Are Equivariant Equilibrium Approximators Beneficial?

Zhijian Duan [1]   **Yunxuan Ma** [1]   **Xiaotie Deng** [1 2]

Abstract

Recently, remarkable progress has been made by approximating Nash equilibrium (NE), correlated equilibrium (CE), and coarse correlated equilibrium (CCE) through function approximation that trains a neural network to predict equilibria from game representations. Furthermore, equivariant architectures are widely adopted in designing such equilibrium approximators in normal-form games. In this paper, we theoretically characterize the benefits and limitations of equivariant equilibrium approximators. For the benefits, we show that they enjoy better generalizability than general ones and can achieve better approximations when the payoff distribution is permutation-invariant. For the limitations, we discuss their drawbacks in terms of equilibrium selection and social welfare. Together, our results help to understand the role of equivariance in equilibrium approximators.

1. Introduction

The equivariant equilibrium property states that, given a Nash Equilibrium (NE) solution of a game, the permuted solution is also an NE for the game whose actions of representation are permuted in the same way. The same property also holds in correlated equilibrium (CE) and coarse correlated equilibrium (CCE), as well as the approximate solutions for all three solution concepts.

In this paper, we are interested in understanding the equivariant equilibrium property in designing neural networks that predict equilibria from game payoffs, following such recent approaches in designing equivariant equilibrium approximators (Feng et al., 2021; Marris et al., 2022) in normal-form games. Informally, such equivariant approximators keep the same permutation of the output strategies (represented as

vectors or tensors) when the input game representations (e.g., the game payoff tensors) are permuted. We will provide a formal definition of equivariance equilibrium approximators in Section 4. While equivariant approximators achieved empirical success, little work has theoretically discussed whether they are beneficial.

We theoretically characterize the benefits and limitations of equivariant NE, CE, and CCE approximators. For the benefits, we first show that equivariant approximators enjoy better generalizability, where we evaluate the approximators using the maximum exploitability (Lockhart et al., 2019; Goktas & Greenwald, 2022) over all players. To get such a result, we derive the generalization bounds and the sample complexities of the NE, CE, and CCE approximators: The generalization bounds offer confidence intervals on the expected testing approximations based on the empirical training approximations; The sample complexities describe how many training samples the equilibrium approximators need to achieve desirable generalizability. Both the generalization bounds and sample complexities include the covering numbers (Shalev-Shwartz & Ben-David, 2014), which measure the representativeness of the approximators' function classes. We then prove that the equivariant approximators have lower covering numbers than the general models, therefore have lower generalization bounds and sample complexities. Additionally, we also show that the equivariant approximators can achieve better approximation when the payoff distribution is permutation-invariant.

As for the limitations, we find the equivariant approximators unable to find all the equilibria of some normal-form games. Such a result is caused by the limited representativeness of the equivariant approximators' function class. Besides, we find that the equivariant NE approximator may lose social welfare. Specifically, in an example we constructed, while the maximum NE social welfare is significant, the maximum social welfare of NEs that the equivariant NE approximators could find can be arbitrarily close to zero. Such a negative result inspires us to balance generalizability and social welfare when we design the approximators' architectures.

In summary, we make the following contributions:

- We theoretically characterize the benefits of equivariant equilibrium approximators: Compared to general

[1]Center on Frontiers of Computing Studies, School of Computer Science, Peking University, Beijing, China [2]Center for Multi-Agent Research, Institute for AI, Peking University, Beijing, China. Correspondence to: Xiaotie Deng <xiaotie@pku.edu.cn>.

approximators, equivariant ones offer improved generalizability and, in some instances, achieve better approximation performance.

- We demonstrate a limitation of equivariant equilibrium approximators through constructive examples: they cannot identify all equilibria in some games.

- We illustrate that equivariant NE approximators may find NE with arbitrarily small maximum social welfare in extreme cases. In contrast, the actual maximum social welfare achievable by NEs can be significantly larger.

The rest of our paper is organized as follows: In Section 2 we discuss further related works. In Section 3 we introduce the preliminary of game theory and equilibrium approximators. In Section 4 we formally define the equivariance of equilibrium approximators. We present our theoretical analysis of benefits in Section 5 and limitations in Section 6. We conclude and point out the future work in Section 7.

2. Related Work

Solving (approximate) NE, CE, and CCE for a single game has been extensively studied in the literature (Fudenberg et al., 1998; Cesa-Bianchi & Lugosi, 2006; Tsaknakis & Spirakis, 2008; Fu & Lin, 2020; Deligkas et al., 2022). However, in many practical scenarios and multi-agent learning algorithms, there is a need to solve many similar games (Marris et al., 2021; Liu et al., 2022). For instance, in repeated traffic routing games (Sessa et al., 2020), the payoffs of games depend on the capacity of the underlying network, which can vary with time and weather conditions. In repeated sponsored search auctions, advertisers value different keywords based on the current marketing environment (Nekipelov et al., 2015). In multi-agent reinforcement learning (Deng et al., 2023) algorithms such as Nash Q-learning (Hu & Wellman, 2003), Correlated-Q learning (Greenwald et al., 2003), V-learning (Jin et al., 2022) and PSRO (Lanctot et al., 2017), an NE, CE or CCE of a normal-form game need to be solved in every update step.

In these settings, it is preferred to accelerate the speed of game solving by function approximation: Duan et al. (2023) propose the construction of a neural network for approximating NE in normal-form games and analyze the PAC learnability of such an NE approximator; Marris et al. (2022) introduce a neural equilibrium approximator to approximate CE and CCE in n-player normal-form games; Feng et al. (2021) propose a neural NE approximator within the context of PSRO (Lanctot et al., 2017); Wu & Lisser (2022) design a CNN-based NE approximator specifically for zero-sum bimatrix games. Differentiable approximators have also been developed to learn QREs (Ling et al., 2018), NE in chance-

constrained games (Wu & Lisser, 2023), and opponent's strategy (Hartford et al., 2016).

Equivariance is an ideal property of the equilibrium approximator. After formally defining them, we will discuss the literates of equivariant approximators in Section 4.2.

3. Preliminary

In this section, we introduce the preliminary and notations of our paper. We also provide a brief introduction to equilibrium approximators.

3.1. Game Theory

Normal-Form Game Let a normal-form game with joint payoff u be $\Gamma_u = (n, \mathcal{A}, u)$, in which

- $n \in \mathbb{N}_{\geq 2}$ is the number of players. Each player is represented by the index $i \in [n]$.

- $\mathcal{A} = \times_{i \in [n]} \mathcal{A}_i$ is the product action space of all players, where $\mathcal{A}_i = \{1, 2, \ldots, m_i\}$. For player $i \in [n]$, let $a_i \in \mathcal{A}_i$ be a specific action of i (An action is also referred to as a pure strategy). A joint action $a = (a_1, a_2, \ldots, a_n) \in \mathcal{A}$ represents one play of the game in which the player i takes action a_i. The action space $\mathcal{A}$ is a Cartesian product that contains all possible joint actions. We have $|\mathcal{A}| = \prod_{i \in [n]} |\mathcal{A}_i| = \prod_{i \in [n]} m_i$.

- $u = (u_i)_{i \in [n]}$ is the joint payoff or utility of the game. u_i is an n-dimensional tensor (or matrix if $n = 2$) describing player i's payoff on each joint action. In our paper, following previous literatures (Tsaknakis & Spirakis, 2008; Deligkas et al., 2022), we normalize all the elements of payoff into $[0, 1]$.

A joint (mixed) strategy is a distribution over $\mathcal{A}$. Let $\sigma = \times_{i \in [n]} \sigma_i$ be a product strategy and $\pi \in \Delta \mathcal{A}$ be a joint (possibly correlated) strategy. Denote π_i as the marginal strategy of player i in π. The expected utility of player i under π is

$$u_i(\pi) = \mathbb{E}_{a \sim \pi}[u_i(a)] = \sum_{a \in \mathcal{A}} \pi(a) u_i(a).$$

Besides, on behalf of player i, the other players' joint strategy is denoted as π_{-i}, so as a_{-i} and σ_{-i}.

Nash Equilibrium (NE) We say a product strategy $\sigma^* = (\sigma_1^*, \sigma_2^*, \ldots, \sigma_n^*)$ is a NE if each player's strategy is the best response given the strategies of others, i.e.,

$$u_i(\sigma_i, \sigma_{-i}^*) \leq u_i(\sigma_i^*, \sigma_{-i}^*), \quad \forall i, \forall \sigma_i. \quad \text{(NE)}$$

Computing NE for even general 2-player or 3-player games is PPAD-hard (Chen et al., 2009; Daskalakis et al., 2009),

which leads to research on approximate solutions. For arbitrary $\epsilon > 0$, we say a product strategy $\hat{\sigma}$ is an ϵ-*approximate Nash equilibrium* (ϵ-NE) if no one can achieve more than ϵ utility gain by deviating from her current strategy. Formally,

$$u_i(\sigma_i, \hat{\sigma}_{-i}) \leq u_i(\hat{\sigma}_i, \hat{\sigma}_{-i}) + \epsilon, \quad \forall i, \forall \sigma_i. \qquad (\epsilon\text{-NE})$$

The definition of ϵ-NE reflects the idea that players might not be willing to deviate from their strategies when the amount of utility they could gain by doing so is tiny (not more than ϵ).

Coarse Correlated Equilibrium (CCE) We say a joint (possibly correlated) strategy π^* is a CCE if no player can receive a higher payoff by acting independently, i.e.,

$$u_i(\sigma_i, \pi^*_{-i}) \leq u_i(\pi^*), \quad \forall i, \forall \sigma_i \in \Delta\mathcal{A}_i, \qquad (\text{CCE})$$

and we say $\hat{\pi}$ is an ϵ-approximate coarse correlated equilibrium (ϵ-CCE) for $\epsilon > 0$ if

$$u_i(\sigma_i, \hat{\pi}_{-i}) \leq u_i(\hat{\pi}) + \epsilon, \quad \forall i, \forall \sigma_i \in \Delta\mathcal{A}_i, \quad (\epsilon\text{-CCE})$$

The difference between NE and CCE is that in an NE, players execute their strategy individually in a decentralized way, while in a CCE, players' strategies are possibly correlated. A standard technique to correlate the strategy is sending each player a signal from a centralized controller (Shoham & Leyton-Brown, 2008).

Correlated Equilibrium (CE) CE is similar to CCE, except that in a CE, each player can observe her recommended action before she acts. Thus, player i deviates her strategy through *strategy modification* $\phi_i : \mathcal{A}_i \to \mathcal{A}_i$. ϕ_i maps actions in $\mathcal{A}_i$ to possibly different actions in $\mathcal{A}_i$. Based on strategy modification, we say a joint (possibly correlated) strategy π^* is a CE if

$$\sum_{a \in \mathcal{A}} \pi^*(a) u_i(\phi_i(a_i), a_{-i}) \leq u_i(\pi^*), \quad \forall i, \forall \phi_i, \quad (\text{CE})$$

and a joint strategy $\hat{\pi}$ is an ϵ-approximate correlated equilibrium (ϵ-CE) for $\epsilon > 0$ if

$$\sum_{a \in \mathcal{A}} \hat{\pi}(a) u_i(\phi_i(a_i), a_{-i}) \leq u_i(\hat{\pi}) + \epsilon, \quad \forall i, \forall \phi_i, \quad (\epsilon\text{-CE})$$

Note that for a finite n-player normal-form game, at least one NE, CE, and CCE must exist. This is because NE always exists (Nash et al., 1950) and NE $\subseteq$ CE $\subseteq$ CCE.

Equilibrium Approximation To evaluate the quality of a joint strategy to approximate an equilibrium, we define approximation based on exploitability (Lockhart et al., 2019; Goktas & Greenwald, 2022).

Definition 3.1 (Exploitability and Approximation). Given a joint strategy π, the *exploitability* (or regret) $\mathcal{E}_i(\pi, u)$ of player i is the maximum payoff gain of i by deviating from her current strategy, i.e.,

$$\mathcal{E}_i(\pi, u) := \max_{\sigma'_i} u_i(\sigma'_i, \pi_{-i}) - u_i(\pi)$$
$$= \max_{a'_i} u_i(a'_i, \pi_{-i}) - u_i(\pi)$$

and the exploitability under strategy modification $\mathcal{E}_i^{\text{CE}}(\pi, u)$ of player i is the maximum payoff gain of i by deviating through strategy modification, i.e.,

$$\mathcal{E}_i^{\text{CE}}(\pi, u) := \max_{\phi_i} \sum_{a \in \mathcal{A}} \pi(a) u_i(\phi_i(a_i), a_{-i}) - u_i(\pi).$$

The *equilibrium approximation* is defined as the maximum exploitability over all players [1], i.e.,

$$\mathcal{E}(\pi, u) := \begin{cases} \max_{i \in [n]} \mathcal{E}_i(\pi, u) & \text{, for NE and CCE} \\ \max_{i \in [n]} \mathcal{E}_i^{\text{CE}}(\pi, u) & \text{, for CE} \end{cases}$$

Based on approximation, we can restate the definition of solution concepts. A product strategy σ is an NE of game Γ_u if $\mathcal{E}(\sigma, u) = 0$ and is an ϵ-NE if $\mathcal{E}(\sigma, u) \leq \epsilon$. A joint strategy π is a (C)CE of Γ_u if $\mathcal{E}(\pi, u) = 0$ and is an ϵ-(C)CE if $\mathcal{E}(\pi, u) \leq \epsilon$.

3.2. Equilibrium Approximator

The equilibrium approximators, including NE, CE, and CCE approximators, aim to predict solution concepts from game representations. In our paper, we fix the number of players n and action space $\mathcal{A}$. We denote by $\mathcal{U}$ the set of all the possible game payoffs. The NE approximator $f^{\text{NE}} : \mathcal{U} \to \times_{i \in [n]} \Delta\mathcal{A}_i$ maps a game payoff to a product strategy, where $f^{\text{NE}}(u)_i \in \Delta\mathcal{A}_i$ is player i's predicted strategy. We define $\mathcal{F}^{\text{NE}}$ as the function class of the NE approximator. Similarly, we define (C)CE approximator as $f^{(\text{C})\text{CE}} : \mathcal{U} \to \Delta\mathcal{A}$ and (C)CE approximator class as $\mathcal{F}^{(\text{C})\text{CE}}$.

An equilibrium approximator can be learned through machine learning paradigms based on empirical data. For instance, Feng et al. (2021) learn the NE approximator using the game payoffs generated in the learning process of PSRO, and optimize the approximator by gradient descent in exploitability. Marris et al. (2022) learn the CE and CCE approximators using the games i.i.d. generated from a manually designed distribution, and optimize the approximators using maximum welfare minimum relative entropy loss. Such a loss balances the equilibrium approximation,

---

[1] A similar metric of equilibrium approximation is called Nikaido-Isoda function (Nikaidô & Isoda, 1955) or NASH-CONV (Lockhart et al., 2019), which is the sum of exploitability over all players, i.e., $\sum_{i \in [n]} \mathcal{E}_i(\pi, u)$.

Algorithm 1 Example: learning NE/CCE approximator via minibatch SGD

1: **Input:** Training set S
2: **Parameters:** Number of iterations $T > 0$, batch size $B > 0$, learning rate $\eta > 0$, initial parameters $w_0 \in \mathbb{R}^d$ of the approximator model.
3: **for** $t = 0$ **to** T **do**
4:     Receive minibatch $S_t = \{u^{(1)}, \ldots, u^{(B)}\} \subset S$
5:     Compute the empirical average approximation of S_t:
6:       $L_{S_t}(f^{w_t}) \leftarrow \frac{1}{B} \sum_{i=1}^{B} \mathcal{E}(f^{w_t}(u^{(i)}), u^{(i)})$
7:     Update model parameters:
8:       $w_{t+1} \leftarrow w_t - \eta \nabla_{w_t} L_{S_t}(f^{w_t})$
9: **end for**

---

the social welfare, and the relative entropy of the joint strategy. Additionally, another simple way to learn NE or CCE equilibrium approximator is to apply minibatch stochastic gradient descent (SGD) on approximation. Specifically, we denote $w \in \mathbb{R}^d$ as the d-dimensional parameter of the approximator, such as the weights of the neural network. We can optimize w by the standard minibatch SGD algorithm on approximation (See Algorithm 1).

4. Equivariant Equilibrium Approximator

In this section, we introduce the equivariance of the equilibrium approximators and the way how we apply orbit averaging (Elesedy & Zaidi, 2021) to construct equivariant approximators. Equivariant approximator has been developed in many literatures (Hartford et al., 2016; Feng et al., 2021; Marris et al., 2022; Wu & Lisser, 2022), which we will discuss later.

4.1. Definition of Equivariance

We first define the permutation of a game. Let $\rho_i : \mathcal{A}_i \to \mathcal{A}_i$ be a permutation function of player i, which is a bijection from $\mathcal{A}_i$ to $\mathcal{A}_i$ itself. Define $\mathcal{G}_i \ni \rho_i$ as the class of permutation function for player i, which forms an Abelian group.

Definition 4.1 (Permutation of a game). For a normal-form game $\Gamma_u = (n, u, \mathcal{A})$, we define the ρ_i-permutation of payoff tensor u as $\rho_i u = (\rho_i u_j)_{j \in [n]}$, in which

$$(\rho_i u_j)(a_i, a_{-i}) = u_j(\rho_i^{-1}(a_i), a_{-i}), \quad \forall a \in \mathcal{A}.$$

We also define the ρ_i-permutation of joint strategy π as $\rho_i \pi$, where

$$(\rho_i \pi)(a_i, a_{-i}) = \pi(\rho_i^{-1}(a_i), a_{-i}), \quad \forall a \in \mathcal{A},$$

and the ρ_i-permutation of product strategy σ as $\rho_i \sigma = (\rho_i \sigma_j)_{j \in [n]}$, where

$$\forall a_j \in \mathcal{A}_j, \rho_i \sigma_j(a_j) = \begin{cases} \sigma_j(a_j) & \text{, if } j \neq i \\ \sigma_i(\rho_i^{-1} a_i) & \text{, if } j = i \end{cases}$$

Equivariant NE Approximator Considering the relationship of ρ_i-permuted game and ρ_i-permuted product strategy, we have the following result for NE:

Lemma 4.2. *In a normal-form game* $\Gamma_u = (n, u, \mathcal{A})$, *for arbitrary player* $i \in [n]$ *and any* $(\epsilon\text{-})$*NE strategy* $\sigma = (\sigma_i, \sigma_{-i})$, $\rho_i \sigma = (\rho_i \sigma_i, \sigma_{-i})$ *is also an* $(\epsilon\text{-})$*NE for the* ρ_i-*permuted game* $\Gamma_{\rho_i u}$.

Lemma 4.2 tells the unimportance of action order with respect to NE approximation. Based on it, we can summarize two ideal properties for NE approximators, which are defined as follows:

Definition 4.3 (Player-Permutation-Equivariance, PPE). We say an NE approximator f^{NE} satisfies *player i permutation-equivariant (i-PE)* if for arbitrary permutation function $\rho_i \in \mathcal{G}_i$ we have

$$f^{\mathrm{NE}}(\rho_i u)_i = \rho_i f^{\mathrm{NE}}(u)_i, \qquad (i\text{-PE})$$

Moreover, we say f^{NE} is *player-permutation-equivariant (PPE)* if f^{NE} satisfies i-PE for all player $i \in [n]$.

Definition 4.4 (Opponent-Permutation-Invariance, OPI). We say an NE approximator f^{NE} is *opponent i permutation-invariant (i-PI)* if for any other player $j \in [n] - \{i\}$ and arbitrary permutation function $\rho_i \in \mathcal{G}_i$ we have

$$f^{\mathrm{NE}}(\rho_i u)_j = f^{\mathrm{NE}}(u)_j, \forall j \neq i \qquad (i\text{-PI})$$

and we say f^{NE} is *opponent-permutation-invariant (OPI)* if f^{NE} satisfies i-PI for all player $i \in [n]$.

Equivariant (C)CE approximator Considering the relationship of ρ_i-permuted game and ρ_i-permuted joint strategy, we have a similar result for CE and CCE:

Lemma 4.5. *In a normal-form game* $\Gamma_u = (n, u, \mathcal{A})$, *for an arbitrary player* $i \in [n]$ *and any* $(\varepsilon\text{-})$*CE or* $(\epsilon\text{-})$*CCE strategy* π, $\rho_i \pi$ *is also an* $(\varepsilon\text{-})$*CE or* $(\epsilon\text{-})$*CCE for the* ρ_i-*permuted game* $\Gamma_{\rho_i u}$.

Inspired by Lemma 4.5, we can also summarize an ideal property for CE and CCE approximators defined as follows.

Definition 4.6 (Permutation-Equivariance,PE). We say an (C)CE approximator $f^{\mathrm{(C)CE}}$ is *player i permutation-equivariant (i-PE)* if for permutation function $\rho_i \in \mathcal{G}_i$ we have

$$f^{\mathrm{(C)CE}}(\rho_i u) = \rho_i f^{\mathrm{(C)CE}}(u),$$

and we say $f^{\mathrm{(C)CE}}$ is *permutation-equivariant (PE)* if $f^{\mathrm{(C)CE}}$ satisfies i-PE for all player $i \in [n]$.

4.2. Equivariant Approximators in Literature

For two-player games, Feng et al. (2021) propose an MLP-based NE approximator that satisfies both PPE and OPI for

zero-sum games. Additionally, they also design a Conv1d-based NE approximator that satisfies PPE only. Hartford et al. (2016) give a PPE approximator to predict players' strategies. The traditional algorithms Tsaknakis & Spirakis (2008) and Deligkas et al. (2022), which approximate NE by optimization, are also PPE and OPI to payoffs and the initial strategies. For n-player general games, Marris et al. (2022) provide a permutation-equivariant approximator to approximate CE and CCE. Equivariant architectures are also adopted in optimal auction design (Rahme et al., 2021; Duan et al., 2022; Ivanov et al., 2022), and Qin et al. (2022) theoretically characterize the benefits of permutation-equivariant in auction mechanisms. We follow the rough idea of Qin et al. (2022) when we analyze the benefits of equivariant equilibrium approximators.

4.3. Orbit Averaging

Orbit averaging is a well-known method to enforce equivariance or invariance for a function (Schulz-Mirbach, 1994). It averages the inputs of a function over the orbit of a group (e.g., the permutation group in our paper).

Orbit Averaging for NE Approximator For an NE approximator f^{NE} and any player $i \in [n]$, we can construct a i-PI or i-PE NE approximator by averaging with respect to all the permutations of player i. Specifically, we construct an i-PI NE approximator by operator $\mathcal{O}_i$ with

$$(\mathcal{O}_i f^{\text{NE}})(u)_j = \begin{cases} f^{\text{NE}}(u)_i & \text{, if } j = i \\ \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} f^{\text{NE}}(\rho_i u)_j & \text{, otherwise} \end{cases}$$

and we construct an i-PE NE approximator by operator $\mathcal{P}_i$ with:

$$(\mathcal{P}_i f^{\text{NE}})(u)_j = \begin{cases} \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} \rho_i^{-1} f^{\text{NE}}(\rho_i u)_i & \text{, if } j = i \\ f^{\text{NE}}(u)_j & \text{, otherwise} \end{cases}$$

Lemma 4.7. $\mathcal{O}_i f^{\text{NE}}$ *is i-PI and $\mathcal{P}_i f^{\text{NE}}$ is i-PE. Specially, if f^{NE} is already i-PI or i-PE, then we have $\mathcal{O}_i f^{\text{NE}} = f^{\text{NE}}$ or $\mathcal{P}_i f^{\text{NE}} = f^{\text{NE}}$, respectively.*

To construct a PPE or OPI NE approximator, we composite the operators with respect to all players. Let $\mathcal{O} = \mathcal{O}_1 \circ \mathcal{O}_2 \circ \cdots \circ \mathcal{O}_n$ and $\mathcal{P} = \mathcal{P}_1 \circ \mathcal{P}_2 \circ \cdots \circ \mathcal{P}_n$, we get the following corollary:

Lemma 4.8. $\mathcal{O} f^{\text{NE}}$ *is OPI and $\mathcal{P} f^{\text{NE}}$ is PPE. If f^{NE} is already OPI or PPE, we have $\mathcal{O} f^{\text{NE}} = f^{\text{NE}}$ or $\mathcal{P} f^{\text{NE}} = f^{\text{NE}}$, respectively.*

Furthermore, we can also compose $\mathcal{P} \circ \mathcal{O}$ to construct a NE approximator with both PPE and OPI.

Orbit Averaging for (C)CE Approximator For CE or CCE approximator f, we define $\mathcal{Q}_i$-project for player $i \in$

$[n]$ to construct an i-PE approximator, which averages with respect to all the permutations of player i.

$$(\mathcal{Q}_i f^{\text{(C)CE}})(u) = \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} \rho_i^{-1} f^{\text{(C)CE}}(\rho_i u)$$

Similarly, we define $\mathcal{Q} = \mathcal{Q}_1 \circ \mathcal{Q}_2 \circ \cdots \circ \mathcal{Q}_n$ as the composite operator.

Lemma 4.9. $\mathcal{Q}_i f^{\text{(C)CE}}$ *is i-PE and $\mathcal{Q} f^{\text{(C)CE}}$ is PE. Specifically, If $f^{\text{(C)CE}}$ is already i-PE or PE, then we have $\mathcal{Q}_i f^{\text{(C)CE}} = f^{\text{(C)CE}}$ or $\mathcal{Q} f^{\text{(C)CE}} = f^{\text{(C)CE}}$, respectively.*

Combined with Lemma 4.7, Lemma 4.8 and Lemma 4.9, we can have the following corollary directly.

Corollary 4.10. $\mathcal{O}^2 = \mathcal{O}, \mathcal{P}^2 = \mathcal{P}, \mathcal{Q}^2 = \mathcal{Q}$.

The benefit of using orbit averaging is shown in the following lemma:

Lemma 4.11. *Denote $\mathcal{X}$ as an idempotent operator, i.e. $\mathcal{X}^2 = \mathcal{X}$ (e.g. $\mathcal{O}, \mathcal{P}$ or $\mathcal{Q}$). For function class $\mathcal{F}$ of NE, CE or CCE approximator, let $\mathcal{F}_{\mathcal{X}}$ be any subset of $\mathcal{F}$ that is closed under $\mathcal{X}$, then $\mathcal{X} \mathcal{F}_{\mathcal{X}}$ is the largest subset of $\mathcal{F}_{\mathcal{X}}$ that is invariant under $\mathcal{X}$.*

According to Lemma 4.8, Lemma 4.9 and Lemma 4.11, $\mathcal{O} \mathcal{F}^{\text{NE}}$ (or $\mathcal{P} \mathcal{F}^{\text{NE}} / \mathcal{Q} \mathcal{F}^{\text{(C)CE}}$) is the largest subset of $\mathcal{F}^{\text{NE}}$ (or $\mathcal{F}^{\text{NE}} / \mathcal{F}^{\text{(C)CE}}$) with the corresponding property (OPI, PPE, or PE) if $\mathcal{F}^{\text{NE}}$ (or $\mathcal{F}^{\text{NE}} / \mathcal{F}^{\text{(C)CE}}$) is closed operator under $\mathcal{O}$ (or $\mathcal{P} / \mathcal{Q}$). The result tells that the orbit averaging operators, while enforcing the operated function to be equivariance or invariance, keep as large a capacity of the function class as possible. Therefore, we believe that orbit averaging is an ideal approach to constructing equivariant or invariant functions.

5. Theoretical Analysis of Benefits

In this section, we theoretically analyze the benefits of equivariant approximators with respect to generalizability and approximation.

5.1. Benefits for Generalization

We first derive the generalization bound and sample complexity for general approximator classes, and then we show the benefits of equivariant approximators by applying orbit averaging to the approximators.

The representativeness of an approximator class is measured by the covering numbers (Shalev-Shwartz & Ben-David, 2014) under ℓ_∞-distance, which are defined as follows:

Definition 5.1 (ℓ_∞-distance)**.** The ℓ_∞-distance between two equilibrium approximators f, g is:

$$\ell_\infty(f, g) = \max_{u \in \mathcal{U}} \|f(u) - g(u)\|,$$

where we define the distance of two product strategies σ and σ' as

$$\|\sigma^1 - \sigma^2\| = \max_{i \in [n]} \sum_{a_i \in \mathcal{A}_i} |\sigma_i^1(a_i) - \sigma_i^2(a_i)|,$$

and the distance of two joint strategy π and π' as

$$\|\pi^1 - \pi^2\| = \sum_{a \in \mathcal{A}} |\pi^1(a) - \pi^2(a)|.$$

Definition 5.2 (r-covering number). For $r > 0$, we say function class $\mathcal{F}_r$ r-covers another function class $\mathcal{F}$ under ℓ_∞-distance if for all function $f \in \mathcal{F}$, there exists $f_r \in \mathcal{F}_r$ such that $\|f - f_r\|_\infty \leq r$. The r-covering number $\mathcal{N}_\infty(\mathcal{F}, r)$ of $\mathcal{F}$ is the cardinality of the smallest function class $\mathcal{F}_r$ that r-covers $\mathcal{F}$ under ℓ_∞-distance.

Based on covering numbers, we provide the generalization bounds of NE, CE and CCE approximators. The bounds describe the difference between the expected testing approximation and empirical training approximation.

Theorem 5.3. *[Generalization bound] For function class $\mathcal{F}$ of NE, CE or CCE approximator, with probability at least $1 - \delta$ over draw of the training set S (with size m) from payoff distribution $\mathcal{D}$, for all approximator $f \in \mathcal{F}$ we have*

$$\mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}(f(u), u)] - \frac{1}{m} \sum_{u \in S} \mathcal{E}(f(u), u)$$

$$\leq 2 \cdot \inf_{r > 0}\{\sqrt{\frac{2 \ln \mathcal{N}_\infty(\mathcal{F}, r)}{m}} + Lr\} + 4\sqrt{\frac{2 \ln(4/\delta)}{m}},$$

where $L = 2n$ for NE approximator, and $L = 2$ for CE and CCE approximators.

To get the theorem, we first show that all three equilibrium approximations are Lipschitz continuous with respect to strategies. Afterward, we derive the Rademacher complexity (Bartlett & Mendelson, 2002) of the expected approximation based on the Lipschitz continuity and covering numbers. See Appendix B.1 for the detailed proof.

We can see from Theorem 5.3 that, with a large enough training set, the generalization gaps of equilibrium approximators go to zero if the covering number $\mathcal{N}_\infty(\mathcal{F}, r)$ is bounded. As a result, we can estimate the expected testing performance through the empirical training performance.

We can also derive the sample complexities of equilibrium approximators to achieve the desirable generalizability.

Theorem 5.4. *[Sample complexity] For $\epsilon, \delta \in (0, 1)$, function class $\mathcal{F}$ of NE, CE or CCE approximator and distribution $\mathcal{D}$, with probability at least $1 - \delta$ over draw of the training set S with*

$$m \geq \frac{9}{2\epsilon^2}\left(\ln \frac{2}{\delta} + \ln \mathcal{N}_\infty(\mathcal{F}, \frac{\epsilon}{3L})\right)$$

games sampled from $\mathcal{D}$, $\forall f \in \mathcal{F}$ we have

$$\mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}(f(u), u)] \leq \frac{1}{m} \sum_{u \in S} \mathcal{E}(f(u), u) + \epsilon,$$

where $L = 2n$ for NE approximators, and $L = 2$ for CE and CCE approximators.

The proof is based on the Lipschitz continuity of approximation, uniform bound, and concentration inequality. See Appendix B.2 for details. Theorem 5.4 is also called the uniform convergence of function class $\mathcal{F}$, which is a sufficient condition for agnostic PAC learnable (Shalev-Shwartz & Ben-David, 2014).

As for the benefits of equivariant approximators for generalizability, the following result indicates that the projected equilibrium approximators have smaller covering numbers.

Theorem 5.5. *The $\mathcal{O}$-projected, $\mathcal{P}$-projected and $\mathcal{Q}$-projected approximator classes have smaller covering numbers, i.e., $\forall r > 0$ we have*

$$\mathcal{N}_\infty(\mathcal{O}\mathcal{F}^{\mathrm{NE}}, r) \leq \mathcal{N}_\infty(\mathcal{F}^{\mathrm{NE}}, r),$$
$$\mathcal{N}_\infty(\mathcal{P}\mathcal{F}^{\mathrm{NE}}, r) \leq \mathcal{N}_\infty(\mathcal{F}^{\mathrm{NE}}, r),$$
$$\mathcal{N}_\infty(\mathcal{Q}\mathcal{F}^{(\mathrm{C})\mathrm{CE}}, r) \leq \mathcal{N}_\infty(\mathcal{F}^{(\mathrm{C})\mathrm{CE}}, r).$$

The proof is done by showing all the operators are contraction mappings. See Appendix B.3 for details.

Both the generalization bounds in Theorem 5.3 and the sample complexities in Theorem 5.4 decrease with the decrease of covering numbers $\mathcal{N}_\infty(\mathcal{F}, r)$. Thus, we can see from Theorem 5.5 that both PPE and OPI can improve the generalizability of NE approximators, and PE can improve the generalizability of CE and CCE approximators.

5.2. Benefits for Approximation

We then show the benefits of equivariance for approximation when the payoff distribution is invariant under permutation. The permutation-invariant distribution holds when the action is anonymous or indifferent or when we pre-train the equilibrium approximators using a manually designed distribution (Marris et al., 2022).

(C)CE Approximator The following theorem tells the benefit of permutation-equivariance in decreasing the exploitability of (C)CE approximators.

Theorem 5.6. *When the payoff distribution $\mathcal{D}$ is invariant under the permutation of payoffs, the $\mathcal{Q}$-projected (C)CE approximator has a smaller expected equilibrium approximation. Formally, for all $f^{(\mathrm{C})\mathrm{CE}} \in \mathcal{F}^{(\mathrm{C})\mathrm{CE}}$ and permutation-invariant distribution $\mathcal{D}$, we have*

$$\mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}(\mathcal{Q}f^{(\mathrm{C})\mathrm{CE}}(u), u)] \leq \mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}(f^{(\mathrm{C})\mathrm{CE}}(u), u)],$$

The proof is done by using the convexity of approximation. See Appendix B.4 for details. We can see from Theorem 5.6 that, when payoff distribution is invariant under permutation, it is beneficial to use equivariant architecture as the CE or CCE approximators.

NE Approximator As for NE approximators, we have similar results.

Theorem 5.7. *For bimatrix constant-sum games, when the payoff distribution $\mathcal{D}$ is invariant under the permutation of payoffs, then the $\mathcal{X}$-projected ($\mathcal{X} \in \{\mathcal{O}, \mathcal{P}\}$) NE approximator has a smaller expected exploitability. Formally, for all $f^{\mathrm{NE}} \in \mathcal{F}^{\mathrm{NE}}$ and permutation-invariant distribution $\mathcal{D}$ for bimatrix constant-sum games, we have*

$$
\mathbb{E}_{u \sim \mathcal{D}}\left[\sum_{i=1}^{n} \mathcal{E}_i((\mathcal{X}f^{\mathrm{NE}})(u), u)\right]
$$
$$
\leq \mathbb{E}_{u \sim \mathcal{D}}\left[\sum_{i=1}^{n} \mathcal{E}_i(f^{\mathrm{NE}}(u), u)\right].
$$

Theorem 5.8. *When the payoff distribution $\mathcal{D}$ is invariant under the permutation of payoffs, and f^{NE} satisfies OPI, then the $\mathcal{P}$-projected NE approximator has a smaller expected NE approximation. Formally, for all $f^{\mathrm{NE}} \in \mathcal{F}^{\mathrm{NE}}$ that is OPI and permutation-invariant distribution $\mathcal{D}$, we have*

$$
\mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}((\mathcal{P}f^{\mathrm{NE}})(u), u)] \leq \mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}(f^{\mathrm{NE}}(u), u)].
$$

Theorem 5.9. *For bimatrix games, when the payoff distribution $\mathcal{D}$ is invariant under the permutation of payoffs, and f^{NE} satisfies PPE, then the $\mathcal{O}$-projected NE approximator has a smaller expected NE approximation. Formally, for all $f^{\mathrm{NE}} \in \mathcal{F}^{\mathrm{NE}}$ that is PPE and permutation-invariant distribution $\mathcal{D}$ of bimatrix games, we have*

$$
\mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}((\mathcal{O}f^{\mathrm{NE}})(u), u)] \leq \mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}(f^{\mathrm{NE}}(u), u)].
$$

Theorem 5.8 and Theorem 5.9 tell that PPE and OPI approximators can achieve better approximation than ones with only PPE or OPI. Meanwhile, we can see from Theorem 5.7 that for bimatrix constant-sum games (such as zero-sum games), it can be preferred to introduce PPE or OPI to the architectures.

6. Theoretical Analysis of Limitations

As we discussed in Section 5, equivariant approximators enjoy better generalizability and better approximation sometimes. However, as we will show, they have some limitations regarding equilibrium selection and social welfare. Such limitations attribute to the limited representativeness caused by equivariance.

6.1. Equilibrium Selection

We first show that there may be equilibria points that equivariant approximators will never find. We illustrate such limitation in permutation-invariant games, which is defined as follows:

Definition 6.1 (Permutation-ρ-Invariant Game)**.** We say a game Γ_u is permutation-ρ-invariant, where $\rho = \circ_{i \in [n]} \rho_i$, if the payoff u is permutation-invariant with respect to ρ. That is, $\rho u = u$.

Permutation-ρ-invariance indicates that one cannot distinguish joint action a from ρa using only the payoff u. We'd like to provide an example to show more insight of permutation-ρ-invariant games:

Example 6.2. *For a 2-player game $\Gamma_u = (2, u = (u_1, u_2), \mathcal{A} = ([m_1], [m_2]))$, Let $\rho_i = (m_i, m_i - 1, \dots, 1)$ and $\rho = \rho_1 \circ \rho_2$. If one of the following conditions holds, then u is permutation-ρ-invariant:*

1. *u_1 and u_2 are symmetric and persymmetric (i.e., symmetric with respect to the northeast-to-southwest diagonal) squares.*

2. *Both u_1 and u_2 are centrosymmetric, i.e., $u_i(x, y) = u_i(m_1 + 1 - x, m_2 + 1 - y)$ for $i \in \{1, 2\}, x \in [m_1]$ and $y \in [m_2]$.*

For permutation $\rho = (\circ_{i \in [n]} \rho_i)$ and player $k \in [n]$, we denote the set of non-fixed actions of player k under ρ_k as

$$
V(\rho_k) := \{a_k | a_k \in \mathcal{A}_k, \rho_k(a_k) \neq a_k\}.
$$

Based on $V(\rho_k)$, we find some equilibria points of permutation-ρ-invariant games that any equivariant approximators will never find.

Theorem 6.3. *For a permutation-ρ-invariant game Γ_u. if there is a pure NE $a^* = (a_i^*)_{i \in [n]}$ and at least one player $k \in [n]$ such that $a_k^* \in V(\rho_k)$, then a^* will never be found by any NE approximator with both PPE and OPI. Besides, a^* (as a pure CE or CCE) will also never be found by any CE or CCE approximator with PE.*

We illustrate Theorem 6.3 by the following example:

Example 6.4. *Consider a bimatrix game with identity utility*

$$
u = \begin{bmatrix} \mathbf{1,1} & 0,0 \\ 0,0 & \mathbf{1,1} \end{bmatrix},
$$

There are two pure NE (bolded in the above matrix) and one mixed NE of $\sigma_1 = (0.5, 0.5)$ and $\sigma_2 = (0.5, 0.5)$. Let ρ_i be the unique permute function (except for identity function) of player $i \in [2]$, and $\rho = \rho_1 \circ \rho_2$. The game is permutation-ρ-invariant.

Case 1: *Let f be a permutation-equivariant CE or CCE approximator, and denote $\pi = f(u)$. We have*

$$\pi = f(u) \overset{(a)}{=} f(\rho u) \overset{(b)}{=} \rho f(u),$$

where (a) holds by permutation-ρ-invariance of u, and (b) holds by PE of f. Thus, we have $\pi_{1,1} = \pi_{2,2} \in [0, \frac{1}{2}]$ and $\pi_{1,2} = \pi_{2,1} \in [0, \frac{1}{2}]$. As a result, the two pure (C)CEs cannot be found.

Case 2: *Let f be a NE approximator that holds PPE and OPI. Denote $f(u) = (\sigma_1, \sigma_2)$, where $\sigma_1 = (p_1, 1 - p_1)$ and $\sigma_2 = (p_2, 1 - p_2)$. By PPE and OPI of f, we have*

$$f(u)_1 = (p_1, 1 - p_1) \overset{(a)}{=} f(\rho_1 \rho_2 u)_1 \overset{(b)}{=} \rho_1 f(\rho_2 u)_1$$
$$\overset{(c)}{=} \rho_1 f(u)_1 = (1 - p_1, p_1),$$

where (a) holds by permutaion-ρ-invariance of u, (b) holds by PPE of f, and (c) holds by OPI of f. As a result, the only NE that f could find is the mixed NE.

As we can see from the example and Theorem 6.3, the equivariance, while introducing inductive bias to the approximator architecture, is also a strong constraint. Such a constraint is why the equivariant approximators cannot find all the equilibria points.

6.2. Social Welfare

The social welfare of a joint strategy π is defined as the sum of all players' utilities, i.e.,

$$\mathrm{SW}(\pi, u) = \sum_{i \in [n]} u_i(\pi).$$

The equilibrium with higher social welfare is usually preferred (Marris et al., 2022).

To analyze the social welfare of equivariant approximators, we define the worst social welfare ratio as follows:

Definition 6.5. For any $N, M \geq 2$ and two NE (or CE/CCE) approximator classes $\mathcal{F}_1, \mathcal{F}_2$ that target on games with the number of players $n \leq N$ and $|\mathcal{A}_i| \leq M$, we define the worst social welfare ratio of $\mathcal{F}_1$ over $\mathcal{F}_2$ as:

$$\mathrm{SWR}_{\mathrm{N,M}}(\mathcal{F}_1, \mathcal{F}_2) := \inf_{\mathcal{D}} \frac{\max_{f_1 \in \mathcal{F}_1} \mathbb{E}_{u \sim \mathcal{D}} \mathrm{SW}(f_1(u), u)}{\max_{f_2 \in \mathcal{F}_2} \mathbb{E}_{u \sim \mathcal{D}} \mathrm{SW}(f_2(u), u)}.$$

$\mathrm{SWR}_{\mathrm{N,M}}(\mathcal{F}_1, \mathcal{F}_2)$ measures the relative representativeness of $\mathcal{F}_1$ over $\mathcal{F}_2$ in terms of social welfare. Based on that, we have the following result for equivariant CE and CCE approximator classes:

Theorem 6.6. *Given $N, M \geq 2$, let $\mathcal{F}_{\mathrm{PE}}^{\mathrm{(C)CE}}$ be the function class (target on games with the number of players $n \leq N$ and $|\mathcal{A}_i| \leq M$) of all the (C)CE approximators with*

PE. *Denote by $\mathcal{F}_{\mathrm{general}}^{\mathrm{(C)CE}}$ the function class of all the (C)CE approximators. Then we have*

$$\mathrm{SWR}_{\mathrm{N,M}}(\mathcal{F}_{\mathrm{PE}}^{\mathrm{(C)CE}}, \mathcal{F}_{\mathrm{general}}^{\mathrm{(C)CE}}) = 1.$$

Theorem 6.6 tells that, while the permutation-equivariant (C)CE approximator class may not be able to find all the (C)CE in a game, it can keep the social welfare of the output solutions.

However, when considering equivariant NE approximators, we have the following negative result:

Theorem 6.7. *Given $N, M \geq 2$, let $\mathcal{F}_{\mathrm{OPI}}^{\mathrm{NE}}, \mathcal{F}_{\mathrm{PPE}}^{\mathrm{NE}}$ and $\mathcal{F}_{\mathrm{both}}^{\mathrm{NE}}$ be the function classes (target on games with number of players $n \leq N$ and $|\mathcal{A}_i| \leq M$) of all the NE approximators with OPI, PPE and both. Denote the function class of all the NE approximators as $\mathcal{F}_{\mathrm{general}}^{\mathrm{NE}}$. Then we have*

$$\mathrm{SWR}_{\mathrm{N,M}}(\mathcal{F}_{\mathrm{OPI}}^{\mathrm{NE}}, \mathcal{F}_{\mathrm{general}}^{\mathrm{NE}}) = \frac{1}{M^{N-1}}, \qquad (1)$$

$$\mathrm{SWR}_{\mathrm{N,M}}(\mathcal{F}_{\mathrm{PPE}}^{\mathrm{NE}}, \mathcal{F}_{\mathrm{general}}^{\mathrm{NE}}) \leq \frac{1}{M}, \qquad (2)$$

$$\mathrm{SWR}_{\mathrm{N,M}}(\mathcal{F}_{\mathrm{both}}^{\mathrm{NE}}, \mathcal{F}_{\mathrm{general}}^{\mathrm{NE}}) = \frac{1}{M^{N-1}}. \qquad (3)$$

Additionally, when $M \geq 3$, denote by $\widetilde{\mathcal{F}}_{\mathrm{both}}^{\mathrm{NE}}$ the function class of all the NE oracles (functions that always output exact NE solutions of the input games) with both PPE and OPI, and by $\widetilde{\mathcal{F}}_{\mathrm{general}}^{\mathrm{NE}}$ the function class of all the NE oracles. Then we have

$$\mathrm{SWR}_{\mathrm{N,M}}(\widetilde{\mathcal{F}}_{\mathrm{both}}^{\mathrm{NE}}, \widetilde{\mathcal{F}}_{\mathrm{general}}^{\mathrm{NE}}) = 0. \qquad (4)$$

The proof is done by construction (See Appendix C.3 for details). As an illustration of Equation (4), consider a bimatrix game with the following payoff:

$$u = \begin{bmatrix} 1, 1 & 0, 0 & 0, 1/2 + \varepsilon \\ 0, 0 & 1, 1 & 0, 1/2 + \varepsilon \\ 1/2 + \varepsilon, 0 & 1/2 + \varepsilon, 0 & \varepsilon, \varepsilon \end{bmatrix},$$

for $\epsilon \in (0, 1/2)$. The maximum NE (the upper-left corner of u) social welfare is 2, which can be found by at least one NE oracle in $\widetilde{\mathcal{F}}_{\mathrm{general}}^{\mathrm{NE}}$. However, the only NE (the lower-right corner of u) that the NE oracles in $\widetilde{\mathcal{F}}_{\mathrm{both}}^{\mathrm{NE}}$ could find only has a social welfare of 2ϵ. As a result,

$$\mathrm{SWR}_{2,3}(\widetilde{\mathcal{F}}_{\mathrm{both}}^{\mathrm{NE}}, \widetilde{\mathcal{F}}_{\mathrm{general}}^{\mathrm{NE}}) \leq \frac{2\epsilon}{2} = \epsilon,$$

which goes to zero as $\epsilon \to 0$. Recall that we always have $\mathrm{SWR}_{N,M} \geq 0$, thus Equation (4) holds when $N = 2$ and $M = 3$.

Theorem 6.7 tells that equivariant NE approximators may lose some social welfare while enjoying better generalizability. Such a result inspires us to balance generalizability and social welfare when designing the NE approximator architecture.

7. Conclusion and Future Work

In this paper, we theoretically analyze the benefits and limitations of equivariant equilibrium approximators, including player-permutation-equivariant (PPE) and opponent-permutation-invariant (OPI) NE approximator, and permutation-equivariant (PE) CE and CCE approximators. For the benefits, we first show that these equivariant approximators enjoy better generalizability. To get the result, we derive the generalization bounds and sample complexities based on covering numbers, and then we prove that the symmetric approximators have lower covering numbers. We then show that the equivariant approximators can decrease the exploitability when the payoff distribution is invariant under permutation. For the limitations, we find the equivariant approximators may fail to find some equilibria points due to their limited representativeness caused by equivariance. Besides, while equivariant (C)CE approximators can keep the social welfare, the equivariant NE approximators reach a small worst social welfare ratio compared to the general approximators. Such a result indicates that equivariance may reduce social welfare; therefore, we'd better balance the generalizability and social welfare when we design the architectures of NE approximators.

As for future work, since in our paper we assume the training and testing payoff distribution are the same, an interesting topic is to study the benefits of equivariant approximators under the payoff distribution shift. Moreover, since we consider fixed and discrete action space, another interesting future direction is to analyze the benefits of equivariant approximators in varying or continuous action space.

Acknowledgements

References

Bartlett, P. L. and Mendelson, S. Rademacher and gaussian complexities: Risk bounds and structural results. *Journal of Machine Learning Research*, 3(Nov):463–482, 2002.

Cesa-Bianchi, N. and Lugosi, G. *Prediction, learning, and games*. Cambridge university press, 2006.

Chen, X., Deng, X., and Teng, S.-H. Settling the complexity of computing two-player Nash equilibria. *Journal of the ACM (JACM)*, 56(3):1–57, 2009.

Daskalakis, C., Goldberg, P. W., and Papadimitriou, C. H. The complexity of computing a Nash equilibrium. *SIAM Journal on Computing*, 39(1):195–259, 2009.

Deligkas, A., Fasoulakis, M., and Markakis, E. A polynomial-time algorithm for 1/3-approximate Nash equilibria in bimatrix games. In *30th Annual European Symposium on Algorithms, ESA*, 2022.

Deng, X., Li, N., Mguni, D., Wang, J., and Yang, Y. On the complexity of computing markov perfect equilibrium in general-sum stochastic games. *National Science Review*, 10(1):nwac256, 2023.

Duan, Z., Tang, J., Yin, Y., Feng, Z., Yan, X., Zaheer, M., and Deng, X. A context-integrated transformer-based neural network for auction design. In *International Conference on Machine Learning*, pp. 5609–5626. PMLR, 2022.

Duan, Z., Huang, W., Zhang, D., Du, Y., Wang, J., Yang, Y., and Deng, X. Is Nash equilibrium approximator learnable? In *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems*, pp. 233–241, 2023.

Dütting, P., Feng, Z., Narasimhan, H., Parkes, D., and Ravindranath, S. S. Optimal auctions through deep learning. In *International Conference on Machine Learning*, pp. 1706–1715. PMLR, 2019.

Elesedy, B. and Zaidi, S. Provably strict generalisation benefit for equivariant models. In *International Conference on Machine Learning*, pp. 2959–2969. PMLR, 2021.

Feng, X., Slumbers, O., Wan, Z., Liu, B., McAleer, S., Wen, Y., Wang, J., and Yang, Y. Neural auto-curricula in two-player zero-sum games. *Advances in Neural Information Processing Systems*, 34:3504–3517, 2021.

Fu, H. and Lin, T. Learning utilities and equilibria in non-truthful auctions. *Advances in Neural Information Processing Systems*, 33:14231–14242, 2020.

Fudenberg, D., Drew, F., Levine, D. K., and Levine, D. K. *The theory of learning in games*, volume 2. MIT press, 1998.

Goktas, D. and Greenwald, A. Exploitability minimization in games and beyond. In *Advances in Neural Information Processing Systems*, 2022.

Greenwald, A., Hall, K., Serrano, R., et al. Correlated Q-learning. In *ICML*, volume 3, pp. 242–249, 2003.

Hartford, J. S., Wright, J. R., and Leyton-Brown, K. Deep learning for predicting human strategic behavior. *Advances in neural information processing systems*, 29, 2016.

Hu, J. and Wellman, M. P. Nash Q-learning for general-sum stochastic games. *Journal of machine learning research*, 4(Nov):1039–1069, 2003.

Ivanov, D., Safiulin, I., Filippov, I., and Balabaeva, K. Optimal-er auctions through attention. In *Advances in Neural Information Processing Systems*, 2022.

Jin, C., Liu, Q., Wang, Y., and Yu, T. V-learning – a simple, efficient, decentralized algorithm for multiagent RL. In *ICLR 2022 Workshop on Gamification and Multiagent Solutions*, 2022.

Lanctot, M., Zambaldi, V., Gruslys, A., Lazaridou, A., Tuyls, K., Pérolat, J., Silver, D., and Graepel, T. A unified game-theoretic approach to multiagent reinforcement learning. *Advances in neural information processing systems*, 30, 2017.

Ling, C., Fang, F., and Kolter, J. Z. What game are we playing? End-to-end learning in normal and extensive form games. In *IJCAI*, pp. 396–402, 2018.

Liu, S., Lanctot, M., Marris, L., and Heess, N. Simplex neural population learning: Any-mixture bayes-optimality in symmetric zero-sum games. In *International Conference on Machine Learning, ICML*, 2022.

Lockhart, E., Lanctot, M., Pérolat, J., Lespiau, J.-B., Morrill, D., Timbers, F., and Tuyls, K. Computing approximate equilibria in sequential adversarial games by exploitability descent. In Kraus, S. (ed.), *IJCAI*, pp. 464–470. ijcai.org, 2019.

Marris, L., Muller, P., Lanctot, M., Tuyls, K., and Graepel, T. Multi-agent training beyond zero-sum with correlated equilibrium meta-solvers. In *International Conference on Machine Learning*, pp. 7480–7491. PMLR, 2021.

Marris, L., Gemp, I., Anthony, T., Tacchetti, A., Liu, S., and Tuyls, K. Turbocharging solution concepts: Solving NEs, CEs and CCEs with neural equilibrium solvers. In *Advances in Neural Information Processing Systems*, 2022.

Nash, J. F. et al. Equilibrium points in n-person games. *Proceedings of the national academy of sciences*, 36(1): 48–49, 1950.

Nekipelov, D., Syrgkanis, V., and Tardos, E. Econometrics for learning agents. In *Proceedings of the sixteenth acm conference on economics and computation*, pp. 1–18, 2015.

Nikaidô, H. and Isoda, K. Note on non-cooperative convex games. *Pacific Journal of Mathematics*, 5(S1):807–815, 1955.

Qin, T., He, F., Shi, D., Huang, W., and Tao, D. Benefits of permutation-equivariance in auction mechanisms. In *Advances in Neural Information Processing Systems*, 2022.

Rahme, J., Jelassi, S., Bruna, J., and Weinberg, S. M. A permutation-equivariant neural network architecture for auction design. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021.

Schulz-Mirbach, H. Constructing invariant features by averaging techniques. In *Proceedings of the 12th IAPR International Conference on Pattern Recognition, Vol. 3-Conference C: Signal Processing (Cat. No. 94CH3440-5)*, volume 2, pp. 387–390. IEEE, 1994.

Sessa, P. G., Bogunovic, I., Krause, A., and Kamgarpour, M. Contextual games: Multi-agent learning with side information. *Advances in Neural Information Processing Systems*, 33:21912–21922, 2020.

Shalev-Shwartz, S. and Ben-David, S. *Understanding machine learning: From theory to algorithms*. Cambridge university press, 2014.

Shoham, Y. and Leyton-Brown, K. *Multiagent systems: Algorithmic, game-theoretic, and logical foundations*. Cambridge University Press, 2008.

Tsaknakis, H. and Spirakis, P. G. An optimization approach for approximate nash equilibria. *Internet Mathematics*, 5 (4):365–382, 2008.

Wu, D. and Lisser, A. Using CNN for solving two-player zero-sum games. *Expert Systems with Applications*, pp. 117545, 2022.

Wu, D. and Lisser, A. CCGnet: A deep learning approach to predict Nash equilibrium of chance-constrained games. *Information Sciences*, 2023.

A. Omitted Proofs in Section 4

A.1. Useful Lemma

We first introduce a lemma, which will be frequently used in the following proofs.

Lemma A.1. $\forall i, j \in [n], \rho_i \in \mathcal{G}_i$ we have $(\rho_i u)_j(\sigma_i, \sigma_{-i}) = u_j(\rho_i^{-1}\sigma_i, \sigma_{-i})$ and $(\rho_i u)_j(\pi) = u_j(\rho_i^{-1}\pi)$.

Proof. Define $\widehat{a}_i := \rho_i^{-1} a_i$. For product strategy $\sigma = (\sigma_i)_{i \in [n]}$,

$$
\begin{aligned}
(\rho_i u)_j(\sigma_i, \sigma_{-i}) &= \sum_{a_i \in \mathcal{A}_i} \sum_{a_{-i} \in \mathcal{A}_{-i}} (\rho_i u)_j(a_i, a_{-i}) \cdot \sigma_i(a_i) \cdot \sigma_{-i}(a_{-i}) \\
&= \sum_{a_i \in \mathcal{A}_i} \sum_{a_{-i} \in \mathcal{A}_{-i}} u_j(\rho_i^{-1}a_i, a_{-i}) \cdot \sigma_i(a_i) \cdot \sigma_{-i}(a_{-i}) \\
&= \sum_{a_i \in \mathcal{A}_i} \sum_{a_{-i} \in \mathcal{A}_{-i}} u_j(\rho_i^{-1}a_i, a_{-i}) \cdot (\rho_i^{-1}\sigma_i)(\rho_i^{-1}a_i) \cdot \sigma_{-i}(a_{-i}) \\
&= \sum_{\widehat{a}_i \in \mathcal{A}_i} \sum_{a_{-i} \in \mathcal{A}_{-i}} u_j(\widehat{a}_i, a_{-i}) \cdot (\rho_i^{-1}\sigma_i)(\widehat{a}_i) \cdot \sigma_{-i}(a_{-i}) \\
&= u_j(\rho_i^{-1}\sigma_i, \sigma_{-i}).
\end{aligned}
$$

For joint strategy π,

$$
\begin{aligned}
(\rho_i u)_j(\pi) &= \sum_{a_i \in \mathcal{A}_i} \sum_{a_{-i} \in \mathcal{A}_{-i}} (\rho_i u_j)(a_i, a_{-i}) \cdot \pi(a_i, a_{-i}) \\
&= \sum_{a_i \in \mathcal{A}_i} \sum_{a_{-i} \in \mathcal{A}_{-i}} u_j(\rho_i^{-1}a_i, a_{-i}) \cdot \pi(a_i, a_{-i}) \\
&= \sum_{a_i \in \mathcal{A}_i} \sum_{a_{-i} \in \mathcal{A}_{-i}} u_j(\rho_i^{-1}a_i, a_{-i}) \cdot (\rho_i^{-1}\pi)(\rho_i^{-1}a_i, a_{-i}) \\
&= \sum_{\widehat{a}_i \in \mathcal{A}_i} \sum_{a_{-i} \in \mathcal{A}_{-i}} u_j(\widehat{a}_i, a_{-i}) \cdot (\rho_i^{-1}\pi)(\widehat{a}_i, a_{-i}) \\
&= u_j(\rho_i^{-1}\pi).
\end{aligned}
$$

$\square$

A.2. Proof of Lemma 4.2

Lemma 4.2. *In a normal-form game $\Gamma_u = (n, u, \mathcal{A})$, for arbitrary player $i \in [n]$ and any (ϵ-)NE strategy $\sigma = (\sigma_i, \sigma_{-i})$, $\rho_i \sigma = (\rho_i \sigma_i, \sigma_{-i})$ is also an (ϵ-)NE for the ρ_i-permuted game $\Gamma_{\rho_i u}$.*

Proof. For player i, we have

$$
\begin{aligned}
\mathcal{E}_i(\rho_i \sigma, \rho_i u) &= \max_{a_i \in \mathcal{A}_i} \rho_i u_i(a_i, \rho_i \sigma_{-i}) - \rho_i u_i(\rho_i \sigma) = \max_{a_i \in \mathcal{A}_i} \rho_i u_i(a_i, \sigma_{-i}) - \rho_i u_i(\rho_i \sigma_i, \sigma_{-i}) \\
&= \max_{a_i \in \mathcal{A}_i} u_i(\rho_i^{-1}a_i, \sigma_{-i}) - u_i(\rho_i^{-1}\rho_i \sigma_i, \sigma_{-i}) \stackrel{(a)}{=} \max_{a_i \in \mathcal{A}_i} u_i(a_i, \sigma_{-i}) - u_i(\sigma_i, \sigma_{-i}) = \mathcal{E}_i(\sigma, u),
\end{aligned}
$$

where (a) holds since ρ_i is a bijection on $\mathcal{A}_i$. For player $j \neq i$, we have

$$
\begin{aligned}
\mathcal{E}_j(\rho_i \sigma, \rho_i u) &= \max_{a_j \in \mathcal{A}} \rho_i u_j(a_j, \rho_i \sigma_{-j}) - \rho_i u_j(\rho_i \sigma) = \max_{a_j \in \mathcal{A}_j} u_j(a_j, \rho_i^{-1}\rho_i \sigma_{-j}) - u_j(\rho_i^{-1}\rho_i \sigma) \\
&= \max_{a_j \in \mathcal{A}_j} u_j(a_j, \sigma_{-j}) - u_j(\sigma) = \mathcal{E}_j(\sigma, u).
\end{aligned}
$$

From above, we have $\mathcal{E}(\rho_i \sigma, \rho_i u) = \mathcal{E}(\sigma, u)$, thus if σ is a ε-NE of Γ_u, then $\rho_i \sigma$ must be a ε-NE of $\Gamma_{\rho_i u}$. $\square$

A.3. Proof of Lemma 4.5

Lemma 4.5. *In a normal-form game $\Gamma_u = (n, u, \mathcal{A})$, for an arbitrary player $i \in [n]$ and any $(\varepsilon\text{-})CE$ or $(\epsilon\text{-})CCE$ strategy π, $\rho_i \pi$ is also an $(\varepsilon\text{-})CE$ or $(\epsilon\text{-})CCE$ for the ρ_i-permuted game $\Gamma_{\rho_i u}$.*

CCE   For player i, we have

$$
\begin{aligned}
\mathcal{E}_i(\rho_i \pi, \rho_i u) &= \max_{a_i \in \mathcal{A}_i} (\rho_i u_i)(a_i, (\rho_i \pi)_{-i}) - (\rho_i u_i)(\rho_i \pi_i) \\
&= \max_{a_i \in \mathcal{A}_i} (\rho_i u_i)(a_i, (\rho_i \pi)_{-i}) - u_i(\rho_i^{-1} \rho_i \pi_i) \\
&= \max_{a_i \in \mathcal{A}_i} (\rho_i u_i)(a_i, (\rho_i \pi)_{-i}) - u_i(\pi_i) \\
&= \max_{a_i \in \mathcal{A}_i} \sum_{b \in \mathcal{A}} (\rho_i u_i)(a_i, b_{-i}) \cdot (\rho_i \pi)(b) - u_i(\pi_i) \\
&= \max_{a_i \in \mathcal{A}_i} \sum_{b_i \in \mathcal{A}_i} \sum_{b_{-i} \in \mathcal{A}_{-i}} u_i(\rho_i^{-1} a_i, b_{-i}) \cdot \pi(\rho_i^{-1} b_i, b_{-i}) - u_i(\pi_i) \\
&= \max_{a_i \in \mathcal{A}_i} \sum_{b_i \in \mathcal{A}_i} \sum_{b_{-i} \in \mathcal{A}_{-i}} u_i(a_i, b_{-i}) \cdot \pi(b_i, b_{-i}) - u_i(\pi_i) \qquad , \rho_i \text{ is a bijection on } \mathcal{A}_i. \\
&= \mathcal{E}_i(\pi, u).
\end{aligned}
$$

For player $j \neq i$, we have

$$
\begin{aligned}
\mathcal{E}_j(\rho_i \pi, \rho_i u) &= \max_{a_j \in \mathcal{A}_j} (\rho_i u_j)(a_j, (\rho_i \pi)_{-j}) - (\rho_i u_j)(\rho_i \pi_j) \\
&= \max_{a_j \in \mathcal{A}_j} (\rho_i u_j)(a_j, (\rho_i \pi)_{-j}) - u_j(\rho_i^{-1} \rho_i \pi_j) \\
&= \max_{a_j \in \mathcal{A}_j} (\rho_i u_j)(a_j, (\rho_i \pi)_{-j}) - u_j(\pi_j) \\
&= \max_{a_j \in \mathcal{A}_j} \sum_{b \in \mathcal{A}} (\rho_i u_j)(a_j, b_{-j}) \cdot (\rho_i \pi)(b) - u_j(\pi_j) \\
&= \max_{a_j \in \mathcal{A}_j} \sum_{b_i \in \mathcal{A}_i} \sum_{b_{-i} \in \mathcal{A}_{-i}} u_j(a_j, (b_{-j})_{-i}, \rho_i^{-1} b_i) \cdot \pi(\rho_i^{-1} b_i, b_{-i}) - u_j(\pi_j) \\
&= \max_{a_j \in \mathcal{A}_j} \sum_{b_i \in \mathcal{A}_i} \sum_{b_{-i} \in \mathcal{A}_{-i}} u_j(a_j, (b_{-j})_{-i}, b_i) \cdot \pi(b_i, b_{-i}) - u_j(\pi_j) \qquad , \rho_i \text{ is a bijection on } \mathcal{A}_i. \\
&= \mathcal{E}_j(\pi, u).
\end{aligned}
$$

Thus, we have $\mathcal{E}(\rho_i \pi, \rho_i u) = \mathcal{E}(\pi, u)$. Thus, if π is a ε-CCE of Γ_u, then $\rho_i \pi$ must be a ε-CCE of $\Gamma_{\rho_i u}$.

CE   For player $j \neq i$, we have

$$
\begin{aligned}
\mathcal{E}_j^{\mathrm{CE}}(\rho_i \pi, \rho_i u) &= \max_{\phi_j : \mathcal{A}_j \to \mathcal{A}_j} \sum_{a \in \mathcal{A}} (\rho_i \pi)(a) \cdot (\rho_i u_j)(\phi_j(a_j), a_{-j}) - (\rho_i u_j)(\rho_i \pi) \\
&= \max_{\phi_j : \mathcal{A}_j \to \mathcal{A}_j} \sum_{a \in \mathcal{A}} \pi(\rho_i^{-1} a_i, a_{-i}) \cdot u_j(\phi_j(a_j), a_{-i,j}, \rho_i^{-1} a_i) - u_j(\pi) \\
&= \max_{\phi_j : \mathcal{A}_j \to \mathcal{A}_j} \sum_{a \in \mathcal{A}} \pi(a_i, a_{-i}) \cdot u_j(\phi_j(a_j), a_{-i,j}, a_i) - u_j(\pi) \qquad , \rho_i \text{ is a bijection on } \mathcal{A}_i. \\
&= \mathcal{E}_j^{\mathrm{CE}}(\pi, u).
\end{aligned}
$$

For player i, we define operator $\bar{\rho}_i$ as $(\bar{\rho}_i \phi_i)(a_i) = \rho_i^{-1} \phi_i(\rho_i a_i)$. We can verify that $\bar{\rho}_i$ is a bijection on $\{\phi_i : \mathcal{A}_i \to \mathcal{A}_i\}$, because $\bar{\ }$ is a homomorphism in the sense that $\overline{\rho_i^1} \circ \overline{\rho_i^2} = \overline{\rho_i^2 \rho_i^1}$ and $\bar{\ }$ maps the identity mapping of $\mathcal{A}_i$ to the identity mapping of $\{\mathcal{A}_i \to \mathcal{A}_i\}$. Specifically,

$$
\overline{\rho_i^1} \circ \overline{\rho_i^2} \phi_i(a_i) = (\rho_i^1)^{-1} (\overline{\rho_i^2} \phi_i)(\rho_i^1 a_i) = (\rho_i^1)^{-1} (\rho_i^2)^{-1} \phi_i(\rho_i^2 \rho_i^1 a_i) = \overline{\rho_i^2 \rho_i^1} \phi_i(a_i),
$$

and

$$\overline{e_i}\phi_i(a_i) = e_i^{-1}\phi_i(e_i a_i) = \phi_i(a_i).$$

Based on $\bar{\rho}_i$, we have

$$
\begin{aligned}
&\mathcal{E}_i^{\mathrm{CE}}(\rho_i\pi, \rho_i u)\\
&= \max_{\phi_i:\mathcal{A}_i\to\mathcal{A}_i} \sum_{a\in\mathcal{A}} (\rho_i\pi)(a) \cdot (\rho_i u_i)(\phi_i(a_i), a_{-i}) - u_i(\pi)\\
&= \max_{\phi_i:\mathcal{A}_i\to\mathcal{A}_i} \sum_{a\in\mathcal{A}} \pi(\rho_i^{-1}a_i, a_{-i}) u_i(\rho_i^{-1}\phi_i(a_i), a_{-i}) - u_i(\pi)\\
&= \max_{\phi_i:\mathcal{A}_i\to\mathcal{A}_i} \sum_{a\in\mathcal{A}} \pi(\rho_i^{-1}a_i, a_{-i}) u_i(\rho_i^{-1}\phi_i(\rho_i(\rho_i^{-1}a_i)), a_{-i}) - u_i(\pi)\\
&= \max_{\phi_i:\mathcal{A}_i\to\mathcal{A}_i} \sum_{a\in\mathcal{A}} \pi(a_i, a_{-i}) u_i(\rho_i^{-1}\phi_i(\rho_i a_i), a_{-i}) - u_i(\pi) && , \rho_i \text{ is a bijection on } \mathcal{A}_i.\\
&= \max_{\phi_i:\mathcal{A}_i\to\mathcal{A}_i} \sum_{a\in\mathcal{A}} \pi(a_i, a_{-i}) u_i((\bar{\rho}_i\phi_i)(a_i), a_{-i}) - u_i(\pi)\\
&= \max_{\phi_i:\mathcal{A}_i\to\mathcal{A}_i} \sum_{a\in\mathcal{A}} \pi(a_i, a_{-i}) u_i(\phi_i(a_i), a_{-i}) - u_i(\pi) && , \bar{\rho}_i \text{ is a bijection on } \{\mathcal{A}_i\to\mathcal{A}_i\}.\\
&= \mathcal{E}_i^{\mathrm{CE}}(\pi, u).
\end{aligned}
$$

Thus, we have $\mathcal{E}(\rho_i\pi, \rho_i u) = \mathcal{E}(\pi, u)$, thus if π is a ε-CE of Γ_u, then $\rho_i\pi$ must be a ε-CE of $\Gamma_{\rho_i u}$.

A.4. Proof of Lemma 4.7 to Lemma 4.9

Lemma 4.7. $\mathcal{O}_i f^{\mathrm{NE}}$ *is i-PI and $\mathcal{P}_i f^{\mathrm{NE}}$ is i-PE. Specially, if f^{NE} is already i-PI or i-PE, then we have $\mathcal{O}_i f^{\mathrm{NE}} = f^{\mathrm{NE}}$ or $\mathcal{P}_i f^{\mathrm{NE}} = f^{\mathrm{NE}}$, respectively.*

Proof. $\forall j \neq i, \rho_0 \in \mathcal{G}_i$, for operator $\mathcal{O}_i$ we have

$$(\mathcal{O}_i f^{\mathrm{NE}})(\rho_0 u)_j = \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i\in\mathcal{G}_i} f^{\mathrm{NE}}(\rho_i\rho_0 u)_j \overset{(a)}{=} \frac{1}{|\mathcal{A}_i|!} \sum_{\widehat{\rho}_i\in\mathcal{G}_i} f^{\mathrm{NE}}(\widehat{\rho}_i u)_j = (\mathcal{O}_i f^{\mathrm{NE}})(u)_j,$$

where in (a) we define $\widehat{\rho}_i = \rho_i\rho_0$, and (a) holds since ρ_0 is a bijection on $\mathcal{G}_i$. As a result, $\mathcal{O}_i f^{\mathrm{NE}}$ is i-PI.

For operator $\mathcal{P}_i$ we have

$$
\begin{aligned}
(\mathcal{P}_i f^{\mathrm{NE}})(\rho_0 u)_i &= \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i\in\mathcal{G}_i} \rho_i^{-1} f^{\mathrm{NE}}(\rho_i\rho_0 u)_j = \rho_0 \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i\in\mathcal{G}_i} \rho_0^{-1}\rho_i^{-1} f^{\mathrm{NE}}(\rho_i\rho_0 u)_j\\
&= \rho_0 \frac{1}{|\mathcal{A}_i|!} \sum_{\widehat{\rho}_i\in\mathcal{G}_i} \widehat{\rho}_i^{-1} f^{\mathrm{NE}}(\widehat{\rho}_i u)_j = \rho_0(\mathcal{P}_i f^{\mathrm{NE}})(u)_i,
\end{aligned}
$$

therefore $\mathcal{P}_i f^{\mathrm{NE}}$ is i-PE.

If f^{NE} is already i-PI, $\forall j \neq i$ we have

$$\mathcal{O}_i f^{\mathrm{NE}}(u)_j = \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i\in\mathcal{G}_i} f^{\mathrm{NE}}(\rho_i u)_j = \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i\in\mathcal{G}_i} f^{\mathrm{NE}}(u)_j = f^{\mathrm{NE}}(u)_j,$$

and $\mathcal{O}_i f^{\mathrm{NE}}(u)_i = f^{\mathrm{NE}}(u)_i$ according to definition of $\mathcal{O}_i$. Therefore, $\mathcal{O}_i f^{\mathrm{NE}} = f^{\mathrm{NE}}$ for i-PI f^{NE}.

If f^{NE} is already i-PE, we have

$$\mathcal{P}_i f^{\mathrm{NE}}(u)_i = \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i\in\mathcal{G}_i} \rho_i^{-1} f^{\mathrm{NE}}(\rho_i u)_i = \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i\in\mathcal{G}_i} \rho_i^{-1}\rho_i f^{\mathrm{NE}}(u)_i = \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i\in\mathcal{G}_i} f^{\mathrm{NE}}(u)_i = f^{\mathrm{NE}}(u)_i,$$

and $\forall j \neq i, \mathcal{P}_i f^{\mathrm{NE}}(u)_j = f^{\mathrm{NE}}(u)_j$ according to definition of $\mathcal{P}_i$. Therefore, $\mathcal{P}_i f^{\mathrm{NE}} = f^{\mathrm{NE}}$ for i-PE f^{NE}.

$\square$

Lemma 4.8. $\mathcal{O} f^{\mathrm{NE}}$ *is OPI and* $\mathcal{P} f^{\mathrm{NE}}$ *is PPE. If* f^{NE} *is already OPI or PPE, we have* $\mathcal{O} f^{\mathrm{NE}} = f^{\mathrm{NE}}$ *or* $\mathcal{P} f^{\mathrm{NE}} = f^{\mathrm{NE}}$, *respectively.*

Proof. A direct inference from Lemma 4.7 $\square$

Lemma 4.9. $\mathcal{Q}_i f^{(C)CE}$ *is* i-PE *and* $\mathcal{Q} f^{(C)CE}$ *is PE. Specifically, If* $f^{(C)CE}$ *is already* i-PE *or PE, then we have* $\mathcal{Q}_i f^{(C)CE} = f^{(C)CE}$ *or* $\mathcal{Q} f^{(C)CE} = f^{(C)CE}$, *respectively.*

Proof. $\forall \rho_0 \in \mathcal{G}_i$, we have

$$(\mathcal{Q}_i f^{(\mathrm{C})\mathrm{CE}})(\rho_0 u) = \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} \rho_i^{-1} f^{(\mathrm{C})\mathrm{CE}}(\rho_i \rho_0 u) = \rho_0 \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} \rho_0^{-1} \rho_i^{-1} f^{(\mathrm{C})\mathrm{CE}}(\rho_i \rho_0 u)$$

$$= \rho_0 \frac{1}{|\mathcal{A}_i|!} \sum_{\widehat{\rho}_i \in \mathcal{G}_i} \widehat{\rho}_i^{-1} f^{(\mathrm{C})\mathrm{CE}}(\widehat{\rho}_i u) = \rho_0 (\mathcal{Q}_i f^{(\mathrm{C})\mathrm{CE}})(u).$$

If $f^{(\mathrm{C})\mathrm{CE}}$ is already i-PE, we have

$$\mathcal{Q}_i f^{(\mathrm{C})\mathrm{CE}}(u) = \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} \rho_i^{-1} f^{(\mathrm{C})\mathrm{CE}}(\rho_i u) = \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} \rho_i^{-1} \rho_i f^{(\mathrm{C})\mathrm{CE}}(u) = \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} f^{(\mathrm{C})\mathrm{CE}}(u) = f^{(\mathrm{C})\mathrm{CE}}(u).$$

$\square$

A.5. Proof of Lemma 4.11

Lemma 4.11. *Denote* $\mathcal{X}$ *as an idempotent operator, i.e.* $\mathcal{X}^2 = \mathcal{X}$ *(e.g.* $\mathcal{O}, \mathcal{P}$ *or* $\mathcal{Q}$*). For function class* $\mathcal{F}$ *of NE, CE or CCE approximator, let* $\mathcal{F}_\mathcal{X}$ *be any subset of* $\mathcal{F}$ *that is closed under* $\mathcal{X}$, *then* $\mathcal{X} \mathcal{F}_\mathcal{X}$ *is the largest subset of* $\mathcal{F}_\mathcal{X}$ *that is invariant under* $\mathcal{X}$.

Proof. We prove the three claims below.

1. $\mathcal{X} \mathcal{F}_\mathcal{X} \subseteq \mathcal{F}_\mathcal{X}$.

2. $\mathcal{X}^2 \mathcal{F}_\mathcal{X} = \mathcal{X} \mathcal{F}_\mathcal{X}$.

3. If $\mathcal{X} \mathcal{Y} = \mathcal{Y} \subseteq \mathcal{F}_\mathcal{X}$, then $\mathcal{Y} \subseteq \mathcal{X} \mathcal{F}_\mathcal{X}$.

The first claim holds because $\mathcal{F}_\mathcal{X}$ is closed under $\mathcal{X}$, and the second claim holds because $\mathcal{X}$ is idempotent. For the third claim, from $\mathcal{Y} \subseteq \mathcal{F}_\mathcal{X}$ we know $\mathcal{X} \mathcal{Y} \subseteq \mathcal{X} \mathcal{F}_\mathcal{X}$, then $\mathcal{Y} = \mathcal{X} \mathcal{Y} \subseteq \mathcal{X} \mathcal{F}_\mathcal{X}$.

We immediately know $\mathcal{X} \mathcal{F}_\mathcal{X}$ is the largest subset of $\mathcal{F}_\mathcal{X}$ that is invariant under $\mathcal{X}$. $\square$

B. Omitted Proofs in Section 5

B.1. Proof of Theorem 5.3

Theorem 5.3. *[Generalization bound] For function class* $\mathcal{F}$ *of NE, CE or CCE approximator, with probability at least* $1 - \delta$ *over draw of the training set* S *(with size* m*) from payoff distribution* $\mathcal{D}$, *for all approximator* $f \in \mathcal{F}$ *we have*

$$\mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}(f(u), u)] - \frac{1}{m} \sum_{u \in S} \mathcal{E}(f(u), u)$$

$$\leq 2 \cdot \inf_{r>0} \left\{ \sqrt{\frac{2 \ln \mathcal{N}_\infty(\mathcal{F}, r)}{m}} + Lr \right\} + 4 \sqrt{\frac{2 \ln(4/\delta)}{m}},$$

where $L = 2n$ for NE approximator, and $L = 2$ for CE and CCE approximators.

Some of the proof techniques come from Dütting et al. (2019) and Duan et al. (2023). We first introduce some useful lemmas. Denote $\ell : \mathcal{F} \times \mathcal{U} \rightarrow \mathbb{R}$ as the loss function (such as $\ell(f, u) := \mathcal{E}(f(u), u)$). We measure the capacity of the composite function class $\ell \circ \mathcal{F}$ using the empirical Rademacher complexity (Bartlett & Mendelson, 2002) on the training set S, which is defined as:

$$\mathcal{R}_S(\ell \circ \mathcal{F}) := \frac{1}{m} \mathbb{E}_{\boldsymbol{x} \sim \{+1, -1\}^m} \left[\sup_{f \in \mathcal{F}} \sum_{i=1}^m x_i \cdot \ell(f, u^{(i)}) \right],$$

where $\boldsymbol{x}$ is distributed i.i.d. according to uniform distribution in $\{+1, -1\}$. We have

Lemma B.1 (Shalev-Shwartz & Ben-David (2014)). *Let S be a training set of size m drawn i.i.d. from distribution $\mathcal{D}$ over $\mathcal{U}$. Then with probability at least $1 - \delta$ over draw of S from $\mathcal{D}$, for all $f \in \mathcal{F}$,*

$$\mathbb{E}_{u \sim \mathcal{D}}[\ell(f, u)] - \frac{1}{m} \sum_{u \in S} \ell(l, u) \leq 2\mathcal{R}_S(\ell \circ \mathcal{F}) + 4\sqrt{\frac{2\ln(4/\delta)}{m}}.$$

Lemma B.2. *If $|\ell(\cdot)| \leq c$ for constant $c > 0$ and $\forall f, f' \in \mathcal{F}, |\ell(f, u) - \ell(f', u)| \leq L\|f - f'\|_\infty$, then we have*

$$\mathbb{E}_{u \sim \mathcal{D}}[\ell(f, u)] - \frac{1}{m} \sum_{u \in S} \ell(l, u) \leq 2 \inf_{r > 0} \left\{ c\sqrt{\frac{2\ln \mathcal{N}_\infty(\mathcal{F}, r)}{m}} + Lr \right\} + 4\sqrt{\frac{2\ln(4/\delta)}{m}}.$$

Proof. For function class $\mathcal{F}$, let $\mathcal{F}_r$ with $|\mathcal{F}_r| = \mathcal{N}_\infty(\mathcal{F}, r)$ be the function class that r-covers $\mathcal{F}$ for some $r > 0$. Similarly, $\forall f \in \mathcal{F}$, denote $f_r \in \mathcal{F}_r$ be the function that r-covers f. We have

$$
\begin{aligned}
\mathcal{R}_S(\ell \circ \mathcal{F}) =& \frac{1}{m} \mathbb{E}_{\boldsymbol{x}} \left[\sup_{f \in \mathcal{F}} \sum_{i=1}^m x_i \cdot \ell(f, u^{(i)}) \right] \\
=& \frac{1}{m} \mathbb{E}_{\boldsymbol{x}} \left[\sup_{f \in \mathcal{F}} \sum_{i=1}^m x_i \cdot \left(\ell(f_r, u^{(i)}) + \ell(f, u^{(i)}) - \ell(f_r, u^{(i)}) \right) \right] \\
\leq& \frac{1}{m} \mathbb{E}_{\boldsymbol{x}} \left[\sup_{f_r \in \mathcal{F}_r} \sum_{i=1}^m x_i \cdot \ell(f_r, u^{(i)}) \right] + \frac{1}{m} \mathbb{E}_{\boldsymbol{x}} \left[\sup_{f \in \mathcal{F}} \sum_{i=1}^m |x_i \cdot Lr| \right] \quad, |\ell(f, u) - \ell(f_r, u)| \leq L\|f - f_r\|_\infty = Lr. \\
\leq& \sup_{f_r \in \mathcal{F}_r} \sqrt{\sum_{i=1}^m \ell^2(f_r, u^{(i)})} \cdot \frac{\sqrt{2\ln \mathcal{N}_\infty(\mathcal{F}, r)}}{m} + \frac{Lr}{m} \mathbb{E}_{\boldsymbol{x}} \|\boldsymbol{x}\| \quad\quad \text{, the first term holds by Massart's lemma.} \\
\leq& \sqrt{c^2 m} \cdot \frac{\sqrt{2\ln \mathcal{N}_\infty(\mathcal{F}, r)}}{m} + \frac{Lr}{m} \mathbb{E}_{\boldsymbol{x}} \|\boldsymbol{x}\| \\
\leq& c\sqrt{\frac{2\ln \mathcal{N}_\infty(\mathcal{F}, r)}{m}} + Lr,
\end{aligned}
$$
(5)

Combining Lemma B.1 and Equation (5), we get

$$\mathbb{E}_{u \sim \mathcal{D}}[\ell(f, u)] - \frac{1}{m} \sum_{u \in S} \ell(l, u) \leq 2 \inf_{r > 0} \left\{ c\sqrt{\frac{2\ln \mathcal{N}_\infty(\mathcal{F}, r)}{m}} + Lr \right\} + 4\sqrt{\frac{2\ln(4/\delta)}{m}}.$$

$\square$

B.1.1. NE APPROXIMATOR

Lemma B.3. *For arbitrary product mixed strategy σ and σ', we have*

$$|\mathcal{E}(\sigma, u) - \mathcal{E}(\sigma', u)| \leq 2n\|\sigma - \sigma'\|.$$

Proof. $\forall \sigma, \sigma'$, we define $y_{-j} := (\sigma_1, \ldots, \sigma_{j-1}, \sigma'_{j+1}, \ldots, \sigma'_n)$. Then, $\forall i \in [n]$ we have

$$
\begin{aligned}
|u_i(\sigma) - u_i(\sigma')| =& |u_i(\sigma_1, \sigma_2, \ldots, \sigma_n) - u_i(\sigma', \sigma'_2, \ldots, \sigma'_n)| \\
=& \Big| \sum_{j=1}^{n} \Big(u_i(\sigma_1, \ldots, \sigma_j, \sigma'_{j+1}, \ldots, \sigma'_n) - u_i(\sigma_1, \ldots, \sigma'_j, \sigma'_{j+1}, \ldots, \sigma'_n) \Big) \Big| \\
=& \Big| \sum_{j=1}^{n} \Big(u_i(\sigma_j, y_{-j}) - u_i(\sigma'_j, y_{-j}) \Big) \Big| \\
=& \Big| \sum_{j=1}^{n} \sum_{a_j} (\sigma_j(a_j) - \sigma'_j(a_j)) \sum_{a_{-j}} u_i(a_j, a_{-j}) y_{-j}(a_{-j}) \Big| \\
\leq& \sum_{j=1}^{n} \sum_{a_j} \Big| \sigma_j(a_j) - \sigma'_j(a_j) \Big| \sum_{a_{-j}} u_i(a_j, a_{-j}) y_{-j}(a_{-j}) \\
\leq& \sum_{j=1}^{n} \sum_{a_j} \Big| \sigma_j(a_j) - \sigma'_j(a_j) \Big| \sum_{a_{-j}} y_{-j}(a_{-j}) \qquad\qquad , u_i(\cdot) \in [0,1] \\
\leq& \sum_{j=1}^{n} \sum_{a_j \in A_j} \Big| \sigma_j(a_j) - \sigma'_j(a_j) \Big| \leq n \max_{j \in [n]} \sum_{a_j \in A_j} \Big| \sigma_j(a_j) - \sigma'_j(a_j) \Big| \\
=& n \|\sigma - \sigma'\|,
\end{aligned}
$$

Therefore, $\forall a_i \in A_i$,

$$
\begin{aligned}
u_i(a_i, \sigma_{-i}) - u_i(\sigma) =& u_i(a_i, \sigma_{-i}) - u_i(a_i, \sigma'_{-i}) + u_i(a_i, \sigma'_{-i}) - u_i(\sigma') + u_i(\sigma') - u_i(\sigma) \\
\leq& n \|\sigma - \sigma'\| + \mathcal{E}(\sigma', u) + n \|\sigma - \sigma'\| \\
=& \mathcal{E}(\sigma', u) + 2n \|\sigma - \sigma'\|.
\end{aligned}
$$

Based on that, we get

$$
\mathcal{E}(\sigma, u) = \max_{i \in N, a_i \in A_i} [u_i(a_i, \sigma_{-i}) - u_i(\sigma)] \leq \mathcal{E}(\sigma', u) + 2n \|\sigma - \sigma'\|.
$$

Similarly, we also have

$$
\mathcal{E}(\sigma', u) \leq \mathcal{E}(\sigma, u) + 2n \|\sigma - \sigma'\|.
$$

$\square$

Based on Lemma B.3, $\forall f, f' \in \mathcal{F}^{\mathrm{NE}}$, we have

$$
\mathcal{E}(f(u), u) - \mathcal{E}(f'(u), u) \leq 2 \|f(u) - f'(u)\| \leq 2 \|f - f'\|_\infty.
$$

Considering that $|\mathcal{E}(\cdot)| \leq 1$, according to Lemma B.2, we have:

$$
\mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}(f^{\mathrm{NE}}(u), u)] - \frac{1}{m} \sum_{u \in S} \mathcal{E}(f^{\mathrm{NE}}(u), u) \leq 2 \cdot \inf_{r > 0} \left\{ \sqrt{\frac{2 \ln \mathcal{N}_\infty(\mathcal{F}^{\mathrm{NE}}, r)}{m}} + 2nr \right\} + 4 \sqrt{\frac{2 \ln(4/\delta)}{m}}.
$$

B.1.2. CCE APPROXIMATOR

Lemma B.4. *For arbitrary joint mixed strategy π and π', we have*

$$
|\mathcal{E}(\pi, u) - \mathcal{E}(\pi', u)| \leq 2 \|\pi - \pi'\|,
$$

Proof. $\forall \pi, \pi', \forall i \in [n]$ we have

$$
|u_i(\pi) - u_i(\pi')| = \sum_{a \in \mathcal{A}} (\pi(a) - \pi'(a)) u_i(a) \overset{(a)}{\leq} \sum_{a \in \mathcal{A}} |\pi(a) - \pi'(a)| = \|\pi - \pi'\|, \tag{6}
$$

where (a) holds since $u_i(\cdot) \in [0, 1]$. Therefore, $\forall a_i \in A_i$,

$$
\begin{aligned}
u_i(a_i, \pi_{-i}) - u_i(\pi) =& u_i(a_i, \pi_{-i}) - u_i(a_i, \pi'_{-i}) + u_i(a_i, \pi'_{-i}) - u_i(\pi') + u_i(\pi') - u_i(\pi) \\
\leq& \|\pi - \pi'\| + \mathcal{E}(\pi', u) + \|\pi - \pi'\| \\
=& \mathcal{E}(\pi', u) + 2\|\pi - \pi'\|.
\end{aligned}
$$

Based on that, we get

$$
\mathcal{E}(\pi, u) = \max_{i \in N, a_i \in A_i} [u_i(a_i, \pi_{-i}) - u_i(\pi)] \leq \mathcal{E}(\pi', u) + 2\|\pi - \pi'\|.
$$

Similarly, we also have

$$
\mathcal{E}(\pi', u) \leq \mathcal{E}(\pi, u) + 2\|\pi - \pi'\|.
$$

$\square$

Based on Lemma B.4, $\forall f, f' \in \mathcal{F}^{\mathrm{CCE}}$, we have

$$
\mathcal{E}(f(u), u) - \mathcal{E}(f'(u), u) \leq 2\|f(u) - f'(u)\| \leq 2\|f - f'\|_\infty.
$$

Considering that $|\mathcal{E}(\cdot)| \leq 1$, according to Lemma B.2, we have:

$$
\mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}(f^{\mathrm{CCE}}(u), u)] - \frac{1}{m} \sum_{u \in S} \mathcal{E}(f^{\mathrm{CCE}}(u), u) \leq 2 \cdot \inf_{r > 0} \left\{ \sqrt{\frac{2 \ln \mathcal{N}_\infty(\mathcal{F}^{\mathrm{CCE}}, r)}{m}} + 2r \right\} + 4\sqrt{\frac{2 \ln(4/\delta)}{m}}.
$$

B.1.3. CE APPROXIMATOR

Lemma B.5. *For arbitrary joint mixed strategy π and π', we have*

$$
|\mathcal{E}^{\mathrm{CE}}(\pi, u) - \mathcal{E}^{\mathrm{CE}}(\pi', u)| \leq 2\|\pi - \pi'\|.
$$

Proof. $\forall a_i \in A_i, \forall \phi_i$, we have

$$
\begin{aligned}
\sum_{a \in \mathcal{A}} \pi(a) u_i(\phi(a_i), a_{-i}) - u_i(\pi) =& \sum_{a \in \mathcal{A}} \pi(a) u_i(\phi(a_i), a_{-i}) - \sum_{a \in \mathcal{A}} \pi'(a) u_i(\phi(a_i), a_{-i}) \\
&+ \sum_{a \in \mathcal{A}} \pi'(a) u_i(\phi(a_i), a_{-i}) - u_i(\pi') + u_i(\pi') - u_i(\pi) \\
\leq& \|\pi - \pi'\| + \mathcal{E}^{\mathrm{CE}}(\pi', u) + \|\pi - \pi'\| \\
=& \mathcal{E}^{\mathrm{CE}}(\pi', u) + 2\|\pi - \pi'\|.
\end{aligned}
$$

Based on that, we get

$$
\mathcal{E}^{\mathrm{CE}}(\pi, u) = \max_{i \in N} \max_{\phi_i} \sum_{a \in \mathcal{A}} \pi(a) u_i(\phi(a_i), a_{-i}) - u_i(\pi) \leq \mathcal{E}^{\mathrm{CE}}(\pi', u) + 2\|\pi - \pi'\|.
$$

Similarly, we also have

$$
\mathcal{E}^{\mathrm{CE}}(\pi', u) \leq \mathcal{E}^{\mathrm{CE}}(\pi, u) + 2\|\pi - \pi'\|.
$$

$\square$

Based on Lemma B.4, $\forall f, f' \in \mathcal{F}^{\mathrm{CE}}$, we have

$$
\mathcal{E}^{\mathrm{CE}}(f(u), u) - \mathcal{E}^{\mathrm{CE}}(f'(u), u) \leq 2\|f(u) - f'(u)\| \leq 2\|f - f'\|_\infty.
$$

Considering that $|\mathcal{E}(\cdot)| \leq 1$, according to Lemma B.2, we have:

$$
\mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}^{\mathrm{CE}}(f^{\mathrm{CE}}(u), u)] - \frac{1}{m} \sum_{u \in S} \mathcal{E}^{\mathrm{CE}}(f^{\mathrm{CE}}(u), u) \leq 2 \cdot \inf_{r > 0} \left\{ \sqrt{\frac{2 \ln \mathcal{N}_\infty(\mathcal{F}^{\mathrm{CE}}, r)}{m}} + 2r \right\} + 4\sqrt{\frac{2 \ln(4/\delta)}{m}}.
$$

B.2. Proof of Theorem 5.4

Theorem 5.4. *[Sample complexity] For $\epsilon, \delta \in (0,1)$, function class $\mathcal{F}$ of NE, CE or CCE approximator and distribution $\mathcal{D}$, with probability at least $1 - \delta$ over draw of the training set S with*

$$m \geq \frac{9}{2\epsilon^2} \left(\ln \frac{2}{\delta} + \ln \mathcal{N}_\infty(\mathcal{F}, \frac{\epsilon}{3L}) \right)$$

games sampled from $\mathcal{D}$, $\forall f \in \mathcal{F}$ we have

$$\mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}(f(u), u)] \leq \frac{1}{m} \sum_{u \in S} \mathcal{E}(f(u), u) + \epsilon,$$

where $L = 2n$ for NE approximators, and $L = 2$ for CE and CCE approximators.

Proof. For function class $\mathcal{F}$ of NE, CE or CCE approximators, according to Lemma B.3, Lemma B.4 and Lemma B.5, $\forall f, g \in \mathcal{F}$ we have
$$\mathcal{E}^{(\text{CE})}(f(u), u) - \mathcal{E}^{(\text{CE})}(g(u), u) \leq L\|f(u) - g(u)\| \leq L\|f - g\|_\infty, \tag{7}$$
where $L = 2n$ for NE approximators, and $L = 2$ for CE and CCE approximators.

For simplicity, we denote $L_S(f) = \frac{1}{m} \sum_{u \in S} \mathcal{E}^{(\text{CE})}(f(u), u)$ and $L_\mathcal{D}(f) = \mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}^{(\text{CE})}(f(u), u)]$. let $\mathcal{F}_r$ with $|\mathcal{F}_r| = \mathcal{N}_\infty(\mathcal{F}, r)$ be the function class that r-covers $\mathcal{F}$ for some $r > 0$. $\forall \epsilon \in (0,1)$, by setting $r = \frac{\epsilon}{3L}$ we have

$$\mathbb{P}_{S \sim \mathcal{D}^m}\left[\exists f \in \mathcal{F}, \left| L_S(f) - L_\mathcal{D}(f) \right| > \epsilon \right]$$

$$\leq \mathbb{P}_{S \sim \mathcal{D}^m}\left[\exists f \in \mathcal{F}, \left| L_S(f) - L_S(f_r) \right| + \left| L_S(f_r) - L_\mathcal{D}(f_r) \right| + \left| L_\mathcal{D}(f_r) - L_\mathcal{D}(f) \right| > \epsilon \right]$$

$$\overset{(a)}{\leq} \mathbb{P}_{S \sim \mathcal{D}^m}\left[\exists f \in \mathcal{F}, Lr + \left| L_S(f_r) - L_\mathcal{D}(f_r) \right| + Lr > \epsilon \right]$$

$$\leq \mathbb{P}_{S \sim \mathcal{D}^m}\left[\exists f_r \in \mathcal{F}_r, \left| L_S(f_r) - L_\mathcal{D}(f_r) \right| > \epsilon - 2Lr \right]$$

$$\overset{(b)}{\leq} \mathcal{N}_\infty(\mathcal{F}, r) \mathbb{P}_{S \sim \mathcal{D}^m}\left[\left| L_S(f) - L_\mathcal{D}(f) \right| > \epsilon - 2Lr \right]$$

$$\overset{(c)}{\leq} 2\mathcal{N}_\infty(\mathcal{F}, r) \exp(-2m(\epsilon - 2Lr)^2),$$

$$= 2\mathcal{N}_\infty(\mathcal{F}, \frac{\epsilon}{3L}) \exp(-\frac{2}{9}m\epsilon^2),$$

where (a) holds by Equation (7), (b) holds by union bound, and (c) holds by Hoeffding inequality. As a result, when $m \geq \frac{9}{2\epsilon^2} \left(\ln \frac{2}{\delta} + \ln \mathcal{N}_\infty(\mathcal{F}, \frac{\epsilon}{3L}) \right)$, we have $\mathbb{P}_{S \sim \mathcal{D}^m}\left[\exists f \in \mathcal{F}, \left| L_S(f) - L_\mathcal{D}(f) \right| > \epsilon \right] < \delta$. $\qquad\square$

B.3. Proof of Theorem 5.5

Theorem 5.5. *The $\mathcal{O}$-projected, $\mathcal{P}$-projected and $\mathcal{Q}$-projected approximator classes have smaller covering numbers, i.e., $\forall r > 0$ we have*

$$\mathcal{N}_\infty(\mathcal{O}\mathcal{F}^{\text{NE}}, r) \leq \mathcal{N}_\infty(\mathcal{F}^{\text{NE}}, r),$$
$$\mathcal{N}_\infty(\mathcal{P}\mathcal{F}^{\text{NE}}, r) \leq \mathcal{N}_\infty(\mathcal{F}^{\text{NE}}, r),$$
$$\mathcal{N}_\infty(\mathcal{Q}\mathcal{F}^{(\text{C})\text{CE}}, r) \leq \mathcal{N}_\infty(\mathcal{F}^{(\text{C})\text{CE}}, r).$$

We first provide an auxiliary lemma.

Lemma B.6. *For function class $\mathcal{F}$ and orbit averaging operator $\mathcal{X}$, if $\forall f, g \in \mathcal{F}, \ell_\infty(\mathcal{X}f, \mathcal{X}g) \leq \ell_\infty(f, g)$, then $\mathcal{N}_\infty(\mathcal{X}\mathcal{F}, r) \leq \mathcal{N}_\infty(\mathcal{F}, r)$ for any $r > 0$.*

Proof. $\forall r > 0$, Denote $\mathcal{F}_r$ as the smallest r-covering set that covers $\mathcal{F}$ with size $\mathcal{N}_\infty(\mathcal{F}, r)$. $\forall f \in \mathcal{F}$, let $f_r \in \mathcal{F}_r$ be the function that r-covers f. We have $\ell_\infty(\mathcal{X}f_r, \mathcal{X}f) \leq \ell_\infty(f_r, f) \leq r$. Therefore, $\mathcal{X}\mathcal{F}_r$ is a r-covering set of $\mathcal{X}\mathcal{F}$, and we have $\mathcal{N}_\infty(\mathcal{X}\mathcal{F}, r) \leq |\mathcal{X}\mathcal{F}_r| \leq |\mathcal{F}_r| = \mathcal{N}_\infty$. $\qquad\square$

Proof of Theorem 5.5. For player $i \in [n]$ and $\forall f^{\mathrm{NE}}, g^{\mathrm{NE}} \in \mathcal{F}^{\mathrm{NE}}$, assuming $\mathcal{U}$ is closed under any $\rho_i \in \mathcal{G}_i$. For $\mathcal{O}_i$,

$$
\begin{aligned}
l_\infty(\mathcal{O}_i f^{\mathrm{NE}}, \mathcal{O}_i g^{\mathrm{NE}}) &= \max_{u \in \mathcal{U}} \|\mathcal{O}_i f^{\mathrm{NE}}(u) - \mathcal{O}_i g^{\mathrm{NE}}(u)\| \\
&= \max_{j \in [n]} \max_{u \in \mathcal{U}} \|(\mathcal{O}_i f^{\mathrm{NE}})(u)_j - (\mathcal{O}_i g^{\mathrm{NE}})(u)_j\| \\
&= \max \left\{ \max_{u \in \mathcal{U}} \|f^{\mathrm{NE}}(u)_i - g^{\mathrm{NE}}(u)_i\|, \ \max_{j \neq i} \max_{u \in \mathcal{U}} \|\frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} (f^{\mathrm{NE}}(\rho_i u)_j - g^{\mathrm{NE}}(\rho_i u)_j)\| \right\} \\
&\leq \max \left\{ \max_{u \in \mathcal{U}} \|f^{\mathrm{NE}}(u)_i - g^{\mathrm{NE}}(u)_i\|, \ \max_{j \neq i} \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} \max_{u \in \mathcal{U}} \|f^{\mathrm{NE}}(\rho_i u)_j - g^{\mathrm{NE}}(\rho_i u)_j\| \right\} \\
&= \max \left\{ \max_{u \in \mathcal{U}} \|f^{\mathrm{NE}}(u)_i - g^{\mathrm{NE}}(u)_i\|, \ \max_{j \neq i} \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} \max_{u \in \mathcal{U}} \|f^{\mathrm{NE}}(u)_j - g^{\mathrm{NE}}(u)_j\| \right\} \\
&= \max \left\{ \max_{u \in \mathcal{U}} \|f^{\mathrm{NE}}(u)_i - g^{\mathrm{NE}}(u)_i\|, \ \max_{j \neq i} \max_u \|f^{\mathrm{NE}}(u)_j - g^{\mathrm{NE}}(u)_j\| \right\} \\
&= l_\infty(f^{\mathrm{NE}}, g^{\mathrm{NE}}).
\end{aligned}
$$

Since $\mathcal{O} = \mathcal{O}_1 \circ \cdots \circ \mathcal{O}_n$, we have

$$
\ell_\infty(\mathcal{O} f^{\mathrm{NE}}, \mathcal{O} g^{\mathrm{NE}}) \leq \ell_\infty(f^{\mathrm{NE}}, g^{\mathrm{NE}}). \tag{8}
$$

For $\mathcal{P}_i$,

$$
\begin{aligned}
l_\infty(\mathcal{P}_i f^{\mathrm{NE}}, \mathcal{P}_i g^{\mathrm{NE}}) &= \max_{u \in \mathcal{U}} \max_{j \in [n]} \|(\mathcal{P}_i f^{\mathrm{NE}})(u)_j - (\mathcal{P}_i g^{\mathrm{NE}})(u)_j\| \\
&= \max \left\{ \max_{j \neq i} \max_u \|f^{\mathrm{NE}}(u)_j - g^{\mathrm{NE}}(u)_j\|, \ \max_u \|\frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} \rho_i^{-1}(f^{\mathrm{NE}}(\rho_i u)_i - g^{\mathrm{NE}}(\rho_i u)_i)\| \right\} \\
&= \max \left\{ \max_{j \neq i} \max_u \|f^{\mathrm{NE}}(u)_j - g^{\mathrm{NE}}(u)_j\|, \ \max_u \|\frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} (f^{\mathrm{NE}}(\rho_i u)_i - g^{\mathrm{NE}}(\rho_i u)_i)\| \right\} \\
&\leq \max \left\{ \max_{j \neq i} \max_u \|f^{\mathrm{NE}}(u)_j - g^{\mathrm{NE}}(u)_j\|, \ \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} \max_u \|f^{\mathrm{NE}}(\rho_i u)_i - g^{\mathrm{NE}}(\rho_i u)_i\| \right\} \\
&= \max \left\{ \max_{j \neq i} \max_u \|f^{\mathrm{NE}}(u)_j - g^{\mathrm{NE}}(u)_j\|, \ \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} \max_u \|f^{\mathrm{NE}}(u)_i - g^{\mathrm{NE}}(u)_i\| \right\} \\
&= \max \left\{ \max_{j \neq i} \max_u \|f^{\mathrm{NE}}(u)_j - g^{\mathrm{NE}}(u)_j\|, \ \max_u \|f^{\mathrm{NE}}(u)_i - g^{\mathrm{NE}}(u)_i\| \right\} \\
&= l_\infty(f^{\mathrm{NE}}, g^{\mathrm{NE}}).
\end{aligned}
$$

Since $\mathcal{P} = \mathcal{P}_1 \circ \cdots \circ \mathcal{P}_n$, we have

$$
\ell_\infty(\mathcal{P} f^{\mathrm{NE}}, \mathcal{P} g^{\mathrm{NE}}) \leq \ell_\infty(f^{\mathrm{NE}}, g^{\mathrm{NE}}). \tag{9}
$$

For CE or CCE approximator $f^{(\mathrm{C})\mathrm{CE}} \in \mathcal{F}^{(\mathrm{C})\mathrm{CE}}$ and $\mathcal{Q}_i$, we have

$$
\begin{aligned}
l_\infty(\mathcal{Q}_i f^{(\mathrm{C})\mathrm{CE}}, \mathcal{Q}_i g^{(\mathrm{C})\mathrm{CE}}) &= \max_{u \in \mathcal{U}} \|(\mathcal{Q}_i f^{(\mathrm{C})\mathrm{CE}})(u) - (\mathcal{Q}_i g^{(\mathrm{C})\mathrm{CE}})(u)\| \\
&= \max_u \|\frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} \rho_i^{-1}(f^{(\mathrm{C})\mathrm{CE}}(\rho_i u) - g^{(\mathrm{C})\mathrm{CE}}(\rho_i u))\| \\
&\leq \max_u \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} \|\rho_i^{-1}(f^{(\mathrm{C})\mathrm{CE}}(\rho_i u) - g^{(\mathrm{C})\mathrm{CE}}(\rho_i u))\| \\
&\leq \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} \max_u \|\rho_i^{-1}(f^{(\mathrm{C})\mathrm{CE}}(\rho_i u) - g^{(\mathrm{C})\mathrm{CE}}(\rho_i u))\| \\
&= \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} \max_u \|f^{(\mathrm{C})\mathrm{CE}}(\rho_i u) - g^{(\mathrm{C})\mathrm{CE}}(\rho_i u)\| \\
&= \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} \max_u \|f^{(\mathrm{C})\mathrm{CE}}(u) - g^{(\mathrm{C})\mathrm{CE}}(u)\| \\
&= l_\infty(f^{(\mathrm{C})\mathrm{CE}}, g^{(\mathrm{C})\mathrm{CE}}).
\end{aligned}
$$

Since $\mathcal{Q} = \mathcal{Q}_1 \circ \cdots \circ \mathcal{Q}_n$, we have

$$
\ell_\infty(\mathcal{Q} f^{(\mathrm{C})\mathrm{CE}}, \mathcal{Q} g^{(\mathrm{C})\mathrm{CE}}) \leq \ell_\infty(f^{(\mathrm{C})\mathrm{CE}}, g^{(\mathrm{C})\mathrm{CE}}). \tag{10}
$$

Combing Lemma B.6, Equation (8), Equation (9) and Equation (10), we finish the proof. $\qquad\square$

B.4. Proof of Theorem 5.6

Theorem 5.6. *When the payoff distribution $\mathcal{D}$ is invariant under the permutation of payoffs, the $\mathcal{Q}$-projected (C)CE approximator has a smaller expected equilibrium approximation. Formally, for all $f^{(\mathrm{C})\mathrm{CE}} \in \mathcal{F}^{(\mathrm{C})\mathrm{CE}}$ and permutation-invariant distribution $\mathcal{D}$, we have*

$$
\mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}(\mathcal{Q} f^{(\mathrm{C})\mathrm{CE}}(u), u)] \leq \mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}(f^{(\mathrm{C})\mathrm{CE}}(u), u)],
$$

We first prove a lemma about the property of $\mathcal{E}_i(\pi, u)$ and $\mathcal{E}_i^{\mathrm{CE}}(\pi, u)$.

Lemma B.7. $\mathcal{E}_i(\pi, u)$ *and* $\mathcal{E}_i^{\mathrm{CE}}(\pi, u)$ *are convex on π, i.e.*

$$
p\mathcal{E}_i^{(CE)}(\pi_1, u) + (1-p)\mathcal{E}_i^{(CE)}(\pi_2, u) \geq \mathcal{E}_i^{(CE)}(p\pi_1 + (1-p)\pi_2, u), \quad \forall p \in [0, 1].
$$

Proof. We recall the definition $\mathcal{E}_i(\pi, u) = \max_{a_i \in \mathcal{A}_i} u_i(a_i, \pi_{-i}) - u_i(\pi)$ for CCE approximator and $\mathcal{E}_i^{\mathrm{CE}}(\pi, u) = \max_{\phi_i \in \mathcal{A}_i \to \mathcal{A}_i} \sum_a \pi(a) u_i(\phi_i(a_i), a_{-i}) - u_i(\pi)$ for CE approximator. $u_i(a_i, \pi_{-i})$ is linear on π. Given ϕ, $\sum_a \pi(a) u_i(\phi_i(a_i), a_{-i})$ is also linear on π. Moreover, the maximum operator on a set of linear functions will induce a convex function. $\qquad\square$

Proof of Theorem 5.6. For $f \in \mathcal{F}^{(C)CE}$ and $\forall i, j \in [n]$,

$$\mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}_i^{(CE)}(\mathcal{Q}_j f(u), u)] = \mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}_i^{(CE)}(\frac{1}{|\mathcal{A}_j|!} \sum_{\rho_j \in \mathcal{G}_j} \rho_j^{-1} f(\rho_j u), u)] \qquad \text{, by definition.}$$

$$\leq \frac{1}{|\mathcal{A}_j|!} \sum_{\rho_j \in \mathcal{G}_j} \mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}_i^{(CE)}(\rho_j^{-1} f(\rho_j u), u)] \qquad \text{, by convexity.}$$

$$= \frac{1}{|\mathcal{A}_j|!} \sum_{\rho_j \in \mathcal{G}_j} \mathbb{E}_{v \sim \mathcal{D}}[\mathcal{E}_i^{(CE)}(\rho_j^{-1} f(v), \rho_j^{-1} v)] \qquad \text{, let } v = \rho_j u.$$

$$= \frac{1}{|\mathcal{A}_j|!} \sum_{\rho_j \in \mathcal{G}_j} \mathbb{E}_{v \sim \mathcal{D}}[\mathcal{E}_i^{(CE)}(f(v), v)] \qquad \text{, invariance of } \mathcal{E}_i^{(CE)}(\pi, u) \text{ under } \rho_j^{-1} \in \mathcal{G}_j.$$

$$= \mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}_i^{(CE)}(f(u), u)].$$

Since $\mathcal{Q} = \circ_i \mathcal{Q}_i$ and $\mathcal{E} = \max_i \mathcal{E}_i$, we have

$$\mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}(\mathcal{Q}f(u), u)] \leq \mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}(f(u), u)].$$

$\square$

B.5. Proof of Theorem 5.7

Theorem 5.7. *For bimatrix constant-sum games, when the payoff distribution $\mathcal{D}$ is invariant under the permutation of payoffs, then the $\mathcal{X}$-projected ($\mathcal{X} \in \{\mathcal{O}, \mathcal{P}\}$) NE approximator has a smaller expected exploitability. Formally, for all $f^{NE} \in \mathcal{F}^{NE}$ and permutation-invariant distribution $\mathcal{D}$ for bimatrix constant-sum games, we have*

$$\mathbb{E}_{u \sim \mathcal{D}} \left[\sum_{i=1}^{n} \mathcal{E}_i((\mathcal{X}f^{NE})(u), u) \right]$$

$$\leq \mathbb{E}_{u \sim \mathcal{D}} \left[\sum_{i=1}^{n} \mathcal{E}_i(f^{NE}(u), u) \right].$$

Proof. We only prove for the $\mathcal{P}$-projected case; the proof for $\mathcal{O}$-projected case is similar and therefore omitted.

Recall

$$\mathcal{E}_i(\sigma, u) = \max_{a_i \in \mathcal{A}_i} u_i(a_i, \sigma_{-i}) - u_i(\sigma).$$

Denote $u_1(\sigma) + u_2(\sigma) \equiv c$, then

$$\sum_{i=1}^{n} \mathcal{E}_i(\sigma, u) = \max_{a_1 \in \mathcal{A}_1, a_2 \in \mathcal{A}_2} u_1(a_1, \sigma_2) + u_2(a_2, \sigma_1) - c.$$

Then we have

$$\mathbb{E}_{u \sim \mathcal{D}}[\sum_{i=1}^{n} \mathcal{E}_i((\mathcal{P}f^{NE})(u), u)] = \mathbb{E}_{u \sim \mathcal{D}}[\max_{a_1, a_2} u_1(a_1, (\mathcal{P}f^{NE})(u)_2) + u_2(a_2, (\mathcal{P}f^{NE})(u)_1) - c]$$

$$= \mathbb{E}_{u \sim \mathcal{D}}[\max_{a_1} u_1(a_1, (\mathcal{P}f^{NE})(u)_2)] + \mathbb{E}_{u \sim \mathcal{D}}[\max_{a_2} u_2(a_2, (\mathcal{P}f^{NE})(u)_1)] - c.$$

For the first term,

$$
\begin{aligned}
\mathbb{E}_{u\sim\mathcal{D}}[\max_{a_1} u_1(a_1, (\mathcal{P}f^{\mathrm{NE}})(u)_2)] =& \mathbb{E}_{u\sim\mathcal{D}}[\max_{a_1} u_1(a_1, \frac{1}{|\mathcal{A}_2|!}\sum_{\rho_2\in\mathcal{G}_2}\rho_2^{-1}f^{\mathrm{NE}}(\rho_2 u)_2)] \\
\leq& \frac{1}{|\mathcal{A}_2|!}\sum_{\rho_2\in\mathcal{G}_2}\mathbb{E}_{u\sim\mathcal{D}}[\max_{a_1} u_1(a_1, \rho_2^{-1}f^{\mathrm{NE}}(\rho_2 u)_2)] \\
=& \frac{1}{|\mathcal{A}_2|!}\sum_{\rho_2\in\mathcal{G}_2}\mathbb{E}_{v\sim\mathcal{D}}[\max_{a_1}(\rho_2^{-1}v)_1(a_1, \rho_2^{-1}f^{\mathrm{NE}}(v)_2)] \\
=& \frac{1}{|\mathcal{A}_2|!}\sum_{\rho_2\in\mathcal{G}_2}\mathbb{E}_{v\sim\mathcal{D}}[\max_{a_1} v_1(a_1, f^{\mathrm{NE}}(v)_2)] \\
=& \mathbb{E}_{u\sim\mathcal{D}}[\max_{a_1} u_1(a_1, f^{\mathrm{NE}}(u)_2)].
\end{aligned}
$$

Similarly, for the second term,

$$
\mathbb{E}_{u\sim\mathcal{D}}[\max_{a_2} u_2(a_2, (\mathcal{P}f^{\mathrm{NE}})(u)_1)] \leq \mathbb{E}_{u\sim\mathcal{D}}[\max_{a_2} u_2(a_2, f^{\mathrm{NE}}(u)_1)].
$$

Above all,

$$
\begin{aligned}
\mathbb{E}_{u\sim\mathcal{D}}[\sum_{i=1}^n \mathcal{E}_i((\mathcal{P}f^{\mathrm{NE}})(u), u)] =& \mathbb{E}_{u\sim\mathcal{D}}[\max_{a_1} u_1(a_1, (\mathcal{P}f^{\mathrm{NE}})(u)_2)] + \mathbb{E}_{u\sim\mathcal{D}}[\max_{a_2} u_2(a_2, (\mathcal{P}f^{\mathrm{NE}})(u)_1)] - c \\
\leq& \mathbb{E}_{u\sim\mathcal{D}}[\max_{a_1} u_1(a_1, f^{\mathrm{NE}}(u)_2)] + \mathbb{E}_{u\sim\mathcal{D}}[\max_{a_2} u_2(a_2, f^{\mathrm{NE}}(u)_1)] - c \\
=& \mathbb{E}_{u\sim\mathcal{D}}[\sum_{i=1}^n \mathcal{E}_i(f^{\mathrm{NE}}(u), u)].
\end{aligned}
$$

$\square$

B.6. Proof of Theorem 5.8

Theorem 5.8. *When the payoff distribution $\mathcal{D}$ is invariant under the permutation of payoffs, and f^{NE} satisfies OPI, then the $\mathcal{P}$-projected NE approximator has a smaller expected NE approximation. Formally, for all $f^{\mathrm{NE}} \in \mathcal{F}^{\mathrm{NE}}$ that is OPI and permutation-invariant distribution $\mathcal{D}$, we have*

$$
\mathbb{E}_{u\sim\mathcal{D}}[\mathcal{E}((\mathcal{P}f^{\mathrm{NE}})(u), u)] \leq \mathbb{E}_{u\sim\mathcal{D}}[\mathcal{E}(f^{\mathrm{NE}}(u), u)].
$$

We first introduce a useful lemma. It is about the property of $\mathcal{E}_i(\sigma, u)$

Lemma B.8. $\mathcal{E}_i(\sigma, u)$ *is*

1. *Linear on σ_i, i.e.*

$$
p\mathcal{E}_i((\sigma_i^1, \sigma_{-i}), u) + (1-p)\mathcal{E}_i((\sigma_i^2, \sigma_{-i}), u) = \mathcal{E}_i((p\sigma_i^1 + (1-p)\sigma_i^2, \sigma_{-i}), u), \ \forall p \in [0, 1].
$$

2. *Convex on σ_j, i.e.*

$$
p\mathcal{E}_i((\sigma_j^1, \sigma_{-j}), u) + (1-p)\mathcal{E}_i((\sigma_j^2, \sigma_{-j}), u) \geq \mathcal{E}_i((p\sigma_j^1 + (1-p)\sigma_j^2, \sigma_{-j}), u), \ \forall p \in [0, 1], j \neq i.
$$

Proof. We recall the definition $\mathcal{E}_i(\sigma, u) = \max_{a_i\in\mathcal{A}_i} u_i(a_i, \sigma_{-i}) - u_i(\sigma)$. Notice that $u_i(\sigma)$ is linear on σ_k for all $k \in [n]$, thus both $u_i(a_i, \sigma_{-i})$ and $u_i(\sigma)$ are linear on σ_k for any $k \in [n]$. Moreover, the maximum operator on a set of linear functions will induce a convex function.

$\square$

Proof of Theorem 5.8. We prove the theorem in two steps.

Step 1 First, we show that

$$\mathbb{E}_{u\sim\mathcal{D}}[\mathcal{E}_i((\mathcal{P}_i f^{\mathrm{NE}})(u), u)] = \mathbb{E}_{u\sim\mathcal{D}}[\mathcal{E}_i(f^{\mathrm{NE}}(u), u)], \quad \forall f^{\mathrm{NE}} \in \mathcal{F}^{\mathrm{NE}}.$$

By definition,

$$\mathbb{E}_{u\sim\mathcal{D}}[\mathcal{E}_i(\mathcal{P}_i f^{\mathrm{NE}}(u), u)]$$

$$=\mathbb{E}_{u\sim\mathcal{D}}[\mathcal{E}_i((\frac{1}{|\mathcal{A}_i|!}\sum_{\rho_i\in\mathcal{G}_i}\rho_i^{-1}f(\rho_i u)_i, f(u)_{-i}), u)]$$

$$=\frac{1}{|\mathcal{A}_i|!}\sum_{\rho_i\in\mathcal{G}_i}\mathbb{E}_{u\sim\mathcal{D}}[\mathcal{E}_i((\rho_i^{-1}f(\rho_i u)_i, f(u)_{-i}), u)] \qquad\qquad \text{, by linearity of } \mathcal{E}_i(\sigma, u) \text{ on } \sigma_i.$$

$$=\frac{1}{|\mathcal{A}_i|!}\sum_{\rho_i\in\mathcal{G}_i}\mathbb{E}_{v\sim\mathcal{D}}[\mathcal{E}_i((\rho_i^{-1}f(v)_i, f(\rho_i^{-1}v)_{-i}), \rho_i^{-1}v)] \qquad \text{, let } v=\rho_i u \text{ and use the invariance of } \mathcal{D}.$$

$$=\frac{1}{|\mathcal{A}_i|!}\sum_{\rho_i\in\mathcal{G}_i}\mathbb{E}_{v\sim\mathcal{D}}[\mathcal{E}_i((\rho_i^{-1}f(v)_i, f(v)_{-i}), \rho_i^{-1}v)] \qquad\qquad \text{, OPI of } f.$$

$$=\frac{1}{|\mathcal{A}_i|!}\sum_{\rho_i\in\mathcal{G}_i}\mathbb{E}_{u\sim\mathcal{D}}[\mathcal{E}_i((f(u)_i, f(u)_{-i}), u)] \qquad\qquad\quad \text{, invariance of } \mathcal{E}_i(\sigma, u). \text{ under } \rho_i^{-1}\in\mathcal{G}_i.$$

$$=\mathbb{E}_{u\sim\mathcal{D}}[\mathcal{E}_i(f^{\mathrm{NE}}(u), u)].$$

Step 2 Then we show that

$$\mathbb{E}_{u\sim\mathcal{D}}[\mathcal{E}_j((\mathcal{P}_i f^{\mathrm{NE}})(u), u)] \le \mathbb{E}_{u\sim\mathcal{D}}[\mathcal{E}_j(f^{\mathrm{NE}}(u), u)], \quad \forall f^{\mathrm{NE}} \in \mathcal{F}^{\mathrm{NE}}, j \ne i.$$

$$\mathbb{E}_{u\sim\mathcal{D}}[\mathcal{E}_j((\mathcal{P}_i f^{\mathrm{NE}})(u), u)]$$

$$=\mathbb{E}_{u\sim\mathcal{D}}[\mathcal{E}_j((\frac{1}{|\mathcal{A}_i|!}\sum_{\rho_i\in\mathcal{G}_i}\rho_i^{-1}f(\rho_i u)_i, f(u)_{-i}), u)]$$

$$\le\frac{1}{|\mathcal{A}_i|!}\sum_{\rho_i\in\mathcal{G}_i}\mathbb{E}_{u\sim\mathcal{D}}[\mathcal{E}_j((\rho_i^{-1}f(\rho_i u)_i, f(u)_{-i}), u)] \qquad\qquad \text{, by convexity of } \mathcal{E}_j(\sigma, u) \text{ on } \sigma_i.$$

$$=\frac{1}{|\mathcal{A}_i|!}\sum_{\rho_i\in\mathcal{G}_i}\mathbb{E}_{v\sim\mathcal{D}}[\mathcal{E}_j((\rho_i^{-1}f(v)_i, f(\rho_i^{-1}v)_{-i}), \rho_i^{-1}v)] \qquad \text{, let } v=\rho_i u \text{ and use the invariance of } \mathcal{D}.$$

$$=\frac{1}{|\mathcal{A}_i|!}\sum_{\rho_i\in\mathcal{G}_i}\mathbb{E}_{v\sim\mathcal{D}}[\mathcal{E}_j((\rho_i^{-1}f(v)_i, f(v)_{-i}), \rho_i^{-1}v)] \qquad\qquad \text{, OPI of } f.$$

$$=\frac{1}{|\mathcal{A}_i|!}\sum_{\rho_i\in\mathcal{G}_i}\mathbb{E}_{u\sim\mathcal{D}}[\mathcal{E}_j((f(u)_i, f(u)_{-i}), u)] \qquad\qquad\quad \text{, invariance of } \mathcal{E}_j(\sigma, u) \text{ under } \rho_i^{-1}\in\mathcal{G}_i.$$

$$=\mathbb{E}_{u\sim\mathcal{D}}[\mathcal{E}_j(f^{\mathrm{NE}}(u), u)].$$

Since $\mathcal{P} = \circ_i \mathcal{P}_i$ and $\mathcal{E} = \max_i \mathcal{E}_i$, we have

$$\mathbb{E}_{u\sim\mathcal{D}}[\mathcal{E}((\mathcal{P}f^{\mathrm{NE}})(u), u)] \le \mathbb{E}_{u\sim\mathcal{D}}[\mathcal{E}(f^{\mathrm{NE}}(u), u)].$$

$\square$

B.7. Proof of Theorem 5.9

Theorem 5.9. *For bimatrix games, when the payoff distribution $\mathcal{D}$ is invariant under the permutation of payoffs, and f^{NE} satisfies PPE, then the $\mathcal{O}$-projected NE approximator has a smaller expected NE approximation. Formally, for all $f^{\mathrm{NE}} \in \mathcal{F}^{\mathrm{NE}}$ that is PPE and permutation-invariant distribution $\mathcal{D}$ of bimatrix games, we have*

$$\mathbb{E}_{u\sim\mathcal{D}}[\mathcal{E}((\mathcal{O}f^{\mathrm{NE}})(u), u)] \le \mathbb{E}_{u\sim\mathcal{D}}[\mathcal{E}(f^{\mathrm{NE}}(u), u)].$$

Proof. We prove the theorem in two steps, similar to the proof of Theorem 5.8.

Step 1 First we show that for player $i \in \{1, 2\}$, let $\{j\} = \{1, 2\} \backslash \{i\}$,

$$\mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}_i((\mathcal{O}_i f^{\mathrm{NE}})(u), u)] \leq \mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}_i(f^{\mathrm{NE}}(u), u)].$$

This is because

$$\mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}_i((\mathcal{O}_i f^{\mathrm{NE}})(u), u)] = \mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}_i((f^{\mathrm{NE}}(u)_i, \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} f^{\mathrm{NE}}(\rho_i u)_j), u)]$$

$$\leq \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} \mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}_i((f^{\mathrm{NE}}(u)_i, f^{\mathrm{NE}}(\rho_i u)_j), u)] \qquad \text{, by convexity of } \mathcal{E}_i \text{ on } \sigma_j.$$

$$= \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} \mathbb{E}_{v \sim \mathcal{D}}[\mathcal{E}_i((f^{\mathrm{NE}}(\rho_i^{-1} v)_i, f^{\mathrm{NE}}(v)_j), \rho_i^{-1} v)] \qquad \text{, let } v = \rho_i u.$$

$$= \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} \mathbb{E}_{v \sim \mathcal{D}}[\mathcal{E}_i((\rho_i^{-1} f^{\mathrm{NE}}(v)_i, f^{\mathrm{NE}}(v)_j), \rho_i^{-1} v)] \qquad \text{, by PPE of } f^{\mathrm{NE}}.$$

$$= \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} \mathbb{E}_{v \sim \mathcal{D}}[\mathcal{E}_i((f^{\mathrm{NE}}(v)_i, f^{\mathrm{NE}}(v)_j), v)] \qquad \text{, invariance of } \mathcal{E}_i(\sigma, u) \text{ under } \rho_i^{-1} \in \mathcal{G}.$$

$$= \mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}_i((f^{\mathrm{NE}})(u), u)].$$

Step 2 Then we show that if $j \neq i$ and $\{i, j\} = \{1, 2\}$

$$\mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}_j((\mathcal{O}_i f^{\mathrm{NE}})(u), u)] = \mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}_j(f^{\mathrm{NE}}(u), u)].$$

This is because

$$\mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}_j((\mathcal{O}_i f^{\mathrm{NE}})(u), u)] = \mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}_j((f^{\mathrm{NE}}(u)_i, \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} f^{\mathrm{NE}}(\rho_i u)_j), u)]$$

$$= \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} \mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}_j((f^{\mathrm{NE}}(u)_i, f^{\mathrm{NE}}(\rho_i u)_j), u)] \qquad \text{, by linearity of } \mathcal{E}_j \text{ on } \sigma_j.$$

$$= \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} \mathbb{E}_{v \sim \mathcal{D}}[\mathcal{E}_j((f^{\mathrm{NE}}(\rho_i^{-1} v)_i, f^{\mathrm{NE}}(v)_j), \rho_i^{-1} v)] \qquad \text{, let } v = \rho_i u.$$

$$= \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} \mathbb{E}_{v \sim \mathcal{D}}[\mathcal{E}_j((\rho_i^{-1} f^{\mathrm{NE}}(v)_i, f^{\mathrm{NE}}(v)_j), \rho_i^{-1} v)] \qquad \text{, by PPE of } f^{\mathrm{NE}}.$$

$$= \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} \mathbb{E}_{v \sim \mathcal{D}}[\mathcal{E}_j((f^{\mathrm{NE}}(v)_i, f^{\mathrm{NE}}(v)_j), v)] \qquad \text{, invariance of } \mathcal{E}_j(\sigma, u) \text{ under } \rho_i^{-1} \in \mathcal{G}_i.$$

$$= \mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}_j(f^{\mathrm{NE}}(u), u)].$$

Since $\mathcal{O} = \circ_i \mathcal{O}_i$ and $\mathcal{E} = \max_i \mathcal{E}_i$, we have

$$\mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}(\mathcal{O} f^{\mathrm{NE}}(u), u)] \leq \mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}(f^{\mathrm{NE}}(u), u)].$$

$\square$

C. Omitted Proofs in Section 6

C.1. Proof of Theorem 6.3

Theorem 6.3. *For a permutation-ρ-invariant game Γ_u. if there is a pure NE $a^* = (a_i^*)_{i \in [n]}$ and at least one player $k \in [n]$ such that $a_k^* \in V(\rho_k)$, then a^* will never be found by any NE approximator with both PPE and OPI. Besides, a^* (as a pure CE or CCE) will also never be found by any CE or CCE approximator with PE.*

Proof. Let f be a PPE and OPI NE approximator. Denote $f(u) = (\sigma_i)_{i \in [n]}$. For player k that $a_k^* \in V(\rho_k)$, we get

$$\sigma_k = f(u)_k \overset{(a)}{=} f(\rho u)_k \overset{(b)}{=} f(\rho_k u)_k \overset{(c)}{=} \rho_k f(u)_k = \rho_k \sigma_k, \tag{11}$$

where (a) holds since u is permutable w.r.t. ρ, (b) holds by OPI of f, and (c) holds by PPE of f. If a^* can be found by f, we will have $1 = \sigma_k(a_k^*) \overset{(d)}{=} \rho_k \sigma_k(a_k^*) = \sigma_k(\rho_k^{-1}(a_k^*))$, where (d) holds by Equation (11). However, such result leads to a contradiction, because $a_k^* \neq \rho_k^{-1}(a_k)$ but $\sigma_k(a_k^*) = \sigma(\rho_k^{-1}(a_k^*)) = 1$.

Let f be a PE (C)CE approximator. Denote $f(u) = \pi$, we have

$$\pi = f(u) \overset{(a)}{=} f(\rho u) \overset{(b)}{=} \rho f(u) = \rho \pi, \tag{12}$$

where (a) holds since u is permutable w.r.t. ρ, (b) holds by PE of f. If a^* can be found by f, we will have $1 = \pi(a^*) \overset{(c)}{=} \rho \pi(a^*) = \pi(\rho^{-1} a^*) = \pi(\rho_1^{-1} a_1^*, \cdots, \rho_n^{-1} a_n^*)$, where (c) holds by Equation (12). However, from $a_k^* \in V(\rho_k)$ we know $\rho_k^{-1}(a_k^*) \neq a_k^*$, then $\rho^{-1} a^* \neq a^*$, but $\pi(a^*) = \pi(\rho^{-1} a^*) = 1$. $\qquad \square$

C.2. Proof of Theorem 6.6

Theorem 6.6. *Given $N, M \geq 2$, let $\mathcal{F}_{\mathrm{PE}}^{\mathrm{(C)CE}}$ be the function class (target on games with the number of players $n \leq N$ and $|\mathcal{A}_i| \leq M$) of all the (C)CE approximators with PE. Denote by $\mathcal{F}_{\mathrm{general}}^{\mathrm{(C)CE}}$ the function class of all the (C)CE approximators. Then we have*

$$\mathrm{SWR}_{\mathrm{N,M}}(\mathcal{F}_{\mathrm{PE}}^{\mathrm{(C)CE}}, \mathcal{F}_{\mathrm{general}}^{\mathrm{(C)CE}}) = 1.$$

Proof. Assume $f \in \mathcal{F}_{\mathrm{general}}^{\mathrm{(C)CE}}$ is an (C)CE approximator that always finds the strategy that maximizes the social welfare. Afterward, we construct another f_0 that satisfies PE and always finds the strategy that maximizes social welfare. f_0 is constructed by orbit averaging:

$$f_0(u) = \mathcal{Q}f(u),$$

thus f_0 is PE.

Denote $\mathcal{D}$ as an arbitrary payoff distribution of u such that $\mathcal{D}$ is invariant under permutation and the cardinality of its support is finite. We have

$$\begin{aligned}
\mathbb{E}_{u \sim \mathcal{D}} \mathrm{SW}(\mathcal{Q}_i f(u), u) &= \mathbb{E}_{u \sim \mathcal{D}} \mathrm{SW}\left(\frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} \rho_i^{-1} f(\rho_i u), u\right) \\
&= \mathbb{E}_{u \sim \mathcal{D}} \sum_{i=1}^{n} u_i \left(\frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} \rho_i^{-1} f(\rho_i u)\right) \\
&= \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} \mathbb{E}_{u \sim \mathcal{D}} \sum_{i=1}^{n} u_i(\rho_i^{-1} f(\rho_i u)) \\
&= \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} \mathbb{E}_{v \sim \mathcal{D}} \sum_{i=1}^{n} (\rho_i^{-1} v)_i(\rho_i^{-1} f(v)) \qquad\qquad \text{, let } v = \rho_i u. \\
&= \frac{1}{|\mathcal{A}_i|!} \sum_{\rho_i \in \mathcal{G}_i} \mathbb{E}_{v \sim \mathcal{D}} \sum_{i=1}^{n} v_i(f(v)) \\
&= \mathbb{E}_{u \sim \mathcal{D}} \sum_{i=1}^{n} u_i(f(u)) \\
&= \mathbb{E}_{u \sim \mathcal{D}} \mathrm{SW}(f(u), u).
\end{aligned}$$

Due to that $\mathcal{Q} = \mathcal{Q}_1 \circ \cdots \circ \mathcal{Q}_n$, we have

$$\mathbb{E}_{u \sim \mathcal{D}} \mathrm{SW}(f_0(u), u) = \mathbb{E}_{u \sim \mathcal{D}} \mathrm{SW}(f(u), u).$$

Due to the arbitrariness of $\mathcal{D}$, we know that f_0 maximizes the social welfare w.r.t. any u.

From above, we immediately know

$$\text{SWR}_{\text{N,M}}(\mathcal{F}_{\text{PE}}^{\text{(C)CE}}, \mathcal{F}_{\text{general}}^{\text{(C)CE}}) = 1.$$

$\square$

C.3. Proof of Theorem 6.7

Theorem 6.7. *Given $N, M \geq 2$, let $\mathcal{F}_{\text{OPI}}^{\text{NE}}, \mathcal{F}_{\text{PPE}}^{\text{NE}}$ and $\mathcal{F}_{\text{both}}^{\text{NE}}$ be the function classes (target on games with number of players $n \leq N$ and $|\mathcal{A}_i| \leq M$) of all the NE approximators with OPI, PPE and both. Denote the function class of all the NE approximators as $\mathcal{F}_{\text{general}}^{\text{NE}}$. Then we have*

$$\text{SWR}_{\text{N,M}}(\mathcal{F}_{\text{OPI}}^{\text{NE}}, \mathcal{F}_{\text{general}}^{\text{NE}}) = \frac{1}{M^{N-1}}, \tag{1}$$

$$\text{SWR}_{\text{N,M}}(\mathcal{F}_{\text{PPE}}^{\text{NE}}, \mathcal{F}_{\text{general}}^{\text{NE}}) \leq \frac{1}{M}, \tag{2}$$

$$\text{SWR}_{\text{N,M}}(\mathcal{F}_{\text{both}}^{\text{NE}}, \mathcal{F}_{\text{general}}^{\text{NE}}) = \frac{1}{M^{N-1}}. \tag{3}$$

Additionally, when $M \geq 3$, denote by $\widetilde{\mathcal{F}}_{\text{both}}^{\text{NE}}$ the function class of all the NE oracles (functions that always output exact NE solutions of the input games) with both PPE and OPI, and by $\widetilde{\mathcal{F}}_{\text{general}}^{\text{NE}}$ the function class of all the NE oracles. Then we have

$$\text{SWR}_{\text{N,M}}(\widetilde{\mathcal{F}}_{\text{both}}^{\text{NE}}, \widetilde{\mathcal{F}}_{\text{general}}^{\text{NE}}) = 0. \tag{4}$$

C.3.1. PROOF OF EQUATION (1) AND EQUATION (3)

We first prove the theorem with respect to $\mathcal{F}_{\text{OPI}}^{\text{NE}}$ and $\mathcal{F}_{\text{both}}^{\text{NE}}$

Step 1 On the one part, we prove

$$\left.\begin{array}{l} \text{SWR}_{\text{N,M}}(\mathcal{F}_{\text{OPI}}^{\text{NE}}, \mathcal{F}_{\text{general}}^{\text{NE}}) \\ \text{SWR}_{\text{N,M}}(\mathcal{F}_{\text{both}}^{\text{NE}}, \mathcal{F}_{\text{general}}^{\text{NE}}) \end{array}\right\} \leq \frac{1}{M^{N-1}}.$$

We prove this by construction.

Consider a game with N player and $\mathcal{A}_i = [M]$ for $i \in [N]$. $\forall a \in \mathcal{A}, i \in [N]$, define the payoff $\bar{u}$ as follows:

$$\bar{u}_i(a) = \begin{cases} 1 & \text{, if } a_1 = a_2 = \cdots = a_N. \\ 0 & \text{, otherwise.} \end{cases}$$

Define $U = \{u' | u' = \circ_i \rho_i \bar{u}, \rho_i \in \mathcal{G}_i\}$ and $\mathcal{D}$ as a uniform distribution on U. Easy to certify that $\mathcal{D}$ is a permutation-invariant distribution.

Let $\tilde{f} \in \tilde{\mathcal{F}}_{\text{general}}^{\text{NE}}$ be the NE oracle that $\tilde{f}(\bar{u})_i = 1$ and for any $u' = \circ_i \rho_i \bar{u} \in U$, $\tilde{f}(u')_i = \rho_i(1)$. Intuitively, the oracle will choose the action that will provide all players with revenue 1, leading to a social welfare of N. Since each player has got her maximum possible utility, we have

$$\max_{f \in \mathcal{F}_{\text{general}}^{\text{NE}}} \mathbb{E}_{u \sim \mathcal{D}} \text{SW}(f(u), u) = \max_{\tilde{f} \in \tilde{\mathcal{F}}_{\text{general}}^{\text{NE}}} \mathbb{E}_{u \sim \mathcal{D}} \text{SW}(\tilde{f}(u), u) = N. \tag{13}$$

For any $j_1, j_2 \in [M]$ and $j_1 < j_2$, let $\rho_i^{(j_1,j_2)} = (1, \ldots, j_2, \ldots, j_1, \ldots, M)$ for all player $i \in [N]$ be the swap permutation that swaps actions j_1 and j_2 and keeps other actions still. Then $\circ_{i \neq j} \rho_i^{(j_1,j_2)} \bar{u} = \rho_j^{(j_1,j_2)} \bar{u}$ for player j. For $f \in \mathcal{F}_{\text{OPI}}^{\text{NE}}$, we have $f(\bar{u})_j = f(\circ_{i \neq j} \rho_i^{(j_1,j_2)} \bar{u})_j = f(\rho_j^{(j_1,j_2)} \bar{u})_j$ for arbitrary swap permutation $\rho_j^{(j_1,j_2)}$. Since any permutation can be

achieved by composition of swap permutations, we have $\forall \rho_j \in \mathcal{G}_j$, $f(\bar{u})_j = f(\rho_j \bar{u})_j$. Based on that, and by OPI of f, $\forall \rho = \circ_{i \in [N]} \rho_i$ we have $f(\bar{u})_j = f(\rho \bar{u})_j$, i.e. f is a constant function on U. Without loss of generality, we denote $f(u) \equiv \sigma$ for all $u \in U$. Then

$$\mathbb{E}_{u \sim \mathcal{D}} \mathrm{SW}(f(u), u) = \frac{1}{|U|} \sum_{u' \in U} \mathrm{SW}(\sigma, u') = \frac{1}{(M!)^{N-1}} \mathrm{SW}(\sigma, \sum_{u' \in U} u').$$

Additionally, we have $(\sum_{u' \in U} u')(a) = ((M-1)!)^{N-1}$ for any $a \in \mathcal{A}$. Based on that, we have

$$\max_{f \in \mathcal{F}_{\mathrm{OPI}}^{\mathrm{NE}}} \mathbb{E}_{u \sim \mathcal{D}} \mathrm{SW}(f(u), u) = \frac{1}{(M!)^{N-1}} \cdot N((M-1)!)^{N-1} = \frac{N}{M^{N-1}}. \tag{14}$$

Combining Equation (13) and Equation (14), we have

$$\mathrm{SWR}_{\mathrm{N,M}}(\mathcal{F}_{\mathrm{OPI}}^{\mathrm{NE}}, \mathcal{F}_{\mathrm{general}}^{\mathrm{NE}}) \leq \frac{1}{M^{N-1}}.$$

Due to $\mathcal{F}_{\mathrm{both}}^{\mathrm{NE}} \subseteq \mathcal{F}_{\mathrm{OPI}}^{\mathrm{NE}}$, we immediately know

$$\mathrm{SWR}_{\mathrm{N,M}}(\mathcal{F}_{\mathrm{both}}^{\mathrm{NE}}, \mathcal{F}_{\mathrm{general}}^{\mathrm{NE}}) \leq \frac{1}{M^{N-1}}.$$

Step 2 On the other part, we prove

$$\left.\begin{array}{l} \mathrm{SWR}_{\mathrm{N,M}}(\mathcal{F}_{\mathrm{OPI}}^{\mathrm{NE}}, \mathcal{F}_{\mathrm{general}}^{\mathrm{NE}}) \\ \mathrm{SWR}_{\mathrm{N,M}}(\mathcal{F}_{\mathrm{both}}^{\mathrm{NE}}, \mathcal{F}_{\mathrm{general}}^{\mathrm{NE}}) \end{array}\right\} \geq 1/M^{N-1}.$$

Define the maximum possible utility (MPU) for player i with respect to utility u_i and action a_i as

$$\mathrm{MPU}(u_i, a_i) := \max_{a_{-i} \in \mathcal{A}_{-i}} u_i(a_i, a_{-i}). \tag{15}$$

Define the set of maximum possible utility best response for player i w.r.t. u_i as

$$\mathcal{B}_i(u_i) := \{a_i \in \mathcal{A}_i : \mathrm{MPU}(u_i, a_i) = \max_{a_i' \in \mathcal{A}_i} \mathrm{MPU}(u_i, a_i')\}.$$

We first conduct some simplification to the target.

$$\mathrm{SWR}_{\mathrm{N,M}}(\mathcal{F}_{\mathrm{both}}^{\mathrm{NE}}, \mathcal{F}_{\mathrm{general}}^{\mathrm{NE}}) = \inf_{\mathcal{D}} \frac{\max_{f \in \mathcal{F}_{\mathrm{both}}^{\mathrm{NE}}} \mathbb{E}_{u \sim \mathcal{D}} \mathrm{SW}(f(u), u)}{\max_{f \in \mathcal{F}_{\mathrm{general}}^{\mathrm{NE}}} \mathbb{E}_{u \sim \mathcal{D}} \mathrm{SW}(f(u), u)} \geq \inf_{\mathcal{D}} \frac{\max_{f \in \mathcal{F}_{\mathrm{both}}^{\mathrm{NE}}} \mathbb{E}_{u \sim \mathcal{D}} \mathrm{SW}(f(u), u)}{\mathbb{E}_{u \sim \mathcal{D}} \max_{\sigma} \mathrm{SW}(\sigma, u)}.$$

Then we constrain u to be a cooperation game. For a normal form game Γ_u, we define $\tilde{u} = (\tilde{u}_i)_{i \in [n]}$ in which $\tilde{u}_i = \frac{1}{n} \sum_{i=1}^{n} u_i$. Then we have $\mathrm{SW}(\sigma, u) = \mathrm{SW}(\sigma, \tilde{u})$, which means that constraining u to be a cooperation game will induce the same social welfare. Then

$$\inf_{\mathcal{D}} \frac{\max_{f \in \mathcal{F}_{\mathrm{both}}^{\mathrm{NE}}} \mathbb{E}_{u \sim \mathcal{D}} \mathrm{SW}(f(u), u)}{\mathbb{E}_{u \sim \mathcal{D}} \max_{\sigma} \mathrm{SW}(\sigma, u)} = \inf_{\mathcal{D}} \frac{\max_{f \in \mathcal{F}_{\mathrm{both}}^{\mathrm{NE}}} \mathbb{E}_{u \sim \mathcal{D}} \mathrm{SW}(f(u), \tilde{u})}{\mathbb{E}_{u \sim \mathcal{D}} \max_{\sigma} \mathrm{SW}(\sigma, \tilde{u})}.$$

Denote f_0 be the approximator that always outputs uniform strategy on $\mathcal{B}_i(\tilde{u}_i)$ for player i. It's obvious that f_0 is both OPI and PPE because the operations from u to $f_0(u)$ are all permutation-equivariant. Then,

$$\inf_{\mathcal{D}} \frac{\max_{f \in \mathcal{F}_{\mathrm{both}}^{\mathrm{NE}}} \mathbb{E}_{u \sim \mathcal{D}} \mathrm{SW}(f(u), \tilde{u})}{\mathbb{E}_{u \sim \mathcal{D}} \max_{\sigma} \mathrm{SW}(\sigma, \tilde{u})} \geq \inf_{\mathcal{D}} \frac{\mathbb{E}_{u \sim \mathcal{D}} \mathrm{SW}(f_0(u), \tilde{u})}{\mathbb{E}_{u \sim \mathcal{D}} \max_{\sigma} \mathrm{SW}(\sigma, \tilde{u})}.$$

Ignore the infimum and the expectation operator, consider $\frac{\text{SW}(f_0(u),\tilde{u})}{\max_\sigma \text{SW}(\sigma,\tilde{u})}$ for arbitrary $\tilde{u}$, denote b be the maximum element appeared in $\tilde{u}$, then the denominator equals Nb. But for the numerator, for player i, no matter what action $a_i \in \mathcal{B}_i(\tilde{u}_i)$ she chooses, she always has probability at least $\prod_{j\neq i} \frac{1}{|\mathcal{B}_j|} \geq \frac{1}{M^{N-1}}$ to achieve revenue b, therefore inducing $\text{SW}(f_0(u),\tilde{u}) \geq \frac{Nb}{M^{N-1}}$.

Then, $\frac{\text{SW}(f_0(u),\tilde{u})}{\max_\sigma \text{SW}(\sigma,\tilde{u})} \geq \frac{1}{M^{N-1}}$, so as $\inf_\mathcal{D} \frac{\mathbb{E}_{u\sim\mathcal{D}}\text{SW}(f_0(u),\tilde{u})}{\mathbb{E}_{u\sim\mathcal{D}}\max_\sigma \text{SW}(\sigma,\tilde{u})}$, $\text{SWR}_{\text{N,M}}(\mathcal{F}^{\text{NE}}_{\text{both}})$ and $\text{SWR}_{\text{N,M}}(\mathcal{F}^{\text{NE}}_{\text{OPI}})$.

Above all,

$$\left.\begin{array}{r} \text{SWR}_{\text{N,M}}(\mathcal{F}^{\text{NE}}_{\text{OPI}}, \mathcal{F}^{\text{NE}}_{\text{general}}) \\ \text{SWR}_{\text{N,M}}(\mathcal{F}^{\text{NE}}_{\text{both}}, \mathcal{F}^{\text{NE}}_{\text{general}}) \end{array}\right\} = \frac{1}{M^{N-1}}.$$

C.3.2. PROOF OF EQUATION (2)

We next prove the theorem with respect to $\mathcal{F}^{\text{NE}}_{\text{PPE}}$ that

$$\text{SWR}_{\text{N,M}}(\mathcal{F}^{\text{NE}}_{\text{PPE}}, \mathcal{F}^{\text{NE}}_{\text{general}}) \leq \frac{1}{M}.$$

Consider a bimatrix game and $\mathcal{A}_i = [M]$ for $i \in [2]$. $\forall a \in \mathcal{A}, i \in [2]$, define the payoff $\bar{u}$ as follows:

$$\bar{u}_i(a) = \begin{cases} 1 & \text{, if } a_1 = a_2. \\ 0 & \text{, otherwise.} \end{cases}$$

Define $U := \{u' | u' = \rho_1 \rho_2 \bar{u}, \rho_i \in \mathcal{G}_i\}$ and $\mathcal{D}$ as a uniform distribution on U. Easy to certify that $U = \{u' | u' = \rho_1 \bar{u}, \rho_1 \in \mathcal{G}_1\} = \{u' | u' = \rho_2 \bar{u}, \rho_2 \in \mathcal{G}_2\}$ and $\mathcal{D}$ is a permutation-invariant distribution.

Let $\tilde{f} \in \tilde{\mathcal{F}}^{\text{NE}}_{\text{general}}$ be the NE oracle that $\tilde{f}(\bar{u})_i = 1$ and for any $u' = \circ_i \rho_i \bar{u} \in U$, $\tilde{f}(u')_i = \rho_i(1)$. Intuitively, the oracle will choose the action that will provide all players with revenue of 1, leading to a social welfare of 2.

For a permutation ϱ on $[M]$, let $P_\varrho \in \{0,1\}^{M\times M}$ be the corresponding permutation matrix. Denote $\mathcal{P}$ as the set of all permutation matrice. As a result, $\forall u \in U, \forall \rho_1 \in \mathcal{G}_1, \rho_1 u = (P_{\rho_1} u_1, P_{\rho_1} u_2) =: P_{\rho_1} u$ and $\forall \rho_2 \in \mathcal{G}_2, \rho_2 u = (u_1 P^T_{\rho_2}, u_2 P^T_{\rho_2}) =: u P^T_{\rho_2}$. Specially, we have $P_\varrho \bar{u} P^T_\varrho = \bar{u}$. For $f \in \mathcal{F}^{\text{NE}}_{\text{PPE}}$, Denote $f(\bar{u}) = \sigma = (\sigma_1, \sigma_2)$. For permutation ϱ in $[M]$ and payoff $u' = P_\varrho \bar{u} = \bar{u}(P^T_\varrho)^{-1}$, by PPE of f, we have $f(u')_1 = f(P_\varrho \bar{u})_1 = P_\varrho \sigma_1 = \varrho\sigma_1$, and $f(u')_2 = f(\bar{u}(P^T_\varrho)^{-1})_2 = (P_\varrho)^{-1}\sigma_2 = \varrho^{-1}\sigma_2$. Then we have

$$\text{SW}(f(u'), u') = \sum_{i=1}^n (P_\varrho \bar{u})_i(\varrho\sigma_1, \varrho^{-1}\sigma_2) = \sum_{i=1}^n \bar{u}_i(\sigma_1, \varrho^{-1}\sigma_2) = \sum_{i=1}^n (\bar{u}P^T_\varrho)_i(\sigma_1, \sigma_2) = \text{SW}(f(\bar{u}), \bar{u}P^T_\varrho).$$

Therefore

$$\mathbb{E}_{u\sim\mathcal{D}}\text{SW}(f(u), u) = \frac{1}{|U|} \sum_{u'\in U} \text{SW}(f(u'), u')$$

$$= \frac{1}{|U|} \sum_{P_\varrho \in \mathcal{P}} \text{SW}(f(\bar{u}), \bar{u}P^T_\varrho)$$

$$= \frac{1}{|U|} \sum_{u=\bar{u}(P^T_\varrho)\in U} \text{SW}(f(\bar{u}), u)$$

$$= \frac{1}{|U|}\text{SW}(\sigma, \sum_{u'\in U} u').$$

Since $|U| = \frac{1}{M!}$ and $\sum_{u'\in U} u'$ is a tensor with all elements equal to $(M-1)!$. Thus $\mathbb{E}_{u\sim\mathcal{D}}\text{SW}(f(u), u) = \frac{2}{M}$ and

$$\text{SWR}_{\text{N,M}}(\mathcal{F}^{\text{NE}}_{\text{PPE}}, \mathcal{F}^{\text{NE}}_{\text{general}}) \leq \frac{1}{M}.$$

C.3.3. PROOF OF EQUATION (4)

Consider a 3×3 game as follows, where $\epsilon \in (0, 1/2)$:

$$u = \begin{bmatrix} \mathbf{1,1} & 0,0 & 0,1/2+\varepsilon \\ 0,0 & \mathbf{1,1} & 0,1/2+\varepsilon \\ 1/2+\varepsilon,0 & 1/2+\varepsilon,0 & \varepsilon,\varepsilon \end{bmatrix}.$$

It is obvious that $\max_{\sigma^* \subseteq \text{NE}(\Gamma_u)} \text{SW}(\sigma^*, u) = 2$, and the corresponding strategy has been bolded. However, for NE oracles with both PPE and OPI, it can only output a unique NE with a pure strategy that induces utility $(\varepsilon, \varepsilon)$.

Let $\rho_1 = \rho_2 = (2, 1, 3)$, we have $\rho_1\rho_2 u = u$. From the analysis above we know if $f^{\text{NE}} \in \widetilde{\mathcal{F}}^{\text{NE}}_{\text{both}}$ and $f^{\text{NE}}(u) = (\sigma_1, \sigma_2)$, then $\sigma_1(1) = \sigma_1(2), \sigma_2(1) = \sigma_2(2)$. We integrate the first two actions of player 1 and player 2 into a new action that will choose randomly between the first two actions, then we form the utility matrix below:

$$\overline{u} = \begin{bmatrix} 1/2, 1/2 & 0,1/2+\varepsilon \\ 1/2+\varepsilon,0 & \boldsymbol{\varepsilon},\boldsymbol{\varepsilon} \end{bmatrix}.$$

There is a unique NE in this Prisoner's Dilemma, which has been bolded. The game $\overline{u}$ is the same with the game u under the assumption that $\sigma_1(1) = \sigma_1(2)$ and $\sigma_2(1) = \sigma_2(2)$ in u. Then $\max_{f \in \widetilde{\mathcal{F}}^{\text{NE}}_{\text{both}}} \text{SW}(f(u), u) = 2\varepsilon$. Since ε can be arbitrarily small, we have $\text{SWR}_{2,3}(\widetilde{\mathcal{F}}^{\text{NE}}_{\text{both}}, \widetilde{\mathcal{F}}^{\text{NE}}_{\text{general}}) = 0$. As a result, we have $\text{SWR}_{N,M}(\widetilde{\mathcal{F}}^{\text{NE}}_{\text{both}}, \widetilde{\mathcal{F}}^{\text{NE}}_{\text{general}}) = 0$ for all $N \geq 2$ and $M \geq 3$.

D. Experiments of NE Approximation Performance

In this section, we present experimental results that assess the approximation performance of orbit-averaged NE approximators. Our primary objective is to verify the validity of Theorem 5.7, 5.8, and 5.9. Additionally, we examine the social welfare performance observed in each experiment.

D.1. Experimental Setup

We employ a 5-layer fully connected neural network as the NE approximator, with 512 nodes in each hidden layer. ReLU serves as the activation function for each hidden layer, and batch normalization is applied before activation. During training, the Nash approximation loss function is utilized, and the model is optimized using the Adam optimizer. A batch size of 1024 is employed, with 65536 data points used for training and 1000 data points for testing purposes. Data is generated on-site for each training epoch. The initial learning rate is set to 5×10^{-4}, with a decay ratio of $\gamma = 0.3$ at the 60th and 80th epochs.

Following the training phase, we proceed to test four different versions of the model: Raw, PPE, OPI, and BOTH. The Raw model represents the original network, while the PPE model applies operator $\mathcal{P}$ to the Raw model. Similarly, the OPI model applies operator $\mathcal{O}$ to the Raw model, and the BOTH model applies both operator $\mathcal{P}$ and $\mathcal{O}$ to the Raw model.

Due to the exponentially increasing computational complexity associated with orbit averaging, we limit the game size to be sufficiently small. In our experiments, we select a bimatrix game of size 4×4.

D.2. Experimental Results

Different distribution. In this experiment, we assess the performance of the NE approximator on games with different payoff distributions, namely Uniform, Gaussian, and Exponential distributions. In the Uniform distribution, each element in u is independently and identically sampled from a uniform distribution over the interval $[0, 1]$. For the Gaussian Games, each element in u is independently and identically sampled from a Gaussian distribution with mean 0 and standard deviation 1. In the Exponential Games, each element in u is independently and identically distributed according to an exponential distribution with rate parameter 1. The experimental results are presented in Table 1. We observe that the NE approximations of the PPE and OPI models are consistently higher than those of the BOTH model, while the BOTH model yields higher NE approximations compared to the RAW model. Furthermore, the social welfare decreases after orbit averaging in comparison to the RAW model, aligning with our theoretical findings. Overall, the experimental results align with Theorem 5.8 and Theorem 5.9.

	Uniform		Gaussian		Exponential	
	Approximation	Social Welfare	Approximation	Social Welfare	Approximation	Social Welfare
Raw	0.0142±0.0005	1.5211±0.0048	0.0470±0.0010	1.7695±0.0295	0.0422±0.0022	3.8628±0.0410
PPE	0.0224±0.0012	1.4905±0.0086	0.0725±0.0017	1.6859±0.0281	0.0658±0.0021	3.7801±0.0448
OPI	0.0227±0.0012	1.4905±0.0089	0.0714±0.0026	1.6876±0.0298	0.0645±0.0020	3.7792±0.0448
Both	0.0196±0.0005	1.4908±0.0087	0.0630±0.0011	1.6868±0.0309	0.0581±0.0017	3.7773±0.0466

Table 1. Approximation and social welfare of games from different payoff distributions. Each experiment is repeated five times, and the average results along with the standard deviation are presented.

	General		Cooperation		Zero-Sum		Permutation-ρ-Invariant	
	Approximation	Social Welfare	Approximation	Social Welfare	Approximation	Social Welfare	Approximation	Social Welfare
Raw	0.0142±0.0005	1.5211±0.0048	0.0035±0.0002	1.8226±0.0033	0.0138±0.0006	0.0000	4.6404E-5±1.3778E-5	0.9996±0.0044
PPE	0.0224±0.0012	1.4905±0.0086	0.1660±0.0081	1.4238±0.0100	0.0120±0.0007	0.0000	3.5154E-6±1.2956E-6	0.9996±0.0044
OPI	0.0227±0.0012	1.4905±0.0089	0.1657±0.0064	1.4247±0.0085	0.0119±0.0004	0.0000	4.4566E-5±1.2887E-5	0.9996±0.0044
Both	0.0196±0.0005	1.4908±0.0087	0.1056±0.0007	1.4224±0.0064	0.0116±0.0005	0.0000	2.9783E-7±1.2921E-7	0.9996±0.0044

Table 2. Approximation and social welfare of different game classes. Each experiment is repeated five times, and the average results along with the standard deviation are presented.

Different Game Type. In this experiment, we evaluate the performance of the NE approximator on different game classes, including General, Cooperation, Zero-Sum, and Permutation-ρ-Invariant games. In General Games, each element in u is independently and identically distributed from a uniform distribution over the interval $[0, 1]$. In Cooperation Games, we first generate u_1, and then set $u_2 = u_1$. In Zero-Sum Games, we first generate u_1, and then set $u_2 = -u_1$. In Permutation-ρ-Invariant Games, both u_1 and u_2 are set to be permutation-ρ-invariant, where $\rho = \rho_1 \rho_2$ and $\rho_i \sim U(\mathcal{R})$, with $\mathcal{R}$ being the set of circulations of $[n]$ and $n = 4$. The choice of these game classes ensures that each game distribution is permutation-invariant, which aligns with the assumptions of our theorems. The BOTH model consistently achieves smaller approximations compared to the PPE and OPI models. In Zero-Sum games, both the PPE and OPI models yield smaller approximations than the RAW model, which aligns with the findings of Theorem 5.7. In Permutation-ρ-Invariant Games, all models converge to the ordinary solution $\sigma_1 = \sigma_2 = (\frac{1}{n}, \ldots, \frac{1}{n})$, resulting in almost zero error, regardless of the inputs. These experimental results provide empirical support for the theoretical findings presented in our theorems.

Different Training Mode. In this part, we train the four versions of the model, and in testing, we also transform the trained model into the four versions. By doing so, we get $4 \times 4 = 16$ different experimental results. Due to the exponential computation costs, we set the game to be 3×3. Moreover, we set the hidden size to be 128 and the data size per epoch to be 20000 to deduce costs. The results are presented in Table 2, where the diagonal elements, representing the same training and testing approach, are highlighted in bold. We observed that in RAW training mode, both the PPE and OPI models showed larger approximations compared to the RAW model. However, when trained in their respective modes, both the PPE and OPI models outperformed the RAW model in terms of approximation. These findings highlight the impact of training mode on the model's performance and demonstrate that both the PPE and OPI models, in their respective training modes, achieved better approximation results compared to the RAW model in RAW training mode.

E. Experiments of Social Welfare

In Theorem 6.7, we discussed the limitation of NE approximators in terms of social welfare. To prove this theorem, we constructed a bimatrix game with the following payoff matrix:

$$\begin{bmatrix} 1,1 & 0,0 & 0,1/2+\varepsilon \\ 0,0 & 1,1 & 0,1/2+\varepsilon \\ 1/2+\varepsilon,0 & 1/2+\varepsilon,0 & \varepsilon,\varepsilon \end{bmatrix},$$

where $\varepsilon \in (0, 1/2)$. We showed that although the NE with the largest social welfare in this game can achieve a welfare of 2, our proof demonstrated that the maximum social welfare of the exact NEs that any equivariant (PPE and OPI) NE approximators can find is only 2ε.

Train / Test	Raw $\mathcal{E}$	Raw SW	PPE $\mathcal{E}$	PPE SW	OPI $\mathcal{E}$	OPI SW	BOTH $\mathcal{E}$	BOTH SW
Raw	**0.0122**	**1.4395**	0.0515	1.4311	0.0462	1.4334	0.0543	1.4397
PPE	0.0148	1.4291	**0.0091**	**1.4324**	0.0421	1.4318	0.0323	1.4353
OPI	0.0145	1.4288	0.0427	1.4314	**0.0091**	**1.4304**	0.0278	1.4339
BOTH	0.0125	1.4283	0.0078	1.4324	0.0078	1.4323	**0.0053**	**1.4355**

Table 3. Approximation and social welfare under different training and testing mode.

Model	Social Welfare	Approximation
General	1.9930	9.9655×10^{-4}
Equivariant	0.1854	9.7611×10^{-4}

Table 4. Social welfare and corresponding approximation on the constructed games in the proof of Theorem 6.7.

On top of that, we conduct experiments on the game to empirically verify the findings of Theorem 6.7.

E.1. Experimental Setup

We set our general model as a 5-layer fully connected neural network with 256 nodes in each hidden layer. ReLU is used as the activation function for each hidden layer, and batch normalization is added before the activation. The equivariant model we use is the (PPE and OPI) orbit-averaged general model. Both models are trained using a constrained optimization framework, where the learning problem is formulated as follows:

$$\max_{\theta} \quad \mathbb{E}_{u \sim \mathcal{D}}[SW(f_{\theta}(u), u)],$$
$$\text{s.t} \quad \mathbb{E}_{u \sim \mathcal{D}}[\mathcal{E}(f_{\theta}(u), u)] = 0.$$

In this formulation, we introduce an objective to maximize social welfare while maintaining the constraint of minimizing approximation. This encourages both models to identify NE solutions with high social welfare.

Similar to previous works (Dütting et al., 2019; Duan et al., 2022), we employ the Augmented Lagrangian Multiplier Method to optimize our model. The loss function with respect to the training set S is defined as follows:

$$L_{\rho}(\theta; \lambda) = -\frac{1}{m} \sum_{u \in S} SW(f_{\theta}(u), u) + \frac{\lambda}{m} \sum_{u \in S} \mathcal{E}(f_{\theta}(u), u) + \frac{\rho}{2} \left(\frac{1}{m} \sum_{u \in S} \mathcal{E}(f_{\theta}(u), u) \right)^2,$$

where $\lambda > 0$ is the Lagrange multiplier, m is the data size, and $\rho > 0$ is a hyperparameter controlling the penalty weight for constraint violation. During optimization, we alternate between updating the model parameters using gradient descent and updating the Lagrange multipliers using the following equation:

$$\lambda^{new} = \lambda^{old} + \rho \cdot \frac{1}{m} \sum_{u \in S} \mathcal{E}(f_{\theta}(u), u).$$

The initial value of the Lagrange multiplier λ is set to 1, and the initial value of ρ is set to 1 and increased by 5 after each epoch.

We utilize the Adam optimizer with a batch size of 512 to train our model. The initial learning rate is set to 5×10^{-4}. Each epoch involves training on 4096 data points and testing on 1000 data points. Data is generated on-site for each training epoch. Specifically, the training and testing data are sampled from a fixed distribution, where the ε parameter of the game is sampled from a uniform distribution with a minimum of 0.005 and a maximum of 0.01 (i.e., $U(0.005, 0.01)$).

E.2. Experimental Results

The results of our experiments are presented in Table 4. Both approaches demonstrate good approximations, with an approximation loss of less than 1e-3. However, there is a notable difference in the achieved social welfare between the

general model and the equivariant model. While the general model achieves nearly the highest social welfare, the equivariant model can only achieve significantly lower social welfare. It is important to highlight that both models were encouraged to find NE solutions with higher social welfare. The reason why the equivariant model achieves higher social welfare than 2ϵ is that it finds an approximate NE rather than an exact NE. This result suggests that the equivariant model may sometimes prioritize other factors over social welfare.

In summary, our findings demonstrate that even in this small-scale setting, equivariant approximators can exhibit limitations in terms of social welfare, particularly in extreme cases.