $N_{i_1}(\mathcal{P}_{\underline{i}}^r) \rightarrow A^p \otimes A^{1-p}$ ;  $N_{i_2}(\mathcal{P}_{\underline{i}}^r) \rightarrow A^p$ ;  $\dots$ ;  $N_{i_{\tau-3}} \rightarrow A^p \otimes A^{1-p}$  for all  $r = 3, 4, \dots, d_{\underline{i}}$ ;  
 $N_{i_{\tau+3}}(\mathcal{P}_{\underline{i}}^r) \rightarrow A^p \otimes A^{1-p}$ ;  $N_{i_{\tau+4}}(\mathcal{P}_{\underline{i}}^r) \rightarrow A^p$ ;  $\dots$ ;  $N_{i_{d_{\underline{i}}}} \rightarrow A^p \otimes A^{1-p}$  for all  $r = 3, 4, \dots, d_{\underline{i}}$ ;  
 All other players play strategies in either  $A^p$  or  $A^{1-p}$ .  
 $t = 3$  For the same reasons above, we have  $N_{i_{\tau-3}}(\mathcal{P}_{\underline{i}}^r) \cup N_{i_{\tau-2}}(\mathcal{P}_{\underline{i}}^r) \cup N_{i_{\tau-1}} \cup N_{i_{\tau}} \cup N_{i_{\tau+1}} \cup N_{i_{\tau+2}} \cup N_{i_{\tau+3}}(\mathcal{P}_{\underline{i}}^r) \rightarrow A^p$  for all  $r = 3, 4, \dots, d_{\underline{i}}$ ;  
 $N_{i_1}(\mathcal{P}_{\underline{i}}^r) \rightarrow A^p$ ;  $N_{i_2}(\mathcal{P}_{\underline{i}}^r) \rightarrow A^p \otimes A^{1-p}$ ;  $\dots$ ;  $N_{i_{\tau-4}}(\mathcal{P}_{\underline{i}}^r) \rightarrow A^p \otimes A^{1-p}$  for all  $r = 3, 4, \dots, d_{\underline{i}}$ ;  
 $N_{i_{\tau+4}}(\mathcal{P}_{\underline{i}}^r) \rightarrow A^p \otimes A^{1-p}$ ;  $N_{i_{\tau+5}}(\mathcal{P}_{\underline{i}}^r) \rightarrow A^p$ ;  $\dots$ ;  $N_{i_{d_{\underline{i}}}} \rightarrow A^p$  for all  $r = 3, 4, \dots, d_{\underline{i}}$ ;  
 All other players play strategies in either  $A^p$  or  $A^{1-p}$ .  
 --- --- And so on, until the entire network eventually plays strategies in  $A^p$  ---

Thus, mutations by all players in any  $N_{i_{\tau}}(\mathcal{P}_{\underline{i}}^r)$  are sufficient to trigger an exit from the basin of attraction of any  $W \subseteq \mathcal{C}(\mathbf{A})$  to some state in  $\mathbf{x} \in \mathbf{A}^p$ . We can pick  $N_{i_{\tau}}(\mathcal{P}_{\underline{i}})$  with cardinality  $\underline{n}_{i_1}(\mathcal{P}_{\underline{i}}) = \min_{\tau, r \geq 3} |N_{i_{\tau}}(\mathcal{P}_{\underline{i}}^r)|$  so that  $C(\mathbf{x}, \mathbf{A}^p) \leq \underline{n}_{i_1}(\mathcal{P}_{\underline{i}})$ , which proves the inequality in (B.1).

To prove the first equality in (B.1), we show that when  $p \leq \underline{\eta}(G)$ , mutations by all players in some  $N_{i_{\tau}}(\mathcal{P}_{\underline{i}}^r)$  to strategies in  $A^{1-p}$  need not trigger exit from the basin of attraction of  $W \subseteq \mathcal{C}(\mathbf{A})$ . Let states  $\mathbf{x}, \mathbf{z} \in W \subseteq \mathcal{C}(\mathbf{A})$  be as defined above but replace  $\underline{i}$  by  $i$ . Starting from  $\mathbf{z}$ , let all players in  $N_{i_{\tau}}(\mathcal{P}_i^r)$  mutate to play strategies in  $A^{1-p}$  at  $t = 1$  instead of strategies in  $A^p$  which are their best responses. If  $d_i = 3$ , then we let players in  $N_{i_2}(\mathcal{P}_i^r)$  mutate to strategies in  $A^p$  so that, at  $t = 1$ ,  $i \rightarrow A^p \otimes A^{1-p}$ ,  $N_{i_1} \rightarrow A^p \otimes A^{1-p}$ ,  $N_{i_2}(\mathcal{P}_i^3) \rightarrow A^{1-p}$ ,  $N_{i_3} \rightarrow A^p \otimes A^{1-p}$ ; and all other players not in  $N_{i_2}(\mathcal{P}_i^3)$  also play either strategies in  $A^p$  or  $A^{1-p}$ . From  $t = 2$  onward, we have  $i \cup N_{i_1} \cup N_{i_2} \cup N_{i_3} \rightarrow A^p \otimes A^{1-p}$ .

Similarly, for  $d_i = 4$ , we have that at  $t = 1$ ,  $i \rightarrow A^p \otimes A^{1-p}$ ,  $N_{i_1} \rightarrow A^p \otimes A^{1-p}$ ,  $N_{i_2}(\mathcal{P}_i^3) \rightarrow A^{1-p}$ ,  $N_{i_3} \rightarrow A^p \otimes A^{1-p}$ ,  $N_{i_4} \rightarrow A^p$ ; and all other players not in  $N_{i_2}(\mathcal{P}_i^4)$  also play either strategies in

$A^p$  or  $A^{1-p}$ . From  $t = 2$  onward, we have  $i \cup N_{i_1} \cup N_{i_2} \cup N_{i_3} \cup N_{i_4} \rightarrow A^p \otimes A^{1-p}$ . If  $d_i \geq 5$ , the following iterative process unfolds:

$$\begin{array}{l|l}
t = 1 & \begin{array}{l} i \cup N_{i_1}(\mathcal{P}_i^r) \rightarrow A^p \otimes A^{1-p}; N_{i_2}(\mathcal{P}_i^r) \rightarrow A^p; \dots; N_{i_{\tau-2}}(\mathcal{P}_i^r) \rightarrow A^p \text{ for all } r = 3, 4, \dots, d_i; \\ N_{i_{\tau-1}}(\mathcal{P}_i^r) \rightarrow A^p \otimes A^{1-p}, N_{i_\tau}(\mathcal{P}_i^r) \rightarrow A^{1-p}, N_{i_{\tau+1}}(\mathcal{P}_i^r) \rightarrow A^p \otimes A^{1-p} \text{ for all } r = 3, 4, \dots, d_i; \\ N_{i_{\tau+2}}(\mathcal{P}_i^r) \rightarrow A^p; N_{i_{\tau+3}}(\mathcal{P}_i^r) \rightarrow A^p \otimes A^{1-p}; \dots; N_{i_{d_i}} \rightarrow A^p \text{ for all } r = 3, 4, \dots, d_i; \\ \text{All other players play strategies in either } A^p \text{ or } A^{1-p}. \end{array} \\
t = 2 & \begin{array}{l} i \cup N_{i_1}(\mathcal{P}_i^r) \rightarrow A^p; N_{i_2}(\mathcal{P}_i^r) \rightarrow A^p \otimes A^{1-p}; \dots; N_{i_{\tau-2}}(\mathcal{P}_i^r) \rightarrow A^p \otimes A^{1-p} \text{ for all } r = 3, 4, \dots, d_i; \\ N_{i_{\tau-1}}(\mathcal{P}_i^r) \rightarrow A^p \otimes A^{1-p}, N_{i_\tau}(\mathcal{P}_i^r) \rightarrow A^p \otimes A^{1-p}, N_{i_{\tau+1}}(\mathcal{P}_i^r) \rightarrow A^p \text{ for all } r = 3, 4, \dots, d_i; \\ N_{i_{\tau+2}}(\mathcal{P}_i^r) \rightarrow A^p \otimes A^{1-p}; N_{i_{\tau+3}}(\mathcal{P}_i^r) \rightarrow A^p; \dots; N_{i_{d_i}} \rightarrow A^p \otimes A^{1-p} \text{ for all } r = 3, 4, \dots, d_i; \\ \text{All other players play strategies in either } A^p \text{ or } A^{1-p}. \end{array} \\
t = 3 & \begin{array}{l} i \cup N_{i_1}(\mathcal{P}_i^r) \rightarrow A^p \otimes A^{1-p}; N_{i_2}(\mathcal{P}_i^r) \rightarrow A^p; \dots; N_{i_{\tau-2}}(\mathcal{P}_i^r) \rightarrow A^p \text{ for all } r = 3, 4, \dots, d_i; \\ N_{i_{\tau-1}}(\mathcal{P}_i^r) \rightarrow A^p \otimes A^{1-p}, N_{i_\tau}(\mathcal{P}_i^r) \rightarrow A^p, N_{i_{\tau+1}}(\mathcal{P}_i^r) \rightarrow A^p \otimes A^{1-p} \text{ for all } r = 3, 4, \dots, d_i; \\ N_{i_{\tau+2}}(\mathcal{P}_i^r) \rightarrow A^p; N_{i_{\tau+3}}(\mathcal{P}_i^r) \rightarrow A^p \otimes A^{1-p}; \dots; N_{i_{d_i}} \rightarrow A^p \text{ for all } r = 3, 4, \dots, d_i; \\ \text{All other players play strategies in either } A^p \text{ or } A^{1-p}. \end{array} \\
t \geq 4 & \text{--- --- --- The process } (A, U, N, G, P) \text{ reverts to an absorbing cycle } W \subseteq \mathcal{C}(\mathbf{A}) \text{ --- --- ---}
\end{array}$$

Thus, mutations by all players in some  $N_{i_\tau}(\mathcal{P}_i^r)$  to strategies in  $A^{1-p}$  need not trigger exit from the basin of attraction of  $W \subseteq \mathcal{C}(\mathbf{A})$ . The radius of the basin of attraction of  $W \subseteq \mathcal{C}(\mathbf{A})$  then corresponds to the number of mutations to strategies in  $A^p$  needed to trigger an exit to some state  $\mathbf{x} \in \mathbf{A}^p$ , so that for any  $\mathbf{x} \in W \subseteq \mathcal{C}(\mathbf{A})$ , we have  $R(W) = C(\mathbf{x}, \mathbf{A}^p) \leq \underline{n}_{i_1}(\mathcal{P}_i) \leq C(\mathbf{x}, \mathbf{A}')$ , where  $\mathbf{A}' \not\subseteq \mathcal{C}(\mathbf{A})$ .

Recall the definition of the modified coradius of an absorbing set  $\mathbf{A}^p$  as

$$CR^*(\mathbf{A}^p) = \max_{\mathbf{x} \in \mathbf{X} \setminus \mathbf{A}^p} C^*(\mathbf{x}, \mathbf{A}^p) = \max_{\mathbf{x} \in \mathbf{X} \setminus \mathbf{A}^p} \min_{(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T) \in S(\mathbf{x}, \mathbf{A}^p)} c^*(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T) \quad (\text{B.2})$$

where  $c^*(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T) = (c(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T) - \sum_{l=2}^{q-1} R(Z_l))$ . From the above analysis, we see that if the process  $(A, U, N, G, P)$  starts from some  $\mathbf{x} \in \mathbf{X} \setminus \mathbf{A}^p$  and passes through some absorbing cycle  $W \subseteq \mathcal{C}(\mathbf{A})$ , then for  $\mathbf{x} = \mathbf{x}_1$  and  $\mathbf{x}_T \in \mathbf{A}^p$ , we have

$$\begin{aligned}
C^*(\mathbf{x}, \mathbf{A}^p) &= \min_{(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T) \in S(\mathbf{x}, \mathbf{A}^p)} c^*(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T) \\
&= \min_{(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{T'}) \in S(\mathbf{x}, W)} c(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{T'}) \\
&\quad + \min_{(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T) \in S(W, \mathbf{A}^p)} c(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T) - R(W)
\end{aligned} \quad (\text{B.3})$$

We know from (B.1) that for any  $\mathbf{x} \in W \subseteq \mathcal{C}(\mathbf{A})$ ,  $R(W) = C(\mathbf{x}, \mathbf{A}^p)$ . This implies that  $C(W, \mathbf{A}^p) = \min_{(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T) \in S(W, \mathbf{A}^p)} c(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T) = R(W)$ . The expression on the right hand

side of (B.3) then reduces to

$$C^*(\mathbf{x}, \mathbf{A}^p) = \min_{(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{T'}) \in S(\mathbf{x}, W)} c(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{T'}) = C(\mathbf{x}, W) \leq \underline{b}_1 \quad (\text{B.4})$$

Since (B.4) is true for all  $\mathbf{x} \notin \mathbf{A}^p$ , the coradius of the basin of attraction of  $\mathbf{A}^p$  is then bounded from above by

$$CR^*(\mathbf{A}^p) = \max_{\mathbf{x} \in \mathbf{X} \setminus \mathbf{A}^p} C^*(\mathbf{x}, \mathbf{A}^p) \leq \underline{b}_1 \quad (\text{B.5})$$

□

### Lower bound for the radius of $\mathbf{A}^p$

We now derive the lower bound for  $R(\mathbf{A}^p)$ . The intuition of the proof is as follows. Starting from any  $\mathbf{x} \in \mathbf{A}^p$ , if  $b_{i_1}$  players, for any  $i \in N$ , mutate to strategies in  $A^{1-p}$ , then under some conditions we derive below, strategies in  $A^{1-p}$  cannot spread to players in  $N_{i_2}(\mathcal{P}_i^3) \cup N_{i_3} \cup N_{i_4} \cup \dots \cup N_{i_{d_i}}$ ; and after a sufficient number of iterations, the evolutionary process reverts to a state in  $\mathbf{A}^p$ . This in turn implies that the number of mutations needed to exit the basin of attraction of  $\mathbf{A}^p$  is greater than  $b_{i_1}$ , and hence,  $R(\mathbf{A}^p) \geq b_{i_1} + \iota$ , where  $\iota \geq 1$  is some positive integer. We start with networks with  $\underline{d}(G) \geq 4$  followed by networks with  $\underline{d}(G) = 3$ .

*The case of  $\underline{d}(G) \geq 4$*

Consider an undirected, unweighted and strongly connected network  $G$  with  $\underline{d}(G) \geq 4$ . The following claim holds.

**Claim 2.** *Given that  $\underline{d}(G) \geq 4$ , pick any  $i \in N$ . Then for every  $h \in N_{i_\tau}(\mathcal{P}_i^r)$  with  $\tau \geq 2$ ,  $r \geq 4$  and  $\tau + 2 \leq r \leq d_i$ , there exists some  $j \in N_{i_r}$  for whom:*

$$(1 - p) \geq (1 - \alpha_h(B_{j_{r-\tau-1}})) \geq \alpha_h(B_{i_{\tau-1}}) \geq \underline{\eta}(G) \quad (\text{B.6})$$

*Proof.* For any  $h \in N_{i_\tau}(\mathcal{P}_i^r)$  with  $\tau \geq 2$ ,  $r \geq 4$  and  $\tau + 2 \leq r \leq d_i$ , there exists some  $j \in N_{i_r}$  for whom  $\alpha_h(B_{j_{r-\tau-1}}) \leq (1 - \alpha_h(B_{i_{\tau-1}}))$ . To see why, recall that  $\alpha_h(B_{i_{\tau-1}}) = |N_{h_1} \cap B_{i_{\tau-1}}|/n_h$ . Some of  $h$ 's first-order neighbours may belong within  $N_{i_\tau}$ , and hence, not all players in  $N_{h_1} \setminus \{N_{h_1} \cap B_{i_{\tau-1}}\}$  belong to  $N_{i_{\tau+1}}$ . This implies that:

$$\alpha_h(N_{i_{\tau+1}}) = |N_{h_1} \cap N_{i_{\tau+1}}|/n_h \leq 1 - |N_{h_1} \cap B_{i_{\tau-1}}|/n_h = (1 - \alpha_h(B_{i_{\tau-1}})) \quad (\text{B.7})$$

Some of the players in  $N_{h_1} \cap N_{i_{\tau+1}}$  do not lie along the paths  $\mathcal{P}_{hj}^{r-\tau}$ , which implies that  $N_{h_1} \cap B_{j_{r-\tau-1}} \subseteq N_{h_1} \cap N_{i_{\tau+1}}$ , and hence,

$$\alpha_h(B_{j_{r-\tau-1}}) \leq \alpha_h(N_{i_{\tau+1}}) = |N_{h_1} \cap N_{i_{\tau+1}}|/n_h \leq 1 - |N_{h_1} \cap B_{i_{\tau-1}}|/n_h = (1 - \alpha_h(B_{i_{\tau-1}})) \quad (\text{B.8})$$

Now, recall that by definition of contagion threshold,  $\alpha_h(B_{j_{r-\tau-1}}) \geq \alpha_j \geq \underline{\eta}(G)$  for all  $h \in N_{j_{r-\tau}}$ . Thus, since  $p \leq \underline{\eta}(G) \leq \alpha_j \leq \alpha_h(B_{j_{r-\tau-1}}) \leq (1 - \alpha_h(B_{i_{\tau-1}}))$ , for  $\tau = 2, 3, \dots, r-2$  and  $r \geq 4$ , we have:

$$(1-p) \geq (1-\underline{\eta}(G)) \geq (1-\alpha_j) \geq (1-\alpha_h(B_{j_{r-\tau-1}})) \geq \alpha_h(B_{i_{\tau-1}}) \geq \alpha_i \geq \underline{\eta}(G), \quad (\text{B.9})$$

which proves the Lemma.  $\square$

The implication of Claim 2 is as follows. Let  $p \leq \underline{\eta}(G)$  and  $\underline{d}(G) \geq 4$ . Starting from any  $\mathbf{x} \in \mathbf{A}^p$ , let all players in  $B_{i_1}$ , for any  $i \in N$ , simultaneously mutate to strategies in  $A^{1-p}$  at  $t = 1$ . We know from claim 2 that when  $d_i \geq 4$ ,  $(1-p) \geq \alpha_h(B_{i_{\tau-1}}) \geq \underline{\eta}(G)$  for all  $h \in N_{i_\tau}(\mathcal{P}_i^r)$  with  $r \geq 4$  and  $2 \leq \tau \leq r-2$ . This implies that for all players in  $N_{i_2}(\mathcal{P}_i^r)$  strategies in  $A^{1-p}$  are not best responses since they are best responses only when they are played by more than  $1-p$  of neighbours.

This in turn implies that from  $t = 2$  onward, players in  $N_{i_2}(\mathcal{P}_i^r) \cup N_{i_3}(\mathcal{P}_i^r) \cup N_{i_4} \cup \dots \cup N_{i_{d_i}}$  play strategies in  $A^p$ . It also implies that there exists at least one  $j \in N_{i_{d_i}}$  for whom players in  $B_{j_2}$  all play strategies in  $A^p$ . Since  $p \leq \underline{\eta}(G) \leq \alpha_j \leq \alpha_h(B_{j_{\tau-1}})$  for all  $h \in N_{j_\tau}$  and all  $\tau = 3, 4, \dots, d_j$ , it follows that from  $t = 2$  onward,  $B_{j_3} \rightarrow A^p$ ; from  $t = 3$  onward  $B_{j_4} \rightarrow A^p$ ; and so on, until the entire network reverts to strategies in  $A^p$ .

Thus, when  $p \leq \underline{\eta}(G)$  and  $\underline{d}(G) \geq 4$ , more than  $b_{i_1}$ , for any  $i \in N$ , mutations are needed to trigger an exit from the basin of attraction of  $\mathbf{A}^p$ , and hence,  $R(\mathbf{A}^p) \geq b_{i_1} + \iota$ , where  $\iota \geq 1$  is some positive integer.

*The case of  $\underline{d}(G) = 3$*

The above proof for the case of  $\underline{d}(G) \geq 4$  relies on the results of Claim 2, especially on the inequality  $(1-p) \geq \alpha_h(B_{i_{\tau-1}})$  for all  $h \in N_{i_\tau}(\mathcal{P}_i^r)$  with  $\tau \geq 2$ ,  $r \geq 4$  and  $\tau+2 \leq r \leq d_i$ . However, this inequality does not hold when  $d_i = 3$ . To see why, notice that when  $d_i = 3$ , for any  $j \in N_{i_3}$ , we have  $N_{i_2}(\mathcal{P}_{ij}^3) \subseteq B_{j_1}$  so that  $\alpha_h(B_{j_0})$  is undefined for all  $h \in N_{i_2}(\mathcal{P}_{ij}^3)$ . It is then possible that  $\underline{\eta}(G) = \alpha_j < \alpha_i \leq \alpha_h(B_{i_1})$  for all  $h \in N_{i_2}$ . Under this scenario, we cannot guarantee that  $(1-p) \geq \alpha_h(B_{i_1})$ , and hence, some or all players in  $N_{i_2}$  may switch to strategies in  $A^{1-p}$  at  $t = 2$ . Thus,  $b_{i_1}$  may trigger an exit from the basin of attraction of  $\mathbf{A}^p$  to some absorbing cycle containing both strategies in  $A^p$  and  $A^{1-p}$ .

To avoid the above situation, we can fix  $(1-p) \geq \bar{\eta}_2(G)$ ; that is, for every  $i \in N$ ,  $(1-p) \geq \alpha_h(B_{i_1})$  for all  $h \in N_{i_2}(\mathcal{P}_i^3)$ . Under this condition, even if all players in  $B_{i_1}$  mutate to strategies in  $A^{1-p}$  starting from some  $\mathbf{x} \in \mathbf{A}^p$ , strategies in  $A^{1-p}$  will not spread to players in  $N_{i_2}(\mathcal{P}_i^3)$  and  $N_{i_3}$  through best response.



Specifically, let players in  $B_{i_1}$  mutate to strategies in  $A^{1-p}$  at  $t = 1$  starting from some  $\mathbf{x} \in \mathbf{A}^p$ . Consider any  $j \in N_{i_3}$ ; all players in  $B_{j_1}$  play strategies in  $A^p$  at  $t = 1$ . At  $t = 2$ , all players in  $B_{j_2}$  play strategies in  $A^p$ . To see why, first notice that all players in  $B_{j_1}$  play strategies in  $A^p$  because  $(1-p) \geq \bar{\eta}_2(G) \geq \alpha_h(B_{i_1})$  for all  $h \in N_{i_2}(\mathcal{P}_{ij}^3) \cap N_{j_1}$  (i.e. strategies in  $A^{1-p}$  will not spread from  $B_{i_1}$  to players in  $B_{j_1}$ ). Second, every  $h \in N_{j_2}$  who does not lie along the paths  $\mathcal{P}_{ij}^3$  (i.e.  $h \notin N_{i_1}(\mathcal{P}_{ij}^3)$ ) plays a strategy in  $A^p$  because they do not have any neighbours playing strategies in  $A^{1-p}$  at  $t = 1$  (i.e. if  $h \in N_{j_2}$  but  $h \notin N_{i_1}(\mathcal{P}_{ij}^3)$ , then  $N_{h_1} \cap B_{i_1} = \emptyset$ ). Third, every  $h \in N_{j_2}$  who lies along the paths  $\mathcal{P}_{ij}^3$  (i.e.  $h \in N_{i_1}(\mathcal{P}_{ij}^3)$ ) also plays a strategy in  $A^p$  because: (i)  $p \leq \underline{\eta}(G) \leq \alpha_j \leq \alpha_h(B_{j_1})$  so that a strategy in  $A^p$  is best responses to such a player; (ii)  $p \leq \underline{\eta}(G) \leq \alpha_j \leq \alpha_h(B_{j_1})$  implies that  $(1-p) \geq (1-\underline{\eta}(G)) \geq (1-\alpha_j) \geq (1-\alpha_h(B_{j_1})) = \alpha_h(B_{i_1})$ , and hence, none of the strategies in  $A^{1-p}$  is a best response to such a player even if all players in  $B_{i_1}$  play strategies in  $A^{1-p}$  at  $t = 1$ . Hence, all players in  $N_{j_2}$  also play strategies in  $A^p$ .

From  $t = 3$  onward, all players in  $B_{i_3}$  play strategies in  $A^p$ , where every  $h \in N_{j_3}$  plays a strategy in  $A^p$  because  $p \leq \underline{\eta}(G) \leq \alpha_j \leq \alpha_h(B_{j_2})$ , and that  $(1-p) \geq (1-\underline{\eta}(G)) \geq (1-\alpha_j) \geq (1-\alpha_h(B_{j_2}))$ . This iterative process continues until the entire network eventually switches back to strategies in  $A^p$ . This implies that, if  $p \leq \underline{\eta}(G)$  and  $(1-p) \geq \bar{\eta}_2(G)$ , or equivalently,  $p \leq 1 - \bar{\eta}_2(G)$ , then  $b_{i_1}$  mutations, for any  $i \in N$ , are not sufficient to trigger an exit from the basin of attraction of  $\mathbf{A}^p$ . Thus, when  $p \leq \min\{\underline{\eta}(G), 1 - \bar{\eta}_2(G)\}$ , the radius  $R(\mathbf{A}^p) \geq b_{i_1} + \iota$ , where  $\iota \geq 1$ .

### Stochastically stable states/strategies

Putting the above results together, we see that:

- (i) For an undirected, unweighted and strongly connected network with  $\underline{d}(G) \geq 4$ , if  $p \leq \underline{\eta}(G)$  then  $CR^*(\mathbf{A}^p) \leq \underline{b}_1$  and  $R(\mathbf{A}^p) \geq b_{i_1} + \iota$ , for  $\iota \geq 1$ . Thus, under these conditions,  $R(\mathbf{A}^p) > CR^*(\mathbf{A}^p)$ , which implies that configurations in  $\mathbf{A}^p$  (strategies in  $A^p$ ) are uniquely stochastically stable.
- (ii) Similarly, for an undirected, unweighted and strongly connected network with  $\underline{d}(G) = 3$ , if  $p \leq \min\{\underline{\eta}(G), 1 - \bar{\eta}_2(G)\}$ , then  $CR^*(\mathbf{A}^p) \leq \underline{b}_1$  and  $R(\mathbf{A}^p) \geq b_{i_1} + \iota$ . In which case  $R(\mathbf{A}^p) > CR^*(\mathbf{A}^p)$ , and hence, configurations in  $\mathbf{A}^p$  (strategies in  $A^p$ ) are uniquely stochastically stable.

## Acknowledgements

This paper was part of my thesis as a PhD student at the School of Economics and Business, Maastricht University. I am currently a Research Associate at the School of Economics, University of Cape Town, South Africa. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

I gratefully acknowledge the comments from Alan Kirman, Vianney Dequiedt, Jean-Jacques Herings, Arkadi Predtetchinski, Théophil Azomahou, Robin Cowan, François Lafond, Giorgio Triulzi and Stefania Innocenti. I also acknowledge the fruitful discussions with seminar participants at Maastricht Lecture Series in Economics, University of Cape Town and University of Pretoria. The usual disclaimer applies.

## References

- Alós-Ferrer, C., Weidenholzer, S., 2006. Imitation, local interactions, and efficiency. *Economics Letters* 93, 163–168.
- 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.
- Blume, L.E., 1995. The statistical mechanics of best-response strategy revision. *Games and Economic Behavior* 11, 111–145.
- Chen, H.C., Chow, Y., Wu, L.C., 2013. Imitation, local interaction, and coordination. *International Journal of Game Theory* 42, 1041–1057.
- 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.*
- Kandori, M., Mailath, G.J., Rob, R., 1993. Learning, mutation, and long run equilibria in games. *Econometrica* 61, 29–56.

- Klaus, B., Newton, J., 2016. Stochastic stability in assignment problems. *Journal of Mathematical Economics* 62, 62–74.
- 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.
- Maruta, T., 2002. Binary games with state dependent stochastic choice. *Journal of Economic Theory* 103, 351–376.
- 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.
- Myatt, D.P., Wallace, C., 2003. A multinomial probit model of stochastic evolution. *Journal of Economic Theory* 113, 286–301.
- Newton, J., Sawa, R., 2015. A one-shot deviation principle for stability in matching problems. *Journal of Economic Theory* 157, 1–27.
- Oyama, D., Sandholm, W.H., Tercieux, O., 2015. Sampling best response dynamics and deterministic equilibrium selection. *Theoretical Economics* 10, 243–281.
- Peski, M., et al., 2010. Generalized risk-dominance and asymmetric dynamics. *Journal of Economic Theory* 145, 216.
- Sandholm, W.H., 2001. Almost global convergence to  $p$ -dominant equilibrium. *International Journal of Game Theory* , 107–116.
- Sandholm, W.H., 2010. Orders of limits for stationary distributions, stochastic dominance, and stochastic stability. *Theoretical Economics* 5, 1–26.
- Sandholm, W.H., Staudigl, M., 2016. Large deviations and stochastic stability in the small noise double limit. *Theoretical Economics* 11, 279–355.
- Sawa, R., Wu, J., 2018. Reference-dependent preferences, super-dominance and stochastic stability. *Journal of Mathematical Economics* .

- Staudigl, M., 2012. Stochastic stability in asymmetric binary choice coordination games. *Games and Economic Behavior* 75, 372–401.
- Staudigl, M., Weidenholzer, S., 2014. Constrained interactions and social coordination. *Journal of Economic Theory* 152, 41–63.
- Tercieux, O., 2006. p-best response set. *Journal of Economic Theory* 131, 45–70.
- Weidenholzer, S., 2012. Long-run equilibria, dominated strategies, and local interactions. *Games and Economic Behavior* 75, 1014–1024.
- 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.