On the Convergence of Fictitious Play: A Decomposition Approach

Yurong Chen[1] , **Xiaotie Deng**[1*] , **Chenchen Li**[2] , **David Mguni**[3] ,
Jun Wang[4] , **Xiang Yan**[5] and **Yaodong Yang**[6]
[1]Center on Frontiers of Computing Studies, School of Computer Science, Peking University
[2]Business Growth BU, JD.com
[3]Huawei R&D
[4]UCL
[5]Huawei TCS Lab
[6]Institute for AI, Peking University
chenyurong@pku.edu.cn, xiaotie@pku.edu.cn, lcc104@qq.com, davidmguni@hotmail.com,
jun.wang@cs.ucl.ac.uk, xyansjtu@163.com, yaodong.yang@pku.edu.cn

Abstract

Fictitious play (FP) is one of the most fundamental game-theoretical learning frameworks for computing Nash equilibrium in n-player games, which builds the foundation for modern multi-agent learning algorithms. Although FP has provable convergence guarantees on zero-sum games and potential games, many real-world problems are often a mixture of both and the convergence property of FP has not been fully studied yet. In this paper, we extend the convergence results of FP to the combinations of such games and beyond. Specifically, we derive new conditions for FP to converge by leveraging game decomposition techniques. We further develop a linear relationship unifying cooperation and competition in the sense that these two classes of games are mutually transferable. Finally, we analyze a non-convergent example of FP, the Shapley game, and develop sufficient conditions for FP to converge.

1 Introduction

Solving Nash equilibrium (NE) [Nash, 1950; Deng *et al.*, 2021] in multi-player games has become a central interest in a variety of fields including but not limited to economics, computer science and artificial intelligence. Among the many NE solvers, fictitious play (FP) [Brown, 1951] is one of the most well-known learning algorithms. In FP, at each iteration, each player takes a best response to the empirical average of the opponent's previous strategies. It is guaranteed that FP dynamics converge to an NE on two-player zero-sum games [Robinson, 1951] and potential games [Monderer and Shapley, 1996a; Monderer and Shapley, 1996b; Mguni *et al.*,], with no need to access the other player's utility information. Thus, the design principle of FP (i.e., the iterative best-response dynamics) has inspired many other approximation solutions to NE. For example, in solving two-player zero-sum games, two representative methods are double oracle (DO) [McMahan *et al.*, 2003; Dinh *et al.*, 2021] and policy space response oracle (PSRO) [Lanctot *et al.*, 2017; Feng *et al.*, 2021; Perez-Nieves *et al.*, 2021; Liu *et al.*, 2021] where a subgame NE is adopted as the best-responding target and multi-agent reinforcement learning (MARL) algorithms [Yang and Wang, 2020] are applied to approximate the best response. Similarly, Heinrich *et al.* [2015] combined fictitious self-play with deep RL methods and demonstrated remarkable performance on Leduc Poker and Limit Texas Holdem at real-world scale. Besides its inspirations for modern MARL algorithms [Muller *et al.*, ; Yang *et al.*, 2018], FP itself has been shown to have good performance of converging to approximate equilibria on some more general games [Candogan *et al.*, 2013; Ostrovski and van Strien, 2014], and is still a popular interest of research to the communities of both game theory and machine learning [Swenson *et al.*, 2018].

However, the assumptions of zero-sum or potential games are rather limited. As Dasgupta and Collins [2019] mentioned, adversarial learning has been modeled as a two-player zero-sum game, but chances are that the learner's loss may not equal the adversary's utility. Similar as zero-sum games characterizing full competition, potential games can be regarded as full cooperation. While in most real applications, there are both competition and cooperation among players. For example, in a market, sellers for the same type of goods are not only players competing for the same group of buyers, but also collaborators attracting more buyers to the market. More general results for FP to converge on combinations of competition and cooperation are needed, considering its current limited positive convergence results.

Our techniques and results. We leverage two game decomposition techniques, Hodge Decomposition [Candogan *et al.*, 2011] and Strategic Decomposition [Hwang and Rey-Bellet, 2020], to study the convergence of FP on mixtures of games modelling full competition and cooperation. The idea of game decomposition is to treat the set of games as a linear space and decompose a game into several simple basis games, whose NEs are easy to characterize. Interestingly, in both decompositions, a game is made up of a competitive part, a co-

*Corresponding Author

operative part and a trivial part. It is known that FP converges on all these basis games. Combining two decompositions enables us to find latent relationships among games and study how a game's dynamics is influenced by each component. We note that, however, there will be no convergence guarantee for FP on arbitrary combinations of these basis games, since they span the whole game space, and FP has been proved to fail to converge on all games [Shapley, 1964].

The contributions of this paper are as follows,

- We prove that FP converges on any linear combination of a harmonic game (competition) and a potential game (cooperation), so long as they sum to be strategically equivalent to a zero-sum game or to an identical interest game. The conditions are polynomial-time checkable.
- We show that, utilizing a linear parameter, games lying in these two equivalent classes can be transformed from one class into another.
- We give a new analysis of the non-convergence of Continuous-time FP (CFP) on the classic example, the Shapley game, from the view of dynamical system and game decomposition, and provide a sufficient condition about initial conditions for it to converge on linear combinations of zero-sum games and identical interest games.

Since many machine learning methods are built on FP, the fact that FP converges in larger game classes provides a guarantee for them to be applied to more real situations safely. We argue that the classes of games considered in our paper are of wide interest and include many important games like load balancing games [Vöcking, 2007], cost and market sharing games with various distribution rules [Gopalakrishnan *et al.*, 2011], and strictly competitive games [Adler *et al.*, 2009].

Related work. Our analysis is mainly based on the techniques of game decomposition. To make the paper self-contained, we discussed related work on game decompositions here, and put the full related work in the Appendix.

Candogan *et al.* [2011] and Hwang and Rey-Bellet [2020] studied decompositions for finite games. Helmholtz Decomposition [Balduzzi *et al.*, 2018], on the other hand, studied the decomposition of continuous games into potential games and Hamiltonian games. But few works are about making use of decomposition to analyze games. To the best of our knowledge, utilizing game decomposition techniques to analyze game dynamics is still a new area. Tuyls *et al.* [2018] decomposed an asymmetric bimatrix game, where two players have the same number of pure actions, into two one population games, from the evolutionary game theory point of view. Cheung and Tao [2021] utilized the canonical decomposition of decomposing a game into a zero-sum game and an identical interest game to study the chaotic behaviors of general-sum games under multiplicative weight dynamics. We study the convergence properties of fictitious play. Instead of checking which component dominates the other, we model the mixture of two basis games by linear combinations and see how the game patterns evolve when the parameter changes smoothly and continuously, which can also be seen as a linear homotopy from the view of homotopy method.

Paper organization. Section 2 introduces necessary preliminaries. Section 3 shows our main convergence results. The illustration of transformations between cooperation and competition is shown in Example 1 and Section 3.2. Section 4 analyzes the Shapley game and give a condition for CFP to converge. Section 5 summarizes and give future work.

2 Background

We use bold lowercase characters to denote vectors. $\mathbf{1}_n$ and $\mathbf{0}_n$ denote n-dimensional all-one and all-zero vectors, respectively, $\mathbf{e}_i$ the vector the i^{th} coefficient of which is 1 and all other coefficients are 0. $[n]$ denotes the set $\{1, \dots, n\}$.

2.1 Games and Nash Equilibrium

We focus on two-player bimatrix games. We use bold uppercase letters to denote games and uppercase letters to denote matrices: A game $\mathbf{G}$ is given in the bimatrix form (A, B), where the first (second) matrix in the pair denotes the payoff matrix of player 1 (player 2, respectively). Both matrices have dimension $m \times n$, i.e., player 1 (player 2) has m (n, respectively) actions.

We call an action i a pure strategy, and a distribution over all actions a mixed strategy. Denote the set of all mixed strategies as Δ_m, where $\Delta_m := \{\mathbf{p} \in \mathbb{R}^m : p_i \geq 0, \forall i \in \{1, \dots, m\}, \sum_{i=1}^m p_i = 1\}$. Given a game $\mathbf{G} = (A, B)$ and two mixed strategies $\mathbf{p} \in \Delta_m, \mathbf{q} \in \Delta_n$, player 1 (player 2)'s utility is $\mathbf{p}^\top A\mathbf{q}$ ($\mathbf{p}^\top B\mathbf{q}$, respectively). We use $\mathrm{BR}_i(\cdot, \mathbf{G})$ to denote player i's best response set:

$$\mathrm{BR}_1(\mathbf{q}, \mathbf{G}) = \arg\max_{i \in [m]} (A\mathbf{q})_i, \mathrm{BR}_2(\mathbf{p}, \mathbf{G}) = \arg\max_{j \in [n]} (\mathbf{p}^\top B)_j.$$

We omit the last variable of BR_i when there is no confusion.

A Nash equilibrium (NE) is a pair of mixed strategies such that no one wants to deviate with the other's strategy fixed:

Definition 1. *Strategy pair $(\mathbf{p}^*, \mathbf{q}^*)$ is a Nash Equilibrium (NE) of game $\mathbf{G} = (A, B)$ if for any $\mathbf{p} \in \Delta_m$ and $\mathbf{q} \in \Delta_n$,*

$$\mathbf{p}^{*\top} A\mathbf{q}^* \geq \mathbf{p}^\top A\mathbf{q}^*, \quad \mathbf{p}^{*\top} B\mathbf{q}^* \geq \mathbf{p}^{*\top} B\mathbf{q}$$

We call an NE pure when strategies in the NE are all pure strategies, and mixed otherwise.

Let $\mathcal{G}$ be the set of all bimatrix games. Given games $\mathbf{G}_1 = (A_1, B_1)$, $\mathbf{G}_2 = (A_2, B_2) \in \mathcal{G}$, define their addition to be $\mathbf{G}_1 + \mathbf{G}_2 := (A_1 + A_2, B_1 + B_2)$. Given a scalar $\alpha \in \mathbb{R}$ and a game $\mathbf{G} = (A, B)$, define the scalar multiplication to be $\alpha\mathbf{G} = (\alpha A, \alpha B)$. Now $\mathcal{G}$ is a linear space and we can consider the combinations and decompositions of games.

2.2 Basic Games and Relations among Games

We introduce subspaces of $\mathcal{G}$ that are basic in this paper.

Definition 2. *Define the following subspaces of $\mathcal{G}$*
*1. **Identical interest** games, $\mathcal{I} := \{(A, B) \in \mathcal{G} : A = B\}$*
*2. **Zero-sum** games, $\mathcal{Z} := \{(A, B) \in \mathcal{G} : A + B = 0\}$*
*3. **Non-strategic** games, $\mathcal{E} := \left\{(A, B) \in \mathcal{G} : A = \mathbf{1}_m \mathbf{u}^\top, B = \mathbf{v}\mathbf{1}_n^\top, \mathbf{u} \in \mathbb{R}^n, \mathbf{v} \in \mathbb{R}^m\right\}$*
*4. **Normalized** games, $\mathcal{N} := \{(A, B) \in \mathcal{G} : \sum_{j=1}^m A_{ji} = 0, \forall i \in [n]; \sum_{j=1}^n B_{ij} = 0, \forall i \in [m]\}$*

Identical interest games and zero-sum games are important games in game theory and machine learning, especially in multi-agent learning and adversarial learning, with the former modeling team cooperation and the latter modeling competition. In a non-strategic game, a player's utility only depends on the other player's strategy, and is not affected by her own strategy at all. Thus any strategy pair of the game is an NE.

Noticing that adding a non-strategic game to a game does not change its original game structure, e.g., the best response structure and NEs, one can define an equivalence relation between two games if they only differ by a non-strategic game (also called strategic equivalence by Hwang and Rey-Bellet [2020]), and mainly focus on the normalized games. In a normalized game, the sum of one player's utilities, with her own strategy changing and the other's fixed, equals to zero.

Different from Hwang and Rey-Bellet [2020], we call two games $\mathbf{G}$ and $\mathbf{G}'$ are **additionally equivalent** if $\mathbf{G}' = \mathbf{G} + \mathbf{E}$ for some $\mathbf{E} \in \mathcal{E}$. Here we introduce a more general equivalence among games.

Definition 3. *Game $\mathbf{G} = (A, B)$ is **strategically equivalent** to game $\mathbf{G}' = (A', B')$, if there exist two positive constants $\alpha,\ \beta \in \mathbb{R}_+$ and a non-strategic game $\mathbf{E} \in \mathcal{E}$ such that*

$$(A', B') = (\alpha A, \beta B) + \mathbf{E},\ \textit{for some } \mathbf{E} \in \mathcal{E}$$

While additional equivalence is a special case of strategic equivalence, the notion of additional equivalence is compatible with the operations of the space. Lemma 1 shows that both equivalences preserve the best response structure.

Lemma 1. *Given two strategically equivalent games $\mathbf{G}$ and $\mathbf{G}'$, we have*

$$\mathrm{BR}_1(\mathbf{q}, \mathbf{G}) = \mathrm{BR}_1(\mathbf{q}, \mathbf{G}'), \quad \mathrm{BR}_2(\mathbf{p}, \mathbf{G}) = \mathrm{BR}_2(\mathbf{p}, \mathbf{G}')$$

We use $\mathcal{S}(\cdot)$ ($\mathcal{A}(\cdot)$) to denote the set of games that are strategically (additionally) equivalent to the games in $\cdot$. In particular, we call games in $\mathcal{S}(\mathcal{Z})$ ($\mathcal{S}(\mathcal{I})$) zero-sum equivalent games (identical interest equivalent games, respectively).

2.3 Discrete-time Fictitious Play (DFP)

In fictitious play, each player regards the empirical distribution over the other player's actions as her belief towards the other player's mixed strategy, and acts myopically to maximize her utility in the next step. Specifically, let $\mathbf{p}(t) \in \Delta_m$ and $\mathbf{q}(t) \in \Delta_n$ be the beliefs of two players' strategies at time step t, then the sequence $(\mathbf{p}(t), \mathbf{q}(t))$ is a discrete-time fictitious play (DFP) if:

$$(\mathbf{p}(0), \mathbf{q}(0)) \in \Delta_m \times \Delta_n$$

and for all t:

$$\begin{aligned} \mathbf{p}(t+1) &\in \frac{t}{t+1}\mathbf{p}(t) + \frac{1}{t+1}\,\mathrm{BR}_1(\mathbf{q}(t)), \\ \mathbf{q}(t+1) &\in \frac{t}{t+1}\mathbf{q}(t) + \frac{1}{t+1}\,\mathrm{BR}_2(\mathbf{p}(t)) \end{aligned} \tag{1}$$

With a specific tie-breaking rule, we can regard $\mathrm{BR}_i(\cdot)$ as a vector and the inclusion above becomes equality.

In the following sections, DFP is simply called FP. With abuse of notation, the term FP refers to the dynamic rules in Eqn. (1) or the sequences generated by the rules, according to the context. We say a game has fictitious play property (FPP) if every FP sequence of it converges. Noticing that games with the same best response structure enjoy the same FPP.

3 Convergence on the Combinations of Cooperation and Competition

In this section, we formally present our first two main results. In Section 3.1, we introduce two game decompositions, and present our results on the linear combinations of basis games. In Section 3.2, we illustrate how the players' relationships transform with respect to the linear parameter.

3.1 Proof of Convergence by Game Decompositions

Before formally stating our first result, we present two important game decompositions:

Theorem 1 (Strategic Decomposition)**.** *[Hwang and Rey-Bellet, 2020] The space of games $\mathcal{G}$ can be decomposed as:*

$$\mathcal{G} = (\mathcal{I} \cap \mathcal{N}) \oplus (\mathcal{Z} \cap \mathcal{N}) \oplus \mathcal{B}.$$

where $\mathcal{B} := (\mathcal{I} + \mathcal{E}) \cap (\mathcal{Z} + \mathcal{E})$ is the set of zero-sum equivalent potential games, $\oplus$ denotes the direct sum of two linear subspaces.

$\mathcal{I} + \mathcal{E}$ is the space of games additionally equivalent to identical interest games, and is actually the space of all potential games. The equivalence between this definition and the one using potential function is shown in Appendix. As shown by Hwang and Rey-Bellet [2020], a two-player zero-sum equivalent potential game $\mathbf{B} \in \mathcal{B}$ has the form

$$\mathbf{B} = (\mathbf{u}\mathbf{1}_n^\top, \mathbf{1}_m\mathbf{v}^\top) + \mathbf{E} = \left(\mathbf{u}\mathbf{1}_n^\top + \mathbf{1}_m\mathbf{x}^\top, \mathbf{1}_m\mathbf{v}^\top + \mathbf{y}\mathbf{1}_n^\top\right) \tag{2}$$

for some $\mathbf{u}, \mathbf{y} \in \mathbb{R}^m$, $\mathbf{v}, \mathbf{x} \in \mathbb{R}^n$ and $\mathbf{E} = (\mathbf{1}_m\mathbf{x}^\top, \mathbf{y}\mathbf{1}_n^\top) \in \mathcal{E}$. That is, a player's utility is not affected by the other's strategy, which can be seen as the opposite of non-strategic games. Each player has a dominant strategy, and a pure NE exists.

Theorem 2 (Hodge Decomposition)**.** *[Candogan* et al.*, 2011] The space of games $\mathcal{G}$ can be decomposed as:*

$$\mathcal{G} = \mathcal{P} \oplus \mathcal{H} \oplus \mathcal{E}.$$

where $\mathcal{P} := \mathcal{N} \cap (\mathcal{I} + \mathcal{E})$ denotes normalized potential games, and $\mathcal{H} := \{(A, B) \in \mathcal{N} : mA + nB = 0\}$ normalized harmonic games. $\mathcal{P} + \mathcal{E}$ is the set of all potential games. $\mathcal{H} + \mathcal{E}$ is the set of all harmonic games.

The definition of normalized harmonic games tells that they are like zero-sum games. Thus both decompositions show that any bimatrix game is made up of a fully cooperative component, a fully competitive component and a component that either has both features or is trivial. Combining the decompositions and equivalences makes us able to study bimatrix games from multiple angles.

Since the basis games in the decompositions generate the whole game space, FP will not converge on all combinations of them (a famous example of non-convergence is the Shapley game [Shapley, 1964]). Notice that on all basis games, however, FP will converge: results on zero-sum games and potential games are known [Monderer and Shapley, 1996a; Robinson, 1951]. When the tie-breaking rule is decided, best responses in FPs are always the same on games in $\mathcal{B}$ and $\mathcal{E}$ — a dominant strategy in $\mathbf{B} \in \mathcal{B}$ and the prescribed strategy by the tie-breaking rule in $\mathbf{E} \in \mathcal{E}$ — thus FP will converge.

$$\begin{bmatrix} -14,21 & -20,30 & -14,21 \\ 18,-27 & 14,-21 & 2,-3 \\ -18,27 & 0,0 & -16,24 \end{bmatrix} = \left(2\begin{bmatrix} -7 & -10 & -7 \\ 9 & 7 & 1 \\ -9 & 0 & -8 \end{bmatrix}, 3\begin{bmatrix} 7 & 10 & 7 \\ -9 & -7 & -1 \\ 9 & 0 & 8 \end{bmatrix} \right) + \begin{bmatrix} 0,0 & 0,0 & 0,0 \\ 0,0 & 0,0 & 0,0 \\ 0,0 & 0,0 & 0,0 \end{bmatrix}$$

Figure 1: Game $\mathbf{G} \in \mathcal{S}(\mathcal{Z})$: $\mathbf{G} = (2Z, -3Z)$, where $Z \in \mathbb{R}^{m\times n}$.

In Appendix, we provide a simple proof to show that FP also converges on harmonic games.

It is then interesting to study under what conditions do combinations of these games preserve FPP. By first conducting experiments on mixtures of normalized harmonic games and normalized potential games, we find out that if they are components of a zero-sum game, then FP converges on any linear combination of them. The following theorem gives the formal explanation for this phenomenon and provide a more general condition for FP to converge: if their sum is either fully competitive or cooperative, then any linear combination of these games has FPP. Recall that the set $\mathcal{S}(\cdot)$ is the set of games strategically equivalent to games in $\cdot$, then we have:

Theorem 3. *For any game $\mathbf{G} \in \mathcal{S}(\mathcal{Z}) \cup \mathcal{S}(\mathcal{I})$ with Hodge Decomposition,*

$$\mathbf{G} = \mathbf{P} + \mathbf{H} + \mathbf{E}$$

where $\mathbf{P} \in \mathcal{P}$, $\mathbf{H} \in \mathcal{H}$, $\mathbf{E} \in \mathcal{E}$. Then for any $\lambda \in \mathbb{R}$, game $\lambda\mathbf{P} + (1-\lambda)\mathbf{H}$ has FPP.

We note that the decomposition of a game that is either strategically equivalent to a zero-sum game or to an identical interest game is non-trivial: It can have all game components, since a game lying in these two classes does not necessarily belong to any basis game class. One can refer to the decomposition of the game in Example 1 in Appendix C.

To prove Theorem 3, we first give a necessary lemma. Lemma 2 states the properties of zero-sum equivalent potential games; it shows that sets $\mathcal{S}(\mathcal{I})$ and $\mathcal{S}(\mathcal{Z})$ are closed under the operation of adding a zero-sum equivalent potential game.

Lemma 2. *If $\mathbf{G} \in \mathcal{S}(\mathcal{Z}) \cup \mathcal{S}(\mathcal{I})$, then $\mathbf{G} + \mathbf{B} \in \mathcal{S}(\mathcal{Z}) \cup \mathcal{S}(\mathcal{I})$ for any $\mathbf{B} \in \mathcal{B}$.*

With Lemma 2, we can now prove Theorem 3. The convergence of games $\lambda\mathbf{P} + (1-\lambda)\mathbf{H}$ are actually the trade-offs between cooperation ($\mathbf{P}$) and competition ($\mathbf{H}$). Though the decompositions of games considered are non-trivial, by further decomposing the components using the other decomposition, we find interesting relations among components of different decompositions, which is a key technique in our proof.

Proof sketch of Theorem 3. We show the proof sketch of the case when $\mathbf{G} \in \mathcal{S}(\mathcal{Z})$ here and the full proof is in Appendix. Let $\mathbf{P} = \mathbf{I} + \mathbf{E} = (I, I) + \mathbf{E}$, where $\mathbf{I} \in \mathcal{I}$, $\mathbf{E} \in \mathcal{E}$, $I \in \mathbb{R}^{m\times n}$. $\mathbf{H}$ can be formulated as $(nZ, -mZ)$ for some $Z \in \mathbb{R}^{m\times n}$.

When $\mathbf{G} = (Z', -\alpha Z') + \mathbf{E}' \in \mathcal{S}(\mathcal{Z})$, where $Z' \in \mathbb{R}^{m\times n}$, $\alpha > 0$, $\mathbf{E}' \in \mathcal{E}$, then we have

$$(I, I) + (nZ, -mZ) = (Z', -\alpha Z') + \mathbf{E}' - \mathbf{E}$$

By letting $\mathbf{E}' - \mathbf{E} = (E_1, E_2)$, where $E_1, E_2 \in \mathbb{R}^{m\times n}$ satisfies $E_1 = \mathbf{1}_m \mathbf{u}^\top$, $E_2 = \mathbf{v}\mathbf{1}_n^\top$, for some $\mathbf{u} \in \mathbb{R}^n$, $\mathbf{v} \in \mathbb{R}^m$, we have I and Z represented as linear combinations of Z', E_1 and E_2. Thus for any $\lambda \in \mathbb{R}$,

$$\lambda\mathbf{P} + (1-\lambda)\mathbf{H} = ((a_1(\lambda)Z', b_1(\lambda)Z') \\ + (a_2(\lambda)E_1 + a_3(\lambda)E_2, b_2(\lambda)E_1 + b_3(\lambda)E_2) + \lambda\mathbf{E}$$

By the definition of non-strategic games and Eqn. (2), $(a_2(\lambda)E_1 + a_3(\lambda)E_2, b_2(\lambda)E_1 + b_3(\lambda)E_2) \in \mathcal{B}$ is a zero-sum equivalent potential game. When $a_1(\lambda)b_1(\lambda) \neq 0$, game $(a_1(\lambda)Z', b_1(\lambda)Z') \in \mathcal{S}(\mathcal{Z}) \cup \mathcal{S}(\mathcal{I})$. By Lemma 2, $\lambda\mathbf{P} + (1-\lambda)\mathbf{H} \in \mathcal{S}(\mathcal{I}) \cup \mathcal{S}(\mathcal{Z})$ and thus has FPP.

When $a_1(\lambda)b_1(\lambda) = 0$, $\lambda\mathbf{P} + (1-\lambda)\mathbf{H}$ has one payoff matrix in the form of $\mathbf{x}\mathbf{1}_n^\top + \mathbf{1}_m\mathbf{y}^\top$ for some $\mathbf{x} \in \mathbb{R}^m$, $\mathbf{y} \in \mathbb{R}^n$. The player with this kind of payoff matrix has a dominant strategy. During each time step of FP, the player will choose her dominant strategy, and the other best responds to that dominant strategy. The sequence will converge to a pure NE. □

Following Theorem 3, define $\mathcal{D}$ to be a new set of games, in which one of the payoff matrices has the form of $\mathbf{x}\mathbf{1}_n^\top + \mathbf{1}_m\mathbf{y}^\top$ for some $\mathbf{x} \in \mathbb{R}^m$, $\mathbf{y} \in \mathbb{R}^n$, then:

Corollary 1. *For any game $\mathbf{G} \in \mathcal{D}$ with Hodge Decomposition*

$$\mathbf{G} = \mathbf{P} + \mathbf{H} + \mathbf{E}$$

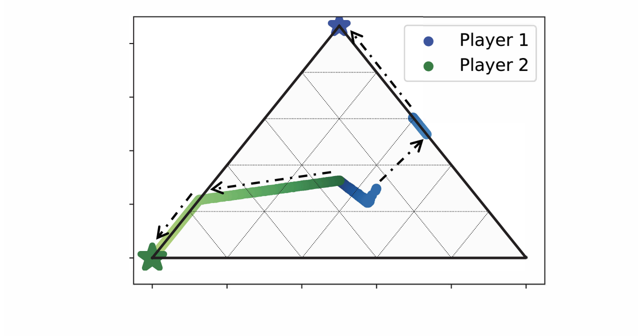
where $\mathbf{P} \in \mathcal{P}$, $\mathbf{H} \in \mathcal{H}$, $\mathbf{E} \in \mathcal{E}$. Then for any $\lambda \in \mathbb{R}$, game $\lambda\mathbf{P} + (1-\lambda)\mathbf{H}$ has FPP.

Conditions in Theorem 3 and Corollary 1 can be checked in linear time, e.g. through the method by Heyman [2019].

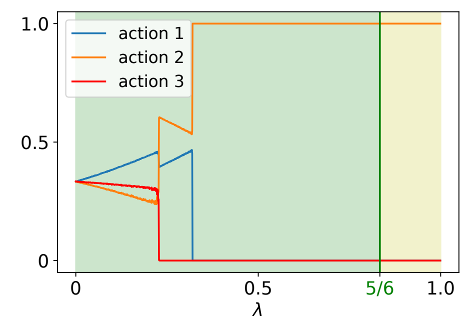
3.2 Transformations from Cooperation to Competition

We use one example to show the non-trivial decompositions of games considered in Section 3.1 and how the proportions of the cooperative component and the competitive component influence the game, when $\lambda \in [0, 1]$ ranges from 0 to 1.

Example 1. *A game in $\mathcal{S}(\mathcal{Z})$ is shown in Figure 1. We decompose this game using Hodge Decomposition into three components $\mathbf{P} \in \mathcal{P}$, $\mathbf{H} \in \mathcal{H}$ and $\mathbf{E} \in \mathcal{E}$, the details of which are in Appendix C, and compute the game $\mathbf{G}(\lambda) = \lambda\mathbf{P} + (1-\lambda)\mathbf{H}$. The changes of NE which FP converges to with λ changing from 0 to 1 is shown in Figure 2. In Figure 2a, the strategy trajectories are presented in the mixed strategy simplex: each vertex of the triangle is an action $\mathbf{v}_i$. We draw mixed strategy $\mathbf{p} = (p_1, p_2, p_3) \in \Delta_3$ on the convex combination of the vertices, $\sum_{i=1}^{3} p_i \mathbf{v}_i$. The green dots represent the strategy trajectory of player 1, and the blue dots that of player 2. The star sign denotes the strategies when $\lambda = 1$. We set the step length to be 0.001. For each λ, we run FP starting from $(\mathbf{v}_1, \mathbf{v}_1)$ for 500,000 rounds. The*



(a) Equilibrium changes in one strategy simplex.



(b) Equilibrium strategy changes of player 1.

Figure 2: How Nash equilibrium of game $\lambda\mathbf{P} + (1-\lambda)\mathbf{H}$ changes as the linear parameter λ increases from 0 to 1. (a) The changes of equilibrium strategies of both players are shown in one strategy simplex. The black arrows tell the direction of the changes. The stars denote the stopping point. When λ is small, the harmonic part dominates. The game shows more competitive patterns, as the equilibrium strategies are mixed in the interior of the strategy simplex. As λ gets larger, the support of equilibrium strategies shrinks. (b) The changes of player 1's strategy in a line chart. In the green area (when $\lambda < \frac{5}{6}$), the game is zero-sum equivalent. In the yellow area (when $\lambda > \frac{5}{6}$), the game is identical interest equivalent. When λ exceeds $\frac{5}{6}$, the game is more cooperative, and the equilibrium stays pure.

strategy trajectories of both players start at the center of a strategy simplex, i.e. the uniform equilibrium of harmonic games [Candogan et al.*, 2011]. As λ increases, the strategies move towards the boundaries, and their supports become small. When λ is large enough, they reach a pure NE (nodes with stars), which is typical of potential games, and no longer moves. Figure 2b shows the changes of player 1's strategies in line chart.*

We utilize the algorithm by Heyman [2019] to decide when will $\mathbf{G}(\lambda) \in \mathcal{S}(\mathcal{Z})$. When $\lambda < \frac{5}{6}$, The competitive part takes over, $\mathbf{G}(\lambda) \in \mathcal{S}(\mathcal{Z})$. When $\lambda > \frac{5}{6}$, $\mathcal{G}(\lambda) \in \mathcal{S}(\mathcal{I})$. When $\lambda = \frac{5}{6}$, $\mathbf{G}(\lambda) \notin \mathcal{S}(\mathcal{Z}) \cup \mathcal{S}(\mathcal{I})$ but belongs to $\mathcal{D}$ instead. We draw the point $\lambda = \frac{5}{6}$ on Figure 2b. One can find out that when λ is less than but close to $\frac{5}{6}$, the equilibrium for FP to converge to already becomes pure and never moves even when λ exceeds $\frac{5}{6}$.

When λs in the above examples reach the threshold, the games enter $\mathcal{D}$ instead of $\mathcal{S}(\mathcal{Z}) \cap \mathcal{S}(\mathcal{I})$. We argue that games in $\mathcal{D}$ show the same dynamic patterns with games in $\mathcal{S}(\mathcal{Z}) \cap \mathcal{S}(\mathcal{I})$. Specifically, we show that $\mathcal{S}(\mathcal{Z}) \cap \mathcal{S}(\mathcal{I}) = \mathcal{A}(\mathcal{Z}) \cap \mathcal{A}(\mathcal{I}) = \mathcal{B}$, and players in games belonging to $\mathcal{B}$ have dominant strategies. In both classes of games, the players will choose a determined strategy and the dynamics converge to the corresponding pure NE directly.

Proposition 1. $\mathcal{S}(\mathcal{Z}) \cap \mathcal{S}(\mathcal{I}) = \mathcal{A}(\mathcal{Z}) \cap \mathcal{A}(\mathcal{I}) = \mathcal{B}$.

The process of transforming from zero-sum equivalent games to identical interest equivalent games with the changes of a linear parameter can be seen as an instinct feature of these two games, and a bridge linking non-cooperative game-theoretic view of cooperation and competition. Using a linear parameter that changes with time, we may be able to model how the relationships of players evolve in real world. Furthermore, when λ is small enough, we can regard the potential part in $\lambda\mathbf{P} + (1-\lambda)\mathbf{H}$ as a small perturbation added up to a harmonic game. Our result shows that when the perturbation satisfies certain conditions, the game with perturbations can still be regarded as a zero-sum equivalent game.

4 Analysis on the Non-convergence Example

In this section, we use game decompositions to give a new analysis of the classic non-convergent example of FP, the Shapley game [Shapley, 1964]. We consider Continuous-time Fictitious Play (CFP), a useful tool to give insights into DFP's dynamics. we present the Shapley game and introduce a novel function, the best response utility function. We give two results about initial conditions based on this function.

4.1 Continuous-time Fictitious Play (CFP)

DFP can be regarded as an update procedure for players' empirical beliefs, where the update rate is 1. Now consider the corresponding continuous version, by rescaling the rate to $\delta > 0$ and letting $\delta \to 0$. This equivalently defines the derivatives of the sequence $(\mathbf{p}(t), \mathbf{q}(t))$ with respect to t : [1]

$$\dot{\mathbf{p}}(t) = \frac{\mathrm{BR}_1(\mathbf{q}(t)) - \mathbf{p}(t)}{t}, \dot{\mathbf{q}}(t) = \frac{\mathrm{BR}_2(\mathbf{p}(t)) - \mathbf{q}(t)}{t}. \quad (3)$$

The detailed derivation from DFP to CFP is in Appendix D.

For CFP, we define another stable state following the studies on dynamical systems.

Definition 4 (Cycle)**.** *A CFP follows a cycle* C *if there is an integer* $K > 0$ *and a sequence of* K *pairs of pure strategies:*

$$\mathrm{C} = \{(i_1, j_1), \ldots, (i_K, j_K)\}$$

s.t., $\exists T > 0$, $\forall t > T$, $(\mathrm{BR}_1(\mathbf{q}(t)), \mathrm{BR}_2(\mathbf{p}(t)))$ *of the play takes the values* $(i_1, j_1), \ldots, (i_K, j_K)$ *periodically.*

It is a special case that a CFP converges to an NE when it follows a cycle. So far, all the known convergence results of CFP and DFP are the same. Intuitively, when discrete time step tends to $+\infty$, so as the denominator of Eqn. (1), the changes of $(\mathbf{p}(t), \mathbf{q}(t))$ at each time step become infinitesimal, and will resemble the derivatives defined by Eqn. (3). Thus, CFP can provide useful insights into original DFP, though the relationships of two dynamics are still not clear.

Now we study the conditions that a CFP with some initial conditions converges to an NE.

[1] One can further rescale the variable t to $t' = \ln t$ (for $t > 0$) to eliminate the the denominator.

$$\underbrace{\begin{bmatrix} 0,0 & 2,1 & 1,2 \\ 1,2 & 0,0 & 2,1 \\ 2,1 & 1,2 & 0,0 \end{bmatrix}}_{\text{the Shapley game}} = \underbrace{\begin{bmatrix} -1,-1 & 0.5,0.5 & 0.5,0.5 \\ 0.5,0.5 & -1,-1 & 0.5,0.5 \\ 0.5,0.5 & 0.5,0.5 & -1,-1 \end{bmatrix}}_{\substack{\text{Potential game} \\ \text{Normalized identical interest game}}} + \underbrace{\begin{bmatrix} 0,0 & 0.5,-0.5 & -0.5,0.5 \\ -0.5,0.5 & 0,0 & 0.5,-0.5 \\ 0.5,-0.5 & -0.5,0.5 & 0,0 \end{bmatrix}}_{\substack{\text{Harmonic game} \\ \text{Normalized zero-sum game}}} + \underbrace{\begin{bmatrix} 1,1 & 1,1 & 1,1 \\ 1,1 & 1,1 & 1,1 \\ 1,1 & 1,1 & 1,1 \end{bmatrix}}_{\substack{\text{Non-strategic game} \\ \text{Zero-sum equivalent potential game}}}$$

Figure 3: Hodge Decomposition and Strategic Decomposition of the Shapley game. It is a coincident that they are the same.

4.2 The Shapley Game

The Shapley game, proposed by Shapley [1964], is a counter-example of FP's convergence: except under certain initial conditions, FP does not converge to NE. Figure 3 shows its payoff matrices and the two decomposition results.

When two players' initial strategies are different, e.g. $(1, 2)$, FP follows cycle $\mathrm{C}_1 = \{(1, 2), (1, 3), (2, 3), (2, 1), (3, 1), (3, 2)\}$, but does not converge: both players' strategies will change periodically but never reach any fixed point. When their initial strategies are the same, e.g. $(1, 1)$, FP follows another cycle $\mathrm{C}_2 = \{(2, 2), (1, 1), (3, 3)\}$, and tends to an NE $(\mathbf{p}, \mathbf{q})$, where $\mathbf{p} = \mathbf{q} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$.

4.3 The Best Response Utility Function

We further consider more general combinations of the decomposed components $\mathbf{G}(\lambda) = \lambda\mathbf{P} + (1-\lambda)\mathbf{H} = \lambda\mathbf{I} + (1-\lambda)\mathbf{Z}$. Define the best response utility (BRU) function for game $\mathbf{G} = (A, B)$ and players' strategies $(\mathbf{p}, \mathbf{q})$ to be

$$U(\mathbf{p}, \mathbf{q}, \mathbf{G}) = \max_{i\in[m]}(A\mathbf{q})_i + \max_{j\in[n]}(\mathbf{p}B)_j.$$

It is the sum of maximal utilities each player can get by her best response to the opponent's strategy. We omit the last variable of U when there is no confusion. Experimental analysis (in Appendix D) on the Shapley game shows that whether DFP converges to NE or not, U always converges.

For the empirical frequencies $(\mathbf{p}(t), \mathbf{q}(t))$ $(t \geq 1)$ obtained by CFP, consider the derivative of $U(\mathbf{p}(t), \mathbf{q}(t))$ ($U(t)$ for short) with respect to t, for almost all t, then for any fixed $t_0 > 0$, we have

$$U(t) = \frac{t_0 U(t_0)}{t} + \frac{1}{t}\int_{t_0}^{t} G_{i(\tau),j(\tau)}\mathrm{d}\tau, \tag{4}$$

where $G = A + B$, $i(\tau)$ and $j(\tau)$ are short for the index of $\mathrm{BR}_1(\mathbf{q}(\tau))$ and $\mathrm{BR}_2(\mathbf{p}(\tau))$. The detailed derivation is in Appendix. This leads to a sufficient condition for the convergence of BRU.

Lemma 3. *For a game $\mathbf{G} = (A, B)$, let $G = A + B$. Assume CFP follows a cycle $\{(i_1, j_1), \ldots, (i_K, j_K)\}$. If $G_{i_1,j_1} = \cdots = G_{i_K,j_K}$, then BRU converges.*

Back to the Shapley game, on cycle C_2, $G_{i_1,j_1} = \cdots = G_{i_3,j_3} = \min G_{ij}$. To be compared, on cycle C_1, $G_{i_1,j_1} = \cdots = G_{i_6,j_6} > \min G_{ij}$. This means BRU converges but not to the minimal possible value, which coincides with the fact that the strategies fail to converge.

One can further find out that when a CFP starts with the same strategies for both players, i.e., when it converges to NE, by the symmetry of the game, $\mathbf{p}(t) = \mathbf{q}(t)$ for all t. Then for all strategy pairs $(\mathbf{p}(t), \mathbf{q}(t))$ related to C_2, $\dot{U}$ satisfies

$$\dot{U} = 2\mathbf{e}(t)I\mathbf{e}(t) - (\mathbf{p}^\top B\mathbf{e}(t) + \mathbf{e}(t)A\mathbf{q}) < 0,$$

where I is the payoff matrix of the Shapley game's normalized identical interest component, and $\mathbf{e}(t) = \mathrm{BR}_1(\mathbf{q}(t)) = \mathrm{BR}_2(\mathbf{p}(t))$. $U(t)$ will keep decreasing till it tends to NE.

Now we consider the Shapley game as a linear combination of a normalized identical interest game and a normalized zero-sum game and we state the reason formally in Theorem 4 why it converges under such initial conditions. Given a cycle C of game $\mathbf{G}$ such that there exist CFPs tending to C, let $P(\mathrm{C}, \mathbf{G})$ be all the mixed strategy pairs that lie on the paths following C, and $\mathrm{Conv}(\mathrm{S})$ be the closed convex hull of set S. Denote the set of $\mathbf{G}$'s NEs as $\mathcal{X}(\mathbf{G})$. We have

Theorem 4. *For game $\mathbf{G}(\lambda) = \lambda\mathbf{I} + \mathbf{Z}$ for some $\mathbf{I} \in (\mathcal{I} \cap \mathcal{N})$ and $\mathbf{G} \in (\mathcal{Z} \cap \mathcal{N})$, $\lambda \in \mathbb{R}$. If from an initial point, CFP on $\mathbf{G}(\lambda)$ enters the same cylce C of $\mathbf{G}(0)$, where C is a cycle that a convergent CFP on $\mathbf{G}(0)$ will tend to, and for all $(\mathbf{p}, \mathbf{q}) \in \mathrm{Conv}(P(\mathrm{C}, \mathbf{G}))\backslash\mathcal{X}(\mathbf{G}(0))$, $\dot{U}(\mathbf{p}, \mathbf{q}) < 0$, then CFPs on $\mathbf{G}(\lambda)$ which tend to C will converge to an NE.*

Theorem 4 gives a sufficient condition for a CFP to converge. Intuitively, when $|\lambda|$ is small enough, the dynamics of $\mathbf{G}(\lambda)$ will resemble that of $\mathbf{G}(0)$. If CFPs on these games enter the same cycle of $\mathbf{G}(0)$, and $\dot{U} < 0$ for all points related to this cycle, then CFPs on $\mathbf{G}(\lambda)$ converge to NE.

5 Conclusion and Future Work

Decomposing the game space into combinations of simple classes enables us to find the new relations among games. In this paper, we use this method to prove a new condition for FP to converge and build a bridge between games modelling full competition and cooperation. We derive an instinct property for them that these two classes of games can be mutually transformed to the other with a simple parameter. This ability of mutual transformation may be applied in the dynamic multi-agent systems in which agents' relationships vary with time, which helps study complex real environments. Furthermore, we analyze the well-known example of FP's non-convergence, the Shapley game, and give a sufficient condition for its continuous version to converge.

As for the future work, the first one is to have more analysis on FP dynamics using game decomposition techniques. We note that it will be of vital importance but great challenge to give a full characterization, and it is impossible to have a convergence guarantee on arbitrary combinations of games. Another interesting problem is to study the dynamical properties of simple games, e.g., zero-sum games, with perturbations.

Acknowledgments

This work is supported by Science and Technology Innovation 2030 –"The Next Generation of Artificial Intelligence" Major Project No. (2018AAA0100901).

References

[Adler *et al.*, 2009] Ilan Adler, Constantinos Daskalakis, and Christos H Papadimitriou. A note on strictly competitive games. In *International Workshop on Internet and Network Economics*, pages 471–474. Springer, 2009.

[Balduzzi *et al.*, 2018] David Balduzzi, Sebastien Racaniere, James Martens, Jakob Foerster, Karl Tuyls, and Thore Graepel. The mechanics of n-player differentiable games. In *International Conference on Machine Learning*, pages 354–363. PMLR, 2018.

[Brown, 1951] George W Brown. Iterative solution of games by fictitious play, 1951. *Activity Analysis of Production and Allocation (TC Koopmans, Ed.)*, pages 374–376, 1951.

[Candogan *et al.*, 2011] Ozan Candogan, Ishai Menache, Asuman Ozdaglar, and Pablo A Parrilo. Flows and decompositions of games: Harmonic and potential games. *Mathematics of Operations Research*, 36(3):474–503, 2011.

[Candogan *et al.*, 2013] Ozan Candogan, Asuman Ozdaglar, and Pablo A Parrilo. Dynamics in near-potential games. *Games and Economic Behavior*, 82:66–90, 2013.

[Cheung and Tao, 2021] Yun Kuen Cheung and Yixin Tao. Chaos of learning beyond zero-sum and coordination via game decompositions. In *ICLR 2021*. OpenReview.net, 2021.

[Dasgupta and Collins, 2019] Prithviraj Dasgupta and Joseph Collins. A survey of game theoretic approaches for adversarial machine learning in cybersecurity tasks. *AI Magazine*, 40(2):31–43, 2019.

[Deng *et al.*, 2021] Xiaotie Deng, Yuhao Li, David Henry Mguni, Jun Wang, and Yaodong Yang. On the complexity of computing markov perfect equilibrium in general-sum stochastic games. *arXiv preprint arXiv:2109.01795*, 2021.

[Dinh *et al.*, 2021] Le Cong Dinh, Yaodong Yang, Zheng Tian, Nicolas Perez Nieves, Oliver Slumbers, David Henry Mguni, Haitham Bou Ammar, and Jun Wang. Online double oracle. *arxiv preprint arxiv:2103.07780*, 2021.

[Feng *et al.*, 2021] Xidong Feng, Oliver Slumbers, Ziyu Wan, Bo Liu, Stephen McAleer, Ying Wen, Jun Wang, and Yaodong Yang. Neural auto-curricula in two-player zero-sum games. *Advances in Neural Information Processing Systems*, 34, 2021.

[Gopalakrishnan *et al.*, 2011] Ragavendran Gopalakrishnan, Jason R Marden, and Adam Wierman. Characterizing distribution rules for cost sharing games. In *International Conference on NETwork Games, Control and Optimization (NetGCooP 2011)*, pages 1–4. IEEE, 2011.

[Heinrich *et al.*, 2015] Johannes Heinrich, Marc Lanctot, and David Silver. Fictitious self-play in extensive-form games. In *Proceedings of The 32nd International Conference on Machine Learning*, pages 805–813, 2015.

[Heyman, 2019] Joseph Lee Heyman. *On the Computation of Strategically Equivalent Games*. ProQuest LLC, Ann Arbor, MI, 2019. Thesis (Ph.D.)–The Ohio State University.

[Hwang and Rey-Bellet, 2020] Sung-Ha Hwang and Luc Rey-Bellet. Strategic decompositions of normal form games: Zero-sum games and potential games. *Games and Economic Behavior*, 2020.

[Lanctot *et al.*, 2017] Marc Lanctot, Vinicius Zambaldi, Audrūnas Gruslys, Angeliki Lazaridou, Karl Tuyls, Julien Pérolat, David Silver, and Thore Graepel. A unified game-theoretic approach to multiagent reinforcement learning. In *NeurIPS*, pages 4193–4206, 2017.

[Liu *et al.*, 2021] Xiangyu Liu, Hangtian Jia, Ying Wen, Yaodong Yang, Yujing Hu, Yingfeng Chen, Changjie Fan, and Zhipeng Hu. Unifying behavioral and response diversity for open-ended learning in zero-sum games. *NeurIPS*, 2021.

[McMahan *et al.*, 2003] H Brendan McMahan, Geoffrey J Gordon, and Avrim Blum. Planning in the presence of cost functions controlled by an adversary. In *ICML-03*, pages 536–543, 2003.

[Mguni *et al.*,] David H Mguni, Yutong Wu, Yali Du, Yaodong Yang, Ziyi Wang, Minne Li, Ying Wen, Joel Jennings, and Jun Wang. Learning in nonzero-sum stochastic games with potentials. In *ICML 2021*, pages 7688–7699. PMLR.

[Monderer and Shapley, 1996a] Dov Monderer and Lloyd S. Shapley. Fictitious play property for games with identical interests. *Journal of Economic Theory*, 68(1):258–265, 1996.

[Monderer and Shapley, 1996b] Dov Monderer and Lloyd S Shapley. Potential games. *Games and economic behavior*, 14(1):124–143, 1996.

[Muller *et al.*,] Paul Muller, Shayegan Omidshafiei, Mark Rowland, Karl Tuyls, Julien Perolat, Siqi Liu, Daniel Hennes, Luke Marris, Marc Lanctot, Edward Hughes, Zhe Wang, Guy Lever, Nicolas Heess, Thore Graepel, and Remi Munos. A generalized training approach for multiagent learning. In *ICLR 2020*.

[Nash, 1950] John F. Nash. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, 36(1):48–49, 1950.

[Ostrovski and van Strien, 2014] Georg Ostrovski and Sebastian van Strien. Payoff performance of fictitious play. *J. Dyn. Games*, 1(4):621–638, 2014.

[Perez-Nieves *et al.*, 2021] Nicolas Perez-Nieves, Yaodong Yang, Oliver Slumbers, David H Mguni, Ying Wen, and Jun Wang. Modelling behavioural diversity for learning in open-ended games. In *ICML*, pages 8514–8524. PMLR, 2021.

[Robinson, 1951] Julia Robinson. An iterative method of solving a game. *Annals of Mathematics*, 54(2):296–301, 1951.

[Shapley, 1964] Lloyd Shapley. Some topics in two-person games. *Advances in game theory*, 52:1–29, 1964.

[Swenson *et al.*, 2018] Brian Swenson, Ryan Murray, and Soummya Kar. On best-response dynamics in potential games. *SIAM Journal on Control and Optimization*, 56(4):2734–2767, 2018.

[Tuyls *et al.*, 2018] Karl Tuyls, Julien Pérolat, Marc Lanctot, Georg Ostrovski, Rahul Savani, Joel Z Leibo, Toby Ord, Thore Graepel, and Shane Legg. Symmetric decomposition of asymmetric games. *Scientific reports*, 8(1):1–20, 2018.

[Vöcking, 2007] Berthold Vöcking. Selfish load balancing. *Algorithmic game theory*, 20:517–542, 2007.

[Yang and Wang, 2020] Yaodong Yang and Jun Wang. An overview of multi-agent reinforcement learning from game theoretical perspective. *arXiv preprint arXiv:2011.00583*, 2020.

[Yang *et al.*, 2018] Yaodong Yang, Rui Luo, Minne Li, Ming Zhou, Weinan Zhang, and Jun Wang. Mean field multi-agent reinforcement learning. In *International Conference on Machine Learning*, pages 5571–5580. PMLR, 2018.