Fundamental Limits of Game-Theoretic LLM Alignment: Smith Consistency and Preference Matching

Anonymous Author(s)
Affiliation
Address
email

Abstract

Nash Learning from Human Feedback (NLHF) is a game-theoretic framework for aligning large language models (LLMs) with human preferences by modeling learning as a two-player zero-sum game. However, using raw preference as the payoff in the game highly limits the potential of the game-theoretic LLM alignment framework. In this paper, we systematically study using what choices of payoff based on the pairwise human preferences can yield desirable alignment properties. We establish necessary and sufficient conditions for Condorcet consistency, diversity through mixed strategies, and Smith consistency. These results provide a theoretical foundation for the robustness of game-theoretic LLM alignment. Further, we show the impossibility of preference matching—i.e., no smooth and learnable mappings of pairwise preferences can guarantee a unique Nash equilibrium that matches a target policy, even under standard assumptions like the Bradley-Terry-Luce (BTL) model. This result highlight the fundamental limitation of game-theoretic LLM alignment.

1 Introduction

Large language models (LLMs), such as OpenAI-o3 (OpenAI, 2025) and DeepSeek-R1 (DeepSeek-AI et al., 2025), have demonstrated impressive capabilities across a wide range of domains, including code generation, data analysis, elementary mathematics, and reasoning (Hurst et al., 2024; Anthropic, 2024; Chowdhery et al., 2023; Touvron et al., 2023; Ji et al., 2025). These models are increasingly being used to tackle previously unsolved mathematical problems, drive scientific and algorithmic discoveries, optimize complex code-bases, and support decision-making processes that were once considered unlikely to be automated in the near future (Bubeck et al., 2023; Eloundou et al., 2024; Novikov et al., 2025).

A key factor behind the popularity and effectiveness of LLMs is alignment: the process by which models learn to interact with human users and accommodate diverse human opinions and values by aligning their outputs with human preferences (Christiano et al., 2017). The traditional method for alignment, reinforcement learning from human feedback (RLHF) (Ouyang et al., 2022; Casper et al., 2023; Dong et al., 2024), typically begins by training a reward model on preference data collected from human labelers, often using the Bradley-Terry-Luce (BTL) model (Bradley and Terry, 1952; Luce, 2012),

$$\mathcal{P}(y \succ y' \mid x) = \frac{\exp(r(x,y))}{\exp(r(x,y)) + \exp(r(x,y'))}, \tag{1.1}$$

where $r(x,y)$ is the reward function and $\mathcal{P}(y \succ y'|x)$ is pairwise human preference, i.e., the fraction of individuals who prefer y over y' under prompt x. In this framework, a higher scalar score assigned

Submitted to 39th Conference on Neural Information Processing Systems (NeurIPS 2025). Do not distribute.

by the reward model to an LLM-generated response indicates a stronger preference by human labelers. The LLM is then fine-tuned through maximizing the reward to produce responses that are more likely to align with these preferences. However, Munos et al. (2024) points out that reward model can not deal with preference with cycles, and proposes an alternative alignment approach called Nash learning from human feedback (NLHF). Unlike the reward-based methods, NLHF directly uses preference data to train a preference model and formulates LLM finetuning as finding Nash equilibrium in a two-player zero-sum game, also known as a von Neumann game (Myerson, 2013). Specifically, for a given prompt x, the LLM's policy $\boldsymbol{\pi}$ competes against an opposing policy $\boldsymbol{\pi}'$ in a pairwise preference contest, where the objective is to find a policy that maximizes its worst-case preference score. Formally, NLHF solves the following min-max optimization problem:

$$\max_{\boldsymbol{\pi}} \min_{\boldsymbol{\pi}'} \mathbb{E}_{x\sim\rho}\left[\mathbb{E}_{y\sim\boldsymbol{\pi}(\cdot|x),y'\sim\boldsymbol{\pi}'(\cdot|x)}\left[\mathcal{P}\left(y \succ y' \mid x\right)\right]\right],$$

where ρ is a given distribution over prompts. However, Munos et al. (2024) does not demonstrate the advantages of using the preference as the payoff in the game.

Recently, criteria from both social choice theory (Conitzer et al., 2024; Dai and Fleisig, 2024; Mishra, 2023) and principles related to diversity (Xiao et al., 2024; Chakraborty et al., 2024) has been increasingly employed to scrutinize the alignment of LLM with human preference. Notably, RLHF has been shown to fail both social choice theory considerations (Noothigattu et al., 2020; Siththaranjan et al., 2024; Ge et al., 2024; Liu et al., 2025) and diversity considerations (Xiao et al., 2024; Chakraborty et al., 2024). In contrast, NLHF has been proved to enjoy these desirable properties. It is shown in Maura-Rivero et al. (2025); Liu et al. (2025) that NLHF is *Condorcet consistent* (see Definition 3.2), meaning that the method always outputs the Condorcet winning response, a response that beats every other alternative response in pairwise majority comparisons, whenever one exists. Further, under a no-tie assumption (see Assumption 2.1), Liu et al. (2025) shows that NLHF is *Smith consistent* (see Definition 4.1), meaning that the method always output responses from the Smith set, the smallest nonempty set of responses that pairwise dominate all alternatives outside the set. Moreover, Liu et al. (2025) shows that when human preference is diverse, i.e., there does not exist a single response that beat every other alternative, NLHF avoid collapsing to a single response by adopting a *mixed strategy*.

Despite these advantages, there is no reason why we must use raw preference to design the payoff in the game-theoretic alignment approach. Using alternative payoffs in the game-theoretic LLM alignment framework might also lead to desirable alignment. In this work, we systematically investigate *the fundamental limits of the game-theoretic LLM alignment framework* by analyzing how various choices of payoff influence its ability to satisfy key alignment criteria. We consider the following general game-theoretic alignment problem, involving a mapping of the preference denoted by Ψ:

$$\max_{\boldsymbol{\pi}} \min_{\boldsymbol{\pi}'} \mathbb{E}_{x\sim\rho}\left[\mathbb{E}_{y\sim\boldsymbol{\pi}(\cdot|x)}\mathbb{E}_{y'\sim\boldsymbol{\pi}'(\cdot|x)}\left[\Psi\left(\mathcal{P}(y \succ y' \mid x)\right)\right]\right]. \tag{1.2}$$

The general problem (1.2) encompasses a range of games. When $\Psi(t) = t$ is the identity mapping, the objective in Equation (1.2) is equivalent to the standard NLHF objective. When $\Psi(t) = \log(t/(1-t))$ and the preference is generated by a BTL model, Equation (1.2) recovers the standard RLHF objective. More importantly, the preference model $\mathcal{P}_\theta$ used in practice is an estimation of true human preference, which can be regarded as a noisy mapping of the ground-truth preference $\mathcal{P}$. Allowing Ψ to be stochastic provides a way to account for the uncertainty and noise inherent in estimating human preferences. It is worthy to note that similar formalism has been proposed for non game-theoretic approach in Azar et al. (2024), which uses an non-decreasing mapping to process the preference.

In this paper, we first discuss when the solution to problem (1.2) is Condorcet consistent and Smith consistent in Section 3 and 4 respectively. Our results show that these desirable properties is insensitive to the exact value of the payoff, revealing the robustness of game-theoretic alignment approaches. As a special case, we discover a natural generalization of RLHF objective that satisfy all these desirable properties. Technically, we develop novel proof techniques that can tackle a general non-symmetric game directly, instead of relying crucially on the symmetric nature of NLHF as in Liu et al. (2025).

In addition, we examine the diversity of the solution by investigating whether the model produces a mixed strategy (Liu et al., 2025), and whether its output can satisfy the criterion of *preference matching* (Xiao et al., 2024), meaning that the model output exactly matches a target policy which

fully accounts for the diversity of human preference. Our findings suggest diversity can be ensured by mixed strategies, but exactly matching a target is difficult for any game-theoretic alignment approach. This reveal a fundamental limitation of game-theoretic alignment approaches.

1.1 Summary of Contributions

We summarize our contributions as follows:

- We show that Condorcet consistency is insensitive to the exact value of the payoff (Theorem 3.1), revealing the robustness of game-theoretic alignment approaches.
- We show that Smith consistency can be ensured by further maintaining the symmetry of the game (Theorem 4.2). Moreover, Smith consistent methods automatically preserve the diversity in human preferences by adopting mixed strategies (Corollary 4.2).
- We show that preserving the diversity in human preference strictly, in the sense of preference matching, is impossible in general (Theorem 5.1). This reveals a fundamental limitation of game-theoretic alignment approaches.

Assuming Ψ is continuous, Table 1 provides a concise summary of our mathematical results.

Table 1: Summary of our mathematical results: the necessary and sufficient conditions on Ψ to guarantee certain desirable alignment properties.

Condorcet consistency	$\Psi(t) \geqslant \Psi(1/2), \forall\, 1/2 \leqslant t \leqslant 1$ and $\Psi(t) < \Psi(1/2), \forall\, 0 \leqslant t < 1/2$
Mixed & Condorcet consistency	$\Psi(t) + \Psi(1-t) \geqslant 2\Psi(1/2), \forall\, 1/2 \leqslant t \leqslant 1$ and $\Psi(t) < \Psi(1/2), \forall\, 0 \leqslant t < 1/2$
Smith consistency	$\Psi(t) + \Psi(1-t) = 2\Psi(1/2), \forall\, 1/2 \leqslant t \leqslant 1$ and $\Psi(t) < \Psi(1/2), \forall\, 0 \leqslant t < 1/2$

1.2 Related Works

A general mapping Ψ is first introduced in Azar et al. (2024) to facilitate the analysis of traditional non game-theoretic LLM alignment methodologies. Their objective function, called ΨPO, applies a general mapping Ψ to the original human preference. In this way, they treat RLHF and DPO as special cases of ΨPO under BTL model and argue that they are prone to overfitting. To avoid overfitting, they take Ψ to be identity and arrive at a new efficient algorithm called IPO. Our problem (1.2) can be regarded as the analogy of ΨPO in the context of game-theoretic LLM alignment. Another difference is that rather than focusing on statistical properties like overfitting, our focus is on the alignment properties such as Smith consistency and preference matching. Moreover, they restrict Ψ to be an non-decreasing map, while we allow Ψ to be arbitrary, even stochastic.

Condorcet consistency is one of the dominant concept in the theory of voting (Gehrlein, 2006; Balinski and Laraki, 2010), and Smith consistency is its natural generalization (Shoham and Leyton-Brown, 2008; Börgers, 2010). They are not studied in the context of LLM alignment until recently (Maura-Rivero et al., 2025; Liu et al., 2025). In Maura-Rivero et al. (2025), the authors show that NLHF with a selection probability that deals with ties is Condorcet consistent. Under a no-tie assumption, Liu et al. (2025) show that NLHF is Condorcet consistent and Smith consistent, whereas RLHF is not unless the preference satisfies a BTL model. Further, they show that the probability that the preference satisfies a BTL model is vanishing under an impartial culture assumption, highlighting a key advantage of the NLHF framework.

Several recent works also focus on aligning LLMs with the diverse human preference (Chakraborty et al., 2024; Xiao et al., 2024; Liu et al., 2025). In Chakraborty et al. (2024), the authors introduce a mixture model to account for the opinion of minority group and arrive at the MaxMin-RLHF method. In Xiao et al. (2024), the authors introduce the concept of preference matching and develop the PM-RLHF objective to pursue this goal. Liu et al. (2025) demonstrates that the original NLHF yields a mixed strategy when no Condorcet winning response exists, whereas standard RLHF produces a deterministic strategy, highlighting a potential advantage of NLHF in preserving the diversity of human preferences.

2 Preliminaries

Consider a general mapping $\Psi : [0, 1] \to \mathbb{R}$. We apply Ψ to the preference and study the max-min problem (1.2) with this generalized payoff. Any solution $\boldsymbol{\pi}$ employed by the first player at the Nash equilibrium,

$$\boldsymbol{\pi} \in \arg\max_{\boldsymbol{\pi}} \min_{\boldsymbol{\pi}'} \mathbb{E}_{x\sim\rho} \left[\mathbb{E}_{y\sim\boldsymbol{\pi}(\cdot|x)}\mathbb{E}_{y'\sim\boldsymbol{\pi}'(\cdot|x)} \left[\Psi\left(\mathcal{P}(y \succ y' \mid x)\right)\right]\right], \tag{2.1}$$

is called a Nash solution to the problem (1.2). The Nash solution is the policy which fully aligned LLMs will perform. Note that the set of Nash solutions remain the same after an overall shift of payoff, that is, changing Ψ to $\Psi + C$ for any constant C will not affect the problem. The original `NLHF` objective (Munos et al., 2024) corresponds to the special case where $\Psi(t) = t$, equivalent to $\Psi(t) = t - 1/2$, and the resulting game is symmetric (Duersch et al., 2012), meaning that the two players are the same. However, for an arbitrary mapping Ψ, the game is usually not symmetric, and we only focus on the Nash solution employed by the first player.

Given a prompt x, we consider the set of all possible responses generated by the LLM: $\{y_1, \ldots, y_n\}$, where n is the total number of possible responses. Without any loss of generality, we drop the dependence on the prompt x from now on. For any two distinct response y and y', recall that $\mathcal{P}(y \succ y')$ denote the preference of y over y', defined as the expected proportion of individuals who prefer y over y'. By definition, human preference satisfies the condition $\mathcal{P}(y \succ y') + \mathcal{P}(y' \succ y) = 1$ and naturally we let $\mathcal{P}(y \succ y) = 1/2$ (Munos et al., 2024). For any distinct pair of responses y and y', we say that y beats y' if $\mathcal{P}(y \succ y') > 1/2$. Additionally, following Liu et al. (2025), we adopt the No-Tie assumption throughout this paper.

Assumption 2.1 (No-Tie)**.** *For any distinct responses y and y', we assume that $\mathcal{P}(y \succ y') \neq 1/2$.*

This assumption is both minimal and practically reasonable. First, if the number of labelers is odd, it automatically holds. Even in cases where a tie occurs, it can always be resolved through a more precise comparison.

Notation. For any set A, we denote its cardinality by $|A|$. For any $n \in \mathbb{N}_+$, we define $[n] := \{1, \ldots, n\}$. We use $\delta_{ij} := \mathbb{1}\{i = j\}$ for $1 \leqslant i, j \leqslant n$. We represent high-dimensional vectors using bold symbols. Any policy $\boldsymbol{\pi}$ over the set of possible responses $\{y_1, \ldots, y_n\}$ can be identified with a vector in $\mathbb{R}^n$, where each entry π_i corresponds to the probability assigned to y_i for $i \in [n]$. We then define the support of a policy $\boldsymbol{\pi}$ as $\mathrm{supp}(\boldsymbol{\pi}) := \{y_i \mid \pi_i > 0\,, i \in [n]\}$. We write $\boldsymbol{\pi} > 0$ if $\pi_i > 0$ for all $i \in [n]$, and similarly, $\boldsymbol{\pi} \geqslant 0$ if $\pi_i \geqslant 0$ for all $i \in [n]$.

3 Condorcet Consistency

In this section, we examine Condorcet consistency—a desirable property for LLM alignment inspired by social choice theory—within the generalized game-theoretic LLM fine-tuning framework (1.2). We begin by defining the Condorcet winning response and Condorcet consistency. We then present Theorem 3.1, which characterizes the necessary and sufficient conditions on the mapping Ψ to guarantee Condorcet consistency. Next, we examine the conditions under which Ψ preserves human preference diversity when no Condorcet winner exists and introduce Theorem 3.2. Finally, we discuss the continuity assumption underlying Theorem 3.2.

Following Liu et al. (2025), a response that is preferred over all others in pairwise comparisons by the preference model is referred to as the Condorcet winning response.

Definition 3.1 (Condorcet Winning Response)**.** A response $y^\star$ is called a Condorcet winning response if $\mathcal{P}(y^\star \succ y) > 1/2$ for all $y \neq y^\star$.

It is clear that there can be at most one Condorcet winning response. When such a response exists, a natural requirement for LLM alignment is that this response should be the output. This property is known as Condorcet consistency.

Definition 3.2 (Condorcet Consistency)**.** Problem (1.2) is Condorcet consistent if when there exists a Condorcet winning response, the Nash solution to (1.2) is unique and corresponds to this Condorcet winning response.

Liu et al. (2025); Maura-Rivero et al. (2025) show that the original NLHF objective, which corresponds to the case where $\Psi(\cdot)$ is identity, is Condorcet consistent. In this paper, we proceed further and investigate the following question:

Which choices of Ψ ensure Condorcet consistency?

We answer this question in Theorem 3.1. The proof is provided in Appendix A.

Theorem 3.1. *Problem (1.2) is Condorcet consistent if and only if $\Psi(\cdot)$ satisfies*

$$\begin{cases} \Psi(t) \geqslant \Psi(1/2)\,, 1 \geqslant t \geqslant 1/2 \\ \Psi(t) < \Psi(1/2)\,, 1/2 > t \geqslant 0 \end{cases}. \tag{3.1}$$

Note that this condition is much weaker than requiring Ψ to be increasing. It only demands that Ψ maps any value greater than 1/2 to some value larger than $\Psi(1/2)$, and any value less than $1/2$ to some value smaller than $\Psi(1/2)$. This implies that a wide range of mapping functions can be used within the game-theoretic LLM alignment framework (1.2) to ensure Condorcet consistency. Furthermore, in practice, we do not have access to the ground-truth preference model. Instead, we parameterize the preference model using a deep neural network, $\mathcal{P}_\theta(y \succ y')$, trained on large-scale preference datasets (Munos et al., 2024). Due to the limitations of the datasets and the optimization process, the learned model only approximates the true human preferences. We can view this approximation as $\Psi(\mathcal{P}(y \succ y'))$ in our framework. In practice, we can enforce the parameterized preference model to satisfy $\mathcal{P}_\theta(y \succ y) = 1/2$, then our results show that as long as this approximation yields the correct pairwise majority comparisons—specifically, that $\mathcal{P}_\theta(y \succ y') \geqslant 1/2 > \mathcal{P}_\theta(y' \succ y)$ whenever y beats y'—then the LLM alignment remains Condorcet consistent. This strongly highlights the robustness of the game-theoretic LLM alignment approach in achieving Condorcet consistency.

When a Condorcet winning response does not exist, human preferences are diverse and there is no single response that is better than others. Therefore, in order to preserve the diversity inherent in human preferences, it is natural to require the Nash solution not to collapse to a single response. This motivation leads to the following characterization of diversity through mixed strategies.

Definition 3.3 (Mixed Strategies)**.** A Nash solution $\boldsymbol{\pi}$ is called a mixed strategy if $|\operatorname{supp}(\boldsymbol{\pi})| > 1$.

Liu et al. (2025) demonstrates that the original NLHF, which corresponds to the case where $\Psi(\cdot)$ is identity, yields a mixed strategy when no Condorcet winning response exists. Assuming that problem (1.2) is Condorcet consistent, we proceed further and investigate:

Which choices of Ψ lead to a mixed strategy in the absence of a Condorcet winning response?

We now focus on mappings Ψ that are continuous at 1/2, a condition commonly encountered in practical learning setups. Under this mild assumption, we answer this question in Theorem 3.2 and the proof is provided in Appendix C.

Theorem 3.2. *Assume that the mapping $\Psi(\cdot)$ is continuous at $1/2$. Assuming the Condorcet consistency of problem (1.2), then any Nash solution is mixed when there is no Condorcet winning response if and only if $\Psi(\cdot)$ satisfies*

$$\Psi(t) + \Psi(1-t) \geqslant 2\Psi(1/2)\,, \forall\, 0 \leqslant t \leqslant 1 \text{ and } \Psi(t) < \Psi(1/2)\,, \forall\, 0 \leqslant t < 1/2\,. \tag{3.2}$$

The first condition arises from the requirement of mixed strategies, while the second condition is a reduction of the condition inherited from Theorem 3.1 under the assumption of Condorcet consistency and the first condition.

Choices of payoff functions are harder to characterize when we relax the continuity assumption. The following example investigate a special piece-wise constant mapping, which does not satisfy the first condition in Theorem 3.2.

Example 3.4. *Let $M_- < \Psi(1/2) \leqslant M_+$ and take*

$$\Psi(t) = \begin{cases} M_-\,, & 0 \leqslant t < 1/2 \\ \Psi(1/2)\,, & t = 1/2 \\ M_+\,, & 1/2 < t \leqslant 1 \end{cases}.$$

Then, any Nash solution is mixed when there is no Condorcet winning response.

The proof of Example 3.4 is deferred to Appendix B. This example implies that choices of payoff functions are considerably richer when we relax the continuity assumption.

4 Smith Consistency

In this section, we extend the discussion of Condorcet consistency to Smith consistency. First, we define the Smith set and Smith consistency. Next, we present Theorem 4.2, which provides the necessary and sufficient condition for the mapping Ψ to ensure Smith consistency. Finally, we highlight that Smith-consistent methods inherently preserve the diversity present in human preferences and discuss the continuity assumption in Theorem 4.2.

Condorcet consistency only ensures the method capture the right response when there exists a Condorcet winning response. In general, when there is no Condorcet winning response, we can expect that there might be a set of responses satisfying similar property, generalizing Definition 3.1. Under Assumption 2.1, Liu et al. (2025) revealed a more detailed decomposition of the preference structure. Specifically, the set of responses can be partitioned into distinct groups $S_1, \ldots, S_k$, where every response in S_i is preferred over all responses in S_j for $i < j$, summarized in the following theorem.

Theorem 4.1 (Liu et al. (2025))**.** *Under Assumption 2.1, the set of responses can be partitioned into disjoint subsets $S_1, \ldots, S_k$ such that:*

1. *Each S_i either forms a Condorcet cycle or is a single response.*
2. *For any $j > i$, any response $y \in S_i$ and $y' \in S_j$, $\mathcal{P}(y \succ y') > \frac{1}{2}$.*

Moreover, this decomposition is unique.

When $|S_1| = 1$, the response in S_1 is exactly the Condorcet winning response. Thus, S_1 is the generalization of Condorcet winning response, and is referred as the Condorcet winning set in Liu et al. (2025). Traditionally, a subset with such property is also known as the Smith set in the literature of social choice theory (Shoham and Leyton-Brown, 2008). Here we choose to adopt the name Smith set to distinguish with the concept of Condorcet winning response. Given this decomposition, it is natural to desire that an aligned LLM adopts a strategy supported exclusively on the top group S_1, as any response outside S_1 is strictly less preferred than any response inside S_1. This desirable property is referred to as Smith consistency:

Definition 4.1 (Smith Consistency)**.** Problem (1.2) is Smith consistent if the support of any Nash solution is contained in the Smith set S_1.

Liu et al. (2025) showed that the original `NLHF` payoff, which corresponds to the case where $\Psi(t) = t$, is Smith consistent. Here, we investigate this question for a general mapping Ψ:

Which choices of Ψ ensure Smith consistency?

Here, similar to Theorem 3.2, we answer this question in Theorem 4.2 for mappings that is continuous at $1/2$. The proof is provided in Appendix E.

Theorem 4.2. *Suppose that the mapping $\Psi(\cdot)$ is continuous at $1/2$, problem (1.2) is Smith consistent if and only if $\Psi(\cdot)$ satisfies*

$$\Psi(t) + \Psi(1-t) = 2\Psi(1/2)\,, \forall\, t \in [0,1] \text{ and } \Psi(t) < \Psi(1/2)\,, \forall\, 0 \leqslant t < 1/2\,.$$

The first condition $\Psi(t) + \Psi(1-t) = 2\Psi(1/2)$ says nothing but the zero-sum game formed by problem (1.2) is equivalent to a symmetric two-player zero-sum game[1] (Duersch et al., 2012). By definition, Smith consistency implies Condorcet consistency because when there is a Condorcet winning response, S_1 is exactly the set whose only element is the Condorcet winning response. Thus, the second condition is just a reduction of the condition in Theorem 3.1 under the first condition. It is easy to see $\Psi(t) = t$ satisfies these conditions, and thus our result generalize Theorem 3.6 in Liu et al. (2025). More interestingly, $\Psi(t) = \log(t/(1-t))$ also satisfies these conditions. This implies that

$$\max_{\boldsymbol{\pi}} \min_{\boldsymbol{\pi}'} \mathbb{E}_{x\sim\rho}\left[\mathbb{E}_{y\sim\boldsymbol{\pi}(\cdot|x)}\mathbb{E}_{y'\sim\boldsymbol{\pi}'(\cdot|x)}\left[\log\left(\frac{\mathcal{P}(y \succ y' \mid x)}{\mathcal{P}(y' \succ y \mid x)}\right)\right]\right],$$

which is a natural generalization of standard `RLHF` when human preferences does not satisfy BTL model, is also Smith consistent.

[1] This can be seen by shifting the payoff by $\Psi(1/2)$, which leaves the Nash solution unchanged.

The set of choices for Ψ that ensure Smith consistency is quite broad. We can easily construct such a Ψ by first defining $\Psi(t)$ on $[0, 1/2]$ to satisfy $\Psi(t) < \Psi(1/2)$ for all $t \in [0, 1/2)$, and then extending it to $[0, 1]$ by setting $\Psi(t) = 2\Psi(1/2) - \Psi(1-t)$ for all $t \in (1/2, 1]$. Moreover, as discussed in Section 3, a practical preference model $\mathcal{P}_\theta(y \succ y')$ can be seen as a mapping of the ground truth preference via Ψ, i.e., $\Psi(\mathcal{P}(y \succ y'))$. Thus, the first condition in Theorem 4.2 requires the preference model to satisfy $\mathcal{P}_\theta(y \succ y') + \mathcal{P}_\theta(y' \succ y) = 1$, with $\mathcal{P}_\theta(y \succ y) = 1/2$ enforced. However, several practically used preference models (Munos et al., 2024; Jiang et al., 2023; Wu et al., 2024) do not guarantee this condition, which may cause the aligned LLM strategy to fail to satisfy Smith consistency.

As any mapping satisfying the condition in Theorem 4.2 also satisfies the condition in Theorem 3.2, we obtain the following corollary:

Corollary 4.2. *Suppose that the mapping $\Psi(\cdot)$ is continuous at $1/2$. Then if problem (1.2) is Smith consistent, any Nash solution is also mixed.*

This shows that when $|S_1| > 1$, the Nash solution to problem (1.2) with any Ψ such that Smith consistency holds will not only support on S_1 but also be a mixed strategy on S_1 without collapsing to a single response. As a conclusion, a Smith consistent method can preserve the diversity inherent in human preferences, at least partially.

Lastly, we discuss what happens if Ψ is not continuous at $1/2$. Choices of mappings Ψ are considerably richer and consequently harder to characterize when we relax the continuity assumption. The following example shows that the piece-wise constant mapping in Example 3.4 also ensures Smith consistency.

Example 4.3. *Let $M_- < \Psi(\frac{1}{2}) < M_+$, and we take*

$$\Psi(t) = \begin{cases} M_- & 0 \leqslant t < 1/2 \\ \Psi(1/2) & t = 1/2 \\ M_+ & 1/2 < t \leqslant 1 \end{cases}.$$

Then problem (1.2) is Smith consistent. The proof is provided in Appendix D.

5 Impossibility of Preference Matching

In this section, we first revisit the definition of preference matching in the BTL model (Xiao et al., 2024). Then we introduce a general theoretical framework of preference matching within the context of game-theoretic LLM alignment, and establish a general impossibility result, as stated in Theorem 5.1. Finally we apply this general result to problem (1.2), concluding that preference matching is impossible.

In previous sections, we have characterized the diversity of alignment result via mixed strategies. In Xiao et al. (2024), the authors propose a more refined criterion for diversity when the preference $\mathcal{P}(y \succ y' \mid x)$ follows a BTL model,

$$\mathcal{P}(y \succ y' \mid x) = \frac{\exp(r(x, y))}{\exp(r(x, y)) + \exp(r(x, y'))},$$

as given by (1.1). They point out that it is unwise to completely disregard any minority opinions in the case that 51% of human labelers prefer y_1 over y_2 for a binary comparison. They suggest that the policy (5.1),

$$\boldsymbol{\pi}^*(y \mid x) = \frac{\exp\left(r(x, y)\right)}{\sum_{y'} \exp\left(r(x, y')\right)}, \tag{5.1}$$

referred to as the preference-matching policy, fully accounts for the diversity in human preferences.

It is easy to see that there exists a Condorcet winning response under BTL model. According to Theorem 3.1, using preference $\Psi(\mathcal{P}(y \succ y' \mid x))$ as payoff with $\Psi(t) = t$ or $\Psi(t) = \log(t/(1-t))$ will lead the Nash solution to collapse to a single response instead of matching with $\boldsymbol{\pi}^*$. This shows that both `RLHF` and `NLHF` do not accounts for the diversity inherent in human preferences from the perspective of preference matching (Xiao et al., 2024; Liu et al., 2025), even under BTL model.

To achieve alignment fully accounting for diversity, we would like to match the Nash solution with the desired policy $\boldsymbol{\pi}^*$. In Xiao et al. (2024), the authors answered this question for `RLHF`. They proposed

the preference matching RLHF (PM-RLHF) method which successfully achieves preference matching, by slightly modifying the RLHF objective. Here, we aim to explore the possibility of designing a new learnable payoff matrix that aligns with the desired strategy in a game-theoretic framework for LLM alignment:

Which choices of Ψ ensure preference matching?

Although it is currently unknown how to generalize the notion of preference matching policy to a general non-BTL preference, to maintain the generality of the discussion and drop the BTL model assumption, we suppose there exists an ideal policy, denoted by $\boldsymbol{\pi}^*$, which captures the diversity of human preferences perfectly.

Given a prompt x, we consider the set of all possible responses generated by the LLM: $\{y_1, \ldots, y_n\}$. We further suppose that the policy $\boldsymbol{\pi}^*$ has full support over these n responses, meaning $\boldsymbol{\pi}^* > 0$, as we exclude responses not supported by $\boldsymbol{\pi}^*$ from consideration. Then our goal is to construct a game, represented by a payoff matrix $\{\alpha_{ij}\}_{i,j=1}^n$, with its Nash solution the given policy $\boldsymbol{\pi}^*$, i.e.,

$$\boldsymbol{\pi}^* = \arg\max_{\boldsymbol{\pi}} \min_{\boldsymbol{\pi}'} \sum_{i=1}^n \sum_{j=1}^n \alpha_{ij} \pi_i \pi'_j \,.$$

To answer this question, we characterize the Nash solution under the given payoff matrix $\{\alpha_{ij}\}_{i,j=1}^n$, which is summarized by the following KKT condition. The proof is deferred to Appendix F.1.

Lemma 5.1 (KKT Condition)**.** *Consider a game with payoff matrix $\{\alpha_{ij}\}_{i,j=1}^n$. Then $\boldsymbol{\pi}^* > 0$ is a Nash solution to the game if and only if there exists $\boldsymbol{u}^* \in \mathbb{R}^n$ with $\boldsymbol{u}^* \geqslant 0$ and $\sum_{i=1}^n u_i^* = 1$, and $t^* \in \mathbb{R}$ such that the following KKT conditions hold:*

$$\begin{cases} \sum_{i=1}^n \pi_i^* \alpha_{ij} - t^* \leqslant 0 & j = 1, \cdots, n \\ u_j^* \left(\sum_{i=1}^n \pi_i^* \alpha_{ij} - t^*\right) = 0 & j = 1, \cdots, n \\ \sum_{j=1}^n \alpha_{ij} u_j^* = t^* & i = 1, \cdots, n \end{cases} .$$

According to Lemma 5.1, it is easy to verify that the payoff matrix

$$\alpha_{ij} = \pi_i^* + \pi_j^* - \delta_{ij} \,, \forall\, 1 \leqslant i, j \leqslant n \,, \tag{5.2}$$

and the payoff matrix

$$\alpha_{ij} = -\frac{\pi_j^*}{\pi_i^*} + n\delta_{ij} \,, \forall\, 1 \leqslant i, j \leqslant n \,, \tag{5.3}$$

both guarantee that $\boldsymbol{\pi}^*$ is a Nash solution (the details are provided in Appendix F.2). However, these payoff matrices do not depend on the given policy $\boldsymbol{\pi}^*$ in a reasonable way. The payoff matrix in Equation (5.2) is symmetric, making it difficult to interpret. Even worse, it depends on the raw value of $\boldsymbol{\pi}^*$. In practice, $\boldsymbol{\pi}^*$ is often only known up to a normalizing constant. For instance, the preference matching policy (5.1) includes a normalizing constant in the denominator that involves summing over n terms. This constant is hard to determine when n is large and unknown, as is often the case in LLMs. The payoff matrix in Equation (5.3) faces a similar issue as it explicitly depends on n, which is an extremely large and unknown value in practice.

In summary, the above two payoff matrices rely on information that is often unavailable in practice, such as n and the raw value of $\boldsymbol{\pi}^*$. What we can obtain in practice for the design of α_{ij} is the preference information between two responses y_i and y_j, which we assume depends solely on the ratio between π_i^* and π_j^*. When the preference satisfies the BTL model (1.1), this assumption is justified by the fact that the preference between any two responses depends solely on the ratio of the values assigned by their corresponding preference matching policies (5.1). From this practical consideration, we assume that the payoff matrix satisfies the following assumptions:

Assumption 5.2. *Given any $\boldsymbol{\pi}^* > 0$, the payoff matrix $\{\alpha_{ij}\}_{i,j=1}^n$ satisfies the following conditions:*

1. *For all $i \in [n]$, $\alpha_{ii} = C$ where C is a constant independent of $\boldsymbol{\pi}^*$ and n. In other words, the diagonal elements are the same constant.*

2. *For all $i, j \in [n]$ with $i \neq j$, $\alpha_{ij} = f\left(\frac{\pi^*_i}{\pi^*_j}\right)$ for some smooth function f that is independent of $\boldsymbol{\pi}^*$ and n. In other words, the off-diagonal elements depend on the ratio $\frac{\pi^*_i}{\pi^*_j}$ in the same way for all pairs (i, j) with $i \neq j$.*

We emphasize that the above two assumptions are crucial for constructing a meaningful and practically learnable payoff matrix. Furthermore, for effective alignment, the payoff matrix should not only ensure $\boldsymbol{\pi}^*$ to be a Nash solution, but $\boldsymbol{\pi}^*$ must be the only Nash solution. The uniqueness requirement excludes trivial payoff matrices such as $\alpha_{ij} = C$, where every $\boldsymbol{\pi}^* > 0$ is a Nash solution.

Unfortunately, in Theorem 5.1, we prove that such a payoff matrix $\{\alpha_{ij}\}_{i,j=1}^n$ does not exist generally. The proof can be found in Appendix F.3.

Theorem 5.1 (Impossibility of Preference Matching for General Payoffs)**.** *There does not exist a payoff matrix $\{\alpha_{ij}\}_{i,j=1}^n$ satisfying Assumption 5.2 such that for any given $\boldsymbol{\pi}^* > 0$, the Nash solution to the game is unique and equals to $\boldsymbol{\pi}^*$.*

Remark 5.3. If we relax Assumption 5.2 and allow the entries of the payoff matrix to depend on n, then the design (5.3) is actually eligible for preference matching.

Theorem 5.1 implies that no simple mapping of the preference can yield a payoff that leads to preference matching. As a special case of Theorem 5.1, under BTL model, the generalized game in Equation (1.2) with a smooth mapping Ψ,

$$\max_{\boldsymbol{\pi}} \min_{\boldsymbol{\pi}'} \mathbb{E}_{x\sim\rho}\left[\mathbb{E}_{y\sim\boldsymbol{\pi}(\cdot|x)}\mathbb{E}_{y'\sim\boldsymbol{\pi}'(\cdot|x)}\left[\Psi\left(\mathcal{P}(y \succ y' \mid x)\right)\right]\right],$$

cannot achieve preference matching. Therefore, we obtain the following corollary.

Corollary 5.4. *Problem (1.2) with smooth mapping Ψ cannot achieve preference matching.*

6 Conclusion

We investigate several properties motivated by social choice theory and diversity considerations within the general game-theoretic LLM alignment framework (1.2), where the payoff is designed as a mapping Ψ of the original preference. We identify the necessary and sufficient conditions on Ψ to guarantee Condorcet consistency and Smith consistency. These conditions allow for a considerably broad class of choices for Ψ, demonstrating that these desirable alignment properties are not sensitive to the exact values of the payoff, thereby providing a theoretical foundation for the robustness of the game-theoretic LLM alignment approach. Additionally, we examine conditions on Ψ that ensure the resulting policy is a mixed strategy, preserving diversity in human preferences. Finally, we prove that achieving exact preference matching is impossible under the general game-theoretic alignment framework with a smooth mapping, revealing fundamental limitations of this approach.

Limitations and Discussion. Our findings suggest several promising directions for future research on LLM alignment. First, while we establish an impossibility result for preference matching under the assumption that Ψ is smooth, it remains an open question whether preference matching can be achieved when Ψ is merely continuous. Second, in practical settings, regularization terms based on the reference model are often added to problem (1.2). Regularization may be crucial for preference matching, for example, Xiao et al. (2024) modify the regularization term in `RLHF` to achieve preference matching. Analyzing the alignment properties of game-theoretic methods with such regularization is another interesting avenue for future work. Furthermore, how to explicitly define a preference-matching policy for general preferences that do not satisfy the BTL model, and how to develop alignment approaches capable of learning such a policy, remain open problems. Finally, our results highlight that practical preference models must satisfy certain anti-symmetry conditions to ensure Smith consistency — conditions that are not guaranteed by several currently used models. Thus, designing preference model architectures that enforce anti-symmetry is an important and interesting future direction.

Broader Impacts. The goal of this paper is to investigate several theoretical properties of the general game-theoretic LLM alignment approach. There are many potential societal consequences of our work, none of which we feel must be specifically highlighted here.

References

Anthropic, A. (2024). The claude 3 model family: Opus, sonnet, haiku. *Claude-3 Model Card*.

Azar, M. G., Guo, Z. D., Piot, B., Munos, R., Rowland, M., Valko, M., and Calandriello, D. (2024). A general theoretical paradigm to understand learning from human preferences. In *International Conference on Artificial Intelligence and Statistics*, pages 4447–4455. PMLR.

Balinski, M. L. and Laraki, R. (2010). *Majority judgment : measuring, ranking, and electing*. MIT Press.

Börgers, C. (2010). *Mathematics of social choice: voting, compensation, and division*. Society for Industrial and Applied Mathematics.

Bradley, R. A. and Terry, M. E. (1952). Rank analysis of incomplete block designs: I. the method of paired comparisons. *Biometrika*, 39(3/4):324–345.

Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., Lee, P., Lee, Y. T., Li, Y., Lundberg, S., Nori, H., Palangi, H., Ribeiro, M. T., and Zhang, Y. (2023). Sparks of artificial general intelligence: Early experiments with gpt-4. *arXiv preprint arXiv:2303.12712*.

Casper, S., Davies, X., Shi, C., Gilbert, T. K., Scheurer, J., Rando, J., Freedman, R., Korbak, T., Lindner, D., Freire, P., Wang, T. T., Marks, S., Segerie, C.-R., Carroll, M., Peng, A., Christoffersen, P. J., Damani, M., Slocum, S., Anwar, U., Siththaranjan, A., Nadeau, M., Michaud, E. J., Pfau, J., Krasheninnikov, D., Chen, X., Langosco, L., Hase, P., Biyik, E., Dragan, A., Krueger, D., Sadigh, D., and Hadfield-Menell, D. (2023). Open problems and fundamental limitations of reinforcement learning from human feedback. *Transactions on Machine Learning Research*.

Chakraborty, S., Qiu, J., Yuan, H., Koppel, A., Huang, F., Manocha, D., Bedi, A. S., and Wang, M. (2024). Maxmin-rlhf: Towards equitable alignment of large language models with diverse human preferences. *arXiv preprint arXiv:2402.08925*.

Chowdhery, A., Narang, S., Devlin, J., Bosma, M., Mishra, G., Roberts, A., Barham, P., Chung, H. W., Sutton, C., Gehrmann, S., Schuh, P., Shi, K., Tsvyashchenko, S., Maynez, J., Rao, A., Barnes, P., Tay, Y., Shazeer, N., Prabhakaran, V., Reif, E., Du, N., Hutchinson, B., Pope, R., Bradbury, J., Austin, J., Isard, M., Gur-Ari, G., Yin, P., Duke, T., Levskaya, A., Ghemawat, S., Dev, S., Michalewski, H., Garcia, X., Misra, V., Robinson, K., Fedus, L., Zhou, D., Ippolito, D., Luan, D., Lim, H., Zoph, B., Spiridonov, A., Sepassi, R., Dohan, D., Agrawal, S., Omernick, M., Dai, A. M., Pillai, T. S., Pellat, M., Lewkowycz, A., Moreira, E., Child, R., Polozov, O., Lee, K., Zhou, Z., Wang, X., Saeta, B., Diaz, M., Firat, O., Catasta, M., Wei, J., Meier-Hellstern, K., Eck, D., Dean, J., Petrov, S., and Fiedel, N. (2023). Palm: Scaling language modeling with pathways. *Journal of Machine Learning Research*, 24(240):1–113.

Christiano, P. F., Leike, J., Brown, T., Martic, M., Legg, S., and Amodei, D. (2017). Deep reinforcement learning from human preferences. In *Advances in Neural Information Processing Systems*, volume 30.

Conitzer, V., Freedman, R., Heitzig, J., Holliday, W. H., Jacobs, B. M., Lambert, N., Mossé, M., Pacuit, E., Russell, S., Schoelkopf, H., Tewolde, E., and Zwicker, W. S. (2024). Position: social choice should guide ai alignment in dealing with diverse human feedback. In *Forty-first International Conference on Machine Learning*.

Dai, J. and Fleisig, E. (2024). Mapping social choice theory to RLHF. In *ICLR 2024 Workshop on Reliable and Responsible Foundation Models*.

DeepSeek-AI, Guo, D., Yang, D., Zhang, H., Song, J.-M., Zhang, R., Xu, R., Zhu, Q., Ma, S., Wang, P., Bi, X., Zhang, X., Yu, X., Wu, Y., Wu, Z. F., Gou, Z., Shao, Z., Li, Z., Gao, Z., Liu, A., Xue, B., Wang, B.-L., Wu, B., Feng, B., Lu, C., Zhao, C., Deng, C., Zhang, C., Ruan, C., Dai, D., Chen, D., Ji, D.-L., Li, E., Lin, F., Dai, F., Luo, F., Hao, G., Chen, G., Li, G., Zhang, H., Bao, H., Xu, H., Wang, H., Ding, H., Xin, H., Gao, H., Qu, H., Li, H., Guo, J., Li, J., Wang, J., Chen, J., Yuan, J., Qiu, J., Li, J., Cai, J., Ni, J., Liang, J., Chen, J., Dong, K., Hu, K., Gao, K., Guan, K., Huang, K., Yu, K., Wang, L., Zhang, L., Zhao, L., Wang, L., Zhang, L., Xu, L., Xia, L., Zhang, M., Zhang, M., Tang, M., Li, M., Wang, M., Li, M., Tian, N., Huang, P., Zhang, P., Wang, Q., Chen, Q., Du, Q.,

Ge, R., Zhang, R., Pan, R., Wang, R., Chen, R. J., Jin, R., Chen, R., Lu, S., Zhou, S., Chen, S., Ye, S., Wang, S., Yu, S., Zhou, S., Pan, S., Li, S. S., Zhou, S., Wu, S.-K., Yun, T., Pei, T., Sun, T., Wang, T., Zeng, W., Zhao, W., Liu, W., Liang, W., Gao, W., Yu, W.-X., Zhang, W., Xiao, W. L., An, W., Liu, X., Wang, X., Chen, X., Nie, X., Cheng, X., Liu, X., Xie, X., Liu, X., Yang, X., Li, X., Su, X., Lin, X., Li, X. Q., Jin, X., Shen, X.-C., Chen, X., Sun, X., Wang, X., Song, X., Zhou, X., Wang, X., Shan, X., Li, Y. K., Wang, Y. Q., Wei, Y. X., Zhang, Y., Xu, Y., Li, Y., Zhao, Y., Sun, Y., Wang, Y., Yu, Y., Zhang, Y., Shi, Y., Xiong, Y., He, Y., Piao, Y., Wang, Y., Tan, Y., Ma, Y., Liu, Y., Guo, Y., Ou, Y., Wang, Y., Gong, Y., Zou, Y.-J., He, Y., Xiong, Y., Luo, Y.-W., mei You, Y., Liu, Y., Zhou, Y., Zhu, Y. X., Huang, Y., Li, Y., Zheng, Y., Zhu, Y., Ma, Y., Tang, Y., Zha, Y., Yan, Y., Ren, Z., Ren, Z., Sha, Z., Fu, Z., Xu, Z., Xie, Z., guo Zhang, Z., Hao, Z., Ma, Z., Yan, Z., Wu, Z., Gu, Z., Zhu, Z., Liu, Z., Li, Z.-A., Xie, Z., Song, Z., Pan, Z., Huang, Z., Xu, Z., Zhang, Z., and Zhang, Z. (2025). Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning. *arXiv preprint arXiv:2501.12948*.

Dong, H., Xiong, W., Pang, B., Wang, H., Zhao, H., Zhou, Y., Jiang, N., Sahoo, D., Xiong, C., and Zhang, T. (2024). RLHF workflow: From reward modeling to online RLHF. *Transactions on Machine Learning Research*.

Duersch, P., Oechssler, J., and Schipper, B. C. (2012). Pure strategy equilibria in symmetric two-player zero-sum games. *International Journal of Game Theory*, 41:553–564.

Eloundou, T., Manning, S., Mishkin, P., and Rock, D. (2024). GPTs are GPTs: Labor market impact potential of LLMs. *Science*, 384(6702):1306–1308.

Ge, L., Halpern, D., Micha, E., Procaccia, A. D., Shapira, I., Vorobeychik, Y., and Wu, J. (2024). Axioms for ai alignment from human feedback. In *Advances in Neural Information Processing Systems*, volume 37, pages 80439–80465.

Gehrlein, W. V. (2006). *Condorcet's paradox*. Springer.

Hurst, A., Lerer, A., Goucher, A. P., Perelman, A., Ramesh, A., Clark, A., Ostrow, A., Welihinda, A., Hayes, A., Radford, A., Mądry, A., Baker-Whitcomb, A., Beutel, A., Borzunov, A., Carney, A., Chow, A., Kirillov, A., Nichol, A., Paino, A., Renzin, A., Passos, A. T., Kirillov, A., Christakis, A., Conneau, A., Kamali, A., Jabri, A., Moyer, A., Tam, A., Crookes, A., Tootoochian, A., Tootoonchian, A., Kumar, A., Vallone, A., Karpathy, A., Braunstein, A., Cann, A., Codispoti, A., Galu, A., Kondrich, A., Tulloch, A., Mishchenko, A., Baek, A., Jiang, A., Pelisse, A., Woodford, A., Gosalia, A., Dhar, A., Pantuliano, A., Nayak, A., Oliver, A., Zoph, B., Ghorbani, B., Leimberger, B., Rossen, B., Sokolowsky, B., Wang, B., Zweig, B., Hoover, B., Samic, B., McGrew, B., Spero, B., Giertler, B., Cheng, B., Lightcap, B., Walkin, B., Quinn, B., Guarraci, B., Hsu, B., Kellogg, B., Eastman, B., Lugaresi, C., Wainwright, C., Bassin, C., Hudson, C., Chu, C., Nelson, C., Li, C., Shern, C. J., Conger, C., Barette, C., Voss, C., Ding, C., Lu, C., Zhang, C., Beaumont, C., Hallacy, C., Koch, C., Gibson, C., Kim, C., Choi, C., McLeavey, C., Hesse, C., Fischer, C., Winter, C., Czarnecki, C., Jarvis, C., Wei, C., Koumouzelis, C., Sherburn, D., Kappler, D., Levin, D., Levy, D., Carr, D., Farhi, D., Mely, D., Robinson, D., Sasaki, D., Jin, D., Valladares, D., Tsipras, D., Li, D., Nguyen, D. P., Findlay, D., Oiwoh, E., Wong, E., Asdar, E., Proehl, E., Yang, E., Antonow, E., Kramer, E., Peterson, E., Sigler, E., Wallace, E., Brevdo, E., Mays, E., Khorasani, F., Such, F. P., Raso, F., Zhang, F., von Lohmann, F., Sulit, F., Goh, G., Oden, G., Salmon, G., Starace, G., Brockman, G., Salman, H., Bao, H., Hu, H., Wong, H., Wang, H., Schmidt, H., Whitney, H., Jun, H., Kirchner, H., de Oliveira Pinto, H. P., Ren, H., Chang, H., Chung, H. W., Kivlichan, I., O'Connell, I., O'Connell, I., Osband, I., Silber, I., Sohl, I., Okuyucu, I., Lan, I., Kostrikov, I., Sutskever, I., Kanitscheider, I., Gulrajani, I., Coxon, J., Menick, J., Pachocki, J., Aung, J., Betker, J., Crooks, J., Lennon, J., Kiros, J., Leike, J., Park, J., Kwon, J., Phang, J., Teplitz, J., Wei, J., Wolfe, J., Chen, J., Harris, J., Varavva, J., Lee, J. G., Shieh, J., Lin, J., Yu, J., Weng, J., Tang, J., Yu, J., Jang, J., Candela, J. Q., Beutler, J., Landers, J., Parish, J., Heidecke, J., Schulman, J., Lachman, J., McKay, J., Uesato, J., Ward, J., Kim, J. W., Huizinga, J., Sitkin, J., Kraaijeveld, J., Gross, J., Kaplan, J., Snyder, J., Achiam, J., Jiao, J., Lee, J., Zhuang, J., Harriman, J., Fricke, K., Hayashi, K., Singhal, K., Shi, K., Karthik, K., Wood, K., Rimbach, K., Hsu, K., Nguyen, K., Gu-Lemberg, K., Button, K., Liu, K., Howe, K., Muthukumar, K., Luther, K., Ahmad, L., Kai, L., Itow, L., Workman, L., Pathak, L., Chen, L., Jing, L., Guy, L., Fedus, L., Zhou, L., Mamitsuka, L., Weng, L., McCallum, L., Held, L., Ouyang, L., Feuvrier, L., Zhang, L., Kondraciuk, L., Kaiser, L., Hewitt, L., Metz, L., Doshi, L., Aflak, M., Simens, M., Boyd, M., Thompson, M., Dukhan, M., Chen,

M., Gray, M., Hudnall, M., Zhang, M., Aljubeh, M., Litwin, M., Zeng, M., Johnson, M., Shetty, M., Gupta, M., Shah, M., Yatbaz, M., Yang, M. J., Zhong, M., Glaese, M., Chen, M., Janner, M., Lampe, M., Petrov, M., Wu, M., Wang, M., Fradin, M., Pokrass, M., Castro, M., de Castro, M. O. T., Pavlov, M., Brundage, M., Wang, M., Khan, M., Murati, M., Bavarian, M., Lin, M., Yesildal, M., Soto, N., Gimelshein, N., Cone, N., Staudacher, N., Summers, N., LaFontaine, N., Chowdhury, N., Ryder, N., Stathas, N., Turley, N., Tezak, N., Felix, N., Kudige, N., Keskar, N., Deutsch, N., Bundick, N., Puckett, N., Nachum, O., Okelola, O., Boiko, O., Murk, O., Jaffe, O., Watkins, O., Godement, O., Campbell-Moore, O., Chao, P., McMillan, P., Belov, P., Su, P., Bak, P., Bakkum, P., Deng, P., Dolan, P., Hoeschele, P., Welinder, P., Tillet, P., Pronin, P., Tillet, P., Dhariwal, P., Yuan, Q., Dias, R., Lim, R., Arora, R., Troll, R., Lin, R., Lopes, R. G., Puri, R., Miyara, R., Leike, R., Gaubert, R., Zamani, R., Wang, R., Donnelly, R., Honsby, R., Smith, R., Sahai, R., Ramchandani, R., Huet, R., Carmichael, R., Zellers, R., Chen, R., Chen, R., Nigmatullin, R., Cheu, R., Jain, S., Altman, S., Schoenholz, S., Toizer, S., Miserendino, S., Agarwal, S., Culver, S., Ethersmith, S., Gray, S., Grove, S., Metzger, S., Hermani, S., Jain, S., Zhao, S., Wu, S., Jomoto, S., Wu, S., Shuaiqi, Xia, Phene, S., Papay, S., Narayanan, S., Coffey, S., Lee, S., Hall, S., Balaji, S., Broda, T., Stramer, T., Xu, T., Gogineni, T., Christianson, T., Sanders, T., Patwardhan, T., Cunninghman, T., Degry, T., Dimson, T., Raoux, T., Shadwell, T., Zheng, T., Underwood, T., Markov, T., Sherbakov, T., Rubin, T., Stasi, T., Kaftan, T., Heywood, T., Peterson, T., Walters, T., Eloundou, T., Qi, V., Moeller, V., Monaco, V., Kuo, V., Fomenko, V., Chang, W., Zheng, W., Zhou, W., Manassra, W., Sheu, W., Zaremba, W., Patil, Y., Qian, Y., Kim, Y., Cheng, Y., Zhang, Y., He, Y., Zhang, Y., Jin, Y., Dai, Y., and Malkov, Y. (2024). Gpt-4o system card. *arXiv preprint arXiv:2410.21276*.

Ji, W., Yuan, W., Getzen, E., Cho, K., Jordan, M. I., Mei, S., Weston, J. E., Su, W. J., Xu, J., and Zhang, L. (2025). An overview of large language models for statisticians. *arXiv preprint arXiv:2502.17814*.

Jiang, D., Ren, X., and Lin, B. Y. (2023). LLM-blender: Ensembling large language models with pairwise ranking and generative fusion. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 14165–14178. Association for Computational Linguistics.

Liu, K., Long, Q., Shi, Z., Su, W. J., and Xiao, J. (2025). Statistical impossibility and possibility of aligning llms with human preferences: From condorcet paradox to nash equilibrium. *arXiv preprint arXiv:2503.10990*.

Luce, R. D. (2012). *Individual choice behavior: A theoretical analysis*. Courier Corporation.

Maura-Rivero, R.-R., Lanctot, M., Visin, F., and Larson, K. (2025). Jackpot! alignment as a maximal lottery. *arXiv preprint arXiv:2501.19266*.

Mishra, A. (2023). Ai alignment and social choice: Fundamental limitations and policy implications. *arXiv preprint arXiv:2310.16048*.

Munos, R., Valko, M., Calandriello, D., Gheshlaghi Azar, M., Rowland, M., Guo, Z. D., Tang, Y., Geist, M., Mesnard, T., Fiegel, C., Michi, A., Selvi, M., Girgin, S., Momchev, N., Bachem, O., Mankowitz, D. J., Precup, D., and Piot, B. (2024). Nash learning from human feedback. In *Forty-first International Conference on Machine Learning*.

Myerson, R. B. (2013). *Game theory*. Harvard university press.

Noothigattu, R., Peters, D., and Procaccia, A. D. (2020). Axioms for learning from pairwise comparisons. In *Advances in Neural Information Processing Systems*, volume 33, pages 17745–17754.

Novikov, A., Vũ, N., Eisenberger, M., Dupont, E., Huang, P.-S., Wagner, A. Z., Shirobokov, S., Kozlovskii, B., Ruiz, F. J. R., Mehrabian, A., Kumar, M. P., See, A., Chaudhuri, S., Holland, G., Davies, A., Nowozin, S., Kohli, P., and Balog, M. (2025). AlphaEvolve: A coding agent for scientific and algorithmic discovery. Technical report, Google DeepMind.

OpenAI (2025). Openai o3 and o4-mini system card. Technical report, OpenAI.

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C. L., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., Schulman, J., Hilton, J., Kelton, F., Miller, L., Simens, M., Askell, A., Welinder, P., Christiano, P., Leike, J., and Lowe, R. (2022). Training language models to follow instructions with human feedback. In *Advances in Neural Information Processing Systems*, volume 35, pages 27730–27744.

Shoham, Y. and Leyton-Brown, K. (2008). *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge University Press.

Siththaranjan, A., Laidlaw, C., and Hadfield-Menell, D. (2024). Distributional preference learning: Understanding and accounting for hidden context in RLHF. In *The Twelfth International Conference on Learning Representations*.

Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E., and Lample, G. (2023). Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*.

Wu, Y., Sun, Z., Yuan, H., Ji, K., Yang, Y., and Gu, Q. (2024). Self-play preference optimization for language model alignment. In *Adaptive Foundation Models: Evolving AI for Personalized and Efficient Learning*.

Xiao, J., Li, Z., Xie, X., Getzen, E., Fang, C., Long, Q., and Su, W. J. (2024). On the algorithmic bias of aligning large language models with rlhf: Preference collapse and matching regularization. *arXiv preprint arXiv:2405.16455*.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The claims in the abstract and introduction accurately reflect the contributions and scope of this paper.

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: The paper discusses the limitations of the work in Section 6 (see the "Limitations and Discussion" paragraph).

   Guidelines:

   - The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [Yes]

Justification: The full set of assumptions is introduced before theoretical result and the complete proof of theoretical result is provided in the appendix.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [NA]

Justification: This paper focuses on theoretical analysis and does not include experiments.

Guidelines:

- The answer NA means that the paper does not include experiments.
- If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
  (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
  (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
  (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
  (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [NA]

Justification: This paper does not include experiments requiring code.

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so "No" is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

Answer: [NA]

Justification: This paper does not include experiments.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

Answer: [NA]

Justification: The paper does not include experiments.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)

- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [NA]

   Justification: The paper does not include experiments.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: Our research conforms with the NeurIPS Code of Ethics.

   Guidelines:

   - The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

    Justification: The paper discusses broader impacts in Section 6 (see the "Broader Impacts" paragraph).

    Guidelines:

    - The answer NA means that there is no societal impact of the work performed.
    - If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.

- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

    Answer: [NA]

    Justification: This paper poses no risk for misuse.

    Guidelines:

    - The answer NA means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [NA]

    Justification: The paper does not use existing assets.

    Guidelines:

    - The answer NA means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.

- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [NA]

    Justification: The paper does not release new assets.

    Guidelines:

    - The answer NA means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [NA]

    Justification: The paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [NA]

    Justification: The paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does not impact the core methodology, scientific rigorousness, or originality of the research, declaration is not required.

Answer: [NA]

Justification: The core method development in this research does not involve LLMs as any important, original, or non-standard components.

Guidelines:

- The answer NA means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy (`https://neurips.cc/Conferences/2025/LLM`) for what should or should not be described.

A Proof of Theorem 3.1

Notation. For simplicity, we denote $\Psi(\mathcal{P}(y_i \succ y_j))$ as Ψ_{ij} for any $1 \leqslant i, j \leqslant n$, and define the payoff matrix as $\boldsymbol{\Psi} := \Psi_{ij}{}_{1 \leqslant i,j \leqslant n}$. We then define the total payoff by:

$$\mathcal{P}_\Psi(\boldsymbol{\pi}_1, \boldsymbol{\pi}_2) := \sum_{i=1}^{n} \sum_{j=1}^{n} \pi_{1,i} \pi_{2,j} \Psi_{ij} \,.$$

We denote by $\boldsymbol{\delta}_i$ the policy supported solely on y_i, i.e., $\mathrm{supp}(\boldsymbol{\delta}_i) = \{y_i\}$. The mixed policy $(\boldsymbol{\delta}_{i_1} + \ldots + \boldsymbol{\delta}_{i_k})/k$ is then defined as the policy $\boldsymbol{\pi}$ such that

$$\pi_i = \begin{cases} 1/k \,, & i \in \{i_1, \cdots, i_k\} \\ 0 \,, & \text{otherwise} \end{cases} \,,$$

for any subset $\{i_1, \ldots, i_k\} \subseteq [n]$.

Proof of Theorem 3.1. Without any loss of generality, we assume that y_1 is the Condorcet winning response. First, we show that a necessary condition that ensures the Condorcet consistency of problem (1.2) is:

$$\begin{cases} \Psi(t) \geqslant \Psi(1/2) \,, 1 \geqslant t \geqslant 1/2 \\ \Psi(t) < \Psi(1/2) \,, 0 \leqslant t < 1/2 \end{cases} \,. \tag{A.1}$$

To show this, we examine the case where $n = 2$. For any $1 \geqslant t > 1/2$, we consider the game with the payoff in Table 2. By the definition of Condorcet consistency, all Nash equilibrium of this game is of the form $(\boldsymbol{\delta}_1, \boldsymbol{\pi}^\star)$ for some $\boldsymbol{\pi}^\star$.

Table 2: Payoff matrix with two responses $\{y_1, y_2\}$.

$\Psi(\mathcal{P}(y \succ y'))$	$y' = y_1$	$y' = y_2$
$y = y_1$	$\Psi(1/2)$	$\Psi(t)$
$y = y_2$	$\Psi(1-t)$	$\Psi(1/2)$

Case 1. If $\Psi(t) > \Psi(1/2)$, we have

$$\boldsymbol{\pi}^\star = \arg\min_{\boldsymbol{\pi}} \mathcal{P}_\Psi(\boldsymbol{\delta}_1, \boldsymbol{\pi}) = \arg\min_{\boldsymbol{\pi}} \{\pi_1 \Psi(1/2) + \pi_2 \Psi(t)\} = \boldsymbol{\delta}_1 \,.$$

Therefore, we have

$$\begin{aligned} \Psi(1/2) = \mathcal{P}_\Psi(\boldsymbol{\delta}_1, \boldsymbol{\delta}_1) = \max_{\boldsymbol{\pi}} \mathcal{P}_\Psi(\boldsymbol{\pi}, \boldsymbol{\delta}_1) \geqslant \mathcal{P}_\Psi(\boldsymbol{\delta}_2, \boldsymbol{\delta}_1) = \Psi_{21} = \Psi(1-t) \,, \\ \Psi(1/2) = \mathcal{P}_\Psi(\boldsymbol{\delta}_1, \boldsymbol{\delta}_1) = \min_{\boldsymbol{\pi}} \mathcal{P}_\Psi(\boldsymbol{\delta}_1, \boldsymbol{\pi}) \leqslant \mathcal{P}_\Psi(\boldsymbol{\delta}_1, \boldsymbol{\delta}_2) = \Psi_{12} = \Psi(t) \,. \end{aligned}$$

Hence, we have $\Psi(1-t) \leqslant \Psi(1/2) < \Psi(t)$. If $\Psi(1/2) = \Psi(1-t)$, notice that

$$\begin{aligned} \mathcal{P}_\Psi(\boldsymbol{\pi}, \boldsymbol{\delta}_1) = \pi_1 \Psi(1/2) + \pi_2 \Psi(1-t) = \Psi(1/2) \Longrightarrow \boldsymbol{\delta}_2 \in \arg\max_{\boldsymbol{\pi}} \mathcal{P}_\Psi(\boldsymbol{\pi}, \boldsymbol{\delta}_1) \,, \\ \mathcal{P}_\Psi(\boldsymbol{\delta}_2, \boldsymbol{\pi}) = \pi_1 \Psi(1-t) + \pi_2 \Psi(1/2) = \Psi(1/2) \Longrightarrow \boldsymbol{\delta}_1 \in \arg\min_{\boldsymbol{\pi}} \mathcal{P}_\Psi(\boldsymbol{\delta}_2, \boldsymbol{\pi}) \,. \end{aligned}$$

Therefore, $(\boldsymbol{\delta}_2, \boldsymbol{\delta}_1)$ is also a Nash equilibrium, which causes a contradiction to the fact that problem (1.2) is Condorcet consistent. Therefore, we have $\Psi(t) > \Psi(1/2) > \Psi(1-t)$ for any $1 \geqslant t > 1/2$.

Case 2. If $\Psi(t) < \Psi(1/2)$, we have

$$\boldsymbol{\pi}^\star = \arg\min_{\boldsymbol{\pi}} \mathcal{P}_\Psi(\boldsymbol{\delta}_1, \boldsymbol{\pi}) = \arg\min_{\boldsymbol{\pi}} \{\pi_1 \Psi(1/2) + \pi_2 \Psi(t)\} = \boldsymbol{\delta}_2 \,.$$

However, notice that

$$\Psi(1/2) = \mathcal{P}_\Psi(\boldsymbol{\delta}_2, \boldsymbol{\delta}_2) \leqslant \max_{\boldsymbol{\pi}} \mathcal{P}_\Psi(\boldsymbol{\pi}, \boldsymbol{\delta}_2) = \mathcal{P}_\Psi(\boldsymbol{\delta}_1, \boldsymbol{\delta}_2) = \Psi(t) < \Psi(1/2) \,,$$

which causes a contradiction.

Case 3. If $\Psi(t)=\Psi(1/2)$. When $\Psi(1-t)=\Psi(1/2)$, any $(\boldsymbol{\pi}_1,\boldsymbol{\pi}_2)$ is a Nash equilibrium, which causes a contradiction to the fact that problem (1.2) is Condorcet consistent. When $\Psi(1-t)>\Psi(1/2)$, note that

$$\mathcal{P}_\Psi(\boldsymbol{\delta}_2,\boldsymbol{\pi})=\pi_1\Psi(1-t)+\pi_2\Psi(1/2)\geqslant\Psi(1/2)\Longrightarrow\boldsymbol{\delta}_2\in\arg\min_{\boldsymbol{\pi}}\mathcal{P}_\Psi(\boldsymbol{\delta}_2,\boldsymbol{\pi})\,,$$
$$\mathcal{P}_\Psi(\boldsymbol{\pi},\boldsymbol{\delta}_2)=\pi_1\Psi(t)+\pi_2\Psi(1/2)=\Psi(1/2)\Longrightarrow\boldsymbol{\delta}_2\in\arg\max_{\boldsymbol{\pi}}\mathcal{P}_\Psi(\boldsymbol{\pi},\boldsymbol{\delta}_2)\,.$$

Therefore, $(\boldsymbol{\delta}_2,\boldsymbol{\delta}_2)$ is a Nash equilibrium, which also causes a contradiction to the fact problem (1.2) is Condorcet consistent. Hence, we have $\Psi(1-t)<\Psi(1/2)$.

In summary, for any $1\geqslant t>1/2$, we have $\Psi(1-t)<\Psi(1/2)\leqslant\Psi(t)$. Hence, (A.1) holds if problem (1.2) is Condorcet consistent. Next, we prove that (A.1) is also sufficient for the Condorcet consistency of problem (1.2). Recall that $\Psi_{i1}=\Psi(\mathcal{P}(y_i\succ y_1))<\Psi(1/2)$, and $\Psi_{1i}=\Psi(\mathcal{P}(y_1\succ y_i))\geqslant\Psi(1/2)$ for any $i\neq 1$. If $(\boldsymbol{\pi}_1^\star,\boldsymbol{\pi}_2^\star)$ is a Nash equilibrium. Notice that

$$\mathcal{P}_\Psi(\boldsymbol{\pi}_1^\star,\boldsymbol{\pi}_2^\star)=\max_{\boldsymbol{\pi}}\mathcal{P}_\Psi(\boldsymbol{\pi},\boldsymbol{\pi}_2^\star)\geqslant\mathcal{P}_\Psi(\boldsymbol{\delta}_1,\boldsymbol{\pi}_2^\star)=\sum_{i=1}^n\pi_{2,i}^\star\Psi_{1i}\,,$$
$$\mathcal{P}_\Psi(\boldsymbol{\pi}_1^\star,\boldsymbol{\pi}_2^\star)=\min_{\boldsymbol{\pi}}\mathcal{P}_\Psi(\boldsymbol{\pi}_1^\star,\boldsymbol{\pi})\leqslant\mathcal{P}_\Psi(\boldsymbol{\pi}_1^\star,\boldsymbol{\delta}_1)=\sum_{i=1}^n\pi_{1,i}^\star\Psi_{i1}\,.$$

Therefore, if $\boldsymbol{\pi}_1^\star\neq\boldsymbol{\delta}_1$, we have

$$\Psi(1/2)\leqslant\sum_{i=1}^n\pi_{2,i}^\star\Psi_{1i}\leqslant\mathcal{P}_\Psi(\boldsymbol{\pi}_1^\star,\boldsymbol{\pi}_2^\star)\leqslant\sum_{i=1}^n\pi_{1,i}^\star\Psi_{i1}<\Psi(1/2)\,,\tag{A.2}$$

which causes a contradiction. Therefore, $\boldsymbol{\pi}_1^\star=\boldsymbol{\delta}_1$, i.e., problem (1.2) is Condorcet consistent. Hence, we conclude our proof. □

B Proof of Example 3.4

Proof of Example 3.4. We prove this conclusion by contradiction. Suppose that the Nash solution is $\boldsymbol{\delta}_{i^\star}$ for some $i^\star\in[n]$, and the Nash equilibrium is $(\boldsymbol{\delta}_{i^\star},\boldsymbol{\pi}^\star)$. As there is no Condorcet winning response, by definition, there exists j' such that $\mathcal{P}(y_{i^\star}\succ y_{j'})<1/2$. Then we have

$$\mathcal{P}_\Psi(\boldsymbol{\delta}_{i^\star},\boldsymbol{\pi}^\star)=\min_{\boldsymbol{\pi}}\mathcal{P}_\Psi(\boldsymbol{\delta}_{i^\star},\boldsymbol{\pi})\leqslant\mathcal{P}_\Psi(\boldsymbol{\delta}_{i^\star},\boldsymbol{\delta}_{j'})=\Psi\left(\mathcal{P}\left(y_{i^\star}\succ y_{j'}\right)\right)=M_-\,.\tag{B.1}$$

However, choosing i' such that $\pi_{i'}^\star>0$, we have

$$\mathcal{P}_\Psi(\boldsymbol{\delta}_{i^\star},\boldsymbol{\pi}^\star)=\max_{\boldsymbol{\pi}}\mathcal{P}_\Psi(\boldsymbol{\pi},\boldsymbol{\pi}^\star)\geqslant\mathcal{P}_\Psi(\boldsymbol{\delta}_{i'},\boldsymbol{\pi}^\star)=\sum_{i=1}^n\pi_i^\star\Psi\left(\mathcal{P}\left(y_{i'}\succ y_i\right)\right)>M_-\,,$$

which causes a contradiction to (B.1). Hence, we conclude our proof. □

C Proof of Theorem 3.2

Proof of Theorem 3.2. First, according to Theorem 3.1, when the Nash solution is Condorcet consistent, we have

$$\begin{cases}\Psi(t)\geqslant\Psi(1/2)\,,1\geqslant t\geqslant 1/2\\ \Psi(t)<\Psi(1/2)\,,1/2>t\geqslant 0\end{cases}\,.\tag{C.1}$$

In addition, we show that $\Psi(\cdot)$ must satisfy $\Psi(t)+\Psi(1-t)\geqslant 2\Psi(1/2)\,,\forall t\in[0,1]$ for ensuring that the Nash solution is mixed when there is no Condorcet winning response. We consider the case where $n=4$ and the game with the payoff in Table 3 for any $t_1,t_2>1/2$. Notice that if $\Psi(t_1)+\Psi(1-t_1)+\Psi(1/2)\leqslant 3\Psi(1-t_2)$, we have

$$\mathcal{P}_\Psi(\boldsymbol{\delta}_4,\boldsymbol{\pi})=(\pi_1+\pi_2+\pi_3)\Psi(1-t_2)+\pi_4\Psi(1/2)\Longrightarrow\frac{\boldsymbol{\delta}_1+\boldsymbol{\delta}_2+\boldsymbol{\delta}_3}{3}\in\arg\min_{\boldsymbol{\pi}}\mathcal{P}_\Psi(\boldsymbol{\delta}_4,\boldsymbol{\pi})\,,$$

and

$$\mathcal{P}_\Psi\left(\boldsymbol{\pi}, \frac{\boldsymbol{\delta}_1+\boldsymbol{\delta}_2+\boldsymbol{\delta}_3}{3}\right) = (\pi_1+\pi_2+\pi_3)\cdot\frac{\Psi(1/2)+\Psi(t_1)+\Psi(1-t_1)}{3}+\pi_4\Psi(1-t_2)$$
$$\Longrightarrow \boldsymbol{\delta}_4 \in \arg\max_{\boldsymbol{\pi}} \mathcal{P}_\Psi\left(\boldsymbol{\pi}, \frac{\boldsymbol{\delta}_1+\boldsymbol{\delta}_2+\boldsymbol{\delta}_3}{3}\right).$$

Therefore, $(\boldsymbol{\delta}_4, (\boldsymbol{\delta}_1+\boldsymbol{\delta}_2+\boldsymbol{\delta}_3)/3)$ is a Nash equilibrium, which causes a contradiction to the fact that the Nash solution is mixed. Hence, we have $\Psi(t_1)+\Psi(1-t_1)+\Psi(1/2) > 3\Psi(1-t_2)$ for any $t_1, t_2 > 1/2$. Let $t_2 \to 1/2$, we have $\Psi(t)+\Psi(1-t) \geqslant 2\Psi(1/2)$ for any $t \in [0,1]$. Hence, combining (C.1), we have shown that the necessary condition for ensuring that the Nash solution is mixed is:

$$\Psi(t)+\Psi(1-t) \geqslant 2\Psi(1/2), \forall t \in [0,1] \text{ and } \begin{cases} \Psi(t) \geqslant \Psi(1/2), 1 \geqslant t \geqslant 1/2 \\ \Psi(t) < \Psi(1/2), 1/2 > t \geqslant 0 \end{cases}. \tag{C.2}$$

Next, we prove that the condition (C.2) is also sufficient. Suppose that $(\boldsymbol{\delta}_{i^\star}, \boldsymbol{\pi}^\star)$ is a Nash equilibrium, then we have

$$\mathcal{P}_\Psi(\boldsymbol{\delta}_{i^\star}, \boldsymbol{\pi}^\star) = \max_{\boldsymbol{\pi}} \mathcal{P}_\Psi(\boldsymbol{\pi}, \boldsymbol{\pi}^\star) \geqslant \mathcal{P}_\Psi(\boldsymbol{\pi}^\star, \boldsymbol{\pi}^\star)$$
$$= \sum_{i=1}^n\sum_{j=1}^n \pi_i^\star\pi_j^\star\Psi_{ij} = \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n \pi_i^\star\pi_j^\star\left(\Psi_{ij}+\Psi_{ji}\right) \geqslant \Psi(1/2).$$

However, notice that for any j, we have

$$\mathcal{P}_\Psi(\boldsymbol{\delta}_{i^\star}, \boldsymbol{\pi}^\star) = \min_{\boldsymbol{\pi}} \mathcal{P}_\Psi(\boldsymbol{\delta}_{i^\star}, \boldsymbol{\pi}) \leqslant \mathcal{P}_\Psi(\boldsymbol{\delta}_{i^\star}, \boldsymbol{\delta}_j) = \Psi_{i^\star j}.$$

As there is no Condorcet winning response, there must exist $j^\star$ such that $\mathcal{P}(y_{i^\star} \succ y_{j^\star}) < 1/2$, thus $\Psi_{i^\star j^\star} < \Psi(1/2)$. Hence, $\Psi(1/2) \leqslant \mathcal{P}_\Psi(\boldsymbol{\delta}_{i^\star}, \boldsymbol{\pi}^\star) \leqslant \Psi_{i^\star j^\star} < \Psi(1/2)$, which causes a contradiction. Therefore, the Nash solution must be mixed. □

D Proof of Example 4.3

Proof of Example 4.3. We prove this conclusion by contradiction. Suppose that the Nash solution is $\boldsymbol{\pi}_1^\star$ that satisfies $\operatorname{supp}(\boldsymbol{\pi}_1^\star) \cap S_1^c \neq \emptyset$, and the Nash equilibrium is $(\boldsymbol{\pi}_1^\star, \boldsymbol{\pi}_2^\star)$.

Case 1. If $\operatorname{supp}(\boldsymbol{\pi}_1^\star) \cap S_1 = \emptyset$, taking $j' \in S_1$, we have

$$\mathcal{P}_\Psi(\boldsymbol{\pi}_1^\star, \boldsymbol{\pi}_2^\star) = \min_{\boldsymbol{\pi}} \mathcal{P}_\Psi(\boldsymbol{\pi}_1^\star, \boldsymbol{\pi}) \leqslant \mathcal{P}_\Psi(\boldsymbol{\pi}_1^\star, \boldsymbol{\delta}_{j'}) = \sum_{i\in S_1^c} \pi_{1,i}^\star\Psi_{ij'} = M_-.$$

However, we have

$$\mathcal{P}_\Psi(\boldsymbol{\pi}_1^\star, \boldsymbol{\pi}_2^\star) = \max_{\boldsymbol{\pi}} \mathcal{P}_\Psi(\boldsymbol{\pi}, \boldsymbol{\pi}_2^\star) \geqslant \mathcal{P}_\Psi(\operatorname{Unif}(S_1), \boldsymbol{\pi}_2^\star) = \sum_{i\in S_1}\sum_{j=1}^n \frac{\pi_{2,j}^\star}{|S_1|}\Psi_{ij} > M_-,$$

which causes a contradiction.

Case 2. If $\operatorname{supp}(\boldsymbol{\pi}_2^\star) \cap S_1 = \emptyset$ and $\operatorname{supp}(\boldsymbol{\pi}_1^\star) \cap S_1 \neq \emptyset$, taking $i' \in \operatorname{supp}(\boldsymbol{\pi}_1^\star) \cap S_1$, we have

$$\mathcal{P}_\Psi(\boldsymbol{\pi}_1^\star, \boldsymbol{\pi}_2^\star) = \max_{\boldsymbol{\pi}} \mathcal{P}_\Psi(\boldsymbol{\pi}, \boldsymbol{\pi}_2^\star) \geqslant \mathcal{P}_\Psi(\boldsymbol{\delta}_{i'}, \boldsymbol{\pi}_2^\star) = \sum_{j\in S_1^c} \pi_{2,j}^\star\Psi_{i'j} = M_+.$$

However, we have

$$\mathcal{P}_\Psi(\boldsymbol{\pi}_1^\star, \boldsymbol{\pi}_2^\star) = \min_{\boldsymbol{\pi}} \mathcal{P}_\Psi(\boldsymbol{\pi}_1^\star, \boldsymbol{\pi}) \leqslant \mathcal{P}_\Psi(\boldsymbol{\pi}_1^\star, \operatorname{Unif}(S_1)) = \sum_{i=1}^n\sum_{j\in S_1} \frac{\pi_{1,i}^\star}{|S_1|}\Psi_{ij} < M_+,$$

which cause a contradiction.

Case 3. If If $\mathrm{supp}(\boldsymbol{\pi}_2^\star) \cap S_1 \neq \emptyset$ and $\mathrm{supp}(\boldsymbol{\pi}_1^\star) \cap S_1 \neq \emptyset$, taking $i_2^\star \in \mathrm{supp}(\boldsymbol{\pi}_2^\star) \cap S_1$, we consider the following strategy $\boldsymbol{\pi}_1'$:

$$\begin{cases} \pi'_{1,i} = 0\,, & i \in S_1^c \\ \pi'_{1,i} = \pi^\star_{1,i}\,, & i \in S_1 \backslash \{i_2^\star\} \\ \pi'_{1,i_2^\star} = \pi^\star_{1,i_2^\star} + \sum_{i \in S_1^c} \pi^\star_{1,i}\,, & i = i_2^\star \end{cases}.$$

Then we have

$$\begin{aligned} \mathcal{P}_\Psi(\boldsymbol{\pi}_1', \boldsymbol{\pi}_2^\star) - \mathcal{P}_\Psi(\boldsymbol{\pi}_1^\star, \boldsymbol{\pi}_2^\star) &= \sum_{i=1}^n \sum_{j=1}^n \left(\pi'_{1,i} - \pi^\star_{1,i}\right) \pi^\star_{2,j} \Psi_{ij} \\ &= -\sum_{i \in S_1^c} \sum_{j=1}^n \pi^\star_{1,i} \pi^\star_{2,j} \Psi_{ij} + \sum_{j=1}^n \sum_{i \in S_1^c} \pi^\star_{1,i} \pi^\star_{2,j} \Psi_{i_2^\star j} \\ &= \sum_{j=1}^n \pi^\star_{2,j} \left[\sum_{i \in S_1^c} \pi^\star_{1,i} \left(\Psi_{i_2^\star j} - \Psi_{ij} \right) \right] > 0\,. \end{aligned} \tag{D.1}$$

where the last inequality follows from the following two facts: for any $i \in S_1^c$,

$$\Psi_{i_2^\star j} - \Psi_{ij} = \begin{cases} M_+ - \Psi_{ij} \geqslant 0\,, & j \in S_1^c \\ \Psi_{i_2^\star j} - M_- \geqslant 0\,, & j \in S_1 \end{cases},$$

and when $j = i_2^\star$,

$$\pi^\star_{2,i_2^\star} \left[\sum_{i \in S_1^c} \pi^\star_{1,i} \left(\Psi_{i_2^\star i_2^\star} - \Psi_{i i_2^\star} \right) \right] = \pi^\star_{2,i_2^\star} \left(\Psi(1/2) - M_- \right) \sum_{i \in S_1^c} \pi^\star_{1,i} > 0\,.$$

However, (D.1) causes a contradiction to the fact that $\mathcal{P}_\Psi(\boldsymbol{\pi}_1', \boldsymbol{\pi}_2^\star) \leqslant \max_{\boldsymbol{\pi}} \mathcal{P}_\Psi(\boldsymbol{\pi}, \boldsymbol{\pi}_2^\star) = \mathcal{P}_\Psi(\boldsymbol{\pi}_1^\star, \boldsymbol{\pi}_2^\star)$.

Hence, in summary, it must hold that $\mathrm{supp}(\boldsymbol{\pi}_1^\star) \cap S_1^c = \emptyset$, i.e., $\mathrm{supp}(\boldsymbol{\pi}_1^\star) \subseteq S_1$. □

E Proof of Theorem 4.2

Proof of Theorem 4.2. First, we show that the necessary condition for ensuring that problem (1.2) is Smith consistent is:

$$\Psi(t) + \Psi(1-t) = 2\Psi(1/2)\,, \forall t \in [0,1] \text{ and } \begin{cases} \Psi(t) \geqslant \Psi(1/2)\,, 1 \geqslant t \geqslant 1/2 \\ \Psi(t) < \Psi(1/2)\,, 0 \leqslant t < 1/2 \end{cases}. \tag{E.1}$$

First, Condorcet consistency must hold when Smith consistency holds. According to Theorem 3.1, we have

$$\begin{cases} \Psi(t) \geqslant \Psi(1/2)\,, 1 \geqslant t \geqslant 1/2 \\ \Psi(t) < \Psi(1/2)\,, 0 \leqslant t < 1/2 \end{cases}$$

Next, we show that when $\Psi(\cdot)$ is continuous at $1/2$, $\Psi(\cdot)$ must satisfy $\Psi(t) + \Psi(1-t) \geqslant 2\Psi(1/2)$ (Lemma E.1) and $\Psi(t) + \Psi(1-t) \leqslant 2\Psi(1/2)$ (Lemma E.2) for any $t \in [0,1]$. Therefore, combining the two results together, we obtain the condition (E.1).

Lemma E.1. *When $\Psi(\cdot)$ is continuous at $1/2$. Achieving Smith consistency only if*

$$\Psi(t) + \Psi(1-t) \geqslant 2\Psi(1/2)\,, \forall t \in [0,1]\,.$$

Proof of Lemma E.1. We consider the case where $n = 4$ and the game with the payoff in Table 3 for any $t_1, t_2 > 1/2$. Notice that if $\Psi(t_1) + \Psi(1-t_1) + \Psi(1/2) \leqslant 3\Psi(1-t_2)$, we have

$$\mathcal{P}_\Psi(\boldsymbol{\delta}_4, \boldsymbol{\pi}) = (\pi_1 + \pi_2 + \pi_3)\Psi(1-t_2) + \pi_4 \Psi(1/2) \Longrightarrow \frac{\boldsymbol{\delta}_1 + \boldsymbol{\delta}_2 + \boldsymbol{\delta}_3}{3} \in \arg\min_{\boldsymbol{\pi}} \mathcal{P}_\Psi(\boldsymbol{\delta}_4, \boldsymbol{\pi})\,,$$

and

$$\mathcal{P}_\Psi\left(\boldsymbol{\pi}, \frac{\boldsymbol{\delta}_1+\boldsymbol{\delta}_2+\boldsymbol{\delta}_3}{3}\right) = (\pi_1+\pi_2+\pi_3)\cdot\frac{\Psi(1/2)+\Psi(t_1)+\Psi(1-t_1)}{3}+\pi_4\Psi(1-t_2)$$
$$\Longrightarrow \boldsymbol{\delta}_4 \in \arg\max_{\boldsymbol{\pi}} \mathcal{P}_\Psi\left(\boldsymbol{\pi}, \frac{\boldsymbol{\delta}_1+\boldsymbol{\delta}_2+\boldsymbol{\delta}_3}{3}\right).$$

Therefore, $(\boldsymbol{\delta}_4, (\boldsymbol{\delta}_1+\boldsymbol{\delta}_2+\boldsymbol{\delta}_3)/3)$ is a Nash equilibrium, which causes a contradiction to the fact that the Nash solution supports on $S_1 := \{y_1, y_2, y_3\}$. Hence, we have $\Psi(t_1)+\Psi(1-t_1)+\Psi(1/2) > 3\Psi(1-t_2)$ for any $t_1, t_2 > 1/2$. Let $t_2 \to 1/2$, we have $\Psi(t)+\Psi(1-t) \geqslant 2\Psi(1/2)$ for any $t \in [0,1]$. □

Table 3: Payoff matrix with four responses $\{y_1, y_2, y_3, y_4\}$.

$\Psi(\mathcal{P}(y \succ y'))$	$y'=y_1$	$y'=y_2$	$y'=y_3$	$y'=y_4$
$y=y_1$	$\Psi(1/2)$	$\Psi(t_1)$	$\Psi(1-t_1)$	$\Psi(t_2)$
$y=y_2$	$\Psi(1-t_1)$	$\Psi(1/2)$	$\Psi(t_1)$	$\Psi(t_2)$
$y=y_3$	$\Psi(t_1)$	$\Psi(1-t_1)$	$\Psi(1/2)$	$\Psi(t_2)$
$y=y_4$	$\Psi(1-t_2)$	$\Psi(1-t_2)$	$\Psi(1-t_2)$	$\Psi(1/2)$

Lemma E.2. *When $\Psi(\cdot)$ is continuous at $1/2$. Achieving Smith consistency only if*

$$\Psi(t)+\Psi(1-t) \leqslant 2\Psi(1/2), \forall t \in [0,1].$$

Proof of Lemma E.2. We consider the case where $n = 6$ and the game with the payoff in Table 4 for any $t_1, t_2 > 1/2$. Notice that if $\Psi(t_1)+\Psi(1/2)+\Psi(1-t_1) > 3\Psi(t_2)(\geqslant 3\Psi(1/2) > 3\Psi(1-t_2))$, there exists positive $\boldsymbol{\mu} = (\mu_1/3, \mu_1/3, \mu_1/3, \mu_2/3, \mu_2/3, \mu_2/3)$ and $\boldsymbol{\mu}' = (\mu_1'/3, \mu_1'/3, \mu_1'/3\mu_2'/3, \mu_2'/3, \mu_2'/3)$ such that $\mu_1+\mu_2 = \mu_1'+\mu_2' = 1$, and

$$\mu_1\left[\Psi(1/2)+\Psi(t_1)+\Psi(1-t_1)-3\Psi(t_2)\right] = \mu_2\left[\Psi(1/2)+\Psi(t_1)+\Psi(1-t_1)-3\Psi(1-t_2)\right],$$
$$\mu_1'\left[\Psi(1/2)+\Psi(t_1)+\Psi(1-t_1)-3\Psi(1-t_2)\right] = \mu_2'\left[\Psi(1/2)+\Psi(t_1)+\Psi(1-t_1)-3\Psi(t_2)\right].$$

Hence, we have

$$\begin{aligned}
&\mu_1(\Psi(1/2)+\Psi(t_1)+\Psi(1-t_1))+3\mu_2\Psi(1-t_2)\\
=&\,\mu_2(\Psi(1/2)+\Psi(t_1)+\Psi(1-t_1))+3\mu_1\Psi(t_2) := 3A,\\
&\mu_1'(\Psi(1/2)+\Psi(t_1)+\Psi(1-t_1))+3\mu_2'\Psi(t_2)\\
=&\,\mu_2'(\Psi(1/2)+\Psi(t_1)+\Psi(1-t_1))+3\mu_1'\Psi(1-t_2) := 3B.
\end{aligned}$$

Thus, we have

$$\begin{aligned}
\mathcal{P}_\Psi(\boldsymbol{\pi}, \boldsymbol{\mu}') = &\,(\pi_1+\pi_2+\pi_3)\left[\frac{\mu_1'}{3}(\Psi(1/2)+\Psi(t_1)+\Psi(1-t_1))+\mu_2'\Psi(t_2)\right]\\
&+(\pi_4+\pi_5+\pi_6)\left[\mu_1'\Psi(1-t_2)+\frac{\mu_2'}{3}(\Psi(1/2)+\Psi(t_1)+\Psi(1-t_1))\right] = B,
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{P}_\Psi(\boldsymbol{\mu}, \boldsymbol{\pi}) = &\,(\pi_1+\pi_2+\pi_3)\left[\frac{\mu_1}{3}(\Psi(1/2)+\Psi(t_1)+\Psi(1-t_1))+\mu_2\Psi(1-t_2)\right]\\
&+(\pi_4+\pi_5+\pi_6)\left[\mu_1\Psi(t_2)+\frac{\mu_2}{3}(\Psi(1/2)+\Psi(t_1)+\Psi(1-t_1))\right] = A.
\end{aligned}$$

Therefore, $\boldsymbol{\mu} \in \arg\max_{\boldsymbol{\pi}} \mathcal{P}_\Psi(\boldsymbol{\pi}, \boldsymbol{\mu}')$, $\boldsymbol{\mu}' \in \arg\min_{\boldsymbol{\pi}} \mathcal{P}_\Psi(\boldsymbol{\mu}, \boldsymbol{\pi})$, which provides that $(\boldsymbol{\mu}, \boldsymbol{\mu}')$ is a Nash equilibrium. However, this causes a contradiction to the fact that the Nash solution supports on $S_1 := \{y_1, y_2, y_3\}$. Thus, it must hold that $\Psi(t_1)+\Psi(1/2)+\Psi(1-t_1) \leqslant 3\Psi(t_2)$ for any $t_1, t_2 > 1/2$. Let $t_2 \to 1/2$, we obtain $\Psi(t)+\Psi(1-t) \leqslant 2\Psi(1/2)$ for any $t \in [0,1]$. □

Table 4: Payoff matrix with six responses $\{y_1, y_2, y_3, y_4, y_5, y_6\}$.

$\Psi(\mathcal{P}(y \succ y'))$	$y' = y_1$	$y' = y_2$	$y' = y_3$	$y' = y_4$	$y' = y_5$	$y' = y_6$
$y = y_1$	$\Psi(1/2)$	$\Psi(t_1)$	$\Psi(1-t_1)$	$\Psi(t_2)$	$\Psi(t_2)$	$\Psi(t_2)$
$y = y_2$	$\Psi(1-t_1)$	$\Psi(1/2)$	$\Psi(t_1)$	$\Psi(t_2)$	$\Psi(t_2)$	$\Psi(t_2)$
$y = y_3$	$\Psi(t_1)$	$\Psi(1-t_1)$	$\Psi(1/2)$	$\Psi(t_2)$	$\Psi(t_2)$	$\Psi(t_2)$
$y = y_4$	$\Psi(1-t_2)$	$\Psi(1-t_2)$	$\Psi(1-t_2)$	$\Psi(1/2)$	$\Psi(t_1)$	$\Psi(1-t_1)$
$y = y_5$	$\Psi(1-t_2)$	$\Psi(1-t_2)$	$\Psi(1-t_2)$	$\Psi(1-t_1)$	$\Psi(1/2)$	$\Psi(t_1)$
$y = y_6$	$\Psi(1-t_2)$	$\Psi(1-t_2)$	$\Psi(1-t_2)$	$\Psi(t_1)$	$\Psi(1-t_1)$	$\Psi(1/2)$

Finally, we prove that the condition (E.1) is also sufficient for Smith consistency. Suppose that $(\boldsymbol{\pi}_1^\star, \boldsymbol{\pi}_2^\star)$ is a Nash equilibrium, notice that

$$\begin{aligned} \mathcal{P}_\Psi(\boldsymbol{\pi}_1^\star, \boldsymbol{\pi}_2^\star) &= \max_{\boldsymbol{\pi}} \mathcal{P}_\Psi(\boldsymbol{\pi}, \boldsymbol{\pi}_2^\star) \geqslant \mathcal{P}_\Psi(\boldsymbol{\pi}_2^\star, \boldsymbol{\pi}_2^\star) = \Psi(1/2)\,, \\ \mathcal{P}_\Psi(\boldsymbol{\pi}_1^\star, \boldsymbol{\pi}_2^\star) &= \min_{\boldsymbol{\pi}} \mathcal{P}_\Psi(\boldsymbol{\pi}_1^\star, \boldsymbol{\pi}) \leqslant \mathcal{P}_\Psi(\boldsymbol{\pi}_1^\star, \boldsymbol{\pi}_1^\star) = \Psi(1/2)\,, \end{aligned} \tag{E.2}$$

which follows from the following fact: for any $\boldsymbol{\pi}$,

$$\mathcal{P}_\Psi(\boldsymbol{\pi}, \boldsymbol{\pi}) = \sum_{i=1}^n \sum_{j=1}^n \pi_i \pi_j \Psi_{ij} = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \pi_i \pi_j \left(\Psi_{ij} + \Psi_{ji}\right) = \Psi(1/2) \sum_{i=1}^n \sum_{j=1}^n \pi_i \pi_j = \Psi(1/2)\,.$$

Thus, from (E.2), we have $\mathcal{P}_\Psi(\boldsymbol{\pi}_1^\star, \boldsymbol{\pi}_2^\star) = \Psi(1/2)$. Then we prove $\operatorname{supp}(\boldsymbol{\pi}_1^\star) \subseteq S_1$. Hence, the Nash solution is Smith consistent, i.e., only supports on S_1.

Case 1. If $\operatorname{supp}(\boldsymbol{\pi}_1^\star) \bigcap S_1 = \emptyset$, taking any $j \in S_1$, we have

$$\mathcal{P}_\Psi(\boldsymbol{\pi}_1^\star, \boldsymbol{\pi}_2^\star) = \min_{\boldsymbol{\pi}} \mathcal{P}_\Psi(\boldsymbol{\pi}_1^\star, \boldsymbol{\pi}) \leqslant \mathcal{P}_\Psi(\boldsymbol{\pi}_1^\star, \boldsymbol{\delta}_j) = \sum_{i=1}^n \pi_{1,i}^\star \Psi_{ij} = \sum_{i \in S_1^c} \pi_{1,i}^\star \Psi_{ij} < \Psi(1/2)\,,$$

which causes a contradiction to the fact that $\mathcal{P}_\Psi(\boldsymbol{\pi}_1^\star, \boldsymbol{\pi}_2^\star) = \Psi(1/2)$.

Case 2. If $\operatorname{supp}(\boldsymbol{\pi}_1^\star) \bigcap S_1 \neq \emptyset$, and $\operatorname{supp}(\boldsymbol{\pi}_1^\star) \bigcap S_1^c \neq \emptyset$, taking $\widetilde{\boldsymbol{\pi}}_2^\star$ as:

$$\widetilde{\pi}_{2,j}^\star = \mathbb{1}\left\{j \in S_1\right\} \cdot \frac{\pi_{1,j}^\star}{\sum_{j \in S_1} \pi_{1,j}^\star}\,.$$

Then we have

$$\begin{aligned} \mathcal{P}_\Psi(\boldsymbol{\pi}_1^\star, \boldsymbol{\pi}_2^\star) &= \min_{\boldsymbol{\pi}} \mathcal{P}_\Psi(\boldsymbol{\pi}_1^\star, \boldsymbol{\pi}) \leqslant \mathcal{P}_\Psi(\boldsymbol{\pi}_1^\star, \widetilde{\boldsymbol{\pi}}_2^\star) \\ &= \sum_{i \in S_1} \sum_{j \in S_1} \pi_{1,i}^\star \widetilde{\pi}_{2,j}^\star \Psi_{ij} + \sum_{i \in S_1^c} \sum_{j \in S_1} \pi_{1,i}^\star \widetilde{\pi}_{2,j}^\star \Psi_{ij} \\ &< \frac{\sum_{i \in S_1} \sum_{j \in S_1} \pi_{1,i}^\star \pi_{1,j}^\star \Psi_{ij}}{\sum_{j \in S_1} \pi_{1,j}^\star} + \Psi(1/2) \sum_{i \in S_1^c} \sum_{j \in S_1} \pi_{1,i}^\star \widetilde{\pi}_{2,j}^\star \\ &= \Psi(1/2) \sum_{i \in S_1} \pi_{1,i}^\star + \Psi(1/2) \sum_{i \in S_1^c} \pi_{1,i}^\star = \Psi(1/2)\,, \end{aligned} \tag{E.3}$$

which follows from the following fact:

$$\begin{aligned} \sum_{i \in S_1} \sum_{j \in S_1} \pi_{1,i}^\star \pi_{1,j}^\star \Psi_{ij} &= \frac{1}{2} \sum_{i \in S_1} \sum_{j \in S_1} \pi_{1,i}^\star \pi_{1,j}^\star \left(\Psi_{ij} + \Psi_{ji}\right) \\ &= \Psi(1/2) \sum_{i \in S_1} \sum_{j \in S_1} \pi_{1,i}^\star \pi_{1,j}^\star = \Psi(1/2) \left(\sum_{i \in S_1} \pi_{1,i}^\star\right) \left(\sum_{j \in S_1} \pi_{1,j}^\star\right) \end{aligned}$$

However, (E.3) also causes a contradiction to the fact that $\mathcal{P}_\Psi(\boldsymbol{\pi}_1^\star, \boldsymbol{\pi}_2^\star) = \Psi(1/2)$.

Therefore, it must hold that $\operatorname{supp}(\boldsymbol{\pi}_1^\star) \bigcap S_1^c = \emptyset$, i.e., $\operatorname{supp}(\boldsymbol{\pi}_1^\star) \subseteq S_1$. We conclude our proof. □

F Proofs of Results in Section 5

F.1 Proof of Lemma 5.1

Proof of Lemma 5.1. Suppose each player has n policies and the payoff matrix is $\{\alpha_{ij}\}_{i=1}^n$. Then,

$$\max_{\boldsymbol{\pi}} \min_{\boldsymbol{\pi}'} \left\{ \sum_{i=1}^n \sum_{j=1}^n \alpha_{ij} \pi_i \pi_j' \right\} = \max_{\boldsymbol{\pi}} \min_{\boldsymbol{\pi}'} \left\{ \sum_{j=1}^n \left(\sum_{i=1}^n \alpha_{ij} \pi_i \right) \pi_j' \right\} = \max_{\boldsymbol{\pi}} \min_{j} \left\{ \sum_{i=1}^n \alpha_{ij} \pi_i \right\}.$$

Let us reformulated it into a convex optimization problem.

$$\begin{aligned} \min_{\boldsymbol{\pi}} \quad & \max_j \sum_{i=1}^n \alpha_{ij} \pi_i \\ \text{subject to} \quad & -\pi_i \leqslant 0, \quad i = 1, \ldots, n \\ & \sum_{i=1}^n \pi_i - 1 = 0 \end{aligned} \tag{P}$$

Let us further reformulate this problem into the epigraph form by introducing a single variable $t \in \mathbb{R}$:

$$\begin{aligned} \min_{\boldsymbol{\pi}, t} \quad & t \\ \text{subject to} \quad & \sum_{i=1}^n \alpha_{ij} \pi_i - t \leqslant 0, \quad j = 1, \ldots, n \\ & -\pi_i \leqslant 0, \quad i = 1, \ldots, n \\ & \sum_{i=1}^n \pi_i - 1 = 0 \end{aligned} \tag{P'}$$

By introducing the dual variables $\boldsymbol{u}^* \in \mathbb{R}^n$, $\tilde{\boldsymbol{u}}^* \in \mathbb{R}^n$ and $v^* \in \mathbb{R}$, the KKT conditions is:

- stationary condition:

$$\sum_{j=1}^n \alpha_{ij} u_j^* - \tilde{u}_i^* = -v^* \quad i = 1, \cdots, n$$

- complementary slackness:

$$u_j^* \left(\sum_{i=1}^n \pi_i^* \alpha_{ij} - t^* \right) = 0 \quad j = 1, \cdots, n$$

$$\tilde{u}_i^* \pi_i^* = 0 \quad i = 1, \cdots, n$$

- primal feasibility:

$$\sum_{i=1}^n \pi_i^* \alpha_{ij} - t^* \leqslant 0$$

$$\boldsymbol{\pi}^* \geqslant 0$$

$$\sum_{i=1}^n \pi_i^* = 1$$

- dual feasibility:

$$\boldsymbol{u}^* \geqslant 0$$

$$\sum_{i=1}^n u_i^* = 1$$

$$\tilde{\boldsymbol{u}}^* \geqslant 0$$

We can easily see that Slater's condition is satisfied for this problem, so the KKT points are equivalent to primal and dual solutions. Then taking $\boldsymbol{\pi}^* > 0$ into account, we have $\tilde{u}_i^* = 0$ by the second complementary slackness condition, and the above equations can be simplified to the following system of equations:

$$\begin{cases} \boldsymbol{u}^* \geqslant 0 \\ \sum_{i=1}^n u_i^* = 1 \\ \sum_{i=1}^n \pi_i^* \alpha_{ij} - t^* \leqslant 0 & j = 1, \cdots, n \\ u_j^* \left(\sum_{i=1}^n \pi_i^* \alpha_{ij} - t^* \right) = 0 & j = 1, \cdots, n \\ \sum_{j=1}^n \alpha_{ij} u_j^* = -v^* & i = 1, \cdots, n \end{cases}$$

Moreover, notice that

$$0 = \sum_{j=1}^n u_j^* \left(\sum_{i=1}^n \pi_i^* \alpha_{ij} - t^* \right) = \sum_{i=1}^n \pi_i^* \sum_{j=1}^n \alpha_{ij} u_j^* - t^* = -v^* - t^*,$$

thus $v^* = -t^*$. Hence, we conclude our proof. □

F.2 Verifying Equation (5.2) and Equation (5.3)

For (5.2), choosing $t^* = -v^* = \sum_{i=1}^n (\pi_i^*)^2$ and $u_i^* = v_i^*$, $\boldsymbol{\pi}^*$ is a Nash solution. For (5.3), choosing $t^* = -v^* = 0$ and $u_j^* = \frac{(\pi_j^*)^{-1}}{\sum_{j=1}^n (\pi_j^*)^{-1}}$, $\boldsymbol{\pi}^*$ is a Nash solution.

F.3 Proof of Theorem 5.1

We first present a useful lemma (Lemma F.1) that further investigates the KKT conditions (Lemma 5.1) when the payoff matrix induces a unique Nash equilibrium.

Lemma F.1. *If a game with the payoff matrix $\{\alpha_{ij}\}_{i,j=1}^n$ has a unique Nash solution $\boldsymbol{\pi}^*$, then for any $j \in [n]$, it must hold $u_j^* > 0$ in the KKT conditions, and*

$$\sum_{i=1}^n \pi_i^* \alpha_{ij} = t^*.$$

Proof of Lemma F.1. Suppose that the KKT conditions provide the unique Nash solution $(\boldsymbol{\pi}^*, \boldsymbol{u}^*, t^*)$. Then we define:

$$\mathcal{J}_0 := \left\{ j \in [n] : u_j^* \neq 0 \right\}, \text{ and } \tilde{\mathcal{J}}_0 := \left\{ j \in [n] : u_j^* = 0 \right\},$$

with $\mathcal{J}_0 \cup \tilde{\mathcal{J}}_0 = [n]$. Since $\boldsymbol{u}^* \geqslant 0$ and $\sum u_j^* = 1$, there exists $j \in [n]$, such that $u_j^* \neq 0$, i.e., $\mathcal{J}_0 \neq \emptyset$. Now, we aim to show $\tilde{\mathcal{J}}_0 = \emptyset$. We prove by contradiction. Suppose $\tilde{\mathcal{J}}_0 \neq \emptyset$, taking $j_0 \in \mathcal{J}_0$, we consider two spaces

$$V_1 := \left\{ \boldsymbol{\pi} \in \mathbb{R}^n : \sum_{i=1}^n \pi_i \left(\alpha_{ij} - \alpha_{ij_0} \right) = 0, \ \forall j \in \mathcal{J}_0 \backslash \{j_0\} \right\},$$

$$V_2 := V_1 \bigcap \left\{ \boldsymbol{\pi} \in \mathbb{R}^n : \sum_{i=1}^n \pi_i (\alpha_{ij} - \alpha_{ij_0}) \leqslant 0, \ \forall j \in \tilde{\mathcal{J}}_0 \right\}.$$

Then we claim that $\boldsymbol{\pi}^* \in V_2$ and $\dim(V_2) \geqslant 2$. For the first claim, by the KKT conditions in Lemma 5.1, for any $j \in \mathcal{J}_0$, we obtain

$$\sum_{i=1}^n \pi_i^* \alpha_{ij} = t^* = \sum_{i=1}^n \pi_i^* \alpha_{ij_0},$$

thus $\boldsymbol{\pi}^* \in V_1$. Moreover, again by the KKT conditions, for any $j \in \tilde{\mathcal{J}}_0$, we have

$$\sum_{i=1}^n \pi_i^* \alpha_{ij} \leqslant t^* = \sum_{i=1}^n \pi_i^* \alpha_{ij_0},$$

which shows that $\boldsymbol{\pi}^* \in V_2$. For the second claim, take $\tilde{j}_0 \in \tilde{\mathcal{J}}_0$ and consider

$$V_3 := \left\{ \boldsymbol{\pi} \in \mathbb{R}^n : \sum_{i=1}^n \pi_i \left(\alpha_{ij} - \alpha_{ij_0}\right) = 0,\ \forall j \in [n] \backslash \{j_0, \tilde{j}_0\} \right\},$$

$$V_4 := V_3 \bigcap \left\{ \boldsymbol{\pi} \in \mathbb{R}^n : \sum_{i=1}^n \pi_i (\alpha_{i\tilde{j}_0} - \alpha_{ij_0}) \leqslant 0 \right\}.$$

We can easily see $V_4 \subseteq V_2$. Note that V_3 can be regarded as a kernel space of a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^{n-2}$. By the dimension theorem in linear algebra, we obtain $\dim(V_3) = n - \dim(\mathrm{Im}(A)) \geqslant n - (n-2) = 2$. For any $\boldsymbol{\pi} \in V_3$, it must hold that $\boldsymbol{\pi} \in V_4$ or $-\boldsymbol{\pi} \in V_4$, so $\dim(V_4) = \dim(V_3) \geqslant 2$. Therefore, we have $\dim(V_1) \geqslant \dim(V_2) \geqslant \dim(V_4) \geqslant 2$.

Thus, we can take another $\tilde{\boldsymbol{\pi}}^* \in V_2$ which is linear independent with $\boldsymbol{\pi}^*$. Note that for any $a, b \in \mathbb{R}_+$, $a\boldsymbol{\pi}^* + b\tilde{\boldsymbol{\pi}}^* \in V_2$. Taking large $a \in \mathbb{R}_+$, we have $a\boldsymbol{\pi}^* + b\tilde{\boldsymbol{\pi}}^* \in V_2$ and $a\boldsymbol{\pi}^* + b\tilde{\boldsymbol{\pi}}^* > 0$, since $\pi^* > 0$. Therefore, there exists $a_1 \in \mathbb{R}_+$, such that $\boldsymbol{\pi}_2^* := \frac{a\boldsymbol{\pi}^* + \tilde{\boldsymbol{\pi}}^*}{a_1} \in V_2$ that satisfies $\boldsymbol{\pi}_2^* \neq \boldsymbol{\pi}^*$, $\boldsymbol{\pi}_2^* > 0$, and $\sum_i \pi_{2,i}^* = 1$. Thus, we obtain another Nash equilibrium $(\boldsymbol{\pi}_2^*, \boldsymbol{u}^*, t^*)$, causing contradiction to the uniqueness of Nash solution. Hence, it must hold that $\tilde{\mathcal{J}}_0 = \emptyset$. □

Next we provide the proof for Theorem 5.1.

Proof of Theorem 5.1. Using Lemma F.1, uniqueness requires us to seek solutions that satisfies

$$\sum_{i=1}^n \pi_i^* \alpha_{ij} = t^*$$

for all $j \in [n]$, where t^* is a constant that may depend on $\boldsymbol{\pi}^*$. Consider $n \geqslant 5$, for any four distinct indices j_1, j_2, k_1, k_2, we have

$$\begin{aligned} &\sum_{i \neq j_1, k_1, k_2} \pi_i f\left(\frac{\pi_i}{\pi_{j_1}}\right) + \pi_{k_1} f\left(\frac{\pi_{k_1}}{\pi_{j_1}}\right) + \pi_{k_2} f\left(\frac{\pi_{k_2}}{\pi_{j_1}}\right) + C\pi_{j_1} \\ =\ &\sum_{i \neq j_2, k_1, k_2} \pi_i f\left(\frac{\pi_i}{\pi_{j_2}}\right) + \pi_{k_1} f\left(\frac{\pi_{k_1}}{\pi_{j_2}}\right) + \pi_{k_2} f\left(\frac{\pi_{k_2}}{\pi_{j_2}}\right) + C\pi_{j_2} \end{aligned} \tag{F.1}$$

Let us consider the infinitesimal variation $\pi_{k_1} \to \pi_{k_1} + \delta$ and $\pi_{k_2} \to \pi_{k_2} - \delta$, keeping others still. We obtain that

$$\begin{aligned} &\sum_{i \neq j_1, k_1, k_2} \pi_i f\left(\frac{\pi_i}{\pi_{j_1}}\right) + (\pi_{k_1} + \delta) f\left(\frac{\pi_{k_1} + \delta}{\pi_{j_1}}\right) + (\pi_{k_2} - \delta) f\left(\frac{\pi_{k_2} - \delta}{\pi_{j_1}}\right) + C\pi_{j_1} \\ =\ &\sum_{i \neq j_2, k_1, k_2} \pi_i f\left(\frac{\pi_i}{\pi_{j_2}}\right) + (\pi_{k_1} + \delta) f\left(\frac{\pi_{k_1} + \delta}{\pi_{j_2}}\right) + (\pi_{k_2} - \delta) f\left(\frac{\pi_{k_2} - \delta}{\pi_{j_2}}\right) + C\pi_{j_2} \end{aligned} \tag{F.2}$$

Subtracting both sides of (F.1) from (F.2), we obtain that

$$\begin{aligned} &(\pi_{k_1} + \delta) f\left(\frac{\pi_{k_1} + \delta}{\pi_{j_1}}\right) + (\pi_{k_2} - \delta) f\left(\frac{\pi_{k_2} - \delta}{\pi_{j_1}}\right) - \pi_{k_1} f\left(\frac{\pi_{k_1}}{\pi_{j_1}}\right) - \pi_{k_2} f\left(\frac{\pi_{k_2}}{\pi_{j_1}}\right) \\ =\ &(\pi_{k_1} + \delta) f\left(\frac{\pi_{k_1} + \delta}{\pi_{j_2}}\right) + (\pi_{k_2} - \delta) f\left(\frac{\pi_{k_2} - \delta}{\pi_{j_2}}\right) - \pi_{k_1} f\left(\frac{\pi_{k_1}}{\pi_{j_2}}\right) - \pi_{k_2} f\left(\frac{\pi_{k_2}}{\pi_{j_2}}\right) \end{aligned} \tag{F.3}$$

i.e., we have

$$\begin{aligned} &(\pi_{k_1} + \delta)\left(f\left(\frac{\pi_{k_1} + \delta}{\pi_{j_1}}\right) - f\left(\frac{\pi_{k_1}}{\pi_{j_1}}\right)\right) + \delta f\left(\frac{\pi_{k_1}}{\pi_{j_1}}\right) \\ &+ (\pi_{k_2} - \delta)\left(f\left(\frac{\pi_{k_2} - \delta}{\pi_{j_1}}\right) - f\left(\frac{\pi_{k_2}}{\pi_{j_1}}\right)\right) - \delta f\left(\frac{\pi_{k_2}}{\pi_{j_1}}\right) \\ =\ &(\pi_{k_1} + \delta)\left(f\left(\frac{\pi_{k_1} + \delta}{\pi_{j_2}}\right) - f\left(\frac{\pi_{k_1}}{\pi_{j_2}}\right)\right) + \delta f\left(\frac{\pi_{k_1}}{\pi_{j_2}}\right) \\ &+ (\pi_{k_2} - \delta)\left(f\left(\frac{\pi_{k_2} - \delta}{\pi_{j_2}}\right) - f\left(\frac{\pi_{k_2}}{\pi_{j_2}}\right)\right) - \delta f\left(\frac{\pi_{k_2}}{\pi_{j_2}}\right). \end{aligned} \tag{F.4}$$

As f is smooth, using

$$\lim_{\delta\to 0}\frac{f\left(\frac{x+\delta}{\pi_j}\right)-f\left(\frac{x}{\pi_j}\right)}{\delta}=\frac{1}{\pi_j}f'\left(\frac{x}{\pi_j}\right),$$

and taking $\delta\to 0$, we obtain the following identity from (F.4),

$$\begin{aligned}&f\left(\frac{\pi_{k_1}}{\pi_{j_1}}\right)+\frac{\pi_{k_1}}{\pi_{j_1}}f'\left(\frac{\pi_{k_1}}{\pi_{j_1}}\right)-f\left(\frac{\pi_{k_2}}{\pi_{j_1}}\right)-\frac{\pi_{k_2}}{\pi_{j_1}}f'\left(\frac{\pi_{k_2}}{\pi_{j_1}}\right)\\&=f\left(\frac{\pi_{k_1}}{\pi_{j_2}}\right)+\frac{\pi_{k_1}}{\pi_{j_2}}f'\left(\frac{\pi_{k_1}}{\pi_{j_2}}\right)-f\left(\frac{\pi_{k_2}}{\pi_{j_2}}\right)-\frac{\pi_{k_2}}{\pi_{j_2}}f'\left(\frac{\pi_{k_2}}{\pi_{j_2}}\right).\end{aligned}\tag{F.5}$$

Thus, we obtain that

$$\begin{aligned}&f\left(\frac{\pi_{k_1}}{\pi_{j_1}}\right)+\frac{\pi_{k_1}}{\pi_{j_1}}f'\left(\frac{\pi_{k_1}}{\pi_{j_1}}\right)-f\left(\frac{\pi_{k_1}}{\pi_{j_2}}\right)-\frac{\pi_{k_1}}{\pi_{j_2}}f'\left(\frac{\pi_{k_1}}{\pi_{j_2}}\right)\\&=f\left(\frac{\pi_{k_2}}{\pi_{j_1}}\right)+\frac{\pi_{k_2}}{\pi_{j_1}}f'\left(\frac{\pi_{k_2}}{\pi_{j_1}}\right)-f\left(\frac{\pi_{k_2}}{\pi_{j_2}}\right)-\frac{\pi_{k_2}}{\pi_{j_2}}f'\left(\frac{\pi_{k_2}}{\pi_{j_2}}\right).\end{aligned}\tag{F.6}$$

Since (F.6) holds for any $\pi>0$, given any $\pi_{j_1}\neq\pi_{j_2}$, for any $x_1,x_2\in(0,1-\pi_{j_1}-\pi_{j_2})$, we have

$$\begin{aligned}&f\left(\frac{x_1}{\pi_{j_1}}\right)+\frac{x_1}{\pi_{j_1}}f'\left(\frac{x_1}{\pi_{j_1}}\right)-f\left(\frac{x_1}{\pi_{j_2}}\right)-\frac{x_1}{\pi_{j_2}}f'\left(\frac{x_1}{\pi_{j_2}}\right)\\&=f\left(\frac{x_2}{\pi_{j_1}}\right)+\frac{x_2}{\pi_{j_1}}f'\left(\frac{x_2}{\pi_{j_1}}\right)-f\left(\frac{x_2}{\pi_{j_2}}\right)-\frac{x_2}{\pi_{j_2}}f'\left(\frac{x_2}{\pi_{j_2}}\right),\end{aligned}$$

which induces the following for any $x\in(0,1-\pi_{j_1}-\pi_{j_2})$,

$$f\left(\frac{x}{\pi_{j_1}}\right)+\frac{x}{\pi_{j_1}}f'\left(\frac{x}{\pi_{j_1}}\right)-f\left(\frac{x}{\pi_{j_2}}\right)-\frac{x}{\pi_{j_2}}f'\left(\frac{x}{\pi_{j_2}}\right)=C(\pi_{j_1},\pi_{j_2}).\tag{F.7}$$

i.e., we have

$$f\left(\frac{x}{\pi_{j_1}}\right)+\frac{x}{\pi_{j_1}}f'\left(\frac{x}{\pi_{j_1}}\right)=C(\pi_{j_1},\pi_{j_2})+f\left(\frac{x}{\pi_{j_2}}\right)+\frac{x}{\pi_{j_2}}f'\left(\frac{x}{\pi_{j_2}}\right).\tag{F.8}$$

Without any loss of generality, we assume $\pi_{j_1}<\pi_{j_2}$, then we obtain

$$\begin{aligned}&f\left(\frac{x}{\pi_{j_1}}\right)+\frac{x}{\pi_{j_1}}f'\left(\frac{x}{\pi_{j_1}}\right)\\&=C(\pi_{j_1},\pi_{j_2})+f\left(\frac{x}{\pi_{j_2}}\right)+\frac{x}{\pi_{j_2}}f'\left(\frac{x}{\pi_{j_2}}\right)\\&=2C(\pi_{j_1},\pi_{j_2})+f\left(\frac{\pi_{j_1}x}{\pi_{j_2}^2}\right)+\frac{\pi_{j_1}x}{\pi_{j_2}^2}f'\left(\frac{\pi_{j_1}x}{\pi_{j_2}^2}\right)\\&=\cdots\cdots\\&=nC(\pi_{j_1},\pi_{j_2})+f\left(\frac{\pi_{j_1}^{n-1}x}{\pi_{j_2}^n}\right)+\frac{\pi_{j_1}^{n-1}x}{\pi_{j_2}^n}f'\left(\frac{\pi_{j_1}^{n-1}x}{\pi_{j_2}^n}\right)\\&=\cdots\cdots.\end{aligned}$$

Taking limit, it must hold $C(\pi_{j_1},\pi_{j_2})=0$, i.e., we have

$$f\left(\frac{x}{\pi_{j_1}}\right)+\frac{x}{\pi_{j_1}}f'\left(\frac{x}{\pi_{j_1}}\right)=f\left(\frac{x}{\pi_{j_2}}\right)+\frac{x}{\pi_{j_2}}f'\left(\frac{x}{\pi_{j_2}}\right).\tag{F.9}$$

Since (F.9) holds for any $\pi>0$, for any $x_1,x_2\in\mathbb{R}_+$, we have

$$f(x_1)+x_1f'(x_1)=f'(x_2)+x_2f'(x_2),$$

thus, for any $x\in\mathbb{R}_+$, we have

$$f(x)+xf'(x)=C_1.\tag{F.10}$$

Solving (F.10), we obtain that

$$f(x) = \frac{C_2}{x} + C_3 \,.$$

Then we obtain that

$$\sum_{i=1}^{n} \pi_i \alpha_{ij} = C\pi_j + \sum_{i \neq j} \pi_i \left(\frac{C_2 \pi_j}{\pi_i} + C_3 \right) = C_3 + (C + (n-1)C_2 - C_3)\pi_j \,,$$

yielding $C_3 = C + (n-1)C_2$, and

$$f(x) = C + C_2 \left(\frac{1}{x} + n - 1 \right),$$

which is contradictory to our assumptions. □