THE UNIVERSITY OF SYDNEY

Economics Working Paper Series

2013 - 02

Coalitions, tipping points and the speed of evolution

Jonathan Newton & Simon D. Angus

October 2013

Coalitions, tipping points and the speed of evolution☆

Jonathan Newton[a,1], Simon D. Angus[b]

[a] *School of Economics, University of Sydney.*
[b] *Department of Economics, Monash University.*

Abstract

This study considers pure coordination games on networks and the waiting time for an adaptive process of strategic change to achieve efficient coordination. Although it is in the interest of every player to coordinate on a single globally efficient norm, coalitional behavior at a local level can greatly slow, as well as hasten convergence to efficiency. For some networks, when one action becomes efficient enough relative to the other, the effect of coalitional behavior changes abruptly from a conservative effect to a reforming effect. These effects are confirmed for a variety of stylized and empirical social networks found in the literature. For coordination games in which the Pareto efficient and risk dominant equilibria differ, polymorphic states can be the only stochastically stable states.

Keywords: Evolution, stochastic stability, learning, coalition, social norm, reform, conservatism, networks, social networks.
JEL: C71, C72, C73

1. Introduction

Why do some innovations spread rapidly and others slowly? Why are some innovations never adopted, even though they are inexpensive and

☆The authors would like to thank Murali Agastya, Suren Basov, Larry Blume, Zhiwei Cui, David Easley, Sung-Ha Hwang, Claudio Mezzetti, Birendra Rai, Ryoji Sawa, Russell Toth and Galina Zudenkova for detailed comments and fruitful questions, as well as seminar audiences at UNSW, University of Melbourne, Boston University, Cornell University, McGill University, University of Lausanne, University of Zurich, Seoul National University, Sogang University and ANU.

[1]Come rain or come shine, I can be reached at jonathan.newton@sydney.edu.au, telephone +61293514429. This work was completed while the author was supported by a Discovery Early Career Researcher Award funded by the Australian Research Council.

	A	B
A	α, α	$0, 0$
B	$0, 0$	$1, 1$

Figure 1: A two player pure coordination game, $\alpha > 1$.

methods of equilibrium selection would point to their universal adoption? An answer to these questions should consider waiting times. It takes time for any novelty or innovation to be adopted. If the expected waiting time to adoption of an innovation is long, it may be superceded and rendered redundant before it has become widespread. The question of whether it would have eventually been adopted, had the world remained the same in all other respects following its invention, is then moot. Given the importance of waiting times, it is natural to query how they are affected by sensible behavioral assumptions. The behavioral assumption we make in the current paper is the following: from time to time, players with interdependent payoffs come together and adjust their actions to their mutual benefit. That is to say, they form temporary coalitions.

This paper examines the effect of coalitional behavior on expected waiting times for processes to reach long run equilibria. We focus on two action pure coordination games with one efficient action A, and one inefficient action B, as illustrated in figure 1. The relative efficiency of the efficient action to the inefficient action is given by the parameter α. The set of players with whom any given player interacts is governed by an underlying network. A long run equilibrium (stochastically stable state) is the state in which every player plays the efficient action. The waiting time for the process to reach this long run equilibrium can thus be understood as the delay before a society converges to an efficient social norm. In line with the theoretical predictions of Olson (1965) and much of the subsequent literature on collective action, we are particularly interested in the effect of joint strategic switching by coalitions which are small relative to the total population size.[2]

Two possible effects of coalitional behavior are discovered, a *reforming* effect and a *conservative* effect. For high values of α, we observe a reforming

[2]See also Poteete and Ostrom (2004). There also exist important provisos to such predictions (Chamberlin, 1974), particularly in the presence of punishment (Mathew and Boyd, 2011; Hwang, 2009).

effect: convergence to the long run equilibrium is much faster when coalitional behavior is allowed. Less obviously, for some networks and low values of α, there is a conservative effect: convergence to the long run equilibrium is much slower in the presence of coalitional behavior. These effects, taken singly or together, imply that coalitional behavior increases the sensitivity of convergence speeds to the relative efficiency of competing norms.

Several network types display tipping point effects. For values of α below some threshold $\underline{\alpha}$, coalitional behavior has a conservative effect. For values of α above some threshold $\bar{\alpha}$, coalitional behavior has a reforming effect. In some instances $\underline{\alpha}$ and $\bar{\alpha}$ take the same value. These speed-of-convergence tipping points are driven by preferences in a similar way to those of Granovetter (1978) and not by informational concerns such as in, for example, Bikhchandani et al. (1992). However, in comparison to Granovetter (1978) or Ellison (1993), additional tipping points are created at the values of α above or below which certain coalitional deviations become optimal. For example, it only makes sense for a group to coordinate a break away from a current norm if the additional payoffs the members of the group generate amongst themselves outweigh their losses from miscoordinating with the rest of the population. The principal results of the paper are as follows:

(i) For any network, a reforming effect is observed for large enough α. Furthermore, for any α, a reforming effect is observed if large enough coalitions can form.

(ii) The notion of a *parochial* set of players is defined recursively, building outwards from some core players who are completely isolated from the network outside of the parochial set. It is shown that parochial sets of players are the only sets which are immune to conservative effects for any α and coalition size.

(iii) If all Nash equilibria for a network involve every player choosing the same action, then there cannot be a conservative effect for any α and coalition size. This set of networks includes the ring network and the complete network.

(iv) We characterize sets of players which are immune to 'contagion' (in the sense of Morris, 2000) by the efficient action. In the absence of coalitional behavior, Morris (2000) gives an external stability condition: such a set of players must be sufficiently insular. In the presence of coalitional behavior, there is also an internal stability condition:

the set must not contain small groups of players which are insular enough to profitably coordinate a switch to the efficient action.

The model can be considered a stylized representation of multiple phenomena. These include the dissemination of ideas and socio-cultural memes, the spread of process innovation within and between firms, and the choice of consumer technology (e.g. mobile phone network providers). It can also be considered a model of online platforms used for purposes such as photo sharing (e.g. Flickr, My Opera, Fotki, Fotolog), microblogging (e.g. Twitter, Tumbler, Plurk, Jaiku), or internet telephony (e.g. Skype, Google talk, Oovoo). Essentially, it can be considered a model of any setting in which (i) the main benefit or cost of the choice of action by a player arises through his interaction with others, and (ii) switching costs between actions are relatively low. For example, the direct cost of changing the political position one advocates are small compared to the costs that arise through social interaction as a consequence of making such a change. The same can sometimes be said for the choice of technology to be used for a specific purpose within a firm.[3]

The theoretical results in the paper are confirmed as robust for non-vanishing parameter values via simulations on both stylized and empirical networks. These include coauthorship networks, workplace networks and social networks, including fragments of friendship networks on FACEBOOK.

The question arises as to what the model of the paper gives when instead of a pure coordination game, the players play a game in which one of the Nash Equilibria is Pareto efficient and the other is risk dominant. In these circumstances, it is no longer the case that the state in which every player chooses the efficient action is always stochastically stable. Neither need the state in which all play the risk dominant action be stochastically stable. In fact, long run equilibria may involve heterogeneous action choices by different parts of a population. Consider a population composed of cliques of players such that players within any given clique are densely connected, but each clique is only loosely connected to other cliques. For some parameter

[3]To give an example, employees may choose to keep track of appointments via a paper diary or via particular software integrated with a system of electronic mail. Another example is the choice faced by academic researchers of whether to use a TEX editor or WYSIWYG software such as Scientific Workplace (or even Microsoft Word). Such a choice generates significant payoff externalities for coauthors, as the default source code generated by such software often has to be adjusted before it can be compiled in a TEX editor.

values, the stochastically stable states of such a population involve medium-size cliques coordinating on the efficient equilibrium action and large cliques coordinating on the risk dominant equilibrium action. This gives a justification for the use of smaller teams in large organizations: efficient behavior may not be long-run stable in large operating units.

The paper is organized as follows. Section 2 relates the paper and its results to existing literatures. Section 3 gives the model. Section 4 introduces the ideas of the paper via an example. Section 5 gives some general results for any network. Section 6 studies conservative effects. Sections 7 and 8 discuss small worlds and random regular networks. Section 9 gives results of simulations. Section 10 analyzes the situation where Pareto efficiency and risk dominance do not coincide. Section 11 concludes. All proofs not in the main body of the text are given in the appendix.

2. Relation to existing literatures

2.1. Perturbed adaptive dynamics and stochastic stability

Whereas concepts of equilibrium stability such as asymptotic stability or evolutionary stable strategies (Smith and Price, 1973) analyze robustness to single errors (mutations) in strategies, Foster and Young (1990), Young (1993) and Kandori et al. (1993) use the methods of Freidlin and Wentzell (1984) to measure the robustness of equilibria of an adaptive strategy revision process to multiple errors in players' choice of strategies. They show that although there may be several stationary states in a dynamic process, some of them may be more robust to such errors than others, and that if the probability of errors becomes very small, then in the long run some nonempty subset of stationary states which are relatively robust to such errors will be observed almost all of the time. These are the *stochastically stable* states. Bergin and Lipman (1996) prove a kind of folk theorem for stochastic stability, that is they show that any stable state of the unperturbed dynamic process can be selected with appropriately chosen state-dependent mutation rates. Therefore, the structure given to error probabilities is crucial to the predictions of the model.[4] Naidu et al. (2010) analyze a model in which transitions between stochastically stable states are driven by errors on the part of the players who stand to gain from the move and arrive at different predictions to those of Young (1998a) for games

[4]van Damme and Weibull (2002) give conditions on error probabilities under which the results of Young are recovered.

of contracting.[5] Logit models, in which more costly errors occur with lower probability, are also common in the literature[6].

A pervasive criticism of stochastic stability as a tool of equilibrium selection has been the large lengths of time it can take for the perturbed process, starting at a non-stochastically stable equilibrium of the unperturbed dynamic, to reach a stochastically stable state (Ellison, 1993). This calls into question the empirical validity of stochastic stability: if it takes a billion periods to reach a stochastically stable state, the concept may not be predictively useful on human timescales. Results on waiting times can depend on network topology, network size, and whether the small error probability limit is analyzed or error probabilities are kept fixed.[7] The current paper finds that the potential problem of long waiting times can be considerably worsened or mitigated by orders of magnitude when coalitional behavior is introduced, and whether a worsening or a mitigation occurs depends on knife edge parameter values. The literature sometimes neglects to mention that whether a waiting time is long or short in absolute terms depends on the interpretation of period length. The current study does not aim to prove that convergence is ever fast or slow in absolute terms: the focus is on the comparison of processes with and without coalitional behavior.

2.2. Coalitional behavior

There exists a large literature in cooperative game theory on the behavior of coalitions.[8] Concepts include 'strong equilibrium' (Aumann, 1959), coalition proof Nash equilibrium (Bernheim et al., 1987), farsighted coalitional stability (Konishi and Ray, 2003), and coalitional rationalizability (Ambrus, 2009). There is a small literature on coalitional behavior in perturbed evolutionary models. Newton (2012a) introduces a model of *coalitional stochastic stability* in which the 'errors' in the dynamic process are actually small probabilities of payoff improving behavior by coalitions of players, and shows that this can lead to significant differences in equilibrium selection when compared to random error driven selection. Sawa (2012) adapts coalitional stochastic stability for logit-style dynamics. The model

[5]See Binmore et al. (2003) for a good survey of results in evolutionary bargaining models.

[6]See for example Blume (1993); Alós-Ferrer and Netzer (2010).

[7]See Ellison (2000); Young (1998b, 2011); Kreindler and Young (2013); Montanari and Saberi (2010). Kreindler and Young contains a concise survey of this literature.

[8]For a survey the reader is referred to Peleg and Sudholter (2003).

of Sawa (2012) also features coalitional behavior as part of the unperturbed dynamic. Serrano and Volij (2008) and Newton (2012b) do similarly, applying stochastic stability to models of coalitional recontracting. Matching models such as those found in Jackson and Watts (2002) and Klaus et al. (2010), in which coalitions are pairs of recontracting agents, also fall into this category.

2.3. Networks and local interaction

Some coalition structures can be considered more reasonable than others. Suggestions have been made in the cooperative game theory literature that subcoalitions of coalitions which are allowed to coordinate should also be allowed to coordinate.[9] Alternatively, it has been suggested that the union of coalitions which are allowed to deviate and have a nonempty intersection should also be allowed to deviate, the justification being that players who belong to both potential coalitions could act as intermediaries.[10]

Networks are a natural way to represent payoff effects in games and are also a natural way to delineate feasible coalition structures, for instance by assuming that any coalitional activity is undertaken by connected subgraphs of a graph representing a wider social network.[11] This is the approach taken in the current paper, in which it is assumed that aside from payoff effects, the network ties represent the potential for communication and thus coalition formation between sets of players. The authors believe that coalitional effects are very natural in a local interaction setting such as those analyzed in Ellison (2000, 1993); Eshel et al. (1998). In fact, often the motivations for players' being connected to one another in a network representing payoff effects can double as reasons why coalitional behavior between the players is plausible. However, although networks can facilitate the formation of coalitions, the two are distinct concepts. To quote from the International Encyclopedia of Civil Society (Anheier and Toepler, 2009) in the context of transnational organization: "Sometimes these networks generate the shared goals, mutual trust, and understanding needed to form coalitions capable of collaborating.... But networks do not necessarily coordinate their actions, nor do they necessarily come to agreement on specific joint actions (as implied by the concept of coalition)."

[9] Algaba et al. (2004).
[10] Algaba et al. (2001).
[11] Myerson (1977), Jackson and Wolinsky (1996), Jackson (2005), Kets et al. (2011).

The current paper studies interaction on given networks, and not network formation. This is another reason that waiting times for convergence to a stochastically stable state are important. For very long waiting times, it may not be plausible to assume that the underlying network structure remains static long enough for the long run equilibrium to be reached. Waiting times in such models without coalitional behavior have been the subject of recent studies by Montanari and Saberi (2010) and Young (2011). Montanari and Saberi (2010) examine the effect of network structure on the order of magnitude of convergence times as networks increase in size. Young (2011) bounds convergence times for fixed, non-vanishing error probabilities.

2.4. Homophily

The paper is related to the literature on 'homophily' - the desire of people to associate with those similar to themselves. This is a well documented phenomenon in the sociology literature. For a survey the reader is referred to McPherson et al. (2001). The economic literature on the topic, presaged by Schelling (1969), has been growing of late. For example, Currarini et al. (2009) explain data on friendships via direct assumptions about people's preference to associate with those of a similar race. The most relevant paper in the homophily literature is that of Golub and Jackson (2012), which defines homophily as the level of preferential linking to vertices of the same colour in a random network, then relates this level to convergence times for a simple averaging process. They conclude that homophily slows convergence. This result follows because for an averaging process the important factor is the size of the channels through which innovation can spread, rather than the probability with which any given innovation gains a foothold in the population. To illustrate this point, consider a complete network. An averaging process will quickly converge on such a network, whereas for a two action coordination game such as the one in the current paper it will take many errors (and therefore in expectation a very long time) for the process to begin a move from an inefficient equilibrium to an efficient one.

From the perspective of the current paper and its emphasis on joint strategic switching, we note that aside from the network formation and informational (different types of player access different information) interpretations of homophily as in Golub and Jackson (2012), there may exist a further effect: players of the same type may find it easier to coordinate their changes in action. This could be due to underlying cultural norms or even the perception of similar 'sunspots'. The location of players of the same type close to one another in a network (i.e. homophily) would then facilitate coalitional behavior. The implications of this for convergence times would be ambiguous and follow the proceeding analysis.

2.5. Poverty traps

Finally, we note that the results of the current paper can be considered to illustrate a potential *poverty trap.* A poverty trap is a persistent institution which is harmful to economic growth. In our paper the relevant institution is the ability of small groups of players to behave coalitionally, or equivalently, the habits of trust and cooperation which facilitate such behavior.[12] Intriguingly, we see that the ability of small groups to coordinate their choice of actions can create a poverty trap by slowing the movement of society towards the efficient equilibrium.

It has been shown that norms of kin-based sharing, which would have been evolutionarily advantageous in the past, can create a poverty trap by preventing members of a kin-group from successfully integrating in a modern economy.[13,14] The persistence of kin-sharing norms in such a setting can be seen as a failure of strategic coordination amongst members of a kin-group.[15] In contrast, the poverty trap in our paper can be caused by *coordination success.* In a model of adaptive behavior and poverty traps, Bowles (2004) shows how inegalitarian and inefficient (though not Pareto inefficient) social norms are sustained due to two classes having different preferences over possible norms. In our model there exists an efficient norm which is (weakly) preferred by every player to every other norm.

3. Model

Let N be a finite set of players. Players are arranged in a network, which we represent as a graph $\mathbf{g}$, where $g_{ij} = 1$ if there exists a link between players i and j, and $g_{ij} = 0$ otherwise. We assume that the graph is undirected:

[12] Such norms can be highly persistent. See, for example, Nunn and Wantchekon (2011) on the effect of the slave trade on trust in Africa.

[13] Hoff and Sen (2006). See also Jakiela and Ozier (2012); Baland et al. (2011) on how kin-sharing norms can lead to less profitable investment and borrowing behavior.

[14] This switch from a norm being advantageous to disadvantageous can also be seen where highly inegalitarian societal norms, which evolved to exploit economies of scale in agriculture, later became a hindrance to development (Engerman and Sokoloff, 2006). See also Acemoglu et al. (2002).

[15] See also Akerlof (1976) on the caste system in India.

$g_{ij} = g_{ji}$. Set $g_{ii} = 0$ for all $i \in N$. The network will affect players in two ways:

(i) The network determines the structure of payoffs, in particular which players' actions impose externalities on other players.

(ii) The network mediates joint action by *coalitions* of players.

Let $x_i^t \in \{A, B\}$ denote player *i's* action at time t. Let $x_S^t = \prod_{i \in N} x_i^t$ denote the action profile of all players in $S \subseteq N$ at time t. In the absence of a subscript, $x^t := x_N^t$. Let x_i and x_S denote representative actions and action profiles respectively. Let N_i denote the set of *neighbors* of player i: $N_i = \{j \in N : g_{ij} = 1\}$. For $S \subseteq N$, let $N_S = \left(\bigcup_{i \in S} N_i\right) \setminus S$. Payoffs of player i in period t are given by:

$$u_i(x^t) = \sum_{j \in N_i} \delta(x_i^t, x_j^t)$$

where:

$$\delta(A, A) = \alpha > 1; \qquad \delta(A, B) = \delta(B, A) = 0; \qquad \delta(B, B) = 1.$$

That is, the players play a pure coordination game on the network. If a player chooses action B, his payoff is the number of his neighbors who play action B. If a player chooses action A, his payoff is the number of his neighbors who play action A multiplied by some constant α which is strictly greater than 1. Effectively, the players play their chosen action against each of their neighbors in the game in figure 1. The constant α can be understood to represent some technological superiority of action A over action B, with the magnitude of α representing the magnitude of this superiority. The model can be understood as a threshold model, with $1/(\alpha + 1)$ being the proportion of a player's neighbors who must play A for the player to want to play A.[16] Settings can be considered for which players' thresholds differ, but for the purpose of the current paper, homogeneous thresholds suffice to obtain rich results.[17,18]

[16] Action A is p-dominant for any $p > 1/(\alpha + 1)$ under the definition of Morris et al. (1995).

[17] Note that in a situation where a regulator enforces interoperability of the two 'technologies', giving some constant nonzero payoff for miscoordination, then payoffs can be rescaled back to this setup, corresponding to a higher value of α and a lower threshold.

[18] For many networks and values of α which are not too high, the one-shot game will have a very large number of pure Nash equilibria. This multiplicity can persist when

The underlying dynamic process of this paper is one in which coalitions of players adjust their actions in a coordinated manner. The sets of players which can do this are determined by the underlying network.

Definition 1. A coalition of players $S \subseteq N$ is *feasible* in g, denoted $S|_g$ if and only if for all $i, j \in S$ there exists $\{s_m\}_{m=1}^{m=l}$ such that $s_1 = i$, $s_l = j$, and $s_m \in S$, $g_{s_m s_{m+1}} = 1$ for $m < l$.

That is, S is a feasible coalition if and only if S is a singleton set or there is a path between any two players in S that only uses edges between players in S. That is, players in a feasible coalition either directly interact with one another, or have interactions mediated by other players in the coalition. Another way of stating this is that the network restricted to players in S forms a connected subgraph. It is assumed that N is feasible: $\mathbf{g}$ is a connected network. This is without loss of generality: if the network comprised more than one component, analysis of each component would proceed independently of the other components.

When a coalition chooses its actions, we mandate that it chooses a *better response*. That is, players in the coalition adjust their actions in a coordinated manner such that no member of the coalition loses from the adjustment. Note that the addition of a Pareto condition to define a form of coalitional best response would complicate definitions without changing the results of the paper. Define the set of better responses for a set of players S:

$$A_S(x^t) := \Big\{ x_S : u_i(x_S, x^t_{N \setminus S}) \geq u_i(x^t) \; \forall \, i \in S \Big\}.$$

Let $G_{A_S(x^t)}(.)$ be a probability distribution over $A_S(x^t)$. $G_{A_S(x^t)}(.)$ will determine the actions chosen by a coalition S when it is called upon to better respond. We assume full support on the set of better responses.

Assumption 1. *Each $G_{A_S(x^t)}(.)$ has full support on $A_S(x^t)$.*

We are particularly interested in the effect of coalitional behavior by coalitions which are small relative to the total size of the population. It is natural to assume that there are limits to how large a coalition can be. Such a limit could be a consequence of higher costs of communication for

k-strong Nash equilibria are considered, by which we mean action profiles which are resistant to deviation by coalitions of at most k players.

larger coalitions. One approach would be to bound the maximum path length between coalition members on the network. Another approach, the one taken here, is to limit the maximum number of players in a coalition. Let $\mathcal{N}(k)$ be the set of feasible coalitions of size k or smaller:

$$\mathcal{N}(k) = \{S \subseteq N : (S|_g \text{ and } |S| \leq k)\}.$$

For given k, let $F_k(.)$ be a distribution on $\mathcal{N}(k)$. $F_k(.)$ will determine which coalition gets the opportunity to update its actions in any given period. The process is one of asynchronous updating: only one coalition at a time updates its actions.[19]

Assumption 2. *$F_k(.)$ has full support on $\mathcal{N}(k)$.*

The process of strategy updating is constructed as follows. Each period, a coalition S is chosen according to $F_k(.)$. The coalition decides on an intended new action profile for its members. Denote this intended action profile by y_S^{t+1}. This profile is chosen from the set of better responses $A_S(x^t)$:

$$y_S^{t+1} \sim G_{A_S(x^t)}(.).$$

Following the decision on which actions to take, each player will play his intended action. This is the unperturbed dynamic. A perturbed dynamic is generated by considering the possibility that a player makes a mistake when attempting to play the action he intends to play. Each player in the coalition, independently of the other players, with a small probability ε makes an *error* and chooses an action at random. That is, independently for each $i \in S$:

$$\text{With probability } 1 - \varepsilon : \quad x_i^{t+1} = y_i^{t+1}$$

$$\text{With probability } \varepsilon : \quad x_i^{t+1} \sim U[\{A, B\}].$$

Finally, all players who are not part of the chosen coalition for period t do not update their actions. For all $i \in N \setminus S$:

$$x_i^{t+1} = x_i^t$$

[19]The general implications of the results of the paper do not change if instead of bounding coalition size we bound diameter, or if we allow some possibility of synchronous updating.

So the change in the action profile is determined by a Markov process, $\Phi_{k,\alpha,\varepsilon}$, on state space $X := \{A, B\}^{|N|}$, with transition probabilities $P_{k,\alpha,\varepsilon}(.,.)$ derived from the above description of the process. Let $P^t_{k,\alpha,\varepsilon}(.,.)$ denote the t-step Markov transition probabilities. Each period, this process involves feasible coalitions of players changing their strategies in a payoff improving manner. When choosing his strategy, each player in the coalition independently makes an error with probability ε and chooses an action at random.[20] The process with $\varepsilon = 0$ corresponds to an unperturbed dynamic in which players do not make errors.

Note that for $\varepsilon > 0$, $\Phi_{k,\alpha,\varepsilon}$ is irreducible and aperiodic and therefore has a unique invariant distribution $\pi_{k,\alpha,\varepsilon}$ which is ergodic. Denote the expected time for the process $\Phi_{k,\alpha,\varepsilon}$ to reach state y starting from x by $W_{k,\alpha,\varepsilon}(x, y)$.

Definition 2.

$$\tau_y = \min\{t \geq 0 : \Phi^t_{k,\alpha,\varepsilon} = y\}; \qquad W_{k,\alpha,\varepsilon}(x, y) = \mathbb{E}[\tau_y | \Phi^0_{k,\alpha,\varepsilon} = x]$$

The focus of the paper is on $W_{k,\alpha,\varepsilon}(B^{|N|}, A^{|N|})$, the expected time for the process to move from an inefficient social norm in which every player plays B to an efficient social norm in which every player plays A. We shall occasionally be interested in the set of absorbing states under the process with $\varepsilon = 0$. Denote this set $\Lambda_{k,\alpha}$.

$$\Lambda_{k,\alpha} := \{x \in X : \ P_{k,\alpha,0}(x, x) = 1\}$$

For expositional brevity, we avoid the existence of absorbing cycles under the process with $\varepsilon = 0$ by making the following assumption which holds for generically all values of α.[21]

Assumption 3.

$$\forall\, z \in \mathbb{N}_+,\ z \leq \max_{i \in N} |N_i| : \quad \alpha z \notin \mathbb{N}_+$$

[20] This describes errors in implementation. If errors were instead made in the process by which a coalition chooses its actions, then errors within a coalition could be perfectly correlated and different results would obtain. However, remarkably, even if a probability ε event leads *all* the members of a coalition to make mistakes, conservative effects are still possible. See Appendix D for just such an example. Some correlation between errors in the process is fine: the results of the paper can be appropriately restated. Again, see Appendix D.

[21] This assumption is weaker than assuming α is not rational.

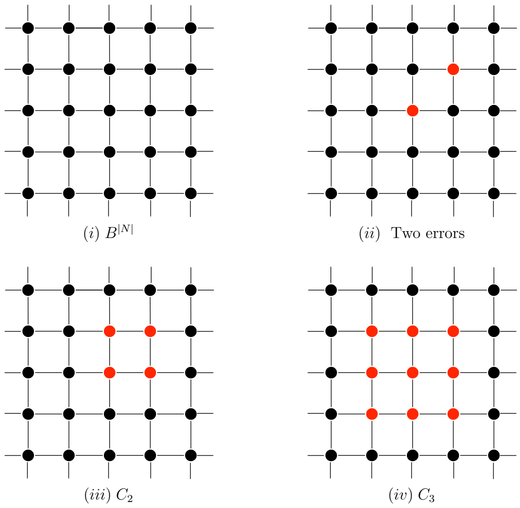


Figure 2: Square lattice in various states. Black vertices play B, red vertices play A.

Throughout the paper, for functions $f(\varepsilon)$, $g(\varepsilon)$, the notation $f(\varepsilon) \in O(g(\varepsilon))$ and $f(\varepsilon) \in \Omega(g(\varepsilon))$ denotes that $f(\varepsilon)$ is asymptotically (as $\varepsilon \to 0$) bounded above or below respectively by some multiple of $g(\varepsilon)$. $f(\varepsilon) \in \Theta(g(\varepsilon))$ denotes that $f(\varepsilon) \in O(g(\varepsilon))$ and $f(\varepsilon) \in \Omega(g(\varepsilon))$.

4. Leading example: the square lattice

Consider the square lattice with von Neumann neighborhood, pictured in figure 2, embedded on a torus as in figure 3.

Consider the benchmark case without coalitional behavior, $k = 1$, $\alpha < 3$. We have that $W_{1,\alpha,\varepsilon}(B^{|N|}, A^{|N|}) \in \Theta(\varepsilon^{-2})$. To see this, consider that two errors (figure 2.ii) are necessary to move to a state C_2 (figure 2.iii) in which a block of four players play A and every other player plays B. From C_2, it takes a single error to move to a state such as C_3 (figure 2.iv) in which a larger block of players plays A. However, at least one error is required to move backwards from C_2 to $B^{|N|}$. That is, the probability of moving to a state in which a larger block of players plays A *conditional on* leaving C_2 is of order 1. This means that the waiting time until $A^{|N|}$ is reached involves the wait for the initial two errors of order ε^{-2}, followed by subsequent waits of order ε^{-1}. These terms combine additively and so the wait for the initial two errors dominates as ε becomes small.[22]

[22]This argument is a slight adaptation of Ellison (2000).

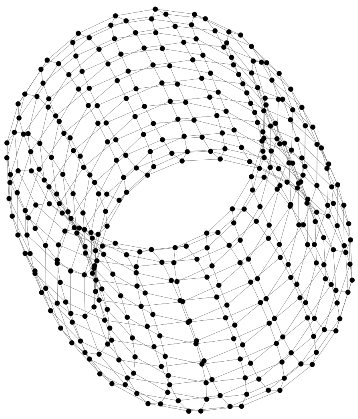


Figure 3: Square lattice on a torus.

Now consider $k = 4$, coalitions including up to four players can form. First, consider the case $\alpha < 3/2$; the relative benefits of the better technology are not so great. Errors are no longer required to leave C_2. If two adjacent A players form a coalition, they can gain by switching to B and achieving a payoff of 3, which is higher than the payoff of 2α they attain in state C_2. In the absence of random errors, C_2 collapses and the process returns to $B^{|N|}$. More than two errors are required to exit the basin of attraction of $B^{|N|}$. Following three errors on a diagonal, the process can attain the state C_3 which has a 3 by 3 block of players playing A. This block of players playing A can expand with the aid of a single error. It is not possible to leave C_3 without the help of errors, no matter how close α is to 1: the players who play A in C_3 form what will later be formally defined as a *parochial* set. To see that errors are required to leave C_3, first consider the player in the centre of the square. He attains his maximum possible payoff of 4α, so he will not intentionally change his action as part of a coalition or otherwise. Secondly, consider the neighbors of the central player. Their payoffs at C_3 are 3α, so they will never intentionally change their action unless the central player also changes his, which he will not. The players at the corners of the block of A players cannot earn more than their C_3 payoff of 2α unless some non-corner player in the square changes his action, which will not occur. Similar arguments to those in the case $k = 1$ lead us to conclude that $W_{4,\alpha,\varepsilon}(B^{|N|}, A^{|N|}) \in \Theta(\varepsilon^{-3})$. Convergence to the efficient social norm is an order of magnitude slower in the presence of coalitional behavior: small outbreaks of innovation are snuffed out as the players involved in the outbreak collaborate to recoordinate with the population as a whole. The possibility of coalitional behavior has led to a *conservative* effect.

Now, consider the case $3/2 \leq \alpha < 2$, $k = 4$. From state C_2, pairs of players who play strategy B in C_2, and who each have a neighbor playing A, can switch together to playing A and attaining a payoff of 2α which is greater than their payoff of 3 in C_2. In this way, the set of players playing A can expand without errors. This speeds up the process of moving to the efficient norm. However, to reach C_2 the initial two errors are still necessary so the order of magnitude of the wait is the same as that for the case without coalitional behavior, $W_{4,\alpha,\varepsilon}(B^{|N|}, A^{|N|}) \in \Theta(\varepsilon^{-2})$.

Finally, consider the case $2 \leq \alpha$, $k = 4$. From state $B^{|N|}$, a square of 4 players can form a coalition. By switching to playing A, thus reaching state C_2, they can attain a payoff of 2α which is greater than their payoff at $B^{|N|}$ of 4. This can happen for any such set of players on the grid.[23] Therefore, the process can move to $A^{|N|}$ without the aid of any errors. $W_{4,\alpha,\varepsilon}(B^{|N|}, A^{|N|}) \in \Theta(1)$. Convergence to the efficient social norm is orders of magnitude faster in the presence of coalitional behavior: coalitions coordinate innovation in the population. The possibility of coalitional behavior has led to a *reforming* effect.

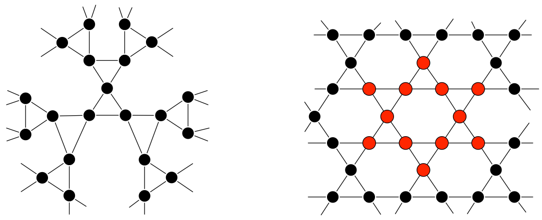
The reasoning of the preceding three paragraphs leads to the following proposition.[24]

Proposition 1. *Let* $\boldsymbol{g}$ *be the* n_1 *by* n_2, $n_1 n_2 = |N|$, *square lattice on a torus with von Neumann neighborhoods, size* $4 \leq k \ll n_1, n_2$. *Then, as* $\varepsilon \to 0$*:*

$$\begin{aligned} \alpha < \frac{3}{2} \quad &\Longrightarrow \quad \frac{W_{k,\alpha,\varepsilon}(B^{|N|}, A^{|N|})}{W_{1,\alpha,\varepsilon}(B^{|N|}, A^{|N|})} \to \infty \\ \frac{3}{2} \leq \alpha < 2 \quad &\Longrightarrow \quad \frac{W_{k,\alpha,\varepsilon}(B^{|N|}, A^{|N|})}{W_{1,\alpha,\varepsilon}(B^{|N|}, A^{|N|})} \in \Theta(1) \\ 2 \leq \alpha \quad &\Longrightarrow \quad \frac{W_{k,\alpha,\varepsilon}(B^{|N|}, A^{|N|})}{W_{1,\alpha,\varepsilon}(B^{|N|}, A^{|N|})} \to 0 \end{aligned}$$

[23] Note that although the underlying game is a potential game with a potential function given by the sum of the payoffs of all the players, local maxima of the potential function are not necessarily absorbing states of the unperturbed dynamic when $k > 1$. For $\alpha < 3$, $B^{|N|}$ is a local maximum of the potential function. The move from $B^{|N|}$ to C_2 changes the potential by $8\alpha - 24$, which is negative for $\alpha < 3$, and yet the move from $B^{|N|}$ to C_2 occurs under the unperturbed dynamic for $k = 4$ and $\alpha \geq 2$.

[24] These results giving existence of reforming and conservative effects dependent on the value of α extend readily to hyper-cubic lattices.



(*i*) Overlapping triangles (*ii*) Kagome lattice at state Kag_3

Figure 4: Kagome lattice and graph of overlapping triangles, with minimum group of A players robust to coalitions highlighted for the Kagome lattice.

Proposition 1 tells us that coalitional behavior can have either a *conservative* or a *reforming* effect. For low values of α, it has a conservative effect: when small groups of players start to play A and form a configuration which is stable under an individual best response dynamic, coalitions of players tear apart the cluster of deviant behavior, taking the process back to the state in which B is played by all. For large values of α, coalitional behavior has a reforming effect: groups of players can coordinate their choice to play A, increasing their payoffs from the change by ensuring that it occurs simultaneously to that of their neighbors. This speeds up the process of convergence to the efficient social norm.[25]

Note that $W_{k,\alpha,\varepsilon}(B^{|N|}, A^{|N|})$ is not necessarily monotonic in k. If $k = |N|$, there is the possibility of the coalition $S = N$ being chosen to respond, and the players in N choosing to play $A^{|N|}$ in a single step without the aid of any errors. Convergence is fast and the formation of the grand coalition has a reforming effect.

Other networks display tipping point effects even starker than those of the square lattice. Consider the regular graph of overlapping triangles with local structure shown in figure 4(i). Action profiles in which a large majority of players play B, but there exist triangles of A players, are stable under the unperturbed dynamic for $k = 1$ when $\alpha < 3$. However, when $k = 2$, $\alpha < {}^3\!/_2$, pairs of A players at the edge of such a set of triangles gain from

[25]Montanari and Saberi (2010) would consider all of these parameter specifications to give fast convergence as the order of magnitude of the waiting time does not increase in population size. Given that period length is undefined, for fixed small ε this could encompass massive differences in actual waiting times. The focus of the current paper is not whether convergence is 'fast' or 'slow' as such (although simulation results can be read this way), but on the effects of coalitional behavior relative to the baseline process without coalitional behavior.

agreeing to switch simultaneously to B to earn payoffs of 3 which are higher than their existing payoffs of 2α. Such a configuration thus unravels and the process returns to $B^{|N|}$. When $\alpha > 3/2$, pairs of B players who have a common neighbor playing A gain from switching to A, so the set of players who play A can expand rapidly following just a single error.

The Kagome lattice is slightly different. Limiting convergence times in this case depend on the number of errors required to attain a configuration such as that in figure 4(ii), which can be considered analagous to the 3 by 3 square of players in the square lattice. Players who play A in this configuration will not switch to B without the aid of random errors. Similar arguments to those for the square lattice show that for both the regular graph of overlapping triangles and the Kagome lattice, coalitional behavior slows convergence to $A^{|N|}$ by an order of magnitude when $\alpha < \frac{3}{2}$ and speeds it by orders of magnitude when $\alpha \geq \frac{3}{2}$. Note that vertices in these networks have clustering coefficients (the proportion of pairs of neighbors who are themselves neighbors) of $1/3$, unlike the square lattice with von-Neumann neighborhoods, for which all vertices have clustering coefficients of 0. This illustrates that the story of reforming and conservative effects is more than just a story about clustering.

The question arises as to how far the results of the preceding paragraphs can be extended to general network architectures. The answer is sometimes unambiguously positive and sometimes not.

5. General results

Here we give some general observations for any network. We start by noting that if α is large enough then there exists a reforming effect: in the limit, the process with coalitional behavior, compared to the process without coalitional behavior, has an infinitely faster transition to $A^{|N|}$. The reasoning behind this result is simple. In state $B^{|N|}$, the players who obtain the highest payoffs are those with the largest number of neighbors. Let one of these players be player i. Let $\alpha > |N_i|$. Then any pair of neighbors in the network can switch to action A and obtain payoffs of α, which is higher than their payoffs in state $B^{|N|}$. So, without any errors occuring, the process can transit to $A^{|N|}$.

Proposition 2. *For any $k \geq 2$, there exists $\bar{\alpha}$ such that, as $\varepsilon \to 0$:*

$$\forall\, \alpha \geq \bar{\alpha} : \quad \frac{W_{k,\alpha,\varepsilon}(B^{|N|}, A^{|N|})}{W_{1,\alpha,\varepsilon}(B^{|N|}, A^{|N|})} \to 0$$

Proof. Choose $\bar{\alpha} = \max_{i \in N} |N_i|$. Then for $\alpha > \bar{\alpha}$, $\Lambda_{k,\alpha} = \{A^{|N|}\}$. □

The value of $\bar{\alpha}$ in the proof of Proposition 2 is an upper bound on the value of α above which a reforming effect is observed. This bound cannot be improved without reference to the network in question or the value of k. To see this, consider the ring network, $k < |N|$. In this setting a single error is necessary and sufficient to transition to $A^{|N|}$ for any $\alpha < 2 = \max_{i \in N} |N_i|$, so there is no reforming effect for α below the bound of Proposition 2.

Take any set of players S which is small enough to engage in coalitional behavior, and is isolated in the sense that every member of S has 'enough' of his neighbors within S. All players in such a set will play A in any absorbing state of the unperturbed process.

Proposition 3. *If $S \subseteq N$, $|S| \leq k$, is such that:*

$$\forall\, i \in S : \quad \alpha > \frac{|N_i|}{|N_i \cap S|}$$

then for all $x \in \Lambda_{k,\alpha}$, $i \in S$, we have that $x_i = A$.

Proof. Assume S is feasible. If not, then analyze each component of S in isolation. If S is chosen to better respond, for any $x^t \in X$, we have $A^{|S|} \in A_S(x^t)$, as either a player is already playing A and cannot be harmed by others switching to A, or is playing B and earning a payoff lower than $|N_i|$ which is lower than the payoff obtained from $A^{|S|}$ which is at least $\alpha|N_i \cap S|$. So a state in which all players in S play A is reached without errors with some probability. The same inequality implies that any player in S would strictly lose were he to switch to B in a later period. □

This result implies that a reforming effect is obtained if k is large enough.[26] If $k = N$, the process can move to $A^{|N|}$ in a single step without

[26]The expression to the right hand side of the inequality in proposition 3 can be understood as a measure of the surface tension of S local to player i (see Vinkovic and Kirman, 2006). That is, coalitions which have low surface tension at every point will rapidly adopt action A. However, the appropriate analogue of the model within the physical sciences is not a dynamic-location—fixed-type model such as Vinkovic and Kirman (2006), which discusses the physical analogue of Schelling (1971). The model is rather a fixed-location—dynamic-type model such as those found in Ising field models (see Blume, 1993; Montanari and Saberi, 2010): both the global topological structure, and individual location within that topology, is fixed for our agents, whilst their type varies under a constantly updating local 'field'.

errors, as every player agrees together to play A. For $k = 1$, at the very least a single error is required to move from $B^{|N|}$ to $A^{|N|}$.

Corollary 1. *For any $\alpha > 1$, there exists $\bar{k}$ such that, as $\varepsilon \to 0$:*

$$\forall\, k \geq \bar{k} : \quad \frac{W_{k,\alpha,\varepsilon}(B^{|N|}, A^{|N|})}{W_{1,\alpha,\varepsilon}(B^{|N|}, A^{|N|})} \to 0$$

Proof. Choose $\bar{k} = |N|$. Choose $S = N$. This gives $|N_i \cap S| = |N_i \cap N| = |N_i| > |N_i|\alpha^{-1}$, so proposition 3 and its proof apply. □

The *basin of attraction* of an absorbing state of the unperturbed dynamic $x \in \Lambda_{k,\alpha}$ is the set of states from which, in the absence of errors, convergence to x is guaranteed. The *basin of possible attraction* is the set of states from which, in the absence of errors, convergence to x is possible. Allowing coalitions of a greater size reduces the set of absorbing states. This is equivalent to the fact that for any game, the set of $(k+1)$-strong Nash equilibria is a subset of the set of k-strong Nash equilibria. Furthermore, an expansion of the set of allowable coalitions also (weakly) expands the basins of possible attraction of absorbing states, as any path which was possible with a lower k, is also possible with a higher k. No such monotonicity exists for basins of attraction.[27]

Definition 3. The basin of attraction of $x \in \Lambda_{k,\alpha}$ is defined as:

$$D_{k,\alpha}(x) = \left\{ y \in X : \; P^t_{k,\alpha,0}(y, x) \to 1 \;\text{ as } t \to \infty \right\}$$

Definition 4. The basin of possible attraction of $x \in \Lambda_{k,\alpha}$ is defined as:

$$\bar{D}_{k,\alpha}(x) = \left\{ y \in X : \; \sum_{t=1}^{\infty} P^t_{k,\alpha,0}(y, x) > 0 \right\}$$

Proposition 4. *Assume $k_1 \leq k_2$. Then $\Lambda_{k_1,\alpha} \supseteq \Lambda_{k_2,\alpha}$. Moreover, $x \in \Lambda_{k_2,\alpha}$ implies that $\bar{D}_{k_1,\alpha}(x) \subseteq \bar{D}_{k_2,\alpha}(x)$.*

[27] By increasing the size of basins of possible attraction, an increase in k will decrease the size of basins of attraction, *as long as the set of stable states remains the same.* If the set of stable states changes, this is no longer the case. Consider the example of section 4 with $\alpha < 3/2$. $C_2 \notin D_{1,\alpha}(B^{|N|})$, but for $k = 4$, C_2 is no longer stable and $C_2 \in D_{4,\alpha}(B^{|N|})$.

Furthermore, we can directly characterize the basin of possible attraction of $A^{|N|}$. This is a result along the lines of Morris (2000) in that it characterizes the states from which 'contagion', the spread of A across the entire population under the unperturbed ($\varepsilon = 0$) dynamic, can occur.[28]

Proposition 5. *For any $x \in X$, $x \in \bar{D}_{k,\alpha}(A^{|N|})$ if and only if there does not exist $S \subseteq \{i : x_i = B\}$ such that:*

$$\forall T \subseteq S,\ |T| \leq k :\ \exists i \in T : \quad \frac{|N_i \setminus S| + |N_i \cap T|}{|N_i| + |N_i \cap T|} < \frac{1}{1+\alpha}$$

That is, contagion cannot occur if there exists a set of players which (i) is insular enough to protect it from being 'infected' by A by the rest of the population but (ii) does not contain subsets of players which are themselves insular enough to coordinate their switch to A. When $k = 1$, T must be a singleton, so $N_i \cap T$ is empty and we have, in essence, the result of Morris (2000). The proof, however, is more similar to that of Easley and Kleinberg (2010). Note that proposition 4 implies that the larger is k, the larger is the basin of possible attraction of $A^{|N|}$, from which contagion can occur. However, the size of the basin of possible attraction is only part of the story, and we have already seen in section 4 that although larger k increases the size of $\bar{D}_{k,\alpha}(A^{|N|})$, it can also increase the waiting time until it is reached. To emphasize: increasing k has a monotonic effect on deterministic contagion in the style of Morris (2000); it can have a non-monotonic effect on the contagion of the current model.

6. Conservative effects

It was seen in section 5 that a reforming effect of coalitional behavior is always possible for large enough values of α. Is a conservative effect similarly always possible? The answer is no, as can be seen if we consider ring networks. In such a network each vertex is connected to m neighbors on either side. For $k = 1$, the only absorbing states of the unperturbed dynamic are $A^{|N|}$ and $B^{|N|}$. No intermediate equilibria exist for any m, α. The only possible effect of coalitional behavior is then to speed the transition. This is a general result in the absence of intermediate equilibria. From any state, without random errors, the process will converge to either

[28] Morris (2000) pays specific attention to countably infinite populations for which the basin of possible attraction contains states in which a finite number of players play A.

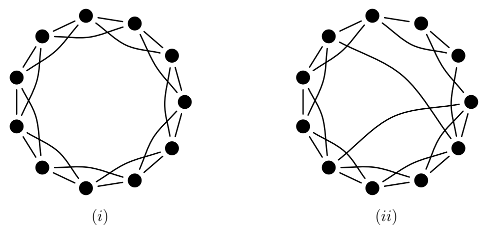


Figure 5: (i) Ring, connected to 2 neighbors on either side. (ii) Ring with rewiring.

$A^{|N|}$ or $B^{|N|}$. Any number of errors that is enough to move the process into the basin of possible attraction of $A^{|N|}$ when $k = 1$ is also enough to move the process into the basin of possible attraction of $A^{|N|}$ when $k > 1$.

Proposition 6. *If for* $k = 1$, $\Lambda_{k,\alpha} = \{A^{|N|}, B^{|N|}\}$, *then for any* $k \geq 1$*:*

$$\frac{W_{k,\alpha,\varepsilon}(B^{|N|}, A^{|N|})}{W_{1,\alpha,\varepsilon}(B^{|N|}, A^{|N|})} \in O(1)$$

So, rings and square lattices give very different results when it comes to predicting the impact of coalitional behavior on adaptive dynamics. This is important as each is a commonly used model of local interaction. Moreover, they are common starting points for the construction of small world networks. The question of whether such small worlds retain the properties derived for square lattices and ring networks is addressed in the next section. Another important network that satisfies the conditions of proposition 6 is the complete network, in which every player neighbors every other player. The complete network and the ring are very different networks: in the class of connected networks with $|N|$ vertices, the complete network has the greatest number of links ($\frac{|N|(|N|-1)}{2}$); the ring with $m = 1$ has $(|N| + 1)$ links, one more than the lowest possible number.

Now we turn our attention to a method for showing the existence of a conservative effect for small enough α. This requires us to define the notion of a *parochial* set of players.

Definition 5. For $S \subseteq N$, define:

$$\begin{aligned} I_0(S) &= \{i \in S : \quad N_i \subseteq S\}, \\ I_m(S) &= \{i \in S : \quad |N_i \setminus S| \leq |N_i \cap I_{m-1}(S)|\}, \qquad m \geq 1. \end{aligned}$$

Note that $I_{m-1}(S) \subseteq I_m(S)$. We say that S is *parochial* if there exists $m \geq 0$ such that $S = I_m(S)$.

Note that unlike the concept of isolation used in the statement of proposition 3, the definition of a parochial set does not depend on the value of α. A parochial set, S, always contains a set of players, $I_0(S)$, who are completely isolated from the network outside of S. $I_1(S)$ is then formed from all of the players in S who have at least as much exposure to $I_0(S)$ as they have to the network outside of S. $I_2(S)$ is similarly defined, and so on. Intuitively, the recursive definition means that every member of a parochial set has at least as many neighbors who are more deeply embedded in the set than he is, than he has neighbors outside of the set. Every member of a parochial set will have at least half of his neighbors within the set, but this fact alone does not suffice to make a set parochial. In fact, any set of players, S, in which every player has at least one neighbor outside of the set cannot be parochial, as $I_0(S)$ will be empty.

Define $\mathcal{P}_A$ as the set of states such that the set of players playing A contains a parochial subset.

Definition 6.

$$\mathcal{P}_A = \{x \in X : \quad \exists\, S \subseteq \{i \in N : \; x_i = A\} \text{ such that } S \text{ is parochial.}\}$$

If S is the set of players playing A and S is not parochial then there exist, for some k, α, nonempty sets of A players who are not in $I_m(S)$ for any m, and who can gain by coordinating a switch back to B. Iterating, the process can return to either $B^{|N|}$, or a state in which the set of players playing A is a parochial set. Conversely, if some parochial set of players is playing A, for any values of k and α, no member of the set will ever switch to B unless at least one member of the set makes an error.

Proposition 7. *$x \in \mathcal{P}_A$ if and only if $\nexists\, k, \alpha$ such that $x \in \bar{D}_{k,\alpha}(B^{|N|})$.*

As the process is time homogeneous and has a finite state space, this implies that under parameters satisfying proposition 7, the process $\Phi_{k,\alpha,0}$ will either enter $\mathcal{P}_A$ or reach state $B^{|N|}$. Leaving the basin of attraction of $B^{|N|}$ implies entering the basin of possible attraction of $\mathcal{P}_A$. This bounds the waiting time to reach $A^{|N|}$ from below by the waiting time to reach $\mathcal{P}_A$. That is, the waiting time to reach $A^{|N|}$ is at least of the order of ε to the power of the number of errors required to reach a state in which some parochial subset of players play A.

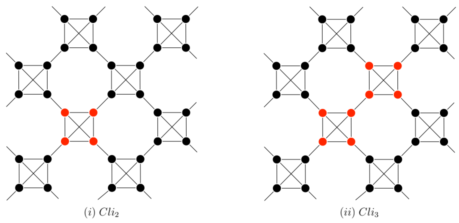


Figure 6: Interconnected cliques.

The use of this result can be seen in that for the square and Kagome lattices, states C_3 and Kag_3 respectively are in the set $\mathcal{P}_A$. For α close to 1, three errors are required to reach these states from $B^{|N|}$. As the waiting times to reach $A^{|N|}$ for these lattices without coalitional behavior is $O(\varepsilon^{-2})$, this immediately implies the existence of a conservative effect. The same cannot be said for the regular network of connected cliques (figure 6). Assume α is close to 1. Consider $k = 4$. Although it is true that $Cli_2 \in \Lambda_{1,\alpha}$ and $Cli_2 \notin \Lambda_{k=4,\alpha}$, there is no conservative effect. The reason for this is that when $k = 1$, two errors are required to reach Cli_2. When $k = 4$, only a single error in which a player plays A is required, following which the three other members of the clique can better respond by switching to A. It then requires only one more error to move the process to $Cli_3 \in \mathcal{P}_A$. From Cli_3, the process can expand via single error driven steps between states in $\Lambda_{k=4,\alpha}$ until $A^{|N|}$ is reached. Therefore the waiting time for $k = 1$ and $k = 4$ is $\Theta(\varepsilon^{-2})$.

Moreover, the introduction of coalitional behavior can completely alter the paths by which A spreads in the population. Take a square lattice on a torus and a clique of ten players and construct a connected network $\mathbf{g}$ via the following step: select a player i on the torus, remove his exiting edges to the right and below, and replace them with edges to player j and l in the clique. The network is shown in figure 7. Assume α is close to 1. When $k = 1$, the fastest paths from $B^{|N|}$ to $A^{|N|}$ involve action A invading the torus (including player i), before errors by four players other than j in the clique make it possible for j to better respond by playing A (for this we require that he has five neighbors playing A). After j switches to A, all of the other players in the clique have 5 neighbors playing A and can individually better respond by playing A. The adoption of A on the clique is the hardest step, requiring 4 errors, and the waiting time until $A^{|N|}$ is reached is $\Theta(\varepsilon^{-4})$. When $k = 9$, the fastest way for A to spread is when

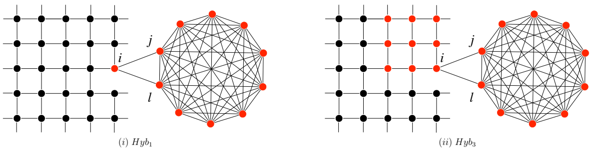


Figure 7: Square lattice on a torus connected to a clique.

it begins in the clique. A single player in the clique who errs and plays A can result in the other players in the clique forming a coalition and better responding by also switching to A. This in turn leads to player i on the torus switching to A. This state, Hyb_1, is in $\mathcal{P}_A$ and it requires at least two errors to leave its basin of attraction. Two such errors can prompt a move to Hyb_3 which is in $\mathcal{P}_A$. From here the process can move between states in $\mathcal{P}_A$ via single errors, culminating at $A^{|N|}$. The waiting time is $\Theta(\varepsilon^{-2})$. So the presence of a single clique on the edge of the network has changed what, in its absence, would have been a conservative effect of coalitional behavior into a reforming effect. It does this by making possible joint deviations by a highly interconnected group of players, amongst whom, in the absence of coalitional behavior, the efficient action A struggles to gain a foothold.

As an interesting aside, we relate the network of figure 7 to the discussion in Granovetter (1973) of 'strong' and 'weak' ties. Recall the assertion in the cited paper that if a player i is linked to j and k by strong ties, then it is likely that j and k will be linked, whereas this is not the case if i is linked to either j or k by a weak tie. An implication of this for the network in figure 7 is that most of the links in the clique are 'strong' ties, and most of the links in the lattice are 'weak' ties. We have seen that, depending on the value of k, contagion by the efficient action can start either in the part of the network rich in weak ties, or the part rich in strong ties. This relates to the debate discussed by Granovetter, citing Becker (1970); Rogers (1962); Coleman et al. (1966), over whether diffusion of innovation starts with 'central' players who have many strong ties, or with 'marginal' players who do not.

Finally, we note an interesting implication of the existence of networks which admit conservative effects and of networks which do not admit conservative effects. Consider two networks of $|N|$ vertices each, the first for which conservative effects are not possible, the second for which they are possible. If we move from the first network to the second network by adding or deleting one link at a time, there must be at least one addition or sub-

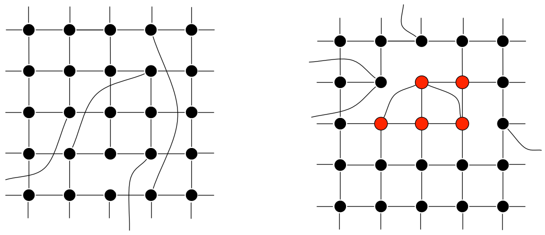


Figure 8: Square lattices with randomly rewired edges.

traction of a link which causes a conservative effect to come into being. That is, there exist 'butterfly effects' – the presence or otherwise of a single link, even in a large network, can lead to the existence or nonexistence of a conservative effect.

7. Small worlds

An important class of networks are small worlds: networks with small average shortest path lengths between players. Social networks often exhibit small world properties, such as the six degrees of separation conjectured in Friyges Karinthy's short story 'Chain Links' and famously examined in Milgram's small world experiments. Other small worlds include neural networks and electric power grids. Two common ways to construct small worlds involve small amounts of rewiring of square lattices or ring networks (see Watts and Strogatz, 1998; Szabo and Fath, 2007).

7.1. Small world, square lattice

Consider small world networks formed by rewiring edges of a square lattice. Specifically, consider square lattices with a small number of edges rewired so that every player still has four neighbors. Such a network will not necessarily exhibit a conservative effect for small α. For some rewirings, a state in $\mathcal{P}_A$ can be reached via only two errors. Such a rewiring is exhibited in figure 8(ii). If, from such a state, there exists an onward path to $A^{|N|}$ between states in $\mathcal{P}_A$ with each step caused by a single error, then a conservative effect will not exist for any α. Networks like this one, in which several rewired edges affect a small neighborhood, will clearly be a small proportion of the class of regular rewired square lattices when the number of rewired edges is small enough relative to the network size. Moreover, if

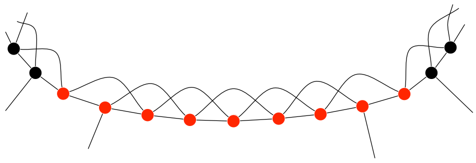


Figure 9: Rewired ring in state Rin_2.

a proportion p of all the edges in the network are rewired, as the population becomes large, the proportion of networks for which a state in $\mathcal{P}_A$ can be reached by two errors or fewer goes to zero. For a fixed neighborhood size, the proportion of networks which include a neighborhood with two or more rewirings that remain locally linked within said neighborhood decreases with $|N|^2$, whereas the number of such neighborhoods in the graph increases with $|N|$.

7.2. Small world, rings

Ring networks do not display limiting conservative effects. Small amounts of relinking will not usually alter this property. Consider the graph of figure 9. A small amount of relinking compared to the network size means that new links are likely to join nodes separated by a significant distance on the original ring. The state Rin_2 illustrated in figure 9 can be reached from $B^{|N|}$ after two errors. Note that although Rin_2 is an intermediate state which did not exist before relinking, $Rin_2 \in \mathcal{P}_A$ so coalitional behavior does not return the process to $B^{|N|}$. At least one error is required to leave Rin_2, but only a single error suffices to continue the spread of A around the ring.

So, the existence or otherwise of a conservative effect of coalitional behavior on square lattices and rings is extended to small worlds constructed from said networks. Simulation results demonstrate that behavior for even large (10 percent) amounts of link rewiring is similar to that obtained for the original networks. How the small world is constructed has important ramifications for predictions regarding the effect of coalitional behavior on convergence times. It is interesting to note that in ring derived graphs, there is a large amount of clustering: the neighbor of a neighbor is likely to be a neighbor. This differs from a square lattice derived graph, for which there is almost no clustering. Yet, in the sense of there being a conservative effect, the square lattice derived graph is more sensitive to coalitional behavior.

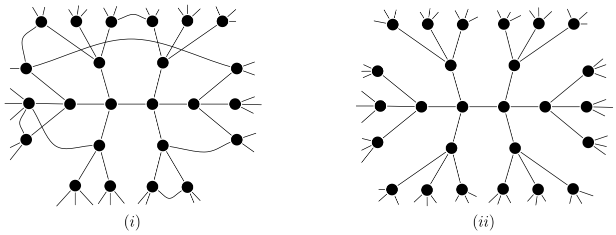


Figure 10: (i) Random regular graph. (ii) Bethe lattice.

8. Random regular graph

As $p \to 1$ the graph in section 7.1 approximates a random regular graph. Note that for all connected regular graphs of degree m, when $\alpha \geq m-1$, a single error will be enough to lead to convergence to $A^{|N|}$. High values of α always lead to fast convergence. The interesting case is when $\alpha < m-1$. Let $Q_{m,|N|}$ be the set of connected m-regular graphs of size $|N|$. Let $\mathcal{Q}_{m,|N|}$ be the uniform distribution over $Q_{m,|N|}$. For any $a \in \mathbb{N}_+$, $k \geq 1$, define:

$$\mathcal{R}^{a,k}_{m,|N|} = \left\{ q \in Q_{m,|N|} : W_{k,\alpha,\varepsilon}(B^{|N|}, A^{|N|}) \in O(\varepsilon^{-a}). \right\}.$$

$\mathcal{R}^{a,k}_{m,|N|}$ is the set of connected m-regular networks which give convergence times which are shorter or of the same order as ε^{-a}. For large $|N|$, random regular graphs are locally like the Bethe lattice. Any small set of players playing action A will then have players on its edge who have only one neighbor playing A. A player in this situation will benefit from switching to B and earning a payoff of $m-1$ rather than a payoff of α. Thus, the set is rolled back from its edges. The order of magnitude of convergence times grows with population size whether or not coalitional behavior is allowed.[29]

Proposition 8. *For all a, k:*

$$\alpha < m-1 \quad \Longrightarrow \quad \lim_{|N| \to \infty} \mathcal{Q}_{m,|N|}(\mathcal{R}^{a,k}_{m,|N|}) \to 0$$

This implies that for large random graphs, convergence will be very slow, with or without coalitional behavior. This does not preclude coalitional

[29]The condition $\alpha < m-1$ imposes the condition that any player with only a single neighbor playing A would like to play B. This plays a role in the proof by ensuring that any state in $\Lambda_{k,\alpha}$ other than $B^{|N|}$ features cycles in the graph on which every player plays A.

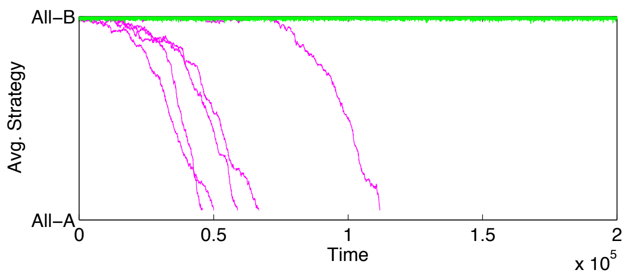


Figure 11: Example runs on square lattice, $\alpha = 1.1$, for $k = 1$ (pink) and $k = 8$ (green).

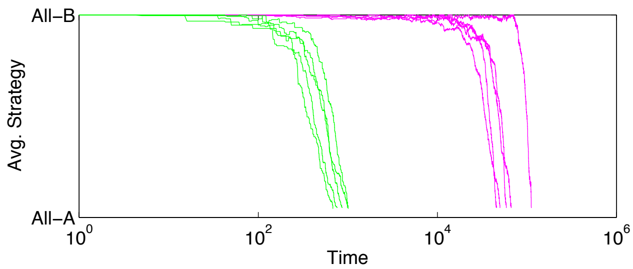


Figure 12: Example runs on square lattice, $\alpha = 2.0$, for $k = 1$ (pink) and $k = 8$ (green).

behavior from slowing or speeding convergence, and we refer you to the simulation results of the following section for random graphs of size $|N| = 256$. What it does indicate is that for any $k \ll |N|$, convergence on random graphs will become qualitatively slow as the population size becomes large, regardless of the presence or otherwise of coalitional behavior.

9. Simulations

Simulations were carried out for a variety of networks to analyze the effect of coalitional behavior on waiting times under small, positive values of ε. The networks fall into two categories. The first category is *generated* networks: squares, rings, associated small worlds and random graphs. The second category, *empirical* networks, includes several examples of real life networks from academic literature.

We apply the model to five empirical networks, visualisations of which are given in Fig. 13.[30] Zachary's karate club describes social relationships

[30]Graph manipulation, measures, visualisation, and sub-graph formation performed with GEPHI: `http://gephi.org/`, an open source graph visualisation and analysis platform.

between 34 club members of a university karate club, observed over the period 1970-1972. In the network, an edge indicates that two members were observed to maintain significant interactions outside of the formal classes and meetings of the club. Zachary interpreted these interactions as indicating friendship outside of club activities. The network is of interest since within the study period, a fission occured in the club's social structure due to a disagreement over tuition fees. As such, the network has been of interest to developers of community identification algorithms for some time (Bagrow and Bollt, 2005; Girvan and Newman, 2002; Newman, 2006b). Alternatively, Newman (2006b,a) compiled a very large scientific collaboration network for the network theory scientific community. Here, we only consider the largest connected component, comprising 379 vertices (authors). Edges indicate that two authors have co-authored a paper. Schwimmer's taro exchange network represents gift-giving, via taro-exchange, amongst 22 households in a Papuan village (Schwimmer, 1973). Kapferer's tailor shop network (Kapferer, 1972) represents observed interactions amongst workers in a Zambian tailor factory during a 10 month period. Kapferer's network is of interest since during the period of observation, the workers collectively negotiated for higher wages. The network we use in this study is the first network (of two) which Kapferer described, after which an unsuccessful strike occured.

Finally, we consider a sub-graph of the FACEBOOK social-network. To obtain a manageable size sub-graph for analysis we consider first the principle, connected, component of the graph '0' from Stanford's SNAP database[31] (McAuley and Leskovec, 2012) (size 324 nodes) before taking three connected communities from this graph, comprising 95 nodes in total as identified by the modularity algorithm of Blondel et al. (2008). The FACEBOOK network was collected by the 'Social circles' FACEBOOK APP (for details, see reference). Taken together, we have a diverse set of real networks amongst human actors for either social, exchange, or collaborative purposes.

The simulations were conducted for $\varepsilon = 0.01$ unless convergence results for comparison could not be readily obtained using the computational resources at hand.[32] In the latter case, a *fast-convergence* setting with $\varepsilon = 0.1$ was used, and is indicated for a row of results in the tables by a superscript '$*$' in the α column. In Tables 1 and 2 we describe the parameters used un-

[31] See: `http://snap.stanford.edu/data/egonets-Facebook.html` .

[32] As expected, this was the case with highly random networks and networks of relatively high average degree.

der each setting. γ is the proportion of B players in the population below which the process is considered to have converged. Simulations were halted when the convergence criterion was met, or 2×10^5 (2×10^6 for FACEBOOK) periods had elapsed.

Table 1: The *default* simulation parameter settings used in the study.

Parameter	**Symbol**	**Value**
Tremble probability	ε	0.01
Convergence limit	γ	5%
Replicates	R	20

Table 2: The *fast-convergence* parameter settings used in the study.

Parameter	**Symbol**	**Value**
Tremble probability	ε	0.10
Convergence limit	γ	20%
Replicates	R	5

In the simulation results tables in the paper, waiting times are given as averages of all replicates for a given experimental condition. Where a result is given as $> x$ this indicates that some fraction of the replicates converged within the maximum waiting time set for the experiment. Those replicates which did not converge were assigned the maximal waiting time prior to averaging across all replicates, and the average, in this case, is treated as a lower bound.

As expected, reforming effects of coalitional behavior were found for all networks for large enough α (Tables 3, 4, 5).[33] Conservative effects for

[33]Simulations, by their nature, do not deal with limiting results. This introduces a measure of imprecision into the reading of the results. Aside from the effects of small ε considered in the theory of the paper, there can also be other effects of coalitional behavior. Higher k could, for example, lead to faster convergence due to multiple switches per period. Conversely, if a single player needs to move to continue on a path to convergence, higher k can reduce the probability of that player being selected in any given period. How any normalization of results for such effects should be carried out is unclear, therefore raw results are presented. Only results displaying order of magnitude differences can be understood to be due to different error requirements on paths to convergence.

small α were also observed for several networks. Consider the small world networks of table 3. A large conservative effect is seen for rewired square lattices, whereas no effect is observed for rewired rings. Random graphs, including scale-free networks, also exhibit strong conservative effects for $\varepsilon = 0.1$, $\alpha = 1.1$ treatments.[34] Results for the empirical networks indicate a conservative effect at small α for all networks other than the 'Network theory coauthorship' graph and 'Schwimmer taro exchange'.

Ratios of convergence times for different α vary markedly in the presence and absence of coalitional behavior. For a square lattice with von-Neumann neighborhood, the ratio of convergence times for $\alpha = 1.1$ and $\alpha = 2.0$ for $k = 1$ is 1, whereas for $k = 8$ the ratio is over 200.[35] Similarly marked differences in the ratios of convergence times ($\alpha = 1.1$ versus $\alpha = 2.0, 3.0$) for the process without and with coalitional behavior are seen for *every* network we test, with the exception of the 'Schwimmer taro exchange'.[36] To emphasize: coalitional behavior makes convergence speeds more sensitive to the relative efficiency of competing norms.

10. Pareto efficiency versus risk dominance

Consider the model adapted so that rather than play a pure coordination game, the players instead play the 'stag hunt' game in figure 14 with each of their neighbors. Note that for this game, there is never a conservative effect of coalitional behavior, as in every situation in which a player would wish to return to playing B as part of a coalition, he would also wish to return to playing B when acting as an individual. As long as $\alpha < 2$,

[34]For $\varepsilon = 0.01$, convergence in these networks was too slow to make meaningful comparisons.

[35]In line with the theory, the effects are predominantly felt early on, when almost every player is playing B. As a comparison, additional simulations were carried out for the square lattice with von-Neumann neighborhood, $\alpha = 1.1$, with the initial state randomized so that each player's action was A or B with probability $1/2$ each. Mean convergence times were then 2245 for $k = 1$ and 1018 for $k = 8$. That is, there was a similar order of magnitude of the convergence times for $k = 1, 8$, and certainly no conservative effect.

[36]This is a small network and we have checked directly that the theory of the paper predicts reforming effects of coalitional behavior for all values of α. Moreover, the theory predicts the same order of magnitude reduction in convergence speeds for $\alpha = 1.1, 2.0, 3.0$, which explains the anomalous result pertaining to this network. The theory predicts a greater reduction in convergence speeds for $\alpha = 1.6$, which can be observed by comparing the ratios of speeds for $k = 1$ and $k = 4$ for fixed values of α.

Table 3: Average wait-times for convergence to the efficient strategy profile for generated networks over selected values of α, with ($k > 1$) and without ($k = 1$) coalitional updating. (*) indicates that *fast-convergence* parameters were used. Note that only a single replicate did not converge for the Regular Square lattice, $k = 1$. See Appendix for details of networks and simulations.

α	**Square lattice**			**Ring**		
	$k = 1$	$k = 4$	$k = 8$	$k = 1$	$k = 4$	$k = 8$
Regular						
1.1	>78,628	>200,000	>200,000	41,500	71,005	>100,179
1.6	>78,628	20,265	43,937	41,500	7,176	9,153
2.0	>78,628	1,467	902	41,500	1,045	499
3.0	3,622	1,134	758	7,006	964	466
Small-world						
1.1*	7,919	61,449	>200,000	8,971	8,368	12,248
1.6	>105,310	24,731	>76,110	59,020	7,834	6,885
2.0	38,527	1,989	1,246	23,632	1,219	674
3.0	3,830	956	552	5,292	755	389
Random						
1.1*	13,358	>200,000	>200,000	13,027	>200,000	>200,000
1.6*	5,260	5,487	>99,806	5,755	4,578	52,596
2.0	19,094	4,628	9,308	19,214	5,033	11,919
3.0	4,992	719	438	4,596	694	425

Table 4: Average wait-times for convergence to the efficient strategy profile for a scale-free network over selected values of α, with ($k > 1$) and without ($k = 1$) coalitional updating. (*) indicates that *fast-convergence* parameters were used. See Appendix for details of networks and simulations.

α	**Scale-Free**		
	$k = 1$	$k = 4$	$k = 8$
1.1*	8,458	>200,000	>200,000
1.6*	5,068	11,465	>200,000
2.0*	3,669	1,100	1,095
3.0*	2,070	456	253

Table 5: Average wait-times for convergence to the efficient strategy profile for real networks over selected values of α, with $(k = 4)$ and without $(k = 1)$ coalitional updating. (*) indicates that *fast-convergence* parameters were used. Size of each real network given in parentheses. The column 'Conv.' gives the convergence rate where not 100%. See Appendix for details of networks and simulations.

α	$k = 1$	Conv.	$k = 4$	Conv.
Zachary's Karate Club ($\lvert N\rvert = 34$)				
1.1	55,758		>182,833	10%
1.6	14,846		15,008	
2.0	4,740		469	
3.0	1,683		152	
Network Theory Co-authorship ($\lvert N\rvert = 379$)				
1.1*	42,366		37,715	
1.6*	16,947		6,037	
2.0*	7,954		1,885	
3.0*	4,243		802	
Schwimmer Taro Exchange ($\lvert N\rvert = 22$)				
1.1	>182425	20%	29,793	
1.6	28,253		438	
2.0	928		160	
3.0	518		34	
Kapferer's Tailor Shop ($\lvert N\rvert = 39$)				
1.1*	>87,908	80%	>200,000	0%
1.6*	2,634		28,685	
2.0*	975		1,791	
3.0*	593		227	
Facebook sub-graph ($\lvert N\rvert = 95$)				
1.1*	444,845		>1,494,273	60%
1.6*	6,699		24,165	
2.0*	2,973		1,646	
3.0*	1,258		365	

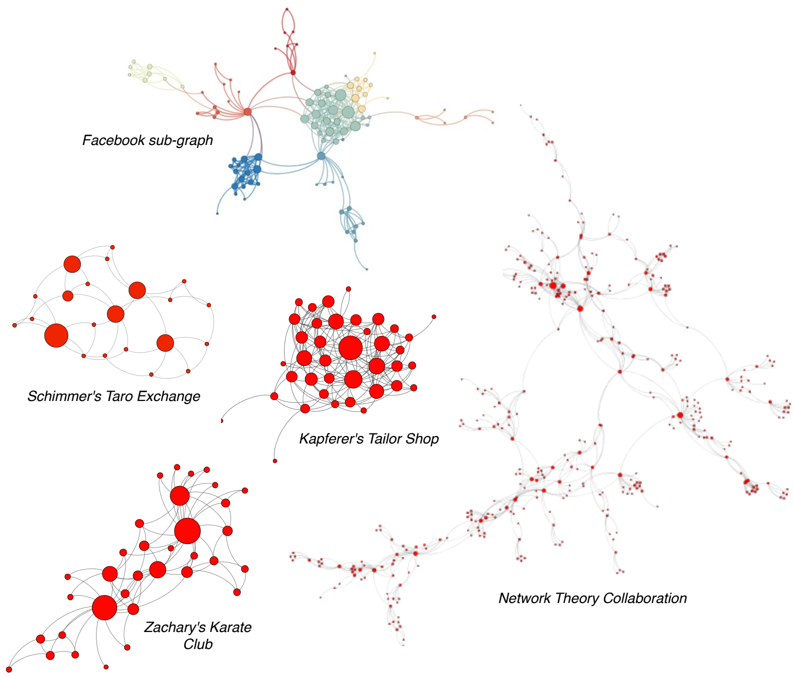


Figure 13: Visualisations of the five real networks considered in this study: the FACEBOOK sub-graph McAuley and Leskovec (2012); Schwimmer's taro exchange network Schwimmer (1973); Kapferer's tailor shop Kapferer (1972); the co-author collaboration network in network theory science Newman (2006a); and Zachary's Karate Club network Zachary (1977).

	A	B
A	α, α	$0, 1$
B	$1, 0$	$1, 1$

Figure 14: A two player coordination game, $1 < \alpha < 2$.

(B, B) is the risk dominant equilibrium of this game, whereas (A, A) is the Pareto efficient equilibrium. Individual adaptive dynamics select the equilibrium in which every player plays the risk dominant action B as a stochastically stable state.[37] This is not the case for the coalitional dynamics of this paper. Instead, it can be the case that all stochastically stable states are polymorphic states in which different players choose different actions depending on their location in the network.

We consider a class of networks Γ_{Cli} composed of interconnected cliques such that every member of a clique is connected to every other member of the clique, and one player outside of the clique. An example in which every clique comprises four players is given in figure 6. Here we allow cliques to be of any size.

Proposition 9. *Consider a network $\boldsymbol{g} \in \Gamma_{Cli}$. For a given clique S, $|S| > k$, in any stochastically stable state x:*

$$|S| \geq \tfrac{(k+1)\alpha}{2-\alpha} \qquad \Longrightarrow \qquad x_S = B^{|S|},$$

$$|S| \leq \tfrac{(k-3)\alpha}{2-\alpha} \qquad \Longrightarrow \qquad x_S = A^{|S|}.$$

So in networks composed of lightly interconnected cliques, stochastically stable states involve medium-size cliques playing A and large cliques playing B. In long run equilibria, efficiency is attained within medium-size cliques, but not within larger ones.[38] This means that on the margin, an additional

[37]Other stochastically stable states may also exist. Consider a two player connected network. A single error suffices to move from $B^{|N|}$ to $A^{|N|}$ or from $A^{|N|}$ to $B^{|N|}$. Both monomorphic states are stochastically stable. See Blume (1996).

[38]Further information about the neighbors of the clique allows these bounds to be tightened. If there is no B player in N_S then the $(k+1)$,$(k-3)$ factors can be replaced by $(k+1)$,$(k-1)$. If the number of A players in N_S is less than k then the relevant factors are $(k-1)$,$(k-3)$. If neither of these hold then the factors are (k),$(k-2)$. Furthermore, if we replace 'in any stochastically stable state' with 'in some stochastically stable state' then the upper and lower bounds on $|S|$ can be tightened by 1.

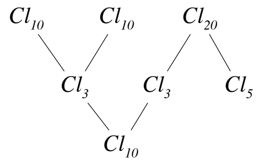


Figure 15: Fragment of network of interconnected cliques. Cl_m indicates a clique of size m.

player being added to a clique may decrease total payoffs. Despite the positive payoff externalities enjoyed by clique members on the addition of an extra player, the long run stable outcome may switch to one in which the clique in question coordinates on the inefficient action. Now consider small cliques such that the whole clique can form a coalition. For a stochastically stable state x, if no player in the clique has a neighbor outside the clique who plays B, then members of the clique all play A in x. If at least one member of the clique has a neighbor who plays B, then all members of the clique will play A if every member does better from this than he would do if he played B. Otherwise, all members of the clique play B.

Proposition 10. *Consider a network* $\boldsymbol{g} \in \Gamma_{Cli}$. *For a given clique* S, $|S| \leq k$, *in any stochastically stable state* x:

$$x_{N_S} = A^{|N_S|} \quad \Longrightarrow \quad x_S = A^{|S|},$$

$$x_{N_S} \neq A^{|N_S|},\ |S| > \tfrac{\alpha}{\alpha-1} \quad \Longrightarrow \quad x_S = A^{|S|},$$

$$x_{N_S} \neq A^{|N_S|},\ |S| < \tfrac{\alpha}{\alpha-1} \quad \Longrightarrow \quad x_S = B^{|S|}.$$

To see how this works in an example, consider the partial network illustrated in figure 15. Cl_m indicates a clique of size m, and the lines between cliques indicate a link between one of the players in each of the cliques. $k = 8$ and $\alpha = \frac{4}{3}$. Then proposition 9 implies that members of a clique of size $m \geq 18$ play B in any stochastically stable state. This applies to the members of the clique of size twenty in figure 15. Similarly, members of a clique of size $8 < m \leq 10$ will play A in any stochastically stable state. This applies to the three Cl_{10} in figure 15. The clique Cl_3 of size three to the left of the figure will then have no neighbors who play B, and thus by proposition 10 will play A in any stochastically stable state. The cliques Cl_3 and Cl_5 of size three and five to the right of the figure have at least one neighbor who plays B, so their behavior depends on whether their size is greater or lower than the value $\alpha/(\alpha - 1) = 4$. So by proposition 10, members of Cl_3 on the right hand side play B and members of Cl_5 play A in any stochastically stable state.

So coalitional behavior can lead to heterogeneous choices by cliques within a population depending on their size. This is not necessarily monotonic, with large cliques playing the risk dominant action, medium-size cliques the Pareto efficient action. In the absence of neighbors, small cliques would also play the Pareto efficient action, but the presence of neighbors playing the risk dominant action pushes them to do likewise. An interpretation of this from the perspective of organizational design could be that teams should not be so large that the internal pressure against risky yet efficient behavior dominates, but neither should they be so small that external pressure has the same effect.

11. Conclusion

In settings in which every member of a population has a common interest in coordinating on a given efficient action, it might be expected that overt cooperation by coalitions of players in their choice of action would facilitate the spread of that action in the population. This paper has shown that this is not always the case. Aside from the existence of such reforming effects of coalitional behavior, there can also exist conservative effects by which coalitions slow the spread of efficient behavior in a population. As the relative efficiency of the efficient action increases, the effect of coalitional behavior on a dynamic can switch abruptly from a conservative effect to a reforming effect. Not all networks exhibit conservative effects and their existence or otherwise depends on the structure of the underlying network.

Simulation results appear consistent with the theory of the paper and indicate that we should expect conservative and reforming effects to manifest themselves in dynamics of social and technological change on a variety of network types, including several empirical social networks analyzed in the literature. The introduction of coalitional behavior is seen to greatly increase the sensitivity of convergence speeds to the relative efficiency of competing technologies. Finally, for games in which the Pareto efficient and risk dominant equilibria differ, we see that play of both actions can coexist in the population in the long run and that the network design determines which actions survive in which parts of the network.

This paper shows the effects of coalitional behavior on networked coordination games to be important and non-obvious. The assumptions of the paper are not strong, the departure from existing literature being that small groups of connected players coordinate their action choice some of the time. Implications for the study of social dynamics and network design clearly exist and merit subsequent study.

References

Acemoglu, D., Johnson, S., Robinson, J.A., 2002. Reversal of fortune: geography and institutions in the making of the modern world income distribution. Quarterly Journal of Economics 117, 1231 – 94.

Akerlof, G., 1976. The economics of caste and of the rat race and other woeful tales. Quarterly Journal of Economics 90, 599 – 617.

Algaba, E., Bilbao, J., Van Den Brink, R., 2004. Cooperative games on antimatroids. Discrete Mathematics 282, 1–15.

Algaba, E., Bilbao, J.M., Borm, P., Lpez, J.J., 2001. The myerson value for union stable structures. Mathematical Methods of Operations Research 54, 359–371.

Alós-Ferrer, C., Netzer, N., 2010. The logit-response dynamics. Games and Economic Behavior 68, 413–427.

Ambrus, A., 2009. Theories of coalitional rationality. Journal of Economic Theory 144, 676 – 695.

Anheier, H., Toepler, S., 2009. International Encyclopedia of Civil Society. Number v. 3 in International encyclopedia of civil society, Springer.

Aumann, R., 1959. Acceptable points in general cooperative n-person games, in contributions to the theory of games iv, (a. w. tucker and r. d. luce, eds.). Princeton University Press , 287–324.

Bagrow, J.P., Bollt, E.M., 2005. Local method for detecting communities. Physical Review E 72, 046108+.

Baland, J.M., Guirkinger, C., Mali, C., 2011. Pretending to be poor: Borrowing to escape forced solidarity in cameroon. Economic Development and Cultural Change 60, 1–16.

Becker, M.H., 1970. Sociometric location and innovativeness: Reformulation and extension of the diffusion model. American Sociological Review 35, 267–282.

Bergin, J., Lipman, B.L., 1996. Evolution with state-dependent mutations. Econometrica 64, 943–56.

Bernheim, B.D., Peleg, B., Whinston, M.D., 1987. Coalition-proof nash equilibria i. concepts. Journal of Economic Theory 42, 1–12.

Bikhchandani, S., Hirshleifer, D., Welch, I., 1992. A theory of fads, fashion, custom, and cultural change as informational cascades. Journal of Political Economy 100, pp. 992–1026.

Binmore, K., Samuelson, L., Young, P., 2003. Equilibrium selection in bargaining models. Games and Economic Behavior 45, 296 – 328.

Blondel, V.D., loup Guillaume, J., Lambiotte, R., Lefebvre, E., 2008. Fast unfolding of communities in large networks.

Blume, L.E., 1993. The statistical mechanics of strategic interaction. Games and Economic Behavior 5, 387 – 424.

Blume, L.E., 1996. Population Games. Working Papers 96-04-022. Santa Fe Institute.

Bowles, S., 2004. Microeconomics: Behavior, institutions, and evolution, isbn 978-0-691-09163-1 608 pp. Princeton University Press .

Chamberlin, J., 1974. Provision of collective goods as a function of group size. The American Political Science Review 68, pp. 707–716.

Coleman, J.S., Katz, E., Menzel, H., 1966. Medical innovation: A diffusion study. Bobbs-Merrill, New York.

Currarini, S., Jackson, M.O., Pin, P., 2009. An economic model of friendship: Homophily, minorities,andsegregation. Econometrica 77, 1003–1045.

van Damme, E., Weibull, J.W., 2002. Evolution in games with endogenous mistake probabilities. Journal of Economic Theory 106, 296–315.

Easley, D., Kleinberg, J., 2010. Networks, Crowds, and Markets: Reasoning About a Highly Connected World. Cambridge University Press, New York, NY, USA.

Ellison, G., 1993. Learning, local interaction, and coordination. Econometrica 61, 1047–71.

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

Engerman, S.L., Sokoloff, K.L., 2006. The persistence of poverty in the Americas. Chapter 2, Poverty Traps (ed. Samuel Bowles, Steven N. Durlauf, Karla Hoff). Princeton University Press.

Eshel, I., Samuelson, L., Shaked, A., 1998. Altruists, egoists, and hooligans in a local interaction model. American Economic Review 88, 157–79.

Foster, D., Young, H.P., 1990. Stochastic evolutionary game dynamics. Theoretical Population Biology 38, 219–232.

Freidlin, M.I., Wentzell, A.D., 1984. Random perturbations of dynamical systems, isbn 9780387983622 430 pp., 2nd ed.(1998). Springer .

Girvan, M., Newman, M.E., 2002. Community structure in social and biological networks. Proceedings of the National Academy of Sciences of the United States of America 99, 7821–7826.

Golub, B., Jackson, M.O., 2012. How homophily affects the speed of learning and best-response dynamics. The Quarterly Journal of Economics 127, 1287–1338.

Granovetter, M., 1973. The Strength of Weak Ties. The American Journal of Sociology 78, 1360–1380.

Granovetter, M., 1978. Threshold models of collective behavior. American Journal of Sociology 83, pp. 1420–1443.

Hoff, K., Sen, A., 2006. The kin system as a poverty trap. Chapter 4, Poverty Traps (ed. Samuel Bowles, Steven N. Durlauf, Karla Hoff). Princeton University Press.

Hwang, S.H., 2009. Larger groups may alleviate collective action problems. Santa Fe Institute, Working Paper .

Jackson, M.O., 2005. Allocation rules for network games. Games and Economic Behavior 51, 128–154.

Jackson, M.O., Watts, A., 2002. The evolution of social and economic networks. Journal of Economic Theory 106, 265–295.

Jackson, M.O., Wolinsky, A., 1996. A strategic model of social and economic networks. Journal of Economic Theory 71, 44–74.

Jakiela, P., Ozier, O., 2012. does africa need a rotten kin theorem? experimental evidence from village economiesn. Working Paper .

Kandori, M., Mailath, G.J., Rob, R., 1993. Learning, mutation, and long run equilibria in games. Econometrica 61, 29–56.

Kapferer, B., 1972. Strategy and transaction in an African factory. Univ. Pr., Manchester.

Kets, W., Iyengar, G., Sethi, R., Bowles, S., 2011. Inequality and network structure. Games and Economic Behavior 73, 215–226.

Klaus, B., Klijn, F., Walzl, M., 2010. Stochastic stability for roommate markets. Journal of Economic Theory 145, 2218 – 2240.

Konishi, H., Ray, D., 2003. Coalition formation as a dynamic process. Journal of Economic Theory 110, 1–41.

Kreindler, G.E., Young, H.P., 2013. Fast convergence in evolutionary equilibrium selection. Games and Economic Behavior 80, 39 – 67.

Mathew, S., Boyd, R., 2011. Punishment sustains large-scale cooperation in prestate warfare. Proceedings of the National Academy of Sciences 108, 11375–11380.

McAuley, J., Leskovec, J., 2012. Learning to discover social circles in ego networks, in: Bartlett, P., Pereira, F., Burges, C., Bottou, L., Weinberger, K. (Eds.), Advances in Neural Information Processing Systems 25, pp. 548–556.

McPherson, M., Smith-Lovin, L., Cook, J.M., 2001. Birds of a feather: Homophily in social networks. Annual Review of Sociology 27, 415–444.

Montanari, A., Saberi, A., 2010. The spread of innovations in social networks. Proceedings of the National Academy of Sciences 107, 20196–20201.

Morris, S., 2000. Contagion. Review of Economic Studies 67, 57–78.

Morris, S., Rob, R., Shin, H.S., 1995. p-dominance and belief potential. Econometrica 63, 145–157.

Myerson, R.B., 1977. Graphs and cooperation in games. Mathematics of Operations Research 2, pp. 225–229.

Naidu, S., Hwang, S.H., Bowles, S., 2010. Evolutionary bargaining with intentional idiosyncratic play. Economics Letters 109, 31 – 33.

Newman, M.E., 2006a. Modularity and community structure in networks. Proc Natl Acad Sci U S A 103, 8577–8582.

Newman, M.E.J., 2006b. Finding community structure in networks using the eigenvectors of matrices. Physical review E 74. Cite arxiv:physics/0605087Comment: 22 pages, 8 figures, minor corrections in this version.

Newton, J., 2012a. Coalitional stochastic stability. Games and Economic Behavior 75, 842–54.

Newton, J., 2012b. Recontracting and stochastic stability in cooperative games. Journal of Economic Theory 147, 364–81.

Nunn, N., Wantchekon, L., 2011. The slave trade and the origins of mistrust in africa. American Economic Review 101, 3221–3252.

Olson, M., 1965. The Logic of Collective Action : Public Goods and the Theory of Groups. Harvard economic studies, v. 124, Harvard University Press. revised 1971 edition.

Peleg, B., Sudholter, P., 2003. Introduction to the theory of cooperative games, isbn 1-4020-7410-7 378 pages. Kluwer Academic, Boston .

Poteete, A.R., Ostrom, E., 2004. Heterogeneity, group size and collective action: The role of institutions in forest management. Development and Change 35, 435–461.

Rogers, E., 1962. Diffusion of Innovations. Free Press, New York. Fifth edition dated

2003.
Sawa, R., 2012. Coalitional stochastic stability in games, networks and markets. Unpublished Manuscript .
Schelling, T.C., 1969. Models of segregation. American Economic Review 59, 488–93.
Schelling, T.C., 1971. Dynamic models of segregation. J. Math. Sociol. 1, 143–186.
Schwimmer, E., 1973. Exchange in the social structure of the Orokaiva; traditional and emergent ideologies in the northern district of Papua, by Erik Schwimmer. C. Hurst, London,.
Serrano, R., Volij, O., 2008. Mistakes in cooperation: the stochastic stability of edgeworth's recontracting. Economic Journal 118, 1719–1741.
Smith, J.M., Price, G.R., 1973. The logic of animal conflict. Nature , 15–18.
Szabo, G., Fath, G., 2007. Evolutionary games on graphs. Physics Reports 446, 97–216.
Vinkovic, D., Kirman, A., 2006. A physical analogue of the schelling model. Proc Natl Acad Sci U S A 103, 19261–5.
Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of 'small-world' networks. Nature 393, 440–442.
Wormald, N.C., 1999. Models of random regular graphs. in j.d. lamb and d.a. preece (eds.). Surveys in Combinatorics 1999, LMS Lecture Note Series 267, 239–298.
Young, H.P., 1993. The evolution of conventions. Econometrica 61, 57–84.
Young, H.P., 1998a. Conventional contracts. Review of Economic Studies 65, 773–92.
Young, H.P., 1998b. Individual strategy and social structure. Princeton University Press.
Young, H.P., 2011. The dynamics of social innovation. Proceedings of the National Academy of Sciences 108 Suppl 4, 21285–91.
Zachary, W.W., 1977. An Information Flow Model for Conflict and Fission in Small Groups. Journal of Anthropological Research 33, 452–473.

Appendix A. Proofs

We use the concepts of and similar notation to Ellison (2000).[39] For $x, y \in X$, define the resistance $r(x, y)$ so that the most probable transition from x to y occurs with probability of order $\varepsilon^{r(x,y)}$.

$$r(x,y) = \min\left\{ r \in \mathbb{R}_+ : \exists\, t \in \mathbb{N}_+ : \lim_{\varepsilon \to 0} \frac{P^t_{k,\alpha,\varepsilon}(x,y)}{\varepsilon^r} > 0 \right\}$$

and for $x_1, \ldots, x_T \in X$, sets $W, Y \subseteq X$:

$$r(x^1, \ldots, x^T) = \sum_{t=1}^{t=T-1} r(x^t, x^{t+1}); \qquad r(W,Y) = \min_{\substack{(x^1,\ldots,x^T) \\ x^1 \in W \\ x^T \in Y}} r(x^1, \ldots x^T).$$

[39]Ellison (2000) cites a no longer extant working paper of Evans as containing the first statements and use of some of these concepts.

Extend the notion of basin of attraction to sets:

$$D_{k,\alpha}(W) = \left\{ y \in X : \; P^t_{k,\alpha,0}(y, W) \to 1 \;\; \text{as } t \to \infty \right\}$$

then the *radius* of a set W is:

$$R(W) = r(W, X \setminus D_{k,\alpha}(W)).$$

The radius is the resistance of the lowest resistance path from W to outside the basin of attraction of W. Define modified resistance:[40]

$$r^*(x^1, \ldots, x^T) = r(x^1, \ldots, x^T) - \sum_{t=2}^{t=T-1} R(x^t)$$

and

$$r^*(x, W) = \min_{\substack{(x^1,\ldots,x^T) \\ x=x^1 \\ x^T \in W}} r^*(x^1, \ldots x^T).$$

Finally, define the *modified coradius*:

$$CR^*(W) = \max_{x \notin W} r^*(x, W)$$

For purposes of comparison, these quantities will sometimes be given subscripts k, α.

Proof of proposition 1. For $k = 1$, $\alpha \geq 3$, a single error (and no fewer) is enough to move the process to $A^{|N|}$. For $\alpha < 3$, two errors move the process to C_2. No fewer than two errors suffice to move the process out of $D_{k,\alpha}(B^{|N|})$. From C_2, a single error then suffices to move the process to C_3, and so on. Noting that from any state not equal to $B^{|N|}$, a single error is

[40] The reason that modified resistances and associated quantities are useful is that when using spanning tree arguments as found in Freidlin and Wentzell (1984), when edges are added to an existing tree, other edges must be deleted for the graph to remain a tree. Theorems using radii and coradii to give sufficient conditions for stochastic stability follow trivially from this observation.

enough to give an expanding set of squares, and that $R_{1,\alpha}(C_i) = 1$, we have:

$$\begin{aligned} \alpha < 3: &\qquad CR^*_{1,\alpha}(A^{|N|}) = 2 \\ \alpha \geq 3: &\qquad CR^*_{1,\alpha}(A^{|N|}) = 1 \end{aligned}$$

For $4 \leq k \ll n_1, n_2$, for $\alpha < 3/2$, two errors do not suffice to leave $D_{k,\alpha}(B^{|N|})$. Three errors, however, can take the process to C_3, from where a single error can take the process to C_4 and so on. Note that $R_{k,\alpha}(C_i) = 1$ for $i \geq 3$. For $3/2 \leq \alpha < 2$, two errors are required to move to $C_2 \in \bar{D}_{k,\alpha}(A^{|N|})$. For $\alpha \geq 2$, $B^{|N|} \in D_{k,\alpha}(A^{|N|})$. We have:

$$\begin{aligned} \alpha < 3/2: &\qquad CR^*_{k,\alpha}(A^{|N|}) = 3 \\ 3/2 \leq \alpha < 2: &\qquad CR^*_{k,\alpha}(A^{|N|}) = 2 \\ \alpha \geq 2: &\qquad CR^*_{k,\alpha}(A^{|N|}) = 0 \end{aligned}$$

Results follow from $W_{k,\alpha,\varepsilon}(B^{|N|}, A^{|N|}) \geq W_{k,\alpha,\varepsilon}(x, A^{|N|})$ for all $x \in X$; and Theorem 2 of Ellison (2000). □

Proof of proposition 4. The set of feasible coalitions is independent of k, so $k_1 \leq k_2$ implies:

$$\mathcal{N}(k_1) \subseteq \mathcal{N}(k_2) \qquad \Longrightarrow \qquad \text{supp}(F_{k_1}) \subseteq \text{supp}(F_{k_2})$$

so for all $x, y \in X$:

$$P_{k_1,\alpha,0}(x,y) > 0 \qquad \Longrightarrow \qquad P_{k_2,\alpha,0}(x,y) > 0$$

and

$$P_{k_1,\alpha,0}(x, X \setminus \{x\}) > 0 \qquad \Longrightarrow \qquad P_{k_2,\alpha,0}(x, X \setminus \{x\}) > 0$$

which implies

$$x \notin \Lambda_{k_1,\alpha} \qquad \Longrightarrow \qquad x \notin \Lambda_{k_2,\alpha}$$

which implies $\Lambda_{k_1,\alpha} \supseteq \Lambda_{k_2,\alpha}$. Furthermore, for $x \in \Lambda_{k_2,\alpha}$:

$$y \in \bar{D}_{k_1,\alpha}(x) \qquad \Longrightarrow \qquad P^t_{k_1,\alpha,0}(y,x) > 0 \text{ for some } t \in \mathbb{N}_+$$

$$\Longrightarrow \qquad P^t_{k_2,\alpha,0}(y,x) > 0 \qquad \Longrightarrow \qquad y \in \bar{D}_{k_2,\alpha}(x).$$

□

Proof of proposition 5. For given $x \in X$, assume there exists $S \subseteq \{i : x_i = B\}$ such that:

$$\forall\, T \subseteq S,\ |T| \leq k :\ \exists\, i \in T : \quad \frac{|N_i \setminus S| + |N_i \cap T|}{|N_i| + |N_i \cap T|} < \frac{1}{1+\alpha}.$$

Take such an S. Then for all $T \subseteq S$, $|T| \leq k$, there exists $i \in T$ such that:

$$\alpha(|N_i \setminus S| + |N_i \cap T|) < |N_i \cap S| \leq u_i(x).$$

The left hand side is the maximum payoff attainable by player i from playing A if all players in $S \setminus T$ play B. So no subset of players in S of size $\leq k$ will switch to A unless some subset of players in S have already switched to A. Therefore $x \notin \bar{D}_{k,\alpha}(A^{|N|})$. This proves the 'only if' part of the proposition.

To prove the 'if' part of the proposition assume that $x \notin \bar{D}_{k,\alpha}(A^{|N|})$. Starting from state $x = x^1$, if there is any feasible coalition $U \subseteq N$, $|U| \leq k$, such that for all $i \in U$, $u_i(x_U = A^{|U|}, x^t_{-U}) \geq u_i(x^t)$, then with some probability U better responds and the state moves to $x^{t+1}_S = A^{|U|}$, $x^{t+1}_{-U} = x^t_{-U}$. Iterate until there is no such subset of players, say at time τ. Let $S = \{i : x^\tau_i = B\}$. This set must be nonempty or else $x^\tau = A^{|N|}$, which would contradict $x \notin \bar{D}_{k,\alpha}(A^{|N|})$. Now, for all $T \subseteq S$, $|T| \leq k$, as at least one player, say player i, in T would be strictly worse off if T switched to action A:

$$\alpha(|N_i \setminus S| + |N_i \cap T|) < u_i(x^\tau) = |N_i \cap S| = |N_i| - |N_i \setminus S|.$$

Rearranging gives:

$$\frac{|N_i \setminus S| + |N_i \cap T|}{|N_i| + |N_i \cap T|} < \frac{1}{1+\alpha}.$$

and we have our result. □

Proof of proposition 6. $\Lambda_{1,\alpha} = \{B^{|N|}, A^{|N|}\}$. Note that $\bar{D}_{1,\alpha}(A^{|N|}) \subseteq \bar{D}_{k,\alpha}(A^{|N|})$ for all $k \geq 1$. So:

$$CR^*_{1,\alpha}(A^{|N|}) = R_{1,\alpha}(B^{|N|}) = r_{1,\alpha}\left(B^{|N|}, \bar{D}_{1,\alpha}(A^{|N|})\right)$$

$$\geq r_{1,\alpha}\left(B^{|N|}, \bar{D}_{k>1,\alpha}(A^{|N|})\right) \geq r_{k>1,\alpha}\left(B^{|N|}, \bar{D}_{k>1,\alpha}(A^{|N|})\right)$$

$$= R_{k>1,\alpha}(B^{|N|}) = CR^*_{k>1,\alpha}(A^{|N|})$$

$\square$

Proof of proposition 7. First the 'only if' part of the statement is addressed. Let $x \in \mathcal{P}_A$ and $S \subseteq \{i \in N : \ x_i = A\}$ be parochial. $I_0(S)$ must be nonempty. For any α, $i \in I_0(S)$, $N \supseteq T \supseteq \{i\}$, as

$$u_i(x) = \alpha|N_i| > |N_i| \geq u_i(\tilde{x}_i = B, \hat{x}_{-i}) \quad \text{for any } \hat{x}_{-i},$$

we have:

$$\tilde{x}_T \in A_T(x) \quad \Longrightarrow \quad \tilde{x}_i = A.$$

Then, by induction, for any $i \in I_m(S)$, $m \geq 1$, $N \supseteq T \supseteq \{i\}$, as

$$u_i(x) \geq \alpha(|N_i| - |N_i \setminus S|) > |N_i| - |N_i \setminus S| \geq |N_i| - |N_i \cap I_{m-1}(S)|$$

$$\geq u_i(\tilde{x}_i = B, \tilde{x}_{I_{m-1}(S)} = A^{|I_{m-1}(S)|}, \hat{x}_{-(I_{m-1}(S)\cup\{i\})}) \quad \text{for any } \hat{x}_{-(I_{m-1}(S)\cup\{i\})},$$

we have:

$$\tilde{x}_T \in A_T(x) \quad \Longrightarrow \quad \tilde{x}_i = A.$$

So, for any $x^t \in \mathcal{P}_A$: $x_S^{t+1} = A^{|S|}$ and therefore $x^{t+1} \in \mathcal{P}_A \not\ni B^{|N|}$. So $x \in \mathcal{P}_A$ implies $x \notin \bar{D}_{k,\alpha}(B^{|N|})$.

Now the 'if' part of the statement is addressed. Assume that $x \notin \bar{D}_{k,\alpha}(B^{|N|})$ for any k, α. If there exists feasible $S \subseteq N$ such that:

$$\begin{aligned} (*) \quad & x_S \neq B^{|S|} \quad \text{and} \\ & \exists\, \underline{\alpha} : \ \forall\, i \in S : \quad u_i(B^{|S|}, x_{-S}) \geq u_i(x), \end{aligned}$$

then, assuming $k \geq |S|$, we have $S \in \text{supp}(F_k)$. Assuming $\alpha < \underline{\alpha}$:

$$B^{|S|} \in A_S(x) \quad \text{so} \quad P_{k,\alpha,0}(x, (B^{|S|}, x_{-S})) > 0.$$

Starting from $x^t \notin \bar{D}_{k,\alpha}(B^{|N|})$, let $x^{t+1} = (B^{|S|}, x_{-S})$ for such an S, and iterate until a state, $\tilde{x}$, is reached such that there does not exist feasible S which satisfies $(*)$. Let $T = \{i \in N : \ \tilde{x}_i = A\}$. Then $I_0(T) \neq \emptyset$ or T would satisfy $(*)$ as there would exist α such that:

$$\forall\, i \in T : \quad u_i(B^{|T|}, \tilde{x}_{-T}) = |N_i| \geq \alpha(|N_i| - 1) \geq u_i(\tilde{x}).$$

If $\nexists\, m$ such that $I_m(T) = T$, then choose m such that $I_m(T) = I_{m-1}(T)$. Then:

$$\forall\, i \in T \setminus I_m(T): \quad |N_i \setminus T| > |N_i \cap I_{m-1}(T)| = |N_i \cap I_m(T)|.$$

But then, for all $i \in T \setminus I_m(T)$, there exists α such that:

$$u_i(B^{|T \setminus I_m(T)|}, \tilde{x}_{-(T \setminus I_m(T))}) \geq |N_i \setminus T| + |N_i \cap (T \setminus I_m(T))|$$

$$> \alpha(|N_i \cap I_m(T)| + |N_i \cap (T \setminus I_m(T))|) = \alpha |N_i \cap T| \geq u_i(\tilde{x})$$

so some feasible subset of $T \setminus I_m(T)$ satisfies $(*)$ and we have a contradiction. Therefore $\exists\, m$ such that $I_m(T) = T$. T is a parochial set such that $x_T = A^{|T|}$. Therefore $x \in \mathcal{P}_A$.

$\square$

Appendix B. Further proofs [ONLINE - NOT FOR PRINT]

Proof of proposition 8. Fix a. Define l_1 as the largest integer l such that $\lceil \frac{l}{k+1} \rceil < a$. Let $l_2 = 3l_1$, $l_3 = 3l_2$. Assume $|N|$ is large and that any cycle with vertex set S, $|S| \leq l_2$, is of a distance at least l_3 on the network from any cycle with vertex set T, $T \neq S$, $|T| \leq l_2$. This is true asymptotically almost surely as $|N| \to \infty$ and follows from the fact that any subgraph with more edges than vertices does not, asymptotically almost surely, appear in a random regular graph as $|N| \to \infty$ (Wormald, 1999).

Then, any cycle with vertex set S, $l_1 < |S| \leq l_2$, requires at least a errors to switch to playing $A^{|S|}$, conditional on $x_{N_S} = B^{|N_S|}$. If $2\alpha > m - 1$ (or $2\alpha > m - 2$ for $k = 1$), assume the process is at the state in which all cycles with vertex set S, $|S| \leq l_1$ play A, all other players play B. Otherwise assume the process is at $B^{|N|}$.

Note that for all $x \in \Lambda_{k,\alpha}$, $k \geq 1$, $\alpha < m - 1$, all $i \in N$ such that $x_i = A$ have at least 2 neighbors playing A. We show that, starting from x, to reach $y \in \Lambda_{k,\alpha}$ such that strictly more players play A in y than in x, requires at least a errors.

For a path of players playing A to form between two cycles with vertex sets S, T, $|S|, |T| \leq l_2$, we require at least $\lceil \frac{l_3-k-1}{k+1} \rceil > a$ errors.

Assume a cycle with vertex set S, $l_1 < |S| \leq l_2$, switches to play A. For paths of players, starting from some player outside S to switch to A and influence more than a single player in S requires at least a errors by definition of l_2. So we can rule out such outside influence. Therefore, by definition of l_1, at least a errors are required for S to switch to play $A^{|S|}$.

Assume a cycle with vertex set S, $l_2 < |S|$, switches to play A. As $l_1 < |S|$, for there to be a chance of this requiring fewer than a errors, the formation of such a cycle must be assisted by existing small cycles of A players. If there are two such cycles, a path of length at least l_3 exists between them, which similarly to above, would require more than a errors to cross. So assume a single such influencing small cycle exists. At most, such a cycle can cause l_1 consecutive players on S to play A. Thus, for S to switch to $A^{|S|}$ we require $\lceil \frac{|S|-l_1-1}{k+1} \rceil \geq \lceil \frac{l_2-l_1}{k+1} \rceil = \lceil \frac{2l_1}{k+1} \rceil > a$ errors. $\square$

Proof of proposition 9. For a clique S, let $W_A = \{x \in X : \ x_S = A^{|S|}\}$, $W_B = \{x \in X : \ x_S = B^{|S|}\}$. Note that $x_S \notin \{A^{|S|}, B^{|S|}\}$ implies $x \notin \Lambda_{k,\alpha}$. Now, from any state in W_A, to escape the basin of attraction of W_A (and enter the basin of attraction of W_B) requires that some player in S play B without making an error. This will be easier when the external neighbor of such a player plays B, and harder when his external neighbor plays A. Note that each player in S has $|S|$ neighbors. This means $R(W_A)$ is bounded below by $\underline{r}_A$, given by:

$$|S| \geq \alpha(|S| - 1 - \underline{r}_A) \quad \Longrightarrow \quad \underline{r}_A \geq \frac{(\alpha - 1)}{\alpha}|S| - 1$$

$$\Longrightarrow \quad \underline{r}_A = \left\lceil \frac{(\alpha - 1)}{\alpha}|S| - 1 \right\rceil$$

and $CR^*(W_B)$ is bounded above by $\bar{cr}_B$, given by:

$$|S| \geq \alpha(|S| - \bar{cr}_B) \quad \Longrightarrow \quad \bar{cr}_B \geq \frac{(\alpha - 1)}{\alpha}|S|$$

$$\Longrightarrow \quad \bar{cr}_B = \left\lceil \frac{(\alpha - 1)}{\alpha}|S| \right\rceil .$$

For moves from W_B to W_A, coalitions of players can coordinate their move to playing A. For each such player this gives an extra $(k-1)$ neighbors who play A, thus reducing resistances for such moves. Such moves are easier when external neighbors play A, and harder when they play B. This means $R(W_B)$ is bounded below by $\underline{r}_B$, given by:

$$\alpha(1 + (k - 1) + \underline{r}_B) \geq |S| \quad \Longrightarrow \quad \underline{r}_B \geq \frac{|S|}{\alpha} - k$$

$$\Longrightarrow \quad \underline{r}_B = \left\lceil \frac{|S|}{\alpha} - k \right\rceil$$

and $CR^*(W_A)$ is bounded above by $\bar{cr}_A$, given by:

$$\alpha((k - 1) + \bar{cr}_A) \geq |S| \quad \Longrightarrow \quad \bar{cr}_A \geq \frac{|S|}{\alpha} - k + 1$$

$$\Longrightarrow \quad \bar{cr}_A = \left\lceil \frac{|S|}{\alpha} - k + 1 \right\rceil .$$

Simple manipulation then shows that:

$$
\begin{aligned}
|S| \geq \frac{(k+1)\alpha}{2-\alpha} \quad &\Longrightarrow \quad \underline{r}_B > \bar{cr}_B \quad \Longrightarrow \quad R(W_B) > CR^*(W_B) \\
|S| \leq \frac{(k-3)\alpha}{2-\alpha} \quad &\Longrightarrow \quad \underline{r}_A > \bar{cr}_A \quad \Longrightarrow \quad R(W_A) > CR^*(W_A)
\end{aligned}
$$

The result follows from Theorem 2 of Ellison (2000). □

Proof of proposition 10. For given x, if $x_{N_S} = A^{|N_S|}$ then $A^{|S|} \in \mathcal{A}_S(x)$. Furthermore, if $x_{N_S} = A^{|N_S|}$ and $x_S = A^{|S|}$ then $\mathcal{A}_S(x) = \{A^{|S|}\}$. Therefore $x_{N_S} = A^{|N_S|}$, $x \in \Lambda_{k,\alpha}$ implies $x_S = A^{|S|}$. If $x_{N_S} \neq A^{|N_S|}$ and $|S| > \frac{\alpha}{\alpha-1}$ then $\alpha(|S|-1) > |S|$ and the same argument applies. If $x_{N_S} \neq A^{|N_S|}$ and $|S| < \frac{\alpha}{\alpha-1}$ then for some player in S with an external neighbor who plays B, his payoff $|S|$ from playing B is higher than his maximum payoff from playing A of $\alpha(|S|-1)$, so he will play B. But then all of the players in S will have a neighbor who plays B, and the same argument will apply to them. □

Appendix C. Simulation methodology [ONLINE - NOT FOR PRINT]

Appendix C.1. Introduction

We describe below the methodology behind the simulation results presented in the paper. Full, commented, code for the simulations is available upon request from the authors.

Appendix C.2. The Main Algorithm

In Algorithm 1 we present pseudo-code to run a single replicate of the model under a given set of parameters as required. After random seed initialisation (see section Appendix C.4), the underlying network is either generated or imported from the empirical network library. Players' actions are initialised in all simulations at $B^{|N|}$ and payoffs calculated as per the model of the paper. The main loop is iterated until either the maximum number of steps (T) is reached or the convergence condition has been attained. Convergence in all simulations required the fraction of 'B' players in the population to be less than γ, the convergence limit.

Appendix C.3. Coalition formation

The problem we address is to identify subsets of up to k agents, $S \subseteq N$, who form connected subgraphs of $\mathbf{g}$. The complexity of enumerating feasible coalitions of size $|S|$ depends strongly on the topology of the network. For example, if $\mathbf{g}$ is the complete graph, then there are $\binom{|N|}{|S|}$ such coalitions, which for the generated networks in this paper ($|N| = 256, k \in \{4, 8\}$) gives rise to over 174 million and over 4×10^{14} ways of forming S respectively. On the other hand, for a $2-$regular ring network, there are only $|N|$ feasible coalitions of size $|S|$, since the topology constrains the composition of S to $|N|$ consecutive index sets, each with a different starting vertex. However, in most cases, even a small amount of density at the local level complicates the picture dramatically.

For this reason, rather than enumerating every feasible coalition of size less than k for a given network $\mathbf{g}$ and choosing S from a pre-defined library each iteration, we perform a run-time algorithm to define S each iteration. The algorithm is as follows: 1) Choose l from a uniform distribution on $\{1, \dots k\}$; 2) Choose $i \in N$ Set $S = \{i\}$; 3) Randomly (uniform) choose a vertex from the set of vertices adjacent to some member of S which are not in S and add this vertex to S. 4) Repeat step 3 until $|S| = l$. In effect, the algorithm guarantees that every possible topological configuration of a given coalition size will be selected with positive probability.

Algorithm 1 One simulation replicate for given parameter values

Require: T, the maximum number of iterations in a simulation; G_{type} or G_{id}, the network type to generate or real network ID to import; $|N|$, the number of agents in the population; α, the efficient coordination payoff; k, the maximum size of a coalition; ε, the tremble probability; p_b coalition branching parameter; γ convergence limit.

{*initialise:*}
seed ← Set and record random seed
G ← Create or import network($G_{type}|G_{id}$)
x ← initialise strategy vector(n)
π ← Update payoffs(G, x, α)
$t \leftarrow 1$
converged? ← FALSE
{*main loop:*}
while ($t < T$) AND (*converged*? = FALSE) **do**
  S ← Get coalition(G, k, p_b)
  x ← Obtain better response(S, G, x, π, α)
  x ← Apply tremble(S, x, ε)
  π ← Update payoffs(G, x, α)
  converged? ← Test convergence(x, γ)
  $t \leftarrow t + 1$
end while
return (*converged*?, t)

Appendix C.4. The random stream

The random stream used in the simulations is the MATLAB stream method `mt19937ar`, which is described by the MATLAB documentation as 'Mersenne Twister with Mersenne prime $2^{19937} - 1$'. In each experiment (unless indicated), 20 replicates were used with this stream, each with an individual seed defined by the number of the replicate (e.g. replicate 1 of 20 passed '1' to the initialisation method) to create the stream[41]. Thus, each replicate in a given experiment was conducted with identical random number stream conditions to the corresponding replicate of another experiment. Such a setup ensures that any differences amongst experiments can be attributed wholly to the conditions of the experiment and not to differing random number streams.

Appendix C.5. Networks

The simulation study considered two general categories of networks: generated networks; and empirical networks. In the first category, we constructed square lattices with Von-Neumann (VN) neighborhoods and ring networks via simple linear algebra methods written by one of the authors (available on request). Next, we adopted the algorithm of Watts and Strogatz (1998) to randomly 'rewire' 10% of the edges of these networks to form so-called 'small-world' (SW) networks, and then 100% of the edges to form random (rand) networks. The figure of 10% rewiring was settled on after inspection of the characteristic path length $L(\mathbf{g})$, and clustering coefficient $C(\mathbf{g})$, of each network type after systematic rewiring from 0% to 100%. A 'small-world' can be said to have been created when $\mathbf{g}$, having $|N|$ vertices and K average degree, has $L(\mathbf{g})$ similar to the equivalent (asymptotic) values for a random network of similar $|N|$ and K, $\mathbf{g}_{R(N,K)}$. The 'small world' will have an equivalently low average path length to a counterpart random network but much higher local edge density.

In Table C.6 and Fig. C.16 we present summary statistics for the networks used in the simulations where $|N|$, E and K are the number of vertices, edges and the average degree of the network respectively, whilst $C(\mathbf{g})$ and $L(\mathbf{g})$ respectively give the clustering coefficiency and characteristic path length of the network. As can be seen, the networks we study cover a diverse region within $\frac{L(\mathbf{g})}{L(\mathbf{g}_{R(N,K)})}$—$\frac{C(\mathbf{g})}{C(\mathbf{g}_{R(N,K)})}$ space.

---

[41] Actual MATLAB code used:
`s1 = RandStream.create('mt19937ar','seed',r);`
`RandStream.setDefaultStream(s1);`, where 'r' is the replicate number.

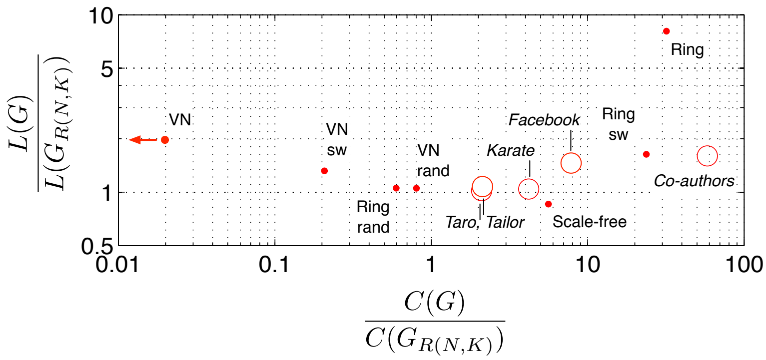


Figure C.16: Properties of networks considered in the simulations. See Table C.6 for details.

Table C.6: Summary statistics for the networks considered in the simulations.

Network Type	$\lvert N \rvert$	E	K	$C(\mathbf{g})$	$L(\mathbf{g})$	$\frac{C(\mathbf{g})}{C(\mathbf{g}_{R(N,K)})}$	$\frac{L(\mathbf{g})}{L(\mathbf{g}_{R(N,K)})}$
Artificial networks							
Square lattice (VN)	256	512	4.0	0.0000	8.0	0.0	2.0
				(na)	(na)		
Square lattice (small world)	256	512	4.0	0.0032	5.3	0.2	1.3
				(0.0026)	(0.1)		
Square lattice (random)	256	512	4.0	0.0126	4.2	0.8	1.1
				(0.0047)	(0.0)		
Ring	256	512	4.0	0.4981	32.4	31.9	8.1
				(na)	(na)		
Ring (small world)	256	512	4.0	0.3701	6.5	23.7	1.6
				(0.0166)	(0.4)		
Ring (random)	256	511	4.0	0.0093	4.2	0.6	1.1
				(0.0041)	(0.0)		
Scale-free	256	510	4.0	0.0809	3.4	5.2	0.9
				(0.0259)	(0.1)		
Real networks							
Zachary's Karate Club (Zachary, 1977)	34	78	4.6	0.571	2.41	4.2	1.04
Network Th. Co-authorship (Newman, 2006b,a)	379	914	4.8	0.741	6.042	58.3	1.6
Schwimmer's Taro Exchange (Schwimmer, 1973)	22	39	3.5	0.339	2.494	2.1	1.02
Kapferer's Tailor Shop (Kapferer, 1972)	39	158	8.1	0.458	1.772	2.2	1.01
Facebook sub-graph (McAuley and Leskovec, 2012)	95	325	6.8	0.559	3.509	7.8	1.5

Appendix D. Further discussion of model assumptions [ONLINE - NOT FOR PRINT]

Assumption 1: Better responses

Assumption 1 states that every best response by a coalition has a positive probability. As a full support assumption, this assumption is weak in the sense that any of the probabilities can be arbitrarily small as long as they are positive. The independence of results with respect to the exact probabilities is a consequence of the theoretical results being given for ε approaching zero. One possible criticism of the use of coalitional better responses is that it allows coalitions to choose action profiles which are Pareto dominated. The addition of a Pareto condition to create coalitional best responses would not change the analysis in this paper, but it should be borne in mind if the theoretical apparatus of the paper is ever used to analyze a different underlying game.

A second possible criticism is that coalitions are allowed heterogeneity in their better responses. That is, a coalition may alter its actions in a way that some members of the coalition play A while others play B. The analyis in the current paper only uses coalitional changes from the status quo whereby the entire coalition switches to play A or the entire coalition switches to play B. Hence, the results of the paper would not change under a restriction that a coalition changing its actions from the status quo must do so homogeneously. The same proviso as in the previous paragraph applies with regards to the further application of the theory. Furthermore, the authors would defend the possibility of a coalition agreeing on heterogeneous actions on the grounds that it highlights the distinction between action choice and payoff determination. Put simply, choosing what to do is different from doing it, and the idea of two people agreeing to act differently is plausible. Consider two cliques joined by a single edge. It is reasonable to imagine a coalition of the two players at either end of this edge agreeing something along the lines of "your other friends all play A so you can play A, but my other friends all play B so I shall play B."

Assumption 2: Full support

This is another full support assumption, stating that the set of coalitions which will with some probability update their strategy profiles is the entire set of feasible coalitions. The assumption is weak in the same sense as assumption 1 discussed in the previous section. Criticisms must therefore be directed at the support of the distribution. The authors regard allowing all coalitions below a certain size to be a fairly natural assumption. However, there will exist domains of application for which different support assumptions, such as forbidding coalitions below a certain size, make greater sense.

Error probabilities

The paper works with errors that occur uniformly and independently among members of the currently active coalition. Here we address several questions that can arise about the error process.

Errors outside a coalition

The first question is: what if players outside the active coalition can also make errors? This does not affect results: the effect of a player outside of an active coalition making an error can be easily replicated within the model of the paper by selecting that player as an active coalition and having him make an error.

Payoff dependent errors

Aside from uniform errors, the other specification common in the literature is of error probabilities which are log-linear in payoff loss. Such a specification would substantially complicate exposition, detracting from a clear description of the forces at play. Moreover, there are good theoretical arguments why uniform error probabilities may be more plausible than log-linear specifications (see van Damme and Weibull, 2002).

Perfectly correlated errors

Consider an adjustment to the model in which errors in a coalition are perfectly correlated, and moreover, they still occur with probability ε. That is, with probability ε, every player in the coalition makes an error. In this way up to k players could make an error together. This is a different model, and a full discussion is beyond the scope of this paper. However, we note that a conservative effect of coalitional behavior can still exist in the model with perfect correlation. Consider the graph of overlapping triangles in figure 4. Then, for any $k \geq 2$, if the graph is large enough relative to k, a conservative effect exists for $\alpha < 3/2$. This can be seen to be the case as even if many errors have occurred to create a large block of A players, pairs of players on the edges of this block can still profitably switch to play B.

Slightly correlated errors

Consider an adjustment to the model in which the players in an active coalition are sequentially given the opportunity to make an error, and errors occur with probability ε if nobody in the coalition has yet made an error, and with probability $\varepsilon^{1-\rho}$, ρ close to zero, if a member of the coalition has already made an error. For small enough ρ, the analysis in the paper will remain substantially the same.