GENERALIZED RISK-DOMINANCE AND ASYMMETRIC DYNAMICS

MARCIN PESKI

Abstract. This paper proposes two (ordinal and cardinal) generalizations of Harsanyi and Selten (1988) risk-dominance to multi-player, multi-action games. There are three reasons why generalized risk-dominance (GR-dominance) is interesting. Extending the logic of risk-dominance, GR-dominant actions can be interpreted as best responses to conjectures that satisfy a certain type of symmetry. Second, in a local interaction game of Ellison (1993), if an action is risk-dominant in individual binary interactions with neighbors, it is also GR-dominant in the large game on a network. Finally, we show that GR-dominant actions are stochastically stable under a class of evolutionary dynamics. The last observation is a corollary to new abstract selection results that applies to a wide class of so-called *asymmetric dynamics*. In particular, I show that a (strictly) ordinal GR-dominant profile is (uniquely) stochastically stable under the approximate best-response dynamics of Kandori, Mailath, and Rob (1993). A (strictly) cardinal GR-dominant equilibrium is (uniquely) stochastically stable under a class of payoff-based dynamics that includes Blume (1993). Among others, this leads to a generalization of a result from Ellison (2000) on the $\frac{1}{2}$-dominant evolutionary selection to all networks and the unique selection to all networks that satisfy a simple, sufficient condition.

1. Introduction

There is a large literature that is concerned with selecting a unique equilibrium in games with with multiple equilibria. Harsanyi and Selten (1988) proposed risk-dominance as an selection criterion that is appropriate for games with two players and two actions. It turned out soon that risk-dominance plays an important role in two different models of equilibrium selection: robustness to incomplete information (Carlsson and Damme (1993) and Kajii and Morris (1997)) and evolutionary learning (Kandori, Mailath, and Rob (1993) and Young (1993)). The connection between the two models can be easily attributed to the relation between two equivalent statement of risk-dominance. In a two-action coordination game,

Date: June 2007.

Key words and phrases. JEL C73, Evolutionary theory, stochastic games.

I am grateful to Eddie Dekel, Jeff Ely, Stephen Morris, the editor and two anonymous referees for all helpful comments. Priscilla Tien Yuek Man provided excellent research assistance. All remaining errors are mine.

Department of Economics, University of Chicago, 1126 E. 59th St., Chicago, IL 60637; Department of Economics, University of Texas at Austin, 1 University Station #C3100, Austin, Texas 78712.

the risk-dominant action is a best response to the conjecture that assigns equal probability to each of the opponent actions. Alternatively, consider a population game in which players are randomly and uniformly matched in pairs to play two-action coordination game. The risk-dominant action is a best response to any population profile in which each action is played by exactly half of the population.

This paper proposes two generalizations of risk-dominance from two-player to multi-player and multi-action games. The first generalization depends on the ordinal, and the second on the cardinal properties of the payoff function. We motivate the generalizations on two levels. First, we show that they preserve the logic of risk-dominance in richer environments. Second, we show that the generalizations are stochastically stable under a class of evolutionary dynamics. These results are corollaries to an abstract selection result about asymmetric dynamics.

1.1. **Generalized risk-dominance.** We present two versions, an ordinal and a cardinal, of a generalized risk-dominance (GR-dominance). Suppose that there are $I \geq 2$ players, and each player i chooses an action $x_i \in A_i$. Fix profile $\mathbf{a} = (a_i) \in \times_i A_i$. Say that two (pure strategy) profiles $\eta, \bar{\eta} \in \times_i A_i$ are $\mathbf{a}$-*associated*, if all players i who do not play a_i in profile η, play a_i in profile $\bar{\eta}$: for all i, either $\eta_i = a_i$ or $\bar{\eta}_i = a_i$. Say that profile $\mathbf{a}$ is *ordinal GR-dominant*, if whenever action a_i is not a best response against profile η, it is a best response against any $\mathbf{a}$-associated profile $\bar{\eta}$:

$$\begin{gathered}\text{For all } \mathbf{a}\text{-associated profiles } \eta \text{ and } \bar{\eta}, \text{ for all players } i, \\ \text{action } a_i \text{ is a best response to either } \eta \text{ or } \bar{\eta}.\end{gathered} \tag{1.1}$$

Let $u_i(x_i, \eta)$ be the payoff of player i when she plays x_i and the other players follow profile η. The definition of ordinal $GR-$dominance depends only on the best-response behavior of the payoff function. Say that profile $\mathbf{a}$ is *cardinal* $GR-$*dominant* if the utility loss from playing a_i rather than the best action $x_i \neq a_i$ to profile η is not higher than the gain from playing a_i rather than the best action $x_i \neq a_i$ to $\mathbf{a}$-associated profile $\bar{\eta}$:

$$\begin{gathered}\text{For all } \mathbf{a}\text{-associated profiles } \eta \text{ and } \bar{\eta}, \text{ for all players } i, \\ \max_{x_i \neq a_i} u_i(x_i, \eta) - u_i(a_i, \eta) \leq u_i(a_i, \bar{\eta}) - \max_{x_i \neq a_i} u_i(x_i, \bar{\eta}).\end{gathered} \tag{1.2}$$

These definitions have appropriate *strict* versions.

Any cardinal or ordinal $GR-$dominant profile is a pure-strategy equilibrium. With two players and two actions, a profile is cardinal $GR-$dominant if and only if it is risk-dominant. Because Harsanyi and Selten (1988)'s definition depends on cardinal properties of payoffs, there is no immediate relation between risk-dominance and ordinal GR-dominance in two-player games. For multi-player games, Morris, Rob, and Shin (1995) define a $(p_1, ..., p_I)$-dominant equilibrium, in which the action of player i is the best response to any conjecture that assigns a probability of at least p_i to the equilibrium action profile. We show in Section 4 that if there are only two actions for each player, then cardinal $GR-$dominance implies $\left(\frac{1}{2}, ..., \frac{1}{2}\right)$-dominance. There is no further logical relationship between $GR-$dominance and $\frac{1}{2}$-dominance in general multi-player games.[1]

1.2. **Motivation.** We argue that $GR-$dominance extends the logic of risk-dominance to multi-player games. It is useful to distinguish among two statements of risk-dominance: belief-based and population-based.[2] According to the belief-based version, profile $\mathbf{a} = (a_1, a_2)$ in a two-player, two-action game is risk-dominant, if, for each player i, action a_i is a best response to any conjecture that assigns at least $\frac{1}{2}$-probability to the opponent playing a_{-i}.[3] In coordination games, it is enough to require that a_i is a best response to any conjecture that is symmetric with respect to (i.e., which assigns equal probability to) actions $A_{-i} = \{a_{-i}, b_{-i}\}$.

Cardinal $GR-$dominant actions are best responses to conjectures that satisfy an analogous type of symmetry in multi-player games. Suppose that each player chooses between two actions, $A_i = \{a_i, b_i\}$ and that actions a_i are complementary, i.e., if a_i is a best response, it remains the best response when more players switch to $\mathbf{a}$. For each profile η, find the unique profile $-\eta$ in which actions of all players are flipped. One can think about $-\eta$ as obtained from η by relabeling all the actions as into bs and vice versa. Say that player i's conjecture about other players' behavior is *symmetric with respect to labels*, if the conjectured probability of profile η is equal to the conjectured probability of profile $-\eta$ for all η. Players with symmetric conjectures are reluctant to assume that actions with different labels are treated differently by other players. We show that $\mathbf{a}$ is cardinally $GR-$dominant if and only if a is a best response of each player to any symmetric conjecture (Lemma 5).

Alternatively, consider the population-based interpretation of risk-dominance. Suppose each player is matched to play a two-player two-action interaction game with an opponent chosen *randomly and uniformly* from a large population (this is the model analyzed in Kandori, Mailath, and Rob (1993)). Profile (a_i, a_j) is risk-dominant in the two-player interaction

[1] In particular, there is no relationship between $\left(\frac{1}{2}, ..., \frac{1}{2}\right)$-dominance and *ordinal* flip-dominance. This is not surprising given the fact that the former is defined with respect to cardinal properties of the payoff function, and the latter is a purely ordinal concept.

[2] I am grateful to the editior for this suggestion.

[3] When there are two players, but more than two actions, the above definition has been introduced in Morris, Rob, and Shin (1995) as $\frac{1}{2}$-dominance. The following discussion applies unchanged if the interaction game π has more than two actions.

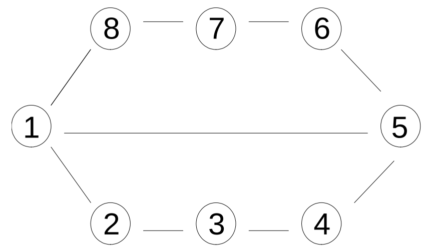

FIGURE 1. Example of a network

game between players i and j, if and only if action a_i is a best response to *any* population profile in which a majority of players chooses a^*.

In more general games, players are not matched according to uniform distribution. For example, in a model of local interactions introduced by Ellison (1993), there are I players located on a network. Player i's payoffs are equal to the sum of payoffs in the interactions with his neighbors,

$$u_i^{\text{local}}(x_i, \eta) := \sum_{j \neq i} g_{ij} \pi(x_i, \eta_j), \tag{1.3}$$

where $g_{ij} \in \{0, 1\}$ denotes the existence of a connection between players i and j and π is a payoff function in the two-player interaction. (1.3) generalizes the global interaction model of Kandori, Mailath, and Rob (1993) where all players are connected, $g_{ij} = 1$ for each i and j.

Consider a network in Figure 1. Player 1 has three neighbors: players 2, 5, and 8. Suppose that action a is best response in the two-player interaction if the opponent plays a with a probability at least $\frac{2}{5}$; b is a best response when the opponent plays a with a probability not more than $\frac{2}{5}$. Then, there are profiles where the majority of players play a, but a is *not* player i's best response. For example, if all players but player 2 and 5 play a, b is player 1's best response even though more than half of players play a. On the other hand, if all players flip their actions from a to b and vice versa, then action a becomes player 1's best response. More generally, Lemmas 4 and 6 show that the profile of actions that are risk-dominant in the two-player interaction π is (cardinal and ordinal) $GR-$dominant in any multi-player game (1.3) on any network. The profiles are strictly ordinal $GR-$dominant, if the network satisfies a simple sufficient condition on the number of neighbors of each player and strictly cardinal dominant if the actions satisfy a version of strict risk-dominance in the two-player interaction.[4]

[4]One could work with an alternative extension of risk-dominance. Say that the profile of actions $\mathbf{a}^*$ is *multi-player risk-dominant*, if action a_i is a best response against any profile of actions in which at least half of the players play a. This is a stronger definition than generalized risk-dominance. The example shows that

1.3. **Stochastic stability.** The main results of the paper extend and generalize various stochastic stability results that the evolutionary literature has traditionally associated with risk-dominance. We discuss these results as corollaries to an abstract selection result described in section 2. We identify a simple sufficient condition for stochastic stability, called *asymmetry of dynamics*. Asymmetry is a condition of its own interest, as it is likely to be satisfied by other dynamics and other generalizations of risk-dominance. Theorem 1 shows that if the dynamics are asymmetric toward profile **a** then profile **a** is stochastically stable. Uniqueness is implied by appropriate versions of asymmetry: robustness and strictness. Theorem 2 shows that the evolutionary dynamics of Kandori, Mailath, and Rob (1993) are (robustly) asymmetric toward any (strictly) ordinal $GR-$dominant profile. Blume (1993) studied an alternative evolutionary dynamics in which the cost of transition is linear in the difference between the payoff from a given action and the best response payoff. Theorem 3 shows that Blume's and related dynamics are (strictly) asymmetric toward any (strictly) cardinal $GR-$dominant profile. Figure 2 presents logical connections between the results in the paper.

Recall that Ellison (1993) and Ellison (2000) establish the stochastic stability of profile **a** of risk-dominant actions in model (1.3), if the network has a particular shape like a circle or torus (see, for example, Lee, Szeidl, and Valentinyi (2003) or Blume and Temzelides (2003) for generalizations of Ellison's argument). The results from Figure 2 imply that **a** is stochastically stable on *all* networks, and uniquely so on networks that satisfy a simple condition on the number of players.

Blume (1993) shows if the interaction game is symmetric and has two actions, then the profile of strictly risk-dominant actions is uniquely stochastically stable on a two-dimensional lattice. This is further in Young (1998) to all local interaction models (1.3)). Blume (1993) and Young (1998) rely on the fact that two-action symmetric coordination games have a potential function. Because of Figure 2, the results from Blume (1993) and Young (1998) can be generalized to, among others, games with multiple actions and games without potential.

Since Rubinstein (1989), Carlsson and Damme (1993), and Morris, Rob, and Shin (1995), it is known that risk- or, more generally, p-dominant outcomes are robust to incomplete information. It has been argued that the connection between the two models is not accidental. For example, Morris (2000) shows that the arguments behind evolutionary selection on some networks closely resemble contagion arguments used in the robustness to incomplete information literature. At this moment, it remains unknown whether a similar relationship

---

such a definition is too strong. In particular, is a is risk-dominant, but not a best response when more than $\frac{3}{5}$ neighbors plays b, then a is not multi-player risk-dominant.

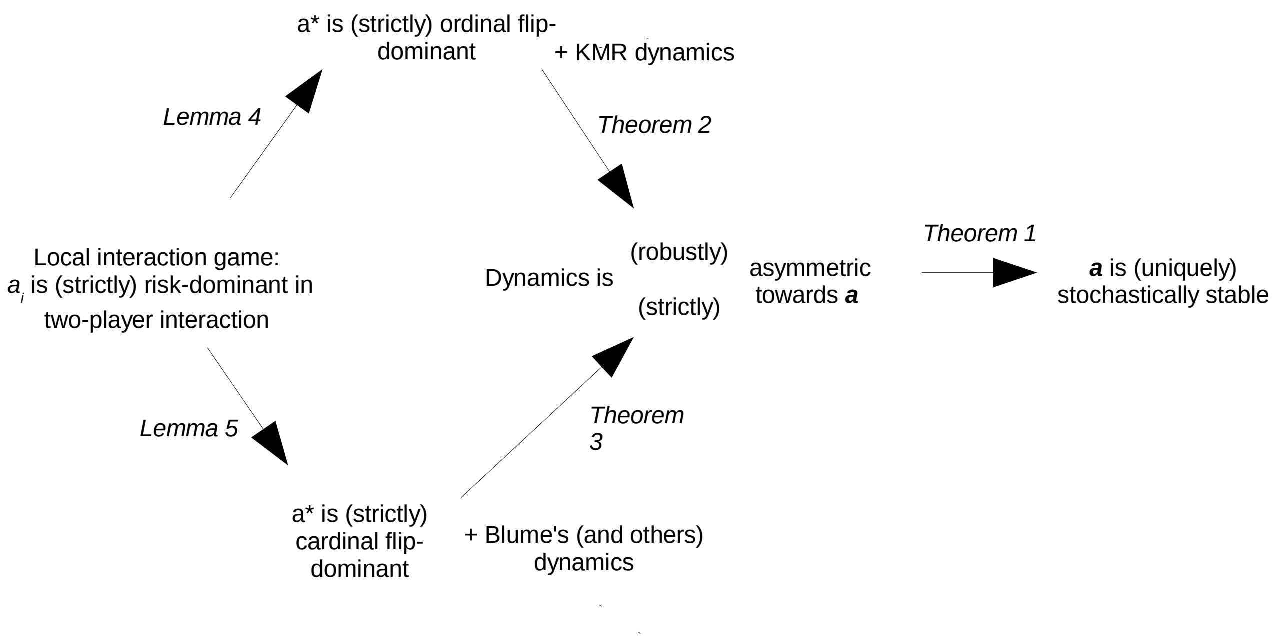

FIGURE 2. Connections between the results of the paper

exists between $GR-$dominance and incomplete information games. The answer to this question would shed light on the connections between the two models. In particular, a negative answer would mean that the connection established in the previous literature is restricted to very special classes of games, such as games on networks with a specific network structure. A positive answer may lead to new results in the global games literature.

2. Asymmetric dynamics

This section identifies an abstract property of evolutionary dynamics, *asymmetry*, that is responsible for the evolutionary selection results discussed later in the paper.

2.1. **Model.** There are $I \geq 2$ players. Each player i chooses an action η_i from a finite set A_i. A state of the population is represented as an action profile $\eta \in \Sigma = \times_i A_i$, and η_{-i} denotes the actions of all players but i.

We describe abstract evolutionary dynamics on the space state Σ following Ellison (2000). The state of the population evolves according to the Markov process P_ε, where $P_\varepsilon(\eta, \eta') \geq 0$ denotes the probability of transition from η to η' and for each η, ε, $\sum_{\eta'} P_\varepsilon(\eta, \eta') = 1$. The parameter $\varepsilon > 0$ is interpreted as the probability of an individual mutation. We assume that limits

$$\lim_{\varepsilon \to 0} \frac{\log P_\varepsilon(\eta, \eta')}{\log \varepsilon} =: c(\eta, \eta') \in [0, \infty] \tag{2.1}$$

exist (note that they might be equal to ∞). Function $c : \Sigma \times \Sigma \to [0, \infty]$ is called a *cost function* if it satisfies standard assumptions: For any η, η' (a) $c(\eta, \eta) = 0$, and (b) there exists a path $\eta = \eta_0, ..., \eta_t = \eta'$ such that $c(\eta_s, \eta_{s+1}) < \infty$ for any $s < t$. Assumption (a) is without a loss of generality, and assumption (b) ensures that for each $\varepsilon > 0$, P_ε is ergodic and there are unique ergodic distributions $\mu_\varepsilon \in \Delta\Sigma$ st. for each η, $\mu_\varepsilon(\eta) = \sum_{\eta'} \mu_\varepsilon(\eta') P_\varepsilon(\eta', \eta)$.

Profile η is *stochastically stable* if $\lim_{\varepsilon \to 0} \mu_\varepsilon(\eta) > 0$. It is *uniquely stochastically stable*, if $\lim_{\varepsilon \to 0} \mu_\varepsilon(\eta) = 1$.

2.2. **Asymmetric dynamics.** Fix profile $\mathbf{a} = (a_i) \in \Sigma$. Profile η is $\mathbf{a}$-*dominated by* η' if for each player i, $\eta_i = a_i \Rightarrow \eta'_i = a_i$. Profile η is *almost* $\mathbf{a}$-*dominated by* η' if there is a player j such that for each $i \neq j$, $\eta_i = a_i \Rightarrow \eta'_i = a_i$. Profiles $\eta, \bar{\eta}$ are $\mathbf{a}-$*associated* if for all players i, either $\eta_i = a_i$ or $\bar{\eta}_i = a_i$. Profiles η and $\bar{\eta}$ are *almost* $\mathbf{a}-$*associated* if they are $\mathbf{a}$-associated or there is player j, such that either $\eta_i = a_i$ or $\bar{\eta}_i = a_i$ for any player $i \neq j$.

To parse these definitions, consider the binary case $A_i = \{0, 1\}$ and $a_i = 1$ for each i. The set $\Sigma = \times_i A_i$ is a lattice with a natural partial order $\leq$. Profile η is $\mathbf{a}$-dominated by η' if and only if $\eta \leq \eta'$. Two profiles η, η' are $\mathbf{a}$-associated if and only if $\eta_i + \eta'_i \geq 1$ for each i. Figure 3 presents an example of such a lattice together with a cost function. Profile $\mathbf{a} = 111$ $\mathbf{a}-$dominates any other profile and is $\mathbf{a}-$associated with any other profile. Profile 011 is $\mathbf{a}$-associated with profile 100 and any other profile that $\mathbf{a}$-dominates 100. Lines between states correspond to transitions. The cost of transition upwards (downwards) between neighboring profiles are written at the top left (bottom right) of a corresponding line. For example, the cost of transition from profile 111 to profile 011 is equal to $c(111, 011) = 5$. The costs of transitions that are not denoted in the Figure are defined as the sums of costs of the lowest cost path between profiles. For example, $c(111, 010) = c(111, 011) + c(011, 010) = 8$.

Definition 1. *Cost function c is* asymmetric toward $\mathbf{a}$, *if for any profiles $\eta, \eta', \bar{\eta}$, such that $\eta, \bar{\eta}$ are $\mathbf{a}$-associated, there exists $\bar{\eta}'$ such that $\bar{\eta}$ is $\mathbf{a}$-dominated by $\bar{\eta}'$, $\bar{\eta}'$ is $\mathbf{a}$-associated with η', and*

$$c(\bar{\eta}, \bar{\eta}') \leq c(\eta, \eta'). \tag{2.2}$$

This definition has a simple intuition. Very informally, if the cost function is asymmetric, then for any transition away from profile $\mathbf{a}$, there is an "associated" transition toward profile $\mathbf{a}$ with at most the same cost. The cost function drawn in Figure 3 is asymmetric. For example, notice that $c(000, 100) = 4 \leq 5 = c(111, 011)$.

Theorem 1 below states that if the cost function is asymmetric toward $\mathbf{a}$, then profile $\mathbf{a}$ is stochastically stable. Unique stochastic stability requires additional conditions. Two variations of the above definition are useful. In what follows, we adopt the convention that $\infty < \infty$.

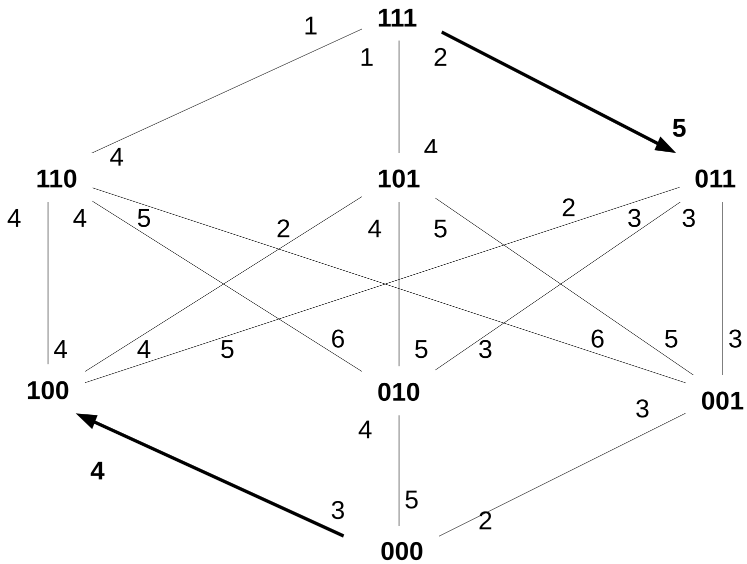

Figure 3. Example of a cost function. The costs of upwards (downwards) transitions are written at the top left (bottom right) of the corresponding lines.

Definition 2. *Cost function c is* strictly asymmetric toward $\mathbf{a}$ *if*

(1) *for any $\eta \neq \mathbf{a}$, $c(\mathbf{a}, \eta) > 0$, and*
(2) *for any profiles $\eta, \eta', \bar{\eta}$, such that η and $\bar{\eta}$ are $\mathbf{a}$-associated, there exists $\bar{\eta}'$ such that $\bar{\eta}$ is $\mathbf{a}$-dominated by $\bar{\eta}'$, $\bar{\eta}'$ is $\mathbf{a}$-associated with η', and*

$$\text{either } c(\bar{\eta}, \bar{\eta}') = 0 \text{ or } c(\bar{\eta}, \bar{\eta}') < c(\eta, \eta').$$

Strict asymmetry is more restrictive than asymmetry. Strict asymmetry requires that inequality (2.2) is strict whenever $c(\bar{\eta}, \bar{\eta}') > 0$. Additionally, strict asymmetry requires that the transitions out of profile $\mathbf{a}$ are costly. We demonstrate below that the cost function from the dynamics by Blume (1993) is robustly asymmetric toward strictly cardinal $GR-$dominant profiles.

The cost function from Figure 3 is strictly asymmetric. In fact, one cannot lower the cost of any transition downwards by 1 or more without violating the strict asymmetry. For example, if the cost of the transition from 111 to 011 changed from 5 to 4, this would violate the inequality $c(000, 100) < c(111, 011)$.

Definition 3. *Cost function c is* robustly asymmetric toward $\mathbf{a}$ *if*

(1) *for any $\eta \neq \mathbf{a}$, $c(\mathbf{a}, \eta) > 0$,*

(2) *for any profiles* $\eta, \eta', \bar{\eta}$, *such that* $\eta, \bar{\eta}$ *are almost* $\mathbf{a}$*-associated, there exists* $\bar{\eta}'$ *such that* $\bar{\eta}$ *is* $\mathbf{a}$*-dominated by* $\bar{\eta}'$, $\bar{\eta}'$ *is almost* $\mathbf{a}$*-associated with* η', *and*

$$c(\bar{\eta}, \bar{\eta}') \leq c(\eta, \eta'),$$

(3) *for any profiles* η, η' *such that* $c(\eta, \eta') > 0$, *if* $\eta, \bar{\eta}$ *are* $\mathbf{a}$*-associated, then there exists* $\bar{\eta}'$ *such that* $\bar{\eta}'$ *is* $\mathbf{a}$*-dominated by* $\bar{\eta}$, $\bar{\eta}'$ *is almost* $\mathbf{a}$*-associated with* η', *and*

$$c(\bar{\eta}, \bar{\eta}') < c(\eta, \eta'),$$

(4) *for any profiles* $\eta, \eta', \bar{\eta}$ *such that* η' *is almost* $\mathbf{a}$*-dominated by* η *and* $\bar{\eta}$ *is* $\mathbf{a}$*-associated with* η', *either* $c(\eta, \eta') = 0$ *or there is* $\bar{\eta}'$ *that is* $\mathbf{a}$*-associated with* η, $\bar{\eta}$ *is* $\mathbf{a}$*-dominated by* $\bar{\eta}'$, *and* $c(\bar{\eta}, \bar{\eta}') = 0$.

Robust asymmetry is more restrictive than asymmetry. We demonstrate below that the cost function from the dynamics by Kandori, Mailath, and Rob (1993) is robustly asymmetric toward strictly ordinal $GR-$dominant profiles.

The main theorem of this section characterizes stochastic stability under Markov dynamics with asymmetric cost functions.

Theorem 1. *If the cost function is asymmetric toward* $\mathbf{a}$, *then* $\mathbf{a}$ *is stochastically stable. If it is strictly or robustly asymmetric toward* $\mathbf{a}$, *then* $\mathbf{a}$ *is uniquely stochastically stable.*

2.3. **Proof of Theorem 1.** The rest of this section discusses the proof of the Theorem. The reader interested in the applications of the above result should go directly to Sections 3 and 4. There are three steps in the proof of the Theorem.

The first step is standard. We rely on the well-known tree technique of Freidlin and Wentzell (1984). A *tree* h is a function $h : \Sigma \to \Sigma$ with a distinguished element $\eta_h \in \Sigma$ called a *root* of the tree, st. (a) $h(\eta_h) = \eta_h$ and (b) for every $\eta \neq \eta_h$ there is a path $\eta, h(\eta), ..., h^m(\eta) = \eta_h$ leading from η to η_h. Let $c : \Sigma \times \Sigma \to [0, \infty]$ be a cost function. A *cost* of tree h is equal to the sum of the costs of its branches: $c(h) = \sum_{\eta \neq \eta_h} c(\eta, h(\eta))$. A tree has minimal cost, if there is no tree with a lower cost. Denote the set of all roots of minimal cost trees as

$$MR(\Sigma, c) = \{\eta : \eta \text{ is a root of a tree } h \text{ st. for any } h', c(h) \leq c(h')\}.$$

It turns out that the stochastic stability of a profile is equivalent to whether this profile is a root of a minimal cost tree.

Lemma 1 (Freidlin and Wentzell (1984))**.** *Profile* $\mathbf{a}$ *is stochastically stable if* $\mathbf{a} \in MR(\Sigma, c)$. *Profile* $\mathbf{a}$ *is uniquely stochastically stable if* $\{\mathbf{a}\} = MR(\Sigma, c)$.

Second, we reduce the proof to the binary case. Let $S = \{0,1\}^I$ be a binary lattice with a partial order defined above. To distinguish between sets Σ and S, we refer to a typical element $v \in S$ as a *state.* Let $\mathbf{1}$ denote a state that consists only of 1s. Let $\psi : S \times S \to [0,\infty]$ be a cost function on S (i.e., it satisfies the assumptions stated above). Let $MR(S,\psi)$ be the set of all roots of minimal cost trees on S with a cost function ψ.

Definitions 1, 2, and 3 extend to cost functions ψ. An additional property of ψ is helpful. A cost function is supermodular if the transitions upwards are less costly if started from a higher state.

Definition 4. *Cost function ψ is* supermodular*, if for any states $v, v', \bar{v}$ st. $\bar{v} \geq v$, there is $\bar{v}' \geq v', \bar{v}$, such that $\psi(\bar{v}, \bar{v}') \leq \psi(v, v')$.*

The next Lemma is proven in Appendix A. It shows that the problem of finding the roots of minimal cost trees on Σ can be reduced to the analogous problem in the binary case.

Lemma 2. *Suppose that cost function $c : \Sigma \times \Sigma \to [0,\infty]$ is asymmetric (strictly asymmetric, robustly asymmetric). There exists a supermodular and asymmetric (strictly asymmetric, robustly asymmetric) cost function $\psi : S \times S \to [0,1]$, such*

$$\mathbf{1} \in MR(S,\psi) \Longrightarrow \mathbf{a} \in MR(\Sigma, c) \text{ and}$$
$$\{\mathbf{1}\} = MR(S,\psi) \Longleftrightarrow \{\mathbf{a}\} = MR(\Sigma, c).$$

The central part of the proof is contained in the following Lemma.

Lemma 3. *If cost function $\psi : S \times S \to [0,1]$ is supermodular and asymmetric, then $\mathbf{1} \in MR(S,\psi)$. Additionally, if it is strictly or robustly asymmetric, then $\{\mathbf{1}\} = MR(S,\psi)$.*

Lemma 3 identifies sufficient conditions on the cost function ψ which guarantee that state $\mathbf{1}$ is a (unique) root of minimal cost trees. Theorem 1 is a corollary to Lemmas 1, 2, and 3.

2.4. **Intuition for Lemma 3.** The proof of Lemma 3 can be found in Appendix B. Here, we use the example from Figure 3 to sketch the main ideas behind the Lemma.

All known algorithms for finding the roots of minimal cost trees are based on a similar idea (Chu and Liu (1965), Edmonds (1967)). Instead of solving the problem (S,ψ) directly, one replaces it by a simpler problem using one of two techniques:

(1) *Subtracting a constant.* One can replace ψ by $\psi^0 = \psi - \alpha$, where $\alpha := \min_{v \neq v'} \psi(v, v')$. It is easy to see that this operation does not change the set of minimal cost trees, $MR(S,\psi) = MR(S,\psi^0)$. The left graph in Figure 4 shows the cost function obtained from Figure 3 by subtracting 1.

(2) *Merging.* Consider the left graph in Figure 4. Notice that the cost of transition from states 110 and 101 to 111 is equal to 0 and each transition out of state 111 has a cost of at least 3. Any tree with a root at one of the states 110 or 101 can be replaced by a tree with a root at 111 with a total cost lower by at least 3. Thus, states 110 and 101 cannot be roots of minimal trees. One can replace the problem (S, ψ') from the left graph in Figure 4 by the problem $\left(S^1, \tilde{\psi}^1\right)$ drawn on the right-hand side of Figure 4, where $S^1 = S \backslash \{110, 101\}$. Some care is required in defining the new cost function $\tilde{\psi}^1$. For each two states $v, v' \in S'$, we define $\tilde{\psi}^1(v, v')$ as the sum of ψ^0-costs of the lowest cost path between v and v'. For example, $\tilde{\psi}^1(010, 111) = \psi^0(010, 101) + \psi^0(101, 111) = 3$. One can verify that the set of the roots of minimal cost trees in problems $\left(S^1, \tilde{\psi}^1\right)$ and (S, ψ^0) (hence, also (S, ψ)) must be equal. This is standard (see, for example, Proposition 1 of Nöldeke and Samuelson (1993) and the Appendix in Young (1993)); Lemma 15 (Appendix B) presents the formal argument.
In order to define merging formally, say that state v is an *attractor*, if for any other state v', if there is a 0-cost path from v to v', then there is a 0-cost path from v' to v.[5] For each state v, let $U(v)$ be the set of all attractors v' such that there is a 0–cost path from v to v'. As a convention, we take that if v is an attractor, then $v \in U(v)$. Clearly, the set of minimal tree roots is contained in the set of all attractors $\bigcup_{v \in S} U(v)$.
Moreover, if v is the minimal tree root, then any other attractor $v' \in U(v)$ is also a minimal tree root.
Lemma 16 shows that, if the cost function is supermodular, then for any state v, the set of attractors $U(v)$ contains its highest element: there is $\bar{v}(v) \in U(v)$ such that for each $v' \in U(v)$, v' is $\mathbf{1}$-dominated by $\bar{v}(v)$. Define $S^1 := \{\bar{v}(v) : v \in S\}$ to be the space of the highest elements in the attractor sets with the interpretation that each state v merges into state $\bar{v}(v)$.
If each attractor set contains only one element, then merging does not change the set of minimal tree roots. If there is an attractor set with more than two elements, some minimal tree roots might be lost. However, the merging guarantees that the new state space always contains at least some of the minimal tree roots of the original problem.

One can continue simplifying the problem by alternating between the two techniques. At each step, subtracting a constant creates 0-costs transitions and merging reduces the size of the state space. The algorithm stops when there is only one state remaining. Of course, the

[5] A *0-cost path* between states v and v' is a sequence of states $v = v_0, ..., v_m = v'$ such that $\psi^0(v_i v_{i+1}) = 0$ for each $i < m$.

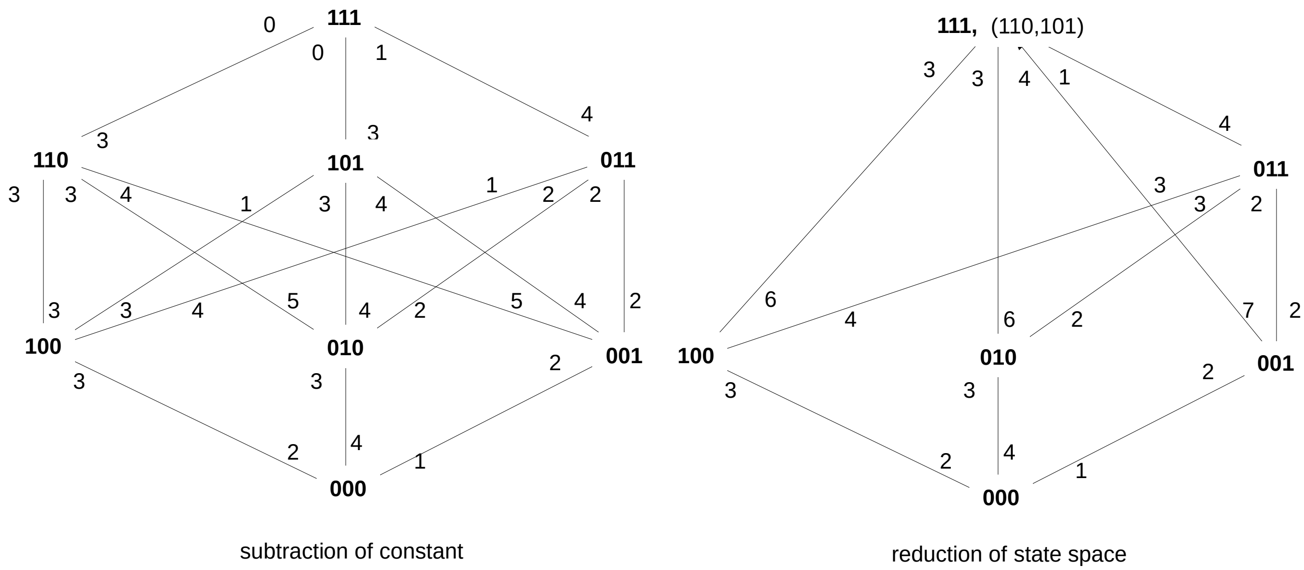

FIGURE 4. Tree techniques

remaining state must be a root of the minimal cost tree. Figure 5 presents the subsequent steps of the algorithm in the example. In this example, one can check that no roots were lost along the steps of the procedure. Because 111 is the only remaining state, it is the unique minimal tree root in the original problem (S, ψ).

The above discussion defines the algorithm. The proof of Lemma 3 shows that, if the initial cost function is supermodular and asymmetric, then state **1** survives all the steps of the algorithm; hence it is a minimal tree root of the initial problem. In addition, if the initial cost function is robustly or strictly asymmetric, then no other minimal tree roots are lost along the algorithm. The idea is based on the following observations: If the original cost function is supermodular and strictly or robustly asymmetric, then

- If the cost function at step $k \geq 1$ of the algorithm is supermodular and (strictly) asymmetric, then it remains supermodular and (strictly) asymmetric also at step $k+1$ (Lemmas 16 and 17). If the original cost function is supermodular and (strictly or robustly asymmetric), then step 1 and all the subsequent steps are supermodular and (strictly) asymmetric. Notice that the cost function from Figure 5 is strictly asymmetric at all steps. For example, consider the transition from state 001 to 000. State 111 is the unique state that is **1-**associated with 001 in state space S^1, and, $\psi^1(001, 000) = 1 > \psi^1(111, 111) = 0$ (the cost of the transition from a state to itself is always equal to 0 by definition).

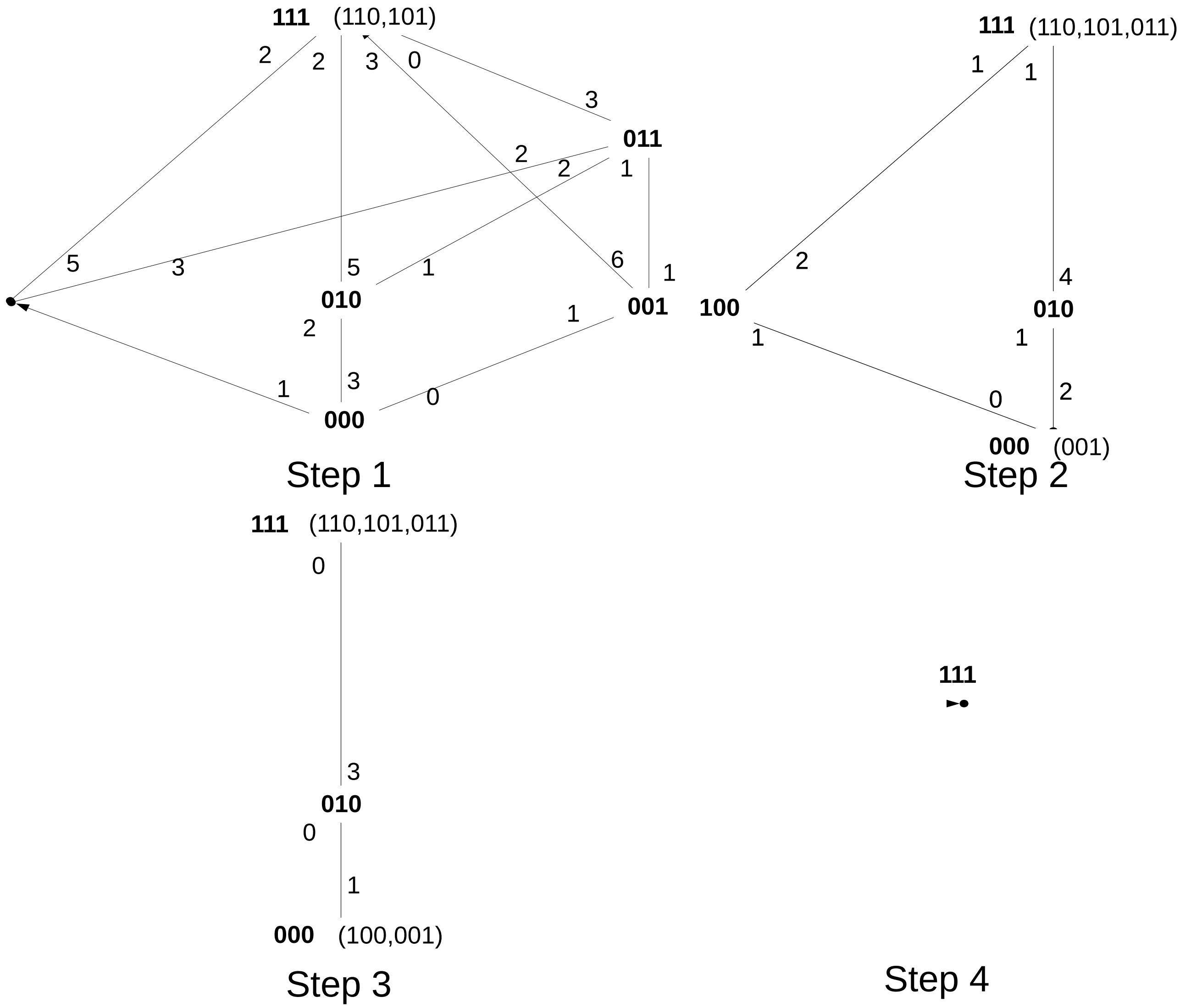

FIGURE 5. Steps of the algorithm

- If step k is supermodular and asymmetric, then state $\mathbf{1}$ is not "merged" into any other state. More precisely, if $\psi^k(\mathbf{1},\mathbf{v}) = 0$ for $v \in S^k, v \neq \mathbf{1}$, then $\psi^k(v,\mathbf{1}) = 0$ and, by the chosen convention of merging states into the upper most candidate, all states v such that $\psi^k(v,\mathbf{1}) = 0$ are merged into state $\mathbf{1}$. In particular, the only state remaining at the end of the algorithm is $\mathbf{1}$. Therefore, $\mathbf{1}$ is the only remaining state when the algorithm stops.
- In addition, if step k is strictly asymmetric, then there is no minimal tree root lost when merging other states into state $\mathbf{1}$. More precisely, if $\psi^k(v,\mathbf{1}) = 0$ for $v \in S^k, v \neq \mathbf{1}$, then $\psi^k(\mathbf{1},v) > 0$. Because this must be true along all the steps of the

algorithm, not only **1** is the only remaining state when the algorithm stops, but also **1** is the unique minimal tree root of the initial problem.

3. Ordinal generalized risk-dominance

This section discusses the ordinal extension of risk-dominance to multi-player games.

3.1. **Definition.** Let $u_i(x_i, \eta_{-i})$ denote a payoff of player i from action x_i given the profile of the opponents' actions η_{-i}. To economize on subscripts, we sometimes write $u_i(x_i, \eta) = u_i(x_i, \eta_{-i})$. Action x_i is a *best response* (*strict best response*) to profile η if for any action $x_i' \in A_i$, $u_i(x_i, \eta) \geq (>)\, u_i(x_i', \eta)$.

Definition 5. *Profile* $\mathbf{a} = (a_i) \in \times_i A_i$ *is* ordinal generalized risk$-$dominant (GR-dominant)*, if (1.1) holds. Profile* $\mathbf{a}$ *is* strictly ordinal $GR-$dominant*, if for any player* i*, any action* $x_i \in A_i, x_i \neq a_i$*, any profiles* $\eta, \bar{\eta}$ *that are* almost $\mathbf{a}$*-associated,* a_i *is a* strict *best response to either* η *or* $\bar{\eta}$*.*

This definition says that if a_i is not a best response of player i to certain profile η, then it becomes a best response if all players j who do not play a_j in profile η switch to a_j, and the other players switch to any action.

Suppose that there are two players, $I = 2$, and each player chooses one of two actions, $|A_i| = 2$ for each i. In this case, any equilibrium profile is ordinal $GR-$dominant. Indeed, suppose that $\mathbf{a} = (a_1, a_2) \in A_1 \times A_2$ is an equilibrium profile, i.e., for each player i, a_i is a best response to a_{-i}. Let $\eta, \eta' \in A_1 \times A$ be two $\mathbf{a}$-associated profiles. Then, either $\eta_{-i} = a_{-i}$ or $\eta'_{-i} = a_{-i}$ and a_i is a best response to one of the profiles η or η'. On the other hand, profile $\mathbf{a}$ is strictly ordinal $GR-$dominant if and only if a_i is a best response to a_{-i} as well as to $x_{-i} \neq a_{-i}$, hence, if and only if a_i is dominant.

3.2. **Example: Local interactions.** Let $\pi : A \times A \to R$ be a payoff function in a symmetric two-player interaction game. Following Morris, Rob, and Shin (1995), say that $a \in A$ is (strictly) p*-dominant* in interaction game π, if it is a (strict) best response to any distribution $\lambda \in \Delta A$ that assigns probability of at least p to action a.

Consider the local interaction model with payoffs u_i^{local} defined in (1.3). Ellison (1993) assumes that $g_{ij} = g_{ji} \in \{0, 1\}$ for each i and j. The next result is formulated for general weights $g_{ij} \geq 0$. Let $\mathbf{a} = (a, ..., a) \in A^I$.

Lemma 4. *If action* a *is* $\frac{1}{2}$*-dominant in two-player interaction game* π*, then* $\mathbf{a}$ *is ordinal* $GR-$*dominant.*
If action a *is strictly* $\left(\frac{1}{2} - \delta^*\right)$*-dominant in interaction game* π*, then* $\mathbf{a}$ *is strictly ordinal*

GR−dominant. Here, $\delta^* = \max_i \delta_i$ is a characteristic of weights (g_{ij}) and δ_i is defined as

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

Proof. We verify only the second statement. Suppose that a is strictly $\left(\frac{1}{2}-\delta^*\right)$-dominant in the interaction game. Take any two almost **a**-associated profiles η and $\bar{\eta}$ and suppose that a is not a strict best response to η. Then,

$$\frac{\sum_{j\neq i:\eta_j=a} g_{ij}}{\sum_{j\neq i} g_{ij}} < \frac{1}{2}-\delta^* \text{ and } \frac{\sum_{j\neq i:\eta_j\neq a} g_{ij}}{\sum_{j\neq i} g_{ij}} > \frac{1}{2}+\delta^*.$$

Because profiles η and $\bar{\eta}$ are almost **a**-associated, there is player $j^*\neq i$, such that

$$\frac{\sum_{j\neq i:\bar{\eta}_j=a} g_{ij}}{\sum_{j\neq i} g_{ij}} \geq \frac{\sum_{j\neq i,j^*:\eta_j\pm a} g_{ij}}{\sum_{j\neq i} g_{ij}} \geq \frac{1}{2}-\delta^*,$$

where the last inequality follows from the definition of δ^*. Hence, a is a strict best response to profile $\bar{\eta}$. □

There are two important applications. In the global interaction model of Kandori, Mailath, and Rob (1993) with uniform matching $g_{ij}^{\text{uniform}} = 1$ for each i and j. It is easy to check that

$$\delta_i^{\text{uniform}} = \delta^{\text{uniform},*} = \begin{cases} \frac{1}{2(I-1)}, & \text{if } I-1 \text{ is odd,} \\ 0, & \text{if } I-1 \text{ is even.}\end{cases}$$

Thus, Lemma 4 implies that (a) if a is risk-dominant in the interaction game, then profile **a** is ordinal $GR-$dominant in the multiplier game u^{uniform}, and (b) if a is strictly risk-dominant in the interaction game, then profile **a** is strictly ordinal $GR-$dominant in the multiplier game u^{uniform} if either the number of players is odd or the number of players is even and sufficiently large.

In the games on a network of Ellison (1993) $g_{ij}^{\text{network}} = g_{ji}^{\text{network}} \in \{0,1\}$ for each i and j. Define $g_i^{\text{network}} = \sum_{j\neq i} g_{ij}^{\text{network}}$ be the number of neighbors of player i. One can check that

$$\delta_i^{\text{network}} = \begin{cases} \frac{1}{2g_i^{\text{network}}}, & \text{if } g_i^{\text{network}} \text{ is odd,} \\ 0, & \text{if } g_i^{\text{network}} \text{ is even,}\end{cases}$$

and $\delta^{\text{network},*} := \max_i \delta_i^{\text{network}}$ is a characteristic of the network. One can interpret $\delta^{\text{network},*}$ as a measure of the fineness of the network. In particular, $\delta^{\text{network},*}$ can be bounded from above by a number that is inversely proportional to the minimum number of neighbors across all players, $\delta^{\text{network},*} \leq \frac{1}{2}\frac{1}{\min_i g_i}$. The sufficient conditions of Lemma 4 relate the fineness of the network to the strength of p-dominance of a in the interaction game. Specifically, if a is $\frac{1}{2}$-dominant, then **a** is ordinal $GR-$dominant on each network, and if a is strictly $\frac{1}{2}$-dominant, then $\mathbf{a}$ is strictly ordinal $GR-$dominant on each network such that

$$\delta^{\text{network},*} < \frac{1}{2} - \inf\{p : a \text{ is strictly } p\text{-dominant}\}. \tag{3.1}$$

Note that for the two networks considered by Ellison, circle and torus, the number of neighbors of each player is always even. Thus, $\delta^{\text{network},*} = 0$, and condition (3.1) is trivially satisfied whenever a is strictly $\frac{1}{2}$-dominant in the interaction game.

3.3. **KMR dynamics.** Consider the evolutionary learning process based on Kandori, Mailath, and Rob (1993). Each period player i is randomly drawn (say with a probability $\frac{1}{N}$) and given an opportunity to change his action. With a probability of order $1-\varepsilon$, she chooses one of the best responses to the current profile in the population; with a probability of order ε, she chooses an action randomly. Formally, assume that the dynamics is a Markov process with a cost function $c_{KMR} : \Sigma \times \Sigma \to [0,\infty]$:

- $c_{KMR}(\eta,\eta) = 0$,
- $c_{KMR}(\eta,\eta') = 0$ if there is player j, such that η'_j is a best response of player j to η_{-j} and $\eta_i = \eta'_i$ for all $i \neq j$,
- $c_{KMR}(\eta,\eta') = 1$ if there is player j, such that η'_j is not a best response of player j to η_{-j} and $\eta_i = \eta'_i$ for all $i \neq j$,
- $c_{KMR}(\eta,\eta') = \infty$ in all other cases.

Theorem 2. *If* $\mathbf{a}$ *is ordinal* $GR-$*dominant, then it is stochastically stable under dynamics* c_{KMR}. *If* $\mathbf{a}$ *is strictly ordinal* $GR-$*dominant, then it is uniquely stochastically stable under dynamics* c_{KMR}.

Proof. We show that if profile $\mathbf{a}$ is ordinal $GR-$dominant (strictly ordinal $GR-$dominant), then cost function c_{KMR} is asymmetric (robustly asymmetric) toward $\mathbf{a}$. The result will follow from Theorem 1.

The proof proceeds in four steps. First, take any profiles $\eta,\eta',\bar{\eta}$ such that $\eta,\bar{\eta}$ are $\mathbf{a}$-associated (almost $\mathbf{a}$-associated). We show that there is a profile $\bar{\eta}'$, such that η' and $\bar{\eta}'$ are $\mathbf{a}$-associated (almost $\mathbf{a}$-associated) and $c_{KMR}(\bar{\eta},\bar{\eta}') \leq c_{KMR}(\eta,\eta')$. W.l.o.g. assume that $c_{KMR}(\eta,\eta') = 0$ (the other cases are trivial). Let j and x_j be a player and an action, such that $\eta'_j = x_j \neq \eta_j$. Hence, x_j is a best response of player j to profile η. Consider two cases:

- If $x_j \neq a_j$, then, by ordinal $GR-$dominance, a_j is a best response of player j to $\bar{\eta}$. Let $\bar{\eta}'$ be such that $\bar{\eta}'_{-j} = \bar{\eta}_{-j}$ and $\bar{\eta}'_j = a_j$. Then, $\bar{\eta}$ is $\mathbf{a}$-dominated by $\bar{\eta}'$, profiles $\eta',\bar{\eta}'$ are $\mathbf{a}$-associated (almost $\mathbf{a}$-associated), and $c_{KMR}(\bar{\eta},\bar{\eta}') = 0$ by the definition of cost function c_{KMR}.

- If $x_j = a_j$, then take $\bar{\eta}' := \bar{\eta}$. Profile $\bar{\eta}$ is **a**-dominated by itself. Since η is **a**-dominated by η', profiles η' and $\bar{\eta}'$ are **a**-associated. By definition, $c_{KMR}(\bar{\eta}, \bar{\eta}') = c_{KMR}(\bar{\eta}, \bar{\eta}) = 0$.

This shows that if **a** is ordinal $GR-$dominant, then cost function c_{KMR} is asymmetric. This also shows that if **a** is strictly ordinal $GR-$dominant, then part 2 of the definition of robustly asymmetric cost function (Definition 3) is satisfied.

Second, notice that if **a** is strictly ordinal $GR-$dominant, then it is a strict Nash equilibrium. Thus, for any $\eta \in \Sigma$, $c_{KMR}(\mathbf{a}, \eta) > 0$ and part 1 of Definition 3 holds.

Third, take any profiles η, η' such that $c_{KMR}(\eta, \eta') > 0$. Assume that $c_{KMR}(\eta, \eta') = 1$ (the case $c_{KMR}(\eta, \eta') = \infty$ is trivial). Let $\bar{\eta}$ be a profile such that η and $\bar{\eta}$ are **a**-associated and take $\bar{\eta}' := \bar{\eta}$. Then, $c_{KMR}(\bar{\eta}, \bar{\eta}') = 0$ and $\bar{\eta}'$ is almost **a**-associated with η' (the latter is a consequence of the fact that there is a player j such that $\eta'_j \neq \eta_j$ and $\eta_i = \eta'_i$ for each $i \neq j$; hence, for each $i \neq j$, either $\bar{\eta}'_i = a_i$ or $\eta'_i = a_i$). Thus, part 3 of Definition 3 holds.

Fourth, take any profiles $\eta, \bar{\eta}$ such that (a) $\eta, \bar{\eta}$ are almost **a**-associated and (b) $c_{KMR}(\eta, \eta'') > 0$ for any profile η'' that strictly dominates η. We show that there is profile $\bar{\eta}'$ such that η and $\bar{\eta}'$ are **a**-associated, $\bar{\eta}$ is **a**-dominated by $\bar{\eta}'$ and $c(\bar{\eta}, \bar{\eta}') = 0$. If η and $\bar{\eta}$ are **a**-associated, then let $\bar{\eta}' := \bar{\eta}$; thus $c_{KMR}(\bar{\eta}, \bar{\eta}) = 0$. If $\eta, \bar{\eta}$ are not **a**-associated, then there is a player j such that $\eta_j \neq a_j$ and $\bar{\eta}_j \neq a_j$. Because of (b), a_j is *not* a best response to η_{-j}. Because **a** is ordinal $GR-$dominant, a_j is a strict best response to $\bar{\eta}_{-i}$. Let $\bar{\eta}'$ be a profile such that $\bar{\eta}'_j = a_j$ and for any $i \neq j$, $\bar{\eta}'_i = \bar{\eta}_i$. Then, $c_{KMR}(\bar{\eta}, \bar{\eta}') = 0$. Thus, if **a** is ordinal $GR-$dominant, then part 4 of Definition 3 holds.

This shows that if **a** is strictly ordinal $GR-$dominant, then cost function c_{KMR} is robustly asymmetric. □

The dynamics c_{KMR} differ from Kandori, Mailath, and Rob (1993) who allow for multiple adjustments in each period. In that version, each period, each player is activated with probability $p \in (0,1)$. Once activated, with a probability of order $1-\varepsilon$, she chooses one of the best responses to the current profile in the population; with a probability of order ε, she chooses an action randomly. The cost function of these dynamics can be defined as

$$c_{KMR}^{\text{multiple}}(\eta, \eta') := \left|\{i : \eta'_i \neq \eta_i \text{ and } \eta_i \text{ is not a best response to } \eta_{-i}\}\right|.$$

This modification does not change the main result: Theorem 2 still holds with c_{KMR} replaced by $c_{KMR}^{\text{multiple}}$. The proof of this fact follows the same lines as the proof of Theorem 2 and, as such, is left out.

Theorem 2 and Lemma 4 imply that

Corollary 1. *If a is $\frac{1}{2}$-dominant in interaction game π, then* $\mathbf{a}$ *is stochastically stable under dynamics c_{KMR}. Additionally, if the network satisfies condition (3.1), then* $\mathbf{a}$ *is uniquely stochastically stable under dynamics c_{KMR}.*

The Corollary extends the results by Ellison (1993) and Ellison (2000) to all networks. The role of condition (3.1) is to avoid integer problems that appear when the number of players is small. To see this, consider two examples. The first example comes from Jackson and Watts (2002).

Example 1 (Star)**.** *Suppose that $A = \{a, b\}$ and π is a payoff in a coordination game with two pure strategy equilibria (a, a) and (b, b). Consider a network with a central player connected to all other players and none of the other players are connected to each other. Then ,both the equilibrium profiles are stochastically stable. To see this, observe that one needs exactly one mistake by the central player to switch between coordinating on one of two actions.*

Example 2. *Suppose that $A = \{a, b\}$, action a is a best response to any profile with at least 40% players playing a, and action b is a best response to any profile with at least 60% playing b. Suppose that there are n players, all connected to each other. If $n = 4$, then the number of neighbors of each player is odd, and it takes exactly two mutations to move the process out of the respective basins of attraction. Hence, both profiles* $\mathbf{a}$ *and* $\mathbf{b}$ *are stochastically stable. If there are $n = 3$ or $n = 5$ players, then the number of neighbors of each player is even. If one (in case $n = 3$) or two (in case $n = 5$) players change their actions to a, then a becomes a strict best response for the remaining players. On the other hand, starting from profile* $\mathbf{a}$*, one needs two (in case $n = 3$) or three (in case $n = 5$) players to change their actions to make b a best response for the remaining players. Hence, with $n = 3, 5$, only* $\mathbf{a}$ *is stochastically stable.*

4. Cardinal generalized risk-dominance

4.1. Definition.

This section presents a cardinal generalization of risk-dominance.

Definition 6. *Profile $\mathbf{a} = (a_i) \in \times_i A_i$ is* cardinal $GR-$dominant*, if condition (1.2) holds. Profile* $\mathbf{a}$ *is* strictly cardinal $GR-$dominant *if for any player i, any profiles $\eta, \bar{\eta}$ that are* $\mathbf{a}$*-associated,*

$$\max_{x_i \neq a_i} u_i(x_i, \eta) - u_i(a_i, \eta) < u_i(a_i, \bar{\eta}) - \max_{x_i \neq a_i} u_i(x_i, \bar{\eta}) \text{ (strict inequality).}$$

This says that for any player i, any pair of $\mathbf{a}$-associated profiles $\eta, \bar{\eta}$, the difference between the payoff from any action $x_i \neq a_i$ and the payoff from a_i against profile η is smaller than the difference between the payoff from a_i and any other action $x_i \neq a_i$ against profile $\bar{\eta}$.

Clearly, cardinal $GR-$dominance implies ordinal $GR-$dominance. However, strict cardinal $GR-$dominance does not necessarily imply strict ordinal $GR-$dominance.

4.2. **Symmetric conjectures.** Suppose that there are only two actions for each player, $A_i = \{a_i, b_i\}$. Let $\mathbf{a} = (a_i) \in \times_i A_i$. Say that payoffs are *complementary*, if for each player i, for each profiles η and η' such that η' **a**-dominates η, $u_i(a_i, \eta) - u_i(b_i, \eta) \leq u_i(a_i, \eta') - u_i(b_i, \eta')$. For each profile η, denote $-\eta$ as the profile with all actions flipped: $\eta_i = a_i$ iff $(-\eta)_i = b_i$. Conjecture $\sigma \in \Delta(\times_i A_i)$ about action profiles is *symmetric with respect to labels* if for each η, $\sigma(\eta) = \sigma(-\eta)$.

Lemma 5. *If payoffs are complementary, then a_i is player i's (strict) best response to any symmetric conjecture for each player i if and only if* **a** *is (strictly) cardinal $GR-$dominant.*

Proof. Suppose that a_i is player i's (strict) best response to any symmetric conjecture for each player i. Take any **a**-associated pair of profiles η and $\bar{\eta}$. Consider a symmetric conjecture σ_η that assigns probability $\frac{1}{2}$ to profile η (and probability $\frac{1}{2}$ to profile $-\eta$). Then,

$$\frac{1}{2}u_i(a_i, \eta) + \frac{1}{2}u_i(a_i, -\eta) = u_i(a_i, \sigma_\eta)$$
$$\geq (>)\, u_i(b_i, \sigma_\eta) = \frac{1}{2}u_i(b_i, \eta) + \frac{1}{2}u_i(b_i, -\eta).$$

Observe that $\bar{\eta}$ **a**-dominates $-\eta$. By complementarity,

$$u_i(a_i, -\eta) \leq u_i(a_i, \bar{\eta}) \text{ and } u_i(b_i, -\eta) \geq u_i(b_i, \bar{\eta}).$$

The inequalities imply that **a** is (strictly) cardinal $GR-$dominant.

Suppose that **a** is (strictly) cardinal $GR-$dominant. Take any symmetric conjecture σ and notice that it is equal to convex combinations of symmetric conjectures that assign positive probability to only two states, $\sigma = \sum \alpha_\eta \sigma_\eta$ where $a_\eta \geq 0$, $\sum \alpha_\eta = 1$ and $\sigma_\eta(\eta) = \sigma_\eta(-\eta) = \frac{1}{2}$. By (strict) cardinal $GR-$dominance,

$$u_i(a_i, \sigma) = \frac{1}{2}\sum a_\eta (u_i(a_i, \bar{\eta}) + u_i(a_i, -\eta))$$
$$\geq (>)\, \frac{1}{2}\sum a_\eta (b_i(a_i, \bar{\eta}) + u_i(b_i, -\eta)) = u_i(b_i, \sigma).$$

□

4.3. **Cardinal $GR-$dominance and $\frac{1}{2}$-dominance.** With two players and two actions, (strict) cardinal $GR-$dominance is equivalent to Harsanyi and Selten (1988)'s (strict) risk dominance. This is a consequence of Lemma 5.[6] A similar argument shows that in two-player, multi-action games, (strict) cardinal GR-dominance implies Morris, Rob, and Shin (1995)'s (strict) $\frac{1}{2}$-dominance.

For multi-player games, Morris, Rob, and Shin (1995) defines a $(p_1, ..., p_I)$-dominant equilibrium, in which the action of player i is the best response to any conjecture that assigns a probability of at least p_i to the equilibrium action profile. If there are only two actions for each player and complementarities, then cardinal $GR-$dominance implies $\left(\frac{1}{2}, ..., \frac{1}{2}\right)$-dominance. This follows from Lemma 5 and the fact that a conjecture according to which each player randomizes equally between the two actions is symmetric.

In general, there is no further logical relationship between $GR-$dominance and $\left(\frac{1}{2}, ..., \frac{1}{2}\right)$-dominance. As we discuss next, there exists a connection between $\frac{1}{2}$-dominance in the two-player interaction game and cardinal $GR-$dominance with local interactions.

4.4. **Example: Local interactions.** Consider the local interaction model described in section 3.2. Let $\pi : A \times A \rightarrow R$ be a symmetric payoff in a two-player interaction. Fix $a \in A$ and let $\mathbf{a} = (a, ..., a)$. Say that action $a \in A$ is *(strictly) strongly $\frac{1}{2}$-dominant* if for any distribution $\lambda \in \Delta A$ such that $\lambda(a) \geq \frac{1}{2}$,

$$\sum_a \pi(a, x) \lambda(x) \geq (>) \sum_a \max_{x' \neq a} \pi(x', x) \lambda(x).$$

Thus, if action is strongly $\frac{1}{2}$-dominant, it is also $\frac{1}{2}$-dominant, but not necessarily the other way.

Lemma 6. *If a is strongly $\frac{1}{2}$-dominant in the interaction game π, then $\mathbf{a}$ is cardinal $GR-$dominant. If a is strictly strongly $\frac{1}{2}$-dominant in the interaction game π, then $\mathbf{a}$ is strictly cardinal $GR-$dominant.*

Proof. We verify only the second statement. Suppose that a is strictly strongly $\frac{1}{2}$-dominant in the interaction game. Take any pair of $\mathbf{a}$-associated profiles $\eta, \bar{\eta}$ and fix player i. Define distribution $\lambda \in \Delta A$: for any $x \in A$

$$\lambda(x) = \frac{1}{g_i} \sum\nolimits_{j \neq i} \frac{1}{2} \left(\mathbf{1}(\eta_j = x) + \mathbf{1}(\bar{\eta}_j = x) \right).$$

[6]Note that, if profile (a_i, a_j) is an equilibrium in a two-player, two-action game, then the payoffs are complementary.

Then, $\lambda(a) \geq \frac{1}{2}$. By strict strong $\frac{1}{2}$-dominance,

$$\begin{aligned}
& u_i(a,\eta) + u_i(a,\bar{\eta}) \\
&= 2\sum_a \pi(a,x)\lambda(x) > 2\sum_a \max_{x' \neq a} \pi(x',x)\lambda(x) \\
&= \frac{1}{g_i}\sum_a \sum\nolimits_{j\neq i} \max_{x'\neq a} \pi(x',x)\mathbf{1}(\eta_j = x) + \frac{1}{g_i}\sum_a \sum\nolimits_{j\neq i} \max_{x'\neq a} \pi(x',x)\mathbf{1}(\bar{\eta}_j = x) \\
&\geq \max_{x_i \neq a_i} u_i(x_i,\eta) + \max_{x_i\neq a_i} u_i(x_i,\bar{\eta}).
\end{aligned}$$

Thus, $\mathbf{a}$ is strictly cardinal $GR-$dominant. □

4.5. **Payoff-based dynamics.** Let $f : [0,\infty) \to [0,\infty)$ be a strictly increasing function such that $f(0) = 0$. Define cost function c_f : for any profiles η, η'

- if $\eta = \eta'$, then $c_f(\eta,\eta') = 0$,
- if there is player j, such that $\eta_i = \eta'_i$ for all $i \neq j$, then

$$c_f(\eta,\eta') = f\left(\max_{x_j} u_j(x_j,\eta) - u(\eta'_j,\eta)\right).$$

- otherwise, $c_f(\eta,\eta') = \infty$.

These dynamics relates the cost of a transition toward x_j to the difference between best response payoff and the payoff from x_j. By assumption, the cost of a transition is increasing in the difference between payoffs. For example, if the cost of a transition is linear in the difference between payoffs, $f(c) = \beta c$, for some $\beta > 0$, then c_f is a cost function of Blume (1993). (If f is increasing rather than strictly increasing and $f(c) = \mathbf{1}\{c \geq 0\}$, then $c_f = c_{KMR}$.)

Theorem 3. *If* $\mathbf{a}$ *is cardinal* $GR-$*dominant, then it is stochastically stable under dynamics* c_f. *If* $\mathbf{a}$ *is strictly cardinal* $GR-$*dominant, then it is uniquely stochastically stable under dynamics* c_f.

Proof. We show that if profile $\mathbf{a}$ is cardinal $GR-$dominant (strictly cardinal $GR-$dominant), then cost function c_f is asymmetric (strictly asymmetric). The result will follow from Theorem 1.

Take any profiles $\eta, \eta', \bar{\eta}$, such that $\eta, \bar{\eta}$ are $\mathbf{a}$-associated. We show that there exists $\bar{\eta}'$ such that $\bar{\eta}$ is $\mathbf{a}$-dominated by $\bar{\eta}'$, η' and $\bar{\eta}'$ are $\mathbf{a}$-associated and either $c_f(\bar{\eta},\bar{\eta}') = 0$ or $c_f(\bar{\eta},\bar{\eta}') < c(\eta,\eta')$. Assume that there is a player j, such that $\eta_i = \eta'_i$ for all $i \neq j$ (the other case is trivial because of convention that $\infty < \infty$). Let $x_j = \eta'_j$. There are a few cases to be considered.

- If $x_j = a_j$ or ($\eta_j \neq a_j$ and $x_j \neq a_j$), then let $\bar{\eta}' := \bar{\eta}$. Profiles η' and $\bar{\eta}'$ are **a**-associated, and, by definition, $c_f(\bar{\eta}, \bar{\eta}') = 0 \leq c_f(\eta, \eta')$.
- If $x_j \neq a_j$, $\eta_j = a_j$ and there is $x'_j \neq a_j$ which is a best response to η, then, by cardinal $GR-$dominance, a_j is a best response to $\bar{\eta}$. Take $\bar{\eta}'$ such that $\bar{\eta}'_{-j} = \bar{\eta}_{-j}$ and $\bar{\eta}'_j = a_j$. Then, $\bar{\eta}$ is **a**-dominated by $\bar{\eta}'$ and profiles $\eta', \bar{\eta}'$ are **a**-associated. Because a_j is a best response to $\bar{\eta}$, $c_f(\bar{\eta}, \bar{\eta}') = 0 \leq c_f(\eta, \eta')$.
- If $x_j \neq a_j$, $\eta_j = a_j$ and a_j is a strict best response to η, then

$$c_f(\eta, \eta') = f(u_j(a_j, \eta) - u(x_j, \eta)) > 0.$$

Take $\bar{\eta}'$ such that $\bar{\eta}'_{-j} = \bar{\eta}_{-j}$ and $\bar{\eta}'_j = a_j$. Then, $\bar{\eta}$ is **a**-dominated by $\bar{\eta}'$ and profiles $\eta', \bar{\eta}'$ are **a**-associated. By (strict) cardinal $GR-$dominance,

$$\begin{aligned}
& u_j(a_j, \eta) - u(x_j, \eta) \\
& \leq u_j(a_j, \eta) - \max_{x'_j \neq a_i} u\left(x'_j, \eta\right) \\
& \leq (<) \max_{x'_j \neq a_i} u\left(x'_j, \bar{\eta}\right) - u_j(a_j, \bar{\eta}) \\
& \leq \max_{x'_j} u\left(x'_j, \bar{\eta}\right) - u_j(a_j, \bar{\eta}),
\end{aligned}$$

which implies that

$$c_f(\bar{\eta}, \bar{\eta}') \leq c_f(\eta, \eta') \ \left(c_f(\bar{\eta}, \bar{\eta}') < c_f(\eta, \eta') \right).$$

This shows that if **a** is cardinal $GR-$dominant, then c_f is asymmetric and it also shows that strict cardinal $GR-$dominance implies part 2 of the definition of strictly asymmetric cost function (Definition 2).

Finally, if **a** is strictly cardinal $GR-$dominant, then for any $\eta \in \Sigma$, $c_f(\mathbf{a}, \eta) > 0$. This is because a_i is a strict best response to profile **a** (note that **a** is **a**-associated with **a**). Thus, part 1 of Definition 2 holds, and c_f is strictly asymmetric. □

It is instructive to compare Theorems 2 and 3. Under c_{KMR}, the probability of a player switching between two actions depends only on whether these actions are myopic best responses. That, in turn, depends only on the ordinal properties of the payoff function. It is not surprising that these dynamics select an equilibrium that satisfies certain ordinal properties. In contrast, the dynamics of Blume (1993), or, more generally, dynamics c_f, depend on the cardinal differences between payoffs from various actions. Thus, such dynamics should select an equilibrium that satisfies certain cardinal properties. This intuition is confirmed by two Theorems.

Corollary 2. *If a is strongly $\frac{1}{2}$-dominant (strictly strongly $\frac{1}{2}$-dominant) in the interaction game π, then* **a** *is stochastically stable (uniquely stochastically stable) under dynamics c_f.*

The Corollary extends the results by Blume (1993) and Young (1998) to all interaction games with strongly $\frac{1}{2}$-dominant action.

References

Blume, A., and T. Temzelides (2003): "On the Geography of Conventions," *Economic Theory*, 22, 863–873.

Blume, L. E. (1993): "The Statistical-Mechanics of Strategic Interaction," *Games and Economic Behavior*, 5(3), 387–424.

Carlsson, H., and E. V. Damme (1993): "Global Games and Equilibrium Selection," *Econometrica*, 61, 989–1018.

Chu, Y., and T. Liu (1965): "On the Shortest Arborescence of a Directed Graph," *Science Sinica*, 14, 1396–1400.

Edmonds, J. (1967): "Optimum Branchings," *Journal of Research of National Bureau of Standards*, B71, 233–240.

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

——— (2000): "Basins of attraction, long-run stochastic stability, and the speed of step-by-step evolution," *Review of Economic Studies*, 67(230), 17.

Freidlin, M. I., and A. D. Wentzell (1984): *Random perturbations of dynamical systems*, Grundlehren der mathematischen Wissenschaften ; Bd. 260. Springer-Verlag, New York.

Harsanyi, J. C., and R. Selten (1988): *A general theory of equilibrium selection in games*. Cambridge, Mass. and London: MIT Press.

Jackson, M. O., and A. Watts (2002): "On the formation of interaction networks in social coordination games," *Games and Economic Behavior*, 41(2), 265–291.

Kajii, A., and S. Morris (1997): "The Robustness of Equilibria to Incomplete Information," *Econometrica*, 65, 1283–1309.

Kandori, M., G. J. Mailath, and R. Rob (1993): "Learning, Mutation, and Long Run Equilibria in Games," *Econometrica*, 61(1), 29.

Lee, I. H., A. Szeidl, and A. Valentinyi (2003): "Contagion and State Dependent Mutations," *Advances in Theoretical Economics*, 3.

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

Morris, S., R. Rob, and H. S. Shin (1995): "P-Dominance and Belief Potential," *Econometrica*, 63(6), 145 – 157.

Nöldeke, G., and L. Samuelson (1993): "An Evolutionary Analysis of Backward and Forward Induction," *Games and Economic Behavior*, 5, 425–454.

Okada, D., and O. Tercieux (2008): "Log-Linear Dynamics and Local Potential," http://www3.grips.ac.jp/ yyasuda/workshop/games2008.files/Okada.pdf.

Rubinstein, A. (1989): "The Electronic Mail Game: Strategic Behavior under "Almost Common-Knowledge."," *American Economic Review*, 79, 385–91.

Young, H. P. (1993): "The Evolution of Conventions," *Econometrica*, 61(1), 57–84.

Young, H. P. (1998): *Individual Strategy and Social Structure*. Princeton University Press, Princeton.

Appendix A. Proof of Lemma 2

We prove Lemma 2 indirectly, as an application of Lemma 1. Take any cost function c. Let $P_{\varepsilon}(\eta,\eta')$ be a family of ergodic Markov processes such that limits (2.1) exist for any pair of states $\eta,\eta'\in\Sigma$. Let μ_{ε} be the stationary distribution of P_{ε}. We construct three auxiliary ergodic Markov processes on state space S, $Q^0_{\varepsilon}, Q^{\alpha}_{\varepsilon}, Q_{\epsilon}(v,v')$, and their stationary distributions $q^0_{\varepsilon}, q^{\alpha}_{\varepsilon}, q_{\varepsilon}\in\Delta S$ such that:

- $\mu_{\varepsilon}(\mathbf{a}) = q^0_{\varepsilon}(\mathbf{1}) = q^{\alpha}_{\varepsilon}(\mathbf{1}) \geq q_{\varepsilon}(\mathbf{1})$ (Lemma 14),
- there exists a sequence $\varepsilon_n\to 0$ such that limits
$$\lim_{n\to\infty}\frac{1}{\log\varepsilon_n}\log Q_{\varepsilon_n}(v,v') = \psi(v,v') \tag{A.1}$$
exist (Lemma 11),
- ψ is supermodular and, if c is (robustly, strictly) asymmetric toward $\mathbf{a}$, then ψ is (robustly, strictly) asymmetric toward $\mathbf{1}$ (Lemmas 9 and 10).

Assume that cost function c is (robustly, strictly) asymmetric toward $\mathbf{a}$. Then,

$$\begin{aligned}
&\mathbf{1}\in MR(S,\psi)\\
&\quad\Rightarrow \liminf_{\varepsilon\to 0} q_{\varepsilon}(\mathbf{1}) > 0\\
&\quad\Rightarrow \liminf_{\varepsilon\to 0} \mu_{\varepsilon}(\mathbf{a}) > 0\\
&\quad\Rightarrow \mathbf{a}\in MR(\Sigma,c).
\end{aligned}$$

The first implication is a consequence of (A.1) and the proof of Lemma 1 (indeed, the proof of Lemma 1 implies that $\lim_{n\to\infty} q^*_{\varepsilon_n}(\mathbf{1}) \approx \lim_{n\to\infty}\varepsilon^{c(h_{\mathbf{1}})}\left(\sum_{\text{all trees } h}\varepsilon^{c(h)}\right)^{-1}$, where $h_{\mathbf{1}}$ is a minimal cost tree with its root at $\mathbf{1}$); the second one follows from the fact that $q_{\varepsilon}^{1}(\mathbf{1}) \leq \mu_{\varepsilon}(\mathbf{a})$; the third is yet another application of Lemma 1. Similarly,

$$\begin{aligned}\{\mathbf{1}\} &= MR(S, \psi) \\ &\Rightarrow \liminf_{\varepsilon \to 0} q_{\varepsilon}(\mathbf{1}) = 1 \\ &\Rightarrow \liminf_{\varepsilon \to 0} \mu_{\varepsilon}(\mathbf{a}) = 1 \\ &\Rightarrow \{\mathbf{a}\} = MR(\Sigma, c).\end{aligned}$$

This ends the proof of the Lemma.

A.1. **Markov processes on lattice** S**.** Define a lattice structure on S. For any $v, v' \in S$, let $v \vee v'$ be the *join* of v and v' : for each i, $(v \vee v')_i = \max(v_i, v_i')$, and let $v \wedge v'$ be the *meet* of v and v' : for each i, $(v \vee v')_i = \max(v_i, v_i')$. Set $E \subseteq S$ is called *upper* if for any $v \in E, v' \in S$, if $v \leq v'$, then $v' \in E$. For all v, v', define $\Lambda(v, v') = \{\tilde{v}' : \tilde{v}' \wedge v = v'\}$.

Let ΔS be the space of probability distributions on S. For distributions $\lambda, \lambda' \in \Delta S$ say that $\lambda \leq \lambda'$ if for each upper $E \subseteq S$, $\lambda(E) \leq \lambda'(E)$. For any distribution $\lambda \in \Delta S$, any Markov process $Q : S \to \Delta S$, let $Q\lambda \in \Delta S$ be defined as $Q\lambda(v) = \sum_{v'} \lambda(v') Q(v', v)$. A Markov process on S is a mapping $Q : S \to \Delta S$. For any two Markov processes $Q, Q' : S \to \Delta S$, write $Q \leq Q'$ if $Q(v) \leq Q'(v)$ for each $v \in S$. One shows that if $Q \leq Q$", then $Q\lambda \leq Q'\lambda$ for any $\lambda \in \Delta S$. A Markov process $Q : S \to \Delta S$ is *monotonic* if for each $v, v' \in S$, if $v \leq v'$, then $Q(v) \leq Q(v')$. One shows that if Q is monotonic and $\lambda \leq \lambda'$, then $Q\lambda \leq Q\lambda'$.

Lemma 7. *Suppose that Q and Q' are ergodic Markov processes on S and q and q' are their respective stationary distribution. If Q is monotonic and $Q \leq Q'$, then $q \leq q'$.*

Proof. (See also Okada and Tercieux (2008).) Observe that

$$\begin{aligned}q &= \lim_{n \to \infty} (Q)^n q = \lim_{n \to \infty} (Q....QQ)_{n \text{ times}}\, q \\ &\leq \lim_{n \to \infty} (Q....QQ')_{n \text{ times}}\, q \leq ... \leq \lim_{n \to \infty} (Q'....Q'Q')_{n \text{ times}}\, q \\ &= \lim_{n \to \infty} (Q')^n q = q'.\end{aligned}$$

□

A.2. **Auxiliary family** Q^0**.** Write $P_{\varepsilon}(\eta, E) = \sum_{\eta' \in F} P_{\varepsilon}(\eta, \eta')$ for any subset $F \subseteq \Sigma$ and any profile η. Let $\sigma : \Sigma \to S$ be a mapping such that for any profile $\eta \in \Sigma$, $(\sigma(\eta))_i = 1$ if and only if $\eta_i = a_i$. For any $v, v' \in S$, define

$$Q_{\varepsilon}^{0}(v, v') := \sum_{\eta \in \sigma^{-1}(v)} \frac{\mu_{\varepsilon}(\eta)}{\mu_{\varepsilon}(\sigma^{-1}(v))} P_{\varepsilon}\left(\eta, \sigma^{-1}(v')\right).$$

This defines Markov process Q^0_ε on S. For any state v, let $q^0_\varepsilon(v) = \sum_{\eta\in\sigma^{-1}(v)} \mu_\varepsilon(\eta)$. Then, $q^0_\varepsilon \in \Delta S$ is the unique stationary distribution of Markov process Q_ε.

Lemma 8. *There exists sequence $\varepsilon_n \to 0$, cost function ψ^0, η_v for each v and $b^0(v,v') > 0$ for each v, v', such that for each $v, v' \in S$,*

$$\lim_{n\to\infty} \frac{\log Q^0_{\varepsilon_n}(v,v')}{\log \varepsilon_n} = 1, \tag{A.2}$$

and, if c is (strictly, robustly) asymmetric toward **a***, then ψ^0 is (strictly, robustly) asymmetric toward* **1***.*

Proof. Find subsequence of ε_n such that for each v,there exists η_v so that

$$\lim_{n\to\infty}\inf \frac{\mu_{\varepsilon_n}(\eta_v)}{\mu_{\varepsilon_n}(\sigma^{-1}(v))} > 0,$$

and limits

$$\psi^0(v,v') := \lim_{n\to\infty} \frac{\log Q^0_{\varepsilon_n}(v,v')}{\log \varepsilon_n} \in [0,\infty]$$

exist for each v and v'. Notice that for each $v \in S$,

$$\begin{aligned} 0 &\leq \limsup_{\varepsilon\to 0} \frac{\log Q^0_\varepsilon(v,v)}{\log \varepsilon} \\ &\leq \limsup_{\eta} \sup_{\varepsilon\to 0} \frac{\log P_\varepsilon(\eta_v,\eta_v)}{\log \varepsilon} = c(\eta_v,\eta_v) = 0. \end{aligned}$$

Thus, $\psi^0(v,v) = 0$. For each $v \neq v'$,

$$\begin{aligned} \frac{\log Q^0_{\varepsilon_n}(v,v')}{\log \varepsilon_n} &\leq (1+\delta_n) \frac{\log P_\varepsilon(\eta_v, \sigma^{-1}(v'))}{\log \varepsilon_n} \leq (1+\delta'_n) \min_{\eta'\in\sigma^{-1}(v')} c(\eta_v,\eta') \\ &\leq (1+\delta'_n) \max_{\eta\in\sigma^{-1}(v)} \min_{\eta'\in\sigma^{-1}(v')} c(\eta,\eta'), \text{ and} \\ \frac{\log Q^0_{\varepsilon_n}(v,v')}{\log \varepsilon_n} &\geq (1+\delta''_n) \min_{\eta\in\sigma^{-1}(v)} \min_{\eta'\in\sigma^{-1}(v')} c(\eta,\eta'). \end{aligned}$$

where $\delta_n, \delta'_n \to 0$. Therefore,

$$\min_{\eta\in\sigma^{-1}(v)} \min_{\eta'\in\sigma^{-1}(v')} c(\eta,\eta') \leq \psi^0(v,v') \leq \max_{\eta\in\sigma^{-1}(v)} \min_{\eta'\in\sigma^{-1}(v')} c(\bar\eta,\bar\eta'). \tag{A.3}$$

(Strict, robust) asymmetry of ψ^0 is implied by (strict, robust) asymmetry of c and inequalities (A.3). We show that the asymmetry asymmetry of cost function c implies the asymmetry of cost function ψ^0. Take any v, v', and $\bar v$ such that v and $\bar v$ are **1**-associated. There are $\eta \in \sigma^{-1}(v), \eta' \in \sigma^{-1}(v')$ such that $\psi^0(v,v') \geq c(\eta,\eta')$. Let $\bar\eta^* \in \sigma^{-1}(\bar v)$ be such that for each $\bar{\eta} \in \sigma^{-1}(\bar{v})$, $\min_{\eta' \in \sigma^{-1}(v')} c(\bar{\eta}, \bar{\eta}') \leq \min_{\eta' \in \sigma^{-1}(v')} c(\bar{\eta}^*, \bar{\eta}')$. Then, η and $\bar{\eta}^*$ are $\mathbf{a}$-associated. By asymmetry of c, there exists $\bar{\eta}'$ that $\mathbf{a}$-dominates $\bar{\eta}^*$, $\bar{\eta}'$ and η' are $\mathbf{a}$-associated, and $c(\bar{\eta}^*, \bar{\eta}') \leq c(\eta, \eta')$. But then, $\psi^0(v, \sigma(\bar{\eta}')) \leq c(\bar{\eta}^*, \bar{\eta}') \leq c(\eta, \eta') \leq \psi^0(v, v')$.

The same argument (with "associated" replaced by "almost associated") implies that if c is robustly asymmetric then part 2 of Definition 3 holds for ψ.

suppose that condition (1) of definitions of strict and robust asymmetry holds for cost function c. Then, because $\sigma^{-1}(\mathbf{1}) = \{\mathbf{a}\}$, for any $v < \mathbf{1}$,

$$\psi^0(\mathbf{1}, v) \geq \min_{\eta' \in \sigma^{-1}(v')} c(\mathbf{a}, \eta') > 0,$$

and condition (1) holds also for cost function ψ^0.

Suppose that c is strictly asymmetric. Take any $v, v', \bar{v}$, such that $v, \bar{v}$ are $\mathbf{1}$-associated and $\psi^0(v, v') > 0$. There are $\eta \in \sigma^{-1}(v), \eta' \in \sigma^{-1}(v')$ such that $\psi^0(v, v') \geq c(\eta, \eta')$. Let $\bar{\eta}^* \in \sigma^{-1}(\bar{v})$ be such that for each $\bar{\eta} \in \sigma^{-1}(\bar{v})$, $\min_{\eta' \in \sigma^{-1}(v')} c(\bar{\eta}, \bar{\eta}') \leq \min_{\eta' \in \sigma^{-1}(v')} c(\bar{\eta}^*, \bar{\eta}')$. Then, η and $\bar{\eta}^*$ are $\mathbf{a}$-associated. By strict asymmetry of c, there exists $\bar{\eta}'$ that $\mathbf{a}$-dominates $\bar{\eta}^*$, $\bar{\eta}'$ and η' are $\mathbf{a}$-associated, and $c(\bar{\eta}^*, \bar{\eta}') < c(\eta, \eta')$. But then, $\psi^0(v, \sigma(\bar{\eta}')) \leq c(\bar{\eta}^*, \bar{\eta}') < c(\eta, \eta') \leq \psi^0(v, v')$. Thus, ψ^0 is strictly asymmetric toward $\mathbf{1}$.

Suppose that c is robustly asymmetric. A similar argument to the one given above shows that condition (3) of Definition 3 holds for ψ^0. We will show that condition (4) of Definition 3 holds for ψ^0. Take v,v', and $\bar{v}$ such that v' is almost $\mathbf{1}$-dominated by v and $\bar{v}$ is $\mathbf{1}$-associated with v'. If $\psi^0(v, v') > 0$, then there are $\eta \in \sigma^{-1}(v), \eta' \in \sigma^{-1}(v')$ such that $c(\eta, \eta') > 0$. By robust asymmetry of c, for each $\bar{\eta} \in \sigma^{-1}(\bar{v})$, there is $\bar{\eta}'$ that is $\mathbf{a}$-associated with η, $\bar{\eta}$ is $\mathbf{a}$-dominated by $\bar{\eta}'$, and $c(\bar{\eta}, \bar{\eta}') = 0$. Thus,

$$\begin{aligned}
&\min_{\bar{v}' \geq \bar{v}, \bar{v}' \text{ is } \mathbf{1}-associated \text{ with } v} \psi^0(\bar{v}, \bar{v}') \\
&\leq \min_{\bar{v}' \geq \bar{v}, \bar{v}' \text{ is } \mathbf{1}-associated \text{ with } v} \max_{\bar{\eta} \in \sigma^{-1}(\bar{v})} \min_{\bar{\eta}' \in \sigma^{-1}(\bar{v}')} c(\bar{\eta}, \bar{\eta}') \\
&\leq \max_{\bar{\eta} \in \sigma^{-1}(\bar{v})} \min_{\bar{\eta}' \ \mathbf{a}\text{-dominates } \bar{\eta},\ \bar{\eta}' \text{ is } \mathbf{a}-associated \text{ with } \eta} c(\bar{\eta}, \bar{\eta}') = 0,
\end{aligned}$$

and ψ^0 is robustly asymmetric. □

A.3. **Cost function ψ.** We define cost function $\psi : S \times S \to R$ on state space S :

$$\psi(v, v') := \max_{\tilde{v} \geq v} \min_{\tilde{v}' \geq \tilde{v} \vee v'} \psi^0(\tilde{v}, \tilde{v}') \text{ for each } v < v',$$

$$\psi(v, v') := \min_{\tilde{v} \geq v} \min_{\tilde{v}' \in \Lambda(v, v')'} \psi^0(\tilde{v}, \tilde{v}') \text{ for each } v < v',$$

$$\psi(v, v) := 0, \text{ and in all other cases}$$

$$\psi(v, v') := \infty.$$

Lemma 9. *ψ is supermodular.*

Proof. Take any $v, v', \bar{v}$, such that $v \leq \bar{v}$. If $v' \leq v$, then take $\bar{v}' = \bar{v}$ and notice that $0 = \psi(\bar{v}, \bar{v}') \leq \psi(v, v')$. If neither $v' \leq v$ nor $v < v'$, then take $\bar{v}' = \mathbf{1}$. By convention $\infty < \infty$, we have $\psi(v, v') = \infty > \psi(\bar{v}, \mathbf{1})$. Finally, suppose that $v < v'$. Then, because $v \leq \bar{v}$,

$$\begin{aligned}\psi(v, v') &= \max_{\tilde{v} \geq v} \min_{\tilde{v}' \geq \tilde{v} \vee v'} \psi^0(\tilde{v}, \tilde{v}') \leq \max_{\tilde{v} \geq \bar{v}} \min_{\tilde{v}' \geq \tilde{v} \vee v'} \psi^0(\tilde{v}, \tilde{v}') \\ &= \min_{\bar{v}' \geq \bar{v} \vee v'} \max_{\tilde{v} \geq \bar{v}} \min_{\tilde{v}' \geq \tilde{v} \vee \bar{v}'} \psi^0(\tilde{v}, \tilde{v}') = \min_{\bar{v}' \geq \bar{v} \vee v'} \psi(\bar{v}, \bar{v}').\end{aligned}$$

Hence, ψ is supermodular. □

Lemma 10. *If c is (strictly, robustly) asymmetric toward* $\mathbf{a}$, *then ψ is (strictly, robustly) asymmetric toward* $\mathbf{1}$.

Proof. Assume that c is asymmetric toward $\mathbf{a}$. By Lemma 8, ψ^0 is asymmetric toward $\mathbf{1}$. Take any v, v', and $\bar{v}$ such that v and $\bar{v}$ are $\mathbf{1}$-associated. If $v' \geq v$, then take $\bar{v}$ and v' are $\mathbf{1}$-associated and $\psi(\bar{v}, \bar{v}) = 0 \leq \psi(v, v')$. If not $v' > v$ nor $v' < v$, then $\psi(v, v') = \infty \geq \psi(\bar{v}, \mathbf{1})$. If not $v' < v$, then $\Lambda(v, v') = \{v'\}$ and find $v^* \geq v$ such that $\psi(v, v') = \psi^0(v^*, v')$. For any $\bar{v}^* \geq \bar{v}$, $\bar{v}^*$ and v^* are $\mathbf{1}$-associated. By asymmetry of ψ^0, there is $\bar{v}' \geq \bar{v}^*$ such that $\bar{v}'$ and v' are $\mathbf{1}$-associated and $\psi^0(\bar{v}^*, \bar{v}') \leq \psi^0(v^*, v')$. Hence,

$$\psi(\bar{v}, v') = \max_{\tilde{v} \geq v} \min_{\tilde{v}' \geq \tilde{v} \vee v'} \psi^0(\tilde{v}, \tilde{v}') \leq \psi^0(v^*, v'),$$

and ψ^0 is asymmetric toward $\mathbf{1}$.

The same argument (with "associated" replaced by "almost associated") implies that if c is robustly asymmetric then part 2 of Definition 3 holds for ψ.

Suppose that c is either strictly (or robustly) asymmetric. Then, by Lemma 8 $\psi^0(\mathbf{1}, v) > 0$ for any $v \neq \mathbf{1}$. This implies that $\psi(\mathbf{1}, v) > 0$ for any $v \neq \mathbf{1}$ and part 1 of Definition 2 (or Definition 3) holds for ψ.

Suppose that c is strictly asymmetric. Take any $v, v', \bar{v}$, such that $v, \bar{v}$ are $\mathbf{1}$-associated and $\psi(v, v') > 0$. If $v < v'$, take $\bar{v}' := \bar{v}$; then, $\psi(\bar{v}, \bar{v}') = 0 < \psi(v, v')$. If not $v < v'$ nor $v > v'$, then $\psi(v, v') = \infty$. By convention that $\infty < \infty$, $\psi(\bar{v}, \mathbf{1}) < \psi(v, v')$. Finally, if $v > v'$, find $\Lambda(v, v') = \{v'\}$ and find $v^* \geq v$ such that $\psi(v, v') = \psi^0(v^*, v')$. For any $\bar{v}^* \geq \bar{v}$, $\bar{v}^*$ and v^* are $\mathbf{1}$-associated. By strict asymmetry of ψ^0, there is $\bar{v}' \geq \bar{v}^*$ such that $\bar{v}'$ and v' are $\mathbf{1}$-associated and $\psi^0(\bar{v}^*, \bar{v}') < \psi^0(v^*, v')$. Hence,

$$\psi(\bar{v}, v') = \max_{\tilde{v} \geq v} \min_{\tilde{v}' \geq \tilde{v} \vee v'} \psi^0(\tilde{v}, \tilde{v}') < \psi^0(v^*, v'),$$

and ψ^0 is strictly asymmetric toward $\mathbf{1}$.

Suppose that c is robustly asymmetric. A similar argument to the one given above shows that condition (3) of Definition 3 holds for ψ. We will show that part 4 of Definition 3 holds for ψ. Take v,v', and $\bar{v}$ such that v' is almost $\mathbf{1}$-dominated by v and $\bar{v}$ is $\mathbf{1}$-associated with v'. If $v > v'$, then, for any $\bar{v}$ that is $\mathbf{1}$-associated with v. $\bar{v}$ is also $\mathbf{1}$-associated with v and $\psi(\bar{v}, \bar{v}) = 0$. If not $v < v'$ nor $v > v'$, then $\psi(v, v') = \infty$, and by convention that $\infty < \infty$, $\psi(\bar{v}, \mathbf{1}) < \psi(v, v')$. Finally, if $v < v'$, and $\psi(v, v') < 0$, then find $v^* \geq v$ such that

$$\psi(v, v') = \min_{v'' \geq v^* \vee v'} \psi^0(v^*, v'') > 0.$$

Find $v'^* > v^*$ such that v^* almost $\mathbf{1}$-dominates v'^*, and $v'^* \geq v^* \vee v'$. Then, $\psi^0(v^*, v'^*) > 0$. For any $\bar{v}^* \geq \bar{v}$, $\bar{v}^*$ is $\mathbf{1}$-associated with v'^*. By robust asymmetry of ψ^0, there is $\bar{v}'^* \geq \bar{v}^*$ such that $\bar{v}'^*$ is $\mathbf{1}$-associated with $v \leq v^*$ and $\psi^0(\bar{v}^*, \bar{v}'^*) = 0$. Thus,

$$\min_{\bar{v}' \geq \bar{v}, \bar{v}' \text{ is } \mathbf{1}-\textit{associated} \text{ with } v} \psi(\bar{v}, \bar{v}') \leq \max_{\bar{v}^* \geq \bar{v}} \min_{\bar{v}'^* \geq \bar{v}^*, \bar{v}'^* \text{ is } \mathbf{1}-\textit{associated} \text{ with } v} \psi^0(\bar{v}^*, \bar{v}'^*) = 0,$$

and ψ satisfies condition (4) of Definition 3. □

A.4. **Auxiliary family Q^α.** Take $\alpha \in \left(0, \frac{1}{|S|}\right)$ and define auxiliary process $Q^\alpha_{\varepsilon_n}$:

$$Q^\alpha_{\varepsilon_n}(v, v') = \alpha Q^0_{\varepsilon_n}(v, v') \text{ for each } v \neq v', \text{ and}$$
$$Q^\alpha_{\varepsilon_n}(v, v) = 1 - \alpha + \alpha Q^0_{\varepsilon_n}(v, v).$$

It is easy to check that $Q^\alpha_{\varepsilon_n}$ is ergodic, the stationary distribution of $Q^\alpha_{\varepsilon_n}$ is equal to $q^\alpha_{\varepsilon_n} = q^0_{\varepsilon_n}$ and

$$\lim_{n\to\infty} \frac{1}{\log \varepsilon_n} \log Q^\alpha_{\varepsilon_n}(v, v') = \psi^0(v, v')$$

For $\alpha \leq \frac{1}{|S|}$, we can decompose $Q^\alpha_{\varepsilon_n}(v) = \frac{1}{2} Q^{\alpha,U}_{\varepsilon_n}(v) + \frac{1}{2} Q^{\alpha,D}_{\varepsilon_n}(v)$ such that $Q^{\alpha,U}_{\varepsilon_n}(v), Q^{\alpha,D}_{\varepsilon_n}(v) \in \Delta S$ and

$$Q^{\alpha,U}_{\varepsilon_n}(v)(\{v' : v' \geq v\}) = 1 - Q^{\alpha,D}_{\varepsilon_n}(v)(\{v' : v' \geq v\}) = 0.$$

A.5. **Auxiliary family Q.** We construct Markov processes $Q_{\varepsilon_n}(v)$ so that (A.1) holds, Q_{ε_n} is monotonic, and $Q_{\varepsilon_n} \leq Q^\alpha_{\varepsilon_n}$. Fix $\gamma < \frac{1}{2|S|}\alpha$ and let $\gamma(v) = \gamma^{\#\{i:v_i=1\}}$ and define

$$Q^U_{\varepsilon_n}(v, v') := \gamma^{\#\{i:v_i=0\}+1} \min_{\tilde{v} \geq v} \max_{\tilde{v}' \geq \tilde{v} \vee v'} Q^{\alpha,U}_{\varepsilon_n}(\tilde{v}, \tilde{v}') \text{ for } v < v',$$
$$Q^U_{\varepsilon_n}(v, v) := 1 - \sum_{v' > v} Q^U_{\varepsilon_n}(v, v'),$$
$$Q^U_{\varepsilon_n}(v, v') := 0 \text{ in all other cases},$$
$$Q^D_{\varepsilon_n}(v, v') := \sum_{\tilde{v} \geq v} \sum_{\tilde{v}' \in \Lambda(v, v')} Q^{\alpha,D}_{\varepsilon_n}(\tilde{v}, \tilde{v}') \text{ for } v' < v,$$
$$Q^D_{\varepsilon_n}(v, v) := 1 - \sum_{v' < v} Q^D_{\varepsilon_n}(v, v')$$
$$Q^D_{\varepsilon_n}(v, v') := 0 \text{ in all other cases}.$$

Because $\alpha, \gamma \leq \frac{1}{|S|}$, $Q^U_{\varepsilon_n}(v,v), Q^D_{\varepsilon_n}(v,v) \geq 0$. Thus, Markov processes $Q^U_{\varepsilon_n}$ and $Q^D_{\varepsilon_n}$are well-defined. Let

$$Q_{\varepsilon_n}(v) = \frac{1}{2}Q^U_{\varepsilon_n}(v) + \frac{1}{2}Q^D_{\varepsilon_n}(v).$$

It is easy to check that the ergodicity of $Q^0_{\varepsilon_n}$ implies ergodicity of $Q_{\varepsilon_n}(v)$. Let q_{ε_n} be the stationary distribution of Q_{e_n}

Lemma 11. *(A.1) holds.*

Proof. If $v < v'$, then

$$\begin{aligned}
&\lim_{n\to\infty} \frac{1}{\log \varepsilon_n} \log Q_{\varepsilon_n}(v,v') \\
&= \lim_{n\to\infty} \frac{1}{\log \varepsilon_n} \log Q^U_{\varepsilon_n}(v,v') \\
&= \lim_{n\to\infty} \frac{1}{\log \varepsilon_n} \left(\log \gamma^{\#\{i:v_i=0\}+1} + \log \min_{\tilde{v}\geq v} \max_{\tilde{v}'\geq \tilde{v}\vee v'} Q^{\alpha,U}_{\varepsilon_n}(\tilde{v},\tilde{v}')\right) \\
&= \min_{\tilde{v}\geq v} \max_{\tilde{v}'\geq \tilde{v}\vee v'} \lim_{n\to\infty} \frac{1}{\log \varepsilon_n} Q^{\alpha,U}_{\varepsilon_n}(\tilde{v},\tilde{v}') \\
&= \max_{\tilde{v}\geq v} \min_{\tilde{v}'\geq \tilde{v}\vee v'} \psi^0(\tilde{v},\tilde{v}') = \psi(v,v'),
\end{aligned}$$

where the second to last equality follows from the fact that limits (A.2) exist. If $v > v'$,

$$\begin{aligned}
&\lim_{n\to\infty} \frac{1}{\log \varepsilon_n} \log Q_{\varepsilon_n}(v,v') \\
&= \lim_{n\to\infty} \frac{1}{\log \varepsilon_n} \log Q^D_{\varepsilon_n}(v,v') \\
&= \lim_{n\to\infty} \frac{1}{\log \varepsilon_n} \log \sum_{\tilde{v}\geq v} \sum_{\tilde{v}'\in\Lambda(v,v')} Q^{\alpha,D}_{\varepsilon_n}(\tilde{v},\tilde{v}') \\
&= \min_{\tilde{v}\geq v} \min_{\tilde{v}'\in\Lambda(v,v')} \psi^0(\tilde{v},\tilde{v}') = \psi(v,v').
\end{aligned}$$

In all the other cases, the result is trivial. □

Lemma 12. *Q_{ε_n} is monotonic.*

Proof. It is enough to show that $Q^U_{\varepsilon_n}$ and $Q^D_{\varepsilon_n}$ are monotonic. First, we show that $Q^U_{\varepsilon_n}$ is monotonic. Take any upper $E \subseteq S$ and $v^0 < v^1$. If $v_1 \in E$, then $Q^U_{\varepsilon_n}(v^0, E) \leq 1 =$

$Q^U_{\varepsilon_n}(v^1, E)$. If $v_1 \notin E$, then

$$\begin{aligned} Q^U_{\varepsilon_n}\left(v^0, E\right) &= \gamma^{\#\left\{i:v^0_i=0\right\}+1} \sum_{v' \in E, v' \geq v^0} \min_{\tilde{v} \geq v^0} \max_{\tilde{v}' \geq \tilde{v} \vee v'} Q^{\alpha,U}_{\varepsilon_n}(\tilde{v}, \tilde{v}') \\ &\leq \gamma^{\#\left\{i:v^0_i=0\right\}+1} \sum_{v' \in E, v' \geq v^0} \min_{\tilde{v} \geq v^1} \max_{\tilde{v}' \geq \tilde{v} \vee v'} Q^{\alpha,U}_{\varepsilon_n}(\tilde{v}, \tilde{v}') \\ &= \gamma^{\#\left\{i:v^0_i=0\right\}+1} \sum_{v' \in E, v' \geq v^0} \min_{\tilde{v} \geq v^1} \max_{\tilde{v}' \geq \tilde{v} \vee (v' \vee v^1)} Q^{\alpha,U}_{\varepsilon_n}(\tilde{v}, \tilde{v}') \\ &\leq \gamma^{\#\left\{i:v^0_i=0\right\}+1} \sum_{v' \in E, (v' \vee v^1) \geq v^1} \min_{\tilde{v} \geq v^1} \max_{\tilde{v}' \geq \tilde{v} \vee (v' \vee v^1)} Q^{\alpha,U}_{\varepsilon_n}(\tilde{v}, \tilde{v}') \\ &\leq \gamma^{\#\left\{i:v^0_i=0\right\}+1} |S| \sum_{v' \in E, v' \geq v^1} \min_{\tilde{v} \geq v^1} \max_{\tilde{v}' \geq \tilde{v} \vee v'} Q^{\alpha,U}_{\varepsilon_n}(\tilde{v}, \tilde{v}') \\ &\leq \gamma^{\#\left\{i:v^1_i=0\right\}} \sum_{v' \in E, v' \geq v^1} \min_{\tilde{v} \geq v^1} \max_{\tilde{v}' \geq \tilde{v} \vee v'} Q^{\alpha,U}_{\varepsilon_n}(\tilde{v}, \tilde{v}') = Q^U_{\varepsilon_n}\left(v^1, E\right). \end{aligned}$$

where the second equality comes from the fact that for each $\tilde{v} \geq v^1$ and $\tilde{v}' \geq \tilde{v} \vee v'$, it must be that $\tilde{v}' \geq \tilde{v} \vee (v' \vee v^1)$; the third inequality comes from the fact that if $v' \in E$ and E is upper, then $v' \vee v^1 \in E$ and that there is at most $|S|$ elements in set E; and the fourth inequality follows from $\gamma |S| \leq 1$ and $\#\{i : v^1_i = 0\} < \#\{i : v^0_i = 0\}$.

Next, we show that $Q^D_{\varepsilon_n}$ is monotonic. Take any upper $E \subseteq S$ and $v^0 < v^1$. If $v^0 \notin E$, then $Q^D_{\varepsilon_n}(v^0, E) = 0 \leq Q^D_{\varepsilon_n}(v^1, E)$. If $v^0 \in E$, then $v^1 \in E$, and

$$\begin{aligned} Q^D_{\varepsilon_n}\left(v^0, S \backslash E\right) &= \sum_{\tilde{v} \geq v^0} \sum_{\substack{\tilde{v}' : \tilde{v}' \wedge v^0 \notin E, \\ \text{and not } \tilde{v}' \geq v^0,}} Q^{\alpha,D}_{\varepsilon_n}(\tilde{v}, \tilde{v}') \\ &\geq \sum_{\tilde{v} \geq v^1} \sum_{\substack{\tilde{v}' : \tilde{v}' \wedge v^1 \notin E, \\ \text{and not } \tilde{v}' \geq v^1,}} = Q^D_{\varepsilon_n}\left(v^1, S \backslash E\right). \end{aligned}$$

The last inequality follows from the fact that if $\tilde{v} \geq v^1$, then $\tilde{v} \geq v^0$, if $\tilde{v}' \wedge v^1 \notin E$, then $\tilde{v}' \wedge v^0 \leq \tilde{v}' \wedge v^0 \notin E$, and because $v^0 \in E$, if $\tilde{v}' \wedge v^1 \notin E$, then it cannot be that $\tilde{v}' \geq v^0$. □

Lemma 13. *$Q_{\varepsilon_n}(v) \leq Q^{\alpha}_{\varepsilon_n}(v)$ for each $v \in S$.*

Proof. It is enough to show that $Q^U_{\varepsilon_n}(v) \leq Q^{\alpha,U}_{\varepsilon_n}(v)$ and $Q^D_{\varepsilon_n}(v) \leq Q^{\alpha,D}_{\varepsilon_n}(v)$. First, we show that for any upper E, any $v \in S$, $Q^U_{\varepsilon_n}(v, E) \leq Q^{\alpha,U}_{\varepsilon_n}(v, E)$. If $v \in E$, then $1 = Q^U_{\varepsilon_n}(v, E) =$

$Q^{\alpha,U}_{\varepsilon_n}(v,E)$. If $v \notin E$, then $\#\{i : v_i = 0\} \geq 1$, and

$$\begin{aligned} Q^{U}_{\varepsilon_n}(v,E) &= \gamma^{\#\{i:v_i=0\}+1} \sum_{v' \in E, v' > v} \min_{\tilde{v} \geq v} \max_{\tilde{v}' \geq \tilde{v} \vee v'} Q^{\alpha,U}_{\varepsilon_n}(\tilde{v},\tilde{v}') \\ &\leq \gamma^2 \sum_{v' \in E, v' > v} \max_{\tilde{v}' \geq v \vee v'} Q^{\alpha,U}_{\varepsilon_n}(v,\tilde{v}') \\ &\leq \gamma^2 |S| \max_{\tilde{v}' > v} Q^{\alpha,U}_{\varepsilon_n}(v,\tilde{v}') \\ &\leq \gamma^2 |S|^2 \sum_{v' \in E, v' > v} Q^{\alpha,U}_{\varepsilon_n}(v,v') \leq Q^{\alpha,U}_{\varepsilon_n}(v,E), \end{aligned}$$

where the last inequality follows from $\gamma |S| \leq 1$ and the fact that set E cannot contain more than $|S|$ elements. Next, we show that for any upper E, any $v \in S$, $Q^{D}_{\varepsilon_n}(v,E) \leq Q^{\alpha,D}_{\varepsilon_n}(v,E)$. If $v \notin E$, then $Q^{D}_{\varepsilon_n}(v,E) = 0 \leq Q^{\alpha,D}_{\varepsilon_n}(v,E)$. If $v \in E$, then

$$\begin{aligned} Q^{D}_{\varepsilon_n}(v,S\backslash E) &= \sum_{v' < v, v' \notin E} \sum_{\tilde{v} \geq v} \sum_{\tilde{v}' \in \Lambda(v,v')} Q^{\alpha,D}_{\varepsilon_n}(\tilde{v},\tilde{v}') \\ &\geq \sum_{v' < v, v' \notin E} \sum_{\tilde{v}' \in \Lambda(v,v')} Q^{\alpha,D}_{\varepsilon_n}(v,\tilde{v}') \\ &= \sum_{v' \notin E} Q^{\alpha,D}_{\varepsilon_n}(v,v') = Q^{\alpha,D}_{\varepsilon_n}(v,S\backslash E). \end{aligned}$$

□

Lemma 14. $\mu_\varepsilon(\mathbf{a}) = q^0_\varepsilon(\mathbf{1}) = q^\alpha_\varepsilon(\mathbf{1}) \geq q_\varepsilon(\mathbf{1})$.

Proof. This follows from Lemmas 7, 12, and 13. □

Appendix B. Proof of Lemma 3

The algorithm of finding minimal tree roots is defined as a

(1) sequence of sets, $V^0 = S \supseteq V^1 \supseteq ... \supseteq V^n$,
(2) sequence of costs functions $\psi^k : V^k \times V^k \to [0,\infty)$, such that $\psi^k(v,v) = 0$ and
(3) sequence of projections $j^k : V^{k-1} \to V^k$, such that $j^k(v) = v$ for any $v \in V^k$.

Let $j^0 := \mathrm{id}_S$ and denote composition $J^k = j^k \circ ... \circ j^0 : V \to V^k$. Thus, $J^k(v) = v$ for each $v \in V^k$.

Below, we describe how to choose sets, cost functions and projections in a way that allows to recover $MR(V^k,\psi^k)$ from $MR(V^{k+1},\psi^{k+1})$. Next, we discuss how to trace properties of cost functions ψ^k along steps k. Finally, we use these properties to prove Lemma 3.

B.1. **Tracing the algorithm - roots.** Take a sequence of sets and projection functions as given. Let $\psi_{\min} = \min_{v\neq v'}\psi(v,v')$. Let $\psi^0(v,v') = \psi(v,v') - \psi_{\min}$ for any $v\neq v'$ and $\psi(v,v)=0$. For any $k\geq 0$, define inductively: for any $v,v'\in V^{k+1}$

$$\begin{aligned}\tilde{\psi}^{k+1}(v,v') &:= \min_{v_0,v_1,\ldots,v_l\in V^k,\text{ st. } v=v_0,\ v'=v_s}\sum\nolimits_{t=0}^{s-1}\psi^k(v_t,v_{t+1})\\ \psi^{k+1}(v,v') &:= \tilde{\psi}^{k+1}(v,v') - \min_{u\neq u'\text{ and } u,u'\in V^{k+1}}\tilde{\psi}^{k+1}(u,u).\end{aligned} \tag{B.1}$$

The cost $\tilde{\psi}^{k+1}$ is equal to the minimum of ψ^{k+1}-costs across all paths in set V^k that link states v and v'. The cost function ψ^{k+1} is obtained from $\tilde{\psi}^{k+1}$ by subtracting a constant chosen so that ψ^{k+1} has 0-cost transitions. From now on, we assume that cost functions are defined as above.

Say that there is a 0-*cost path* between $v,v'\in V^k$ if there is a sequence of states $v_0,v_1,...,v_l\in V^k$ st. $v=v_0$, $v'=v_s$ so that $\sum_{t=0,}^{s-1}\psi^k(v_t,v_{t+1})=0$. A state $v\in V^k$ is called k-attractor if for any $v'\in V^k$, if there is a 0-cost path from v to v', then there is a 0-cost path from v' to v. For any $v\in V^k$, let $U^k(v)$ denote the set of all k-attractors v' such that there is 0-cost path from v to v'.

Lemma 15. *Suppose that for each k, $v\in V^k$, $j^{k+1}(v)\in U^k(v)$. Then,*

$$MR\left(V^k,\psi^k\right) = \bigcup_{v\in MR\left(V^{k+1},\psi^{k+1}\right)} U^k(v). \tag{B.2}$$

If the sequence of projections satisfies the assumption of the Lemma, then formula (B.2) leads to a simple procedure of recovering $MR\left(V^k,\psi^k\right)$ from $MR\left(V^{k+1},\psi^{k+1}\right)$.

Proof. This result is standard and it is satisfied by most known algorithms (Chu and Liu (1965), Edmonds (1967); Proposition 1 of Nöldeke and Samuelson (1993) and the Appendix in Young (1993) contain a version of the Lemma for $k=1$). We present the proof for the sake of completeness. Let $V^{k+1} = \operatorname{Im} j^{k+1}\subseteq\bigcup_v U^k(v)$.

(1) For any k

$$MR\left(V^k,\psi^k\right)\subseteq\bigcup_{v\in V^k}U^k(v).$$

Suppose that h is a tree with a root that is not a k-attractor, $v_h\notin\bigcup_{v\in V^k}U^k(v)$. Then, there is at least one state $v^*\in U^k(v_h)$, such that $\psi(v^*,h(v^*))>0$. There is also a 0-cost path from v_h to v^*. One can modify h into a new tree h^* with the root at v^*, where the only changes are those that are necessary to connect v_h to v^* via the 0-cost path. The cost of tree h^* is lower than the cost of h by $\psi(v^*,h(v^*))$.

(2) For any tree h on $\left(V^k,\psi^k\right)$ with the root at v_h, there is a tree h^* on $\left(V^{k+1},\tilde{\psi}^{k+1}\right)$ with at most the same cost and a root at $j^{k+1}\left(v_h\right)\in U^k\left(v_h\right)$. This is a consequence of the definition of $\tilde{\psi}^{k+1}$.

(3) For any tree h^* on $\left(V^{k+1},\tilde{\psi}^{k+1}\right)$ with a root at v_h^*, for any $v_h\in U^k\left(v_h^*\right)$, there is a tree h on $\left(V^k,\psi^k\right)$ with at most the same cost and the root at v_h. Indeed, let $n:\left\{1,...,\left|V^{k+1}\right|-1\right\}\to V^{k+1}\backslash\left\{v_h^*\right\}$ be an enumeration of set $V^{k+1}\backslash\left\{v_h^*\right\}$ with the property that if a path from v to v_h^* passes through v', then $n^{-1}\left(v\right)<n^{-1}\left(v'\right)$. By induction on n, we can construct a sequence of functions $h^{(n)}:H^{(n)}\to V^k$, where

$$V^k/V^{k+1}\cup n^{-1}\left\{1,...,n\right\}\subseteq H^{(n)}\subseteq V^k$$

and such that for each $v\in H^{(n)}$, there is a unique path along h to some $v'\notin H^{(n)}$ and

$$\sum_{v\in H^{(n)}}\psi\left(v,h^{(n)}\left(v\right)\right)\leq\sum_{m\leq n}\psi\left(v,h^{(n)}\left(v\right)\right)$$

Indeed, let $h^{(0)}$ consist of 0-cost paths which connect each $v\notin V^{k+1}$ with $j^{k+1}\left(v\right)$ and then inductively modify $h^{(m-1)}$ to $h^{(m)}$ by adding lowest cost path (on V^k) between $n\left(m-1\right)$ to $n\left(m\right)$. This gives a tree $h'=h^{\left(\left|V^{k+1}\right|-1\right)}$ on V^k with a cost at most equal to the cost of tree h^* and with a root at some $v'\in U^k\left(v_h^*\right)$. Such a tree can be easily modified to a tree with exactly the same cost and the root at $v_h\in U^k\left(v_h^*\right)$.

(4) By (2) and (3), the cost of minimal cost tree in problem $\left(V^k,\psi^k\right)$ is equal to the cost of the minimal cost tree in problem $\left(V^k,\tilde{\psi}^k\right)$. Let us denote the cost as $c_{\min}^k$. Then, $MR\left(V^k,\psi^k\right)$ is equal to the set of all elements of V^k that are roots of trees with ψ^k-cost $c_{\min}$; similarly, $MR\left(V^{k+1},\tilde{\psi}^{k+1}\right)$ is equal to the set of all elements of V^{k+1} that are roots of trees with $\tilde{\psi}^{k+1}$-cost $c_{\min}$.

(5) Take any $v\in MR\left(V^k,\psi^k\right)$. By (2) and (4), $j^{k+1}\left(v_h\right)\in MR\left(V^{k+1},\tilde{\psi}^{k+1}\right)$. By (1),

$$MR\left(V^k,\psi^k\right)\subseteq\bigcup_{v\in MR\left(V^{k+1},\tilde{\psi}^{k+1}\right)}U^k\left(v\right).$$

(6) Take any $v^*\in MR\left(V^{k+1},\tilde{\psi}^{k+1}\right)$. By (3) and (4), $U^k\left(v^*\right)\subseteq MR\left(V^k,\psi^k\right)$, and

$$\bigcup_{v\in MR\left(V^{k+1},\tilde{\psi}^{k+1}\right)}U^k\left(v\right)\subseteq MR\left(V^k,\psi^k\right)..$$

(7) Notice that any minimal cost tree on $\left(V^{k+1},\tilde{\psi}^{k+1}\right)$ is also a minimal cost tree on $\left(V^{k+1},\psi^{k+1}\right)$ and vice versa (subtracting a constant from the cost function does

not change the comparison of cost between trees.) Thus, $MR\left(V^{k+1}, \tilde{\psi}^{k+1}\right) = MR\left(V^{k+1}, \psi^{k+1}\right)$.

□

B.2. **Tracing the algorithm - supermodularity.** Next, we describe how to choose $j^{k+1}(v) \in U^k(v)$. Step k of the algorithm is *supermodular* if

- for any $v, v' \in S$, if $v \leq v'$, then $J^k(v) \leq J^k(v')$ and
- for any $v, v', \bar{v} \in V^k$, st. $v \leq \bar{v}$, there is $\bar{v}' \geq v', \bar{v}$, such that

$$\psi^k(v, v') \geq \psi^k\left(\bar{v}, J^k(\bar{v}')\right). \tag{B.3}$$

For any $v, v' \in S$, define $v \vee v' = \inf\{v'' : v'' \geq v, v'\}$ as the *joint* of v and v' (i.e., for each player i, $(v \vee v')_i = 1$ iff $v_i = 1$ or $v'_i = 1$). The next result guarantees that supermodularity is inherited along the sequence of steps.

Lemma 16. *If step k is supermodular, then for any $v \in V^k$.*

$$\bigvee\nolimits_{v' \in U^k(v)} v' \in U^k(v),$$

i.e., set $U^k(v)$ contains its largest element. Define

$$j^{k+1}(v) := \bigvee\nolimits_{v' \in U^k(v)} v' \text{ and} \tag{B.4}$$

$$V^{k+1} := j^{k+1}\left(V^k\right) = \left\{\bigvee\nolimits_{v' \in U^k(v)} v' : v \in V^k\right\}.$$

Then, step $k+1$ is supermodular.

Proof. Suppose that $u, u' \in V^k$ are k-attractors that are connected by a 0-cost path in V^k. Let $u = v_0, v_1, ..., v_s = u'$ be a 0-cost path between u and u'. Inductively construct a 0-cost path path $v_0 := v_0^a \leq ... \leq v_s^a$ such that $v_s^a \geq u, u'$. (This can be done as follows: For any $t = 0, ..., s-1$, suppose that v_t^a is constructed and $v_t^a \geq v_t, v_{t-1}^a$. By the supermodularity of kth step, there is $v_{t+1}^a \geq v_{t+1}, v_t^a$, such that $\psi^k\left(v_t^a, v_{t+1}^a\right) = 0$.) Thus, v_s^a is a k–attractor and there is a 0-cost path from v_s^a to both u and u'. Hence, for any two k-attractors that are connected by a 0-cost path there is another k-attractor that is larger than both of them and connected to them by a 0-cost path. Because set V^k is finite, this shows that any subset of k-attractors $U^k(v)$ contains its largest element and demonstrates the first part of the Lemma.

A similar argument shows that if $v, v' \in V^k$ and $J^k(v) \leq J^k(v')$, then $\bigvee_{u \in U^k\left(J^k(v)\right)} u \leq \bigvee_{u \in U^k\left(J^k(v')\right)} u$, which implies that $J^{k+1}(v) \leq J^{k+1}(v')$.

Next, take $v, v', \bar{v} \in V^{k+1}$ st. $v \leq \bar{v}$, and suppose that $v = v_0, ..., v_s \in V^k$ is the minimal cost path from v to $v' = v_s$:

$$\sum_{t=0}^{s-1} \psi^k (v_t, v_{t+1}) = \tilde{\psi}^{k+1} (v, v').$$

The same argument as above leads to the existence of a path $\bar{v} = v_0^a \leq ... \leq v_s^a$ such that $v_t^a \geq \bar{v}, v_t$ for each $t \leq s$ and

$$\psi^k (v_t, v_{t+1}) \geq \psi^k \left(v_t^a, v_{t+1}^a\right) \text{ for any } t = 0, ..., s-1.$$

By construction, there is a 0-cost path from v_s^a to $j^{k+1} (v_s^a) = J^{k+1} (v_s^a)$. Thus,

$$\tilde{\psi}^{k+1} \left(\bar{v}, J^k (\bar{v}')\right) \leq \sum_{t=0}^{s-1} \psi^k \left(v_t^a, v_{t+1}^a\right) = \tilde{\psi}^{k+1} (v, v').$$

Because $v_s^a \geq \bar{v}, v_t$, by the first part of this proof,

$$J^{k+1} (v_s^a) \geq J^{k+1} (\bar{v}) = \bar{v} \text{ and}$$
$$J^{k+1} (v_s^a) \geq J^{k+1} (v_s) = J^{k+1} (v') = v'.$$

□

Because $J^0 = \text{id}_S$, if cost function $\psi^0 = \psi$ is supermodular, then step 0 is supermodular. By induction on k, define $j^{k+1} (v)$ as in (B.4). Then, the Lemma implies that each step k is supermodular.

B.3. **Tracing the algorithm - asymmetry.** Step k is *asymmetric* (*strictly asymmetric*) if

- for any $v, \bar{v} \in S$, if $v, \bar{v}$ are **1**-associated, then $J^k (v), J^k (\bar{v})$ are **1**-associated and
- for any states $v, v', \bar{v} \in V^k$, such that $v, \bar{v}$ are **1**-associated, there is $\bar{v}' \in V^k$, such that $\bar{v} \leq \bar{v}'$, states $v', \bar{v}'$ are **1**-associated and

$$\psi^k (\bar{v}, \bar{v}') \leq \psi^k (v, v') \quad (\ \psi^k (\bar{v}, \bar{v}') < \psi^k (v, v')\ ). \tag{B.5}$$

The next lemma guarantees that the asymmetry is inherited along the sequence of steps:

Lemma 17. *Suppose that step k is supermodular and j^{k+1} is defined by (B.4). If step k is asymmetric (strictly asymmetric), then step $k+1$ is asymmetric (strictly asymmetric).*

Proof. Suppose that $v, \bar{v} \in S$ are **1**-associated. In two steps, we show that $J^{k+1} (v), J^{k+1} (\bar{v})$ are **1**-associated. First, observe that by the asymmetry of kth step, $J^k (v), J^k (\bar{v})$ are **1**-associated. Suppose that $J^k (v) = v_0, ..., v_s = J^{k+1} (v)$ is a 0-cost path from $J^k (v)$ to $J^{k+1} (v) = j^{k+1} \left(J^k (v)\right)$: for each $t < s$, $\psi^k (v_t, v_{t+1}) = 0$. Inductively construct a path $J^k (\bar{v}) := v_0^a \leq ... \leq v_s^a$ such that

- for each $t < s$, $\psi^k \left(v_t^a, v_{t+1}^a\right) = 0$ and
- for each $t \leq s$, v_t, v_t^a are **1**-associated.

(This can be done as follows: For any $t = 0, ..., s-1$, suppose that v_t^a is such that states v_t^a, v_t are **1**-associated. By the asymmetry of kth step, there is $v_{t+1}^a \geq v_t^a$, such that states v_{t+1}^a, v_{t+1} are **1**-associated and $\psi^k\left(v_t^a, v_{t+1}^a\right) = 0$.) Hence, states $v_s^a, v_s = J^{k+1}(v)$ are **1**-associated and there is a 0-cost path from $J^k(\bar{v})$ to $v_s^a \geq J^k(\bar{v})$. By Lemma 16 and the definition of j^{k+1},

$$J^{k+1}(\bar{v}) = j^{k+1}\left(J^k(\bar{v})\right) = J^{k+1}(v_s^a).$$

Second, let $v_s^a = v_0^b, ..., v_{s'}^b = J^{k+1}(\bar{v})$ be a 0-cost path from v_s^a to $J^{k+1}(v_s^a)$. Inductively construct a path $J^{k+1}(v) := v_0^c \leq ... \leq v_{s'}^c$ such that

- for each $t < s'$, $\psi^k\left(v_t^c, v_{t+1}^c\right) = 0$ and
- for each $t \leq s$, v_t^b, v_t^c are **1**-associated.

(This can be done as follows: For any $t = 0, ..., s-1$, suppose that v_t^c is such that states v_t^b, v_t^c are **1**-associated. By the asymmetry of kth step, there is $v_{t+1}^c \geq v_t^c$, such that states v_{t+1}^b, v_{t+1}^c are **1**-associated and $\psi^k\left(v_t^c, v_{t+1}^c\right) = 0$.) Hence, states $J^{k+1}(\bar{v}), v_{s'}^c$ are **1**-associated and there is a 0-cost path from $J^{k+1}(v)$ to $v_s^c \geq J^{k+1}(v)$. Because of Lemma 16 and the definition of j^{k+1}, $J^{k+1}(v) = v_s^c$. Hence, states $J^{k+1}(v), J^{k+1}(\bar{v})$ are **1**-associated.

Next, take $v, v', \bar{v} \in V^{k+1}$ st. $v, \bar{v}$ are **1**-associated and suppose that $v = v_0, ..., v_s \in V^k$ is a minimal cost path from v to $v' = v_s$:

$$\sum_{t=0}^{s-1} \psi^k(v_t, v_{t+1}) = \tilde{\psi}^{k+1}(v, v').$$

Inductively construct path $\bar{v} := v_0^a \leq ... \leq v_s^a$: For any $t = 0, ..., s-1$, suppose that v_t^a is constructed such that $v_t^a \geq v_{t-1}^a$ and v_t^a, v_t are **1**-associated. By the asymmetry of kth step (strict asymmetry of kth step), there is $v_{t+1}^a \geq v_{t+1}$, such that states v_{t+1}^a, v_{t+1} are **1**-associated and such that

$$\psi^k\left(v_t^a, v_{t+1}^a\right) \leq \psi^k(v_t, v_{t+1}) \;\; \left(\; \psi^k\left(v_t^a, v_{t+1}^a\right) < \psi^k(v_t, v_{t+1}) \;\right).$$

It is shown above that, if states v', v_s^a are **1**-associated, then states $v' = J^{k+1}(v'), J^{k+1}(v_s^a)$ are **1**-associated. By construction, there is a 0-cost path from v_s^a to $\bar{v}' := J^{k+1}(v_s^a)$. Hence,

$$\tilde{\psi}^{k+1}(v, v') \leq \tilde{\psi}^{k+1}(\bar{v}, \bar{v}') \;\; (\tilde{\psi}^{k+1}(v, v') < \tilde{\psi}^{k+1}(\bar{v}, \bar{v}')).$$

□

Because $J^0 = \mathrm{id}_S$, if cost function $\psi^0 = \psi$ is asymmetric, then step 0 is asymmetric. The Lemma implies that each step k is asymmetric. The next Lemma shows that if the cost function is strictly or robustly asymmetric, then step 1 is strictly asymmetric.

Lemma 18. *Suppose that step* 0 *is supermodular and* j^1 *is defined by (B.4). If cost function* ψ *is strictly or robustly asymmetric, then step* 1 *is strictly asymmetric.*

Proof. Take any states $v, v', \bar{v} \in V^1$, such that $v, \bar{v}$ are **1**-associated. Because $v \in V^1$ and j^1 is defined by (B.4), there is no state $v'' > v$ such that $\psi(v, v'') = 0$. We show that if ψ is strictly or robustly asymmetric then there is $\bar{v}' \in V^1$, such that $\bar{v} \leq \bar{v}'$, states $v', \bar{v}'$ are **1**-associated and

$$\psi^1(\bar{v}, \bar{v}') < \psi^1(v, v'). \tag{B.6}$$

This will demonstrate that step 1 is strictly asymmetry.

Suppose first that ψ is strictly asymmetric. Let $v = v_0, ..., v_s = v'$ be a minimal cost path between v and v' :

$$\tilde{\psi}^1(v, v') = \sum_{t=0}^{s-1} \psi(v_t, v_{t+1}).$$

Because $v, v' \in V^1$ and $v \neq v'$, it must be that $\tilde{\psi}^1(v, v') > 0$. Thus, there is t st. $\psi(v_t, v_{t+1}) > 0$. Strict asymmetry of ψ allows to inductively construct a path $\bar{v} = \bar{v}_0, ..., \bar{v}_s = \bar{v}'$ such that

- for any $t \leq s$, $v_t, \bar{v}_t$ are **1**-associated,
- for any $t < s$, $\bar{v}_t \leq \bar{v}_{t+1}$,
- for any $t < s$, either $\psi(v_t, v_{t+1}) = \psi(\bar{v}_t, \bar{v}_{t+1}) = 0$, or $\psi(v_t, v_{t+1}) > \psi(\bar{v}_t, \bar{v}_{t+1})$.

Then, $\bar{v} \leq \bar{v}'$, states v' and $\bar{v}'$ are **1**-associated and $\tilde{\psi}^1(\bar{v}, \bar{v}') < \tilde{\psi}^1(v, v')$. The latter implies that (B.6) holds.

Next, suppose that ψ is robustly asymmetric and let $v = v_0, ..., v_s = v'$ be a minimal cost path between v and v'. Because $v, v' \in V^1$ and $v \neq v'$, $\tilde{\psi}^1(v, v') > 0$ and there is $t^* < s$ st. $\psi(v_{t^*}, v_{t^*+1}) > 0$. Using robust asymmetry, construct a path $\bar{v} = \bar{v}_0, .., \bar{v}_{t^*}, .., \bar{v}_{s+1} = \bar{v}'$ such that

- for any $t \leq t^*$, $v_t, \bar{v}_t$ are **1**-associated; for any $t^* < t \leq s$, $v_t, \bar{v}_t$ are almost **1**-associated; $v_s, \bar{v}_{s+1}$ are **1**-associated.
- for any $t < s + 1$, $\bar{v}_t \leq \bar{v}_{t+1}$,
- for any $t \leq t^*$, $\psi(\bar{v}_t, \bar{v}_{t+1}) \leq \psi(v_t, v_{t+1})$. This is due to asymmetry of ψ,
- $0 \leq \psi(\bar{v}_{t^*}, \bar{v}_{t^*+1}) < \psi(v_{t^*}, v_{t^*+1})$. This is possible due to part (3) of Definition 3,
- for any $t^* < t < s$, $\psi(\bar{v}_t, \bar{v}_{t+1}) \leq \psi(v_t, \eta_{t+1})$. This is possible due to part (2) of Definition 3,
- $\psi(\bar{v}_s, \bar{v}_{s+1}) = 0$. This follows from the following argument: Because $\bar{v}_s$ is almost **1**-associated with v_s, there is v'' that is almost **1**-dominated by v_s and v'' is **1**-associated with $\bar{v}_s$. Because $v_s = v' \in V^1$ and V^1 is defined through Lemma 16, $\psi(v_s, v'') > 0$. By part (4) of Definition 3, there is $\bar{v}_{s+1} \geq \bar{v}_s$ that is **1**-associated with v_s and $\psi(\bar{v}_s, \bar{v}_{s+1}) = 0$.

Then, $\bar{v} \leq \bar{v}'$, states v' and $\bar{v}'$ are **1**-associated and $\tilde{\psi}^1(\bar{v}, \bar{v}') < \tilde{\psi}^1(v, v')$. The latter implies that (B.6) holds. □

Suppose that cost function ψ is supermodular and strictly or robustly asymmetric and j^1 is defined by (B.4). Together with Lemma 17, the above implies that all steps $k \geq 1$ are strictly asymmetric.

B.4. **Proof of Lemma 3.** Suppose that cost function ψ is asymmetric and supermodular, cost functions ψ^k are defined by (B.1) and sets and projections are defined by (B.4). Let k^* be the lowest $k \geq 0$, such that $V^k = V^{k+1}$. Since S is finite, k^* is well-defined. Also, $\left|V^{k^*}\right| = 1$. Indeed, there are two states $v, v' \in V^{k^*}$, $v \neq v'$ such that $\psi^{k^*}(v, v') = 0$. Either one of states v, v' is not a k^*-attractor, or they are both k^*-attractors which can be connected by 0-cost paths. In both cases, (B.4) eliminates some of these states and $V^{k^*+1} \subsetneq V^{k^*}$, which is a contradiction.

The supermodularity and asymmetry along the steps of the algorithm can be used to characterize attractors of state $\mathbf{1}$.

Lemma 19. *For each $k \geq 0$, $\mathbf{1} \in U^k(\mathbf{1})$. If cost function ψ is strictly or robustly asymmetric, then $\{\mathbf{1}\} = U^k(\mathbf{1})$ for each $0 \leq k \leq k^*$.*

Proof. By the remarks after Lemmas 16 and 17, all steps are supermodular and asymmetric.

We show first that $\mathbf{1} \in U^0(\mathbf{1})$. Note that states $\mathbf{1}, \mathbf{0}$ are $\mathbf{1}$-associated. Suppose that $\psi^0(\mathbf{1}, v') = 0$ for some $v' \in S$. By asymmetry, $\psi^0(\mathbf{0}, \bar{v}') = 0$ for some $\bar{v}'$ such that $v', \bar{v}'$ are $\mathbf{1}$-associated. By supermodularity, $\psi^0(v', \bar{v}'') = 0$ for some $\bar{v}'' \geq v', \bar{v}'$. Because $v', \bar{v}'$ are $\mathbf{1}$-associated, it must be that $\bar{v}'' = \mathbf{1}$. Hence, $\mathbf{1} = \bigvee_{v' \in U^0(\mathbf{1})} v' \in U^0(\mathbf{1})$.

We show by induction on $k \geq 0$ that $\mathbf{1} \in U^k(\mathbf{1})$. Assume that this is true for $k-1$. Suppose that $\psi^k(\mathbf{1}, v') = 0$ for some $v' \in V^k$. By the asymmetry of kth step, $\psi^k\left(J^k(\mathbf{0}), \bar{v}'\right) = 0$ for some $v' \in V^k$, such that $v', \bar{v}'$ are $\mathbf{1}$-associated. By supermodularity, $\psi^k(v', \bar{v}'') = 0$ for some $\bar{v}'' \in V^k$, such that $\bar{v}'' \geq v', \bar{v}'$. Because $v', \bar{v}'$ are $\mathbf{1}$-associated, it must be that $\bar{v}'' = \mathbf{1}$. Hence, $\mathbf{1} = \bigvee_{v' \in U^k(\mathbf{1})} v' \in U^k(\mathbf{1})$.

Suppose now that the cost function is, in addition, strictly or robustly asymmetric. Then, $\psi(\mathbf{1}, v') > 0$ for any $v' \neq \mathbf{1}$. Thus, $\{\mathbf{1}\} = U^0(\mathbf{1})$.

We show by induction on k that $\{\mathbf{1}\} = U^k(\mathbf{1})$. Assume that $\{\mathbf{1}\} = U^{k-1}(\mathbf{1})$ for $k \geq 1$. By the remarks after Lemma 18, step k is strictly asymmetric. Suppose that $\psi^k(\mathbf{1}, v') = 0$ for some $v' \in V^k$. By strict asymmetry of step k, $\psi^k\left(J^k(\bar{v}), \bar{v}'\right) < 0$ for some $\bar{v}' \in V^k$. A contradiction shows that $\psi^k(\mathbf{1}, v') > 0$ for any $v' \in V^k$ and $\{\mathbf{1}\} = U\left(U^{k-1}(\mathbf{1})\right) = U^k(\mathbf{1})$. □

Trivially, $MR\left(V^{k^*}, \psi^{k^*}\right) = V^{k^*}$. By Lemma 19, $MR\left(V^{k^*}, \psi^{k^*}\right) = \{\mathbf{1}\}$. Lemma 3 follows from Lemmas 15 and Lemma 19.

E-mail address: mpeski@uchicago.edu