Game Theory with Simulation of Other Players

Vojtech Kovarik, **Caspar Oesterheld**, **Vincent Conitzer**
Foundations of Cooperative AI Lab (FOCAL), Carnegie Mellon University
vojta.kovarik@gmail.com, oesterheld@cmu.edu, conitzer@cs.cmu.edu

Abstract

Game-theoretic interactions with AI agents could differ from traditional human-human interactions in various ways. One such difference is that it may be possible to simulate an AI agent (for example because its source code is known), which allows others to accurately predict the agent's actions. This could lower the bar for trust and cooperation. In this paper, we formalize games in which one player can simulate another at a cost. We first derive some basic properties of such games and then prove a number of results for them, including: (1) introducing simulation into generic-payoff normal-form games makes them easier to solve; (2) if the only obstacle to cooperation is a lack of trust in the possibly-simulated agent, simulation enables equilibria that improve the outcome for both agents; and however (3) there are settings where introducing simulation results in strictly worse outcomes for both players.

1 Introduction

Game theory is in principle agnostic as to the nature of the players: besides individual human beings, they can be households, firms, countries, and indeed AI agents. Nevertheless, throughout most of the development of game theory, game theorists have had in mind players that were either humans or entities whose decisions were taken by humans; and as with any theory, the examples one has in mind while developing that theory are likely to affect its focus. If we try to re-develop game theory specifically with AI agents in mind, how might the theory turn out different? Of course, theorems in traditional game theory will not suddenly become false just because of the change in focus. Instead, we would expect any difference to consist in the kinds of settings and phenomena for which we develop models, analysis, and computational tools.

In this paper, we focus on one specific phenomenon that is more pertinent in the context of AI agents: agents being able to *simulate* each other. If an agent's source code is available, another agent can simulate what the former agent will do, which intuitively appears to significantly change the game strategically. We consider settings in which one agent can simulate another, and if they do so, they learn what the other agent will do in the actual game; however, simulating comes at a cost to the simulator, and therefore it is not immediately clear whether and when simulation will actually be used in equilibrium. In particular, we are interested in understanding whether and when the availability of such simulation results in play that is more cooperative. For example, in settings where *trust* is necessary for cooperative behavior [Berg *et al.*, 1995a], one may expect that the ability to simulate the other player can help to establish this trust. But does this in fact happen in equilibrium? And if so, does the ability to simulate foster cooperation in all games, or are there games where it backfires? Are we even able to compute equilibria of games with the ability to simulate?

In terms of related work, our setting is similar to the one of credible commitment [von Stackelberg, 1934], except that one needs to decide whether to pay for allowing the *other* player to commit. Another perspective is that we study a program equilibria [Tennenholtz, 2004], except that only *one* player's program can read the other's source code, and has to pay a cost to do so. For further references and a more detailed discussion, see Section 7.

In the remainder of this introduction, we describe a specific example of a trust game and use it to overview the technical results presented later. We also give several examples that illustrate how simulation can lead to different results when moving beyond trust games. ***For a quick overview, the key takeaways are in Section 1.1, highlighted in italics.***

1.1 Overview and Illustrative Examples

Trust Game As a motivation, consider the following Trust Game (depicted in Fig. 1; our TG is a variation on the traditional one from Berg *et al.* [1995b]). Alice has \$100k in savings, which are currently sitting in her bank account at a 0% interest rate. She is considering hiring an AI assistant from the company Bobble to manage her savings instead. If Bobble and its AI *cooperate* with her, the collaboration generates a profit of \$50k, to be split evenly between her and Bobble. However, Alice is reluctant to *trust* Bobble, which might have instructed

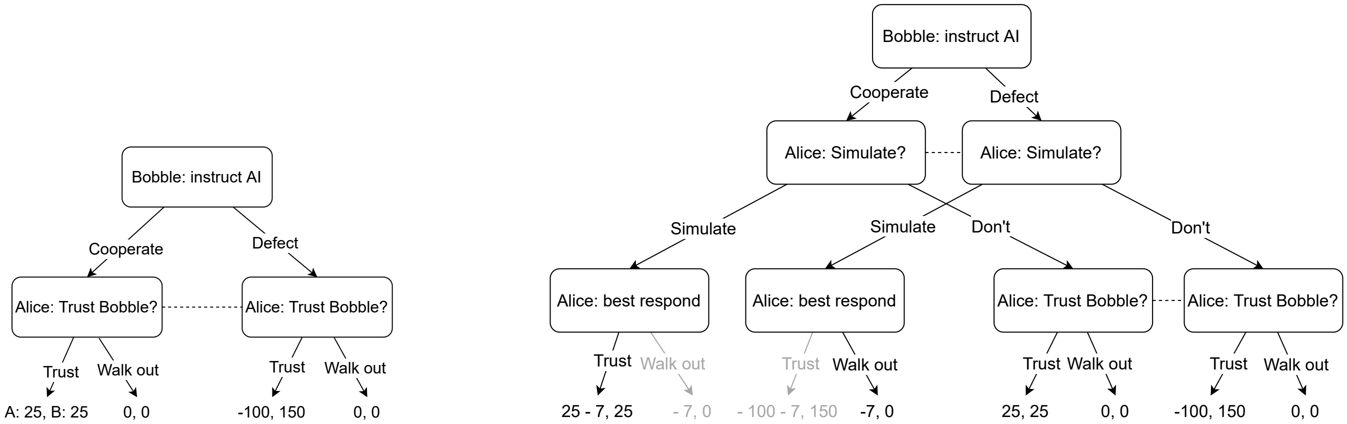


Figure 1: The underlying trust game TG (left) and the corresponding simulation game $\mathrm{TG}_{\mathrm{sim}}$ (right).

the AI to *defect* on Alice by pretending to malfunction, while siphoning off all of the $150k. In fact, the only Nash equilibria of this scenario are ones where Bobble defects on Alice with high probability, and Alice, expecting this, *walks out* on Bobble.

Adding simulation Dismayed by their inability to make a profit, Bobble decides to share with Alice a portion of the AI's source code. This gives Alice the ability to spend $7k on hiring a programmer, to *simulate* the AI in a sandbox and learn whether it is going to cooperate or defect. Crucially, we assume that the AI either does not learn whether it has been simulated or is unable to react to this fact. We might hope that this will ensure that Alice and Bobble can reliably cooperate. However, perhaps Alice will try to save on the *simulation cost* and trust Bobble blindly instead — and perhaps Bobble will bet on this scenario and instruct their AI to defect.

To analyze this modified game $\mathrm{TG}_{\mathrm{sim}}$, note that when Alice simulates, the only sensible followup is to trust Bobble if and only if the simulation reveals they instructed the AI to cooperate. As a result, the normal-form representation of $\mathrm{TG}_{\mathrm{sim}}$ is equivalent to the normal form of the original game TG with a single added action for Alice (Fig. 2). Analyzing $\mathrm{TG}_{\mathrm{sim}}$ reveals that it has two types of Nash equilibria. In one, Bobble defects with high probability and Alice, expecting this, walks out without bothering to simulate. In the other, Bobble still sometimes defects ($\pi_{\mathrm{B}}(\mathrm{D}) = {}^{7}/_{100}$), but not enough to stop Alice from cooperating altogether. In response, Alice simulates often enough to stop Bobble from outright defection ($\pi_{\mathrm{A}}(\mathsf{S}) = 1 - {}^{25}/_{150} = {}^{5}/_{6}$), but also sometimes trusts Bobble blindly ($\pi_{\mathrm{A}}(\mathrm{T}) = {}^{25}/_{150} = {}^{1}/_{6}$). In expectation, this makes Alice and Bobble better off by $16.25k, resp. $25k relative to the (*defect, walk-out*) equilibrium.

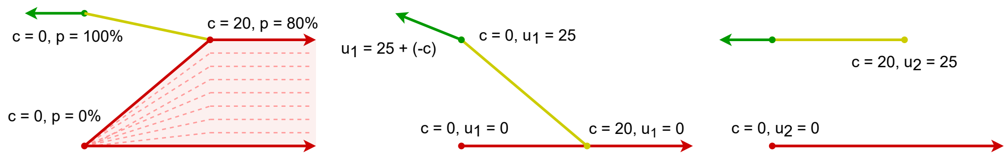
More generally, we can also consider $\mathrm{TG}^{\mathtt{c}}_{\mathrm{sim}}$, a parametrization of $\mathrm{TG}_{\mathrm{sim}}$ where simulation costs some $\mathtt{c} \in \mathbb{R}$. As shown in Figure 2, the equilibria of $\mathrm{TG}^{\mathtt{c}}_{\mathrm{sim}}$ are similar to the special case $\mathtt{c} = 7$ for a wide range of $\mathtt{c}$.

Generalizable properties of the trust game The analysis of Figure 2 illustrates several trends that hold more generally: First, *when simulation is subsidized, the simulation game turns into a "pure commitment game" where the simulated player is the Stackelberg leader* (Prop. 2 (i)). Conversely, *when simulation is prohibitively costly, the simulation game is equivalent to the original game* (Prop. 2 (ii)). Third, the simulation game has a finite number of breakpoints between which individual equilibria change continuously — more specifically, the simulator's strategy does not change at all while the simulated player's strategy changes linearly in $\mathtt{c}$ (Prop. 6). Informally speaking, *simulation games have piecewise constant/linear equilibrium trajectories.* A corollary of this observation is that *it is* not *the case that as simulation gets cheaper, the simulator must use it more and more often* (Fig. 2). Fourth, the indifference principle implies that when the simulator simulates with a nontrivial probability (i.e., neither 0 nor 1), *the value of information of simulating must be precisely equal to the simulation cost.* This also implies that *any pure NE of the original game is also a NE of the simulation game* for any $\mathtt{c} \geq 0$ (Prop. 7). Finally, we saw that *at* $\mathtt{c} = 0$, *the* outcome *of the simulation game becomes deterministic despite the* strategy *of the simulator being stochastic.* (For example, in the NE where Bobble always cooperates, Alice will always end up trusting him — either directly or after first simulating.) In Section 5, we show that this result holds quite generally but not always. Using this result, *we can find the equilibria of generic normal-form games with cheap simulation in linear time* (Thm. 2).

Different effects of simulation There are classes of games in which simulation behaves similarly to the Trust Game above. Indeed, in Theorem 3, we prove that *simulation leads to a strict Pareto improvement in generalized trust games* with generic payoffs (defined in Section 6). However, simulation can also affect games quite differently from what we saw so far. For example, *simulation can benefit either of the players at the cost of the other, or even be harmful to both of them.* Indeed, simulation benefits only the simulator in zero-sum games, benefits only the simulated player in the Commitment

	Coop	Defect
Trust	25, 25	-100, 150
Walk -out	0, 0	0, 0

	Coop	Defect
Trust	25, 25	-100, 150
Walk -out	0, 0	0, 0
Sim	25 - 7, 25	-7, 0



simulation cost	NE of $\mathrm{TG}^{\mathrm{c}}_{\mathrm{sim}}$, given as the probs. of (*simulate, trust, walk-out*) and (*cooperate, defect*)		
$(-\infty, 0)$	(*simulate, cooperate*)		
$(0, 20)$	$(1 - {}^{25}\!/_{150}, {}^{25}\!/_{150}, 0)$ and $(1 - {}^{c}\!/_{100}, {}^{c}\!/_{100})$	*walk-out* and $({}^{c}\!/_{25}, 1 - {}^{c}\!/_{25})$	(*walk-out, defect*)
$(20, \infty)$		*walk-out* and $({}^{100}\!/_{125}, {}^{25}\!/_{125})$	(*walk-out, defect*)

Figure 2: Top left: The normal-form representation of the trust game from Figure 1, before and after adding simulation. Bottom: The extremal equilibria of $\mathrm{TG}^{\mathrm{c}}_{\mathrm{sim}}$. The non-extremal NE are precisely the convex combinations of the last two columns. Top right: The cooperation probability and utilities under each of these NE. The non-extremal NE are light red, the dashed lines illustrate the NE trajectories from Proposition 6. Note that all the red NE (i.e., with $\pi_1(\mathrm{WO}) = 1$) yield $u_1 = u_2 = 0$.

	L	R
U	0, 3	1, 2
D	2, 1	0, 0

	C	C'	D
T	25, 25	-999, -999	-100,0
T'	-999, -999	25, 25	-100,0
WO	0, 0	0, 0	0, 0

Figure 3: Left: Commitment game, where the row player prefers to *not* be able to simulate. For details, see Example 18. Right: A variant of Trust Game with multiple simulation NE.

Game (Fig. 3), and harms both if cooperation is predicated upon the simulated player's ability to maintain privacy (Ex. 19). In fact, there are even *cases where the Pareto optimal outcome requires simulation to be neither free nor prohibitively expensive* (Ex. 30). Finally, with multiple, incompatible ways to cooperate, a *game might admit multiple simulation equilibria* (i.e., multiple NE with $\pi_1(\mathtt{S}) > 0$; cf. Fig. 3).

1.2 Outline

The remainder of the paper is structured as follows. First, we recap the necessary background (Section 2). In Section 3, we formally define simulation games and describe their basic properties. In Section 4, we prove several structural properties of simulation games; while these are instrumental for the subsequent results, we also find them interesting in their own right. Afterwards, we analyze the computational complexity of solving simulation games (Section 5) and the effects of simulation on the players' welfare (Section 6). Finally, we review the most relevant existing work (Section 7), summarize our results, and discuss future work (Section 8). The detailed proofs are presented in the appendix.

2 Background

A two-player **normal-form game** (NFG) is a pair $\mathcal{G} = (\mathcal{A}, u)$ where $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2 \neq \emptyset$ is a finite set of **actions** and $u = (u_1, u_2) : \mathcal{A} \to \mathbb{R}^2$ is the **utility function**. We use P1 and P2 as shorthands for "player one" and "player two". For finite X, $\Delta(X)$ denotes the set of all probability distributions over X. A **strategy** (or policy) **profile** is a pair $\pi = (\pi_1, \pi_2)$ of **strategies** $\pi_i \in \Delta(\mathcal{A}_i)$. We denote the set of all strategies as $\Pi = \Pi_1 \times \Pi_2$. A strategy is **pure** if it has support $\mathrm{supp}(\pi_i)$ of size 1. We identify such strategy with the corresponding action.

For $\pi \in \Pi$, $u_i(\pi) := \sum_{(a,b)\in\mathcal{A}} \pi_1(a)\pi_2(b)u_i(a,b)$ is the **expected utility** of π. π_1 is said to be a **best response** to π_2 if $\pi_1 \in \arg\max_{\pi_1' \in \Pi_1} u_1(\pi_1', \pi_2)$; $\mathrm{br}(\pi_2)$ denotes the set of all pure best responses to π_2. Since the **best-response utility** "$u_1(\mathrm{br}, \cdot)$" is uniquely determined by π_2, we denote it as $u_1(\mathrm{br}, \pi_2) := \max_{a\in\mathcal{A}_1} u_1(a, \pi_2)$. (The analogous definitions apply for P2 and π_1.) A **Nash equilibrium** (NE) is a strategy profile (π_1, π_2) under which each player's strategy is a best response to the strategy of the other player. We use $\mathrm{NE}(\mathcal{G})$ to denote the set of all Nash equilibria of $\mathcal{G}$.

Informally, a **pure-commitment equilibrium** (cf. von Stackelberg [1934]) is a subgame-perfect equilibrium of the game in which the leader first commits to a pure action, after which the follower sees the commitment and best-responds, possibly stochastically. Since our formalism will assume that P1 is the simulator, we naturally encounter situations where P2 acts as the leader. Formally, we will use $\mathrm{SE}^{\mathrm{P2}}_{\mathrm{pure}}(\mathcal{G})$ to denote all pairs (ψ_{br}, b) where the **optimal commitment** $b \in \mathcal{A}_2$ and P1's best-response policy $\psi_{\mathrm{br}} : b' \in \mathcal{A}_2 \mapsto \psi_{\mathrm{br}}(b') \in \Delta(\mathrm{br}(b')) \subseteq \Delta(\mathcal{A}_1)$ satisfy $b \in \arg\max_{b'\in\mathcal{A}_2} \mathbf{E}_{a\sim\psi_{\mathrm{br}}(b')}\, u_1(a, b')$.

As an auxiliary definition, we say that $\mathcal{G}$ admits **no best-response utility tiebreaking** by P1 if for every pure strategy b of P2, any two pure best-responses $a, a' \in \mathrm{br}(b)$ give the same utility $u_2(a, b) = u_2(a', b) =: u_2(\mathrm{br}, b)$. Note that in such a game, any element of $\mathrm{SE}^{\mathrm{P2}}_{\mathrm{pure}}(\mathcal{G})$ can be identified with a pair of pure strategies (a, b) for which $a \in \mathrm{br}(b)$ and $b \in \arg\max_{b\in\mathcal{A}_2} u_2(\mathrm{br}, b)$.

3 Simulation Games

In this section, we formally define simulation games and describe their basic properties. To streamline this initial investigation of simulation games, the remainder of this paper makes two simplifying assumptions: First, we assume that when the simulator learns the other agent's action, they always best-respond to it — in other words, they will not execute non-credible threats [Shoham and Leyton-Brown, 2008]. Since this assumption somewhat limits the applicability of the results, we consider moving beyond it a worthwhile future direction. Second, we consider only a single (possibly stochastic) best-response policy against any pure action of the opponent. This could be justified by the simulator using a particular best-response policy for some external reasons, or by the game not allowing best-response tie-breaking in the first place. These assumptions result in the following formal definition:

Definition 1 (Simulation game)**.** *Let $\mathcal{G}$ be a two-player normal-form game, $\mathsf{c} \in \mathbb{R}$ be a **simulation cost**, and fix some best-response policy $\psi_{\mathrm{br}} : b \in \mathcal{A}_2 \mapsto \psi_{\mathrm{br}}(b) \in \Delta(\mathrm{br}(b)) \subseteq \Delta(\mathcal{A}_1)$. The corresponding **simulation game** $\mathcal{G}^{\mathsf{c}}_{\mathrm{sim}}$ is defined as the NFG that is identical to $\mathcal{G}$, except that* P1 *has an additional "simulate" action that corresponds to utilities $u_1(\mathcal{S}, b) := u_1(\mathrm{br}, b) - \mathsf{c}$, $u_2(\mathcal{S}, b) := \mathbf{E}_{a \sim \psi_{\mathrm{br}}(b)} u_2(a, b)$.*

We refer to P1 as the **simulator** and to P2 as the **simulated player**. We use $\mathcal{G}_{\mathrm{sim}}$ to denote the simulation game with unspecified simulation cost c.

3.1 Basic Properties

The first observation we make (Proposition 2) is that if simulation is too costly, then it is never used and the simulation game $\mathcal{G}_{\mathrm{sim}}$ becomes strategically equivalent to the original game $\mathcal{G}$. Conversely, if simulation is subsidized (i.e., a negative simulation cost), then P1 will always use it, which effectively turns $\mathcal{G}_{\mathrm{sim}}$ into a pure Stackelberg game with P2 moving first. (The situation is similar when simulation is free but not subsidized, except that this allows for additional equilibria where the simulation probability is less than 1.)

Proposition 2 (Equilibria for extreme simulation costs)**.** *In any simulation game $\mathcal{G}_{\mathrm{sim}}$, we have:*

(i) For $\mathsf{c} < 0$, simulating is a strongly dominant action. In particular, $\mathrm{NE}(\mathcal{G}^{\mathsf{c}}_{\mathrm{sim}}) \subseteq \mathrm{SE}^{\mathrm{P2}}_{\mathrm{pure}}(\mathcal{G})$.[1]

(ii) For $\mathsf{c} > \max_{a \in \mathcal{A}_1, b \in \mathcal{A}_2} u_1(a, b) - \max_{\pi_1 \in \Pi_1} \min_{\pi_2 \in \Pi_2} u_1(\pi_1, \pi_2)$, $\mathcal{S}$ is a strictly dominated action. In particular, $\mathrm{NE}(\mathcal{G}^{\mathsf{c}}_{\mathrm{sim}}) = \mathrm{NE}(\mathcal{G}_{\mathrm{sim}})$.

3.2 Information-Value of Simulation

The following definition measures the extra utility that the simulator can gain by using the knowledge of the other player's strategy:

Definition 3 (Value of information of simulation)**.** *The **value of information of simulation** for $\pi_2 \in \Pi_2$ is*

$$\mathit{VoI}_{\mathcal{S}}(\pi_2) := \left(\sum\nolimits_{b \in \mathcal{A}_2} \pi_2(b) \max_{a \in \mathcal{A}_1} u_1(a, b)\right) - \max_{\pi_1 \in \Pi_1} u_1(\pi_1, \pi_2).$$

[1]If we allowed P1 to consider all possible best-response policies, $\mathrm{NE}(\mathcal{G}^{\mathsf{c}}_{\mathrm{sim}}) \subseteq \mathrm{SE}^{\mathrm{P2}}_{\mathrm{pure}}(\mathcal{G})$ would turn into equality.

Lemma 4. $\forall \pi_2 : u_1(\mathcal{S}, \pi_2) = u_1(\mathrm{br}, \pi_2) + \mathit{VoI}_{\mathcal{S}}(\pi_2) - \mathsf{c}$.

Lemma 4 implies that $\mathrm{VoI}_{\mathcal{S}}(\pi_2)$ always lies between 0 and the difference between maximum possible u_1 and P1's maxmin value. Moreover, to make P1 simulate with a non-trivial probability, P2 needs to pick a strategy whose value of information is equal to the simulation cost:

Lemma 5 ($\mathrm{VoI}_{\mathcal{S}}$ is equal to simulation cost)**.** *(1) For any $\pi \in \mathrm{NE}(\mathcal{G}^{\mathsf{c}}_{\mathrm{sim}})$, we have $\pi_1(\mathcal{S}) \in (0, 1) \implies \mathit{VoI}_{\mathcal{S}}(\pi_2) = \mathsf{c}$. (2) Moreover, unless $\mathcal{G}$ admits multiple optimal commitments of* P2 *that do not have a common best-response, any $\pi \in \mathrm{NE}(\mathcal{G}^{0}_{\mathrm{sim}})$ has $\mathit{VoI}_{\mathcal{S}}(\pi_2) = 0$.*

(Where, in (2), a set of actions having a common best-response means that $\bigcap_{b \in B} \mathrm{br}(b) \neq \emptyset$.)

4 Structural Properties

In this section, we review several structural properties that appear in simulation games because of the special nature of the simulation action. These results will prove instrumental when determining the complexity of simulation games (Section 5) and predicting the impact of simulation on the players' welfare (Section 6). Moreover, we find these results interesting in their own right.

The first of these properties is that a change of the simulation cost *typically* results in a very particular change in a Nash equilibrium of the corresponding game: The strategy of the simulating player (P1) doesn't change at all, while the simulated player's strategy changes linearly. However, to be technically accurate, we need to make two disclaimers. First, there is a finite number of "atypical" values of c, called breakpoints, where the nature of the NE strategies changes discontinuously.[2] Second, there can be multiple equilibria, which complicates the formal description of the result.

Proposition 6 (Simulation equilibria trajectories are piecewise constant/linear)**.** *For every $\mathcal{G}$, there is a finite set of simulation-cost breakpoint values $-\infty = e_{-1} < 0 = e_0 < e_1 < \cdots < e_k < e_{k+1} = \infty$ such that the following holds: For every $\mathsf{c}_0 \in (e_l, e_{l+1})$ and every $\pi^{\mathsf{c}_0} \in \mathrm{NE}(\mathcal{G}^{\mathsf{c}_0}_{\mathrm{sim}})$, there is a linear mapping $\mathsf{t}_2 : \mathsf{c} \in [e_l, e_{l+1}] \mapsto \pi^{\mathsf{c}}_2 \in \Pi_2$ such that $\mathsf{t}_2(\mathsf{c}_0) = \pi^{\mathsf{c}_0}_2$ and $(\pi^{\mathsf{c}_0}_1, \mathsf{t}_2(\mathsf{c})) \in \mathrm{NE}(\mathcal{G}^{\mathsf{c}}_{\mathrm{sim}})$ for every $\mathsf{c} \in [e_l, e_{l+1}]$.*

Since we were not able to find any existing result that would immediately imply this proposition, we provide our own proof in Appendix B. However, a related result in the context of parameterized linear programming appears in [Adler and Monteiro, 1992, Prop. 2.3]. To get an intuitive sense for why this result holds, recall that in an equilibrium, each player uses a strategy that makes the other player indifferent between the actions in their support. Since P2's payoffs are not affected by c, P1 should keep their strategy constant to keep P2 indifferent, even when c changes. Similarly, increasing c linearly

[2] While all of the non-breakpoint equilibria extend to the corresponding breakpoints as limits (Definition 8), the breakpoints might also admit additional non-limit equilibria, typically convex combinations of the limits (cf. Figure 2).

decreases P1's payoff for the simulate action, so P2 needs to linearly adjust their strategy to bring P1's payoffs back into equilibrium.

A particular corollary of Proposition 6 is that while one might perhaps expect simulation will gradually get used more and more as it gets more affordable, this is in fact not what happens — instead, the simulation rate is dictated by the need to balance the unchanging tradeoffs of the other player.

The second structural property of simulation games is the following refinement of Proposition 2:

Proposition 7 (Gradually recovering the NE of $\mathcal{G}$)**.** *Let π be a NE of $\mathcal{G}$. Then π, as a strategy in $\mathcal{G}^{\mathsf{c}}_{\text{sim}}$ with $\pi_1(\mathcal{S}) := 0$, is a NE precisely when $\mathsf{c} \geq \mathit{VoI}_{\mathcal{S}}(\pi_2)$.*

In particular, $\mathit{VoI}_{\mathcal{S}}(\pi_2)$ is a breakpoint of $\mathcal{G}$.

Together, these two results imply that with $\mathsf{c} = 0$, $\mathcal{G}^0_{\text{sim}}$ may have no NE in common with $\mathcal{G}$. As we increase c, the NE of $\mathcal{G}$ gradually appear in $\mathcal{G}^{\mathsf{c}}_{\text{sim}}$ as well, while the simulation equilibria of $\mathcal{G}_{\text{sim}}$ (i.e., those with $\pi_1(\mathsf{S}) > 0$) gradually disappear, until eventually $\mathrm{NE}(\mathcal{G}^{\mathsf{c}}_{\text{sim}}) = \mathrm{NE}(\mathcal{G})$.

4.1 Equilibria for Cheap Simulation

By combining the concept of value of information with the piecewise constancy/linearity of simulation equilibria, we are now in a position to give a more detailed description of Nash equilibria of games where simulation is cheap. First, we identify the equilibria of $\mathcal{G}_{\text{sim}}$ with $\mathsf{c} = 0$ that might be connected to the equilibria for $\mathsf{c} > 0$:

Definition 8 (Limit equilibrium of $\mathcal{G}_{\text{sim}}$)**.** *A policy profile π^0 is a **limit equilibrium** at $\mathsf{c} = 0$ of $\mathcal{G}_{\text{sim}}$ (or just "limit equilibrium") if it is a limit of some $\pi^{\mathsf{c}_n} \in \mathrm{NE}(\mathcal{G}^{\mathsf{c}_n}_{\text{sim}})$ where $\mathsf{c}_n \to 0_+$.*

As witnessed by the Trust Game (and Table 2 in particular), not every NE of $\mathcal{G}^0_{\text{sim}}$ is a limit equilibrium. Note that this definition automatically implies a stronger condition:

Lemma 9. *For any limit equilibrium π^0 of $\mathcal{G}_{\text{sim}}$, there is some $e > 0$ and π^e_2 such that for every $\mathsf{c} \in [0, e]$, $(\pi^0_1, (1 - \frac{\mathsf{c}}{e})\pi^0_2 + \frac{\mathsf{c}}{e}\pi^e_2)$ is a NE of $\mathcal{G}^{\mathsf{c}}_{\text{sim}}$.*

The following result shows that cheap-simulation equilibria have a very particular structure. Informally, every such NE corresponds to a "baseline" limit equilibrium π^{B} and P2's "deviation policy" π^{D}_2. As the simulation cost increases, P2 gradually deviates away from their baseline, which forces P1 to randomize between their baseline and simulating. While the technical formulation can seem daunting, all of the conditions in fact have quite intuitive interpretations that can be used for locating the simulation equilibria of small games by hand.

Lemma 10 (Structure of cheap-simulation equilibria)**.** *Let $\mathsf{c}_0 \in (0, e_1)$ and suppose that $\mathcal{G}$ admits no best-response utility tiebreaking by* P1*. Then any $\pi \in \mathrm{NE}(\mathcal{G}^{\mathsf{c}_0}_{\text{sim}})$ with $\pi_1(\mathcal{S}) \in (0, 1)$ is of the form $\pi = (\pi_1, \pi^{\mathsf{c}_0}_2)$, where*

$$\begin{aligned} \pi_1 &= (1 - \pi_1(\mathcal{S})) \cdot \pi^B_1 + \pi_1(\mathcal{S}) \cdot \mathcal{S} \\ \pi^{\mathsf{c}}_2 &= \quad (1 - \alpha\mathsf{c}) \quad \cdot \pi^B_2 + \quad \alpha\mathsf{c} \quad \cdot \pi^D_2, \quad \alpha > 0, \end{aligned}$$

and the following holds:

(i) For every $\mathsf{c} \in [0, e_1]$, $(\pi_1, \pi^{\mathsf{c}}_2) \in \mathrm{NE}(\mathcal{G}^{\mathsf{c}}_{\text{sim}})$.

*(ii) $\pi^B \in \Pi$ is some **baseline policy** that satisfies:*

(B1) every action in the support of π^B_1 is a best-response to every action from $\mathrm{supp}(\pi^B_2)$;

(B2) every action in the support of π^B_2 is an optimal commitment by P2 *conditional on* P2 *only using strategies that satisfy (B1).*

*(iii) $\pi^D_2 \in \Pi_2$ is some **deviation policy** that satisfies:*

(D1) No $a \in \mathrm{supp}(\pi^B_1)$ lies in $\mathrm{br}(d)$ for all $d \in \mathrm{supp}(\pi^D_2)$.

(D2) Every $d \in \mathrm{supp}(\pi^D_2)$ satisfies one of

$$u_2(\pi^B_1, \pi^D_2) > u_2(\pi^B) > u_2(\mathrm{br}, \pi^D_2) \qquad (D^>_2)$$

$$u_2(\pi^B_1, \pi^D_2) = u_2(\pi^B) = u_2(\mathrm{br}, \pi^D_2) \qquad (D^=_2)$$

$$u_2(\pi^B_1, \pi^D_2) < u_2(\pi^B) < u_2(\mathrm{br}, \pi^D_2). \qquad (D^<_2)$$

(D3) If $d \in \mathrm{supp}(\pi^D_2)$ satisfies $(D^>_2)$, resp. $(D^<_2)$, it maximizes the attractiveness ratio *r_d, resp. r_d^{-1}*

$$\frac{u_2(\pi^B_1, d') - u_2(\pi^B)}{u_2(\pi^B) - u_2(\mathrm{br}, d')} \;\; \textit{resp.} \;\; \frac{u_2(\mathrm{br}, d') - u_2(\pi^B)}{u_2(\pi^B) - u_2(\pi^B_1, d')}$$

among all $d' \in \mathcal{A}_2$ that satisfy $(D^>_2)$, resp. $(D^<_2)$.

In a generic game, these conditions even imply that both the baseline and deviation policies are pure. (Recall that a property is said to be generic – i.e., typical – if it holds for almost all elements of a set [Rudin, 1987, 1.35]. In particular, "P being true for a **game with generic payoffs**" means that if each payoff is an i.i.d. sample from the uniform distribution over $[0, 1]$, P holds with probability 1.)

Theorem 1 (Equilibria with binary supports)**.** *Let $\mathcal{G}$ be a game with generic payoffs and $\mathsf{c} \in (0, e_1)$. Then all NE of $\mathcal{G}^{\mathsf{c}}_{\text{sim}}$ are either pure or have supports of size two.*

5 Computational Aspects

In this section, we investigate the difficulty of solving simulation games. Since many of the results hold for multiple solution concepts, we formulate them using the phrase "solving a game", with the understanding that this refers to either finding all Nash equilibria, or a single NE, or a single NE with a specific property (e.g., one with the highest social welfare). For a specific game $\mathcal{G}$, we will also use $-\infty < 0 < e_1 < \cdots < e_k < \infty$ to denote the breakpoints of $\mathcal{G}_{\text{sim}}$ (given by Proposition 6).

The immediate implications of the definition of simulation games (resp. of Proposition 2) are the following two results:

Proposition 11 (Simulation games are no harder than general games)**.** *Solving $\mathcal{G}^{\mathsf{c}}_{\text{sim}}$ is at most as difficult as solving a normal-form game where* P1 *has one more action than in $\mathcal{G}$.*

Proposition 12 (Solving $\mathcal{G}_{\text{sim}}$ for extreme c)**.**

(i) For $\mathsf{c} \in (-\infty, 0)$, the time complexity of solving $\mathcal{G}^{\mathsf{c}}_{\text{sim}}$ is $O(|\mathcal{A}|)$.

(ii) For $\mathsf{c} \in (e_k, \infty)$, the time-complexity of solving $\mathcal{G}^{\mathsf{c}}_{\text{sim}}$ is the same as the time-complexity of solving $\mathcal{G}$.

This result leaves unresolved the situation for the intervals $(0, e_1)$ and (e_1, e_2), …, (e_{k-1}, e_k):

Problem 13. What is the complexity of solving simulation games for c in (e_2, e_3), …, and (e_{k-1}, e_k)?

In contrast with Proposition 12 (ii), finding the equilibria at low simulation costs is, typically, straightforward:

Theorem 2 (Cheap-simulation equilibria in generic games)**.** *Let $\mathcal{G}$ be a NFG with generic payoffs and $\mathsf{c} \in (0, e_1)$. Then the time complexity of finding all equilibria of $\mathcal{G}^{\mathsf{c}}_{\text{sim}}$ is $O(|\mathcal{A}|)$.*

Finally, an important case not covered by Theorem 2 are *extensive* form games, whose normal-form representations have non-generic payoffs even when the EFG itself does have generic payoffs. We consider determining the complexity of finding the limit equilibria of such games to be an interesting open problem:

Problem 14. What is the complexity of finding the limit equilibria of $\mathcal{G}_{\text{sim}}$ (i.e., $\text{NE}(\mathcal{G}^{\mathsf{c}}_{\text{sim}})$ for $\mathsf{c} \in (0, e_1)$) when $\mathcal{G}$ is (i) a general NFG or (ii) an EFG with generic payoffs?

6 Effects on Players' Welfare

In this section, we first confirm the hypothesis that simulation is beneficial in settings where the only obstacle to cooperation is the missing trust in the simulated player. We then observe that in general games, simulation can also benefit either player at the cost of the other, or even be harmful to both.

6.1 Simulation in Generalized Trust Games

We now prove that in settings where the *only* obstacle to cooperation is the lack of trust in the possibly-simulated player, simulation enables equilibria that improve the outcome for both players.

Definition 15 (Generalized trust games)**.** *A game $\mathcal{G}$ is said to be a* ***generalized trust game*** *if any pure-commitment Stackelberg equilibrium (where* P2 *is the leader) is a strict Pareto improvement over any $\pi \in$* NE($\mathcal{G}$).

Theorem 3 (Simulation in trust games helps)**.** *Let $\mathcal{G}$ be a generalized trust game with generic payoffs. Then for all sufficiently low c, $\mathcal{G}^{\mathsf{c}}_{\text{sim}}$ admits a Nash equilibrium with $\pi_1(\mathcal{S}) > 0$ that is a strict Pareto improvement over any NE of $\mathcal{G}$.*

Proof sketch. In Appendix A, we construct an equilibrium where P2 mixes between their optimal commitment b (from the pure-commitment equilibrium corresponding to $\mathcal{G}$) and some deviation d while P1 mixes between their best-response to b and simulating. We show that (a, b) forms the baseline policy of this simulation equilibrium, which implies that as $\mathsf{c} \to 0_+$, the eventually becomes a strict Pareto improvement over any NE of $\mathcal{G}$. (And the fact that (a, b) cannot be a NE of $\mathcal{G}$ ensures that we can find a suitable d.) □

6.2 Simulation in General Games

We now investigate the relationship between simulation cost and the players' payoffs in *general* games. We start by listing the two general trends that we are aware of.

The first of the general results is that for the extreme values of c, the situation is always predictable: For $\mathsf{c} < 0$, P1 always simulates (Prop. 2) and making simulation cheaper will increase their utility without otherwise affecting the outcome. Similarly, when c is already so high that P1 never simulates, any further increase of c makes no additional difference.

Second, if P2 could choose the value of c, they would generally be indifferent between all the values within a specific interval (e_i, e_{i+1}). Indeed, this follows from Proposition 6, which implies that P2's utility remains constant between any two breakpoints of $\mathcal{G}_{\text{sim}}$.

The Examples 16-19 illustrate that the players might both agree and disagree about their preferred value of c, and this value might be both low and high.

Example 16 (Both players prefer cheap simulation)**.** In the Alice and Bobble game from Figure 2, each player's favoured NE exists for $\mathsf{c} = 0$.

Example 17 (Only simulator prefers cheap simulation)**.** Consider the "unfair guess-the-number game" where each player picks an integer between 1 and N. If the numbers match, P2 pays 1 to P1. Otherwise, P1 pays 1 to P2. In this game, P2 clearly prefers simulation to be prohibitively costly while P1 prefers as low c as possible.

Example 18 (Only simulator prefers expensive simulation)**.** In the commitment game (Figure 3), introducing free simulation creates a second NE in which P1 is strictly worse off and stops the original NE from being trembling-hand perfect. If simulation were subsidized, the original simulator-preferred NE would disappear completely. (In fact, with $\mathsf{c} > 0$ that is not prohibitively costly, the situation is similar to the $\mathsf{c} = 0$ case.) In summary, this shows that simulation can hurt the simulator, even when using it is free (or even subsidized) and voluntary.

Example 19 (Both players prefer expensive simulation)**.** Consider a Joint Project game where P1 proposes that P2 collaborates with them on a startup. If P2 accepts, their business will be successful, yielding utilities $u_1 = u_2 = 100$. P2 then picks a secure password ($pw \in \{1, \ldots, 26\}^4$) and puts their profit in a savings account protected by that password. Finally, P1 can either do nothing or try to guess P2's password ($g \in \{1, \ldots, 26\}^4$) and steal their money. Successfully guessing the password would result in utilities $u_1 = 200$, $u_2 = -10$, where the -10 comes from opportunity costs. However, if P1 guesses wrong, they will be caught and sent to jail, yielding utilities $u_1 = -999$, $u_2 = 123$ [Smith *et al.*, 2009].

Without simulation, the NE of this game is for the players to collaborate and for P1 to not attempt to guess the password. However, with cheap enough simulation, P1 would simulate P2's choice of password and steal their money — and P2, expecting this, would not agree to the collaboration in the first place. As a result, both players would prefer simulation to be prohibitively expensive.

Example 20 (The preferences depend on equilibrium selection)**.** Consider various mixed-motive games such as the Threat Game (e.g., [Clifton, 2020, Sec. 3-4]), Battle of the Sexes, or Chicken (e.g., Shoham and Leyton-Brown [2008]). Generally, these games have one pure NE that favours P1, a second pure NE that favours P2, and a mixed NE that is strictly worse than either of the pure equilibria for both P1 and P2. By introducing subsidized simulation into such a game, we eliminate both the simulator-favoured pure NE and the dispreferred mixed NE. This can be bad, neutral, or even good news for the simulator, depending on which of the NE would have been selected in the original game. Somewhat relatedly, introducing subsidized simulation destroys the suboptimal equilibria in Stag Hunt and Coordination Game (e.g., Shoham and Leyton-Brown [2008]).

Beyond the examples above, players might even prefer neither $\mathsf{c} = 0$ nor $\mathsf{c} = \infty$ but rather something inbetween:

Example 21 (The preferred c is non-extreme)**.** Informally, the underlying idea behind the example is that the game should have the potential for a positive-sum interaction, but also be unfair towards P1 if they never simulate and unfair towards P2 if P1 always simulates. If we then give each player the option to opt out, the only way either of the players can profit is if simulation is neither free nor prohibitively expensive. For a detailed proof, see Appendix C.

7 Related Work

In terms of the formal framework, our work is closest to the literature on games with commitment [Conitzer and Sandholm, 2006; von Stengel and Zamir, 2010]. This is typically modelled as a Stackelberg game [von Stackelberg, 1934], where one player commits to a strategy while the other player only selects their strategy after seeing the commitment. In particular, Letchford *et al.* [2014] investigates how much the committing player can gain from committing. Commitment in a Stackelberg game is always observed. (An exception is Korzhyk *et al.* [2011], which assumes a fixed chance of commitment observation.) In contrast, the simulation considered in this paper would correspond to a setting where (1) one player pays for having *the other player* make a (pure) commitment and (2) the latter player does not know whether their commitment is observed, as the probability of it being observed is a parameter controlled by the observer. Ultimately, these differences imply that the Stackelberg game results are highly relevant as inspiration, but they are unlikely to have immediate technical implications for our setting (except for when $\mathsf{c} < 0$).

In terms of motivation, the setting that is the closest to our paper is open-source game theory and program equilibria [McAfee 1984; Howard 1988; Rubinstein 1998, Sect. 10.4; Tennenholtz 2004]. In program games, two (or more) players each choose a program that will play on their behalf in the game, and these programs can read each other. To highlight the connection to the present paper, note that one approach to attaining cooperative play in this formalism is to have the programs simulate each other [Oesterheld, 2019]. The setting of the program equilibrium literature differs from ours in two important ways. First, the program equilibrium literature assumes that both players have access to the other player's strategy. (Much of the literature addresses the difficulties of mutual simulation or analysis, e.g., see Barasz *et al.* [2014]; Critch [2019]; Critch *et al.* [2022]; Oesterheld [2022] in addition to the above.) Second, with the exception of time discounting as studied by Fortnow [2009], the program equilibrium formalism assumes that access to the other player's code is without cost.

Another approach to simulation is game theory with translucent players [Halpern and Pass, 2018]. This framework assumes that the players tentatively settle on some strategy from which they can deviate, but doing so has some chance of being visible to the other player. In our terminology, this corresponds to a setting where each player always performs free but unreliable simulation of the other player.

8 Discussion

In this paper, we considered how the traditional game-theoretic setting changes when one player obtains the ability to run an accurate but costly simulation of the other. We established some basic properties of the resulting simulation games. We saw that (between breakpoint values of which there can be only finitely many), their equilibria change piecewise constantly/linearly (for P1/P2) with the simulation cost. Additionally, the value of information of simulating is often equal to the simulation cost. These properties had strong implications for the equilibria of games with cheap simulation and allowed us to prove several deeper results. Our initial hope was that simulation could counter a lack of trust — and this turned out to be true. However, we also saw that the effects of simulation can be ambiguous, or even harmful to both players. This suggests that before introducing simulation to a new setting (or changing its cost), one should determine whether doing so is likely to be beneficial or not. Fortunately, our analysis revealed that for the very general class of normal-form games with generic payoffs, this can be done cheaply.

The future work directions we find particularly promising are the following: First, the results on generic-payoff NFGs cover the normal-form representations of some, but not all, extensive-form games. Extending these results to EFGs thus constitutes a natural next step. Second, we saw that the cost of simulation that results in the socially-optimal outcome varies between games. It might therefore be beneficial to learn how to tailor the simulation cost to the specific game, and to what value. Third, we assumed that simulation predicts not only the simulated agent's policy, but also the result of any of their randomization — i.e., their precise action. Whether this assumption makes sense depends on the precise setting, but in any case, by considering mixtures over *behavioral* strategies [Halpern and Pass, 2021], it might be possible to go beyond this assumption while recovering most of our results. Finally, our work assumes that simulation is perfectly reliable, captures all parts of the other agent, and is only available to one agent but not the other. Ultimately, it will be necessary to go beyond these assumptions. We hope that progress in this direction can be made by developing a framework that encompasses both our work and some of the formalisms discussed in Section 7 (and in particular the work on program equilibria).

A Proofs

Proposition 2 (Equilibria for extreme simulation costs)**.** *In any simulation game $\mathcal{G}_{\text{sim}}$, we have:*

(i) For $\mathsf{c} < 0$, simulating is a strongly dominant action. In particular, $\text{NE}(\mathcal{G}^{\mathsf{c}}_{\text{sim}}) \subseteq \text{SE}^{\text{P2}}_{\text{pure}}(\mathcal{G})$.[3]

(ii) For $\mathsf{c} > \max_{a \in \mathcal{A}_1, b \in \mathcal{A}_2} u_1(a,b) - \max_{\pi_1 \in \Pi_1} \min_{\pi_2 \in \Pi_2} u_1(\pi_1, \pi_2)$, *$\mathtt{S}$ is a strictly dominated action. In particular,* $\text{NE}(\mathcal{G}^{\mathsf{c}}_{\text{sim}}) = \text{NE}(\mathcal{G}_{\text{sim}})$.

Proof. In (i), the dominance claim hold because $u_1(\mathtt{S}, b) = u_1(\text{br}, b) - \mathsf{c}$ for every $b \in \mathcal{A}_2$. As a result, when P1 simulates with probability 1, P2 gets utility $u_2(\mathtt{S}, b) = u_2(\psi_{\text{br}}(b), b)$ for some best-response policy ψ_{br}. As a result, P2 must select an action which maximises this value. And since P1 realises the utility $u_2(\mathtt{S}, b)$ by playing according to $\psi_{\text{br}}(b)$, we can identify this equilibrium with some pure Stackleberg equilibrium of $\mathcal{G}$ where P2 is the leader.

In (ii), $\mathtt{S}$ is strongly dominated by P1's min-max policy. In particular, $\mathtt{S}$ cannot be played in any NE. □

Lemma 4. $\forall \pi_2 : u_1(\mathtt{S}, \pi_2) = u_1(\text{br}, \pi_2) + \mathit{VoI}_{\mathtt{S}}(\pi_2) - \mathsf{c}$.

Proof. Let π be a policy in $\mathcal{G}$. Recall that $\text{VoI}_{\mathtt{S}}(\pi_2)$ is defined as the difference between $\sum_{b \in \mathcal{A}_2} \pi_2(b) u_1(\text{br}, b)$ and P1's best-response utility against π_2. In other words, we have

$$\begin{aligned} & u_1(\mathtt{S}, \pi_2) \\ & = \sum_{b \in \mathcal{A}_2} \pi_2(b) u_1(\text{br}, b) - \mathsf{c} \\ & = \left(\sum_{b \in \mathcal{A}_2} \pi_2(b) u_1(\text{br}, b) - u_1(\text{br}, \pi_2) \right) + u_1(\text{br}, \pi_2) - \mathsf{c} \\ & = \text{VoI}_{\mathtt{S}}(\pi_2) + u_1(\text{br}, \pi_2) - \mathsf{c} \\ & = u_1(\text{br}, \pi_2) + (\text{VoI}_{\mathtt{S}}(\pi_2) - \mathsf{c}). \end{aligned}$$

□

This result immediately yields the following:

Corollary 22. *Let $\mathcal{G}$ be a game in which $\bigcap_{b \in \mathcal{A}_2} \text{br}(b) = \emptyset$. Then all trembling-hand-perfect NE of $\mathcal{G}^0_{\text{sim}}$ satisfy $\pi_1(\mathtt{S}) = 1$. In particular, the set of trembling-hand-perfect NE of $\mathcal{G}^0_{\text{sim}}$ can be identified with the set of pure Stackelberg equilibria of $\mathcal{G}$.*

Lemma 5 ($\text{VoI}_{\mathtt{S}}$ is equal to simulation cost)**.** *(1) For any $\pi \in \text{NE}(\mathcal{G}^{\mathsf{c}}_{\text{sim}})$, we have $\pi_1(\mathtt{S}) \in (0,1) \implies \mathit{VoI}_{\mathtt{S}}(\pi_2) = \mathsf{c}$. (2) Moreover, unless $\mathcal{G}$ admits multiple optimal commitments of* P2 *that do not have a common best-response, any $\pi \in \text{NE}(\mathcal{G}^0_{\text{sim}})$ has $\mathit{VoI}_{\mathtt{S}}(\pi_2) = 0$.*

Proof. (1) This proposition straightforwardly follows from Lemma 4. Indeed, if $\text{VoI}_{\mathtt{S}}(\pi_2) < \mathsf{c}$, the equation implies that deviating to $\mathtt{S}$ would decrease P1's utility, and thus $\mathtt{S}$ cannot be in the support of π_1. If $\text{VoI}_{\mathtt{S}}(\pi_2) > \mathsf{c}$, simulation would give strictly higher utility (against π_2) than any action from the original game, so simulation would have to be the *only* action in the support of π_1. Consequently, the only case when the support of π_1 can include both $\mathtt{S}$ and some other action is when $\text{VoI}_{\mathtt{S}}(\pi_2) = \mathsf{c}$.

(2) Suppose that π is a NE of $\mathcal{G}^0_{\text{sim}}$ with $\text{VoI}_{\mathtt{S}}(\pi_2) > 0$. By Lemma 4, this means that simulating is a strongly dominant action for P1 and $\pi_1(\mathtt{S}) = 1$. Subsequently, any action from the support of π_2 must be an optimal commitment against ψ_{br}. However, the definition of $\text{VoI}_{\mathtt{S}}$ implies that there can be no single action of P1 which would give maximum utility against all actions b from the support of π_2. In other words, $\mathcal{G}$ must have optimal commitments of P2 that do not share a best response. This concludes the proof. □

Proposition 7 (Gradually recovering the NE of $\mathcal{G}$)**.** *Let π be a NE of $\mathcal{G}$. Then π, as a strategy in $\mathcal{G}^{\mathsf{c}}_{\text{sim}}$ with $\pi_1(\mathtt{S}) := 0$, is a NE precisely when $\mathsf{c} \geq \mathit{VoI}_{\mathtt{S}}(\pi_2)$.*

In particular, $\mathit{VoI}_{\mathtt{S}}(\pi_2)$ is a breakpoint of $\mathcal{G}$.

Proof. Let π be a NE of $\mathcal{G}_{\text{sim}}$. Proposition 5 implies that when $\mathsf{c} < \text{VoI}_{\mathtt{S}}(\pi_2)$, π cannot be a NE of $\mathcal{G}^{\mathsf{c}}_{\text{sim}}$ (since $\mathtt{S}$ is not in the support of π_1). Conversely, when $\mathsf{c} \geq \text{VoI}_{\mathtt{S}}(\pi_2)$, Lemma 4 implies that P1 isn't incentivised to unilaterally switch to $\mathtt{S}$. Moreover, since π is a NE of $\mathcal{G}$, no player is incentivised to switch to any other actions. As a result, π is a NE of $\mathcal{G}^{\mathsf{c}}_{\text{sim}}$ for any $\mathsf{c} \geq 0$. □

Lemma 9. *For any limit equilibrium π^0 of $\mathcal{G}_{\text{sim}}$, there is some $e > 0$ and π^e_2 such that for every $\mathsf{c} \in [0, e]$, $(\pi^0_1, (1 - \frac{\mathsf{c}}{e})\pi^0_2 + \frac{\mathsf{c}}{e}\pi^e_2)$ is a NE of $\mathcal{G}^{\mathsf{c}}_{\text{sim}}$.*

Proof. Let e_1 be the first breakpoint of $\mathcal{G}_{\text{sim}}$ that is higher than 0. Let π^0 be a limit equilibrium of $\mathcal{G}_{\text{sim}}$ and let π^n_2 be a sequence of strategies for which $\pi^0_2 = \lim_n \pi^n_2$, $(\pi^0_1, \pi^n_2) \in \text{NE}(\mathcal{G}^{\mathsf{c}_n}_{\text{sim}})$, $\mathsf{c}_n \to 0_+$. By Proposition 6, each π^n_2 lies on some line segment $t_n : \mathsf{c} \in [0, e_1] \mapsto \pi^n_2 + \delta^n(\mathsf{c} - \mathsf{c}_n)$, where $\delta^n \in \mathbb{R}^{\mathcal{A}_2}$ is the direction the line goes in and $(\pi^0_1, t_n(\mathsf{c})) \in \text{NE}(\mathcal{G}^{\mathsf{c}}_{\text{sim}})$ for each $\mathsf{c} \in [0, e_1]$. The set $\{\delta_n \mid n \in \mathbb{N}\}$ is necessarily bounded in $\mathbb{R}^{\mathcal{A}_2}$ (otherwise $t_n(e)$ would be unbounded in $\mathbb{R}^{\mathcal{A}_2}$ — i.e., it wouldn't lie Π_2). Using a compactness argument, we can assume that δ_n converges to some $\delta_0 \in \mathbb{R}^{\mathcal{A}_2}$. Denote by t_0 the line segment $t_0 : \mathsf{c} \in [0, e_1] \mapsto \pi^0_2 + \delta_0(\mathsf{c} - 0)$. Since the set $\{(\mathsf{c}, \pi) \mid \mathsf{c} \in [0, e_1], \pi \in \text{NE}(\mathcal{G}^{\mathsf{c}}_{\text{sim}})\}$ is closed, t_0 satisfies $(\pi^0_1, t_0(\mathsf{c})) \in \text{NE}(\mathcal{G}^{\mathsf{c}}_{\text{sim}})$ for every $\mathsf{c} \in [0, e_1]$. Denoting $e := e_1$ and $\pi^e_2 := t_0(e)$ concludes the proof. □

Lemma 10 (Structure of cheap-simulation equilibria)**.** *Let $\mathsf{c}_0 \in (0, e_1)$ and suppose that $\mathcal{G}$ admits no best-response utility tiebreaking by* P1*. Then any $\pi \in \text{NE}(\mathcal{G}^{\mathsf{c}_0}_{\text{sim}})$ with $\pi_1(\mathtt{S}) \in (0,1)$ is of the form $\pi = (\pi_1, \pi^{\mathsf{c}_0}_2)$, where*

$$\begin{aligned} \pi_1 &= (1 - \pi_1(\mathtt{S})) \cdot \pi^B_1 + \pi_1(\mathtt{S}) \cdot \mathtt{S} \\ \pi^{\mathsf{c}}_2 &= (1 - \alpha\mathsf{c}) \cdot \pi^B_2 + \alpha\mathsf{c} \cdot \pi^D_2, \quad \alpha > 0, \end{aligned}$$

and the following holds:

(i) For every $\mathsf{c} \in [0, e_1]$, $(\pi_1, \pi^{\mathsf{c}}_2) \in \text{NE}(\mathcal{G}^{\mathsf{c}}_{\text{sim}})$.

*(ii) $\pi^B \in \Pi$ is some **baseline policy** that satisfies:*

(B1) *every action in the support of* π_1^B *is a best-response to every action from* $\mathrm{supp}(\pi_2^B)$*;*

(B2) *every action in the support of* π_2^B *is an optimal commitment by* P2 *conditional on* P2 *only using strategies that satisfy (B1).*

(iii) $\pi_2^D \in \Pi_2$ *is some* ***deviation policy*** *that satisfies:*

(D1) *No* $a \in \mathrm{supp}(\pi_1^B)$ *lies in* $\mathrm{br}(d)$ *for all* $d \in \mathrm{supp}(\pi_2^D)$*.*

(D2) *Every* $d \in \mathrm{supp}(\pi_2^D)$ *satisfies one of*

$$u_2(\pi_1^B, \pi_2^D) > u_2(\pi^B) > u_2(\mathrm{br}, \pi_2^D) \qquad (D_2^>)$$

$$u_2(\pi_1^B, \pi_2^D) = u_2(\pi^B) = u_2(\mathrm{br}, \pi_2^D) \qquad (D_2^=)$$

$$u_2(\pi_1^B, \pi_2^D) < u_2(\pi^B) < u_2(\mathrm{br}, \pi_2^D). \qquad (D_2^<)$$

(D3) *If* $d \in \mathrm{supp}(\pi_2^D)$ *satisfies* $(D_2^>)$*, resp.* $(D_2^<)$*, it maximizes the* attractiveness ratio r_d*, resp.* r_d^{-1}

$$\frac{u_2(\pi_1^B, d') - u_2(\pi^B)}{u_2(\pi^B) - u_2(\mathrm{br}, d')} \text{ resp. } \frac{u_2(\mathrm{br}, d') - u_2(\pi^B)}{u_2(\pi^B) - u_2(\pi_1^B, d')}$$

among all $d' \in \mathcal{A}_2$ *that satisfy* $(D_2^>)$*, resp.* $(D_2^<)$*.*

Proof. Let $\mathcal{G}$, $\mathtt{c}_0$, and π be as in the assumptions of the lemma. To prove the statement, we first identify α and the policies π^{B} and π_2^{D}, and then we show that they have the desired properties.

Finding π_1^{B} is trivial — we simply $\pi_1 =: (1-\pi_1(\mathtt{S}))\cdot\pi_1^{\mathrm{B}} + \pi_1(\mathtt{S})\cdot\mathtt{S}$ and observe that π_1^{B} must be a valid policy for P1. To find α, π_2^{B}, and π_2^{D}, we can use Proposition 6. Indeed, this proposition implies that there is some linear function $\mathtt{c} \in [0, e_1] \mapsto \pi_2^{\mathtt{c}} \in \Pi_2$ for which $(\pi_1, \pi_2^{\mathtt{c}}) \in \mathrm{NE}(\mathcal{G}_{\mathrm{sim}}^{\mathtt{c}})$. (Once we define the desired properties, this proves the condition (i).) To define the baseline policy of P2, we simply set $\pi_2^{\mathrm{B}} := \pi_2^0$. To define the deviation policy, we first project $\pi_2^{e_1}$ onto π_2^{B} by setting $\beta := \max\{\beta' \mid \pi_2^{e_1} - \beta'\pi_2^{\beta} \geq 0\}$ (where the inequality $\geq$ holds pointwise). We then set $\tilde{\pi}_2^{\mathrm{D}} := \pi_2^{e_1} - \beta\pi_2^{\mathrm{B}}$ and define $\pi_2^{\mathrm{D}} := \tilde{\pi}_2^{\mathrm{D}} / \|\tilde{\pi}_2^{\mathrm{D}}\|$. Finally, by setting $\alpha := \beta / e_1$, we get the desired "slope" for which $(1-\alpha\mathtt{c})\cdot\pi_2^{\mathrm{B}} + \alpha\mathtt{c}\cdot\pi_2^{\mathrm{D}} = \pi_2^{\mathtt{c}}$. (This holds because the functions on both sides of this equation are linear and they coincide at $\mathtt{c} = 0$ and $\mathtt{c} = e_1$.)

As a side-product of the previous paragraph, we already have (i). To prove the lemma, it remains to prove that π^{B} satisfies (B1-2) and π_2^{D} satisfies (D1-3).

(ii) (B1): By Lemma 5, $\mathrm{VoI}_{\mathtt{S}}(\pi_2^{\mathrm{B}})$ must be 0. This implies that any action in the support of π_1 must be a best-response to any action from the support of π_2^{B} — i.e., we have (B1).

(B2): Since $\mathcal{G}$ admits no best-response tie-breaking by P1, (B1) implies that playing any b with $\mathrm{supp}(\pi_1^{\mathrm{B}}) \subseteq \mathrm{br}(b)$ is guaranteed to yield the same utility

$$\begin{aligned} u_2(\pi_1, b) &= (1 - \pi_1(\mathtt{S}))u_2(\pi_1^{\mathrm{B}}, b) + \pi_1(\mathtt{S})u_2(\mathtt{S}, b) \\ &= (1 - \pi_1(\mathtt{S}))u_2(\mathrm{br}, b) + \pi_1(\mathtt{S})u_2(\mathrm{br}, b) \\ &= u_2(\mathrm{br}, b). \end{aligned}$$

As a result, any b from the support of π^{B} must satisfy

$$u_2(\mathrm{br}, b) = \max\{u_2(\mathrm{br}, b') \mid \mathrm{br}(b') \supseteq \mathrm{supp}(\pi_1^{\mathrm{B}})\}.$$

Indeed, if it did not, P2 could increase their utility by switching to an action that does satisfy this equality, thus contradicting the fact that π^{B} is a NE of $\mathcal{G}_{\mathrm{sim}}^0$. This shows that π^{B} satisfies (B2).

(iii): To prove this part, consider $\pi^{\mathtt{c}}$ for some $\mathtt{c} \in (0, e_1)$.

(D1): Suppose there was a single action of P1 that was a best response to both all actions from $\mathrm{supp}(\pi_2^{\mathrm{B}})$ and all actions from $\mathrm{supp}(\pi_2^{\mathrm{D}})$. Then P1 could gain utility by unilaterally switching to that action (since $\pi_1(\mathtt{S}) > 0$ and $\mathtt{c} > 0$). This in particular implies that no action from $\mathrm{supp}(\pi_1^{\mathrm{B}})$ can have this property.

(D2): First, suppose some $d \in a_2$ satisfied both $u_2(\pi_1^{\mathrm{B}}, d) > u_2(\pi^{\mathrm{B}})$ and $u_2(\mathrm{br}, d) > u_2(\pi^{\mathrm{B}})$. Then P2 could gain utility by unilaterally deviating to d, contradicting the fact that $\pi^{\mathtt{c}}$ is a Nash equilibrium. The same would be true if some d satisfied these formulas with $=$ and $>$ or with $>$ and $=$. Conversely, any action of P2 that satisfies these formulas with $<$ and $<$, $<$ and $=$, or $=$ and $<$ is dominated (against π_1) by playing π_2^{B} and therefore cannot be played in equilibrium.

(D3): Denote by $\mathcal{A}_2^{\mathrm{LS}}$, resp. $\mathcal{A}_2^{\mathrm{LB}}$ the sets of actions that satisfy $(D_2^<)$, resp. $(D_2^>)$; these are the actions for which P2 would **L**ike P1 to **S**imulate, resp. would **L**ike them to play their **B**aseline strategy.

First, we will consider some $d \in \mathcal{A}_2^{\mathrm{LB}}$ and determine the simulation probability p that would make P1 indifferent between π^{B} and d. To determine p, we first observe that $u_2(\pi_1, \pi_2^{\mathrm{B}}) = u_2(\pi_1^{\mathrm{B}}, \pi_2^{\mathrm{B}}) = u_2(\pi^{\mathrm{B}})$ and $u_2(\pi_1, d) = (1-p)u_2(\pi_1^{\mathrm{B}}, d) + pu_2(\mathrm{br}, d)$. This yields

$$\begin{aligned} & u_2(\pi_1, \pi_2^{\mathrm{B}}) = u_2(\pi_1, d) \\ \iff\ & u_2(\pi^{\mathrm{B}}) = (1-p)u_2(\pi_1^{\mathrm{B}}, d) + pu_2(\mathrm{br}, d) \\ \iff\ & u_2(\pi^{\mathrm{B}}) = u_2(\pi_1^{\mathrm{B}}, d) - p\left(u_2(\pi_1^{\mathrm{B}}, d) - u_2(\mathrm{br}, d)\right) \\ \iff\ & p = \frac{u_2(\pi_1^{\mathrm{B}}, d) - u_2(\pi^{\mathrm{B}})}{u_2(\pi_1^{\mathrm{B}}, d) - u_2(\mathrm{br}, d)} \\ \iff\ & p = \frac{u_2(\pi_1^{\mathrm{B}}, d) - u_2(\pi^{\mathrm{B}})}{\left(u_2(\pi_1^{\mathrm{B}}, d) - u_2(\pi^{\mathrm{B}})\right) + \left(u_2(\pi^{\mathrm{B}}) - u_2(\mathrm{br}, d)\right)}. \end{aligned}$$

Denote the right-hand side of the last line as

$$\bar{p}_d := \frac{u_2(\pi_1^{\mathrm{B}}, d) - u_2(\pi^{\mathrm{B}})}{\left(u_2(\pi_1^{\mathrm{B}}, d) - u_2(\pi^{\mathrm{B}})\right) + \left(u_2(\pi^{\mathrm{B}}) - u_2(\mathrm{br}, d)\right)}. \qquad \text{(A.1)}$$

Clearly, $\bar{p}_d$ is a strictly increasing function of the deviation attractiveness ratio

$$r_d = \frac{u_2(\pi_1^{\mathrm{B}}, d) - u_2(\pi^{\mathrm{B}})}{u_2(\pi^{\mathrm{B}}) - u_2(\mathrm{br}, d}.$$

(Intuitively, r_d captures the tradeoffs P2 faces when deviating and hoping they will not be caught by the simulator.) Since d is of the type that causes P2 to prefer π_1^{B} over simulation, this implies that P2 would deviate to d for $\pi_1(\mathtt{S}) < \bar{p}_d$, be indifferent for $\pi_1(\mathtt{S}) = \bar{p}_d$, and switch to D for $\pi_1(\mathtt{S}) > \bar{p}_d$. Finally, denote $\bar{p}_* := \max\{\bar{p}_d \mid d \in \mathcal{A}_2^{\mathrm{LB}}\}$.

Considering the same equation for $d \in \mathcal{A}_2^{\mathrm{LS}}$, we get

$$u_2(\pi_1, \pi_2^{\mathrm{B}}) = u_2(\pi_1, d) \iff \cdots \iff p = \underline{p}^d, \text{ where}$$

$$\underline{p}^d := \frac{u_2(\pi^{\mathrm{B}}) - u_2(\pi_1^{\mathrm{B}}, d)}{\left(u_2(\pi^{\mathrm{B}}) - u_2(\pi_1^{\mathrm{B}}, d)\right) + \left(u_2(\mathrm{br}, d) - u_2(\pi^{\mathrm{B}})\right)}. \tag{A.2}$$

Clearly $\underline{p}^d$ is a strictly decreasing function of inverse ratio

$$r_d^{-1} = \frac{u_2(\mathrm{br}, d - u_2(\pi^{\mathrm{B}}))}{u_2(\pi^{\mathrm{B}}) - u_2(\pi_1^{\mathrm{B}}, d)}.$$

(In contrast to r_d, this ratio captures the tradeoffs P2 faces when deviating and hoping they *will* be caught by the simulator.) Finally, denote $\underline{p}^* := \min\{\underline{p}^d \mid d \in \mathcal{A}_2^{\mathrm{LS}}\}$.

Using these calculations, we are not only able to conclude the proof, but we have in fact also determined the values of $\pi_1(\mathtt{S})$ that are compatible with π^{B}: If $\pi_1(\mathtt{S})$ was strictly lower than $\bar{p}_*$, P2 would deviate towards some $d \in \mathcal{A}_2^{\mathrm{LB}}$ for which $\bar{p}_d > \pi_1(\mathtt{S})$. If it was strictly higher than $\underline{p}^*$, P2 would deviate towards some $d \in \mathcal{A}_2^{\mathrm{LS}}$ for which $\bar{p}_d < \pi_1(\mathtt{S})$. (In particular, we must have $\bar{p}_* \leq \underline{p}^*$ — otherwise, π^{B} could not be a limit equilibrium in $\mathcal{G}_{\mathrm{sim}}$.) This shows that if there is some action that satisfies $(D_2^{=})$, $\pi_1(\mathtt{S})$ can take any value from $[\bar{p}_*, \underline{p}^*]$. For $\mathrm{supp}(\pi_2^{\mathrm{D}})$ to contain some action $d \in \mathcal{A}_2^{\mathrm{LB}}$, $\pi_1(\mathtt{S})$ must be equal to $\bar{p}_*$ and d must satisfy $\bar{p}_d = \bar{p}_*$. (Which gives the "$>$" part of (D3).) And analogously, for $\mathrm{supp}(\pi_2^{\mathrm{D}})$ to contain some action $d \in \mathcal{A}_2^{\mathrm{LS}}$, $\pi_1(\mathtt{S})$ must be equal to $\underline{p}^*$ and d must satisfy $\underline{p}^d = \underline{p}^*$. (Which gives the "$<$" part of (D3).) This concludes the whole proof. ☐

Theorem 1 (Equilibria with binary supports)**.** *Let $\mathcal{G}$ be a game with generic payoffs and $\mathtt{c} \in (0, e_1)$. Then all NE of $\mathcal{G}_{\mathrm{sim}}^{\mathtt{c}}$ are either pure or have supports of size two.*

Proof. First, we observe several implications of the assumption that $\mathcal{G}$ has generic payoffs.

(0) In a generic game, no two payoffs are the same.

(1) For every action b of P1, P1 only has a single best response. With a slight abuse of notation, we denote this action as $\mathrm{br}(b)$.

(1') The same applies for P2.

(2) By (1), there is a unique action of P2 for which

$$u_2(\mathrm{br}(b), b) = \max_{b' \in \mathcal{A}_2} u_2(\mathrm{br}(b'), b'). \tag{A.3}$$

(3) Any two distinct actions d_1, d_2 of P2 must also have distinct attractiveness ratios r_d, $r_{d'}$ from (D3) of Lemma 10. (Indeed, if the payoffs of $\mathcal{G}$ are i.i.d. samples from the uniform distribution over $[0, 1]$, the probability two of these ratios coinciding is 0.)

(3') From (3), it further follows that the variables $\bar{p}_d$ and $\underline{p}^d$, defined in equations (A.1) and (A.2), will also differ for different actions.

We now proceed with the proof by separately considering the cases $\pi_1(\mathtt{S}) = 1$, $\pi_1(\mathtt{S}) = 0$, and $\pi_1(\mathtt{S} \in (0, 1)$.

$\pi_1(\mathtt{S}) = 1$: If P1 simulated with probability 1, P2 could respond be playing actions that satisfy (A.3). By (2), there is only one such action; call it b. However, this would mean that P1 could gain additional $\mathtt{c}$ utility by switching from $\mathtt{S}$ to $\mathrm{br}(b)$, contradiction the assumption that π is an equilibrium. As a result, a generic game with cheap simulation will never have an equilibrium where P1 simulates with probability 1.

$\pi_1(\mathtt{S}) = 0$: Suppose that π is a NE of $\mathcal{G}_{\mathrm{sim}}^{\mathtt{c}_0}$ for some $\mathtt{c}_0 \in (0, e_1)$. We will show that π must be pure.

Let $\mathtt{c} \in [0, e_1] \mapsto \pi_2^{\mathtt{c}}$ be some linear function (given by Proposition 6) for which $(\pi_1, \pi_2^{\mathtt{c}}) \in \mathrm{NE}(\mathcal{G}_{\mathrm{sim}}^{\mathtt{c}})$ holds for every $\mathtt{c}$. Since the $\mathtt{S}$ is not in the support of π_1, $\mathrm{VoI}_{\mathtt{S}}(\pi_2^0)$ must be equal to 0 (otherwise P1 could gain by deviating to $\mathtt{S}$ for $\mathtt{c} = 0$, and (π_1, π_2^0) would not be a NE of $\mathcal{G}_{\mathrm{sim}}^0$). This means that there must exist some $a \in \mathcal{A}_1$ that is a best-response to every b from $\mathrm{supp}(\pi_2^0)$. However, recall that (1) implies that P1 only has a single best response for every action of P2. As a result, a is the *only* action in the support of π_1. By (1'), this means that for every $\mathtt{c} \in [0, e_1]$ — and for $\mathtt{c}_0$ in particular — $\pi^{\mathtt{c}}$ must also be pure.

$\pi_1(\mathtt{S}) \in (0, 1)$: Let π be a NE of $\mathcal{G}_{\mathrm{sim}}^{\mathtt{c}_0}$ for some $\mathtt{c}_0 \in (0, e_1)$. By Lemma 10, π can be expressed as a convex combination of some baseline policy π_1^{B} and simulation (for P1), resp. of π_2^{B} and some deviation policy π_2^{D} (for P2). Combining the condition (B1-2) from Lemma 10 with (1) and (2), we get that π^{B} must be pure.

Let d be some element of $\mathrm{supp}(\pi_2^{\mathrm{D}})$ and consider the three cases listed in (D2). If d satisfies $(D_2^{=})$, (0) implies that it must be equal to π_2^{B}, and thus not count against the size of $\mathrm{supp}(\pi_2)$. Moreover, to avoid contradicting (D1), $\mathrm{supp}(\pi_2)$ must also contain some other action that does not satisfy $(D_2^{=})$. If d satisfies $(D_2^{>})$ or $(D_2^{<})$, the probability $\pi_1(\mathtt{S})$ must be equal to $\bar{p}_d$, resp. $\underline{p}^d$. (We observed this in the last paragraph of the proof of Lemma 10.) By (3'), it is impossible for this to be true for two different actions $d' \neq d$ at the same time. Together with π^{B} being pure, this shows that $|\mathrm{supp}(\pi_2)| = 2$ and concludes the proof. ☐

Proposition 11 (Simulation games are no harder than general games)**.** *Solving $\mathcal{G}_{\mathrm{sim}}^{\mathtt{c}}$ is at most as difficult as solving a normal-form game where* P1 *has one more action than in $\mathcal{G}$.*

Proof. This trivially follows from the assumption that simulation games are modelled as the original normal-form game $\mathcal{G}$ with the added simulate action $\mathtt{S}$. ☐

Proposition 12 (Solving $\mathcal{G}_{\mathrm{sim}}$ for extreme $\mathtt{c}$)**.**

(i) For $\mathtt{c} \in (-\infty, 0)$, the time complexity of solving $\mathcal{G}_{\mathrm{sim}}^{\mathtt{c}}$ is $O(|\mathcal{A}|)$.

(ii) For $\mathtt{c} \in (e_k, \infty)$, the time-complexity of solving $\mathcal{G}_{\mathrm{sim}}^{\mathtt{c}}$ is the same as the time-complexity of solving $\mathcal{G}$.

Proof. (i): First, suppose that $\mathcal{G}$ is a game with no best-response tie-breaking (i.e., P1's choice of best response never affects P2's utility). By Proposition 2(iii), simulation strongly dominates all other actions when $\mathtt{c} < 0$. Consequently, all that is needed to solve $\mathcal{G}^{\mathtt{c}}_{\text{sim}}$ is for P2 to search through a_2 for the action b with the highest best-response value $u_2(\text{br}, b) = u_2(\mathtt{S}, b)$. As a result, the complexity of solving $\mathcal{G}^{\mathtt{c}}_{\text{sim}}$ is dominated by the complexity of determining the best-response utilities corresponding the simulate action (which is $O(|\mathcal{A}|)$).

If $\mathcal{G}$ allows best-response tie-breaking for P1, the complexity might be higher because P1 could have multiple ways of responding after simulation. However, for the purpose of the paper, we were assuming that this policy (for how to respond after simulation) is fixed. As a result, the argument from the previous paragraph applies to this case as well.

(ii): By Proposition 2(i), $\mathtt{S}$ will never be played (in a NE) for high enough $\mathtt{c}$. As a result, solving $\mathcal{G}_{\text{sim}}$ becomes equivalent to solving $\mathcal{G}$. ☐

Theorem 2 (Cheap-simulation equilibria in generic games)**.** *Let $\mathcal{G}$ be a NFG with generic payoffs and $\mathtt{c} \in (0, e_1)$. Then the time complexity of finding all equilibria of $\mathcal{G}^{\mathtt{c}}_{\text{sim}}$ is $O(|\mathcal{A}|)$.*

Proof. By Theorem 1, all NE of $\mathcal{G}^{\mathtt{c}}_{\text{sim}}$ are either pure or have $|\text{supp}(\pi_1)| = |\text{supp}(\pi_2)| = 2$. (This straightforwardly implies that we could find all NE of $\mathcal{G}^{\mathtt{c}}_{\text{sim}}$ in $O(|\mathcal{A}|^2)$ time, by trying all possible supports of size one and two. The purpose of the theorem is, therefore, to show that the task can even be done in linear time.)

First, note that since $\mathcal{G}$ has generic payoffs, $\mathcal{G}^{\mathtt{c}}_{\text{sim}}$ doesn't have any equilibria with $\pi_1(\mathtt{S}) = 1$ and all of its equilibria with $\pi_1(\mathtt{S}) = 0$ are pure. (This is not hard to see directly. For a detailed argument, see the proof of Theorem 1.) As a result, these two cases can be handled in $O(|\mathcal{A}|)$ time.[4]

Second, consider the case when $\pi \in \text{NE}(\mathcal{G}^{\mathtt{c}}_{\text{sim}})$ satisfies $\pi(\mathtt{S}) \in (0, 1)$. From Theorem 1, we know that P1 will be mixing between some a and $\mathtt{S}$ and P2 will be mixing between some b and d, where (a, b) is the baseline strategy satisfying (B1-2) from Lemma 10 and d is the deviation strategy satisfying (D1-3) from Lemma 10. If a triplet (a, b, d) satisfies (B1-2) and (D1-3), we will call it "suitable".

To find all NE of $\mathcal{G}^{\mathtt{c}}_{\text{sim}}$, we can use the following procedure: (1a) For each $a \in \mathcal{A}_1$, find the (unique) best response of P2. (1b) For each $b \in \mathcal{A}_2$, find the (unique) best response of P1. (2) Find all suitable triplets (a, b, d). (3) For every suitable triplet from (2), find the unique NE with $\text{supp}(\pi_1) = \{a, \mathtt{S}\}$ and $\text{supp}(\pi_2) = \{b, d\}$, or learn that no such NE exists. (The uniqueness follows from (D1) and (D2).) To prove that the steps (1-3) can be performed in $O(|\mathcal{A}|)$ time, we will use the following claims:

[4]Recall that to find all pure NE of an NFG in linear time, we can: First, find all best-responses of P1 to every action of P2. Then find all best-responses of P2 to every action of P1. And finally use these findings to identify all joint actions that form a mutual best response; these coincide with all pure NE.

(I) Each of the steps (1a) and (1b) can be performed in $O(|\mathcal{A}|)$. (II) There are at most $2 \cdot \min\{|\mathcal{A}_1|, |\mathcal{A}_2|\}$ suitable, and it is possible to find all of them in $O(|\mathcal{A}|)$ time. (III) Performing (3) for a single suitable triplet takes $O(|\mathcal{A}_1| + |\mathcal{A}_2|)$ time.

Clearly, the combination of (I), (II), and (II) yields the conclusion of the theorem. Moreover, (I) is elementary and (III) follows from the fact that performing (2) only requires solving the 2-by-2 game with actions $\{a, \mathtt{S}\} \times \{b, d\}$ and checking that none of the remaining actions is a profitable deviation. To prove the theorem, it remains to prove the claim (II). We do this in two steps: First, we show that there are at most $\min\{|\mathcal{A}_1|, |\mathcal{A}_2|\}$ pairs (a, b) that might be a part of some suitable triplet (a, b, d). Second, we show that for any pair (a, b), there are at most two actions for which the triplet (a, b, d) is suitable.

For the first step, note that for every b, the only pair (a, b) that might satisfy the condition (B1) is $(\text{br}(b), b)$. Therefore, there are at most $|\mathcal{A}_2|$ pairs (a, b) that might be a part of some suitable triplet (a, b, d). Moreover, for every a, there will only be a single b that satisfies the condition (B2) (i.e., the condition that b maximizes $u_2(a, b')$ among the actions b' for which $a \in \text{br}(b')$). Therefore, there are at most $|\mathcal{A}_1|$ pairs (a, b) that might be a part of some suitable triplet (a, b, d). Combining the two bounds shows that the number of pairs that might be a part of a suitable triplet is $\min\{|\mathcal{A}_1|, |\mathcal{A}_2|\}$.

For the second step, note that in a generic game, the only action that satisfies $(D_2^{=})$ for $\pi^{\text{B}} = (a, b)$ is b. Moreover, the genericity of $\mathcal{G}$ implies that either the set of actions satisfying $(D_2^{>})$ is empty, or there is exactly one action d that satisfies $(D_2^{>})$ and maximizes the attractiveness r_d ratio from (D3). Analogously, there will be at most one action d that satisfies $(D_2^{<})$ and maximizes the inverse attractiveness ratio r_d^{-1} from (D3). This shows that for any (a, b) from the previous step, there are at most two actions for which the triplet (a, b, d) is suitable. Since this proves completes the proof of (II), we have concluded the whole proof. ☐

Theorem 3 (Simulation in trust games helps)**.** *Let $\mathcal{G}$ be a generalized trust game with generic payoffs. Then for all sufficiently low $\mathtt{c}$, $\mathcal{G}^{\mathtt{c}}_{\text{sim}}$ admits a Nash equilibrium with $\pi_1(\mathtt{S}) > 0$ that is a strict Pareto improvement over any NE of $\mathcal{G}$.*

Proof. Let $\mathcal{G}$ be a generalised trust game with generic payoffs and suppose that $\mathtt{c} > 0$ is sufficiently low (to be specified later in the proof). Denote (a, b) the unique equilibrium of the pure-commitment game corresponding to $\mathcal{G}$.

We need to find a suitable deviation $d \in \mathcal{A}_2$ that makes it possible to construct a NE $(\pi_1, \pi^{\mathtt{c}})$ of $\mathcal{G}^{\mathtt{c}}_{\text{sim}}$ with $\text{supp}(\pi_1) = \{a, \mathtt{S}\}$ and $\text{supp}(\pi_2^{\mathtt{c}}) = \{b, d\}$. Since $\text{supp}(\pi_1)$ is going to be equal to $\{a, \mathtt{S}\}$, any such deviation will need to satisfy either (a) $u_2(a, d) > u_2(a, b)$, or (b) $u_2(\mathtt{S}, d) > u_2(\mathtt{S}, b)$, or (c) $u_2(a, d) = u_2(a, b)$ and $u_2(\mathtt{S}, d) = u_2(\mathtt{S}, b)$. However, because b is an optimal commitment, there is no action that satisfies (b) or (c) (since $u_2(a, b) =$

$u_2(\text{br}, b) = \max_{b' \in \mathcal{A}_2} u_2(\text{br}, b')$). As a result, we need to consider the deviations that satisfy (a):

$$D := \{d \in \mathcal{A}_2 \mid u_2(a, d) > u_2(a, b)\}.$$

Note that the set D has the following properties:

(1) *D is non-empty.*
(Indeed, since $\mathcal{G}$ is a generalized trust game, any NE of $\mathcal{G}$ must give strictly lower u_2 than any pure-commitment equilibrium. In particular, the pure-commitment equilibrium (a, b) cannot be a NE of $\mathcal{G}$, so one of the players must have a profitable deviation. Since P1's action a is already a best response to b, P2 must have some d that satisfies the definition of D.)

(2) *Any $d \in D$ has $u_2(\mathtt{S}, d) < u_2(\mathtt{S}, b)$.*
(Indeed, this holds because (a, b) is the only pure-commitment equilibrium of $\mathcal{G}$, so all elements must give strictly lower $u_2(\text{br}, d)$.)

(3) *Any $d \in D$ has $u_1(\mathtt{S}, d) > u_1(a, d)$.*
(Indeed, this holds because a cannot be a best-response to d — if it was, d would have been a better commitment for P2 than b, contradicting the definition of b.)

We are now ready to define the desired policy. We select d as some element of D that maximizes the deviation attractiveness ratio r_d

$$r_d = \frac{u_2(a, d) - u_2(a, b)}{u_2(\mathtt{S}, b) - u_2(\mathtt{S}, d)}. \tag{A.4}$$

(The property (2) above ensures that the denominator is never equal to 0. The property (1) ensures that some suitable d exists.) We select $\pi_2^{\mathtt{c}}(d) =: q_{\mathtt{c}}$ as the value for which P1 is indifferent between a and $\mathtt{S}$. Since $u_1(\mathtt{S}, d) > u_1(a, d)$ by (3) and $u_1(a, b) = u_1(\text{br}, b) > u_1(br, b) - \mathtt{c} = u_1(\mathtt{S}, b)$, this value exists and is uniquely determined by the following equations.

$$\begin{aligned}
&u_1(a, \pi_2^{\mathtt{c}}) = u_1(\mathtt{S}, \pi_2^{\mathtt{c}}) \\
&\iff (1 - q_{\mathtt{c}}) u_1(a, b) + q_{\mathtt{c}} u_1(a, d) \\
&\qquad = (1 - q_{\mathtt{c}})(u_1(a, b) - \mathtt{c}) + q_{\mathtt{c}}(u_1(\text{br}, d) - \mathtt{c}) \\
&\iff q_{\mathtt{c}} u_1(a, d) = q_{\mathtt{c}} u_1(\text{br}, d) - \mathtt{c} \\
&\iff \mathtt{c} = q_{\mathtt{c}} \left(u_1(\text{br}, d) - q_{\mathtt{c}} u_1(a, d)\right) \\
&\iff q_{\mathtt{c}} = \frac{\mathtt{c}}{u_1(\text{br}, d) - q_{\mathtt{c}} u_1(a, d)}.
\end{aligned}$$

We select $\pi_1(\mathtt{S}) =: p$ as the value for which P2 is indifferent between b and d. (This derivation is analogous to the derivation of $\pi_2^{\mathtt{c}}(d)$, using the facts that $u_2(a, d) > u_2(a, b)$ holds by definition of D and $u_2(\mathtt{S}, b) > u_2(\mathtt{S}, d)$ holds by (2). Since the specific value of $\pi_1(\mathtt{S})$ is not important for us, we do not repeat the full calculation here — for details, see the step (D3) of the proof of Lemma 10.)

To conclude the proof, we show that $(\pi_1, \pi_2^{\mathtt{c}})$ has the desired properties. To see that P2 has no profitable deviation, recall that d was selected as the action that maximizes the attractiveness ratio r_d. By Lemma 10 (or more precisely, by repeating the argument from the step (D3) of the proof of Lemma 10) this ensures that no other element of D will give more utility against π_1. And as we observed earlier, any element of $\mathcal{A}_2 \setminus (D \cup \{b\})$ will give strictly less utility against both a and $\mathtt{S}$ than b. To see that P1 has no profitable deviation, note that P2's deviation probability goes to 0 as $\mathtt{c} \to 0_+$. This means that once $\mathtt{c}$ gets sufficiently low, any profitable deviation $a' \in \mathcal{A}_1 \setminus \{a\}$ of P1 would need to be a best-response to b. However, since $\mathcal{G}$ has generic payoffs, a is the only action with this property. To see that the this policy is a strict Pareto improvement over any NE of $\mathcal{G}$, note that as $\mathtt{c}$ tends to 0, the simulation cost becomes negligible, and $\pi_2^{\mathtt{c}}$ converges to b. This means that both $u_i(a, \pi_2^{\mathtt{c}})$ and $u_i(\mathtt{S}, \pi_2^{\mathtt{c}})$ converge to $u_i(a, b)$. Since (a, b) is a pure-commitment equilibrium and $\mathcal{G}$ is a generalized trust game, this value is guaranteed to be a strict improvement over the utility under any NE of $\mathcal{G}$. This concludes the whole proof. ☐

B Proof of Proposition 6: Linear Adjustment of P1's Payoffs Have Constant/Linear Effects on NE

In the main text, we used the following result:

Proposition 6 (Simulation equilibria trajectories are piecewise constant/linear)**.** *For every $\mathcal{G}$, there is a finite set of simulation-cost breakpoint values $-\infty = e_{-1} < 0 = e_0 < e_1 < \cdots < e_k < e_{k+1} = \infty$ such that the following holds: For every $\mathtt{c}_0 \in (e_l, e_{l+1})$ and every $\pi^{\mathtt{c}_0} \in \text{NE}(\mathcal{G}_{\text{sim}}^{\mathtt{c}_0})$, there is a linear mapping $\mathfrak{t}_2 : \mathtt{c} \in [e_l, e_{l+1}] \mapsto \pi_2^{\mathtt{c}} \in \Pi_2$ such that $\mathfrak{t}_2(\mathtt{c}_0) = \pi_2^{\mathtt{c}_0}$ and $(\pi_1^{\mathtt{c}_0}, \mathfrak{t}_2(\mathtt{c})) \in \text{NE}(\mathcal{G}_{\text{sim}}^{\mathtt{c}})$ for every $\mathtt{c} \in [e_l, e_{l+1}]$.*

Proof. With the exception of the claim about e_0 being equal to 0, the result is an immediate corollary the more general Lemma 24 listed below. The claim about 0 being a breakpoint and about the non-existence of any breakpoints in $(-\infty, 0)$ immediately follows from Proposition 2 (i). ☐

In this section, we prove a more general version of Proposition 6 using the following setting:

Definition 23 (Auxiliary)**.** *A game with **linearly adjustable payoffs** is any pair $(\mathcal{G}, \vec{\alpha})$ where $\mathcal{G} = (\mathcal{A}, \tilde{u})$ is a two-player normal-form game and $\vec{\alpha} = (\alpha_a)_{a \in \mathcal{A}_1} \in \mathcal{A}_1$ is a vector of **adjustments** for* P1*'s actions. For a **cost-scaling** factor $\mathtt{c} \in \mathbb{R}$, $\mathcal{G}_{\vec{\alpha}}^{\mathtt{c}}$ denotes the NFG with actions $\mathcal{A}$ and utilities $u_2 := \tilde{u}_2$, $u_1(a, b) := \tilde{u}_1(a, b) - \mathtt{c}\alpha_a$.*

The connection between this notion and our setting is that any simulation game $\mathcal{G}_{\text{sim}}^{\mathtt{c}}$ can be expressed as $\mathcal{G}_{\text{sim}}^{\mathtt{c}} = (\mathcal{G}')_{(0,\ldots,0,1)}^{\mathtt{c}}$, where $\mathcal{G}'$ is the original game $\mathcal{G}$ with one additional P1 action $\mathtt{S}$ that yields utilities $\tilde{u}_i(\mathtt{S}, b) := u_i(a, b)$ (where we fix some $a \in \text{br}(b)$ for every $b \in \mathcal{A}_2$).

The piecewise constant/linear phenomenon that we observed on the motivating example of simulation in Trust Game (Figure 2) in fact holds more generally — for every game with linearly adjustable payoffs. The goal of this section is to build up to the proof of the following result, which immediately gives our desired result – Proposition 6 as a corollary:

Lemma 24 (Games with linearly adjustable payoffs have piecewise constant/linear NE trajectories)**.** *For every $\mathcal{G}$, there is a finite set of breakpoint values $-\infty = e_{-1} < e_0 < \cdots < e_k < e_{k+1} = \infty$ such that the following holds: For every $\mathsf{c}_0 \in (e_l, e_{l+1})$ and every $\pi^{\mathsf{c}_0} \in \mathrm{NE}(\mathcal{G}^{\mathsf{c}_0}_{\vec{\alpha}})$, there is a linear mapping $t_2 : \mathsf{c} \in [e_l, e_{l+1}] \mapsto \pi_2^{\mathsf{c}} \in \Pi_2$ such that $t_2(\mathsf{c}_0) = \pi_2^{\mathsf{c}_0}$ and $(\pi_1^{\mathsf{c}_0}, t_2(\mathsf{c})) \in \mathrm{NE}(\mathcal{G}^{\mathsf{c}}_{\vec{\alpha}})$ for every $\mathsf{c} \in [e_l, e_{l+1}]$.*

B.1 Background: Linear Programming

Before proceeding with the proof, we recall several results from linear programming. (Since these results are standard, they will be given without a proof. For a detailed exposition of using LPs for solving normal-form games, see for example [Shoham and Leyton-Brown, 2008].)

First, if we can guess the support of a Nash equilibrium, the strategy itself can be found using a linear program:

Observation 25 (Indifference sets of NE)**.** *Every NE π of $\mathcal{G}$ satisfies $\mathrm{supp}(\pi_i) \subseteq \mathrm{br}(\pi_{\text{-}i})$. As a result, we can write the set of NE in $\mathcal{G}$ as a (possibly overlapping) union*

$$\mathrm{NE}(\mathcal{G}) = \bigcup \{\mathrm{NE}(\mathcal{G}, S_1, S_2) \mid \forall i : S_i \subseteq \mathcal{A}_i\},$$
$$\textit{where } \mathrm{NE}(\mathcal{G}, S_1, S_2) :=$$
$$\{\pi \in \mathrm{NE}(\mathcal{G}) \mid \forall i : \mathrm{supp}(\pi_i) \subseteq S_i \subseteq \mathrm{br}(\pi_{\text{-}i})\}.$$

Lemma 26 (NE as solutions of LP)**.** *For any S_1, S_2, the elements of $\mathrm{NE}(\mathcal{G}, S_1, S_2)$ are precisely (the π-parts of) the solutions of the following linear program (with no maximisation objective).*

$$\forall i :$$
$$\sum\nolimits_{a_i \in S_i} \pi_i(a_i) = 1 \tag{B.1}$$
$$u_i(a_i, \pi_{\text{-}i}) = \gamma_i \qquad \textit{for } a_i \in S_i \tag{B.2}$$
$$u_i(a_i, \pi_{\text{-}i}) \geq \gamma_i \qquad \textit{for } a_i \in \mathcal{A}_i \setminus S_i \tag{B.3}$$
$$\textit{where the variables satisfy}$$
$$\pi_i(a_i) = 0 \qquad \textit{for } a_i \in \mathcal{A}_i \setminus S_i \tag{B.4}$$
$$\pi_i(a_i) \geq 0 \qquad \textit{for } a_i \in S_i \tag{B.5}$$
$$\gamma_i \in \mathbb{R} \tag{B.6}$$

Another standard result is that the geometry of the set $\mathrm{NE}(\mathcal{G}, S_1, S_2)$ can be derived from the LP above:

Lemma 27 (Geometry of NE)**.** *(1) For any $\pi, \pi' \in \mathrm{NE}(\mathcal{G}, S_1, S_2)$, we have $(\pi_1, \pi_2'), (\pi_1', \pi_2) \in \mathrm{NE}(\mathcal{G}, S_1, S_2)$. (2) $\mathrm{NE}(\mathcal{G}, S_1, S_2)$ is a convex polytope and its vertices are precisely the basic feasible solutions of the LP from Lemma 26.*

In light of Lemma 27, we can denote

$$\mathrm{NE}(\mathcal{G}, S_1, S_2) := \mathrm{NE}_1(\mathcal{G}, S_1, S_2) \times \mathrm{NE}_2(\mathcal{G}, S_1, S_2).$$

We also use $\mathrm{NE}_i^{\mathrm{ext}}(\mathcal{G}, S_1, S_2)$ to denote the **extremal NE strategies** — i.e., the vertices $\mathrm{NE}_i(\mathcal{G}, S_1, S_2)$.

B.2 Linearity of Simulation Equilibria

The first observation is that since the utilities of P2 do not depend on c, their Nash equilibrium strategy of P1 do not need to change either:

Lemma 28 (WLOG, P1's strategy is constant)**.** *Suppose that π, resp. π', is a solution of the LP from Lemma 26 for $\mathcal{G}^{\mathsf{c}}_{\vec{\alpha}}$ and (S_1, S_2) for c, resp. c'. Then (π_1, π_2') is a solution of the LP for $\mathcal{G}^{\mathsf{c}'}_{\vec{\alpha}}$.*

Proof. To prove the lemma, it suffices to verify that (π_1, π_2') is a feasible solution of the LP from Lemma 26. However, this is trivial once we realise that the utility of P2 does not depend on c. ☐

As a result, it only remains to prove the linearity of $\mathrm{NE}_2^{\mathrm{ext}}(\mathcal{G}^{\mathsf{c}}_{\vec{\alpha}}, S_1, S_2)$ (and then put all the results together).

Proposition 29. *For every $S_1 \subset \mathcal{A}_1$ and $S_2 \subset \mathcal{A}_2$, there is a finite number of breakpoints $-\infty = e_{-1} < \cdots < e_{k+1} = \infty$, such that on any of the intervals (e_i, e_{i+1}), the elements of $\mathrm{NE}_2^{\mathrm{ext}}(\mathcal{G}^{\mathsf{c}}_{\vec{\alpha}}, S_1, S_2)$ change linearly with c.*

Here, "elements of $\mathrm{NE}_2^{\mathrm{ext}}(\mathcal{G}^{\mathsf{c}}_{\vec{\alpha}}, S_1, S_2)$ changing linearly" means that (a) for a fixed i, there is some $N \geq 0$ such that for every (e_i, e_{i+1}), the set $\mathrm{NE}_2^{\mathrm{ext}}(\mathcal{G}^{\mathsf{c}}_{\vec{\alpha}}, S_1, S_2)$ has exactly N elements and (b) there are linear functions $t_2^n : (e_i, e_{i+1}) \to \Pi_2$, $n = 1, \ldots, N$, such that for every $\mathsf{c} \in (e_i, e_{i+1})$, $\mathrm{NE}_2^{\mathrm{ext}}(\mathcal{G}^{\mathsf{c}}_{\vec{\alpha}}, S_1, S_2) = \{t_2^n(\mathsf{c}) \mid n = 1, \ldots, N\}$.

Proof. Let $(\mathcal{G}, \vec{\alpha})$ be a game with linearly adjustable payoffs and suppose that c is such that there exists some NE in $(\mathcal{G}^{\mathsf{c}}_{\vec{\alpha}}, S_1, S_2)$.

As the first step, we rewrite the relevant part of the LP from earlier. Using Lemma 26 (in combination with (1) from Lemma 27), we see that a policy π_2 lies in $\mathrm{NE}_2^{\mathrm{ext}}(\mathcal{G}^{\mathsf{c}}_{\vec{\alpha}}, S_1, S_2)$ if and only if it is a basic feasible solution of the following LP:

$$\sum\nolimits_{b \in S_2} \pi_2(b) = 1 \tag{B.7}$$
$$u_1(a, \pi_2) - \mathsf{c}\alpha_a = \gamma \qquad \text{for } a \in S_1 \tag{B.8}$$
$$u_1(a', \pi_2) - \mathsf{c}\alpha_{a'} \geq \gamma \qquad \text{for } a' \in a_1 \setminus S_1 \tag{B.9}$$
$$\text{where the variables satisfy} \tag{B.10}$$
$$\pi_2(b) = 0 \qquad \text{for } b \in \mathcal{A}_2 \setminus S_2 \tag{B.11}$$
$$\pi_2(b) \geq 0 \qquad \text{for } a \in S_2 \tag{B.12}$$
$$\gamma \in \mathbb{R}. \tag{B.13}$$

We turn all of the inequalities into equalities by introducing slack variables $w_{a'} \geq 0$, $a' \in \mathcal{A}_1 \setminus S_1$:

$$\sum\nolimits_{b \in S_2} \pi_2(b) = 1 \tag{B.14}$$
$$u_1(a, \pi_2) - \mathsf{c}\alpha_a = \gamma \qquad \text{for } a \in S_1 \tag{B.15}$$
$$u_1(a', \pi_2) - \mathsf{c}\alpha_a + w_{a'} = \gamma \qquad \text{for } a' \in \mathcal{A}_1 \setminus S_1 \tag{B.16}$$
$$\text{where the variables satisfy} \tag{B.17}$$
$$\pi_2(b) = 0 \qquad \text{for } b \in \mathcal{A}_2 \setminus S_2 \tag{B.18}$$
$$\pi_2(b) \geq 0 \qquad \text{for } a \in S_2 \tag{B.19}$$
$$\gamma \in \mathbb{R}. \tag{B.20}$$

$$
\left[\begin{array}{ccccccccc|c}
1 & 1 & \cdots & 1 & & \dots & & 0 & 0 & 1 \\
 & & & & & & & & & \\
u_1(a_1,b_1) & u_1(a_1,b_2) & \cdots & u_1(a_1,b_m) & & \dots & & 0 & \text{-1} & \alpha_{a_1} \\
u_1(a_2,b_1) & u_1(a_2,b_2) & \cdots & u_1(a_2,b_m) & & \dots & & 0 & \text{-1} & \alpha_{a_2} \\
\vdots & \vdots & \vdots & \vdots & & & & \vdots & \vdots & \vdots \\
u_1(a_n,b_1) & u_1(a_n,b_2) & \cdots & u_1(a_n,b_m) & & \dots & & 0 & \text{-1} & \alpha_{a_n} \\
u_1(a'_1,b_1) & u_1(a'_1,b_2) & \cdots & u_1(a'_1,b_m) & 1 & 0 & \dots & 0 & \text{-1} & \alpha_{a'_1} \\
u_1(a'_2,b_1) & u_1(a'_2,b_2) & \cdots & u_1(a'_2,b_m) & 0 & 1 & & 0 & \text{-1} & \alpha_{a'_2} \\
\vdots & \vdots & \ddots & \vdots & \vdots & & \ddots & \vdots & \vdots & \vdots \\
u_1(a'_{n'},b_1) & u_1(a'_{n'},b_2) & \cdots & u_1(a'_{n'},b_m) & 0 & 0 & \dots & 1 & \text{-1} & \alpha_{a'_{n'}}
\end{array}\right],
$$

Figure 4: The matrix form $Ax_{\mathsf{c}}^{\mathbf{T}} = y^{\mathsf{c}}$ of the LP (B.14)-(B.20), where the columns are indexed by $x_{\mathsf{c}} = (\pi_2^{\mathsf{c}}(b_1), \dots, \pi_2^{\mathsf{c}}(b_1), w_{a'_1}, \dots, w_{a'_{n'}}, \gamma)$. (The numbers m, n, and n' stand for the size of S_2, S_1, and $\mathcal{A}_1 \setminus S_1$ respectively.) The additional constraints are $\pi_2^{\mathsf{c}}(b_j) \geq 0$ and $w'_a \geq 0$. Note that because P1's utilities are adjusted independently of P2's actions, the adjustments α_a and $\alpha_{a'}$ can be moved to right-hand side of the equation.

Second, we rewrite the LP (B.14)-(B.20) in a matrix form. We denote the relevant variables as $x_{\mathsf{c}} = (\pi_2^{\mathsf{c}}(b_1), \dots, \pi_2^{\mathsf{c}}(b_n), w_{a'_1}, \dots, w_{a'_{n'}}, \gamma)$, where $n := |\mathcal{A}_1|$, $n' := |\mathcal{A}_1 \setminus S_1|$. In this notation, there will be some matrix A and right-hand side $y^{\mathsf{c}} = (1, 0, \dots, 0)$ for which some π_2 is a solution of (B.14)-(B.20) if and only if it satisfies $Ax_{\mathsf{c}}^{\mathbf{T}} = y^{\mathsf{c}}$ and $\pi_2^{\mathsf{c}}(b_j), w_{a'} \geq 0$. However, we can additionally use the fact that the adjustment cost $\mathsf{c}\alpha_a$ that P1 pays does not depend on the action of P2. This allows us the matrix form depicted in Figure 4, where all the $\mathsf{c}\alpha_a$-s have been moved to the right-hand side.

As the third step, we note that while the matrix from Figure 4 might have linearly dependent rows, we can always replace it by a matrix whose rows are linearly independent. To see this, note first that clearly no row corresponding to one of the actions $a' \in \mathcal{A}_1 \setminus S_1$ can be expressed as a linear combination of any other rows, because of the 1-s in the bottom-right corner of the matrix. Second, it *is* possible that one of the rows corresponding to some $a \in S_1$ can be expressed as a linear combination of the other rows corresponding to S_1. Suppose that a is such action and $\lambda_i \in \mathbb{R}$ are the corresponding weights. There are two options: If $\sum_i \lambda_i \alpha_{a_i} = \alpha_a$, the condition corresponding to a can be omitted, since it is already subsumed by the conditions corresponding to $S_1 \setminus \{a\}$. Conversely, if $\sum_i \lambda_i \alpha_{a_i} \neq \alpha_a$, the system of equations from Figure 4 will be unsolvable. However, this case is ruled out by our assumption that c is such that $\mathrm{NE}(\mathcal{G}, S_1, S_2)$ is non-empty. In summary: For the remainder of the proof, we can assume that the rows of the matrix A are linearly independent.

Fourth, we identify the basic solution of the LP given by Figure 4. For the purpose of this step, denote the set of column-indices of A as $\mathcal{I} := S_2 \cup (\mathcal{A}_1 \setminus S_1) \cup \{\gamma\}$. For a "basis" $B \subseteq \mathcal{I}$, we use A_B to denote the sub-matrix of A consisting of the columns indexed by B. By $\mathcal{B}$, we denote the set of all B-s for which the sub-matrix A_B is regular. Finally, for $B \in \mathcal{B}$, we denote by x_{c}^B the basic solution corresponding to B — i.e., the solution of $Ax_{\mathsf{c}}^{\mathbf{T}} = y^{\mathsf{c}}$ for which all the variables indexed by $\mathcal{I} \setminus B$ are equal to 0. By definition of a BFS, the basic feasible solutions of the LP are precisely all the vectors of the form x_{c}^B, $B \in \mathcal{B}$.

Fifth, we show that every basic (not necessarily feasible solution of the LP given by Figure 4 changes linearly with c. To see this, note that each basic (not necessarily feasible) solution x_{c}^B can be written as the vector $A_B^{-1} y^{\mathsf{c}}$, extended by 0-s at the indices $\mathcal{I} \setminus B$. (Since A_B is assumed to be regular, the inverse exists.) Since the matrix A does not depend on c and y^{c} only depends on c linearly, the mapping $\mathsf{c} \mapsto x_{\mathsf{c}}^B$ is linear.

Finally, we conclude the proof. To do this, recall that a basic solution x_{c}^B is feasible if all of the variables $\pi_2^{\mathsf{c}}(b), w_a$ are non-negative. For every x_{c}^B, the set of the values of c for which all of these definitions are satisfied is going to be some (possibly empty or trivial) closed interval $[e_0^B, e_1^B]$. By taking the set $\{e_i^B \mid i = 0, 1, B \in \mathcal{B}\} \cup \{-\infty, \infty\}$ and reordering it as an increasing sequence, we obtain the desired breakpoint set E. This completes the whole proof □

Proof of Lemma 24. Let $(\mathcal{G}, \vec{\alpha})$ be a game with linearly adjustable payoffs. To get the desired sequence of breakpoints, we take – for every pair $S_1 \subseteq \mathcal{A}_1$ and $S_2 \subseteq \mathcal{A}_2$ – some set $E(S_1, S_2)$ of breakpoints given by Proposition 29 and define $E := \bigcup_{S_1 \subseteq \mathcal{A}_1, S_2 \subseteq \mathcal{A}_2} E(S_1, S_2)$. We then enumerate E as a strictly increasing sequence $(e_i)_i$ To prove our result, let $\mathsf{c} \in [e_i, e_{i+1}]$, and $\pi^{\mathsf{c}_0} \in \mathrm{NE}(\mathcal{G}_{\vec{\alpha}}^{\mathsf{c}_0})$. We finding a trajectory t_2 and that satisfies the conclusion of the lemma.

By Observation 25, there are some sets S_1, S_2 for which $\pi^{\mathsf{c}_0} \in \mathrm{NE}(\mathcal{G}_{\vec{\alpha}}^{\mathsf{c}_0}, S_1, S_2)$. By Lemma 27 and the subsequent observation, there are some basic feasible solutions $\beta^1, \dots, \beta^N$ (of the LP from Lemma 26) and convex combination $\lambda_1, \dots, \lambda_N$ such that $\pi_2^{\mathsf{c}_0} = \sum_{i=1}^N \lambda_i \beta_2^i$. By Proposition 29, there are some linear trajectories $t_2^1, \dots, t_2^N$ such that $t_2^i(\mathsf{c}_0) = \beta_2^i$ and for every $\mathsf{c} \in (e_i, e_{i+1})$,

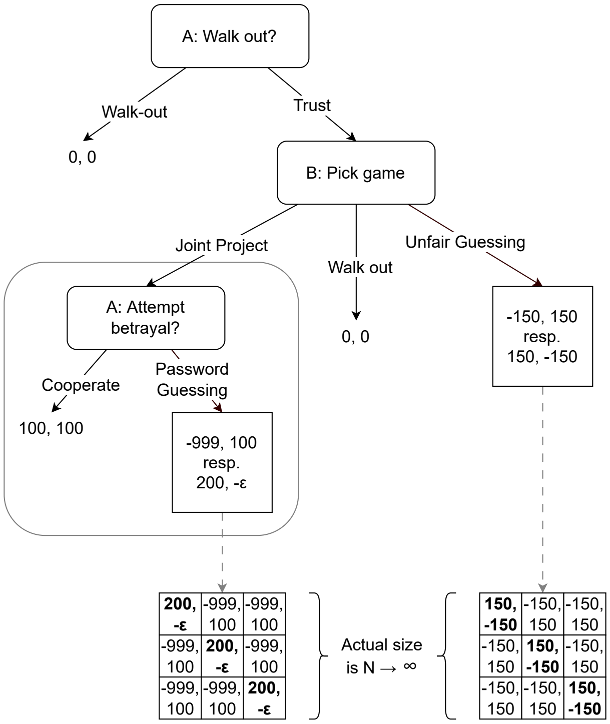


Figure 5: A game where both players prefer simulation to be neither cheap nor prohibitively costly.

$t_2^i(\mathsf{c}) \in \mathrm{NE}_2^{\mathrm{ext}}(\mathcal{G}_{\vec{\alpha}}^{\mathsf{c}}, S_1, S_2)$. Moreover, by Lemma 28, we have $(\pi_1^{\mathsf{c}_0}, t_2^i(\mathsf{c})) \in \mathrm{NE}(\mathcal{G}_{\vec{\alpha}}^{\mathsf{c}}, S_1, S_2)$ for every $\mathsf{c} \in (e_i, e_{i+1})$. By the convexity of NE (Lemma 27), $t_2 := \sum_{i=1}^{N} t_2^i$ is a linear trajectory for which $t_2(\mathsf{c}_0) = \pi^{\mathsf{c}_0}$ and for every $\mathsf{c} \in (e_i, e_{i+1})$, $(\pi^{\mathsf{c}_0}, t_2^i(\mathsf{c})) \in \mathrm{NE}(\mathcal{G}_{\vec{\alpha}}^{\mathsf{c}}, S_1, S_2)$. Moreover, since the utilities in $\mathcal{G}_{\vec{\alpha}}^{\mathsf{c}}$ depend continuously on c, this also implies that $(\pi^{\mathsf{c}_0}, t_2^i(\mathsf{c})) \in \mathrm{NE}(\mathcal{G}_{\vec{\alpha}}^{\mathsf{c}}, S_1, S_2)$ for $\mathsf{c} \in \{e_i, e_{i+1}\}$. (By Observation 25,) this concludes the whole proof. □

C Example: Optimal Simulation Cost is Non-trivial

Example 30 (Optimal simulation cost is non-trivial)**.** Consider the game $\mathcal{G}$ depicted in Figure 5. First, both Alice and Bob have an option to walk out and not play ($u_{\mathrm{A}} = u_{\mathrm{B}} = 0$). If Alice chooses to trust Bob and play, Bob gets to decide which game to play. One option is the Unfair Guessing game (Example 17), where Alice needs to guess an integer that Bob is thinking, else she ends up transferring 150 utility to Bob. If she guesses correctly, Bob transfers 150 to her instead. (This game is parametrized by N, the highest integer that Bob is allowed to pick. Since the game is biased in Bob's favor, the corresponding expected utilities converge to $u_{\mathrm{A}} = -150, u_{\mathrm{B}} = 150$ as $N \to \infty$. Since the precise numbers are not important for our conclusions, we will, for the purpose of this example, treat them as exactly equal to ± 150.) The other option available to Bob is to play the Joint Project game (Example 19). In this game, Alice can either Cooperate with Bob ($u_{\mathrm{A}} = u_{\mathrm{B}} = 100$) or attempt to betray him by guessing his password and stealing all his profits. A successful betrayal results in utilities $u_{\mathrm{A}} = 200$, u_{B}, while an unsuccessful one sends Alice to prison ($u_{\mathrm{A}} = -999$, $u_{\mathrm{B}} = 100$). (This game is also parametrized by N, such that when Bob picks his password uniformly at random, the outcomes converge to $u_{\mathrm{A}} = -999$, $u_{\mathrm{B}} = 100$. To simplify the notation, we treat them as equal to these numbers.)

We first discuss how the game works before simulation enters the picture. In this simulation, Alice would prefer to Bob to pick the Joint Project game and she would cooperate if Bob did pick this game. However, Bob would rather play the Unfair Guessing game, which is virtually guaranteed to make him better off. Realizing this, Alice decides to walks out instead, and the only equilibrium outcome is $u_1 = u_2 = 0$.

Conversely, if simulation is free, the Unfair Guessing game becomes unfavourable to Bob — he would get $u_{\mathrm{B}} = -150$ if he picked it. However, simulating would also allow Alice to betray Bob in the Joint Project game, making him worse off than if he didn't play at all. As a result, when c is equal to 0, Bob always walks out and the only equilibrium outcome is $u_1 = u_2 = 0$, as before.

However, consider the case where $\mathsf{c} = 101$, such that simulation is cheap enough to be justified by the fear of the Unfair Guessing game, but expensive enough to not be justified by the greed in the Joint Project game. Intuitively, we might hope that this will cause Bob to pick the Joint Project game, in which Alice will cooperate — and this is mostly what actually ends up happening. The only wrinkle is that Alice needs to make the decision about simulation before knowing Bob's choice of game — and if she never simulates, Bob would switch to always selecting the Unfair Guessing game instead. As a result, the actual equilibrium (given below) will have Bob sometimes deviating towards Unfair Guessing and Alice sometimes simulating. As a by-product, this will sometimes lead to Alice betraying Bob in the Joint Project Game (when she simulates and he doesn't deviate). However, even with these drawbacks, the resulting outcome is still much better than the default $u_{\mathrm{A}} = u_{\mathrm{B}} = 0$. Indeed, it is not hard to verify that $\mathcal{G}_{\mathrm{sim}}^{101}$ has an equilibrium where Alice simulates with probability $(150 - 100)/(300 - 100) = 1/4$ and trusts Bob otherwise, while Bob picks the Unfair Guessing game with probability $(\mathsf{c} - 100)/(100 + 300) = 1/400$ and selects the Joint Project game otherwise, and the resulting utilities are $u_{\mathrm{A}} \approx 100$, $u_{\mathrm{B}} = (1 - 1/4)100 = 75$. (Note that – somewhat counterintuitively – making the Unfair Guessing game riskier for Bob would make the overall outcome *better* for him, because Alice would not need to simulate with so high probability to disincentivize deviation.)

By adjusting the payoffs in this game, we can obtain examples where the players prefer various values of c that are strictly higher than 0, yet induce equilibria where simulation happens with non-zero probability.

References

Ilan Adler and Renato DC Monteiro. A geometric view of parametric linear programming. *Algorithmica*, 8(1):161–176, 1992.

Mihaly Barasz, Paul Christiano, Benja Fallenstein, Marcello Herreshoff, Patrick LaVictoire, and Eliezer Yudkowsky. Robust cooperation in the prisoner's dilemma: Program equilibrium via provability logic. *arXiv preprint arXiv:1401.5577*, 2014.

Joyce Berg, John Dickhaut, and Kevin McCabe. Trust, reciprocity, and social history. *Games and Economic Behavior*, 10:122–142, 1995.

Joyce Berg, John Dickhaut, and Kevin McCabe. Trust, reciprocity, and social history. *Games and economic behavior*, 10(1):122–142, 1995.

Jesse Clifton. Cooperation, conflict, and transformative artificial intelligence: A research agenda. *Effective Altruism Foundation, March*, 4, 2020.

Vincent Conitzer and Tuomas Sandholm. Computing the optimal strategy to commit to. In *Proceedings of the 7th ACM conference on Electronic commerce*, pages 82–90, 2006.

Andrew Critch, Michael Dennis, and Stuart Russell. Cooperative and uncooperative institution designs: Surprises and problems in open-source game theory. *arXiv preprint arXiv:2208.07006*, 2022.

Andrew Critch. A parametric, resource-bounded generalization of Löb's theorem, and a robust cooperation criterion for open-source game theory. *Journal of Symbolic Logic*, 84(4):1368–1381, 12 2019.

Lance Fortnow. Program equilibria and discounted computation time. In *Proceedings of the 12th Conference on Theoretical Aspects of Rationality and Knowledge*, pages 128–133, 2009.

Joseph Y Halpern and Rafael Pass. Game theory with translucent players. *International Journal of Game Theory*, 47(3):949–976, 2018.

Joseph Y Halpern and Rafael Pass. Sequential equilibrium in games of imperfect recall. *ACM Transactions on Economics and Computation*, 9(4):1–26, 2021.

J. V. Howard. Cooperation in the prisoner's dilemma. *Theory and Decision*, 24:203–213, 5 1988.

Dmytro Korzhyk, Vincent Conitzer, and Ronald Parr. Solving stackelberg games with uncertain observability. In *AAMAS*, pages 1013–1020, 2011.

Joshua Letchford, Dmytro Korzhyk, and Vincent Conitzer. On the value of commitment. *Autonomous Agents and Multi-Agent Systems*, 28(6):986–1016, 2014.

R. Preston McAfee. Effective computability in economic decisions. https://vita.mc4f.ee/PDF/EffectiveComputability.pdf, 1984. Accessed: 2022-12-14.

Caspar Oesterheld. Robust program equilibrium. *Theory and Decision*, 86(1):143–159, 2 2019.

Caspar Oesterheld. A note on the compatibility of different robust program equilibria of the prisoner's dilemma. *arXiv preprint arXiv:2211.05057*, 2022.

Ariel Rubinstein. *Modeling Bounded Rationality.* Zeuthen Lecture Book Series. The MIT Press, 1998.

Walter Rudin. *Real and Complex Analysis.* Mathematics. McGraw.. HiII, 1987.

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

Richard H Smith, Caitlin AJ Powell, David JY Combs, and David Ryan Schurtz. Exploring the when and why of Schadenfreude. *Social and Personality Psychology Compass*, 3(4):530–546, 2009.

Moshe Tennenholtz. Program equilibrium. *Games and Economic Behavior*, 49(2):363–373, 11 2004.

Heinrich von Stackelberg. *Marktform und Gleichgewicht.* Springer, 1934.

Bernhard von Stengel and Shmuel Zamir. Leadership games with convex strategy sets. *Games and Economic Behavior*, 69(2):446–457, 2010.

Acknowledgments

We are grateful to Emanuel Tewolde for pointing out the connection between Proposition 6 and parametrized linear programming, Zuzana Kovarikova for help with Lemma 9, and Lewis Hammond for discussions and feedback on an earlier version of the text. This research was funded by Cooperative AI Foundation.

Contribution Statement