Grouped Satisficing Paths in Pure Strategy Games: a Topological Perspective

Anonymous Author(s)
Affiliation
Address
`email`

Abstract

In game theory and multi-agent reinforcement learning (MARL), each agent selects a strategy, interacts with the environment and other agents, and subsequently updates its strategy based on the received payoff. This process generates a sequence of joint strategies $(s^t)_{t \geq 0}$, where s^t represents the strategy profile of all agents at time step t. A widely adopted principle in MARL algorithms is "win-stay, lose-shift", which dictates that an agent retains its current strategy if it achieves the best response. This principle exhibits a fixed-point property when the joint strategy has become an equilibrium. The sequence of joint strategies under this principle is referred to as a satisficing path, a concept first introduced in [40] and explored in the context of N-player games in [39]. A fundamental question arises regarding this principle: Under what conditions does every initial joint strategy s admit a finite-length satisficing path $(s^t)_{0 \leq t \leq T}$ where $s^0 = s$ and s^T is an equilibrium? This paper establishes a sufficient condition for such a property, and demonstrates that any finite-state Markov game, as well as any N-player game, guarantees the existence of a finite-length satisficing path from an arbitrary initial strategy to some equilibrium. These results provide a stronger theoretical foundation for the design of MARL algorithms.

1 Introduction

Game theory provides a formal framework for analyzing strategic interactions among rational decision-makers. It examines how individuals optimize their decisions to maximize their own payoff while accounting for the actions of others. In most settings, agents act in self-interest, leading to the concept of an equilibrium [24], a strategy profile where no agent can unilaterally deviate to achieve a higher payoff. Due to its generality, game theory has become a cornerstone in machine learning, particularly in modeling competitive and cooperative behaviors, such as multi-agent systems [19], multi-objective reinforcement learning [32], and adversarial learning [18].

Multi-agent reinforcement learning (MARL) extends traditional reinforcement learning frameworks to the situation where multiple autonomous agents interact and make decisions concurrently [26]. In multi-agent systems, each agent learns optimal behavior through an iterative process, that is interacting with both the environment and other agents, receiving the reward based on its actions, and dynamically adapting its policy to maximize long-term returns. This paradigm captures the complex inter-dependencies that emerge when multiple learning agents co-evolve within a shared environment.

From the perspective of game theory, MARL can be modeled as a repeated game [38], where agents iteratively select strategies based on current information, receive rewards from the environment, and update their strategies accordingly. This process generates a strategy path $(s^t)_{t=0}^{T}$, where s^t represents the joint strategy profile at time step t. The concept of equilibrium is particularly crucial

in MARL [23, 12, 14], as it represents a stable case where no agent can improve its payoff through unilateral deviation.

A fundamental question in MARL algorithm design is whether decentralized strategy updates can lead the joint strategy to converge to an equilibrium. While this problem has been extensively studied [36, 22, 27], it remains incompletely resolved [42]. Formally, each agent i updates its strategy according to a revision function: $s_i^{t+1} = f_i(s^u)_{0 \leq u \leq t}$, where f_i maps the history of joint strategies $(s^u)_{0 \leq u \leq t}$ to a new strategy for agent i [5, 28]. A widely adopted principle in MARL algorithms is "win-stay, lose-shift" [9, 41, 16, 29]: If the agent's current strategy is a best response to other agents, it will maintain this strategy, otherwise it will switch to a substitute strategy. The path generated under this principle is termed a satisficing path [40].

This paper studies the satisficing path from a topological perspective, aiming to provide new insights into the following core question:

Question 1. *Under what conditions does the game possess the property that for any initial joint strategy s, there exists a finite-length satisficing path $(s^t)_{0 \leq t \leq T}$ where $s^0 = s$ and s^T is an equilibrium?*

This paper adopts a general theoretical perspective to study the existence conditions of the satisficing path, rather than analyzing specific revision methods or update functions f_i. The primary objective is to establish fundamental theory of the satisficing path, thereby providing essential theoretical supplements to the existing research on convergence problems in MARL.

Contributions. 1) The novel concept of the grouped satisficing path is introduced and formalized in this paper, extending the existing concept of the satisficing path. 2) The structural property of the local minimum in the grouped satisficing path is studied from a topological perspective. 3) The sufficient conditions for the existence of the grouped satisficing path are established, thus solving Question 1. In brief, if the pure strategy game satisfies that a) the strategy set is convex and compact, b) the payoff function is continuous and partially analytic, c) and any sub-game has an equilibrium, then for any initial joint strategy s, there exists a finite-length grouped satisficing path $(s^t)_{0 \leq t \leq T}$ where $s^0 = s$ and s^T is an equilibrium. More details can be found in Theorem 2.

This paper is organized as follows. In Section 2, the theoretical framework is established by formalizing key concepts, and the novel concept of the grouped satisficing path is introduced. In Section 3, the topological properties of the grouped satisficing path are studied, and the conditions for reducing the original game to a sub-game are proposed in Theorem 1. In Section 4, the sufficient conditions for the existence of the grouped satisficing path are established in Theorem 2, and this theorem is applied in various situations, such as continuous games, N-player games and finite-state Markov games. In Section 5, the theoretical implications of these results are examined, some open problems are discussed for future study, and finally there comes to the conclusion of this paper.

Related work. MARL literature contains numerous algorithms employing the iterative strategy adjustment mechanism. Among these, the fictitious play algorithm [7] stands as a paradigmatic example that has profoundly influenced development in this field. While extensive research has investigated the convergence properties of the fictitious play and its variants [25, 33], existing theoretical guarantees remain limited to the specific game structures [3, 34, 35]. The fundamental connection between the fictitious play and the best-response dynamics deserves particular attention. Agents select optimal strategies based on historical joint strategy profiles, and iteratively update their strategies according to the payoff they get. This formulation suggests that the analytical tools from dynamical systems [4, 11, 37] may offer valuable insights for analyzing the convergence of the game and establishing broader convergence guarantees.

The dynamics of the agent strategies in discrete time settings can be formally characterized by the update equation $s_i^{t+1} = f_i(s^t)$, where each agent i's strategy selection at time step $t+1$ is determined by an update rule f_i that depends on the previous joint strategy profile s^t. The seminal work in [20] established a crucial theoretical limitation: for continuous strategy dynamics, if any update function satisfies the regularity conditions and the uncoupled-ness property, the game will be not convergent to an equilibrium generically. Subsequent research [2, 30] has further generalized these impossibility results. However, notable positive convergence results emerge when introducing stochastic elements. For instance, the agent has the probability to update the strategy randomly when it deviates from the best response [21, 15]. The conditional update mechanism has the similar results as the stochastic one [10, 8, 9].

2

The concept of the satisficing path, originally introduced in [40], enforces a key behavior constraint: agents achieving best response must maintain their current strategies, while allowing unconstrained adaptation for other agents. This flexible formulation naturally encompasses a broad spectrum of learning algorithms [1, 6]. Prior research has progressively established existence guarantees for the satisficing path: [16] proved the path existence in two-player games; [40] extended these results to N-player symmetric Markov games; [39] demonstrated universal existence in all N-player games. While these works successfully addressed the existence problem in specific game classes, they could not extend the result in a more general form directly. So it is necessary to establish the theory of the satisficing path. Since this paper studies the topological properties of the satisficing path, the existence results can be successfully extended in a more general form, thus solving Question 1 mentioned above.

2 Formalization

2.1 Pure strategy game

A pure strategy game is defined by the tuple $G = (I, (S_i)_{i \in I}, (g_i)_{i \in I})$, where I represents the set of players, S_i denotes the pure strategy set of player i, and $g_i : \prod_{i \in I} S_i \to \mathbb{R}$ is player i's payoff function. In this paper, only pure strategy games with finite players ($|I| < \infty$) are discussed.

Each player $i \in I$ independently chooses a strategy $s_i \in S_i$ from their respective strategy set. The resulting payoff for player i is determined by the joint strategy profile $(s_1, s_2, \ldots, s_{|I|})$ through the payoff function g_i. Each rational player aims to maximize their individual payoff. However, since the payoff functions depend on all players' simultaneous strategy choices and the players cannot directly control others' strategy selections, the optimal strategic choice reduces to selecting a best response to the current strategies of other players.

Definition 1. *In a pure strategy game G, the ϵ-best response correspondence for player i given opponents' strategy profile s_{-i} is*

$$BR_\epsilon(s_{-i}) = \left\{ s \in S_i \mid g_i(s, s_{-i}) \geq \sup_{t \in S_i} g_i(t, s_{-i}) - \epsilon \right\}. \tag{1}$$

where $s_{-i} \in \prod_{j \neq i, j \in I} S_j$ denotes the joint strategy of all other players. In particular, the notation BR_ϵ will be simplified to BR when $\epsilon = 0$.

In the absence of additional constraints, the best response set $BR(s_{-i})$ may be empty, as the existence of a maximum point in S_i is not guaranteed a priori. When S_i is closed and compact, and g_i is continuous, then $BR(s_{-i})$ is non-empty for all $s_{-i} \in S_{-i}$. For $\epsilon > 0$, the existence condition can be relaxed to the boundedness of g_i.

Definition 2. *In a pure strategy game G, a joint strategy $s \in \prod_{i \in I} S_i$ is called an ϵ-equilibrium if*

$$\forall i \in I, \quad s_i \in BR_\epsilon(s_{-i}). \tag{2}$$

A mixed strategy game can also be defined by the tuple $G = (I, (S_i)_{i \in I}, (g_i)_{i \in I})$. However, each player $i \in I$ can choose a Borel probability measure over S_i as a mixed strategy. Construct a pure strategy game $\tilde{G} = (I, (\Delta S_i)_{i \in I}, (\tilde{g}_i)_{i \in I})$, where ΔS_i denotes the set of all Borel probability measures over S_i, and

$$\tilde{g}_i : \prod_{i \in I} \Delta S_i \to \mathbb{R}, \quad \prod_{i \in I} \sigma_i \mapsto \int_{s \in \prod_{i \in I} S_i} g_i(s) \prod_{i \in I} d\sigma_i(s_i).$$

Then the equilibrium of the pure strategy game $\tilde{G}$ is equivalent to the Nash equilibrium of the original mixed strategy game G. This fundamental equivalence justifies referring to $\tilde{G}$ as the extended game of G.

Remark 1. *The class of pure strategy games exhibits strict representational generality over mixed strategy games. For any mixed strategy game, there exists an equivalent pure strategy game which possesses the same properties of the original game. However, a pure strategy game admits a mixed strategy representation only if each strategy set can be embedded into the set of Borel probability measures over some base space.*

3

2.2 Grouped satisficing path

In a pure strategy game G, the repeated play generates a path $(s^t)_{t \geq 0}$ where s^t represents the joint strategy profile at time step t. In this paper, only discrete time strategy dynamics ($t \in \mathbb{N}$) is discussed.

Definition 3. *In a pure strategy game G, the strategy path $(s^t)_{t \geq 0}$ is called an ϵ-satisficing path if*

$$\forall t \geq 0, \quad \forall i \in I, \quad s_i^t \in BR_\epsilon(s_{-i}^t) \quad \Rightarrow \quad s_i^{t+1} = s_i^t. \tag{3}$$

The definition of the satisficing path is an accurate description of the "win-stay, lose-shift" principle. If the player's strategy is a current best response, it will maintain this strategy in the next time step, otherwise it will explore an alternative strategy to replace it. This principle is fundamental in the MARL algorithm design, since it ensures the stability at the equilibrium.

Definition 4. *In a pure strategy game G, the group set P is a partition of the player set I satisfying:*

    1. $\bigcup_{p \in P} p = I$.

    2. For any $p, q \in P$, if $p \neq q$, then $p \cap q = \varnothing$.

A strategy path $(s^t)_{t \geq 0}$ is called a grouped ϵ-satisficing path with respect to the group set P if

$$\forall t \geq 0, \quad \forall p \in P, \quad (\forall i \in p, s_i^t \in BR_\epsilon(s_{-i}^t)) \quad \Rightarrow \quad (\forall i \in p, s_i^{t+1} = s_i^t). \tag{4}$$

The definition of the grouped satisficing path extends the "win-stay, lose-shift" principle to player groups. The basic decision unit is a group of players rather than individual players. If all players in a group achieve best responses, all of them will maintain their strategies. When each group contains exactly one player ($P = \{\{i\} \mid i \in I\}$), the grouped satisficing path reduces to the classical satisficing path.

Remark 2. *The grouped satisficing path constitutes a proper generalization of the classical satisficing path. Obviously, any satisficing path can be viewed as a grouped satisficing path, while the converse proposition is not always true. The grouped satisficing path serves as an essential technical tool for proving the existence of the satisficing path in finite-state Markov games.*

Definition 5. *In a pure strategy game G with the group set P, for a joint strategy s, the ϵ-best response group count with respect to s is*

$$N_\epsilon(s) = |\{p \in P \mid \forall i \in p, s_i \in BR_\epsilon(s_{-i})\}|. \tag{5}$$

In particular, the notation N_ϵ will be simplified to N when $\epsilon = 0$.

The ϵ-best response group count constitutes an important characteristic of joint strategy profiles. It quantifies the number of player groups where all members simultaneously achieve ϵ-best responses.

Definition 6. *In a pure strategy game G with the group set P, for a joint strategy s, the admissible subsequent joint strategy set with respect to s is*

$$T_\epsilon(s) = \left\{ s' \in \prod_{i \in I} S_i \mid (s, s') \text{ satisfies the grouped } \epsilon\text{-satisficing dynamics} \right\}. \tag{6}$$

$N_\epsilon(s)$ is called a local minimum if

$$N_\epsilon(s) \leq \inf_{t \in T_\epsilon(s)} N_\epsilon(t). \tag{7}$$

$N_\epsilon(s)$ is called a local maximum if

$$N_\epsilon(s) \geq \sup_{t \in T_\epsilon(s)} N_\epsilon(t). \tag{8}$$

In particular, the notation T_ϵ will be simplified to T when $\epsilon = 0$.

The admissible subsequent joint strategy set $T_\epsilon(s)$ is the set of all possible joint strategies following s without violating the property of grouped ϵ-satisficing path. So, a local minimum means that any admissible joint strategy cannot decrease the ϵ-best response group count, while maximum means cannot increase.

2.3   Two critical definitions

Definition 7. *$\{X_i\}_{i \in I}$ is a family of sets, where each X_i is a convex set in a topological vector space. A function $f : \prod_{i \in I} X_i \to \mathbb{R}$ is called a partially analytic function if for any $i \in I$, $p_i, q_i \in X_i$, $x_{-i} \in \prod_{j \neq i, j \in I} X_j$, the function*

$$g : [0,1] \to \mathbb{R}, \quad a \mapsto f(ap_i + (1-a)q_i, x_{-i}).$$

is real analytic with respect to a.

Definition 8. *In a pure strategy game G with the group set P, it is called that any sub-game has an equilibrium, if for any group subset $Q \subset P$ and any joint strategy s, the sub-game $\tilde{G} = (J, (S_i)_{i \in J}, (\tilde{g}_i)_{i \in J})$ admits at least one equilibrium, in which $J = \bigcup_{p \in Q} p$ and*

$$\tilde{g}_i : \prod_{j \in J} S_j \to \mathbb{R}, \quad w \mapsto g_i(s_{-J}, w).$$

where $s_{-J} = (s_k)_{k \in I \setminus J}$ denotes the strategy profile of all players outside J.

The two definitions above represent the two conditions stated in Theorem 2.

3   Topological properties

Two fundamental lemmas are established first and form the basis for proving the main theorem of the topological properties. Lemma 1 provides the necessary conditions for a joint strategy s to be a local minimum of the ϵ-best response group count $N_\epsilon(s)$. Lemma 2 characterizes the local behavior of the best response when a neighborhood of the strategy satisfies optimality conditions. Theorem 1 provides the sufficient conditions under which the original game can be reduced to a well-defined sub-game. Due to space limitation, the detailed proofs of Lemma 1 and Lemma 2 appear in the Appendix.

Lemma 1. *In a pure strategy game G with the group set P, suppose that each S_i is compact and Hausdorff, and each g_i is continuous. For a joint strategy s, if $N_\epsilon(s)$ is a local minimum, then*

1. *either for any joint strategy $t \in T_\epsilon(s)$ in the admissible subsequent joint strategy set, the group which has achieved ϵ-best response in s will still achieve ϵ-best response in t,*

2. *or there exists a non-empty open set $O \subset T_\epsilon(s)$ in the admissible subsequent joint strategy set satisfying:*

   (a) *For any joint strategy $t \in O$, there exists a group which has achieved ϵ-best response in s will not achieve ϵ-best response in t.*
   (b) *There exists a group which has not achieved ϵ-best response in s will achieve ϵ-best response in any joint strategy $t \in O$.*

Lemma 1 is proved in Appendix A.1. In brief, the main idea is to construct a finite family of closed sets covering a given Baire space, and then to discuss different cases about this Baire space. The property of the Baire space will be used in this proof.

Lemma 2. *In a pure strategy game G, suppose that each S_i is a convex set in a topological vector space, and each g_i is partially analytic. If there exists a non-empty open set $U \subset \prod_{i \in I} S_i$ satisfying that the player k achieves best response in any joint strategy $s \in U$, then the player k will achieve best response in any joint strategy $s \in \prod_{i \in I} S_i$.*

Lemma 2 is proved in Appendix A.2. In brief, the main idea is to construct a function to describe the property of the best response, and then to use the analytic extendability to expand the best response set step by step.

Theorem 1. *In a pure strategy game G with the group set P, suppose that each S_i is a convex compact set in a topological vector space, and each g_i is continuous and partially analytic. Then, for a joint strategy s, the best response group count $N(s)$ is a local minimum, if and only if, for any joint strategy $t \in T(s)$ in the admissible subsequent joint strategy set, the group which has achieved best response in s still achieves best response in t.*

Proof. **Sufficiency.** Suppose for any joint strategy $t \in T(s)$ in the admissible subsequent joint strategy set, the group which has achieved best response in s still achieves best response in t. This means that for any joint strategy $t \in T(s)$, there exists $N(s) \leq N(t)$. As a result

$$N(s) \leq \inf_{t \in T(s)} N(t).$$

According to Definition 6, $N(s)$ is a local minimum.

Necessity. Suppose the best response group count $N(s)$ is a local minimum. Since S_i is the set in a topological vector space, S_i is Hausdorff. According to Lemma 1, there are two cases:

1) Either any joint strategy in $T(s)$ makes the group which has achieved best response in s will still achieve,

2) or there exists a non-empty open set $O \subset T(s)$ satisfying: a) There exists one group $p \in P$ which has not achieved best response in s will achieve in any joint strategy in the set O, b) and for any joint strategy in the set O, there exists one group which has achieved best response in s will not achieve.

Suppose case 2) is valid. Assume group $p_1, \ldots, p_n$ do not achieve best response in s. Obviously, $p \in \{p_1, \ldots, p_n\}$. Let $Q = \bigcup_{i=1}^{n} p_i$. Construct a pure strategy game $\tilde{G} = (Q, (S_i)_{i \in Q}, (\tilde{g}_i)_{i \in Q})$. $\tilde{g}_i$ is defined as

$$\tilde{g}_i : \prod_{j \in Q} S_j \to \mathbb{R}, \quad w \mapsto g_i(s_{-Q}, w).$$

where $s_{-Q} = (s_k)_{k \in I \setminus Q}$ denotes the strategy profile of all players outside Q.

Let the projection of the non-empty open set O onto $\prod_{j \in Q} S_j$ be O_p. Obviously, O_p is a non-empty open set, and each player in p achieves best response in any joint strategy in O_p. Each $\tilde{g}_i$ is actually a constrained function of g_i. Since g_i is continuous and partially analytic, $\tilde{g}_i$ is also continuous and partially analytic. According to Lemma 2, for any joint strategy $w \in \prod_{j \in Q} S_j$, each player in p will achieve best response in w. In turn, for any joint strategy $t \in T(s)$, each player in p will achieve best response in t in the game G. However, $s \in T(s)$. This is a contradiction with the fact that group p does not achieve best response in s. So case 2) is invalid.

As a result, for any joint strategy $t \in T(s)$, the group which has achieved best response in s will still achieve best response in t. $\qquad\square$

Remark 3. *This theorem actually establishes a formal reduction principle for games under grouped satisficing dynamics. When there exists a group where all players fix their strategies as their strategies are always best responses in all possible grouped satisficing paths, it can be formally treated as part of the environmental dynamics, reducing the original game to a sub-game.*

4 Existence of paths

In this section, a fundamental existence theorem for the grouped satisficing path in pure strategy games is established. Theorem 2 provides the sufficient conditions for the existence, thus answering Question 1. This theorem yields four corollaries, each addressing distinct game-theoretic scenarios. Due to space limitation, the proofs of these corollaries appear in the Appendix.

Proposition 1. *In a pure strategy game G with the group set P, for any infinite-length path $(s^t)_{t \geq 0}$, there exists an index v satisfying $N_\epsilon(s^v) = \inf_{t \geq 0} N_\epsilon(s^t)$.*

Proposition 1 is proved in Appendix B.

Theorem 2. *In a pure strategy game G with the group set P, suppose that each S_i is a convex compact set in a topological vector space, each g_i is continuous and partially analytic, and any sub-game has an equilibrium. Then for any initial joint strategy s, there exists a finite-length grouped satisficing path $(s^t)_{t=0}^{T}$ where $s^0 = s$ and s^T is an equilibrium.*

Proof. Any finite-length grouped satisficing path $(s^t)_{t=0}^{T}$ where s^T is an equilibrium can be extended to an infinite-length one. Construct a path

$$a^t = \begin{cases} s^t, & 0 \leq t \leq T, \\ s^T, & T < t. \end{cases}$$

Since a^T is an equilibrium, any player will not change their strategies any longer. So $(a^t)_{t\geq 0}$ is a grouped satisficing path.

Construct a set

$$A = \{(s^t)_{t\geq 0} \mid (s^t)_{t\geq 0} \text{ is an infinite-length grouped satisficing path, } s^0 = s\}.$$

According to Proposition 1, each path $(s^t)_{t\geq 0}$ has a minimum of $N(s^t)$. Construct $|P| + 1$ sets

$$A_i = \left\{(s^t)_{t\geq 0} \in A \mid \inf_{t\geq 0} N(s^t) = i\right\}, \quad i = 0, \ldots, |P|.$$

There exists a minimum k satisfying $A_k \neq \varnothing$. Choose a path $(s^t)_{t\geq 0} \in A_k$. According to Proposition 1, there exists $s^x \in (s^t)_{t\geq 0}$ satisfying $N(s^x) = k$.

Reduction to absurdity needs to be used here.

Assume on the contrary that $N(s^x)$ is not local minimum. By Definition 6, there exists a joint strategy u following s^x without violating the property of grouped satisficing path, and satisfying $N(u) \leq N(s^x) - 1$. Construct a path

$$a^t = \begin{cases} s^t, & 0 \leq t \leq x, \\ u, & x < t. \end{cases}$$

It is sure that $(a^t)_{t=0}^{x+1}$ is a grouped satisficing path. For any $t > x$, any player does not change their strategies, which means that the group who has achieved best response will still achieve, and the group who not will still not. So $(a^t)_{t\geq 0}$ is actually a grouped satisficing path. As a result, $(a^t)_{t\geq 0} \in A$, and

$$\inf_{t\geq 0} N(a^t) \leq N(u) \leq N(s^x) - 1 = k - 1.$$

So $(a^t)_{t\geq 0}$ belongs to one of $A_0, \ldots, A_{k-1}$, which is a contradiction with the fact that A_k is the non-empty set with the minimal index.

As a result, $N(s^x)$ is local minimum.

According to Theorem 1, any possible subsequent joint strategy following s^x without violating the property of grouped satisficing path, makes the group which has achieved best response in s^x will still achieve. Assume group $p_1, \ldots, p_n$ do not achieve best response in s^x. Let $J = \bigcup_{i=1}^n p_i$ Construct a pure strategy game $\tilde{G} = (J, (S_i)_{i\in J}, (\tilde{g}_i)_{i\in J})$. $\tilde{g}_i$ is defined as

$$\tilde{g}_i : \prod_{j\in J} S_j \to \mathbb{R}, \quad w \mapsto g_i(s^x_{-J}, w).$$

where $s^x_{-J} = (s^x_k)_{k\in I\setminus J}$ denotes the strategy profile of all players outside J.

By assumption that any sub-game has an equilibrium, there exists an equilibrium w in $\tilde{G}$. Let

$$u = (s^x_{-J}, w) \in \prod_{i\in I} S_i. \tag{9}$$

Construct a finite path

$$a^t = \begin{cases} s^t, & 0 \leq t \leq x, \\ u, & t = x + 1. \end{cases}$$

Obviously, $\{a^t\}_{t=0}^{x+1}$ is a grouped satisficing path, and $a^0 = s$. According to Theorem 1 and Equation 9, the group which has achieved best response in s^x will still achieve in u. Since w is an equilibrium in the sub-game $\tilde{G}$, the group which has not achieved best response in s^x will achieve in u. As a result, a^{x+1} is an equilibrium in G.

So $\{a^t\}_{t=0}^{x+1}$ is a finite-length grouped satisficing path where $a^0 = s$ and a^{x+1} is an equilibrium.

$\square$

This theorem establishes fundamental guarantees for the existence of the grouped satisficing path. While the technical requirements may appear stringent, they are in fact satisfied in numerous practical scenarios.

Corollary 1. *In a pure strategy game G, suppose that each S_i is a convex compact set in a topological vector space and each g_i is continuous and partially analytic. If for any g_i and any $s_{-i} \in \prod_{j \neq i, j \in I} S_j$, the function $g_i(s_i, s_{-i})$ is quasi-convex with respect to s_i, then for any initial joint strategy s, there exists a finite-length satisficing path $(s^t)_{t=0}^T$ where $s^0 = s$ and s^T is an equilibrium.*

Corollary 1 is proved in Appendix B.1. The assumption of this corollary is almost the same as Theorem 2, so only need to check whether any sub-game has an equilibrium.

Corollary 2. *In a mixed strategy game G, suppose that each S_i is a finite set. Then for any initial joint mixed strategy σ, there exists a finite-length satisficing path $(\sigma^t)_{t=0}^T$ where $\sigma^0 = \sigma$ and σ^T is a mixed equilibrium.*

Corollary 2 is proved in Appendix B.2. This corollary is essentially the main result of [39], where the authors directly construct a satisficing path and prove that the equilibrium is a limit point under certain conditions. In contrast, the existence of such a path in N-player games is established as a corollary of Theorem 2 in this paper. Thus, the proof reduces to verifying that the conditions of Theorem 2 hold in this setting.

A stationary mixed strategy stochastic game is defined as $G = (I, (S_i)_{i \in I}, X, P, (g_i)_{i \in I}, (\gamma_i)_{i \in I})$. I is the set of players. S_i is the strategy set of player i. X is the state set. $P : X \times \prod_{i \in I} S_i \to \Delta X$ is the transition probability function, mapping the current state and joint strategy to a probability distribution over the next states. $g_i : X \times \prod_{i \in I} S_i \to \mathbb{R}$ is the payoff function for player i. $\gamma_i \in [0, 1)$ is the discount factor for player i. In this paper, only stationary mixed strategy stochastic games with finite players and finite states ($|I| < \infty, |X| < \infty$) are discussed.

Each player i selects a strategy from S_i according to the probability distribution $\pi_i(x)$. The mapping $\pi_i : X \to \Delta S_i$ is called a stationary mixed strategy, as the player i's strategy depends solely on the current state $x \in X$. Given a joint strategy profile s and the current state x, two events occur: 1) Each player i receives an immediate payoff $g_i(x, s)$, 2) and the game transits to a new state $x' \in X$ with probability $P(x, s)(x')$. As the game progresses, the player i obtains a sequence of payoffs $\{g_i(x^t, s^t)\}_{t \geq 0}$ and aims to maximize the discounted average payoff $\sum_{t \geq 0} \gamma_i^t g_i(x^t, s^t)$.

A joint mixed strategy profile $(\pi_i)_{i \in I}$ constitutes a mixed equilibrium in a stationary mixed strategy stochastic game, if for any initial state $x \in X$ and any player $i \in I$, there exists

$$
\mathbb{E}_{(\pi_j)_{j \in I}} \left[\sum_{t \geq 0} \gamma_i^t g_i(x^t, s^t) \mid x_0 = x \right] = \sup_{\sigma_i \in (\Delta S_i)^X} \mathbb{E}_{\sigma_i, (\pi_j)_{j \neq i, j \in I}} \left[\sum_{t \geq 0} \gamma_i^t g_i(x^t, s^t) \mid x_0 = x \right].
$$

Corollary 3. *In a stationary mixed strategy stochastic game G, suppose that each S_i is a finite set. Then for any initial stationary joint mixed strategy σ, there exists a finite-length satisficing path $(\sigma^t)_{t=0}^T$ where $\sigma^0 = \sigma$ and σ^T is a mixed equilibrium.*

Corollary 3 is proved in Appendix B.3. The proof of this corollary is a bit technically demanding. First, it is necessary to transform the stationary mixed strategy stochastic game with the satisficing path into a pure strategy game with the grouped satisficing path. Since it contains infinite summation, the validity of the operation must be verified. Then, check whether each condition listed in Theorem 2 is satisfied. Contraction mapping theorem will be used in this proof.

Similar to the definition of the stationary mixed strategy stochastic game, a k-step mixed strategy stochastic game can also be defined as $G = (I, (S_i)_{i \in I}, X, P, (g_i)_{i \in I}, (\gamma_i)_{i \in I})$, where all components are defined analogously to the stationary mixed strategy case. However, each player i selects a strategy from S_i according to the probability distribution $\pi_i(x, s^{-1}, \ldots, s^{-k})$, where s^{-t} denotes the joint strategy profile from t steps before the current time step. So $\pi_i : X \times (\prod_{i \in I} S_i)^k \to \Delta S_i$. Similarly, only the case with finite players and finite states ($|I| < \infty, |X| < \infty$) is discussed in this paper.

Corollary 4. *In a k-step mixed strategy stochastic game G, suppose that each S_i is a finite set. Then for any initial stationary joint mixed strategy σ, there exists a finite-length satisficing path $(\sigma^t)_{t=0}^T$ where $\sigma^0 = \sigma$ and σ^T is a mixed equilibrium.*

Corollary 4 is proved in Appendix B.4. The main idea is to find a bijection between the paths in the k-step mixed strategy stochastic game and the paths in a stationary mixed strategy stochastic game, and then to use Corollary 3 to come to the conclusion.

8

Remark 4. *Usually, the game in Corollary 1 is called a continuous game, where the payoff function is continuous. The game in Corollary 2 is termed an N-player game, while the games in Corollary 3 and Corollary 4 are referred to as finite-state Markov games. Corollary 2 establishes the existence of the satisficing path in a general N-player game, Corollary 3 proves its existence in a standard reinforcement learning setting, and Corollary 4 extends the result to the reinforcement learning with historical records.*

5   Discussion

Theorem 1 establishes a framework in which the grouped satisficing paths of the original game can be derived from those of the sub-game. Consequently, the original game can be effectively reduced to its sub-game. Theorem 2 presents sufficient conditions for the existence of the grouped satisficing paths connecting any initial joint strategy to an equilibrium. Thus, Theorem 2 provides a solution to Question 1, or more precisely, it identifies sufficient conditions for the question's resolution.

From a practical standpoint, the corollaries of Theorem 2 may be more valuable. Specifically, these results demonstrate that: $1)$ In any N-player game, there exists satisficing paths from any initial joint mixed strategy to a mixed equilibrium. $2)$ For any finite-state Markov game, whether stationary or finite-step, such satisficing paths always exist from arbitrary initial joint mixed strategy to a mixed equilibrium. This implies that in reinforcement learning with finite states, satisficing paths are guaranteed to exist regardless of whether agents select strategies based on current states or finite history records.

These results remain consistent with the findings in [30, 20, 2], as no restriction is imposed on the update functions in this paper, unlike the regularity or uncoupled-ness conditions required in those studies. Conversely, these results provide theoretical support for stochastic update approaches [21, 15]. The existence of finite-length satisficing paths from any initial joint strategy to an equilibrium implies that: it is possible for stochastic algorithms to start from arbitrary initial joint strategy and to stop at some equilibrium, while keeping the player who has achieved best response not change its strategy.

Open problems. Theorem 2 establishes a sufficient condition for the existence of grouped satisficing paths connecting any initial joint strategy to an equilibrium. This naturally leads to a question: What constitutes a necessary condition for Question 1? More fundamentally, what is a complete necessary and sufficient condition to characterize the existence of such grouped satisficing paths?

Theorem 1 and Theorem 2 are established under the condition of $\epsilon = 0$. This restriction arises because the application of the analytic extendability in Lemma 2 requires the use of certain equations to characterize best responses. So, does Theorem 2 remain valid when $\epsilon > 0$?

Grouped satisficing paths do not impose any neighborhood restriction on strategy selection for the player who has not achieved best response. This raises a question: does Theorem 2 remain valid when non-best-responding players are restricted to select strategies only from a neighborhood of their current strategies? More fundamentally, what restriction can be imposed on non-best-responding players while still preserving the existence of grouped satisficing paths?

Conclusion. In brief, if the pure strategy game satisfies that $a)$ the strategy set is convex and compact, $b)$ the payoff function is continuous and partially analytic, $c)$ and any sub-game has an equilibrium, then for any initial joint strategy s, there exists a finite-length grouped satisficing path $(s^t)_{0 \leq t \leq T}$ where $s^0 = s$ and s^T is an equilibrium. In particular, any N-player game and any finite-state Markov game have the finite-length satisficing paths from any initial joint mixed strategy to some mixed equilibrium.

References

[1] ARSLAN, G., AND YUKSEL, S. Decentralized q-learning for stochastic teams and games. *IEEE Transactions on Automatic Control, 4* (2017), 1545–1558.

[2] BABICHENKO, Y. Completely uncoupled dynamics and nash equilibria. *Games and Economic Behavior 76*, 1 (2012), 1–14.

[3] BAUDIN, L., AND LARAKI, R. Best-response dynamics and fictitious play in identical-interest and zero-sum stochastic games. *International Conference on Machine Learning* (2022), 1664–1690.

[4] BENAIM, M., HOFBAUER, J., AND SORIN, S. Stochastic approximations and differential inclusions. *SIAM Journal on Control and Optimization*, 1 (2005), 328–348.

[5] BLUME, L. E. The statistical mechanics of strategic interaction. *Games and Economic Behavior 5*, 3 (1993), 387–424.

[6] BRIAN, S., CEYHUN, E., SOUMMYA, K., AND ALEJANDRO, R. Distributed inertial best-response dynamics. *IEEE Transactions on Automatic Control 63*, 12 (2018), 4294–4300.

[7] BROWN, G. W. Iterative solution of games by fictitious play. *Activity Analysis of Production and Allocation*, 1 (1951), 374.

[8] CANDOGAN, O., OZDAGLAR, A., AND PARRILO, P. A. Near-potential games: Geometry and dynamics. *ACM Transactions on Economics and Computation*, 2 (2013), 1–32.

[9] CHASPARIS, G. C., ARAPOSTATHIS, A., AND SHAMMA, J. S. Aspiration learning in coordination games. *SIAM Journal on Control and Optimization*, 1 (2013), 465–490.

[10] CHIEN, S., AND SINCLAIR, A. Convergence to approximate nash equilibria in congestion games. *Games and Economic Behavior 71*, 2 (2011), 315–327.

[11] COLLINS, D. S., AND LESLIE, E. J. Individual q-learning in normal form games. *SIAM Journal on Control and Optimization*, 2 (2005), 495–514.

[12] DASKALAKIS, C., FRONGILLO, R., PAPADIMITRIOU, C. H., PIERRAKOS, G., AND VALIANT, G. On learning algorithms for nash equilibria. In *Algorithmic Game Theory - Third International Symposium* (2010), p. 114–125.

[13] FINK, A. Equilibrium in a stochastic n-person game. *Journal of Science of the Hiroshima University 28* (1964), 89–93.

[14] FLOKAS, L., VLATAKIS-GKARAGKOUNIS, E. V., LIANEAS, T., MERTIKOPOULOS, P., AND PILIOURAS, G. No-regreet learning and mixed nash equilibria: They do not mix. *Advances in Neural Information Processing Systems* (2020), 1380–1391.

[15] FOSTER, D., AND YOUNG, H. P. Regret testing: Learning to play nash equilibrium without knowing you have an opponent. *Theoretical Economics 1*, 3 (2006), 341–367.

[16] GERMANO, F., AND LUGOSI, G. Global nash convergence of foster and young's regret testing. *Games and Economic Behavior 60*, 1 (2007), 135–154.

[17] GLICKSBERG, I. L., BURGESS, D. C. J., AND GOCHBERG, I. C. A further generalization of the kakutani fixed point theorem, with application to nash equilibrium points. *Proceedings of the American Mathematical Society 3*, 1 (1952).

[18] GOODFELLOW, I. J., POUGET-ABADIE, J., MIRZA, M., XU, B., WARDE-FARLEY, D., OZAIR, S., COURVILLE, A., AND BENGIO, Y. Generative adversarial nets. *arXiv preprint arXiv:1406.2661* (2014).

[19] GRONAUER, S., AND DIEPOLD, K. Multi-agent deep reinforcement learning: a survey. *Artificial Intelligence Review* (2021), 895–943.

[20] HART, S., AND MAS-COLELL, A. Uncoupled dynamics do not lead to nash equilibrium. *The American economic review 93*, 5 (2003), 1830–1836.

[21] HART, S., AND MAS-COLELL, A. Stochastic uncoupled dynamics and nash equilibrium. *Games and Economic Behavior 57*, 2 (2006), 286–303.

[22] HSIEH, Y. G., ANTONAKOPOULOS, K., AND MERTIKOPOULOS, P. Adaptive learning in continuous games: Optimal regret bounds and convergence to nash equilibrium. *Conference on Learning Theory* (2021), 2388–2422.

[23] JAFARI, A., GREENWALD, A., GONDEK, D., AND ERCAL, G. On no-regret learning, fictitious play, and nash equilibrium. *International Conference on Machine Learning* (2001), 226–233.

[24] LARAKI, R., RENAULT, J., AND SORIN, S. *Mathematical Foundations of Game Theory*. Springer Nature Switzerland AG, Gewerbestrasse 11, 6330 Cham, Switzerland, 2019.

[25] LESLIE, D. S., AND COLLINS, E. J. Generalised weakened fictitious play. *Games and Economic Behavior*, 2 (2006), 285–298.

[26] LI, S. E. *Reinforcement Learning for Sequential Decision and Optimal Control*. Springer Nature Singapore Pte Ltd, 152 Beach Road, 21-01/04 Gateway East, Singapore 189721, Singapore, 2023.

[27] LU, Y. Two-scale gradient descent ascent dynamics finds mixed nash equilibria of continuous games: A mean-field perspective. *International Conference on Machine Learning* (2023), 22790–22811.

[28] MARDEN, J. R., AND SHAMMA, J. S. Revisiting log-linear learning: Asynchrony, completeness and payoff-based implementation. *Games and Economic Behavior 75*, 2 (2012), 788–808.

[29] MARDEN, J. R., YOUNG, H. P., ARSLAN, G., AND SHAMMA, J. S. Payoff-based dynamics for multiplayer weakly acyclic games. *SIAM Journal on Control and Optimization 48*, 1 (2009), 373–396.

[30] MILIONIS, J., PAPADIMITRIOU, C., PILIOURAS, G., AND SPENDLOVE, K. An impossibility theorem in game dynamics. *Proceedings of the National Academy of Sciences 120*, 41 (2023).

[31] NASH, J. F. The bargaining problem. *Econometrica 18*, 2 (1950), 155–162.

[32] ROIJERS, D. M., VAMPLEW, P., WHITESON, S., AND DAZELEY, R. A survey of multi-objective sequential decision-making. *Journal of Artificial Intelligence Research 48*, 1 (2013), 67–113.

[33] SAYIN, M. O. On the global convergence of stochastic fictitious play in stochastic games with turn-based controllers. *arXiv e-prints* (2022).

[34] SAYIN, M. O., PARISE, F., AND OZDAGLAR, A. Best-response dynamics and fictitious play in identical-interest and zero-sum stochastic games. *SIAM Journal on Control and Optimization*, 4 (2022), 2095–2114.

[35] SAYIN, M. O., PARISE, F., AND OZDAGLAR, A. Fictitious play in markov games with single controller. *Proceedings of the 23rd ACM Conference on Economics and Computation* (2022), 919–936.

[36] SINGH, S., KEARNS, M., AND MANSOUR, Y. Nash convergence of gradient dynamics in general-sum games. *Uncertainty in artificial intelligence: Sixteenth conference on uncertainty in artificial intelligence* (2000), 541–548.

[37] SWENSON, B., MURRAY, R., AND KAR, S. On best-response dynamics in potential games. *SIAM Journal on Control and Optimization 56*, 4 (2018), 2734–2767.

[38] TUYLS, K., AND NOWÉ, A. Evolutionary game theory and multi-agent reinforcement learning. *Knowledge Engineering Review 20* (2005), 63–90.

[39] YONGACOGLU, B., ARSLAN, G., PAVEL, L., AND YÜKSEL, S. Paths to equilibrium in games. *38th Conference on Neural Information Processing Systems* (2024).

[40] YONGACOGLU, B., ARSLAN, G., AND YÜKSEL, S. Satisficing paths and independent multiagent reinforcement learning in stochastic games. *SIAM Journal on Mathematics of Data Science*, 3 (2023), 745–773.

[41] YOUNG, H. P., AND PAO, M. L. Y. Achieving pareto optimality through distributed learning. No. 5, p. 2753–2770.

[42] ZHANG, K., YANG, Z., AND BAAR, T. Multi-agent reinforcement learning: A selective overview of theories and algorithms. *Handbook of Reinforcement Learning and Control* (2021).

Appendix

A   Some proofs in Section 3

Definition 9. *In a pure strategy game G, for a joint strategy s, a subset of players $J \subset I$, and a*
player $i \in I \backslash J$, the dual ϵ-best response set of player i over J is

$$BRD_{\epsilon,i}(s, J) = \left\{ x \in \prod_{j \in J} S_j \mid s_i \in BR_\epsilon(s_{-J \cup \{i\}}, x) \right\}. \tag{10}$$

where $s_{-J \cup \{i\}} = (s_k)_{k \in I \backslash (J \cup \{i\})}$ denotes the strategy profile of all players outside $J \cup \{i\}$. In
particular, the notation $BRD_{\epsilon,i}$ will be simplified to BRD_i when $\epsilon = 0$.

The dual ϵ-best response set $BRD_{\epsilon,i}(s, J)$ characterizes the strategic interdependence between
player i and the coalition J. If all players in J adopt strategies $x \in BRD_{\epsilon,i}(s, J)$ while other
players maintain their current strategies $s_{-J \cup \{i\}}$, then player i's current strategy s_i becomes an ϵ-best
response to the resulting strategy profile.

Proposition 2. *In a pure strategy game G, suppose that each g_i is continuous. Then for a joint*
strategy s, a subset of players $J \subset I$, and a player $i \in I \backslash J$, the set $BRD_{\epsilon,i}(s, J)$ is closed.

Proof. By Definition 1 and Definition 9

$$BRD_{\epsilon,i}(s, J) = \left\{ x \in \prod_{j \in J} S_j \mid \forall t \in S_i, \ g_i(s_i, s_{-J \cup \{i\}}, x) \geq g_i(t, s_{-J \cup \{i\}}, x) - \epsilon \right\}.$$

Choose a limit point x of $BRD_{\epsilon,i}(s, J)$, choose any $t \in S_i$, and construct a function

$$h_t : \prod_{j \in J} S_j \to \mathbb{R}, \quad y \mapsto g_i(s_i, s_{-J \cup \{i\}}, y) - g_i(t, s_{-J \cup \{i\}}, y) + \epsilon.$$

Since g_i is continuous, h_t is also continuous. Assume on the contrary that $h_t(x) < 0$, then
$h_t^{-1}(-\infty, 0)$ is an open neighborhood of x. Since x is a limit point of $BRD_{\epsilon,i}(s, J)$, there ex-
ists a point $y \in h_t^{-1}(-\infty, 0) \cap BRD_{\epsilon,i}(s, J)$. However

$$g_i(s_i, s_{-J \cup \{i\}}, y) - g_i(t, s_{-J \cup \{i\}}, y) + \epsilon \geq 0.$$

is a contradiction. As a result, $h_t(x) \geq 0$. Since t is arbitrary

$$\forall t \in S_i, \quad g_i(s_i, s_{-J \cup \{i\}}, x) - g_i(t, s_{-J \cup \{i\}}, x) + \epsilon \geq 0.$$

So $x \in BRD_{\epsilon,i}(s, J)$ which means $BRD_{\epsilon,i}(s, J)$ is a closed set. $\qquad \square$

Proposition 3. *In a pure strategy game G, suppose that each g_i is continuous. Then for any $j \in I$,*
the set

$$A_j = \left\{ s \in \prod_{i \in I} S_i \mid s_j \in BR_\epsilon(s_{-j}) \right\}.$$

is closed.

Proof. This proof is similar to the proof of Proposition 2. Since g_j is continuous, any limit point of
A_j belongs to A_j, which means A_j is closed. $\qquad \square$

A.1   Proof of Lemma 1

Lemma 1. *In a pure strategy game G with the group set P, suppose that each S_i is compact and*
Hausdorff, and each g_i is continuous. For a joint strategy s, if $N_\epsilon(s)$ is a local minimum, then

    1. either for any joint strategy $t \in T_\epsilon(s)$ in the admissible subsequent joint strategy set, the
    group which has achieved ϵ-best response in s will still achieve ϵ-best response in t,

2. *or there exists a non-empty open set $O \subset T_\epsilon(s)$ in the admissible subsequent joint strategy set satisfying:*

    (a) *For any joint strategy $t \in O$, there exists a group which has achieved ϵ-best response in s will not achieve ϵ-best response in t.*

    (b) *There exists a group which has not achieved ϵ-best response in s will achieve ϵ-best response in any joint strategy $t \in O$.*

Proof. Without loss of generality, let $|P| = n + m$, group $p_1, \ldots, p_n$ achieve ϵ-best response in the joint strategy s, while group $p_{n+1}, \ldots, p_{n+m}$ not. Let

$$J = \bigcup_{x=1}^{n} p_x, \quad K = \bigcup_{x=n+1}^{n+m} p_x.$$

By assumption, all players in J achieve ϵ-best response in s. Construct a set

$$A = \bigcap_{j \in J} BRD_{\epsilon,j}(s, K). \tag{11}$$

Since each g_j is continuous, $BRD_{\epsilon,j}(s, K)$ is closed by Proposition 2. A is the intersection of finite closed sets, so A is closed.

Let

$$V = \prod_{k \in K} S_k. \tag{12}$$

Since each S_k is compact and Hausdorff, the product space V is also compact and Hausdorff. According to Baire category theorem, V is a Baire space. The set $V \backslash A$ is the complement of A in V, so $V \backslash A$ is open. The open set in a Baire space is also a Baire space when it is non-empty, so the topological subspace $V \backslash A$ is a Baire space when it is non-empty.

Actually, any possible joint strategy following s without violating the property of grouped ϵ-satisficing path, has the same component of s with index set J, and the entirety of their components with index set K is V, namely

$$T_\epsilon(s) = s_J \times V.$$

As a result, $s_J \times A \subset s_J \times V$ is all admissible subsequent joint strategies which keep players in J achieving ϵ-best response, in other words, keep group $p_1, \ldots, p_n$ achieving ϵ-best response. Since $N_\epsilon(s)$ is a local minimum, any joint strategy $t \in s_J \times (V \backslash A)$ must let at least one group $p \in \{p_{n+1}, \ldots, p_{n+m}\}$ achieve ϵ-best response. Otherwise $N_\epsilon(t) \leq N_\epsilon(s) - 1$, which is a contradiction with Definition 6.

Construct m sets

$$B_{n+i} = \{v \in V \mid w = (s_J, v), \forall x \in p_{n+i}, w_x \in BR_\epsilon(w_{-x})\}, \quad i = 1, \ldots, m. \tag{13}$$

which means any joint strategy in $s_J \times B_{n+i}$ lets group p_{n+i} achieve ϵ-best response. Obviously

$$B_{n+i} = \bigcap_{x \in p_{n+i}} \{v \in V \mid w = (s_J, v), w_x \in BR_\epsilon(w_{-x})\}, \quad i = 1, \ldots, m.$$

According to Proposition 3, each set in the right-hand-side is closed, so B_{n+i} is closed. So $B_{n+i} \cap (V \backslash A)$ is closed in the topological subspace $V \backslash A$, and

$$\bigcup_{i=1}^{m} B_{n+i} \cap (V \backslash A) = V \backslash A. \tag{14}$$

Two cases need to be discussed here.

Case 1. $V \backslash A = \varnothing$. Then $A = V$ which means any possible joint strategy following s without violating the property of grouped ϵ-satisficing path, keeps group $p_1, \ldots, p_n$ achieve ϵ-best response.

Case 2. $V \backslash A \neq \varnothing$. According to Equation 14, $B_{n+i} \cap (V \backslash A)$ is closed and $V \backslash A$ is a Baire space, there exists one $B_{n+i} \cap (V \backslash A)$ has a non-empty interior. So there exists a non-empty open set C satisfying

$$C \subset B_{n+i} \cap (V \backslash A).$$

13

By Equation 13, any joint strategy in $s_J \times C$ lets group p_{n+i} achieve ϵ-best response. By Equation 11, Equation 12 and $C \subset V \backslash A$, any joint strategy in $s_J \times C$ makes at least one group $p \in \{p_1, \ldots, p_n\}$ no longer achieve ϵ-best response. Obviously, $s_J \times C$ is an open set in $T_\epsilon(s)$.

$\square$

Remark 5. *This lemma is very critical for Theorem 1. Usually, one point is defined as a closed set, so a closed set does not necessarily have a non-empty interior. In this lemma, Baire space and Baire theorem provide a guarantee to find a non-empty interior in finite closed sets when they satisfy some conditions.*

A.2   Proof of Lemma 2

Lemma 2. *In a pure strategy game G, suppose that each S_i is a convex set in a topological vector space, and each g_i is partially analytic. If there exists a non-empty open set $U \subset \prod_{i \in I} S_i$ satisfying that the player k achieves best response in any joint strategy $s \in U$, then the player k will achieve best response in any joint strategy $s \in \prod_{i \in I} S_i$.*

Proof. Without loss of generality, let $I = \{1, \ldots, n\}$, and $k = 1$. Since U is open and non-empty, there exists a non-empty open set O

$$O = \prod_{i \in I} O_i \subset U. \tag{15}$$

where O_i is open in S_i. Choose a joint strategy

$$s = (s_i)_{i \in I}, \quad \forall i \in I, \quad s_i \in O_i.$$

Then $s \in O$ which means the joint strategy s makes player k achieve best response.

Induction needs to be used here.

Head. For any $t_k \in S_k$, construct a function

$$f : [0, 1] \to \mathbb{R}, \quad x \mapsto g_k((1-x)s_k + xt_k, s_{-k}) - g_k(s_k, s_{-k}). \tag{16}$$

Obviously, f is an analytic function with respect to x according to the assumption of g_k and Definition 7. Consider

$$t : [0, 1] \to \prod_{i \in I} S_i, \quad x \mapsto ((1-x)s_k + xt_k, s_{-k}).$$

Since each S_i is a convex set in a topological vector space, the finite product space $\prod_{i \in I} S_i$ is a subspace in some topological vector space also. Because $t(0) = s \in O$, O is open and $t(x)$ is a linear mapping, there exists an open neighborhood $(a, b) \subset \mathbb{R}$ of 0 which satisfies

$$\forall x \in [0, b), \quad t(x) \in O.$$

According to Equation 15 and the assumption of U

$$\forall x \in [0, b), \quad (1-x)s_k + xt_k \in BR(s_{-k}).$$

As a result

$$\forall x \in [0, b), \quad g_k((1-x)s_k + xt_k, s_{-k}) = g_k(s_k, s_{-k}) \geq \sup_{r_k \in S_k} g_k(r_k, s_{-k}).$$

So for any $x \in [0, b)$, $f(x) = 0$. Since f is analytic, then

$$\forall x \in [0, 1], \quad f(x) = 0.$$

So $t_k \in BR(s_{-k})$. Due to arbitrary t_k, $S_k = BR(s_{-k})$. Construct a set

$$V_1 = S_k \times \prod_{i=2}^{n} O_i.$$

So each joint strategy $s \in V$ makes player k achieve best response.

Recursion. Assume

$$V_m = S_k \times \prod_{i=2}^{m} S_i \times \prod_{i=m+1}^{n} O_i.$$

509 Each joint strategy $s \in V_m$ makes player k achieve best response.

510 For any $t_k \in S_k$, $t_i \in S_i$, $i = 2, \ldots, m+1$, construct a function

$$
\begin{aligned}
h : &[0,1] \to \mathbb{R}, \\
x \mapsto{} & g_k(s_k, t_2, \ldots, t_m, (1-x)s_{m+1} + x t_{m+1}, s_{m+2}, \ldots, s_n) \\
& - g_k(t_k, t_2, \ldots, t_m, (1-x)s_{m+1} + x t_{m+1}, s_{m+2}, \ldots, s_n).
\end{aligned}
\tag{17}
$$

Since each part of h is analytic, h is analytic. Since O_{m+1} is open, there exists an open neighborhood $(c, d) \subset \mathbb{R}$ of 0 which satisfies

$$\forall x \in [0, d), \quad (1-x)s_{m+1} + x t_{m+1} \in O_{m+1}.$$

So for any $x \in [0, d)$, $h(x) = 0$ according to the assumption of recursion. Due to analytic property, for any $x \in [0, 1]$, $h(x) = 0$. So

$$\forall t_k \in S_k, \quad g_k(t_k, t_2, \ldots, t_m, t_{m+1}, s_{m+2}, \ldots, s_n) = g_k(s_k, t_2, \ldots, t_m, t_{m+1}, s_{m+2}, \ldots, s_n).$$

So

$$S_k \subset BR(t_2, \ldots, t_m, t_{m+1}, s_{m+2}, \ldots, s_n).$$

Construct a set

$$V_{m+1} = S_k \times \prod_{i=2}^{m+1} S_i \times \prod_{i=m+2}^{n} O_i.$$

511 Each joint strategy $s \in V_{m+1}$ makes player k achieve best response.

512 By induction, each joint strategy $s \in \prod_{i \in I} S_i$ makes player k achieve best response.

513 $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

514 **Remark 6.** *The function f in Equation 16 is not similar to h in Equation 17. f is used to extend the*
515 *best response property along S_k, while h to extend the gap between different strategies in S_k along*
516 *S_{m+1}. The continuity of scalar multiplication in topological vector spaces and the extendability of*
517 *analytic functions play a important role in the proof of this lemma.*

518 B Some proofs in Section 4

519 **Proposition 1.** *In a pure strategy game G with the group set P, for any infinite-length path $(s^t)_{t \geq 0}$,*
520 *there exists an index v satisfying $N_\epsilon(s^v) = \inf_{t \geq 0} N_\epsilon(s^t)$.*

Proof. Construct $|P| + 1$ sets

$$A_i = \{t \mid N_\epsilon(s^t) = i\}, \quad i = 0, \ldots, |P|.$$

521 There exists a minimum k making $|A_k| \neq \varnothing$. Choose an element $t \in A_k$. For finite joint strategies
522 $s_0, s_1, \ldots, s_t$, there exists the first index v making $N_\epsilon(s^v) = k$. So $N_\epsilon(s^v) = \inf_{i \geq 0} N_\epsilon(s^i)$. $\quad\square$

523 B.1 Proof of Corollary 1

524 **Corollary 1.** *In a pure strategy game G, suppose that each S_i is a convex compact set in a topological*
525 *vector space and each g_i is continuous and partially analytic. If for any g_i and any $s_{-i} \in \prod_{j \neq i, j \in I} S_j$,*
526 *the function $g_i(s_i, s_{-i})$ is quasi-convex with respect to s_i, then for any initial joint strategy s, there*
527 *exists a finite-length satisficing path $(s^t)_{t=0}^{T}$ where $s^0 = s$ and s^T is an equilibrium.*

528 *Proof.* There exists a theorem with respect to continuous games [17]. In a pure strategy game
529 $G = (I, (S_i)_{i \in I}, (g_i)_{i \in I})$, suppose that each S_i is a convex compact set in a topological vector space,
530 and each g_i is continuous. If for each g_i and any $s_{-i} \in \prod_{j \neq i, j \in I} S_j$, the function $g_i(s_i, s_{-i})$ is
531 quasi-convex with respect to s_i, then G admits at least one equilibrium.

As a result, for any subset $J \subset I$ and any joint strategy $s \in \prod_{i \in I} S_i$, construct a sub-game $\tilde{G} = (J, (S_i)_{i \in J}, (\tilde{g}_i)_{i \in J})$ in which

$$\tilde{g}_i : \prod_{i \in J} S_i \to \mathbb{R}, \quad w \mapsto g_i(s_{-J}, w).$$

532   where $s_{-J} = (s_i)_{i \in I \setminus J}$ denotes the strategy profile of all players outside J.

533   Function $\tilde{g}_i$ is also continuous, and for any $s_{-i} \in \prod_{j \neq i, j \in J} S_j$, the function $g_i(s_i, s_{-i})$ is quasi-
534   convex with respect to s_i as well. So $\tilde{G}$ has an equilibrium according to the theorem with respect to
535   continuous games.

536   So all conditions of Theorem 2 are satisfied. As a result, for any initial joint strategy s in G, there
537   exists a finite-length satisficing path $(s^t)_{t=0}^T$ where $s^0 = s$ and s^T is some equilibrium.   □

538   ## B.2   Proof of Corollary 2

Proposition 4. *Suppose that real numbers $a_1, \ldots, a_n, b_1, \ldots, b_n$ belong to $[0, 1]$, then*

$$\left| \prod_{i=1}^n a_i - \prod_{i=1}^n b_i \right| \leq \sum_{i=1}^n |a_i - b_i|.$$

539   *Proof.* Obviously, the proposition is true when $n = 1$. Assume the proposition is true when $n = k$,
540   then

$$\left| \prod_{i=1}^{k+1} a_i - \prod_{i=1}^{k+1} b_i \right| = \left| \prod_{i=1}^{k+1} a_i - b_1 \prod_{i=2}^{k+1} a_i + b_1 \prod_{i=2}^{k+1} a_i - \prod_{i=1}^{k+1} b_i \right|$$

$$\leq |a_1 - b_1| \left| \prod_{i=2}^{k+1} a_i \right| + |b_1| \left| \prod_{i=2}^{k+1} a_i - \prod_{i=2}^{k+1} b_i \right|$$

$$\leq |a_1 - b_1| + \left| \prod_{i=2}^{k+1} a_i - \prod_{i=2}^{k+1} b_i \right|.$$

By assumption

$$\left| \prod_{i=1}^{k+1} a_i - \prod_{i=1}^{k+1} b_i \right| \leq |a_1 - b_1| + \sum_{i=2}^{k+1} |a_i - b_i|.$$

541   So the proposition is true when $n = k + 1$.   □

542   **Corollary 2.** *In a mixed strategy game G, suppose that each S_i is a finite set. Then for any initial*
543   *joint mixed strategy σ, there exists a finite-length satisficing path $(\sigma^t)_{t=0}^T$ where $\sigma^0 = \sigma$ and σ^T is a*
544   *mixed equilibrium.*

Proof. Construct a pure strategy game $\tilde{G} = (I, (\Delta S_i)_{i \in I}, (\tilde{g}_i)_{i \in I})$ where ΔS_i is the set of all Borel probability measures over S_i, and

$$\tilde{g}_i : \prod_{i \in I} \Delta S_i \to \mathbb{R}, \quad \prod_{i \in I} \sigma_i \mapsto \sum_{s \in \prod_{i \in I} S_i} g_i(s) \prod_{i \in I} \sigma_i(s_i).$$

545   Since S_i is a finite set, ΔS_i is actually a $(|S_i| - 1)$-dimensional simplex in $\mathbb{R}^{|S_i|}$. So ΔS_i is a convex
546   compact set in $\mathbb{R}^{|S_i|}$.

547   **Prove that $\tilde{g}_i$ is continuous.**

g_i is bounded by

$$M = \max_{s \in \prod_{j \in I} S_j} |g_i(s)|.$$

16

The space $\prod_{i \in I} \Delta S_i$ can be endowed with a metric topology.

$$|\sigma - \eta| = \sum_{i \in I} |\sigma_i - \eta_i| = \sum_{i \in I} \sum_{s \in S_i} |\sigma_i(s) - \eta_i(s)|, \quad \sigma, \eta \in \prod_{i \in I} \Delta S_i.$$

Since the topology constructed by metric $\sum_{s \in S_i} |\sigma_i(s) - \eta_i(s)|$ is the same as the topology in $\mathbb{R}^{|S_i|}$, the definition above is reasonable.

For any $\epsilon > 0$, there exists $\delta = \epsilon/M$ such that for any $|\sigma - \eta| < \delta$

$$|\tilde{g}_i(\sigma) - \tilde{g}_i(\eta)| \leq \sum_{s \in \prod_{i \in I} S_i} M \left| \prod_{i \in I} \sigma_i(s_i) - \prod_{i \in I} \eta_i(s_i) \right|.$$

According to proposition 4

$$|\tilde{g}_i(\sigma) - \tilde{g}_i(\eta)| \leq M \sum_{s \in \prod_{i \in I} S_i} \sum_{i \in I} |\sigma_i(s_i) - \eta_i(s_i)| = M|\sigma - \eta| < \epsilon.$$

So $\tilde{g}_i$ is continuous.

Prove that $\tilde{g}_i$ is partially analytic.

For any $j \in I$, $\sigma_j, \eta_j \in \Delta S_j$, $\sigma_{-j} \in \prod_{l \neq j, l \in I} \Delta S_l$, then

$$\tilde{g}_i(x\sigma_j + (1-x)\eta_j, \sigma_{-j}) = \sum_{s \in \prod_{i \in I} S_i} g_i(s)(x\sigma_j(s_j) + (1-x)\eta_j(s_j))\sigma_{-j}(s_{-j}).$$

Obviously, $\tilde{g}_i(x\sigma_j + (1-x)\eta_j, \sigma_{-j})$ is a linear polynomial with respect to x, which means an analytic function with respect to x.

Prove that any sub-game has an equilibrium.

For any subset $J \subset I$ and any joint strategy $\sigma \in \prod_{i \in I} \Delta S_i$, construct a sub-game $\tilde{H} = (J, (\Delta S_i)_{i \in J}, (\tilde{h}_i)_{i \in J})$ where

$$\tilde{h}_i : \prod_{i \in J} \Delta S_i \to \mathbb{R}, \quad \eta \mapsto \tilde{g}_i(\sigma_{-J}, \eta).$$

Construct functions

$$h_i : \prod_{j \in J} S_j \to \mathbb{R}, \quad w \mapsto \sum_{v \in \prod_{j \in I \setminus J} S_j} g_i(w, v)\sigma_{-J}(v).$$

So

$$\tilde{h}_i(\eta) = \sum_{s \in \prod_{i \in J} S_i} h_i(s) \prod_{i \in J} \eta_i(s_i).$$

which means that $\tilde{H}$ is the extended pure strategy game of a mixed strategy game $H = (J, (S_i)_{i \in J}, (h_i)_{i \in J})$.

There exists a theorem with respect to mixed strategy games [31]. Each N-player game has a mixed equilibrium.

As a result, H has a mixed equilibrium and $\tilde{H}$ has an equilibrium.

So all conditions of Theorem 2 are satisfied. As a result, for any initial joint mixed strategy σ in G, there exists a finite-length satisficing path $(\sigma^t)_{t=0}^T$ where $\sigma^0 = \sigma$ and σ^T is a mixed equilibrium.

$\square$

B.3 Proof of Corollary 3

Proposition 5. *For the real number $c \in (0,1)$ and the positive integer p, there exists*

$$\sum_{q \geq p, q \in \mathbb{N}} c^q \binom{q}{p} \leq \frac{2\sqrt{2\pi} p^{1/2}}{(-\ln c)^p}.$$

when p goes to infinity.

Proof. Since $\binom{q}{p} = q(q-1)\cdots(q-p+1)/p! \le q^p/p!$

$$\sum_{q \ge p, q \in \mathbb{N}} c^q \binom{q}{p} \le \frac{1}{p!} \sum_{q \ge p, q \in \mathbb{N}} c^q q^p.$$

The function $f(q) = c^q q^p$ has

$$\frac{d}{dq} f = c^q q^{p-1}(q \ln c + p).$$

Since $f(0^+) = f(+\infty) = 0$ and $f(q) \ge 0$, f has only one maximal point $-p/\ln c$. $df/dq > 0$ when $0 < q < -p/\ln c$, and $df/dq < 0$ when $-p/\ln c < q$. So

$$\sum_{q \ge p, q \in \mathbb{N}} c^q q^p \le \int_0^\infty f(q)dq + 2f\left(-\frac{p}{\ln c}\right).$$

The form of the first term in the right-hand-side is similar to Γ function.

$$\Gamma(n+1) = \int_0^\infty t^n e^{-t} dt = n!.$$

Actually

$$\int_0^\infty f(q)dq = \left(-\frac{1}{\ln c}\right)^{p+1} \int_0^\infty e^{-t} t^p dq = \left(-\frac{1}{\ln c}\right)^{p+1} p!.$$

So

$$\sum_{q \ge p, q \in \mathbb{N}} c^q \binom{q}{p} \le \left(-\frac{1}{\ln c}\right)^{p+1} + 2\left(-\frac{1}{\ln c}\right)^p \frac{e^{-p}p^p}{p!} = \left(-\frac{1}{\ln c}\right)^p \left(-\frac{1}{\ln c} + 2\frac{e^{-p}p^p}{p!}\right).$$

According to Stirling's formula $p! \sim \sqrt{2\pi p} p^p e^{-p}$, there exists $e^{-p} p^p/p! \sim \sqrt{2\pi p}$. So

$$\sum_{q \ge p, q \in \mathbb{N}} c^q \binom{q}{p} \le \left(-\frac{1}{\ln c}\right)^p 2\sqrt{2\pi} p^{1/2}, \quad p \to \infty.$$

565

$\square$

566 **Corollary 3.** *In a stationary mixed strategy stochastic game G, suppose that each S_i is a finite set.*
567 *Then for any initial stationary joint mixed strategy σ, there exists a finite-length satisficing path*
568 $(\sigma^t)_{t=0}^T$ *where $\sigma^0 = \sigma$ and σ^T is a mixed equilibrium.*

Proof. According to the definition of g_i and the finiteness of $X \times \prod_{i \in I} S_i$, g_i is bounded by

$$M = \max_{x \in X, s \in \prod_{i \in I} S_i} |g_i(x, s)|.$$

Construct a pure strategy game $H = ((I, X), (T_{i,x})_{(i,x) \in (I,X)}, (h_{i,x})_{(i,x) \in (I,X)})$. (I, X) means a double index set, namely $(I, X) = \{(i, x) \mid i \in I, x \in X\}$. Since $|I| < \infty$ and $|X| < \infty$, (I, X) is finite.

$$T_{i,x} = \Delta S_i.$$

Since S_i is finite, $T_{i,x}$ is a simplex in some finite-dimensional Euclidean space. So $T_{i,x}$ is convex and compact.

$$h_{i,x} : \prod_{(j,y) \in (I,X)} T_{j,y} \to \mathbb{R}, \quad \prod_{(j,y) \in (I,X)} \pi_j(y) \mapsto \mathbb{E}_{(\pi_j)_{j \in I}}\left[\sum_{t \ge 0} \gamma_i^t g_i(x^t, s^t) \mid x_0 = x\right].$$

So

$$h_{i,x}(\pi) = \mathbb{E}_{s \sim \pi}[g_i(x, s)] + \gamma_i \mathbb{E}_{s \sim \pi, y \sim P(x,s)}[h_{i,y}(\pi)].$$

If $\pi \in \prod_{(j,y) \in (I,X)} T_{j,y}$ is considered as a constant, $h_{i,x}(\pi)$ will be viewed as a function with respect to x, denoted as

$$h_i(\pi) : X \to \mathbb{R}, \quad x \mapsto h_{i,x}(\pi).$$

18

The function space $\mathbb{R}^X$ is endowed with uniform metric topology, using

$$|f| = \sup_{x \in X} |f(x)|, \quad f \in \mathbb{R}^X.$$

as a norm in $\mathbb{R}^X$. Since $\mathbb{R}$ is complete, the uniform metric topology in $\mathbb{R}^X$ is complete.

Construct an operator

$$op_\pi : \mathbb{R}^X \to \mathbb{R}^X, \quad f \mapsto \left\{ (x, \mathbb{E}_{s \sim \pi}[g_i(x, s)] + \gamma_i \mathbb{E}_{s \sim \pi, y \sim P(x,s)}[f(y)]) \mid x \in X \right\}.$$

For any $x \in X$, $p, q \in \mathbb{R}^X$, there exists

$$|op_\pi(p)(x) - op_\pi(q)(x)| = \gamma_i |\mathbb{E}_{s \sim \pi, y \sim P(x,s)}[p(y) - q(y)]| \leq \gamma_i \sup_{y \in X} |p(y) - q(y)| = \gamma_i |p - q|.$$

So

$$|op_\pi(p) - op_\pi(q)| = \sup_{x \in X} |op_\pi(p)(x) - op_\pi(q)(x)| \leq \gamma_i |p - q|.$$

Since $0 \leq \gamma_i < 1$, op_π is a contraction mapping.

According to contraction mapping theorem, there exists a unique fixed point. That is

$$h_i(\pi) = op_\pi(h_i(\pi)). \tag{18}$$

So $h_i(\pi)$ exists and is unique. As a result, $h_{i,x}$ is a well-defined function.

Obviously, $h_{i,x}(\pi)$ is bounded by

$$|h_{i,x}(\pi)| \leq \sum_{t \geq 0} \gamma_i^t M \leq \frac{M}{1 - \gamma_i}.$$

This conclusion is always true no matter what x or π is. So for any $\pi \in \prod_{(j,y) \in (I,X)} T_{j,y}$

$$|h_i(\pi)| \leq \frac{M}{1 - \gamma_i}. \tag{19}$$

Prove that $h_{i,x}$ is continuous.

Since each $T_{i,x}$ is a metric topological space, the finite product space $\prod_{(i,x) \in (I,X)} T_{i,x}$ is also a metric topological space. The metric in this space is denoted as $|\pi - \sigma|$ where $\pi, \sigma \in \prod_{(i,x) \in (I,X)} T_{i,x}$.

For any $x \in X$, $p, q \in \mathbb{R}^X$, $\pi, \sigma \in \prod_{(i,x) \in (I,X)} T_{i,x}$, there exists

$$\begin{aligned}|op_\pi(p)(x) - op_\sigma(q)(x)| \leq &|\mathbb{E}_{s \sim \pi}[g_i(x, s)] - \mathbb{E}_{s \sim \sigma}[g_i(x, s)]| \\ &+ \gamma_i |\mathbb{E}_{s \sim \pi, y \sim P(x,s)}[p(y)] - \mathbb{E}_{s \sim \sigma, y \sim P(x,s)}[q(y)]|.\end{aligned}$$

Since

$$\begin{aligned}|\mathbb{E}_{s \sim \pi}[g_i(x, s)] - \mathbb{E}_{s \sim \sigma}[g_i(x, s)]| &= \left| \sum_{s \in \prod_{j \in I} S_j} g_i(x, s)(\pi(x)(s) - \sigma(x)(s)) \right| \\ &\leq \sum_{s \in \prod_{j \in I} S_j} M \left| \prod_{k \in I} \pi_k(x)(s_k) - \prod_{k \in I} \sigma_k(x)(s_k) \right| \\ &\leq M \sum_{s \in \prod_{j \in I} S_j} \sum_{k \in I} |\pi_k(x)(s_k) - \sigma_k(x)(s_k)| \\ &\leq MC |\pi - \sigma|.\end{aligned}$$

where C is an independent constant. The second $\leq$ inequality is guaranteed by Proposition 4.

$$|\mathbb{E}_{s\sim\pi,y\sim P(x,s)}[p(y)] - \mathbb{E}_{s\sim\sigma,y\sim P(x,s)}[q(y)]|$$

$$= \left| \sum_{s\in\prod_{j\in I}S_j,y\in X} (p(y)\pi(x)(s)P(x,s)(y) - q(y)\sigma(x)(s)P(x,s)(y)) \right|$$

$$\leq \sum_{s\in\prod_{j\in I}S_j,y\in X} |\pi(x)(s) - \sigma(x)(s)||p(y)|P(x,s)(y) + |p(y) - q(y)|\sigma(x)(s)P(x,s)(y)$$

$$\leq \sum_{y\in X} \left(\sum_{s\in\prod_{j\in I}S_j} |\pi(x)(s) - \sigma(x)(s)||p| \right) + |p-q| \sum_{s\in\prod_{j\in I}S_j,y\in X} \sigma(x)(s)P(x,s)(y)$$

$$\leq C|\pi-\sigma||p||X| + |p-q|.$$

580 So

$$|op_\pi(p)(x) - op_\sigma(q)(x)| \leq (MC + \gamma_i C|p||X|)|\pi-\sigma| + \gamma_i|p-q|. \tag{20}$$

581 Here consider $|h_i(\pi) - h_i(\sigma)|$. According to Equation 18, Equation 19 and Equation 20

$$|h_i(\pi) - h_i(\sigma)| = |op_\pi(h_i(\pi)) - op_\sigma(h_i(\sigma))|$$

$$\leq \left(MC + \gamma_i C\frac{M}{1-\gamma_i}|X| \right)|\pi-\sigma| + \gamma_i|h_i(\pi) - h_i(\sigma)|.$$

So

$$|h_i(\pi) - h_i(\sigma)| \leq \frac{MC}{1-\gamma_i}\left(1 + \frac{\gamma_i}{1-\gamma_i}|X| \right)|\pi-\sigma|.$$

582 Since each element in the coefficient of the right-hand-side is independent of π and σ, h_i is continuous
583 about π. As a result, $h_{i,x}$ is continuous about π.

584 **Prove that $h_{i,x}$ is partially analytic.**

585 For any $(j,y) \in (I,X)$, $\pi_j(y), \sigma_j(y) \in T_{j,y}$, $\eta \in \prod_{(k,z)\neq(j,y),(k,z)\in(I,X)} T_{k,z}$, let $\theta_a = (a\pi_j(y) +$
586 $(1-a)\sigma_j(y), \eta)$, then

$$h_{i,x}(\theta_a) = \mathbb{E}_{\theta_a}\left[\sum_{t\geq 0} \gamma_i^t g_i(x^t, s^t) \mid x_0 = x \right]$$

$$= \sum_{t\geq 0} \gamma_i^t \mathbb{E}_{x^0=x,\dots,x^u\sim P(x^{u-1},s^{u-1}),s^u\sim\theta_a(x^u),\dots,s^t\sim\theta_a(x_t)} g_i(x^t, s^t).$$

Considering the boundary of the k-th term $c_k a^k$ in the polynomial with respect to a

$$|c_k a^k| \leq M \sum_{t\geq k} \gamma_i^{t-1}\binom{t}{k}(2a)^k = RHS.$$

587 Attention, the right-hand-side is the boundary of absolute sum of all possible coefficients of a^k
588 appeared in $h_{i,x}(\theta_a)$.

589 When $\gamma_i = 0$, $h_{i,x}$ is obviously analytic with respect to a.

When $0 < \gamma_i < 1$, according to Proposition 5

$$RHS \leq M(2a)^k \gamma_i^{-1}\frac{2\sqrt{2\pi}k^{1/2}}{(-\ln\gamma_i)^k} = \left(\frac{2a(2M)^{1/k}\gamma_i^{-1/k}(2\pi k)^{1/(2k)}}{-\ln\gamma_i} \right)^k, \quad k\to\infty.$$

590 Obviously, $(2M)^{1/k}\gamma_i^{-1/k}(2\pi k)^{1/(2k)} \to 1$ when $k\to\infty$. So the radius of convergence is $-\ln\gamma_i/2$,
591 which is independent of $\pi_j(y)$ and $\sigma_j(y)$. This means that when $|a| < -\ln\gamma_i/2$, $h_{i,x}(\theta_a)$ is an
592 analytic function with respect to a. Since the overall analyticity can be decomposed into the analyticity
593 of segments, $h_{i,x}$ is analytic in the whole $[0,1]$.

20

594 **Prove that any sub-game has an equilibrium.**

Construct a group set

$$A_i = \{(i,x) \mid \forall x \in X\}, \quad A = \{A_i \mid \forall i \in I\}.$$

595 So A_i is actually a real player in the original stochastic game who need choose a probability
596 distribution over S_i for each state in X.

For any joint strategy $\pi \in \prod_{(i,x)\in(I,X)} T_{i,x}$, a subset $\Omega \subset A$, there exists $J \subset I$ satisfying

$$\bigcup_{\omega \in \Omega} \omega = (J,X).$$

Construct a sub-game $U = ((J,X), (T_{i,x})_{(i,x)\in(J,X)}, (u_{i,x})_{(i,x)\in(J,X)})$ where

$$u_{i,x} : (T_{i,x})_{(i,x)\in(J,X)} \to \mathbb{R}, \quad \sigma \mapsto h_{i,x}(\pi_{-(J,X)}, \sigma).$$

So

$$u_{i,x}(\sigma) = \sum_{t \geq 0} \gamma_i^t \mathbb{E}_{x^0=x,\dots,x^u \sim P(x^{u-1}, s^{u-1}), s^u \sim (\pi_{(I\backslash J, X)}, \sigma)(x^u),\dots,s^t \sim (\pi_{(I\backslash J, X)}, \sigma)(x_t)} g_i(x^t, s^t).$$

Construct a stationary mixed strategy stochastic game $V = (J, (S_i)_{i\in J}, X, Q, (v_i)_{i\in I}, (\gamma_i)_{i\in I})$.

$$\forall s \in \prod_{i\in J} S_i, \quad \forall x \in X, \quad \forall y \in X, \quad Q(x,s)(y) = \sum_{w \in \prod_{i\in I\backslash J} S_i} \pi_{(I\backslash J, X)}(x)(w) P(x,s,w)(y).$$

$$\forall i \in J, \quad \forall s \in \prod_{i\in J} S_i, \quad \forall x \in X, \quad v_i(x,s) = \sum_{w \in \prod_{i\in I\backslash J} S_i} \pi_{(I\backslash J, X)}(x)(w) g_i(x,s,w).$$

597 Induction needs to be used here.

598 **Head.** Choose $\sigma \in (T_{i,x})_{(i,x)\in(J,X)}$.

$$\mathbb{E}_{x^0=x,s^0\sim\sigma(x^0)} v_i(x^0, s^0) = \sum_{s^0 \in \prod_{i\in J} S_i} \sigma(x^0)(s^0) v_i(x^0, s^0)$$

$$= \sum_{s^0 \in \prod_{i\in J} S_i} \sigma(x^0)(s^0) \sum_{w \in \prod_{i\in I\backslash J} S_i} \pi_{(I\backslash J, X)}(x^0)(w) g_i(x^0, s^0, w)$$

$$= \mathbb{E}_{x^0=x, r^0 \sim (\pi_{(I\backslash J, X)}, \sigma)(x^0)} g_i(x^0, r^0)$$

599 where $r^0 = (w, s^0)$.

Recursion. Assume for any $x^0 \in X$, there exists

$$\mathbb{E}_{\dots, x^u \sim Q(x^{u-1}, s^{u-1}), s^u \sim \sigma(x^u)} v_i(x^u, s^u) = \mathbb{E}_{\dots, x^u \sim P(x^{u-1}, r^{u-1}), r^u \sim (\pi_{(I\backslash J, X)}, \sigma)(x^u)} g_i(x^u, r^u).$$

Consider the case $u+1$. To simplify notations, let

$$\xi^u(x) = \mathbb{E}_{x^0=x,\dots,x^u \sim Q(x^{u-1}, s^{u-1}), s^u \sim \sigma(x^u)} v_i(x^u, s^u).$$

$$\zeta^u(x) = \mathbb{E}_{x^0=x,\dots,x^u \sim P(x^{u-1}, r^{u-1}), r^u \sim (\pi_{(I\backslash J, X)}, \sigma)(x^u)} g_i(x^u, r^u).$$

600 So

$$\xi^{u+1}(x) = \sum_{s^0 \in \prod_{i\in J} S_i} \sigma(x)(s^0) \sum_{x^1 \in X} Q(x, s^0)(x^1) \xi^u(x^1)$$

$$= \sum_{s^0 \in \prod_{i\in J} S_i} \sigma(x)(s^0) \sum_{x^1 \in X} \xi^u(x^1) \sum_{w \in \prod_{i\in I\backslash J} S_i} \pi_{(I\backslash J, X)}(x)(w) P(x, s^0, w)(x^1)$$

$$= \sum_{s^0 \in \prod_{i\in J} S_i} \sum_{w \in \prod_{i\in I\backslash J} S_i} \sum_{x^1 \in X} \sigma(x)(s^0) \pi_{(I\backslash J, X)}(x)(w) P(x, s^0, w)(x^1) \zeta^u(x^1)$$

$$= \zeta^{u+1}(x).$$

21

601 So the case $u + 1$ is valid no matter what $x^0 = x$ is.

As a result

$$u_{i,x}(\sigma) = \sum_{t \geq 0} \gamma_i^t \mathbb{E}_{x^0=x,\ldots,x^u \sim Q(x^{u-1},s^{u-1}),s^u \sim \sigma(x^u),\ldots,s^t \sim \sigma(x_t)} v_i(x^t, s^t).$$

602 This means that the game U is actually the extended form of the game V.

603 There exists a theorem with respect to stationary mixed strategy stochastic games [13]. Each finite-
604 state Markov game has a mixed equilibrium.

605 As a result, the game V has a mixed equilibrium, which means that the game U also has an
606 equilibrium.

607 To sum up, all conditions of Theorem 2 are satisfied. So in the game H, for any initial strategy π,
608 there exists a finite-length grouped satisficing path $(\sigma^t)_{t=0}^T$ where $\sigma^0 = \pi$ and σ^T is an equilibrium.

609 Since the group set A actually views a player in different state as a group, if a group achieves best
610 response, the corresponding player will achieve best response in any state. Otherwise there exists a
611 state where this player does not achieve best response.

612 As a result, for any initial stationary joint mixed strategy σ in G, there exists a finite-length satisficing
613 path $(\sigma^t)_{t=0}^T$ where $\sigma^0 = \sigma$ and σ^T is a mixed equilibrium.

614 $\square$

615 ## B.4 Proof of Corollary 4

616 **Corollary 4.** *In a k-step mixed strategy stochastic game G, suppose that each S_i is a finite set. Then*
617 *for any initial stationary joint mixed strategy σ, there exists a finite-length satisficing path $(\sigma^t)_{t=0}^T$*
618 *where $\sigma^0 = \sigma$ and σ^T is a mixed equilibrium.*

Proof. Construct a stationary mixed strategy stochastic game H.

$$H = (I, (S_i)_{i \in I}, Y, Q, (h_i)_{i \in I}, (\gamma_i)_{i \in I}).$$

$$Y = X \times \left(\prod_{i \in I} S_i \right)^k.$$

619 Obviously, Y is finite, $|Y| < \infty$.

$$Q : Y \times \prod_{i \in I} S_i \to \Delta Y,$$

$$\left(x, (s^{-t})_{t=1}^k \right) \times s$$

$$\mapsto \left\{ \left(\left(y, (u^{-t})_{t=1}^k \right), P(x,s)(y) I[u^{-1} = s] \prod_{t=2}^k I[u^{-t} = s^{-t+1}] \right) \mid \left(y, (u^{-t})_{t=1}^k \right) \in Y \right\}.$$

where $I[\cdot]$ equals 1 if $[\cdot]$ is true else 0.

$$h_i : Y \times \prod_{i \in I} S_i \to \mathbb{R}, \quad \left(x, (s^{-t})_{t=1}^k \right) \times s \mapsto g_i(x,s).$$

620 So the game H is the extended form of the game G. When the runtime of the game G exceeds k time
621 steps and the first k steps of any path in G are not be considered, each path in G can be project onto a
622 certain and unique path in H, and this projection is actually bijective.

623 According to Corollary 3, for any initial stationary joint mixed strategy σ in H, there exists a finite-
624 length satisficing path $(\sigma^t)_{t=0}^T$ where $\sigma^0 = \sigma$ and σ^T is a mixed equilibrium in H. In turn, this
625 conclusion is also true in G.

626 $\square$

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: Primary findings are established in Theorem 1 and Theorem 2. Four corollaries provide complementary perspectives for applying these main theorems in different contexts.

   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: In Section 5, the theoretical implications and limitations of the main findings are analyzed, and their connections to existing literature are discussed.

   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 proofs of Theorem 1 and Theorem 2 appear in the main text, while those of lemmas and corollaries are deferred to 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 does not contain 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 contain experiments.

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 contain 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: This paper does not contain 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: This paper does not contain 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: This paper strictly adheres to the ethical guidelines established by NeurIPS in all aspects.

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: [NA]

Justification: While this paper establishes a theoretical foundation for MARL algorithm design, its broader societal implications remain beyond the scope of current assessment.

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 does not pose such risks.

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: This 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: This 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: This 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: This 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 paper 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.