How Complex Are Networks Playing Repeated Games?*

In-Koo Cho[1] and Hao Li[2]

[1] Department of Economics, Brown University, Providence, Rhode Island 02912

[2] School of Economics and Finance, University of Hong Kong, Hong Kong

Received: revised version:

Summary. This paper examines implications of complexity cost in implementing repeated game strategies through networks with finitely many classifiers. A network consists of individual classifiers that summarize history of repeated play according to a weighted sum of empirical frequency of the outcomes of the stage game, and a decision unit that chooses an action in each period based on the summaries of the classifiers. Each player maximizes his long run average payoff, while minimizing the complexity cost of implementing his strategy through a network, measured by its number of classifiers. We examine locally stable equilibria where the selected networks are robust against small perturbations. In any locally stable equilibrium, no player uses a network with more than a single classifier. Moreover, the set of locally stable equilibrium payoff vectors lies on two line segments in the payoff space of the stage game.

JEL Classification Number: C72.

* We are grateful for insightful comments by the anonymous referee, which helped us improve the paper substantially. Financial support for the first author from the National Science Foundation and the Alfred P. Sloan Foundation, and for the second author from the Alfred P. Sloan Foundation Dissertation Fellowship is gratefully acknowledged.

1. Introduction

The potential complexity of equilibrium strategies in dynamic economic models raises suspicion about the assumption of rationality of economic agents. In response, models of bounded rationality are proposed to recover the conclusions of equilibrium models through "simple" strategic plans. For example, in fictitious play (e.g., Fudenberg and Kreps [1993]), each player computes the empirical distribution of actions by his opponent, and chooses a myopically optimal action based on this summary statistic. In 2×2 games, fictitious play dictates an action according to a single threshold rule and emulates a Nash equilibrium outcome in the long run regardless of the initial conditions. One can find a similar spirit in Bray [1983] in the context of rational expectations equilibrium. Each agent forms a forecast of market price according to the average price in the past and chooses his optimal action based on the forecast. Under certain conditions, the agent's forecast converges to the actual price distribution in the long run, even though the agent does not have perfect foresight.

A common thread between fictitious play and adaptive expectations is that a player's action in each period depends upon a simple summary statistic of past history. Thus, to implement this kind of strategy, we need not assume excessively sophisticated information processing capabilities as we often do in equilibrium analysis. The main exercise in these works is to demonstrate that the "simple" rule leads to an equilibrium in the long run, and to verify the stability of the limit outcomes. The convergence result is important in showing the robustness of equilibrium analysis against bounded rationality of agents, and also useful in selecting interesting equilibria.

This paper views bounded rationality as a consequence of the "complexity cost" in implementing a repeated game strategy rather than an exogenous constraint in processing information. Complex strategies require economic agents to respond to new information in a more sophisticated way, and are therefore more costly to implement. The kind of complexity cost we have in mind is best understood as computational cost, since a more complex decision procedure requires more computation in processing information. Imagine agents engaged in long term competition who are concerned with the cost of implementing complex strategies as well as their long term payoffs. Under this interpretation, the complexity of an equilibrium strategy is determined endogenously in balancing the trade-off between higher payoffs and lower complexity cost, for achieving higher payoff generally necessitates a more complex strategy. We choose infinitely repeated games as the environment to study the trade-off, because the results of infinitely repeated games are well-established for comparison. The objective of this paper is to understand the long run implications of the complexity cost on equilibrium strategies and equilibrium outcomes in infinitely repeated games.

The cost approach to bounded rationality is pioneered by Rubinstein [1986]. Rubinstein models a repeated game strategy by a finite automaton, and measures the complexity of the automaton by counting the number of its states. His result imposes restrictions on the set of equilibrium outcomes in the repeated prisoner's dilemma game. Since an automaton with infinitely many states is needed to compute an average from a history in an infinitely repeated game, finite automata do not capture the simplicity of fictitious play and adaptive expectation. In this paper we use networks with finitely many classifiers to represent repeated game strategies.[1] Roughly speaking, we take fictitious play as the basic unit of information processing, called classifier, and connect multiple classifiers into a network. Each classifier discriminates histories according to a weighted average of the empirical distribution of the stage game outcomes, and the network chooses an action in each period based on the summaries of its classifiers. Since networks with more classifiers can discriminate histories of the repeated game play in a finer way, we measure the complexity cost of a network by the number of its classifiers.

One might think that we admit more general computing machines than finite automata because a strategy represented by a network can be implemented only by an automaton with infinitely many states. This is misleading. A finite automaton admits an arbitrary transition rule between the states, but a network imposes a rigid, yet intuitive, transition rule in that the empirical frequency of outcomes must be updated each period according to simple averaging.[2] Due to this difference in transition rules, networks cannot be said to

[1] Such networks are a version of neural networks used in artificial intelligence as model of human brain. For reference, see Weisbuch [1990].

[2] The general automaton is often criticized on the ground that the transition rule of the state often

be more general than finite automata. As a result, the analysis of Rubinstein [1986] and Abreu and Rubinstein [1988] does not apply to network strategies. We believe that the simplicity of the transition rule in a network strategy makes networks with finitely many classifiers an attractive alternative to finite automata as a model of bounded rationality.

We posit the following question: how complex are equilibrium networks in the long run when complexity cost is small but positive? To emphasize the impact of complexity cost, we model the trade-off between repeated game payoff and complexity cost by a lexicographical preference. The primary objective of each player is to maximize his repeated game payoff. Only when two networks achieve the same long run payoff, the player opts for the simpler one.

We imagine a decision maker who has to delegate his strategic plan to a network of information processors, but is also concerned about the possibility of small perturbations that can push the play off the equilibrium path. Due to the limited computational capability, the decision maker cannot describe the perturbations precisely. In this situation, it would be natural to require the long run outcome of the network to be robust against small mistakes. That is, even after someone makes a mistake, the network should be "stable" enough to recover what it is supposed to achieve in the long run. We call this requirement local stability. Local stability requires first that the equilibrium networks be time-consistent, so that along the equilibrium path, no player has incentives to change his network to achieve a higher payoff or the same payoff with a network with fewer classifiers. Furthermore, local stability requires that no such incentives exist in a small neighborhood of the equilibrium path. Without the requirement, it would be difficult for a player with bounded computational capability to evaluate the long run performance of a network exposed to small "unexpected" shocks.

Since we are primarily interested in networks that survive long run competition, we will focus on equilibrium networks that induce well-defined long run average frequency of outcomes. Except for an outcome path where players choose a pure strategy Nash

lacks intuition and structure. As demonstrated by Banks and Sundaram [1990], virtually all equilibria constructed by Rubinstein [1986] and by Abreu and Rubinstein [1988] collapse once players give minimal consideration to the complexity cost of the transition rule.

equilibrium of the component game in each period, each player has to alternate between different actions according to the equilibrium path. To identify the time to switch to another action, the player has to process the information contained in the history of plays. Given that such computation is costly, each player is eager to adopt a simpler network as long as he can achieve the same long run payoff. In a locally stable equilibrium, no player can have such incentives in a small neighborhood of the equilibrium path.

We demonstrate that complexity of equilibrium network unravels under the unilateral incentives to reduce complexity cost. In any locally stable equilibrium of an infinitely repeated 2×2 game, no player uses a network with more than a single classifier. Thus, there is a tight upper bound on the complexity of equilibrium strategy, even though players give only secondary consideration to the complexity cost. This result stands in sharp contrast to Rubinstein [1986] and Abreu and Rubinstein [1988] where the number of states of equilibrium automata is not bounded.

The set of locally stable equilibrium payoff vectors lies on two line segments in the payoff space of the stage game. In the prisoner's dilemma game, for example, an individually rational payoff vector can be sustained by a locally stable equilibrium if and only if it is located on the two diagonals in the feasible payoff space of the stage game. This paper achieves a significant extension of Rubinstein [1986] to a class of computing machines with infinitely many states and a simple transition rule. Despite the fact that our solution concept is a refinement of semi-perfect equilibrium (which requires only time-consistency), we end up with a larger set of equilibrium payoffs than Rubinstein [1986], while maintaining a uniform upper bound on the complexity of equilibrium strategies.

The paper is organized as follows. Section 2 formally describes the 2×2 stage game, the network, the preference, and the solution concept. Section 3 states the main result of this paper. The result is established in two major steps. First, we show that in any locally stable equilibrium no player uses a network with more than a single classifier. Second, we present necessary and sufficient conditions for a payoff vector to be sustained as a locally stable equilibrium payoff vector by a pair of networks with at most a single classifier. Section 3 presents the second step only, since the main intuitions behind our result can be found in the second step. The proof of the first step is relegated to the appendix. Section 4 concludes.

2. Formal Description

Consider a 2×2 game. Let $S_1 = S_2 = \{C, D\}$ be the set of actions of each player, and $S = S_1 \times S_2$ be the set of outcomes. Let $u_i(s)$ be player i's payoff from an outcome $s \in S$. Define for each $i \neq j \in \{1, 2\}$,

$$\underline{v}_i = \min_{s_j} \max_{s_i} u_i(s_i, s_j)$$

as the (pure) security payoff of player i. Between the two actions of player j, we identify D as the minmax strategy of the stage game:

$$\underline{v}_i = \max_{s_i} u_i(s_i, D).$$

A payoff vector $v = (v_1, v_2)$ is individually rational if $v_i \geq \underline{v}_i$ for each i, and strictly individually rational if $v_i > \underline{v}_i$ for each i.

We assume that the set of strictly individually rational payoff vectors is non-empty. This assumption excludes some games like matching-penny game

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{bmatrix} 1,-1 & -1,1 \\ -1,1 & 1,-1 \end{bmatrix}, \end{array}$$

but still admits a wide range of 2×2 games, including the prisoner's dilemma

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{bmatrix} 3,3 & 0,4 \\ 4,0 & 1,1 \end{bmatrix}, \end{array} \tag{2.1}$$

and the battle of sexes

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{bmatrix} 0,0 & 1,4 \\ 4,1 & 0,0 \end{bmatrix}. \end{array}$$

Let $G = \langle S, u_1, u_2 \rangle$ be a 2×2 game that has a non-empty set of strictly individually rational payoff vectors. Consider the infinitely repeated game of G. A history at the beginning of period $T+1$ $(T \geq 0)$ is $h^T = (s^1, \ldots, s^T)$ where $s^t \in S$ for each $t \leq T$. A repeated game strategy of player i is a mapping from a history to an action $s_i \in S_i$.

We now formally describe a neural network, or simply a network, which implements a repeated game strategy for player i. The network consists of one decision unit and K_i classifiers. At the beginning of the game, the decision unit selects an action s_i^1. Let $h^T = (s^1, \ldots, s^T)$ be a history at the beginning of period $T+1$ $(T \geq 1)$. Define

$$f(s : h^T) = \frac{1}{T} \#\{t \leq T : s^t = s\}$$

as the empirical frequency of s in h^T. The k-th classifier $(k = 1, \ldots, K_i)$, represented by a vector $\alpha_i^k \in \mathbb{R}^4$, summaries the history into

$$\sum_{s \in S} \alpha_i^k(s) f(s : h^T)$$

and reports to the decision unit $b^k = 1$ if the summary statistic is positive and $b^k = 0$ if otherwise. Based on the binary vector $b_i = (b^1, \ldots, b^{K_i})$ of the reports from K_i classifiers, the decision unit chooses an action $s_i^{T+1} \in S_i$ according to a decision function

$$B_i : \{0,1\}^{K_i} \rightarrow \{C, D\}.$$

We can represent player i's network by

$$\langle \alpha_i^1, \ldots, \alpha_i^{K_i}; B_i; s_i^1 \rangle.$$

Let $\Phi_i^{K_i}$ be the collection of networks with K_i classifiers. Let $\Phi_i = \cup_{K_i \geq 0} \Phi_i^{K_i}$ be the collection of all networks with finitely many classifiers.

What we have just described is known as a perceptron with K_i linear classifiers, which has been used to model how our brain processes information to identify objects (Weisbuch [1990]). This network does not adjust the parameters $\alpha_i^1, \ldots, \alpha_i^{K_i}$ in response to outcomes of information processing. In this sense, it does not "learn" to play. Rather, it is designed to implement a strategy by taking the empirical frequency of outcomes in each period and processing this information to take an action.[3] [4]

The network has a natural interpretation as a hierarchical information-processing organization consisting of a tier of accountants and a decision maker.[5] The input to the organization is a summary of the history of the repeated game play. After processing the input information, the organization outputs a decision C or D. Each classifier can be thought of as an accountant, whose job is to collect the input information (determine the empirical frequency of the outcomes from the history) and to process it by summarizing it into a binary variable according to a weighted average rule with pre-specified weights. The decision unit can be viewed as a decision maker, whose job is to process the binary vector of reports from accounts and transform it into a decision C or D. With this interpretation of networks as hierarchical organizations, the paper can be viewed as an exercise to determine the size and scope of organizations that survive in the long run competition.

Let $f(h^T) \in \Delta^4$ be the vector of empirical frequency of outcomes in history h^T, where Δ^4 is the unit simplex in $\mathbb{R}^4$. For each $k = 1, \ldots, K_i$, let

$$\mathbf{H}_i^k = \left\{ f \in \Delta^4 : \sum_{s \in S} \alpha_i^k(s) f(s) = 0 \right\}$$

be the hyperplane in Δ^4 determined by α_i^k. Then the k-th classifier can be thought of as computing the empirical frequency $f(h^T)$ following h^T and reporting whether $f(h^T)$ is "above" or "below" $\mathbf{H}_i^k$. Let $\overline{\mathbf{H}}_i$ and $\underline{\mathbf{H}}_i$ be the upper and lower open half spaces determined by $\mathbf{H}_i$. We often identify the k-th classifier with the hyperplane $\mathbf{H}_i^k$. The hyperplanes $\mathbf{H}_i^1, \ldots, \mathbf{H}_i^{K_i}$ divide Δ^4 into at most 2^{K_i} "cells". The decision unit assigns an action $s_i^{T+1} \in S_i$ to each cell according to the decision function B_i.

[3] A repeated game played by a neural network with learning capability is extensively studied in Cho [1996b].

[4] Hornik, Stinchcombe and White [1989] show that one can approximate any measurable function by a sequence of single layered neural networks. Thus, in a static game, a decision maker who can use any network with finitely many classifiers can approximate virtually all strategies. This result clearly holds for fintely repeated games. Extending their result to infinitely repeated games would be a worthwhile exercise for future research.

[5] There has been recent growth in the literature of organizational design where organizations are modeled as information processing networks. See, e.g., Radner [1993]. Van Zandt [1996] surveys the most recent works in this literature.

Given a pair $\varphi = (\varphi_1, \varphi_2) \in \Phi_1 \times \Phi_2$, let $\{\sigma^t(\varphi)\}_{t=1}^T$ be the sequence of outcomes induced by φ up to period T. For each i, define

$$v_i(\varphi) = \liminf_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} u_i(\sigma^t(\varphi))$$

as player i's long run average payoff from the repeated game. Let $c(\varphi_i)$ be the "complexity cost" of the network φ_i. For simplicity, we assume that $c(\varphi_i)$ is equal to the number of classifiers K_i in φ_i.

The number of classifiers naturally represents the sophistication of the network, or the complexity of the repeated game strategy implemented by the network. A network with no classifier can only implement the strategies that play C or D following every history, while a network with many classifiers is more complex because it can divide the frequency simplex Δ^4 into many cells and discriminate the history of repeated game play in a finer way.

The complexity cost can be regarded as the cost of maintaining the classifiers in the network. This interpretation parallels the number-of-states measure for finite automata used by Rubinstein [1986] and Abreu and Rubinstein [1988]. The idea is that in each period the player has to pay a fee for each classifier used in his network. Because each classifier is involved in choosing the action in each period, it is necessary to maintain all classifiers, and therefore, the maintenance cost of a network can be measured according to the number of classifiers. An alternative interpretation of the complexity cost in our model is the computational cost related to processing the reports from the classifiers. This interpretation is not valid in the finite automaton model: since in each period only one state is involved in the computation before an action is chosen, a finite automaton with more states does not have a greater computational cost than one with fewer states. In contrast, due to the parallel computation nature of our networks, in each period all classifiers are involved in computation before an action is chosen. Therefore, in our network model, the maintenance cost interpretation and the computational cost interpretation of the complexity cost are equally valid. If we view the network as a hierarchical organization with K_i as the number of accountants, then the complexity cost of the network represents the sum of wages paid to the accountants.[6]

The difference between the number-of-states measure of complexity cost in the finite automata literature and our number-of-classifier measure needs to be stressed. Both the finite automaton and our network are attempts to model procedural aspects of decision-making in competitive situations. A finite automaton takes a "sequential" approach by emphasizing the dynamics of transition of the states. By contrast, a network is "parallel" in the sense that it emphasizes the classification of "the state space" (the space of the outcome frequency) and recognition of the state (the frequency of realized outcomes in the history) in the state space in each period. We can imagine that the outcome frequency space is divided into cells by the classifiers and associated with an action by the decision unit. The difference between the number-of-states measure and our number-of-classifier measure reflects the difference in the two approaches. A priori, one cannot say which approach is more appropriate in modeling the procedural aspects of decision-making.

To illustrate the different measures of complexity cost used in finite automaton model and our network model, consider the following example. Suppose that two decision makers play the infinitely repeated prisoner's dilemma game, whose component game is (2.1). Consider the following outcome path

$$(C,C),(C,C),(D,D),(D,D),(C,C),(C,C),(D,D),(D,D),\ldots \tag{2.2}$$

where the two players synchronize their actions in a two-period cycle. In order to sustain (2.2), it is necessary for the machine to know in which phase of the cycle the game is. This can be achieved by a finite automaton by assigning different states for each C, for example, so that the machine knows whether the game is at the beginning or at the end phase of C cycle. Indeed, following Abreu and Rubinstein [1988], one can show that the outcome path can be induced by a pair of 4-state automata. In contrast, a pair of networks with a single linear classifier cannot distinguish the same outcome within one phase, and therefore cannot induce the outcome path (2.2). However, the long run frequency of outcomes in

6 This measure ignores the computational complexity of summarizing information provided by multiple accountants. For a discussion of the implications of this complexity to organizational design, see Li [1995].

(2.2) is the same as the following outcome path:

$$(C, C), (D, D), (C, C), (D, D), \ldots$$

which can be induced by a pair of networks with a single linear classifier. For example, for each $i = 1, 2$ and $\forall s_i \in \{C, D\}$, we can choose

$$\alpha_i(C, s_i) = 1, \quad \alpha_i(D, s_i) = -1,$$

with the initial action $s_i^1 = C$ and the decision rule

$$B_i(z) = \begin{cases} D & \text{if } z = \sum_{s \in S} \alpha_i(s) f(s : h^T) > 0 \\ C & \text{otherwise.} \end{cases}$$

We assume that each player i has the lexicographic ordering $\prec_i$ between $v_i(\varphi)$ and $c(\varphi_i)$: given any two pairs of networks $\varphi = (\varphi_1, \varphi_2)$ and $\varphi' = (\varphi_1', \varphi_2')$, $\varphi_i \prec_i \varphi_i'$ if $v_i(\varphi) < v_i(\varphi')$ or $v_i(\varphi) = v_i(\varphi')$ and $c(\varphi_i) > c(\varphi_i')$. Under the lexicographical ordering, each player's first objective is to maximize his average payoff from the repeated game using a network with finitely many classifiers. Throughout the paper, we assume that the complexity cost of the network is given a secondary consideration. This assumption is made to emphasize the impact of any slight consideration of complexity cost on equilibrium strategies in repeated games. But our conclusions hold even if players' preferences exhibit non-trivial trade-offs between the long run average payoff and the complexity cost, as long as the consideration of complexity cost is sufficiently insignificant compared to the long run average payoff.

A network game is a normal form game

$$G^n = \langle \Phi_1^1, \Phi_2^1; \prec_1, \prec_2 \rangle,$$

where each player delegates his repeated game strategy to a network at the beginning of the game. To be a solution of this game, the pair of networks must form a Nash equilibrium.

Definition 2.1. A pair of networks $(\varphi_1^*, \varphi_2^*) \in \Phi_1^1 \times \Phi_2^2$ is a Nash equilibrium of G^n if $\forall i \neq j = 1, 2$, $\forall \varphi_i \in \Phi_i$, $\varphi_i^* \not\prec_i \varphi_i$.

We imagine a decision maker who delegates his repeated game strategy to a network of classifiers. If the network never makes mistakes in carrying out the instructions, Nash equilibrium would be an appropriate solution concept. If small mistakes are possible but the decision maker knows the precise probability distribution of the mistakes, he can calculate the optimal network at the beginning of the game. In such a case, Nash equilibrium is again appropriate.

However, when small mistakes by the network are possible and when the decision maker has bounded computational capability, it would be extremely difficult to calculate the probability distribution of all future mistakes of a network, which may depend on states of the networks in each period. Instead of a decision maker with unlimited computational capability, we assume that he is aware of the existence of small mistakes, but cannot calculate the entire probability distribution of the mistakes. Also, the decision maker is aware that once he delegates the repeated game strategy to a network, he will have no further opportunity to modify it once the repeated game starts. In such an environment, it is natural to consider a locally stable equilibrium: if a decision maker cannot identify the small mistakes, then he should choose a network that remains optimal even after "unexpected" shocks push the future play into a small neighborhood of the equilibrium path.

To illuminate the importance of the local stability, consider a pair of "grim trigger" strategy:

$$s_i^1 = C, \quad \alpha_i(C,C) = 0 \quad \text{and} \quad \alpha_i(s) = 1 \quad \forall s \neq (C,C). \tag{2.3}$$

Each player chooses D in period T following history h^T if and only if the summary statistic $\sum_s \alpha_i(s) f(s : h^T) > 0$. Note that as long as both players choose C, the summary statistic of each network remains 0, and each player continues to play C. However, as soon as a player deviates from C, the statistic becomes positive and remains positive for the rest of the game so that each player chooses D forever. Thus, it takes only a single mistake to push the long run outcome to (D, D). Given that the decision maker has bounded computational capability and therefore cannot identify future shocks when he delegates the game to a network, such extreme sensitivity to perturbations would make it very difficult for him to evaluate the long run outcome.

Since we are primarily interested in networks that survive long run competition, we will focus on equilibrium networks that induce well-defined long run average frequency. Under this assumption, the notion of "local stability" can be easily formalized.[7] For each $s \in S$, we require the existence of

$$f_\varphi(s) = \lim_{T \to \infty} f(s : \{\sigma(\varphi)\}_{t=1}^T).$$

For any $\epsilon > 0$ and $f \in \Delta^4$, let

$$N_\epsilon(f) = \left\{ f' \in \Delta^4 : \sum_{s \in S} |f(s) - f'(s)| < \epsilon \right\}$$

be the ϵ neighborhood of f.

Definition 2.2. A Nash equilibrium $\varphi^* = (\varphi_1^*, \varphi_2^*)$ of G^n is a locally stable equilibrium of G^n if $\exists \bar{\epsilon} > 0$ such that $\forall \epsilon \in (0, \bar{\epsilon})$, $\forall T \geq 1$ and $\forall h^T$ such that $f(h^T) \in N_\epsilon(f_{\varphi^*})$, $\forall i = 1, 2$, and for all $\varphi_i \in \Phi_i^1$, player i cannot receive more than $v_i(\varphi^*)$ by switching to φ_i in period T, and receives strictly less than $v_i(\varphi^*)$ by switching to any network with classifiers fewer than $c(\varphi_i^*)$.

Thus, our solution concept is stronger than the semi-perfect equilibrium of Rubinstein [1986], which requires only optimality of equilibrium finite automata along the equilibrium path. He offers two justifications for the restriction of semi-perfection: natural decay of unused part of finite automaton, and dynamic consistency. Both justifications are applicable in our setup here. However, since our goal is to examine the complexity of networks that survive in long run competition, dynamic consistency is necessary but not sufficient. We believe that networks viable in the long run should have the local stability property that a few rounds of mis-coordination between opposing networks due to small mistakes should not provide incentives to for players to scrap their networks. The equilibrium of finite automata constructed by Rubinstein is not locally stable in our sense. One mistake

[7] This assumption has strong analytical implications as well. In particular, under the incentives to save on the complexity cost, the assumption implies that in equilibrium all hyperplanes that represent classifiers in a network in the frequency space must pass through the long run limit frequency. Unfortunately, we are unable to state this assumption in terms of restrictions on the networks that are allowed. The results of this paper do not exclude the existence of equilibrium where long-run averages do not converge.

by one of the two equilibrium automata is sufficient to induce the two automata to punish each other forever. Incidentally, the equilibrium constructed in our results is "globally" stable: the restriction in Definition 2.2 is satisfied for all time period T and arbitrarily large values of ϵ.[8] An additional reason for imposing local stability is that the restriction simplifies analysis. Without the restriction, the dynamics of transition of state variable can be too complex. Together with the assumption that long run average frequencies converge in equilibrium, the restriction of local stability allows us to focus on the long run implications of strategies implemented by a pair of networks. This will be explained in the next section.

3. Analysis

3.1. The Main Results

Our first result shows that only simple strategies can survive the strategic pressure.

THEOREM 3.1. *No locally stable equilibrium network has more than a single classifier.*

PROOF. See the appendix.

Theorem 3.1 imposes a *uniform* bound on the complexity of locally stable equilibrium networks. This result should be contrasted with Rubinstein [1986] and Abreu and Rubinstein [1988] where there is no such bound on the complexity of equilibrium automata.

The idea of this result is best described geometrically. Recall that each classifier can be viewed as a hyperplane in the unit simplex in $\mathbb{R}^4$. One can imagine the empirical frequency as a particle moving around in $\mathbb{R}^4$ and passing through the hyperplanes induced by different classifiers. It is easy to see that the empirical frequency must pass through *every* hyperplane infinitely often. If there is a hyperplane that the empirical frequency does not visit after a certain period, a decision maker can achieve the same long run

8 But we do not know whether there exists equilibrium with $\epsilon = 0$ in Definition 2.2. That is, we do not know whether there are semi-perfect equilibria of the network game as defined by Rubinstein [1986].

outcome even after eliminating the associated classifier, which contradicts the equilibrium requirement of minimal computational cost.

Note that the increment of the empirical frequency converges to 0 at the rate of $1/t$. If the limit of the empirical frequency is well defined, it must be the intersection of all hyperplanes. Because all hyperplanes pass through a single point, the state space can be divided into different cones, each of which has its vertex at the limit frequency. One can view local stability as the requirement that starting from any point in a small neighborhood of the limit frequency, no player can improve his long run payoff from the game or achieve the same payoff with a simpler network.

The rub of the proof is to show that against any network of the opponent, a player can always find a hyperplane induced by the opponent's network and a initial point in an arbitrarily small neighborhood of the limit point such that he can achieve the same long run outcome by using a network with 3 classifiers.[9] This argument shows that in any locally stable equilibrium, no player needs a network with more than 3 classifiers. Then, we reduce the number of classifiers in equilibrium networks one by one given the bound on complexity of equilibrium networks. In the end, we are left with single-classifier networks only. Networks with single classifier are examined in our next result.

THEOREM 3.2. *An individually rational payoff vector v^* can be supported by a locally stable equilibrium if and only if [1] v^* corresponds to a Nash equilibrium of the stage game, or [2] v^* is strictly individually rational and satisfies*

$$v^* \in \{\lambda u(C,C) + (1-\lambda)u(D,D) | \lambda \in (0,1)\}, \tag{3.1}$$

or [3] v^* is strictly individually rational and satisfies

$$v^* \in \{\lambda u(C,D) + (1-\lambda)u(D,C) | \lambda \in (0,1)\}. \tag{3.2}$$

Because the proof of Theorem 3.2 reveals the key idea of the proof of Theorem 3.1, we state a complete proof for Theorem 3.2 while delegating the proof for Theorem 3.1 to the appendix.

[9] The set of such initial points contains an open set in the neighborhood of the limit point.

The conditions (3.1) and (3.2) in Theorem (3.2) reduce the set of locally stable equilibrium payoff vectors to at most two line segments in the payoff space of the stage game. With networks with a single classifier, it might seem improbable that any individually rational payoff vector can be sustained by a locally stable equilibrium. But, Cho [1995] demonstrates that networks with single classifier can in fact sustain any individually rational payoff vector, *if* neither player considers the complexity cost of implementing the network.

It is instructive to review the construction of the single classifier equilibrium strategies in Cho [1995]. Fix an individually rational payoff vector $v^* = (v_1^*, v_2^*)$. For each $s \in S$ and $i \neq j \in \{1, 2\}$, let $\alpha_i \in \mathbb{R}^4$ be the weights of the single classifier of network used by player i, determined according to

$$\alpha_i(s) = u_j(s) - v_j^*, \tag{3.3}$$

and let B be a function that maps 1 to D and 0 to C. Player i's network is given by

$$\langle \alpha_i; B; C \rangle.$$

Note that player i punishes player j whenever player j has accumulated an average payoff higher than v_j^*. On the other hand, if player j's average payoff falls below his equilibrium payoff, player i plays C so that player j can improve his payoff. As a result, this pair of networks induces a Nash equilibrium following any history.

In (3.3), the locally stable equilibrium payoff of player i is "enforced" by player j in the sense that whenever player i's payoff exceeds his target, player j's punishment is triggered and whenever player i's payoff falls below his target, player j chooses C. In such a case, player i can simplify the complexity of the network by eliminating the single classifier and consequently, most equilibria with a single classifier collapse.

We know that with no classifier, a player can implement a repeated game strategy that dictates D following every history, or C following every history. With such strategies, one can sustain only the stage game pure strategy Nash equilibria as the long run outcome. The difficult part in proving Theorem 3.2 is to show that one can also sustain the individually rational payoff vectors in (3.1) and (3.2).

3.2. Necessity

The crucial first step in proving that the conditions stated in Theorem 3.2 are necessary for an individually rational payoff vector to be sustained by a locally stable equilibrium is to show that the weights determined by (3.3), although quite special, are in fact necessary for a pair of networks with single classifier to be a locally stable equilibrium.

LEMMA 3.3. *Fix a locally stable equilibrium payoff vector* $v = (v_1, v_2)$ *where each player is using a network with single classifier. For each* $i = 1, 2$, *let* $\alpha_i(s)$ *be the weight assigned to* $s \in S$ *by the single classifier of the network used by player* i, *and let the decision function map 1 to* D *and 0 to* C. *Then there exists* $k_i > 0$ *such that for any outcome* s *that is realized with positive frequency in equilibrium and for* $j \neq i = 1, 2$,

$$\alpha_i(s) = k_i(u_j(s) - v_j). \tag{3.4}$$

The proof of this lemma is given in the appendix of Cho [1996a] for the case of the prisoner's dilemma game. Extension to any 2×2 stage game with non-empty set of strictly individually rational payoff vector is straightforward; detailed proof is available from the authors.

We are ready to show the necessity of the conditions stated in Theorem 3.2. In what follows, we assume without loss of generality that the decision function of each player i's network maps 1 to D and 0 to C so that player i chooses D if and only if

$$\sum_{s \in S} \alpha_i(s) f(s : h^T) > 0$$

where $\alpha_i(s)$ is the weight assigned to $s \in S$ by player i's single classifier.

Fix a locally stable equilibrium φ^* such that $v^* = v(\varphi^*)$. Individual rationality of v^* is obvious. Suppose that v^* does not satisfy (3.1) or (3.2). Then either there is s_i^* such that $\sum_{s_j \in S_j} f_{\varphi^*}(s_i^*, s_j) = 1$ $(j \neq i \in \{1, 2\})$, or at least 3 outcomes are realized with positive limit frequency. We will show that in the first case both players must be using a network with no classifier and v^* corresponds to a Nash equilibrium payoff vector of the stage game, and that in the second case v^* cannot be sustained by locally stable equilibrium payoffs by a pair of networks with single classifier. We state the two cases as the following two lemmas.

Lemma 3.4. *If* $\exists s_i^*$ *such that* $\sum_{s_j \in S_j} f_{\varphi^*}(s_i^*, s_j) = 1$, *then in equilibrium player* i *uses a network with no classifier.*

Proof of Lemma 3.4. Without loss of generality, let us assume that $i = 1$ and $j = 2$. First suppose that $f_{\varphi^*}(s_1^*, C) = 1$. Then we must have $\alpha_2(s_1^*, C) \le 0$. Otherwise since player 2 plays C only when $\sum_{s \in S} \alpha_2(s) f(s : h^T) \le 0$, whenever (s_1^*, C) is played, some other outcome must be realized within finite periods, and the long run frequency of (s_1^*, C) cannot be 1. But if $\alpha_2(s_1^*, C) \le 0$, player 1 can enforce $v_1^* = u_1(s_1^*, C)$ by switching to the s_1^*-forever strategy as soon as (s_1^*, C) is played along the path.

Similarly, if $f_{\varphi^*}(s_1^*, D) = 1$, we have $\alpha_2(s_1^*, D) \ge 0$, and player 1 can enforce v_1^* by switching to the s_1^*-forever strategy along the path.

Finally, suppose that $f_{\varphi^*}(s_1^*, C) + f_{\varphi^*}(s_1^*, D) = 1$ with $f_{\varphi^*}(s_1^*, C), f_{\varphi^*}(s_1^*, D) > 0$. Since v_1^* is individually rational, $v_1^* \ge u_1(s_1^*, D)$. It follows that $u_1(s_1^*, C) \ge v_1^* \ge u_1(s_1^*, D)$. Lemma 3.3 then implies that there is $k_2 > 0$ such that

$$\alpha_2(s_1^*, C) = k_2(u_1(s_1^*, C) - v_1^*) \ge 0,$$
$$\alpha_2(s_1^*, D) = k_2(u_1(s_1^*, D) - v_1^*) \le 0.$$

Suppose that player 1 uses the C-forever strategy. Then only (C, C) and (C, D) will be realized, and

$$\sum_{s \in S} \alpha_2(s) f(s : h^T) = k_2 \left[\sum_{t=1}^{T} u_1(C, s_2^t)/T - v_1^* \right],$$

where $s_2^t \in \{C, D\}$ is player 2's action in period t. Since player 2 plays D if the above sum is positive and C otherwise, and since $\alpha_2(C, D) \le 0$ and $\alpha_2(C, C) \ge 0$,

$$\alpha_2(C, D) \le \sum_{t=1}^{T} \alpha_2(C, s_2^t) = k_2 \left[\sum_{t=1}^{T} u_1(C, s_2^t) - v_1^* T \right] \le \alpha_2(C, C).$$

Dividing the above by T, we obtain v_1^* as the long run average payoff for player 1.

Q.E.D.

It is clear that if player i uses the s_i^*-forever strategy, which can be implemented by a network with no classifier, the other player can achieve his equilibrium long run average payoff by a network with single classifier. Thus, both players must be using a network with no classifier. Furthermore, the equilibrium payoff vector v^* must correspond to a Nash equilibrium of the stage game.

Lemma 3.5. *There is no locally stable equilibrium where at least three outcomes are realized with positive limit frequency.*

Proof of Lemma 3.5. The key of the proof is to use Lemma 3.3 to identify a player who can achieve his equilibrium payoff through the D-forever strategy or C-forever strategy. Since the argument for each case considered below follows the same logic as the last case in the proof of Lemma 3.4, we only identify the player and the deviation strategy with which he can achieve his equilibrium payoff.

First, suppose that only (C,D), (D,D) and (C,C) are played with positive frequency. By individual rationality, $v_1^* \geq \underline{v}_1 = \max(u_1(C,D), u_1(D,D))$. In order for (C,C), (C,D), and (D,D) to support v_1^*, $u_1(C,C) \geq v_1^*$. By Lemma 3.3, $\alpha_2(C,D) \leq 0$ and $\alpha_2(C,C) \geq 0$. Player 1 can then achieve v_1^* by using the C-forever strategy.

Suppose that only (C,D), (D,D) and (C,C) are played with positive frequency. By individual rationality, $v_1^* \geq \underline{v}_1 = \max(u_1(C,D), u_1(D,D))$. Thus, $u_1(D,C) \geq v_1^*$. By Lemma 3.3, $\alpha_2(D,D) \leq 0$ and $\alpha_2(D,C) \geq 0$. Player 1 can achieve v_1^* by using the D-forever strategy.

Next, suppose that only (C,C), (D,C) are (D,D) are played with positive frequency. By individual rationality, $v_2^* \geq \underline{v}_2 \geq \max(u_2(D,D), u_2(D,C))$. Thus, $u_2(C,C) \geq v_2^*$. By Lemma 3.3, $\alpha_2(D,C) \leq 0$ and $\alpha_2(C,C) \geq 0$. In this case, player 2 can enforce v_2^* by playing C forever.

Finally, suppose that (C,C), (D,C) and (C,D) are played with positive frequency (including the case where all four outcomes are realized with positive frequency). If $u_i(C,C) \leq v_i^*$ for each $i = 1,2$, Lemma 3.3 implies that $\alpha_i(C,C) \leq 0$. Recall that each player plays C when $\sum_{s \in S} \alpha_2(s) f(s : h^T) \leq 0$. Then once (C,C) is realized, each player will play C for the rest of the game, contradicting the assumption that (C,C), (D,C) and (C,D) are played with positive frequency. Thus, there is i such that $v_1^* < u_i(C,C)$, say $i = 1$. By individual rationality, $v_1^* \geq \underline{v}_1 \geq u_1(C,D)$, and Lemma 3.3 implies that $\alpha_2(C,C) \geq 0$ and $\alpha_2(C,D) \geq 0$. Player 1 can then achieve v_1^* by using the C-forever strategy.

Since in each of the cases above, there is a player that can achieve his equilibrium payoff through a network with single classifier, there is no locally stable equilibrium of a pair of networks with single classifiers where at least 3 outcomes are realized with positive limit frequency. *Q.E.D.*

It remains to show the strict individual rationality of v^* that satisfies (3.1) or (3.2). Suppose that there exists $\lambda \in (0,1)$ such that v^* satisfies $v^* = \lambda u(C,C) + (1-\lambda)u(D,D)$ and $v_1^* = \underline{v}_1 = \max(u_1(C,D), u_1(D,D))$. Since $\lambda \in (0,1)$, if $u_1(C,D) \leq u_1(D,D)$, we must have $u_1(C,C) = u_1(D,D)$. By the definition of $\underline{v}_1$, $u_1(D,C) \geq u_1(C,C) = u_1(D,D)$. But then by using the D-forever strategy, player 1 can obtain at least v_1^*. On the other hand, if $u_1(C,D) > u_1(D,D)$, to support v_1^* by $u_1(D,D)$ and $u_1(C,C)$, we must have $u_1(C,C) > u_1(C,D) > u_1(D,D)$. But then player 1 can obtain at least v_1^* by using the C-forever strategy. In either case, v^* cannot be sustained by a pair of networks with single classifier. *Q.E.D.*

3.3. Sufficiency

The analysis in the last subsection seems to suggest that no pair of networks with single classifier constitutes a locally stable equilibrium, since at least one player has incentives to replace his network with another one with no classifier. However, only two outcomes need to be realized with positive frequency to support the payoff vectors that satisfy (3.1) and (3.2), and Lemma 3.3 imposes no restrictions on outcomes realized with zero frequency. By exploiting this freedom of choosing weights, we can construct a pair of networks with single classifier such that neither player has incentives to switch to a network with no classifier to save the complexity cost, to take advantage of the simplicity of his opponent's network. Moreover, the pair of networks satisfies the local stability condition.[10] The construction then proves the sufficiency of the conditions in Theorem 3.2.

For each $s \in S$, let f_s be the unit vector in Δ^4 that is concentrated on s, and for any two vectors $f, f' \in \Delta^4$, let

$$[f, f'] = \{\lambda f + (1-\lambda) f' | \lambda \in [0,1]\}$$

[10] In fact, it satisfies a stronger global stability condition, as in Smale [1980].

be the closed line segment that connects f and f'.

[1] If v^* can be sustained by a Nash equilibrium of the stage game, we can support v^* by a pair of networks with no classifier that always plays the Nash equilibrium action following every history. Clearly, this pair of networks is locally stable.

[2] Fix a strictly individually rational v^* that satisfies (3.1). Then $\exists \lambda \in (0,1)$ such that

$$f^* = \lambda f_{C,C} + (1-\lambda) f_{D,D},$$

and for each $i = 1, 2$,

$$v_i^* = \sum_{s \in S} u_i(s) f^*(s).$$

Since v^* is strictly individually rational, $\forall i = 1,2$, $u_i(C,C) > v_i^* > u_i(D,D)$. Furthermore, strict individual rationality implies that there is $\lambda_1 \in (0,1)$ close enough to 0 such that $\forall i = 1, 2$,

$$\sum_{s \in S} u_i(s) f^1(s) < v_i^*,$$

where $f^1 = \lambda_1 f_{D,D} + (1-\lambda_1) f_{D,C}$. Similarly, there is $\lambda_2 \in (0,1)$ such that $\forall i = 1, 2$,

$$\sum_{s \in S} u_i(s) f^1(s) < v_i^*,$$

where $f^2 = \lambda_2 f_{D,D} + (1-\lambda_2) f_{C,D}$.

Let $\mathbf{H}^*$ be the hyperplane spanned by f^*, f^1 and f^2:

$$\mathbf{H}^* = \left\{ \lambda^1 f^1 + \lambda^2 f^2 + \lambda^3 f^* \middle| \lambda^k \geq 0, \sum_{k=1}^{3} \lambda^k = 1 \right\}.$$

Choose the directional vector of $\mathbf{H}^*$ so that

$$f_{D,D} \in \overline{\mathbf{H}}^*.$$

For each player i, let the weights α_i of the single classifier of his network and the decision function B_i be chosen such that following any history h^T player i plays C if and only if $f(h^T) \in \overline{\mathbf{H}}^*$. The initial action s_i^1 of each player can be arbitrary.

We claim that the pair of candidate networks with single classifier constructed above forms a locally stable equilibrium. This will follow from the following three claims: (i) if both players use the candidate networks, the long run frequency of outcomes converges to f^* following any history; (ii) against the candidate network of his opponent and following any history, no player i can obtain a long run average payoff that exceeds v_i^* with any repeated game strategy; and (iii) following any history, each player i gets a strictly lower payoff than v_i^* if he uses a network with no classifier.

By the construction of $\mathbf{H}^*$, $f_{C,C} \in \underline{\mathbf{H}}^*$ and $f_{D,D} \in \overline{\mathbf{H}}^*$. Given any history, if both players use the candidate networks thereafter, only (C, C) and (D, D) will played. Since by construction (C, C) is played in $\overline{\mathbf{H}}^*$ and (D, D) is played in $\underline{\mathbf{H}}^*$, the long run frequency converges to the intersection between H^* and $[f_{C,C}, f_{D,D}]$, which is just f^*. This establishes the first claim.

To see that against the candidate network of player i, player j ($j \neq i = 1, 2$) can get at most v_j^* with any repeated game strategy, note that by construction for any $f \in \overline{\mathbf{H}}^*$, $\sum_{s \in S} u_j(s) f(s) \leq v_j^*$. Therefore, if there exists T such that $\sum_{s \in S} u_j(s) f(s : h^T) > v_j^*$, then $f(s : h^T) \in \underline{\mathbf{H}}^*$. Thus, player i plays D, and

$$\sum_{s \in S} u_j(s) f_i^T(s) \leq \sum_{s \in S} u_j(s) f_i^{T+1}(s).$$

Since the increment of the average payoff is of the order of $1/T$,

$$\sum_{s \in S} u_j(s) f_i^T(s) \leq v_j^* + \frac{\max_{s \in S} |u_j(s)|}{T}.$$

Therefore,

$$\limsup_{T \to \infty} \sum_{s \in S} u_j(s) f_i^T(s) \leq v_j^*.$$

Finally, we verify that if following any history player i uses a network with no classifier, then player i's long run average payoff is strictly less than v_i^*. Note that by the construction of $\mathbf{H}^*$, $f_{D,C}, f_{C,D} \in \underline{\mathbf{H}}^*$. If player 1, say, uses the C-forever strategy, then only (C, D) and (C, C) will be realized with positive frequency. Since player 2 plays C in $\overline{\mathbf{H}}^*$ but $f_{C,C} \in \underline{\mathbf{H}}$, and he plays D in $\underline{\mathbf{H}}^*$ and $f_{C,D} \in \underline{\mathbf{H}}^*$, only (C, D) will realized with positive frequency. By strictly individual rationality, $v_1^* > u_1(C, D)$. Thus, player 1 is strictly worse off by using the C-forever strategy. On the other hand, if player 1 uses D-forever strategy, only (D, D) and (D, C) will be realized with positive frequency. Since (D, D) is realized in $\underline{\mathbf{H}}^*$ but $f_{D,D} \in \overline{\mathbf{H}}$, and (D, C) is realized in $\overline{\mathbf{H}}^*$ but $f_{D,C} \in \underline{\mathbf{H}}$, the long run frequency of outcome must be the intersection between H^* and $[f_{D,D}, f_{D,C}]$, which is f^1. By the definition of f^1, player 1 gets a strictly lower long run payoff than v_1^* by using the D-forever strategy.

[3] Fix a strictly individually rational v^* that satisfies (3.2). Then, $\exists \lambda \in (0, 1)$ such that

$$f^* = \lambda f_{D,C} + (1 - \lambda) f_{C,D},$$

and for each $i = 1, 2$,

$$v_i^* = \sum_{s \in S} f^*(s) u_i(s).$$

Given a positive integer M, define

$$f^M = \frac{1}{M+1} \left[f_{C,C} + M f_{D,D} \right].$$

By strict individual rationality, for each $i = 1, 2$, $v_i^* > u_i(D, D)$. Therefore, we can choose M sufficiently large so that for each $i = 1, 2$,

$$\sum_{s \in S} u_i(s) f^M(s) < v_i^*.$$

Let $\mathbf{H}^o$ be the hyperplane spanned by f^M and $[f_{C,D}, f_{D,C}]$:

$$\mathbf{H}^o = \left\{ \lambda f^M + (1 - \lambda) f | \exists \lambda \in [0, 1], f \in [f_{C,D}, f_{D,C}] \right\}.$$

By definition, $\mathbf{H}^o$ separates $f_{C,C}$ from $f_{D,D}$. Choose the directional vector so that

$$f_{D,D} \in \overline{\mathbf{H}}^o.$$

Since v^* is strictly individually rational, we can find $\lambda_1 \in (0, 1)$ such that

$$\sum_{s \in S} u_2(s) f^1(s) < v_2^*,$$

where $f^1 = \lambda_1 f_{D,C} + (1 - \lambda_1) f_{C,C}$. Similarly, choose $\lambda_2 \in (0, 1)$ such that

$$\sum_{s \in S} u_1(s) f^2(s) < v_1^*.$$

where $f^2 = \lambda_2 f_{C,D} + (1-\lambda_2) f_{C,C}$. For each $i = 1, 2$, let $\mathbf{H}^i$ be the hyperplane spanned by f^*, f^M and f^i, and choose the directional vector so that

$$f_{D,D} \in \overline{\mathbf{H}}^i.$$

Note that by construction, for each $i = 1, 2$ and $\forall f \in \overline{\mathbf{H}}^i$,

$$\sum_{s \in S} u_i(s) f(s) < v_i^*.$$

For player 1, let the weights α_1 of the single classifier of his network and the decision function B_1 be chosen such that following any history h^T player 1 plays C if $f(h^T) \in \overline{\mathbf{H}}^1$ and D otherwise. For player 2, let the weights α_2 and the decision function B_2 be such that player 2 plays D if $f(h^T) \in \underline{\mathbf{H}}^2$ and C otherwise. Notice a slight asymmetry between the two candidate networks. Recall that we have defined $\overline{\mathbf{H}}$ and $\underline{\mathbf{H}}$ as the open half spaces above and below $\mathbf{H}$. The candidate networks are constructed to make sure that (D, C) is played if $f(h^T) \in \mathbf{H}^1 \cap \mathbf{H}^2$.

Following the argument for the case of the main diagonal, we can show that against player i's linear strategy constructed above, player j cannot receive more than v_j^* through any repeated game strategy, and that following any history, each player i is strictly worse off by using a network with no classifier. To prove that the candidate networks form a locally stable equilibrium, we only need to verify that the long run outcome is exactly f^* if both players use the candidate networks.

Suppose that (C, C) and (D, D) are played only finitely many times. Then, $\exists T^*$ such that $\forall T \geq T^*$, $s^T = (C, D)$ or $s^T = (D, C)$. The long run frequency must be located on $[f_{C,D}, f_{D,C}]$. Moreover, by the construction of the candidate networks, (C, D) is played following T if $f(h^T) \in \overline{\mathbf{H}}^1 \cap \underline{\mathbf{H}}^2$ but $f_{C,D}$ is located in $\underline{\mathbf{H}}^1 \cap \overline{\mathbf{H}}^2$, and (D, C) is played following T if $f(h^T) \in \underline{\mathbf{H}}^1 \cap \overline{\mathbf{H}}^2$ but $f_{C,D}$ is located in $\overline{\mathbf{H}}^1 \cap \underline{\mathbf{H}}^2$. Thus, the long run frequency must be located on $\mathbf{H}^1 \cap \mathbf{H}^2$, and hence on $\mathbf{H}^1 \cap \mathbf{H}^2 \cap [f_{C,D}, f_{D,C}]$, which is precisely f^*.

To establish the desired conclusion, it suffices to show that following any history, (C, C) and (D, D) are played only finitely many times. Note that (C, C) is played following h^t only if $f(h^t) \in \overline{\mathbf{H}}^o$. By construction, $f(h^t) \in \overline{\mathbf{H}}^o$ if and only if $f(D, D : h^t)/f(C, C : h^t) > M$, or equivalently,

$$F(D, D : h^t) > MF(C, C : h^t),$$

where $F(s : h^t) = tf(s : h^t)$ for each $s \in S$. Thus, if (C, C) is played infinitely many times, then $\exists T \geq 1$ such that $F(D, D : h^T) < MF(C, C : h^T)$ and $s^T = (D, D)$. Such T obviously exists if (D, D) is played infinitely many times. Define

$$k = MF(C, C : h^T) - F(D, D : h^T).$$

Then, for each $t = T, \ldots, T + k$, $F(D, D : h^t) < MF(C, C : h^t)$, and (C, C) is never played between T and $T + k$. At $T + k$,

$$F(D, D : h^{T+k}) = MF(C, C : h^{T+k}).$$

By the definition of $\mathbf{H}^o$, following any history h^t such that $f(h^t) \in \mathbf{H}^o$, either (C, D) or (D, C) is played. Since $[f_{C,D}, f_{D,C}] \subset \mathbf{H}^o$, $\forall t \geq T + k$, $f(h^t) \in \mathbf{H}^o$. Thus, following $T + k$, (C, C) or (D, D) is never played, which shows that (C, C) and (D, D) are played only finitely many times. *Q.E.D.*

If a strategic plan is implemented by a finite automaton, then the set of equilibrium outcomes of the automaton game may be sensitive to whether the stage game is a normal form or an extensive form (Piccione and Rubinstein [1993]). For example, if the stage game is a 2×2 normal form game, then the set of Nash equilibrium outcomes of the automaton game is roughly the two diagonals. However, if the stage game is an extensive form game that has the same normal form game, then the set of Nash equilibrium outcomes of the automaton game collapses to the set of Nash equilibrium outcomes of the stage game. The networks considered in this paper summarize a history of play by counting the number of each outcome realized in the history, and ignores how such an outcome has occurred or how the opponent should have played off the equilibrium path. As a result, our model is not sensitive to the formal structure of the stage game, and the same analysis applies when the stage game is an extensive form game.

4. Concluding Remarks

This paper examines the implications of the complexity cost in implementing repeated game strategies through networks with finitely many classifiers. We examine locally stable equilibria where no player has an incentives to change his network to achieve a higher than equilibrium payoff or the same equilibrium payoff with a network with fewer classifiers. We demonstrate that in the locally stable equilibrium no player uses a network with more than a single classifier. Moreover, the set of locally stable equilibrium payoff vectors lies on two line segments in the payoff space of the stage game.

It must be emphasized that the results hinge on the fact that the network in our model does not adjust itself to changing environment. However, if the network can adjust itself in a simple "back propagation" process, a complex network can survive and even outperform a simple network. Cho [1996b] studies a repeated game played by simple neural networks with a single classifier. It remains to investigate how more complex networks with learning capability evolve over time.

Appendix

This appendix presents the proof of Theorem 3.1. We reduce the complexity of equilibrium networks in three steps. First we show that in any equilibrium, each player needs to use a network with at most three classifiers. This is the crucial step. By exploiting the assumption that the equilibrium long run frequencies of outcomes are well-defined, we construct a network with three classifiers that guarantees a player his equilibrium payoff regardless of the complexity of the network used by his opponent. In the second step we construct a network with two classifiers for a player to obtain his equilibrium payoff against any network of his opponent with three classifiers. In the third and last step we reduce the number of classifiers to one. These three steps are in section A.2, A.3 and A.4 respectively. Our proof involves mainly geometric arguments. In section A.1, we transform the "state space" of networks in order to use geometric arguments more effectively.

A.1. Transformation of States

Since a network chooses an action in each period based on the empirical frequencies of the outcomes, the natural state space of a network is Δ^4. However, we find it more convenient to use a state space that is explicitly three-dimensional. Fix an equilibrium φ^* with long run payoffs v_i^* to player i. By assumption, the equilibrium long run frequency f_φ^* is well-defined. Given any history h^T, define

$$v_f(h^T) = (v_1(h^T) - v_1^*, v_2(h^T) - v_2^*, f(D, D : h^T) - f_\varphi^*(D, D))$$

as the "average state variable," where

$$v_i(h^T) = \frac{1}{T}\sum_{t=1}^{T} u_i(s^t).$$

By the "average state space" V_f, we mean the collection of all average state variables. Clearly, V_f is a convex compact set spanned by 4 extreme points $\{v_f(s) : s \in S\}$ where

$$v_f(s) = \left(u_1(s) - v_1^*, u_2(s) - v_2^*, \mathbf{1}_{\{s=(D,D)\}} - f_\varphi^*(D, D)\right).$$

Whenever the meaning is clear from the context, we write average state variable as $(v_1^T, v_2^T, f^T(D, D))$ instead of $(v_1(h^T), v_2(h^T), f(D, D : h^T))$.

In section 2.2 we have interpreted a classifier as a hyperplane in Δ^4. It is easy to see that the hyperplane becomes a plane in the average state space. To ease the notational burden, we continue to refer to the classifier as the hyperplane in the average state space. In the average state space, the equilibrium is represented by the origin. We also refer to it as the "target." It follows directly from the definition of locally stable equilibrium that the hyperplane associated with each classifier in an equilibrium network must pass through the target in the average state space.

Often, it is more convenient to work with the "gross" instead of average state variable. The gross state variable $w_f(h^T)$ for given h^T is simply the average state $v_f(h^T)$ multiplied by T. We often suppress h^T to simplify notation. The increment of the gross state variable is $v_f(s)$ after each realization of $s \in S$. Let W_F^T be the gross state space following h^T. One can easily verify that W_F^T is the convex hull of $\{Tv_f(s) : s \in S\}$ and the origin. Since each extreme point $\{Tv_f(s) : s \in S\}$ of the gross state space expands "outward", the gross state space grows in size as $T \to \infty$.

Given a hyperplane **H** in the average state space, we can obtain the corresponding hyperplane in the gross state by "expanding" it with T. Because our analysis relies mainly on the orientation of the hyperplanes, we will use the same notation **H** to represent the corresponding hyperplane in the gross state space. To emphasize the geometric nature of the network, we call a network with K classifiers a K-hyperplane strategy, and use hyperplanes and classifiers interchangeably. We will also need the following notations. Let $[x, y]$ denote the closed line segment connecting $x, y \in \mathbb{R}^3$, and $\Delta(s, s', s'')$ denote the triangle spanned by $v_f(s)$, $v_f(s')$ and $v_f(s'')$. We call $\Delta(s, s', s'')$ a "surface" of V_f.

A.2. No More than 3 Hyperplanes

PROPOSITION A.1. *Fix any equilibrium φ^* with $f^*_\varphi(s) > 0$ for each $s \in \{C, D\}^2$. No equilibrium network entails more than 3 hyperplanes. Moreover, if $\exists s \in S$ such that $f^*_\varphi(s) = 0$, then no equilibrium network has more than 2 hyperplanes.*

PROOF. Fix the equilibrium network of player 2. We assume that $f^*_\varphi(s) > 0$ for each $s \in \{C, D\}^2$. In this case, the target is in the interior of the average state space. The other case can be handled similarly. Let $\mathcal{R}_2(C)$ be the region in the average state space where the player 2's network dictates C.

LEMMA A.2. $\mathcal{R}_2(C) \cap [v_f(D, D), v_f(C, D)] \neq \emptyset$

PROOF. Suppose the contrary. Consider the triangle formed by 0, $v_f(D, D)$, and $v_f(C, D)$. Since $\mathcal{R}_2(C) \cap [v_f(D, D), v_f(C, D)] = \emptyset$, player 2 always plays D when the state variable is located on the triangle. Thus, starting from a point on the triangle and arbitrarily close to the target in the average state space, the continuation game outcome is either (C, D) or (D, D). As a result, the state variable will drift away to $[v_f(C, D), v_f(D, D)]$. This contradicts local stability. *Q.E.D.*

LEMMA A.3. $\mathcal{R}_2(C) \subset \{v_f : v_1 \leq 0\}$.

PROOF. Suppose that $\mathcal{R}_2(C) \setminus \{v_f : v_1 \leq 0\} \neq \emptyset$. Then $\mathcal{R}_2(C) \setminus \{v_f : v_1 \leq 0\}$ must intersect at least one surface of the average state space V_f. We shall show that player 1 has a strategy that generates a higher payoff than his target, which contradicts the hypothesis that $\mathcal{R}_2(C)$ is part of an equilibrium. The construction of the deviation strategy differs slightly depending upon the surface which $\mathcal{R}_2(C) \setminus \{v_f : v_1 \leq 0\}$ intersects. We will describe only the case where

$$[\mathcal{R}_2(C) \setminus \{v_f : v_1 \leq 0\}] \cap \Delta(C, C; D, D; C, D) \neq \emptyset.$$

The remaining cases follow from the same logic.

Note that by individual rationality, player 1 receives the highest payoff at $v_f(C, C)$ in $\Delta(C, C; D, D; C, D)$. If $v_f(C, C) \in \mathcal{R}_2(C)$, player 1 can obtain $v_1(C, C)$ by sticking to C once the state variable in on $\Delta(C, C; D, D; C, D)$. Thus,

$$v_f(C, C) \notin \mathcal{R}_2(C).$$

Then, there exists $v_f^* \in \mathcal{R}_2(C) \cap \Delta(C, C; D, D; C, D)$ such that $v_f^* \neq v_f(C, C)$ and player 1 receives a higher payoff in v_f^* than in any $v_f \in \mathcal{R}_2(C) \cap \Delta(C, C; D, D; C, D)$. By the definition of v_f^*, $[v_f(C, C), v_f^*]$ has no intersection with $\mathcal{R}_2(C)$ other than v_f^* itself. Let $\mathbf{H}^1$ be the boundary hyperplane of $\mathcal{R}_2(C)$ so that $v_f^* \in \mathbf{H}^1$. Choose the directional vector of $\mathbf{H}^1$ such that

$$v_f(C, C) \in \underline{\mathbf{H}}^1.$$

In the neighborhood of $\mathbf{H}^1$, player 2's strategy is completely determined by $\mathbf{H}^1$: player 2 plays D in $\underline{\mathbf{H}}^1$ and C in $\overline{\mathbf{H}}^1$. We will show that player 1 has a strategy that forces the average state variable to converge to $v_f^* \in \mathbf{H}^1$, which contradicts the assumption that $\mathcal{R}_2(C)$ is part of player 2's equilibrium strategy.

Let $\mathbf{H}(v_f^*)$ be the hyperplane spanned by v_f^*, $v_f(C, C)$ and 0. Clearly, $\mathbf{H}(v_f^*)$ separates $v_f(D, D)$ from $v_f(C, D)$, and

$$v_f(D, D) \in \overline{\mathbf{H}}(v_f^*).$$

Consider player 1's strategy that dictates D if the state variable is in $\underline{\mathbf{H}}(v_f^*) \cap \underline{\mathbf{H}}^1$, and D otherwise. In $\overline{\mathbf{H}}^1$, (C, C) is played. Since $v_f(C, C) \in \underline{\mathbf{H}}^1$, the state variable moves toward

$\underline{\mathbf{H}}^1$. In $\underline{\mathbf{H}}^1$, (D, D) and (C, D) are played. The state variable converges to $\mathbf{H}(v_f^*)$ and moves toward $\overline{\mathbf{H}}^1$, since (C, D) is played in $\overline{\mathbf{H}}(v_f^*)$ while $v_f(C, D) \in \underline{\mathbf{H}}(v_f^*)$, and (D, D) is played in $\underline{\mathbf{H}}(v_f^*)$ while $v_f(D, D) \in \overline{\mathbf{H}}(v_f^*)$. Therefore, the state variable converges to $\mathbf{H}^1 \cap \mathbf{H}(v_f^*)$. Since only (C, C), (D, D) and (C, D) are realized, the state variable converges to v_f^* on $\Delta(C, C; D, D; C, D)$. *Q.E.D.*

[Insert Figure 1 Here]

Construct two 2 hyperplanes in the average state space as depicted in Figure 1. Let $\mathbf{H}^2$ be the hyperplane determined by 0, $v_f(D, C)$, and $v_f(C, C)$. Let $\mathbf{H}^3$ be determined by 0, $v_f(D, D)$ and $v_f(C, D)$. Note that by construction,

$$v_f(D, D) \in \overline{\mathbf{H}}^2, v_f(C, D) \in \underline{\mathbf{H}}^2;$$

$$v_f(D, C) \in \overline{\mathbf{H}}^3, v_f(C, C) \in \underline{\mathbf{H}}^3.$$

Define $v_D = \mathbf{H}^3 \cap [v_f(C, C), v_f(D, C)]$ and $v_C = \mathbf{H}^2 \cap [v_f(D, D), v_f(C, D)]$. By individual rationality, at v_C, player 1's payoff is below the target. It follows that at v_D, player 1's payoff is above the target. By Lemma A.3, $v_D \notin \mathcal{R}_2(C)$. Therefore, there is a hyperplane $\mathbf{H}^1$ of player 2's network such that player 2 plays D in its v_D side and C in its v_C side. Choose the directional vector of $\mathbf{H}^1$ such that $v_C \in \overline{\mathbf{H}}^1$. In the neighborhood of $\mathbf{H}^1$, player 2's strategy is entirely determined by $\mathbf{H}^1$. Consider the following strategy of player 1: play D if and only if the state is in $(\underline{\mathbf{H}}^3 \cap \overline{\mathbf{H}}^1) \cup (\underline{\mathbf{H}}^2 \cap \underline{\mathbf{H}}^1)$, as depicted in the right of Figure 1. Note that this strategy can be carried out by a network with three classifiers. We want to show that given the two strategies of player 1 and player 2, there is "self-containing" neighborhood in the gross state space in that once the state variable falls into the neighborhood, it will stay there for the rest of the game.

Fix the gross state variable following a history. When the gross state variable is in $\underline{\mathbf{H}}^1$, (D, D) and (C, D) are realized. Note that $v_f(D, D) \in \overline{\mathbf{H}}^2$, but (D, D) is realized in a subset of $\underline{\mathbf{H}}^2$. Thus, as (D, D) is realized, the gross state variable must move toward $\mathbf{H}^2$ in $\underline{\mathbf{H}}^2$. Similarly, $v_f(C, D) \in \underline{\mathbf{H}}^2$, but (C, D) is realized in a subset of $\overline{\mathbf{H}}^2$. As (C, D) is realized, the gross state variable moves toward $\mathbf{H}^2$ in $\overline{\mathbf{H}}^2$. Since $\mathbf{H}^3$ is spanned by 0, $v_f(D,D)$ and $v_f(C,D)$, the gross state variable maintains the same distance to $\mathbf{H}^3$ as (C,D) and (D,D) are realized. Since $v_C = \mathbf{H}^2 \cap [v_f(D,D), v_f(C,D)]$, as (C,D) and (D,D) are realized, the average state variable must move toward v_C which is located in $\overline{\mathbf{H}}^1$. Thus, in a finite number of periods (C,D) and (D,D), the gross state variable enters $\overline{\mathbf{H}}^1$.

We can apply a symmetric argument in $\overline{\mathbf{H}}^1$ to show that as (C,C) and (D,C) are realized, the state variable moves toward v_D in the average state space, while the gross state variable maintains the same distance from $\mathbf{H}^2$. Each time the gross state variable passes through $\mathbf{H}^1$, it must move toward $\mathbf{H}^2$ or $\mathbf{H}^3$. Therefore, there exists a self-containing neighborhood in the gross state space. The size of the neighborhood in the gross state space is completely determined by the orientation of $\mathbf{H}^1$. The gross state space expands as the game continues. When the gross state variable falls into the self-containing neighborhood, player 1 can switch to the 3-hyperplane strategy. Then gross state variable will remain in the neighborhood, implying that player 1 obtains his equilibrium payoff with the 3-hyperplane strategy. This completes the proof of Proposition A.1. *Q.E.D.*

A.3. No More than 2 Hyperplanes

PROPOSITION A.4. *There is at least one player who uses at most two hyperplanes in any equilibrium.*

PROOF. Local stability implies that in any equilibrium, $v_f(D,D) \in \mathcal{R}_1(C) \cup \mathcal{R}_2(C)$. To see this, consider the line segment $[0, v_f(D,D)]$. Since $\mathcal{R}_i(C)$ is a cone for each i, if $v_f(D,D) \notin \mathcal{R}_1(C) \cup \mathcal{R}_2(C)$, the line is contained in a region where both players play D. Once the state variable is in the region, (D,D) will be played for the rest of the game and the long run outcome is (D,D), which contradicts local stability. Without loss of generality, we assume for the rest of the paper that

$$v_f(D,D) \in \mathcal{R}_2(C). \tag{A.1}$$

By Proposition A.1, $\mathcal{R}_2(C)$ is formed by at most 3 hyperplanes. Let $\mathbf{H}^{2j}$ be the boundary hyperplanes where $j = 1, 2, 3$. Choose the directional vector such that $v_f(D, D) \in \overline{\mathbf{H}}^{2j}$ for each j. By (A.1),

$$v_f(D, D) \in \bigcap_{j=1}^{3} \overline{\mathbf{H}}^{2j} \subset \mathcal{R}_2(C).$$

LEMMA A.5. *There exists at least one hyperplane, say $\mathbf{H}^2$, of player 2's network such that* $[v_f(D, D), v_f(C, D)] \subset \overline{\mathbf{H}}^2$.

PROOF. Suppose not. Then, for each $j = 1, 2, 3$, $v_f(C, D) \in \underline{\mathbf{H}}^{2j}$. Then, $\exists \lambda' > 0$ such that $\forall \lambda \in [1, \lambda']$, $(1 - \lambda) v_f(C, D) \in \overline{\mathbf{H}}^{2j}$, for each $j = 1, 2, 3$. Thus, $(1 - \lambda) v_f(C, D) \in \mathcal{R}_2(C)$. But, by individual rationality $u_1(C, D) \leq v_1^*$, player 1 obtains a payoff greater than his target payoff at $(1 - \lambda) v_f(C, D)$. This contradicts Lemma A.3. *Q.E.D.*

We will show that player 1 has a strategy with 2 hyperplanes to achieve his equilibrium payoff. One hyperplane is $\mathbf{H}^2$. The other one is $\mathbf{H}^*$ with appropriate choice of directional vector such that

$$0 \in \mathbf{H}^*,$$

$$v_f(D, C) \in \underline{\mathbf{H}}^*, v_f(D, D), v_f(C, D), v_f(C, C) \in \overline{\mathbf{H}}^*,$$

$$\mathbf{H}^* \cap [v_f(C, C), v_f(D, C)] \in \underline{\mathbf{H}}^2,$$

$$\mathbf{H}^* \cap \mathbf{H}^2 \cap \Delta(C, C; C, D; D, C) \neq \emptyset.$$

(First, choose the triangle spanned by 0, $v_f(D, D)$ and $v_f(C, D)$ in the average state space. Next, move this triangle by slightly shifting the vertex at $v_f(D, D)$ toward $v_f(D, C)$ along $[v_f(D, C), v_f(D, D)]$, and then by slightly shifting another vertex at $v_f(C, D)$ toward $v_f(D, C)$ along $[v_f(D, C), v_f(C, D)]$. Let $\mathbf{H}^*$ be the hyperplane that embeds resulting triangle.) Consider the following strategy of player 1 determined by $\mathbf{H}^2$ and $\mathbf{H}^*$: play D if and only if the state variable is in $\overline{\mathbf{H}}^2 \cap \overline{\mathbf{H}}^*$. We now show that player 1 achieves his target payoff with the above strategy.

In the neighborhood of $\mathbf{H}^2 \cap \mathbf{H}^*$, player 2 plays C in $\overline{\mathbf{H}}^2$ and D in $\underline{\mathbf{H}}^2$. There exists a self-containing neighborhood of $\mathbf{H}^2 \cap \mathbf{H}^*$. To see this, note that only 3 outcomes, (C, C),

(D,C), and (C,D), are realized in the neighborhood of $\mathbf{H}^2 \cap \mathbf{H}^*$. (C,C) and (D,C) are realized in $\overline{\mathbf{H}}^2$. Since $v_f(C,C) \in \overline{\mathbf{H}}^*$ and $v_f(D,C) \in \underline{\mathbf{H}}^*$, the state variable converges to $\mathbf{H}^* \cap \overline{\mathbf{H}}^2$. But since $\mathbf{H}^* \cap [v_f(C,C), v_f(D,C)] \in \underline{\mathbf{H}}^2$, the state variable moves toward $\mathbf{H}^2 \cap \mathbf{H}^*$. (C,D) is realized in $\underline{\mathbf{H}}^2$, but since $v_f(C,D) \in \overline{\mathbf{H}}^2$, the state variable moves toward $\mathbf{H}^2 \cap \mathbf{H}^*$. Therefore, for the all three possible orientations of $\mathbf{H}^2$, there is a self-containing neighborhood of $\mathbf{H}^2 \cap \mathbf{H}^*$. Moreover, since (D,D) is not realized, the average state variable converges to $\mathbf{H}^2 \cap \mathbf{H}^* \cap \Delta(C,C;C,D;D,C)$, which gives player 1 his equilibrium long run payoff. This completes the proof of Proposition A.4. *Q.E.D.*

A.4. No More Than a Single Hyperplane

This is the third and last step in the proof of Theorem 3.1. We will first show that there is at least one player who can achieve his equilibrium long run payoff with a single hyperplane, if the target is in the interior of the average state space. To explain the key idea of the proof, we first examine a highly stylized case. Assume that each player is using a single hyperplane strategy. By Lemma 3.3, player i's hyperplane $\mathbf{H}_i$ must be such that he plays D if and only if his opponent's *payoff* is below the target. Since each player's strategy is conditioned only on the average payoff, we can describe the strategy in the payoff space, a subset of $\mathbb{R}^2$ as depicted in Figure 2.

[Insert Figure 2 Here]

%put $v_f(C,D)$ [rb] at -0.1 8.1

One can easily check that following any history, the average payoff vector induced by this pair of strategies converges to the target payoff vector.[11] We examine how the frequency of outcomes change when we rotate $\mathbf{H}_1$ "counterclockwise" around the target payoff to $\mathbf{H}^*$ as depicted in Figure 2. Notice that as we rotate $\mathbf{H}_1$ to $\mathbf{H}^*$, the regions where (C,C) and (D,D) are realized are "squeezed". Let $\lambda > 0$ be the sharp angle between $\mathbf{H}^*$

[11] For precise logic, see Cho [1995].

and $\mathbf{H}_2$. It is easy to show that for any $\lambda > 0$, the limit payoffs induced by the pair of the single hyperplane strategies are still the target payoff vector. For the following lemma, consider a hyperplane passing through the $\mathbf{H}^* \cap \mathbf{H}_2$, and parallel to $[v_f(C,D), v_f(D,C)]$. Denote this hyperplane as $\mathbf{H}^o$.

LEMMA A.6. *[1] If $[v_f(C,D), v_f(D,C)] \in \underline{\mathbf{H}}^o$, then $\exists\ \overline{\lambda} > 0$ such that $\forall \lambda \in (0, \overline{\lambda})$, the limit frequency of (C,C) is 0. [2] If $[v_f(C,D), v_f(D,C)] \in \overline{\mathbf{H}}^o$, then $\exists\ \overline{\lambda} > 0$ such that $\forall \lambda \in (0, \overline{\lambda})$, the limit frequency of (D,D) is 0.*

PROOF. We only examine [1] which is depicted in Figure 2. The other case follows from the same logic. Suppose that in period T, the gross state variable is in $\underline{\mathbf{H}}^o$, and in period $T+k$ $(k \geq 1)$, (C,C) is played. Then, there exists non-negative $k' < k$ such that in period $T + k'$ (D,D) is played and the state variable enters $\overline{\mathbf{H}}^o$. This is because if (C,D) or (D,C) is played in $\underline{\mathbf{H}}^o$, the state variable cannot approach $\overline{\mathbf{H}}^o$, and (C,C) is played in a subset of $\overline{\mathbf{H}}^o$. We will show that if $\lambda > 0$ is sufficiently small, (C,C) is never realized after $T + k'$. This then implies that starting with any initial condition in $\underline{\mathbf{H}}^o$, (C,C) will not be realized in the continuation of the play. The lemma then follows from combining this result with the fact that starting with any initial condition in $\overline{\mathbf{H}}^o$, the state variable drifts into $\underline{\mathbf{H}}^0$ in finite number of periods because $v_f(C,D)$, $v_f(D,C)$ and $v_f(C,C) \in \underline{\mathbf{H}}^*$.

Note that if (C,C) is not played after $T + k'$, the state variable must return to $\underline{\mathbf{H}}^o$ in finite number, say N, of periods as (C,D) and (D,C) are played. Both $v_f(C,D)$ and $v_f(D,C)$ are in $\underline{\mathbf{H}}^o$. Furthermore, we can choose N as a uniform upper bound. To see this, note that when the state variable enters $\overline{\mathbf{H}}^o$, its distance from $\mathbf{H}^o$ is bounded by the size of the increment of the gross state variable. When either (C,D) or (D,C) is played in $\overline{\mathbf{H}}^o$, the distance between the state variable and $\mathbf{H}^o$ decreases.

Let us determine the area from which the state variable can move into $\overline{\mathbf{H}}^o$ after (D,D) is played. First, find the hyperplane such that if it is shifted by the vector $v_f(D,D)$, it coincides with $\mathbf{H}^o$. The triangle formed by this hyperplane and the area where (D,D) is played is what we are looking for. Now, shift this triangle by $v_f(D,D)$. If this triangle has no intersection with the area where (C,C) is played, (C,C) does not immediately follow (D,D). In this case, the triangle must be contained entirely in either the (C,D) or

(D,C) area, i.e., either (C,D) or (D,C) follows (D,D). Then we can shift the triangle by $v_f(C,D)$ or $v_f(D,C)$. If there exists an intersection with the (C,C) area after the initial shift, we reduce $\lambda > 0$ until this intersection disappears. Then, the resulting triangle is again contained entirely in either the (C,D) or (D,C) area. Either (C,D) or (D,C) is played, and we can shift it accordingly. After we repeat the same exercise N times, the triangle must be entirely contained in $\underline{\mathbf{H}}^o$, since the state variable returns to $\underline{\mathbf{H}}^o$ after (C,D) or (D,C) is played N times in $\overline{\mathbf{H}}^0$. The resulting λ is the $\overline{\lambda}$ stated in the lemma. It is clear that for any smaller $\lambda > 0$, the shifted triangles still do not intersect with the (C,C) area. *Q.E.D.*

LEMMA A.7. *Suppose that $0 \in \Delta(C,D;D,C;s)$ with $s \in \{(C,C),(D,D)\}$. Then if there is a hyperplane of player i which separates $v_f(C,D)$ from $v_f(D,D)$, player $j \neq i$ has a strategy with a single hyperplane that enforces his equilibrium payoff. [2] Suppose that $0 \in \Delta(C,C;D,D;s)$ with $s \in \{(C,D),(D,C)\}$. Then if there is a hyperplane of player i that separates $v_f(C,C)$ from $v_f(D,D)$, player $j \neq i$ has a single hyperplane strategy that enforces his equilibrium payoff.*

PROOF. We only prove the case where

$$0 \in \Delta(C,D;D,C;C,C). \tag{A.2}$$

The other cases can be shown with similar arguments. By Proposition A.1, each player uses a strategy with two hyperplanes. Suppose that player 1's strategy is formed by $\mathbf{H}^1$ and $\mathbf{H}^3$, and player 2's strategy is formed by $\mathbf{H}^2$ and $\mathbf{H}^4$. It follows from Lemma A.3 that each $\mathcal{R}_i(C)$ is a convex cone formed by the two hyperplanes. To prove the lemma, we assume that

$$v_f(C,D) \in \underline{\mathbf{H}}^1, \quad v_f(D,C) \in \overline{\mathbf{H}}^1, \tag{A.3}$$

and show that player 2 has a single hyperplane strategy to achieve his target.

By (A.2) and (A.3), we can construct a hyperplane $\mathbf{H}^*$ (by rotating $\mathbf{H}^1$ slightly as we did in Lemma A.6) such that

$$\mathbf{H}^* \cap \mathbf{H}^1 \cap \Delta(C,C;C,D;D,C) = \{0\};$$

$$v_f(D,C) \in \underline{\mathbf{H}}^*;$$

$$v_f(C,D) \in \overline{\mathbf{H}}^*;$$

$$\mathbf{H}^* \cap [v_f(C,D), v_f(D,C)] \in \overline{\mathbf{H}}^1.$$

Since $0 \in \Delta(C,D;D,C;C,C)$, there is a hyperplane $\mathbf{H}^o$ such that with an appropriate directional vector for $\mathbf{H}^o$, we have $\mathbf{H}^* \cap \mathbf{H}^1 \subset \mathbf{H}^o$ and $v_f(D,C), v_f(C,D) \in \overline{\mathbf{H}}^o$. Invoking the same logic as the proof of Lemma A.6, we can show that the long run frequency induced by $\mathbf{H}^1$ and $\mathbf{H}^*$ must be 0 if the sharp angle $\lambda > 0$ between the two hyperplane is sufficiently small. *Q.E.D.*

PROPOSITION A.8. *If the target is in the interior of the average state space, then at least one player uses a single hyperplane strategy in equilibrium.*

PROOF. By (A.1), we can assume that $v_f(D,D) \in \mathcal{R}_2(C) = \overline{\mathbf{H}}^2 \cap \overline{\mathbf{H}}^4$. Since 0 is in the interior of V_f, by Lemma A.3,

$$\ell = \mathbf{H}^2 \cap \mathbf{H}^4 \subset \{v_f : v_1 = 0\}. \tag{A.4}$$

Since V_f is convex, and ℓ passes through a point in the interior of V_f, ℓ must has an intersection with exactly two surfaces of V_f. For the rest of this proof, by v_f^x for some superscript x, we mean $(v_1^x, v_2^x, f^x) \in V_f$. For $v_f, v_f' \in V_f$, define

$$[v_f, v_f'] = \{\lambda v_f + (1-\lambda)v_f' : \lambda \in [0,1]\};$$

$$(v_f, v_f') = \{\lambda v_f + (1-\lambda)v_f' : \lambda \in (0,1)\}.$$

First, suppose for some $s \in \{(C,C),(D,C)\}$,

$$\ell \cap \Delta(C,D;D,D;s) \neq \emptyset.$$

Define $v_f^o = \ell \cap \Delta(C,D;D,D;s)$. By (A.4), $v_1^o = 0$. Since v_1^o is an equilibrium payoff,

$$v_1^o \geq \underline{v}_1 \geq \max\left[u_1(D,D), u_1(C,D)\right].$$

If $\exists v_f \in \Delta(C,D;D,D;s)$ such that $v_1 \geq \underline{v}_1$, then it is necessary that $u_1(s) \geq v_1^o$. By Lemma A.3, there exists a boundary hyperplane, say $\mathbf{H}^2$, such that

$$v_f^\theta = \mathbf{H}^2 \cap [v_f(D,D), v_f(s)] \neq \emptyset. \tag{A.5}$$

By Lemma A.3, $v_1^\theta \leq v_1^o = 0$. Moreover,

$$v_f(D,C) \in \underline{\mathbf{H}}^2. \tag{A.6}$$

For $\lambda \in \mathbb{R}$, define $v_f^\lambda = \lambda v_f^o + (1-\lambda)v_f^\theta$. Then, $\forall \lambda \geq 1$,

$$v_1^\lambda \geq v_1^o = 0. \tag{A.7}$$

Since $v_f^o, v_f^\theta \in \mathbf{H}^2$, $v_f^\theta = \{v_f^\lambda : \lambda \in \mathbb{R}\} \cap [v_f(D,D), v_f(s)]$. Define

$$\lambda' = \max\{\lambda \geq 1 : v_f^\lambda \in \Delta(C,D;D,D;s)\}.$$

Since $v_f^\theta \in [v_f(D,D), v_f(s)]$, $v_f^{\lambda'}$ must be in either $[v_f(C,D), v_f(D,D)]$ or $[v_f(C,D), v_f(s)]$. By (A.7), at $v_f^{\lambda'}$, $v_1^{\lambda'} \geq v_1^o$. But, $\forall v_f \in (v_f(D,D), v_f(C,D))$, $v_1 < \underline{v}_1 \leq v_1^o$. Therefore, $v_f^{\lambda'}$ cannot be located in $(v_f(C,D), v_f(D,D))$, which implies that

$$v_f^{\lambda'} \in [v_f(C,D), v_f(s)]. \tag{A.8}$$

If $s = (C,C)$, then by (A.5), $v_f(C,C) \in \underline{\mathbf{H}}^2$. By Lemma A.7, player 1 can enforce $v_f^o \in \Delta(D,D;C,D;C,C)$. If $s = (D,C)$, then by (A.8) and (A.6), $v_f(C,D) \in \overline{\mathbf{H}}^2$. By Lemma A.7, player 1 can enforce $v_f^o \in \Delta(D,D;C,D;D,C)$.

It remains to prove the proposition when ℓ intersects neither $\Delta(C,D;D,D;C,C)$ nor $\Delta(C,D;D,D;D,C)$. Since ℓ intersects exactly 2 surfaces of V_f, this case is equivalent to

$$\begin{aligned} v_f^u =& \ell \cap \Delta(C,C;D,C;D,D) \neq \emptyset \\ v_f^d =& \ell \cap \Delta(C,C;D,C;C,D) \neq \emptyset. \end{aligned} \tag{A.9}$$

By (A.4), $v_1^u = v_1^d = 0$. If there exists $\mathbf{H}^{2j}$ such that $v_f(C,C) \in \underline{\mathbf{H}}^{2j}$, Lemma A.7 implies that player 1 can enforce v_f^u by a single hyperplane.

Let us assume for the rest of the proof that $v_f(C,C) \in \overline{\mathbf{H}}^{2j}$ for $j = 1,2$, which is equivalent to

$$[v_f(D,D), v_f(C,C)] \subset \overline{\mathbf{H}}^2 \cap \overline{\mathbf{H}}^4.$$

Under this assumption, Lemma A.3 implies that

$$v_1^d > u_1(C,C) \tag{A.10}$$

which is possible only if $u_1(D,C) = \max_{s \in S} u_1(s) \geq u_1(C,C)$. It then follows that $v_f(D,C) \in \{v_f : v_1 \geq 0\}$, and that there exists a boundary hyperplane, say $\mathbf{H}^2$, such that $v_f^\theta = \mathbf{H}^2 \cap [v_f(D,D), v_f(D,C)]$. By Lemma A.3, $v_1^\theta \leq v_1^u$. For each $\lambda \in \mathbb{R}$, define

$$v_f^\lambda = \lambda v_f^u + (1-\lambda) v_f^\theta.$$

Clearly, $\forall \lambda \geq 1$, $v_1^\lambda \geq v_1^u$. Define

$$\lambda^* = \max\{\lambda \geq 1 : v_f^\lambda \in \Delta(C,C;D,D;D,C)\}.$$

Since $v_1^\theta \leq v_1^u$, $v_1^{\lambda^*} \geq v_1^u = v_1^d$. By (A.10), $v_f^{\lambda^*} \in [v_f(C,C), v_f(D,C)]$. Since $v_f^{\lambda^*} \in \mathbf{H}^2$ and $v_f^{\lambda^*} \geq v_1^d$, Lemma A.3 implies that $v_f^{\lambda^*} \in \underline{\mathbf{H}}^4$. Combining this observation with (A.9), we conclude that

$$v_f' = \mathbf{H}^4 \cap [v_f(C,C), v_f(D,C)] \in [v_f(C,C), v_f^{\lambda^*}]. \tag{A.11}$$

By Lemma A.3, $v_1' \leq v_1^d = 0$. Define

$$\lambda'' = \max\{\lambda \geq 1 : \lambda v_f^d + (1-\lambda) v_f' \in \Delta(D,C;C,C;C,D)\}.$$

Since $\lambda'' \geq 1$ and $v_1' \leq v_1^d$,

$$v_f'' = \lambda'' v_f^d + (1-\lambda'') v_f' \in \{v_f : v_1 \geq 0\}.$$

By construction, v_f'' must be in either $[v_f(C,C), v_f(C,D)]$ or $[v_f(C,D), v_f(D,C)]$. Since $v_1 < v_1(C,C) < v_1^d = 0$ for all $(v_1, v_2, f) \in (v_f(C,C), v_f(C,D))$,

$$v_f'' \notin (v_f(C,C), v_f(C,D)).$$

Therefore,

$$v_f'' \in [v_f(C,D), v_f(D,C)].$$

By (A.11), $v_f(D,C) \in \underline{\mathbf{H}}^4$. Since $v_f'' \in \mathbf{H}^4$, we conclude that $v_f(C,D) \in \overline{\mathbf{H}}^4$. Lemma A.7 implies that player 1 can enforce v_f^d by a single hyperplane. *Q.E.D.*

A similar argument applies if the target is on a surface of the average state space. It is straightforward to show that if in equilibrium one player uses a single hyperplane strategy, then the other player also uses a single hyperplane strategy. This completes the proof of Theorem 3.1. *Q.E.D.*

References

Dilip Abreu and Ariel Rubinstein [1988]: "The Structure of Nash Equilibrium in Repeated Games with Finite Automata," *Econometrica*, 56, 1259-1282.

Jeffrey S. Banks and Rangarajan Sundaram [1990]: "Repeated Games, Finite Automata and Complexity" *Games and Economic Behavior*, 2, 97-117.

Margaret Bray [1983]: "Convergence to Rational Expectations Equilibrium," INDIVIDUAL FORECASTS AND AGGREGATE OUTCOMES edited by Roman Frydman and Edmund Phelps, Cambridge University Press.

In-Koo Cho [1995]: "Perceptrons Play the Repeated Prisoners' Dilemma" *Journal of Economic Theory*, 67, 266–284

In-Koo Cho [1996a]: "Perceptrons Play Repeated Games with Imperfect Monitoring" *Games and Economic Behavior*, 16, 1996: 22-53

In-Koo Cho [1996b]: "Learning to Coordinate in Repeated Games" mimeo, Brown University

Drew Fudenberg and David M. Kreps [1993]: "Learning Mixed Equilibria," *Games and Economic Behavior*, 5, 320-367.

Kurt Hornik, Maxell Stinchcombe and Halbert White [1989]: "Multi-layer Feedforward Networks are Universal Approximators," mimeo, University of California, San Diego

Hao Li [1997]: "Hierarchies and Information Processing Organizations," mimeo, University of Hong Kong.

Michele Piccione and Ariel Rubinstein [1993]: "Finite Automata Play a Repeated Extensive Game" *Journal of Economic Theory*, 61, 160-168.

Roy Radner [1993]: "Organization of Decentralized Information Processing," *Econometrica*, 61, 1109-1146.

Ariel Rubinstein [1986]: "Finite Automata Play the Repeated Prisoner's Dilemma," *Journal of Economic Theory*, 39. 83-96.

Steve Smale [1980]: "The Prisoner's Dilemma and Dynamical Systems Associated with Non-Cooperative Games," *Econometrica*, 48, 1617-1634.

Timothy Van Zandt [1996]: "Decentralized Information Processing in the Theory of Organizations," forthcoming in *Contemporary Economic Development Reviewed, Volume 4: The Enterprise and its Environment*, edited by Murat Sertel, London: MacMillan Press Ltd.

Gérard Weisbuch [1990]: COMPLEX SYSTEM DYNAMICS: AN INTRODUCTION TO AUTOMATA NETWORKS, Lecture Note Volume II, Santa Fe Institute Studies in the Sciences of Complexity, Addison-Wesley.