Fictitious Play in Network Aggregative Games

AH, FB

April 23, 2021

1 Paper Structure

1. Introduction: Explain the notion of a Network Aggregative Game, and Fictitious Play with literature on the state of the art. Assumptions made as well as summary of results
2. Maybe? Relation with the model of Grammatico
3. Preliminaries
   - The Network Aggregative Game model
   - Fictitious Play, as well as the CTFP property
4. Convergence of Fictitious Play
   - Existence of the Nash Equilibrium
   - Existence of a CTFP
   - Convergence of CTFP to a fixed point
   - Corrolary: when $w_{ii} = 0$ (i.e. the interaction graph is simple) CTFP converges to a NE
   - Convergence of CTFP to the NA-CCE set
5. Numerical Experiments
   - Convergence rates w.r.t. connectivity and size
   - Non-convergent examples
   - Agent based simulations
6. Discussion of Experiments
7. Concluding Remarks

2 Introduction

2.1 Contributions

The main contribution of this work is to study the behaviour of the Fictitious Play Learning Algorithm, in continuous time, when applied on a Network Aggregative Matrix Game. FB: *why FP and NAG are important?*

We first show that a Nash Equilibrium exists and that Fictitious Play admits solutions in this setting. In particular, we study zero-sum games and show that Fictitious Play converges to a fixed point, which for a simple FB: *define simple* network corresponds to a Nash Equilibrium. In addition, we find that, for games which are not zero-sum, agents following Continuous Time Fictitious Play are able to achieve no regret.

Finally, we explore the Fictitious Play algorithm through numerical simulations to understand how convergence rates depend on the network size and connectivity. We also show that it does not necessarily converge for more complex games and can, in fact, lead to more complex dynamics. In addition, we experiment the Fictitious Play algorithm on an agent based model for (resource allocation? voting?). Finally, our experiments document how noise affects the convergence of fictitious play, whose theoretical analysis we point out as an interesting direction for future work.

AH: *Possible additions: 1) Show that the flow is unique almost everywhere in CTFP. 2) Show the same results (i.e. existence and convergence) for Discrete Time Fictitious Play.*

2.2 Related Work

Main articles: Ewerhart (FP in Networks), Grammatico (Nash Equilibrium seeking in NA games), Perrin (FP in mean field games), Harris (FP in 2 player games).

3 Preliminaries

3.1 Network Aggregative Games

The model we consider is that there is a set $\mathcal{N} = 1, \ldots, N$ of agents who are connected through an underlying interaction graph. This graph is given by the tuple $(\mathcal{N}, (\mathcal{E}, W))$ in which $\mathcal{E}$ is the set of all pairs $(\mu, \nu) \in \mathcal{N} \times \mathcal{N}$ such that agent μ and agent ν are connected. FB:

Definition (Interaction Graph)**.** Given a set $\mathcal{N}$ of agents, an *interaction graph* $I = (\mathcal{N}, (\mathcal{E}, W))$ such that

- $\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$. Then, the set of neighbours of agent μ is denoted as $N^\mu = \{\nu \in \mathcal{N} : (\mu, \nu) \in \mathcal{E}\}$.
- $W \in M_N(\mathbb{R})$ is the weighted adjacency matrix whose elements $w^{\mu\nu}$ expresses the importance that agent μ places on agent ν. If $(\mu, \nu) \notin \mathcal{E}$ then $w^{\mu\nu} = 0$; $w^{\mu\nu} \in (0, 1]$ otherwise.

As a network game, each agent μ has an associated set of pure strategies S^μ which has cardinality $|S^\mu| = n$. FB: *what is a strategy then? how can we define them if we didn't introduce actions?*

Then we can construct the FB: *unit-simplex ?* on this action set as $\Delta_\mu := \{x \in \mathbb{R}^n : \sum_i x_i^\mu = 1\}$ where x_i^μ is the probability with which agent μ plays the pure strategy i. As such we refer to x^μ as the *state* of agent μ (this is often referred to as μ's mixed strategy). Also associated with each agent is a utility function $u^\mu(x^\mu, x^{-\mu})$ in which we use the standard notation $-\mu$ to refer to all agents other than μ. Notice that this requires that each agent plays the same strategy against all of their neighbours.

With all of the above definitions, we can formalise the network aggregative (NA) game as the tuple $\Gamma = (\mathcal{N}, (\mathcal{E}, W), (S^\mu, u^\mu)_{\mu \in \mathcal{N}})$. FB: *preferably use a definition environment.*

What is unique about the NA game is the structure of the payoffs themselves. In this format, each agent is receives a reference σ which is a weighted sum of each of their neighbours state. Formally

$$\sigma^\mu = \sum_{\nu \in N^\mu} w^{\mu\nu} x^\nu. \tag{1}$$

Then, the agent's must optimise their payoff with respect to this reference vector. Thus, instead of having to consider the actions of the entire population, or play individual games against each of their neighbours, the agent only has to consider σ^μ as a 'measurement' of the local aggregate state and optimise with respect to this measurement. This allows us to make the reduction $u^\mu(x^\mu, x^{-\mu}) = u^\mu(x^\mu, \sigma^\mu)$. In particular, we consider that the agent is engaged in a matrix game against the reference vector so that

$$u^\mu(x^\mu, \sigma^\mu) = x^\mu \cdot A^\mu \sigma^\mu = x^\mu \cdot A^\mu \sum_{\nu \in N^\mu} w^{\mu\nu} x^\nu. \tag{2}$$

where A^μ is the payoff matrix associated to agent μ. Note that this means we can write the game Γ with the payoff matrices $A^m u$ in place of the utility functions u^μ. The NA game allows for the reduction of a multiplayer game into a series of two-player games. The agent's goal is to maximise their payoff with respect to the reference vector. As such, we define the best response correspondence BR^μ which maps any σ^μ the set $\arg\max_{y \in \Delta_\mu} u^\mu(y, \sigma^\mu)$. This leads naturally to the definition of a Nash Equilibrium in an NA Game as

Definition. (NE) The set of vectors $\{\bar{x}^\mu\}_{\mu \in \mathcal{N}}$ is an NE if, for all μ,

$$\bar{x}^\mu \in \arg\max_{x \in \Delta_\mu} u^\mu(x, w^{\mu\mu} x + \sum_{\nu \in N^\mu \setminus \{\bar{\mu}\}} w^{\mu\nu} \bar{x}^\nu).$$

FB: *It this a definition or rather a lemma?*

We will show that such an equilibrium exists in .... Finally we note that an NA game is *zero-sum* if the utilities of each agent sum to zero for any strategy set $\{x^\mu\}_{\mu \in \mathcal{N}}$. Formally

$$\sum_\mu u^\mu(x^\mu, \sum_{\nu \in N^\mu} w^{\mu\nu} x^\nu) = 0. \tag{3}$$

3.2 Continuous Time Fictitious Play

AH: *For Francesco: is it worth giving a quick introduction to Fictitious Play by talking about two player games?*
FB: *I'd rather give the definition first and then the example.*

Fictitious Play requires that, at the current time, each agent considers the average behaviour of their opponent in the past and play a best response to that strategy. This is best illustrated in the two-player setting. Consider the two player normal form game $\Gamma_2 = (\{1, 2\}, ((S^1, A), (S^2, B)))$ so that A and B are the payoff matrices of agent 1 and 2 respectively. As a remark, note that we can write the two player normal form game as an NA game simply by requiring that $\mathcal{E} = \{(1, 2), (2, 1)\}$ and W is a 2x2 matrix with zeros on its leading diagonal and ones on the off diagonal. We write the time-average of both agents' strategies as

$$\alpha^i = \frac{1}{t} \int_0^T x^i(t)\, dt \text{ for } i \in \{1, 2\} \tag{4}$$

FB: *why using an integral rather than a simple summation?*

FB: *decide whether you start enumerations from 1 or 0.*

In this manner, the $\alpha^\mu(t)$ denotes the time average of the strategies played by agent μ up to time t. Then, fictitious play requires that the agents update their strategy as $x^1(t) \in BR^1(\alpha^2(t))$ and $x^2(t) \in BR^2(\alpha^1(t))$.

We extend this naturally to the NA game Γ by requiring that each agent update their strategy according to the time average of their reference vector σ^μ. To formalise this, we write

$$\alpha^\mu_\sigma = \frac{1}{t} \int_0^t \sigma^\mu(s)\, ds. \tag{5}$$

Using this, we follow in the footsteps of Ewerhart [] and Harris [] to define Fictitious Play in continuous time.

Definition (Continuous Time Fictitious Play (CTFP))**.** A CTFP is defined as a measurable map m with components m_i such that for all i and all $t \geq 1$, $m_i : [0, \infty) \to \Delta_i$ satisfies $m_i(t) \in BR_i(\alpha_{\sigma_i})$ for all $t \geq 1$.

We can think of this definition as saying that the player plays some arbitrary strategy before $t = 1$, but beyond this it must play a best response to the time average of its reference signal.

3.3 Assumption

With the above preliminaries in place, we can state the assumptions that we make in this study.

Assumption 1. The weighted adjacency matrix W is constant and *row stochastic* meaning that the sum elements in each row of W is equal to one.

Assumption 2. The payoffs are given through matrix games and, therefore, are bilinear.

Assumption 3. The cardinality of each action set $|S^\mu|$ is equal for all agents.

Assumption 4. The NA game is zero-sum in the sense that $\sum_\mu u^\mu(x^\mu, \sum_{\nu\in N^\mu} w^{\mu\nu}x^\nu)$ for any set of states $(x^\mu)_{\mu\in N^\mu}$

FB: *discuss how restrictive each of these assumptions is.*

4 Convergence of Fictitious Play

4.1 Existence of the Nash Equilibrium

Note that the NE Condition requires

$$\begin{aligned}\bar{x}^\mu &\in \arg\max_{x\in\Delta_\mu} u^\mu(x, w^{\mu\mu}x + \sum_{\nu\in N^\mu} w^{\mu\nu}\bar{x}^\nu) \\ &=: \arg\max_{x\in\Delta_i} \bar{u}^\mu(x, \sum_{\nu\in N^\mu} w^{\mu\nu}\bar{x}^\nu) \qquad (6)\end{aligned}$$

where we can find $\bar{u}_i$ through the following argument

$$\begin{aligned}u^\mu(x, w^{\mu\mu}x + \sum_{\nu\in N^\nu} w_{\mu\nu}\bar{x}^\nu) &= x\cdot A^\mu(w^{\mu\mu}x + \sum_{\nu\in N^\mu} w_{\mu\nu}\bar{x}^\nu) \qquad (7)\\ &= x\cdot(w^{\mu\mu}A^\mu)x + \sum_{\nu\in N^\mu} u^{\mu\nu}(x,\bar{x}^\nu) \qquad (8)\\ &=: \bar{u}^\mu(x, \sum_{\nu\in N^\mu} w^{\mu\nu}\bar{x}^\nu),\end{aligned}$$

where $u^{\mu\nu}(x^\mu, x^\nu) = x^\mu\cdot A^\mu x^\nu$. Note that, in order to get this formulation, we had to use the assumption of payoffs being bilinear so that we could separate out the term in the weighted sum involving x from $\bar{x}^\nu$.

To show existence of an NE we will need the following definition and theorem.

Definition (Upper Semi-Continuous)**.** A compact-valued correspondence $\Phi : A \to B$ is *upper semi-continuous* at a point a if $g(a)$ is non-empty and if, for every sequence $a_n \to a$ and every sequence (b_n) such that $b_n \in g(a_n)$ for all n, there exists a convergent subsequence of (b_n) whose limit point b is in $g(a)$.

Theorem 1 (Kakutani)**.** *Let $K \subset \mathbb{R}^n$ be compact and convex and $\Phi : K \to K$ be closed or upper semi-continuous, with nonempty, convex and compact values. Then Φ has a fixed point.*

Note that, when acting on a simplex, the function $\Phi : \Delta \to \mathbf{P}(\Delta)$, where $\mathbf{P}(\Delta)$ denotes the nonempty, closed and convex subsets of Δ only has to satisfy the upper-semi continuity condition to admit a fixed point.

Theorem 2 (Existence of NE)**.** *Under the assumption (II), namely that the payoff function achieves a bilinear property, a Nash Equilibrium $\{\bar{x}^\mu\}_{\mu\in\mathcal{N}}$ exists.*

Proof. We begin by rewriting the NE condition by concatenating the set of NE vectors into one long colunn vector. This gives

$$\begin{bmatrix}\bar{x}^1\\ \cdot\\ \cdot\\ \cdot\\ \bar{x}^N\end{bmatrix} \in \begin{bmatrix}\arg\max_{y\in\Delta_1}\bar{u}^1(y, \sum_{\nu\in N^1} w^{1\nu}\bar{x}^\nu)\\ \cdot\\ \cdot\\ \cdot\\ \arg\max_{y\in\Delta_N}\bar{u}^1(y, \sum_{\nu\in N^N} w^{N\nu}\bar{x}^\nu)\end{bmatrix}. \qquad (9)$$

We can rewrite the weighted summation for μ as

$$\sum_{\nu \in N^\mu} w^{\mu\nu} \bar{x}^\nu = (w_{-\mu}^T \otimes I_n) \begin{bmatrix} \bar{x}^1 \\ . \\ . \\ . \\ \bar{x}^N \end{bmatrix}, \tag{10}$$

in which $w_{-\mu}$ is a column vector containing $w^{\mu\nu}$ in the ν'th element for all $j \in N^\mu$ and 0 everywhere else (including in the μ'th slot), I_n is the $n \times n$ identity matrix and $\otimes$ is the kronecker product. For example, the form for agent 1, in the case of 2-action game is given by

$$(w_{-1}^T \otimes I_2) = [0, w_{12}, w_{13}, ..., w_{1n}] \otimes I_2 = \begin{bmatrix} 0 & 0 & w_{12} & 0 & ... & w_{1n} & 0 \\ 0 & 0 & 0 & w_{12} & ... & 0 & w_{1n} \end{bmatrix} \tag{11}$$

Returning to the N-player NA game our condition becomes

$$\begin{bmatrix} \bar{x}^1 \\ . \\ . \\ . \\ \bar{N} \end{bmatrix} \in \begin{bmatrix} \arg\max_{y \in \Delta_1} \bar{u}^1(y, (w_{-1}^T \otimes I_n) \begin{bmatrix} \bar{x}^1 \\ . \\ . \\ . \\ \bar{x}^N \end{bmatrix}) \\ . \\ . \\ . \\ \arg\max_{y \in \Delta_N} \bar{u}_N(y, (w_{-N}^T \otimes I_n) \begin{bmatrix} \bar{x}^1 \\ . \\ . \\ . \\ \bar{x}^N \end{bmatrix})) \end{bmatrix}. \tag{12}$$

This means that we achieve a Nash Equilibrium iff $(\bar{x}^\mu)_{\mu \in \mathcal{N}}$ is a fixed point of the map

$$\begin{bmatrix} \arg\max_{y \in \Delta_1} \bar{u}^1(y, (w_{-1}^T \otimes I_n)(\cdot)) \\ . \\ . \\ . \\ \arg\max_{y \in \Delta_N} \bar{u}^N(y, (w_{-N}^T \otimes I_n)(\cdot)) \end{bmatrix}. \tag{13}$$

Now, since the modified payoff $\bar{u}^\mu$ shares the same bilinear property as u^μ, we can assert that it is continuous is its second argument. Therefore, the above vector valued map can be asserted to be continuous and so is upper semi-continuous. As such, we can apply Kakutani's Fixed Point Theorem to assert that the above map admits a fixed point.

□

4.2 Existence of CTFP

Theorem 3. There exists a path m which satisfies the property that, for all μ, $m^\mu \in BR^\mu(\alpha_\sigma^\mu)$ for almost all $t \geq 1$.

Proof. Recall the definition of $\alpha^\mu(t)$

$$\alpha^\mu(t) = \frac{1}{t} \int_0^t m^\mu(s)\, ds$$

Then

$$\begin{aligned} \frac{d}{dt}\alpha^\mu(t) &= \frac{d}{dt}\frac{1}{t}\int_0^t m^\mu(s) ds \\ &= \frac{1}{t}m^\mu(t) - \frac{1}{t}\alpha^\mu(t) \end{aligned} \tag{14}$$

Now we assert that $m^\mu(t) \in BR^\mu(\alpha_\sigma^\mu) = \arg\max_{y\in\Delta_\mu} u^\mu(y, w^{\mu\mu}\alpha^\mu(t) + \sum_{\nu\in N^\mu} w_{ij}\alpha^\nu(t))$.
Let us then define $\alpha(t)$ as the concatenation of all $\alpha_i(t)$

$$\alpha(t) = \left[\alpha^1(t)^T, \dots, \alpha^N(t)^T\right]^T \subset \mathbb{R}^{Nn} \tag{15}$$

We can, therefore, write the aggregation as

$$(W \otimes I_n)\alpha(t) = \begin{bmatrix} \sum_{\nu\in N^1} w^{1\nu}\alpha^\nu(t) \\ \cdot \\ \cdot \\ \cdot \\ \sum_{\nu\in N^N} w^{N\nu}\alpha^\nu(t) \end{bmatrix} \subset \mathbb{R}^{Nn} \tag{16}$$

I will write $(W \otimes I_n)\alpha(t)$ as α_W for convenience. Furthermore, using the bilinear property of the game, we can say that

$$\begin{bmatrix} \arg\max_{y_1\in\Delta_1} y \cdot A^1(\alpha_W(t))_1 \\ \cdot \\ \cdot \\ \cdot \\ \arg\max_{y_N\in\Delta_N} y \cdot A^N(\alpha_W(t))_N \end{bmatrix} = \arg\max_{y\in\Delta} y \cdot \begin{bmatrix} A^1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & A^2 & \cdot & \cdot & \cdot & 0 \\ & & \cdot & & & \\ & & & \cdot & & \\ & & & & \cdot & \\ 0 & 0 & \cdot & \cdot & \cdot & A^N \end{bmatrix} \alpha_W(t) = \arg\max_{y\in\Delta} y \cdot \Lambda(W \otimes I_n)\alpha(t) \tag{17}$$

in which $\Delta = \times_i \Delta_i$ and Λ is the block diagonal matrix containing each A^i.
We can now write the differential inclusion as

$$\dot{\alpha}(t) \in \frac{1}{t}(\arg\max_{y\in\Delta}\{y \cdot \Lambda(W \otimes I_n)\alpha(t)\} - \alpha(t)) \tag{18}$$

with the initial condition $\alpha(1) = x_0$. Then, the path m is a CTFP if its corresponding $\alpha(t; m)$ is a solution to the above differential inclusion. Following the definition of (Harris 1998), a solution $\alpha(t)$ must satisfy

1. $\alpha(t)$ is locally Lipschitz
2. $\alpha(1) = x_0$
3. $\dot{\alpha}(t) \in \frac{1}{t}(\arg\max_{y\in\Delta}\{y \cdot \Lambda(W \otimes I_n)\alpha(t)\} - \alpha(t))$ for almost all $t \in [1, \infty)$

The question then remains, does the differential inclusion admit a solution? Using the results of Aubin and Cellina (see Harris 1998, paragraph below Prop. 6), can use the fact that the $\arg\max$ function is non-empty, compact and convex valued, bounded (all of these follow since the function acts from Δ to Δ) and upper semi-continuous. With these facts in place, we can say that there is a CTFP m for any initial value x_0. Further, the solution $m(t)$ is Lipschitz for almost all $t \geq 0$.

□

With the existence of the CTFP in place, we can show that it converges to a fixed point. In particular, let $\Omega(\alpha)$ be the set of all limit points for $\alpha(t)$. Then, a CTFP path is said to have converged if $\Omega(\alpha)$ is contained within the set of Nash Equilibria of the game. If this is for any such CTFP path, then the game is said to have the *CTFP property*. We adapt the techniques of (Ewerhart 2020) to prove that zero-sum NA games have the CTFP property. AH: *I also want to attempt this with the techniques of Harris (which is quite similar) to see if it yields any new results*.

Theorem 4. Any zero-sum NA game has the property that, for any CTFP path m, the corresponding $\alpha(t; m)$ converges to a set of fixed points.

Proof.

$$m^\mu(t) \in BR^\mu\left(\frac{1}{t}\int_0^t \sigma^\mu(s)ds\right). \tag{19}$$

By the definition of Continuous Time Fictitious Play (CTFP) given by Ewerhart, we require that $m : [0, \infty) \to \times_\mu \Delta_\mu$ is a measurable mapping such that, for each μ, $m^\mu(t) \in BR^\mu\left(\frac{1}{t}\int_0^t \sigma^\mu(t')dt'\right)$. Now notice,

$$m^\mu(t) \in BR^\mu\left(\frac{1}{t}\int_0^t \sigma^\mu(t')dt'\right) \iff \in x^\mu \in BR^\mu\left(\frac{1}{t}\int_0^t [w^{\mu\mu}m^\mu(s) + \sum_{\nu\in N^\mu} w^{\mu\nu}m^\nu(s)]\, ds\right)$$

Let us assume that u^μ takes the form $x \cdot A^\mu \sigma^\mu$ where A^μ is the payoff matrix associated with agent μ. Then,

$$\begin{aligned}
& m^\mu \in \arg\max_{x\in\Delta^\mu} u^\mu(x, \frac{1}{t}\int_0^t [w^{\mu\mu}m^\mu(t') + \sum_{\nu\in N^\mu} w^{\mu\nu}m^\nu(s)]\, ds) \\
\iff & m^\mu \in \arg\max_{x\in\Delta_\mu} x \cdot A^\mu \left(\frac{1}{t}\int_0^t [w^{\mu\mu}m^\mu(t') + \sum_{\nu\in N^\mu} w^{\mu\nu}m^\nu(s)]\, ds\right) \\
\iff & m^\mu \in \arg\max_{x\in\Delta_\mu} x \cdot (w^{\mu\mu}A^\mu)\left(\frac{1}{t}\int_0^t m^\mu(s)ds\right) + \sum_{\nu\in N^\mu} x \cdot (w^{\mu\nu}A^\nu)\left(\frac{1}{t}\int_0^t m^\nu(s)ds\right) \\
\iff & m^\mu \in \arg\max_{x\in\Delta_i} x \cdot A^{\mu\mu}\alpha^\mu(t;m) + \sum_{\nu\in N^\mu} x \cdot A^{\mu\nu}\alpha^\nu(t;m). \qquad (20)
\end{aligned}$$

where each $A^{\mu\nu} = w^{\mu\nu}A^\mu$ and $\alpha^\mu(t;m) = \frac{1}{t}\int_0^t m^\mu(s)ds$ as defined by Ewerhart. We can, therefore, think of the NA game as a network game in which each agent plays the same strategy against each of its neighbours, itself included. As such, any m which satisfies the CTFP property for the equivalent network game also satisfies the CTFP requirement for the network aggregative game.

Continuing with this approach, let us look at the case where $\sum_\mu A^\mu = \mathbf{0}_{n\times n}$ (i.e. a zero sum game). This means that $\sum_i u_i = 0$. Let $x^\mu_\infty = \lim_{t\to\infty}\alpha^\mu(t;m)$. More specifically, x^μ_∞ belongs to the set of accumulation points of $\alpha^\mu(\cdot;m)$ (which is set valued since the BR^μ map is set valued). To say that the game has 'converged' we require that every $x^\mu_\infty = (x^1_\infty, ...x^N_\infty)$ is an NE. AH: *I'll follow through the proof of Ewerhart for the sake of completeness:*

Let us take the Lyapunov function

$$L(x) = \sum_\mu \max_{y\in\Delta_\mu}\{u^\mu(y, w^{\mu\mu}x^\mu + \sum_{\nu\in N^\mu} w^{\mu\nu}x^\nu) - u^\mu(x^\mu, w^{\mu\mu}x^\mu + \sum_{\nu\in N^\mu} w^{\mu\nu}x^\nu)\} \qquad (21)$$

Using the same transformation as before:

$$\begin{aligned}
u^m u(y, w^{\mu\mu}x^\mu + \sum_{\nu\in N^\mu} w_{\mu\nu}x^\nu) &= y \cdot A^\mu(w^{\mu\mu}x^\mu + \sum_{\nu\in N^\mu} w^{\mu\nu}x^\nu) \\
&= y \cdot (w^{\mu\mu}A^\mu)x^\mu + \sum_{\nu\in N^\mu} y \cdot (w^{\mu\nu}A^\mu)x^\nu \\
&= \sum_{\nu\in N^\mu\cup\{\mu\}} y \cdot A^{\mu\nu}x^\nu
\end{aligned}$$

Then,

$$\begin{aligned}
L(x) &= \sum_\mu \max_{y\in\Delta_\mu} \sum_{\nu\in N^\mu\cup\{\mu\}} y \cdot A^{\mu\nu}x^\nu - u^m u(x^\mu, w^{\mu\mu}x^\mu + \sum_{\nu\in N^\mu} w^{\mu\nu}x^\nu)\} \\
&= \sum_\mu \max_{y\in\Delta_\mu}\{\sum_{\nu\in N^\mu\cup\{\mu\}} y \cdot A^{\mu\nu}x^\nu\} - \sum_\mu u^\mu(x^\mu, w^{\mu\mu}x^\mu + \sum_{\nu\in N^\mu} w^{\mu\nu}x^\nu) \\
&= \sum_\mu \max_{y\in\Delta_\mu}\{\sum_{\nu\in N^\mu\cup\{\mu\}} y \cdot A^{\mu\nu}x^\nu\}
\end{aligned}$$

where the last equality holds because $\sum_\mu u^\mu = 0$. Now, in parallel with (Ewerhart 2020) we recall that we defined $m^\mu(t)$ as the best response to $\alpha(t)$. As such,

$$L(\alpha(t)) = \sum_{\mu} \sum_{\nu \in N^{\mu} \cup \{\mu\}} m^{\mu}(t) \cdot A^{\mu\nu} \alpha^{\nu}(t) \tag{22}$$

$$tL(\alpha(t)) = \sum_{\mu} \sum_{\nu \in N^{\mu} \cup \{\mu\}} m^{\mu}(t) \cdot A^{\mu\nu} t\alpha^{\nu}(t) \tag{23}$$

$$tL(\alpha(t)) = \sum_{\mu} \sum_{\nu \in N^{\mu} \cup \{\mu\}} m^{\mu}(t) \cdot A^{\mu\nu} \int_0^t m^{\nu}(s) ds \tag{24}$$

Similarly, since we know that $m^{\mu}(t')$ is a best response to $\alpha(t')$ at time t' we have

$$\sum_{\nu \in N^{\mu} \cup \{\mu\}} m^{\mu}(t') \cdot A^{\mu\nu} \alpha^{\nu}(t') \geq \sum_{\nu \in N^{\mu} \cup \{\mu\}} m^{\mu}(t) \cdot A^{\mu\nu} \alpha^{\nu}(t') \tag{25}$$

$$t' \sum_{\mu} \sum_{\nu \in N^{\mu} \cup \{\mu\}} m^{\mu}(t') \cdot A^{\mu\nu} \alpha^{\nu}(t') \geq \sum_{\mu} \sum_{\nu \in N^{\mu} \cup \{\mu\}} m^{\mu}(t) \cdot A^{\mu\nu} \int_0^{t'} m^{\nu}(s) ds \tag{26}$$

$$t' L(\alpha(t')) \geq \sum_{\mu} \sum_{\nu \in N^{\mu} \cup \{\mu\}} m^{\mu}(t) \cdot A^{\mu\nu} \int_0^{t'} m^{\nu}(s) ds \tag{27}$$

Combining the above equations in $tL(\alpha(t))$ and $t'L(\alpha(t'))$, we get

$$tL(\alpha(t)) - t'L(\alpha(t')) \leq \sum_{\mu} \sum_{\nu \in N^{\mu} \cup \{\mu\}} m^{\mu}(t) \cdot A^{\mu\nu} \int_{t'}^{t} m^{\nu}(s) ds \tag{28}$$

We can divide this expression by $t - t'$ and take the limit for $t' \to t$. This yields the derivative (in particular the upper right Dini derivative)

$$\limsup_{t' \to t, t' < t} \frac{tL(\alpha(t)) - t'L(\alpha(t'))}{t - t'} \leq 0 \tag{29}$$

As this derivative is (weakly) negative, and we know that $tL(\alpha(t))$ is continuous in t, we have the result that $tL(\alpha(t))$ is monotone decreasing. Therefore, it is bounded, i.e. there is some $C \geq 0$ such that $L(\alpha(t)) \leq C/t$. In addition, we know that each term in the summation in the original expression of the Lyapunov function is non-negative and so we can say that

$$\max_{y \in \Delta_{\mu}} \{u^{\mu}(y, w^{\mu\mu} x^{\mu} + \sum_{\nu \in N^{\mu}} w^{\mu\nu} x^{\nu}) - u^{\mu}(x^{\mu}, w^{\mu\mu} x^{\mu} + \sum_{\nu \in N^{\mu}} w^{\mu\nu} x^{\nu})\} \leq \frac{C}{t} \tag{30}$$

$$\implies u^{\mu}(y, w^{\mu\mu} x^{\mu} + \sum_{\nu \in N^{\mu}} w^{\mu\nu} x^{\nu}) - u^{\mu}(x^{\mu}, w^{\mu\mu} x^{\mu} + \sum_{\nu \in N^{\mu}} w^{\mu\nu} x^{\nu}) \leq \frac{C}{t} \; \forall y \in \Delta_n \tag{31}$$

Now let us take some $x^{\mu}_{\infty} \in \Omega(\alpha)$. Since this is a limit point, there exists a sequence $\{t_n\}_{n=0}^{\infty} \to \infty$ such that $\{\alpha^{\mu}(t_n)\}_{n=0}^{\infty} \to x^{\mu}_{\infty}$. So if we take the limit as $t_n \to \infty$ we get

$$u^{\mu}(y, w^{\mu\mu} x^{\mu}_{\infty} + \sum_{\nu \in N^{\mu}} w^{\mu\nu} x^{\nu}_{\infty}) - u^{\mu}(x^{\mu}_{\infty}, w^{\mu\mu} x^{\mu}_{\infty} + \sum_{\nu \in N^{\mu}} w^{\mu\nu} x^{\nu}_{\infty}) \leq 0 \tag{32}$$

This means that $(x^{\mu}_{\infty})_{\mu \in \mathcal{N}}$ is a best response to itself and so is a fixed point of the NA-CTFP dynamic. □

We now point out that, if we choose $w^{\mu\mu}$ to be zero for all μ, then the final inequality yields that, for all μ

$$u^{\mu}(y, \sum_{\nu \in N^{\mu}} w^{\mu\nu} x^{\nu}_{\infty}) \leq u^{\mu}(x^{\mu}_{\infty}, \sum_{\nu \in N^{\mu}} w^{\mu\nu} x^{\nu}_{\infty}) \tag{33}$$

which is precisely the NA-Nash Equilibrium condition. This leads to the next result

Corollary 1. With the additional assumption that, for all agents μ, all zero-sum NA games have the CTFP property

5 Convergence to the CCE set

In this section we aim to find some convergence structure for the case in which the NA game is not zero-sum. In particular, we show that the NA-CTFP process converges to the set of coarse correlated equilibria. So first I will define what this means.

The notion of the CCE set is best considered through *regret* which, for agent μ we define as

$$R^{\mu} = \max_{x^{\mu}_{i'} \in S^{\mu}} \left\{ \frac{1}{T} \int_0^T u^{\mu}(x^{\mu}_{i'}(t), \sigma(t)) - u^{\mu}(m^{\mu}(t), \sigma(t))\, dt \right\} \tag{34}$$

Note, this is the *average regret* for the agent μ and, of course, can be related to the *cumulative regret* which is used for analysis in, e.g. (Leonardos and Piliouras, Cesa-Bianchi). To illustrate the average regret, let us consider the case where each agent has only two actions. Then $u^{\mu}(x^{\mu}(t), \sigma(t))$ is given by

$$u^{\mu}(x^{\mu}(t), \sigma(t)) = \sum_{ij} a_{ij} x^{\mu}_i \sigma^{\mu}_j = a_{11} x^{\mu}_1 \sigma^{\mu}_1 + a_{12} x^{\mu}_1 \sigma^{\mu}_2 + a_{21} x^{\mu}_2 \sigma^{\mu}_1 + a_{22} x^{\mu}_2 \sigma^{\mu}_2 \tag{35}$$

On the other hand, let us consider that agent μ's first strategy maximises $u^{\mu}(x^{\mu}_{i'}(t), \sigma(t))$, then

$$\max_{x^{\mu}_{i'} \in S^{\mu}} u^{\mu}(x^{\mu}_{i'}(t), \sigma(t)) = \sum_{ij} a_{1j} x^{\mu}_i \sigma^{\mu}_j = a_{11} x^{\mu}_1 \sigma^{\mu}_1 + a_{12} x^{\mu}_1 \sigma^{\mu}_2 + a_{11} x^{\mu}_2 \sigma^{\mu}_1 + a_{12} x^{\mu}_2 \sigma^{\mu}_2 \tag{36}$$

By comparing the two expanded expressions, we can see that the latter gives the reward that agent μ would have received had they played action 1 throughout the entire play, assuming that the behaviour of the other agents (encoded in σ) does not change. As such, this is a measure of agent μ's regret, in hindsight, for not playing action 1 the entire time. An agent achives *no regret* if R^{μ} is non-positive.

Theorem 5. Assuming that $w^{\mu\mu} = 0$, then for any choice of payoff matrix, agents following the NA-FP process achieve zero regret in the limit, i.e.

$$\lim_{T \to \infty} \max_{x^{\mu}_{i'} \in S^{\mu}} \left\{ \frac{1}{T} \int_0^T u^{\mu}(x^{\mu}_{i'}(t), \sigma(t)) - u^{\mu}(m^{\mu}(t), \sigma(t))\, dt \right\} = 0 \tag{37}$$

Proof. We will adapt the techniques of Ostrovski and van Strein for this proof. Their result was found for a two player game, we will leverage the fact that the NA process means that we can reduce an N player game down to a series of two player games.

For any agent μ define

$$\bar{u}^{\mu}(\alpha^{\mu}_{\sigma}) := max_{y \in \Delta_{\mu}} u^{\mu}(y, \alpha^{\mu}_{\sigma}) \tag{38}$$

Now, if the agents are following the NA-FP process, then we know that $m^{\mu}(t) \in \Delta_{\mu}$ is the strategy which maximises $u^{\mu}(\cdot, \alpha^{\mu}_{\sigma})$ so we can write

$$\bar{u}^{\mu}(\alpha^{\mu}_{\sigma}) := max_{y \in \Delta_{\mu}} u^{\mu}(y, \alpha^{\mu}_{\sigma}) = m^{\mu}(t) \cdot A^{\mu}(\alpha^{\mu}_{\sigma}) \tag{39}$$

Then, applying the *envelope theorem* we have

$$\frac{d}{d\alpha^{\mu}_{\sigma}} \bar{u}^{\mu}(\alpha^{\mu}_{\sigma}) = \frac{d}{d\alpha^{\mu}_{\sigma}} x^{\mu} \cdot A^{\mu}(\alpha^{\mu}_{\sigma}) \Big|_{x^{\mu} = m^{\mu}} = m^{\mu} \cdot A^{\mu} \tag{40}$$

Which gives that

$$\frac{d}{dt} \bar{u}^{\mu}(\alpha^{\mu}_{\sigma}(t)) = m^{\mu} \cdot A^{\mu} \frac{d\alpha^{\mu}_{\sigma}(t)}{dt} \tag{41}$$

Then,

$$\frac{d}{dt}(t \bar{u}^{\mu}(\alpha^{\mu}_{\sigma}(t))) = \bar{u}^{\mu}(\alpha^{\mu}_{\sigma}(t)) + t m^{\mu} \cdot A^{\mu} \frac{d\alpha^{\mu}_{\sigma}(t)}{dt} \tag{42}$$

$$= m^{\mu} \cdot A^{\mu} (\sum_{\nu \in N(\mu)} w^{\mu\nu} \alpha^{\nu}(t) + t(m^{\mu} \cdot A^{\mu} \sum_{\nu \in N(\mu)} w^{\mu\nu} \frac{d}{dt} \alpha^{\nu}(t)) \tag{43}$$

Now, we know that the differential inclusion for the NA-FP process can be written as

$$\frac{d}{dt}\alpha^\mu(t) = \frac{1}{t}(m^\mu(t) - \alpha^\mu(t)) \tag{44}$$

We insert this into the previous equation to yield

$$\frac{d}{dt}(t\bar{u}^\mu(\alpha^\mu_\sigma(t))) = m^\mu \cdot A^\mu(\sum_{\nu\in N(\mu)} w^{\mu\nu}\alpha^\nu(t)) + t(m^\mu \cdot A^\mu \sum_{\nu\in N(\mu)} w^{\mu\nu}\frac{1}{t}(m^\nu(t) - \alpha^\nu(t))) \tag{45}$$

$$= m^\mu \cdot A^\mu(\sum_{\nu\in N(\mu)} w^{\mu\nu}\alpha^\nu(t) + m^\mu \cdot A^\mu \sum_{\nu\in N(\mu)} w^{\mu\nu}(m^\nu(t) - \alpha^\nu(t)) \tag{46}$$

$$= m^\mu \cdot A^\mu \sum_{\nu\in N(\mu)} w^{\mu\nu} m^\nu(t) = m^\mu \cdot A^\mu \sigma^\mu(t) \tag{47}$$

Now if we integrate both sides with respect to t in the bound $[1, T]$ (on which the NA-FP) process is defined, we get

$$T\bar{u}^\mu(\alpha^\mu_\sigma(T)) - \bar{u}^\mu(\alpha^\mu_\sigma(1)) = \int_1^T u^\mu(m^\mu(t), \sigma^\mu(t))\, dt \tag{48}$$

$$\implies \bar{u}^\mu(\alpha^\mu_\sigma(T)) - \frac{1}{T}\bar{u}^\mu(\alpha^\mu_\sigma(1)) = \frac{1}{T}\int_1^T u^\mu(m^\mu(t), \sigma^\mu(t))\, dt \tag{49}$$

$$\implies \lim_{T\to\infty} \bar{u}^\mu(\alpha^\mu_\sigma(T)) - \frac{1}{T}\int_0^T u^\mu(m^\mu(t), \sigma^\mu(t))\, dt = 0 \tag{50}$$

Now notice

$$\bar{u}^\mu(\alpha^\mu_\sigma(T)) = \max_{x^\mu_{i'}\in S^\mu} \sum_j a_{i'j}(\alpha^\mu_\sigma(T))_j = \max_{x^\mu_{i'}\in S^\mu} \sum_j a_{i'j}(\frac{1}{T}\int_0^T \sigma^\mu(t)\, dt)_j \tag{51}$$

$$= \max_{x^\mu_{i'}\in S^\mu} \sum_j a_{i'j}\frac{1}{T}\int_0^T \sum_i (m^\mu(t))_i(\sigma^\mu(t))_j\, dt \tag{52}$$

$$= \max_{x^\mu_{i'}\in S^\mu} \frac{1}{T}\int_0^T \sum_{ij} a_{i'j}(m^\mu(t))_i(\sigma^\mu(t))_j\, dt \tag{53}$$

$$= \max_{x^\mu_{i'}\in S^\mu} \frac{1}{T}\int_0^T u^\mu(x^\mu_{i'}, \sigma^\mu(t)) \tag{54}$$

So we recover

$$\lim_{T\to\infty} \max_{x^\mu_{i'}\in S^\mu} \left\{\frac{1}{T}\int_0^T u^\mu(x^\mu_{i'}(t), \sigma(t)) - u^\mu(m^\mu(t), \sigma(t))\, dt\right\} = 0 \tag{55}$$

which was precisely the regret condition that we required.

AH: *Note: I'm just figuring out how best to formulate the notion of the NA-CCE set and then I can add it in as a definition to say that NA-CTFP converges to this set.* □

6 Numerical Experiments

7 Convergence rates w.r.t connectivity and size

8 Non-convergent examples

9 Agent based simulations

10 Addition of Noise

11 Discussion

12 Concluding remarks