The University of Adelaide
School of Economics

Research Paper No. 2011-13
March 2011

Information, Matching and Outcome Selection

Virginie Masson

Information, Matching and Outcome Selection *

Virginie Masson †
School of Economics
University of Adelaide

March 7, 2011

Abstract

We consider a finite population of agents who exchange information and are paired every period to play a game with tension between risk dominance and Pareto efficiency. Agents sample past plays and corresponding payoffs from their information neighborhood, and choose one of two possible actions using either best response or imitation. Information exchanges and possible matchings each constitutes a network. We first provide a complete description of the medium run outcomes and show that in the medium run only information matters. We then identify the conditions whereby either the risk dominant or the Pareto efficient convention is stochastically stable, and show how efficiency in the long run depends on the matching network.

JEL Classification C73, D85.

Keywords Contagion · Networks · Coordination games · Best response · Imitation

---

*I am particularly grateful to Fabrice Collard for his interest in this paper, the multiple discussions we had and his very valuable support. I also would like to thank Alexander Matros, Dmitriy Kvasov, Kieron Meagher and audiences at UNSW, ANU and Massey University for their insight and numerous suggestions. Also thanks to Steven Hail, Mark Dodd and Markus Brueckner for their comments at different stages of this manuscript.

†Corresponding author. School of Economics, University of Adelaide, Adelaide, SA 5005, Australia. *Email*: virginie.masson@adelaide.edu.au Phone: +61 8 8303 4930. Fax: +61 8 8223 1460.

1 Introduction

A conventional way of representing information exchanges and interactions is to construct a unique network that accounts for both. The graphs of such networks require every edge to simultaneously capture information exchange and interaction. However this is a restrictive representation of reality, as we will argue later. A few studies on evolutionary games in networks offer some alternatives by modeling distinct information and interaction networks. While the information network depicts the exchanges of information between agents, the interaction network captures who plays with whom. In these studies however, both networks are assumed to be embedded in one another. Either interactions always occur between information neighbors, as in Alós-Ferrer and Weidenholzer [1] and Mengel [9], or information is only gathered from the interaction neighborhood, as in Durieu and Solal [4]. The purpose of this paper is to study the potential implications of such embedded networks. In particular, it provides a better understanding of the role of information versus interaction in shaping the medium and long run outcomes of a standard evolutionary game.

To illustrate why embedded networks may be unsuitable, consider a modified version of the original Stag-Hunt. A group of hunters may consist of mutual friends and their invited guests. Invited guests are friends with their hosts but unfamiliar to others. The group of mutual friends know about each other's past hunting behaviors from previous conversations, even though some never hunted together. This illustrates how information can be gathered without interaction. Once the hunt commences, the group dissociates into pairs. No communication can occur during the hunt to avoid alerting the animal in pursuit. While some pairs are formed between mutual friends, others are formed among strangers. In this last instance, interaction occurs without information exchange.

Our analysis unfolds within a discrete time framework. We consider a finite population whose agents exchange information. Each period, these agents interact in pairs to play a 2×2 coordination game that exhibits some tension between risk dominance and Pareto efficiency. Agents sample past plays and corresponding payoffs from their information neighborhood, and choose one of two possible actions using either best response or imitation. The details of the information exchanges and the possible matchings each constitutes a network, illustrated by a graph with agents as the vertices. We first study the stability of the system in a mistake free environment and then extend our analysis to a perturbed version of the process.

Like previous studies by Mengel [9] and Alós-Ferrer and Weidenholzer [1] we allow information to be sampled outside the interaction neighborhood. We also allow interactions to occur outside the information neighborhood like Durieu and Solal [4]. To simultaneously capture both scenarios, we use a different concept to the interaction network, and consider a matching network. It captures the potential pairings between agents rather than identifying for each agent an unchanging subset of partners with whom each period the game is played. Although we lose the multiplicity of interactions per period, we offer a more flexible environment where rare or unlikely encounters are accounted for. These infrequent matchings may be more the exception than the rule, but they play a major role in identifying the conditions by which an outcome is stochastically stable. Furthermore, by untwining the two networks, we can show that the matching network plays no role in shaping medium run outcomes.

The first related study was developed by Durieu and Solal [4]. In their paper, they consider a finite population of n agents placed on a non-oriented ring. Each agent interacts with $2k < n$ agents in each period to play a game. The game is a symmetric coordination game that exhibits tension between risk dominance and Pareto efficiency. Agents choose their actions according to a spatial sampling procedure similar to the one introduced by Young [11] [12]. More precisely, each agent samples the actions played by $r < 2k$ of their interaction neighbors in the previous period and plays best-response to the frequency of actions in the sample. Within this framework, Durieu and Solal [4] show that the risk dominant action prevails in the long-run in most cases. Our sampling procedure bears some similarity with theirs, the game is identical and we also consider best reply as a possible decision rule. However, our paper differs from theirs as we study arbitrary information networks with possible non-reciprocal information exchanges, and also investigate imitation as a decision rule.

Mengel [9] also considers agents on a non-oriented ring. Agents choose their action using either a payoff-biased or a conformist-biased imitation rule. In her study, she shows that when agents can access information beyond their interaction neighborhood and use a payoff-biased imitation rule, a unique stochastic stable outcome is identified where everyone chooses to defect in a prisoner's dilemma type game. This contrasts with the results obtained by Eshel, Samuelson and Shaked [7], who consider identical information and interaction neighborhoods. However, if agents use a conformist-biased imitation rule with a strong enough bias, full cooperation emerges regardless of whether agents can access information beyond their interaction neighborhood. Although we also consider imitation, we look at a different rule whereby agents imitate the most successful action in the sample they draw. The game we study differs as it contains tension between risk dominance and Pareto efficiency, and as stated earlier, we look at arbitrary networks and best-response.

The closest study to our own is the one by Alós-Ferrer and Weidenholzer [1]. They consider arbitrary information networks and assume that agents interact only with a subset of their information neighbors. Each period, agents sample information from their information neighbors and choose the most successful action in the sample as the action to play in the current period. The authors focus on whether the Pareto efficient action is contagious. The notion of contagion they use is adopted from Morris [10] whereby an action is said to be contagious if it spreads to the entire population when it is initially played by only a finite subset of agents.The combination of local interactions with imitation is what, the authors conclude, leads to the contagion of the Pareto efficient action. They extend their results to non-reciprocal information exchanges, however this extension only applies to those neighbors with whom one does not interact. We also consider arbitrary networks and imitation, and allow for non-reciprocal information exchanges. The way we introduce asymmetric information differs however due to a different approach in modeling interactions. In particular, we account for interactions without imposing reciprocal information exchange. In our modified Stag-Hunt example, this represents the scenario where, in a pair of hunters, only one has information about the other 's past hunting behavior. Furthermore, we provide a thorough analysis into the role played by information and interactions in both, the medium, or mistake free environment, and the long run. In particular, we provide a detailed characterization of medium run outcomes and identify the conditions by which the risk dominant convention is stochastically stable, even with imitators. This contrasts with Alós-Ferrer and Weidenholzer [1] whose focus on the contagion of the Pareto efficient action allows them to bypass the medium run analysis, and neglect the risk dominant uniform convention as a candidate for stochastic stability.

Our results provide, for any arbitrary network, a complete characterization of the medium run outcomes and show that the matching network plays no role in identifying them, whether agents are best repliers or imitators. We then identify the characteristics of the information network, matching network and sample size under which stochastic stability arises. We show that the risk dominant uniform convention emerges as stochastically stable when either the sample size is within some bounds and agents are best repliers, or when agents imitate and both the sample size and the number of strongly connected components of the information network equal one. Finally, given some conditions on the sample size relative to the number of strongly connected components of the information network, we find that the Pareto efficient uniform convention is stochastically stable when agents imitate. However, the conditions required for the sample size are sensitive to variations of the matching network.

In Section 2, we describe the networks, the game played by the agents and the decision rules used. Absorbing sets and stochastically stable sets are presented in Section 3. Section 4 concludes.

2 The model

In what follows, we first introduce the information network and decompose its graph into strongly connected components. These components play a major role in describing the absorbing sets of the system. We then describe the matching network and insure that the pairing of all agents is always possible. Finally, we explain how agents choose their actions, using either imitation or best response, when playing the game. This last subsection ends with an illustration of the different roles played by each of the networks depending on the decision rule in use.

2.1 Networks

2.1.1 Information Network

We represent the information network as a directed graph over the set of agents $I = \{1, 2, ..., 2n\}$ with adjacency matrix G. An element g_{ij} of the adjacency matrix is therefore equal to 1 if there is an information flow from i to j and 0 otherwise. More precisely, $i \rightarrow j$ means that agent j can access agent i's information in which case agent i is agent j's neighbor; the reverse is however not true, and if $g_{ji} \neq 1$, agent j is not a neighbor of agent i. This allows us to capture possible non-reciprocal information exchanges.

In the case where agent i accesses information from agent j and provides it to agent k, we have a flow from agent j to agent k. What matters in our framework is who accesses whose information rather than the means by which the information is accessed. We can differentiate between these cases by saying that agent j is a direct neighbor of agent i but an indirect neighbor of agent k. However this distinction will only be used when referring to the matching network and the possible matchings it accounts for.

For every agent $i \in I$ we then define a set of neighbors $N_i = \{j \in I, g_{ji} = 1\}$. We assume that agent i accesses his own information, and hence $g_{ii} = 1$ and $i \in N_i$. All agents have a finite memory containing their m most recent actions and associated payoffs. These two elements, actions and corresponding payoffs, constitute what we refer to as information, and this is what agents can potentially observed from one another.

Since we consider *any* possible information network we need to introduce some ways to decompose the network into tractable objects. In order to do so, we first introduce the notion of directed path:

Definition 1 *A directed path between two vertices i_0 and i_k denoted $\mathcal{P}_{i_0 i_k}$ is a non-empty directed graph with distinct vertices $\{i_0, i_1, ..., i_k\}$ for which $g_{i_{m-1} i_m} = 1$ for all i_m, $m = 1, ...k$.*

A directed path is therefore a sequence of directed links from one agent to another. It captures how information can potentially be transmitted between two agents. When $\mathcal{P}_{ij}$ exists, it does not guarantee however that agent j observes agent i's information. It only captures the potential influence agent i can exert over agent j through indirect transmission of information.

Consider now a group of agents. If a directed path exists between any pair of agents within that group, it forms a strongly connected component (SCC):

Definition 2 *A strongly connected component $C \subseteq I$ of an information network is a maximal set of vertices such that for all $(i, j) \in C$, with $i \neq j$, $\mathcal{P}_{ij}$ and $\mathcal{P}_{ji}$ exist.*

As every directed graph is a directed acyclic graph of its strongly connected components it is possible to uniquely decompose an information network into its strongly connected components.[1] Furthermore, we can identify and separate two main types of SCCs: source (no incoming edge) and sink (no outgoing edge). This allows us to capture the channels by which information disseminates through the population. More formally,

[1]See S. Dasgupta, C.H. Papadimitriou, and U.V. Vazirani [3]

Definition 3 *A strongly connected component $C \subseteq I$ of an information network is said to be source if there is no $i \in I \backslash C$ and there is no $j \in C$ such that $g_{ij} = 1$*

Hence, agents within a source SCC never receive information from agents outside the SCC. Similarly,

Definition 4 *A strongly connected component $C \subseteq I$ of an information network is said to be sink if there is no $i \in I \backslash C$ and there is no $j \in C$ such that $g_{ji} = 1$*

The difference between this definition and the previous one resides in the change from $g_{ij} = 1$ to $g_{ji} = 1$. Therefore, agents in a sink SCC never provide information to agents outside their SCC. Finally, a SCC whose agents never provide nor receive information to and from other SCCs is called disconnected. Formally,

Definition 5 *A strongly connected component $C \subseteq I$ of an information network is said to be disconnected if it is source and sink.*

For the rest of the paper, we denote $\mathcal{C} = \{C_1, ..., C_k, ...\}$ the set of all SCCs of the information network, and $\mathcal{S}$ the set of all source SCCs of the information network.

Example: In order to provide a better understanding of the abovementioned definitions, Figure 1 reports an information network featuring nine SCCS. Figure 1 Panel A illustrates the network at the agent level, while Figure 1 Panel B exemplifies the decomposition of the network into SCCs. For clarity purposes, we did not draw edges from an agent to himself and this convention is adopted throughout the paper. Three of the SCCs are source: C_1, C_7 and C_9; six of the SCCs are sink : C_2, C_3, C_4, C_6, C_8, and C_9; one is neither source nor sink: C_5; one is identified as being both, source and sink, and is therefore disconnected: C_9.

2.1.2 Matching Network

The matching network over I describes the potential pairings between agents and is assumed to be distinct from the information network. All possible

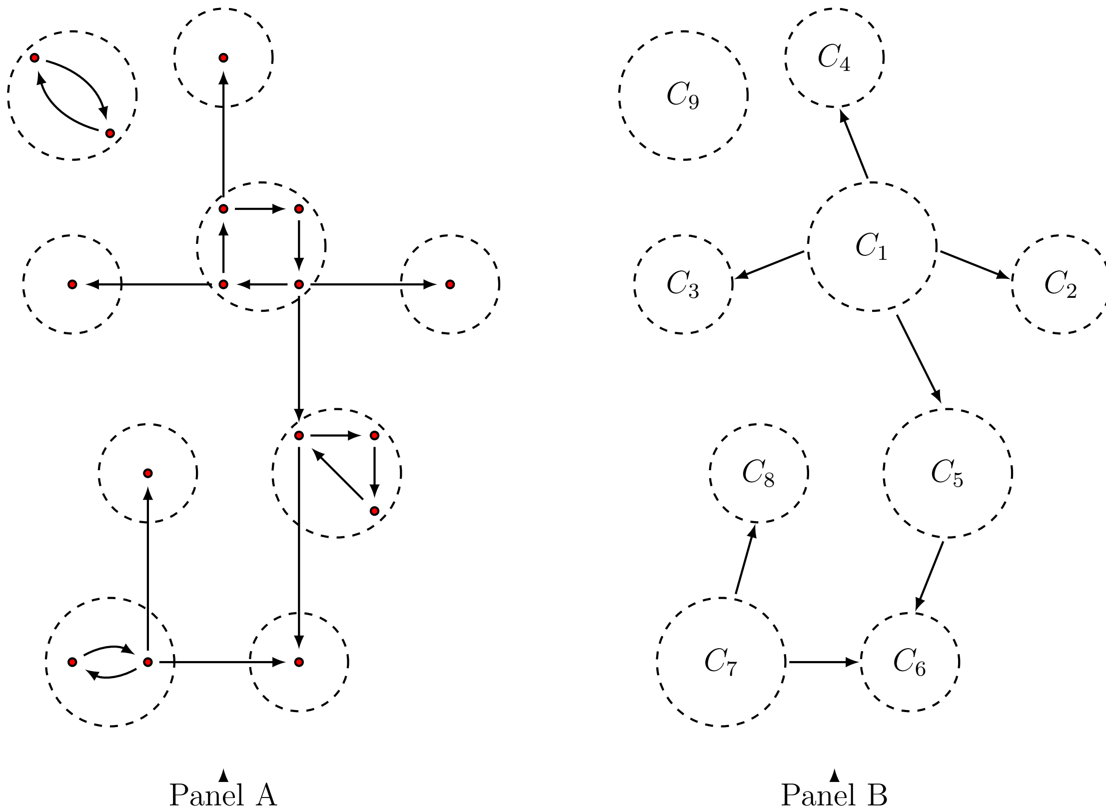

Figure 1: Decomposition of an Information Network into the Directed Acyclic Graph of its Strongly Connected Components

matchings are described by an undirected graph with agents as vertices and an edge (undirected link) between two agents represents the possibility that these agents could be paired to play the game. Denote M the adjacency matrix of the graph. An element m_{ij} of M is equal to 1 if there is a positive probability that agents i and j can be paired, and 0 otherwise. Of course, $m_{ii} = 0$ and $m_{ij} = m_{ji}$ for all $i \in I$.

Without further requirements on the matching network, there may be situations where the pairing of all agents is impossible. To avoid this, we thus make the following two assumptions.

First, we want to make sure that all agents can be paired every period so that every agent has someone to play with. This means that we should be able to separate the population into two distinct groups, and every agent in one group could potentially be matched with at least one agent in the other group. This is captured by Assumption 1.

Assumption 1 *There exists a bi-partition of the set of agents I into I_1 and I_2 such that the three following conditions are satisfied:*

(1) $I_1 \cap I_2 = \emptyset$

(2) $I_1 \cup I_2 = I$

(3) $m_{ij} = 1$*, for all* $i \in I_1$ *and for all* $j \in I_2$*, and 0 otherwise.*

Furthermore, if two agents can potentially be matched, i.e. if $m_{ij} = 1$, then it should be possible to pair the $2n - 2$ remaining agents. Otherwise m_{ij} should be 0, as stated in Assumption 2.

Assumption 2 *If* $m_{ij} = 1$ *then there exists a bi-partition of* I *satisfying assumption 1.*

Example: To illustrate these assumptions and the role they play, we consider four agents numbered from 1 to 4. Figure 2 illustrates a matching network where $m_{12} = m_{13} = m_{24} = 1$ but $m_{14} = m_{23} = m_{34} = 0$. In this case, the pairing between agent 1 and agent 2 can never be realized as agents 3 and 4 cannot be paired. Hence, the only matching that can arise consists in pairing agents 1 and 3 together and agents 2 and 4 together. The role of our second assumption is therefore to restrict m_{12} to 0.

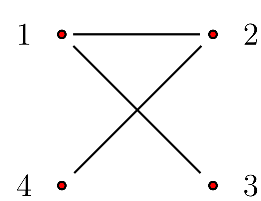

Figure 2: Matching network violating assumption 2

By defining the matching network this way we can for example model that direct neighbors are more likely to be paired; or we can restrict some agents from ever being matched with one another. We can further tie the matching network to the information network and impose that only neighbors can be paired, and hence interactions only happen when some information exchange is possible. Or by contrast, we can assume that both networks are completely independent from one another.

In the next section, we introduce the game played every period and describe in details the decision rules agents follow.

2.2 Dynamics of the game

We consider a standard game of the literature which normal form is represented by the following payoff matrix:

	1	0
1	a, a	b, c
0	c, b	d, d

where $a > c$ and $d > b$ so that both, $(1, 1)$ and $(0, 0)$ are strict Nash equilibrium.

Let $q = \frac{d-b}{a-b-c+d}$ be the probability associated with playing action 1 in the mixed strategy Nash Equilibrium of this game. If $q \geqslant 1/2$, then action 0 is said to be *risk dominant* as defined by Harsanyi and Selten [8]. The higher the value of q, the more risk dominant action 0 is. Furthermore, we assume that $a > d$ so that the equilibrium $(1, 1)$ is *Pareto efficient*. We hence refer to 1 as the Pareto efficient action and to 0 as the risk dominant action.

Agents choose their action relative to the information they sample from their neighbors and the decision rule they use. We consider two decision rules: *best response* (BR) and *imitate-the-best* (IM). The choice of these two rules is motivated by the different perspectives they offer. With BR, the frequency of occurrence of each action matters while payoffs are overlooked. In our framework, every agent samples s actions from his neighbors and computes the frequency of occurrence of each action within that sample. When the frequency of action 1 in the sample is strictly less than q, the agent chooses action 0; while if it is exactly q, the agent is indifferent and chooses his action at random.

When agents use IM however, they base their decision on the payoffs incurred by each action. More precisely, every agent samples s actions and corresponding payoffs from his neighbors and selects the action that gave the highest payoff in that sample. Again, when an agent is indifferent between the two actions, he picks an action at random.

Contrary to BR, IM is fundamentally shaped by payoffs. Thus, information and matchings do matter, while with BR, only information seem to matter as illustrated by the following example.

Example: Consider a population $I = \{1, 2, 3, 4\}$ whose agents play the game presented in Section 2.2. Let $a = 6$, $b = 2$, $c = 5$ and $d = 4$ and hence $q = \frac{2}{3}$.

The information network is such that $N_1 = N_2 = \{1, 2\}$ and $N_3 = N_4 = \{3, 4\}$. In other words, agents 1 and 2 share information, and so do agents 3 and 4, but agents 1 and 2 do not share any information with either 3 and/or 4. We represent this population and its information network in Figure 3. Each plain line with an arrow represents a directed link of the information network. We have,

$$\text{Information Network: } g_{12} = g_{21} = g_{34} = g_{43} = 1 \text{ and } g_{ij} = 0 \text{ otherwise.}$$

We consider two possible matching networks. The first network, labeled Network A, assumes that agent 1 can only be matched with agent 3 and agent 2 can only be matched with agent 4. The second network, labeled Network B, assumes that agent 1 can only be paired with agent 2, and agent 3 can only be paired with agent 4. Network A and Network B are represented by the dashed lines in panel (a) and panel (b) of Figure 3 respectively. This therefore amounts to assume the following:

$$\begin{array}{ll} \text{Matching Network A:} & m_{13} = m_{24} = 1 \text{ and } m_{ij} = 0 \text{ otherwise.} \\ \text{Matching Network B:} & m_{12} = m_{34} = 1 \text{ and } m_{ij} = 0 \text{ otherwise.} \end{array}$$

We assume that every agent remembers their most recent action and corresponding payoff ($m = 1$) and samples two observations ($s = 2$). In this example, it means that agent 1 (resp. 2) always samples from himself and agent 2 (resp. 1), and that agent 3 (resp. 4) always samples from himself and agent 4 (resp. 3). We also assume that, in period t, agents 1 and 3 played action 1, while agents 2 and 4 played action 0. Hence, in period $t + 1$, the memory of all agents for each of the interaction networks is therefore:

$$\begin{array}{ll} \text{Agents' memories with Network A:} & ((1, 6), (0, 4), (1, 6), (0, 4)) \\ \text{Agents' memories with Network B:} & ((1, 2), (0, 5), (1, 2), (0, 5)) \end{array}$$

Figure 3: Varying Matching Networks

(a) Matching Network A (b) Matching Network B

where the first element represents the memory of agent 1, the second element represents the memory of agent 2 and so on. Let us now consider, for each of the matching networks and each decision rule, the agents' actions next periods .

Consider first the case where agents use BR. Recall that in this case, agents base their decision on the relative frequency of each action in the sample they drew from their neighbors. Since the relative frequency of action 1 in all samples is $1/2$, which is less than $2/3$, all agents will choose action 0 next period. Consequently, when agents use BR, they will all play action 0 in period $t+1$, and any period thereafter, independently of which matching network is in place.

This is not however the case when agents use IM. In that case, if matchings are characterized by Matching Network A, all agents will choose action 1 next period, as all agents will sample the most successful action-payoff pair, that is $(1, 6)$. On the other hand, if matchings are captured by Matching Network B, all agents will choose action 0 next period as, in all samples, action 0 provided a payoff of 5 which is higher than the payoff of 2 obtained by action 1. Hence, when agents use IM, the matching network plays a major role in shaping agents' decisions.

This example illustrates the respective roles of the information and matching networks. We now investigate in what follows this issue more thoroughly.

3 Stability of the system

3.1 Absorbing sets of the Process

Define Θ, the state of the system at a given period, as the collection of the m most recent actions of every agent. The dynamics described above define a finite Markov process over the state space for which standard techniques apply.

We first want to characterize the absorbing sets of the process. An absorbing set is defined as a minimal set of states that has positive probability to be reached but cannot be left. A singleton absorbing set is referred to as an absorbing state. These sets capture the stability of the system when agents make no mistakes, i.e. when agents choose their action according to one of the decision rules described above. Furthermore, these sets are also the only candidates for stochastic stability, which we will study later.

The characterization of all absorbing sets necessitate that we introduce a number of concepts, the first of which is the notion of local state. Similarly to Θ, denote θ_{C_k}, the local state within $C_k \in \mathcal{C}$, as the ordered collection of the m most recent actions of every agent i in C_k. If all agents within C_k have the same m most recent actions, then θ_{C_k} is called a local convention. It is denoted θ^p, where $p = 1$ [resp. 0] if all agents in C_k played action 1 [resp. 0] for the last m periods. Furthermore, denote L_{C_k}, the set of source SCCs which provide some information to C_k. Formally,

$$L_{C_k} = \{C_l \in \mathcal{S} \text{ s.t } \exists\, i \in C_l \text{ and } \exists\, j \in C_k \text{ for which } \mathcal{P}_{ij} \text{ exists } \}.$$

We can now state Proposition 1.

Proposition 1 *Consider an arbitrary information network and a matching network satisfying Assumptions 1 and 2. All agents play IM or BR. If $s \leq m/2$, a set of states is absorbing if and only if both conditions are satisfied:*

(1) $\forall\, C_k \in \mathcal{S}$, either $\theta_{C_k} = \theta^1$ or $\theta_{C_k} = \theta^0$

(2) $\forall\, C_k \in \mathcal{C} \setminus \mathcal{S}$, $\theta_{C_k} = \theta^p$ if $\forall\, C_l \in L_k$, $\theta_{C_l} = \theta^p$, with p either 0 or 1.

Proof: See Appendix.

As stated in Proposition 1, absorbing states are characterized entirely by the structure of the information network. They are independent of the matching network or the decision rule. This is because, when looking at absorbing sets, we are only interested in finding the conditions for which agents maintain their action choice. Therefore, if all the information available to an agent comprises only one action, then this agent will choose this action independently of the payoff generated by the action and the decision rule in place. This in turns means that only the samples drawn, and hence the information network, matter.

Intuitively, source SCCs play a regulatory role in the stability of the system as their agents never receive information from outside agents. If a source SCC follows a local convention, i.e. all agents within the SCC played the same action for the last m periods, it will never leave it. This is captured by the first condition of Proposition 1.

For non-source SCCs however, the origin of the information sampled by their agents matters. To understand further the role played by L_k, consider the following. If all source SCCs within L_k follow the same local convention, then it has to be that all the SCCs from which agents in C_k can potentially receive information also follow the same local convention. Hence, if C_k follows the same local convention as all the source SCCs in L_{C_k}, then the local convention within C_k will not change as no information related to the other action can be sampled. This is captured by the second condition of Proposition 1. However, when SCCs in L_k do not all follow the same local convention, C_k may never settle in a local convention: the samples drawn by the agents in C_k potentially contain attractive information for both actions. This can trigger changes in the choice of actions of some agents. Although this situation cannot be found in an absorbing state, it can be part of an absorbing set.

The following example is an illustration of Proposition 1.

Example: Consider the information network presented in Figure 1. We identify L_k for each of the nine SCCs: $L_{C_2} = L_{C_3} = L_{C_4} = L_{C_5} = \{C_1\}$, $L_{C_6} = \{C_1, C_7\}$ and $L_{C_8} = \{C_7\}$. Note that by definition, $L_{C_1} = L_{C_7} = L_{C_9} = \emptyset$.

Given the structure of the information network considered, there are eight possible absorbing sets: the two uniform conventions, and the 6 absorbing sets described in Table 1 below.

In Table 1, each column represents the local state within each of the SCCs. More precisely, the first column represents the local state within C_1, the second column represents the local state within C_2 and so on. This leads to each line characterizing an absorbing set. The notation θ stands for any local state within the SCC, including θ^1 and θ^0.

Table 1: Characterization of absorbing sets

	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9
Abs set # 1	θ^1	θ^1	θ^1	θ^1	θ^1	θ^1	θ^1	θ^1	θ^0
Abs set # 2	θ^1	θ^1	θ^1	θ^1	θ^1	θ	θ^0	θ^0	θ^0
Abs set # 3	θ^1	θ^1	θ^1	θ^1	θ^1	θ	θ^0	θ^0	θ^1
Abs set # 4	θ^0	θ^0	θ^0	θ^0	θ^0	θ^0	θ^0	θ^0	θ^1
Abs set # 5	θ^0	θ^0	θ^0	θ^0	θ^0	θ	θ^1	θ^1	θ^0
Abs set # 6	θ^0	θ^0	θ^0	θ^0	θ^0	θ	θ^1	θ^1	θ^1

To relate our results to the literature, we look at the case where information links are double-sided (symmetric information) and matchings only occur between neighbors. Both assumptions are formalized below. First, information is symmetric:

Assumption 3 *For any $i, j \in I$, $g_{ij} = g_{ji}$*

Then, matchings occur only between neighbors:

Assumption 4 *For any $i,\ j \in I$, $m_{ij} = 1$ only if $g_{ij} = 1$*

One of the consequences of Assumption 3 is that the information network presents only disconnected SCCs. Furthermore, it is worth noting that assumptions 1, 2 and 4 eradicate the possibility of neighborless agent. Indeed, Assumptions 1 and 2 insure that the pairing of all agents is always possible, while Assumption 4 restricts matchings to occur only between neighbors.

As suggested by Proposition 1, the role played by the matching network in determining the absorbing sets of the system is irrelevant. Hence results presented in the following Corollary hold with or without Assumption 4:

Corollary 1 *Consider an arbitrary information network satisfying Assumption 3 and a matching network satisfying Assumptions 1, 2 and 4. All agents play IM or BR. A set of sets is absorbing if and only if $s \leq m/2$ and $\theta_{C_k} = \theta^p$ where $p = 1$ or 0 for all $C_k \in \mathcal{S}$.*

Proof: An information network satisfying Assumption 3 only possesses disconnected SCCs. Hence, all SCCs are source SCCs and all follow a local convention if and only if the system is in an absorbing set. *End of Proof*

In other words, when information is symmetric, the system presents only absorbing states, where each of the SCCs follows a local convention. Our next step consists in perturbing the system so as to study its stochastic stability.

3.2 Stochastic stability

We consider a perturbed version of the Markov process described above and study the stochastic stability of the system. The perturbed dynamics are generated by assuming that each player independently and with probability $\epsilon > 0$ in each period picks a strategy at random from a uniform distribution. Using Ellison's method of radius-coradius [6], we then determine which absorbing sets are sochastically stable. Although we are fully aware of the criticism by Bergin and Lipman [2] regarding the uniformity of the perturbations, we avoid addressing this issue here and keep our focus on the role played by each network in characterizing long-run outcomes.

The radius-coradius method requires first the definitions of the following concepts. Denote $B(\Theta)$ the basin of attraction of an absorbing set Θ. The basin of attraction of Θ is the set of initial states from which the unperturbed Markov process converges to Θ with probability one. The radius of the basin of attraction of Θ, denoted $R(\Theta)$, is the mimimal number of mistakes needed to leave $B(\Theta)$ when play begins in Θ. The coradius of $B(\Theta)$, denoted $CR(\Theta)$, is the maximum over all other states of the minimum number of mitakes needed to reach $B(\Theta)$. A state Θ is said to be stochastically stable if $R(\Theta) > CR(\Theta)$.

In what follows, we start by explaining in more details the concepts of radius and coradius. We then look at the stochastic stability of the system depending on the decision rule in place.

3.2.1 Radius and Coradius

The Radius-Coradius method requires the computation of the number of mistakes needed to leave or enter an absorbing set's basin of attraction. In our framework, it relies on the understanding of the importance played by source SCCs in determining absorbing sets.

Consider an absorbing set Θ and C_k a source SCC. If C_k leaves the local convention it follows under Θ for a local state that could lead it to switch convention, then we have left the basin of attraction of Θ. The number of mistakes needed to achieve such result is the Radius of Θ.

For the computation of the coradius of Θ, we first need to introduce $\overline{\Theta}$, the *furthest* possible state from Θ. Formally, consider $C_k \in \mathcal{S}$ and let $\theta_{C_k}^{\Theta}$ be the local state of C_k in state Θ. Then define $\overline{\Theta}$ such that for all $C_k \in \mathcal{S}$:

$$\theta_{C_k}^{\overline{\Theta}} = \begin{cases} \theta^1 & \text{if } \theta_{C_k}^{\Theta} = \theta^0; \\ \theta^0 & \text{if } \theta_{C_k}^{\Theta} = \theta^1. \end{cases}$$

The state $\overline{\Theta}$ is the mirror image of Θ. It requires that all source SCCs switch away from the local convention they follow under Θ. This incurs the largest number of changes in local conventions. The Coradius of Θ is therefore the minimum number of mistakes needed to go from $\overline{\Theta}$ to $B(\Theta)$.

In the next sections, we study the stochastic stability of the system first with BR, and then with IM. The reason for such split resides in the role, or absence of role, played by the matching network in computing radius and coradius.

3.2.2 BR and Stochastic Stability: The Non Role of the Matching Network

Since BR depends only on frequency of plays and ignores payoffs, the matching network is irrelevant in determining stochastic stability. Hence, under BR, the stochastic stability of the system is relatively straightforward and entirely deducted from the properties of the information network. Denote $\lceil x \rceil$ the value of x if x is an integer, or the integer part of $x+1$ if x is not an integer.

Proposition 2 *Consider an arbitrary information network and a matching network satisfying Assumptions 1 and 2. All agents play BR. Denote K the number of source SCCs which follow θ^1 under Θ. We have,*

$$R(\Theta) = \begin{cases} \lceil (1-q)s \rceil & \text{if } \Theta \neq \Theta^0; \\ \lceil qs \rceil & \text{if } \Theta = \Theta^0. \end{cases}$$

and

$$CR(\Theta) = K \lceil qs \rceil + (Card(\mathcal{S}) - K) \lceil (1-q)s \rceil$$

Thus, if $s \leq m/2$ and $Card(\mathcal{S}) < \frac{\lceil qs \rceil}{\lceil (1-q)s \rceil}$ the uniform convention Θ^0 is stochastically stable.

Proof: See Appendix.

Proposition 2 states that given some restrictions on the sample size with respect to the risk dominant coefficient, the memory size of the agents and the number of SCCs, we find that when agents use best reply, the risk dominant uniform convention is stochastically stable. In particular, when $Card(\mathcal{S}) = 1$, i.e. when the information network contains only one strongly connected component, the expression $Card(\mathcal{S}) < \frac{\lceil qs \rceil}{\lceil (1-q)s \rceil}$ is always true. Hence, when looking at connected undirected information graphs, such a rings, if agents use best reply, the risk dominant uniform convention prevails in the long run. This corroborates results by Ellison [5] and Durieu and Solal [4].

Furthermore, since the matching network plays no role in determining stochastic stability with BR, whether or not matchings occur only between neighbors is irrelevant. Hence, with BR, we can have a complete independence of the two networks and still draw some conclusion regarding the stochastic stability of the system. This is not however the case with IM.

3.2.3 IM and Stochastic Stability: The Interdependence of the Networks

When all agents play IM, some assumptions tying up the two networks are more often than not required to get analytical results. One possibility, as seen earlier, is to restrict matchings to occur only between neighbors, but this would curtail our analysis. Another option is therefore to relax slightly this assumption to accommodate an environment with asymmetric information. In this case, matchings would only occur between agents who exchange some information, i.e. they do not have to both considered each other as neighbors. But this would still impose quite a strong dependence between the two networks and we would rather try keeping this dependence at a minimum.

In order to achieve this, we introduce the notion of information hub. The information hub of a given source SCC, say C_k, is the set of SCCs, including C_k, whose agents can be reached through a directed path from some agents in C_k. Define $\mathcal{P}_{C_k}$ the set of such SCCs, excluding C_k. Formally,

$$P_{C_k} = \{C_l \in \mathcal{C} \setminus C_k \text{ s.t } \exists\, i \in C_k \text{ and } \exists\, j \in C_l \text{ for which } \mathcal{P}_{ij} \text{ exists}\}.$$

The information hub of C_k is therefore defined as:

$$\mathcal{H}_{C_k} = P_{C_k} \cup C_k$$

In the example presented in Figure 1 Panel B, there are three information hubs: $\mathcal{H}_{C_1} = \{C_1, C_2, C_3, C_4, C_5, C_6\}$, $\mathcal{H}_{C_7} = \{C_6, C_7, C_8\}$ and $\mathcal{H}_{C_9} = \{C_9\}$. In what follows, we use this example to illustrate how the existence or nonexistence of matchings between agents from different information hubs can affect the computation of radii.

Suppose that the system presented in Figure 1 Panel B is in an absorbing state Θ^* where C_9 follows θ^0 while all other SCCs follow θ^1. This is absorbing set # 1 in Table 1. To leave $B(\Theta^*)$, some mistakes from agents in either C_1, C_7 or C_9 are required, so that one of these three SCCs can depart from its current local convention and possibly never return to it.

We first consider C_9. The number of mistakes needed from some agents in C_9 to leave $B(\Theta^*)$ depends on the matchings that can be realised. If agents from C_9 cannot be paired with agents from other hubs, then two simultaneous mistakes made by two paired agents in C_9 are necessary. This allows the highest payoff to be sampled, and hence, makes it possible for the system to leave $B(\Theta^*)$. If however an agent from C_9 can be matched with an agent outside C_9 and makes the mistake of playing action 1, one mistake is enough. This agent from C_9 then gets the highest possible payoff, which can then be sampled by other agents in C_9 and prompt them to switch action, thus allowing the system to depart from $B(\Theta^*)$.

Another possibility for the system to depart from $B(\Theta^*)$ is to consider agents in either C_1 or C_7. Consider C_1. For C_1 to possibly lead the system out of $B(\Theta^*)$, it requires s mistakes from one of its agents. Agents within C_1 can then draw samples which only contain action 0 and thus switch to action 0. This is enough for the system to depart from $B(\Theta^*)$. It is worth noting that in this last case the number of mistakes is independent of the matching network. This is because the payoffs generated by action 0 cannot surpass those from action 1.

From our example, it is clear that $R(\Theta^*)$ is contingent upon two things: the sample size, and the ties existing between the information network and the matching network. If $s = 1$, then $R(\Theta^*) = 1$, no matter the ties between the matching network and the information network. If $s > 1$ however, $R(\Theta^*)$ is either 1 or 2, depending on whether agents from distinct information hubs can be matched.

We therefore consider two possible scenarios: one that allows matchings within and between information hubs and one that restricts matchings within information hubs. These two scenarios are formally captured by the following two assumptions.

In Assumption 5, agents from different information hubs can be paired; however, these kind of matchings can remain sparse and infrequent.

Assumption 5 *For any $\mathcal{H}_{C_k}$ and $\mathcal{H}_{C_l}$, there exist agent $i \in C_k$ and agent $j \in \mathcal{H}_{C_l}$ such that $m_{ij} = 1$*

With Assumption 5, the propagation of mistakes is made easier, as illustrated earlier when agents from C_9 could be matched with agents from other information hubs. The radius of absorbing sets computed under this assumption are therefore expected to be lower than those computed under the next assumption.

Assumption 6 captures the possibility that some segregation exists between information hubs, as exemplified earlier when agents from C_8 could only be paired with themselves. Agents from different information hubs cannot be paired. Furthermore, if two information hubs happen to have common SCCs, then agents within these SCCs can only be matched with themselves. This makes it harder for a source SCC to leave its current local convention, and hence makes it harder for the system to leave the basin of attraction of its current absorbing set.

Assumption 6 *For any $\mathcal{H}_{C_k}$ and $\mathcal{H}_{C_l}$, if $\mathcal{H}_{C_k} \cap \mathcal{H}_{C_l} = \emptyset$, there is no agent $i \in \mathcal{H}_{C_k}$ and no agent $j \in \mathcal{H}_{C_l}$ such that $m_{ij} = 1$. Furthermore, if $\mathcal{H}_{C_k} \cap \mathcal{H}_{C_l} \neq \emptyset$ then for any agent $i \in \{\mathcal{H}_{C_k} \cap \mathcal{H}_{C_l}\}$, $m_{ij} = 1$ only if agent $j \in \{\mathcal{H}_{C_k} \cap \mathcal{H}_{C_l}\}$.*

We are now in a position to compute the radius of every absorbing set, and we start by computing the radius of both uniform conventions. The reason for the separation between these two states and other absorbing sets is based on the fact that when the system follows a uniform convention, the matching network is irrelevant when computing the radius of either convention.

Lemma 1 *Consider an arbitrary information network and a matching network satisfying Assumptions 1 and 2. All agents use IM. Then, $R(\Theta^0) = 2$ and $R(\Theta^1) = s$.*

Proof: Consider Θ^0 and $C_k \in \mathcal{S}$. We have $\theta_{C_k} = \theta^0$. If agent $i \in C_k$, and the agent is paired with simultaneously make the mistake of playing action 1, then agent i's memory contains the highest possible payoff. Agent i's neighbors can then sample this payoff and thus choose action 1. This also applies to the neighbors of agent i's neighbors and so on. This is enough to leave $B(\Theta^0)$. Hence, $R(\Theta^0) = 2$.
Similarly, we compute the radius of Θ^1. Consider $C_k \in \mathcal{S}$ which follows a local convention θ^1. The system can leave $B(\Theta^1)$, if one agent in C_k makes s successive mistakes of playing action 0 rather than action 1. This is because the neighbors of this agent can then draw samples uniquely composed of action 0 thus choosing action 0 themselves. Applying this reasoning recursively to the neighbors of these agents, and then the neighbors of the neighbors of these agents and so on, we have left $B(\Theta^1)$. Hence, $R(\Theta^1) = s$.
End of Proof.

We now compute, under Assumption 5, the radius of absorbing sets that are distinct from uniform conventions:

Lemma 2 *Consider an arbitrary information network and a matching network satisfying Assumptions 1, 2 and 5. All agents use IM. Then,* $R(\Theta) = 1$ *for any absorbing set* $\Theta \neq \Theta^0, \Theta^1$.

Proof: Since $\Theta \neq \Theta^1, \Theta^0$, there exist $C_k \in \mathcal{S}$ and $C_l \in \mathcal{S}$ such that $\theta_{C_k} = \theta^0$ and $\theta_{C_l} = \theta^1$. Also, from Assumption 5, there exist $i \in C_k$ and $j \in C_l$ such that $m_{ij} = 1$. Hence, there is a positive probability that agents i and j are paired and that agent i mistakenly chooses action 1, thus recording the highest possible payoff in his memory. Applying a similar reasoning to the one developed in the proof of Lemma 1, we have left $B(\Theta)$. Hence, $R(\Theta) = 1$.
End of Proof.

Then, under Assumption 6:

Lemma 3 *Consider an arbitrary information network and a matching network satisfying Assumptions 1, 2 and 6. All agents use IM. Then, for any absorbing set* $\Theta \neq \Theta^0, \Theta^1$

$$R(\Theta) = \begin{cases} 2 & \textit{if } s \geq 2; \\ 1 & \textit{if } s = 1. \end{cases}$$

Proof: Since $\Theta \neq \Theta^1, \Theta^0$, there exist $C_k \in \mathcal{S}$ and $C_l \in \mathcal{S}$ such that $\theta_{C_k} = \theta^0$ and $\theta_{C_l} = \theta^1$. From Assumption 6, agents in C_k [resp. C_l] cannot be matched with agents from a SCC which follows θ^1 [resp. θ^0]. To leave $B(\Theta)$, one agent from C_k and the agent is matched with can make two simultaneous mistakes. Or, an agent in C_l can make s successive mistakes. Using the arguments presented in the proof of Lemma 1, we can see that the value of the radius of Θ depends then on the sample size. It is 2 if $s \geq 2$, 1 otherwise.
End of Proof.

A couple of remarks regarding the possible candidates for stochastic stability are in order. From Lemma 1, we can see that if $s = 1$, the unit value of $R(\Theta^1)$ makes it impossible for Θ^1 to satisfy the condition $R(\Theta) > CR(\Theta)$. Under Assumption 5, the same conclusion arises from Lemma 2 for any s and any $\Theta \neq \Theta^0, \Theta^1$. Finally, Lemma 3 shows that if $s = 1$, no stochastically stable sets can be identified under Assumption 6.

Although we could have narrowed down our focus to a reduced number of candidates for stochastic stability, we chose to present here the general results for the computation of coradii.

Denote by $Card(\mathcal{S})$ the total number of source SCCs and by $0 \leq K \leq Card(\mathcal{S})$ the number of source SCCs which follow θ^1 under Θ. By definition, K is also the number of source SCCs which follow θ^0 under $\overline{\Theta}$.

Lemma 4 *Consider an arbitrary information network and a matching network satisfying Assumptions 1 and 2. All agents use IM. Then, for any absorbing set* Θ

$$CR(\Theta) = \begin{cases} sCard(\mathcal{S}) + K(1-s) & \text{if Assumption 5 is satisfied;} \\ sCard(\mathcal{S}) + K(2-s) & \text{if Assumption 6 is satisfied.} \end{cases}$$

Proof: Consider Θ and denote K the number of source SCCs which follow θ^1 under Θ. The number of source SCCs which follow the local convention θ^0 under Θ is therefore equal to $Card(\mathcal{S}) - K$. We need to compute the umber of mistakes needed to go from $\overline{\Theta}$ to Θ. This means that $Card(\mathcal{S}) - K$ source SCCs need to switch convention from θ^1 to θ^0, and K source SCC should go from θ^1 to θ^0.

When Assumption 5 is satisfied, only one mistake per source SCC following Θ^0 is needed to switch local convention, and s mistakes are needed per source SCCs following Θ^1 to have them also switch their local convention. Hence, $CR(\Theta) = s(Card(\mathcal{S}) - K) + K = sCard(\mathcal{S}) + K(1 - s)$ if Assumption 6 is satisfied.

With Assumption 6, since source SCCs following θ^0 cannot be matched with SCCs where agents play action 1. Hence, each of these SCCs will require two mistakes to switch local convention. Furthermore, s mistakes per source SCC following θ^1 is needed. This number is independent of the possible matchings. Therefore, $CR(\Theta) = s(Card(\mathcal{S}) - K) + 2K = sCard(\mathcal{S}) + K(2 - s)$ if Assumption 6 is satisfied.
End of Proof.

Note that in particular, the computation of $CR(\Theta^0)$ is invariant to matching changes.

The following result is probably our most general result, in terms of networks, as it does not require further assumptions than Assumption 1 and 2. It states that when the sample size is exactly 1, and the information network presents only one SCC, the risk dominant convention is stochastically stable.

Proposition 3 *Consider an arbitrary information network and a matching network satisfying Assumptions 1 and 2. All agents use IM. If* $Card(\mathcal{S}) = s = 1$ *and* $s \leq m/2$, *then* Θ^0 *is stochastically stable.*

Proof: From Lemma 1, $R(\Theta^0) = 2$. Furthermore, If $Card(\mathcal{S}) = 1$ and $s = 1$ then $CR(\Theta^0) = 1$, thus $R(\Theta^0) > CR(\Theta^0)$.
End of Proof.

Proposition 3 corroborates the statement made by Alós-Ferrer and Weidenholzer [1] who acknowledge that their efficiency result relies heavily on the fact that agents always observe the information of agents they interact with.

As the coradius of any absorbing set other than Θ^0 depends on the matching network, we now consider the stochastic stability of the system under Assumption 5 and Assumption 6 respectively. First, with Assumption 5, we already noted that the only absorbing set worth looking at, other than Θ^0, is Θ^1. This is because the radius of any absorbing set other than the two uniform conventions is equal to one. Hence the value of their coradius cannot be lower. Given some conditions on the sample size, it turns out that Θ^1 is stochastic stable when matchings between agents from different hubs are allowed.

Proposition 4 *Consider an arbitrary information network and a matching network satisfying Assumptions 1, 2 and 5. All agents use IM. If* $Card(\mathcal{S}) < s < m/2$, *then* Θ^1 *is stochastically stable.*

Proof: From Lemma 1, $R(\Theta^1) = s$. Furthermore, from Lemma 4, under Assumption 5, $CR(\Theta^1) = Card(\mathcal{S})$. Hence, if $s > Card(\mathcal{S}) = 1$ then $R(\Theta^1) > CR(\Theta^1)$.
End of Proof.

In Proposition 4, the matching between agents from different information hubs makes it easier for SCCs to switch local conventions. This leads to a decrease in coradius, but it also leads to a decrease in the radii of absorbing sets that are not uniform conventions, thus precluding them from being stochastically stable. The stochastic stability of the uniform efficient convention is confirmed when Assumption 5 is replaced with Assumption 6. However, the requirements regarding the sample size differ and the allowed range becomes narrower in most cases.

Proposition 5 *Consider an information network and a matching network satisfying Assumptions 1, 2 and 6. All agents use IM. If* $2Card(\mathcal{S}) < s < m/2$, *then* Θ^1 *is stochastically stable.*

Proof: From Lemma 1, $R(\Theta^1) = s$. Furthermore, from lemma 4, under Assumption 6, $CR(\Theta^1) = 2Card(\mathcal{S})$. Hence, if $s > 2Card(\mathcal{S}) = 1$ then $R(\Theta^1) > CR(\Theta^1)$.
End of Proof.

When agents imitate, and given some conditions on the sample size, the Pareto efficient convention is stochastically stable, whether matchings between information hubs are allowed or not, as shown in Proposition 5. Although these results tend to confirm the conclusions made by Alós-Ferrer and Weidenholzer [1], they also show the sensitivity of the requirements regarding the sample size to the matching network. Broader matchings, i.e matchings between agents who belong to different information hubs, widen the sample size's possible range, thus making the stochastic stability of the efficient uniform convention more likely.

4 Conclusion

This paper models information exchanges and possible matchings between a finite number of agents using two distinct networks. We fully characterize medium run outcomes for arbitrary networks, and show that these results are independent of the matching network and the decision rule used by the agents. In other words, for medium run predictions, only information matters.

In the long run, we show that the risk dominant uniform convention prevails when agents are best repliers, and that it may also arise when agents imitate. The first case, illustrated by Proposition 2, requires the sample size to be within some bounds characterized by the memory of the agents, the coefficient of risk dominance and the number of strongly connected components. The latter case, as shown by Proposition 3, requires a sample size equal to one and an information network containing only one strongly connected component. These results are insensitive to variations of the matching network.

Finally, the uniform efficient convention is proven to be stochastically stable when agents imitate and the sample size belongs to some range which depends once again on the memory of the agents as well as the number of strongly connected components of the information network. The long run efficiency however, as stated in Proposition 4 and Proposition 5, is sensitive to the matching network. Broader matching opportunities seem to promote efficiency by expanding the range of values the sample size can take.

References

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

[2] Bergin, J., and Lipman, B. L. Evolution with state-dependent mutations. *Econometrica 64* (1996), 943–956.

[3] Dasgupta, S., C. P., and Vazirani, U. *Algorithms.* McGraw-Hill, 2006.

[4] Durieu, J., and Solal, P. Adaptative play with spatial sampling. *Games and Economic Behavior 43* (2003), 189–195.

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

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

[7] Eshel I., L. S., and Shaked, A. Altruists, egoists, and hooligans in a local interaction model. *The American Economic Review 88* (1998), 157–179.

[8] Harsanyi, J., and Selten, R. *A General Theory of Equilibrium in Games*. Cambridge: MIT Press, 1988.

[9] Mengel, F. Conformism and cooperation in a local interaction model. *Journal of Evolutionary Economics 19* (2009), 397–415.

[10] Morris, S. Contagion. *The Review of Economic Studies 67* (2000), 57–78.

[11] Young, P. The evolution of conventions. *Econometrica 61* (1993), 57–84.

[12] Young, P. *Individual Strategy and Social Structure*. Princeton University Press, 1998.

5 Appendix

5.1 Proof Proposition 1

We define R_k, the set of SCCs containing agents who are neighbors to some agents in C_k, as follows:

$$R_k = \{C_l \in \mathcal{C} \setminus C_k \text{ s.t } \exists\, i \in C_l \text{ and } \exists\, j \in C_k \text{ with } g_{ij} = 1\}.$$

Without loss of generality, let all source SCCs be of order 1. Denote O_{C_k} the order of $C_k \in \mathcal{C}$ and define it as:

$$O_{C_k} = \max_{C_l \in R_k} O_{C_l} + 1$$

This creates a partial order on the SCCs of the information network. This partial order is used to describe how a given action propagates from one SCC to another. In what follows, we first show that the states described in Proposition 1 are either absorbing or part of an absorbing set. We then demonstrate that these absorbing states are reached with positive probability from any arbitrary state.

First, consider $C_k \in \mathcal{S}$. By definition, $\nexists\, i \in C_k$ and $\nexists\, j \in I \setminus C_k$ such that $g_{ji} = 1$. Hence if $\theta_{C_k} = \theta^p$ with p either 0 or 1, then agents in C_k can only sample action p, thus maintaining the status quo. Therefore, when a source SCC follows a local convention, it never leaves it.

Consider now $C_k \in \mathcal{C}$, with $O_{C_k} = 2$. By definition, $\exists\, i \in R_k$ and $\exists\, j \in C_k$ such that $g_{ij} = 1$. This means that some agents in C_k can sample the actions of some agents in R_k. If $\theta_{C_l} = \theta^p$ for all $C_l \in R_k$, then agents in C_k can only sample action p outside of C_k. Since $\theta_{C_k} = \theta^p$ as per condition 2 of Proposition 1, no agents in C_k can sample an action different from p and hence, agents keep playing the same action. Applying this reasoning recursively to all SCCs of order greater than 2, one can see that if condition 2 is satisfied, then no agents will change their action.

Therefore, when both conditions are satisfied, the system is in an absorbing state or set. Note that nothing is mentioned in Proposition 1 when $\exists\, C_l \in R_k$ and $\exists\, C_m \in R_k$ such that $\theta_{C_l} \neq \theta_{C_m}$. In this case, agents in C_k may sample different actions outside of C_k and hence C_k may not follow a local convention. This can happen in an absorbing set.

The reasoning presented above holds whether all agents play BR or IM. We now need to show that these absorbing sets can be reached with positive probability from any arbitrary state.

Consider an arbitrary state Θ and without loss of generality let $C_1 \in \mathcal{S}$. Pick agent i in C_1. With positive probability, since $s \leq m/2$, agent i can sample the same set of actions for s periods. Agent i thus plays a unique strategy p^* for his s most recent plays. From here, it is possible that agent i has a memory uniquely composed of p^*, as he can sample from himself. If agent i is the only agent in C_1, then C_1 follows a local convention.

If $card(C_1) > 1$, let $A_i = \{j \in C_1 \ s.t \ g_{ij} = 1\}$. We apply the same reasoning as the one presented above to all agents in $A_i \cup C_1$. We then do so recursively to all agents whose neighbors belong to $A_i \cup C_1$ and so on. Thus, C_1 can follow a local convention. Whether $\theta_{C_1} = \theta^1$ or $\theta_{C_1} = \theta^0$ depends on agent i's initial sample.

By applying a similar argument to all $C_k \in \mathcal{S}$, we show that there is a positive probability that from any arbitrary local state, all source SCCs can follow a local convention. Once this is the case, we know that the local state within each of the source SCCs will not change.

Consider now $C_k \in \mathcal{C} \backslash \mathcal{S}$, with $O_{C_k} = 2$. By definition, $\exists \ i \in C_k$ and $\exists \ j \in L_k$ such that agent i can sample the same action p for m periods from an agent in one of the source SCCs. In which case, the memory of this agent in C_k who samples information from outside C_k contains only action p. Following the reasoning described above, all agents in C_k can then have a memory containing only action p. Furthermore, if all $C_l \in L_{C_k}$ follow the same local convention θ^p, then $\theta_{C_k} = \theta^p$ for all periods after that.

To complete the proof, the steps described above are repeated to SCCs of order 3, then to SCCs of order 4, and so on. This shows that the only absorbing sets are those described in Proposition 1.

5.2 Proof Proposition 2

Consider $C_k \in \mathcal{S}$ with $\theta_{C_k} = \theta^1$. With positive probability, one agent within C_k, agent i, can make $\lceil (1-q)s \rceil$ successive mistakes. All agents in $N_i \cap C_k$ can then sample the string of mistakes and play 0 for s periods. This is enough to create a memory containing only action 0 for all agents in $N_i \cap C_k$. The same reasoning can be applied to the neighbors of these agents, and the neighbors of the neighbors of these agents within C_k until $\theta_{C_k} = \theta^0$.

A similar argument applies to $C_k \in \mathcal{S}$ with $\theta_{C_k} = \theta^0$. In this case, $\lceil qs \rceil$ mistakes are needed to switch from θ^0 to θ^1.

Since we consider games where $q \geq 1/2$, we have $\lceil qs \rceil \geq \lceil (1-q)s \rceil$. Hence,

$$R(\Theta) = \begin{cases} \lceil (1-q)s \rceil & \text{if } \Theta \neq \Theta^0; \\ \lceil qs \rceil & \text{if } \Theta = \Theta^0. \end{cases}$$

Consider Θ and denote K the number of source SCCs which follow θ^1 under Θ. The number of source SCCs which follow the local convention θ^0 is therefore equal to $Card(\mathcal{S}) - K$. Thus, under $\overline{\Theta}$, K source SCCs follow θ^0, and $Card(\mathcal{S}) - K$ source SCCs follow Θ^1. Using the above arguments, we find that:

$$CR(\Theta) = K \lceil qs \rceil + (Card(\mathcal{S}) - K) \lceil (1-q)s \rceil$$

Finally, the condition $R(\Theta) > CR(\Theta)$ is satisfied when $\lceil (1-q)s \rceil > K \lceil qs \rceil + (Card(\mathcal{S}) - K) \lceil (1-q)s \rceil$ for $\Theta \neq \Theta^0$, and $\lceil qs \rceil > K \lceil qs \rceil + (Card(\mathcal{S}) - K) \lceil (1-q)s \rceil$ for $\Theta = \Theta^0$.

But since $\frac{\lceil qs \rceil}{\lceil (1-q)s \rceil} > 1$ and $\frac{k+1-Card(\mathcal{S})}{k} \leq 1$, no state $\Theta \neq \Theta^0$ satisfies the condition $R(\Theta) > CR(\Theta)$. However, when $\Theta = \Theta^0$ and hence $K = 0$, the condition becomes $Card(\mathcal{S}) < \frac{\lceil qs \rceil}{\lceil (1-q)s \rceil}$ and can be satisfied for some s and $Card(\mathcal{S})$.