On the relationship between p-dominance and stochastic stability in network games [*]

Daniel C. Opolot

University of Cape Town

Abstract

This paper examines the properties of networks that determine the uniqueness of long-run equilibria emerging from symmetric coordination games when players are myopic best responders. We identify the *contagion threshold* and the *network diameter* as two measures of finite networks that determine when strategies in the *minimal p-best response set* of a coordination game are uniquely stochastically stable. We show that when the contagion threshold is greater or equal to p, strategies in the minimal p-best response set are uniquely stochastically stable in strongly connected networks with diameter greater or equal to seven. The contagion threshold and the network diameter are easy to compute and their values are unique for every strongly connected network.

Keywords: Evolutionary dynamics, stochastic stability, networks, p-best response set, contagion threshold

JEL Classification: 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,[1] 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 examining the process by which conventions (Nash equilibria of coordination games) become established using evolutionary models with persistent randomness. In this framework, persistent randomness ensures that all outcomes

---

[*] Daniel C. Opolot: opodanchris@gmail.com.

[1]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.

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. Recent studies show that the predictions of stochastic stability strictly depend on the interaction structure (Alós-Ferrer and Weidenholzer, 2007; Peski, 2010).[2] And for most coordination games, there can be multiple stochastically stable outcomes on a given interaction structure. Even for 2×2 coordination games, it is easy to construct an interaction structure where uniform adoption of either strategy is stochastically stable.

This paper seeks to identify properties of the interaction structure that robustly determine the 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*. The social network is modeled as a graph consisting of the set of players as nodes and a set of edges/links connecting different pairs of players. For any pair of players i and j, a directed link from i to j implies that i observes strategies chosen by j. The central behavioural assumptions of the BRM model are: (*i*) players are *myopic* in that their strategy choices at any period depend on the profile of strategies chosen by their opponents (i.e., the set of players with whom a player directly interacts) in the previous period; (*ii*) players choose strategies that are best responses to opponents' strategy profile with a high probability, and with a small probability, independent across players and time, they experiment (mutate) and choose any strategy at random. These behavioural assumptions capture some aspects of bounded rationality of decision makers (e.g., limited attention and the inability to make farsighted decisions) often observed in economics and psychology.[3]

We develop a new network measure, the *contagion threshold*, which together with the *network diameter* determine when strategies in the *minimal p-best response set* are uniquely stochastically stable.[4] The minimal p-best response set is a generalization of risk-

---

[2]This result has both strengths and weaknesses. On the one hand, it implies that, where possible, the interaction structure can be manipulated to obtain a desired stochastically stable outcome. On the other hand, it implies that when computing the stochastically stable outcome of an evolutionary model on arbitrary networks, the modeller must keep track of the identity of players. The computational time of stochastic stability algorithms, however, increases exponentially with the population size in such scenarios. This could in turn limit the applicability of evolutionary models.

[3]See DellaVigna (2009) for the survey of the economics and psychology literature on bounded rationality.

[4]Throughout this paper, a set of strategies, A', is said to be uniquely stochastically stable if the stochastically stable outcomes consist of states where all players play only strategies in A'.

dominance to games with more than two strategies. A set of strategies of a coordination game 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). The *minimal p-best response set* is the p-best response set that does not contain another smaller subset that itself is a p-best response set. If the minimal p-best response set contains only one strategy, then that strategy is a *p-dominant equilibrium* (Morris et al., 1995). In a coordination game with only two strategies, a $\frac{1}{2}$-dominant equilibrium is also a risk-dominant equilibrium.

To define the contagion threshold, consider an *unweighted* and *strongly connected* network, and denote it by G – a network is unweighted if every link between any pair of players carries a weight of one, and it is strongly connected if every two players are connected through some path. Denote the set of players by N and define the *r-order neighbours* of any player i as the set of players located at a distance of r steps from i (i.e. player j is an r-order neighbour of i if the shortest path from i to j consists of r consecutive links). Define also the *rth-neighbourhood* of i as the set of all players within r steps from i, with i included; and d_i as the maximum shortest distance from i to any other player. The contagion threshold, $\underline{\eta}(G)$, of network G is the minimum over all $i \in N$, all $r = 2, 3, \cdots, d_i$ and all r-order neighbours of i, of the proportion of the first-order neighbours of each r-order neighbour of i that are contained in the $(r-1)$th-neighbourhood of i.[5] We extend this definition of the contagion threshold for unweighted networks to weighted networks in Section 5. Every strongly connected network has a unique contagion threshold bounded from above by $1/2$ and from below by $1/\Delta(G)$, where $\Delta(G)$ is the maximum number of first-order neighbours of any player in the network.

The diameter, $d(G)$, of network G is the maximum shortest distance between any two players (i.e., the maximum over all $i \in N$ of d_i). The value of the network diameter is unique for every strongly connected network and it captures the *density* and *centralization* of the network. Centralized networks exhibit a *core-periphery* structure where a subset of densely interconnected players form the core, 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). The network density is the total number of connections relative to the number of players. For a fixed number of players, the lower the network density the larger the network diameter.

We show that an absorbing set of states containing only strategies in the minimal p-best response set is uniquely stochastically stable in undirected, unweighted and strongly connected networks with the contagion threshold greater or equal to p and diameter greater or equal to seven. We discuss in Section 5 how this result directly extends to

---

[5]Formally, given network G, let N_{i_r} and B_{i_r} denote the r-order neighbours and the rth-neighbourhood of i respectively. Let $\alpha_j(B_{i_r})$ denote the proportion of player j's first-order neighbours that are contained in set B_{i_r}. Then the contagion threshold of G is defined as $\underline{\eta}(G) = \min_{i \in N} \min_{r \in \{2, 3, \cdots, d_i\}} \min_{j \in N_{i_r}} \alpha_j(B_{i_{r-1}})$.

directed, weighted and strongly connected networks.

The proof of this result relies on the properties of contagion dynamics in networks. We distinguish between three forms of contagion dynamics: *full contagion*, *partial contagion*, and *step-by-step contagion*. Full contagion is the spread of strategies to the whole network through best response starting from a group of players whose size is less than half the population size. Partial contagion is the spread of strategies through best response from a group of players, say N', whose size, n', is less than half the population size, to a group of players whose size is larger than n'. Step-by-step contagion is the spread of strategies to the whole network through iterative application of partial contagion.[6]

The contagion threshold and the network diameter determine when strategies in the minimal p-best response set are uniquely stochastically stable because, firstly, when $\underline{\eta}(G) \geq p$ and $d(G) \geq 7$, strategies in the minimal p-best response set spread through full and step-by-step contagion from the first-neighbourhood of any player. This implies that the minimum number of mutations that trigger the spread of strategies in the minimal p-best response set through full and step-by-step contagion is at most the size of the smallest first-neighbourhood of G. Secondly, when $\underline{\eta}(G) \geq p$ and $d(G) \geq 7$, the cost of reaching the basin of attraction of an absorbing set of states containing only strategies in the minimal p-best response set from any other state is bounded from above by the minimum number of mutations that trigger full and step-by-step contagion.[7] Thirdly, when $\underline{\eta}(G) \geq p$ and $d(G) \geq 7$, the minimum number of mutations needed to leave the basin of attraction of an absorbing set of states containing only strategies in the minimal p-best response set is greater than the size of the smallest first-neighbourhood of G. Following Ellison (2000, Theorems 1 & 2), these three observations imply that an absorbing set of states containing only strategies in the minimal p-best response set is uniquely stochastically stable whenever $\underline{\eta}(G) \geq p$ and $d(G) \geq 7$.[8]

The network measures that have previously been identified as determinants of the uniqueness of stochastically stable outcomes include the sizes of the *smallest odd* and the *largest* first-order neighbourhoods, and the *maximum group cohesion* of the network (Alós-Ferrer and Weidenholzer, 2008; Peski, 2010; Opolot, 2020). Let $\delta_0(G)$ be the size

---

[6]More formally, a set of strategies, A', is step-by-step contagious if, starting from any state $\mathbf{x}$, there exists a sequence of absorbing sets $W_1, W_2, \cdots, W_J$, with $W_J = \mathbf{A}'$ being the absorbing set of states containing only strategies in A', and a corresponding sequence of strategy sets, $A_2, A_3, \cdots, A_J$, such that, for any $\mathbf{y} \in W_j$, strategies in A_{j+1} spread through partial contagion after n_{j+1} mutations to strategies in A_{j+1}, where, for all $j = 1, \cdots, J-1$, n_{j+1} is less than half the population size. Along the sequence $W_1, W_2, \cdots, W_J$, the evolutionary process converges to W_{j+1} from any $\mathbf{y} \in W_j$. When $n_{j+1} \leq n^*$ for all $j = 1, \cdots, J-1$, we say that n^* mutations trigger step-by-step contagion of strategies in A' starting from $\mathbf{x}$.

[7]That is, the *coradius* and *modified coradius* (Ellison, 2000) of an absorbing set of states containing only strategies in the minimal p-best response set are bounded from above by the size of the smallest first-neighbourhood of the network.

[8]Ellison (2000, Theorems 1 & 2) states that if the cost of leaving the basin of attraction of an absorbing set of states is greater than the cost of reaching it from any other state, then that absorbing set is uniquely stochastically stable.

of the smallest odd first-order neighbourhood of network G. Peski (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 $\frac{1}{2}\left(1 - 1/\delta_0(G)\right)$. In comparison, our results offer
more accurate predictions of stochastically stable outcomes in sparsely connected networks
(i.e. where $\Delta(G)$ is small so that the lower bound of the contagion threshold, $1/\Delta(G)$, is
large), while Peski (2010, Theorem 2 & Corollary 1) offers more accurate predictions in
densely connected networks (i.e., where $\delta_0(G)$ is large so that $\frac{1}{2}\left(1 - 1/\delta_0(G)\right)$ is close to
one half). For example, when $\delta_0(G) = 1$, Peski (2010, Theorem 2 & Corollary 1) fails to
predict the stochastically stable outcome.

Alós-Ferrer and Weidenholzer (2008) show that in a BRM model, a $1/\Delta(G)$-dominant
strategy is uniquely stochastically stable.[9] Our results generalize this special case from
Alós-Ferrer and Weidenholzer (2008) since the lower bound of the contagion threshold of
any network is $1/\Delta(G)$.[10] Opolot (2020) finds that for a BRM model on cyclic regular
networks where the size of the first-order neighbourhoods is k, strategies in the *smallest
iterated p-best response set* (i.e., a set of strategies that remain after iterative application
of the notion of p-best response sets) are uniquely stochastically stable whenever p is less
or equal to one minus the maximum group cohesion.[11] Opolot (2020) provides a counter
example showing that the maximum group cohesion fails to predict the uniqueness of
stochastically stable outcomes in some arbitrary networks. Since every strongly connected
arbitrary network has a unique contagion threshold greater or equal to $1/\Delta$, our result
provides a prediction of when the minimal p-best response set is uniquely stochastically
stable in any strongly connected arbitrary network.

Besides the aforementioned papers, many other papers have examined long-run sta-
bility of the BRM model in networks.[12] Ellison (1993, 2000) and Weidenholzer (2012)

---

[9]More generally, Alós-Ferrer and Weidenholzer (2008) study an evolutionary model where players
imitate strategies that yield the highest payoff (i.e. "imitate-the-best") and find that a payoff dominant
strategy of a 2×2 coordination game is the unique stochastically stable outcome. Related papers in this
regard include Alós-Ferrer and Weidenholzer (2006) and Chen et al. (2013).

[10]Sandholm (2001) and Oyama et al. (2015) also show that a $1/\Delta$-dominant strategy is the long-
run stable outcome of a deterministic evolutionary model where players choose strategies that are best
responses to the empirical distribution of at most Δ others sampled randomly from an infinitely large
population. Although the solution concept of this model – *almost global convergence* – is different from
stochastic stability, the result is comparable to the stochastic stability of a $1/\Delta$-dominant equilibrium in
a BRM model.

[11]A group of players in a given network is *c-cohesive* if every player in that group has at least proportion
c of her neighbours within that group. The maximum group cohesion of the network is then the maximum
cohesiveness of any group of players.

[12]There are many papers that also study alternative classes of noisy best response choice rules to the
BRM model considered in this paper (Maruta, 2002; Myatt and Wallace, 2003; Staudigl, 2012; Staudigl
and Weidenholzer, 2014; Newton and Sawa, 2015; Sandholm and Staudigl, 2016; Sawa and Wu, 2018).
These papers however focus on global interactions, which makes comparison to our findings non-trivial.
Few papers have also examined the BRM model with asymmetric coordination games. For example,
Staudigl (2012) finds that a risk-dominant strategy is stochastically stable in 2×2 asymmetric coordination
games; and Neary (2012) examines the effects of payoff asymmetries caused by differences in preferences
of mutiple intereacting groups of players. We instead focus on symmetric coordination games where it is
relatively easier to isolate the effects of the network structure on stochasitcally stable outcomes.

show that a $\frac{1}{2}$-dominant strategy is uniquely stochastically stable in cyclic regular networks where each player has an even number of neighbours; Blume (1995), Young (1998), Lee and Valentinyi (2000) and Lee et al. (2003) show that the risk-dominant strategy is uniquely stochastically stable in 2-dimensional grid lattices; and Alós-Ferrer and Weidenholzer (2007) show that a globally pairwise risk-dominant strategy of a 3×3 coordination game is stochastically stable in a cyclic network where each player has two neighbours. In contrast to the present paper, these papers focus on specific (mostly regular) networks and do not identify the general network properties that determine the uniqueness of stochastically stable outcomes.

Finally, this paper is related to Morris (2000), Oyama and Takahashi (2015) and Azomahou and Opolot (2018) who study contagion in networks when players are myopic best responders. Morris (2000) finds that a p-dominant strategy of a 2×2 coordination game is fully contagious on an unbounded network if p is less or equal to the contagion threshold. We discuss in Section 3.1 the difference between the contagion threshold for finite networks defined in this paper and the contagion threshold for unbounded networks as defined by Morris (2000). Oyama and Takahashi (2015) derive the conditions under which the risk-dominant and Pareto-dominant strategies of a 3×3 coordination game containing a dominated strategy are fully contagious on unbounded networks; they also compare networks in terms of their power of inducing contagion in general supermodular games. Azomahou and Opolot (2018) show that a p-dominant strategy of an m-strategy coordination game is fully contagious on a finite network whenever p is less or equal to the contagion threshold. We build on the ideas of contagion dynamics developed in these papers and use them to derive the conditions under which the minimal p-best response set is uniquely stochastically stable.

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 discusses how contagion determines stochastically stable outcomes. Section 4 states the main results and discusses the intuition of the proof. Section 5 discusses the implications of our results and how they relate to the literature. Concluding remarks are offered in Section 6 and proofs are contained in the Appendix.

2   An evolutionary model of best response with mutations

We consider a finite set of players, $N = \{1, 2, \cdots, n\}$, connected through a social network. Each player plays a symmetric coordination game against her direct neighbours. Players revise their strategies simultaneously and independently at discrete times, $t = 1, 2, \cdots$. We start with a description of an unperturbed evolutionary model where, at

each t, each player chooses a pure strategy randomly from the set of pure strategies that are best responses to the profile of strategies played by the opponents at $t-1$. Players do not play mixed strategies. For the characterization of stochastically stable outcomes, we consider an evolutionary model of *best response with mutations* (BRM) where, with a fixed small probability ε, independent across players and across time, players choose any pure strategy at random, and with the complementary probability $1-\varepsilon$, they choose a strategy from the best response set. The following subsections formally outline these concepts.

2.1 The coordination game

We consider m-strategy 2-player symmetric strict coordination games with a set of pure strategies denoted by $A = \{a_1, \cdots, a_j, \cdots, a_m\}$, and identical to all players. Let U be an $m \times m$ payoff matrix with the elements $u(a_j, a_l)$, for all $a_j, a_l \in A$, being the payoff to a player playing strategy a_j against an opponent playing strategy a_l. The matrix U constitutes a *symmetric* coordination game if $u(a_j, a_l)$, for all $a_j, a_l \in A$, is identical for all players. And it constitutes a *strict* coordination game if $u(a_j, a_j) > u(a_l, a_j)$ for all $a_j, a_l \in A$ and $a_l \neq a_j$. We refer to the double, (A, U), where U is symmetric and strict, as a *symmetric strict coordination game.*

Let Σ be the set of all distributions over A whereby, for any $a_j \in A$ and $\sigma \in \Sigma$, $\sigma(a_j)$ is the mass that σ places on a_j, and $\sum_{a_k \in A} \sigma(a_k) = 1$. We consider linear payoffs where the payoff to a player playing pure strategy a_j against distribution σ is given by $U(a_j \mid \sigma) = \sum_{a_k \in A} \sigma(a_k) u(a_j, a_k)$. The set of pure strategy best responses to σ 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 network game

The social network is represented by an exogenously given and fixed over time graph $G(N, E)$, with N representing the set of players and E the set of edges linking different pairs of players.[13] Let N_i be the set of *first-order neighbours* of i – the set of players that i directly interacts with in $G(N, E)$ – and let n_i be the cardinality of N_i. A directed *path* from player i to j is a finite sequence of players $(i_1, i_2, \cdots, i_r)$, with $i = i_1$ and $i_r = j$, that are connected through *distinct* directed links (i.e., no link appears more than once).

We focus on *unweighted, undirected* and *strongly connected* networks. A network is unweighted if a link between any pair of players $i, j \in N$ has a weight of one; it is undirected if the existence of a link from i to j implies existence of a reverse link from j to i; and it is strongly connected if there exists a directed path connecting any pair of players $i, j \in N$ and $i \neq j$.

---

[13]Where no confusion arises, we simply write G as a shorthand for $G(N, E)$.

Focusing on unweighted and undirected networks is for simplicity and we discuss in Section 5 how our results extend to weighted and directed networks. Requiring the network to be strongly connected is necessary for our results to hold. Networks that are not strongly connected may consist of disconnected subgroups, which affects convergence of an evolutionary process. By avoiding issues related to convergence, we can focus on examining structural properties of the network that determine uniqueness of stochastically stable outcomes. One of the network measures that is relevant to our analysis is the *diameter*. Let d_{ij} be the length of the shortest directed path from i to j (i.e. the distance from i to j), and $d_i = \max_{j \in N} d_{ij}$ be the maximum shortest distance from i to any other player. The diameter of $G(N, E)$ is $d(G) = \max_{i \in N} d_i$, which is the maximum shortest distance between any two players.

In analogy to the linear payoffs defined above, let $\sigma_i = (\sigma_i(a_1), \cdots, \sigma_i(a_m))$ be the distribution over A that represents the proportion of i's direct neighbours playing each pure strategy. Then the total payoff that i receives from playing pure strategy a_j against distribution σ_i is (players do not play mixed strategies)[14]

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

We refer to the quadruple (A, U, N, G), where each player's payoff is given by (1), for all $\sigma_i \in \Sigma$, as a *network game*.

2.3 Unperturbed evolutionary process

The unperturbed evolutionary process is a *myopic best response* process where players revise their strategies simultaneously and independently over discrete times, $t = 1, 2, \cdots$. At each t, each player plays the symmetric strict coordination game against the same set of direct neighbours and selects a pure strategy randomly from a set of strategies that maximize (1). Players are myopic in that the strategies chosen at t are those that maximize the payoffs of the profiles of strategies chosen by their neighbours at $t - 1$.

Formally, let $\mathbf{x} = (x^1, \cdots, x^n)$ denote a profile/configuration of pure strategies, where x^i is the pure strategy played by the ith 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 $\sigma_i(\mathbf{x}) = (\sigma_i(a_1; \mathbf{x}), \cdots, \sigma_i(a_m; \mathbf{x}))$ be the distribution over A representing the proportions of i's neighbours playing different strategies in profile $\mathbf{x}$. Let also $\mathbf{x}(t)$ be the strategy profile

---

[14]Since the network is unweighted, the linear total payoff represented by (1) implies that player i attaches a weight of $\frac{1}{n_i}$ to each neighbour's strategy. That is, if $u(a_j, x^k)$ is the payoff to a player playing strategy a_j against a neighbour, k, playing strategy x^k, then (1) can equivalently be expressed as $U_i(a_j \mid (x_1, \cdots, x_{n_i})) = \frac{1}{n_i} \sum_{k \in N_i} u(a_j, x^k)$, where $(x_1, \cdots, x_{n_i})$ is a vector of strategies adopted by i's direct neighbours.

at t, and $x^i(t)$ the respective ith strategy in profile $\mathbf{x}(t)$. In analogy to (1), $U_i\big(a_j \mid \sigma_i(\mathbf{x}_t)\big) = \sum_{a_l \in A} \sigma_i(a_l; \mathbf{x}_t) u(a_j, a_l)$ is player i's payoff from playing pure strategy a_j against the profile of strategies chosen by the neighbours at t. Accordingly, $BR(\sigma_i(\mathbf{x}(t)))$ is the best response set of $\sigma_i(\mathbf{x}(t))$ (i.e. the set of all strategies $a_j \in A$ that satisfy the inequality $U_i\big(a_j \mid \sigma_i(\mathbf{x}_t)\big) \geq U_i\big(a_l \mid \sigma_i(\mathbf{x}_t)\big)$, for all $a_l \in A$). Then, at period $t+1$, player i chooses strategy $a_j \in A$ with probability $P\big(x^i(t+1) = a_j | \sigma_i(\mathbf{x}(t))\big) = \frac{1}{|BR(\sigma_i(\mathbf{x}(t)))|}$ 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 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 evolutionary framework described above is a finite time homogeneous Markov chain (i.e., the associated transition probability matrix is independent of time) on the state space $\mathbf{X}$. It is time homogeneous because the four components, A, U, N and G, that make up the network game are independent of time. 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 step. We refer to the tuple (A, U, N, G, P) as an *unperturbed evolutionary process* on network G. The equilibrium behaviour of (A, U, N, G, P) is fully described by its *absorbing sets*. A subset of states, $W \subseteq \mathbf{X}$, is an absorbing set of (A, U, N, G, P) 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. Absorbing sets 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.

2.4 Perturbed evolutionary process and stochastic stability

Following the literature, a *perturbed evolutionary process* is derived from the unperturbed evolutionary process by adding rare mutations. We follow the evolutionary framework of best response with mutations (BRM) by Young (1993) and Ellison (2000), whereby, with a fixed small probability $\varepsilon > 0$, independent across players and time, a player chooses a strategy uniformly at random, and with a complementary 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 $P_\varepsilon\big(x^i(t+1) = a_j | \sigma_i(\mathbf{x}(t))\big) = \frac{\varepsilon}{m} + \frac{1-\varepsilon}{|BR(\sigma_i(\mathbf{x}(t)))|}$ if $a_j \in BR(\sigma_i(\mathbf{x}(t)))$, and $P_\varepsilon\big(x^i(t+1) = a_j | \sigma_i(\mathbf{x}(t))\big) = \frac{\varepsilon}{m}$ if $a_j \notin BR(\sigma_i(\mathbf{x}(t)))$.

Let P_ε be the transition probability matrix of the associated Markov chain, so that $P_\varepsilon(\mathbf{x}, \mathbf{y})$ is the probability that profile $\mathbf{x}$ is followed by $\mathbf{y}$. Specifically, let $c(\mathbf{x}, \mathbf{y})$ be the number of players for whom y^i, the ith 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 $b_i(\mathbf{x}) = |BR(\sigma_i(\mathbf{x}))|$.[15]

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

The parameter $c(\mathbf{x}, \mathbf{y})$, for any pair $\mathbf{x}, \mathbf{y} \in \mathbf{X}$, is also commonly referred to as the cost function of the stochastic evolutionary process. The cost, $c(\mathbf{x}, \mathbf{y})$, reflects how unlikely the direct transition from $\mathbf{x}$ to $\mathbf{y}$ is when ε is small. For this reason, $c(\mathbf{x}, \mathbf{y})$ is formally defined as

$$c(\mathbf{x}, \mathbf{y}) = -\lim_{\varepsilon \to 0} \varepsilon \ln P_\varepsilon(\mathbf{x}, \mathbf{y}) \qquad (3)$$

Substituting for the expression of the transition probabilities from (2), we see that the limit on the right hand side of (3) exists. The definition of the cost function in (3) is not limited to the BRM model. Sandholm (2010) applies this definition to derive the cost functions of evolutionary models where the transition probabilities assume a logit and probit probability structures.

The tuple $(A, U, N, G, P_\varepsilon)$ is here referred to as the *perturbed evolutionary process* on network G. The equilibrium behaviour of $(A, U, N, G, P_\varepsilon)$ is fully described by its *stationary distribution*, π_ε, which is the probability distribution over the state space, $\mathbf{X}$, 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 that $(A, U, N, G, P_\varepsilon)$ spends in $\mathbf{x}$ in the long run. The stationary distribution of $(A, U, N, G, P_\varepsilon)$ exists, and is unique, because $(A, U, N, G, P_\varepsilon)$ is *ergodic* (i.e., it is possible to get from every state to every other state with positive probability). The ergodocity of $(A, U, N, G, P_\varepsilon)$ follows because the incorporation of mutations that are independent across players and time ensures that $P_\varepsilon(\mathbf{x}, \mathbf{y}) > 0$, for all pairs $\mathbf{x}, \mathbf{y} \in \mathbf{X}$.

We are interested in identifying the long-run or stochastically stable set of states of $(A, U, N, G, P_\varepsilon)$, which is a set of states that maximize π_ε when ε is small. However, computing π_ε is difficult, and the standard approach is to focus on the *limit stationary distribution*, π^*, defined as $\pi^* = \lim_{\varepsilon \to 0} \pi_\varepsilon$. The reason for this is that π^*, which is easier to compute, provides an approximation to π_ε when ε is small. The limit stationary distribution exists and the set of states $\mathbf{A}^* \subset \mathbf{X}$, with $\pi^*(\mathbf{A}^*) > 0$, is called a *stochastically stable set* (Young, 1993; Ellison, 2000). It is known that the stochastically stable set of $(A, U, N, G, P_\varepsilon)$ is contained in $\mathbf{A}$ (Young, 1993). If $\mathbf{A}^* = \{\mathbf{x}^*\}$, then $\mathbf{x}^*$ is a stochastically stable state. Accordingly, a set of strategies $A^* \subseteq A$ played in $\mathbf{A}^*$ are said to be stochastically stable strategies.

To compute the stochastically stable set of $(A, U, N, G, P_\varepsilon)$, we employ the method

---

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

developed by Ellison (2000), which involves computing the *radius*, *coradius* and *modified coradius* of the *basins of attraction* of absorbing sets. The basin of attraction, $D(W)$, of $W \in \mathbf{A}$, is the set of initial states from which the unperturbed evolutionary process, (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, coradius and modified coradius are all cost functions that depend on the total costs of paths between states. Define a *path* from subset Z to W as a finite sequence of distinct states $(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_T)$ with $\mathbf{x}_1 \in Z$, $\mathbf{x}_T \in W$ and $\mathbf{x}_\tau \notin W$ for $2 \leq \tau \leq T - 1$. Let $S(Z, W)$ be the set of all paths from Z to W. The cost $c(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_T)$ of path $(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_T)$ is the sum of the costs of the transitions between pairs of states along $(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_T)$. That is,

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

Recall that each $c(\mathbf{x}_\tau, \mathbf{x}_{\tau+1})$, defined in (3), reflects how unlikely the direct transition from $\mathbf{x}_\tau$ to $\mathbf{x}_{\tau+1}$ is when ε is small. Accordingly, $c(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_T)$ reflects how unlikely it is for $(A, U, N, G, P_\varepsilon)$ to reach $\mathbf{x}_T$ from $\mathbf{x}_1$ through the path $(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_T)$ when ε is small.

The radius, $R(W)$, of W, is the minimum number of mutations needed to exit the basin of attraction of W. Put differently, it is the minimum cost of any path from W out of the basin of attraction of W. That is, $R(W) = \min_{(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_T) \in S(W, \mathbf{X} - D(W))} c(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_T)$. The greater the value of $R(W)$, the less likely it is for $(A, U, N, G, P_\varepsilon)$ to leave the basin of attraction of W when ε is small.

Let $C(Z, W) = \min_{(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_T) \in S(Z, W)} c(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_T)$ be the cost of the minimum cost path from Z to W. The *coradius*, $CR(W)$, of W is the maximum over all other states of the minimum number of mutations required to reach W. Put differently, it is the maximum over all $\mathbf{x} \notin W$ of the minimum cost of all paths from $\mathbf{x}$ to W. That is,

$$CR(W) = \max_{\mathbf{x} \notin W} \min_{(\mathbf{x}_1, \cdots, \mathbf{x}_T) \in S(\mathbf{x}, W)} c(\mathbf{x}_1, \cdots, \mathbf{x}_T) = \max_{\mathbf{x} \notin W} C(\mathbf{x}, W)$$

Note that since the cost of reaching W from any state within the basin of attraction of W is zero, the coradius of W can also be expressed as $CR(W) = \max_{\mathbf{x} \notin D(W)} C(\mathbf{x}, D(W))$. The smaller the value of $CR(W)$, the more likely it is for $(A, U, N, G, P_\varepsilon)$ to reach W from any other state when ε is small. Ellison (2000, Theorem 1) shows that the set of states, $\mathbf{A}^* \in \mathbf{A}$, with $R(\mathbf{A}^*) > CR(\mathbf{A}^*)$ (i.e. a set that is reachable from every other state at a smaller cost than is needed to leave its basin of attraction) is the unique stochastically stable set of $(A, U, N, G, P_\varepsilon)$.

To define the modified coradius, let $Z_1, Z_2, \cdots, Z_q$, with $\mathbf{x} \in D(Z_1)$, $Z_q \subseteq W$ and $Z_l \nsubseteq W$ for $l < q$, be a sequence of absorbing sets through which some path, $(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_T)$, from $\mathbf{x} = \mathbf{x}_1$ to $\mathbf{x}_T \in Z_q$ traverses consecutively. Define the *modified cost*, $c^*(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_T)$, of the path $(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_T)$ as the cost, $c(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_T)$, minus the radius of the intermediate limit sets through which the path $(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_T)$ passes. That is,

$$c^*(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_T) = c(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_T) - \sum_{l=2}^{q-1} R(Z_l) \qquad (4)$$

The modified coradius, $CR^*(W)$, of W is the maximum over all $\mathbf{x} \notin W$ of the minimum modified cost of all paths from $\mathbf{x}$ to W. That is,

$$CR^*(W) = \max_{\mathbf{x} \notin W} \min_{(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_T) \in S(\mathbf{x}, W)} c^*(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_T) = \max_{\mathbf{x} \notin W} C^*(\mathbf{x}, W) \qquad (5)$$

where $C^*(\mathbf{x}, W) = \min_{(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_T) \in S(\mathbf{x}, W)} c^*(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_T)$ is the *minimum modified cost* of all paths from $\mathbf{x}$ to W.

Unlike $CR(W)$ which is derived from the probabilities of the paths from any other state to W, the modified coradius, $CR^*(W)$, is derived from the waiting time of reaching W from any other state. Ellison (2000) shows that, when ε is small, the waiting time to reach any $W \subset \mathbf{X}$ from any other state is bounded from above by $\kappa \varepsilon^{-CR^*(W)}$, where κ is some constant. And the waiting time to exit the basin of attraction of $\mathbf{A}^* \subseteq \mathbf{A}$ is bounded from below by $\kappa_1 \varepsilon^{-R(\mathbf{A}^*)}$ for some constant κ_1. By relating the waiting time to the stationary distribution, Ellison (2000, Theorem 2) then shows that a set $\mathbf{A}^* \subseteq \mathbf{A}$ with $R(\mathbf{A}^*) > CR^*(\mathbf{A}^*)$ is uniquely stochastically stable.

3 How contagion affects stochastically stable outcomes

3.1 *p*-best response sets, contagion and the contagion threshold

We aim to establish the relationship between *p-best response/p-dominance* and stochastic stability of the evolutionary process $(A, U, N, G, P_\varepsilon)$. We derive the conditions under which the *minimal p-best response set* is uniquely stochastically stable on a given network. 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. And for the network game (A, U, N, G), 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. 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 σ 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 $\sigma_{A'} \geq p$, $BR(\sigma) \subseteq A'$. A'*

is a minimal p-best response set if A′ does not contain any proper subset that is a p-best response set.

The definition of a p-best response set in Definition 1 implies that the strategy set, A, is, trivially, a p-best response set of (A, U) for all $p \in [0, 1]$. A coordination game can have more than one p-best response set, and the minimal p-best response set is the one that does not contain another smaller subset that itself is a p-best response set.[16] We denote by A^p the minimal p-best response set of (A, U). Every coordination game has a unique minimal p-best response set when $p \leq \frac{1}{2}$ (Tercieux, 2006, Theorem 1). If, for some $p \leq \frac{1}{2}$, $A^p = \{a\}$, then a is a p-*dominant equilibrium*.

Although the predictions of stochastic stability will be coarse when A^p contains more than one strategy, it is useful to examine when strategies in A^p are stochastically stable because, firstly, many coordination games do not have a p-dominant equilibrium but every coordination game has a unique minimal p-best response set for $p \leq \frac{1}{2}$. Secondly, if only one strategy is indeed uniquely stochastically stable, then we would know that that strategy is in A^p. A more precise method can then be used to isolate the unique stochastically stable strategy from A^p. Thirdly, there are situations, such as technology adoption, where one may be interested in the diffusion of a combination of strategies. A firm that owns two technologies may be interested in the long-run market survival of both technologies if they are complementary or if producing both is cost-effective.

Consider the two examples of symmetric coordination games in Figure 1. For the game in Figure 1a, set $A' = \{a_2, a_3\}$ is a p-best response set for $p > \frac{5}{18}$, and for $p > \frac{3}{8}$, both $A' = \{a_2, a_3\}$ and $A'' = \{a_3\}$ are p-best response sets. The minimal p-best response set is $A^p = A''$, which is also a p-dominant equilibrium since it is a singleton set. The game in Figure 1b does not contain a p-dominant equilibrium for all $p \leq \frac{1}{2}$, but it contains a minimal p-best response set $A^p = \{a_2, a_3\}$, for $p > \frac{1}{3}$. For this scenario, our analysis below establishes the conditions under which the unique stochastically stable strategy is contained in $A^p = \{a_2, a_3\}$.

Given $A^p \subseteq A$, let $\mathbf{A}^p \subseteq \mathbf{A}$ be the set of all 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 $R(\mathbf{A}^p) > CR(\mathbf{A}^p), CR^*(\mathbf{A}^p)$) by exploiting the properties of the process of *contagion* in networks. We distinguish between two forms of contagion that are both used in the derivation of the main results: *full contagion* and *step-by-step contagion*. Full contagion is the spread of strategies in $A' \subset A$ to the whole network through best response starting from a group of players whose size is less than half the population size. Step-by-step contagion is derived by iteratively applying *partial contagion*, which is the spread of strategies in $A' \subset A$ through best response from a group

---

[16]The notion of p-best response also 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.

	a_1	a_2	a_3			a_1	a_2	a_3
a_1	10	0	0		a_1	10	0	0
a_2	6	11	8		a_2	6	11	8
a_3	5	8	13		a_3	0	8	13
		(a)					(b)	

Figure 1: (a) A 3×3 symmetric coordination game with strategy a_3 as the p-dominant equilibrium, for $p > \frac{3}{8}$; (b) A 3×3 symmetric coordination game with $A' = \{a_2, a_3\}$ as the minimal p-best response set, for $p > \frac{1}{3}$.

of players, say $N' \subset N$, whose size, n', is less than half the population size, to a group of players whose size is larger than n'. These three concepts are formally developed through the following series of definitions.

Definition 2. *A sequence of strategy profiles $\{\mathbf{x}_t\}_{t=0}^{\bar{t}}$ of (A, U, N, G, P), for some $\bar{t} \geq 2$, is a best response sequence if it satisfies the following properties: (i) for all $1 \leq t < \bar{t}$, there exists at least one $i \in N$ such that $x_t^i \neq x_{t-1}^i$; (ii) if $x_t^i \neq x_{t-1}^i$, then $x_t^i \in BR(\sigma_i(\mathbf{x}_{t-1}))$.*

According to Definition 2, a sequence of states is a best response sequence if, (i) for any pair of consecutive states $\mathbf{x}_t$ and $\mathbf{x}_{t+1}$ along the sequence, at least one player must play a strategy in state $\mathbf{x}_{t+1}$ that is different from a strategy played in state $\mathbf{x}_t$; (ii) players switch strategies through best response, that is, if i's strategy in state $\mathbf{x}_{t+1}$ is different from the strategy played in state $\mathbf{x}_t$, then x_{t+1}^i is a best response to $\sigma_i(\mathbf{x}_t)$.

Definition 3. *Let (A, U, N, G, P) start from some $\mathbf{x} \in \mathbf{X}$. Strategies in $A' \subset A$ spread through full contagion from a subgroup of players, $\mathcal{M}(A'; \mathbf{x}) \subset N$, if there exists some $\bar{t} \geq 2$ such that every best response sequence $\{\mathbf{x}_t\}_{t=0}^{\bar{t}}$ with $\mathbf{x}_0 = \mathbf{x}$ and $x_1^i \in A'$ for all $i \in \mathcal{M}(A'; \mathbf{x})$ satisfies $x_{\bar{t}}^i \in A'$ for all $i \in N$.*

According to Definition 3, when (A, U, N, G, P) starts from $\mathbf{x}$, strategies in A' spread through full contagion from a subset $\mathcal{M}(A'; \mathbf{x})$ if they spread to the whole network through best response after all players in $\mathcal{M}(A'; \mathbf{x})$ play strategies in A'. Let $\mu(A'; \mathbf{x})$ be the cardinality of $\mathcal{M}(A'; \mathbf{x})$ and let $\mu^*(A'; \mathbf{x})$ be the size of the smallest group of players from which strategies in A' can spread through full contagion starting from $\mathbf{x}$. Equivalently, $\mu^*(A'; \mathbf{x})$ is the minimum number of mutations needed to trigger evolution of (A, U, N, G, P) from $\mathbf{x}$ to some state in $\mathbf{A}'$, a set of absorbing sets containing only strategies in A'.

Definition 4. *Strategies in $A' \subset A$ are fully contagious on G if, starting from every $\mathbf{x} \notin \mathbf{A}'$, they can spread through full contagion from a group of $\mu^*(A'; \mathbf{x}) < \frac{n}{2}$ players.*

When $\max_{\mathbf{x} \notin \mathbf{A}^p} \mu^*(A^p; \mathbf{x}) \leq \mu^*$, we say that μ^* mutations to strategies in A^p trigger full contagion. From the definition of the coradius in Section 2.4, it follows from Definition

14

4 that when strategies in A^p are fully contagious on a given network, the coradius of $\mathbf{A}^p$ is bounded from above by $\max_{\mathbf{x} \notin \mathbf{A}^p} \mu^*(A^p; \mathbf{x})$. In the analysis that follows, we develop a network measure that determines when strategies in A^p can spread through full contagion. Once such a measure is defined, it is then possible to establish an upper bound of $\max_{\mathbf{x} \notin \mathbf{A}^p} \mu^*(A^p; \mathbf{x})$, and hence, the coradius of $\mathbf{A}^p$.

Although the notion of full contagion enables us to derive the upper bound of the coradius of $\mathbf{A}^p$, the coradius is a less accurate measure of the cost of reaching $\mathbf{A}^p$ from any other state compared to the modified coradius. In some networks, the value of $\max_{\mathbf{x} \notin \mathbf{A}^p} \mu^*(A^p; \mathbf{x})$, and hence, the coradius of $\mathbf{A}^p$, can be large (i.e. close to $\frac{n}{2}$), which in turn requires imposing strict assumptions on the network structure to ensure that $R(\mathbf{A}^p) > CR(\mathbf{A}^p)$. Using the modified coradius can mitigate this problem because in most cases, the modified coradius is less than the coradius of an absorbing set.

Recall that the modified coradius of $\mathbf{A}^p$ is the cost of the minimum cost path that passes along absorbing sets that are different from $\mathbf{A}^p$. The notion of step-by-step contagion captures the costs of evolution along sequences of absorbing sets, and can thus be used to place an upper bound on the modified coradius of $\mathbf{A}^p$. Let $N(A'; \mathbf{x})$ be the set of players playing strategies in A' in configuration $\mathbf{x}$, and let $n(A'; \mathbf{x})$ be the cardinality of $N(A'; \mathbf{x})$. Then partial and step-by-step contagion are formally defined as follows.

Definition 5. *Let (A, U, N, G, P) start from some $\mathbf{x} \in \mathbf{X}$ and let $\mathbf{x} \in D(W')$, where $W' \in \mathbf{A}$ and $W' \neq \mathbf{A}'$. Strategies in $A' \subset A$ spread through partial contagion from a subgroup of players, $\mathcal{M}(A'; \mathbf{x}; W) \subset N$, to some absorbing set $W \neq W'$, if there exists some $\bar{t} \geq 2$ such that every best response sequence $\{\mathbf{x}_t\}_{t=0}^{\bar{t}}$ with $\mathbf{x}_0 = \mathbf{x}$ and $x_1^i \in A'$ for all $i \in \mathcal{M}(A'; \mathbf{x}; W)$ satisfies $\mathbf{x}_{\bar{t}} \in W$ and $n(A'; \mathbf{x}_{\bar{t}}) > n(A'; \mathbf{x})$.*

Let $\mu(A'; \mathbf{x}; W)$ be the cardinality of $\mathcal{M}(A'; \mathbf{x}; W)$. Definition 5 states that strategies in A^p spread through partial contagion from a group $\mathcal{M}(A^p; \mathbf{x}; W)$ of players, where $\mathbf{x} \in D(W')$, to an absorbing set $W \neq W'$ if $\mu(A^p; \mathbf{x}; W)$ mutations to strategies in A^p trigger an iterative process of best response where (A, U, N, G, P) eventually converges to W. Let $\mu^*(A^p; \mathbf{x}; W)$ be the size of the smallest group of players from which strategies in A^p can spread through partial contagion starting from $\mathbf{x}$ to W. Equivalently, $\mu^*(A^p; \mathbf{x}; W)$ is the minimum number of mutations needed to trigger evolution of (A, U, N, G, P) from $\mathbf{x}$ to some state in W.

Definition 6. *Strategies in $A' \subset A$ are step-by-step contagious on G starting from $\mathbf{x} \notin \mathbf{A}'$, if there exists a sequence of absorbing sets $W_1, W_2, \cdots, W_J$, with $\mathbf{x} \in D(W_1)$ and $W_J = \mathbf{A}'$, and a corresponding sequence of strategy sets, $A_2, A_3, \cdots, A_J$, such that, for any $\mathbf{y} \in W_j$, strategies in A_{j+1} spread through partial contagion from $\mu^*(A_{j+1}; \mathbf{y}; W_{j+1}) < \frac{n}{2}$ players.*

Definition 6 states that strategies in A^p are step-by-step contagious on a given network if they spread to the whole network through iterative application of partial contagion. That is, starting from $\mathbf{x} \in D(W)$, there exists a sequence of absorbing sets

$W_1, \cdots, W_J$, with $W = W_1$ and $W_J = \mathbf{A}^p$, and a corresponding sequence of strategy sets $A_2, A_3, \cdots, A_J$, with $A_J = A^p$, such that along $W_1, \cdots, W_J$, (A, U, N, G, P) evolves from $\mathbf{x}$ to W_2 through best response once $\mu^*(A_2; \mathbf{x}; W_2) < \frac{n}{2}$ players play (mutation to) strategies in A_2; and from any $\mathbf{y} \in W_2$ to W_3 through best response once $\mu^*(A_3; \mathbf{y}; W_3) < \frac{n}{2}$ players play (mutation to) to strategies in A_3; and so on, until it eventually reaches $\mathbf{A}^p$. When $\mu^*(A_{j+1}; \mathbf{y}; W_{j+1}) \leq \mu^*$ for all $j = 1, \cdots, J - 1$, we say that μ^* mutations trigger step-by-step contagion of strategies in A^p starting from $\mathbf{x}$.

We develop a new network measure, the contagion threshold, that not only determines when strategies in A^p can spread through full contagion but also when they can spread through partial, and hence, step-by-step contagion. The definition of the contagion threshold is based on the following reinterpretation of the definitions of full and partial contagion in Definitions 3 and 5. Given a network G and any subset of players $S \subset N$, let $N_i(S) = N_i \cap S$ be the set of i's direct neighbours within G that belong to subgroup S. Let $n_i(S)$ be the corresponding cardinality of $N_i(S)$ and $\alpha_i(S) = \frac{n_i(S)}{n_i}$ be the proportion of i's neighbours in S.

From Definition 3, strategies in A^p spread through full contagion if, for every best response sequence $\{\mathbf{x}_\tau\}_{\tau=0}^{\bar{t}}$ satisfying the conditions in Definition 3, there exists a sequence of players $S_1, S_2, \cdots, S_{\bar{t}}$, with $S_1 = \mathcal{M}(A^p; \mathbf{x})$ and $\cup_{\tau=1}^{\bar{t}} S_\tau = N$, such that for each $j \in S_\tau$, and all $\tau = 2, 3, \cdots, \bar{t}$, $\alpha_j(\cup_{k=1}^{\tau-1} S_k) \geq p$. That is, every player in S_τ has at least proportion p of their neighbours in $\cup_{k=1}^{\tau-1} S_k$. The inequality $\alpha_j(\cup_{k=1}^{\tau-1} S_k) \geq p$, for all $j \in S_\tau$ and $\tau = 2, 3, \cdots, \bar{t}$, ensures that once all players in $S_1 = \mathcal{M}(A^p; \mathbf{x})$ play strategies in A^p, all players in $S_1 \cup S_2$ subsequently play strategies in A^p since each $j \in S_2$ has at least proportion p of her neighbours in S_1; all players in $S_1 \cup S_2 \cup S_2$ subsequently play strategies in A^p since each $j \in S_2 \cup S_3$ has at least proportion p of her neighbours in $S_1 \cup S_2$; this iterative process continues until the whole network eventually plays strategies in A^p.

Similarly, from Definition 5, strategies in A^p spread through partial contagion from a group of players, $\mathcal{M}(A^p; \mathbf{x}; W')$, where $\mathbf{x} \in D(W)$, to $W' \neq W$ if, for every best response sequence $\{\mathbf{x}_\tau\}_{\tau=0}^{\bar{t}}$ satisfying the conditions in Definition 5, there exists a sequence of players $S_1, S_2, \cdots, S_{\bar{t}}$, with $S_1 = \mathcal{M}(A^p; \mathbf{x}; W') = N(A^p; \mathbf{x}_1)$, $S_2 = N(A^p; \mathbf{x}_2), \cdots, S_{\bar{t}} = N(A^p; \mathbf{x}_{\bar{t}})$, such that for each $j \in S_\tau$, and all $\tau = 2, 3, \cdots, \bar{t}$, $\alpha_j(S_{\tau-1}) \geq p$. These sequences of players have two main properties: first, the sets S_τ, for $\tau = 1, \cdots, \bar{t}$, need not be disjoint sets, that is, for any pair S_τ and $S_{\tau+1}$ along the sequence, it is not necessary for $S_\tau \cap S_{\tau+1} = \emptyset$ to hold; second, $S_{\bar{t}} \subseteq N$ (i.e. $S_{\bar{t}}$ is a subset of N) because for partial contagion, strategies in A^p need not spread to the whole network.

Thus, the first step in defining the network measure that determines when strategies in A^p can spread though full and step-by-step contagion involves choosing a set of players from which contagion can be triggered. We identify the *first-neighbourhoods* of players (i.e. the set of direct neighbours of each $i \in N$, with i included) as groups of players from

which strategies in A^p can spread through full and partial contagion.[17] For each $i \in N$, let B_{i_r} be the *rth-neighbourhood* of i (i.e. the set of all players within r steps from i, with i included) and N_{i_r} be the *r-order neighbours* of i (i.e. all players that are located at r steps from i). These definitions imply that $B_{i_r} = i \cup N_{i_1} \cup N_{i_2} \cup \cdots \cup N_{i_r}$. Let b_{i_r} and n_{i_r} be the respective cardinalities of B_{i_r} and N_{i_r}.

Strategies in A^p can spread from B_{i_1} of any $i \in N$ through either full or partial contagion if the corresponding sequence, $B_{i_1}, N_{i_2}, \cdots, N_{i_{d_i}}$, satisfies the conditions outlined in the above reinterpretations of Definitions 3 and 5. That is, starting from any $\mathbf{x} \notin D(\mathbf{A}^p)$, b_{i_1} mutations to strategies in A^p by all players in B_{i_1} can trigger full contagion, if, along the sequence $B_{i_1}, N_{i_2}, \cdots, N_{i_{d_i}}$, $\alpha_j(B_{i_{\tau-1}}) \geq p$ for each $j \in N_{i_\tau}$ and all $\tau = 2, \cdots, d_i$ (i.e. $\min_{\tau \in [2,d_i]} \min_{j \in N_{i_\tau}} \alpha_j(B_{i_{\tau-1}}) \geq p$). Similarly, b_{i_1} mutations to strategies in A^p by all players in B_{i_1} can trigger partial contagion, if, along the sequence, $S_1, S_2, \cdots, S_{\bar{t}}$ with $S_1 = B_{i_1}$, $S_2 \subseteq B_{i_2}, \cdots, S_{\bar{t}} \subseteq B_{i_{\bar{t}}}$, $\alpha_h(S_{\tau-1}) \geq p$ for each $h \in S_\tau$ and all $\tau = 2, \cdots, \bar{t}$. The following lemma shows that such sequences are guaranteed to exist whenever $\min_{\tau \in [2,d_i]} \min_{j \in N_{i_\tau}} \alpha_j(B_{i_{\tau-1}}) \geq p$.

Lemma 1. *Let G be unweighted and strongly connected. If $\min_{\tau \in [2,d_i]} \min_{j \in N_{i_\tau}} \alpha_j(B_{i_{\tau-1}}) \geq p$ for all $i \in N$, then there exist sequences of the form $S_1 = B_{i_1}$, $S_2 \subseteq B_{i_2}, S_3 \subseteq B_{i_3}, \cdots, S_{\bar{t}} \subseteq B_{i_{\bar{t}}}$, such that $\alpha_h(S_{\tau-1}) \geq p$ for all $h \in S_\tau$ and all $\tau = 2, \cdots, \bar{t}$.*

Proof. See Appendix B

$\square$

Thus, the sufficient condition for strategies in A^p to spread through either full or partial contagion from B_{i_1} is $\min_{\tau \in [2,d_i]} \min_{j \in N_{i_\tau}} \alpha_j(B_{i_{\tau-1}}) \geq p$. We then define the contagion threshold, $\underline{\eta}(G)$, as the minimum $\min_{\tau \in [2,d_i]} \min_{j \in N_{i_\tau}} \alpha_j(B_{i_{\tau-1}})$ over all $i \in N$ such that strategies in A^p spread from B_{i_1} of any $i \in N$ through either full or partial contagion whenever $\underline{\eta}(G) \geq p$.

Definition 7. *The contagion threshold of a strongly connected network, G, is defined as $\underline{\eta}(G) = \min_{i \in N} \min_{r \in [2,d_i]} \min_{j \in N_{i_r}} \alpha_j(B_{i_{r-1}})$.*

The definition of the contagion threshold in Definition 7 is specific to finite networks but shares some similarities with the definition of the contagion threshold for unbounded networks according to Morris (2000). Morris (2000) defines the contagion threshold of an unbounded network as the maximum p such that a p-dominant strategy spreads to the whole network through best response starting from a finite group of players. The

---

[17]The first-neighbourhood is not unique in this role. For example, *cohesive subgroups* (i.e. groups of players where each member of the group has at least half of her interactions with other group members) can also be used as groups of players from which strategies in A^p spread through full and partial contagion (see Morris (2000) and Opolot (2020) for a formal definition of group cohesion and how group cohesion determines when strategies can spread through contagion).

similarity between this definition and Definition 7 is that they both determine when a p-dominant strategy is fully contagious. The difference concerns the definition of the group of players from which a p-dominant strategy spreads contagiously. To be able to establish the bounds of the coradius, modified coradius and radius of $\mathbf{A}^p$ in finite networks, it is necessary to precisely define the set and the number of players from which contagion can be triggered. It is not sufficient, as in the case of unbounded networks, to require strategies in A^p to spread contagiously from a finite group of players without defining the composition and the size of such a group.

3.2 Contagion, coradius, modified coradius and the radius of $\mathbf{A}^p$

This subsection discusses how full and step-by-step contagion determine the upper bounds of the coradius and modified coradius of $\mathbf{A}^p$, and the lower bound of the radius of $\mathbf{A}^p$. By definition of the contagion threshold, when $\underline{\eta}(G) \geq p$, strategies in A^p spread from the first-neighbourhood of any player through either full or partial contagion. When full contagion is feasible on an undirected, unweighted and strongly network G, the coradius of $\mathbf{A}^p$ is bounded from above by $\max_{\mathbf{x} \notin \mathbf{A}^p} \mu^*(A^p; \mathbf{x})$. And since full contagion can be triggered from b_{i_1} of any $i \in N$, it follows that $CR(\mathbf{A}^p) \leq \max_{\mathbf{x} \notin \mathbf{A}^p} \mu^*(A^p; \mathbf{x}) \leq \min_{i \in N} b_{i_1} = \underline{b}_1$ for any G where full contagion of strategies in A^p is feasible.

To see the connection between step-by-step contagion and the modified coradius of $\mathbf{A}^p$, consider any $\mathbf{x} \in D(W)$ and $W \neq \mathbf{A}^p$. If $W_1, \cdots, W_J$, with $W_1 = W$ and $W_J = \mathbf{A}^p$, is the sequence of absorbing sets along which some path, $(\mathbf{x}_1, \cdots, \mathbf{x}_T)$, from $\mathbf{x} = \mathbf{x}_1$ to $\mathbf{x}_T \in \mathbf{A}^p$ passes, then $(\mathbf{x}_1, \cdots, \mathbf{x}_T)$ can be partitioned into sub-paths between absorbing sets. That is, $(\mathbf{x}_1, \cdots, \mathbf{x}_T)$ can be partitioned into $\bigcup_{j=1}^{J-1}(\mathbf{x}_{j_1}, \cdots, \mathbf{x}_{j_{T'}})$, where, for $j = 1, \cdots, J - 1$, $(\mathbf{x}_{j_1}, \cdots, \mathbf{x}_{j_{T'}})$ is the sub-path of $(\mathbf{x}_1, \cdots, \mathbf{x}_T)$ that starts from $\mathbf{x}_{j_1} \in W_j$ and ends at $\mathbf{x}_{j_{T'}} \in W_{j+1}$.[18] From (4), the modified cost $c^*(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_T)$ of path $(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_T)$ can then be rewritten as:

$$c^*(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_T) = c(\mathbf{x}_{1_1}, \cdots, \mathbf{x}_{1_{T'}}) + \sum_{j=2}^{J-1} c(\mathbf{x}_{j_1}, \cdots, \mathbf{x}_{j_{T'}}) - \sum_{j=2}^{J-1} R(W_j)$$

Let $\Gamma(\mathbf{x}, \mathbf{A}^p)$ be the set of all sequences of absorbing sets along which paths from $\mathbf{x}$ to

---

[18]To avoid notational clutter, we write $(\mathbf{x}_{j_1}, \cdots, \mathbf{x}_{j_{T'}})$ for a typical path starting from W_j and ending at W_{j+1} whenever the sequence $W_1, \cdots, W_J$ is clearly defined. This notation should not be interpreted to mean that all paths from W_j to W_{j+1}, and for all $j = 1, \cdots, J - 1$, are of the same length, T'. These paths can have different lengths. A more explicit notation would involve adding a quantifier for each T', but this would lead to notational clutter.

$\mathbf{A}^p$ traverse. Then the minimum modified cost of all paths from $\mathbf{x}$ to $\mathbf{A}^p$ becomes

$$C^*(\mathbf{x}, \mathbf{A}^p) = \min_{(W_1, \cdots, W_J) \in \Gamma(\mathbf{x}, \mathbf{A}^p)} \left[\min_{(\mathbf{x}_{1_1}, \cdots, \mathbf{x}_{1_{T'}}) \in S(\mathbf{x}, W_2)} c(\mathbf{x}_{1_1}, \cdots, \mathbf{x}_{1_{T'}}) \right.$$
$$\left. + \sum_{j=2}^{J-1} \min_{(\mathbf{x}_{j_1}, \cdots, \mathbf{x}_{j_{T'}}) \in S(W_j, W_{j+1})} c(\mathbf{x}_{j_1}, \cdots, \mathbf{x}_{j_{T'}}) - \sum_{j=2}^{J-1} R(W_j) \right] \qquad (6)$$

By definition of partial contagion above, we have

$$\min_{(\mathbf{x}_{j_1}, \cdots, \mathbf{x}_{j_{T'}}) \in S(W_j, W_{j+1})} c(\mathbf{x}_{j_1}, \cdots, \mathbf{x}_{j_{T'}}) = \min_{\mathbf{y} \in W_j} \mu^*(A_{j+1}; \mathbf{y}; W_{j+1})$$

$$C^*(\mathbf{x}, \mathbf{A}^p) = \min_{(W_1, \cdots, W_J) \in \Gamma(\mathbf{x}, \mathbf{A}^p)} \left[\mu^*(A_2; \mathbf{x}; W_2) + \sum_{j=2}^{J-1} \left(\min_{\mathbf{y} \in W_j} \mu^*(A_{j+1}; \mathbf{y}; W_{j+1}) - R(W_j) \right) \right]$$
$$\qquad (7)$$

We demonstrate in Section 4 below that when strategies in A^p spread through step-by-step contagion, then, along the sequence of absorbing sets, $W_1, W_2, \cdots, W_J$, that the minimum modified cost path from $\mathbf{x} \in D(W_1)$ to $\mathbf{A}^p$ traverses, $R(W_j) = \min_{\mathbf{y} \in W_j} \mu^*(A_{j+1}; \mathbf{y}; W_{j+1})$. More specifically, we show that when strategies in A^p are not fully contagious, there are two other possible categories of paths from any $\mathbf{x} \in D(W)$ to $\mathbf{A}^p$. The first category of paths traverse through a sequence of only three absorbing sets, $W, W', \mathbf{A}^p$, where W' is an absorbing cycle containing both strategies in A^p and $A \backslash A^p$. Along this sequence, $R(W) = \min_{\mathbf{y} \in W} \mu^*(A^p; \mathbf{y}; W')$ and $R(W') = \min_{\mathbf{y} \in W'} \mu^*(A^p; \mathbf{y}; \mathbf{A}^p)$. The second category of paths traverse through a sequence of four absorbing sets, $W, W', W'', \mathbf{A}^p$, where W' and W'' are absorbing cycles containing both strategies in A^p and $A \backslash A^p$. Along this sequence, $R(W) = \min_{\mathbf{y} \in W} \mu^*(A^p; \mathbf{y}; W')$, $R(W') = \min_{\mathbf{y} \in W'} \mu^*(A \backslash A^p; \mathbf{y}; W'')$ and $R(W'') = \min_{\mathbf{y} \in W''} \mu^*(A^p; \mathbf{y}; \mathbf{A}^p)$. For both categories of paths, it follows from (7) that $C^*(\mathbf{x}, \mathbf{A}^p) = \min_{W' \neq W} \mu^*(A^p; \mathbf{x}; W') \leq \underline{b}_1$, where the inequality follows because the partial contagion of strategies in A^p can be triggered from the first-neighbourhood of any player. The modified coradius of $\mathbf{A}^p$ then becomes

$$CR^*(\mathbf{A}^p) = \max_{\mathbf{x} \notin D(\mathbf{A}^p)} C^*(\mathbf{x}, \mathbf{A}^p) = \max_{\mathbf{x} \notin D(\mathbf{A}^p)} \min_{W' \neq W} \mu^*(A^p; \mathbf{x}; W') \leq \underline{b}_1 \qquad (8)$$

Contagion not only determines the upper bounds of the coradius and modified coradius of $\mathbf{A}^p$ but also the lower bound of its radius. We demonstrate that when strategies in A^p are fully or step-by-step contagious, the diameter of the network determines when the lower bound of the radius of $\mathbf{A}^p$ is greater than $\underline{b}_1$. Specifically, we show that when (A, U, N, G, P) starts from any $\mathbf{x} \in \mathbf{A}^p$, and $d(G) \geq 7$, b_{i_1} mutations to strategies in $A \backslash A^p$ by all players in B_{i_1} of any $i \in N$ are not sufficient to trigger an exit from the basin of

attraction of $\mathbf{A}^p$.

The underlying intuition is that when $d(G) \geq 7$, each $i \in N$ has $d_i \geq 4$ (see Lemma 2 for the proof), which implies that for all $j \in N_{i_4} \cup N_{i_5} \cup \cdots \cup N_{i_{d_i}}$, $B_{j_2} \cap B_{i_1} = \emptyset$ (i.e. there is no overlap between the first-neighbourhood of i and the second-neighbourhood of j). This implies that, starting from any $\mathbf{x} \in \mathbf{A}^p$, if all players in B_{i_1} mutate to strategies in $A \backslash A^p$, there exists at least one player $j \in N_{i_4} \cup N_{i_5} \cup \cdots \cup N_{i_{d_i}}$ for whom all players in B_{j_2} play strategies in A^p. The definition of the contagion threshold implies that when $\underline{\eta}(G) \geq p$, (A, U, N, G, P) converges to $\mathbf{A}^p$ from any configuration where all players in B_{j_2} of any $j \in N$ play strategies in A^p (i.e., b_{j_2} mutations to strategies in A^p sufficiently trigger full contagion on any unweighted and strongly connected network).[19] Thus, when $\underline{\eta}(G) \geq p$ and $d(G) \geq 7$, more than b_{i_1} mutations are required to trigger an exit from the basin of attraction of $\mathbf{A}^p$ so that $R(\mathbf{A}^p) > \underline{b}_1 \geq CR(\mathbf{A}^p), CR^*(\mathbf{A}^p)$.

4 Stochastic stability of the minimal p-best response set

This section presents the main results of this paper. The following Theorem provides the conditions under which strategies in A^p are uniquely stochastically stable.

Theorem 1. *Given an evolutionary process of best response with mutations, $(A, U, N, G, P_\varepsilon)$, on an undirected, unweighted and strongly connected network $G(N, E)$, the minimal p-best response set $A^p \subseteq A$ is uniquely stochastically stable whenever $\underline{\eta}(G) \geq p$ and $d(G) \geq 7$.*

Proof. The following steps provide a sketch and the intuition of the proof of Theorem 1. A detailed exposition is provided in Appendix C. The proof involves deriving the upper bounds of the coradius and modified coradius, and the lower bound of the radius of $\mathbf{A}^p$. It exploits the implications of the process of full and step-by-step contagion discussed in Section 3.

We show that the upper bound of both the coradius and modified coradius of $\mathbf{A}^p$ is $\underline{b}_1 = \min_{i \in N} b_{i_1}$. To prove this result, we first show that when (A, U, N, G, P) starts from any $\mathbf{x} \notin \mathbf{A}^p$, b_{i_1} mutations to strategies in A^p, for any $i \in N$, trigger either full contagion of strategies in A^p or partial contagion where (A, U, N, G, P) converges to an absorbing cycle of states containing both strategies in A^p and $A \backslash A^p$. For networks where b_{i_1} mutations trigger full contagion, we have $C(\mathbf{x}, \mathbf{A}^p) = \mu^*(A^p; \mathbf{x}) \leq \min_{i \in N} b_{i_1} = \underline{b}_1$ and $CR(\mathbf{A}^p) = \max_{\mathbf{x} \in \mathbf{X} \backslash \mathbf{A}^p} C(\mathbf{x}, \mathbf{A}^p) \leq \underline{b}_1$.

---

[19]To see why, let (A, U, N, G, P) start, at $t = 0$, from any $\mathbf{x} \notin D(\mathbf{A}^p)$ where all players in B_{j_2} of some $j \in N$ play strategies in A^p. Then, from $t = 1$ onward, all players in B_{j_3} play strategies in A^p since $\underline{\eta}(G) \geq p$ implies that each $h \in B_{j_3}$ has $\alpha_h(B_{j_2}) \geq p$. From $t = 2$ onward, all players in B_{j_4} play strategies in A^p because each $h \in B_{j_4}$ has $\alpha_h(B_{j_3}) \geq p$; from $t = 3$ onward, all players in B_{j_5} play strategies in A^p because each $h \in B_{j_5}$ has $\alpha_h(B_{j_4}) \geq p$; this iterative process continues until the whole network eventually plays strategies in A^p.

Now, let $\mathcal{C}(\mathbf{A})$ be the set of all absorbing cycles of states containing both strategies in A^p and some strategies in $A \backslash A^p$. Let $N_{i_\tau}^r$, for $\tau < r$, $\tau = 1, 2, \cdots, d_i - 1$ and $r = 2, \cdots, d_i$, be the set of players in N_{i_τ} that lie along the paths of length greater or equal to r. Put differently, $N_{i_\tau}^r$ is the set of all $h \in N_{i_\tau}$ with $N_{h_{r-\tau}} \cap N_{i_r} \neq \emptyset$ (i.e. the set of players in N_{i_τ} with at least one $(r - \tau)$-order neighbour in N_{i_r}). For example, $N_{i_1}^3$ is the set of all $h \in N_{i_1}$ with at least one second-order neighbour in N_{i_3} (i.e. $N_{h_2} \cap N_{i_3} \neq \emptyset$).[20] Let $\bar{N}_{i_\tau}^r = N_{i_\tau} \backslash N_{i_\tau}^r$ be the set of all players in N_{i_τ} that are not contained in $N_{i_\tau}^r$, and $n_{i_\tau}^r$ and $\bar{n}_{i_\tau}^r$ be the respective cardinalities of $N_{i_\tau}^r$ and $\bar{N}_{i_\tau}^r$.

We show that when b_{i_1} mutations to strategies in A^p fail to trigger full contagion, (A, U, N, G, P) converges to an absorbing cycle, $W \in \mathcal{C}(\mathbf{A})$, where players in $N_{i_1}^3 \cup N_{i_3} \cup N_{i_5} \cup \cdots \cup N_{i_{d_i}}$ and $i \cup N_{i_2}^3 \cup N_{i_4} \cup N_{i_6} \cup \cdots \cup N_{i_{d_i-1}}$ alternate between strategies in A^p and some strategies in A (i.e., including some strategies in $A \backslash A^p$), and the rest play some strategies in A. We then show that, firstly, $n_{i_1}^r$ mutations to strategies in A^p, for any $4 \leq r \leq d_i$, sufficiently trigger an exit from the basin of attraction of W to some state in $\mathbf{A}^p$. Since this result holds for any $i \in N$, including player $\underline{i} = i \in \arg\min b_{i_1}$ with the fewest number of direct neighbours, it follows that $C(W, \mathbf{A}^p) \leq \underline{n}_1^* \leq \underline{n}_1^r$, where $\underline{n}_1^r = \min_{i \in N} n_{i_1}^r$ and $\underline{n}_1^* = \min_{i \in N} \min_{r \in [4, d_i]} n_{i_1}^r$.

Secondly, if (A, U, N, G, P) exits the basin of attraction of W after less or equal to n_1^* mutations to strategies in $A \backslash A^p$, then it converges to an absorbing cycle where players in $N_{i_1}^4 \cup N_{i_3}^4 \cup N_{i_5} \cup \cdots \cup N_{i_{d_i}}$ and $i \cup N_{i_2}^4 \cup N_{i_4} \cup N_{i_6} \cup \cdots \cup N_{i_{d_i-1}}$ alternate between strategies in A^p and some strategies in A, and the rest play some strategies in A. Denote this absorbing cycle by $W' \in \mathcal{C}(\mathbf{A})$. This implies that either $R(W) = C(W, \mathbf{A}^p)$ (i.e. when $C(W, \mathbf{A}^p) \leq C(W, W')$), or $R(W) = C(W, W')$ (i.e. when $C(W, W') \leq C(W, \mathbf{A}^p)$).

Thirdly, we show that $n_{i_1}^4$ mutations to strategies in A^p sufficiently trigger an exit from the basin of attraction of W' to $\mathbf{A}^p$. However, $n_{i_1}^4$ mutations to strategies in $A \backslash A^p$ cannot to trigger an exit from the basin of attraction of W'. This implies that $R(W') = C(W', \mathbf{A}^p) \leq \underline{n}_1^4$.

Thus, there are three possible minimum cost paths from any $\mathbf{x} \notin D(\mathbf{A}^p)$ to $\mathbf{A}^p$. First, $\underline{b}_1$ mutations to strategies in A^p trigger full contagion from any $\mathbf{x} \notin D(\mathbf{A}^p)$ so that $CR(\mathbf{A}^p) \leq \underline{b}_1$. Second, (A, U, N, G, P) converges to some absorbing cycle $W \in \mathcal{C}(\mathbf{A})$ after $\underline{b}_1$ mutations to strategies in A^p, and that the number of mutations to strategies in $A \backslash A^p$ needed to trigger an exit from $D(W)$ is greater than $\underline{n}_1^*$. For this scenario,

---

[20]The definition of $N_{i_\tau}^r$ implies that, for $\tau < s \leq r$, $\tau = 1, \cdots, d_i - 1$, $s = 2, \cdots, d_i$ and $r = 2, \cdots, d_i$, $N_{i_\tau}^r \subseteq N_{i_\tau}^s$. For example, all $j \in N_{i_1}^4$ are also contained in $N_{i_1}^3$ and $N_{i_1}^2$, but not vice versa.

$R(W) = C(W, \mathbf{A}^p)$ so that

$$
\begin{aligned}
C^*(\mathbf{x}, \mathbf{A}^p) &= \min_{W \in \mathcal{C}(\mathbf{A})} \Bigg[\min_{(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_{T'}) \in S(\mathbf{x}, W)} c(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_{T'}) \\
&\quad + \min_{(\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_T) \in S(W, \mathbf{A}^p)} c(\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_T) - R(W) \Bigg] \\
&= \min_{W \in \mathcal{C}(\mathbf{A})} \Big[C(\mathbf{x}, W) + C(W, \mathbf{A}^p) - R(W) \Big] \\
&= \min_{W \in \mathcal{C}(\mathbf{A})} C(\mathbf{x}, W) \le \underline{b}_1
\end{aligned}
\tag{9}
$$

Third, (A, U, N, G, P) converges to some absorbing cycle $W \in \mathcal{C}(\mathbf{A})$ after $\underline{b}_1$ mutations to strategies in A^p, but the number of mutations to strategies in $A \backslash A^p$ needed to trigger an exit from $D(W)$ is less or equal to $\underline{n}_1^*$. For this scenario, there exists another absorbing cycle, $W' \in \mathcal{C}(\mathbf{A})$ and $W' \ne W$, where $C(W, W') \le C(W, \mathbf{A}^p)$ and $R(W) = C(W, W')$. And following the above discussion, $R(W') = C(W', \mathbf{A}^p)$ so that

$$
\begin{aligned}
C^*(\mathbf{x}, \mathbf{A}^p) &= \min_{W, W' \in \mathcal{C}(\mathbf{A})} \Bigg[\min_{(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_{T'}) \in S(\mathbf{x}, W)} c(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_{T'}) \\
&\quad + \min_{(\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_T) \in S(W, W')} c(\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_T) - R(W) \\
&\quad + \min_{(\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_T) \in S(W', \mathbf{A}^p)} c(\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_T) - R(W') \Bigg] \\
&= \min_{W, W' \in \mathcal{C}(\mathbf{A})} \Big[C(\mathbf{x}, W) + C(W, W') - R(W) + C(W', \mathbf{A}^p) - R(W') \Big] \\
&= \min_{W \in \mathcal{C}(\mathbf{A})} C(\mathbf{x}, W) \le \underline{b}_1
\end{aligned}
\tag{10}
$$

Since equations (9) and (10) hold for any $\mathbf{x} \notin D(\mathbf{A}^p)$ for which $\underline{b}_1$ cannot trigger full contagion, it follows that $CR^*(\mathbf{A}^p) = \max_{\mathbf{x} \notin D(\mathbf{A}^p)} \min_{W \in \mathcal{C}(\mathbf{A})} C(\mathbf{x}, W) \le \underline{b}_1$. The following examples help to elaborate these steps.
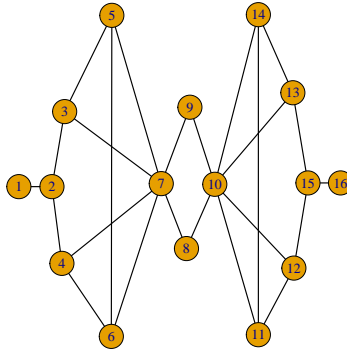

Figure 2: An example of a network with contagion threshold (computed using the matrix-based steps in Appendix A) of $\underline{\eta}(G) = \frac{2}{7}$ and $d(G) = 8$.

Example 1. *Consider the network depicted in Figures 2, with the contagion threshold*

and diameter of $\underline{\eta}(G) = \frac{2}{7}$ and $d(G) = 8$ respectively. The players with the smallest first-neighbourhood are 1 and 17, with $b_{1_1} = b_{17_1} = \underline{b}_1 = 2$. Let $p \leq \frac{2}{7}$, and let (A, U, N, G, P) start from some configuration $\mathbf{x} \notin D(\mathbf{A}^p)$ (i.e. a configuration containing both strategies in A^p and $A \backslash A^p$). If 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:

$t = 1$	*All players in $B_{1_1} = \{1, 2\}$ play strategies in A^p; all other players play strategies in A.*
$t = 2$	*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 $\alpha_2(B_{1_1}) = \frac{1}{3} > p$. All $j \in N_{1_2} = \{3, 4\}$ also play strategies in A^p because each has $\alpha_j(B_{1_1}) \geq \frac{1}{3} > \frac{2}{7} \geq p$.[21] All other players play strategies in A.*
$t = 3$	*All $j \in B_{1_3} = \{1, 2, 3, 4, 5, 6, 7\}$ play strategies in A^p because each has $\alpha_j(B_{1_2}) \geq \frac{2}{7} \geq p$ and that all players in B_{1_2} play strategies in A^p at $t = 2$. All other players play strategies in A.*
$--$	$-------------------------------$
$t = 8$	*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.*

Thus, for this network, $\underline{b}_1$ mutations to strategies in A^p sufficiently trigger full contagion, which implies that $C(\mathbf{x}, \mathbf{A}^p) \leq \underline{b}_1$ for any $\mathbf{x} \notin D(\mathbf{A}^p)$, and that $CR(\mathbf{A}^p) = \max_{\mathbf{x} \notin D(\mathbf{A}^p)} C(\mathbf{x}, \mathbf{A}^p) \leq \underline{b}_1$.
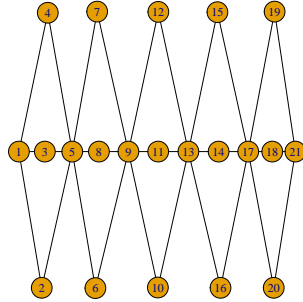

Figure 3: An example of a network with contagion threshold of $\underline{\eta}(G) = \frac{1}{2}$ and $d(G) = 10$.

Example 2. *Consider the network in Figure 3, with the contagion threshold of $\underline{\eta}(G) = \frac{1}{2}$ and diameter of $d(G) = 10$. Let (A, U, N, G, P) start from some configuration $\mathbf{x}$ where all players play strategies in $A \backslash A^p$, and let players in $B_{3_1} = \{1, 3, 5\}$ mutate to strategies in A^p at $t = 1$. For $p = \frac{1}{2} = \underline{\eta}(G)$, (A, U, N, G, P) evolves from $t = 2$ onward as follows:*

---

[21]Note also 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 $A \backslash A^p$ 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 a player's neighbours.

$t=2$	Players in $3 \cup N_{3_2} = \{2,3,4,6,7,8\}$ play strategies in A^p since each has at least proportion $\frac{1}{2} = p$ of neighbours in B_{3_1}. All $j \in N_{3_1} = \{1,5\}$ switch back to strategies in $A \backslash A^p$ because each has $\alpha_j(B_{3_1}) = \frac{1}{3} < p$. All other players play strategies in $A \backslash A^p$.
$t=3$	Players in $3 \cup N_{3_2}$ play strategies in $A \backslash A^p$; players in N_{3_1} play strategies in A^p; a player in $N_{3_3} = \{9\}$ plays a strategy in A^p; all other players play strategies in $A \backslash A^p$.
$t=4$	Players in $3 \cup N_{3_2} \cup N_{3_4}$ play strategies in A^p; players in $N_{3_1} \cup N_{3_3}$ play strategies in $A \backslash A^p$; all other players play strategies in $A \backslash A^p$.
— —	— —
$t=9$	Players in $3 \cup N_{3_2} \cup N_{3_4} \cup N_{3_6} \cup N_{3_8}$ play strategies in $A \backslash A^p$; and players in $N_{3_1} \cup N_{3_3} \cup N_{3_5} \cup N_{3_7} \cup N_{3_9}$ play strategies in A^p.

Thus, for this network, (A, U, N, G, P) converges to an absorbing cycle $W \in \mathcal{C}(\mathbf{A})$ where players in $3 \cup N_{3_2} \cup N_{3_4} \cup N_{3_6} \cup N_{3_8}$ and $N_{3_1} \cup N_{3_3} \cup N_{3_5} \cup N_{3_7} \cup N_{3_9}$ alternate between strategies in A^p and $A \backslash A^p$.[22]

We now show that $\underline{n}_1^* = 1$ mutations[23] to strategies in A^p sufficiently trigger an exit from $D(W)$ to $\mathbf{A}^p$, but $\underline{n}_1^*$ mutations to strategies in $A \backslash A^p$ cannot trigger exit from $D(W)$. Let (A, U, N, G, P) start (at $t=0$) from configuration $\mathbf{y} \in W$ where players in $3 \cup N_{3_2} \cup N_{3_4} \cup N_{3_6} \cup N_{3_8}$ play strategies in $A \backslash A^p$ and players in $N_{3_1} \cup N_{3_3} \cup N_{3_5} \cup N_{3_7} \cup N_{3_9}$ play strategies in A^p. At $t=1$, let a player in $N_{3_1}^9 = \{5\}$ mutate to a strategy in A^p. Then (A, U, N, G, P) evolves from $t=1$ onward as follows:

$t=1$	Players in $3 \cup N_{3_1}^9 \cup N_{3_2} \cup N_{3_4} \cup N_{3_6} \cup N_{3_8}$ play strategies in A^p; players in $\bar{N}_{3_1}^9 \cup N_{3_3} \cup N_{3_5} \cup N_{3_7} \cup N_{3_9}$ play strategies in $A \backslash A^p$, where $\bar{N}_{3_1}^9 = \{1\}$.
$t=2$	Players in $3 \cup N_{3_1} \cup N_{3_2} \cup N_{3_3} \cup N_{3_5} \cup N_{3_7} \cup N_{3_9}$ play strategies in A^p, where players in $3 \cup N_{3_1} \cup N_{3_2} \cup N_{3_3}$ all play strategies in A^p because each has at least $\frac{1}{2} = p$ of their neighbours in $\{3,5\} \cup N_{3_2}$, all of whom play strategies in A^p at $t=1$. Players in $N_{3_4} \cup N_{3_6} \cup N_{3_8}$ play strategies in $A \backslash A^p$.
$t=3$	Players in $B_{3_4} \cup N_{3_6} \cup N_{3_8}$ play strategies in A^p; players in $N_{3_5} \cup N_{3_7} \cup N_{3_9}$ play strategies in $A \backslash A^p$.
— —	— —
$t=7$	Players in B_{3_8} play strategies in A^p; players in N_{3_9} play strategies in $A \backslash A^p$.
$t=8$	All players in $B_{3_9} = N$ play strategies in A^p.

Thus, $\underline{n}_1^*$ mutations to strategies in A^p sufficiently trigger an exit from $D(W)$, which implies that $C(W, \mathbf{A}^p) \leq \underline{n}_1^*$. Now, let (A, U, N, G, P) instead start (at $t=0$) from

---

[22]Note that the composition of this absorbing cycle can equivalently be stated as: players in $3 \cup N_{3_2}^3 \cup N_{3_4} \cup N_{3_6} \cup N_{3_8}$ and $N_{3_1}^3 \cup N_{3_3} \cup N_{3_5} \cup N_{3_7} \cup N_{3_9}$ alternate between strategies in A^p and $A \backslash A^p$, and the rest, consisting of players in $\bar{N}_{3_1}^3 \cup \bar{N}_{3_2}^3$, play strategies in A.

[23]Where $\underline{n}_1^* = 1$ because $N_{3_1}^3 = N_{3_1}^4 = \cdots = N_{3_1}^9 = \{5\}$.

configuration $\mathbf{z} \in W$ *where players in* $3 \cup N_{3_2} \cup N_{3_4} \cup N_{3_6} \cup N_{3_8}$ *play strategies in* A^p *and players in* $N_{3_1} \cup N_{3_3} \cup N_{3_5} \cup N_{3_7} \cup N_{3_9}$ *play strategies in* $A \backslash A^p$.[24] *At* $t = 1$, *let a player in* $N_{3_1}^9 = \{5\}$ *mutate to a strategy in* $A \backslash A^p$. *Then* (A, U, N, G, P) *evolves from* $t = 1$ *onward as follows.*

$t = 1$	*Players in* $\bar{N}_{3_1}^9 \cup N_{3_3} \cup N_{3_5} \cup N_{3_7} \cup N_{3_9}$ *play strategies in* A^p; *players in* $3 \cup N_{3_1}^9 \cup N_{3_2} \cup N_{3_4} \cup N_{3_6} \cup N_{3_8}$ *play strategies in* $A \backslash A^p$.
$t = 2$	*Players in* $3 \cup N_{3_2} \cup N_{3_4} \cup N_{3_6} \cup N_{3_8}$ *play strategies in* A^p. *Player i and players in* $\bar{N}_{3_2}^4$ *all play strategies in* A^p *because each has at least* $\frac{1}{2} = p$ *of their neighbours in* $\bar{N}_{3_1}^9$ *playing strategies in* A^p *at* $t = 1$. *Players in* $N_{3_2}^4 = \{6, 7, 8\}$ *play strategies in* A^p *because for each* $k \in N_{3_2}^4$, *there is a* $j \in N_{3_4} = \{10, 11, 12\}$ *where* $N_{k_1} \cap N_{j_1} \subseteq N_{3_3}^4$. *This implies that, for each* $k \in N_{3_2}^4$, $\alpha_k(N_{3_3}^4) \geq \alpha_k(N_{j_1}) = \alpha_k(B_{j_1}) \geq \min_{l \in N_{j_2}} \alpha_l(B_{j_1}) \geq \underline{\eta}(G) = p = \frac{1}{2}$. *Since all players in* $N_{3_3}^4 \subseteq N_{3_3}$ *play strategies in* A^p *at* $t = 1$, *strategies in* A^p *are best responses to all* $k \in N_{3_2}^4$. *And since* $N_{3_2} = \bar{N}_{3_2}^4 \cup N_{3_2}^4$, *it follows that all players in* N_{3_2} *play strategies in* A^p *at* $t = 2$. *Players in* $N_{3_1} \cup N_{3_3} \cup N_{3_5} \cup N_{3_7} \cup N_{3_9}$ *play strategies in* $A \backslash A^p$.
$t = 3$	*Players in* $N_{3_1} \cup N_{3_3} \cup N_{3_5} \cup N_{3_7} \cup N_{3_9}$ *play strategies in* A^p; *and players in* $3 \cup N_{3_2} \cup N_{3_4} \cup N_{3_6} \cup N_{3_8}$ *play strategies in* $A \backslash A^p$.

Thus, (A, U, N, G, P) *reverts to W after* $\underline{n}_1^*$ *mutations to strategies in* $A \backslash A^p$. *Exiting the basin of attraction of W through mutations to strategies in* $A \backslash A^p$ *requires more than* $\underline{n}_1^*$ *mutations, which implies that* $R(W) = C(W, \mathbf{A}^p) \leq \underline{n}_1^*$. *Following the above discussion, and equation* (9), *it follows that* $C^*(\mathbf{x}, \mathbf{A}^p) = \min_{W \in \mathcal{C}(\mathbf{A})} C(\mathbf{x}, W) \leq \underline{b}_1$.

The second part of the proof of Theorem 1 establishes a lower bound of the radius of $\mathbf{A}^p$. Following the implications of full and partial contagion discussed above, we show in Appendix C that when $\underline{\eta}(G) \geq p$ and $d(G) \geq 7$, $\underline{b}_1$ mutations to strategies in $A \backslash A^p$ can not trigger an exit from the basin of attraction of $\mathbf{A}^p$, and hence, $R(\mathbf{A}^p) \geq \underline{b}_1 + \iota$, where ι is some integer greater or equal to one. The underlying intuition is that if $d(G) \geq 7$, which, according to Lemma 2, also implies that $d_i \geq 4$ for all $i \in N$, then there is no overlap between the first-neighbourhood of any $i \in N$ and the second-neighbourhood of any $j \in N_{i_4} \cup \cdots \cup N_{i_{d_i}}$ (i.e. $B_{i_1} \cap B_{j_2} = \emptyset$). As discussed in Section 3.1 (see also Footnote 19), when $\underline{\eta}(G) \geq p$, (A, U, N, G, P) converges to $\mathbf{A}^p$ from any configuration consisting of at least one $j \in N$ with all players in B_{j_2} playing strategies in A^p. This implies that when (A, U, N, G, P) starts from any state in $\mathbf{A}^p$ and players in $B_{\underline{i}_1}$ mutate to strategies in $A \backslash A^p$, (A, U, N, G, P) will revert to $\mathbf{A}^p$ whenever $\underline{\eta}(G) \geq p$ and $d(G) \geq 7$, so that $R(\mathbf{A}^p) \geq \underline{b}_1 + \iota$.

---

[24]Starting from configuration $\mathbf{z}$ rather than $\mathbf{y}$ represents the best possible starting point of exiting the basin of attraction of W with $\underline{n}_1^*$ mutations to strategies in $A \backslash A^p$.

The above dynamics does not directly extend to networks with $d(G) \leq 6$. Firstly, networks with $d(G) \leq 4$ contain at least one $i \in N$ with $d_i = 2$. For these networks, when all players in B_{i_1} mutate to strategies in $A \backslash A^p$ starting from any $\mathbf{x} \in \mathbf{A}^p$, all players in N_{i_2} can switch to strategies in $A \backslash A^p$ at $t = 2$. Thus, b_{i_1} mutations to strategies in $A \backslash A^p$ can trigger an exit from $D(\mathbf{A}^p)$ to an absorbing state or cycle containing both strategies in A^p and $A \backslash A^p$ or only strategies in $A \backslash A^p$. Secondly, networks with $5 \leq d(G) \leq 6$ contain at least one $i \in N$ with $d_i = 3$. For each $i \in N$ with $d_i = 3$, $N_{j_2} \subseteq B_{i_1}$ for all $j \in N_{i_3}$. This implies that when all players in B_{i_1} mutate to strategies in $A \backslash A^p$ starting from any $\mathbf{x} \in \mathbf{A}^p$, the resulting configuration consists of players in B_{j_1}, for all $j \in N_{i_3}$, playing strategies in A^p and the rest play strategies in $A \backslash A^p$. From such a configuration, (A, U, N, G, P) can converge to an absorbing cycle containing both strategies in A^p and $A \backslash A^p$ (see the proof of the upper bounds of $CR(\mathbf{A}^p)$ and $CR^*(\mathbf{A}^p)$ above). The following counter example helps to further illustrate this point.

Example 3. *Consider an example of the network in Figure 4 with $\underline{\eta}(G) = \frac{2}{5}$ and consisting of at least one player with $d_i = 3$. Let $p \leq \underline{\eta}(G)$, and let (A, U, N, G, P) start from any $\mathbf{x} \in \mathbf{A}^p$. If all players in $B_{6_1} = \{4, 5, 6\}$, where player 6 has $d_6 = 3$, simultaneously mutate to strategies in $A \backslash A^p$ at $t = 1$, then at $t = 2$, all players in $\{2, 3, 6, 7, 8\}$ switch to strategies in $A \backslash A^p$ because each has at least proportion $\frac{2}{3} > \frac{3}{5} = (1 - p)$ of their neighbours playing strategies in $A \backslash A^p$ 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 $A \backslash A^p$. A total of b_{6_1} mutations to strategies in $A \backslash A^p$ are therefore sufficient to trigger an exit from $D(\mathbf{A}^p)$ to an absorbing cycle containing both strategies in A^p and $A \backslash A^p$.*
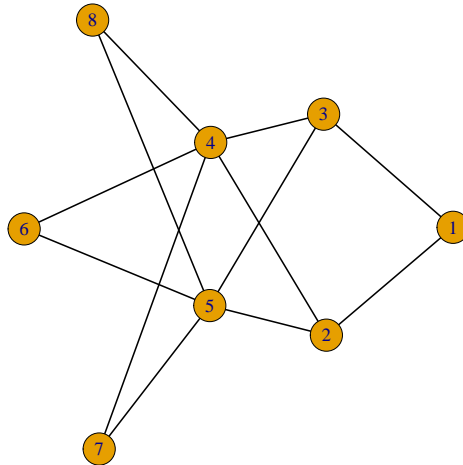

Figure 4: An example of a network with $\underline{\eta}(G) = \frac{2}{5}$ and $d(G) = 3$.

To conclude the proof of Theorem 1, we see from the above discussion that when $p \leq \underline{\eta}(G)$ and $d(G) \geq 7$, $CR(\mathbf{A}^p) \leq \underline{b}_1$, $CR^*(\mathbf{A}^p) \leq \underline{b}_1$ and $R(\mathbf{A}^p) \geq \underline{b}_1 + \iota$. Thus,

the inequalities $R(\mathbf{A}^p) > CR(\mathbf{A}^p)$ and $R(\mathbf{A}^p) > CR^*(\mathbf{A}^p)$ hold, which implies that the absorbing set $\mathbf{A}^p$, and hence, the subset $A^p \subseteq A$, is uniquely stochastically stable.

□

5 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. The results of Theorem 1 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, also hold for directed networks. It follow from the proof of Lemma 2 that, for a directed, unweighted and strongly connected network $G(N, E)$, if $d(G) \geq 7$, then $d_i \geq 4$ for all $i \in N$. Thus, the steps undertaken to derive the upper bounds of the coradius and modified coradius, and the lower bound of the radius of $\mathbf{A}^p$ 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 definition of the contagion threshold. Specifically, for any $j \in N_{i_r}$, let $\phi_j(B_{i_{r-1}})$ be the proportion of j's interactions with players in $B_{i_{r-1}}$. That is, if w_{jk} is the weight of the directed link from j to k, then $\phi_j(B_{i_{r-1}}) = \sum_{k \in B_{i_{r-1}}} w_{jk} / \sum_{l \in N_j} w_{jl}$. The contagion threshold of a weighted network G is then defined as $\underline{\eta}(G) = \min_{i \in N} \min_{r \in [2, d_i]} \min_{j \in N_{i_r}} \phi_j(B_{i_{r-1}})$. To account for the link weights in the total payoffs, the payoff to player i from playing strategy a_j against configuration $\mathbf{x}$ can be expressed as $U_i\left(a_j \mid \mathbf{x}\right) = \sum_{k \in N_i} w_{ik} u(a_j, x^k)$, where x^k is the strategy played by player k in configuration $\mathbf{x}$. These definitions of the contagion threshold and the total payoffs ensure that strategies in A^p will be fully and step-by-step contagious in a weighted and strongly connected network G whenever $\underline{\eta}(G) \geq p$, and that Theorem 1 holds for weighted and strongly connected network.[25]

Unlike directedness and weightedness, requiring a network to be strongly connected is necessary for Theorem 1 to hold. When a network is not strongly connected, there is no guarantee that strategies in A^p can spread contagiously from neighbourhood B_{i_1} of every $i \in N$. Similarly, there is no guarantee that $d(G) \geq 7$ implies that $d_i \geq 4$ for all $i \in N$. The proof, and hence validity of Theorem 1, relies on these two conditions. Note, however, that if an undirected network is not strongly connected, then it consists of strongly connected components which can be analysed as separate networks. More specifically, if $\mathcal{G}$ is the set of all strongly connected components of G and G_k is a typical element of $\mathcal{G}$, then, when $d(G_k) \geq 7$ and $\underline{\eta}(G_k) \geq p$ for all $G_k \in \mathcal{G}$, strategies in A^p are

---

[25]Specifically, these definitions of the contagion threshold and the total payoffs ensure that the derivation of the upper bounds of $CR(\mathbf{A}^p)$ and $CR^*(\mathbf{A}^p)$ and the lower bound of $R(\mathbf{A}^p)$ for a weighted network follows the same steps outlined in Appendix C.

uniquely stochastically stable in each component and in turn the whole network.

The second condition is a restriction on the diameter, $d(G) \geq 7$. This condition is sufficient but not necessary for Theorem 1 to hold. It is easy to construct examples of networks with $d(G) < 7$ where strategies in A^p are uniquely stochastically stable. The inequality, $d(G) \geq 7$, 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 condition does not limit the applicability of the results of Theorem 1. This is because real-world networks (e.g. social media networks, collaboration networks in academia, etc.) are large and sufficiently localized that their diameters are greater than seven.

The last condition relates p to the contagion thresholds $\underline{\eta}(G)$. The inequality, $p \leq \underline{\eta}(G)$, is sufficient but not necessary for strategies in A^p to be uniquely stochastically stable. It is easy to construct an example where strategies in A^p are uniquely stochastically stable in a network with $p > \underline{\eta}(G)$, but such examples cannot easily be generalized.

Our analysis and results are related to Opolot (2020) and Peski (2010) who also identify network measures that determine which and when strategies are uniquely stochastically stable. Opolot (2020) shows that for regular cyclic networks, if *the maximum group cohesion* is less or equal to $(1 - p)$, then there exists a threshold network diameter above which strategies that form the *smallest iterated p-best response set* (i.e., a set of strategies that remain after iterative application of the notion of a p-best response set) are uniquely stochastically stable. And when the network diameter is greater or equal to two, strategies in the minimal p-best response set are uniquely stochastically stable whenever the maximum group cohesion is less or equal to $(1 - p)$. Morris (2000) and Opolot (2020) define the maximum group cohesion as follows.

For some $Z \subset N$ and $i \in Z$ in network G, define $\eta_i(Z, G)$ as the proportion of i's neighbours in Z. That is,

$$\eta_i(Z, G) = \frac{|N_i \cap Z|}{n_i}$$

A group $Z \subset N$ of players is α-cohesive if $\eta(Z, G) = \min_{i \in Z} \eta_i(Z, G) \geq \alpha$. The maximum group cohesion in G is then $\eta(G) = \max_{Z \subset N} \eta(Z, G)$.

Opolot (2020) also demonstrates with an example that although the maximum group cohesion predicts unique stochastically stable outcomes in regular cyclic networks, it fails to do so in some arbitrary networks. In contrast, since $1/\Delta(G) \leq \underline{\eta}(G) \leq \frac{1}{2}$ for all strongly connected networks with $d(G) \geq 7$, where $\Delta(G) = \max_{i \in N} n_i$, is the size of the largest first-order neighbourhood in G, Theorem 1 states that strategies in A^p, for $1/\Delta(G) \leq p \leq \frac{1}{2}$, are uniquely stochastically stable in any undirected, unweighted and strongly connected arbitrary network with $\underline{\eta} \geq p$ and $d(G) \geq 7$.[26]

---

[26]To see why the upper bound of $\underline{\eta}(G)$ is $\frac{1}{2}$, first notice that for any $i \in N$ and each $h \in N_{i_\tau}^r$, for all $\tau = 2, 3 \cdots, d_i - 2$, $r = 2, 3 \cdots, d_i$ and $\tau \leq r + 2$, there exists at least one $j \in N_{i_r}$, with $N_{h_1} \cap N_{j_1} \subseteq N_{i_{\tau+1}}^r$,

The contagion threshold can also be defined using alternative reference groups. For example, the cohesive subgroups, rather than the first-neighbourhoods of players, can be used as reference groups from which strategies in A^p spread. That is, let $\mathcal{Z}(G)$ denote the set of all groups of players in G that are $\frac{1}{2}$-cohesive. For each $Z \in \mathcal{Z}(G)$, let N_{Z_1} be the set of players in $N \backslash Z$ with at least one neighbour in Z; N_{Z_2} is the set of players in $N \backslash \{Z \cup N_{Z_1}\}$ with at least one neighbour in N_{Z_1}; and more generally, N_{Z_r} is the set of players in $N \backslash \{Z \cup N_{Z_1} \cup \cdots \cup N_{Z_{r-1}}\}$ with at least one neighbour in $N_{Z_{r-1}}$. Let $d_Z(G)$ be the diameter of Z in network G (i.e., the value of r at which $N_{Z_r} \neq \emptyset$ but $N_{Z_{r+1}} = \emptyset$). For any $j \in N_{Z_r}$, let $\eta'_j(N_{Z_{r-1}})$ be the proportion of j's neighbours in $N_{Z_{r-1}}$. Then the contagion threshold of G can also be defined as

$$\eta'(G) = \min_{Z \in \mathcal{Z}(G)} \min_{r \in [1, d_Z]} \min_{j \in N_{Z_r}} \eta'_j(N_{Z_{r-1}}) \tag{11}$$

This definition of the contagion threshold ensure that strategies in A^p can spread on G through either full or partial contagion starting from any $Z \in \mathcal{Z}(G)$ whenever $\eta'(G) \geq p$. From the proof of Theorem 1 and the discussion in Opolot (2020, Section 5), we conjecture that when $\eta'(G) \geq p$, there exists a threshold value of the minimum group diameter (i.e., the minimum $d_Z(G)$ over all $Z \in \mathcal{Z}(G)$) above which $CR(\mathbf{A}^p) \leq \min_{Z \in \mathcal{Z}(G)} |Z|$, and $R(\mathbf{A}^p) \geq \min_{Z \in \mathcal{Z}(G)} |Z| + \iota$, where $\iota \geq 1$ is some positive integer (i.e., the threshold diameter at which strategies in A^p are uniquely stochastically stable). The exact proof of this conjecture is beyond the scope of this paper and is left as an avenue for future research.

Peski (2010, Theorem 2 & Corollary 1) shows that for the BRM model, a $\left(\frac{1}{2} - \delta^*\right)$-dominant strategy is uniquely stochastically stable under local interactions, where δ^* is defined below. That is, let strategy $a \in A$ be $\left(\frac{1}{2} - \delta^*\right)$-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

$$\delta_i = \inf \left\{ \delta : \forall_{S \subseteq N \backslash i} \forall_{j^* \in S} \text{ if } \frac{\sum_{j \in S} g_{ij}}{\sum_{j \neq i} g_{ij}} > \frac{1}{2} + \delta, \text{ then } \frac{\sum_{j \in S \backslash j^*} g_{ij}}{\sum_{j \neq i} g_{ij}} \geq \frac{1}{2} - \delta \right\} \tag{12}$$

If we label N_i, the first-order neighbours of player $i \in N$, as $N_i = \{1, 2, \cdots, n_i\}$, then the set S satisfying (12) is such that when n_i is even, $S = \{1, 2, \cdots, \frac{n_i}{2}\}$, and when n_i is odd,

---

for whom $1 - \alpha_h(B_{i_{\tau-1}}) \geq \alpha_h(N^r_{i_{\tau+1}}) \geq \alpha_h(N_{j_1}) = \alpha_h(B_{j_1}) \geq \min_{l \in N_{j_2}} \alpha_l(B_{j_1}) \geq \underline{\eta}(G)$. This implies that if $\alpha_h(B_{i_{\tau-1}}) = \alpha_h(N^r_{i_{\tau-1}}) \geq \frac{1}{2}$ for each $h \in N^r_{i_\tau}$, for all $\tau = 2, 3 \cdots, d_i - 2$ and $r = 2, 3 \cdots, d_i$, which would imply that $\min_{\tau \in [2, d_i]} \min_{l \in N_{i_\tau}} \alpha_l(B_{i_{\tau-1}}) > \frac{1}{2}$, then $\underline{\eta}(G) \leq 1 - \alpha_h(B_{i_{\tau-1}}) \leq \frac{1}{2}$.

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

$$
\delta_i = \begin{cases} 0 & \text{if } n_i \text{ is even} \\ \frac{1}{2n_i} & \text{if } n_i \text{ is odd} \end{cases} \tag{13}
$$

where the equality for n_i odd follows because when $S = \{1, 2, \cdots, \lceil \frac{n_i}{2} \rceil\}$, 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[27] so that $\delta^* = \frac{1}{2\delta_0(G)}$, then the results in Peski (2010, Theorem 2 & Corollary 1) essentially state that a p-dominant strategy is uniquely stochastically stable in network G if $p \leq \frac{1}{2}\left(1 - \frac{1}{\delta_0(G)}\right)$.[28] 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 (2010) offer more accurate predictions for evolutionary dynamics in densely connected networks (i.e. where $\delta_0(G)$ is large) – by more accurate predictions we mean the upper bound of p at which a p-dominant strategy is uniquely stochastically stable is as close to $\frac{1}{2}$ as it can be for a given network. Our results instead offer more accurate predictions in sparsely connected networks. For example, when the smallest odd neighbourhood size is 3, Peski (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 3 above with $\delta_0(G) = 3$, Theorem 1 predicts that a $\frac{1}{2}$-dominant strategy is uniquely stochastically stable. Peski (2010, Theorem 2 & Corollary 1) fails to predict the 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 2 above, where Theorem 1 instead predicts that a $\frac{2}{7}$-dominant strategy is 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

---

[27]That is, if $\mathbb{O}$ is the set of all odd integers, then $\delta_0(G) = \min_{i \in N; n_i \in \mathbb{O}} n_i$

[28]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 $A \backslash a$ for the best response to be within $A \backslash a$ 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.

of a unique p-dominant strategy, the next best option may be to isolate strategies that form the minimal p-best response set.

6   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 2×2 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 the minimal p-best response set are uniquely stochastically stable in networks with the 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.

A   Matrix-based steps for computing the contagion threshold

The contagion threshold of a network can be computed through matrix-based representation of networks. The network $G(N, E)$ can be represented by an adjacency matrix, which, with a slight abuse of notation, is here denoted G. The adjacency matrix is a zero-one matrix with elements $g_{ij} = 1$ if there is a link from i to j, and zero otherwise. Recall that the contagion thresholds is the minimum $\alpha_j(B_{i_{\tau-1}})$ over all $j \in N_{i_\tau}$, $\tau = 2, \cdots, d_i$, and $i \in N$, where $\alpha_j(B_{i_{\tau-1}})$ is the proportion of player j's neighbours that are contained in $B_{i_{\tau-1}}$. For $\tau = 2, 3, \cdots, d(G)$, the values of $\alpha_j(B_{i_{\tau-1}})$, for all $j \in N_{i_\tau}$ and $i \in N$, can be written in matrix form, here denoted by $W(\tau)$. That is, $W(\tau)$ is a matrix with each element $w_{ij}(\tau) = \alpha_j(B_{i_{\tau-1}})$ if $j \in N_{i_\tau}$ and zero otherwise. For this representation, the contagion threshold is given by $\underline{\eta}(G) = \min_{i \in N} \min_{\tau \in [2, d(G)]} \min_{j \neq i} w_{ij}(\tau)$. For any network with adjacency matrix G, the corresponding $W(\tau)$ matrices, for $\tau = 2, \cdots, d_i$, can be computed from G through the following matrix-based algorithmic step. These

algorithmic steps involve matrix multiplication and addition, which makes them *tractable* – matrix multiplication can be executed in polynomial time of order three (Bini et al., 1979).

Let I denote the identity matrix, and $D(G)$, or simply D where no confusion arises, be the diagonal matrix with the diagonal elements being the number of direct neighbours of each player. 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. For any arbitrary matrix M, let $M^{[1]}$ be a matrix derived from M by replacing all elements of M that are greater than one with one, and let $M^{[0]}$ be a matrix derived from M by replacing all negative elements of M with zero. These two matrix operations can be extended to sums and products of matrices. For example, $[H \times M]^{[1]}$ is a matrix derived from the product $H \times M$ by replacing all elements of $H \times M$ that are greater than one with one, and $[H + M]^{[1]}$ is a matrix derived from $H + M$ by replacing all elements of $H + M$ that are greater than one with one. The third operator that we use in the computation of $W(\tau)$ matrices is $\text{RepZero}[M; Z]$, which takes two equal-sized square matrices M and H, and replaces the element m_{ij} of M with zero if $h_{ij} = 0$, otherwise $m_{ij} = m_{ij}$ if $h_{ij} \neq 0$. For $\tau = 2, 3, \cdots d(G)$, each $W(\tau)$ is computed as follows.

To compute $W(2)$, first notice that the entries of $G \times G$ are the number of two-step paths leading up to a given node (i.e. the ijth entry of $G \times G$ is the number of paths that start from i and end at j in two steps). Thus, if all elements of $G \times G$ that are greater than one are replaced by 1, then the ijth element of $[G \times G]^{[1]}$ is one if j is reachable in two steps from i, and zero otherwise. Note, however, that, for each $i \in N$, we aim to compute the values of $\alpha_j(B_{i_1})$ for all $j \in N_{i_2}$, where N_{i_2} is the set of the second-order neighbours of i (i.e. players that are two steps away from i and are not contained in B_{i_1}). Note also that the ijth element of matrix $I + G$ is one if j is in B_{i_1} and zero otherwise. Thus, if we subtract $I + G$ from $[G \times G]^{[1]}$, we end up with $[[G \times G]^{[1]} - (I + G)]^{[0]}$ whose ijth entry is equal to one if j is contained in N_{i_2} and zero otherwise. Now, note that the ijth entry of $G \times G$, which is the number of two-step paths from i to j, is equal to the number of direct neighbours of j that are contained in B_{i_1}. This implies that the ijth entry of $\text{RepZero}\Big[G \times G; [[G \times G]^{[1]} - (I + G)]^{[0]}\Big]$ is the number of neighbours of $j \in N_{i_2}$ that are contained in B_{i_1} and zero otherwise. The $W(2)$ matrix is then equal to $\text{RepZero}\Big[G \times G; [[G \times G]^{[1]} - (I + G)]^{[0]}\Big]$ with each ijth entry normalized by the number of neighbours of j. That is

$$W(2) = \text{RepZero}\Big[G \times G; [[G \times G]^{[1]} - (I + G)]^{[0]}\Big] \times D(G)^{-1}$$

The matrix $W(3)$ is computed through the same steps above. To shorten notation, let $M(2) = [[G \times G]^{[1]} - (I + G)]^{[0]}$. The product $M(2) \times G$ is the number of three-step paths between pairs of players (i.e. the ijth entry is the number of paths that start from i and

32

end at j in three steps); and the ijth element of $[M(2) \times G]^{[1]}$ is one if j is reachable in three steps from i, and zero otherwise. The ijth element of matrix $I + G + M(2)$ is one if j is in B_{i_2}, and zero otherwise. Thus, the ijth entry of $\left[[M(2) \times G]^{[1]} - (I + G + M(2)) \right]^{[0]}$ is equal to one if $j \in N_{i_3}$. The matrix $W(3)$ is then given by

$$W(3) = \text{RepZero}\left[M(2) \times G; \left[[M(2) \times G]^{[1]} - (I + G + M(2)) \right]^{[0]} \right] \times D(G)^{-1}$$

Letting $M(3) = \left[[M(2) \times G]^{[1]} - (I + G + M(2)) \right]^{[0]}$, the product $M(3) \times G$ is the number of four-step paths between pairs of players, so that the ijth entry of $[M(3) \times G]^{[1]}$ is one if j is reachable in four steps from i, and zero otherwise. The ijth element of matrix $I + G + M(2) + M(3)$ is one if j is in B_{i_3}, and zero otherwise. Thus, the ijth entry of $\left[[M(3) \times G]^{[1]} - (I + G + M(2) + M(3)) \right]^{[0]}$ is equal to one if $j \in N_{i_4}$. The matrix $W(4)$ is then given by

$$W(4) = \text{RepZero}\left[M(3) \times G; \left[[M(3) \times G]^{[1]} - (I + G + M(2) + M(3)) \right]^{[0]} \right] \times D(G)^{-1}$$

Let $M(\tau) = \left[[M(\tau-1) \times G]^{[1]} - \sum_{r=0}^{\tau-1} M(r) \right]^{[0]}$, where $M(0) = I$ and $M(1) = G$. Then, following the same steps in the above iterative process, the $W(\tau)$ matrix, for $\tau = 2, 3, \cdots, d(G)$ is given by

$$W(\tau) = \text{RepZero}\left[M(\tau-1) \times G; \left[[M(\tau-1) \times G]^{[1]} - \sum_{r=0}^{\tau-1} M(r) \right]^{[0]} \right] \times D(G)^{-1}$$

B Proof of Lemma 1

To prove Lemma 1, we first show that the following relation holds.

$$\min_{r \in \{\tau, \tau+1\}} \min_{j \in B_{i_r}} \alpha_j(B_{i_{r-1}}) = \min\left\{ \min_{j \in N_{i_{\tau-1}}} \alpha_j(B_{i_{\tau-1}}), \min_{j \in N_{i_\tau}} \alpha_j(B_{i_{\tau-1}}), \min_{j \in N_{i_{\tau+1}}} \alpha_j(B_{i_\tau}) \right\} \quad (14)$$

To prove (14), first notice that

$$\min_{r \in \{\tau, \tau+1\}} \min_{j \in B_{i_r}} \alpha_j(B_{i_{r-1}}) = \min\left\{ \min_{j \in B_{i_\tau}} \alpha_j(B_{i_{\tau-1}}), \min_{j \in B_{i_{\tau+1}}} \alpha_j(B_{i_\tau}) \right\} \quad (15)$$

.

Second, since all direct neighbours of any $h \in B_{i_{\tau-2}}$ are contained within $B_{i_{\tau-1}}$, it follows that $\alpha_h(B_{i_{\tau-1}}) = 1$ for all $h \in B_{i_{\tau-2}}$. And if $\tau + 1 \le d_i$, so that $N_{i_{\tau+1}} \ne \emptyset$, then there exists at least one $k \in N_{i_{\tau-1}} \cup N_{i_\tau}$ with $\alpha_k(B_{i_{\tau-1}}) < 1$, so that

$$\min_{j \in B_{i_\tau}} \alpha_j(B_{i_{\tau-1}}) = \min_{j \in N_{i_{\tau-1}} \cup N_{i_\tau}} \alpha_j(B_{i_{\tau-1}}) = \min\left\{ \min_{j \in N_{i_{\tau-1}}} \alpha_j(B_{i_{\tau-1}}), \min_{j \in N_{i_\tau}} \alpha_j(B_{i_{\tau-1}}) \right\} \quad (16)$$

Similarly, $\min_{j \in B_{i_{\tau+1}}} \alpha_j(B_{i_\tau}) = \min \left\{ \min_{j \in N_{i_\tau}} \alpha_j(B_{i_\tau}), \min_{j \in N_{i_{\tau+1}}} \alpha_j(B_{i_\tau}) \right\}$.

Third, note that $\alpha_k(B_{i_\tau}) \geq \alpha_k(B_{i_{\tau-1}})$ for all $k \in N_{i_\tau}$. This is because some of the direct neighbours of $k \in N_{i_\tau}$ are contained within N_{i_τ}, and that $N_{i_\tau} \subseteq B_{i_\tau}$ but $N_{i_\tau} \not\subseteq B_{i_{\tau-1}}$. This inequality implies that

$$\min \left\{ \min_{j \in B_{i_{\tau-1}}} \alpha_j(B_{i_{\tau-1}}), \min_{j \in B_{i_{\tau+1}}} \alpha_j(B_{i_\tau}) \right\} = \min \left\{ \min_{j \in N_{i_{\tau-1}}} \alpha_j(B_{i_{\tau-1}}), \min_{j \in N_{i_\tau}} \alpha_j(B_{i_{\tau-1}}), \min_{j \in N_{i_{\tau+1}}} \alpha_j(B_{i_\tau}) \right\}$$

This proves equation (14).

Now, considering the minimum $\min_{j \in B_{i_r}} \alpha_j(B_{i_{r-1}})$ over all $r = 3, \cdots, d_i$, it follows from (14) that

$$\min_{r \in [3,d_i]} \min_{j \in B_{i_r}} \alpha_j(B_{i_{r-1}}) = \min \left\{ \min_{j \in N_{i_2}} \alpha_j(B_{i_2}), \min_{r \in [3,d_i]} \min_{j \in N_{i_r}} \alpha_j(B_{i_{r-1}}) \right\}$$

$$\geq \min_{r \in [2,d_i]} \min_{j \in N_{i_r}} \alpha_j(B_{i_{r-1}}) \tag{17}$$

where the inequality on the right hand side of (17) follows because $\alpha_k(B_{i_2}) \geq \alpha_k(B_{i_1})$ for all $k \in N_{i_2}$. Thus, when $\min_{r \in [2,d_i]} \min_{j \in N_{i_r}} \alpha_j(B_{i_{r-1}}) \geq p$, $\min_{r \in [3,d_i]} \min_{j \in B_{i_r}} \alpha_j(B_{i_{r-1}}) \geq p$, and since $S_\tau \subseteq B_{i_\tau}$ for all $\tau = 2, \cdots, d_i$, it follows that for any sequence $S_3 \subseteq B_{i_3}, \cdots, S_{d_i} \subseteq B_{i_{d_i}}$, $\alpha_h(S_{\tau-1}) \geq p$ for each $h \in S_\tau$ and all $\tau = 3, \cdots, d_i$.

It then remains to show that when $\min_{r \in [2,d_i]} \min_{j \in N_{i_r}} \alpha_j(B_{i_{r-1}}) \geq p$, there exists some $S_2 \subseteq B_{i_2}$ with $\alpha_h(S_1) = \alpha_h(B_1) \geq p$ for each $h \in S_2$. Indeed, from (16), $\min_{j \in B_{i_2}} \alpha_j(B_{i_1}) = \min \left\{ \min_{j \in N_{i_1}} \alpha_j(B_{i_1}), \min_{j \in N_{i_2}} \alpha_j(B_{i_1}) \right\}$. Since $\min_{r \in [2,d_i]} \min_{j \in N_{i_r}} \alpha_j(B_{i_{r-1}}) \geq p$ implies that $\min_{j \in N_{i_2}} \alpha_j(B_{i_1}) \geq p$, it follows that there exists some subgroup of players S_2, where $N_2 \subseteq S_2 \subseteq B_{i_2}$, with $\alpha_h(B_1) \geq p$ for each $h \in S_2$.

Thus, when $\min_{r \in [2,d_i]} \min_{j \in N_{i_r}} \alpha_j(B_{i_{r-1}}) \geq p$, for any $i \in N$, there exist sequences of players, $S_1, S_2, S_3, \cdots, S_{\bar{t}}$, with $S_1 = B_{i_1}$, $S_2 \subseteq B_{i_2}, S_3 \subseteq B_{i_3}, \cdots, S_{\bar{t}} \subseteq B_{i_{\bar{t}}}$, such that $\alpha_h(S_{\tau-1}) \geq p$ for all $h \in S_\tau$ and all $\tau = 2, \cdots, \bar{t}$, where $\bar{t} \leq d_i$.

C   Proof of Theorem 1

Before deriving the upper bounds of the cordadius and modified coradius and the lower bound of the radius of $\mathbf{A}^p$, we first show in the following lemma that, for an undirected, unweighted and strongly connected network G, $d(G) \geq 7$ implies that $d_i(G) \geq 4$ for all $i \in N$.

Recall that a directed path from i to j is a finite sequence $(i_1, i_2, \cdots, i_r)$, with $i = i_1$ and $i_r = j$, that is connected through distinct directed links; d_{ij} is the length of the maximum shortest path from i to j; $d_i = \max_{j \neq i} d_{ij}$ is the length of the shortest path from i to any other player; and $d(G) = \max_{i \in N} d_i$ is the diameter of G.

Lemma 2. *For an undirected, unweighted and strongly connected network G, if $d(G) \geq 7$, then $d_i \geq 4$ for all $i \in N$.*

Proof. We first show that, for an undirected, unweighted and strongly connected network G, if $d(G)$ is even, then $d_i \geq \frac{d(G)}{2}$ for all $i \in N$, and if $d(G)$ is odd, then $d_i \geq \frac{d(G)+1}{2}$ for all $i \in N$.

Consider any directed path that forms a diameter of G (there can be many such paths), and label the elements along this path as $i = i_1, i_2, \cdots, i_{d(G)+1} = j$. Since $i, j \in N$ are the end points of this path, and that G is strongly connected and undirected, it follows that $j \in N_{i_{d_i}}$, $i \in N_{j_{d_j}}$ and $d_{ij} = d_i = d_j = d(G)$. If $d(G)$ is even, then the distance (i.e. the number of steps) from $i_{\frac{d(G)}{2}+1}$ (the player at halfway point along $(i_1, i_2, \cdots, i_{d(G)+1})$) to players i and j is exactly $\frac{d(G)}{2}$, That is, $d_{i_{\frac{d(G)}{2}+1}i} = d_{i_{\frac{d(G)}{2}+1}j} = \frac{d(G)}{2}$. If $d(G)$ is odd, then $i_{\frac{d(G)+1}{2}}$ and $i_{\frac{d(G)+3}{2}}$ are the players that halfway point along $i_1, i_2, \cdots, i_{d(G)+1}$. And the distances from $i_{\frac{d(G)+1}{2}}$ to j and $i_{\frac{d(G)+3}{2}}$ to i is exactly $\frac{d(G)+1}{2}$. That is, $d_{i_{\frac{d(G)+1}{2}}j} = d_{i_{\frac{d(G)+3}{2}}i} = \frac{d(G)+1}{2}$.

This implies that for every player, h, along $(i_1, i_2, \cdots, i_{d(G)+1})$, $d_h \geq \frac{d(G)}{2}$ 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 the shortest path connecting every other player, $k \neq h$, that does not lie along $(i_1, i_2, \cdots, i_{d(G)+1})$, to some h along $(i_1, i_2, \cdots, i_{d(G)+1})$. 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. Thus, when $d(G)$ is even, $d_i \geq \frac{d(G)}{2}$ for all $i \in N$, and when $d(G)$ is odd, $d_i \geq \frac{d(G)+1}{2}$ for all $i \in N$, and hence, when $d(G) \geq 7$, $d_i \geq 4$ for all $i \in N$. $\qquad\square$

The upper bounds of the coradius and modified coradius of $\mathbf{A}^p$

We aim to show that, for evolutionary processes (A, U, N, G, P) and $(A, U, N, G, P_\varepsilon)$ on an undirected, unweighted and strongly connected G, if $p \leq \underline{\eta}(G)$ and $d(G) \geq 7$, then $CR(\mathbf{A}^p) \leq \underline{b}_1$ and $CR^*(\mathbf{A}^p) \leq \underline{b}_1$, where $\underline{b}_1 = \min_{i \in N} b_{i_1}$. The proof follows in four steps. First, we show that when (A, U, N, G, P) starts from any $\mathbf{x} \notin D(\mathbf{A}^p)$, $\underline{b}_1$ mutations to strategies in A^p trigger either full contagion of these strategies or an iterative process of best response where (A, U, N, G, P) ultimately converges to an absorbing cycle of states containing both strategies in A^p and $A \backslash A^p$.

Second, for a scenario where $\underline{b}_1$ mutations to strategies in A^p fail to trigger full contagion, (A, U, N, G, P) converges to an absorbing cycle $W \in \mathcal{C}(\mathbf{A})$ where players in $N_{i_1}^3 \cup N_{i_3} \cup N_{i_5} \cup \cdots \cup N_{i_{d_i}}$ and $i \cup N_{i_2}^3 \cup N_{i_4} \cup N_{i_6} \cup \cdots \cup N_{i_{d_i-1}}$ alternate between strategies in A^p and some strategies in A, and the rest play some strategies in A. Third, at most $\underline{n}_1^*$ mutations to strategies in A^p sufficiently trigger an exit from the basin of

35

attraction of $W \in \mathcal{C}(\mathbf{A})$, so that $C(W, \mathbf{A}^p) \leq \underline{n}_1^*$. This implies that if the number of mutations to strategies in $A \backslash A^p$ that trigger an exit from $D(W)$ is greater than $\underline{n}_1^*$, then $R(W) = C(W, \mathbf{A}^p)$.

Fourth, we show that in a scenario where the number of mutations to strategies in $A \backslash A^p$ that trigger an exit from the basin of attraction of W is less or equal to $\underline{n}_1^*$, there exists another absorbing cycle $W' \in \mathcal{C}(\mathbf{A})$ and $W' \neq W$, where $R(W) = C(W, W') \leq \underline{n}_1^*$. For any such $W' \in \mathcal{C}(\mathbf{A})$, $\underline{n}_1^4$ mutations to strategies in A^p trigger an exit from $D(W')$ to $\mathbf{A}^p$, but $\underline{n}_1^4$ mutations to strategies in $A \backslash A^p$ cannot trigger exit from $D(W')$, which implies that $R(W') = C(W', \mathbf{A}^p) \leq \underline{n}_1^4$. Following the discussion in Section 4, together with equations (9) and (10), we then demonstrate that these four results imply that $CR(\mathbf{A}^p) \leq \underline{b}_1$ and $CR^*(\mathbf{A}^p) \leq \underline{b}_1$.

Let (A, U, N, G, P) start from any $\mathbf{x} \notin D(\mathbf{A}^p)$, and at $t = 1$, let all players in B_{i_1} of any $i \in N$ mutate to strategies in A^p. Then from $t = 2$ onward, (A, U, N, G, P) evolves as follows, where we write $i \to A'$ and $N' \to A'$ to mean i and respectively each player in N' plays a strategy in A'.

$t = 1$	$B_{i_1} \to A^p$;
	$N_{i_2} \cup N_{i_3} \cup \cdots \cup N_{i_{d_i}} \to A$.
$t = 2$	$i \cup N_{i_2} \to A^p$, where i plays a strategy in A^p because all $j \in N_{i_1}$ play strategies in A^p at $t = 1$. Players in N_{i_2} play strategies in A^p because each $h \in N_{i_2}$ has $\alpha_h(B_{i_1}) \geq \underline{\eta}(G) \geq p$, and that all $j \in N_{i_1}$ play strategies in A^p at $t = 1$. Note that no $h \in N_{i_2}^3$ will play a strategy in $A \backslash A^p$ even if all players in N_{i_2} and N_{i_3} play strategies in $A \backslash A^p$ at $t = 1$. This is because for each $h \in N_{i_2}^3$, $\alpha_h(B_{i_1}) \geq \underline{\eta}(G) \geq p$ implies that $(1 - p) \geq (1 - \underline{\eta}(G)) \geq (1 - \alpha_h(B_{i_1})) = \alpha_h(N_{i_2} \cup N_{i_3})$. Since $(1 - p) \geq \alpha_h(N_{i_2} \cup N_{i_3})$, strategies in $A \backslash A^p$ are not best responses to any such h because they are best responses only when played by more than proportion $1 - p$ of neighbours.
	$N_{i_1} \cup N_{i_3} \cup N_{i_4} \cup \cdots \cup N_{i_{d_i}} \to A$. Note, however, that some players in N_{i_1} will play strategies in A^p; for example, all $j \in \bar{N}_{i_1}^2$ (i.e. all $j \in N_{i_1}$ with $N_{j_1} \cap N_{i_2} = \emptyset$) play strategies in A^p because each has all their direct neighbours within B_{i_1}, all of whom play strategies in A^p at $t = 1$. But since $A^p \subset A$, we simply write $N_{i_1} \to A$. Moreover, for the rest of the remaining analysis, it is sufficient to only keep track the strategies played by the players in $i, N_{i_1}^3, N_{i_1}^4, N_{i_2}^3, N_{i_2}^4, N_{i_3}, N_{i_4}, \cdots, N_{i_{d_i}}$.

$t = 3$	$N_{i_1}^3 \cup N_{i_3} \to A^p$. All $h \in N_{i_1}^3$ play strategies in A^p because, firstly, each has at least one second-order neighbours in N_{i_3} (i.e. $N_{h_2} \cap N_{i_3} \neq \emptyset$), which implies that, for each $h \in N_{i_1}^3$, there exists some $j \in N_{i_3}$ with $N_{h_1} \cap N_{j_1} = N_{h_1} \cap B_{j_1} \neq \emptyset$ and $N_{h_1} \cap N_{i_2} \bigcap N_{j_1} \cap N_{i_2} \neq \emptyset$. Since for each $h \in N_{i_1}^3$ and $j \in N_{i_3}$, $N_{h_1} \cap N_{j_1} \subseteq N_{i_2}^3$ (i.e. all direct neighbours of $h \in N_{i_1}^3$ that are also direct neighbours of $j \in N_{i_3}$ are contained within $N_{i_2}^3$), it follows that $\alpha_h(N_{i_2}^3) \geq \alpha_h(N_{j_1}) = \alpha_h(B_{j_1}) \geq \min_{k \in N_{j_2}} \alpha_k(B_{j_1}) \geq \underline{\eta}(G) \geq p$. The first inequality, $\alpha_h(N_{i_2}^3) \geq \alpha_h(N_{j_1})$, holds because, if there is more than one $j \in N_{i_3}$, then for each $h \in N_{j_1}^3$, it is not necessary for all $N_{h_1} \subset N_{i_2}^3$ (i.e. direct neighbours of h that are contained in N_{i_2}) to also be contained in N_{j_1}. Secondly, since all players in $N_{i_2}^3 \subseteq N_{i_2}$ play strategies in A^p at $t = 2$, and by the above inequality, $\alpha_h(N_{i_2}^3) \geq \underline{\eta}(G) \geq p$ for each $h \in N_{j_1}^3$, strategies in A^p are best responses to all $h \in N_{j_1}^3$. All players in N_{i_3} play strategies in A^p because each $l \in N_{i_3}$ has $\alpha_l(N_{i_2}) = \alpha_l(B_{i_2}) \geq \underline{\eta}(G) \geq p$ and all players in N_{i_2} play strategies in A^p at $t = 2$. $i \cup \bar{N}_{i_1}^3 \cup N_{i_2} \cup N_{i_4} \cup N_{i_5} \cup \cdots \cup N_{i_{d_i}} \to A$.
$t = 4$	$i \cup N_{i_2}^3 \cup N_{i_4} \to A^p$. Player i plays a strategy in A^p because for each $j \in N_{i_2}^3$, $N_{i_1} \cap N_{j_1} \subseteq N_{i_1}^3$ (i.e. all direct neighbours of $j \in N_{i_2}^3$ that are also direct neighbours of i are contained within $N_{i_1}^3$). This implies that, for any $j \in N_{i_2}^3$, $\alpha_i(N_{i_1}^3) \geq \alpha_i(N_{j_1}) = \alpha_i(B_{j_1}) \geq \min_{l \in N_{j_2}} \alpha_l(B_{j_1}) \geq \underline{\eta}(G) \geq p$. And since all players in $N_{i_1}^3$ play strategies in A^p at $t = 3$, it follows that a strategy in A^p is a best response to player i. Players in $N_{i_2}^3$ play strategies in A^p for the similar reason. That is, for each $k \in N_{i_2}^3$, $N_{k_1} \cap N_{i_1} \subseteq N_{i_1}^3$, which implies that $\alpha_k(N_{i_1}^3) = \alpha_k(N_{i_1}) = \alpha_k(B_{i_1}) \geq \underline{\eta}(G) \geq p$, and hence, strategies in A^p are best responses to all $k \in N_{i_2}^3$. Players in N_{i_4} play strategies in A^p because each $l \in N_{i_4}$ has $\alpha_l(N_{i_3}) = \alpha_l(B_{i_3}) \geq \underline{\eta}(G) \geq p$, and that all players in N_{i_3} play strategies in A^p at $t = 3$. $N_{i_1} \cup \bar{N}_{i_2}^3 \cup N_{i_3} \cup N_{i_5} \cup N_{i_6} \cup \cdots \cup N_{i_{d_i}} \to A$.
$t = 5$	$N_{i_1}^3 \cup N_{i_3} \cup N_{i_5} \to A^p$. Players in $N_{i_1}^3 \cup N_{i_3}$ play strategies in A^p for the same reasons outlined in $t = 3$ above. Players in N_{i_5} play strategies in A^p because each $l \in N_{i_5}$ has $\alpha_l(N_{i_4}) = \alpha_l(B_{i_4}) \geq \underline{\eta}(G) \geq p$, and that all players in N_{i_4} play strategies in A^p at $t = 4$. $i \cup \bar{N}_{i_1}^3 \cup N_{i_2} \cup N_{i_4} \cup N_{i_6} \cup N_{i_7} \cup \cdots \cup N_{i_{d_i}} \to A$.
$t = 6$	Following the same steps above, $i \cup N_{i_2}^3 \cup N_{i_4} \cup N_{i_6} \to A^p$. $N_{i_1} \cup \bar{N}_{i_2}^3 \cup N_{i_3} \cup N_{i_5} \cup N_{i_7} \cup \cdots \cup N_{i_{d_i}} \to A$.
$t = 7$	$N_{i_1}^3 \cup N_{i_3} \cup N_{i_5} \cup N_{i_7} \to A^p$. $i \cup \bar{N}_{i_1}^3 \cup N_{i_2} \cup N_{i_4} \cup N_{i_6} \cup N_{i_7} \cup \cdots \cup N_{i_{d_i}} \to A$.
$-\!-$	$-\,-$
$t = d_i$	If d_i is an odd number, then: $N_{i_1}^3 \cup N_{i_3} \cup N_{i_5} \cup N_{i_7} \cup \cdots \cup N_{i_{d_i}} \to A^p$ $i \cup \bar{N}_{i_1}^3 \cup N_{i_2} \cup N_{i_4} \cup N_{i_6} \cup N_{i_8} \cup \cdots \cup N_{i_{d_i-1}} \to A$. If d_i is an even number, then: $i \cup N_{i_2}^3 \cup N_{i_4} \cup N_{i_6} \cup N_{i_8} \cup \cdots \cup N_{i_{d_i}} \to A^p$. $N_{i_1} \cup \bar{N}_{i_2}^3 \cup N_{i_3} \cup N_{i_5} \cup N_{i_7} \cup N_{i_9} \cup \cdots \cup N_{i_{d_i-1}} \to A$.

Thus, from $t = d_i$ iterations onward, if $d_i \geq 3$, 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 some strategies in $A \backslash A^p$. Specifically, if players in $i \cup \bar{N}_{i_1}^3 \cup N_{i_2} \cup N_{i_4} \cup N_{i_6} \cup \cdots \cup N_{i_{d_i}-1}$, for d_i odd, and players in $N_{i_1} \cup \bar{N}_{i_2}^3 \cup N_{i_3} \cup N_{i_5} \cup \cdots \cup N_{i_{d_i}-1}$, for d_i even, all play strategies in $A^p \subset A$ after $t = d_i$ iterations, then (A, U, N, G, P) converges to an absorbing set containing only strategies in A^p. For this scenario, b_{i_1} mutations, for any $i \in N$, to strategies in A^p sufficiently trigger evolution from any $\mathbf{x} \notin D(\mathbf{A}^p)$ to $\mathbf{A}^p$ so that $CR(\mathbf{A}^p) = \max_{\mathbf{x} \notin D(\mathbf{A}^p)} C(\mathbf{x}, \mathbf{A}^p) \leq \underline{b}_1$.

We use the following definitions and notation in the following steps where we examine the evolution of (A, U, N, GP) out of absorbing cycles containing both strategies in A^p and some strategies in $A \backslash A^p$.

$\bar{N}_{i_1}^r[N_{i_1}^r]$	For $r = 3, 4, \cdots, d_i$, $\bar{N}_{i_1}^r[N_{i_1}^r]$ is the set of players in $\bar{N}_{i_1}^r$ (i.e. the set of players in N_{i_1} that are not contained in $N_{i_1}^r$) that are one step away from at least one player in $N_{i_1}^r$. That is, for each $j \in \bar{N}_{i_1}^r[N_{i_1}^r]$, there exists at least one $h \in N_{i_1}^r$ where $j \in N_{h_1}$. For example, in the network of Figure 5, $\bar{N}_{1_1}^3[N_{1_1}^3] = \{7\}$
$\bar{N}_{i_2}^r[N_{i_1}^r]$ $\bar{N}_{i_2}^r[N_{i_2}^r]$	For $r = 3, 4, \cdots, d_i$, $\bar{N}_{i_2}^r[N_{i_2}^r]$ is the set of players in $\bar{N}_{i_2}^r$ that are one step away from at least one player in $N_{i_2}^r$. That is, for each $j \in \bar{N}_{i_2}^r[N_{i_2}^r]$, there exists at least one $h \in N_{i_2}^r$ where $j \in N_{h_1}$. For example, in the network of Figure 5, $\bar{N}_{1_2}^3[N_{1_2}^3] = \{3\}$. Similarly, $\bar{N}_{i_2}^r[N_{i_1}^r]$ is the set of players in $\bar{N}_{i_2}^r$ that are one step away from at least one player in $N_{i_1}^r$. That is, for each $j \in \bar{N}_{i_2}^r[N_{i_1}^r]$, there exists at least one $h \in N_{i_1}^r$ where $j \in N_{h_1}$. For example, in the network of Figure 5, $\bar{N}_{1_2}^3[N_{1_1}^3] = \{3, 14\}$.
$\bar{N}_{i_1}^r[\bar{N}_{i_1}^r[N_{i_1}^r]]$	For $r = 3, 4, \cdots, d_i$, $\bar{N}_{i_1}^r[\bar{N}_{i_1}^r[N_{i_1}^r]]$ is the set of players in $\bar{N}_{i_1}^r$ that are one step away from at least one player in $i \cup \bar{N}_{i_1}^r[N_{i_1}^r]$. That is, for each $j \in \bar{N}_{i_1}^r[\bar{N}_{i_1}^r[N_{i_1}^r]]$, where $j \in \bar{N}_{i_1}^r$ but $j \notin i \cup \bar{N}_{i_1}^r[N_{i_1}^r]$, there exists at least one $h \in i \cup \bar{N}_{i_1}^r[N_{i_1}^r]$ where $j \in N_{h_1}$. For example, in the network of Figure 5, $\bar{N}_{1_1}^3[\bar{N}_{1_1}^3[N_{1_1}^3]] = \{8, 10\}$.

Now, let $W \in \mathcal{C}(\mathbf{A}^p)$ be an absorbing cycle where players in $N_{i_1}^3 \cup N_{i_3} \cup N_{i_5} \cup \cdots \cup N_{i_{d_i}}$ and $i \cup N_{i_2}^3 \cup N_{i_4} \cup N_{i_6} \cup \cdots \cup N_{i_{d_i}-1}$ alternate between strategies in A^p and some strategies in A, and the rest play some strategies in A. More specifically, consider any two configurations $\mathbf{y}, \mathbf{z} \in W \in \mathcal{C}(\mathbf{A}^p)$ with $P(\mathbf{y}, \mathbf{z}) = 1$, where $\mathbf{y}$ consist of all players in $N_{i_1}^3 \cup N_{i_3} \cup N_{i_5} \cup \cdots \cup N_{i_{d_i}}$ playing strategies in A^p, and all players in $i \cup \bar{N}_{i_1}^3 \cup N_{i_2} \cup N_{i_4} \cup N_{i_6} \cup \cdots \cup N_{i_{d_i}-1}$ play strategies in A; and configuration $\mathbf{z}$ consists of all players in $i \cup N_{i_2}^3 \cup N_{i_4} \cup N_{i_6} \cup \cdots \cup N_{i_{d_i}-1}$ playing strategies in A^p, while all players in $N_{i_1} \cup \bar{N}_{i_2}^3 \cup N_{i_3} \cup N_{i_5} \cup \cdots \cup N_{i_{d_i}}$ play strategies in A. Let (A, U, N, G, P) start from $\mathbf{y}$, and let all players in $N_{i_1}^r$, for $4 \leq r \leq d_i$, mutate to strategies in A^p at $t = 1$. Then for $d_{i_1} \geq 4$, (A, U, N, G, P) evolves from $t = 1$ onward as follows:
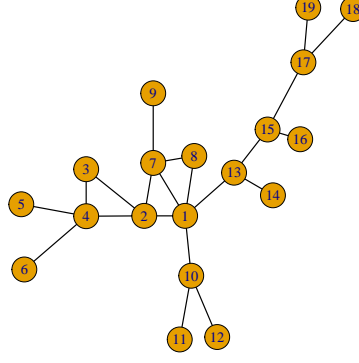
Figure 5: For player 1, the above defined subsets have the following composition: $N_{1_1}^3 = \{2, 13\}$; $\bar{N}_{1_1}^3[N_{1_1}^3] = \{7\}$; $\bar{N}_{1_2}^3[N_{1_1}^3] = \{3, 14\}$; and $\bar{N}_{1_2}^4[N_{1_1}^4] = \{14\}$.

$t = 1$	$i \cup N_{i_1}^r \cup N_{i_2}^3 \cup N_{i_4} \cup N_{i_6} \cup \cdots \cup N_{i_{d_i-1}} \to A^p$;
	$\bar{N}_{i_1}^r \cup \bar{N}_{i_2}^3 \cup N_{i_3} \cup N_{i_5} \cup \cdots \cup N_{i_{d_i}} \to A$.
$t = 2$	$i \cup N_{i_1}^3 \cup N_{i_2}^r \cup \bar{N}_{i_2}^3[N_{i_2}^3] \cup N_{i_3} \cup N_{i_5} \cup \cdots \cup N_{i_{d_i}} \to A^p$. Player i plays a strategy in A^p because for each $k \in N_{i_2}^r$, $N_{k_1} \cap N_{i_1} \subseteq N_{i_1}^r$, which implies that, for any $k \in N_{i_2}^r$, $\alpha_i(N_{i_1}^r) \geq \alpha_i(N_{k_1}) = \alpha_i(B_{k_1}) \geq \min_{l \in N_{k_2}} \alpha_l(B_{k_1}) \geq \underline{\eta}(G) \geq p$. And since all players in $N_{i_1}^r$ play strategies in A^p at $t = 1$, a strategy in A^p must be a best response to player i.

Players in $N_{i_1}^3$ play strategies in A^p because for each $k \in N_{i_1}^3$, there exists at least one $j \in N_{i_3}$ with $N_{k_1} \cap N_{j_1} \subseteq N_{i_2}^3$, which implies that $\alpha_k(N_{i_2}^3) \geq \alpha_k(N_{j_1}) = \alpha_k(B_{j_1}) \geq \min_{l \in N_{j_2}} \alpha_l(B_{j_1}) \geq \underline{\eta}(G) \geq p$. And since all players in $N_{i_2}^3$ play strategies in A^p at $t = 1$, strategies in A^p are best responses to all $k \in N_{i_1}^3$.

Players in $N_{i_2}^r$ play strategies in A^p because for each $k \in N_{i_2}^r$, $N_{k_1} \cap N_{i_1} \subseteq N_{i_1}^r$, which implies that $\alpha_k(N_{i_1}^r) = \alpha_k(N_{i_1}) = \alpha_k(B_{i_1}) \geq \underline{\eta}(G) \geq p$. And since all players in $N_{i_1}^r$ play strategies in A^p at $t = 1$, strategies in A^p are best responses to all $k \in N_{i_2}^r$.

To see why players in $\bar{N}_{i_2}^3[N_{i_2}^3]$ play strategies in A^p, first observe that no $k \in \bar{N}_{i_2}^3[N_{i_2}^3]$ is a direct neighbour of any $j \in N_{i_3}$. That is, if $k \in \bar{N}_{i_2}^3[N_{i_2}^3]$, then $k \notin N_{j_1}$ for all $j \in N_{i_3}$. This observation implies that, for each $k \in \bar{N}_{i_2}^3[N_{i_2}^3]$, there exists at least one $j \in N_{i_3}$ where $N_{k_1} \cap N_{j_1} \subseteq N_{i_2}^3$, so that $\alpha_k(N_{i_2}^3) \geq \alpha_k(N_{j_1}) = \alpha_k(B_{j_1}) \geq \min_{l \in N_{j_2}} \alpha_l(B_{j_1}) \geq \underline{\eta}(G) \geq p$, and hence, strategies in A^p are best responses.

Players in $N_{i_3} \cup N_{i_5} \cup \cdots \cup N_{i_{d_i}}$ play strategies in A^p because each $h \in N_{i_3}$ and $l \in N_{i_r}$, for $r = 5, 7, \cdots, d_i$, has $\alpha_h(N_{i_2}^3) = \alpha_h(N_{i_2}) = \alpha_h(B_{i_2}) \geq \underline{\eta}(G) \geq p$ and $\alpha_l(N_{i_{r-1}}) = \alpha_l(B_{i_{r-1}}) \geq \underline{\eta}(G) \geq p$ respectively. Since all players in $N_{i_2}^3$ and $N_{i_{r-1}}$, for all $r = 5, 7, \cdots, d_i$, play strategies in A^p at $t = 1$, strategies in A^p are best responses to all players in $N_{i_3} \cup N_{i_5} \cup \cdots \cup N_{i_{d_i}}$.

Finally $\bar{N}_{i_1}^3 \cup \bar{N}_{i_2}^r \backslash \bar{N}_{i_2}^3[N_{i_2}^3] \cup N_{i_4} \cup N_{i_6} \cup \cdots \cup N_{i_{d_i-1}} \to A$.

$t = 3$	$i \cup N^r_{i_1} \cup \bar{N}^3_{i_1}[N^3_{i_1}] \cup N^3_{i_2} \cup \bar{N}^3_{i_2}[N^3_{i_1}]\backslash \left\{ \bar{N}^3_{i_2}[N^3_{i_1}] \cap \bar{N}^3_{i_2}[N^3_{i_2}] \right\} \cup \bar{N}^3_{i_2}[N^r_{i_2}] \cup N^r_{i_3} \cup N_{i_4} \cup$ $N_{i_6} \cup \cdots \cup N_{i_{d_i-1}} \to A^p$. Players in $i \cup N^r_{i_1} \cup N^3_{i_2} \cup N_{i_4} \cup N_{i_6} \cup \cdots \cup N_{i_{d_i-1}}$ all play strategies in A^p for the same reasons outlined in $t = 2$ above.

To see why players in $\bar{N}^3_{i_1}[N^3_{i_1}] \cup \bar{N}^3_{i_2}[N^3_{i_1}]\backslash \left\{ \bar{N}^3_{i_2}[N^3_{i_1}] \cap \bar{N}^3_{i_2}[N^3_{i_2}] \right\}$ play strategies in A^p, first observe that for each $k \in \bar{N}^3_{i_1}[N^3_{i_1}] \cup \bar{N}^3_{i_2}[N^3_{i_1}]\backslash \left\{ \bar{N}^3_{i_2}[N^3_{i_1}] \cap \bar{N}^3_{i_2}[N^3_{i_2}] \right\}$, where $k \notin \bar{N}^3_{i_2}[N^3_{i_2}]$, and any $j \in N^3_{i_2}$, we have $N_{k_1} \cap N_{j_1} \subseteq N^3_{i_1} \cup \bar{N}^3_{i_2}[N^3_{i_2}]$ (i.e., all direct neighbours of any $k \in \bar{N}^3_{i_1}[N^3_{i_1}] \cup \bar{N}^3_{i_2}[N^3_{i_1}]$, that are also direct neighbours of some $j \in N^3_{i_2}$, are contained within $N^3_{i_1} \cup \bar{N}^3_{i_2}[N^3_{i_2}]$). This implies that for each $k \in \bar{N}^3_{i_1}[N^3_{i_1}] \cup \bar{N}^3_{i_2}[N^3_{i_1}]\backslash \left\{ \bar{N}^3_{i_2}[N^3_{i_1}] \cap \bar{N}^3_{i_2}[N^3_{i_2}] \right\}$, there exists at least one $j \in N^3_{i_2}$ such that $\alpha_k \left(N^3_{i_1} \cup \bar{N}^3_{i_2}[N^3_{i_2}] \right) \geq \alpha_k(N_{j_1}) = \alpha_k(B_{j_1}) \geq \min_{l \in N_{j_2}} \alpha_l(B_{j_1}) \geq \underline{\eta}(G) \geq p$. And since all players in $N^3_{i_1} \cup \bar{N}^3_{i_2}[N^3_{i_2}]$ play strategies in A^p at $t = 2$, strategies in A^p are best responses to all $k \in \bar{N}^3_{i_1}[N^3_{i_1}] \cup \bar{N}^3_{i_2}[N^3_{i_1}]\backslash \left\{ \bar{N}^3_{i_2}[N^3_{i_1}] \cap \bar{N}^3_{i_2}[N^3_{i_2}] \right\}$.

Players in $\bar{N}^3_{i_2}[N^r_{i_2}]$ play strategies in A^p because, firstly, no $k \in \bar{N}^3_{i_2}[N^r_{i_2}]$ is a direct neighbour of any $j \in N^r_{i_3}$. That is, if $k \in \bar{N}^3_{i_2}[N^r_{i_2}]$, then $k \notin N_{j_1}$ for all $j \in N^r_{i_3}$. This observation implies that, for each $k \in \bar{N}^3_{i_2}[N^r_{i_2}]$, there exists at least one $j \in N^r_{i_3}$ where $N_{k_1} \cap N_{j_1} \subseteq N^r_{i_2}$, so that $\alpha_k(N^r_{i_2}) \geq \alpha_k(N_{j_1}) = \alpha_k(B_{j_1}) \geq \min_{l \in N_{j_2}} \alpha_l(B_{j_1}) \geq \underline{\eta}(G) \geq p$. Secondly, since all players in $N^r_{i_2}$ play strategies in A^p at $t = 2$, strategies in A^p are best responses to all $k \in \bar{N}^3_{i_2}[N^r_{i_2}]$.

Players in $N^r_{i_3}$ play strategies in A^p because for each $h \in N^r_{i_3}$, $N_{h_1} \cap N_{i_2} \subseteq N^r_{i_2}$, which implies that $\alpha_h(N^r_{i_2}) = \alpha_h(N_{i_2}) = \alpha_h(B_{i_2}) \geq \underline{\eta}(G) \geq p$. And since all players in $N^r_{i_2}$ play strategies in A^p at $t = 2$, strategies in A^p are best responses to all $h \in N^r_{i_3}$.

$\left\{ \bar{N}^r_{i_1}\backslash \bar{N}^3_{i_1}[N^3_{i_1}] \right\} \cup \bar{N}^3_{i_2}\backslash \left\{ \bar{N}^3_{i_2}[N^r_{i_2}] \cup \bar{N}^3_{i_2}[N^3_{i_1}]\backslash \left\{ \bar{N}^3_{i_2}[N^3_{i_1}] \cap \bar{N}^3_{i_2}[N^3_{i_2}] \right\} \right\} \cup \bar{N}^r_{i_3} \cup N_{i_5} \cup$ $N_{i_7} \cup \cdots \cup N_{i_{d_i}} \to A$.

$t = 4$	$i \cup N_{i_1} \cup N^r_{i_2} \cup \bar{N}^3_{i_2}[N^3_{i_2}] \cup \bar{N}^3_{i_2}[N^r_{i_1}] \setminus \left\{ \bar{N}^3_{i_2}[N^3_{i_1}] \cap \bar{N}^3_{i_2}[N^3_{i_2}] \right\} \cup N_{i_3} \cup N_{i_5} \cup N_{i_7} \cup \cdots \cup N_{i_{d_i}} \to$ A^p. Players in $i \cup N^r_{i_2} \cup \bar{N}^3_{i_2}[N^3_{i_2}] \cup \bar{N}^3_{i_2}[N^r_{i_1}] \setminus \left\{ \bar{N}^3_{i_2}[N^3_{i_1}] \cap \bar{N}^3_{i_2}[N^3_{i_2}] \right\} \cup N_{i_3} \cup N_{i_5} \cup$ $N_{i_7} \cup \cdots \cup N_{i_{d_i}}$ all play strategies in A^p for the same reasons outlined in $t = 2$ and $t = 3$ above. To see why all players in N_{i_1} play strategies in A^p, first notice that $N^3_{i_1} \cup \bar{N}^3_{i_1}[N^3_{i_1}] \cup$ $\bar{N}^3_{i_1}[\bar{N}^3_{i_1}[N^3_{i_1}]] = N_{i_1}$; and similarly, $N^r_{i_1} \cup \bar{N}^r_{i_1}[N^r_{i_1}] \cup \bar{N}^r_{i_1}[\bar{N}^r_{i_1}[N^r_{i_1}]] = N_{i_1}$. Now, players in $N^3_{i_1}$ play strategies in A^p for the same reasons outlined in $t = 2$. Players in $\bar{N}^r_{i_1}[N^r_{i_1}]$ play strategies in A^p because, firstly, for each $k \in \bar{N}^r_{i_1}[N^r_{i_1}]$, and any $j \in N^r_{i_2}$, we have $N_{k_1} \cap N_{j_1} \subseteq N^r_{i_1} \cup \bar{N}^r_{i_2}[N^r_{i_2}]$, where $\bar{N}^r_{i_2}[N^r_{i_2}] \subseteq \bar{N}^3_{i_2}[N^r_{i_2}] \cup N^3_{i_2} \setminus N^r_{i_2}$. This implies that for each $k \in \bar{N}^r_{i_1}[N^r_{i_1}]$, there exists at least one $j \in N^r_{i_2}$ such that $\alpha_k \left(N^r_{i_1} \cup \bar{N}^r_{i_2}[N^r_{i_2}] \right) \geq \alpha_k(N_{j_1}) = \alpha_k(B_{j_1}) \geq \min_{l \in N_{j_2}} \alpha_l(B_{j_1}) \geq \underline{\eta}(G) \geq p$. Secondly, since all players in $N^r_{i_1} \cup \bar{N}^r_{i_2}[N^r_{i_2}]$ play strategies in A^p at $t = 3$, strategies in A^p are best responses to all $k \in \bar{N}^r_{i_1}[N^r_{i_1}]$. Players in $\bar{N}^3_{i_1}[\bar{N}^3_{i_1}[N^3_{i_1}]]$ play strategies in A^p because, firstly, each $k \in \bar{N}^3_{i_1}[\bar{N}^3_{i_1}[N^3_{i_1}]]$ is not directly connected to any player in $N^3_{i_1}$, but there exists at least one $j \in N^3_{i_1}$ where $N_{k_1} \cap N_{j_1} \subseteq \bar{N}^3_{i_1}[N^3_{i_1}] \cup \bar{N}^3_{i_2}[N^3_{i_2}]$. This implies that for each $k \in \bar{N}^3_{i_1}[\bar{N}^3_{i_1}[N^3_{i_1}]]$, there exists at least one $j \in N^3_{i_1}$ where $\alpha_k \left(\bar{N}^3_{i_1}[N^3_{i_1}] \cup \bar{N}^3_{i_2}[N^3_{i_2}] \right) \geq \alpha_k(N_{j_1}) = \alpha_k(B_{j_1}) \geq \min_{l \in N_{j_2}} \alpha_l(B_{j_1}) \geq \underline{\eta}(G) \geq p$. Secondly, since all players in $\bar{N}^3_{i_1}[N^3_{i_1}] \cup \bar{N}^3_{i_2}[N^r_{i_1}]$ play strategies in A^p at $t = 3$, strategies in A^p are best responses to all $k \in \bar{N}^3_{i_1}[\bar{N}^3_{i_1}[N^3_{i_1}]]$. Finally, since $N^r_{i_1} \subseteq N^3_{i_1}$ and $\bar{N}^r_{i_1}[\bar{N}^r_{i_1}[N^r_{i_1}]] \subseteq \bar{N}^3_{i_1}[\bar{N}^3_{i_1}[N^3_{i_1}]] \cup N^3_{i_1} \setminus N^r_{i_1}$, it follows that all players in $N^r_{i_1} \cup \bar{N}^r_{i_1}[N^r_{i_1}] \cup \bar{N}^r_{i_1}[\bar{N}^r_{i_1}[N^r_{i_1}]] = N_{i_1}$ play strategies in A^p at $t = 4$. $\bar{N}^r_{i_2} \setminus \left\{ \bar{N}^3_{i_2}[N^3_{i_2}] \cup \bar{N}^3_{i_2}[N^r_{i_1}] \setminus \left\{ \bar{N}^3_{i_2}[N^3_{i_1}] \cap \bar{N}^3_{i_2}[N^3_{i_2}] \right\} \right\} \cup N_{i_4} \cup N_{i_6} \cup N_{i_8} \cup \cdots \cup N_{i_{d_i-1}} \to$ A.
$t = 5$	$i \cup N_{i_1} \cup N_{i_2} \cup N^r_{i_3} \cup N_{i_4} \cup N_{i_6} \cup N_{i_8} \cup \cdots \cup N_{i_{d_i-1}} \to A^p$. Players in $i \cup N_{i_1} \cup N^r_{i_3} \cup N_{i_4} \cup N_{i_6} \cup N_{i_8} \cup \cdots \cup N_{i_{d_i-1}}$ all play strategies in A^p for the same reasons outlined in $t = 2$, $t = 3$ and $t = 4$ above. Players in N_{i_2} play strategies in A^p because each $k \in N_{i_2}$ has $\alpha_k(B_{i_1}) \geq \underline{\eta}(G) \geq p$, and that all players in $B_{i_1} = i \cup N_{i_1}$ play strategies in A^p at $t = 4$. $\bar{N}^r_{i_3} \cup N_{i_5} \cup N_{i_7} \cup N_{i_9} \cup \cdots \cup N_{i_{d_i}} \to A$.
$t = 6$	$B_{i_1} \cup N_{i_2} \cup N_{i_3} \cup N_{i_5} \cup N_{i_7} \cup N_{i_9} \cup \cdots \cup N_{i_{d_i}} \to A^p$. $N_{i_4} \cup N_{i_6} \cup N_{i_8} \cup N_{i_{10}} \cup N_{i_{12}} \cup \cdots \cup N_{i_{d_i-1}} \to A$.
$t = 7$	$B_{i_4} \cup N_{i_6} \cup N_{i_8} \cup \cdots \cup N_{i_{d_i-1}} \to A^p$. $N_{i_5} \cup N_{i_7} \cup \cdots \cup N_{i_{d_i}} \to A$.
$--$	$-------------------$
$t = d_i + 2$	$B_{i_{d_i-1}} \to A^p$; and $N_{i_{d_i}} \to A$.
$t = d_i + 3$	$B_{i_{d_i}} = N \to A^p$.

Thus, $n^r_{i_1}$ mutations to strategies in A^p sufficiently trigger an exit from the basin of attraction of $W \in \mathcal{C}(\mathbf{A})$ to $\mathbf{A}^p$. Since this holds for all $i \in N$, it follows that $\underline{n}^*_1 \leq \underline{n}^r_1$ mutations to strategies in A^p also sufficiently trigger an exit from $D(W)$ to $\mathbf{A}^p$, so that

$C(W, \mathbf{A}^p) \leq \underline{n}_1^*$.

Now, consider again the two configurations $\mathbf{y}, \mathbf{z} \in W \in \mathcal{C}(\mathbf{A})$ defined above, and let $\mathbf{y}$ be a subsequent configuration to $\mathbf{z}$ so that $P(\mathbf{z}, \mathbf{y}) = 1$. Let (A, U, N, G, P) start from $\mathbf{z}$ at $t = 0$, and let a subset of players $S_{i_1}^r \subseteq N_{i_1}^r$, for $4 \leq r \leq d_i$, mutate to strategies in $A \backslash A^p$ at $t = 1$. Given that $d_i \geq 4$ for all $i \in N$, (A, U, N, G, P) evolves from $t = 1$ onward as follows.

$t = 1$	$N_{i_1}^3 \backslash S_{i_1}^r \cup N_{i_3} \cup N_{i_5} \cup N_{i_7} \cup \cdots \cup N_{i_{d_i}} \to A^p$.
	$S_{i_1}^r \to A \backslash A^p$.
	$i \cup \bar{N}_{i_1}^3 \cup N_{i_2} \cup N_{i_4} \cup \cdots \cup N_{i_{d_i - 1}} \to A$.
$t = 2$	$N_{i_2}^4 \cup N_{i_4} \cup N_{i_6} \cup N_{i_8} \cup \cdots \cup N_{i_{d_i - 1}} \to A^p$. Players in $N_{i_2}^4$ play strategies in A^p because for each $k \in N_{i_2}^4$, there exists at least one $j \in N_{i_4}$ with $N_{k_1} \cap N_{j_1} \subseteq N_{i_3}^4$, which implies that $\alpha_k(N_{i_3}^4) \geq \alpha_k(N_{j_1}) = \alpha_k(B_{j_1}) \geq \min_{l \in N_{j_2}} \alpha_l(B_{j_1}) \geq \underline{\eta}(G) \geq p$. And since all players in $N_{i_3}^4 \subseteq N_{i_3}$ play strategies in A^p at $t = 1$, strategies in A^p are best responses to all $k \in N_{i_2}^4$.
	Players in N_{i_4} play strategies in A^p because for each $k \in N_{i_4}$, $\alpha_k(N_{i_3}^4) = \alpha_k(B_{i_3}) \geq \underline{\eta}(G) \geq p$. And since all players in $N_{i_3}^4 \subseteq N_{i_3}$ play strategies in A^p at $t = 1$, strategies in A^p are best responses to all $k \in N_{i_4}$.
	$i \cup N_{i_1} \cup \bar{N}_{i_2}^4 \cup N_{i_3} \cup N_{i_5} \cup \cdots \cup N_{i_{d_i}} \to A$. Players in N_{i_1}, including players in $S_{i_1}^r$, play strategies in A because they are surrounded by direct neighbours in $i \cup S_{i_1}^r \cup \bar{N}_{i_1}^3 \cup N_{i_2}$ that all play strategies in A at $t = 1$. Similarly, there is no guarantee that players in $\bar{N}_{i_2}^4$, or even players in $N_{i_2}^3 \backslash N_{i_2}^4$, will play strategies in A^p. For example, all players in $\bar{N}_{i_2}^4[N_{i_1}^4]$ play strategies in A because, although all their direct neighbours in $N_{i_3} \cup N_{i_1}^3 \backslash N_{i_1}^4$ play strategies in A^p at $t = 1$, all their direct neighbours in $N_{i_1}^4$ play strategies in $A \backslash A^p$.
$t = 3$	$N_{i_1}^4 \cup N_{i_3}^4 \cup N_{i_5} \cup N_{i_7} \cup \cdots \cup N_{i_{d_i}} \to A^p$. Players in $N_{i_1}^4$ play strategies in A^p because for each $k \in N_{i_1}^4$, there exists at least one $j \in N_{i_3}^4$ with $N_{k_1} \cap N_{j_1} \subseteq N_{i_2}^4$, which implies that $\alpha_k(N_{i_2}^4) \geq \alpha_k(B_{j_1}) \geq \min_{l \in N_{j_2}} \alpha_l(B_{j_1}) \geq \underline{\eta}(G) \geq p$. And since all players in $N_{i_2}^4$ play strategies in A^p at $t = 2$, strategies in A^p are best responses to all $k \in N_{i_1}^4$.
	All $k \in N_{i_3}^4$ play strategies in A^p because each has $\alpha_k(N_{i_2}^4) = \alpha_k(B_{i_2}) \geq \underline{\eta}(G) \geq p$, and that all players in $N_{i_2}^4$ play strategies in A^p at $t = 2$.
	$i \cup \bar{N}_{i_1}^4 \cup N_{i_2} \cup \bar{N}_{i_3}^4 \cup N_{i_4} \cup N_{i_6} \cup \cdots \cup N_{i_{d_i - 1}} \to A$.
$t = 4$	$i \cup N_{i_2}^4 \cup N_{i_4} \cup N_{i_6} \cup \cdots \cup N_{i_{d_i - 1}} \to A^p$. Player i plays a strategy in A^p because for each $k \in N_{i_2}^4$, $N_{k_1} \cap N_{i_1} \subseteq N_{i_1}^4$, which implies that, for any $k \in N_{i_2}^4$, $\alpha_i(N_{i_1}^4) \geq \alpha_i(B_{k_1}) \geq \underline{\eta}(G) \geq p$. And since all players in $N_{i_1}^4$ play strategies in A^p at $t = 1$, a strategy in A^p must be a best response to player i.
	All $k \in N_{i_2}^4$ play strategies in A^p because each has $\alpha_k(N_{i_1}^4) = \alpha_k(B_{i_1}) \geq \underline{\eta}(G) \geq p$, and that all players in $N_{i_1}^4$ play strategies in A^p at $t = 3$.
	$N_{i_1} \cup \bar{N}_{i_2}^4 \cup N_{i_3} \cup N_{i_5} \cup \cdots \cup N_{i_{d_i}} \to A$.

$$t = 5 \quad \left| \begin{array}{l} N_{i_1}^4 \cup N_{i_3}^4 \cup N_{i_5} \cup N_{i_7} \cup \cdots \cup N_{i_{d_i}} \to A^p. \\ i \cup \bar{N}_{i_1}^4 \cup N_{i_2} \cup \bar{N}_{i_3}^4 \cup N_{i_4} \cup N_{i_6} \cup \cdots \cup N_{i_{d_i-1}} \to A. \end{array} \right.$$

Thus, if (A, U, N, G, P) exits the basin of attraction of W after less or equal to n_1^* mutations (i.e., after a subset of players in $N_{i_1}^r$, for any $i \in N$, mutate) to strategies in $A \backslash A^p$, then it converges to an absorbing cycle, $W' \in \mathcal{C}(\mathbf{A}^p)$, where players in $N_{i_1}^4 \cup N_{i_3}^4 \cup N_{i_5} \cup \cdots \cup N_{i_{d_i}}$ and $i \cup N_{i_2}^4 \cup N_{i_4} \cup N_{i_6} \cup \cdots \cup N_{i_{d_i-1}}$ alternate between strategies in A^p and some strategies in A, and the rest play some strategies in A. Depending on the network structure (e.g., when $N_{i_1}^3 = N_{i_1}^4$, $N_{i_2}^3 = N_{i_2}^4$ and $N_{i_3} = N_{i_3}^4$), it is also possible for (A, U, N, G, P) to revert to $W \in \mathcal{C}(\mathbf{A}^p)$, an absorbing cycle where players in $N_{i_1}^3 \cup N_{i_3} \cup N_{i_5} \cup \cdots \cup N_{i_{d_i}}$ and $i \cup N_{i_2}^3 \cup N_{i_4} \cup N_{i_6} \cup \cdots \cup N_{i_{d_i-1}}$ alternate between strategies in A^p and some strategies in A, in which case $R(W) = C(W, \mathbf{A}^p)$ (see the above discussion that).

Now, let $\mathbf{u}, \mathbf{v} \in W' \in \mathcal{C}(\mathbf{A}^p)$, with $P(\mathbf{u}, \mathbf{v}) = 1$, be two subsequent configurations where $\mathbf{u}$ consist of all players in $N_{i_1}^4 \cup N_{i_3}^4 \cup N_{i_5} \cup \cdots \cup N_{i_{d_i}}$ playing strategies in A^p, and all players in $i \cup \bar{N}_{i_1}^4 \cup N_{i_2} \cup \bar{N}_{i_3}^4 \cup N_{i_4} \cup N_{i_6} \cup \cdots \cup N_{i_{d_i-1}}$ play strategies in A; and configuration $\mathbf{v}$ consists of all players in $i \cup N_{i_2}^4 \cup N_{i_4} \cup N_{i_6} \cup \cdots \cup N_{i_{d_i-1}}$ playing strategies in A^p, while all players in $N_{i_1} \cup \bar{N}_{i_2}^4 \cup N_{i_3} \cup N_{i_5} \cup \cdots \cup N_{i_{d_i}}$ play strategies in A. Let (A, U, N, G, P) start from $\mathbf{u}$, and let all players in $N_{i_1}^4$, mutate to strategies in A^p at $t = 1$. Then (A, U, N, G, P) evolves from $t = 1$ onward as follows:

$$t = 1 \quad \left| \begin{array}{l} i \cup N_{i_1}^4 \cup N_{i_2}^4 \cup N_{i_4} \cup N_{i_6} \cup \cdots \cup N_{i_{d_i-1}} \to A^p; \\ \bar{N}_{i_1}^4 \cup \bar{N}_{i_2}^4 \cup N_{i_3} \cup N_{i_5} \cup \cdots \cup N_{i_{d_i}} \to A. \end{array} \right.$$

$t = 2$	$i \cup N_{i_1}^4 \cup N_{i_2}^4 \cup \bar{N}_{i_2}^4[N_{i_2}^4] \cup N_{i_3}^4 \cup N_{i_5} \cup \cdots \cup N_{i_{d_i}} \to A^p$. Player i plays a strategy in A^p because for each $k \in N_{i_2}^4$, $N_{k_1} \cap N_{i_1} \subseteq N_{i_1}^4$, which implies that, for any $k \in N_{i_2}^4$, $\alpha_i(N_{i_1}^4) \geq \alpha_i(N_{k_1}) = \alpha_i(B_{k_1}) \geq \min_{l \in N_{k_2}} \alpha_l(B_{k_1}) \geq \underline{\eta}(G) \geq p$. And since all players in $N_{i_1}^4$ play strategies in A^p at $t = 1$, a strategy in A^p must be a best response to player i.

Players in $N_{i_1}^4$ play strategies in A^p because for each $k \in N_{i_1}^4$, there exists at least one $j \in N_{i_3}^4$ with $N_{k_1} \cap N_{j_1} \subseteq N_{i_2}^4$, which implies that $\alpha_k(N_{i_2}^4) \geq \alpha_k(N_{j_1}) = \alpha_k(B_{j_1}) \geq \min_{l \in N_{j_2}} \alpha_l(B_{j_1}) \geq \underline{\eta}(G) \geq p$. And since all players in $N_{i_2}^4$ play strategies in A^p at $t = 1$, strategies in A^p are best responses to all $k \in N_{i_1}^4$.

Players in $N_{i_2}^4$ play strategies in A^p because for each $k \in N_{i_2}^4$, $N_{k_1} \cap N_{i_1} \subseteq N_{i_1}^4$, which implies that $\alpha_k(N_{i_1}^4) = \alpha_k(N_{i_1}) = \alpha_k(B_{i_1}) \geq \underline{\eta}(G) \geq p$. And since all players in $N_{i_1}^4$ play strategies in A^p at $t = 1$, strategies in A^p are best responses to all $k \in N_{i_2}^4$. Similarly, players in $N_{i_3}^4$ play strategies in A^p because for each $k \in N_{i_3}^4$, $N_{k_1} \cap N_{i_2} \subseteq N_{i_2}^4$, which implies that $\alpha_k(N_{i_2}^4) = \alpha_k(B_{i_2}) \geq \underline{\eta}(G) \geq p$. And since all players in $N_{i_2}^4$ play strategies in A^p at $t = 1$, strategies in A^p are best responses to all $k \in N_{i_3}^4$.

Players in $N_{i_5} \cup N_{i_7} \cup \cdots \cup N_{i_{d_i}}$ play strategies in A^p because each $l \in N_{i_r}$, for $r = 5, 7, \cdots, d_i$, has $\alpha_l(N_{i_{r-1}}) = \alpha_l(B_{i_{r-1}}) \geq \underline{\eta}(G) \geq p$. And since all players in $N_{i_{r-1}}$, for all $r = 5, 7, \cdots, d_i$, play strategies in A^p at $t = 1$, strategies in A^p are best responses to all players in $N_{i_5} \cup N_{i_7} \cup \cdots \cup N_{i_{d_i}}$.

$\bar{N}_{i_1}^4 \cup \bar{N}_{i_2}^4 \cup \bar{N}_{i_3}^4 \cup N_{i_4} \cup N_{i_6} \cup \cdots \cup N_{i_{d_i-1}} \to A$.

$t = 3$	$i \cup N_{i_1}^4 \cup N_{i_2}^4 \cup N_{i_3}^4 \cup \bar{N}_{i_3}^4[N_{i_3}^4] \cup N_{i_4} \cup N_{i_6} \cup \cdots \cup N_{i_{d_i-1}} \to A^p$. Players in $i \cup N_{i_1}^4 \cup N_{i_2}^4 \cup N_{i_3}^4 \cup N_{i_6} \cup \cdots \cup N_{i_{d_i-1}}$ play strategies in A^p for the same reasons outlined in $t = 2$.

To see why players in $\bar{N}_{i_3}^4[N_{i_3}^4]$ play strategies in A^p, first observe that no $k \in \bar{N}_{i_3}^4[N_{i_3}^4]$ is a direct neighbour of any $j \in N_{i_4}$. This observation implies that, for each $k \in \bar{N}_{i_3}^4[N_{i_3}^4]$, there exists at least one $j \in N_{i_4}$ where $N_{k_1} \cap N_{j_1} \subseteq N_{i_3}^4$, so that $\alpha_k(N_{i_3}^4) \geq \alpha_k(N_{j_1}) = \alpha_k(B_{j_1}) \geq \min_{l \in N_{j_2}} \alpha_l(B_{j_1}) \geq \underline{\eta}(G) \geq p$. Secondly, since all players in $N_{i_3}^4$ play strategies in A^p at $t = 2$, strategies in A^p are best responses to all $k \in \bar{N}_{i_3}^4[N_{i_3}^4]$. Similarly, players in N_{i_4} all play strategies in A^p because for each $k \in N_{i_4}$, $N_{k_1} \cap N_{i_3} \subseteq N_{i_3}^4$, which implies that $\alpha_h(N_{i_3}^4) = \alpha_h(B_{i_3}) \geq \underline{\eta}(G) \geq p$.

$\bar{N}_{i_1}^4 \cup \bar{N}_{i_2}^4 \cup \left\{ \bar{N}_{i_3}^4 \backslash \bar{N}_{i_3}^4[N_{i_3}^4] \right\} \cup N_{i_5} \cup N_{i_7} \cup \cdots \cup N_{i_{d_i}} \to A$.

$t = 4$	$i \cup N_{i_1}^4 \cup N_{i_2}^4 \cup \bar{N}_{i_2}^4[N_{i_2}^4] \cup N_{i_3}^4 \cup \bar{N}_{i_3}^4[N_{i_3}^4] \cup \bar{N}_{i_3}^4[N_{i_2}^4] \cup N_{i_4} \cup N_{i_5} \cup N_{i_7} \cup \cdots \cup N_{i_{d_i}} \to A^p$. Players in $i \cup N_{i_1}^4 \cup N_{i_2}^4 \cup N_{i_3}^4 \cup \bar{N}_{i_3}^4[N_{i_3}^4] \cup N_{i_4} \cup N_{i_5} \cup N_{i_7} \cup \cdots \cup N_{i_{d_i}}$ play strategies in A^p for the same reasons outlined in $t = 2$ and $t = 3$. Players in $\bar{N}_{i_2}^4[N_{i_2}^4]$ and $\bar{N}_{i_3}^4[N_{i_2}^4] \setminus \left\{ \bar{N}_{i_3}^4[N_{i_2}^4] \cap \bar{N}_{i_3}^4[N_{i_3}^4] \right\}$ play strategies in A^p because for each $k \in \bar{N}_{i_2}^4[N_{i_2}^4] \cup \bar{N}_{i_3}^4[N_{i_2}^4] \setminus \left\{ \bar{N}_{i_3}^4[N_{i_2}^4] \cap \bar{N}_{i_3}^4[N_{i_3}^4] \right\}$, there exists at least one $j \in N_{i_3}^4$ with $N_{k_1} \cap N_{j_1} \neq \emptyset$. For any $j \in N_{i_3}^4$ satisfying the preceding property, $N_{k_1} \cap N_{j_1} \subseteq N_{i_2}^4 \cup \bar{N}_{i_3}^4[N_{i_3}^4]$, which implies that $\alpha_k \left(N_{i_2}^4 \cup \bar{N}_{i_3}^4[N_{i_3}^4] \right) \geq \alpha_k(N_{j_1}) = \alpha_k(B_{j_1}) \geq \min_{l \in N_{j_2}} \alpha_l(B_{j_1}) \geq \underline{\eta}(G) \geq p$. And since all players in $N_{i_2}^4 \cup \bar{N}_{i_3}^4[N_{i_3}^4]$ play strategies in A^p at $t = 3$, strategies in A^p are best responses to all players in $\bar{N}_{i_2}^4[N_{i_2}^4]$ and $\bar{N}_{i_3}^4[N_{i_2}^4] \setminus \left\{ \bar{N}_{i_3}^4[N_{i_2}^4] \cap \bar{N}_{i_3}^4[N_{i_3}^4] \right\}$. $\bar{N}_{i_1}^4 \cup \left\{ \bar{N}_{i_2}^4 \setminus \bar{N}_{i_2}^4[N_{i_2}^4] \right\} \cup \left\{ \bar{N}_{i_3}^4 \setminus \left\{ \bar{N}_{i_3}^4[N_{i_2}^4] \cup \bar{N}_{i_3}^4[N_{i_3}^4] \right\} \right\} \cup N_{i_6} \cup N_{i_8} \cup \cdots \cup N_{i_{d_i-1}} \to A$.
$t = 5$	$i \cup N_{i_1}^4 \cup \bar{N}_{i_1}^4[N_{i_1}^4] \cup N_{i_2}^4 \cup \bar{N}_{i_2}^4[N_{i_2}^4] \cup \bar{N}_{i_2}^4[N_{i_1}^4] \cup N_{i_3}^4 \cup \bar{N}_{i_3}^4[N_{i_3}^4] \cup \bar{N}_{i_3}^4[N_{i_2}^4] \cup N_{i_4} \cup N_{i_5} \cup N_{i_6} \cup N_{i_8} \cup \cdots \cup N_{i_{d_i-1}} \to A^p$. Players in $i \cup N_{i_1}^4 \cup N_{i_2}^4 \cup \bar{N}_{i_2}^4[N_{i_2}^4] \cup N_{i_3}^4 \cup \bar{N}_{i_3}^4[N_{i_3}^4] \cup \bar{N}_{i_3}^4[N_{i_2}^4] \cup N_{i_4} \cup N_{i_5} \cup N_{i_6} \cup N_{i_8} \cup \cdots \cup N_{i_{d_i-1}}$ play strategies in A^p for the same reasons outlined in $t = 2$, $t = 3$ and $t = 4$. Players in $\bar{N}_{i_1}^4[N_{i_1}^4]$ and $\bar{N}_{i_2}^4[N_{i_1}^4] \setminus \left\{ \bar{N}_{i_2}^4[N_{i_1}^4] \cap \bar{N}_{i_2}^4[N_{i_2}^4] \right\}$ play strategies in A^p because for each $k \in \bar{N}_{i_1}^4[N_{i_1}^4] \cup \bar{N}_{i_2}^4[N_{i_1}^4] \setminus \left\{ \bar{N}_{i_2}^4[N_{i_1}^4] \cap \bar{N}_{i_2}^4[N_{i_2}^4] \right\}$, there exists at least one $j \in N_{i_2}^4$ with $N_{k_1} \cap N_{j_1} \neq \emptyset$. For any $j \in N_{i_2}^4$ satisfying the preceding property, $N_{k_1} \cap N_{j_1} \subseteq N_{i_1}^4 \cup \bar{N}_{i_2}^4[N_{i_2}^4]$, which implies that $\alpha_k \left(N_{i_1}^4 \cup \bar{N}_{i_2}^4[N_{i_2}^4] \right) \geq \alpha_k(N_{j_1}) = \alpha_k(B_{j_1}) \geq \min_{l \in N_{j_2}} \alpha_l(B_{j_1}) \geq \underline{\eta}(G) \geq p$. And since all players in $N_{i_1}^4 \cup \bar{N}_{i_2}^4[N_{i_2}^4]$ play strategies in A^p at $t = 3$, strategies in A^p are best responses to all players in $\bar{N}_{i_1}^4[N_{i_1}^4]$ and $\bar{N}_{i_2}^4[N_{i_1}^4] \setminus \left\{ \bar{N}_{i_2}^4[N_{i_1}^4] \cap \bar{N}_{i_2}^4[N_{i_2}^4] \right\}$. $\left\{ \bar{N}_{i_1}^4 \setminus \bar{N}_{i_1}^4[N_{i_1}^4] \right\} \cup \left\{ \bar{N}_{i_2}^4 \setminus \left\{ \bar{N}_{i_2}^4[N_{i_2}^4] \cup \bar{N}_{i_2}^4[N_{i_1}^4] \right\} \right\} \cup \left\{ \bar{N}_{i_3} \setminus \left\{ \bar{N}_{i_3}^4[N_{i_2}^4] \cup \bar{N}_{i_3}^4[N_{i_3}^4] \right\} \right\} \cup N_{i_7} \cup N_{i_9} \cup \cdots \cup N_{i_{d_i}} \to A$.
$t = 6$	$i \cup N_{i_1} \cup N_{i_2}^4 \cup \bar{N}_{i_2}^4[N_{i_2}^4] \cup \bar{N}_{i_2}^4[N_{i_1}^4] \cup N_{i_3}^4 \cup \bar{N}_{i_3}^4[N_{i_3}^4] \cup \bar{N}_{i_3}^4[N_{i_2}^4] \cup N_{i_4} \cup N_{i_5} \cup N_{i_6} \cup N_{i_7} \cup N_{i_9} \cup \cdots \cup N_{i_{d_i}} \to A^p$. Players in $i \cup N_{i_2}^4 \cup \bar{N}_{i_2}^4[N_{i_2}^4] \cup \bar{N}_{i_2}^4[N_{i_1}^4] \cup N_{i_3}^4 \cup \bar{N}_{i_3}^4[N_{i_3}^4] \cup \bar{N}_{i_3}^4[N_{i_2}^4] \cup N_{i_4} \cup N_{i_5} \cup N_{i_6} \cup N_{i_7} \cup N_{i_9} \cup \cdots \cup N_{i_{d_i}}$ play strategies in A^p for the same reasons outlined in $t = 2$, $t = 3$, $t = 4$ and $t = 5$. To see why players in N_{i_1} all play strategies in A^p, first recall that $N_{i_1} = N_{i_1}^4 \cup \bar{N}_{i_1}^4[N_{i_1}^4] \cup \bar{N}_{i_1}^4[\bar{N}_{i_1}^4[N_{i_1}^4]]$. Players in $N_{i_1}^4 \cup \bar{N}_{i_1}^4[N_{i_1}^4]$ play strategies in A^p for the same reasons outlined in $t = 5$. Players in $\bar{N}_{i_1}^4[\bar{N}_{i_1}^4[N_{i_1}^4]]$ play strategies in A^p because, firstly, each $k \in \bar{N}_{i_1}^4[\bar{N}_{i_1}^4[N_{i_1}^4]]$ is not directly connected to any player in $N_{i_1}^4$, but there exists at least one $j \in N_{i_1}^4$ where $N_{k_1} \cap N_{j_1} \subseteq \bar{N}_{i_1}^4[N_{i_1}^4] \cup \bar{N}_{i_2}^4[N_{i_1}^4]$. This implies that for each $k \in \bar{N}_{i_1}^4[\bar{N}_{i_1}^4[N_{i_1}^4]]$, there exists at least one $j \in N_{i_1}^4$ where $\alpha_k \left(\bar{N}_{i_1}^4[N_{i_1}^4] \cup \bar{N}_{i_2}^4[N_{i_1}^4] \right) \geq \alpha_k(N_{j_1}) = \alpha_k(B_{j_1}) \geq \min_{l \in N_{j_2}} \alpha_l(B_{j_1}) \geq \underline{\eta}(G) \geq p$. Secondly, since all players in $\bar{N}_{i_1}^4[N_{i_1}^4] \cup \bar{N}_{i_2}^4[N_{i_1}^4]$ play strategies in A^p at $t = 5$, strategies in A^p are best responses to all $k \in \bar{N}_{i_1}^4[\bar{N}_{i_1}^4[N_{i_1}^4]]$. $\left\{ \bar{N}_{i_2}^4 \setminus \left\{ \bar{N}_{i_2}^4[N_{i_2}^4] \cup \bar{N}_{i_2}^4[N_{i_1}^4] \right\} \right\} \cup \left\{ \bar{N}_{i_3}^4 \setminus \left\{ \bar{N}_{i_3}^4[N_{i_2}^4] \cup \bar{N}_{i_3}^4[N_{i_3}^4] \right\} \right\} \cup N_{i_8} \cup N_{i_{10}} \cup \cdots \cup N_{i_{d_i-1}} \to A$.

$t=7$	$i \cup N_{i_1} \cup N_{i_2} \cup N_{i_3}^4 \cup \bar{N}_{i_3}^4[N_{i_3}^4] \cup \bar{N}_{i_3}^4[N_{i_2}^4] \cup N_{i_4} \cup N_{i_5} \cup N_{i_6} \cup N_{i_7} \cup N_{i_8} \cup N_{i_{10}} \cup$ $\cdots \cup N_{i_{d_i-1}} \to A^p$, where players in N_{i_2} all play strategies in A^p because for each $k \in N_{i_2}$, $\alpha_k(B_{i_1}) \geq \underline{\eta}(G) \geq p$, and that all players in B_{i_1} play strategies in A^p at $t=6$. $\left\{ \bar{N}_{i_3}^4 \setminus \left\{ \bar{N}_{i_3}^4[N_{i_2}^4] \cup \bar{N}_{i_3}^4[N_{i_3}^4] \right\} \right\} \cup N_{i_9} \cup N_{i_{11}} \cup \cdots \cup N_{i_{d_i}} \to A.$
$t=8$	$B_{i_2} \cup N_{i_3} \cup N_{i_4} \cup N_{i_5} \cup N_{i_6} \cup N_{i_7} \cup N_{i_8} \cup N_{i_9} \cup N_{i_{11}} \cup \cdots \cup N_{i_{d_i}} \to A^p$, where players in N_{i_3} all play strategies in A^p because for each $k \in N_{i_3}$, $\alpha_k(B_{i_2}) \geq \underline{\eta}(G) \geq p$, and that all players in B_{i_2} play strategies in A^p at $t=7$. $N_{i_{10}} \cup N_{i_{12}} \cup \cdots \cup N_{i_{d_i-1}} \to A.$
$t=9$	$B_{i_{10}} \cup N_{i_{12}} \cup N_{i_{14}} \cup \cdots \cup N_{i_{d_i-1}} \to A^p.$ $N_{i_{11}} \cup N_{i_{13}} \cup \cdots \cup N_{i_{d_i}} \to A.$
$--$	$-----------------$
$t=d_i-2$	$B_{i_{d_i-1}} \to A^p$; and $N_{i_{d_i}} \to A.$
$t=d_i-1$	$B_{i_{d_i}} = N \to A^p.$

Thus, $\underline{n}_1^4$ mutations to strategies in A^p sufficiently trigger an exit from the basin of attraction of W' to $\mathbf{A}^p$, so that $C(W', \mathbf{A}^p) \leq \underline{n}_1^4$. However, as we demonstrated in the dynamics of (A, U, N, G, P) from W to W' above, if, starting from W', $\underline{n}_1^4$ players mutate to strategies in $A \setminus A^p$, (A, U, N, G, P) will revert to W'. Specifically, let (A, U, N, G, P) start from $\mathbf{v}$ at $t=0$, and let all players in $N_{i_1}^4$ mutate to strategies in $A \setminus A^p$ at $t=1$.[29] Given that $d_i \geq 4$ for all $i \in N$, (A, U, N, G, P) evolves from $t=1$ onward as follows.

$t=1$	$N_{i_3}^4 \cup N_{i_5} \cup N_{i_7} \cup \cdots \cup N_{i_{d_i}} \to A^p.$ $i \cup N_{i_1} \cup N_{i_2} \cup N_{i_4} \cup N_{i_6} \cup \cdots \cup N_{i_{d_i-1}} \to A.$
$t=2$	$N_{i_2}^4 \cup N_{i_4} \cup N_{i_6} \cup N_{i_8} \cup \cdots \cup N_{i_{d_i-1}} \to A^p$. Players in $N_{i_2}^4$ play strategies in A^p because for each $k \in N_{i_2}^4$, there exists at least one $j \in N_{i_4}$ with $N_{k_1} \cap N_{j_1} \subseteq N_{i_3}^4$, which implies that $\alpha_k(N_{i_3}^4) \geq \alpha_k(N_{j_1}) = \alpha_k(B_{j_1}) \geq \min_{l \in N_{j_2}} \alpha_l(B_{j_1}) \geq \underline{\eta}(G) \geq p$. And since all players in $N_{i_3}^4 \subseteq N_{i_3}$ play strategies in A^p at $t=1$, strategies in A^p are best responses to all $k \in N_{i_2}^4$. Players in N_{i_4} play strategies in A^p because for each $k \in N_{i_4}$, $\alpha_k(N_{i_3}^4) = \alpha_k(B_{i_3}) \geq \underline{\eta}(G) \geq p$. And since all players in $N_{i_3}^4 \subseteq N_{i_3}$ play strategies in A^p at $t=1$, strategies in A^p are best responses to all $k \in N_{i_4}$. $i \cup N_{i_1} \cup \bar{N}_{i_2}^4 \cup N_{i_3} \cup N_{i_5} \cup \cdots \cup N_{i_{d_i}} \to A.$

---

[29]Starting from $\mathbf{v}$ rather than $\mathbf{u}$ presents the best possible starting point of exiting the basin of attraction of W' after $n_{i_1}^4$ mutations to strategies in $A \setminus A^p$.

$t = 3$ | $N_{i_1}^4 \cup N_{i_3}^4 \cup N_{i_5} \cup N_{i_7} \cup \cdots \cup N_{i_{d_i}} \to A^p$. Players in $N_{i_1}^4$ play strategies in A^p because for each $k \in N_{i_1}^4$, there exists at least one $j \in N_{i_3}^4$ with $N_{k_1} \cap N_{j_1} \subseteq N_{i_2}^4$, which implies that $\alpha_k(N_{i_2}^4) \geq \alpha_k(B_{j_1}) \geq \min_{l \in N_{j_2}} \alpha_l(B_{j_1}) \geq \underline{\eta}(G) \geq p$. And since all players in $N_{i_2}^4$ play strategies in A^p at $t = 2$, strategies in A^p are best responses to all $k \in N_{i_1}^4$.

All $k \in N_{i_3}^4$ play strategies in A^p because each has $\alpha_k(N_{i_2}^4) = \alpha_k(B_{i_2}) \geq \underline{\eta}(G) \geq p$, and that all players in $N_{i_2}^4$ play strategies in A^p at $t = 2$.

$i \cup \bar{N}_{i_1}^4 \cup N_{i_2} \cup \bar{N}_{i_3}^4 \cup N_{i_4} \cup N_{i_6} \cup \cdots \cup N_{i_{d_i-1}} \to A$.

$t = 4$ | $i \cup N_{i_2}^4 \cup N_{i_4} \cup N_{i_6} \cup \cdots \cup N_{i_{d_i-1}} \to A^p$. Player i plays a strategy in A^p because for each $k \in N_{i_2}^4$, $N_{k_1} \cap N_{i_1} \subseteq N_{i_1}^4$, which implies that, for any $k \in N_{i_2}^4$, $\alpha_i(N_{i_1}^4) \geq \alpha_i(B_{k_1}) \geq \underline{\eta}(G) \geq p$. And since all players in $N_{i_1}^4$ play strategies in A^p at $t = 1$, a strategy in A^p must be a best response to player i.

All $k \in N_{i_2}^4$ play strategies in A^p because each has $\alpha_k(N_{i_1}^4) = \alpha_k(B_{i_1}) \geq \underline{\eta}(G) \geq p$, and that all players in $N_{i_1}^4$ play strategies in A^p at $t = 3$.

$N_{i_1} \cup \bar{N}_{i_2}^4 \cup N_{i_3} \cup N_{i_5} \cup \cdots \cup N_{i_{d_i}} \to A$.

$t = 5$ | $N_{i_1}^4 \cup N_{i_3}^4 \cup N_{i_5} \cup N_{i_7} \cup \cdots \cup N_{i_{d_i}} \to A^p$.

$i \cup \bar{N}_{i_1}^4 \cup N_{i_2} \cup \bar{N}_{i_3}^4 \cup N_{i_4} \cup N_{i_6} \cup \cdots \cup N_{i_{d_i-1}} \to A$.

Thus, $n_{i_1}^4 \geq \underline{n}_1^4$, for any $i \in N$, mutations to strategies in $A \backslash A^p$ are not sufficient to trigger an exit from the basin of attraction of W'. This implies that $R(W') = C(W', \mathbf{A}^p) \leq \underline{n}_1^4$.

Putting the above three scenarios together, we see that, for any $\mathbf{x} \notin D(\mathbf{A}^p)$, $\underline{b}_1$ mutations can trigger evolution from $\mathbf{x}$ to $\mathbf{A}^p$ so that $C(\mathbf{x}, \mathbf{A}^p) \leq \underline{b}_1$. Otherwise, there exists at least one absorbing cycle $W \in \mathcal{C}(\mathbf{A})$, with $C(\mathbf{x}, W) \leq \underline{b}_1$, where players in $N_{i_1}^3 \cup N_{i_3} \cup N_{i_5} \cup \cdots \cup N_{i_{d_i}}$ and $i \cup N_{i_2}^3 \cup N_{i_4} \cup N_{i_6} \cup \cdots \cup N_{i_{d_i-1}}$ alternate between strategies in A^p and some strategies in A, and the rest play some strategies in A. For this scenario, the number of mutations to strategies in A^p that trigger evolution from W to $\mathbf{A}^p$ is less or equal to $\underline{n}_1^*$, so that if the number of mutations to strategies in $A \backslash A^p$ needed to trigger an exit from $D(W)$ is greater than $\underline{n}_1^*$, then $R(W) = C(W, \mathbf{A}^p)$. However, if the number of mutations to strategies in $A \backslash A^p$ needed to trigger an exit from the basin of attraction of W is less or equal to $\underline{n}_1^*$, then there exists another absorbing cycle $W' \in \mathcal{C}(\mathbf{A})$ and $W' \neq W$, where $R(W) = C(W, W') \leq \underline{n}_1^*$, and $R(W') = C(W', \mathbf{A}^p) \leq \underline{n}_1^4$. Since these three scenarios hold for any $\mathbf{x} \notin D(\mathbf{A}^p)$, it follows from the discussion in Section 4, together with equations (9) and (10), that $CR(\mathbf{A}^p) \leq \underline{b}_1$ and

$$CR^*(\mathbf{A}^p) = \max_{\mathbf{x} \notin D(\mathbf{A}^p)} C^*(\mathbf{x}, \mathbf{A}^p) = \max_{\mathbf{x} \notin D(\mathbf{A}^p)} \min_{W \in \mathcal{C}(\mathbf{A})} C(\mathbf{x}, W) \leq \underline{b}_1.$$

The lower bound of the radius of $\mathbf{A}^p$

We now show that when $p \leq \underline{\eta}(G)$ and $d(G) \geq 7$, the radius of $\mathbf{A}^p$ is bounded from below by $R(\mathbf{A}^p) \geq \underline{b}_1 + \iota$, where $\iota \geq 1$ is some positive integer. Specifically, for an undirected, unweighted and strongly connected network G with $d(G) \geq 7$, which also means that $d_i \geq 4$ for all $i \in N$, if $p \leq \underline{\eta}(G)$, then $\underline{b}_1$ mutations to strategies in $A \backslash A^p$ cannot trigger an exit from $D(\mathbf{A}^p)$.

We first show that if (A, U, N, G, P) starts, at $t = 0$, from any configuration, $\mathbf{x} \notin D(\mathbf{A}^p)$, where all players in B_{i_2} of any $i \in N$ play strategies in A^p, then it will converge to $\mathbf{A}^p$. Indeed, since, by definition of the contagion threshold, each $h \in B_{i_2} \cup N_{i_3}$ has $\alpha_h(B_{i_2}) \geq \underline{\eta}(G) \geq p$ (i.e. each $h \in B_{i_2} \cup N_{i_3}$ has at least $\underline{\eta}(G) \geq p$ of her direct neighbours within B_{i_2}), and all players in B_{i_1} have all their direct neighbours within B_{i_2}, it follows that all players in B_{i_3} will play strategies in A^p from $t = 2$ onward. From $t = 3$ onward, all players in B_{i_4} play strategies in A^p because each $h \in B_{i_3} \cup N_{i_4}$ has $\alpha_h(B_{i_3}) \geq \underline{\eta}(G) \geq p$. This iterative process continues until $t = d_i - 1$ when all players in $B_{i_{d_i}} = N$ play strategies in A^p because each $h \in B_{i_{d_i - 1}} \cup N_{i_{d_i}}$ has $\alpha_h(B_{i_{d_i - 1}}) \geq \underline{\eta}(G) \geq p$.

Now, for any $i \in N$, if $d_i \geq 4$, then for all $j \in N_{i_4} \cup N_{i_5} \cup \cdots \cup N_{i_{d_i}}$, we have $B_{j_2} \cap B_{i_1} = \emptyset$ (i.e. there is no overlap between the first-neighbourhood of i and the second-neighbourhood of j). This implies that, if (A, U, N, G, P) starts from any $\mathbf{x} \in \mathbf{A}^p$ at $t = 0$, and all players in B_{i_1} mutate to strategies in $A \backslash A^p$ at $t = 1$, then (A, U, N, G, P) will revert to $\mathbf{A}^p$ (i.e. converge to some state within $\mathbf{A}^p$) because: firstly, there exists at least one player $j \in N_{i_4} \cup N_{i_5} \cup \cdots \cup N_{i_{d_i}}$ for whom all players in B_{j_2} play strategies in A^p at $t = 1$; secondly, (A, U, N, G, P) converges to $\mathbf{A}^p$ from any configuration $\mathbf{x} \notin D(\mathbf{A}^p)$ where all players in B_{j_2} of any $j \in N$ play strategies in A^p.

Thus, when $\underline{\eta}(G) \geq p$ and $d(G) \geq 7$, more than b_{i_1} mutations are required to trigger an exit from the basin of attraction of $\mathbf{A}^p$. Since this holds for all $i \in N$, it follows that $R(\mathbf{A}^p) \geq \underline{b}_1 + \iota$, for $\iota \geq 1$.

Stochastically stable set of states/strategies

Putting the above results together, we see that, for an undirected, unweighted and strongly connected network with $d(G) \geq 7$, 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 $\iota \geq 1$, which implies that $R(\mathbf{A}^p) > CR^*(\mathbf{A}^p)$, and hence, $\mathbf{A}^p$, and strategies in A^p, are uniquely stochastically stable.

Acknowledgements

References

Alós-Ferrer C, Weidenholzer S (2006) Imitation, local interactions, and efficiency. Economics Letters 93(2):163–168

Alós-Ferrer C, Weidenholzer S (2007) Partial bandwagon effects and local interactions. Games and Economic Behavior 61(2):179–197

Alós-Ferrer C, Weidenholzer S (2008) Contagion and efficiency. Journal of Economic Theory 143(1):251–274

Azomahou TT, Opolot D (2018) Strategic diffusion in networks through contagion. Available at SSRN 3323076

Bini D, Capovani M, Romani F, Lotti G (1979) O (n2. 7799) complexity for n× n approximate matrix multiplication. Information Processing Letters 8(5):234–235

Blume LE (1995) The statistical mechanics of best-response strategy revision. Games and Economic Behavior 11(2):111–145

Chen HC, Chow Y, Wu LC (2013) Imitation, local interaction, and coordination. International Journal of Game Theory 42(4):1041–1057

DellaVigna S (2009) Psychology and economics: Evidence from the field. Journal of Economic literature 47(2):315–72

Ellison G (1993) Learning, local interaction, and coordination. Econometrica 61(5):1047–1071

Ellison G (2000) Basins of attraction, long-run stochastic stability, and the speed of step-by-step evolution. Review of Economic Studies 67(1):17–45

Foster D, Young P (1990) Stochastic evolutionary game dynamics. Theoret Population Biol

Kandori M, Mailath GJ, Rob R (1993) Learning, mutation, and long run equilibria in games. Econometrica 61(1):29–56

Lee IH, Valentinyi A (2000) Noisy contagion without mutation. Review of Economic Studies 67(1):47–56

Lee IH, Szeidl A, Valentinyi A (2003) Contagion and State Dependent Mutations. The BE Journal of Theoretical Economics 3(1):1–29

Maruta T (2002) Binary games with state dependent stochastic choice. Journal of Economic Theory 103(2):351–376

Morris S (2000) Contagion. Review of Economic Studies 67(1):57–78

Morris S, Rob R, Shin HS (1995) p-Dominance and Belief Potential. Econometrica 63(1):145–57

Myatt DP, Wallace C (2003) A multinomial probit model of stochastic evolution. Journal of Economic Theory 113(2):286–301

Neary PR (2012) Competing conventions. Games and Economic Behavior 76(1):301–328

Newton J, Sawa R (2015) A one-shot deviation principle for stability in matching problems. Journal of Economic Theory 157:1–27

Opolot D (2020) Contagion, network cohesion and long run stability in evolutionary games. Available at SSRN 3678191

Oyama D, Takahashi S (2015) Contagion and uninvadability in local interaction games: The bilingual game and general supermodular games. Journal of Economic Theory 157:100–127

Oyama D, Sandholm WH, Tercieux O (2015) Sampling best response dynamics and deterministic equilibrium selection. Theoretical Economics 10(1):243–281

Peski M (2010) Generalized risk-dominance and asymmetric dynamics. Journal of Economic Theory 145(1):216–248

Sandholm WH (2001) Almost global convergence to p-dominant equilibrium. International Journal of Game Theory pp 107–116

Sandholm WH (2010) Orders of limits for stationary distributions, stochastic dominance, and stochastic stability. Theoretical Economics 5(1):1–26

Sandholm WH, Staudigl M (2016) Large deviations and stochastic stability in the small noise double limit. Theoretical Economics 11(1):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(1):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(1):45–70

Weidenholzer S (2012) Long-run equilibria, dominated strategies, and local interactions. Games and Economic Behavior 75(2):1014–1024

Young HP (1993) The evolution of conventions. Econometrica 61(1):57–84

Young HP (1998) Individual Strategy and Social Structure: An Evolutionary Theory of Institutions. Princeton University Press