Making a Nash Equilibrium Resilient to Coalitions

IVAN GEFFNER, Carnegie Mellon University, USA
MOSHE TENNENHOLTZ, Technion - Israel Institute of Technology, Israel

Nash equilibrium is often heralded as a guiding principle for rational decision-making in strategic interactions. However, it is well-known that Nash equilibrium sometimes fails as a reliable predictor of outcomes, with two of the most notable issues being the fact that it is not resilient to collusion and that there may be multiple Nash equilibria in a single game. In this paper, we show that a game designer can get around these two issues for free by expanding the action sets of the original game. More precisely, given a normal-form or Bayesian game Γ and a Nash equilibrium $\vec{\sigma}$ in Γ, a game designer can construct a new game $\Gamma^{\vec{\sigma}}$ by expanding the action set of each player and defining appropriate utilities in the action profiles that were not already in the original game. We show that the designer can construct $\Gamma^{\vec{\sigma}}$ in such a way that (a) $\vec{\sigma}$ is a semi-strong Nash equilibrium of $\Gamma^{\vec{\sigma}}$, and (b) $\vec{\sigma}$ Pareto-dominates or quasi Pareto-dominates all other Nash equilibria of $\Gamma^{\vec{\sigma}}$.

CCS Concepts: • **Theory of computation** → *Randomness, geometry and discrete structures*; **Solution concepts in game theory**; • **Mathematics of computing** → *Discrete mathematics.*

Additional Key Words and Phrases: Coalitions, Equilibrium Implementation, Game Design

ACM Reference Format:
Ivan Geffner and Moshe Tennenholtz. 2024. Making a Nash Equilibrium Resilient to Coalitions. In *Proceedings of the 25th ACM Conference on Economics and Computation (EC '24), July 8–11, 2024, New Haven, CT, USA.* ACM, New York, NY, USA, 26 pages. https://doi.org/10.1145/123456.654321

1 INTRODUCTION

Suppose that three pirates want to share a treasure consisting of 300 coins. Each pirate can individually choose if it cooperates or defects. If exactly zero, one, or three pirates cooperate, each pirate gets 100 coins. However, if exactly two pirates defect, they can keep the treasure for themselves and get 150 coins each while the remaining pirate gets 0. Clearly, all three pirates cooperating is a Nash equilibrium. However, this equilibrium is not resilient to coalitions since two pirates can get additional utility by defecting together. Now, suppose that these three pirates are in an identical situation with the exception that there is an additional action available for them: *Betray*. Betraying works as follows. If the number of pirates that play either *Defect* or *Betray* is different than two, nothing happens and all three pirates split the treasure equally. Otherwise, we have the following cases depending on what the two non-cooperative pirates play:

- **(Defect, Defect)**: The two defecting pirates get 150 coins each while the remaining pirate gets 0.
- **(Betray, Defect)**: The pirate that betrays kills the pirate that defects and gets all 300 coins.
- **(Betray, Betray)**: Both pirates kill each other. The remaining pirate gets all 300 coins.

The work by Ivan Geffner and Moshe Tennenholtz was supported by funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement 740435).

EC '24, July 8–11, 2024, New Haven, CT, USA

ACM ISBN 123-4-5678-1234-7/23/07
https://doi.org/10.1145/123456.654321

As in the previous game, the strategy profile in which all three pirates cooperate is a Nash equilibrium that is not resilient to collusion. Nevertheless, we can argue that in this scenario it is never in the interest of the pirates to collude. If two pirates decide to play **(Defect, Defect)**, one could easily betray the other in order to get a bigger share. In general, this type of Nash equilibrium is called *semi-strong* [Penn et al., 2005]. In a semi-strong Nash equilibrium, no coalition can increase the utility of all of its members while simultaneously satisfying that no member of the coalition can further increase its utility by defecting (inside of the coalition).

In the game with three pirates described above, we showed that the addition of the *Betray* action converted a Nash equilibrium that was vulnerable to coalitions to a semi-strong one. A natural question that follows is if a game designer can make any Nash equilibrium semi-strong by adding appropriate actions and utilities to the original game. In this paper, we give a positive answer to this question. Moreover, in normal-form games or Bayesian games with independent types, we show that we can add actions in such a way that the original equilibrium is not only semi-strong, but also Pareto-dominates all other Nash equilibria of the extended game. In arbitrary Bayesian games, we show that the original equilibrium quasi Pareto-dominates all other equilibria, which means that the total welfare of any subset S of players with $|S| \geq 2$ is always greater in the original equilibrium. This implies that the game designer include additional structure to the game for free since the newly added actions would never be played in practice.

Perhaps not surprisingly, strengthening a pure Nash equilibrium $\vec{\sigma}$ is relatively simple. We can do so by simply adding a *report* action to each player and defining the new utilities in such a way that players that defect from $\vec{\sigma}$ while being reported get punished (note that if two players play *report*, they both are defecting while being reported), and players who report defecting players get rewarded. The real difficulty of the problem arises when $\vec{\sigma}$ is a mixed strategy, since in this case it is hard to tell if a player i is defecting or if the action played by i is a legitimate outcome from randomization. Additional complexity is added when players have (possibly correlated) private information and when the game is infinite (i.e., when the sets of possible actions are infinite). In fact, these cases require carefully calibrated mechanisms, which serve as the main technical contribution of our paper. At a high level, our approach is to allow players to *bet* or *predict* what actions or action profiles are going to be played by other players (besides playing an action in the original game). If they guess correctly, they get rewarded and the player whose action was predicted gets punished. If they guess wrong, they get punished. We show that we can calibrate the payoffs of such bets in such a way that, if a coalition defects and they all win, there exists a player in the coalition that has incentives to bet on other members of the coalition or there exists a player in the coalition that has incentives to modify its strategy because of the bets placed by other players.

1.1 Equilibrium Selection

An issue that may arise when adding actions to an existing game is the possibility of creating new Nash equilibria, and therefore that players face an equilibrium selection problem. Unfortunately, our construction does not guarantee that the desired equilibrium is unique (see Section 4 for a counterexample), but it does guarantee that it Pareto-dominates or quasi Pareto-dominates all other Nash equilibria of the game. Even though this does not theoretically guarantee that players will converge in the intended equilibrium, there is overwhelming empirical evidence that suggests that rational players act this way [Bardsley et al., 2010, Gold and Colman, 2020]. One possible explanation for this phenomenon is called *team reasoning* [Bacharach, 1999, 2006, Sugden, 1993], which suggests that, when the players' incentives are aligned, they prefer the alternatives that benefit the whole group the most. Another possible explanation is given by evidential decision theory [Gibbard and Harper, 1978]. If each player believes that each other player is going to think the same way she does, it is optimal to select the Pareto-optimal strategy profile.

1.2 Related Work

The work in this paper intersects with three large branches of the economics and computer science literature. The first one is the study of **collusion-resilient protocols** in multi-agent systems. Ben-Or [Ben-Or, 1983] and Rabin [Rabin, 1983] provided protocols that implement consensus in the presence of t byzantine faults if the number of players n satisfies $n > 3t$. These protocols were later generalized by Ben-Or, Goldwasser and Widgerson [Ben-Or et al., 1988] who showed that, if $n > 3t$, we can implement synchronous multiparty secure computation of any function f on a finite domain in the presence of at most t byzantine faults. Many subsequent authors added strategic incentives into the mix [Abraham et al., 2011, 2008, Aiyer et al., 2005, Geffner and Halpern, 2023a, Ranchal-Pedrosa and Gramoli, 2021, Vilaça et al., 2011]. Bernheim, Peleg and Whinston [Bernheim et al., 1987] introduced a recursive definition for *coalition-proof equilibrium*. Intuitively, a strategy profile $\vec{\sigma}$ is coalition-proof if there is no coalition that can defect from the main strategy to another coalition-proof strategy in such a way that all of its members get a better utility. Penn, Polukarov and Tennenholtz [Penn et al., 2005] defined *semi-strong Nash equilibrium* in a similar way, but with only one level of recursion (i.e., for a coalition to succeed, their new strategy must also be a Nash equilibrium). This implies that semi-strong Nash equilibria are also coalition-proof. Abraham, Dolev, Gonen and Halpern [Abraham et al., 2006] adapted Aumann's definition of *k-resilience* [Aumann, 1959] to cheap-talk games, and showed that every synchronous mediator can be implemented by a k-resilient cheap-talk strategy (i.e., a strategy in which coalitions of at most k agents do not benefit from defecting from the proposed strategy) if $n > 2k$ and if there is an appropriate way to punish defecting players. Most of these results have been generalized to the asynchronous setting as well [Ben-Or et al., 1993, Bracha, 1984, Bracha and Toueg, 1985, Geffner and Halpern, 2021, 2023b]. In particular, Abraham, Dolev, Geffner and Halpern [Abraham et al., 2019] provided a k-resilient implementation of asynchronous mediators if $n > 3k$ and there exists a punishment strategy. Although most of the results in this area require that the ratio between the number of players n and the maximum size t or k of the defecting coalition satisfies a certain threshold, our results show that, if we only need semi-strongness, we can take the protocol used for $t = 1$ or $k = 1$ and expand the set of actions in such a way that the same protocol tolerates deviations of coalitions of arbitrary size (according to the semi-strong definition). For example, this means that if a game designer can modify the action sets of the players, we provide a way to design semi-strong protocols that implement mediators in synchronous or asynchronous systems that require only four or five players, respectively.

The second closely related branch is **implementation theory** [Maskin and Sjöström, 2002, Palfrey, 2002]. Here, players do not have concrete utilities but instead have preferences over a fixed set of possible outcomes. These preferences may also depend on the state of the world, which is unknown to the mechanism designer. The mechanism designer is given a social choice rule (SCR), which is a map from states of the world to non-empty subsets of outcomes, and her aim is to design the sets of strategies of the players and to provide a strategy profile that follows a desired notion of equilibrium (e.g., Nash equilibrium or dominant strategies) such that the resulting SCR obtained from playing the given profile is the one provided. The core problems in this area consist of finding the necessary and sufficient conditions for SCRs to be implementable by the mechanism designer (see [Maskin, 1999], for example). This setting is quite similar to the one studied in this paper since, in both cases, there is a *designer* that can manipulate the action sets and there is a target outcome that the designer wants to implement in the resulting game. However, the main difference between these settings is that, in the implementation theory literature, the preferences of the players are fixed and the designer can only provide the structure by which the players strategically convey their information about the state of the world (which is unknown to the mechanism designer). By

contrast, in our setting, players have concrete utilities and the game designer can freely manipulate the relative preferences of the players by including new actions with new utilities into the existing game.

Our work is also related to the study of **mediators** and their influence in multi-agent interactions. Aumann[Aumann, 1974] considered the inclusion of a cheap-talk mediator in normal-form and Bayesian games. More precisely, given a normal-form game Γ, Aumann constructed a game Γ_d where players can communicate via cheap-talk with a mediator d before playing an action in Γ. The Nash equilibria of the resulting games were defined as correlated equilibria. Forges [Forges, 1986] considered the same extension in Bayesian games and gave a geometric characterization of the resulting equilibria, which were labelled as *communication equilibria*. Maskin [Maskin, 1999] introduced the *implementation problem* and showed that a social choice rule can be Nash-implemented if it satisfies monotonicity and no veto power. Monderer and Tennenholtz [Monderer and Tennenholtz, 2009] showed that we can guarantee the existence of strong Nash equilibria in games with several classes of mediators. In their setting, some of these mediators can go beyond sending messages or signals to the players, such as for instance playing on their behalf when permitted. In another paper, Monderer and Tennenholtz [Monderer and Tennenholtz, 2003] considered a setting where the mediator has the power to *promise payments* to the players (i.e., to increase their utility on given outcomes) but cannot communicate. They study the minimum payment that is necessary to influence the behavior of the players. Monderer and Tennenholtz showed that, given any normal-form game Γ and any pure Nash equilibrium $\vec{\sigma}$ in Γ, there exists a mediator d that promises payments only outside of the equilibrium path in such a way that $\vec{\sigma}$ is not only a Nash equilibrium but also a dominant strategy in the newly generated game. Tennenholtz [Tennenholtz, 2004] later considered a type of normal-form games where players are programs and a third-party mediator is able to access and reveal the source code of other players before they commit to an action. Tennenholtz showed that the set of Nash equilibria in this setting (or *program equilibria*, as defined in [Tennenholtz, 2004]) is richer than without the inclusion of the mediator. For example, there exist program equilibria in the one-shot Prisoner's Dilemma where both players cooperate. Oesterheld et al. [Oesterheld et al., 2022] refined this idea and defined *similarity-based coopaeration*, where each player only has access to a similarity parameter instead of the whole source code of other players. They showed that, in their setting, there also exist equilibria in which both players cooperate in the Prisoner's Dilemma. Our setting can also be viewed as including a mediator into a normal-form or Bayesian game in order to influence the outcome. In this case, the mediator cannot communicate with the players but can extend the game by adding new actions with arbitrary utilities.

2 BASIC DEFINITIONS

2.1 Games and Equilibria

A normal-form game is a tuple (P, A, U) where $P = \{1, 2, \ldots, n\}$ is the set of players, $A := A_1 \times A_2 \times \ldots \times A_n$ is the set of possible actions profiles and $U = (u_1, u_2, \ldots, u_n)$ is a tuple of utility functions such that $u_i : A \times \mathbb{R}$ outputs the utility of player i given the action profile played. In a normal form game, a (mixed) strategy σ_i for player i is simply a distribution of actions of A_i (we denote by ΔA_i the set of such possible distributions). A strategy profile $\vec{\sigma} = (\sigma_1, \ldots, \sigma_n)$ is a Nash Equilibrium of Γ if no player can increase its utility by switching to another strategy. More precisely, $\vec{\sigma}$ is a Nash equilibrium if $u_i(\vec{\sigma}) \geq u_i(\vec{\sigma}_{-i}, \tau_i)$ for all players i and all strategies τ_i for i (where $\vec{\sigma}_{-i}$ denotes that all players but i play their part of strategy profile $\vec{\sigma}$).

A Bayesian game is a generalization of normal-form games where players have private types and the utility of each player depends not only on the action profile played but also on its own type.

More formally, a Bayesian game is a tuple (P, T, q, A, U) where $T = T_1 \times T_2 \times \ldots \times T_n$ is the set of possible type profiles and q is the prior distribution over T that is common knowledge among the players. P, A and U denote the same concepts as in normal-form games, except that the domain of each utility function u_i is $T_i \times A$ instead of just A. In Bayesian games, strategies are functions from types to distributions over action profiles. More precisely, in a Bayesian game $\Gamma = (P, T, q, A, U)$, a strategy σ_i for player i maps elements from T_i to distributions over A. As in normal-form games, a strategy profile $\vec{\sigma}$ is a Nash equilibrium if $u_i(\vec{\sigma}) \geq u_i(\vec{\sigma}_{-i}, \tau_i)$ for all players i and all strategies τ_i for i. For simplicity, we will assume that no player has a type that has 0 probability of being sampled.

Given a normal-form or Bayesian game Γ, we say that Γ is *finite* if the set of actions is finite, and we say that Γ is *bounded* if all utility functions are bounded. Note that if a game is finite, it is also bounded.

2.2 Resilience to coalitions

There are well-known notions of equilibria that are resilient to collusion and joint deviations. The first of such notions is *strong Nash equilibrium* [Aumann, 1959], which occurs if no possible coalition can strictly increase the payoff of all its members by jointly deviating. More precisely, in a normal-form or Bayesian game $\Gamma = (P, T, q, A, U)$, a strategy profile $\vec{\sigma}$ is a strong Nash equilibrium if, for all subsets $S \subseteq P$ and all strategy profiles $\vec{\tau}_S$ for players in S, there exists a player $i \in S$ such that $u_i(\vec{\sigma}) \geq u_i(\vec{\sigma}_{-S}, \vec{\tau}_S)$.

In this paper we focus on a variant of strong Nash equilibrium introduced by Penn, Polukarov and Tennenholtz [Penn et al., 2005] in which, for a joint deviation to be successful, not only all members must strictly benefit from colluding, but also the strategy performed by the members of the coalition must be a Nash equilibrium inside of the larger game (in which the players outside of the coalition simply play their part of $\vec{\sigma}$). The following definition makes this precise.

Definition 2.1. Let Γ be a normal-form or a Bayesian game and $\vec{\sigma}$ be a strategy profile for Γ. Then, $\vec{\sigma}$ is a semi-strong Nash equilibrium if, for all subsets S and all strategy profiles $\vec{\tau}_S$ for S, there exists a player $i \in S$ and a strategy τ_i' for i such that

$$u_i(\vec{\sigma}) \geq u_i(\vec{\sigma}_{-S}, \vec{\tau}_S) \quad \text{or} \quad u_i(\vec{\sigma}_{-S}, \vec{\tau}_{S\setminus\{i\}}, \tau_i') > u_i(\vec{\sigma}_{-S}, \vec{\tau}_S)$$

Note that if the first condition is satisfied for some i, then $\vec{\tau}_S$ is not profitable for i, while the second condition implies that i can further deviate from $\vec{\tau}_S$ to increase its utility from the one it would get in the coalition.

2.3 Game extensions

The results of this paper involve games that result from extending the set of actions of simpler games. We say that a Bayesian game $\Gamma' = (P', T', q', A', U')$ *extends* another Bayesian game $\Gamma = (P, T, q, A, U)$ if $P' = P$, $T' = T$, $q' = q$, $A_i \subseteq A_i'$ for all $i \in P$, and $u_i(\vec{a}) = u_i'(\vec{a})$ for all $\vec{a} \in A$ and $i \in P$. Intuitively, game extensions must have the same players, action profiles, and utilities as the original game, but they may also include additional actions for the players that may give arbitrary utilities when played. The definition for normal-form games is analogous.

For future reference, note that if Γ' extends Γ, we can view strategy profiles of Γ as strategy profiles in Γ' in a natural way: each player i plays each action in A_i with the same probability as in Γ, and plays each action in $A_i' \setminus A_i$ with probability 0. Unless it is specified, for the rest of the paper we will always use such convention.

2.4 Comparing Nash Equilibria

Given a Bayesian or Normal-form game Γ and two Nash equilibria $\vec{\sigma}$ and $\vec{\tau}$, we say that $\vec{\sigma}$ *Pareto-dominates* $\vec{\tau}$ if $u_i(\vec{\sigma}) \geq u_i(\vec{\tau})$ for all $i \in P$. Moreover, we say that $\vec{\sigma}$ *quasi Pareto-dominates* $\vec{\tau}$ if, for

all subsets $S \subseteq P$ with $|S| \geq 2$,

$$\sum_{i \in S} u_i(\vec{\sigma}) \geq \sum_{i \in S} u_i(\vec{\tau}).$$

Note that, if $|P| \geq 2$ and $\vec{\sigma}$ quasi Pareto-dominates $\vec{\tau}$,

- At most one player i satisfies $u_i(\vec{\tau}) > u_i(\vec{\sigma})$.
- The total welfare of the players is greater with $\vec{\sigma}$ than with $\vec{\tau}$.

3 MAIN RESULTS

The main contribution of this paper can be summarized by the following Theorems.

Theorem 3.1. *Let $\Gamma = (P, T, q, A, U)$ be a bounded Bayesian game with independent types (i.e., that q samples the types independently). Then, for each Nash equilibrium $\vec{\sigma}$ of Γ there exists an extension $\Gamma^{\vec{\sigma}} = (P, T, q, A^{\vec{\sigma}}, U^{\vec{\sigma}})$ of Γ such that:*

(a) *$\vec{\sigma}$ is a semi-strong Nash equilibrium of $\Gamma^{\vec{\sigma}}$.*

(b) *$\vec{\sigma}$ Pareto-dominates all other Nash equilibria of $\Gamma^{\vec{\sigma}}$.*

Moreover, if A is finite, $A^{\vec{\sigma}}$ satisfies that $|A_i^{\vec{\sigma}}| \leq |A_i| \cdot |P| \cdot |A|_{max}$ for all $i \in P$, where $|A|_{max} = \max_{i \in P} |A_i|$.

Intuitively, part (a) of Theorem 3.1 states that given any game and any Nash equilibrium, we can expand the game such that the original Nash equilibrium becomes semi-strong. At the same time, part (b) implies that none of the newly added actions will ever be played if all players play optimally. The last part of the Theorem states that, if Γ is finite, then the size of the resulting extension is polynomial over the size of Γ. The following Theorem extends Theorem 3.1 to arbitrary Bayesian games.

Theorem 3.2. *Let $\Gamma = (P, T, q, A, U)$ be a bounded Bayesian game. Then, for each Nash equilibrium $\vec{\sigma}$ of Γ there exists an extension $\Gamma^{\vec{\sigma}} = (P, T, q, A^{\vec{\sigma}}, U^{\vec{\sigma}})$ of Γ such that:*

(a) *$\vec{\sigma}$ is a semi-strong Nash equilibrium of $\Gamma^{\vec{\sigma}}$.*

(b) *$\vec{\sigma}$ quasi Pareto-dominates all other Nash equilibria of $\Gamma^{\vec{\sigma}}$.*

Moreover, if A is finite, $A^{\vec{\sigma}}$ satisfies that $|A_i^{\vec{\sigma}}| = |A_i| + |A|$ for all $i \in P$.

The only two differences between Theorem 3.1 and Theorem 3.2 are the size of the extension and the fact that, in Theorem 3.2, the resulting equilibrium quasi Pareto-dominates all other Nash equilibria of $\Gamma^{\vec{\sigma}}$ instead of Pareto-dominating. Note that Theorem 3.1 and Theorem 3.2 do not require that Γ is finite. In fact, as shown in Section 7, the statements still hold when the action sets are uncountable, as long as Γ is bounded.

4 SKETCH OF THE PROOFS

As in the *three pirates* example from the introduction, the idea behind the construction of $\Gamma^{\sigma} = (P, A^{\vec{\sigma}}, U^{\vec{\sigma}})$ is that players can play special actions that reward them if another player goes off the equilibrium path. This construction can be very simple for pure Nash equilibria: suppose that $\vec{\sigma}$ is a pure Nash equilibrium where each player i plays some action a_i with probability 1. Then, we can set $A_i^{\vec{\sigma}} = A_i \cup \{\alpha\}$ and define the payoffs as follows when some player i plays α:

- If another player j plays α, all players k that didn't play a_k get $m - 1$ utility (including those who played α), where m is the minimum utility of the game. All other players get the same utility as in equilibrium.
- If no other player plays α,
  - If all other players j play a_j, they all get the same utility as in equilibrium and i gets $m - 1$.

– Otherwise, i gets the maximum possible utility plus 1, all players j that played a_j get the same utility as in equilibrium, and all other players get $m - 1$.

It is straightforward to check that $\Gamma^{\vec{\sigma}}$ satisfies the conditions of Theorem 3.1: First, no player can increase its utility by defecting from $\vec{\sigma}$. Second, if a coalition defects from $\vec{\sigma}$ and they all increase their utility, it means that none of them is playing α. However, this implies that a player can further increase its utility by switching its action to α. Lastly, an analogous reasoning shows that $\vec{\sigma}$ is Pareto-optimal: if another strategy profile $\vec{\tau}$ gives a better utility than $\vec{\sigma}$ to all players, it means that none of them is playing α. Then, again, this would not be a Nash equilibrium since a player can increase its utility by playing α instead.

If we consider mixed Nash equilibria, the construction becomes more complicated. For instance, there could be Nash equilibria where all actions have non-zero chance of being played, as in the following example.

Example 4.1. Consider a normal-form game $\Gamma = (P, A, U)$ with $P = \{1, 2, \ldots, n\}$, $A_i = \{0, 1\}$ for all $i \in P$, and such that

$$u_i(a_1, \ldots, a_n) = \begin{cases} 1 & \text{if } \#a_i(a_1, \ldots, a_n) \geq \#(1 - a_i)(a_1, \ldots, a_n) \\ 0 & \text{otherwise,} \end{cases}$$

where $\#a_i(a_1, \ldots, a_n)$ denotes the number of players who played action a_i. Denote by $\vec{\sigma}$ the strategy profile in which each player chooses its action uniformly at random.

Intuitively, i gets 1 utility if its action is played by the majority of the players and i gets 0 utility otherwise. Clearly, if all other players randomize their action, it doesn't matter what action i chooses since both give the same utility in expectation. Therefore, randomizing its own action is also a best response for i, which means that $\vec{\sigma}$ is a Nash equilibrium. However, $\vec{\sigma}$ is not semi-strong: if two or more players collude, it is in their interest to vote on the same value. For instance, all of them playing 0 or all of them playing 1 strictly increases their expected utility and no agent in the coalition can further increase it by playing anything else.

In order to create an extension of Γ such that $\vec{\sigma}$ is a semi-strong Nash equilibrium we need a different approach than the one given for pure Nash equilibria. The main issue is that it is not easy to tell if a player i is defecting from the proposed strategy or not since, regardless of what i plays, its action could be the result of playing σ_i. They key idea is that, in equilibrium, i must play 0 and 1 with $\frac{1}{2}$ probability each. By contrast, if i defects, one of these actions must be played with probability greater than $\frac{1}{2}$. We can make use of this fact as follows. Suppose that, for each player i and each action $a_i \in \{0, 1\}$, we give each player j the possibility of *betting* that i will play a_i. If i plays a_i, then j gets x utility and i loses x' utility. Otherwise, j gets $-x$ utility. Clearly, if i follows σ_i, there is no incentive for j to place such a bet. However, if i defects, betting on one of i's actions gives j a strictly positive expected utility. We can adjust the values of x and x' such that it is never in the interest of i to collude and defect. The construction given in Section 5.1 follows a similar approach, except for the fact that the utility won by winning a bet and the utility lost by losing a bet may be different. More precisely, the bets are calibrated in such a way that, if i plays some action a with the same probability as in σ_i, the expected payoff of betting on (i, a) is 0, while if i plays a with more probability than in σ_i the expected payoff of betting on (i, a) is strictly positive.

The construction of $\Gamma^{\vec{\sigma}}$ for Bayesian games is more complicated. The main issue is that, in normal-form games, whenever a player i defects from σ_i to a different strategy τ_i, it is guaranteed that there exists an action that is more likely to be played with τ_i than with σ_i. However, this is not true for Bayesian games. For example, consider a game such that $T_i = A_i = \{0, 1\}$ for all players i, and such that all types are sampled independently and uniformly from $\{0, 1\}$. If σ_i is the strategy in which player i plays whatever its type is and τ_i is the strategy in which it plays the opposite of its

type, it is indistinguishable to the other players if i is playing σ_i or if i is playing τ_i. In particular, we cannot design bets that give a strictly positive expected utility when i plays τ_i and simultaneously give 0 expected utility when i plays σ_i. The key observation is that, if $\vec{\tau}_S$ is a profitable deviation for some coalition S and types are independent, each player $i \in S$ must be able to tell the strategy τ_j of at least one other player $j \in S$. Otherwise, i would get the same utility in $(\vec{\sigma}_{-i}, \tau_i)$. This observation implies that the construction used for normal-form games can be easily generalized to also cover Bayesian games with independent types.

Unfortunately, the construction used for normal-form games does not quite work in general for Bayesian games, as the following example shows.

Example 4.2. Consider a Bayesian game $\Gamma = (P, T, q, A, U)$ such that $P = \{1, 2, 3\}$, $T_i = A_i = \{0, 1\}$ for all $i \in P$, and q is the uniform distribution over the subset of T such that the sum of coordinates is odd (i.e., $\{(t_1, t_2, t_3) \in T \mid t_1 + t_2 + t_3 \equiv 1 \bmod 2\}$). Given an action profile (a_1, a_2, a_3), all players get 1 utility if $a_1 + a_2 + a_3$ is odd. Otherwise, they all get 0 utility.

Consider two strategy profiles $\vec{\sigma}$ and $\vec{\tau}$. In $\vec{\sigma}$, each player independently plays 0 and 1 with equal probability. In $\vec{\tau}$, each player plays its own type.

In Example 4.2, it is straightforward to check that $\vec{\sigma}$ and $\vec{\tau}$ are both Nash equilibria where each player plays 0 and 1 with equal probability. However, players get $1/2$ expected utility with $\vec{\sigma}$ and 1 expected utility with $\vec{\tau}$. The main difference between the strategy profiles is that, even though each player has $1/2$ chance of having type 0 and $1/2$ chance of having type 1, their types are correlated in such a way that their sum is always an odd number. Moreover, we can easily check that players' types are **pairwise independent**. This means that the probability that a player i plays a certain action $a \in \{0, 1\}$ with τ_i is exactly the same as the probability that i plays a with σ_i, even if we condition the probability on another player j's type. Therefore, j cannot guess i's action with $\vec{\tau}$ any better than with $\vec{\sigma}$, which implies that the construction used for normal-form games fails in this case.

In order to deal with situations as the one presented in Example 4.2, we modify the construction $\Gamma^{\vec{\sigma}}$ used for normal form games in such a way that players can bet on action profiles played by other players instead of individual actions. It can be shown that if bets are calibrated in a similar fashion, $\Gamma^{\vec{\sigma}}$ satisfies the properties of Theorem 3.2.

To extend the proofs of Theorems 3.1 and 3.2 to infinite games, we need a slightly different construction, as shown by the following example.

Example 4.3. Consider a normal-form game $\Gamma = (P, A, U)$ with $P = \{1, 2\}$, $A_1 = A_2 = \mathbb{N}$, and

$$u_1(a_1, a_2) = u_2(a_1, a_2) = \begin{cases} 0 & \text{if } a_1 = 0 \text{ or } a_2 = 0 \\ 1 & \text{otherwise.} \end{cases}$$

Denote by $\vec{\sigma}$ the strategy profile in which each player plays 0 with probability 1, and by $(\vec{\tau})_N$ the strategy profile in which both players sample their action uniformly at random from $\{1, 2, \ldots, N\}$.

It is easy to check that $\vec{\sigma}$ and $(\vec{\tau})_N$ are Nash equilibria of Γ. Suppose that we construct $\Gamma^{\vec{\sigma}}$ in the same way as in normal-form games and have each player lose C utility whenever the other player guesses its action correctly. By playing $\vec{\sigma}$, each player gets 0 expected utility. However, if they play $(\vec{\tau})_N$, each player has at most $1/N$ chance of guessing the other player's action correctly. Therefore, regardless of how they place their bets, each player can get at least $1 - \frac{C}{N}$ expected utility in $\Gamma^{\vec{\sigma}}$. If N is large enough, this value is larger than the utility that they would get with $\vec{\sigma}$.

The main issue presented in Example 4.3 is that players can spread their actions really thin, avoiding this way being caught by other players. Even though this doesn't always lead to a better equilibrium (e.g., it may happen that no matter how player i spreads her actions, it is always better

for i to spread them even thinner), this situation is clearly non-desirable. In order to avoid it, the construction in Section 7 allows players to bet on arbitrary sets of actions (or arbitrary sets of action profiles when dealing with Bayesian games) instead of just individual actions. Again, these bets are calibrated in such a way that no player benefits from betting if all other players play their part of $\vec{\sigma}$. However, if a player i defects from $\vec{\sigma}$, betting on the set of actions that i is more likely to play after defecting always gives a strictly positive payoff. It can be shown that the rewards and punishments from betting can be calibrated in such a way that it is always better to follow the proposed equilibrium $\vec{\sigma}$.

In general, one may wonder if the equilibrium obtained with these constructions not only Pareto or quasi-Pareto dominates all other equilibria of $\Gamma^{\vec{\sigma}}$, but is also the only equilibrium. Unfortunately, this is not always the case. For example, consider an instance of Example 4.1 for three players where all players choose their action at random, but additionally player 1 bets that player 3 will play 0 and player 2 bets that player 3 will play 1. It is easy to check that player 3 gets the same (negative) utility by playing 0 or 1, and hence randomizing its own action is also a best response. Therefore, this strategy profile is also a Nash equilibrium of $\Gamma^{\vec{\sigma}}$ where players 1 and 2 get exactly the same expected utility as in $\vec{\sigma}$ and player 3 gets strictly less utility.

5 PROOF OF THEOREM 3.1 FOR FINITE GAMES

In this section we prove Theorem 3.1 for the case in which Γ is finite. We extend this proof for infinite games in Section 7. We will also restrict the proof to finite normal-form games since the proof for finite Bayesian games with independent types is analogous (we simply have the utilities in the construction to depend on the player's type as well). The full proof can be found in Appendix A or in the full version of the paper (see [Geffner and Tennenholtz, 2023]).

5.1 Construction of Γ^{σ}

In this section we give the precise construction of $\Gamma^{\vec{\sigma}}$. Define

$$A_i^{\vec{\sigma}} := A_i \cup \left(A_i \times \bigcup_{j \in P} \left(\{j\} \times A_j \right) \right).$$

Elements of $A_i^{\vec{\sigma}}$ are either actions of A_i or a tuple (a_i, j, a_j) that indicates that i is playing action a_i but is also betting that player j is playing action a_j. Suppose that players play actions $(a_1, b_1), (a_2, b_2), \ldots, (a_n, b_n)$ where, for convenience, we use the notation $b_i := \perp$ to denote that player i placed no bet (i.e., if i played an action in A_i). The utility $u_i^{\vec{\sigma}}$ of player i is given by

$$\begin{aligned} u_i^{\vec{\sigma}}((a_1, b_1), (a_2, b_2), \ldots, (a_n, b_n)) \quad &:= \quad u_i(a_1, a_2, \ldots, a_n) \\ &+ \quad w_i^{\vec{\sigma}}((a_1, b_1), (a_2, b_2), \ldots, (a_n, b_n)) \\ &+ \quad \ell_i^{\vec{\sigma}}((a_1, b_1), (a_2, b_2), \ldots, (a_n, b_n)), \end{aligned}$$

where w_i is the additional utility that i gets by placing a bet on another player's action, and ℓ_i is the additional (negative) utility that i gets if another player bets on i's action and wins. If $p_{(j,a_j)}^{\vec{\sigma}}$ is the probability that strategy σ_j assigns to action a_j, then

$$w_i^{\vec{\sigma}}((a_1, b_1), (a_2, b_2), \ldots, (a_n, b_n)) := \begin{cases} 0 & \text{if } b_i = \perp \\ 1 - p_{(j,a_j)}^{\vec{\sigma}} & \text{if } b_i \text{ is of the form } (j, a_j) \\ -p_{(j,a_j)}^{\vec{\sigma}} & \text{if } b_i \text{ is of the form } (j, a) \text{ with } a \neq a_j. \end{cases}$$

It is important to note that, if j does not defect and i bets on a_j, then its expected additional utility is given by $p_{(j,a_j)}^{\vec{\sigma}} \left(1 - p_{(j,a_j)}^{\vec{\sigma}}\right) + \left(1 - p_{(j,a_j)}^{\vec{\sigma}}\right) \left(-p_{(j,a_j)}^{\vec{\sigma}}\right) = 0$.

Finally, we define $\ell_i^{\vec{\sigma}} := \sum_{j \in P} \ell_{(i,j)}^{\vec{\sigma}}$, where

$$\ell_{(i,j)}^{\vec{\sigma}}((a_1, b_1), (a_2, b_2), \ldots, (a_n, b_n)) := \begin{cases} 0 & \text{if } b_j = \perp \text{ or } b_j \text{ is of the form } (i, a) \text{ with } a \neq a_i \\ -C & \text{otherwise.} \end{cases}$$

The constant C used in the previous definition is simply a large value that serves as a deterrent for players to defect from $\vec{\sigma}$: if a player i follows a different strategy τ_i, then there exists some action $a \in A_i$ such that a is played with strictly more probability in τ_i than in σ_i. If this is the case, because of the way in which we defined w_j, it is in the interest of other players j to bet on (i, a_i), which decreases i's overall utility. In Section 5.2, we show that we can choose a value for C so that conditions (a) and (b) of Theorem 3.1 are satisfied.

5.2 Proof of Correctness

5.2.1 Part (a). We start by showing that $\vec{\sigma}$ is a semi-strong Nash equilibrium if $\Gamma^{\vec{\sigma}}$. First, as observed in the previous section, it is easy to check that our construction of $\Gamma^{\vec{\sigma}}$ satisfies the following property:

Lemma 5.1. *Given a strategy profile $\vec{\tau}$ for $\Gamma^{\vec{\sigma}}$, a player $j \in P$, and an action $a_j \in A_j$, let $p_{(j,a_j)}^{\vec{\tau}}$ denote the probability that j plays an action in $A_j^{\vec{\sigma}}$ such that its first component is a_j. Let $x_{(j,a_j)}^{\vec{\tau}}$ denote the expected additional utility of any other player when betting that j plays action a_j. Then*

$$x_{(j,a_j)}^{\vec{\tau}} = p_{(j,a_j)}^{\vec{\tau}} - p_{(j,a_j)}^{\vec{\sigma}}$$

Proof. The expected additional utility can be computed as follows:

$$x_{(j,a_j)}^{\vec{\tau}} = p_{(j,a_j)}^{\vec{\tau}} \left(1 - p_{(j,a_j)}^{\vec{\sigma}}\right) + \left(1 - p_{(j,a_j)}^{\vec{\tau}}\right) \left(-p_{(j,a_j)}^{\vec{\sigma}}\right) = p_{(j,a_j)}^{\vec{\tau}} - p_{(j,a_j)}^{\vec{\sigma}}.$$

□

With this, it is straightforward to check that, if $C \geq 1$, $\vec{\sigma}$ is a (standard) Nash equilibrium of $\Gamma^{\vec{\sigma}}$: Fix a player i and assume all other players play $\vec{\sigma}_{-i}$. Since $\vec{\sigma}$ is a Nash equilibrium of Γ, regardless of the bets i places, it is a best response for i to distribute the first component of its action in the same way as σ_i. It just remains to check that i gets no utility from placing bets. By Lemma 5.1, i gets no utility from betting on other players. Moreover, if $C \geq 1$, i gets negative expected utility by betting on itself. This shows that σ_i is indeed a best response for i.

Now suppose that $\vec{\sigma}$ does not satisfy the conditions of Definition 2.1. This would imply that there exists a coalition $S \subseteq P$ and a strategy profile $\vec{\tau}_S$ for players in S such that (a1) all players $i \in S$ get a strictly better expected utility than $u_i(\vec{\sigma})$, and (b1) no player $i \in S$ is better off defecting from $\vec{\tau}_S$, assuming that the rest of the players play $(\vec{\sigma}_{-S}, \vec{\tau}_{S \setminus \{i\}})$. Given a strategy τ_i for some player $i \in P$ in $\Gamma^{\vec{\sigma}}$, denote by $\tilde{\tau}_j$ the strategy for i in Γ obtained by projecting the actions of $A_j^{\vec{\sigma}}$ onto their first component (i.e., by getting rid of the bets). Then, we have the following:

Proposition 1. *If S and $\vec{\tau}_S$ satisfy (a1), there exist at least one player $i \in S$ such that $\tilde{\tau}_i \neq \sigma_i$.*

Proof. If all players $i \in S$ satisfy $\tilde{\tau}_i = \sigma_i$. Then, by Lemma 5.1, none of the players in S can increase their utility by placing a bet. This would imply that their expected utilities are at most the ones they would get when playing $\vec{\sigma}_S$, which contradicts the (a1) assumption.

□

Proposition 1 implies that if a coalition S with strategy $\vec{\tau}_S$ satisfies (a1), there exist at least one player $i \in S$ and an action $a_i \in A_i$ such that i is more likely to play a_i in $\vec{\tau}_S$ than in $\vec{\sigma}_S$. By Lemma 5.1, betting that i is going to play a_i gives a strictly positive utility to all other players. In particular, if S and $\vec{\tau}_S$ satisfy (b1), all other players in S must place bets on i or other players in S with probability 1

since there exist at least one bet with strictly positive expected utility. Since there is a finite number of possible bets, we have the following result.

Proposition 2. *Let $q^{(j,\tau_j)}_{(i,a_i)}$ be the probability that player j bets on (i,a_i) when playing τ_j. Then, if S and $\vec{\tau}_S$ satisfy (a1) and (b1), there exists a player $i \in S$ and an action $a_i \in A_i$ such that*

$$\sum_{j \in S} q^{(j,\tau_j)}_{(i,a_i)} \geq \frac{1}{|A|_{max}},$$

where $|A|_{max} = \max_i\{|A_i|\}$.

Proof. By proposition 1, there exists a player $i \in S$ such that $\tilde{\tau}_i \neq \sigma_i$. Because of (b1), if i is the only player in S such that $\tilde{\tau}_i \neq \sigma_i$, all other players in S must bet on i with probability 1. Since $|S| \geq 2$, it follows that there exist an action $a \in A_i$ such that at least $\frac{1}{|A_i|}$ players bet on (i,a_i) in expectation.

If there are at least two players $i \in S$ such that $\tilde{\tau}_i \neq \sigma_i$ and S and $\vec{\tau}_S$ satisfy (a1) and (b1), then all players in S must place bets on other players in S with probability 1. This means that, for a fixed $j \in S$,

$$\sum_{i \in S,\ a_i \in A_i} q^{(j,\tau_j)}_{(i,a_i)} = 1.$$

If we add the equation above for all $j \in S$ we get that

$$\sum_{i,j \in S,\ a_i \in A_i} q^{(j,\tau_j)}_{(i,a_i)} = |S|.$$

Rearranging the terms of the previous sum, we get

$$\sum_{i \in S,\ a_i \in A_i} \left(\sum_{j \in S} q^{(j,\tau_j)}_{(i,a_i)} \right) = |S|.$$

Since the outer sum has at most $|S| \cdot |A|_{max}$ elements, there exists $i \in S$ and $a_i \in A_i$ such that

$$\sum_{j \in S} q^{(j,\tau_j)}_{(i,a_i)} \geq \frac{|S|}{|S| \cdot |A|_{max}} = \frac{1}{|A|_{max}},$$

as desired. □

Proposition 2 shows that, if S and $\vec{\tau}_S$ satisfy (a1) and (b1), there exists a player $i \in P$ and an action $a_i \in A_i$ such that at least a constant fraction of the players are betting on it in expectation. We next show that, if $C > \max\{M \cdot |A|_{max}, 1\}$, i can strictly increase its utility by decreasing the probability of playing a_i. This would contradict the fact that S and $\vec{\tau}_S$ satisfy (b1) and prove Theorem 3.1.

By Proposition 1, the probability $p^{\vec{\tau}}_{(i,a_i)}$ that i plays action a_i with strategy $\tilde{\tau}_i$ must satisfy $p^{\vec{\tau}}_{(i,a_i)} > p^{\vec{\sigma}}_{(i,a_i)}$ since, otherwise, players could get a better utility by betting somewhere else. The existence of a_i implies that there exists another action $a'_i \in A_i$ such that $p^{\vec{\tau}}_{(i,a'_i)} < p^{\vec{\sigma}}_{(i,a'_i)}$ (i.e., an action that is played with less probability in $\tilde{\tau}_i$ than in σ_i). Since S and $\vec{\tau}_S$ satisfy (b1), it is guaranteed that no player will ever bet on (i,a'_i) since, by Lemma 5.1, betting on it gives a strictly negative expected utility. Consider a strategy τ'_i in which i plays in the same way as in τ_i, except that i plays a'_i whenever it would play a_i in τ_i.

Then, we have that

$$\begin{aligned} u^{\vec{\sigma}}_i(\vec{\sigma}_{-S}, \vec{\tau}_{S\setminus\{i\}}, \tau'_i) - u^{\vec{\sigma}}_i(\vec{\sigma}_{-S}, \vec{\tau}_S) &= u_i(\vec{\sigma}_{-S}, \tilde{\tau}_{S\setminus\{i\}}, \tilde{\tau}'_i) - u_i(\vec{\sigma}_{-S}, \tilde{\tau}_S) \\ &+ w^{\vec{\sigma}}_i(\vec{\sigma}_{-S}, \vec{\tau}_{S\setminus\{i\}}, \tau'_i) - w^{\vec{\sigma}}_i(\vec{\sigma}_{-S}, \vec{\tau}_S) \\ &+ \ell^{\vec{\sigma}}_i(\vec{\sigma}_{-S}, \vec{\tau}_{S\setminus\{i\}}, \tau'_i) - \ell^{\vec{\sigma}}_i(\vec{\sigma}_{-S}, \vec{\tau}_S). \end{aligned}$$

Let $p^{\tilde{\tau}_i}_{(i,a_i)}$ be the probability that i plays action a_i when following strategy $\tilde{\tau}_i$. Then, we can bound the first term in the sum by

$$u_i(\vec{\sigma}_{-S}, \tilde{\tau}_{S\setminus\{i\}}, \tilde{\tau}'_i) - u_i(\vec{\sigma}_{-S}, \tilde{\tau}_S) \geq -p^{\tilde{\tau}_i}_{(i,a_i)} M$$

since the difference in utility by switching from a_i to a'_i can be at most $-M$. Since τ'_i bets exactly in the same way as τ_i does, we have that $w^{\vec{\sigma}}_i(\vec{\sigma}_{-S}, \vec{\tau}_{S\setminus\{i\}}, \tau'_i) - w^{\vec{\sigma}}_i(\vec{\sigma}_{-S}, \vec{\tau}_S) = 0$. Finally, since the probability that another player bets on (i, a_i) is at least $\frac{1}{|A_i|}$, the following inequality holds:

$$\ell^{\vec{\sigma}}_i(\vec{\sigma}_{-S}, \vec{\tau}_{S\setminus\{i\}}, \tau'_i) - \ell^{\vec{\sigma}}_i(\vec{\sigma}_{-S}, \vec{\tau}_S) \geq p^{\tilde{\tau}_i}_{(i,a_i)} \frac{C}{|A_i|}.$$

In the equation above, we are using the fact that the only difference between τ_i and τ'_i is that i is switching from a_i to another action a'_i that no player ever bets on. If $C > M \cdot |A|_{max}$, then

$$u^{\vec{\sigma}}_i(\vec{\sigma}_{-S}, \vec{\tau}_{S\setminus\{i\}}, \tau'_i) - u^{\vec{\sigma}}_i(\vec{\sigma}_{-S}, \vec{\tau}_S) \geq p^{\tilde{\tau}_i}_{(i,a_i)} \frac{C}{|A_i|} - p^{\tilde{\tau}_i}_{(i,a_i)} M > 0.$$

This would imply that i can increase its utility by switching from τ_i to τ'_i, which contradicts our original assumption that S and $\vec{\tau}_S$ satisfy (b1).

5.2.2 Part (b). The proof of part (b) follows from the following proposition.

Proposition 3. *If $\vec{\tau}$ is a Nash equilibrium of Γ^σ, then $\tilde{\tau}_i = \sigma_i$ for all $i \in P$.*

Intuitively, this proposition says that all Nash equilibria of $\Gamma^{\vec{\sigma}}$ are essentially $\vec{\sigma}$ with some bets on top of it. Note that if $\tilde{\tau}_i = \sigma_i$ for all $i \in P$, w_i gives exactly 0 utility for all strategies of i. Since ℓ_i is a negative function, Proposition 3 implies that all Nash equilibria of $\Gamma^{\vec{\sigma}}$ are Pareto-dominated by $\vec{\sigma}$, as desired.

Proof of Proposition 3. Suppose that there exists a player i such that $\tilde{\tau}_i \neq \sigma_i$. An analogous reasoning to the one used for proving Proposition 2 shows that there exists a player $i \in P$ and an action $a_i \in A_i$ such that

$$\sum_{j \in S} q^{\tau_j}_{(i,a_i)} \geq \frac{1}{|A|_{max}}.$$

Then, analogously to the proof of part (a), it can be shown that i can strictly increase its utility by switching from a_i to another action a'_i, contradicting the fact that $\vec{\tau}$ is a Nash equilibrium of Γ^σ. □

6 PROOF OF THEOREM 3.2 FOR FINITE BAYESIAN GAMES

The analysis of Example 4.2 given in Section 4 shows that the construction used for normal-form games and for Bayesian games with independent types does not quite work in general for Bayesian games. However, note that if all three players play $\vec{\tau}$ in Example 4.2, even though player 1 cannot tell if player 2 or player 3 are going to play 0 or 1, she knows the following. If her type is 0, then players 2 and 3 are going to play either $(0, 1)$ or $(1, 0)$. Otherwise, they are going to play $(0, 0)$ or $(1, 1)$. This means that, even if i cannot guess the action of another player j with higher probability in τ_j than in σ_j, she can guess the action profile of all other players with higher probability. This suggests that, instead of allowing players to bet on which action another player will play, they should be able to bet on action profiles of other players. This motivates the following construction. Define $\Gamma^{\vec{\sigma}} := (P, T, q, A^{\vec{\sigma}}, U^{\vec{\sigma}})$. We define the action set $A^{\vec{\sigma}}_i$ of each player as $A^{\vec{\sigma}}_i := A_i \cup A$. As in Section 5.1, we can view each element of $A^{\vec{\sigma}}_i$ with a pair (a, b), where a is the actual action that i plays and $b \in A_{-i}$ is i's bet over the action profile of other players. Again, for convenience, whenever i plays an action a_i without placing a bet, we will use the pair $(a_i, \perp)$ instead of a_i. The utility of each player i is defined as follows:

$$\begin{aligned} u_i^{\vec{\sigma}}((a_1,b_1),(a_2,b_2),\ldots,(a_n,b_n)) &:= u_i(a_1,a_2,\ldots,a_n) \\ &+ w_i^{\vec{\sigma}}((a_1,b_1),(a_2,b_2),\ldots,(a_n,b_n)) \\ &+ \ell_i^{\vec{\sigma}}((a_1,b_1),(a_2,b_2),\ldots,(a_n,b_n)), \end{aligned}$$

where $w_i^{\vec{\sigma}}$ and $\ell_i^{\vec{\sigma}}$ are defined similarly to their counterparts in Section 5.1. In this case, let $p^{\vec{\sigma}}_{i,t_i,\vec{a}_{-i}}$ denote the probability that the action profile played by P_{-i} is $\vec{a}_{-i}$, conditioned to the fact that all players play $\vec{\sigma}$ and that player i has type t_i. Then, we can define $w_i^{\vec{\sigma}}$ as follows:

$$w_i^{\vec{\sigma}}((a_1,b_1),(a_2,b_2),\ldots,(a_n,b_n)) := \begin{cases} 0 & \text{if } b_i = \bot \\ 1 - p^{\vec{\sigma}}_{(i,t_i,\vec{a}_{-i})} & \text{if } b_i \text{ is of the form } \vec{a}_{-i} \\ -p^{\vec{\sigma}}_{(i,t_i,\vec{a}_{-i})} & \text{if } b_i \text{ is of the form } \vec{a}'_i \text{ with } \vec{a}'_i \neq \vec{a}_{-i}. \end{cases}$$

Finally, we define $\ell_i^{\vec{\sigma}} := \sum_{j\neq i} \ell^{\vec{\sigma}}_{(i,j)}$, where

$$\ell^{\vec{\sigma}}_{(i,j)}((a_1,b_1),(a_2,b_2),\ldots,(a_n,b_n)) := \begin{cases} 0 & \text{if } b_j = \bot \text{ or } b_j \text{ is of the form } \vec{a}'_{-j} \text{ with } \vec{a}'_{-j} \neq \vec{a}_{-j} \\ -C & \text{otherwise,} \end{cases}$$

and $C = 2|A| \cdot (M+1)$. Note that, in this case, $\ell_{(i,j)}$ does not depend on i.

To prove that this construction satisfies the properties of Theorem 3.1 we need a few preliminary lemmas. Before stating the first of these, we need the following definition. Given a player $i \in P$, a type $t_i \in T_i$, and a strategy profile $\vec{\tau}$, define $\Delta^{\vec{\tau}}_{(i,t_i)}$ as

$$\Delta^{\vec{\tau}}_{(i,t_i)} = \sum_{\vec{a}_{-i}\in A_{-i}} \left| p^{\vec{\sigma}}_{(i,t_i,\vec{a}_{-i})} - p^{\vec{\tau}}_{(i,t_i,\vec{a}_{-i})} \right|.$$

Intuitively, $\Delta^{\vec{\tau}}_{(i,t_i)}$ can be seen as the "Manhattan distance" between the distributions over actions played by P_{-i} that are generated when players play $\vec{\sigma}$ and $\vec{\tau}$ respectively, conditioned to the fact that i has type t_i. Our first lemma states that, if players play $\vec{\tau}$, the expected utility of a player i with type t_i is bounded by its expected utility when players play $\vec{\sigma}$ plus $\Delta^{\vec{\tau}}_{(i,t_i)} \cdot M$.

Lemma 6.1. *Given a player $i \in P$, its type $t_i \in T_i$, and a strategy profile $\vec{\tau}$ for Γ, denote by $u_i(t_i,\vec{\tau})$ the expected utility of i when players play $\vec{\tau}$, conditioned to the fact that i has type t_i. Then,*

$$u_i(t_i,\vec{\tau}) \leq u_i(t_i,\vec{\sigma}) + \Delta^{\vec{\tau}}_{(i,t_i)} \cdot M.$$

Proof. Player i's expected utility when players play $\vec{\tau}$ conditioned to the fact that i has type t_i is given by

$$u_i(t_i,\vec{\tau}) = \sum_{a_i\in A_i} \tau_i(t_i,a_i) \left(\sum_{a_{-i}\in A_{-i}} p^{\vec{\tau}}_{(i,t_i,\vec{a}_{-i})} u_i(t_i,\vec{a}) \right),$$

where $\tau_i(t_i,a_i)$ denotes the probability that i plays a_i when following strategy τ_i, conditioned to the fact that it has type t_i.

We can rewrite the sum as

$$u_i(t_i,\vec{\tau}) = \sum_{a_i\in A_i} \tau_i(t_i,a_i) \left(\sum_{a_{-i}\in A_{-i}} p^{\vec{\sigma}}_{(i,t_i,\vec{a}_{-i})} u_i(t_i,\vec{a}) + (p^{\vec{\tau}}_{(i,t_i,\vec{a}_{-i})} - p^{\vec{\sigma}}_{(i,t_i,\vec{a}_{-i})}) u_i(t_i,\vec{a}) \right).$$

Note that, since $\vec{\sigma}$ is a Nash equilibrium, we have that

$$\sum_{a_i\in A_i} \tau_i(t_i,a_i) \left(\sum_{a_{-i}\in A_{-i}} p^{\vec{\sigma}}_{(i,t_i,\vec{a}_{-i})} u_i(t_i,\vec{a}) \right) \leq u_i(t_i,\vec{\sigma}).$$

It remains to show that

$$\sum_{a_i \in A_i} \tau_i(t_i, a_i) \left(\sum_{a_{-i} \in A_{-i}} (p^{\vec{\tau}}_{(i,t_i,\vec{a}_{-i})} - p^{\vec{\sigma}}_{(i,t_i,\vec{a}_{-i})}) u_i(t_i, \vec{a}) \right) \leq \Delta^{\vec{\tau}}_{(i,t_i)} \cdot M.$$

Let M_{min} be the minimum utility of any player in Γ and M_{max} be the maximum utility of any player in Γ (recall that $M = M_{max} - M_{min}$). Then,

$$\sum_{a_{-i} \in A_{-i}} (p^{\vec{\tau}}_{(i,t_i,\vec{a}_{-i})} - p^{\vec{\sigma}}_{(i,t_i,\vec{a}_{-i})}) u_i(t_i, \vec{a}) = \sum_{a_{-i} \in A_{-i}} (p^{\vec{\tau}}_{(i,t_i,\vec{a}_{-i})} - p^{\vec{\sigma}}_{(i,t_i,\vec{a}_{-i})})(M_{min} + (u_i(t_i, \vec{a}) - M_{min})).$$

Since $\sum_{a_{-i} \in A_{-i}} p^{\vec{\tau}}_{(i,t_i,\vec{a}_{-i})} = \sum_{a_{-i} \in A_{-i}} p^{\vec{\sigma}}_{(i,t_i,\vec{a}_{-i})} = 1$, we have that

$$\begin{aligned}
& \textstyle\sum_{a_{-i} \in A_{-i}} (p^{\vec{\tau}}_{(i,t_i,\vec{a}_{-i})} - p^{\vec{\sigma}}_{(i,t_i,\vec{a}_{-i})})(M_{min} + (u_i(t_i, \vec{a}) - M_{min})) \\
= & \textstyle\sum_{a_{-i} \in A_{-i}} (p^{\vec{\tau}}_{(i,t_i,\vec{a}_{-i})} - p^{\vec{\sigma}}_{(i,t_i,\vec{a}_{-i})})(u_i(t_i, \vec{a}) - M_{min}) \\
\leq & \textstyle\sum_{a_{-i} \in A_{-i}} \left| p^{\vec{\tau}}_{(i,t_i,\vec{a}_{-i})} - p^{\vec{\sigma}}_{(i,t_i,\vec{a}_{-i})} \right| \cdot (M_{max} - M_{min}) \\
= & \Delta^{\vec{\tau}}_{(i,t_i)} \cdot M.
\end{aligned}$$

Therefore,

$$\begin{aligned}
& \textstyle\sum_{a_i \in A_i} \tau_i(t_i, a_i) \left(\sum_{a_{-i} \in A_{-i}} (p^{\vec{\tau}}_{(i,t_i,\vec{a}_{-i})} - p^{\vec{\sigma}}_{(i,t_i,\vec{a}_{-i})}) u_i(t_i, \vec{a}) \right) \\
\leq & (\Delta^{\vec{\tau}}_{(i,t_i)} \cdot M) \textstyle\sum_{a_i \in A_i} \tau_i(t_i, a_i) \\
= & \Delta^{\vec{\tau}}_{(i,t_i)} \cdot M,
\end{aligned}$$

as desired.

□

The second preliminary lemma gives the expected utility that a player can get by betting on the other players' actions.

Lemma 6.2. *Given a player $i \in P$, its type $t_i \in T_i$, and a strategy profile $\vec{\tau}$ for $\Gamma^{\vec{\sigma}}$. The expected utility that i gets when betting on action profile $\vec{a}_{-i}$ is given by $p^{\vec{\tau}}_{(i,t_i,\vec{a}_{-i})} - p_{(i,t_i,\vec{a}_{-i})}$.*

The proof of Lemma 6.2 is equivalent to that of Lemma 5.1. The following corollary is a direct implication.

Corollary 6.3. *Given a player $i \in P$, its type $t_i \in T_i$, and a Nash equilibrium $\vec{\tau}$ for $\Gamma^{\vec{\sigma}}$. The maximum expected utility $w_i(t_i, \vec{\tau}_i)$ that i can get by betting when players play $\vec{\tau}$ and i has type t_i satisfies*

$$w_i(t_i, \vec{\tau}_i) \leq \Delta^{\vec{\tau}}_{(i,t_i)}$$

Proof. By Lemma 6.2, the maximum utility that i can get by betting, conditioned to the fact that i has type t_i and players play $\vec{\tau}$ is given by

$$\max_{\vec{a}_{-i} \in A_{-i}} \{p^{\vec{\tau}}_{(i,t_i,\vec{a}_{-i})} - p^{\vec{\sigma}}_{(i,t_i,\vec{a}_{-i})}\}.$$

By definition, this value is at most $\Delta^{\vec{\tau}}_{(i,t_i)}$. □

The third and final preliminary lemma gives a lower bound on the expected losses of other players incurred by i's bets.

Lemma 6.4. *Given a player $i \in P$, its type $t_i \in T_i$, and a Nash equilibrium $\vec{\tau}$ for $\Gamma^{\vec{\sigma}}$, denote by $\ell^{\vec{\sigma}}_{j,i}(t_i, \vec{\tau})$ the expected utility that j loses because of i's bet, conditioned to the fact that i plays τ_i with type t_i. Then,*

$$\ell^{\vec{\sigma}}_{j,i}(t_i, \vec{\tau}) \leq -\frac{\Delta^{\vec{\tau}}_{(i,t_i)} C}{2|A_{-i}|}.$$

Proof. If $\Delta^{\vec{\tau}}_{(i,t_i)} = 0$, the lemma is automatically true since $\ell^{\vec{\sigma}}_{j,i}(t_i, \vec{\tau}) \leq 0$. Therefore, for the rest of the proof we will assume that $\Delta^{\vec{\tau}}_{(i,t_i)} > 0$, and thus that there exists an action profile $\vec{a}_{-i}$ such that $p^{\vec{\tau}}_{(i,t_i,\vec{a}_{-i})} - p^{\vec{\sigma}}_{(i,t_i,\vec{a}_{-i})} > 0$. By Lemma 6.2, this implies that there exists at least one bet for i that gives her strictly positive utility, and therefore that i will always bet with probability 1 since $\vec{\tau}$ is a Nash equilibrium.

Let

$$X := \sum_{\substack{\vec{a}_{-i} \in A_{-i} \\ p^{\vec{\tau}}_{(i,t_i,\vec{a}_{-i})} - p_{(i,t_i,\vec{a}_{-i})} \geq 0}} (p^{\vec{\tau}}_{(i,t_i,\vec{a}_{-i})} - p^{\vec{\sigma}}_{(i,t_i,\vec{a}_{-i})})$$

and

$$Y := \sum_{\substack{\vec{a}_{-i} \in A_{-i} \\ p^{\vec{\tau}}_{(i,t_i,\vec{a}_{-i})} - p_{(i,t_i,\vec{a}_{-i})} < 0}} (p^{\vec{\tau}}_{(i,t_i,\vec{a}_{-i})} - p^{\vec{\sigma}}_{(i,t_i,\vec{a}_{-i})}).$$

We have that

$$X + Y = \sum_{\vec{a}_{-i} \in A_{-i}} p^{\vec{\tau}}_{(i,t_i,\vec{a}_{-i})} - \sum_{\vec{a}_{-i} \in A_{-i}} p^{\vec{\sigma}}_{(i,t_i,\vec{a}_{-i})} = 0.$$

Moreover, by definition, $X - Y = \Delta^{\vec{\tau}}_{(i,t_i)}$. Putting the last two equations together we get that $X = \frac{\Delta^{\vec{\tau}}_{(i,t_i)}}{2}$. Therefore, since X is a sum of at most $|A_{-i}|$ terms, there exists an action profile $\vec{a}_{-i} \in A_{-i}$ such that

$$p^{\vec{\tau}}_{(i,t_i,\vec{a}_{-i})} - p^{\vec{\sigma}}_{(i,t_i,\vec{a}_{-i})} \geq \frac{1}{|A_{-i}|} \cdot \frac{\Delta^{\vec{\tau}}_{(i,t_i)}}{2} = \frac{\Delta^{\vec{\tau}}_{(i,t_i)}}{2|A_{-i}|}.$$

We show next that the loss of every other player when i bets on an action that maximizes $p^{\vec{\tau}}_{(i,t_i,\vec{a}_{-i})} - p^{\vec{\sigma}}_{(i,t_i,\vec{a}_{-i})}$ is upper-bounded by $-\frac{\Delta^{\vec{\tau}}_{(i,t_i)} C}{2|A_{-i}|}$. Note that this completes the proof since, by Lemma 6.2, i will always bet on actions that maximize $p^{\vec{\tau}}_{(i,t_i,\vec{a}_{-i})} - p^{\vec{\sigma}}_{(i,t_i,\vec{a}_{-i})}$ since $\vec{\tau}$ is a Nash equilibrium.

Let $\vec{a}_{-i}$ be an action profile that maximizes $p^{\vec{\tau}}_{(i,t_i,\vec{a}_{-i})} - p^{\vec{\sigma}}_{(i,t_i,\vec{a}_{-i})}$. The expected loss $\ell^{\vec{\sigma}}_{j,i}(t_i, \vec{a}_{-i})$ of every other player j when i bets on $\vec{a}_{-i}$ and i has type t_i is given by

$$\ell^{\vec{\sigma}}_{j,i}(t_i, \vec{a}_{-i}) = p^{\vec{\tau}}_{(i,t_i,\vec{a}_{-i})} \cdot (-C).$$

Since $p^{\vec{\tau}}_{(i,t_i,\vec{a}_{-i})} \geq p^{\vec{\sigma}}_{(i,t_i,\vec{a}_{-i})} + \frac{\Delta^{\vec{\tau}}_{(i,t_i)}}{2|A_{-i}|} \geq \frac{\Delta^{\vec{\tau}}_{(i,t_i)}}{2|A_{-i}|}$, we get that

$$\ell^{\vec{\sigma}}_{j,i}(t_i, \vec{a}_{-i}) \leq -\frac{\Delta^{\vec{\tau}}_{(i,t_i)} C}{2|A_{-i}|},$$

and this completes the proof. □

With Lemmas 6.1 and 6.4, and Corollary 6.3, we are ready to prove Theorem 3.1. Let S be a coalition of players and $\vec{\tau}_S$ be a strategy profile for players in S such that (a1) $u^{\vec{\sigma}}_i(\vec{\sigma}_{-S}, \vec{\tau}_S) > u^{\vec{\sigma}}_i(\vec{\sigma})$

for all $i \in S$, and (b1) $u_i^{\vec{\sigma}}(\vec{\sigma}_{-S}, \vec{\tau}_S) \geq u_i^{\vec{\sigma}}(\vec{\sigma}_{-S}, \vec{\tau}_{S\setminus\{i\}}, \tau_i')$ for all $i \in S$ and all strategies τ_i' for i. By construction, we have that

$$\begin{aligned} u_i^{\vec{\sigma}}(\vec{\sigma}_{-S}, \vec{\tau}_S) &= u_i(\vec{\sigma}_{-S}, \tilde{\tau}_S) \\ &+ w_i^{\vec{\sigma}}(\vec{\sigma}_{-S}, \vec{\tau}_S) \\ &+ \ell_i^{\vec{\sigma}}(\vec{\sigma}_{-S}, \vec{\tau}_S). \end{aligned}$$

If we denote by $p(i, t_i)$ the probability that i gets type t_i according to distribution p, can rewrite the previous identity as

$$\begin{aligned} u_i^{\vec{\sigma}}(\vec{\sigma}_{-S}, \vec{\tau}_S) &= \textstyle\sum_{t_i \in T_i} p(i, t_i) u_i(t_i, (\vec{\sigma}_{-S}, \tilde{\tau}_S)) \\ &+ \textstyle\sum_{t_i \in T_i} p(i, t_i) w_i^{\vec{\sigma}}(t_i, (\vec{\sigma}_{-S}, \vec{\tau}_S)) \\ &+ \textstyle\sum_{j \in S\setminus\{i\}} \ell_{(i,j)}^{\vec{\sigma}}(\vec{\sigma}_{-S}, \vec{\tau}_S). \end{aligned}$$

Applying Lemma 6.1 and Corollary 6.3 we get that

$$\begin{aligned} u_i^{\vec{\sigma}}(\vec{\sigma}_{-S}, \vec{\tau}_S) &\leq \textstyle\sum_{t_i \in T_i} p(i, t_i)(u_i(t_i, \vec{\sigma}) + \Delta_{(i,t_i)}^{(\vec{\sigma}_{-S}, \vec{\tau}_S)} \cdot M) \\ &+ \textstyle\sum_{t_i \in T_i} p(i, t_i) \Delta_{(i,t_i)}^{(\vec{\sigma}_{-S}, \vec{\tau}_S)} \\ &+ \textstyle\sum_{j \in S\setminus\{i\}} \ell_{(i,j)}^{\vec{\sigma}}(\vec{\sigma}_{-S}, \vec{\tau}_S), \end{aligned}$$

which can be written as

$$\begin{aligned} u_i^{\vec{\sigma}}(\vec{\sigma}_{-S}, \vec{\tau}_S) &\leq \textstyle\sum_{t_i \in T_i} p(i, t_i)(u_i(t_i, \vec{\sigma}) + \Delta_{(i,t_i)}^{(\vec{\sigma}_{-S}, \vec{\tau}_S)} \cdot (M+1)) \\ &+ \textstyle\sum_{j \in S\setminus\{i\}} \ell_{(i,j)}^{\vec{\sigma}}(\vec{\sigma}_{-S}, \vec{\tau}_S), \end{aligned}$$

By linearity of expectation, this implies that

$$\begin{aligned} \textstyle\sum_{i \in S} u_i^{\vec{\sigma}}(\vec{\sigma}_{-S}, \vec{\tau}_S) &\leq \textstyle\sum_{i \in S} \sum_{t_i \in T_i} p(i, t_i)(u_i(t_i, \vec{\sigma}) + \Delta_{(i,t_i)}^{(\vec{\sigma}_{-S}, \vec{\tau}_S)} \cdot (M+1)) \\ &+ \textstyle\sum_{i \in S} \sum_{j \in S\setminus\{i\}} \ell_{(i,j)}^{\vec{\sigma}}(\vec{\sigma}_{-S}, \vec{\tau}_S). \end{aligned} \tag{1}$$

Note that $\sum_{i \in S} \sum_{j \in S\setminus\{i\}} \ell_{(i,j)}^{\vec{\sigma}}(\vec{\sigma}_{-S}, \vec{\tau}_S) = \sum_{i \in S} \sum_{j \in S\setminus\{i\}} \ell_{(j,i)}^{\vec{\sigma}}(\vec{\sigma}_{-S}, \vec{\tau}_S)$. Moreover, we can write $\ell_{(j,i)}^{\vec{\sigma}}(\vec{\sigma}_{-S}, \vec{\tau}_S)$ as $\sum_{t_i \in T_i} p(i, t_i) \ell_{(j,i)}^{\vec{\sigma}}(t_i, (\vec{\sigma}_{-S}, \vec{\tau}_S))$. Therefore, by Lemma 6.4, we get that

$$\ell_{(j,i)}^{\vec{\sigma}}(\vec{\sigma}_{-S}, \vec{\tau}_S) \leq \sum_{t_i \in T_i} -p(i, t_i) \cdot \frac{\Delta_{(i,t_i)}^{(\vec{\sigma}_{-S}, \vec{\tau}_S)} C}{2|A_{-i}|}.$$

This implies that

$$\sum_{i \in S} \sum_{j \in S\setminus\{i\}} \ell_{(j,i)}^{\vec{\sigma}}(\vec{\sigma}_{-S}, \vec{\tau}_S) \leq \sum_{i \in S} \sum_{t_i \in T_i} -p(i, t_i) \cdot (|S| - 1) \cdot \frac{\Delta_{(i,t_i)}^{(\vec{\sigma}_{-S}, \vec{\tau}_S)} C}{2|A_{-i}|}$$

Inserting this into Equation 1 gives that

$$\sum_{i \in S} u_i^{\vec{\sigma}}(\vec{\sigma}_{-S}, \vec{\tau}_S) \leq \sum_{i \in S} \sum_{t_i \in T_i} p(i, t_i) \left[(u_i(t_i, \vec{\sigma}) + \Delta_{(i,t_i)}^{(\vec{\sigma}_{-S}, \vec{\tau}_S)} \cdot (M+1)) - (|S| - 1) \cdot \frac{\Delta_{(i,t_i)}^{(\vec{\sigma}_{-S}, \vec{\tau}_S)} C}{2|A_{-i}|} \right].$$

Since $C = 2|A| \cdot (M+1)$, it follows that

$$(|S| - 1) \cdot \frac{\Delta_{(i,t_i)}^{(\vec{\sigma}_{-S}, \vec{\tau}_S)} C}{2|A_{-i}|} \geq \Delta_{(i,t_i)}^{(\vec{\sigma}_{-S}, \vec{\tau}_S)} \cdot (M+1),$$

and therefore

$$\sum_{i\in S} u_i^{\vec{\sigma}}(\vec{\sigma}_{-S}, \vec{\tau}_S) \leq \sum_{i\in S}\sum_{t_i\in T_i} p(i,t_i)u_i(t_i,\vec{\sigma}) = \sum_{i\in S} u_i(\vec{\sigma}).$$

This implies that there exists at least one player $i \in S$ such that $u_i^{\vec{\sigma}}(\vec{\sigma}_{-S}, \vec{\tau}_S) \leq u_i(\vec{\sigma})$, which contradicts the (a1) assumption. This proves Theorem 3.1 part (a). Part (b) follows from the same argument with $S = P$.

7 PROOF OF THEOREMS 3.1 AND 3.2, INFINITE CASE

As stated in Section 3, the proof of Theorem 3.1 can be generalized to games in which the sets of actions and types are infinite. We'll prove this generalization in the case where Γ is a normal-form game. The generalization in the case of Bayesian games is identical, except that we use the proofs and constructions from Section 6 as a starting point (as opposed to the proofs and constructions of Section 5.2).

For ease of exposition, we will begin by generalizing the constructions given in previous sections to games with a countable number of actions. The proof for arbitrary sets is given in Appendix B and follows the same lines.

7.1 Countable Case

As suggested by Example 4.3, the proof given in Section 5.1 does not quite work when the sets of actions are infinite. In fact, the construction for finite games uses the fact that if a coalition S defects, there exists a player i and an action a such that $\frac{1}{|A|_{max}}$ of the players bet on (i,a) in expectation. This allowed us to adjust C so that it is never beneficial for i to play a. If the game is infinite, there is no lower bound of the expected number of players that bet on a particular action, and therefore there is no value of C that does the trick.

We can avoid this problem by slightly modifying the bet space. Suppose that instead of betting on pairs (i,a) with $i \in P$ and $a \in A_i$, players can bet on pairs (i,X) with $i \in P$ and $X \subseteq A_i$. Given the actions and bets $(a_1,b_1),\ldots,(a_n,b_n)$ of the players, we define the utilities of $\Gamma^{\vec{\sigma}}$ as in Section 5.1, except that

$$w_i^{\vec{\sigma}}((a_1,b_1),(a_2,b_2),\ldots,(a_n,b_n)) := \begin{cases} 0 & \text{if } b_i = \bot \\ 1 - p^{\vec{\sigma}}_{(j,X)} & \text{if } b_i \text{ is of the form } (j,X) \text{ with } a_j \in X \\ -p^{\vec{\sigma}}_{(j,X)} & \text{if } b_i \text{ is of the form } (j,X) \text{ with } a_j \notin X \end{cases}$$

where $p^{\vec{\sigma}}_{(j,X)}$ is the probability that the first component of j's action is in X, assuming that j uses strategy σ_j, and

$$\ell^{\vec{\sigma}}_{(i,j)}((a_1,b_1),(a_2,b_2),\ldots,(a_n,b_n)) := \begin{cases} 0 & \text{if } b_j = \bot \text{ or } b_j \text{ is of the form } (i,X) \text{ with } a \notin X \\ -M-1 & \text{otherwise,} \end{cases}$$

where M is the difference between the supremum and the infimum utility that a player can get in Γ. An analogous reasoning to that of Lemma 5.1 shows that the expected utility $x^{\tau_j}_{(j,a_j)}$ that a player gets by betting on (j,X) conditioned to the fact that j is playing strategy τ_j, is given by

$$x^{\tau_j}_{(j,a_j)} := p^{\tau_j}_{(j,X)} - p^{\sigma_j}_{(j,X)}.$$

This implies that if player j plays some action a with more probability in τ_j than in σ_j, including action a in X strictly increases the expected utility of betting on (j,X). This proves the following lemma.

Lemma 7.1. *Given a normal-form game $\Gamma = (P, A, U)$, let $S \subset P$ be a subset of players and $\vec{\tau}_S$ be a strategy for players in S in $\Gamma^{\vec{\sigma}}$ such that $u_i^{\vec{\sigma}}(\vec{\sigma}_{-S}, \vec{\tau}_S) \geq u_i^{\vec{\sigma}}(\vec{\sigma}_{-S}, \vec{\tau}_{S\setminus\{i\}}, \tau_i')$ for all $i \in S$ and all strategies τ_i' for i. For each player $i \in S$, let $X_i^{>}$ (resp., $X_i^{<}$) be the set of all actions $a \in A_i$ such that the probability that i plays a with $\tilde{\tau}_i$ is strictly greater (resp., smaller) than the probability that i plays a with σ_i. Then, all bets of the form (i, X) that may occur with positive probability when playing $(\vec{\sigma}_{-S}, \vec{\tau}_S)$ must satisfy that $X_i^{>} \subseteq X$ and $X_i^{<} \cap X = \emptyset$.*

With Lemma 7.1, we can use a similar argument to the one used in the finite case. Suppose that a coalition S defects from the proposed strategy following strategy profile $\vec{\tau}_S$ in such a way that (a1) all players in S get more utility with $\vec{\tau}_S$ than with $\vec{\sigma}$ and (b1) no player in S can further increase its utility by defecting from $\vec{\tau}_S$. Then, it is straightforward to check that, if (a1) is satisfied, all players in S bet with probability 1. Therefore, there exists a player $i \in S$ such that at least one other player bets on i in expectation. By Lemma 7.1, all of these bets must include all elements in $X_i^{>}$ and none of them includes any element of $X_i^{<}$. Therefore, if i reduces the probability of playing any action in $X_i^{>}$ by δ and increases the probability of playing any action in $X_i^{<}$ by δ, an analogous reasoning to the one at the end of Section 5.2 shows that the utility of i increases at least by

$$\delta(M+1) - \delta M,$$

which contradicts the (b1) assumption. The proof for Bayesian games follows the same generalization (betting on sets of action profiles rather than individual action profiles), using as a basis the construction in 6.

7.2 Uncountable Case

The idea used for the countable case can be generalized for any probability space, although the proof is much more technical. See Appendix B for a full proof.

8 CONCLUSION AND OPEN PROBLEMS

This paper shows that if a game designer has the power to add new actions to a normal-form or Bayesian game, it can shield a given Nash equilibrium from collusion and from equilibrium selection issues. More precisely, given a normal-form (resp., Bayesian) game Γ and a Nash equilibrium $\vec{\sigma}$ of Γ, we showed that exists an extension $\Gamma^{\vec{\sigma}}$ of Γ such that (a) $\vec{\sigma}$ is a semi-strong Nash equilibrium of Γ and (b) $\vec{\sigma}$ Pareto-dominates (resp., quasi Pareto-dominates) all other equilibria of $\Gamma^{\vec{\sigma}}$. This result opens the door for three important follow-up problems.

- **Can we also get Pareto-dominance for arbitrary Bayesian games?** We proved that the construction of $\Gamma^{\vec{\sigma}}$ given in Section 6 satisfies that $\vec{\sigma}$ quasi Pareto-dominates all other Nash equilibria of $\Gamma^{\vec{\sigma}}$. We conjecture that, in fact, $\vec{\sigma}$ also Pareto-dominates all other Nash equilibria in this construction.
- **Minimizing promises.** As discussed in earlier sections, we can view the newly added actions in $\Gamma^{\vec{\sigma}}$ as *promises* made by the game designer to the players. In practice, if players are rational and have full information, these newly actions should never be played since every Nash equilibrium that includes them is Pareto-dominated (or quasi Pareto-dominated) by $\vec{\sigma}$. Thus, the game designer can simply bluff about the existence of these new actions (e.g., by promising to pay certain amount to a given player). Nevertheless, it is still an open question to see exactly how much utility must the game designer promise in order to construct a game $\Gamma^{\vec{\sigma}}$ that satisfies the properties of Theorem 3.1 and 3.2.
- **Minimum size of $\Gamma^{\vec{\sigma}}$.** Theorems 3.1 and 3.2 give upper bounds on the minimum number of actions that must be added in order to construct $\Gamma^{\vec{\sigma}}$. It is still an open problem to verify if there exists a smaller construction.

REFERENCES

Ittai Abraham, Lorenzo Alvisi, and Joseph Y Halpern. 2011. Distributed computing meets game theory: combining insights from two fields. *Acm Sigact News* 42, 2 (2011), 69–76.

Ittai Abraham, Danny Dolev, Ivan Geffner, and Joseph Y Halpern. 2019. Implementing mediators with asynchronous cheap talk. In *Proceedings of the 2019 ACM Symposium on Principles of Distributed Computing*. 501–510.

Ittai Abraham, Danny Dolev, Rica Gonen, and Joseph Y Halpern. 2006. Distributed computing meets game theory: robust mechanisms for rational secret sharing and multiparty computation. In *Proceedings of the twenty-fifth annual ACM symposium on Principles of distributed computing*. 53–62.

Ittai Abraham, Danny Dolev, and Joseph Y Halpern. 2008. Lower bounds on implementing robust and resilient mediators. In *Theory of Cryptography: Fifth Theory of Cryptography Conference, TCC 2008, New York, USA, March 19-21, 2008. Proceedings 5*. Springer, 302–319.

Amitanand S Aiyer, Lorenzo Alvisi, Allen Clement, Mike Dahlin, Jean-Philippe Martin, and Carl Porth. 2005. BAR fault tolerance for cooperative services. In *Proceedings of the twentieth ACM symposium on Operating systems principles*. 45–58. https://doi.org/10.1145/1095810.1095816

Robert J Aumann. 1959. Acceptable points in general cooperative n-person games. *Contributions to the Theory of Games* 4 (1959), 287–324.

Robert J Aumann. 1974. Subjectivity and correlation in randomized strategies. *Journal of mathematical Economics* 1, 1 (1974), 67–96.

Michael Bacharach. 1999. Interactive team reasoning: A contribution to the theory of co-operation. *Research in economics* 53, 2 (1999), 117–147.

Michael Bacharach. 2006. *Beyond individual choice: teams and frames in game theory*. Princeton University Press.

Nicholas Bardsley, Judith Mehta, Chris Starmer, and Robert Sugden. 2010. Explaining focal points: Cognitive hierarchy theory versus team reasoning. *The Economic Journal* 120, 543 (2010), 40–79.

Michael Ben-Or. 1983. Another advantage of free choice (extended abstract) completely asynchronous agreement protocols. In *Proceedings of the second annual ACM symposium on Principles of distributed computing*. 27–30.

Michael Ben-Or, Ran Canetti, and Oded Goldreich. 1993. Asynchronous secure computation. In *Proceedings of the twenty-fifth annual ACM symposium on Theory of computing*. 52–61.

M. Ben-Or, S. Goldwasser, and A. Wigderson. 1988. Completeness theorems for non-cryptographic fault-tolerant distributed computation. In *Proceedings of the 20th ACM Symposium on Theory of Computing*. 1–10.

B Douglas Bernheim, Bezalel Peleg, and Michael D Whinston. 1987. Coalition-proof nash equilibria i. concepts. *Journal of economic theory* 42, 1 (1987), 1–12.

Gabriel Bracha. 1984. An asynchronous [(n-1)/3]-resilient consensus protocol. In *Proceedings of the third annual ACM symposium on Principles of distributed computing*. 154–162.

Gabriel Bracha and Sam Toueg. 1985. Asynchronous consensus and broadcast protocols. *Journal of the ACM (JACM)* 32, 4 (1985), 824–840.

Francoise Forges. 1986. An approach to communication equilibria. *Econometrica: Journal of the Econometric Society* (1986), 1375–1385.

Ivan Geffner and Joseph Y Halpern. 2021. Security in asynchronous interactive systems. In *International Symposium on Stabilizing, Safety, and Security of Distributed Systems*. Springer, 123–140.

Ivan Geffner and Joseph Y Halpern. 2023a. Communication games, sequential equilibrium, and mediators. *arXiv preprint, arXiv:2309.14618* (2023).

Ivan Geffner and Joseph Y Halpern. 2023b. Lower Bounds on Implementing Mediators in Asynchronous Systems with Rational and Malicious Agents. *J. ACM* 70, 2 (2023), 1–21.

Ivan Geffner and Moshe Tennenholtz. 2023. Strengthening Nash Equilibria. *arXiv preprint arXiv:2312.14745* (2023).

Allan Gibbard and William L Harper. 1978. Counterfactuals and two kinds of expected utility. In *Ifs: Conditionals, belief, decision, chance and time*. Springer, 153–190.

Natalie Gold and Andrew M Colman. 2020. Team reasoning and the rational choice of payoff-dominant outcomes in games. *Topoi* 39, 2 (2020), 305–316.

Eric Maskin. 1999. Nash equilibrium and welfare optimality. *The Review of Economic Studies* 66, 1 (1999), 23–38.

Eric Maskin and Tomas Sjöström. 2002. Implementation theory. *Handbook of social Choice and Welfare* 1 (2002), 237–288.

Dov Monderer and Moshe Tennenholtz. 2003. k-Implementation. In *Proceedings of the 4th ACM conference on Electronic Commerce*. 19–28.

Dov Monderer and Moshe Tennenholtz. 2009. Strong mediated equilibrium. *Artificial Intelligence* 173, 1 (2009), 180–195.

Caspar Oesterheld, Johannes Treutlein, Roger Grosse, Vincent Conitzer, and Jakob Foerster. 2022. Similarity-based cooperation. *arXiv preprint, arXiv:2211.14468* (2022).

Thomas R Palfrey. 2002. Implementation theory. *Handbook of game theory with economic applications* 3 (2002), 2271–2326.

Michal Penn, Maria Polukarov, and Moshe Tennenholtz. 2005. Congestion games with failures. In *Proceedings of the 6th ACM Conference on Electronic Commerce*. 259–268.

Michael O Rabin. 1983. Randomized byzantine generals. In *24th annual symposium on foundations of computer science (sfcs 1983)*. IEEE, 403–409.

Alejandro Ranchal-Pedrosa and Vincent Gramoli. 2021. Rational agreement in the presence of crash faults. In *2021 IEEE International Conference on Blockchain*. IEEE, 470–475.

Robert Sugden. 1993. Thinking as a team: Towards an explanation of nonselfish behavior. *Social philosophy and policy* 10, 1 (1993), 69–89.

Moshe Tennenholtz. 2004. Program equilibrium. *Games and Economic Behavior* 49, 2 (2004), 363–373.

Xavier Vilaça, Joao Leitao, Miguel Correia, and Luís Rodrigues. 2011. N-party BAR transfer. In *Proceedings of the 2011 Conference on Principles of Distributed Systems*. Springer, 392–408.

A PROOF OF THEOREM 3.1 FOR FINITE BAYESIAN GAMES WITH INDEPENDENT TYPES

As shown in Section 4, the construction of $\Gamma^{\vec{\sigma}}$ used for normal-form games does not work in general for Bayesian games. However, if types are independent, we have the following proposition that states that, if a coalition can defect in such a way that they all increase their expected utility with respect to $\vec{\sigma}$, each player in the coalition must be able to tell the strategy of another member j of the coalition from σ_j.

Proposition 4. *Let $\Gamma = (P, T, q, A, U)$ be a Bayesian game with independent types. Let S be a subset of players and $\vec{\tau}_S$ be a strategy profile for players in S such that $u_i(\vec{\sigma}_{-S}, \vec{\tau}_S) > u_i(\vec{\sigma})$ for all $i \in S$. Then, there exists a player $i \in S$ and an action $a_i \in A_i$ such that*

$$p^{\tau_i}_{(i,a_i)} > p^{\sigma_i}_{(i,a_i)},$$

where $p^{\tau_i}_{(i,a_i)}$ (resp., $p^{\sigma_i}_{(i,a_i)}$) is the probability that player i plays action a_i when following strategy τ_i (resp., σ_i).

Proof. Suppose that $p^{\tau_i}_{(i,a_i)} = p^{\sigma_i}_{(i,a_i)}$ for all $i \in S$ and all $a_i \in A_i$. Then, since types are independent, if we fix a player $i \in S$ and a type t_i of i, the distribution over action profiles $\vec{a}_{-i}$ played by the other players j conditioned to the fact that i has type t_i is identical when players in S play $\vec{\tau}_S$ and when players in S play $\vec{\sigma}_S$. Therefore, we would have that

$$u_i(\vec{\sigma}_{-S}, \vec{\tau}_S) = u_i(\vec{\sigma}_{-i}, \tau_i)$$

for all $i \in S$, and thus

$$u_i(\vec{\sigma}_{-i}, \tau_i) > u_i(\vec{\sigma}),$$

which contradicts the fact that $\vec{\sigma}$ is a Nash equilibrium of Γ. □

Note that the proof of Proposition 4 requires that the types of the players are independent. In fact, if they were not, Example 4.2 shows that players can defect in such a way that they all benefit while no player i can tell between the strategy of another player j and σ_j. However, in the case of independent types, Proposition 4 shows that we can use the same approach as in the case of normal-form games. Define $A^{\vec{\sigma}}$ as in Section 5.1, and let

$$\begin{aligned} u_i^{\vec{\sigma}}(t_i, ((a_1, b_1), (a_2, b_2), \ldots, (a_n, b_n))) \quad &:= \quad u_i(t_i, (a_1, a_2, \ldots, a_n)) \\ &+ \quad w_i^{\vec{\sigma}}((a_1, b_1), (a_2, b_2), \ldots, (a_n, b_n)) \\ &+ \quad \ell_i^{\vec{\sigma}}((a_1, b_1), (a_2, b_2), \ldots, (a_n, b_n)), \end{aligned}$$

where w_i and ℓ_i are defined in the same way as in Section 5.1 and $C > \max\{M \cdot |A|_{\max}, 1\}$, as in Section 5.2. An analogous reasoning to that of the proof of Proposition 2 shows that, if S and $\vec{\tau}_S$ satisfy that $u_i^{\vec{\sigma}}(\vec{\sigma}_{-S}, \vec{\tau}_S) > u_i^{\vec{\sigma}}(\vec{\sigma})$ for all $i \in S$, and that $u_i^{\vec{\sigma}}(\vec{\sigma}_{-S}, \vec{\tau}_S) \geq u_i^{\vec{\sigma}}(\vec{\sigma}_{-S}, \vec{\tau}_{S\setminus\{i\}}, \tau_i')$ for all $i \in S$ and all strategies τ_i' for i, then there exists a player i and an action a_i such that, regardless of i's

type, at least $\frac{1}{|A|_{max}}$ players bet on (i, a_i) in expectation. The rest of the proof proceeds as in the case of normal-form games.

B PROOF OF THEOREMS 3.1 AND 3.2, UNCOUNTABLE CASE

We begin by giving the basic notions on probability theory, and then move to the desired generalization.

B.1 Basic Definitions

Given a set X, a *σ-algebra* on X is a collection Σ of subsets of X that is closed under complement, countable unions, and countable intersections. This means that,

- $Y \in \Sigma \implies \overline{Y} \in \Sigma$, where $\overline{Y}$ denotes the complement of Y in X (i.e., $X \setminus Y$).
- If $\{X_i\}_{i=1}^{\infty}$ is a countable collection of elements of Σ, then both $\bigcup_{i\geq 1} X_i$ and $\bigcup_{i\geq 1} X_i$ are also elements of Σ.

A pair (X, Σ) consisting of a set X and a σ-algebra Σ on X is called a *measurable space*. Given a measurable space (X, Σ), a measure on (X, Σ) is a function $\mu : \Sigma \longrightarrow \mathbb{R} \cup \{-\infty, +\infty\}$ that satisfies

- $\mu(\emptyset) = 0$.
- **Non-negativity:** $\mu(Y) \geq 0$ for all $Y \in \Sigma$.
- **Countable additivity:** For all countable collections $\{X_i\}_{i=1}^{\infty}$ of pairwise disjoint sets in Σ, $\mu(\bigcup_{i=1}^{\infty} X_i) = \sum_{i=1}^{\infty} \mu(X_i)$.

Intuitively, a measure is a non-negative function that assigns 0 to the empty set and such that the measure of a countable union disjoint sets is the sum of their measures.

With these definitions, we can define a probability space. A probability space is a tuple (X, Σ, P), where (X, Σ) is a σ-algebra, and P is a measure such that the image of all elements of Σ is in $[0, 1]$, and such that the measure of the whole space is 1. More precisely, P must satisfy that

- $P(E) \in [0, 1]$ for all $E \in \Sigma$.
- $P(X) = 1$.

In the next sections we will assume that, in a normal-form game Γ, the action sets A_i of each player have an embedded σ-algebra Σ_i, and the strategy of each player i consists of choosing a measure Q_i such that (A_i, Σ_i, Q_i) is a probability space. Intuitively, this means that i cannot choose which sets of A_i are measurable, but it can choose what probability it is assigned to each of these sets (i.e., with what probability each action is sampled). Given the strategy P_i of each player, the probability space (A, Σ, Q) that gives the probability of each subset of action profiles when each player plays Q_i is precisely the product of (A_i, Σ_i, P_i) for all $i \in P$, in which $A := A_1 \times A_2 \times \ldots \times A_n$, $\Sigma := \sigma(\Sigma_1 \otimes \ldots \otimes \Sigma_n)$ (i.e., the σ-algebra generated by sets of the form $(X_1, \ldots, X_n)$ with $X_i \in \Sigma_i$ for all i), and $Q := Q_1 \times \ldots \times Q_n$, which is the unique measure function such that $Q(X_1, \ldots, X_n) = Q_1(X_1)Q_2(X_2)\ldots Q_n(X_n)$. Fixing (A, Σ, Q), the expected utility of a player i is given by the Lebesgue integral

$$u_i(Q) := \int_A u_i(X)\, dQ.$$

In order to avoid confusion between the strategy profile $\vec{\sigma}$ of Theorems 3.1 and 3.2 and the σ used in the definition of σ-algebras, in the next sections we will use Q instead of $\vec{\sigma}$ to denote the Nash equilibrium of Theorems 3.1 and 3.2.

B.2 Construction of Γ^Q

The construction of Γ^Q for arbitrary probability spaces is quite similar to the one given in Section 7.1. In fact, we define $A_i^Q := A_i \cup \left(A_i \times \bigcup_{j\in P}(\{j\} \times \Sigma_j)\right)$. Intuitively, i can either play an action in A_i,

or play an action in A_i and simultaneously bet that the action that another player j plays is in some measurable subset $S \in \Sigma_j$. The σ-algebra of A_i^Q is the one spanned by Σ_i and $\sigma(\Sigma_i \otimes \mathcal{P}(\{j\} \times \Sigma_j))$ for each $j \in P$ (where $\mathcal{P}(S)$ denotes the power set of a set S).

We define U^Q exactly in the same way as in Section 7.1. Note that $p^Q_{(j,X)}$ is well defined since, by construction, X must belong to Σ_j. In fact, $p^Q_{(j,X)} = Q_j(X)$.

B.3 High Level Proof

A similar argument to the one given in Lemma 5.1 shows that, if players follow strategy profile Q', the expected utility that a player i gets when betting on (j, S) for some $S \in \Sigma_j$ is given by $\tilde{Q}'_j(S) - Q_j(S)$ (where, again, $\tilde{Q}'$ is the projection of Q' onto its first component). Therefore, it is optimal for player i to bet on sets that maximize $\tilde{Q}'_j(S) - Q_j(S)$ with probability 1. However, it is not clear at first if such a bet always exists, since it could be the case that the supremum of $\{\tilde{Q}'_j(S) - Q_j(S)\}_{S \in \Sigma_j}$ is never attained in Σ_j. The first part of the proof is precisely showing that such a bet always exists.

Proposition 5. *Let Γ^Q be the game constructed in Section B.2 and let $j \in P$ be a player that plays strategy Q'_j. Then, for all other players $i \in P$, there exists a maximal and a minimal bet of the form (j, X).*

The proof of Proposition 5 is somewhat technical and will be shown in Section B.4. Once we have Proposition 5, the second part of the proof is showing that, among all sets X that maximize or minimize the bets on j, there exists one that minimizes $Q(X)$.

Proposition 6. *Let Γ^Q be the game constructed in Section B.2 and let $j \in P$ be a player that plays strategy Q'_j. Let $\Sigma^+_{Q'_j}$ and $\Sigma^-_{Q'_j}$ be the subsets of Σ_j that maximize and minimize $\tilde{Q}'_j(S) - Q_j(S)$, respectively. Then, there exists a set $X^+ \in \Sigma^+_{Q'_j}$ and a set $X^- \in \Sigma^-_{Q'_j}$ such that $X^+ \cap X^- = \emptyset$, $Q(X^+) \leq Q(X)$ for all $X \in \Sigma^+_{Q'_j}$ and $Q(X^-) \leq Q(X)$ for all $X \in \Sigma^-_{Q'_j}$.*

The proof of Proposition 6 is quite similar to that of Proposition 5 and is also shown in Section B.4. With these propositions we are ready to give complete the proof of correctness of the construction given in Section B.2. Suppose that a subset S of players with $|S| \geq 2$ defects from Q_S to another strategy profile Q'_S in such a way that (a1) $u_i^Q(Q_{-S}, Q'_S) > u_i^Q(Q)$ for all $i \in S$, and (b1) $u_i^Q(Q_{-S}, Q'_{S\setminus\{i\}}, Q''_i) \leq u_i^Q(Q_{-S}, Q'_S)$ for all $i \in S$ and all strategies Q''_i for i. Then, there must exist a player $i \in S$ and a subset $X \in \Sigma_i$ such that $\tilde{Q}'_i(X) \neq Q_i(X)$ since, otherwise, all bets would give 0 expected utility and all players would get at most the expected utility they'd get with Q, which contradicts (a). If $\tilde{Q}'_i(X) \neq Q_i(X)$, it means that either $\tilde{Q}'_i(X) - Q_i(X) > 0$ or $\tilde{Q}'_i(\overline{X}) - Q_i(\overline{X}) > 0$, which implies that all other players in S must place bets with probability 1 (otherwise, it contradicts (b1)). It is easy to check that, if this is the case, there exists a player j such that the expected number of other players that bet on j is at least 1. Let X^+ and X^- be the subsets given in Proposition 6 (for player j) and consider a strategy Q''_j bets in the same way as in Q'_j but that intuitively does the following in order to select an action in A_j:

1. **Step 1:** j samples an action a according to Q'_j.
2. **Step 2:** If $a \notin X^+$, j plays a.
3. **Step 3:** Otherwise, j sequentially samples an action a according to Q_j until $a \in X^-$. Then, j plays a.

Note that this algorithm always terminates since $\tilde{Q}'_j(X^-) - Q_j(X^-) < 0$, which means that $Q_j(X^-) > 0$. More precisely, what j is doing is selecting an action in A_i according to the probability

measure $\tilde{Q}''_j$ on (A_j, Σ_j) defined by

$$\tilde{Q}''_j(X) := \tilde{Q}'_j(X \setminus X^+) + \frac{\tilde{Q}'_j(X^+)}{Q_j(X^-)} Q_j(X \cap X^-).$$

It is straightforward to check that $\tilde{Q}''_j(X)$ is indeed a probability measure on (A_j, Σ_j):

- $\tilde{Q}''_j(\emptyset) = \tilde{Q}'_j(\emptyset) + \frac{\tilde{Q}'_j(X^+)}{Q_j(X^-)} Q_j(\emptyset) = 0.$
- $\tilde{Q}''_j(A_j) = \tilde{Q}'_j(A_j) - \tilde{Q}'_j(X^+) + \frac{\tilde{Q}'_j(X^+)}{Q_j(X^-)} Q_j(X^-) = 1.$
- If $\{X_i\}_{i\geq 1}$ is a countable sequence of pairwise disjoint subsets in Σ_j, we have that $\left(\bigcup_{i\geq 1} X_i\right) \setminus X^+ = \bigcup_{i\geq 1}(X_i \setminus X^+)$, and $\left(\bigcup_{i\geq 1} X_i\right) \cap X^- = \bigcup_{i\geq 1}(X_i \cap X^-)$. Moreover, the subsets of the form $X_i \setminus X^+$ are pairwise disjoint, and so are those of the form $X_i \cap X^-$. Therefore,

$$\tilde{Q}''_j\left(\bigcup_{i\geq 1} X_i\right) = \sum_{i\geq 1} \tilde{Q}'_j(X_i \setminus X^+) + \frac{\tilde{Q}'_j(X^+)}{Q_j(X^-)} \sum_{i\geq 1} Q_j(X_i \cap X^-) = \sum_{i\geq 1} \tilde{Q}''_j(X_i).$$

We show next that j strictly improves its utility by switching from Q'_j to Q''_j. Recall that the expected utility $u_j^Q(Q_{-S}, Q'_{S\setminus\{j\}}, Q''_j)$ that j gets when playing strategy Q''_j can be decomposed as

$$\begin{aligned} u_j^Q(Q_{-S}, Q'_{S\setminus\{j\}}, Q''_j) &= u_j(Q_{-S}, \tilde{Q}'_{S\setminus\{j\}}, \tilde{Q}''_j) \\ &+ w_j(Q_{-S}, Q'_{S\setminus\{j\}}, Q''_j) \\ &+ \ell_j(Q_{-S}, Q'_{S\setminus\{j\}}, Q''_j). \end{aligned}$$

As in previous sections, $w_j(Q_{-S}, Q'_{S\setminus\{j\}}, Q''_j) = w_j(Q_{-S}, Q'_S)$ since, by construction, j places the same bets in Q'_j and Q''_j. We finish with the following lemma, whose proof is given in Section B.4.

Lemma B.1.

$$\begin{aligned} u_j(Q_{-S}, \tilde{Q}'_{S\setminus\{j\}}, \tilde{Q}''_j) - u_j(Q_{-S}, \tilde{Q}'_S) &\leq Q'_j(X^+)M \\ \ell_j(Q_{-S}, Q'_{S\setminus\{j\}}, Q''_j) - \ell_j(Q_{-S}, Q'_S) &> Q'_j(X^+)M. \end{aligned}$$

Note that this lemma implies that $u_j^Q(Q_{-S}, Q'_{S\setminus\{j\}}, Q''_j) > u_j(Q_{-S}, Q'_S)$, contradicting the (b1) assumption. This completes the proof of Theorem 3.1 part (a). A similar argument can be used to prove part (b), as done in Section 5.2.2.

B.4 Proof of the Technical Results

In this section, we prove the technical propositions and lemmas given in the previous section.

We begin with Proposition 5. Its proof follows from the following lemma.

Lemma B.2. *Let Σ be a σ-algebra and let $f : \Sigma \longrightarrow \mathbb{R}$ be a bounded function. Then, f has a global minimum and a global maximum on Σ.*

Note that Proposition 5 is an immediate consequence of Lemma B.2 since the expected utility of a bet is bounded between 1 and -1. To prove Lemma B.2 we need the following well-known result.

Lemma B.3. *Let Σ be a σ-algebra on some set Ω, let $\{X_n\}_{n\geq 1}$ be a sequence of elements of Σ, and let f be a function from Σ to $\mathbb{R}$ that satisfies countable additivity and such that*

(a) *$X_n \subseteq X_{n+1}$ for all $n \in \mathbb{N}$, and*
(b) *$\lim_{n\to\infty} f(X_n) = L$, for some $L \in \mathbb{R}$.*

Then $f\left(\bigcup_{n\geq 1} X_n\right) = L$.

Proof. Let $Y_n := X_{n+1} \setminus X_n$. Then, by (a), X_1 and all sets Y_n are pairwise disjoint. Therefore, since f satisfies countable additivity, we have that

$$f\left(\bigcup_{i=1}^{n} X_i\right) = f(X_n) = f(X_1) + \sum_{i=1}^{n-1} f(Y_i).$$

Since $\lim_{n\to\infty} f(X_n) = L$, it means that the sum is convergent and the limit is equal to $L - f(X_1)$. Therefore,

$$f\left(\bigcup_{n\geq 1} X_n\right) = f(X_1) + \sum_{i=1}^{\infty} f(Y_i) = L.$$

□

To prove Lemma B.2 we also need the following observation.

Lemma B.4. *Let Σ be a σ-algebra on some set Ω and let f be a function from Σ to $\mathbb{R}$ that satisfies countable additivity and such that $f(X) \leq L$ for all $X \in \Sigma$. Let $a_1, a_2, \ldots, a_n$ be a finite sequence of positive real numbers and $X_1, X_2, \ldots, X_n$ be a finite sequence of elements of Σ such that $f(X_i) > L - a_i$. Then,*

$$f\left(\bigcap_{i=1}^{n} X_i\right) > L - \sum_{i=1}^{n} a_i$$

.

Proof. We will only show this result for $n = 2$. The general case follows by induction.

Given two sets X and Y in Σ such that $f(X) > L - a$ and $f(Y) > L - b$, let $X' = X \setminus Y$ and let $Y' = Y \setminus X$. Then, we have that

$$\begin{array}{lcl} f(X') + f(X \cap Y) & > & L - a \\ f(Y') + f(X \cap Y) & > & L - b \\ f(X \cup Y) & \leq & L \end{array}$$

Writing $X \cup Y$ as $X' \cup Y' \cup (X \cap Y)$, we can add the first two equations and substract the third one to get the desired result. □

With these lemmas we are ready to prove Lemma B.2.

Proof of Lemma B.2. We will only show that f has a global maximum since, if X is a global maximum of f, then $\overline{X}$ is a global minimum (and viceversa). Let L be the supremum of $\{f(X)\}_{X\in\Sigma}$ and let $X_1, X_2, \ldots$ be a sequence of elements of Σ such that $f(X_n) > L - \frac{1}{2^n}$. Let $Y_m^n = \bigcap_{i=m}^{n} X_i$. By Lemma B.4 we have that $f(Y_m^n) > L - \sum_{i=m}^{n} \frac{1}{2^i} = L - (\frac{1}{2^{m-1}} - \frac{1}{2^n})$.

If we take the limit when $n \to \infty$, we have that $f(Y_m^\infty) > L - \frac{1}{2^{m-1}}$ (note that $Y_m^\infty \in \Sigma$ since Σ is closed under countable intersections). By construction, $Y_m^\infty \subseteq Y_{m+1}^\infty$ for all $m \in \mathbb{N}$. Moreover,

$$\lim_{m\to\infty} f(Y_m^\infty) = \lim_{m\to\infty} L - \frac{1}{2^{m-1}} = L.$$

By Lemma B.3, this means that

$$f\left(\bigcup_{m=1}^{\infty} Y_m^\infty\right) = L.$$

This completes the proof. □

The proof of Proposition 6 is analogous to that of Proposition 5. The key observation is that the set of maximal and the set of minimal bets are closed under countable unions and intersections, which means that we can do an analogue of the proof of Lemma B.2 restricted to $\Sigma_{Q_j^+}$ and to $\Sigma_{Q_j^-}$. To prove this, note that if f is a countably additive bounded function defined on a σ-algebra Σ, we have that $f(X) = f(X \cap Y) + f(X \setminus Y)$ and $f(Y) = f(X \cap Y) + f(Y \setminus X)$ for any two sets $X, Y \in \Sigma$. If Both X and Y maximize f, then $f(Y \setminus X) = f(X \setminus Y) = 0$. Otherwise, $f(X \cap Y)$ or $f(X \cup Y)$ would be strictly greater than $f(X)$. This implies that $f(X \cap Y) = f(X \cup Y) = f(X) = f(Y)$.

It is important to note that if we restrict the proof of Lemma B.2 to $\Sigma_{Q_j^+}$ and to $\Sigma_{Q_j^-}$, finding the minimum does not reduce to finding the maximum (as we did in Lemma B.2) since $\Sigma_{Q_j^+}$ and $\Sigma_{Q_j^-}$ are not closed under complements. However, we can reproduce an analogous proof using analogues to Lemma B.3 and Lemma B.4 for intersections and infimums, respectively. Finally, note that we can assume w.l.o.g. that the sets X^+ and X^- in Proposition 6 are disjoint since, if $X^+ \cap X^- \neq \emptyset$, then we can use the same argument as above to show that $Q'_j(X^{\cap}X^-) = Q_j(X^{\cap}X^-) = 0$. If this happens, we can take $X^+ \setminus X^-$ and $X^- \setminus X^+$ instead of X^+ and X^-.

It remains to show Lemma B.1.

Proof of Lemma B.1. We prove each equation separately.

First equation: $u_j(Q_{-S}, \tilde{Q}'_{S\setminus\{j\}}, \tilde{Q}''_j) - u_j(Q_{-S}, \tilde{Q}'_S) \leq Q'_j(X^+)M$**.**

For simplicity, let R denote the strategy profile where players in S play $\tilde{Q}'_S$ and the rest of the players play Q_{-S}.

By definition,

$$u_j(R_{-j}, \tilde{Q}''_j) = \int_A u_j(X)\ d(R_{-j}, \tilde{Q}''_j).$$

We can split A into $\Omega_1 := X^+ \times A_{-i}$ and $\Omega_2 := (A_i \setminus X^+) \times A_{-i}$ to get

$$u_j(R_{-j}, \tilde{Q}''_j) = \int_{\Omega_1} u_j(X)\ d(R_{-j}, \tilde{Q}''_j) + \int_{\Omega_2} u_j(X)\ d(R_{-j}, \tilde{Q}''_j).$$

By construction, $Q''_j(X) = 0$ for all $X \subseteq X^+$, which means that $\int_{\Omega_1} u_j(X)\ d(R_{-j}, \tilde{Q}''_j) = 0$. Additionally, if we define $\Omega_3 := X^- \times A_{-i}$, we have that

$$\int_{\Omega_2} u_j(X)\ d(R_{-j}, \tilde{Q}''_j) = \int_{\Omega_2} u_j(X)\ dR + \frac{\tilde{Q}'_j(X^+)}{Q_j(X^-)} \int_{\Omega_3} u_j(X)\ d(R_{-j}, Q_j).$$

Note that $u_j(R) = \int_{\Omega_2} u_j(X)\ dR + \int_{\Omega_1} u_j(X)\ dR$, which implies that

$$u_j(R_{-j}, \tilde{Q}''_j) - u_j(R) = \frac{\tilde{Q}'_j(X^+)}{Q_j(X^-)} \int_{\Omega_3} u_j(X)\ d(R_{-j}, Q_j) - \int_{\Omega_1} u_j(X)\ dR.$$

If $M_{\max}$ is the supremum of u_j and $M_{\min}$ is its infimum, we can bound $u_j(R_{-j}, \tilde{Q}''_j) - u_j(R)$ by

$$u_j(R_{-j}, \tilde{Q}''_j) - u_j(R) \leq \frac{\tilde{Q}'_j(X^+)}{Q_j(X^-)} M_{\max} \cdot (R_{-j}, Q_j)(\Omega_3) - M_{\min} \cdot R(\Omega_1).$$

Note, however, that $(R_{-j}, Q_j)(\Omega_3) = Q_j(X^-)$ and $R(\Omega_1) = \tilde{Q}'_j(X^+)$. Therefore,

$$u_j(R_{-j}, \tilde{Q}''_j) - u_j(R) \leq Q'_j(X^+)(M_{\max} - M_{\min}),$$

as desired.

Second equation: $\ell_j(Q_{-S}, Q'_{S\setminus\{j\}}, Q''_j) - \ell_j(Q_{-S}, Q'_S) > Q'_j(X^+) \cdot M$**.**

The second equation follows from the observation we made in the proof of Proposition 6: if (j, X) is a maximal bet on j, then $Q_j(X \setminus X^+) = Q'_j(X \setminus X^+) = 0$. Since $\tilde{Q}''_j(X^+) = 0$, it follows that $\ell_j(Q_{-S}, Q'_{S\setminus\{j\}}, Q''_j) = 0$. Thus, it remains to show that $\ell_j(Q_{-S}, Q'_S) < -Q'_j(X^+)\cdot M$. This follows from the same observation since then $\ell_j(Q_{-S}, Q'_S) = \mathbb{E}[\#\text{ of other players betting on j}]\cdot Q'_j(X^+)\cdot(-M-1)$. Since we assumed that at least one other player bets on j in expectation, it follows that $\ell_j(Q_{-S}, Q'_S) < -Q'_j(X^+)\cdot M$. □