

Maximizing Social Welfare with Side Payments

Abstract

We examine normal-form games in which players may *pre-commit* to outcome-contingent transfers before choosing their actions. In the one-shot version of this model, Jackson and Wilkie showed that side contracting can backfire: even a game with a Pareto-optimal Nash equilibrium can devolve into inefficient equilibria once unbounded, simultaneous commitments are allowed. The root cause is a prisoner’s dilemma effect, where each player can exploit her commitment power to reshape the equilibrium in her favor, harming overall welfare.

To circumvent this problem we introduce a *staged-commitment* protocol. Players may pledge transfers only in small, capped increments over multiple rounds, and the phase continues only with unanimous consent. We prove that, starting from any finite normal-form game with a non-degenerate Nash equilibrium $\vec{\sigma}$, this protocol implements every welfare-maximizing payoff profile that *strictly* Pareto-improves $\vec{\sigma}$. Thus, gradual and bounded commitments restore the full efficiency potential of side payments while avoiding the inefficiencies identified by Jackson and Wilkie.

1 Introduction

Nash equilibrium is a central tool in game theory, yet it abstracts away an important feature of many real-world settings: players often negotiate before they move. Outside of the “vacuum” assumed in normal-form games, agents can coordinate and steer away from inefficient outcomes, frequently by promising side payments that realign incentives and support mutually preferred strategies. Climate policy provides a canonical example: subsidies, rebates, and performance-based transfers compensate firms for adopting greener technologies. On digital platforms, “pre-play” incentives are equally common: marketplaces offer seller subsidies for meeting reliability targets, and employers use bonuses contingent on performance to implement behaviors that would not be an equilibrium without commitment/contracting. For a more concrete example, consider a two-player Prisoner’s Dilemma whose payoffs appear in the left panel of Example 1.

Example 1.

	C	D
C	(3, 3)	(0, 5)
D	(5, 0)	(1, 1)

	C	D
C	(3, 3)	(2, 3)
D	(3, 2)	(1, 1)

As usual, the unique Nash equilibrium is mutual defection (D, D) , even though (D, D) is Pareto dominated by mutual cooperation (C, C) . Moreover, a unilateral deviation to D can increase the deviator’s payoff by at most 2 while decreasing the other player’s payoff by at least 3. This gap suggests that it can be beneficial to commit to paying the other player 2 units if they cooperate: such a commitment removes the other player’s incentive to deviate. If both players make this commitment, incentives are reshaped and, in the modified game (right), (C, C) becomes a Nash equilibrium. One may wonder if this example can be generalized to arbitrary normal-form games, and if the inclusion of side payments always results in an increase of social welfare. The intuition is clear: if player i ’s gain from some action a is smaller than the loss it imposes on others, those

others can offer compensation below their loss yet sufficient to induce i to switch. Unfortunately, as shown by Jackson and Wilkie [35], such contracts are not necessarily best responses when all players can write them. In fact, the negotiation described above is not subgame perfect: if player 1 offers 2 utility to player 2 in exchange for cooperation, it is better for player 2 to offer $2 + \epsilon$ utility in (C, C) and only $(1 + \epsilon)$ utility in (C, D) , resulting in the table below. Since C is a dominant strategy for player 1, the only Nash equilibrium of the resulting game is (C, D) which, even though it is an inefficient outcome, it gives $4 - \epsilon$ utility to player 2 instead of just 3.

	C	D
C	$(3 + \epsilon, 3 - \epsilon)$	$(1 + \epsilon, 4 - \epsilon)$
D	$(3, 2)$	$(1, 1)$

Jackson and Wilkie's results are actually more negative. They showed that, even after including side payments, there is no subgame perfect strategy in Example 1 in which players converge on (C, C) (see [35]; Theorem 2). The root of the problem is that players can use these contracts selfishly to condition opponents' behavior. This can be seen clearly in games like the *game of chicken* (left table of Example 2), where player 1 can promise to pay 20 utility to player 2 if player 1 swerves, making going straight a dominant strategy for itself, and forcing player 2 to swerve instead (right table of Example 2).

Example 2.

	<i>Swerve</i>	<i>Straight</i>
<i>Swerve</i>	$(0, 0)$	$(0, 1)$
<i>Straight</i>	$(1, 0)$	$(-10, -10)$

	<i>Swerve</i>	<i>Straight</i>
<i>Swerve</i>	$(-20, 20)$	$(-20, 21)$
<i>Straight</i>	$(1, 0)$	$(-10, -10)$

The underlying issue is that side payments let players drastically reshape the game's equilibrium set without any safeguard against defection or selfish play during the commitment stage. We can therefore ask which safeguards guarantee efficiency. One natural device is a rollback option: after the commitment phase, a second pre-play stage is held in which each player votes either to accept the proposed side payments or to void them all. If at least one player votes to roll back, all contracts are annulled and the original game is played; otherwise, the transfers remain in force. Under this rollback model, one can implement every welfare-maximizing utility profile that Pareto-improves on at least one equilibrium of the base game. For instance, in Example 1, we can construct a subgame perfect equilibrium in which players end up playing (C, C) and get 3 utility each. Players proceed as follows.

1. During the commitment phase, offer to pay 2 utility to the other if they cooperate.
2. If players make the offers described above, play (C, C) . Otherwise, vote to void the contracts and play (D, D) .

We can verify that this strategy profile is indeed a subgame perfect equilibrium. Moreover, it is easy to check that the same construction can be generalized to any finite normal-form game for any number of players. The idea is quite simple: players must commit to pay each other the necessary amounts in order to converge to the desired utility profile since, otherwise, all commitments will be reversed and players will end up playing a Nash equilibrium that leaves everyone worse off (see Section 3 for a formal statement and proof.)

The argument above shows that rollback power is sufficient to avoid the inefficiencies associated with side payments. Unfortunately, having rollback power is often too unrealistic, as it presumes that players can nullify others' binding contracts. Ideally, we would want players to have the option to stop malicious commitments without being able to void existing ones. For this purpose, we propose

1.1 Related Work

Our work is closest to Jackson and Wilkie [35]. They study a single round of simultaneous, unbounded commitments before actions are chosen in the base game and do not allow money burning. Within this model, they characterize the set of supportable equilibrium payoffs and show that side payments can reduce social welfare: there are games with Pareto-efficient outcomes in which every equilibrium induced by side payments is inefficient. Sauerberg and Oosterheld [49] analyze a closely related setting in which players commit to transfers while simultaneously choosing their actions (rather than acting after observing others' commitments), and Gupta and Schewe [30] consider a similar two-player model where the first mover commits both to an action and to a *bribe* vector before the second moves. Although these modeling choices may appear minor, they lead to substantially different equilibrium sets from ours. In a more recent paper, Liu, Farina and Ozdaglar [41] consider a setting close to that of Jackson and Wilkie, based on *self-enforcing transfer equilibria* (SETE), in which each payment is voided if its recipient deviates from a prescribed strategy. For polymatrix games they show that any socially optimal profile is a SETE, and that every SETE is a Nash equilibrium of the *agent normal form* of the induced two-stage game. Efficiency in arbitrary finite games is achieved only by adding a mediator that makes payments and recommendations binding; in the general case, the obstruction found by Jackson and Wilkie still holds.

A separate line of work, beginning with Guttman [32, 31], Danziger and Schnytzer [16], Guttman and Schnytzer [33], and Varian [54, 53], achieves efficiency in public-good environments via matching schemes. The available contracts there are much more restrictive (for example, payments proportional to particular types of actions), and thus differ substantially from the side-payment flexibility we study.

Another closely related literature examines *disarmament games*. These also involve iterative commitments, but players directly shrink their own strategy sets rather than transferring utility. As in our setting, participants cannot reduce “too much” at once without becoming vulnerable, yet Deng and Conitzer [18, 19] show that gradual, small-step commitments can achieve the desired targets. This aligns with the results presented in this paper.

Payments used to shape outcomes also arise in several other directions. In mechanism design, payments are chosen to induce truthful reports, but the underlying game is designer-specific rather than arbitrary. For instance, in Bayesian allocation problems that maximize reported total value, the VCG mechanism [55, 13, 29] makes truthful reporting optimal. The principal-agent literature [39, 17, 50, 11, 20, 28, 42] studies settings in which one player (the principal) can commit to pay another player (the agent) in a way that depends on what action the agent takes. However, payments typically go only one way and this literature focuses on issues absent from our setting, such as imperfect observability of the agent's action or differing risk preferences between agent and principal. Meanwhile, other results involve commitments to outcome-dependent payments for arbitrary games. In Monderer and Tennenholtz's *k*-implementation paper [44], an outside party commits to outcome-dependent payments to the players in the game in order to transform a Nash equilibrium into a dominant strategy equilibrium. Their approach is similar to ours in the sense that these payments are always promised on outcomes outside of the resulting equilibrium, which means that the payer ultimately does not have to put her own resources on the line. In later work on *internal implementation* [4], the party with the ability to commit to such payments is modeled as one of the players in the game.

Many relaxations of Nash equilibrium introduce a *mediator*. The standard version communicates only via *signals*: correlated equilibrium [7] can be viewed as a mediator sending (possibly correlated) recommendations [40, 6, 5], and this perspective also underlies *Bayesian persuasion* [37, 5]. More general models allow multi-round communication with a mediator [21, 12, 2, 3, 1, 23, 25, 22]. Stronger mediators can do more than signal: empowering the mediator to act on players' behalf [47, 45] or to expand action sets [24] yields equilibria resilient to coalitional deviations. *Program equilibrium* [43, 34, 48, 52, 36, 46, 15, 14] can likewise be seen as mediated interaction enabling players to

commit to and share strategies. In a way, this is reminiscent of our setting since it requires players to commit (in this case, to a strategy instead of side payments) and allows players to converge on more efficient equilibria that would otherwise be impossible. In particular, it can be shown that, under this model, players can reach cooperation on the prisoner's dilemma. Relatedly, mediators that permit simulating opponents can produce similar effects [38]. In this light, our model can be interpreted as a mediated protocol: players submit side payments to a mediator, who can terminate the commitment phase (triggering play in the base game) and enforce transfers at the realized outcome.

The remainder of the paper is organized as follows. Section 2 formally introduces the model, notation, and basic definitions. Section 3 states our main results, and Section 4 provides proofs for the simpler cases. We conclude in Section 5. Additional technical material appears in Appendices B, C, D, and E.

2 Basic Definitions and Notation

2.1 Normal-form games

Given a normal-form game $\Gamma = (P, A, U)$, we denote the set of players by $P := \{1, 2, \dots, n\}$, the set of action profiles by $A := A_1 \times \dots \times A_n$, and the utility function by $U := (u_1, \dots, u_n)$, where $u_i : A \rightarrow \mathbb{R}$ denotes the utility function of player i . Unless stated otherwise, we will always label the player i 's actions as $A_i := \{a_i^1, a_i^2, \dots, a_i^{N_i}\}$ for some $N_i \in \mathbb{N}$. Given a strategy profile $\vec{\sigma}$, we denote by σ_i^j the probability that action a_i^j is played in σ_i , by $\sigma_i(A_i)$ the set of actions of player i that are played with non-zero probability in σ_i , by $u_i(\vec{\sigma})$ the expected utility that player i gets when $\vec{\sigma}$ is played, and by $w(\vec{\sigma})$ the expected social welfare of the players, which is defined as $w(\vec{\sigma}) := \sum_{i=1}^n u_i(\vec{\sigma})$. We also denote by $w_{\max}(\Gamma)$ the maximum social welfare of the game, which is defined as $w_{\max}(\Gamma) := \sup_{\vec{\sigma} \in \Delta A} \{w(\vec{\sigma})\}$. Note that, since $w(\vec{\sigma})$ is a weighted average of the social welfare of pure strategy profiles, the maximum is always achieved by a pure strategy profile. Thus, we have that $w_{\max}(\Gamma) := \max_{\vec{a} \in A} \{w(\vec{a})\}$.

Given two normal form games $\Gamma = (P, A, U)$ and $\Gamma' = (P', A', U')$, we define their distance $d(\Gamma, \Gamma')$ as follows:

$$d(\Gamma, \Gamma') := \begin{cases} +\infty & \text{if } P \neq P' \text{ or } A \neq A' \\ \max_{i \in P, \vec{a} \in A} \{|u_i(\vec{a}) - u'_i(\vec{a})|\} & \text{otherwise.} \end{cases}$$

Intuitively, the distance between normal-form games is infinite if they have different structures (i.e., different number of players or different action sets). Otherwise, it is the maximum difference of utilities that a player can receive on a single outcome. Note that this means that we can view the set of normal-form games as a metric space, and therefore as a topological space too. For the rest of the paper, we will be implicitly assuming that the metric and the topology of the space of normal-form games that we consider are the ones induced by the distance d above.

2.2 δ -commitment games

Given a normal-form game Γ and a strictly positive real number δ , we denote by Γ^δ the δ -commitment extension of Γ . An instance of Γ^δ runs as follows.

- **Step 1:** Each player decides how much utility it commits to pay for each possible outcome. The recipients of the paid utility can be either other players or $\perp$ (the utility in this case is

burned and is not given to other players). The total utility paid in each outcome cannot exceed δ .

- **Step 2:** The payments are made public and the utility functions $u_1, \dots, u_n$ are updated according to the promised payments.
- **Step 3:** Each player votes to end the commitment phase or not. If all players agree on continuing with the commitment phase, they go back to Step 1. Otherwise, they proceed to Step 4.
- **Step 4:** Each player i plays an action a_i in the resulting game and gets $u_i(a_1, \dots, a_n)$ utility (note that these utilities are updated every time that Step 1 is performed). The game ends.

In Γ^δ , a strategy σ_i for player i indicates, given the current state of the game, which payments i should make in each round of the commitment phase, if it should vote to end the commitment phase or not, and what action to play in Step 4. We say that a strategy profile $\vec{\sigma}$ is a δ -*commitment equilibrium* of Γ if $\vec{\sigma}$ is a subgame perfect equilibrium of Γ^δ . For future reference, we consider also an extension $\Gamma_\perp^\delta$ of Γ which is identical to Γ^δ except that players can only burn utility in Step 2 (they cannot transfer it to other players), and a rollback variant Γ_R in which there is a single round of unbounded commitments followed by a vote where, if any player votes against the commitments, they are all voided and players proceed to play the game unaltered. We say that a strategy profile $\vec{\sigma}$ is a $(\delta, \perp)$ -*commitment equilibrium* of Γ (resp., R -*commitment equilibrium* of Γ) if $\vec{\sigma}$ is a subgame perfect equilibrium of $\Gamma_\perp^\delta$ (resp., Γ_R).

2.3 Punishable and Non-Degenerate Equilibria

At a high level, our aim in this paper is proving that, given a normal-form game Γ , we can improve any Nash equilibrium that does not attain maximum social welfare by extending it to a δ -commitment game, for some value $\delta \in \mathbb{R}$. Unfortunately, this result does not quite hold for all games. For instance, consider a normal-form game for two players with the following payoff matrix.

	A	B	C
A	(5, 5)	(0, 5)	(0, 0)
B	(5, 0)	(9, 2)	(7, 1)
C	(0, 0)	(1, 7)	(6, 6)

In this example, (A, A) is a Nash equilibrium that gives 5 utility to both players and (C, C) is a welfare-maximizing outcome that gives 6 utility instead, but it is not a Nash equilibrium. Ideally, we would like players to converge on (C, C) . However, consider a δ -commitment extension of Γ and suppose that player 1 commits to pay δ utility to player 2 in (A, B) and δ utility to $\perp$ in (A, A) . Then, regardless of what player 2 does, action B will be a dominant action for at least one of the players (it will be for player 1 unless player 2 pays δ utility to player 1 in (A, A) , but in this case it will be for player 2). This means that, if player 1 performs this move and then stops the commitment phase right after, she can guarantee that the only Nash equilibrium of the resulting game will be (B, B) , which gives her at least 9 utility. Since player 1 can never expect to get more than 9 utility in (C, C) (even with side payments), it is always optimal for player 1 to run this strategy. This gives 2 utility to player 2 and, therefore, the expected utility that player 2 gets with commitments is strictly less than without commitments.

The issue with the previous example is that player 1 can make the (A, A) equilibrium disappear in her first move, forcing player 2 to converge on (B, B) . For instance, if $\delta < 0.25$ and (A, A) gave 5.5 utility to each player instead, players can play a strategy in which they decrease by δ all utilities except the ones in (A, A) and (C, C) in every step of the commitment phase until (C, C) becomes

a Nash equilibrium. If any player does something different, they stop the commitment phase and play (A, A) . It is straightforward to check that such a strategy is a subgame perfect equilibrium. Since (A, A) remains an equilibrium even if a player defects, the threat of playing (A, A) persists and deters any player from defecting.

This analysis suggests that, for the commitment process to work as intended, it is necessary that players cannot modify the original equilibrium with a single move. In particular, we require the following notion of equilibrium.

Definition 1 (Punishable equilibrium). *Let $\Gamma = (P, A, U)$ be a normal-form game, $\vec{\sigma}$ be a Nash equilibrium of Γ , and let ϵ and δ be two strictly positive real numbers. We say that $\vec{\sigma}$ is an (ϵ, δ) -punishable equilibrium if, whenever $d(\Gamma, \Gamma') < \delta$ for some game Γ' , there exists a Nash equilibrium $\vec{\sigma}'$ of Γ' such that $u_i(\vec{\sigma}') - u_i(\vec{\sigma}) < \epsilon$ for all $i \in P$. If furthermore there exists a Nash equilibrium $\vec{\sigma}'$ that also satisfies that $\sigma_i(A_i) = \sigma'_i(A_i)$, we say that $\vec{\sigma}$ is (ϵ, δ) -strongly punishable. Moreover, we say that $\vec{\sigma}$ is punishable (resp., strongly punishable) if, for all $\epsilon > 0$ there exists a $\delta > 0$ such that $\vec{\sigma}$ is (ϵ, δ) -punishable (resp., (ϵ, δ) -strongly punishable).*

Intuitively, a Nash equilibrium $\vec{\sigma}$ is punishable if, after applying small perturbations on the utilities of the game, we can always find an equilibrium $\vec{\sigma}'$ in the new game such that the utilities obtained by the players are not much higher than in $\vec{\sigma}$ (played in the original game). We will refer to these strategy profiles $\vec{\sigma}'$ as *punishment equilibria*. Since it is generally hard to determine if a given Nash equilibrium is punishable, we will restrict our analysis to equilibria that satisfy a much (computationally) simpler property that implies strong punishability. We call this property *non-degeneracy*, and the vast majority of equilibria satisfy it. To introduce this concept and to show why it implies strong punishability, it is better to start with an example. Consider the following game for two players in which $A_1 = A_2 = \{a_1, a_2, a_3\}$ and in which the payoff matrix is as follows.

	a_1	a_2	a_3
a_1	(5, 5)	(1, 1)	(1, 0)
a_2	(1, 1)	(5, 5)	(0, 1)
a_3	(0, 1)	(1, 0)	(2, 2)

Suppose that we want to determine if there exists a Nash equilibrium $\vec{\sigma}$ such that $\sigma_1(A_1) = \sigma_2(A_2) = \{a_1, a_2\}$. Then, $\vec{\sigma}$ must satisfy the following system of equations:

$$\begin{aligned}
\sigma_1^1 + \sigma_1^2 &= 1 && \text{(Probabilities must add up to 1.)} \\
\sigma_2^1 + \sigma_2^2 &= 1 && \text{(Probabilities must add up to 1.)} \\
u_1(a_1, \sigma_2) - u_1(a_2, \sigma_2) &= 0 && \text{(Player 1 is indifferent between } a_1 \text{ and } a_2.) \\
u_1(a_1, \sigma_2) - u_1(a_3, \sigma_2) &\geq 0 && \text{(Player 1 weakly prefers } a_1 \text{ to } a_3.) \\
u_2(a_1, \sigma_1) - u_2(a_2, \sigma_1) &= 0 && \text{(Player 2 is indifferent between } a_1 \text{ and } a_2.) \\
u_2(a_1, \sigma_1) - u_2(a_3, \sigma_1) &\geq 0 && \text{(Player 2 weakly prefers } a_1 \text{ to } a_3.)
\end{aligned}$$

This system is composed by four linear equations and two linear inequalities. In particular, the four linear equations can be written in matricial form as

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 4 & -4 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 4 & -4 \end{pmatrix} \begin{pmatrix} \sigma_1^1 \\ \sigma_1^2 \\ \sigma_2^1 \\ \sigma_2^2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

This system gives a unique solution in which $\sigma_1^1 = \sigma_1^2 = \sigma_2^1 = \sigma_2^2 = \frac{1}{2}$. It is easy to check that this solution also strictly satisfies the two inequalities, and therefore that it is indeed a Nash equilibrium.

Suppose now that we add small perturbations to the utilities of the players. If these perturbations are small enough and we repeat the same process to find a similar Nash equilibrium, the resulting matrix will still be invertible, its unique solution will still strictly satisfy the two inequalities, and the expected utilities of the players will not be too far off from the original ones. This shows that $\vec{\sigma}$ is a strongly punishable Nash equilibrium.

Note that the two main ingredients that we need in order to argue about the strong punishability of $\vec{\sigma}$ are the fact that (a) the matrix obtained from the system of equations in order to compute $\vec{\sigma}$ is invertible, and (b) that the inequalities are **strictly** satisfied (i.e., with no equality). We can generalize this idea for any two-player game, and later for multi-player games. Consider a normal-form game $\Gamma = (P, A, U)$ for two players and a Nash equilibrium $\vec{\sigma}$. Without loss of generality assume that $\sigma_1(A_1) = \{a_1^1, a_1^2, \dots, a_1^{M_1}\}$ and $\sigma_2(A_2) = \{a_2^1, a_2^2, \dots, a_2^{M_2}\}$. Then, we define the characteristic function of $\vec{\sigma}$ as

$$f_{\vec{\sigma}}(p_1^1, p_1^2, \dots, p_1^{M_1}, p_2^1, p_2^2, \dots, p_2^{M_2}) := \begin{cases} p_1^1 + p_1^2 + \dots + p_1^{M_1} \\ p_2^1 + p_2^2 + \dots + p_2^{M_2} \\ u_1(a_1^1, \vec{p}_2) - u_1(a_1^2, \vec{p}_2) \\ u_1(a_1^1, \vec{p}_2) - u_1(a_1^3, \vec{p}_2) \\ \vdots \\ u_1(a_1^1, \vec{p}_2) - u_1(a_1^{M_1}, \vec{p}_2) \\ u_2(a_2^1, \vec{p}_1) - u_2(a_2^2, \vec{p}_1) \\ u_2(a_2^1, \vec{p}_1) - u_2(a_2^3, \vec{p}_1) \\ \vdots \\ u_2(a_2^1, \vec{p}_1) - u_2(a_2^{M_2}, \vec{p}_1), \end{cases}$$

where we identify $\vec{p}_i$ as the strategy in which player i plays action 1 with probability p_i^1 , plays action 2 with probability p_i^2 , etc. We also define the residual function of $\vec{\sigma}$ as

$$r_{\vec{\sigma}}(p_1^1, p_1^2, \dots, p_1^{M_1}, p_2^1, p_2^2, \dots, p_2^{M_2}) := \begin{cases} u_1(a_1^1, \vec{p}_2) - u_1(a_1^{M_1+1}, \vec{p}_2) \\ u_1(a_1^1, \vec{p}_2) - u_1(a_1^{M_1+2}, \vec{p}_2) \\ \vdots \\ u_1(a_1^1, \vec{p}_2) - u_1(a_1^{N_1}, \vec{p}_2) \\ u_2(a_2^1, \vec{p}_1) - u_2(a_2^{M_2+1}, \vec{p}_1) \\ u_2(a_2^1, \vec{p}_1) - u_2(a_2^{M_2+2}, \vec{p}_1) \\ \vdots \\ u_2(a_2^1, \vec{p}_1) - u_2(a_2^{N_2}, \vec{p}_1) \end{cases}$$

As in the example, $f_{\vec{\sigma}}$ is a linear function from $[0, 1]^{M_1+M_2}$ to $\mathbb{R}^{M_1+M_2}$. It is straightforward to check that condition (a) from the example is equivalent to $|Df_{\vec{\sigma}}(\vec{\sigma})| \neq 0$ (where $|Df_{\vec{\sigma}}(\vec{\sigma})|$ denotes the determinant of the Jacobian of $f_{\vec{\sigma}}$ evaluated at $\vec{\sigma}$), and condition (b) is equivalent to $r_{\vec{\sigma}}(\vec{\sigma}) > 0$ (i.e., all its components are strictly positive when evaluated at $\vec{\sigma}$). This gives the following definition.

Definition 2 (Non-Degenerate Equilibrium). *If Γ is a normal-form game and $\vec{\sigma}$ is a Nash equilibrium of Γ , we say that $\vec{\sigma}$ is non-degenerate if $|Df_{\vec{\sigma}}(\vec{\sigma})| \neq 0$ and $r_{\vec{\sigma}}(\vec{\sigma}) > 0$.*

Note that we haven't specified that Γ must be a two-player game. This is because the construction of $f_{\vec{\sigma}}$ and $r_{\vec{\sigma}}$ can be easily generalized for any number of players. Consider a normal-form game $\Gamma = (P, A, U)$ for n players and a Nash equilibrium $\vec{\sigma}$ for Γ . Suppose that $\sigma_i(A_i) = \{a_i^1, \dots, a_i^{M_i}\}$. Then,

$$f_{\vec{\sigma}}(p_1^1, \dots, p_n^{M_n}) := \begin{cases} p_1^1 + p_1^2 + \dots + p_1^{M_1} \\ p_2^1 + p_2^2 + \dots + p_2^{M_2} \\ \vdots \\ p_n^1 + p_n^2 + \dots + p_n^{M_n} \\ u_1(a_1^1, \vec{p}_{-1}) - u_1(a_1^2, \vec{p}_{-1}) \\ u_1(a_1^1, \vec{p}_{-1}) - u_1(a_1^3, \vec{p}_{-1}) \\ \vdots \\ u_n(a_n^1, \vec{p}_{-n}) - u_n(a_n^{M_n}, \vec{p}_{-n}), \end{cases}$$

where $\vec{p}_{-i}$ denotes the strategy profile in which each player $j \neq i$ plays a_k^j with probability p_k^j . We can also define $r_{\vec{\sigma}}$ as

$$r_{\vec{\sigma}}(p_1^1, \dots, p_{M_n}^n) := \begin{cases} u_1(a_1^1, \vec{p}_{-1}) - u_1(a_1^{M_1+1}, \vec{p}_{-1}) \\ u_1(a_1^1, \vec{p}_{-1}) - u_1(a_1^{M_1+2}, \vec{p}_{-1}) \\ \vdots \\ u_n(a_n^1, \vec{p}_{-n}) - u_n(a_n^{N_n}, \vec{p}_{-n}). \end{cases}$$

If $n > 2$, the components of both $f_{\vec{\sigma}}$ and $r_{\vec{\sigma}}$ are no longer linear. However, perhaps surprisingly, non-degeneracy implies strong punishability even in this case (see Appendix E for the proof.)

Theorem 1. *If Γ is a normal-form game and $\vec{\sigma}$ is a non-degenerate Nash equilibrium of Γ , then $\vec{\sigma}$ is strongly punishable.*

3 Main Results

The aim of this paper is to determine if the inclusion of sequential capped commitments allows players to converge on efficient and fair equilibria. As discussed in the introduction, we will use as benchmark the rollback model, whose set of implementable outcomes can be characterized as follows.

Theorem 2. *Let $\Gamma = (P, A, U)$ be a normal-form game for n players and let $\vec{\sigma}$ be a Nash equilibrium of Γ . Let $x_1, \dots, x_n$ be n real numbers such that $\sum_{i=1}^n x_i = w_{\max}(\Gamma)$ and such that $x_i > u_i(\vec{\sigma})$ for all $i \in P$. Then, there exists an R -commitment equilibrium $\vec{\tau}$ of Γ such that $u_i(\vec{\tau}) = x_i$ for all $i \in P$.*

Proof. Let $\vec{a}$ be a welfare-maximizing outcome of Γ , and let $\vec{t}$ be a commitment profile such that, after applying $\vec{t}$, (a) each player i receives exactly x_i utility on $\vec{a}$, and (b) each player i receives strictly less utility than x_i anywhere else. Consider the full strategy profile in which each player i must place commitments according to t_i . If all players do so, they play $\vec{a}$. Otherwise, they vote to rollback all commitments and play $\vec{\sigma}$. It is easy to check that this strategy profile is subgame perfect: each player i must commit to t_i since otherwise she will get $u_i(\vec{\sigma})$ utility, which is strictly worse. By construction, after committing to $\vec{t}$, $\vec{a}$ is a Nash equilibrium. $\square$

Note that Theorem 2 implies that, under the rollback model, we can implement any utility profile that (a) achieves maximum social welfare, and (b) Pareto improves the expected utilities of any Nash equilibrium $\vec{\sigma}$ of the original game. It turns out that the set of welfare-maximizing Pareto-improving profiles that are implementable under sequential commitments is identical to the set

of utility profiles implementable under the rollback model whenever the original equilibrium $\vec{\sigma}$ is *non-degenerate*³. This is the main contribution of this paper.

Theorem 3. *Let $\Gamma = (P, A, U)$ be a normal-form game for n players and let $\vec{\sigma}$ be a non-degenerate Nash equilibrium of Γ . Let $x_1, \dots, x_n$ be n real numbers such that $\sum_{i=1}^n x_i = w_{\max}(\Gamma)$ and such that $x_i > u_i(\vec{\sigma})$ for all $i \in P$. Then, for some $\delta > 0$, there exists a δ -commitment equilibrium $\vec{\tau}$ of Γ such that $u_i(\vec{\tau}) = x_i$ for all $i \in P$.*

Theorem 3 is stated for non-degenerate Nash equilibria. This is a standard regularity condition: for finite games it holds generically under arbitrarily small payoff perturbations. Non-degeneracy is used to guarantee robustness of the equilibrium and of a fixed set of strict inequalities along the payoff perturbations induced by commitments. In degenerate games these inequalities may be tight and equilibria may fail to vary smoothly with payoffs, so additional case distinctions (or different techniques) would be required.

Our final result states that, if players cannot commit to payments between them and can only burn money, they can still converge on all outcomes that Pareto-improve their utility (even though these outcomes may not always exist).

Theorem 4. *Let $\Gamma = (P, A, U)$ be a normal-form game for n players, let $\vec{\sigma}$ be a Nash equilibrium of Γ , and let $\vec{a}^{PI} \in A$ be an outcome of Γ that strictly Pareto improves $\vec{\sigma}$. Then, there exists $D > 0$ such that, for all $\delta \in (0, D)$, there exists a $(\delta, \perp)$ -commitment equilibrium of Γ in which the utility of $\vec{a}^{PI}$ is unchanged and players end up playing $\vec{a}^{PI}$ with probability 1.*

Note that in this setting, even if Theorem 4 shows that players can still strictly Pareto improve their utility whenever there exists an outcome that does so, it may happen that all the welfare-maximizing outcomes do not strictly Pareto improve the original equilibrium. One may wonder why Theorem 4 requires *strict* Pareto improvements and why Theorem 3 requires strict inequalities. It turns out that, without strict gains in utility, players are not always incentivized to cooperate. For instance, consider the following normal-form game:

	A	B	C
A	(3, 2)	(0, 3)	(0, 0)
B	(2, 1)	(1, 0)	(2, 0)
C	(0, 6)	(0, 0)	(1.5, 5)

Since action C is dominated by actions A and B for both players, it is easy to check that both players uniformly randomizing between A and B is the only Nash equilibrium of the game. Moreover, this equilibrium is non-degenerate and gives player 1 $\frac{3}{2}$ utility. However, decreasing x to player 1's utility in (A, A) yields 1 an expected utility of $\frac{3-x}{2-x}$ in the resulting equilibrium. Therefore, if $\delta < 1$, player 2 cannot break the domination of A and B over C in a single move, and it cannot make B dominate A or vice-versa. Therefore, player 1 can burn $\delta/2$ utility in (A, A) and end the commitment phase immediately afterwards, guaranteeing strictly more utility than in (C, C) or in the original equilibrium of the game.

3.1 Tractability and Computability of the Solutions

The protocols underlying Theorem 3 and Theorem 4 use $O(n \cdot \delta^{-1} U_{\max} A_{\max})$ commitment rounds, where n is the number of players, $U_{\max}$ is the maximum utility attainable in the original game, and $A_{\max} = \max_{i \in P} |A_i|$. Fixing δ , our construction is explicit on the equilibrium path. Off the

³Informally, non-degeneracy means that the equilibrium is isolated and the relevant best-response inequalities are strict; see Section 2.3.

equilibrium path, the strategy profile invokes a punishment equilibrium after a deviation. Such a punishment equilibrium is guaranteed to exist and to have the same support as the original equilibrium $\vec{\sigma}$; thus its computation reduces to solving a system of multivariate polynomial equations.

On the choice of δ : Our proofs establish the existence of a positive step size δ via the implicit function theorem. However, they do not in general yield a tight or easily computable value. This difficulty is inherent: δ acts as a *robustness radius* that must simultaneously ensure (i) that the relevant equilibrium remains non-degenerate and (ii) that a prescribed collection of strict best-response inequalities stays strict along the constructed path of perturbed games. Certifying a maximal admissible δ therefore requires global control over how the game equilibria vary with payoffs, which is not generally available in closed form.

That said, computing the punishment equilibria in the special case of two-player games reduces to solving a system of linear equations. This allows us to provide an explicit (albeit conservative and highly convoluted) lower bound on the step size δ . See Appendix F for more details.

The proofs of the main theorems are done in the following order. First we prove Theorem 4 in the case where $\vec{\sigma}$ has partial support (Section 4.1). Then, we prove it for two-player games where $\vec{\sigma}$ has full support and each player has at least three actions (Section 4.2.1), followed by a proof for arbitrary normal-form games (Appendix B) and for two-player games with binary actions (Appendix C). We then extend these techniques to prove Theorem 3 as well (Appendix D).

4 Proof of Theorem 4

For simplicity, we will prove a weaker version of Theorem 4 that states that, given Γ , $\vec{\sigma}$ and $\vec{a}^{PI}$, there exists $\delta > 0$ and a $(\delta, \perp)$ -commitment equilibrium of Γ in which the utility of $\vec{a}^{PI}$ is unchanged and players end up playing $\vec{a}^{PI}$ with probability 1. It is easy to check that the same constructions that we present in this section are also valid for all smaller values of δ as well.

As mentioned in Section 1, the key idea for proving Theorem 4 is showing that we can always find a $\delta > 0$ and a commitment protocol $\vec{\pi}^\delta$ in $\Gamma_\perp^\delta$ in which (a) after each commitment step, for all perturbations that modify the entries of the payoff matrix by at most δ , there exists a punishment equilibrium in the resulting game that gives each player less utility in expectation than $\vec{a}^{PI}$, and (b) eventually $\vec{a}^{PI}$ becomes a Nash equilibrium. If we can construct such a protocol, we can easily extend it to a subgame perfect equilibrium $\vec{\tau}^{\vec{\pi}^\delta}$ in $\Gamma_\perp^\delta$. In this extension, each player i does as follows. At each step of the commitment phase, i runs π_i^δ to perform its commitments. After the commitment step, if $\vec{a}^{PI}$ is a Nash equilibrium, i votes to stop the commitment phase and plays a_i^{PI} . Otherwise, there are two options. If no player defected during the previous commitment step, i votes to continue the commitment phase and repeats this process. Otherwise, i votes to stop the commitment phase and plays its part of the punishment equilibrium described in (a). If a player ever votes to stop the commitment phase before $\vec{a}^{PI}$ becomes a Nash equilibrium, i plays its part of the punishment equilibrium. To check that this is indeed a Nash equilibrium, note that a player can only defect in the following ways:

- By defecting from π during the commitment phase.
- By voting to continue the commitment phase when it should vote to stop.
- By voting to stop when it should vote to continue the commitment phase.
- By playing a different action after the commitment phase.

If i defects from π during the commitment phase, it guarantees that the remaining players will vote to stop the commitment phase right afterwards and play their part of the punishment equilibrium,

which always exists because of property (a). This ensures that i gets a lower utility than in $\vec{\tau}^{\pi^\delta}$. This also happens if i votes to stop the commitment phase before $\vec{a}^{PI}$ becomes a Nash equilibrium. If i votes to continue whenever it should vote to stop, all the remaining players will vote to stop anyways and i 's vote will not matter. To see that i does not benefit from playing a different action note that, regardless of what happens during the game, the remaining players will end up playing either their part of the punishment equilibrium or their part of $\vec{a}^{PI}$. Since both are Nash equilibria when played, it is a best response for i to play its part too. Altogether this shows the following.

Lemma 1. *If π^δ is a commitment protocol in $\Gamma_\perp^\delta$ that satisfies properties (a) and (b), then it can be extended to a subgame perfect equilibrium $\vec{\tau}^{\pi^\delta}$ of $\Gamma_\perp^\delta$.*

In some cases (namely the ones with partial support), we will design commitment protocols that satisfy a property that is slightly stronger than (a). Let $L_\vec{\sigma}^{\vec{a}^{PI}}$ be the minimum additional utility that a player gets from playing $\vec{a}^{PI}$ instead of playing $\vec{\sigma}$. More precisely,

$$L_\vec{\sigma}^{\vec{a}^{PI}} := \min_{i \in P} \{u_i(\vec{a}^{PI}) - u_i(\vec{\sigma})\}.$$

We define the property (a1) as follows.

(a1) $\vec{\sigma}$ remains a $(L_\vec{\sigma}^{\vec{a}^{PI}}, \delta)$ -strongly punishable Nash equilibrium after each commitment step.

It is easy to check that (a1) implies (a). In fact, if a commitment protocol satisfies (a1), it means that $\vec{\sigma}$ is a Nash equilibrium throughout the whole process. Moreover, since the utilities on given outcomes can only decrease, the expected utilities of the players when playing $\vec{\sigma}$ can only decrease as well. This means that the maximum utility that a player i can get is bounded by $u_i(\vec{\sigma}) + L_\vec{\sigma}^{\vec{a}^{PI}}$, which by definition is lower than $u_i(\vec{a}^{PI})$. This gives the following Corollary.

Corollary 1. *If π^δ is a commitment protocol in $\Gamma_\perp^\delta$ that satisfies properties (a1) and (b), then it can be extended to a subgame perfect equilibrium $\vec{\tau}^{\pi^\delta}$ of $\Gamma_\perp^\delta$.*

To see how to construct such a protocol π^δ we will analyze a couple of simple examples that cover the cases in which $\vec{\sigma}$ has partial support, which means that there exist action profiles that are played with probability 0. Some of the techniques used in these examples will be later extended in Section 4.2 for arbitrary normal-form games and non-degenerate Nash equilibria.

4.1 Cases with partial support

We start this case analysis with a couple of simple examples that can be easily generalized. Consider a normal-form game Γ for two players in which the action set of each player is $\{a_1, a_2, a_3, a_4\}$ and the utility matrix is as follows.

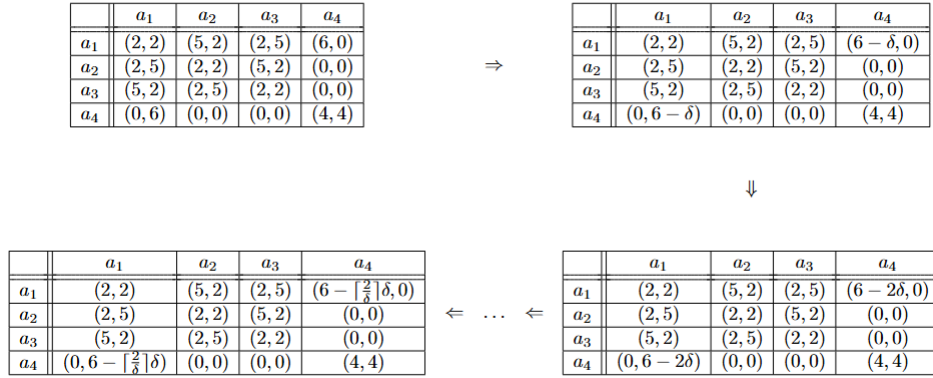
Example 3.

	a_1	a_2	a_3	a_4
a_1	(2, 2)	(5, 2)	(2, 5)	(6, 0)
a_2	(2, 5)	(2, 2)	(5, 2)	(0, 0)
a_3	(5, 2)	(2, 5)	(2, 2)	(0, 0)
a_4	(0, 6)	(0, 0)	(0, 0)	(4, 4)

It is easy to check that the strategy profile $\vec{\sigma}$ in which both players choose at random between a_1, a_2 and a_3 is a non-degenerate Nash equilibrium that gives 3 utility to each player in expectation. To see

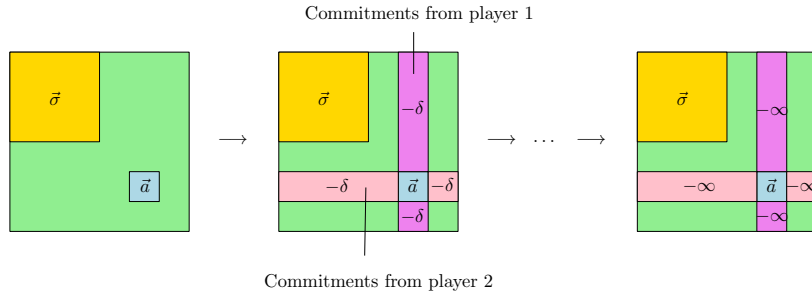
this, note that the game obtained by erasing a_4 from the action set of both players is equivalent to rock-paper-scissors, that the utility that players get by playing a_4 when the other player randomizes between a_1, a_2 and a_3 is 2 (which is strictly less than the utility they would get randomizing), and that Γ satisfies the properties of Definition 2. Therefore, by Theorem 1, there exists a $\delta > 0$ such that $\vec{\sigma}$ is $(1, \delta)$ -strongly punishable, which means that any change made to the payoff matrix in which utilities change at most δ results in a game where there exists a Nash equilibrium with the same support as $\vec{\sigma}$ in which no player gets more than 4 utility.

Hence, if players want to converge on (a_4, a_4) in $\Gamma_{\perp}^{\delta}$, they can make commitments following the diagram below and play (a_4, a_4) afterwards.



Since $f_{\vec{\sigma}}$ is preserved in each commitment step and $r_{\vec{\sigma}}$ decreases, we can deduce that $\vec{\sigma}$ satisfies property (a1). Moreover, it satisfies property (b) by constructions. Therefore, by Corollary 1, this commitment protocol can be extended to a subgame perfect equilibrium.

The main insight used for this argument is the fact that decreasing the values outside of the support of $\vec{\sigma}$ and $\vec{a}^{PI}$ results in a game that satisfies (a1) and (b). This can be generalized to any normal-form game in which $\vec{a}^{PI}$ satisfies that $\sigma_i^{PI} = 0$ for all players i (i.e., when the probability of playing a_i^{PI} with σ_i is 0 for all $i \in P$). In this case, during each commitment step, each player i should decrease by δ the utility of all outcomes in $(A \setminus \{a_i^{PI}\}) \times \{\vec{a}_{-i}^{PI}\}$ until $\vec{a}^{PI}$ becomes a Nash equilibrium. This is illustrated in the diagram below, where $-\infty$ represents a sufficiently negative value.



A slightly more complicated case occurs when $\vec{a}^{PI}$ is not in the support of $\vec{\sigma}$ but, for some player i , a_i^{PI} is in the support of σ_i . For example, consider a small variant of Example 3.

Example 4.

	a_1	a_2	a_3	a_4
a_1	(2, 2)	(5, 2)	(2, 5)	(2, 2)
a_2	(2, 5)	(2, 2)	(5, 2)	(0, 0)
a_3	(5, 2)	(2, 5)	(2, 2)	(1, 1)
a_4	(2, 2)	(2.5, 2)	(4, 4)	(0, 0)

As in Example 3, the strategy profile $\vec{\sigma}$ in which each player chooses between a_1, a_2 and a_3 uniformly at random is a Nash equilibrium. However, if players want to converge on (a_4, a_3) they must be careful. Taking the same approach as in Example 3 could potentially “break” $\vec{\sigma}$ since some of the commitments would lie in its support. Nevertheless, the following commitment preserves $f_{\vec{\sigma}}$.

	a_1	a_2	a_3	a_4
a_1	(2, 2)	(5, 2)	(2, 5)	(2, 2)
a_2	(2, 5)	(2, 2)	(5, 2)	(0, 0)
a_3	(5, 2)	(2, 5)	(2, 2)	(1, 1)
a_4	(2, 2)	(2.5, 2)	(4, 4)	(0, 0)

→

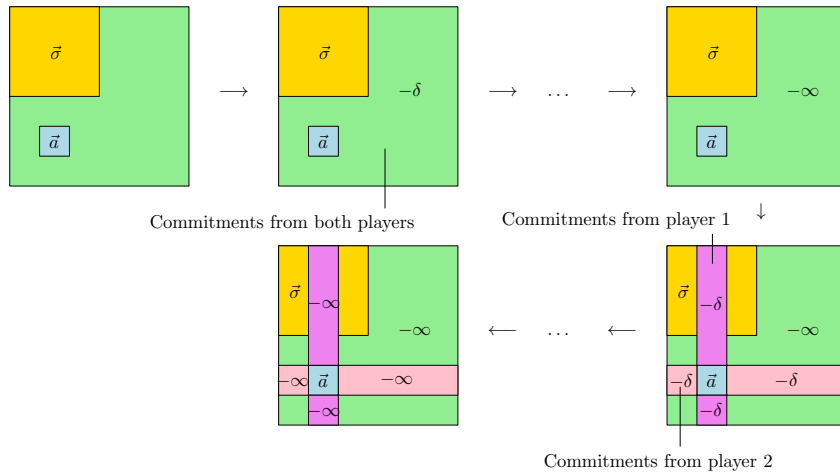
	a_1	a_2	a_3	a_4
a_1	(2, 2)	(5, 2)	(2 - δ , 5)	(2, 2)
a_2	(2, 5)	(2, 2)	(5 - δ , 2)	(0, 0)
a_3	(5, 2)	(2, 5)	(2 - δ , 2)	(1, 1)
a_4	(2, 2)	(2.5, 2)	(4, 4)	(0, 0)

Indeed, the same approach as in Example 3 would work if it was not the case that $r_{\vec{\sigma}}$ increases with each iteration. In fact, if $\delta > 0.5$ we would have that $u_1(a_4, \sigma_2) > u_1(\vec{\sigma})$ and therefore that $\vec{\sigma}$ is no longer a Nash equilibrium after just one iteration. To fix this issue, the players should first start by decreasing all their utilities outside of the support of $\vec{\sigma}$ except for (a_4, a_3) . This way, they decrease all the components of $r_{\vec{\sigma}}$ to the point in which $r_{\vec{\sigma}}$ remains negative after running the whole procedure used in Example 3. For instance, in Example 4, it suffices that player 1 decreases by 1 the value of $u_1(a_4, a_1), u_1(a_4, a_2)$ and $u_1(a_4, a_4)$ (this may not be done necessarily in a single step) before running the procedure used in Example 3.

In short, the commitment protocol in this case goes as follows.

1. First, each player i decreases its utility in all outcomes outside of $\vec{a}^{PI}$ and outside of the support of $\vec{\sigma}$. It repeats this process until all such values are sufficiently low.
2. Then, each player i decreases its utility in all outcomes in $(A_i \setminus \{a_i\}) \times \vec{a}_{-i}$.

An analogous reasoning to the one used in the analysis of Example 3 shows that this protocol satisfies properties (a1) and (b), and therefore it can also be extended to a subgame perfect equilibrium. In the case of two players, this is illustrated in the diagram below.



The remaining case is the one in which $\vec{a}^{PI}$ is in the support of $\vec{\sigma}$, but $\vec{\sigma}$ does not have full support. Assume without loss of generality that $|A_i| \geq 2$ for all players i (we can view this as ignoring all the players that cannot choose). With this assumption, there exists an action profile $\vec{a}$ such that $a_i \neq a_i^{PI}$ for all players i and such that $\vec{a}$ is not in the support of $\vec{\sigma}$. Then, players can commit as follows.

1. First, they decrease their utilities in $\vec{a}$ until $\vec{a}^{PI}$ strictly Pareto-dominates $\vec{a}$.
2. Then, they run the protocol used in the analysis of Example 4 in order to make $\vec{a}$ a Nash equilibrium while satisfying properties (a1) and (b). During the second part of the protocol, each player i decreases the utilities in $(A_i \setminus \{a_i^{PI}\}) \times \vec{a}_{-i}^{PI}$ until $u_i(\vec{a}) > u_i(a'_i, \vec{a}_{-i}) + \delta$ for all players i and all actions $a'_i \in A_i$.
3. After terminating the protocol above, they run again the protocol in order to make $\vec{a}^{PI}$ a Nash equilibrium, but taking $\vec{a}$ as input instead of $\vec{\sigma}$ (note that $\vec{a}$ is a Nash equilibrium at this point.)

Note that this protocol does not satisfy property (a1). However, it still satisfies properties (a) and (b). It clearly satisfies (b) by construction. To see that it satisfies (a), note that players can use $\vec{\sigma}$ as a punishment equilibrium during the first two steps, and they can use $\vec{a}$ during the third one. This proves Theorem 4 in the cases in which $\vec{\sigma}$ does not have full support.

4.2 Cases with Full Support

The case in which $\vec{\sigma}$ has full support is considerably harder than the previous cases. The main issue is that players are forced to modify $f_{\vec{\sigma}}$ in each commitment step and, under these circumstances, it is difficult to design a protocol that satisfies (a1). In order to illustrate the main obstacles, suppose that $\vec{\sigma}$ is a $(L_{\vec{\sigma}}^{\vec{a}^{PI}}, \delta)$ -punishable Nash equilibrium. If we wanted to design a commitment protocol

$$\Gamma = \Gamma_0 \rightarrow \Gamma_1 \rightarrow \Gamma_2 \rightarrow \dots \rightarrow \Gamma_k$$

that satisfies (a1), it is necessary that

- (P1) $d(\Gamma_i, \Gamma_{i+1}) < \delta$ for all $i < k$ (using the metric defined in Section 2).
- (P2) $u_i^j(\vec{a}) \geq u_i^{j+1}(\vec{a})$ for all $i \in P$ and all $j < k$, where u_i^j denotes the utility of player i in Γ_j .
- (P3) $u_i(\vec{a}^{PI}) = u_i^k(\vec{a}^{PI})$ for all $i \in P$.
- (P4) $\vec{\sigma}$ is a non-degenerate Nash equilibrium of Γ_i for all $i \leq k$.
- (P5) $\vec{a}^{PI}$ is a Nash equilibrium of Γ_k .

Designing such a commitment protocol does not seem that complicated since it basically reduces to finding a way to decrease the coefficients of $f_{\vec{\sigma}}$ outside of $\vec{a}^{PI}$ by at most δ , in each step such that $f_{\vec{\sigma}}(\vec{\sigma}) = \vec{0}$ and $|Df_{\vec{\sigma}}(\vec{\sigma})| \neq 0$ throughout the whole process. Unfortunately, this is not enough for our purposes. In fact, the non-degeneracy property guarantees that, in each game Γ_i , there exists a value $\delta_i > 0$ such that $\vec{\sigma}$ is $(L_{\vec{\sigma}}^{\vec{a}^{PI}}, \delta_i)$ -strongly punishable. However, $\vec{\sigma}$ may not necessarily be $(L_{\vec{\sigma}}^{\vec{a}^{PI}}, \delta)$ -strongly punishable. We can try to fix the protocol by setting δ to the minimum of all such δ_i , but then some of the payments involved in the process may become too large.

The key idea to circumvent this issue is making the distance between Γ_i and Γ_{i+1} “infinitesimally small”. This means constructing a path that goes from the original game $\Gamma = (P, A, U)$ to a new game $\Gamma' = (P, A, U')$ in which the following conditions are satisfied:

- (P1') γ is continuous (using the distance defined in Section 2 as the metric) and finite.
- (P2') The utility of the players (weakly) decreases through γ .
- (P3') $u_i(\vec{a}^{PI}) = u'_i(\vec{a}^{PI})$ for all $i \in P$ (note that this, combined with the property above, implies that the utilities in outcome $\vec{a}^{PI}$ are constant in γ .)
- (P4') For each game $\Gamma'' \in \gamma$, $\vec{\sigma}$ is a non-degenerate Nash equilibrium of Γ'' .
- (P5') $\vec{a}^{PI}$ is a Nash equilibrium of Γ' .

The next proposition shows that these conditions are sufficient for our purposes.

Proposition 1. *If there exists a path γ that satisfies (P1'), (P2'), (P3'), (P4') and (P5'), then, for some $\delta > 0$, it can be extended to a subgame perfect Nash equilibrium in $\Gamma^\delta_\perp$ in which players end up playing $\vec{a}^{PI}$.*

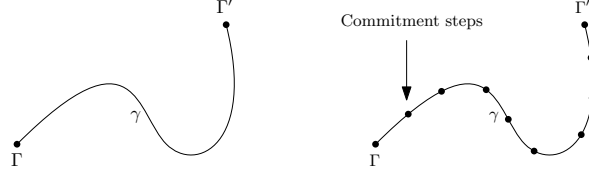
Before proving Proposition 1 we need the following lemma, which is an analogue of the Heine-Cantor theorem for normal-form games.

Lemma 2. *Let X_n^A be the metric space of normal-form games for n players with action set A , where the distance between two elements is the one given in Section 2. Let $\vec{\sigma}$ be a strategy profile over A and let $C \subseteq X_n^A$ be a compact set such that $\vec{\sigma}$ is a non-degenerate Nash equilibrium for all $\Gamma \in C$. Then, for all $\epsilon > 0$, there exists $\delta > 0$ such that $\vec{\sigma}$ is (ϵ, δ) -strongly punishable in all games $\Gamma \in C$.*

The proof of Lemma 2 can be found in Appendix A and mainly uses a compactness argument. With this we can prove Proposition 1.

Proof of Proposition 1. Suppose that such a path exists. By property (P4') and Theorem 1 we have that, for every game $\Gamma'' \in \gamma$, there exists $\delta_{\Gamma''} > 0$ such that $\vec{\sigma}$ is a $(L_{\vec{\sigma}}^{a^{PI}}, \delta_{\Gamma''})$ -strongly punishable Nash equilibrium of Γ'' . By property (P1') we have that γ is closed and bounded, and hence, by the Heine-Borel theorem, it is also a compact set. Therefore, by Lemma 2, there exists $\delta > 0$ such that $\vec{\sigma}$ is a $(L_{\vec{\sigma}}^{a^{PI}}, \delta)$ -strongly punishable equilibrium in all $\Gamma'' \in \gamma$. Note that this means that, in any game Γ''' at distance at most δ from some game $\Gamma'' \in \gamma$, there exists a Nash equilibrium $\vec{\sigma}'$ in which each player gets at most $L_{\vec{\sigma}}^{a^{PI}}$ additional utility than the one they would get by playing $\vec{\sigma}$ in Γ'' . By property (P2'), we have that this utility is bounded by $u_i(\vec{\sigma}) + L_{\vec{\sigma}}^{a^{PI}}$, which by definition is at most $u'_i(\vec{\sigma})$.

This property ensures that the following strategy is a subgame perfect equilibrium in $\Gamma^\delta_\perp$. During each step of the commitment phase, each player commits to pay in such a way that the resulting game remains in the path γ . Intuitively, we can view this as if players perform small leaps over γ until they reach Γ' . If a player ever defects from the proposed strategy, all players should immediately vote to terminate the commitment phase and play the punishment equilibrium. If no one defects and they eventually reach Γ' , they play $\vec{a}^{PI}$. It is easy to check that this strategy profile is indeed a subgame perfect equilibrium. If a player i defects from the main strategy, it guarantees that the remaining players will vote to terminate and that the remaining players will play $\vec{\sigma}'$. In this case, it is also a best response for i to vote to terminate (it doesn't change the outcome) and to play σ'_i since it is a Nash equilibrium. However, as shown in the previous paragraph, doing this guarantees her at most $u'_i(\vec{\sigma})$ utility which, by property (P3'), is less than the utility that i would get if no one defects. Moreover, if no player defects, playing a_i^{PI} is also a best response since, by property (P5'), $\vec{a}^{PI}$ is a Nash equilibrium of Γ' .



□

The argument above shows that Theorem 4 reduces to construct a path γ that satisfies the desired conditions. In the remaining part of this section we will show how to construct such a path by concatenating linear segments. In order to provide more intuition, We will begin our analysis with two-player games and later generalize it for arbitrary games with any players.

4.2.1 Two-Player Games

Suppose that $\Gamma = (P, A, U)$ is a two-player game and $\vec{\sigma}$ is a non-degenerate Nash equilibrium of Γ with full support. Assume without loss of generality that the target outcome is (a_1^1, a_2^1) . Since Γ is a two-player game, we have that, after rearranging some of its rows, $Df_{\vec{\sigma}}$ can be written as

$$Df_{\vec{\sigma}} = \left(\begin{array}{c|c} X_1 & 0_{N \times N} \\ \hline 0_{M \times M} & X_2 \end{array} \right)$$

with

$$X_1 = \begin{pmatrix} 1 & 1 & \dots & 1 \\ u_1(a_1^1, a_2^1) - u_1(a_1^2, a_2^1) & u_1(a_1^1, a_2^2) - u_1(a_1^2, a_2^2) & \dots & u_1(a_1^1, a_2^{N_2}) - u_1(a_1^2, a_2^{N_2}) \\ u_1(a_1^1, a_2^1) - u_1(a_1^3, a_2^1) & u_1(a_1^1, a_2^2) - u_1(a_1^3, a_2^2) & \dots & u_1(a_1^1, a_2^{N_2}) - u_1(a_1^3, a_2^{N_2}) \\ \vdots & \vdots & \ddots & \vdots \\ u_1(a_1^1, a_2^1) - u_1(a_1^{N_1}, a_2^1) & u_1(a_1^1, a_2^2) - u_1(a_1^{N_1}, a_2^2) & \dots & u_1(a_1^1, a_2^{N_2}) - u_1(a_1^{N_1}, a_2^{N_2}) \end{pmatrix}$$

and

$$X_2 = \begin{pmatrix} 1 & 1 & \dots & 1 \\ u_2(a_1^1, a_2^1) - u_2(a_1^1, a_2^2) & u_2(a_1^2, a_2^1) - u_2(a_1^2, a_2^2) & \dots & u_2(a_1^{N_1}, a_2^1) - u_2(a_1^{N_1}, a_2^2) \\ u_2(a_1^1, a_2^1) - u_2(a_1^1, a_2^3) & u_2(a_1^2, a_2^1) - u_2(a_1^2, a_2^3) & \dots & u_2(a_1^{N_1}, a_2^1) - u_2(a_1^{N_1}, a_2^3) \\ \vdots & \vdots & \ddots & \vdots \\ u_2(a_1^1, a_2^1) - u_2(a_1^1, a_2^{N_2}) & u_2(a_1^2, a_2^1) - u_2(a_1^2, a_2^{N_2}) & \dots & u_2(a_1^{N_1}, a_2^1) - u_2(a_1^{N_1}, a_2^{N_2}) \end{pmatrix}.$$

This means that $|Df_{\vec{\sigma}}| \neq 0$ if and only if

- $N = M$.
- X_1 and X_2 are $N \times N$ square matrices such that $|X_1| \neq 0$ and $|X_2| \neq 0$.

The high level idea used for the construction of the path γ is the fact that adding a multiple of a row to another row does not change the determinant of the matrix. Following this intuition, our aim is

to find a commitment protocol for the players such that each commitment induces a row operation on X_1 or X_2 such that

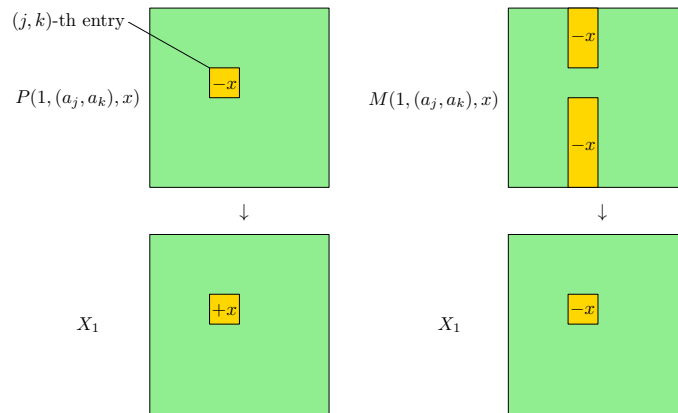
- The row operation consists of adding a multiple of a row to a different row.
- The first row (i.e., the row with all entries equal to 1) is not involved in the operation.
- Neither player modifies the utility of (a_1^1, a_2^1) .
- After a certain number of operations, the lowest entry in the first column of X_1 and X_2 is increased by a constant amount.

Note that, since the equation that corresponds to the first row - i.e., the one that states that the probabilities should add up to 1 - is the only one that has an independent term, any row operation that does not involve the first row preserves the solution of the linear system. Therefore, the second condition guarantees that $\vec{\sigma}$ remains a Nash equilibrium after each commitment step. The third condition ensures that the utility of both players in outcome (a_1^1, a_2^1) is preserved, and the fourth condition guarantees that, eventually, all entries of the first column of X_1 and X_2 will be positive, which implies that a_1 is eventually becomes a best response to a_1 for both players, and thus that (a_1^1, a_2^1) eventually becomes a Nash equilibrium. Thus, if players can manage to commit in such a way that each commitment step performs a row operation on X_1 or X_2 that satisfies the properties above, we can use these commitments in succession until (a_1^1, a_2^1) becomes a Nash equilibrium in the resulting game. The piece-wise linear path formed by linearly extending these row operations forms the desired path γ . More precisely, if a player's commitment performs a row operation on X_1 or X_2 that consists of replacing row r by row $r + c \cdot r'$, where r' is another row of the matrix, then we can connect the original and the resulting matrix with the linear segment defined by $r + \lambda \cdot c \cdot r'$, with $\lambda \in [0, 1]$. The first property guarantees that the determinants of X_1 and X_2 are preserved. In particular, the determinant of $Df_{\vec{\sigma}}$ remains non-zero.

In order to perform these row operations, consider the following two types of commitments:

- $P(i, \vec{a}, x)$: Player i decreases her utility by x in outcome $\vec{a}$.
- $M(i, \vec{a}, x)$: Player i decreases her utility by x in all outcomes of the form $(a', \vec{a}_{-i})$ with $a' \in A_i \setminus \{a_i\}$.

By construction, if $\vec{a} := (a_1^j, a_2^k)$ and $j > 1$, $P(1, \vec{a}, x)$ increases the value of the (j, k) -th entry of X_1 by x and $M(1, \vec{a}, x)$ decreases its value by x instead (for player 2, the entry that is modified is (k, j) instead). These two operations are represented in the diagram below.



Combining these types of commitments, players can perform any type of row operation on X_1 and X_2 . In particular, if the number of available actions N satisfies $N > 2$, each player i can run the following protocol.

1. **Step 1:** Check all rows of X_i , if there is any row except the first one in which the first entry is strictly positive, add positive multiples of this row to all other rows of X_i except the first one. Repeat this process until all the entries of the first column of X_i are positive.
2. **Step 2:** If there is no row in X_i such that the first entry is strictly positive, check if there is one in which the first entry is strictly negative. If it exists, add a negative multiple of this row to all other rows except the first one. Repeat this process until the first entry of all other rows is positive. Then, run Step 1 using any other row as pivot.

Clearly, if all players follow the protocol above, the first columns of X_1 and X_2 will eventually be non-negative, and therefore (a_1^1, a_2^1) will eventually become a Nash equilibrium of the resulting game. Note that the protocol can be performed in $\Gamma_\perp^\delta$ since players can always choose to add multiples of the rows that are small enough for their coordinates to be bounded by δ in absolute value. Also note that since we are always adding rows that have a positive first entry, we are not modifying the players' utilities in (a_1^1, a_2^1) , as desired. Altogether, this proves Theorem 4 in the case of two players and $N > 2$. The case of two players and binary actions is quite special and must be treated separately (see Appendix C). The proof for the case of three or more players uses similar techniques and can be found in Appendix B. The high-level idea is that players can commit in a way that they modify the components of $f_{\vec{\sigma}}$ while preserving the image of $\vec{\sigma}$ and the gradient of that component (and therefore the non-degeneracy property). Thus, they can slowly converge to the desired outcome by applying these types of transformations. To prove Theorem 3, players first commit in a way that results in an outcome that gives all players the desired utility profile, and then they use the techniques of the proof of Theorem 4 to make it a Nash equilibrium of the resulting game (see Appendix D for a full proof.)

5 Conclusion

Given a normal-form game, we showed that giving players the chance to sequentially commit to small side payments allows them to implement all welfare-maximizing utility profiles that strictly Pareto improve any Nash equilibria of the original game. Moreover, if there already exists an outcome that strictly Pareto improves any of these Nash equilibria, that outcome can be implemented via *money burning*. This aligns with Coase's theorem and comes in contrast to Jackson and Wilkie's analysis, who show that this is not always the case for a single round of unbounded commitments. Two natural questions follow from these results.

First, finding out if these positive results can be extended to a Bayesian setting where players have private information. Most of the techniques used in the construction of the commitment protocol make use of the fact that players have full information about the payoffs of the game. Thus, it would seem that the analysis of the Bayesian setting would require a different approach.

Second, constructing a more realistic lower bound for δ , even if it holds only for specific classes of games. In Section 3.1 we provide an upper bound on the number of commitment rounds used in the constructions for Theorems 3 and Theorem 4 as a function of the parameter δ . However, we noted that the lower bound for δ that we would get with a generic estimate obtained by backtracking the proof of Theorem 1 would be in general significantly lower than the actual lower bound for a vast majority of games.

References

- [1] Ittai Abraham, Danny Dolev, and Joseph Y Halpern. “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. 2008, pp. 302–319.
- [2] Ittai Abraham et al. “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*. 2006, pp. 53–62.
- [3] Ittai Abraham et al. “Implementing mediators with asynchronous cheap talk”. In: *Proceedings of the 2019 ACM Symposium on Principles of Distributed Computing*. 2019, pp. 501–510.
- [4] Ashton Anderson, Yoav Shoham, and Alon Altman. “Internal implementation”. In: *Proceedings of the Ninth International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*. Toronto, Canada, 2010, pp. 191–198.
- [5] Itai Arieli, Yakov Babichenko, and Fedor Sandomirskiy. “Bayesian persuasion with mediators”. In: *arXiv preprint arXiv:2203.04285* (2022).
- [6] Itai Arieli, Ivan Geffner, and Moshe Tennenholtz. “Mediated cheap talk design”. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 37. 5. 2023, pp. 5456–5463.
- [7] Robert J Aumann. “Subjectivity and correlation in randomized strategies”. In: *Journal of mathematical Economics* 1.1 (1974), pp. 67–96.
- [8] Michael Bacharach. *Beyond individual choice: teams and frames in game theory*. Princeton University Press, 2006.
- [9] Michael Bacharach. “Interactive team reasoning: A contribution to the theory of co-operation”. In: *Research in economics* 53.2 (1999), pp. 117–147.
- [10] Nicholas Bardsley et al. “Explaining focal points: Cognitive hierarchy theory versus team reasoning”. In: *The Economic Journal* 120.543 (2010), pp. 40–79.
- [11] John M. Barron and Glen R. Waddell. “Executive rank, pay and project selection”. In: *Journal of Financial Economics* 67 (2003), pp. 305–349.
- [12] Elchanan Ben-Porath. “Cheap talk in games with incomplete information”. In: *Journal of Economic Theory* 108.1 (2003), pp. 45–71.
- [13] Edward H Clarke. “Multipart pricing of public goods”. In: *Public choice* (1971), pp. 17–33.
- [14] Emery Cooper, Caspar Oesterheld, and Vincent Conitzer. “Characterising Simulation-Based Program Equilibria”. In: *Proceedings of the Thirty-Ninth Annual AAAI Conference on Artificial Intelligence*. 2025.
- [15] Andrew Critch. “A Parametric, Resource-Bounded Generalization of Löb’s Theorem, and a Robust Cooperation Criterion for Open-Source Game Theory”. In: *Journal of Symbolic Logic* 84.4 (Dec. 2019), pp. 1368–1381. DOI: 10.1017/jsl.2017.42.
- [16] Leif Danziger and Adi Schnytzer. “Implementing the Lindahl voluntary-exchange mechanism”. In: *European Journal of Political Economy* 7.1 (1991), pp. 55–64.
- [17] Joel S. Demski and David E. M. Sappington. “Delegated Expertise”. In: *Journal of Accounting Research* 25.1 (1987), pp. 68–89.
- [18] Yuan Deng and Vincent Conitzer. “Disarmament Games”. In: *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*. San Francisco, CA, USA, 2017, pp. 473–479.
- [19] Yuan Deng and Vincent Conitzer. “Disarmament Games with Resources”. In: *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence*. New Orleans, LA, USA, 2018.
- [20] Eberhard Feess and Markus Walzl. “Delegated expertise – when are good projects bad news?” In: *Econ. Lett.* 82.1 (Jan. 2004), pp. 77–82.

- [21] Francoise Forges. “An approach to communication equilibria”. In: *Econometrica: Journal of the Econometric Society* (1986), pp. 1375–1385.
- [22] Ivan Geffner and Joseph Y Halpern. “Communication games, sequential equilibrium, and mediators”. In: *Journal of Economic Theory* 221 (2024), p. 105890.
- [23] Ivan Geffner and Joseph Y Halpern. “Lower bounds on implementing mediators in asynchronous systems with rational and malicious agents”. In: *Journal of the ACM* 70.2 (2023), pp. 1–21.
- [24] Ivan Geffner and Moshe Tennenholtz. “Making a Nash Equilibrium Resilient to Coalitions”. In: *Proceedings of the 25th ACM Conference on Economics and Computation*. 2024, pp. 213–238.
- [25] Dino Gerardi. “Unmediated communication in games with complete and incomplete information”. In: *Journal of Economic Theory* 114.1 (2004), pp. 104–131.
- [26] Allan Gibbard and William L Harper. “Counterfactuals and two kinds of expected utility”. In: *Ifs: Conditionals, belief, decision, chance and time*. Springer, 1978, pp. 153–190.
- [27] Natalie Gold and Andrew M Colman. “Team reasoning and the rational choice of payoff-dominant outcomes in games”. In: *Topoi* 39.2 (2020), pp. 305–316.
- [28] Denis Gromb and David Martimort. “Collusion and the organization of delegated expertise”. In: *Journal of Economic Theory* 137.1 (Nov. 2007), pp. 271–299.
- [29] Theodore Groves. “Incentives in teams”. In: *Econometrica: Journal of the Econometric Society* (1973), pp. 617–631.
- [30] Anshul Gupta and Sven Schewe. “It pays to pay in bi-matrix games-a rational explanation for bribery”. In: (2015).
- [31] Joel M Guttman. “A non-Cournot model of voluntary collective action”. In: *Economica* (1987), pp. 1–19.
- [32] Joel M Guttman. “Understanding collective action: matching behavior”. In: *The American Economic Review* 68.2 (1978), pp. 251–255.
- [33] Joel M Guttman and Adi Schnytzer. “A solution of the externality problem using strategic matching”. In: *Social Choice and Welfare* 9.1 (1992), pp. 73–88.
- [34] J. V. Howard. “Cooperation in the Prisoner’s Dilemma”. In: *Theory and Decision* 24 (May 1988), pp. 203–213. DOI: 10.1007/BF00148954.
- [35] Matthew O Jackson and Simon Wilkie. “Endogenous games and mechanisms: Side payments among players”. In: *The Review of Economic Studies* 72.2 (2005), pp. 543–566.
- [36] Adam Tauman Kalai et al. “A commitment folk theorem”. In: *Games and Economic Behavior* 69.1 (2010), pp. 127–137.
- [37] Emir Kamenica and Matthew Gentzkow. “Bayesian persuasion”. In: *American Economic Review* 101.6 (2011), pp. 2590–2615.
- [38] Vojtěch Kovařík, Caspar Oesterheld, and Vincent Conitzer. “Game theory with simulation of other players”. In: *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*. 2023, pp. 2800–2807.
- [39] Richard A. Lambert. “Executive Effort and Selection of Risky Projects”. In: *The RAND Journal of Economics* 17.1 (1986), pp. 77–88. DOI: 10.2307/2555629.
- [40] Ehud Lehrer, Dinah Rosenberg, and Eran Shmaya. “Signaling and mediation in games with common interests”. In: *Games and Economic Behavior* 68.2 (2010), pp. 670–682.
- [41] Mingyang Liu, Gabriele Farina, and Asuman Ozdaglar. “Equilibrium with Internal Transfers”. In: *arXiv preprint arXiv:2606.20960* (2026).
- [42] James M Malcomson. “Principal and Expert Agent”. In: *The B.E. Journal of Theoretical Economics* 9.1 (2009). DOI: 10.2202/1935-1704.1528. URL: <https://doi.org/10.2202/1935-1704.1528>.

- [43] R. Preston McAfee. “Effective Computability in Economic Decisions”. May 1984. URL: <https://www.mcafee.cc/Papers/PDF/EffectiveComputability.pdf>.
- [44] Dov Monderer and Moshe Tennenholtz. “K-Implementation”. In: *Journal of Artificial Intelligence Research* 21 (2004), pp. 37–62.
- [45] Dov Monderer and Moshe Tennenholtz. “Strong mediated equilibrium”. In: *Artificial Intelligence* 173.1 (2009), pp. 180–195.
- [46] Caspar Oesterheld. “Robust Program Equilibrium”. In: *Theory and Decision* 86 (2019). DOI 10.1007/s11238-018-9679-3, pp. 143–159.
- [47] Ola Rozenfeld and Moshe Tennenholtz. “Routing Mediators”. In: *Proceedings of the Twentieth International Joint Conference on Artificial Intelligence (IJCAI)*. Hyderabad, India, 2007, pp. 1488–1493.
- [48] Ariel Rubinstein. *Modeling Bounded Rationality*. Ed. by Karl Gunnar Persson. Zeuthen Lecture Book Series. The MIT Press, 1998.
- [49] Nathaniel Sauerberg and Caspar Oesterheld. “Computing Optimal Commitments to Strategies and Outcome-Conditional Utility Transfers”. In: *Proceedings of the Twenty-Third International Conference on Autonomous Agents and Multiagent Systems (AAMAS-24)*. 2024, pp. 1654–1663.
- [50] Neal M. Stoughton. “Moral Hazard and the Portfolio Management Problem”. In: *The Journal of Finance* 48.5 (Dec. 1993), pp. 2009–2028.
- [51] Robert Sugden. “Thinking as a team: Towards an explanation of nonselfish behavior”. In: *Social philosophy and policy* 10.1 (1993), pp. 69–89.
- [52] Moshe Tennenholtz. “Program Equilibrium”. In: *Games and Economic Behavior* 49 (2004), pp. 363–373.
- [53] Hal R Varian. “A solution to the problem of externalities when agents are well-informed”. In: *The American Economic Review* (1994), pp. 1278–1293.
- [54] Hal R Varian. “Sequential contributions to public goods”. In: *Journal of Public Economics* 53.2 (1994), pp. 165–186.
- [55] William Vickrey. “Counterspeculation, auctions, and competitive sealed tenders”. In: *The Journal of finance* 16.1 (1961), pp. 8–37.

A Proof of Lemma 2

Let $\Gamma = (P, A, U)$ be any game in C and let $\epsilon > 0$ be any positive real number. Since $\vec{\sigma}$ is a non-degenerate Nash equilibrium of Γ , by Theorem 1 there exists $\delta > 0$ such that $\vec{\sigma}$ is $(\epsilon/2, \delta)$ -strongly punishable. Let $\Gamma' = (P, A, U')$ be any game in C at distance at most $\min(\epsilon, \delta)/2$ from Γ . Since $d(\Gamma, \Gamma') < \epsilon/2$, we have that $|u'(\vec{\sigma}) - u(\vec{\sigma})| < \epsilon/2$. Moreover, we have that all normal-form games Γ'' at distance at most $\delta/2$ from Γ' are at most at distance δ from Γ , and therefore satisfy that there always exists a Nash equilibrium $\vec{\sigma}''$ with the same support as $\vec{\sigma}$ such that $u''_i(\vec{\sigma}'') - u_i(\vec{\sigma}) < \epsilon/2$ for all $i \in P$. By the triangular inequality, this means that $u''_i(\vec{\sigma}'') - u'_i(\vec{\sigma}) < \epsilon$ for all $i \in P$, and therefore that all games Γ' in a ball of radius $\min(\epsilon, \delta)/2$ centered at Γ satisfy that $\vec{\sigma}$ is a $(\delta/2, \epsilon)$ -strongly punishable Nash equilibrium on those games. More generally, this means that, for each game $\Gamma \in C$, there exists an open set B^Γ and a positive real number $\delta^\Gamma > 0$ such that $\vec{\sigma}$ is $(\epsilon, \delta^\Gamma)$ -punishable on all $\Gamma' \in B^\Gamma$. Since C is a compact set and $\{B_\Gamma\}_{\Gamma \in C}$ is an open cover of C , there exists a finite sub-cover $\{B_{\Gamma_1}, B_{\Gamma_2}, \dots, B_{\Gamma_N}\}$. Let $\delta := \min_{i \leq N} \{\delta_{\Gamma_i}\}$, we show next that $\vec{\sigma}$ is an (ϵ, δ) -strongly punishable equilibrium in all games $\Gamma \in C$. Since $\{B_{\Gamma_1}, B_{\Gamma_2}, \dots, B_{\Gamma_N}\}$ is a cover of C , Γ must be in B_{Γ_k} for some $k \in \{1, 2, \dots, N\}$. Therefore, by definition, $\vec{\sigma}$ is an $(\epsilon, \delta_{\Gamma_k})$ -strongly punishable equilibrium, which implies that it also is (ϵ, δ) -strongly punishable since $\delta \leq \delta_{\Gamma_k}$.

B Proof of Theorem 4 for Arbitrary Games

The case with several players works differently than the case with only two players presented in Section 4.2.1. The main difference is that $f_{\vec{\sigma}}$ is no longer a linear function, and thus some of the arguments used in the previous section do not apply anymore. In particular, doing a $P(i, \vec{a}, x)$ or a $M(i, \vec{a}, x)$ commitment does not increase or decrease the value of a single entry of $Df_{\vec{\sigma}}$ by x . In order to see why, consider the following example.

Example 5. Consider a game $\Gamma = (P, A, U)$ with $|P| = 3$, $A_1 = A_2 = A_3 = \{0, 1\}$ and such that there exists a Nash equilibrium $\vec{\sigma}$ with full support.

In this example, if we use the notation $d_i(\vec{a}_{-i}) := u_i(0, \vec{a}_{-i}) - u_i(1, \vec{a}_{-i})$, we have that

$$f_{\vec{\sigma}}(p_0^1, p_1^1, p_0^2, p_1^2, p_0^3, p_1^3) := \begin{cases} p_0^1 + p_1^1 \\ p_0^2 + p_1^2 \\ p_0^3 + p_1^3 \\ d_1(0, 0) \cdot p_0^2 p_0^3 + d_1(0, 1) \cdot p_0^2 p_1^3 + d_1(1, 0) \cdot p_1^2 p_0^3 + d_1(1, 1) \cdot p_1^2 p_1^3 \\ d_2(0, 0) \cdot p_0^1 p_0^3 + d_2(0, 1) \cdot p_0^1 p_1^3 + d_2(1, 0) \cdot p_1^1 p_0^3 + d_2(1, 1) \cdot p_1^1 p_1^3 \\ d_3(0, 0) \cdot p_0^1 p_0^2 + d_3(0, 1) \cdot p_0^1 p_1^2 + d_3(1, 0) \cdot p_1^1 p_0^2 + d_3(1, 1) \cdot p_1^1 p_1^2 \end{cases}$$

where p_j^i is a variable that denotes the probability that player i plays action j . To see how commitments affect $Df_{\vec{\sigma}}$, note that if player 1 performs a $P(1, (1, 0, 0), x)$ commitment, the only change in $f_{\vec{\sigma}}$ is that $d_1(0, 0)$ increases by x . This induces an increment of $(0, 0, x \cdot p_0^3, 0, x \cdot p_0^3, 0)$ in the fourth row of $Df_{\vec{\sigma}}$. More generally, it can be checked that, in this example, the $P(i, \vec{a}, x)$ and $M(i, \vec{a}, x)$ commitments modify exactly two values of a single row of $Df_{\vec{\sigma}}$.

Even though this property can make us believe that the desired protocol is much more difficult to construct with respect to the case of two players, it allows us to take the following approach. As in the case of two players, we construct a piece-wise linear path. However, each piece of the path, instead of inducing a row operation on $Df_{\vec{\sigma}}$ (as in the case of two players), it will not modify $Df_{\vec{\sigma}}$ at all. To see how this works, we will introduce this technique with the example above.

However, before going into the details, it will be useful to introduce some notation. First, denote by $f_{\vec{\sigma}}^j$ the j th component of $f_{\vec{\sigma}}$. For instance, in Example 5, we have that

$$f_{\vec{\sigma}}^4 = d_1(0, 0) \cdot p_0^2 p_0^3 + d_1(0, 1) \cdot p_0^2 p_1^3 + d_1(1, 0) \cdot p_1^2 p_0^3 + d_1(1, 1) \cdot p_1^2 p_1^3.$$

Using our usual notation where $A_1 = (a_1^1, a_1^2, a_1^3, \dots, a_1^{N_1})$, $A_2 = (a_2^1, a_2^2, a_2^3, \dots, a_2^{N_2})$, $\dots$, $A_n = (a_n^1, a_n^2, a_n^3, \dots, a_n^{N_n})$, there exists a natural lexicographical ordering of action profiles. It can be defined recursively as follows:

- $a_i^j \prec a_i^k \iff j < k$
- $(a_{i_1}^{j_1}, a_{i_2}^{j_2}, \dots, a_{i_\ell}^{j_\ell}) \prec (a_{i_1}^{k_1}, a_{i_2}^{k_2}, \dots, a_{i_\ell}^{k_\ell})$ if and only if $j_1 < k_1$ or $j_1 = k_1$ and $(a_{i_2}^{j_2}, \dots, a_{i_\ell}^{j_\ell}) \prec (a_{i_2}^{k_2}, \dots, a_{i_\ell}^{k_\ell})$.

This also induces a total order between the coefficients of $f_{\vec{\sigma}}^j$ for $j > n$. All these coefficients are of the form $(u_i(a_i^1, \vec{a}_{-i}) - u_i(a_i^k, \vec{a}_{-i}))$ for some fixed player i and some fixed action index k . Then we can define the following order between coefficients:

$$(u_i(a_i^1, \vec{a}_{-i}) - u_i(a_i^k, \vec{a}_{-i})) \prec (u_i(a_i^1, \vec{a}'_{-i}) - u_i(a_i^k, \vec{a}'_{-i})) \iff \vec{a}_{-i} \prec \vec{a}'_{-i}.$$

In Example 5, this order is given by the following relation (note that the first action index is 0 instead of 1):

$$d_i(\vec{a}_{-i}) \prec d_i(\vec{a}'_{-i}) \iff \vec{a}_{-i} \prec \vec{a}'_{-i}.$$

In particular, in $f_{\vec{\sigma}}^4$, we would have the following order:

$$d_1(0,0) \prec d_1(0,1) \prec d_1(1,0) \prec d_1(1,1).$$

For simplicity, we will assume that the coefficients of $f_{\vec{\sigma}}^i$ are always ordered following this relation. Given a positive integer k and an array of real values $\vec{x} = (x_1, x_2, \dots)$ with the appropriate dimension, we denote by $R(j, \vec{x})$ the commitment by which a player modifies the coefficients of $f_{\vec{\sigma}}^j$ by $x_1, x_2, \dots$, respectively. More precisely, by performing a $R(j, \vec{x})$ commitment, we increase the first coefficient of $f_{\vec{\sigma}}^j$ by x_1 , the second one by x_2 , etc. Note that this commitment can always be performed by adding together commitments of the form $P(i, \vec{a}', x)$ and $M(i, \vec{a}', x)$. For instance, in Example 5, if x_1, x_2, x_3 and x_4 are positive real values, we have that

$$\begin{aligned} R(4, (x_1, -x_2, -x_3, x_4)) &= P(1, (1, 0, 0), x_1) \\ &\oplus M(1, (1, 0, 1), x_2) \\ &\oplus M(1, (1, 1, 0), x_3) \\ &\oplus P(1, (1, 1, 1), x_4), \end{aligned}$$

where $\oplus$ indicates the concatenation of commitments (i.e., performing those commitments at the same time). We also denote by $f_{\vec{\sigma}}^j \oplus R(j, \vec{x})$ the resulting function obtained from performing an $R(j, \vec{x})$ commitment on $f_{\vec{\sigma}}^j$. In Example 5, if we set $x_1 = \lambda \cdot p_1^2 p_1^3$, $x_2 = \lambda \cdot p_1^2 p_0^3$, $x_3 = \lambda \cdot p_0^2 p_1^3$ and $x_4 = \lambda \cdot p_0^2 p_0^3$, we have that

$$\begin{aligned} \nabla(f_{\vec{\sigma}}^4 \oplus R(4, (x_1, -x_2, -x_3, x_4))) (\vec{p}) - \nabla f_{\vec{\sigma}}^4 (\vec{p}) &= \lambda \cdot p_2^1 p_3^1 \cdot \nabla(p_2^0 p_3^0) \\ &- \lambda \cdot p_2^1 p_3^0 \cdot \nabla(p_2^0 p_3^1) \\ &- \lambda \cdot p_2^0 p_3^1 \cdot \nabla(p_2^1 p_3^0) \\ &+ \lambda \cdot p_2^0 p_3^0 \cdot \nabla(p_2^1 p_3^1). \end{aligned} \tag{1}$$

The factor by which λ is multiplied can be expanded as

$$\begin{aligned} & \begin{pmatrix} 0, & 0, & p_2^1 p_3^1 p_3^0, & 0, & p_2^1 p_3^1 p_2^0, & 0 \end{pmatrix} \\ + & \begin{pmatrix} 0, & 0, & -p_2^1 p_3^0 p_3^1, & 0, & 0, & -p_2^1 p_3^0 p_2^0 \end{pmatrix} \\ + & \begin{pmatrix} 0, & 0, & 0, & -p_2^0 p_3^1 p_3^0, & -p_2^1 p_3^1 p_2^1, & 0 \end{pmatrix} \\ + & \begin{pmatrix} 0, & 0, & 0, & p_2^0 p_3^0 p_3^1, & 0, & p_2^0 p_3^0 p_2^1 \end{pmatrix}, \end{aligned}$$

which shows that

$$\nabla(f_{\vec{\sigma}}^4 \oplus R(4, (x_1, -x_2, -x_3, x_4))) (\vec{p}) - \nabla f_{\vec{\sigma}}^4 (\vec{p}) = \vec{0}$$

for all $\lambda \in \mathbb{R}$. Moreover, we also have the following.

$$\begin{aligned} (f_{\vec{\sigma}}^4 \oplus R(4, (x_1, -x_2, -x_3, x_4))) (\vec{p}) - f_{\vec{\sigma}}^4 (\vec{p}) &= \lambda \cdot p_2^1 p_3^1 \cdot (p_2^0 p_3^0) \\ &- \lambda \cdot p_2^1 p_3^0 \cdot (p_2^0 p_3^1) \\ &- \lambda \cdot p_2^0 p_3^1 \cdot (p_2^1 p_3^0) \\ &+ \lambda \cdot p_2^0 p_3^0 \cdot (p_2^1 p_3^1). \end{aligned} \tag{2}$$

which is the same as equation 1 but without the gradient operator. It can be easily checked that equation 2 is also equal to 0. This means that player 1 can modify the values of the coefficients of

$f_{\vec{\sigma}}^4$ following the linear path defined by $R(4, \lambda \cdot (p_2^1 p_3^1, p_2^1 p_3^0, p_2^0 p_3^1, p_2^0 p_3^0))$ without modifying $Df_{\vec{\sigma}}(\vec{p})$ or $f_{\vec{\sigma}}(\vec{p})$. In particular, if we take $\vec{p} = \vec{\sigma}$, we have that player 1 can modify the coefficients of $f_{\vec{\sigma}}^4$ following this path while simultaneously satisfying that $Df_{\vec{\sigma}}(\vec{\sigma})$ remains constant and that $\vec{\sigma}$ remains a Nash equilibrium in the resulting game. Suppose that $\vec{a}^{PI} = (0, 0, 0)$. If we choose positive values of λ , we can also guarantee that the utility of player 1 in $\vec{a}^{PI}$ remains untouched and that eventually 0 becomes a best response when the remaining players play $(0, 0)$. Having players 2 and 3 follow similar commitment patterns gives us the desired path γ .

We next extend this argument to arbitrary games. Let $\Gamma = (P, A, U)$ with $P = \{1, 2, \dots, n\}$, $n \geq 3$, and $A_i = \{a_i^1, a_i^2, \dots, a_i^{N_i}\}$ for each $i \in P$. Assume without loss of generality that $|A_i| \geq 2$ for all $i \in P$ and that $\vec{a}^{PI} = (a_1^1, a_2^1, \dots, a_n^1)$. Our aim is to show that, for each component $f_{\vec{\sigma}}^j$ with $j > n$ (i.e., a component that is not of the form $p_i^1 + p_i^2 + \dots + p_i^{N_i}$), there always exists an array $\vec{x}^j$ such that

$$(Q1) \quad (f_{\vec{\sigma}}^j \oplus R(j, \vec{x}^j))(\vec{\sigma}) - f_{\vec{\sigma}}^j(\vec{\sigma}) = 0.$$

$$(Q2) \quad \nabla(f_{\vec{\sigma}}^j \oplus R(j, \vec{x}^j))(\vec{\sigma}) - \nabla f_{\vec{\sigma}}^j(\vec{\sigma}) = \vec{0}.$$

$$(Q3) \quad \text{The first coordinate of } \vec{x} \text{ is strictly positive.}$$

Suppose that we can indeed always find such an array $\vec{x}^j$. Then, players can construct a path from Γ to a game in which $\vec{a}^{PI}$ is a Nash equilibrium by successively performing $R(j, \lambda \cdot \vec{x}^j)$ commitments for different values of λ and j . By property (Q3), eventually the first coefficient of each component of $f_{\vec{\sigma}}$ becomes positive. When this happens, $\vec{a}^{PI}$ becomes a Nash equilibrium of the resulting game (note that we are assuming that $\vec{a}^{PI} = (a_1^1, a_2^1, \dots, a_n^1)$). Moreover, property (Q2) guarantees that the determinant of $Df_{\vec{\sigma}}$ remains unaltered, and property (Q1) guarantees that $\vec{\sigma}$ remains a Nash equilibrium throughout the whole process. This would prove Theorem 4 for $n \geq 3$.

Thus, it remains to show that an array $\vec{x}^j$ always exists. We will show this for $j = n + 1$, which has the following expression:

$$\begin{aligned} f_{\vec{\sigma}}^{n+1}(\vec{\sigma}) &= (u_1(a_1^1, a_2^1, \dots, a_n^1) - u_1(a_1^2, a_2^1, \dots, a_n^1)) \cdot p_2^1 p_3^1 \dots p_n^1 \\ &+ (u_1(a_1^1, a_2^1, \dots, a_n^2) - u_1(a_1^2, a_2^1, \dots, a_n^2)) \cdot p_2^1 p_3^1 \dots p_n^2 \\ &+ (u_1(a_1^1, a_2^{N_2}, \dots, a_n^{N_n}) - u_1(a_1^2, a_2^{N_2}, \dots, a_n^{N_n})) \cdot p_2^{N_2} p_3^{N_3} \dots p_n^{N_n} \end{aligned}$$

The proof for the other components is analogous. Our aim, thus, is to find an array

$$\vec{x} = (x_{(1,1,\dots,1)}, x_{(1,1,\dots,2)}, \dots, x_{(N_2, N_3, \dots, N_n)})$$

such that $R(n + 1, \vec{x})$ satisfies (Q1), (Q2) and (Q3). This is done in the following proposition.

Proposition 2. *Let $\vec{x}$ be an array defined as follows.*

$$x_{(i_2, i_3, \dots, i_n)} = \begin{cases} 0 & \text{if } i_j > 2 \text{ for some } j \in \{2, 3, \dots, n\} \\ (-1)^{(i_2 + \dots + i_n + n + 1)} \cdot \sigma_2^{(3-i_2)} \sigma_3^{(3-i_3)} \dots \sigma_n^{(3-i_n)} & \text{otherwise.} \end{cases}$$

Then, $R(n + 1, \vec{x})$ satisfies (Q1), (Q2) and (Q3).

Proof. (Q1) We have that

$$(f_{\vec{\sigma}}^{n+1} \oplus R(n + 1, \vec{x}))(\vec{\sigma}) - f_{\vec{\sigma}}^{n+1}(\vec{\sigma}) = \sum_{(i_2, i_3, \dots, i_n) \in \{1, 2\}^{n-1}} (-1)^{(i_2 + \dots + i_n + n + 1)} \cdot \sigma_2^{(3-i_2)} \sigma_3^{(3-i_3)} \dots \sigma_n^{(3-i_n)} \sigma_2^{i_2} \sigma_3^{i_3} \dots \sigma_n^{i_n},$$

which can be rearranged into

$$(-1)^{n+1} \left(\prod_{j=2}^n \sigma_j^1 \sigma_j^2 \right) \left(\sum_{(i_2, i_3, \dots, i_n) \in \{1, 2\}^{n-1}} (-1)^{(i_2 + \dots + i_n)} \right).$$

It is easy to check that the rightmost term is 0.

(Q2) This reduces to check that

$$\frac{\partial ((f_{\vec{\sigma}}^{n+1} \oplus R(n+1, \vec{x})) - f_{\vec{\sigma}}^{n+1})}{\partial p_i^j}(\vec{\sigma}) = 0$$

for all $i > 1$ and $j \in \{1, 2\}$. We show this for $i = 2$ and $j = 1$. Since the function is symmetric on both i and j , the general result will follow.

We have that

$$\frac{\partial ((f_{\vec{\sigma}}^{n+1} \oplus R(n+1, \vec{x})) - f_{\vec{\sigma}}^{n+1})}{\partial p_2^1}(\vec{\sigma}) = \sum_{(i_3, \dots, i_n) \in \{1, 2\}^{n-2}} (-1)^{(i_3 + \dots + i_n + n + 2)} \sigma_2^2 \sigma_3^{(3-i_3)} \dots \sigma_n^{(3-i_n)} \sigma_3^{i_3} \sigma_4^{i_4} \dots \sigma_n^{i_n},$$

which can be rearranged into

$$(-1)^n \sigma_2^2 \left(\prod_{j=3}^n \sigma_j^1 \sigma_j^2 \right) \left(\sum_{(i_3, \dots, i_n) \in \{1, 2\}^{n-2}} (-1)^{(i_3 + \dots + i_n)} \right),$$

which is also equal to 0.

(Q3) By construction, the sign of $x_{(1,1,\dots,1)}$ is given by $(-1)^{2n}$, which is positive. $\square$

This completes the proof of Theorem 4 in the case of three or more players.

C Proof of Theorem 4 for Two-Player Games with Binary Actions

The last remaining case is the one with two players and binary actions. For simplicity, assume that $A_1 = A_2 = \{0, 1\}$.

We begin the analysis with a characterization of all 2×2 games in which there exists a Nash equilibrium with full support.

Proposition 3. *Let $\Gamma = (P, A, U)$ be a normal-form game for two players in which $A_1 = A_2 = \{0, 1\}$. Then, there exists a Nash equilibrium with full support if and only if, for each player i , one of the following conditions holds.*

- (i) *Player i is indifferent between 0 and 1 (i.e., $u_i(0_i, 0_{-i}) = u_i(1_i, 0_{-i})$ and $u_i(0_i, 1_{-i}) = u_i(1_i, 1_{-i})$).*
- (ii) *Player i strictly prefers 0 when $-i$ (i.e., the other player) prefers 0 and i strictly prefers 1 when $-i$ prefers 1.*
- (iii) *Player i strictly prefers 0 when $-i$ prefers 1 and i strictly prefers 1 when $-i$ prefers 0.*

Proof. As argued in Section 2.3, Γ has a Nash equilibrium with full support if and only if there exists a solution in $(0, 1)^4$ of the following system of equations

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ d_1(0) & d_1(1) & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & d_2(0) & d_2(1) \end{pmatrix} \begin{pmatrix} p_2^0 \\ p_2^1 \\ p_1^0 \\ p_1^1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix},$$

where $d_i(a)$ denotes $u_i(0_i, a_{-i}) - u_i(1_i, a_{-i})$. This is equivalent to the following systems having solutions in $(0, 1)^2$.

$$\begin{pmatrix} 1 & 1 \\ d_1(0) & d_1(1) \end{pmatrix} \begin{pmatrix} p_2^0 \\ p_2^1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 1 \\ d_2(0) & d_2(1) \end{pmatrix} \begin{pmatrix} p_1^0 \\ p_1^1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

We will restrict our analysis to the first two coordinates of the solution, since the reasoning for the latter two is analogous. Since the probabilities are strictly positive, in order to have a solution in $(0, 1)^2$, $d_1(0)$ and $d_1(1)$ must be either both zero or both non-zero and of different signs. In the first case, any solution in $(0, 1)^2$ is valid. In the latter one we have that p_2^0 and p_2^1 are given by

$$\begin{aligned} p_2^0 &= \frac{d_1(1)}{d_1(1) - d_1(0)} \\ p_2^1 &= \frac{-d_1(0)}{d_1(1) - d_1(0)}. \end{aligned}$$

It is easy to check that these values are indeed in $(0, 1)$ if $d_1(0)$ and $d_1(1)$ have different signs. $\square$

This characterization shows that, if there exists a Nash equilibrium in Γ with full support, then players must be of type (i) or (iii) (if any of them was of type (ii), $\vec{\sigma}$, would be degenerate). We next provide a commitment protocol $\pi_{2 \times 2}^\delta$ that allows players to converge on $\vec{a}^{PI}$. For simplicity, assume that $\vec{a}^{PI} = (0, 0)$ and that $\delta < \min\{|u_i(1_i, 0_{-i}) - u_i(0_i, 0_{-i})|, |u_i(0_i, 1_{-i}) - u_i(1_i, 1_{-i})|\}$ for all players i of type (i) or (iii). For completion, we will also say that a player is of type (iv) if it does not fall into any of the other three categories.

Protocol $\pi_{2 \times 2}^\delta$:

Step 1: Each player i commits to pay δ in $(0_i, 1_{-i})$ and $(1_i, 1_{-i})$ in successive commitment steps, until their utility function satisfies $u_i(0_i, 0_{-i}) > \max\{u_i(0_i, 1_{-i}), u_i(1_i, 1_{-i})\}$.

Step 2: If player i is of type (iii), in each commitment step, it commits to pay the following:

$$\begin{aligned} &\min\{\delta, u_i(1_i, 0_{-i}) - u_i(0_i, 0_{-i}) - \delta\} && \text{in } (1, 0) \\ &\min\{\delta, u_i(0_i, 1_{-i}) - u_i(1_i, 1_{-i}) - \delta\} && \text{in } (0, 1). \end{aligned}$$

Each player of type (iii) repeats this process until all players of this type satisfy that

$$u_i(1_i, 0_{-i}) - u_i(0_i, 0_{-i}) = u_i(0_i, 1_{-i}) - u_i(1_i, 1_{-i}) = \delta.$$

Step 3: Each player i of type (iii) commits to pay δ on $(1, 0)$ and $(0, 1)$. Then, they vote to terminate the commitment phase.

Since players can only be of type (i) or (ii) on termination, it is easy to check that $(0, 0)$ is a Nash equilibrium by the end of the protocol. It only remains to define what the players should do in case

someone defects. The next proposition shows that, immediately after one of the players defects, they can always find a Nash equilibrium in which both of them are worse than in $(0, 0)$. This shows that $\pi_{2 \times 2}$ can be extended to a subgame perfect equilibrium in Γ^δ in which players end up playing $(0, 0)$.

Proposition 4. *Let $\Gamma = (P, A, U)$ be a normal-form for two players with $A_1 = A_2 = \{0, 1\}$ in which there exists a Nash equilibrium $\vec{\sigma}$ with full support and in which $(0, 0)$ Pareto improves $\vec{\sigma}$. Suppose that players run protocol $\pi_{2 \times 2}^\delta$ and a player defects from the proposed strategy at some commitment step. Then, there exists a Nash equilibrium $\vec{\sigma}'$ in the resulting game $\Gamma' = (P, A, U')$ such that $u'_i(\vec{\sigma}') \leq u_i(0, 0)$ for all players i .*

Proof. Note that the utility function of both players must satisfy that $u_i(0_i, 1_{-i}) \leq u_i(0_i, 0_{-i})$, since otherwise it would contradict the fact that $(0, 0)$ Pareto improves $\vec{\sigma}$. This means that, if after some commitment step, there exists a Nash equilibrium with full support, this Nash equilibrium is going to give less utility to the players than playing $(0, 0)$ in the original game.

By construction, the type of the players is constant throughout the whole protocol with the exception of players of type (iii), who change to type (ii) at Step 3. Moreover, since $\delta < \min\{|u_i(1_i, 0_{-i}) - u_i(0_i, 0_{-i})|, |u_i(0_i, 1_{-i}) - u_i(1_i, 1_{-i})|\}$ for all players i of type (i) or (iii), even if a player defects, she cannot change her type by doing so. Again, the only exception is at Step 3, where a player of type (iii) can choose to remain of type (iii) or convert to type (iv) instead of converting to type (ii). By Proposition 3, this means that, if a player defects before Step 3 or if a player defects in Step 3 without converting to type (iv), there always exists a Nash equilibrium $\vec{\sigma}'$ with full support in the resulting game, which by our previous argument satisfies $u'_i(\vec{\sigma}') \leq u_i(0, 0)$. The only remaining case is when a player i defects in Step 3 and converts to Step (iv). In this case, because of Step 1 of $\pi_{2 \times 2}^\delta$, one of the following must hold:

- $(0, 0)$ is the outcome that gives the most utility to i .
- $u_i(0_i, 0_{-i}) < u_i(1_i, 0_{-i})$ and $u_i(0_i, 1_{-i}) = u_i(1_i, 1_{-i})$.

In the first case, any Nash equilibrium of the resulting game will suffice. In the latter, it can be easily checked that $(1, 1)$ is a Nash equilibrium of the resulting game that satisfies $u'_i(1, 1) < u_i(0, 0)$. □

D Proof of Theorem 3

The proof of Theorem 3 is based on the one of Theorem 4. In fact, the commitment protocol that we use for the proof consists of two stages. First, players select a maximizing welfare action $\vec{a}^{SW}$ and commit to transfer utilities in such a way that, by the end of this stage, each player i gets exactly x_i utility in $\vec{a}^{SW}$ and $\vec{\sigma}$ remains a Nash equilibrium of the resulting game that is Pareto improved by $\vec{a}^{SW}$. Second, players use the techniques of Section 4 to make $\vec{a}^{SW}$ a Nash equilibrium of the resulting game.

If we follow this approach, the proof of Theorem 3 reduces to construct a commitment protocol that implements the first stage. Moreover, analogously to Proposition 1, constructing such a commitment protocol further reduces to find a path as follows.

Proposition 5. *Suppose that there exists a path γ from Γ to another game $\Gamma' = (P, A, U')$ that satisfies the following conditions.*

(Q1) γ is continuous and finite.

(Q2) For all action profiles $\vec{a}$, $w(\vec{a})$ (i.e., the social welfare on $\vec{a}$) weakly decreases through γ .

(Q3) There exists an action profile $\vec{a}^{SW}$ such that $u_i(\vec{a}^{SW}) = x_i$ for all $i \in P$.

(Q4) For all $\Gamma'' = (P, A, U'') \in \gamma$, $\vec{\sigma}$ is an equilibrium of Γ'' .

(Q5) For each player $i \in P$, the expected utility of playing $\vec{\sigma}$ weakly decreases through γ .

Then, there exists $D' > 0$ such that, for all $\delta \in (0, D')$, there exists a commitment protocol that goes from Γ to Γ' that satisfies the following conditions.

(i) At the beginning and after each commitment step, $\vec{\sigma}$ is a $(L_{\vec{\sigma}}, \delta)$ -strongly punishable Nash equilibrium of the resulting game, where $L_{\vec{\sigma}} := \max_{i \in P} \{x_i - u_i(\vec{\sigma})\}$.

(ii) Each player pays at most δ on each outcome in each commitment step.

The proof of Proposition 5 is pretty much identical to that of Proposition 1. We next that Theorem 3 reduces to finding such a path γ . Suppose that there exists a path from Γ to some other game Γ' that satisfies (Q1), (Q2), (Q3) and (Q4) and (Q5). Since $\vec{a}^{SW}$ strictly Pareto-improves $\vec{\sigma}$ in Γ' , we can apply Theorem 4 to deduce that there exists $D > 0$ such that, for all $\delta \in (0, D)$, there exists a subgame perfect Nash equilibrium in $(\Gamma')_{\perp}^{\delta}$ which players start on Γ' and each player i ends up receiving exactly x_i utility. It is easy to check that this is also a subgame perfect Nash equilibrium in $(\Gamma')^{\delta}$. Similarly, by Proposition 5, there exists $D' > 0$ such that, for all $\delta \in (0, D)$, there is a commitment protocol that satisfies (i) and (ii). This means that, for all $\delta \in (0, \min\{D, D'\})$, players can run the protocol obtained from Proposition 5 to go from Γ to Γ' , and then the protocols used in the proof of Theorem 4 to converge on a payoff profile that gives x_i utility to each player i . It only remains to show that there always exists a punishment equilibrium, even if a player defects from the main protocol. Property (ii) of Proposition 5 guarantees its existence between Γ and Γ' . After Γ' , the punishment equilibrium is provided by the protocols used in Section 4.

The argument above shows that Theorem 3 reduces to find a path γ that satisfies the properties of Proposition 5. The following proposition shows how to construct such a path.

Proposition 6. Suppose without loss of generality that $\vec{a}^{SW} := (a_1^1, a_2^2, \dots, a_n^2)$ is a welfare maximizing outcome of Γ . Consider the path $\gamma(\lambda)$ where the utilities of the players are defined by

$$(U(\lambda))_i(\vec{a}) := \begin{cases} u_i(\vec{a}) + \lambda \cdot (x_i - u_i(\vec{a}^{SW})) & \text{if } u_i(\vec{a}^{SW}) > x_i \text{ or } \vec{a} \text{ is of the form } (a_i^j, \vec{a}_{-i}^{SW}) \\ u_i(\vec{a}) - \lambda \frac{\sigma_{i+1}^1}{\sigma_{i+1}^2} (x_i - u_i(\vec{a}^{SW})) & \text{if } u_i(\vec{a}^{SW}) < x_i \text{ and } \vec{a} \text{ is of the form } (a_i^j, a_{i+1}^2, \vec{a}_{-\{i, i+1\}}^{SW}) \\ u_i(\vec{a}) & \text{Otherwise} \end{cases}$$

for $\lambda \in [0, 1]$, where we identify player $n+1$ with player 1. This path satisfies (Q1), (Q2), (Q3), (Q4) and (Q5).

Proof. It is easy to check that properties (Q1) and (Q3) are satisfied by construction. To see that it also satisfies (Q2), note that the only outcomes $\vec{a}$ that may increase their social welfare are those of the form $(a_i^j, \vec{a}_{-i}^{SW})$ for some $i \in P$ and $j \leq N_i$. In order to check that the social welfare on these outcomes do not increase, note that the social welfare on a given outcome is a linear function, which means that we just have to check that the social welfare on these outcomes when $\lambda = 1$ is less or equal than when $\lambda = 0$. If we denote by $S(\vec{a})$ the set of players k such that $u_k(\vec{a}^{SW}) < x_k$ and $\vec{a}$ is of the form $(a_i^j, \vec{a}_{-i}^{SW})$, we have that

$$\begin{aligned} \omega(1)(\vec{a}) - \omega(0)(\vec{a}) &= \sum_{\{k : u_k(\vec{a}^{SW}) > x_k\}} (x_k - u_k(\vec{a}^{SW})) + \sum_{k \in S(\vec{a})} (x_k - u_k(\vec{a}^{SW})) \\ &\leq \sum_{k \in P} (x_k - u_k(\vec{a}^{SW})) \\ &= \left(\sum_{k \in P} x_k \right) - \left(\sum_{k \in P} u_k(\vec{a}^{SW}) \right), \end{aligned}$$

and both sums are equal since $\vec{a}^{SW}$ is a welfare maximizing outcome.

It is straightforward to check that $f_{\vec{\sigma}}$ is invariant, which implies that γ satisfies (Q4). To see this, note that, if $u_i(\vec{a}^{SW}) > x_i$, then i loses utility on all outcomes at the same rate. Meanwhile, if $u_i(\vec{a}^{SW}) < x_i$, i 's utility is only altered in outcomes of the form $(a_i^j, a_{i+1}^1, \vec{a}_{\{i, i+1\}}^{SW})$ and $(a_i^j, a_{i+1}^2, \vec{a}_{\{i, i+1\}}^{SW})$. It only remains to show that γ satisfies (Q5). If i is a player such that $u_i(\vec{a}^{SW}) > x_i$, (Q5) is satisfied automatically since her utility in all outcomes decreases with λ . If i is a player such that $u_i(\vec{a}^{SW}) < x_i$, we have that

$$(U(\lambda)_i)(\vec{\sigma}) = u_i(\vec{\sigma}) + \lambda \cdot (x_i - u_i(\vec{a}^{SW})) \cdot \prod_{j \neq i} \sigma_j^1 - \lambda \frac{\sigma_{i+1}^1}{\sigma_{i+1}^2} (x_i - u_i(\vec{a}^{SW})) \cdot \sigma_{i+1}^2 \cdot \prod_{j \notin \{i, i+1\}} \sigma_j^1 = u_i(\vec{\sigma}),$$

which is invariant. $\square$

E Proof of Theorem 1

Before proving Theorem 1 we need the following notation. Given a game Γ , we denote by $\vec{x}_{\Gamma}$ be the ordered vector of all utilities in the game. For instance, if Γ is a two-player game in which $A_1 = A_2 = \{0, 1\}$, then

$$\vec{x}_{\Gamma} = (u_1(0, 0), u_1(0, 1), u_1(1, 0), u_1(1, 1), u_2(0, 0), u_2(0, 1), u_2(1, 0), u_2(1, 1)).$$

Moreover, given a point $\vec{x} \in \mathbb{R}^N$, we denote by $\|x\|_2$, $\|x\|_1$ and $\|x\|_{\infty}$ its Euclidean, Manhattan and infinite norm, respectively. These are defined as follows:

$$\begin{aligned} \|x\|_2 &= (x_1^2 + x_2^2 + \dots + x_N^2)^{\frac{1}{2}} \\ \|x\|_1 &= |x_1| + |x_2| + \dots + |x_N| \\ \|x\|_{\infty} &= \max\{|x_1|, |x_2|, \dots, |x_N|\} \end{aligned}$$

An important property of these norms is that they are *equivalent*. This means that, for each pair of norms $\|\cdot\|_A$ and $\|\cdot\|_B$ from the ones above, there exists a constant $C_{A,B}$ such that $\|\vec{x}\|_A \leq C_{A,B} \cdot \|\vec{x}\|_B$ for all $\vec{x} \in \mathbb{R}^N$.

With this we proceed with the proof. We must also note that, to ease the notation, we will often use $\vec{\sigma}$ to denote the vector of probabilities by which players play their respective actions from their action sets when playing $\vec{\sigma}$ (as in the definition of $f_{\vec{\sigma}}$ and $r_{\vec{\sigma}}$ in Section 2.3).

Proof. Consider the function $F_{\vec{\sigma}}(\vec{x}, \vec{p}) : \mathbb{R}^{(n \cdot \prod_{i=1}^n |A_i| + \sum_{i=1}^n |\sigma_i(A_i)|)} \rightarrow \mathbb{R}^{\sum_{i=1}^n |\sigma_i(A_i)|}$, which is defined identically to $f_{\vec{\sigma}}$ except that we view the utilities of the players as additional variables, as opposed to fixed parameters (note that there are $n \cdot \prod_{i=1}^n |A_i|$ of such variables). By construction, we have that $F_{\vec{\sigma}}(\vec{x}_{\Gamma}, \vec{p}) = f_{\vec{\sigma}}(\vec{p})$, and in particular $F_{\vec{\sigma}}(\vec{x}_{\Gamma}, \vec{\sigma}) = 0$. Since $|Df_{\vec{\sigma}}(\vec{\sigma})| \neq 0$, by the implicit function theorem, there exists a C^{∞} function $g : \mathbb{R}^{(n \cdot \prod_{i=1}^n |A_i|)} \rightarrow \mathbb{R}^{\sum_{i=1}^n |\sigma_i(A_i)|}$ such that $g(\vec{x}_{\Gamma}) = \vec{\sigma}$ and $F_{\vec{\sigma}}(\vec{x}, g(\vec{x})) = 0$ for all $\vec{x}$ in some neighborhood V around $(\vec{x}_{\Gamma}, \vec{\sigma})$.

To show that $\vec{\sigma}$ is strongly robust, fix $\epsilon > 0$. Since the players' utilities is a continuous function of $\vec{\sigma}$, we have that there exists $\epsilon' > 0$ such that $\|\vec{\sigma}' - \vec{\sigma}\|_2 < \epsilon' \implies |u_i(\vec{\sigma}) - u_i(\vec{\sigma}')| < \epsilon/2$. Let $\delta_1 > 0$ be a value such that $\|\vec{x}_{\Gamma} - \vec{x}\|_2 < \delta_1 \implies \|g(\vec{x}_{\Gamma}) - g(\vec{x})\|_2 < \epsilon'$ (which exists since g is continuous), and let $\delta_2 > 0$ be another value such that $\|\vec{x}\|_2 < \delta_2 \implies \|\vec{x}\|_1 < \epsilon/2$. Let $\delta' = \min(\delta_1, \delta_2)$ and let $\delta > 0$ be a value such that $\|\vec{x}\|_{\infty} < \delta \implies \|\vec{x}\|_2 < \delta'$. Note that δ_2 and δ exist since the infinite, Manhattan and Euclidean norms are equivalent.

Let Γ' be a normal-form game such that $d(\Gamma, \Gamma') < \delta$ and let u'_i denote the utility of player i in Γ' . Since Γ' satisfies that $d(\Gamma, \Gamma') < \delta$, we have that $\|\vec{x}_\Gamma - \vec{x}_{\Gamma'}\|_\infty < \delta$, and therefore that $\|\vec{x}_\Gamma - \vec{x}_{\Gamma'}\|_2 < \delta'$. This in turn implies that $\|g(\vec{x}_\Gamma) - g(\vec{x}_{\Gamma'})\|_2 < \epsilon'$, which is equivalent to $\|\vec{\sigma} - g(\vec{x}_{\Gamma'})\|_2 < \epsilon'$, and therefore $|u_i(\vec{\sigma}) - u_i(g(\vec{u}_{\Gamma'}))| < \epsilon/2$. It follows that

$$|u_i(\vec{\sigma}) - u'_i(g(\vec{u}_{\Gamma'}))| \leq |u_i(\vec{\sigma}) - u_i(g(\vec{u}_{\Gamma'}))| + |u_i(g(\vec{u}_{\Gamma'})) - u'_i(g(\vec{u}_{\Gamma'}))| < \frac{\epsilon}{2} + |(u_i - u'_i)(g(\vec{u}_{\Gamma'}))|.$$

To complete the proof, we have to show that $|(u_i - u'_i)(g(\vec{u}_{\Gamma'}))| \leq \epsilon/2$. Note that

$$|(u_i - u'_i)(g(\vec{u}_{\Gamma'}))| \leq \|\vec{x}_\Gamma - \vec{x}'_\Gamma\|_1$$

since all the components of $\vec{\sigma}$ are values between 0 and 1. Since $\|\vec{x}_\Gamma - \vec{x}'_\Gamma\|_2 < \delta \leq \delta_2$, we have that $\|\vec{x}_\Gamma - \vec{x}'_\Gamma\|_1 < \epsilon/2$, which implies that

$$|(u_i - u'_i)(g(\vec{u}_{\Gamma'}))| \leq \epsilon/2,$$

as desired. $\square$

F Computing a Lower Bound for δ in the Case of Two-Player Games

Our goal in this section is to derive an explicit lower bound on the safety threshold δ . Recall that δ must be chosen small enough that a single commitment step cannot eliminate the punishment equilibrium that supports our off-path incentives. We begin with the case where the equilibrium $\vec{\sigma}$ has *partial support*, and then extend the argument to the remaining cases. The partial-support regime is technically simpler because, along our commitment construction, the characteristic function $f_{\vec{\sigma}}$ associated with $\vec{\sigma}$ remains invariant, whereas the residual term $r_{\vec{\sigma}}$ evolves monotonically under successive commitments. As a consequence, it suffices to choose δ so that the desired safety conditions hold in the base game; the same δ then guarantees these conditions throughout the entire commitment scheme.

Using the notation of Section 2.3, if there are only two players, the characteristic function $f_{\vec{\sigma}}$ and the residual function $r_{\vec{\sigma}}$ can be expressed as follows.

$$f_{\vec{\sigma}}(p_1^1, p_1^2, \dots, p_1^{M_1}, p_2^1, p_2^2, \dots, p_2^{M_2}) := \begin{cases} p_1^1 + p_1^2 + \dots + p_1^{M_1} \\ p_2^1 + p_2^2 + \dots + p_2^{M_2} \\ u_1(a_1^1, \vec{p}_2) - u_1(a_1^2, \vec{p}_2) \\ u_1(a_1^1, \vec{q}) - u_1(a_1^3, \vec{p}_2) \\ \vdots \\ u_1(a_1^1, \vec{p}_2) - u_1(a_1^{M_1}, \vec{p}_2) \\ u_2(a_2^1, \vec{p}_1) - u_2(a_2^2, \vec{p}_1) \\ u_2(a_2^1, \vec{p}_1) - u_2(a_2^3, \vec{p}_1) \\ \vdots \\ u_2(a_2^1, \vec{p}_1) - u_2(a_2^{M_2}, \vec{p}_1), \end{cases}$$

$$r_{\vec{\sigma}}(p_1^1, p_1^2, \dots, p_1^{M_1}, p_2^1, p_2^2, \dots, p_2^{M_2}) := \begin{cases} u_1(a_1^1, \vec{p}_2) - u_1(a_1^{M_1+1}, \vec{p}_2) \\ u_1(a_1^1, \vec{p}_2) - u_1(a_1^{M_1+2}, \vec{p}_2) \\ \vdots \\ u_1(a_1^1, \vec{p}_2) - u_1(a_1^{N_1}, \vec{p}_2) \\ u_2(a_2^1, \vec{p}_1) - u_2(a_2^{M_2+1}, \vec{p}_1) \\ u_2(a_2^1, \vec{p}_1) - u_2(a_2^{M_2+2}, \vec{p}_1) \\ \vdots \\ u_2(a_2^1, \vec{p}_1) - u_2(a_2^{N_2}, \vec{p}_1), \end{cases}$$

It is easy to check that, if $\vec{\sigma}$ is non-degenerate, M_1 must be equal to M_2 (we will use M to denote this common value). Moreover, note that both $f_{\vec{\sigma}}$ and $r_{\vec{\sigma}}$ are linear, since $u_1(a_1^i, \vec{p}_2) = \sum_{j=1}^{M_2} u_1(a_1^i, a_2^j) \cdot p_2^j$ and $u_2(a_2^i, \vec{p}_1) = \sum_{j=1}^{M_1} u_2(a_2^i, a_1^j) \cdot p_1^j$. Therefore, we can express $f_{\vec{\sigma}}(p_1^1, \dots, p_2^M)$ and $r_{\vec{\sigma}}(p_1^1, \dots, p_2^M)$ as $A \cdot \vec{p}$ and $R \cdot \vec{p}$, where A is a square $2M \times 2M$ matrix and R is a rectangular $(N_1 + N_2 - 2M) \times 2M$ matrix.

Recall that, by construction, the non-zero probabilities in the support of $\vec{\sigma}$ are precisely the ones obtained by solving $A \cdot \vec{p} = \vec{y}$, with $\vec{y} = (1, 1, 0, 0, 0, \dots, 0)$. Therefore, it is sufficient to find a value of δ such that, for all matrices A_δ and R_δ that are obtained by modifying the coefficients of A and R by at most δ (except the ones in the first two rows of A), the solution of $A_\delta \cdot \vec{p} = \vec{y}$ satisfies that

- (a) $p_i > 0$ for all $i \in [2M]$
- (b) $(R \cdot \vec{p})_i > 0$ for all $i \in [2M]$.

Note that for (a) it is sufficient to ensure that $|p_i - \sigma_i| < \sigma_{\min}$ for all $i \in [2M]$, where $\sigma_{\min}$ is the minimum non-zero probability in the support of $\vec{\sigma}$. For (b), denote by Δ the minimum coordinate of $R \cdot \vec{\sigma}$, and by $R_{\max}$ the maximum coefficient of R . We want δ to be small enough so that that $|R_i^\delta \cdot \vec{p} - R_i \cdot \vec{\sigma}| < \Delta$. We can write $|R_i^\delta \cdot \vec{p} - R_i \cdot \vec{\sigma}|$ as follows.

$$|R_i^\delta \cdot \vec{p} - R_i \cdot \vec{\sigma}| = |(R_i^\delta - R_i) \cdot \vec{p} + R_i \cdot (\vec{p} - \vec{\sigma})| \leq \delta \sum_{i=1}^{2M} p_i + |R_i \cdot (\vec{p} - \vec{\sigma})|.$$

We have that $\sum_{i=1}^{2M} p_i = 2$, and $|R_i \cdot (\vec{p} - \vec{\sigma})| \leq |R_{\max}| \sum_{i=1}^{2M} |p_i - \sigma_i|$. Therefore, in order to satisfy (b), it is sufficient that $\delta < \frac{\Delta}{4}$ and that $|p_i - \sigma_i| < \frac{\Delta}{4M \cdot R_{\max}}$ for all $i \in [2M]$.

The argument above shows that we can reduce satisfying (a) and (b) to finding δ such that $|p_i - \sigma_i| < \min\{\sigma_{\min}, \frac{\Delta}{4M \cdot R_{\max}}\}$. Since $\vec{p} = A_\delta^{-1} \cdot \vec{y}$, we have that $\vec{p}$ is the sum of the first two columns of A_δ^{-1} , and therefore that $|p_i - \sigma_i| < 4M \cdot a_\delta$, where a_δ is the maximum difference between a coefficient in A_δ^{-1} and A^{-1} . The coefficients of A_δ^{-1} can be computed using the fact that it is the transpose of the cofactor matrix of A_δ , divided by the determinant of A . More precisely,

$$A_\delta^{-1}(i, j) = \frac{(-1)^{i+j} \det(A_\delta^{(j,i)})}{\det(A_\delta)}.$$

Therefore, we have that

$$|A_\delta^{-1}(i, j) - A^{-1}(i, j)| = \frac{(\det(A_\delta^{(j,i)}) - \det(A^{(j,i)})) \det(A) + (\det(A) - \det(A_\delta)) \det(A^{(j,i)})}{\det(A_\delta) \det(A)}.$$

Using the fact that $\det(A) = \sum_{\tau \in S_{2M}} \text{sgn}(\tau) \cdot \prod_{i=1}^{2M} a_{i, \tau(i)}$, we can bound

$$|\det(A_\delta) - \det(A)| \leq (2M)! ((a_{\max} + \delta)^{2M} - a_{\max}^{2M}),$$

where $a_{\max}$ is the absolute value of the maximum coefficient of A (in absolute value). We can use this to bound $|A_\delta^{-1}(i, j) - A^{-1}(i, j)|$.

$$|A_\delta^{-1}(i, j) - A^{-1}(i, j)| \leq (2M)! ((a_{\max} + \delta)^{2M} - a_{\max}^{2M}) \left| \frac{\det(A) + \det(A^{(j,i)})}{\det(A_\delta) \det(A)} \right|.$$

Select δ such that $|\det(A_\delta) - \det(A)| < \frac{|\det(A)|}{2}$. From the inequalities above, it can be easily checked that it is sufficient to pick δ such that

$$\delta < \left(\frac{|\det(A)|}{2 \cdot (2M)!} + a_{\max}^{2M} \right)^{1/2M} - a_{\max}.$$

Then, we get that

$$|A_\delta^{-1}(i, j) - A^{-1}(i, j)| \leq (2M)! ((a_{\max} + \delta)^{2M} - a_{\max}^{2M}) \cdot 2C_A,$$

where C_A is the maximum value of $\left| \frac{\det(A) + \det(A^{(j,i)})}{\det(A)^2} \right|$ (between all i, j). Combining all our arguments so far, for our purposes it suffices to choose δ such that $|A_\delta^{-1}(i, j) - A^{-1}(i, j)| < \frac{1}{4M} \min\{\sigma_{\min}, \frac{\Delta}{4M \cdot R_{\max}}\}$ while simultaneously satisfying that $|\det(A_\delta) - \det(A)| < \frac{|\det(A)|}{2}$ and $\delta < \frac{\Delta}{4}$. We have already found a lower bound for the latter two conditions. For the first one, if we define $\epsilon := \frac{1}{4M} \min\{\sigma_{\min}, \frac{\Delta}{4M \cdot R_{\max}}\}$, we can use the bound on $|A_\delta^{-1}(i, j) - A^{-1}(i, j)|$ to get

$$\delta < \left(\frac{\epsilon}{(2M)! \cdot 2C_A} + a_{\max}^{2M} \right)^{1/2M} - a_{\max}.$$

The case when $\vec{\sigma}$ has full support is analogous, with the only difference being that the parameters used in the previous computation (e.g., $a_{\max}$) are modified throughout the procedure. Note, however, that the path γ that players should follow is determined beforehand, and does not depend on δ . Therefore, the computation of the lower bound for δ is identical, except for the fact that we should use the minimum/maximum value of these parameters in γ . This is computationally tractable since, in our construction, γ is piecewise linear.