Monitoring versus Discounting in Repeated Games*

Takuo Sugaya Alexander Wolitzky

Stanford GSB MIT

February 2, 2023

Abstract

We study how discounting and monitoring jointly determine whether cooperation is possible in repeated games with imperfect (public or private) monitoring. Our main result provides a simple bound on the strength of players' incentives as a function of discounting, monitoring precision, and on-path payoff variance. We show that the bound is tight in the low-discounting/low-monitoring double limit, by establishing a folk theorem where the discount factor and the monitoring structure can vary simultaneously.

Keywords: repeated games, monitoring precision, blind game, occupation measure, χ^2-divergence, variance decomposition, folk theorem, frequent actions

JEL codes: C72, C73

*Previous title: "Informational Requirements for Cooperation." We thank many seminar participants and the editor and referees for helpful comments. Wolitzky acknowledges financial support from the NSF.

1 Introduction

Supporting non-static Nash outcomes in long-run relationships requires two ingredients. Players' actions must be monitored, so that future play can depend on current behavior. And players must be patient, so that variation in future play can provide incentives. The current paper asks how to measure these ingredients, and how much of each is required. We find that if the ratio of the discount rate and the "detectability" of deviations is large, then all repeated-game Nash outcomes are static ε-correlated equilibria (Theorem 1); and if the ratio of discounting and detectability is small, then all payoff vectors that Pareto-dominate static Nash payoffs can be attained as perfect equilibria in the repeated game (Theorem 2).

Our paper is in the tradition of the folk theorem for repeated games with imperfect public monitoring (Fudenberg, Levine, and Maskin, 1994; henceforth FLM), but we allow arbitrary (possibly private) monitoring and study the tradeoff between discounting and monitoring, rather than the classical limit where discounting vanishes for fixed monitoring. A similar tradeoff between discounting and monitoring arises in repeated games with frequent actions (e.g., Abreu, Milgrom, and Pearce, 1991; Sannikov and Skrzypacz, 2010; henceforth SS), but we do not parameterize the game by an underlying continuous-time signal process, and instead view the frequent-action limit as a particular instance of a low-discounting/low-monitoring double limit. Our results do have implications for games with frequent actions, as well as other applications. These include games with *many players*, where a large population of players are monitored by a noisy aggregate signal; and the question of the *rate of convergence* of the equilibrium payoff set as discounting and monitoring vary. We discuss these applications at the end of the paper, and pursue them further in companion papers (Sugaya and Wolitzky, 2022a,b).

Our negative result (Theorem 1) involves some new ideas. First, we focus on the *amount of information* conveyed by a monitoring structure, rather than the *distribution of information* among the players. We capture this notion by considering the *blind game* Γ^B associated to any repeated game Γ, where the signals that were observed by the players in Γ are instead observed by a neutral mediator. We interpret Γ^B as the repeated game where society has the same amount of information as in Γ, but this information is distributed so as to support a maximally wide range of equilibrium outcomes. Theorem 1 provides a necessary condition for cooperation in Γ^B. A fortiori, the same condition applies for Γ itself, as well as for any other repeated game where the same amount of information is distributed differently—that is, for any repeated game with the same blind game.

Second, we measure the *average* strength of a player's incentives over all histories that arise in the course of the game. This notion is captured by a player's maximum deviation gain at the *occupation measure over actions* induced by an equilibrium. Here our approach contrasts with earlier work that analyzes incentives history-by-history (e.g., Fudenberg, Levine, and Pesendorfer, 1998; al-Najjar and Smorodinsky, 2000, 2001; Awaya and Krishna, 2016, 2019). It leads to sharper results, because sometimes an equilibrium can be constructed that provides strong incentives at a particular history by letting continuation play depend disproportionately on behavior at that history, but such a construction necessarily provides weaker incentives at other histories.

Third, we measure the detectability of a deviation by the χ^2-*divergence*—the variance of the likelihood ratio difference—between the signal distribution under equilibrium play as compared to that under the deviation. The χ^2-divergence is a standard measure of statistical distance.[1] It arises in our analysis because minimizing continuation payoff variance subject to incentive constraints requires making continuation payoffs proportional to likelihood ratio differences, with a constant of proportionality equal to the (inverse) χ^2-divergence.

In total, Theorem 1 may be summarized as stating that, for any repeated game Γ, any Nash equilibrium outcome in the associated blind game Γ^B, and any possible deviation by any player, we have

$$\text{deviation gain} \leq \sqrt{\frac{\delta}{1-\delta}\,(\text{detectability})\,(\text{payoff variance})},$$

where the deviation gain, detectability (measured by χ^2-divergence), and payoff variance are all assessed at the equilibrium occupation measure. The proof is based on a simple but novel

[1]Several other standard measures (e.g., total variation distance, Kullback-Leibler divergence) are equivalent to χ^2-divergence under a non-moving support assumption, and hence are equally valid for characterizing the tradeoff between discounting and monitoring in the low-discounting/low-monitoring double limit. However, our proofs and intuition rely on χ^2-divergence, and our non-asymptotic results are strongest under this measure. See footnote 8 for details on this point.

variance decomposition argument. The idea is that, if deviating from non-static Nash play is unprofitable, then signals must vary significantly with actions, and continuation payoffs must vary significantly with signals; and, moreover, this payoff variation must be delivered relatively quickly due to discounting. Theorem 1 shows that recursively decomposing the variance of a player's continuation payoffs across periods yields a relatively tight bound on the average strength of her incentives.[2]

Our positive result (Theorem 2) is a partial converse to Theorem 1. It shows that the tradeoff between discounting and monitoring expressed in Theorem 1 is tight up to constant factors in the low-discounting/low-monitoring double limit. Theorem 2 is an extension of the folk theorems of FLM, Kandori and Matsushima (1998; henceforth KM), and SS. It generalizes FLM and KM by letting discounting and monitoring vary simultaneously, and it generalizes SS by considering the general low-discounting/low-monitoring double limit, rather than parameterizing monitoring by an underlying continuous-time signal process.[3] A limitation of Theorem 2 is that it assumes that monitoring has a *product structure*. This assumption facilitates an easy comparison with Theorem 1, but it is overly strong from the perspective of the prior work on repeated games such as FLM, KM, and SS. However, we prove Theorem 2 as a corollary of a more general result, Theorem 3, which we present in the appendix, and which does not assume product structure monitoring.

The tradeoff we find between discounting and monitoring has a clear interpretation. In probability theory, the sum of the conditional variances of a martingale's increments is often a useful measure of the "intrinsic time" experienced by the martingale (e.g., Dubins and Savage, 1965; Freedman, 1975). Analogously, our results show precisely that repeated-game equilibrium play is approximately myopic if players are impatient, and a folk theorem holds if players are patient, where patience is measured relative to the intrinsic time experienced by a martingale with likelihood ratio difference increments, rather than calendar time.

[2] We emphasize that this bound applies for any repeated-game Nash equilibrium, regardless of whether monitoring is public or private, despite the well-known fact that the equilibrium payoff set with private strategies or monitoring generally lacks a useful recursive structure (Kandori, 2002).

[3] We also allow any number of players, while SS consider two-player games.

2 Preliminaries

The Repeated Game. We consider discounted repeated games with imperfect monitoring.

A *stage game* $G = (I, A, u)$ consists of a finite set of players $I = \{1, \ldots, N\}$, a finite product set of actions $A = \times_{i \in I} A_i$, and a payoff function $u_i : A \to \mathbb{R}$ for each $i \in I$. Let $\bar{u} > 0$ denote an upper bound on the range and magnitude of any player's stage-game payoff: e.g., take $\bar{u} = \max_{i,a} 2 |u_i(a)|$. We denote a (possibly correlated) distribution over action profiles by $\alpha \in \Delta(A)$, and denote the set of such distributions resulting from independent mixing by $\Delta^*(A) = \times_{i \in I} \Delta(A_i)$. For any action profile distribution $\alpha \in \Delta(A)$, we let $u_i(\alpha) := \mathbb{E}_{a \sim \alpha}[u_i(a)]$ and $V_i(\alpha) := \mathrm{Var}_{a \sim \alpha}(u_i(a))$ denote the mean and variance of player i's payoff under α.

A *monitoring structure* (Y, p) consists of a finite product set of possible signal realizations $Y = \times_{i \in I} Y_i$ and a family of conditional probability distributions $p(y|a)$, which we assume have common support $\bar{Y} \subseteq Y$: that is, for each y, a, we have $p(y|a) > 0$ iff $y \in \bar{Y}$. This non-moving support assumption excludes perfect monitoring (where $y_i = a$ with probability 1 for all i). Throughout, whenever we take a sum over signals y, this sum should be read as being taken over $\bar{Y}$ rather than Y, so that 0-probability signal profiles are excluded.

A *repeated game* $\Gamma = (G, Y, p, \delta)$ is described by a stage game, a monitoring structure, and a discount factor $\delta \in (0, 1)$. In each period $t = 1, 2, \ldots$, (i) the players take actions $(a_i)_i$, (ii) the signal $y = (y_i)_i$ is drawn according to $p((y_i)_i \,|\, (a_i)_i)$, and (iii) each player i observes y_i. Players remember their own past actions, so a *history* for player i takes the form $h_i^t = (a_{i,t'}, y_{i,t'})_{t'=1}^{t-1}$, and a *strategy* σ_i for player i maps histories h_i^t to distributions over actions $a_{i,t}$. Players maximize discounted expected payoffs with discount factor δ.

An *outcome* μ of the repeated game is a distribution over paths of actions and signals, $(A \times Y)^\infty$. Each strategy profile σ induces a unique outcome μ.

The monitoring structure is said to be *public* if $y_i = y_j$ for all $y \in \bar{Y}, i, j \in I$. When considering public monitoring, we omit the player subscript on y. In addition, public monitoring has a *product structure* if there exist sets $(Y^i)_{i \in N}$ and conditional probability distributions $(p^i(\cdot|a_i))_{i \in N}$ on $(Y^i)_{i \in N}$ such that $Y = \times_{i \in N} Y^i$ and $p(y|a) = \prod_{i \in N} p^i(y^i|a^i)$ for all $y \in Y, a \in A$. Thus, with public, product structure monitoring, y^i is a conditionally independent signal of player i's action a_i (not to be confused with the signal observed by player i in a general monitoring structure, which is denoted by y_i).

The Blind Game. For any repeated game Γ, the set of outcomes μ that are induced by any Nash equilibrium σ (or moreover by any *communication equilibrium*, as in Forges, 1986) is smaller than the set of outcomes that are induced by a Nash equilibrium in the corresponding *blind game*. The blind game, which we denote by Γ^B, is a variant of Γ where (i) the players have access to a neutral mediator, (ii) at the beginning of each period, the mediator privately recommends an action $r_i \in A_i$ to each player i, and (iii) at the end of each period, the mediator observes the signal y (which continues to be drawn according to $p\left((y_i)_i \mid (a_i)_i\right)$), while the players observe nothing. Players remember their own past actions, while the mediator does not observe the players' actions. Thus, a history for player i in the blind repeated game Γ^B takes the form $h_i^t = \left((r_{i,t'}, a_{i,t'})_{t'=1}^{t-1}, r_{i,t}\right)$, and a history for the mediator takes the form $h_0^t = \left((r_{i,t'})_i, (y_{i,t'})_i\right)_{t'=1}^{t-1}$. A strategy σ_i for player i maps histories h_i^t to distributions over actions $a_{i,t}$; a strategy σ_0 for the mediator maps histories h_0^t to distributions over recommendation profiles $(r_{i,t})_i$. By standard arguments (similar to Forges, 1986), any outcome μ that is induced by a Nash or communication equilibrium in Γ is also induced by a Nash equilibrium in Γ^B where the players follow the mediator's recommendations on path. Our necessary conditions for cooperation (Theorem 1) apply for Γ^B, and hence apply a fortiori for Γ.

Occupation Measures. Given an outcome μ, let $\alpha_t^\mu \in \Delta(A)$ denote the marginal distribution of period-t action profiles under μ, and define $\alpha^\mu \in \Delta(A)$, the *occupation measure over action profiles induced by* μ, by

$$\alpha^\mu(a) = (1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}\alpha_t^\mu(a) \qquad \text{for all } a \in A.$$

The occupation measure α^μ describes the "discounted expected fraction of periods" where each action profile is played in the course of the repeated game. Note that the payoffs under an outcome μ are determined by its occupation measure α^{μ}, as

$$(1-\delta)\sum_t \delta^{t-1}\sum_a \alpha_t^{\mu}(a)\,u(a)=\sum_a (1-\delta)\sum_t \delta^{t-1}\alpha_t^{\mu}(a)\,u(a)=\sum_a \alpha^{\mu}(a)\,u(a)=u(\alpha^{\mu}).$$

In other words, the occupation measure is a sufficient statistic for the players' payoffs.

Manipulations. A *manipulation* for a player i is a mapping $s_i : A_i \to \Delta(A_i)$. The interpretation is that when player i is recommended action a_i, she instead plays $s_i(a_i)$.

The *gain* from a manipulation s_i at an action profile distribution $\alpha \in \Delta(A)$ is

$$g_i(s_i,\alpha)=\sum_a \alpha(a)\left(u_i(s_i(a_i),a_{-i})-u_i(a)\right).$$

For any $\varepsilon > 0$, an action profile distribution α is a *static ε-correlated equilibrium* if $g_i(s_i,\alpha) \leq \varepsilon$ for all i and s_i.

For any $\alpha \in \Delta(A)$, let $p(y|\alpha) = \sum_a \alpha(a)\,p(y|a)$. We define the *detectability* of a manipulation s_i at an action profile distribution α as

$$\chi_i^2(s_i,\alpha)=\sum_{a,y}\alpha(a)\,p(y|a)\left(\frac{p(y|a)-p(y|s_i(a_i),a_{-i})}{p(y|a)}\right)^2.$$

When $\alpha(a) = 1$ for some $a \in A$, our detectability measure is the χ^2-*divergence* between the probability distributions $p(\cdot|a)$ and $p(\cdot|s_i(a_i),a_{-i})$. (The measure extends linearly for non-degenerate α.) The χ^2-divergence is a standard measure of statistical distance. Note that it is well-defined by our non-moving support assumption.[4]

We emphasize that manipulations, gain, and detectability are all "static" concepts, in that they are defined relative to a single action profile distribution and (for detectability) a single draw from the monitoring structure.

Remark 1 *Why does χ^2-divergence arise in our analysis? The χ^2-divergence equals the variance of the likelihood ratio difference between $p(\cdot|a)$ and $p(\cdot|s_i(a_i),a_{-i})$. The likelihood ratio difference $(p(y|a)-p(y|\tilde{a}_i,a_{-i}))/p(y|a)$ determines the "strength of incentives" provided by*

[4] Recall that we have also assumed that Y is finite. Theorem 1 goes through when Y is infinite, provided that $\chi_i^2(s_i,\alpha)$ is finite for all i, s_i, α.

rewards or punishments that are conditioned on the arrival of signal y (Mirrlees, 1975; Holmström, 1979). Since the expected likelihood ratio difference $\sum_y p(y|a)\left(\left(p(y|a)-p(y|\tilde{a}_i,a_{-i})\right)/p(y|a)\right)$ equals 0, the likelihood ratio difference is "often large"—so the signal is a useful basis for incentives—if and only if its variance is large.

More concretely, χ^2-divergence arises in Theorem 1 by applying the Cauchy-Schwarz inequality to an expression similar to

$$\begin{aligned} &\sum_y \left(p(y|a)-p(y|s_i(a_i),a_{-i})\right) w_i(y) \\ = &\sum_y p(y|a)\left(\frac{p(y|a)-p(y|s_i(a_i),a_{-i})}{p(y|a)}\right)\left(w_i(y)-\mathbb{E}\left[w_i(\tilde{y})\right]\right), \end{aligned}$$

where $w_i(y)$ denotes player i's continuation payoff following signal y. This expression is the loss in player i's expected continuation payoff when she manipulates according to s_i at action profile a. For the inner product $\langle X,Y\rangle_a=\sum_y p(y|a)X(y)Y(y)$, Cauchy-Schwarz upper-bounds this loss by

$$\sqrt{\chi_i^2(s_i,a)\operatorname{Var}(w_i(y))}.$$

This observation shows that χ^2-divergence and continuation payoff variance must both be large to deter manipulations. It also suggests that, as we will see, χ^2-divergence is a useful metric for analysis based on decomposing the variance of continuation payoffs.

Conversely, χ^2-divergence arises in Theorem 2 because the smallest χ^2-divergence $\chi_i^2(s_i,a)$ among manipulations s_i that always disobey the recommendation a_i is equal to the amount of slack in a standard statistical identifiability condition for the folk theorem with imperfect public monitoring.[5]

An intuition for why Theorems 1 and 2 are near-converses is that Cauchy-Schwarz is tight when the likelihood ratio differences $\left(p(y|a)-p(y|s_i(a_i),a_{-i})\right)/p(y|a)$ and the continuation payoffs $w_i(y)$ are co-linear, and making these quantities co-linear minimizes continuation

[5] See Lemma 3 in Appendix A for a formal statement.

payoff variance subject to incentive-compatibility. That is, the solution to the program

$$\min_{w_i:\bar{Y}\to\mathbb{R}} \sum_y p(y|a)\left(w_i(y) - \mathbb{E}\left[w_i(\tilde{y})\right]\right)^2$$

$$s.t. \qquad \sum_y \left(p(y|a) - p(y|s_i(a_i), a_{-i})\right) w_i(y) \geq \frac{1-\delta}{\delta}\left(u_i(s_i(a_i), a_{-i}) - u_i(a)\right)$$

is

$$w_i(y) = \left(u_i(s_i(a_i), a_{-i}) - u_i(a)\right) \times \frac{p(y|a) - p(y|a_i', a_{-i})}{p(y|a)} \times \frac{1-\delta}{\delta} \times \frac{1}{\chi_i^2(s_i, a_i)}.$$

In turn, minimizing continuation payoff variance maximizes efficiency when continuation payoff movements are small and are approximately confined to the boundary of a smooth set of payoffs, which is the most efficient way to provide incentives in repeated games (e.g., FLM; Sannikov, 2007).[6]

3 Bounding Equilibrium Incentives

Our main result bounds a player's gain from a manipulation as a function of the discount factor, the detectability of the manipulation, and the variance of the player's payoff, where gain, detectability, and variance are all assessed at the equilibrium occupation measure. As a consequence, every repeated-game equilibrium occupation measure is a static ε-correlated equilibrium, and every repeated-game equilibrium payoff vector is a static ε-correlated equilibrium payoff vector, for $\varepsilon > 0$ given by the bound.

Theorem 1 *For any Nash equilibrium outcome μ in Γ^B, any player i, and any manipulation s_i, we have*

$$g_i(s_i, \alpha^\mu) \leq \sqrt{\frac{\delta}{1-\delta}\chi_i^2(s_i, \alpha^\mu) V_i(\alpha^\mu)}. \tag{1}$$

In particular, α^μ is a static ε-correlated equilibrium (and hence payoffs under μ are static

[6] The χ^2-divergence is closely related to the *Fisher information.* If a_i were a continuous variable, the Fisher information would be defined as $\sum_y p(y|a)\left(\frac{\partial}{\partial a_i} p(y|a_i, a_{-i}) / p(y|a)\right)^2$, which is a local χ^2-divergence. Fisher information arises in moral hazard problems with quadratic utility (Jewitt, Kadan, and Swinkels, 2008; Hébert, 2018) or frequent actions (Sadzik and Stacchetti, 2015), as well as in some career concerns models (Dewatripont, Jewitt, and Tirole, 1999), because these problems likewise involve minimizing the variance of rewards subject to incentive compatibility.

ε-correlated equilibrium payoffs), for

$$\varepsilon = \max_{i,s_i} \sqrt{\frac{\delta}{1-\delta} \chi_i^2 \left(s_i, \alpha^\mu\right) V_i \left(\alpha^\mu\right)}.$$

Theorem 1 precludes cooperation when players are too impatient, monitoring is too imprecise, or on-path payoff variance is too small. It permits cooperation if $\delta \to 1$ for any fixed positive detectability, consistent with FLM's folk theorem. It also permits cooperation with vanishing on-path payoff variance if detectability is high enough, consistent with folk theorems under perfect monitoring (which we admit as a limit case). We emphasize that the theorem covers all Nash equilibria, whether signals are observed publicly or privately, by either the players or a mediator.

Theorem 1 implies that cooperation is impossible if detectability is much smaller than discounting. We record this implication as a corollary.

Corollary 1 *For any stage game G and any $\varepsilon > 0$, there exists $k > 0$ such that the following holds:*

For any monitoring structure (Y, p), any discount factor δ satisfying

$$\frac{\max_{i,s_i,a} \chi_i^2 \left(s_i, a\right)}{1-\delta} < k, \tag{2}$$

and any Nash equilibrium outcome μ in the repeated game $\Gamma = (G, Y, p, \delta)$ (or in the blind game Γ^B), the induced occupation measure over actions α^μ is a static ε-correlated equilibrium.

An important feature of Theorem 1 is that the deviation gain is bounded by a multiple of $(1-\delta)^{-1/2}$, rather than $(1-\delta)^{-1}$. This is somewhat surprising, as continuation payoffs are weighted by $(1-\delta)^{-1}$, and it is essential for characterizing the tradeoff between discounting and monitoring (e.g., for establishing Corollary 1). The key idea behind this property is bounding incentives on average, not at each history. In particular, the proof of Theorem 1 shows that if (1) is violated, then there exists a period t such that it is profitable for player i to follow the equilibrium until period t and then manipulate according to s_i. However, this deviation may be profitable only for certain choices of t—it may be unprofitable for a period t that gets disproportionate weight in determining continuation payoffs. Put differently, an incentive bound of order $(1-\delta)^{-1}$ results when no restrictions are placed on continuation payoffs beyond feasibility, while we instead recursively bound the variance of continuation payoffs, which results in an incentive bound of order $(1-\delta)^{-1/2}$.

Remark 2 *Prior results that bound incentives in repeated games as a function of discounting and monitoring precision do so history-by-history, and hence obtain bounds of order* $(1-\delta)^{-1}$ *(e.g., Fudenberg, Levine, and Pesendorfer, 1998, Proposition 1; al-Najjar and Smorodinsky, 2001, Theorem 1; Pai, Roth, and Ullman, 2017, Theorem 3.1). Awaya and Krishna (2016, 2019) derive a bound based on deterring a permanent deviation to a fixed action, which is also of order* $(1-\delta)^{-1}$*.*[7] *In our own prior work, (Sugaya and Wolitzky, 2017, 2018), we derived bounds that hold independently of monitoring precision; these are again of order* $(1-\delta)^{-1}$*.*

We illustrate Theorem 1 with an example.

Example 1 (Prisoner's Dilemma with Binary Product Structure Monitoring) *Consider the prisoner's dilemma with payoff matrix*

	C	D
C	$1,1$	$-1,2$
D	$2,-1$	$0,0$

and symmetric product structure monitoring with precision $\pi \in (1/2,1)$*, so that* $Y = \{C,D\} \times \{C,D\}$*, where each signal component equals the corresponding player's action with probability* π*, independently across players.*

We bound the equilibrium probability of cooperation by applying (1) for the manipulation that always defects. For any equilibrium outcome μ*, the gain from this deviation evaluated*

[7] See, e.g., Awaya and Krishna (2019, Proposition 4.1). Unlike our bound, their bound for each player i depends only on the marginal of p on Y_{-i}, so our bound and theirs are non-nested. Their bound is tighter for monitoring structures where the impact of a player's action on the distribution of y is much greater than its impact on the distribution of y_{-i}. Such monitoring structures play an important role in their analysis. It may be possible to use our techniques to improve their bound; this is left for future research.

at the occupation measure α^μ equals $\alpha^\mu_{CC}+\alpha^\mu_{CD}$, while its detectability evaluated at α^μ equals

$$\begin{aligned}&\left(\alpha^\mu_{CC}+\alpha^\mu_{CD}\right)\left(\pi\left(\frac{\pi-(1-\pi)}{\pi}\right)^2+(1-\pi)\left(\frac{(1-\pi)-\pi}{1-\pi}\right)^2\right)+\left(\alpha^\mu_{DC}+\alpha^\mu_{DD}\right)(0)\\ =\;&\left(\alpha^\mu_{CC}+\alpha^\mu_{CD}\right)\frac{(2\pi-1)^2}{\pi(1-\pi)}.\end{aligned}$$

Thus, (1) gives

$$\alpha^\mu_{CC}+\alpha^\mu_{CD}\leq\frac{\delta}{1-\delta}\frac{(2\pi-1)^2}{\pi(1-\pi)}V_1(\alpha^\mu),\tag{3}$$

where $V_1(\alpha^\mu)=\alpha^\mu_{CC}+4\alpha^\mu_{DC}+\alpha^\mu_{CD}-\left(\alpha^\mu_{CC}+2\alpha^\mu_{DC}-\alpha^\mu_{CD}\right)^2$.

Inequality (3) can be further simplified to bound the players' average equilibrium payoff, $w:=(u_1(\alpha^\mu)+u_2(\alpha^\mu))/2$. Note that $w\leq(1+\alpha^\mu_{CC})/2$ and, for a given value for α^μ_{CC}, $\min_i V_i(\alpha^\mu)$ is maximized by taking $\alpha_{CD}=\alpha_{DC}=(1-\alpha^\mu_{CC})/2$, which gives $V_i(\alpha^\mu)=(5/2)(1-\alpha^\mu_{CC})-(1/4)(1-\alpha^\mu_{CC})^2\leq(5/2)(1-\alpha^\mu_{CC})$. Inequality (3) now implies that

$$\alpha^\mu_{CC}\leq\frac{5}{2}\frac{\delta}{1-\delta}\frac{(2\pi-1)^2}{\pi(1-\pi)}\left(1-\alpha^\mu_{CC}\right)\implies\alpha^\mu_{CC}\leq\max\left\{1-\frac{2}{5}\frac{1-\delta}{\delta}\frac{\pi(1-\pi)}{(2\pi-1)^2},0\right\}.$$

We thus obtain the payoff bound

$$w\leq\max\left\{\frac{1}{2},1-\frac{1}{5}\frac{1-\delta}{\delta}\frac{\pi(1-\pi)}{(2\pi-1)^2}\right\}.\tag{4}$$

In Section 4, we will quantify the tightness of this bound in the frequent-action limit where the prisoner's dilemma converges to Sannikov's (2007) continuous-time partnership game, where each player controls the drift of a Brownian motion.

There are three steps in the proof of Theorem 1. First, if manipulating according to s_i is unprofitable in period t, then the conditional variance of player i's period-$t+1$ continuation payoff must be sufficiently large compared to $(1-\delta)^2$ times the ratio of the (squared) gain from this manipulation in period t and the detectability of this manipulation in period t (equation (7) below). Second, applying this lower bound on conditional variance recursively using the law of total variance, we show that a discounted sum of the variances of player i's stage-game payoffs (times $1-\delta$) must exceed a discounted sum of the conditional variance bounds (equation (8), which obtains after canceling a $1-\delta$ term). Finally, by Jensen's inequality, this inequality relating a discounted sum of payoff variances and a discounted sum of ratios of the deviation gain and detectability of s_i in each period implies a corresponding inequality relating the payoff variance and the ratio of the deviation gain and detectability of s_i evaluated at the equilibrium occupation measure, which simplifies to (1).

We also mention a tighter (but more complicated) bound than that given in Theorem 1, which applies for any communication equilibrium outcome μ in Γ, but not necessarily for any equilibrium outcome in Γ^B. This is the bound that results when the mediator must rely on self-reported signals, so that detectability is now measured with respect to a player's opponents' signals and her own self-report. Specifically, a manipulation for player i would now consist of a pair (s_i, ρ_i), where $s_i : A_i \to \Delta(A_i)$ describes the mixed action $s_i(a_i)$ taken by player i when she is recommended a_i, and $\rho_i : A_i \times A_i \times Y_i \to Y_i$ describes the signal $\rho_i(a_i, \hat{a}_i, y_i)$ reported by player i when she is recommended a_i, takes $\hat{a}_i$, and observes y_i. One can then define the gain from a manipulation (s_i, ρ_i) as above (noting that this depends only on s_i), and define the detectability of a manipulation (s_i, ρ_i) at an action profile distribution α as

$$\tilde{\chi}_i^2(s_i, \rho_i, \alpha) = \sum_{a,y} \alpha(a) p(y|a) \left(\frac{p(y|a) - \sum_{\hat{a}_i, y_i'} s_i(a_i)[\hat{a}_i] p(y_i', y_{-i}|\hat{a}_i, a_{-i}) \rho_i(a_i, \hat{a}_i, y_i')[y_i]}{p(y|a)} \right)^2.$$

Note that

$$\chi_i^2(s_i, \alpha) \geq \tilde{\chi}_i^2(s_i, \alpha) := \min_{\rho_i} \tilde{\chi}_i^2(s_i, \rho_i, \alpha) \quad \text{for all } i, s_i, \alpha,$$

as this inequality holds with equality when $\rho_i(a_i, \hat{a}_i, y_i) = y_i$ for all $a_i, \hat{a}_i, y_i$. Theorem 1 holds for any communication equilibrium outcome μ in Γ with $\tilde{\chi}_i^2(s_i, \alpha^\mu)$ in place of $\chi_i^2(s_i, \alpha^\mu)$, by essentially the same proof.

3.1 Proof of Theorem 1

We first introduce some notation. Given a path of action profiles $a^\infty = (a^1, a^2, \ldots)$, let $u_i^t = u_i(a^t)$, and denote player i's continuation payoff at the beginning of period t by

$$w_i^t = (1-\delta)\sum_{t'=t}^{\infty} \delta^{t'-t} u_i^{t'}.$$

Denote a history of actions and signals at the beginning of period t by $h^t = (a^t, y^t)$.

Fix a Nash equilibrium outcome μ in Γ^B, a player i, and a manipulation s_i. Let H^t denote the set of period-t histories h^t that are reached with positive probability under μ, and define a H^t-measurable random variable $W_i^t : H^t \to \mathbb{R}$ by $W_i^t(h^t) = \mathbb{E}[w_i^t | h^t]$ for all $h^t \in H^t$. By the law of total variance (e.g., Billingsley, 1995, Problem 34.10(b)), we have

$$\operatorname{Var}\left(W_i^{t+1}\right) = \operatorname{Var}\left(\mathbb{E}\left[W_i^{t+1}|h^t\right]\right) + \mathbb{E}\left[\operatorname{Var}\left(W_i^{t+1}|h^t\right)\right]. \tag{5}$$

Similarly, define $U_i^t : H^t \to \mathbb{R}$ by $U_i^t(h^t) = \mathbb{E}[u_i^t | h^t]$ for all $h^t \in H^t$.

In what follows, we suppress the dependence of $g_i(s_i, \alpha)$ and $\chi_i^2(s_i, \alpha)$ on s_i.

Lemma 1 *For each period t, we have*

$$\operatorname{Var}\left(\mathbb{E}\left[W_i^{t+1}|h^t\right]\right) \geq \frac{1}{\delta}\operatorname{Var}\left(W_i^t\right) - \frac{1-\delta}{\delta}\operatorname{Var}\left(U_i^t\right) \qquad \textit{and} \tag{6}$$

$$\mathbb{E}\left[\operatorname{Var}(W_i^{t+1}|h^t)\right] \geq \left(\frac{1-\delta}{\delta}\right)^2 \frac{g_i\left(\alpha_t^\mu\right)^2}{\chi_i^2\left(\alpha_t^\mu\right)}, \tag{7}$$

where in (7) we follow the convention $0/0 = 0$. In particular, (7) implies that $\chi_i^2(\alpha_t^\mu) = 0 \implies g_i(\alpha_t^\mu) = 0$.

Proof. For (6), since $w_i^t = (1-\delta)u_i^t + \delta w_i^{t+1}$, for every history $h^t \in H^t$ we have

$$W_i^t\left(h^t\right) = (1-\delta)U_i^t\left(h^t\right) + \delta\mathbb{E}\left[W_i^{t+1}|h^t\right].$$

Therefore,

$$\operatorname{Var}\left(W_i^t\right) = \operatorname{Var}\left((1-\delta)U_i^t + \delta\mathbb{E}\left[W_i^{t+1}|h^t\right]\right) \leq (1-\delta)\operatorname{Var}\left(U_i^t\right) + \delta\operatorname{Var}\left(\mathbb{E}\left[W_i^{t+1}|h^t\right]\right).$$

Dividing by δ and rearranging yields (6).

For (7), let $\mu(h^t, a)$ denote the probability that history h^t is reached in period t and then action profile a is played. Since μ is an equilibrium outcome, we have

$$\frac{1-\delta}{\delta} g_i(\alpha_t^\mu) \leq \sum_{h^t,a,y} \mu\left(h^t, a\right)\left(p(y|a) - p(y|s_i(a_i), a_{-i})\right) W_i^{t+1}\left(h^t, a, y\right).$$

This holds because, if she follows the equilibrium until period t and then manipulates according to s_i—which is a feasible deviant strategy, albeit perhaps not an optimal one—player i can guarantee an expected continuation payoff of $\sum_{h^t,a,y} \mu(h^t, a) p(y|s_i(a_i), a_{-i}) W_i^{t+1}(h^t, a, y)$ by following the mediator's recommendations from period $t+1$ onward. (In other words, in the continuation game player i plays as if her period-t action were a_i rather than $s_i(a_i)$. This continuation play may not be optimal, but we are only giving a necessary condition.) Therefore,

$$\begin{aligned}
\frac{1-\delta}{\delta} g_i(\alpha_t^\mu) &\leq \sum_{h^t,a,y} \mu\left(h^t, a\right)\left(p(y|a) - p(y|s_i(a_i), a_{-i})\right) W_i^{t+1}\left(h^t, a, y\right) \\
&= \sum_{h^t,a,y} \mu\left(h^t, a\right) p(y|a) \left(\frac{p(y|a) - p(y|s_i(a_i), a_{-i})}{p(y|a)}\right)\left(W_i^{t+1}\left(h^t, a, y\right) - \mathbb{E}\left[W_i^{t+1}|h^t\right]\right) \\
&\leq \sqrt{\textstyle\sum_{h^t,a,y} \mu(h^t, a) p(y|a) \left(\frac{p(y|a)-p(y|s_i(a_i),a_{-i})}{p(y|a)}\right)^2} \\
&\quad \times \sqrt{\textstyle\sum_{h^t,a,y} \mu(h^t, a) p(y|a) \left(W_i^{t+1}(h^t, a, y) - \mathbb{E}\left[W_i^{t+1}|h^t\right]\right)^2} \\
&= \sqrt{\chi_i^2(\alpha_t^\mu) \mathbb{E}\left[\mathrm{Var}\left(W_i^{t+1}|h^t\right)\right]},
\end{aligned}$$

where the second inequality follows from Cauchy-Schwarz. Finally, if $\chi_i^2(\alpha_t^\mu) > 0$ then squaring both sides and rearranging yields (7); if instead $\chi_i^2(\alpha_t^\mu) = 0$ then we have $g_i(\alpha_t^\mu) = 0$, and (7) reduces to $\mathbb{E}\left[\mathrm{Var}(W_i^{t+1}|h^t)\right] \geq 0$, which holds as variance is non-negative. ■

(The different orders in $1-\delta$ in (6) and (7) are important and can be given an intuitive explanation. In (6), $\mathrm{Var}(U_i^t)$ is weighted by $1-\delta$ because current-period payoffs have weight $1-\delta$, and variance is maximized when current payoffs and continuation payoffs are perfectly correlated, in which case the weight comes out of variance linearly. In (7), $g_i(\alpha_t^\mu)^2 / \chi_i^2(\alpha_t^\mu)$ is weighted by $(1-\delta)^2$ because we bounded the current-period deviation gain $(1-\delta) g_i(\alpha_t^\mu)$

using Cauchy-Schwarz and squared both sides of the resulting inequality.)

By (5), (6), and (7), for each period t, we have

$$\mathrm{Var}\left(W_i^{t+1}\right) \geq \frac{1}{\delta}\mathrm{Var}\left(W_i^t\right) - \frac{1-\delta}{\delta}\mathrm{Var}\left(U_i^t\right) + \left(\frac{1-\delta}{\delta}\right)^2 \frac{g_i\left(\alpha_t^\mu\right)^2}{\chi_i^2\left(\alpha_t^\mu\right)}.$$

Recursively applying this inequality and using $\mathrm{Var}\left(W_i^1\right) = 0$, for each $T \in \mathbb{N}$ we have

$$\delta^T \mathrm{Var}\left(W_i^{T+1}\right) \geq (1-\delta)\sum_{t=1}^{T} \delta^{t-1}\left(\frac{1-\delta}{\delta}\frac{g_i\left(\alpha_t^\mu\right)^2}{\chi_i^2\left(\alpha_t^\mu\right)} - \mathrm{Var}\left(U_i^t\right)\right).$$

As payoffs are bounded, the left-hand side of this inequality converges to 0 as $T \to \infty$, while (since $\chi_i^2\left(\alpha_t^\mu\right)$ is also bounded) the right-hand side converges to

$$(1-\delta)\sum_{t} \delta^{t-1}\left(\frac{1-\delta}{\delta}\frac{g_i\left(\alpha_t^\mu\right)^2}{\chi_i^2\left(\alpha_t^\mu\right)} - \mathrm{Var}\left(U_i^t\right)\right).$$

Therefore,

$$\delta\sum_{t} \delta^{t-1}\mathrm{Var}\left(U_i^t\right) \geq (1-\delta)\sum_{t} \delta^{t-1}\frac{g_i\left(\alpha_t^\mu\right)^2}{\chi_i^2\left(\alpha_t^\mu\right)}. \tag{8}$$

At this point we are almost done, because inequality (8) actually implies the desired inequality, (1). This observation relies on the following lemma.

Lemma 2 *Let* $\Delta^+(A) = \{\alpha \in \Delta(A) : \forall i,\ \chi_i^2(\alpha) = 0 \Rightarrow g_i(\alpha) = 0\}$. *The function* $f_i : \Delta^+(A) \to \mathbb{R}_+$ *defined by*

$$f_i(\alpha) = \frac{g_i(\alpha)^2}{\chi_i^2(\alpha)} \qquad \text{for all } \alpha \in \Delta(A),$$

with convention $0/0 = 0$, *is convex.*

Proof. Fix any $\alpha, \alpha' \in \Delta^+(A)$ and $\beta \in [0,1]$, and let

$$a = g_i(\alpha), \qquad b = \chi_i^2(\alpha), \qquad c = g_i(\alpha'), \qquad d = \chi_i^2(\alpha').$$

By linearity of g_i and χ_i^2, we have

$$\beta f_i(\alpha) + (1-\beta) f_i(\alpha') - f_i(\beta\alpha + (1-\beta)\alpha') = \beta\frac{a^2}{b} + (1-\beta)\frac{c^2}{d} - \frac{(\beta a + (1-\beta)c)^2}{\beta b + (1-\beta)d} \geq 0,$$

so f_i is convex. To see why the last inequality holds, note that if $b = 0$ then $a = a^2/b = 0$ (by $\alpha \in \Delta^+(A)$ and the $0/0 = 0$ convention), so the inequality is trivial, and similarly if $d = 0$. If instead b and d are both strictly positive, then we have

$$\beta\frac{a^2}{b} + (1-\beta)\frac{c^2}{d} - \frac{(\beta a + (1-\beta)c)^2}{\beta b + (1-\beta)d} = \frac{\beta(1-\beta)(ad-bc)^2}{(\beta b + (1-\beta)d)\,bd} \geq 0.$$

■

We also use the fact that

$$\begin{aligned}
\frac{\delta}{1-\delta} V_i(\alpha^\mu) &= \frac{\delta}{1-\delta}\sum_a (1-\delta)\sum_t \delta^{t-1}\alpha_t^\mu(a)\left(u_i(a) - u_i(\alpha^\mu)\right)^2 \\
&= \delta\sum_t \delta^{t-1}\sum_a \alpha_t^\mu(a)\left(u_i(a) - u_i(\alpha^\mu)\right)^2 \\
&\geq \delta\sum_t \delta^{t-1}\sum_a \alpha_t^\mu(a)\left(u_i(a) - u_i(\alpha_t^\mu)\right)^2 \\
&= \delta\sum_t \delta^{t-1}\mathrm{Var}\left(u_i^t\right) \geq \delta\sum_t \delta^{t-1}\mathrm{Var}\left(U_i^t\right),
\end{aligned}$$

where the first inequality follows because $\mathbb{E}\left[(X-x)^2\right] \geq \mathbb{E}\left[(X - \mathbb{E}[X])^2\right]$ for any random variable X and number x, and the second inequality follows from the law of total variance. By (8), we thus have

$$\frac{\delta}{1-\delta} V_i(\alpha^\mu) \geq (1-\delta)\sum_t \delta^{t-1}\frac{g_i(\alpha_t^\mu)^2}{\chi_i^2(\alpha_t^\mu)} \geq \frac{\left((1-\delta)\sum_t \delta^{t-1} g_i(\alpha_t^\mu)\right)^2}{(1-\delta)\sum_t \delta^{t-1}\chi_i^2(\alpha_t^\mu)} = \frac{g_i(\alpha^\mu)^2}{\chi_i^2(\alpha^\mu)},$$

where the second inequality follows from Lemma 2 and Jensen's inequality. Rearranging and taking square roots yields (1).

4 Tightness of the Bound

We now show that the bound derived in Theorem 1 is tight up to constant factors in the low-discounting/low-monitoring double limit. We establish this result for public, product structure monitoring. We then discuss extensions to more general monitoring structures.

The bound in Theorem 1 applies for any Nash equilibrium. For a converse result establishing the possibility of cooperation (a kind of "folk theorem"), a more restrictive solution concept is preferable. Since our folk theorem assumes public monitoring, we use the standard solution concept for such games: *perfect public equilibrium (PPE)*, which is a strategy profile that forms a Nash equilibrium conditional on any public history $h^t = (y_{t'})_{t'=1}^{t-1}$. For any stage game G, public monitoring structure (Y, p), and discount factor δ, we denote the set of PPE payoff vectors in the repeated game $\Gamma = (G, Y, p, \delta)$ by $E(\Gamma)$.

For any stage game G, let $F = \text{co}\left(\{u(a)\}_{a \in A}\right) \subseteq \mathbb{R}^N$ denote the set of feasible payoffs, and let $F^* \subseteq F$ denote the set of feasible payoffs that weakly Pareto-dominate a convex combination of static Nash payoffs: that is, $v \in F^*$ if $v \in F$ and there exist a collection of static Nash equilibria (α_n) and non-negative weights (β_n) such that $v \geq \sum_n \beta_n u(\alpha_n)$ and $\sum_n \beta_n = 1$. Our folk theorem gives conditions under which $E(\Gamma)$ covers almost all of F^*, excepting points very close to the boundary. We make the standard assumption that the "target" payoff set (in our case F^*) is full dimensional: $\dim F^* = N$.

Theorem 2 *For any stage game G satisfying* $\dim F^* = N$ *and any $\varepsilon > 0$, there exists $k > 0$ such that the following holds:*

For any public, product monitoring structure (Y, p) and any discount factor δ satisfying

$$\frac{\min_{i, s_i, a : s_i(a_i)[a_i]=0} \chi_i^2(s_i, a)}{1-\delta} > k \qquad \textit{and} \tag{9}$$

$$\frac{p(y|a)}{1-\delta} > k \qquad \textit{for all } y, a, \tag{10}$$

and for any $v \in \text{int}F^*$ such that the Euclidean distance between v and the boundary of F^* is greater than ε, we have $v \in E(G, Y, p, \delta)$.

We compare Corollary 1 and Theorem 2. Heuristically, Corollary 1 says that for any repeated game, cooperation is impossible if detectability is much smaller than discounting;

and Theorem 2 says that for any repeated game with public, product structure monitoring that satisfies a full-dimensionality condition on payoffs, cooperation is possible if detectability is much larger than discounting. Note that the constant k differs between the two results: k must be taken to be sufficiently small in Corollary 1 and sufficiently large in Theorem 2.[8]

Two further points bear emphasis. First, the premise of Corollary 1 is that detectability is small for any recommended action profile a and any manipulation s_i, while the premise of Theorem 2 is that detectability is large for any recommended action profile a and any manipulation s_i that always disobeys the recommendation a_i. (Compare equations (2) and (9).) Without the latter requirement, Theorem 2 would be vacuous, because detectability is always small for a manipulation that rarely disobeys the recommendation. Second, Theorem 2 additionally requires that no signal is exceedingly rare relative to the discount rate (equation (10)). We explain the role of this requirement shortly.

Theorem 2 can also be compared to the classical folk theorems of FLM and KM. The key difference with these results is that Theorem 2 lets the discount factor and the monitoring structure vary simultaneously, while standard folk theorems fix the monitoring structure and show that cooperation is possible when the discount factor is high enough. Formally, the difference is that in Theorem 2 the constant k is uniform over monitoring structures (Y, p) satisfying equations (9) and (10), while standard folk theorems show only that for each monitoring structure (Y, p) that satisfies certain identifiability conditions, there exists a sufficiently high k (or equivalently, a sufficiently high δ) that supports cooperation.[9]

Theorem 2 is also related to SS's folk theorem for repeated games with frequent actions (their Theorem 2). SS consider a model where signals are parameterized by an underlying continuous-time Lévy process (a sum of Brownian and Poisson signals), and players interact every Δ units of time, with real-time discount rate r (so $\delta = e^{-r\Delta}$, and hence $1-\delta \approx r\Delta$).

[8] If the hypothesis that $p(y|a)/(1-\delta) > k$ in Theorem 2 is strengthened to a uniform lower bound on $p(y|a)$ (independent of δ), then χ^2-divergence can be replaced with various other divergences in the statement of Theorem 2, as these divergences are all equivalent when probabilities are bounded away from zero. For example, if $p(y|a) \geq \varepsilon$ for all y, a, then the total variation distance TV satisfies $\chi^2 \geq 4TV^2 \geq \varepsilon\chi^2$, and the Kullback-Leibler divergence KL satisfies $\chi^2 \geq 2\varepsilon KL \geq \varepsilon^2\chi^2$. For inequalities implying these bounds and many more, see e.g. Sason and Verdú (2016).

[9] FLM and KM do not assume product structure monitoring, FLM's folk theorem (e.g., their Theorem 6.1) imposes identifiability conditions only at certain action profiles, and KM's folk theorem (their Theorem 1) is a minmax threat folk theorem. (FLM also proved a minmax folk theorem under additional assumptions.) Our theorem can be extended in these regards, as we discuss below.

SS prove a folk theorem for the double limit where $\Delta \to 0$ and $r \to 0$. To compare with our results, observe that for Brownian signals (with the space of signal realizations partitioned into arbitrary fixed intervals) we have

$$\frac{\left(p\left(y|a\right)-p\left(y|a_i',a_{-i}\right)\right)^2}{p\left(y|a\right)} \approx \frac{\Delta}{1} = \Delta \qquad \text{and} \qquad p\left(y|a\right) \approx 1,$$

and for Poisson signals we have

$$\frac{\left(p\left(y|a\right)-p\left(y|a_i',a_{-i}\right)\right)^2}{p\left(y|a\right)} \approx \frac{\Delta^2}{\Delta} = \Delta \qquad \text{and} \qquad p\left(y|a\right) \approx \Delta.$$

Hence, under SS's information structure equations (9) and (10) both reduce to

$$r < \frac{1}{k}.$$

Theorem 2 therefore implies SS's result, up to some minor differences.[10] Relative to their result, our main contribution is dispensing with their parameterization by an underlying Lévy process. That is, we prove a general folk theorem for discrete-time repeated games in the low-discounting/low-monitoring double limit, which implies the folk theorem for repeated games with frequent actions (which assumes an underlying continuous-time parameterization) as a special case. Another significant difference is that SS assume that $N = 2$: this assumption seems important for their proof, which relies on parameterizing the boundary of the equilibrium payoff set as a 1-dimensional curve.[11]

We preview the key ideas of the proof of Theorem 2. We prove Theorem 2 as a corollary of the more general Theorem 3, which we state in the appendix. Theorem 3 is more general than Theorem 2 because it requires only a version of KM's "pairwise identifiability" condition, rather than product structure monitoring.[12] The proof of Theorem 3 builds on

[10] SS prove a pure-strategy minmax-threat folk theorem (rather than a Nash-threat folk theorem) and do not assume product structure monitoring; however, they assume bounded likelihood ratios $p\left(y|a_i',a_{-i}\right)/p\left(y|a\right)$. Our Theorem 3 likewise does not assume product structure monitoring, and under bounded likelihood ratios it can be adapted to give a pure-strategy minmax-threat folk theorem.

[11] We also mention Fudenberg and Levine (2007), who establish (in)efficiency results in a frequent-action game with one patient player and a myopic opponent, in contrast to SS's model with two patient players, or our model with N patient players.

[12] However, Theorem 3 is not immediately comparable to Corollary 1, which is why we relegate it to the

FLM, KM, and SS. Similarly to FLM and KM, the goal is to show that for any $v \in \text{int}F^*$, a sufficiently small ball B around v is self-generating (cf. Definition 1 in Appendix A). In the $\delta \to 1$ limit considered by FLM and KM, this follows because payoff vectors in B can be enforced with continuation payoff movements of magnitude $O(1-\delta)$, so since the set B is smooth, taking continuation payoffs to lie on translated tangent hyperplanes in B results in a vanishing efficiency loss. In contrast, when discounting and monitoring vary together, equation (9) implies that payoff vectors in B can be enforced with continuation payoffs of variance $o(1-\delta)$, while equation (10) additionally implies that continuation payoff movements can be taken to have magnitude $o(1)$ (but not necessarily $O(1-\delta)$): this follows because the solution to the variance-minimization program in Remark 1 is

$$\sum_y p(y|a)\left(w_i(y) - \mathbb{E}\left[w_i(\tilde{y})\right]\right)^2 = \left(u_i\left(s_i(a_i), a_{-i}\right) - u_i(a)\right)^2 \times \left(\frac{1-\delta}{\delta}\right)^2 \times \frac{1}{\chi_i^2(s_i, a_i)},$$

which is $o(1-\delta)$ when $\delta \to 1$ faster than $\chi_i^2(s_i, a_i) \to 0$, and which implies that $w_i(y) - \mathbb{E}[w_i(\tilde{y})] = o(1)$ for all y when $p(y|a) > 1-\delta$ and $\delta \to 1$ faster than $\chi_i^2(s_i, a_i) \to 0$. A key lemma (Lemma 6) shows that under these conditions, requiring continuation payoffs to lie in B again results in a vanishing efficiency loss. The intuition is that larger continuation payoff movements must land farther in the interior of B, resulting in greater inefficiency; but since the continuation payoff variance is small, these large movements are infrequent enough that the ex ante expected inefficiency is small.

Comparing equations (2) and (9) shows that the tradeoff between dectectability and discounting expressed in Theorem 1 is tight up to constant terms (i.e., the difference in the constant k in the two equations) in the low-discounting/low-monitoring double limit. One may wonder how much slack the constant factors are hiding. This is a tricky question to answer in general, because the calculations involved in proving Theorem 2 are somewhat intricate. However, we can give a clear answer for a frequent-action version of the prisoner's

appendix. While Theorem 3 is more general than Theorem 2, it still assumes pairwise identifiability rather than "individual identifiability." (See Appendix A for the definitions of these terms.) We conjecture that Theorem 3 remains valid under individual identifiability for public monitoring or for private monitoring in the presence of a mediator, but proving either of these results would involve complications similar to those in the literature on the folk theorem with private monitoring (e.g., Sugaya, 2022). These complications are orthogonal to the current paper's focus on monitoring precision, and would necessitate a much longer proof. This conjecture is thus left for future research.

dilemma considered in Example 1.

Example 2 (Prisoner's Dilemma Redux) *Consider again Example 1, now parameterized by a triple (Δ, r, π), where $\Delta > 0$, $r > 0$, and $\pi \in (1/2, 1)$. The payoff matrix is as in Example 1. The discount factor is $\delta = e^{-r\Delta}$. Monitoring is public with a symmetric product structure, where $Y^i = \{-1, 1\}$ and*

$$p^i\left(1|a^i\right) = \begin{cases} \frac{1}{2} + \left(\pi - \frac{1}{2}\right)\sqrt{\Delta/\gamma} & \text{if } a^i = C, \\ \frac{1}{2} - \left(\pi - \frac{1}{2}\right)\sqrt{\Delta/\gamma} & \text{if } a^i = D, \end{cases}$$

where $\gamma := 4\pi(1-\pi)$. We let the players access a public randomization device.

With this parameterization, as $\Delta \to 0$ the process

$$X_\tau^i = \frac{1}{\sqrt{\Delta}} \sum_{t=1}^{\lfloor \tau/\Delta \rfloor} y_t^i$$

converges in distribution to a Brownian motion with drift $2\pi - 1$ (resp., $-(2\pi-1)$) when $a_\tau^i = C$ (resp., D) and variance γ.[13] *Thus, for small Δ the game is almost the same as the continuous-time partnership game studied by Sannikov (2007).*

For any Δ, the detectability χ^2 of a manipulation that always disobeys the recommendation equals

$$\frac{(2\pi-1)^2\Delta}{\pi(1-\pi) - \left(\pi - \frac{1}{2}\right)^2\Delta}.$$

Thus, for sufficiently small Δ,

$$\frac{\chi^2}{1-\delta} = \frac{1}{e^{-r\Delta}} \frac{(2\pi-1)^2\Delta}{\pi(1-\pi) - \left(\pi - \frac{1}{2}\right)^2\Delta} \approx \frac{1}{r}\frac{(2\pi-1)^2}{\pi(1-\pi)}.$$

For any sufficiently small Δ and any $\varepsilon \leq 1/2$, inequality (4) derived in Section 3 now implies that, for the players to attain an average equilibrium payoff of $w = 1 - \varepsilon$, we must have

$$\frac{\chi^2}{1-\delta} \geq \frac{1}{5\varepsilon}. \tag{11}$$

[13] This follows from the functional central limit theorem (e.g., Billingsley, 1995, Theorem 37.8).

At the same time, by adapting the proof of Theorem 2, we can establish the following result.[14]

***Proposition 1** For any sufficiently small Δ and any $\varepsilon \leq \frac{54}{23} - \frac{6}{23}\sqrt{58} \approx .361$, if*

$$\frac{\chi^2}{1-\delta} \geq \frac{1}{\left(\frac{18\sqrt{10}}{125} - \frac{4}{25}\right)\varepsilon^2} \approx \frac{1}{0.295\varepsilon^2}, \tag{12}$$

then there exists a PPE with payoff vector $(1-\varepsilon, 1-\varepsilon)$.

The constants on the right-hand sides of inequalities (11) and (12) correspond to the required values of k in Corollary 1 and Theorem 2, respectively, so the ratio of these constants quantifies the constant-factor slack between our necessary and sufficient conditions for cooperation. Observe that there are two sources of slack: an absolute constant factor of $5/0.295$, and a factor of ε (the distance between the target payoff vector $(1-\varepsilon, 1-\varepsilon)$ and the efficient payoff vector $(1,1)$). Some absolute constant-factor slack is to be expected, since (11) comes from a general theorem (Theorem 1) that does not make use of the structure of the prisoner's dilemma stage game, the symmetric product structure monitoring, or the perfection requirement imposed by the PPE solution concept. On the other hand, the ε-factor slack arises because Theorem 1 tightly characterizes the relationship between discounting and monitoring, but not the relationship between these variables and the distance to the boundary of the feasible payoff set. We derive tighter results on the rate of convergence to the boundary of the feasible payoff set in a companion paper (Sugaya and Wolitzky, 2022b).

5 Discussion

This paper has established general results on the tradeoff between discounting and monitoring for supporting cooperation in repeated games. We conclude by discussing some applications.

We have already mentioned implications of our results for repeated games with frequent actions, where the interaction frequency $1/\Delta$ goes to infinity while the real-time discount rate r is fixed, and signals are parameterized by an underlying continuous-time stochastic

[14] A similar result can be obtained by applying Sannikov's characterization of the set of PPE payoffs in the continuous-time limit game (his Theorem 2). However, it is not straightforward to show that this set approximates the PPE payoff set in the discrete-time game.

process. This "frequent action limit" is a particular type of low-discounting/low-monitoring double limit where discounting and monitoring vanish at the same rate, which corresponds to the edge case in between our necessary and sufficient conditions for cooperation. This edge case is interesting and important, but it is also perhaps somewhat special and detail-dependent.[15] When specialized to frequent-action games where both $\Delta \to 0$ and $r \to 0$, our folk theorem generalizes that of Sannikov and Skrzypacz (2010) to games with more than two players.

Another type of low-discounting/low-monitoring double limit arises in large-population repeated games, where many patient players are monitored by a noisy aggregate signal, which provides little information about each individual player's action. This type of model was studied by Green (1980) and Sabourian (1990) under a continuity condition on the mapping from action profiles to signals, and by Fudenberg, Levine, and Pesendorfer (1998) and al-Najjar and Smorodinsky (2000, 2001) under the assumption that each player's action is hit by independent, individual-level noise. In a companion paper (Sugaya and Wolitzky, 2022a), we derive necessary and sufficient conditions for cooperation in large-population repeated games with individual-level noise, as a function of the population size, the discount factor, and the *channel capacity* (the maximum expected entropy reduction) of the monitoring structure. These results extend those in the current paper by introducing individual-level noise and letting the stage game—and in particular the number of players—vary together with the discount factor and the monitoring structure.

Our negative result (Theorem 1) can be extended to show that for any fixed imperfect monitoring structure, the Nash equilibrium payoff set cannot converge to the boundary of the feasible payoff set at a rate faster than $(1-\delta)^{1/2+\varepsilon}$ for any $\varepsilon > 0$. Since the rate of convergence of the PPE payoff set with imperfect public monitoring is known to be $(1-\delta)^{1/2}$, this result shows that allowing private strategies and monitoring cannot significantly increase the rate of convergence, which resolves in the negative a question posed by Hörner and Takahashi (2016). Moreover, by accounting for monitoring precision as well as discounting,

[15] Sadzik and Stacchetti (2015) study the frequent action limit of repeated principal-agent problems with one-dimensional actions and concave preferences, where the signal process converges to a Brownian motion. Our results cover repeated principal-agent problems, and complement Sadzik and Stacchetti's by providing necessary and sufficient conditions for cooperation for more general games and monitoring structures.

this bound can be refined to show that the distance between the equilibrium payoff set and the boundary of the feasible payoff set must exceed $\left((1-\delta)/\max_{i,s_i,a}\chi_i^2(s_i,a)\right)^{1/2+\varepsilon}$. This is another result where the relevant timescale is the intrinsic time experienced by a martingale with likelihood ratio difference increments. We present these results in a second companion paper (Sugaya and Wolitzky, 2022b).

Appendix

A Proof of Theorem 2

We prove Theorem 2 as a corollary of a more general result, Theorem 3. In this appendix, we assume either product structure monitoring or the genericity condition that each player has a strict incentive to follow an action profile that maximizes her own payoff.

Assumption 1 Monitoring is public, and one of the following holds:

1. Monitoring has a product structure.
2. For each player i, there exists an action profile $a^i \in \operatorname{argmax}_{a\in A} u_i(a)$ satisfying $u_i(a^i) > u_i\left(a_i, a_{-i}^i\right)$ for all $a_i \neq a_i^i$.

We introduce some notation. For each i and a, let

$$P_i(a) = \bigcup_{a_i' \neq a_i} \left(\frac{p(y|a_i', a_{-i})}{p(y|a)}\right)_{y\in\bar{Y}}.$$

That is, $P_i(a)$ is the set of vectors of likelihood ratios that can arise when player i deviates from a_i while the remaining players take a_{-i}. Also, for each a, define the inner product $\langle\cdot,\cdot\rangle_a$ on $\mathbb{R}^{|\bar{Y}|}$ by

$$\langle\tilde{p},\tilde{q}\rangle_a = \sum_y p(y|a)\,\tilde{p}(y)\,\tilde{q}(y) \qquad \text{for all } \tilde{p},\tilde{q} \in \mathbb{R}^{|\bar{Y}|},$$

and let $\|\cdot\|_a$ denote the associated norm. Let $\mathbf{1}$ denote the vector of 1's in $\mathbb{R}^{|\bar{Y}|}$.

For any $\eta \in [0,1]$, we say that monitoring satisfies *η-individual identifiability* if for any action profile a and player i, the following two conditions hold:

1. There exists a vector $z \in \mathbb{R}^{|\bar{Y}|}$ such that $\|z\|_a = 1$ and

$$\langle z, \mathbf{1} \rangle_a > \langle z, p \rangle_a + \eta \qquad \text{for all } a \text{ and } p \in P_i(a). \tag{13}$$

2. We have

$$p(y|a) > \eta^2 \qquad \text{for all } y \in \bar{Y}. \tag{14}$$

Intuitively, (13) says that the vector of likelihood ratios under equilibrium play is sufficiently different from the corresponding vector under any deviation by player i. When $\eta = 0$, η-individual identifiability is weaker than FLM's individual full rank condition. If η-individual identifiability holds under product structure monitoring, then for any action profile a and player i, there exists a unit vector $z \in \mathbb{R}^{|\bar{Y}|}$ that satisfies (13) and (14) as well as

$$z\left(y^i, y^{-i}\right) = z\left(y^i, \tilde{y}^{-i}\right) \qquad \text{for all } y^i, y^{-i}, \tilde{y}^{-i}. \tag{15}$$

For any $\eta \in [0,1]$, we say that monitoring satisfies *η-pairwise identifiability* if for any action profile a and any pair of distinct players i and j, the following three conditions hold:

1. There exists a vector $z \in \mathbb{R}^{|\bar{Y}|}$ such that $\|z\|_a = 1$ and

$$\langle z, \mathbf{1} \rangle_a > \langle z, p \rangle_a + \eta \qquad \text{for all } p \in P_i(a) \cup P_j(a). \tag{16}$$

2. There exists a vector $z \in \mathbb{R}^{|\bar{Y}|}$ such that $\|z\|_a = 1$ and

$$\left\langle z, p^i \right\rangle_a - \eta > \langle z, \mathbf{1} \rangle_a > \left\langle z, p^j \right\rangle_a + \eta \qquad \text{for all } p^i \in P_i(a) \text{ and } p^j \in P_j(a). \tag{17}$$

3. (14) holds.

Intuitively, (16) says that the vector of likelihood ratios under equilibrium play is sufficiently different from the corresponding vector under any deviation by player i or j, and (17) says that the vector of likelihood ratios under any deviation by i is sufficiently different from the vector under any deviation by j. When $\eta = 0$, η-individual identifiability coincides with KM's assumptions (A2) and (A3), which are weaker than FLM's pairwise full rank condition.[16] In general, η-pairwise identifiability says that KM's assumptions hold with η slack.

η-pairwise identifiability implies η-individual identifiability. Conversely, under product structure monitoring, η-individual identifiability implies $(\eta/2)$-pairwise identifiability (Lemma 4).

Theorem 2 follows easily from the following result, which is a generalization of the folk theorems of FLM, KM, and SS.[17]

Theorem 3 *Assume that* $\dim F^* = N$ *and Assumption 1 holds. For any* $v \in \text{int} F^*$*, there exists* $c > 0$ *such that, for any* $\eta > 0$*, any monitoring structure* (Y, p) *that satisfies* η*-pairwise identifiability, and any* $\delta > 1 - c\eta^2$*, we have* $v \in E(G, Y, p, \delta)$*.*

To prove Theorem 2 from Theorem 3, we use two simple lemmas.

Lemma 3 *For any action profile* a *and player* i*, the following are equivalent:*

1. *There exists a vector* $z \in \mathbb{R}^{|\bar{Y}|}$ *satisfying* $\|z\|_a = 1$ *and (13).*
2. $\min_{s_i : s_i(a_i)[a_i]=0} \chi_i^2 (s_i, a) > \eta^2$*.*

Proof. By the separating hyperplane theorem, the former condition is equivalent to $\min_{\tilde{p} \in \text{co}(P_i(a))} \|\mathbf{1} - \tilde{p}\|_a > \eta$, or equivalently

$$\min_{\tilde{p} \in \text{co}(P_i(a))} \sqrt{\sum_y p(y|a) \left(\frac{p(y|a) - \tilde{p}(y)}{p(y|a)} \right)^2} > \eta.$$

The result follows because

$$\min_{s_i : s_i(a_i)[a_i]=0} \chi_i^2 (s_i, a) = \min_{\tilde{p} \in \text{co}(P_i(a))} \sum_y p(y|a) \left(\frac{p(y|a) - \tilde{p}(y)}{p(y|a)} \right)^2.$$

■

[16] When $\eta = 0$, (16) coincides with KM's assumption (A2), and (17) coincides with KM's assumption (A3).

[17] To prove Theorem 2, it suffices to prove Theorem 3 under Assumption 1.1. However, Assumption 1.1 is unduly restrictive and the proof of Theorem 3 under Assumption 1.2 is almost identical, so we include it for completeness.

Remark 3 *When $\chi_i^2(s_i, a) > \eta^2$ for all s_i such that $s_i(a_i)[a_i] = 0$, the vector*

$$z = \frac{\mathbf{1} - \hat{p}}{\sqrt{\min_{s_i : s_i(a_i)[a_i]=0} \chi_i^2(s_i, a)}},$$

where $\hat{p} \in \arg\min_{\tilde{p} \in \mathrm{co}(P_i(a))} \|\mathbf{1} - \tilde{p}\|_a$, satisfies $\|z\|_a = 1$ and (13). Note that this vector is co-linear with the vector of likelihood ratio differences $((p(y|a) - \hat{p}(y))/p(y|a))_{y \in \bar{Y}}$.

Lemma 4 *If a product monitoring structure satisfies η-individual identifiability, then it satisfies $(\eta/2)$-pairwise identifiability.*

Proof. Under product structure monitoring and η-individual identifiability, for any action profile a and any pair of distinct players i and j, there exist unit vectors z^i and z^j that satisfy (13) and (15). Define $z := (z^i + z^j) / \|z^i + z_j\|_a$. For any $p^j \in P_j(a)$, let $a_j' \in A^j$ satisfy $p^j = \left(p\left(y|a_j', a_{-j}\right)/p(y|a)\right)_{y \in \bar{Y}}$. Note that, for any $\bar{y}^{-i}$,

$$\begin{aligned}
\left\langle z, p(a) - p^j \right\rangle_a &= \frac{1}{\|z^i + z^j\|_a}\left(\sum_y z^i(y)\left(p(y|a) - p\left(y|a_j', a_{-j}\right)\right) + \left\langle z^j, p(a) - p^j\right\rangle_a\right) \\
&= \frac{1}{\|z^i + z^j\|_a}\left(\sum_{y^i} z^i\left(y^i, \bar{y}^{-i}\right)\sum_{y^{-i}}\left(p\left(y^i, y^{-i}|a\right) - p\left(y^i, y^{-i}|a_j', a_{-j}\right)\right) + \left\langle z^j, p(a) - p^j\right\rangle_a\right) \\
&= \frac{1}{\|z^i + z^j\|_a}\left\langle z^j, p(a) - p^j\right\rangle_a.
\end{aligned}$$

where the second line follows by (15), and the third line follows because $\sum_{y^{-i}} p\left(y^i, y^{-i}|a\right) = \sum_{y^{-i}} p\left(y^i, y^{-i}|a_j', a_{-j}\right)$ under product structure monitoring. Thus, since $\|z^i + z^j\|_a \leq 2$, the vector z satisfies (16) with $\eta/2$ in place of η. Similarly, the vector $(-z^i + z_j)/\|-z^i + z^j\|_a$ satisfies (17) with $\eta/2$ in place of η. ■

Proof of Theorem 2. Fix c such that the conclusion of Theorem 3 holds, fix any $\hat{c} < c$, and let $k = 4/\hat{c}$. By Lemma 3, (9) and (10) imply $2\sqrt{(1-\delta)/\hat{c}}$-individual identifiability. Since monitoring has a product structure, Lemma 4 now implies that η-pairwise identifiability holds for $\eta = \sqrt{(1-\delta)/\hat{c}}$, and $\delta = 1 - \hat{c}\eta^2 > 1 - c\eta^2$. Then $v \in E(G, Y, p, \delta)$ by Theorem 3.

To complete the proof, it remains to show that the constant c in the statement of Theorem 3 can be chosen uniformly for all payoff vectors v at distance at least ε from the boundary of F^*. This last claim follows immediately from the proof of Theorem 3, where the constant

c is explicitly constructed as a function of the distance between v and the boundary of F^*.

■

A.1 Proof of Theorem 3

Fix $v \in \operatorname{int} F^*$. If Assumption 1.1 holds, let $\varepsilon_u = \infty$ and fix $a^i \in \operatorname{argmax}_{a \in A} u_i(a)$ arbitrarily. If Assumption 1.2 holds, let $\varepsilon_u > 0$ be such that, for each player i, there exists an action profile $a^i \in \operatorname{argmax}_{a \in A} u_i(a)$ satisfying $u_i(a^i) \geq u_i\left(a_i, a^i_{-i}\right) + \varepsilon_u$ for all $a_i \neq a^i_i$. (If both assumptions hold, either definition works.)

Let $\varepsilon_v > 0$ denote the Euclidean distance between v and the boundary of F^*, let $\varepsilon = \min\{\varepsilon_u, \varepsilon_v, \bar{u}/4\} \in (0, \bar{u})$, and let $\hat{u} = \bar{u} + \varepsilon$. Let $B = \{v' : d(v, v') \leq \varepsilon/2\}$, the closed ball of radius $\varepsilon/2$ centered at v. We will find $c > 0$ such that if (Y, p) satisfies η-pairwise identifiability and $\delta > 1 - c\eta^2$, then $B \subseteq E(G, Y, p, \delta)$, and hence $v \in E(G, Y, p, \delta)$.

The following definition and lemma are due to Abreu, Pearce, and Stacchetti (1990).

Definition 1 *A bounded set $W \subseteq \mathbb{R}^N$ is* self-generating *if for all $\hat{v} \in W$, there exist $\alpha \in \Delta^*(A)$ and $w : \bar{Y} \to \mathbb{R}^N$ satisfying*

1. Promise keeping (PK): $\hat{v} = (1-\delta) u(\alpha) + \delta \sum_y p(y|\alpha) w(y)$.

2. Incentive compatibility (IC): $\operatorname{supp}(\alpha_i) \subseteq \operatorname{argmax}_{a_i} (1-\delta) u_i(a_i, \alpha_{-i}) + \delta \sum_y p(y|a_i, \alpha_{-i}) w_i(y)$ *for all i.*

3. Self-generation (SG): $w(y) \in W$ *for all y.*

When (PK), (IC), and (SG) hold, we say that the pair (α, w) decomposes *$\hat{v}$* on *W.*

Lemma 5 *Any bounded, self-generating set W is contained in $E(\Gamma)$.*

Our key lemma (Lemma 6) will provide a sufficient condition for B to be self-generating, and hence contained in $E(\Gamma)$. It is based on the following definition, where $\|\cdot\|$ denotes the Euclidean norm in $\mathbb{R}^N$, $\Lambda = \{\lambda \in \mathbb{R}^N : \|\lambda\| = 1\}$, and for each $\lambda \in \Lambda$, $\|\lambda_+\| = \sqrt{\sum_{\lambda_n > 0} (\lambda_n)^2}$ and $\|\lambda_-\| = \sqrt{\sum_{\lambda_n < 0} (\lambda_n)^2}$.

Definition 2 *The* maximum score in direction $\lambda \in \Lambda$ with reward bound $X > 0$ *is defined as*

$$k(\lambda, X) := \sup_{\alpha \in \Delta^*(A), x: \bar{Y} \to \mathbb{R}^N} \lambda \cdot \left(u(\alpha) + \sum_y p(y|\alpha) x(y) \right)$$

subject to

1. Incentive compatibility with ε slack (ICε): *For each player i, either (i) $u_i(\alpha) \geq u_i(a_i, \alpha_{-i})$ for all a_i and $\sum_{y:x(y)=x} p(y|a) = \sum_{y:x(y)=x} p(y|a_i', a_{-i})$ for all $x \in \mathbb{R}$, a, and a_i', or (ii) for all $a_i \notin \operatorname{supp}(\alpha_i)$,*

$$u_i(\alpha) + \sum_y p(y|\alpha) x_i(y) \geq u_i(a_i, \alpha_{-i}) + \sum_y p(y|a_i, \alpha_{-i}) x_i(y) + \varepsilon \mathbf{1}\{\lambda_i \geq 0\}.$$

2. Half-space decomposability with reward bound X (HSX):

$$\lambda \cdot x(y) \leq 0 \text{ and } \frac{\left\| x(y) - \sum_{y'} p(y'|\alpha) x(y') \right\|}{\|\lambda_+\|} \leq X\bar{u} \quad \text{for all } y, \quad \text{and}$$
$$\sum_y p(y|\alpha) \frac{\left\| x(y) - \sum_{y'} p(y'|\alpha) x(y') \right\|^2}{\|\lambda_+\|^2} \leq X\bar{u}^2.$$

The following is our key lemma:[18]

Lemma 6 *If there exists $X > 0$ such that*

$$k(\lambda, X) \geq \max_{v' \in B} \lambda \cdot v' + \frac{\varepsilon}{4} \quad \text{for all } \lambda \in \Lambda, \quad \text{and} \tag{18}$$
$$\max\{X, N\} \leq \frac{\delta}{1-\delta} \frac{\varepsilon^2}{2^{12}\bar{u}^2}, \tag{19}$$

then B is self-generating.

Proof. See Appendix A.2. ■

[18] If $\varepsilon = 0$ and $X = \infty$ then $k(\lambda, X)$ equals $k^*(\lambda)$, the maximum score in direction λ as defined by Fudenberg and Levine (1994). Fudenberg and Levine showed that B is self-generating for all sufficiently high δ if $k^*(\lambda) \geq \max_{v' \in B} \lambda \cdot v'$ for all λ. Lemma 6 extends their result to show that B is self-generating for a given value of δ if $k(\lambda, X) \geq \max_{v' \in B} \lambda \cdot v' + \varepsilon/4$ for all λ, where the magnitude and the variance of the normalized rewards $x(y)$ are bounded by a constant multiple of $(1-\delta)^{-1}$.

To complete the proof, we find $c > 0$ such that if (Y, p) satisfies η-pairwise identifiability and $\delta > 1 - c\eta^2$, then there exists $X > 0$ that satisfies (18) and (19). To define c and X, we first introduce one more constant, denoted $\underline{\lambda} \in (0, 1)$, which we will use to partition the set of directions $\lambda \in \Lambda$ in a manner similar to FLM and KM.

For any $\lambda \in \Lambda$, let $i(\lambda) \in \operatorname{argmax}_{n \in I} \lambda_n$ denote a player with the highest Pareto weight under λ (choosing arbitrarily in case of a tie); let $m(\lambda) = \lambda_{i(\lambda)} = \max_n \lambda_n$ denote the corresponding Pareto weight; and let $M(\lambda) = \max_{n \neq i} |\lambda_n|$ denote the highest Pareto weight in absolute value terms of any player other than $i(\lambda)$.

Lemma 7 *Let $\underline{\lambda} > 0$ satisfy*

$$N\bar{u} \max \left\{ \underline{\lambda}, \frac{1 - \sqrt{1 - N\underline{\lambda}^2}}{\sqrt{1 - N\underline{\lambda}^2}} \right\} \leq \frac{\varepsilon}{4}. \tag{20}$$

1. *For all $\lambda \in \Lambda$, if $m(\lambda) \geq M(\lambda)/\underline{\lambda}$, then $\lambda \cdot u\left(a^{i(\lambda)}\right) \geq \max_{v' \in B} \lambda \cdot v' + \varepsilon/4$.*

2. *For all $\lambda \in \Lambda$, if $m(\lambda) \leq \underline{\lambda}$, then there exists a static Nash equilibrium α^{NE} such that $\lambda \cdot u\left(\alpha^{NE}\right) \geq \max_{v' \in B} \lambda \cdot v' + \varepsilon/4$.*

Proof. See Appendix A.3 ■

Now we fix the constants

$$\bar{X} = \frac{4N^2\hat{u}^2}{\underline{\lambda}^4\bar{u}^2} \quad \text{and} \quad c = \frac{\underline{\lambda}^4\varepsilon^2}{2^{13}N^2\hat{u}^2}. \tag{21}$$

Lemma 8 *If $\eta < 1$ and $\delta > 1 - c\eta^2$, then*

$$\max\left\{\frac{\bar{X}}{\eta^2}, N\right\} < \frac{\delta}{1-\delta}\frac{\varepsilon^2}{2^{12}\bar{u}^2}.$$

Proof. Note that $c < 1/2$ and hence $\delta > 1/2$, as $\varepsilon < \bar{u}$, $\underline{\lambda} < 1$, and $N \geq 1$. Hence, we have $\delta > 1 - c\eta^2 > 1 - c > 1 - \frac{\underline{\lambda}^4\varepsilon^2}{2^{13}N^2\hat{u}^2}$, and so $\frac{\bar{X}}{\eta^2} < \frac{\bar{X}c}{1-\delta} = \frac{\varepsilon^2}{(1-\delta)2^{13}\bar{u}^2} < \frac{\delta\varepsilon^2}{(1-\delta)2^{12}\bar{u}^2}$ and $N < \frac{\varepsilon^2}{(1-\delta)2^{13}\bar{u}^2} < \frac{\delta\varepsilon^2}{(1-\delta)2^{12}\bar{u}^2}$. ■

We henceforth assume that (Y, p) satisfies η-pairwise identifiability and $\delta > 1 - c\eta^2$. By Lemmas 6 and 8, to complete the proof it suffices to show that $k\left(\lambda, \bar{X}/\eta^2\right) \geq \max_{v' \in B} \lambda \cdot v' + \varepsilon/4$ for all $\lambda \in \Lambda$.

We first observe that for each pair of players $i \neq j$ and each action profile a, we can define rewards $\left(x_i^{j,-}(y;a)\right)_{y\in\bar{Y}}$ and $\left(x_i^{j,+}(y;a)\right)_{y\in\bar{Y}}$ with mean 0 and variance at most $\hat{u}^2/\eta^2$ that induce player i to take a_i when her opponents take a_{-i}; and that have the property that for player j, taking a_j maximizes the expectation of $x_i^{j,-}(y;a)$ and minimizes the expectation of $x_i^{j,+}(y;a)$, for each y. This is a direct implication of η-pairwise identifiability.

Lemma 9 *For each pair of players $i \neq j$ and action profile $a \in A$, there exist $\left(x_i^{j,\zeta}(y;a)\right)_{y\in\bar{Y},\zeta\in\{-1,+1\}}$ such that, for each $\zeta \in \{-1,+1\}$, we have*

$$
\begin{aligned}
\sum_y p(y|a)\, x_i^{j,\zeta}(y;a) &= 0, && (22)\\
\sum_y p(y|a_i', a_{-i})\, x_i^{j,\zeta}(y;a) &\leq -\hat{u} \quad \text{for all } a_i' \neq a_i, && (23)\\
\zeta \times \sum_y p\left(y|a_j', a_{-j}\right) x_i^{j,\zeta}(y;a) &\geq 0 \quad \text{for all } a_j' \neq a_j, \quad \text{and} && (24)\\
\sum_y p(y|a)\, x_i^{j,\zeta}(y;a)^2 &\leq \frac{\hat{u}^2}{\eta^2}. && (25)
\end{aligned}
$$

Moreover, if Assumption 1.1 holds then $x_i^{j,\zeta}(\cdot;a)$ can be taken to depend only on y^i.

Proof. Fix any i, j, and a. We first construct $\left(x_i^{j,-1}(y;a)\right)_{y\in\bar{Y}}$. Let $z \in \mathbb{R}^{|\bar{Y}|}$ satisfy $\|z\|_a = 1$ and (16). By the definitions of $\langle\cdot,\cdot\rangle_a$, P_i, and P_j, we have

$$
\begin{aligned}
\sum_y \left(p(y|a) - p(y|a_i', a_{-i})\right) z(y) &\geq \eta \quad \text{for all } a_i' \neq a_i, \quad \text{and}\\
\sum_y \left(p(y|a) - p\left(y|a_j', a_{-j}\right)\right) z(y) &\geq \eta \quad \text{for all } a_j' \neq a_j.
\end{aligned}
$$

Defining

$$
x_i^{j,-1}(y;a) = \frac{\hat{u}}{\eta}\left(z(y) - \sum_{\tilde{y}} p(\tilde{y}|a)\, z(\tilde{y})\right) \quad \text{for all } y,
$$

conditions (22)–(24) hold by construction, and condition (25) holds because

$$\begin{aligned}\sum_y p(y|a)\left(x_i^{j,-1}(y;a)\right)^2 &= \frac{\hat{u}^2}{\eta^2}\sum_y p(y|a)\left(z(y)-\sum_{\tilde{y}} p(\tilde{y}|a)\, z(\tilde{y})\right)^2 \\ &\leq \frac{\hat{u}^2}{\eta^2}\sum_y p(y|a)\, z(y)^2 = \frac{\hat{u}^2}{\eta^2}.\end{aligned}$$

To construct $\left(x_i^{j,+1}(y;a)\right)_{y\in\bar{Y}}$, let $z \in \mathbb{R}^{|\bar{Y}|}$ satisfy $\|z\|_a = 1$ and (17), and proceed as in the construction of $\left(x_i^{j,-1}(y;a)\right)_{y\in\bar{Y}}$.

Under product structure monitoring, by (13) and (15), there exists x^i such that $\sum_y p(y|a)\, x^i(y;a) = 0$, $\sum_y p(y|a_i', a_{-i})\, x^i(y;a) \leq -\hat{u}$ for all $a_i' \neq a_i$, $\sum_y p(y|a)\, x^i(y;a)^2 \leq \hat{u}^2/\eta^2$, and $x^i(y^i, y^{-i}; a) = x^i(y^i, \tilde{y}^{-i}; a)$ for all $y^i, y^{-i}, \tilde{y}^{-i}, a$. Defining $x_i^{j,-1}(y;a) = x_i^{j,+1}(y;a) = x^i(y;a)$ also satisfies (24) for each ζ, since the distribution of y^i (and hence x^i) is independent of a_{-i}. ■

Finally, we show that $k\left(\lambda, \bar{X}/\eta^2\right) \geq \max_{v'\in B} \lambda\cdot v' + \varepsilon/4$ for all $\lambda \in \Lambda$. We partition Λ into three cases: (1) $\underline{\lambda} < m(\lambda) < M(\lambda)/\underline{\lambda}$; (2) $m(\lambda) \geq M(\lambda)/\underline{\lambda}$; and (3) $m(\lambda) \leq \underline{\lambda}$.

Case 1: $\underline{\lambda} < m(\lambda) < M(\lambda)/\underline{\lambda}$. Fix any $a^\lambda \in \operatorname{argmax}_a \lambda\cdot u(a)$. Fix $i = i(\lambda)$, and fix some $j \neq i$ such that $\lambda_i/|\lambda_j| < 1/\underline{\lambda}$. For each y, define $x(y) = (x_n(y))_{n\in I}$ by

$$x_n(y) = \begin{cases} x_i^{j,\operatorname{sign}(\lambda_j\lambda_i)}\left(y;a^\lambda\right) - \sum_{n'\neq i} \frac{\lambda_{n'}}{\lambda_i} x_{n'}^{i,\operatorname{sign}(\lambda_i\lambda_{n'})}\left(y;a^\lambda\right) & \text{if } n = i, \\ x_j^{i,\operatorname{sign}(\lambda_i\lambda_j)}\left(y;a^\lambda\right) - \frac{\lambda_i}{\lambda_j} x_i^{j,\operatorname{sign}(\lambda_j\lambda_i)}\left(y;a^\lambda\right) & \text{if } n = j, \\ x_n^{i,\operatorname{sign}(\lambda_i\lambda_n)}\left(y;a^\lambda\right) & \text{if } n \neq i,j, \end{cases}$$

where $\operatorname{sign}(\beta) = -1$ for $\beta \leq 0$ and $\operatorname{sign}(\beta) = +1$ for $\beta > 0$. Note that $\lambda\cdot x(y) = 0\ \forall y$, and hence $\lambda\cdot\left(u\left(a^\lambda\right) + \sum_y p\left(y|a^\lambda\right) x(y)\right) = \lambda\cdot u\left(a^\lambda\right) \geq \max_{v'\in B}\lambda\cdot v' + \varepsilon/4$. Moreover, (IC$\varepsilon$) holds because, for each player, (23) and (24) imply that the expected loss in continuation payoff from deviating is at least $\hat{u}$, which exceeds the maximum difference between any two stage game payoffs by at least ε. Finally, for (HS$X/\eta^2$), by $|\lambda_n|/\lambda_i \leq 1/\underline{\lambda}\ \forall n$, $\lambda_i/|\lambda_j| \leq 1/\underline{\lambda}$, (22), and (25), we have

$$\left\|x(y) - \sum_{\tilde{y}} p\left(\tilde{y}|a^\lambda\right) x(\tilde{y})\right\|^2 \leq \frac{\hat{u}^2}{\eta^2}\underbrace{\left(\left(1-\sum_{n\neq i}\lambda_n/\lambda_i\right)^2 + \left(1-\lambda_i/\lambda_j\right)^2 + N - 2\right)}_{\leq(1+(N-1)/\underline{\lambda})^2+(1+1/\underline{\lambda})^2+N-2} \leq \frac{4N^2\hat{u}^2}{\eta^2\underline{\lambda}^2}.$$

Hence, (HS$X/\eta^2$) holds, because we have

$$\frac{\left\| x(y) - \sum_{\tilde{y}} p\left(\tilde{y}|a^\lambda\right) x(\tilde{y}) \right\|}{\|\lambda_+\|} \leq \frac{2N\hat{u}}{\eta^2 \underline{\lambda}^2} = \frac{\sqrt{\bar{X}}}{\eta^2}\bar{u} \leq \frac{\bar{X}}{\eta^2}\bar{u}, \tag{26}$$

and, since $\|\lambda_+\|^2 \geq \lambda_i^2 \geq \underline{\lambda}^2$, we have

$$\sum_y \frac{p\left(y|a^\lambda\right) \left\| x(y) - \sum_{\tilde{y}} p\left(\tilde{y}|a^\lambda\right) x(\tilde{y}) \right\|^2}{\|\lambda_+\|^2} \leq \frac{4N^2\hat{u}^2}{\eta^2 \underline{\lambda}^4} = \frac{\bar{X}}{\eta^2}\bar{u}^2.$$

Case 2: $m(\lambda) \geq M(\lambda)/\underline{\lambda}$. Fix $i = i(\lambda)$ and let $\alpha = a^i$. For each y, define $x(y) = (x_n(y))_{n\in I}$ by

$$x_n(y) = \begin{cases} -\sum_{n'\neq i} \frac{\lambda_{n'}}{\lambda_i} x_{n'}^{i,\mathrm{sign}(\lambda_i\lambda_{n'})}(y;a^i) & \text{if } n = i, \\ x_n^{i,\mathrm{sign}(\lambda_i\lambda_n)}(y;a^i) & \text{if } n \neq i. \end{cases}$$

Note that $\lambda \cdot x(y) = 0\ \forall y$, and hence $\lambda \cdot \left(u(a^i) + \sum_y p(y|a^i)\, x(y)\right) = \lambda \cdot u(a^i) \geq \max_{v'\in B} \lambda \cdot v' + \varepsilon/4$, by Lemma 7.1. We verify (ICε) and (HS$X/\eta^2$).

For (ICε), for player i, note that $\sum_y p(y|a^i)\, x_i(y) = 0$ by (23), and $\sum_y p\left(y|a_i, a_{-i}^i\right) x_i(y) \leq 0\ \forall a_i \neq a_i^i$ by (24). If Assumption 1.1 holds, then $x_i(y)$ depends only on $(y^n)_{n\neq i}$, so (ICε)(i) holds. If Assumption 1.2 holds, then $u_i(a^i) - u_i\left(a_i, a_{-i}^i\right) \geq \varepsilon\ \forall a_i \neq a_i^i$, so (IC$\varepsilon$)(ii) holds. Next, for any player $n \neq i$, (23) implies that the expected loss in continuation payoff from deviating is at least $\hat{u}$, which exceeds the maximum difference between any two stage game payoffs by at least ε.

For (HS$X/\eta^2$), by $|\lambda_n|/\lambda_i \leq M(\lambda)/m(\lambda) \leq \underline{\lambda}\ \forall n$, (22), and (25), we have

$$\left\| x(y) - \sum_{\tilde{y}} p\left(\tilde{y}|a^i\right) x(\tilde{y}) \right\|^2 \leq \frac{\hat{u}^2}{\eta^2} \underbrace{\left(\left(\sum_{n\neq i} \lambda_n/\lambda_i \right)^2 + N - 1 \right)}_{\leq (N-1)^2\underline{\lambda}^2 + N - 1} \leq \frac{N^2\hat{u}^2}{\eta^2\underline{\lambda}^4}.$$

Hence, (HS$X/\eta^2$) holds, because we have (26) with a^i in place of a^λ, and, since $\|\lambda_+\| \geq 1 - N\underline{\lambda} \geq 1 - \varepsilon/2\bar{u} \geq 1/2$ (arguing as in Case 1 of the proof of Lemma 7 and applying (20)), we have

$$\sum_y \frac{p(y|a^i)\left\|x(y)-\sum_{\tilde{y}} p(\tilde{y}|a^i)\,x(\tilde{y})\right\|^2}{\|\lambda_+\|^2} \leq \frac{4N^2\hat{u}^2}{\eta^2\underline{\lambda}^4} = \frac{\bar{X}}{\eta^2}\bar{u}^2.$$

Case 3: $m(\lambda) \leq \underline{\lambda}$. For α^{NE} satisfying Lemma 7.2, taking $\alpha = \alpha^{NE}$ and $x(y) = 0\ \forall y$ attains a score greater than $\max_{v' \in B} \lambda \cdot v' + \varepsilon/4$ and trivially satisfies (ICε) and (HS$X/\eta^2$).

A.2 Proof of Lemma 6

To show that B is self-generating, it suffices to show that the extreme points of any ball $B' \subseteq B$ of radius $\varepsilon/4$ are decomposable on B'.

Lemma 10 *Suppose that for any ball $B' \subseteq B$ with radius $\varepsilon/4$ and any direction $\lambda \in \Lambda$, the point $\hat{v} = \operatorname{argmax}_{v' \in B'} \lambda \cdot v'$ is decomposable on B'. Then B is self-generating.*

Proof. Fix any $v_0 \in B$. Since the radius of B is $\varepsilon/2$, there exists a ball $B' \subseteq B$ with radius $\varepsilon/4$ such that v_0 lies on the boundary of B'. There then exists a direction λ_0 such that $v_0 = \operatorname{argmax}_{v' \in B'} \lambda_0 \cdot v'$. By hypothesis, v_0 is decomposable on B'. Since $B' \subseteq B$, this implies that v_0 is decomposable on B. Hence, B is self-generating. ■

We thus fix a ball $B' \subseteq B$ with radius $\varepsilon/4$ and a direction $\lambda \in \Lambda$, and let $\hat{v} = \operatorname{argmax}_{v' \in B'} \lambda \cdot v'$. We construct (α, w) that decompose $\hat{v}$ on B'.

Since $k(\lambda, X) \geq \max_{v' \in B} \lambda \cdot v' + \varepsilon/4$ by hypothesis, there exist α and $x : \bar{Y} \to \mathbb{R}^N$ that satisfy (ICε), (HSX), and

$$\lambda \cdot \left(u(\alpha) + \sum_y p(y|\alpha)\,x(y)\right) \geq \max_{v' \in B} \lambda \cdot v' + \varepsilon/5 \geq \max_{v' \in B'} \lambda \cdot v' + \varepsilon/5. \tag{27}$$

Fix any such α and x. Define

$$X_y = \frac{\left\|x(y) - \sum_{y'} p(y'|\alpha)\,x(y')\right\|^2}{\|\lambda_+\|^2\,\bar{u}^2}.$$

Note that, by (HSX) and (19), we have

$$\frac{1-\delta}{\delta}\sqrt{X_y}\bar{u}=\frac{1-\delta}{\delta}\frac{\left\|x\left(y\right)-\sum_{y'}p\left(y'|\alpha\right)x\left(y'\right)\right\|}{\left\|\lambda_{+}\right\|}\leq\frac{1-\delta}{\delta}X\bar{u}\leq\frac{\varepsilon}{64},\quad\text{and}\tag{28}$$

$$\frac{1-\delta}{\delta}\sum_{y}p\left(y|\alpha\right)X_y\bar{u}^2\leq\frac{1-\delta}{\delta}X\bar{u}^2\leq\frac{\varepsilon^2}{640}.\tag{29}$$

To construct w, let

$$\xi_i\left(y\right)=-\frac{64\lambda_iX_y\bar{u}^2}{\varepsilon}\mathbf{1}\left\{\lambda_i\geq 0\right\}\qquad\text{for all } i,y,\tag{30}$$

and let $\xi\left(y\right)=\left(\xi_i\left(y\right)\right)_{i\in I}$. Note that $\xi_i\left(y\right)\leq 0$ for all i,y. Finally, for each y, let

$$w\left(y\right)=\hat{v}+\frac{1-\delta}{\delta}\left(\hat{v}-u\left(\alpha\right)+x\left(y\right)-\sum_{y'}p\left(y'|\alpha\right)x\left(y'\right)\right)+\left(\frac{1-\delta}{\delta}\right)^2\left(\xi\left(y\right)-\sum_{y'}p\left(y'|\alpha\right)\xi\left(y'\right)\right).$$

We show that (α,w) decomposes $\hat{v}$ on B' by verifying in turn (PK), (IC), and (SG) (with $W=B'$).

(PK): This holds by construction: we have $\sum_y p\left(y|\alpha\right)w\left(y\right)=\left(1/\delta\right)\left(\hat{v}-\left(1-\delta\right)u\left(\alpha\right)\right)$, and hence $\left(1-\delta\right)u\left(\alpha\right)+\delta\sum_y p\left(y|\alpha\right)w\left(y\right)=\hat{v}$.

(IC): Setting aside the constant terms in $w\left(y\right)$, we see that an action a_i maximizes $\left(1-\delta\right)u_i\left(a_i,\alpha_{-i}\right)+\delta\sum_y p\left(y|a_i,\alpha_{-i}\right)w_i\left(y\right)$ iff it maximizes $u_i\left(a_i,\alpha_{-i}\right)+\sum_y p\left(y|a_i,\alpha_{-i}\right)\left(x_i\left(y\right)+\frac{1-\delta}{\delta}\xi_i\left(y\right)\right)$. If $\sum_{y:x(y)=x}p\left(y|\tilde{a}_i,a_{-i}\right)$ is independent of $\tilde{a}_i$, then the distribution of $\xi_i\left(y\right)$ is independent of $\tilde{a}_i$, and hence (ICε)(i) implies (IC). Otherwise, (ICε)(ii) holds, and hence (IC) holds because, for all $a_i\notin\operatorname{supp}\alpha_i$ we have

$$\begin{aligned}
&u_i\left(\alpha\right)+\sum_y p\left(y|\alpha\right)\left(x_i\left(y\right)+\frac{1-\delta}{\delta}\xi_i\left(y\right)\right)-u_i\left(a_i,\alpha_{-i}\right)-\sum_y p\left(y|a_i,\alpha_{-i}\right)\left(x_i\left(y\right)+\frac{1-\delta}{\delta}\xi_i\left(y\right)\right)\\
\geq\ &\varepsilon\mathbf{1}\left\{\lambda_i\geq 0\right\}+\sum_y p\left(y|\alpha\right)\frac{1-\delta}{\delta}\xi_i\left(y\right)\qquad\text{by (IC}\varepsilon\text{) and }\xi_i\left(y\right)\leq 0\ \forall y\\
\geq\ &\mathbf{1}\left\{\lambda_i\geq 0\right\}\left(\varepsilon-\frac{64\lambda_i}{\varepsilon}\frac{1-\delta}{\delta}\sum_y p\left(y|\alpha\right)X_y\bar{u}^2\right)\qquad\text{by (30)}\\
\geq\ &0\qquad\text{by (29) and }\lambda_i\leq 1.
\end{aligned}$$

(SG): We start with a standard geometric observation.

Lemma 11 *For each $w \in \mathbb{R}^N$, we have $w \in B'$ if $\lambda \cdot (\hat{v} - w) \geq 0$ and*

$$\|\hat{v} - w\| \leq \sqrt{\frac{\varepsilon}{4} \lambda \cdot (\hat{v} - w)}. \tag{31}$$

Proof. (31) implies that $\lambda \cdot (\hat{v} - w) \leq \sqrt{\frac{\varepsilon}{4} \lambda \cdot (\hat{v} - w)}$, and hence $0 \leq \lambda \cdot (\hat{v} - w) \leq \frac{\varepsilon}{4}$. Let $x := \hat{v} - w - \lambda \cdot (\hat{v} - w) \lambda$. Since $\|x\|^2 = \|\hat{v} - w\|^2 - (\lambda \cdot (\hat{v} - w))^2 \leq \|\hat{v} - w\|^2$, (31) implies that $\|x\|^2 \leq \frac{\varepsilon}{4} \lambda \cdot (\hat{v} - w)$. Denote the center of B' by $o = \hat{v} - \frac{\varepsilon}{4} \lambda$. We have

$$\begin{aligned} \|w - o\| &= \|w - o + x - x\| = \|\hat{v} - o - (\lambda \cdot (\hat{v} - w)) \lambda - x\| = \|(\lambda \cdot (w - o)) \lambda - x\| \\ &= \sqrt{\|\lambda \cdot (w - o) \lambda\|^2 + \|x\|^2} \leq \sqrt{\frac{\varepsilon}{4} \lambda \cdot (w - o) + \frac{\varepsilon}{4} \lambda \cdot (\hat{v} - w)} = \frac{\varepsilon}{4}, \end{aligned}$$

where the third equality is by $\hat{v} - o - (\lambda \cdot \hat{v}) \lambda = \frac{\varepsilon}{4} \lambda - \left(\lambda \cdot \left(o + \frac{\varepsilon}{4} \lambda\right)\right) \lambda = -(\lambda \cdot o) \lambda$, the fourth equality is by $\lambda \cdot x = 0$, the inequality is by $\lambda \cdot (w - o) = \lambda \cdot (\hat{v} - o) - \lambda (\hat{v} - w) \in \left[0, \frac{\varepsilon}{4}\right]$ and $\|x\|^2 \leq \frac{\varepsilon}{4} \lambda \cdot (\hat{v} - w)$, and the final equality is by $\lambda \cdot (\hat{v} - o) = \frac{\varepsilon}{4}$. Hence, $w \in B'$. ■

We thus show that, for each y, $w(y)$ satisfies $\lambda \cdot (\hat{v} - w(y)) \geq 0$ and (31). Note that

$$\hat{v} - w(y) = \frac{1 - \delta}{\delta} \Delta(y) - \left(\frac{1 - \delta}{\delta}\right)^2 \xi(y) + \left(\frac{1 - \delta}{\delta}\right)^2 \sum_{y'} p(y'|\alpha) \xi(y'),$$

where $\Delta(y) = u(\alpha) - \hat{v} + \sum_{y'} p(y'|\alpha) x(y') - x(y)$. By (HS$X$) and (27), we have $\lambda \cdot \Delta(y) \geq \frac{1-\delta}{\delta} \frac{\varepsilon}{5}$ and

$$\|\Delta(y)\| \leq \|u(\alpha) - \hat{v}\| + \left\| \sum_{y'} p(y'|\alpha) x(y') - x(y) \right\| \leq \sqrt{N} \bar{u} + \|\lambda_+\| \sqrt{X_y} \bar{u}.$$

By (HSX) and the definition of ξ (cf. (30)),

$$\begin{aligned} -\lambda \cdot \xi(y) &\geq \|\lambda_+\|^2 \frac{64 X_y}{\varepsilon} \bar{u}^2, \qquad \|\xi(y)\| \leq \|\lambda_+\| \frac{64 X_y}{\varepsilon} \bar{u}^2, \qquad \text{and} \\ \left\| \sum_{y'} p(y'|\alpha) \xi(y') \right\| &= \left\| \left(\lambda_i \mathbf{1}\{\lambda_i \geq 0\} \frac{64 \sum_y p(y|\alpha) X_y \bar{u}^2}{\varepsilon} \right)_i \right\| \leq \|\lambda_+\| \frac{64}{\varepsilon} \sum_y p(y|\alpha) X_y \bar{u}^2. \end{aligned}$$

Therefore,

$$\begin{aligned}
\lambda \cdot (\hat{v} - w(y)) &= \frac{1-\delta}{\delta}\lambda \cdot \Delta(y) - \left(\frac{1-\delta}{\delta}\right)^2 \lambda \cdot \xi(y) + \left(\frac{1-\delta}{\delta}\right)^2 \lambda \cdot \left(\sum_{y'} p(y'|\alpha)\xi(y')\right) \\
&\geq \frac{1-\delta}{\delta}\frac{\varepsilon}{5} + \left(\frac{1-\delta}{\delta}\right)^2 \|\lambda_+\|^2 \frac{64X_y\bar{u}^2}{\varepsilon} - \left(\frac{1-\delta}{\delta}\right)^2 \|\lambda_+\| \frac{64}{\varepsilon}\sum_y p(y|\alpha) X_y \bar{u}^2 \\
&\geq \frac{1-\delta}{\delta}\frac{\varepsilon}{10} + \left(\frac{1-\delta}{\delta}\right)^2 \|\lambda_+\|^2 \frac{64X_y\bar{u}^2}{\varepsilon} \geq 0, \qquad (32)
\end{aligned}$$

where the last line follows since $\|\lambda_+\| \leq 1$ and (29) imply that $\frac{1-\delta}{\delta}\|\lambda_+\|\frac{64}{\varepsilon}\sum_y p(y|\alpha) X_y \bar{u}^2 \leq \frac{\varepsilon}{10}$. Thus, we have

$$\sqrt{\frac{\varepsilon}{4}\lambda \cdot (\hat{v} - w(y))} \geq 4\frac{1-\delta}{\delta}\max\left\{\sqrt{\frac{1}{640}\frac{\delta}{1-\delta}\varepsilon}, \|\lambda_+\|\sqrt{X_y}\bar{u}\right\}.$$

Similarly, we have

$$\begin{aligned}
&\|\hat{v} - w(y)\| \leq \frac{1-\delta}{\delta}\|\Delta(y)\| + \left(\frac{1-\delta}{\delta}\right)^2 \|\xi(y)\| + \left(\frac{1-\delta}{\delta}\right)^2 \left\|\textstyle\sum_{y'} p(y'|\alpha)\xi(y')\right\| \\
\leq\ & \frac{1-\delta}{\delta}\left(\sqrt{N}\bar{u} + \|\lambda_+\|\sqrt{X_y}\bar{u}\right) + \left(\frac{1-\delta}{\delta}\right)^2 \|\lambda_+\|\frac{64X_y}{\varepsilon}\bar{u}^2 + \left(\frac{1-\delta}{\delta}\right)^2 \|\lambda_+\|\frac{64}{\varepsilon}\sum_y p(y|\alpha) X_y\bar{u}^2 \\
\leq\ & 2\frac{1-\delta}{\delta}\left(\sqrt{N} + \|\lambda_+\|\sqrt{X_y}\right)\bar{u} \leq 4\frac{1-\delta}{\delta}\max\left\{\sqrt{N}\bar{u}, \|\lambda_+\|\sqrt{X_y}\bar{u}\right\}, \qquad (33)
\end{aligned}$$

where the third inequality follows since $\|\lambda_+\| \leq 1$ and (29) imply that $\frac{1-\delta}{\delta}\frac{64}{\varepsilon}\sum_y p(y|\alpha) X_y \bar{u}^2 \leq \frac{1}{10}\varepsilon \leq \bar{u}$, and (28) implies that $\frac{1-\delta}{\delta}\frac{64X_y}{\varepsilon}\bar{u}^2 \leq \sqrt{X_y}\bar{u}$.

Comparing (32) and (33), we see that $w(y)$ satisfies (31) whenever $\sqrt{N}\bar{u} \leq \sqrt{\frac{1}{640}\frac{\delta}{1-\delta}\varepsilon}$, which holds by (19).

A.3 Proof of Lemma 7

Case 1: $m(\lambda) \geq M(\lambda)/\underline{\lambda}$. Let $i = i(\lambda)$. Since $m(\lambda) \leq 1$ and $|\lambda_n| \leq M(\lambda) \leq m(\lambda)\underline{\lambda}$ $\forall n \neq i$, we have $|\lambda_n| \leq \underline{\lambda}$ $\forall n \neq i$, and hence $|\lambda_i| \geq 1 - N\underline{\lambda}$ (since $\|\lambda\| = 1$). Since $\lambda_i = m(\lambda) \geq M(\lambda)/\underline{\lambda} > 0$, we have $\lambda_i \geq 1 - N\underline{\lambda}$. Since $2|u_i(a)| \leq \bar{u}$ $\forall i, a$, we have, for all $v' \in F^*$ and $\lambda \in \Lambda$, $|(\lambda - e_i)\cdot v'| \leq \sum_n |\lambda_n - e_{i,n}|\bar{u}/2 \leq ((N-1)\underline{\lambda} + |(1 - N\underline{\lambda}) - 1|)\bar{u}/2 \leq N\underline{\lambda}\bar{u}$.

Therefore, for $a^i \in \arg\max_{a\in A} e_i \cdot u(a)$, we have $\lambda \cdot u(a^i) \geq e_i \cdot u(a^i) - N\underline{\lambda}\bar{u} \geq \max_{v'\in F^*} e_i \cdot v' - N\underline{\lambda}\bar{u} \geq \max_{v'\in F^*} \lambda \cdot v' - 2N\underline{\lambda}\bar{u} \geq \max_{v'\in B} \lambda \cdot v' + \varepsilon/2 - 2N\underline{\lambda}\bar{u} \geq \max_{v'\in B} \lambda \cdot v' + \varepsilon/4$, where the last inequality is by (20).

Case 2: $m(\lambda) \leq \underline{\lambda}$. Let $\lambda'_n = \min\{\lambda_n, 0\} / \|\lambda_-\|$ and let $\lambda' = (\lambda'_n)_{n\in I} \in \Lambda$. We claim that $\sum_n |\lambda'_n - \lambda_n| \leq \varepsilon/2\bar{u}$. To see this, note that if $\lambda_n \geq 0$ then $|\lambda_n| \leq \underline{\lambda}$, and hence $|\lambda'_n - \lambda_n| = |0 - \lambda_n| = |\lambda_n| \leq \underline{\lambda}$. If instead $\lambda_n \leq 0$, then

$$|\lambda'_n - \lambda_n| = \left|\frac{\lambda_n - \|\lambda_-\|\,\lambda_n}{\|\lambda_-\|}\right| \leq \frac{1 - \|\lambda_-\|}{\|\lambda_-\|} \leq \frac{1 - \sqrt{1 - N\underline{\lambda}^2}}{\sqrt{1 - N\underline{\lambda}^2}},$$

where the first inequality follows because $|\lambda_n| \leq 1$, and the second inequality follows because, since $\sum_{n'} (\lambda_{n'})^2 = 1$ and $\lambda_n \leq m(\lambda) \leq \underline{\lambda}\ \forall n$, we have $\|\lambda_-\| = \sum_{n':\lambda_{n'}<0} (\lambda_{n'})^2 \geq 1 - N\underline{\lambda}^2$. In total, we have

$$\sum_n |\lambda'_n - \lambda_n| \leq N \max\left\{\underline{\lambda}, \frac{1 - \sqrt{1 - N\underline{\lambda}^2}}{\sqrt{1 - N\underline{\lambda}^2}}\right\} \leq \frac{\varepsilon}{2\bar{u}} \qquad \text{by (20).}$$

Since $\lambda' \leq 0$, by definition of F^* there exists a static Nash equilibrium α^{NE} such that $\lambda' \cdot u\left(\alpha^{NE}\right) \geq \max_{v'\in F^*} \lambda' \cdot v'$. Since $\sum_n |\lambda'_n - \lambda_n| \leq \varepsilon/2\bar{u}$, $2|u_i(a)| \leq \bar{u}\ \forall i, a$, and the distance from B to the boundary of F^* is greater than $\varepsilon/2$, we have $\lambda \cdot u\left(\alpha^{NE}\right) \geq \lambda' \cdot u\left(\alpha^{NE}\right) - \varepsilon/4 \geq \max_{v'\in F^*} \lambda' \cdot v' - \varepsilon/4 \geq \max_{v'\in B} \lambda' \cdot v' + \varepsilon/4$, as desired.

B Proof of Proposition 1

We start by deriving a sufficient condition to self-generate the convex hull of the union of a ball B centered on the 45° line and the mutual defection payoff $(0,0)$.[19] For any compact set of payoffs $V \subseteq F^*$, we define the *minimum inefficiency of* V as

$$\begin{aligned}
\rho(V) &= \min_{\lambda\in\Lambda_+} \max_{v\in F^*} \min_{v'\in V} \lambda \cdot (v - v'), \qquad \text{where} \\
\Lambda_+ &= \left\{\lambda \in \Lambda : (0,0) \notin \arg\max_{v\in F^*} \min_{v'\in V} \lambda \cdot (v - v')\right\}.
\end{aligned}$$

[19] A heuristic motivation for this proof approach is that the set of PPE payoffs in the continuous-time limit game has a shape that resembles the convex hull of the union of a ball centered on the 45° line and the mutual defection payoff. Compare Figure 1 below and Figure 2 of Sannikov (2007).

That is, $\rho(V)$ is the minimum distance between the boundaries of V and F^* in any direction λ, excluding directions where the minimizing boundary point of V is the mutual defection payoff $(0,0)$. Also, for any ball $B \subseteq F^*$ with center $o = (v_o, v_o)$ lying on the 45-degree line and curvature κ, let ψ denote the slope of a tangent line to B passing through $(0,0)$, which is given by

$$\psi = \frac{v_o^2\kappa^2 - \sqrt{2v_o^2\kappa^2 - 1}}{v_o^2\kappa^2 - 1}. \tag{34}$$

Note that $B \subseteq F^*$ implies $v_o^2\kappa^2 \geq 1$, with strict inequality if $B \cap \partial F^* = \emptyset$. See Figure 1.

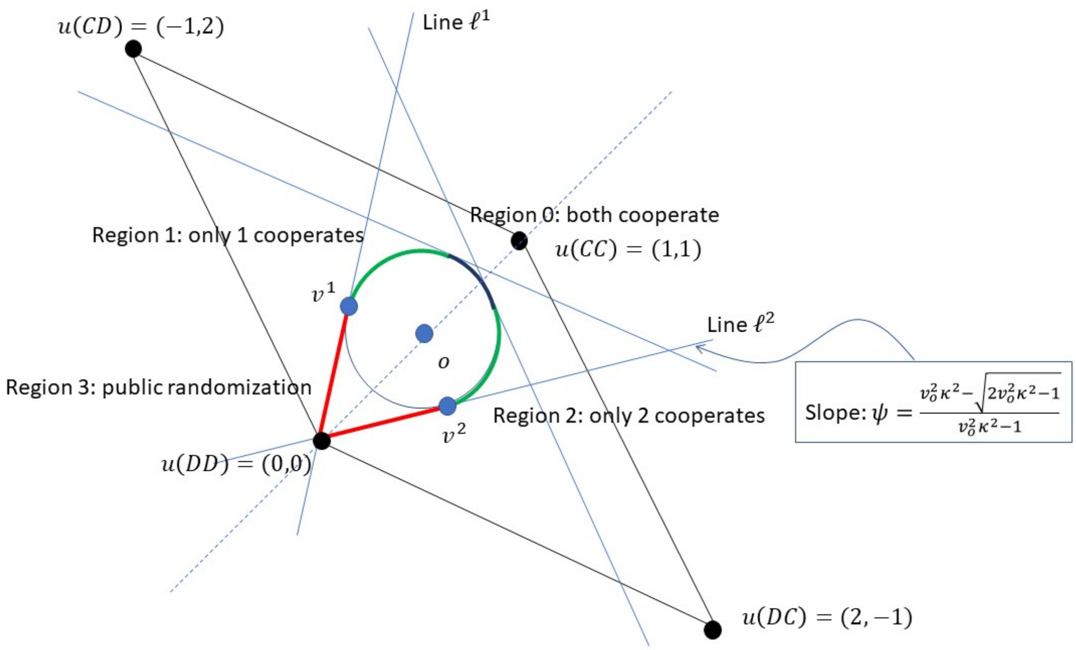

Figure 1: Self-generating the set B^0. The set B is the ball centered at o passing through v^1 and v^2. The set B^0 is the convex hull of the union of B and the point $(0,0)$. To self-generate ∂B^0, both players cooperate to generate points on the blue portion of the bondary (Region 0), one player cooperates to generate points on the green portions (Regions 1 and 2), and the players randomize between v^1 or v^2 and mutual defection to generate points on the red portion (Region 3).

Lemma 12 *Let $B \subseteq F^*$ be a ball with center $o = (v_o, v_o)$ and curvature κ, and let $B^0 = \operatorname{co}((0,0) \cup B)$. If*

$$\frac{1}{r}\frac{(2\pi - 1)^2}{\pi(1-\pi)} > \frac{\kappa}{2\rho(B^0)\min\left\{\psi^2, \frac{4}{25}\right\}}, \tag{35}$$

then B^0 is self-generating for all sufficiently small Δ.

Proof. Let ℓ^1 and ℓ^2 denote the tangent lines to B passing through $(0,0)$, and let v^1 and v^2 denote the corresponding tangency points on ∂B, with $v_1^1 < v_2^1$. (See Figure 1.) Note that ψ is the slope of ℓ^2.[20]

We divide the boundary of B^0 into four regions. Region 0 is the set of $v \in \partial B^0$ where the outward unit normal vector $\lambda(v) = (\lambda_1, \lambda_2)$ at v satisfies $\lambda_1 > 0$, $\lambda_2 > 0$, and $\lambda_1/\lambda_2 \in [1/2, 2]$. For $i \in \{1,2\}$, Region i is the set of $v \in \partial B^0$ where (λ_1, λ_2) satisfies $-1/\psi \leq \lambda_i/\lambda_j \leq 1/2$ and $\lambda_j > 0$. (Throughout, i and j denote distinct players.) Region 3 is the rest of ∂B^0.

For $v \in \partial B^0$, let $a(v) = (C, C)$ for v in Region 0, $(a_i(v), a_j(v)) = (C, D)$ for v in Region i, and $a(v) = (D, D)$ for v in Region 3. Note that $a(v) \in \arg\max_{a \in A} \min_{v \in B^0} \lambda \cdot (u(a) - v)$, and $a(v)$ is the unique maximizer unless v is the boundary of two regions. By the definition of $\rho(B^0)$, for each v in Region 0, 1, or 2, we have

$$\lambda(v) \cdot (u(a(v)) - v) \geq \rho\left(B^0\right). \tag{36}$$

Since public randomization is available, it suffices to show that for sufficient small Δ, each $v \in \partial B^0$ is decomposable on B^0. We prove this for $v \in \partial B^0$ in each of the four regions in turn. We require some notation: first, let

$$\beta_i\left(y^i\right) = \frac{\frac{1}{2} - \left(\pi - \frac{1}{2}\right)\sqrt{\Delta/\gamma}}{(2\pi - 1)\sqrt{\Delta/\gamma}} \mathbf{1}\left\{y^i = 1\right\} - \frac{\frac{1}{2} + \left(\pi - \frac{1}{2}\right)\sqrt{\Delta/\gamma}}{(2\pi - 1)\sqrt{\Delta/\gamma}} \mathbf{1}\left\{y^i = -1\right\}.$$

We will use $\beta_i(y^i)$ as a "reward function" to induce player i to take C: in particular, we will use the properties

$$\mathbb{E}\left[\beta_i\left(y^i\right)|C, a_j\right] - \mathbb{E}\left[\beta_i\left(y^i\right)|D, a_j\right] = 1 \text{ for all } a_j, \qquad \text{and} \qquad \mathbb{E}\left[\beta_i\left(y^i\right)|C, C\right] = 0.$$

Second, we fix $\hat{\varepsilon} > 0$ such that (the existence follows from (35))

$$r\frac{\pi(1-\pi)}{(2\pi - 1)^2}(1 + \hat{\varepsilon})^2 < \frac{2\rho(B^0)\min\left\{\psi^2, \frac{4}{25}\right\}}{\kappa}. \tag{37}$$

[20] Equation (34) for ψ is derived by solving the equations $\left\|v^2 - o\right\| = 1/\kappa$ and $\left(v^2 - o\right) \cdot v^2 = 0$ for $v^2 = \left(v_1^2, v_2^2\right)$, and taking $\psi = v_2^2/v_1^2$.

Region 0: Fix $v \in \partial B^0$ in Region 0. We implement $a = (C, C)$. Let λ be the outward unit normal vector of ∂B^0 at v. Define $x(y)$ such that

$$x_i(y) = (1+\hat{\varepsilon})\beta_i(y^i) - \frac{\lambda_j}{\lambda_i}(1+\hat{\varepsilon})\beta_j(y^j) \text{ for each } i.$$

Since v is in Region 0, we have $|\lambda_1/\lambda_2| \in [1/2, 2]$, which together with $\|\lambda\| = 1$ implies that $\frac{1}{\lambda_1^2} + \frac{1}{\lambda_2^2} \leq \frac{25}{4}$. Thus, we have

$$u_i(a) + \mathbb{E}[x_i(y)|a] = 0, \tag{38}$$

$$u_i(a) + \mathbb{E}[x_i(y)|a] - u_i(a_i') - \mathbb{E}[x_i(y)|a_i', a_j] \geq \hat{\varepsilon} \text{ for } a_i' \neq a_i, \tag{39}$$

$$\lim_{\Delta \to 0} \mathbb{E}\left[\|x(y)\|^2 |a\right] - \mathbb{E}\left[\|x(y)\|^2 |a_i', a_j\right] = -(1+\hat{\varepsilon})^2, \tag{40}$$

$$\lim_{\Delta \to 0} \frac{1-\delta}{\delta}\mathbb{E}\left[\|x(y)\|^2 |a\right] = \left(\frac{1}{\lambda_1^2} + \frac{1}{\lambda_2^2}\right) r \frac{\pi(1-\pi)}{(2\pi-1)^2}(1+\hat{\varepsilon})^2 \leq \frac{25}{4} r \frac{\pi(1-\pi)}{(2\pi-1)^2}(1+\hat{\varepsilon})^2, \tag{41}$$

$$\max_y \sqrt{\frac{1-\delta}{\delta}}\|x(y)\| \leq \sqrt{\frac{1-\delta}{\delta}}\sqrt{\frac{1}{\lambda_1^2} + \frac{1}{\lambda_2^2}}(1+\hat{\varepsilon})\frac{\frac{1-\sqrt{\Delta/\gamma}}{2} + \pi\sqrt{\Delta/\gamma}}{(2\pi-1)\sqrt{\Delta/\gamma}} = O(1). \tag{42}$$

The above limits are all uniform in λ, since $\lambda_1, \lambda_2 \geq 1/4$ in Region 0.

Define continuation payoffs

$$w(y) = v + \frac{1-\delta}{\delta}(v - u(a) + x(y)) - \frac{\kappa}{2}\left(\frac{1-\delta}{\delta}\right)^2 \left(\|x(y)\|^2 - \mathbb{E}\left[\|x(y)\|^2 |a\right]\right)\lambda.$$

With these continuation payoffs, (PK) holds because, since $\mathbb{E}[x(y)|a] = 0$, we have

$$(1-\delta)u(a) + \delta\mathbb{E}[w(y)|a] = v + \frac{1-\delta}{\delta}\mathbb{E}[x(y)|a] = v.$$

(IC) holds because

$$\begin{aligned} & (1-\delta)\left(u(a) + \frac{\delta}{1-\delta}\mathbb{E}[w(y)|a] - u(a_i', a_j) + \frac{\delta}{1-\delta}\mathbb{E}[w(y)|a_i', a_j]\right) \\ = \; & (1-\delta)\left(\hat{\varepsilon} + \frac{\kappa}{2}\frac{1-\delta}{\delta}\left(\mathbb{E}\left[\|x(y)\|^2 |a\right] - \mathbb{E}\left[\|x(y)\|^2 |a_i', a_j\right]\right)\right), \end{aligned}$$

and, by (40),

$$\lim_{\Delta\to 0}\frac{1-\delta}{\delta}\left(\mathbb{E}\left[\|x(y)\|^2|a\right]-\mathbb{E}\left[\|x(y)\|^2|a_i',a_j\right]\right)=0.$$

Moreover, this convergence is uniform over v in Region 0, as λ_i is bounded away from 0 in Region 0. Thus, for sufficiently small Δ, (IC) holds. Finally, by the Pythagorean theorem, (SG) holds if

$$\left((\lambda_2,-\lambda_1)^\top\cdot(v-w(y))\right)^2\le\left(\frac{1}{\kappa}\right)^2-\left(\frac{1}{\kappa}-\lambda\cdot(v-w(y))\right)^2\qquad\text{for all } y.$$

Substituting $w(y)$ and simplifying terms, this inequality is equivalent to

$$\begin{aligned}
&\frac{1-\delta}{\delta}\left((\lambda_2,-\lambda_1)^\top\cdot(u(a)-v)\right)^2-2\left((\lambda_2,-\lambda_1)^\top\cdot(u(a)-v)\right)\frac{1-\delta}{\delta}\|x(y)\|\\
\le\ &\frac{2}{\kappa}\lambda\cdot(u(a)-v)-\left(\frac{1-\delta}{\delta}\right)\mathbb{E}\left[\|x(y)\|^2|CC\right]\\
&-\left(\sqrt{\frac{1-\delta}{\delta}}\lambda\cdot(u(a)-v)-\frac{\kappa}{2}\left(\frac{1-\delta}{\delta}\right)^{\frac{3}{2}}\mathbb{E}\left[\|x(y)\|^2|CC\right]+\frac{\kappa}{2}\left(\frac{1-\delta}{\delta}\right)^{\frac{3}{2}}\|x(y)\|^2\right)^2.
\end{aligned}$$

By (41) and (42), all terms in the first and last lines converge to 0 uniformly in Δ. Hence, by (41), for small Δ this inequality is implied by

$$2\lambda\cdot(u(a)-v)-\kappa r\frac{25}{4}\frac{\pi(1-\pi)}{(2\pi-1)^2}(1+\hat{\varepsilon})^2>0.$$

Since (36) implies $\lambda\cdot(u(a)-v)\ge\rho(B^0)$, this inequality follows from (37).

Region $i\in\{1,2\}$: Fix $v\in\partial B^0$ in Region i. We implement $(a_i,a_j)=(C,D)$. With λ defined as above, we have $|\lambda_j|\ge\psi$. Define

$$\begin{aligned}
x_i(y)&=(1+\hat{\varepsilon})\beta_i\left(y^i\right),\qquad x_j(y)=-(1+\hat{\varepsilon})\frac{\lambda_i}{\lambda_j}\beta_i\left(y^i\right),\text{ and}\\
w(y)&=v+\frac{1-\delta}{\delta}(v-u(a)+x(y))-\frac{\kappa}{2}\left(\frac{1-\delta}{\delta}\right)^2\left(\|x(y)\|^2-\mathbb{E}\left[\|x(y)\|^2|a\right]\right)\lambda.
\end{aligned}$$

Note that only player i is incentivized by continuation payoffs. The rest of the proof is the same as in Region 0, except for the following: For (IC) for player j, player j has a strict incentive to take D because the distribution of any function of $x(y)$ is independent of her action. For (SG), we have

$$\lim_{\Delta\to 0}\frac{1-\delta}{\delta}\mathbb{E}\left[\|x(y)\|^2|a\right]=\frac{1}{\lambda_j^2}r\frac{\pi(1-\pi)}{(2\pi-1)^2}(1+\hat{\varepsilon})^2\leq\frac{1}{\psi^2}r\frac{\pi(1-\pi)}{(2\pi-1)^2}(1+\hat{\varepsilon})^2,$$

so the argument is as in the Region 0 case with ψ^2 in place of $4/25$.

Region 3: Each $v\in\partial B^0$ in Region 3 is a convex combination of $(0,0)$ and either v^1 or v^2. The point $(0,0)$ is a static Nash payoff and hence is trivially decomposable on B^0, and we have seen that each v^i is decomposable on B^0, since v^i is in Region i. Hence, each $v\in\partial B^0$ in Region 3 is decomposable on B^0 given public randomization. ■

Proof of Proposition 1: Take B with center $(1-3\varepsilon/2,1-3\varepsilon/2)$ and curvature $\sqrt{2}/\varepsilon$. Note that $(1-\varepsilon,1-\varepsilon)\in\partial B$. Elementary calculations (given below) yield $\psi\geq 2/5$ and $\rho(B^0)=\left(\frac{9}{10}\sqrt{5}-\frac{1}{\sqrt{2}}\right)\varepsilon$. Hence, if (12) holds then so does (35), so by Lemma 12 B^0 is self-generating, and thus $(1-\varepsilon,1-\varepsilon)\in E$.

The required calculations are as follows: By definition of ψ, we have

$$\psi=\frac{(2-3\varepsilon)^2-4\varepsilon\sqrt{1-3\varepsilon+2\varepsilon^2}}{4-12\varepsilon+7\varepsilon^2}.$$

This expression equals $2/5$ when $\varepsilon=\frac{54}{23}-\frac{6}{23}\sqrt{58}$, and it is decreasing in ε for $\varepsilon\in\left(0,\frac{54}{23}-\frac{6}{23}\sqrt{58}\right)$. Hence, $\psi\geq 2/5$ iff $\varepsilon\leq\frac{54}{23}-\frac{6}{23}\sqrt{58}$.

Next, note that for each $\lambda\in\Lambda_+$, there exists $v\in\partial B^0$ in Region 0, 1, or 2 such that λ is the outward unit normal vector at v. Note that $\|(-1,2)-v^1\|=\|(2,-1)-v^2\|\geq 1$, which is greater than $\left(\frac{9}{10}\sqrt{5}-\frac{1}{\sqrt{2}}\right)\varepsilon$ when $\varepsilon\leq\frac{54}{23}-\frac{6}{23}\sqrt{58}$. Thus, to show that $\rho(B^0)=\left(\frac{9}{10}\sqrt{5}-\frac{1}{\sqrt{2}}\right)\varepsilon$, it suffices to show that the distance d from o to the line ℓ with equation $y=(3-x)/2$ (i.e., the line through $(-1,2)$ and $(1,1)$) is $\frac{9}{10}\sqrt{5}\varepsilon$ as the radius of B^0 is $\varepsilon/\sqrt{2}$). To see this, note that the vector $(1,2)/\sqrt{5}$ is normal to ℓ, so the point $z\in\ell$ that minimizes the distance to o is $z=o+d(1,2)/\sqrt{5}$. At the same time, since $z\in\ell$, we have

$$1-\frac{3}{2}\varepsilon+\frac{2d}{\sqrt{5}}=\frac{1}{2}\left(3-\left(1-\frac{3}{2}\varepsilon+\frac{d}{\sqrt{5}}\right)\right).$$

Solving for d gives $d=\frac{9}{10}\sqrt{5}\varepsilon$, as desired.

References

[1] Abreu, Dilip, David Pearce, and Ennio Stacchetti. "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring." *Econometrica* 58.5 (1990): 1041-1063.

[2] Abreu, Dilip, Paul Milgrom, and David Pearce. "Information and Timing in Repeated Partnerships." *Econometrica* 59.6 (1991): 1713-1733.

[3] Al-Najjar, Nabil I., and Rann Smorodinsky. "Pivotal Players and the Characterization of Influence." *Journal of Economic Theory* 92.2 (2000): 318-342.

[4] Al-Najjar, Nabil I., and Rann Smorodinsky. "Large Nonanonymous Repeated Games." *Games and Economic Behavior* 37.1 (2001): 26-39.

[5] Awaya, Yu, and Vijay Krishna. "On Communication and Collusion." *American Economic Review* 106.2 (2016): 285-315.

[6] Awaya, Yu, and Vijay Krishna. "Communication and Cooperation in Repeated Games." *Theoretical Economics* 14.2 (2019): 513-553.

[7] Billingsley, Patrick. *Probability and Measure*, 3rd Ed. (1995). Wiley.

[8] Dewatripont, Mathias, Ian Jewitt, and Jean Tirole. "The Economics of Career Concerns, Part I: Comparing Information Structures." *Review of Economic Studies* 66.1 (1999): 183-198.

[9] Dubins, Lester E., and Leonard J. Savage. *How to Gamble if You Must: Inequalities for Stochastic Processes* (1965). Dover.

[10] Forges, Francoise. "An Approach to Communication Equilibria." *Econometrica* 54.6 (1986): 1375-1385.

[11] Freedman, David A. "On Tail Probabilities for Martingales." *Annals of Probability* 3.1 (1975): 100-118.

[12] Fudenberg, Drew, and David K. Levine. "Efficiency and Observability with Long-Run and Short-Run Players." J*ournal of Economic Theory* 62.1 (1994): 103-135.

[13] Fudenberg, Drew, and David K. Levine. "Continuous Time Limits of Repeated Games with Imperfect Public Monitoring." *Review of Economic Dynamics* 10.2 (2007): 173-192.

[14] Fudenberg, Drew, David Levine, and Eric Maskin. "The Folk Theorem with Imperfect Public Information." *Econometrica* 62.5 (1994): 997-1039.

[15] Fudenberg, Drew, David Levine, and Wolfgang Pesendorfer. "When are Nonanonymous Players Negligible?" *Journal of Economic Theory* 79.1 (1998): 46-71.

[16] Green, Edward J. "Noncooperative Price Taking in Large Dynamic Markets." *Journal of Economic Theory* 22.2 (1980): 155-182.

[17] Hébert, Benjamin. "Moral Hazard and the Optimality of Debt." *Review of Economic Studies* 85.4 (2018): 2214-2252.

[18] Holmström, Bengt. "Moral Hazard and Observability." *Bell Journal of Economics* 10.1 (1979): 74-91.

[19] Hörner, Johannes, and Satoru Takahashi. "How Fast do Equilibrium Payoff Sets Converge in Repeated Games?" *Journal of Economic Theory* 165 (2016): 332-359.

[20] Jewitt, Ian, Ohad Kadan, and Jeroen M. Swinkels. "Moral Hazard with Bounded Payments." *Journal of Economic Theory* 143.1 (2008): 59-82.

[21] Kandori, Michihiro. "Introduction to Repeated Games with Private Monitoring." *Journal of Economic Theory* 102.1 (2002): 1-15.

[22] Kandori, Michihiro, and Hitoshi Matsushima. "Private Observation, Communication and Collusion." *Econometrica* 66.3 (1998): 627-652.

[23] Mirrlees, James A. "The Theory of Moral Hazard and Unobservable Behaviour: Part I." *Working Paper* (1975) (published in *Review of Economic Studies* 66.1 (1999): 3-21).

[24] Pai, Mallesh M., Aaron Roth, and Jonathan Ullman. "An Antifolk Theorem for Large Repeated Games." *ACM Transactions on Economics and Computation (TEAC)* 5.2 (2016): 1-20.

[25] Sabourian, Hamid. "Anonymous Repeated Games with a Large Number of Players and Random Outcomes." *Journal of Economic Theory* 51.1 (1990): 92-110.

[26] Sadzik, Tomasz, and Ennio Stacchetti. "Agency Models with Frequent Actions." *Econometrica* 83.1 (2015): 193-237.

[27] Sannikov, Yuliy. "Games with Imperfectly Observable Actions in Continuous Time." *Econometrica* 75.5 (2007): 1285-1329.

[28] Sannikov, Yuliy, and Andrzej Skrzypacz. "The Role of Information in Repeated Games with Frequent Actions." *Econometrica* 78.3 (2010): 847-882.

[29] Sason, Igal, and Sergio Verdú. "f-Divergence Inequalities." *IEEE Transactions on Information Theory* 62.11 (2016): 5973-6006.

[30] Sugaya, Takuo, and Alexander Wolitzky, "Bounding Equilibrium Payoffs in Repeated Games with Private Monitoring." *Theoretical Economics* 12 (2017): 691-729.

[31] Sugaya, Takuo, and Alexander Wolitzky, "Bounding Payoffs in Repeated Games with Private Monitoring: n-Player Games." *Journal of Economic Theory* 175 (2018): 58-87.

[32] Sugaya, Takuo, and Alexander Wolitzky, "Repeated Games with Many Players." *Working Paper* (2022a).

[33] Sugaya, Takuo, and Alexander Wolitzky, "Rate of Convergence in Repeated Games: A Universal Speed Limit." *Working Paper* (2022b).