Informational Requirements for Cooperation*

Takuo Sugaya — Stanford GSB
Alexander Wolitzky — MIT

August 16, 2022

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 payoff variance. We show that this bound is tight in the low-discounting/low-monitoring double limit, by establishing a folk theorem where the discount factor and the monitoring structure vary simultaneously.

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

JEL codes: C72, C73

*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 for other applications that we consider in companion papers: these include games with *many players*, where a large population of players are monitored by an aggregate signal; and the *rate of convergence* of the equilibrium payoff set as discounting and monitoring vary. We discuss these applications at the end of the paper.

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 (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 in a particular period by letting continuation play depend disproportionately on behavior in that period, 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. Several other well-known measures (e.g., total variation distance, Kullback-Leibler divergence) are equivalent to χ^2-divergence under our assumptions and hence are equally valid for characterizing the asymptotic tradeoff between discounting and monitoring; however, our proofs rely on χ^2-divergence, and our non-asymptotic results are strongest under this measure.

In total, Theorem 1 may be summarized as stating that, for any repeated game Γ, any 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 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 *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. We show that recursively decomposing the variance of a player's continuation payoffs yields a tight bound on the average strength of her incentives, despite the well-known fact that the set of equilibrium payoffs with private strategies or monitoring generally lacks a useful recursive structure (Kandori, 2002).

Our positive result (Theorem 2) is a partial converse to Theorem 1. 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. It generalizes SS by considering the general low-discounting/low-monitoring double limit, rather than assuming that monitoring is parameterized by an underlying continuous-time signal process.[1]

Like the folk theorems of FLM, KM, and SS, Theorem 2 assumes *pairwise identifiability*. This is a standard assumption, but it makes it harder to directly compare Theorems 1 and 2, because Theorem 1 does not require this assumption. However, pairwise identifiability is unnecessary if monitoring has a *product structure*. In this case, a direct comparison of Theorems 1 and 2 is possible. This comparison (Corollary 1) confirms that Theorems 1 and 2 tightly characterize the asymptotic tradeoff between discounting and monitoring required for cooperation, at least for games with public, product structure monitoring.[2]

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.

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

[2] We conjecture that Theorem 2 holds even without pairwise identifiability, so that Theorems 1 and 2 characterize the asymptotic tradeoff between discounting and monitoring in general. However, proving this result would involve complications beyond the scope of the current paper.

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., $\bar{u} = \max_{i,a} 2 |u_i(a)|$. 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 $(a_i)_{i \in I}$, (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 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 $p\left((y_i)_i \,|\, (a_i)_i\right)$), 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. Our necessary conditions for cooperation (Theorem 1) apply for Γ^B, and hence apply a fortiori for Γ.

The Public Game. We also define another variant of a repeated game Γ: the corresponding *public game*, denoted Γ^P, where all players observe the entire signal vector y at the end of a period. Note that Γ^P is a repeated game with imperfect public monitoring, as defined by FLM.[3] A *perfect public equilibrium (PPE)* in such a game is a strategy profile that forms a Nash equilibrium conditional on any public history $h^t = (y_{t'})_{t'=1}^{t-1}$. For a given stage game G, let $E(Y, p, \delta)$ denote the set of PPE payoff vectors in the repeated game Γ^P. Note that the set of Nash equilibrium payoffs in Γ^P is smaller than that in Γ^B, and the set $E(Y, p, \delta)$ is smaller still. Our sufficient conditions for cooperation (Theorem 2) apply for $E(Y, p, \delta)$, and hence apply a fortiori for the set of Nash equilibrium payoffs in Γ^B.

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 measures the "expected discounted fraction of periods" where each action profile is played in the course of the repeated game. Note that the payoffs under

[3]However, we do not need to impose FLM's assumption that a player's payoff is determined by her own action and the signal y.

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

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

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

Next, 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

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

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 arises 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*

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

ratio difference $\left(p\left(y|\tilde{a}_i, a_{-i}\right) - p\left(y|a\right)\right)/p\left(y|a\right)$ determines the "strength of incentives" provided by rewards or punishments that are conditioned on the arrival of signal y (Mirrlees, 1975; Holmström, 1979). Since the expected likelihood ratio difference $\sum_y p\left(y|a\right)\left(\left(p\left(y|\tilde{a}_i, a_{-i}\right) - p\left(y|a\right)\right)/p\left(y|a\right)\right)$ 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.[5]

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

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

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

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

This observation shows that χ^2-divergence and continuation payoff variance must both be sufficiently 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 this result requires that the Euclidean distance between the (suitably normalized) vectors of signal probabilities $\left(p\left(y|a\right)\right)_{y\in\bar{Y}}$ and $\left(p\left(y|s_i\left(a_i\right), a_{-i}\right)\right)_{y\in\bar{Y}}$ is sufficiently large, which we will see is equivalent to the χ^2-divergence being sufficiently large. The Euclidean distance requirement in turn relates to quadratic approximation arguments in the spirit of FLM, KM, and SS.

[5] 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 $\sum_y p\left(y|a\right)\left(\frac{\partial}{\partial a_i} p\left(y|a_i, a_{-i}\right)/p\left(y|a\right)\right)^2$, which is a local χ^2-divergence. It has previously been observed that the 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, cf. their inequality 2.4).

While our proofs rely on χ^2-divergence and our non-asymptotic results are strongest under this measure, our asymptotic results hold equally for other divergences (e.g., total variation distance, Kullback-Leibler divergence) which are equivalent to χ^2-divergence up to constant factors under our assumptions. We discuss this point further in Section 4, following Corollary 1.

3 Necessary Conditions for Cooperation

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 ε-correlated equilibrium payoffs), for

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

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 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).

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 the bound to be tight. 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 $(1-\delta)^{-1/2}$.

Some prior results bound incentives in repeated games as a function of discounting and monitoring precision, but they do so history-by-history, and hence obtain bounds of order $(1-\delta)^{-1}$ (e.g., Fudenberg, Levine, and Pesendorfer, 1998; al-Najjar and Smorodinsky, 2000, 2001; Pai, Roth, and Ullman, 2017). Awaya and Krishna (2016, 2019) derive a bound based on deterring permanent deviations to a particular action, which is also of order $(1-\delta)^{-1}$.[6] 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 some examples.

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.

[6] See, e.g., Proposition 4.1 of Awaya and Krishna (2019). Unlike our bound, their incentive bound for each player i depends only on the marginal of p on Y_{-i}; so, formally, 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}. These monitoring structures play an important role in their analysis. However, it may be possible to use our techniques to tighten their bound; this is left for future research.

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

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

Thus, (1) gives

$$\alpha^{\mu}_{CC}+\alpha^{\mu}_{CD}\leq\frac{\delta}{1-\delta}\frac{(2\pi-1)^{2}}{\pi(1-\pi)}V_{i}\left(\alpha^{\mu}\right),$$

where $V_{i}\left(\alpha^{\mu}\right)=\alpha^{\mu}_{CC}+4\alpha^{\mu}_{DC}+\alpha^{\mu}_{CD}-\left(\alpha^{\mu}_{CC}+2\alpha^{\mu}_{DC}-\alpha^{\mu}_{CD}\right)^{2}$*. This simple bound applies for any Nash equilibrium, whether the signals* (y_1,y_2) *are observed publicly or privately, by either the players or a mediator.*

Example 2 (Prisoner's Dilemma with Poisson Product Structure Monitoring) *Consider the above prisoner's dilemma payoff matrix with payoffs scaled by a positive number* $\Delta>0$*, let* $\delta=e^{-r\Delta}$ *for a constant* $r>0$*, and let* $Y=\{0,1\}\times\{0,1\}$*, where* $\Pr\left(y_i=1|a\right)=\Delta\lambda_{a_i}$ *for constants* $\lambda_C,\lambda_D\in[0,1]$*, independently across players. The interpretation is that the players interact every* Δ *units of time with a fixed real-time discount rate* r *and a fixed underlying Poisson signal intensity.*

Again, we bound the equilibrium probability of cooperation by applying (1) for the manipulation that always defects. The gain from this deviation equals $\left(\alpha^{\mu}_{CC}+\alpha^{\mu}_{CD}\right)\Delta$*, while its detectability equals*

$$\begin{aligned} & \left(\alpha^{\mu}_{CC}+\alpha^{\mu}_{CD}\right)\left(\Delta\lambda_C\left(\frac{\Delta\lambda_C-\Delta\lambda_D}{\Delta\lambda_C}\right)^{2}+(1-\Delta\lambda_C)\left(\frac{\Delta\lambda_C-\Delta\lambda_D}{1-\Delta\lambda_C}\right)^{2}\right) \\ = & \left(\alpha^{\mu}_{CC}+\alpha^{\mu}_{CD}\right)\frac{\Delta\left(\lambda_C-\lambda_D\right)^{2}}{\lambda_C\left(1-\Delta\lambda_C\right)}. \end{aligned}$$

For Δ *sufficiently small so that* $\delta\approx 1-r\Delta$*, (1) gives*

$$\alpha^{\mu}_{CC}+\alpha^{\mu}_{CD}\leq\frac{1}{r}\frac{\left(\lambda_C-\lambda_D\right)^{2}}{\min\left\{\lambda_C,\lambda_D\right\}}V_{i}\left(\alpha^{\mu}\right),$$

where again $V_i(\alpha) = \alpha_{CC} + 4\alpha_{DC} + \alpha_{CD} - (\alpha_{CC} + 2\alpha_{DC} - \alpha_{CD})^2$ *(i.e., payoff variance normalized by* $1/\Delta^2$*).*

Notice that this bound is independent of the interaction frequency $1/\Delta$*. This might appear to contradict the observation of Abreu, Milgrom, and Pearce (1991) that taking* $\Delta \to 0$ *destroys cooperation under "good news" Poisson signals (i.e.,* $\lambda_C > \lambda_D$*). The resolution is that Abreu, Milgrom, and Pearce consider a monitoring structure that violates pairwise identifiability (a single Poisson signal, rather than a Poisson product structure) and, more importantly, restrict attention to strongly symmetric equilibria. The bound implied by Theorem 1 is not tight under these restrictions: that is, our converse result (Theorem 2) does not apply.*

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 the gain from this manipulation in period t and the (inverse of the) detectability of this manipulation in period t (equation (4) 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 must exceed a discounted sum of the conditional variance bounds (equation (5)). Finally, by Jensen's inequality, this inequality relating a discounted sum of payoff variances and a discounted sum of bounds that depend on the deviation gain and detectability of s_i in each period implies a corresponding inequality relating the payoff variance, deviation gain, and detectability evaluated at the equilibrium occupation measure, which simplifies to (1).

We also mention a tighter (but slightly 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

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

Note that

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

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 $\tilde{\chi}_i (s_i, \alpha^\mu)$ in place of $\chi_i (s_i, \alpha^\mu)$, 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, we have

$$\mathrm{Var} \left(W_i^{t+1} \right) = \mathrm{Var} \left(\mathbb{E} \left[W_i^{t+1} | h^t \right] \right) + \mathbb{E} \left[\mathrm{Var} \left(W_i^{t+1} | h^t \right) \right] . \qquad (2)$$

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*

$$\mathrm{Var}\left(\mathbb{E}\left[W_i^{t+1}|h^t\right]\right) \geq \frac{1}{\delta}\mathrm{Var}\left(W_i^t\right) - \frac{1-\delta}{\delta}\mathrm{Var}\left(U_i^t\right) \quad \textit{and} \tag{3}$$

$$\mathbb{E}\left[\mathrm{Var}(W_i^{t+1}|h^t)\right] \geq \left(\frac{1-\delta}{\delta}\right)^2 \frac{g_i\left(\alpha_t^\mu\right)^2}{\chi_i^2\left(\alpha_t^\mu\right)}, \tag{4}$$

where in (4) we follow the convention $0/0 = 0$.

Proof. For (3), since $w_i^t = (1-\delta)\, u_i^t + \delta w_i^{t+1}$, for every history $h^t \in H^t$ we have

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

Therefore,

$$\begin{aligned}
\mathrm{Var}\left(\mathbb{E}\left[W_i^{t+1}|h^t\right]\right) &= \mathrm{Var}\left(\frac{1}{\delta}W_i^t - \frac{1-\delta}{\delta}U_i^t\right) \\
&= \frac{1}{\delta^2}\mathrm{Var}\left(W_i^t\right) + \left(\frac{1-\delta}{\delta}\right)^2 \mathrm{Var}\left(U_i^t\right) - 2\frac{1-\delta}{\delta^2}\mathrm{Cov}\left(U_i^t, W_i^t\right) \\
&\geq \frac{1}{\delta^2}\mathrm{Var}\left(W_i^t\right) + \left(\frac{1-\delta}{\delta}\right)^2 \mathrm{Var}\left(U_i^t\right) - \frac{1-\delta}{\delta^2}\mathrm{Var}\left(U_i^t\right) - \frac{1-\delta}{\delta^2}\mathrm{Var}\left(W_i^t\right) \\
&= \frac{1}{\delta}\mathrm{Var}\left(W_i^t\right) - \frac{1-\delta}{\delta}\mathrm{Var}\left(U_i^t\right).
\end{aligned}$$

For (4), let $\mu\left(h^t, a\right)$ 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

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

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 $\sum_{h^t,a,y} \mu\left(h^t, a\right) p\left(y|s_i\left(a_i\right), a_{-i}\right) W_i^t\left(h^t, a, y\right)$ 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 $s_i\left(a_i\right)$. This continuation play may not be optimal, but we are only giving a necessary condition.)

Therefore,

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

where the second inequality follows from Cauchy-Schwarz. Finally, if $\chi_i^2\left(\alpha_t^\mu\right) > 0$ then squaring both sides and rearranging yields (4); if instead $\chi_i^2\left(\alpha_t^\mu\right) = 0$ then we have $g_i\left(\alpha_t^\mu\right) = 0$, and (4) reduces to $\mathbb{E}\left[\operatorname{Var}(W_i^{t+1}|h^t)\right] \geq 0$, which holds as variance is non-negative. ■

By (2), (3), and (4), for each period t, we have

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

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

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

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

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

Therefore,

$$\delta\sum_t \delta^{t-1}\operatorname{Var}\left(U_i^t\right) \geq (1-\delta)\sum_t \delta^{t-1}\frac{g_i\left(\alpha_t^\mu\right)^2}{\chi_i^2\left(\alpha_t^\mu\right)}. \tag{5}$$

At this point we are almost done, because inequality (5) actually implies the desired inequality, (1). This observation relies on the following lemma.

Lemma 2 *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 $a^2/b = 0$ (by $a = 0$ 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

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

■

We also use the fact that

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

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 (5), we thus have

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

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

4 Sufficient Conditions for Cooperation

This section presents our second result (Theorem 2), which generalizes the classical folk theorems of FLM and KM by letting discounting and monitoring vary simultaneously. Theorem 2 (together with Corollary 1) shows that the tradeoff between discounting and monitoring implied by Theorem 1 is tight in the low-discounting/low-monitoring double limit.

We impose the simplifying assumption that each player has a strict incentive to follow an action profile that maximizes her own payoff.[7]

Assumption 1 For each player i, there exists an action profile $a^i \in \operatorname{argmax}_{a\in A} u_i(a)$ such that $u_i\left(a^i\right) > u_i\left(a_i, a_{-i}^i\right)$ for all $a_i \neq a_i^i$.

Assumption 1 holds for generic payoffs. Imposing this assumption and fixing such an a^i for each i, we let $\varepsilon_u \in (0, \bar{u})$ be such that $u_i\left(a^i\right) > u_i\left(a_i, a_{-i}^i\right) + \varepsilon_u$ for all i and $a_i \neq a_i^i$.

We require some notation. Let $F = \operatorname{co}\left(\{u(a)\}_{a\in A}\right) \subseteq \mathbb{R}^N$ denote the set of feasible payoffs, and let $F^* \subseteq F$ denote the set of feasible payoffs that weakly Pareto-dominate a payoff which is 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 $\sum_n \beta_n = 1$.

For each i and a, let

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

[7] This assumption is helpful because our construction involves continuation payoffs that lie slightly off translated tangent hyperplanes, in contrast to FLM and KM.

That is, $p(a)$ is the vector of signal probabilities at action profile a, and $P_i(a)$ is the set of such vectors that can arise when player i deviates from a_i while the remaining players take a_{-i}, where these probabilities are all normalized by $\sqrt{p(y|a)}$. 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 $x \in \mathbb{R}^{|\bar{Y}|}$ such that $\|x\| = 1$ and
$$x \cdot p(a) > x \cdot p + \sqrt{\eta} \qquad \text{for all } p \in P_i(a) \cup P_j(a). \tag{6}$$

2. There exists a vector $x \in \mathbb{R}^{|\bar{Y}|}$ such that $\|x\| = 1$ and
$$x \cdot p^i - \sqrt{\eta} > x \cdot p(a) > x \cdot p^j + \sqrt{\eta} \qquad \text{for all } p^i \in P_i(a) \text{ and } p^j \in P_j(a). \tag{7}$$

3. We have
$$p(y|a) > \eta \qquad \text{for all } y \in \bar{Y}. \tag{8}$$

When $\eta = 0$, our definition of pairwise identifiability coincides with KM's assumptions (A2) and (A3), which are weaker than FLM's pairwise full rank condition.[8] Thus, η-pairwise identifiability says that KM's assumptions hold with $\sqrt{\eta}$ slack (after normalizing the signal probabilities by $\sqrt{p(y|a)}$). We comment below on the role of (8); for now, note that this requirement has force only together with the other conditions, because with public monitoring one can always identify distinct signals to suitably coarsen the signal space (Kandori, 1992).

We can now state our "folk theorem."

Theorem 2 *Assume that* $\dim F^* = N$*. 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* $\delta > 1 - c\eta$*, we have* $v \in E(Y, p, \delta)$*.*

Theorem 2 extends Theorem 6.1 of FLM and Theorem 1 of KM by relating the minimum discount factor required to conclude that $v \in E(Y, p, \delta)$ to the precision of monitoring. For

[8] When $\eta = 0$, (6) coincides with KM's assumption (A2), and (7) coincides with KM's assumption (A3). KM do not normalize $p(a)$ and $P_i(a)$ by $\sqrt{p(y|a)}$, but this immaterial since in KM the function $x(y)$ can be scaled by $\sqrt{p(y|a)}$ to offset this normalization.

comparison, a version of these earlier results would state that for any $v \in \text{int}F^*$, any $\eta > 0$, and any monitoring structure (Y, p) that satisfies η-pairwise identifiability, there exists $c > 0$ such that, for every $\delta > 1 - c$, we have $v \in E(Y, p, \delta)$. The current result is stronger because c depends only on the target payoff v, rather than depending on both v and the monitoring structure (Y, p) as in FLM and KM.[9]

Theorem 2 is also related to SS's folk theorem for repeated games with frequent actions (their Theorem 2). In their model, signals are parameterized by an underlying continuous-time Lévy process (a mixture of Brownian and full-support Poisson signals), and players interact every Δ units of time, with real-time discount rate r (so $\delta = e^{-r\Delta}$). They define a set M_-, which under pairwise identifiability and individual full rank generically equals the set of feasible and individually rational (FIR) payoffs (SS, Proposition 1; Appendix O-d), and show that for any $v \in M_-$, there exists $\bar{\Delta}$ and $\bar{r}$ such that, for any $\Delta < \bar{\Delta}$ and $r < \bar{r}$, the payoff vector v arises in a PPE. Observe that for Brownian signals (with the space of signal realizations partitioned into arbitrary fixed intervals) we have

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

and for Poisson signals we have

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

Hence, SS's information structure satisfies Δ-pairwise identifiability. Moreover, for small Δ, we have $e^{-r\Delta} > 1 - c\Delta$ iff $r < c$, so Theorem 2 implies SS's result, with the difference that they support all FIR payoffs as PPE, rather than only those payoffs in F^*. As we discuss below, this difference arises because SS's assumptions on the signal structure imply that the likelihood ratio $p(y|a_i', a_{-i}) / p(y|a)$ is bounded, an assumption we do not impose. Our proof builds on SS (as well as FLM and KM). Relative to their result, our main contribution is dispensing with their parameterization by an underlying Lévy process. That is, we

[9] On the other hand, FLM's Theorem 6.1 requires pairwise identifiability only at certain action profiles, and KM's Theorem 1 is a minmax threat folk theorem. (FLM also proved a minmax folk theorem under additional assumptions.) However, our theorem can be similarly extended, as we discuss below.

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.[10]

To compare our Theorems 1 and 2, consider relaxing η-pairwise identifiability to η-*individual identifiability*: the condition that, for any action profile a and player i, there exists a vector $x \in \mathbb{R}^{|\bar{Y}|}$ such that $\|x\| = 1$ and

$$x \cdot p(a) > x \cdot p + \sqrt{\eta} \qquad \text{for all } a \text{ and } p \in P_i(a), \tag{9}$$

and $p(y|a) > \eta$ for all y.[11] Note that (9) is equivalent to $d(p(a), \text{co}(P_i(a))) > \sqrt{\eta}$ for all a (where $d(\cdot, \cdot)$ denotes Euclidean distance), which in turn is equivalent to

$$\chi_i^2(s_i, a) > \eta \qquad \text{for all } a \text{ and all } s_i \text{ with } s_i(a_i)[a_i] = 0.$$

That is, (9) holds if and only if, at any action profile a, the detectability of a deviation to any mixed action that puts zero weight on a_i is at least η. Thus, Theorem 1 implies that if the ratio of the discount rate $1 - \delta$ and detectability is large for all manipulations, then all equilibrium outcomes in the repeated game Γ^B are static ε-correlated equilibria; while Theorem 2 under η-individual identifiability would imply that if the ratio of discounting and detectability is large for all manipulations, then all payoff vectors $v \in \text{int}F^*$ are attainable in PPE in the repeated game Γ^P.

The comparison between Theorems 1 and 2 is tight in games with *public, product structure monitoring*, where the signal y is public and there exist a family of 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

[10] 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.

[11] We define a slightly more permissive version of η-individual identifiability in our companion paper (Sugaya and Wolitzky, 2022a).

$p(y|a) = \prod_{i \in N} p^i(y^i|a^i)$ for all y, a.[12] This result follows because individual identifiability implies pairwise identifiability under product structure monitoring (but not in general). More precisely, we have the following corollary of Theorems 1 and 2.

Corollary 1 *Fix a stage game G satisfying $\dim F^* = N$. For any $\varepsilon > 0$, there exists a constant $k > 0$ such that the following hold:*

1. *For any public, product monitoring structure (Y, p) and any discount factor δ satisfying $\max_{i,s_i,a} \chi_i^2(s_i, a) < (1-\delta)k$, and for any Nash equilibrium outcome μ in the repeated game $\Gamma = (G, Y, p, \delta)$, the induced occupation measure over actions α^μ is a static ε-correlated equilibrium.*

2. *For any public, product monitoring structure (Y, p) and any discount factor δ satisfying $\min_{i,s_i,a:s_i(a_i)[a_i]=0} \chi_i^2(s_i, a) > (1-\delta)/k$ and $p(y|a) > (1-\delta)/k$ for all y, a, 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 $v \in E(Y, p, \delta)$.*

Proof. Part 1 is an immediate implication of Theorem 1. Indeed, Part 1 holds without the assumption that $\dim F^* = N$, without the restriction to public, product structure monitoring, and with Γ^B in place of Γ.

For Part 2, note that if $\min_{i,s_i,a:s_i(a_i)[a_i]=0} \chi_i^2(s_i) > (1-\delta)/k$ and $p(y|a) > (1-\delta)/k$ for all y, a, then $(1-\delta)/k$-individual identifiability holds. Thus, fixing any $c > 0$ and taking $k < c$, we have that η-individual identifiability holds for $\eta = (1-\delta)/k$, and $\delta > 1 - c\eta$. Next, note that η-individual identifiability implies $(\eta/2)$-pairwise identifiability under product structure monitoring. This follows because, by η-individual identifiability and product structure monitoring, for each player i there exists a vector $x^i \in \mathbb{R}^{|\bar{Y}|}$ that satisfies (9) as well as $x^i(y) = x^i(\tilde{y})$ for all $y, \tilde{y}$ such that $y^{-i} = \tilde{y}^{-i}$, and taking $x = x^i/\sqrt{2} - x^j/\sqrt{2}$ satisfies (7) with $\eta/2$ in place of η. To complete the proof, it remains to show that the constant c in the statement of Theorem 2 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

[12] Thus, y^i is a signal of player i's action a_i. This is not to be confused with the signal observed by player i in a general monitoring structure, which we have denoted by y_i.

2, where the constant c is explicitly constructed as a function of the distance between v and the boundary of F^*. ■

If the hypothesis that $p(y|a) > (1-\delta)/k$ in Part 2 of Corollary 1 is strengthened to a uniform lower bound on $p(y|a)$ (independent of δ), then the χ^2-divergence can be replaced with various other divergences in the statement of Corollary 1, 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 $\chi^2 \geq 2\varepsilon KL \geq \varepsilon^2\chi^2$.[13]

For general monitoring structures, there is a gap between Theorems 1 and 2, because we prove Theorem 2 for public monitoring satisfying η-pairwise identifiability rather than private monitoring satisfying η-individual identifiability. We believe that Theorem 2 likely remains valid under η-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. We therefore content ourselves with the public monitoring, pairwise identifiability version of Theorem 2.

We preview the key ideas of the proof of Theorem 2. 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 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, requiring continuation payoffs to lie in B results in a vanishing efficiency loss. In contrast, when discounting and monitoring vary together and $(1-\delta)/\eta \to 0$ (but we may have $\eta \to 0$ as well as $\delta \to 1$), conditions (6) and (7) imply that payoffs vectors in B can be enforced with continuation payoff movements of variance $o(1-\delta)$, while condition (8) additionally implies that the continuation payoff movements can be taken to have magnitude $o(1)$ (but not necessarily $O(1-\delta)$). A key lemma (Lemma 4) shows that under these conditions, requiring continuation payoffs to lie in B again results in vanishing efficiency loss. The intuition is that forcing "large" payoff movements (greater than $O(1-\delta)$)

[13]For inequalities implying these bounds and many more, see, e.g. Sason and Verdú (2016).

into B incurs a significant loss, but that since the continuation payoff movement variance is small, these large movements are infrequent enough that the ex ante expected loss is small.

We can also explain why we give a "Nash threat" rather than a "minmax threat" folk theorem: that is, why we support only payoffs in F^* rather than the entire FIR set. If we try to support payoffs below a Nash threat point, then the "significant loss" from forcing large payoff movements into B mentioned above becomes a "significant gain" for those players whose payoffs are being minimized at the target payoff vector. We just argued that these large continuation payoff movements are infrequent in equilibrium, but they could become much more frequent following a deviation. When a large continuation payoff movement constitutes a gain for some players (e.g., when the target payoff is below a Nash threat point) and in addition the likelihood ratio $p(y|a'_i, a_{-i})/p(y|a)$ is unbounded (so deviating could make a large payoff movement much more likely), such a deviation may be impossible to deter. This issue is the only obstacle to extending our proof to cover payoff vectors below a Nash threat point, so if we imposed the additional assumption that $p(y|a'_i, a_{-i})/p(y|a)$ is bounded for all y, a, a'_i, we could strengthen Theorem 2 to a minmax threat result.[14]

4.1 Proof of Theorem 2

Fix $v \in \text{int}F^*$. Let $\varepsilon_v > 0$ denote the Euclidean distance between v and the boundary of F^*, and let $\varepsilon = \min\{\varepsilon_u, \varepsilon_v\} \in (0, \bar{u})$. 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 $\delta > 1 - c\eta$, then $B \subseteq E(Y, p, \delta)$, and hence $v \in E(Y, p, \delta)$.

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

[14] This issue explains why SS are able to give a minmax threat folk theorem: they consider Poisson signals with bounded likelihood ratios, as well as Brownian signals which are truncated to give bounded likelihood ratios in their equilibrium construction.

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 3 *Any bounded, self-generating set W is contained in $E(Y, p, \delta)$.*

Our key lemma (Lemma 4) will provide a sufficient condition for B to be self-generating, and hence contained in $E(Y, p, \delta)$. It is based on the following definition, where we let $\|\cdot\|$ denote the Euclidean norm, let $\Lambda = \{\lambda \in \mathbb{R}^N : \|\lambda\| = 1\}$, and for each $\lambda \in \Lambda$, let $\|\lambda_+\| = \sqrt{\sum_{\lambda_n > 0} (\lambda_n)^2}$ 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 all i and $a_i \notin$* supp(α_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):

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

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; henceforth FL). The idea of tightening IC and bounding rewards is inspired by SS. The following is our key lemma:

Lemma 4 *If there exists $X > 0$ such that*

$$k(\lambda, X) \geq \max_{v' \in B} \lambda \cdot v' + \frac{\varepsilon}{4} \quad \textit{for all } \lambda \in \Lambda, \quad \textit{and} \tag{10}$$

$$\max\left\{X, N\bar{u}^2\right\} \leq \frac{\delta}{1-\delta}\frac{\varepsilon^2}{2^{12}}, \tag{11}$$

then B is self-generating.

FL 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. The logic is that B is locally linear and thus accommodates arbitrarily large continuation value transfers when $\delta \to 1$. Lemma 4 extends FL's 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 reward $x(y)$ 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 $\delta > 1 - c\eta$, then there exists $X > 0$ that satisfies (10) and (11). 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 $i(\lambda) \in \text{argmax}_{n \in I} \lambda_n$ 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 5 *Let $\underline{\lambda} > 0$ satisfy*

$$N\bar{u} \max\left\{\underline{\lambda}, \frac{1 - \sqrt{1 - N\underline{\lambda}^2}}{\sqrt{1 - N\underline{\lambda}^2}}\right\} \leq \frac{\varepsilon}{16}. \tag{12}$$

For all $\lambda \in \Lambda$,

1. *If $m(\lambda) \leq \underline{\lambda}$, then there exists a static Nash equilibrium α^{NE} such that*

$$\lambda \cdot u\left(\alpha^{NE}\right) \geq \max_{v' \in B} \lambda \cdot v' + \varepsilon/4. \tag{13}$$

2. If $m(\lambda) \geq M(\lambda)/\underline{\lambda}$, then

$$\lambda \cdot u\left(a^{i(\lambda)}\right) \geq \max_{v' \in B} \lambda \cdot v' + \varepsilon/4. \tag{14}$$

Intuitively, in the first case all Pareto weights λ_n are small or negative, so the maximum score is approximated by a static NE; and in the second case player $i(\lambda)$'s Pareto weight is close to 1, so the maximum score is approximated by $a^{i(\lambda)}$.

Now we fix the constants

$$\bar{X} = \max\left\{\frac{8N^2\bar{u}^2}{\underline{\lambda}^4}, 1\right\} \quad \text{and} \quad c = \frac{\varepsilon^2\underline{\lambda}^4}{2^{16}N^2\bar{u}^2}.$$

Lemma 6 *If $\eta < 1$ and $\delta > 1 - c\eta$, then*

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

Proof. Note that $c < \min\left\{\frac{1}{2^{16}}, \frac{\varepsilon^2}{2^{13}N\bar{u}^2}\right\}$, as $\varepsilon < \bar{u}$, $\underline{\lambda} < 1$, and $N \geq 1$. Hence, we have $\delta > 1 - c\eta > 1 - c > 1 - \min\left\{\frac{1}{2^{16}}, \frac{\varepsilon^2}{2^{13}N\bar{u}^2}\right\}$, and so $\frac{\bar{X}}{\eta} \leq \frac{\bar{X}c}{1-\delta} = \max\left\{\frac{\varepsilon^2}{(1-\delta)2^{13}}, \frac{c}{1-\delta}\right\} \leq \frac{\delta\varepsilon^2}{(1-\delta)2^{12}}$, and $N\bar{u}^2 \leq \frac{\varepsilon^2}{(1-\delta)2^{13}} \leq \frac{\delta\varepsilon^2}{(1-\delta)2^{12}}$. ■

We henceforth assume that (Y, p) satisfies η-pairwise identifiability and $\delta > 1 - c\eta$. By Lemmas 4 and 6, to complete the proof it suffices to show that $k\left(\lambda, \bar{X}/\eta\right) \geq \max_{v' \in B} \lambda \cdot v' + \varepsilon/4$ 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 $4\bar{u}^2/\eta$ 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 conditions (6) and (7).

Lemma 7 *For each pair of players $i \neq j$ and $a \in A$, there exist $\left(x_i^{j,-}(y;a)\right)_{y \in \bar{Y}}$ such that*

$$\sum_y p(y|a)\, x_i^{j,-}(y;a) = 0, \tag{15}$$

$$\sum_y p(y|a_i', a_{-i})\, x_i^{j,-}(y;a) \leq -2\bar{u} \quad \textit{for all } a_i' \neq a_i, \tag{16}$$

$$\sum_y p\left(y|a_j', a_{-j}\right) x_i^{j,-}(y;a) \leq -2\bar{u} \quad \textit{for all } a_j' \neq a_j, \quad \textit{and} \tag{17}$$

$$\sum_y p(y|a)\, x_i^{j,-}(y;a)^2 \leq \frac{4\bar{u}^2}{\eta}; \tag{18}$$

and there exist $\left(x_i^{j,+}(y|a)\right)_{y \in \bar{Y}}$ such that

$$\sum_y p(y|a)\, x_i^{j,+}(y;a) = 0, \tag{19}$$

$$\sum_y p(y|a_i', a_{-i})\, x_i^{j,+}(y;a) \leq -2\bar{u} \quad \textit{for all } a_i' \neq a_i, \tag{20}$$

$$\sum_y p\left(y|a_j', a_{-j}\right) x_i^{j,+}(y;a) \geq 2\bar{u} \quad \textit{for all } a_j' \neq a_j, \quad \textit{and} \tag{21}$$

$$\sum_y p(y|a)\, x_i^{j,+}(y;a)^2 \leq \frac{4\bar{u}^2}{\eta}. \tag{22}$$

Finally, we show that $k\left(\lambda, \bar{X}/\eta\right) \geq \max_{v' \in B} \lambda \cdot v' + \varepsilon/4$ for all $\lambda \in \Lambda$. As in FLM, we partition Λ into three cases: (1) $m(\lambda) \leq \underline{\lambda}$; (2) $m(\lambda) \geq M(\lambda)/\underline{\lambda}$; and (3) $\underline{\lambda} < m(\lambda) < M(\lambda)/\underline{\lambda}$. In Case 1, we take α to be a static Nash equilibrium that satisfies (13) and set $x(y) = 0$ for all y. In Case 2, for $i = i(\lambda)$, we take $\alpha = a^i$ and set $x(y)$ so that players $n \neq i$ have correct incentives and player i's reward "balances the budget" (i.e., satisfies $\sum_{n \in I} \lambda_n x_n(y) = 0$ for all y). In Case 3, we take any $\alpha \in \operatorname{argmax}_a \lambda \cdot u(a)$, and for $i = i(\lambda)$ and for any $j \neq i$ such that $\lambda_i / |\lambda_j| < 1/\underline{\lambda}$, we set $x(y)$ so that all players have correct incentives and players i and j's rewards balance the budget. The details are relatively straightforward and are deferred to the appendix.

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 seen some implications of our results for repeated games with frequent actions. We showed how these games can be viewed as an instance of a general low-discounting/low-monitoring double limit. When specialized to frequent-actions games, our folk theorem extends SS's by allowing more than two players and unbounded likelihood ratios. We also note that the standard frequent-action limit—where the interaction frequency $1/\Delta \to \infty$ for a fixed real-time discount rate $r > 0$, with signals parameterized by an underlying Lévy process—corresponds to the edge case between our positive and negative theorems, where discounting and monitoring vanish at the same rate. This edge case is interesting and important, but also perhaps somewhat special and detail-dependent.

Another type of low-discounting/low-monitoring double limit arises in large-population repeated games, where many patient players are monitored by an aggregate signal, which conveys 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 distributions to signal distributions, and by Fudenberg, Levine, and Pesendorfer (1996) 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* (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 can also be extended to show that, for any fixed imperfect monitoring structure, the Nash equilibrium payoff set cannot converge to the boundary of the FIR payoff set at a rate faster than $(1-\delta)^{1/2+\varepsilon}$ for any $\varepsilon > 0$. (With $\varepsilon = 0$, this is known to be the rate of convergence for PPE with imperfect public monitoring). This rate-of-convergence bound holds even if one allows private strategies and private monitoring, which answers in the negative a question posed by Hörner and Takahashi (2016). Moreover, by accounting for monitoring precision as well as discounting, this bound can be refined to show that the distance between the equilibrium payoff set and the boundary of the FIR payoff set must exceed $\left((1-\delta)/\max_{i,s_i}\chi_i^2(s_i)\right)^{1/2+\varepsilon}$. 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).

Finally, our results also apply to repeated principal-agent problems: that is, two-player games where one player's strategy is fixed in advance. Sadzik and Stacchetti (2015) provide a detailed analysis of repeated principal-agent problems with one-dimensional actions and concave preferences in the frequent-action limit (the edge case between our positive and negative). Our results complement theirs by providing necessary and sufficient conditions for cooperation in a more general class of games, where such detailed analysis may be infeasible.

A Appendix: Omitted Steps in Proof of Theorem 2

A.1 Proof of Lemma 4

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 8 *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

$$\lambda \cdot \left(u(\alpha) + \sum_y p(y|\alpha) x(y) \right) \geq \max_{v' \in B} \lambda \cdot v' + \varepsilon/5 \geq \max_{v' \in B'} \lambda \cdot v' + \varepsilon/5. \tag{23}$$

Fix any such α and x. Define

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

Note that, by (HSX) and (11), we have

$$\frac{1-\delta}{\delta} \sqrt{X_y} \leq \frac{\varepsilon}{64} \tag{24}$$

and

$$\frac{1-\delta}{\delta} \sum_y p(y|\alpha) X_y \leq \frac{\varepsilon^2}{640}. \tag{25}$$

To construct w, let

$$\xi_i(y) = -\frac{64\lambda_i X_y}{\varepsilon}\mathbf{1}\{\lambda_i \geq 0\} \qquad \text{for all } i, y, \tag{26}$$

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

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

Here the first term in parentheses captures orthogonal continuation payoff movements with respect to normal vector λ (as in FL), while the second term is an adjustment that will keep $w(y)$ in B' even when the orthogonal component of $w(y) - \hat{v}$ is large.

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 $\sum_y p(y|\alpha)\, w(y) = (1/\delta)(\hat{v} - (1-\delta)\, u(\alpha))$, and hence $(1-\delta)\, u(\alpha) + \delta \sum_y p(y|\alpha)\, w(y) = \hat{v}$.

(IC): Setting aside the constant terms in $w(y)$, we see that an action a_i maximizes $(1-\delta)\, u_i(a_i, \alpha_{-i}) + \delta \sum_y p(y|a_i, \alpha_{-i})\, w_i(y)$ iff it maximizes $u_i(a_i, \alpha_{-i}) + \sum_y p(y|a_i, \alpha_{-i}) \left(x_i(y) + \frac{1-\delta}{\delta}\xi_i(y)\right)$. Now note that, for all $a_i \notin \operatorname{supp} \alpha$, we have

$$\begin{aligned}
& u_i(\alpha) + \sum_y p(y|\alpha)\left(x_i(y) + \frac{1-\delta}{\delta}\xi_i(y)\right) - u_i(a_i, \alpha_{-i}) - \sum_y p(y|a_i, \alpha_{-i})\left(x_i(y) + \frac{1-\delta}{\delta}\xi_i(y)\right) \\
\geq \; & \varepsilon \mathbf{1}\{\lambda_i \geq 0\} + \sum_y p(y|\alpha)\frac{1-\delta}{\delta}\xi_i(y) \qquad \text{by (IC}\varepsilon\text{) and } \xi_i(y) \leq 0 \; \forall y \\
\geq \; & \mathbf{1}\{\lambda_i \geq 0\}\left(\varepsilon - \frac{64\lambda_i}{\varepsilon}\frac{1-\delta}{\delta}\sum_y p(y|\alpha)\, X_y\right) \qquad \text{by (26)} \\
\geq \; & 0 \qquad \text{by (25) and } \lambda_i \leq 1.
\end{aligned}$$

This establishes (IC).

(SG): We start with a standard geometric observation: if a payoff vector $w \in \mathbb{R}^N$ satisfies $\lambda \cdot (\hat{v} - w) \geq 0$ and the Euclidean distance between $\hat{v}$ and w is sufficiently small compared to the distance between $\hat{v}$ and w in direction λ, then $w \in B'$.

Lemma 9 *If* $w \in \mathbb{R}^N$ *satisfies* $\lambda \cdot (\hat{v} - w) \geq 0$ *and*

$$\|\hat{v} - w\| \leq \sqrt{\frac{\varepsilon}{4}\lambda \cdot (\hat{v} - w)}, \tag{27}$$

then $w \in B'$.

Proof. (27) 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 $\|x\|^2 = \|\hat{v} - w\|^2 - (\lambda \cdot (\hat{v} - w))^2 \leq \|\hat{v} - w\|^2$, (27) 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 (27). Note that

$$\begin{aligned}
\hat{v} - w(y) &= \frac{1-\delta}{\delta}\Delta(y) - \left(\frac{1-\delta}{\delta}\right)^2 \xi(y) + \left(\frac{1-\delta}{\delta}\right)^2 \sum_{y'} p(y'|\alpha)\,\xi(y'), \\
\text{where } \Delta(y) &= u(\alpha) - \hat{v} + \sum_{y'} p(y'|\alpha)\,x(y') - x(y).
\end{aligned}$$

By (HSX) and (23),

$$\begin{aligned}
\lambda \cdot \Delta(y) &\geq \frac{1-\delta}{\delta}\frac{\varepsilon}{5}, \qquad \text{and} \\
\|\Delta(y)\| &\leq \|u(\alpha) - \hat{v}\| + \left\|\sum_{y'} p(y'|\alpha)\,x(y') - x(y)\right\| \leq \sqrt{N}\bar{u} + \|\lambda_+\|\sqrt{\frac{X_y}{\varepsilon}}.
\end{aligned}$$

By (HSX) and the definition of ξ (cf. (26)),

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

Therefore,

$$
\begin{aligned}
\lambda \cdot (\hat{v} - w(y)) &= \frac{1-\delta}{\delta} \lambda \cdot \Delta(y) - \left(\frac{1-\delta}{\delta}\right)^2 \lambda \cdot \xi(y) + \left(\frac{1-\delta}{\delta}\right)^2 \lambda \cdot \left(\sum_{y'} p(y'|\alpha)\, \xi(y') \right) \\
&\geq \frac{1-\delta}{\delta} \frac{\varepsilon}{5} + \left(\frac{1-\delta}{\delta}\right)^2 \|\lambda_+\|^2 \frac{64 X_y}{\varepsilon} - \left(\frac{1-\delta}{\delta}\right)^2 \|\lambda_+\| \frac{64}{\varepsilon} \sum_y p(y|\alpha) X_y \\
&\geq \frac{1-\delta}{\delta} \frac{\varepsilon}{10} + \left(\frac{1-\delta}{\delta}\right)^2 \|\lambda_+\|^2 \frac{64 X_y}{\varepsilon} \geq 0, \qquad (28)
\end{aligned}
$$

where the last line follows since $\|\lambda_+\| \leq 1$ and (25) imply that $\frac{1-\delta}{\delta} \|\lambda_+\| \frac{64}{\varepsilon} \sum_y p(y|\alpha) X_y \leq \frac{\varepsilon}{10}$. Thus, we have

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

Similarly, we have

$$
\begin{aligned}
& \|\hat{v} - w(y)\| \\
\leq\ & \frac{1-\delta}{\delta} \|\Delta(y)\| + \left(\frac{1-\delta}{\delta}\right)^2 \|\xi(y)\| + \left(\frac{1-\delta}{\delta}\right)^2 \left\| \textstyle\sum_{y'} p(y'|\alpha)\, \xi(y') \right\| \\
\leq\ & \frac{1-\delta}{\delta} \left(\sqrt{N} \bar{u} + \|\lambda_+\| \sqrt{\frac{X_y}{\varepsilon}} \right) + \left(\frac{1-\delta}{\delta}\right)^2 \|\lambda_+\| \frac{64 X_y}{\varepsilon} + \left(\frac{1-\delta}{\delta}\right)^2 \|\lambda_+\| \frac{64}{\varepsilon} \sum_y p(y|\alpha) X_y \\
\leq\ & 2 \frac{1-\delta}{\delta} \left(\sqrt{N} \bar{u} + \|\lambda_+\| \sqrt{X_y} \right) \leq 4 \frac{1-\delta}{\delta} \max \left\{ \sqrt{N} \bar{u}, \|\lambda_+\| \sqrt{X_y} \right\}, \qquad (29)
\end{aligned}
$$

where the third inequality follows since $\|\lambda_+\| \leq 1$ and (25) imply that $\frac{1-\delta}{\delta} \|\lambda_+\| \frac{64}{\varepsilon} \sum_y p(y|\alpha) X_y \leq 1$, and (24) implies that $\frac{1-\delta}{\delta} \frac{64 X_y}{\varepsilon} \leq \sqrt{X_y}$.

Comparing (28) and (29), we see that $w(y)$ satisfies (27) whenever $\sqrt{N} \bar{u} \leq \sqrt{\frac{1}{640} \frac{\delta}{1-\delta}} \varepsilon$, which is implied by (11).

A.2 Proof of Lemma 5

Case 1: $m(\lambda) \leq \underline{\lambda}$. Let $\lambda'_n = \min\{\lambda_n, 0\} / \|\lambda_-\|$ and let $\lambda' = (\lambda'_n)_{n \in I} \in \Lambda$. We claim that $\sum_n |\lambda'_n - \lambda_n| \leq \varepsilon / 4\bar{u}$. 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

$$\sum_n |\lambda'_n - \lambda_n| \leq N \max \left\{ \underline{\lambda}, \frac{1 - \sqrt{1 - N\underline{\lambda}^2}}{\sqrt{1 - N\underline{\lambda}^2}} \right\} \leq \frac{\varepsilon}{4\bar{u}} \qquad \text{by (12).}$$

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 $\sum_n |\lambda'_n - \lambda_n| \leq \varepsilon / 4\bar{u}$, $|u_i(a)| \leq \bar{u}$ $\forall i, a$, and the distance from B to the boundary of F^* is greater than $\varepsilon/2$, we have $\lambda \cdot u\left(\alpha^{NE}\right) \geq \lambda' \cdot u\left(\alpha^{NE}\right) - \varepsilon/4 \geq \max_{v' \in F^*} \lambda' \cdot v' - \varepsilon/4 \geq \max_{v' \in B} \lambda' \cdot v' + \varepsilon/4$, establishing (13).

Case 2: $m(\lambda) \geq M(\lambda) / \underline{\lambda}$. 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 $\|\lambda\| = 1$). By (12) and $\varepsilon \leq 1$, this implies that $|\lambda_i| \geq 3/4$, which, since $i \in \operatorname{argmax} \lambda_n$, implies that $\lambda_i > 0$, and hence $\lambda_i \geq 1 - N\underline{\lambda}$. Since $|u_i(a)| \leq \bar{u}$ $\forall i, a$, we have, for all $v' \in F^*$ and $\lambda \in \Lambda$, $|(\lambda - e_i) \cdot v'| \leq \sum_n |\lambda_n - e_{i,n}| \bar{u} \leq ((N-1)\underline{\lambda} + |(1 - N\underline{\lambda}) - 1|) \bar{u} \leq 2N\underline{\lambda}\bar{u}$. Therefore, for $a^i \in \arg\max_{a \in A} e_i \cdot u(a)$, we have $\lambda \cdot u(a^i) \geq e_i \cdot u(a^i) - 2N\underline{\lambda}\bar{u} \geq \max_{v' \in F^*} e_i \cdot v' - 2N\underline{\lambda}\bar{u} \geq \max_{v' \in F^*} \lambda \cdot v' - 4N\underline{\lambda}\bar{u} \geq \max_{v' \in B} \lambda \cdot v' + \varepsilon/2 - 4N\underline{\lambda}\bar{u} \geq \max_{v' \in B} \lambda \cdot v' + \varepsilon/4$, where the last inequality is by (12).

A.3 Proof of Lemma 7

Fix any i, j, and a. We first construct $\left(x_i^{j,-}(y;a)\right)_{y\in\bar{Y}}$. Let $x \in \mathbb{R}^{|\bar{Y}|}$ satisfy $\|x\| = 1$ and (6). By definition of P, P_i, and P_j, we have

$$\begin{aligned}
\left(\frac{p(y|a) - p(y|a_i', a_{-i})}{\sqrt{p(y|a)}}\right)_y \cdot (x(y))_y &\geq \sqrt{\eta} \quad \text{for all } a_i' \neq a_i, \quad \text{or equivalently} \\
(p(y|a) - p(y|a_i', a_{-i}))_y \cdot \left(\frac{x(y)}{\sqrt{p(y|a)}}\right)_y &\geq \sqrt{\eta} \quad \text{for all } a_i' \neq a_i; \quad \text{and similarly} \\
\left(p(y|a) - p\left(y|a_j', a_{-j}\right)\right)_y \cdot \left(\frac{x(y)}{\sqrt{p(y|a)}}\right)_y &\geq \sqrt{\eta} \quad \text{for all } a_j' \neq a_j.
\end{aligned}$$

Defining

$$x_i^{j,-}(y;a) = 2\bar{u}\left(\frac{x(y)}{\sqrt{p(y|a)\,\eta}} - \sum_{\tilde{y}} p(\tilde{y}|a)\frac{x(\tilde{y})}{\sqrt{p(\tilde{y}|a)\,\eta}}\right) \quad \text{for all } y,$$

conditions (15)–(17) hold by construction, and condition (18) holds because

$$\sum_y p(y|a)\left(x_i^{j,-}(y;a)\right)^2 = \frac{4\bar{u}^2}{\eta}\sum_y p(y|a)\left(\frac{x(y)}{\sqrt{p(y|a)}} - \sum_{\tilde{y}} p(\tilde{y}|a)\frac{x(\tilde{y})}{\sqrt{p(\tilde{y}|a)}}\right)^2 = \frac{4\bar{u}^2}{\eta}\sum_y x(y)^2 = \frac{4\bar{u}^2}{\eta}.$$

To construct $\left(x_i^{j,+}(y;a)\right)_{y\in\bar{Y}}$, let $x \in \mathbb{R}^{|\bar{Y}|}$ satisfy $\|x\| = 1$ and (7), and proceed as in the construction of $\left(x_i^{j,-}(y;a)\right)_{y\in\bar{Y}}$.

A.4 Completion of the Proof of Theorem 2

We show that $k\left(\lambda, \bar{X}/\eta\right) \geq \max_{v'\in B} \lambda \cdot v' + \varepsilon/4$ for all $\lambda \in \Lambda$.

Case 1: $m(\lambda) \leq \underline{\lambda}$. For α^{NE} that satisfies (13), 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$X/\eta$).

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

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

where $\text{sign}(z) = -$ for $z \leq 0$ and $\text{sign}(z) = +$ for $z > 0$. 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 (14). It remains to verify (ICε) and (HS$X/\eta$).

For (ICε), for player i, note that $\sum_y p(y|a^i)\, x_i(y) = 0$ by (16) and (20), 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 (17) and (21). Together with $u_i(a^i) - u_i\left(a_i, a^i_{-i}\right) \geq \varepsilon \ \forall a_i \neq a^i_i$, this implies (IC$\varepsilon$) for player i. Next, for any player $n \neq i$, (16) and (20) imply that the expected reduction in continuation payoff from deviating is at least $2\bar{u}$, which exceeds the maximum difference between any two stage game payoffs by at least $\bar{u} > \varepsilon$.

For (HS$X/\eta$), since $|\lambda_n| / \lambda_i \leq M(\lambda) / m(\lambda) \leq \underline{\lambda} \ \forall n$ (by hypothesis), $\|\lambda_+\| \geq 1 - N\underline{\lambda} \geq 1 - \varepsilon/16$ (arguing as in Case 1 of the proof of Lemma 5 and applying (12)), and

$$\sum_y p\left(y|a^i\right) \left(x_n^{i,\text{sign}(\lambda_i \lambda_n)}\left(y|a^i\right) - \sum_{\tilde{y}} p\left(\tilde{y}|a^i\right) x_n^{i,\text{sign}(\lambda_i \lambda_n)}\left(\tilde{y}; a^i\right) \right)^2 \leq \frac{4\bar{u}^2}{\eta} \quad \forall n \tag{30}$$

(by (15), (18), (19) and (22)), we have

$$\begin{aligned} \sum_y \frac{p(y|a) \left\| x(y) - \sum_{\tilde{y}} p(\tilde{y}|a)\, x(\tilde{y}) \right\|^2}{\|\lambda_+\|^2} &\leq \frac{4\bar{u}^2 \left(\left(\sum_{n \neq i} \lambda_n / \lambda_i \right)^2 + N - 1 \right)}{\eta \|\lambda_+\|^2} \\ &\leq \frac{4\bar{u}^2 \left((N-1)^2 \underline{\lambda}^2 + N - 1 \right)}{\eta (1 - \varepsilon/16)^2} \\ &\leq \frac{8\bar{u}^2 N^2}{\eta \underline{\lambda}^4} = \frac{\bar{X}}{\eta}. \end{aligned}$$

Since $p(y|a) \geq \eta$ for all y, we have

$$\frac{\left\| x(y) - \sum_{\tilde{y}} p(\tilde{y}|a)\, x(\tilde{y}) \right\|^2}{\|\lambda_+\|^2} \leq \frac{\bar{X}}{\eta^2} \Rightarrow \frac{\left\| x(y) - \sum_{\tilde{y}} p(\tilde{y}|a)\, x(\tilde{y}) \right\|}{\|\lambda_+\|} \leq \frac{\sqrt{\bar{X}}}{\eta} \leq \frac{\bar{X}}{\eta}.$$

Case 3: $\underline{\lambda} < m(\lambda) < M(\lambda) / \underline{\lambda}$. Fix any $a^\lambda \in \text{argmax}_a \lambda \cdot u(a)$. 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

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

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, (16) and (20) imply that the expected reduction in continuation payoff from deviating is at least $2\bar{u}$, which exceeds the maximum difference between any two stage game payoffs by at least $\bar{u} > \varepsilon$. Finally, (HS$X/\eta$) is satisfied because, since $|\lambda_n| / \lambda_i \leq 1/\underline{\lambda} \; \forall n$, $\lambda_i / |\lambda_j| \leq 1/\underline{\lambda}$, $\|\lambda_+\|^2 \geq \underline{\lambda}^2$ (as $\lambda_i > \underline{\lambda}$), and (30) holds with a^λ in place of a^i, we have

$$\begin{aligned} \sum_y \frac{p(y|a) \left\| x(y) - \sum_{\tilde{y}} p\left(\tilde{y}|a^\lambda\right) x(\tilde{y}) \right\|^2}{\|\lambda_+\|^2} &\leq \frac{4\bar{u}^2 \left(\left(1 - \sum_{n \neq i} \lambda_n / \lambda_i\right)^2 + (1 - \lambda_i/\lambda_j)^2 + N - 2 \right)}{\eta \|\lambda_+\|^2} \\ &\leq \frac{4\bar{u}^2 \left((1 + (N-1)/\underline{\lambda})^2 + (1 + 1/\underline{\lambda})^2 + N - 2 \right)}{\eta \underline{\lambda}^2} \\ &\leq \frac{8\bar{u}^2 N^2}{\eta \underline{\lambda}^4} = \frac{\bar{X}}{\eta}. \end{aligned}$$

Since $p(y|a) \geq \eta$ for all y, we have

$$\frac{\left\| x(y) - \sum_{\tilde{y}} p(\tilde{y}|a) x(\tilde{y}) \right\|^2}{\|\lambda_+\|^2} \leq \frac{\bar{X}}{\eta^2} \Rightarrow \frac{\left\| x(y) - \sum_{\tilde{y}} p(\tilde{y}|a) x(\tilde{y}) \right\|}{\|\lambda_+\|} \leq \frac{\sqrt{\bar{X}}}{\eta} \leq \frac{\bar{X}}{\eta}.$$

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] 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.

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

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

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

[11] 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.

[12] 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.

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

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

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

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

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

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

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

[20] Kandori, Michihiro. "The Use of Information in Repeated Games with Imperfect Monitoring." *Review of Economic Studies* 59.3 (1992): 581-593.

[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, and Andrzej Skrzypacz. "Impossibility of Collusion under Imperfect Monitoring with Flexible Production." *American Economic Review* 97.5 (2007): 1794-1823.

[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).