

Rate of Convergence in Repeated Games: A Universal Speed Limit

Takuo Sugaya

Alexander Wolitzky

Stanford GSB

MIT

July 14, 2022

Abstract

Prior work has established that the rate of convergence of the perfect, public equilibrium (PPE) payoff set to its limit in repeated games with full-support public monitoring is at most $(1 - \delta)^{1/2}$, and that this rate is tight for strictly individually rational payoff vectors. We show that essentially the same bound applies for the *Nash* equilibrium payoff set under *any* full-support imperfect monitoring structure, be it public or private. Consequently, while private strategies or monitoring can outperform PPE for a fixed discount factor, they cannot yield a faster rate of convergence to the efficient payoff frontier when the folk theorem holds for PPE. The rate-of-convergence bound also extends to the case where the discount factor and the monitoring structure vary simultaneously.

Keywords: repeated games, imperfect monitoring, rate of convergence, martingales

JEL codes: C72, C73

1 Introduction

Repeated games with imperfect monitoring and impatient players are a natural and widely applicable model of long-run relationships. Unfortunately, these games can be rather intractable, and there are few general results beyond the fixed-point characterization of perfect public equilibria (PPE) under public monitoring due to Abreu, Pearce, and Stacchetti (1990, henceforth APS). The patient limit—where the discount factor δ is close to 1—is much more tractable, as shown by the folk theorem for PPE due to Fudenberg, Levine, and Maskin (1994, henceforth FLM). However, in reality players may not be this patient, so it is also useful to consider less extreme approaches that retain some of the tractability of the $\delta \rightarrow 1$ limit. These include analyzing double limits where monitoring degrades simultaneously with $\delta \rightarrow 1$ (as in the “frequent action” limit considered by, e.g., Abreu, Milgrom, and Pearce, 1991; see also Sugaya and Wolitzky, 2022, henceforth SW), and analyzing the rate of convergence of the equilibrium payoff set as $\delta \rightarrow 1$.

The latter question was thoroughly analyzed by Hörner and Takahashi (2016, henceforth HT) for PPE under public monitoring (as well as for perfect monitoring). Building on FLM’s analysis, HT gave an example where the rate of convergence of the PPE payoff set to its limit—which, under the sufficient conditions for the folk theorem, equals the set of all feasible and individually rational payoff vectors—is at most $(1 - \delta)^{1/2}$, and they showed that, generically, this rate is tight for strictly individually rational payoff vectors.

Left open by HT is the question of whether the “speed limit” of $(1 - \delta)^{1/2}$ is a special feature of PPE or public monitoring, or whether it is instead a more fundamental property of imperfect-monitoring repeated games. As it is well-known that sequential equilibria where players use private strategies can outperform PPE for a fixed discount factor (e.g., Kandori and Obara, 2006), HT speculated that the former might be the case, writing (p. 357), “It is certainly possible that regarding imperfect monitoring, allowing equilibria in private strategies could accelerate the rate of convergence beyond the results that we have derived. . . This is left for future research.”

The current paper shows that, in fact, essentially the same rate of $(1 - \delta)^{1/2}$ is an upper bound on the rate of convergence of the *Nash* equilibrium payoff set under *any* full-support imperfect monitoring structure. (Full support is clearly necessary for this result—e.g., for perfect monitoring, the limit payoff set is often exactly attained at $\delta < 1$.) More precisely,

we show that for any pure action profile a^* such that $u(a^*)$ is an extreme point of the feasible payoff set and a^* is not a stage-game Nash equilibrium, and for any constant $\varepsilon > 0$, the rate of convergence of the repeated-game Nash equilibrium payoff set toward the point $u(a^*)$ is at most $(1 - \delta)^{1/2+\varepsilon}$. A remarkable implication of this result is that, while private strategies can outperform PPE for a fixed discount factor, they cannot yield a faster rate of convergence to the efficient payoff frontier when the folk theorem holds for PPE. That is, when PPE suffice for the folk theorem, they also give the best rate of convergence.

We actually prove a more general version of this “speed limit” theorem, where we let the discount factor and the monitoring structure vary simultaneously. (The proof is not too much harder than the one for the standard limit where $\delta \rightarrow 1$ for fixed monitoring.) Our main point is doing so is to reinforce the idea introduced in our companion paper, SW, that the relevant measure of discounting in repeated games is often not the discount rate $1 - \delta$ per se, but rather the ratio of the discount rate and a notion of “monitoring precision.” The difference between these metrics is that the former measures discounting relative to calendar time, while the latter measures discounting relative to the “intrinsic time” experienced by the martingale with increments equal to the likelihood ratio difference between the signal at equilibrium play as compared to that under a possible deviation. We show that, when discounting and monitoring vary simultaneously, the speed limit of $(discounting)^{1/2+\varepsilon}$ holds for this intrinsic discounting metric. In the special case where $\delta \rightarrow 1$ for fixed monitoring, the discounting metric simplifies to the discount rate $1 - \delta$, implying the $(1 - \delta)^{1/2+\varepsilon}$ bound.

Our proof approach is very different from HT. HT’s analysis relies on the well-known recursive structure of the PPE payoff set uncovered by APS and FLM. That approach does not apply once private strategies or monitoring are allowed, as the equilibrium payoff set no longer has a tractable recursive structure (Kandori, 2002). Our approach is instead based on general considerations concerning the amount of variation in continuation payoffs required to provide incentives for non-static Nash play. We introduced related ideas in SW; here, we extend these ideas and combine them with martingale methods. We discuss this approach in more detail following the theorem statement.

2 The “Speed Limit”

A repeated game (G, Y, p, δ) consists of: (1) a finite stage game $G = (I, A, u)$, where $I = \{1, \dots, N\}$ is a finite set of players, $A = \times_{i \in I} A_i$ is a finite product set of actions, and $u_i : A \rightarrow \mathbb{R}$ is player i 's payoff function; (2) a monitoring structure (Y, p) , where Y is a finite set of signal realizations and $p(y|a)$ is probability of signal realization y given action profile a ; and (3) a discount factor $\delta \in (0, 1)$. We assume that monitoring satisfies full support: $p(y|a) > 0$ for all y, a .

In each period $t = 1, 2, \dots$, players take actions a , a signal y is drawn from $p(y|a)$, and the players observe something about y . If all players observe y , the game has *public monitoring*, but we need not assume this. For example, our theorem applies equally if y is a vector $(y_i)_{i \in I}$ and each player i observes only the signal component y_i (in which case the game has *private monitoring*); or if y is observed only by a mediator who then sends arbitrary private messages to the players (in which case we have what we called a *blind game* in SW). A player's strategy is a mapping from histories of her own past actions and observations to distributions over her current-period action.

Let $\text{co}(\cdot)$ denote convex hull, and let $d(\cdot, \cdot)$ denotes Euclidean distance in \mathbb{R}^N . We prove the following:

Theorem 1 *Fix a stage game G , and fix any action profile $a^* \in A$ such that $u(a^*) \notin \text{co}(\{u(a)\}_{a \neq a^*})$ and a^* is not a stage-game Nash equilibrium. Consider any player i and any action a'_i such that $u_i(a'_i, a^*_{-i}) > u_i(a^*)$. (Such i and a'_i exist as a^* is not a static equilibrium). For any full-support monitoring structure (Y, p) , let*

$$\begin{aligned} \rho &= \max_{a \in A} \left| \frac{p(y|a'_i, a_{-i}) - p(y|a)}{p(y|a)} \right|, \quad \text{and} \\ \chi^2 &= \max_{a \in A} \sum_y p(y|a) \left(\frac{p(y|a'_i, a_{-i}) - p(y|a)}{p(y|a)} \right)^2. \end{aligned}$$

Then, for any $\varepsilon > 0$, there exists $k > 0$ such that, for every triple (Y, p, δ) such that (Y, p) satisfies full support and $\rho + \sqrt{\chi^2 / (1 - \delta)} \geq k$, and for every Nash equilibrium payoff vector v in the repeated game (G, Y, p, δ) , we have

$$d(v, u(a^*)) \geq \left(\rho + \sqrt{\frac{\chi^2}{1 - \delta}} \right)^{-(1+\varepsilon)}.$$

Note that ρ is the maximum likelihood ratio difference between the signal at equilibrium play as compared to that under a deviation to a'_i , and χ^2 is the variance of this likelihood ratio difference: the latter is a standard measure of statistical distance known as the χ^2 -divergence. Theorem 1 implies that, for any sequence of discount factors and monitoring structures satisfying $(1 - \delta) / \chi^2 \rightarrow 0$ (i.e., “discounting vanishes faster than monitoring”) and $\rho \not\rightarrow \infty$ (i.e., “full support holds uniformly”), the rate of convergence of the Nash equilibrium payoff set toward the extreme point $u(a^*)$ is at most $((1 - \delta) / \chi^2)^{1/2+\varepsilon}$.

When the monitoring structure (Y, p) is also held fixed (in addition to the stage game G), Theorem 1 implies the desired upper bound for the rate of convergence as $\delta \rightarrow 1$.

Corollary 1 *Fix a stage game G and a full-support monitoring structure (Y, p) , and fix any action profile $a^* \in A$ such that $u(a^*) \notin \text{co}(\{u(a)\}_{a \neq a^*})$ and a^* is not a stage-game Nash equilibrium. For any $\varepsilon > 0$, there exists $\bar{\delta} < 1$ such that, for every $\delta > \bar{\delta}$ and every Nash equilibrium payoff vector v in the repeated game (G, Y, p, δ) , we have*

$$d(v, u(a^*)) \geq (1 - \delta)^{\frac{1}{2} + \varepsilon}.$$

Proof. By Theorem 1, for any $\varepsilon' \in (0, 2\varepsilon)$, there exists $k > 0$ such that, whenever δ satisfies $\rho + \sqrt{\chi^2 / (1 - \delta)} \geq k$, we have

$$d(v, u(a^*)) \geq \left(\rho + \sqrt{\frac{\chi^2}{1 - \delta}} \right)^{-(1+\varepsilon')} \quad \text{for all } \delta.$$

Since $\varepsilon' < 2\varepsilon$ and ρ and χ^2 are fixed, for all sufficiently high δ , we have

$$\left(\rho + \sqrt{\frac{\chi^2}{1 - \delta}} \right)^{-(1+\varepsilon')} \geq (1 - \delta)^{\frac{1}{2} + \varepsilon},$$

completing the proof. ■

We briefly describe the logic of Theorem 1 and its proof. The basic intuition is that if a Nash equilibrium gives payoffs close to $u(a^*)$, and $u(a^*)$ is an extreme point of the feasible payoff set, then a^* must be played almost all the time along the equilibrium path of play. Since monitoring satisfies full support, this implies that a^* must still be played almost all the time even after low-probability (but still on-path) signal realizations. This in turn implies that, on average, equilibrium continuation play does not vary much with the signal

realization. But then, if a^* is not a static equilibrium, we can conclude that δ must be so high that even small variations in continuation play can provide strong incentives.

The proof of Theorem 1 essentially fills in this intuitive argument with enough precision to yield the desired bound. A first key idea is the notion of the *occupation measure* induced by an equilibrium. Given an equilibrium distribution over paths of action profiles and signals $\mu \in \Delta((A \times Y)^\infty)$ with marginal over period- t action profiles $\alpha_t^\mu \in \Delta(A)$, the corresponding occupation measure $\alpha^\mu \in \Delta(A)$ is defined as

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

We also define the *gain* for player i from deviating to action a_i at action profile distribution α as

$$g_i(a_i, \alpha) = u_i(a_i, \alpha) - u_i(\alpha).$$

With these definitions, it is fairly straightforward to show that, for any equilibrium distribution μ , any player's deviation gain at the occupation measure distribution is bounded by a discounted sum of the product of a likelihood ratio difference and a measure of the variation in payoffs at the occupation measure.¹

Next, since the expected likelihood ratio difference equals 0 on-path—that is,

$$\sum_y p(y|a) \left(\frac{p(y|a) - p(y|a'_i, a_{-i})}{p(y|a)} \right) = 0 \quad \text{for all } a, i, a'_i,$$

where a is an equilibrium action profile and a'_i is a potential deviation—the discounted sum of likelihood ratios

$$\sum_{t=1}^{\infty} \delta^{t-1} \frac{p(y_t|a_t) - p(y_t|a'_i, a_{-i,t})}{p(y_t|a_t)}$$

is a martingale. Recalling that the χ^2 -divergence is the variance of the likelihood ratio difference, we can apply a large deviations bound for martingales (*Freedman's inequality*) to conclude that, for any $x > 0$,

$$\Pr \left(\left| \sum_{t=1}^{\infty} \delta^{t-1} \frac{p(y_t|a_t) - p(y_t|a'_i, a_{-i,t})}{p(y_t|a_t)} \right| \geq x \right) \leq 2 \exp \left(- \frac{x^2}{2 \left(\rho x + \frac{\chi^2}{1-\delta} \right)} \right).$$

¹A related result is the main theorem in SW.

The above arguments imply that any player's deviation gain at any equilibrium occupation measure must be small compared to the product of $\rho + \chi^2/(1 - \delta)$ and the payoff variation at the occupation measure. Now, if equilibrium payoffs are close to some $u(a^*) \notin \text{co}(\{u(a)\}_{a \neq a^*})$, then a^* must be played almost all the time at the occupation measure, so payoff variation is small. Moreover, if a^* is not a static equilibrium, then some player has a positive deviation gain at the occupation measure. Since we have established that this (positive) deviation gain must be small compared to the product of $\rho + \chi^2/(1 - \delta)$ and the (small) payoff variation, we can conclude that $\rho + \chi^2/(1 - \delta)$ must be large.

We mention two possible extensions of Theorem 1. First, while the established rate of convergence in $1 - \delta$ is essentially the best possible, the rate of convergence in ρ likely involves some slack. That is, our bound is probably slack for sequences of monitoring structures that violate full support in the limit. The slack comes from the martingale inequality we apply. Intuitively, the problem is that the martingale (the discounted sum of likelihood ratios) can overshoot a given threshold x by an amount up to ρ , which introduces some slack when bounding the probability that the martingale crosses the threshold. This aspect of our bound could likely be improved at the cost of some additional complexity.

A more significant extension would allow an infinite set of signal realizations with unbounded likelihood ratios: for example, perhaps a player's action is observed with normal noise. In this case, the rate of convergence could be faster than $(1 - \delta)^{1/2}$. The intuition is that, as $\delta \rightarrow 1$, it becomes possible to base incentives on rarer but more informative signal realizations (e.g., "tail tests"), so there is a sense in which increasing δ endogenously increases monitoring precision. We might conjecture that whether this effect actually increases the rate of convergence beyond $(1 - \delta)^{1/2}$ depends on the tail behavior of the signal distribution. We leave this as an open question.

3 Proof of Theorem 1

Fix G , a^* , i , a'_i , and ε as in the statement. For any $\alpha \in \Delta(A)$, let $d^\alpha = d(u(a^*), u(\alpha))$. Note that if v is the payoff vector corresponding to an outcome $\mu \in \Delta((A \times Y)^\infty)$, then $v = u(\alpha^\mu)$ (where α^μ is the occupation measure induced by μ , defined above), so $d(v, u(a^*)) = d^{\alpha^\mu}$. Thus, the desired conclusion can be rephrased as: for every Nash equilibrium outcome μ , we have $d^{\alpha^\mu} \geq \left(\rho + \sqrt{\chi^2/(1 - \delta)}\right)^{-(1+\varepsilon)}$.

Let $\gamma = g_i(a'_i, a^*) > 0$, let $\bar{u} = \max_{a, a'} u_i(a) - u_i(a')$, and let $\eta > 0$ satisfy $d(u(a^*), \text{co}(\{u(a)\}_{a \neq a^*})) = \eta \bar{u}$. (Such η exists as $u(a^*) \notin \text{co}(\{u(a)\}_{a \neq a^*})$.)

We start with a preliminary lemma: for any $\alpha \in \Delta(A)$, both the total variation distance between payoffs under α and a^* and the difference in i 's deviation gain between α and a are bounded by a factor of d^α .

Lemma 1 *For any action profile distribution $\alpha \in \Delta(A)$, we have*

$$\sum_a \alpha(a) |u_i(a) - u_i(a^*)| \leq \frac{1}{\eta} d^\alpha \quad \text{and} \\ g_i(a'_i, \alpha) \geq \gamma - \frac{\gamma + \bar{u}}{\eta \bar{u}} d^\alpha.$$

Proof. The lemma is trivial if $\alpha(a^*) = 1$, so assume that $\alpha(a^*) < 1$. Since $d(u(a^*), \text{co}(\{u(a)\}_{a \neq a^*})) = \eta \bar{u}$, we have

$$d^\alpha = (1 - \alpha(a^*)) d\left(u(a^*), \sum_{a \neq a^*} \frac{\alpha(a)}{1 - \alpha(a^*)} u(a)\right) \geq (1 - \alpha(a^*)) \eta \bar{u},$$

and hence $\alpha(a^*) \geq 1 - d^\alpha / \eta \bar{u}$. Therefore,

$$\sum_a \alpha(a) |u_i(a) - u_i(a^*)| \leq (1 - \alpha(a^*)) \max_{a \neq a^*} |u_i(a) - u_i(a^*)| \leq \frac{d^\alpha}{\eta \bar{u}} \bar{u} = \frac{d^\alpha}{\eta}, \quad \text{and} \\ g_i(a'_i, \alpha) \geq \alpha(a^*) g_i(a'_i, a^*) - (1 - \alpha(a^*)) \bar{u} \geq \left(1 - \frac{d^\alpha}{\eta \bar{u}}\right) \gamma - \frac{d^\alpha}{\eta \bar{u}} \bar{u} = \gamma - \frac{\gamma + \bar{u}}{\eta \bar{u}} d^\alpha.$$

■

Next, let $h = (a^\infty, y^\infty)$ denote an infinite history of actions and signals, and let $v(h)$ denote the realized payoff vector at history h .

Lemma 2 *For any outcome μ , we have*

$$\sum_h \mu(h) |v(h) - u_i(a^*)| \leq \sum_a \alpha^\mu(a) |u_i(a) - u_i(a^*)|.$$

Proof. For each $h = (a^\infty, y^\infty)$, define $\alpha^h \in \Delta(A)$ by

$$\alpha^h(a) = (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} \mathbf{1}\{a_t = a\} \quad \text{for all } a \in A.$$

Note that, for all $a \in A$,

$$\begin{aligned}\alpha^\mu(a) &= (1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} \alpha_t^\mu(a) = (1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} \sum_h \mu(h) \mathbf{1}\{a_t = a\} \\ &= \sum_h \mu(h) (1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} \mathbf{1}\{a_t = a\} = \sum_h \mu(h) \alpha^h(a).\end{aligned}$$

We then have

$$\begin{aligned}\sum_h \mu(h) |v(h) - u_i(a^*)| &= \sum_h \mu(h) |u_i(\alpha^h) - u_i(a^*)| \\ &= \sum_h \mu(h) \left| \sum_a \alpha^h(a) (u_i(a) - u_i(a^*)) \right| \\ &\leq \sum_h \mu(h) \sum_a \alpha^h(a) |u_i(a) - u_i(a^*)| \\ &= \sum_a \sum_h \mu(h) \alpha^h(a) |u_i(a) - u_i(a^*)| = \sum_a \alpha^\mu(a) |u_i(a) - u_i(a^*)|,\end{aligned}$$

where the inequality is by Jensen. ■

Now fix an equilibrium outcome μ , with marginal over period- t action profiles given by $\alpha_t^\mu \in \Delta(A)$. Let $g_t^\mu = g_i(a'_i, \alpha_t^\mu)$ and $g^\mu = g_i(a'_i, \alpha^\mu)$. Denote a history of actions and signals at the beginning of period t by $h^t = (a_{t'}, y_{t'})_{t'=1}^{t-1}$, where h^1 is the trivial history at the beginning of the game. Let

$$\begin{aligned}w_t &= (1-\delta) \sum_{t'=t}^{\infty} \delta^{t'-t} u_{i,t}(a_{t'}), \\ v_t &= \mathbb{E} \left[(1-\delta) \sum_{t'=1}^{\infty} \delta^{t'-1} u_i(a_{t'}) | h^t \right], \quad \text{and} \\ L_t &= \frac{p(y_t | a_t) - p(y_t | a'_i, a_{-i,t})}{p(y_t | a_t)} \mathbf{1}\{g_t^\mu \geq 0\}.\end{aligned}$$

where $\mathbb{E}[\cdot | \cdot]$ denotes conditional expectation under μ . (Thus, v_t and L_t depend on μ ; we suppress this dependence to ease notation.) Note that $v_1 = u(\alpha^\mu)$, the ex ante expected payoff vector under μ ; v_∞ is the realized payoff vector at the infinite history h^∞ ; and $(v_t)_t$ and $(L_t)_t$ are martingales adapted to filtration $(h^t)_t$ (by iterated expectation).

The next lemma uses player i 's period- t incentive constraint to relate g_t^μ , L_t , and $v_\infty -$

$u_i(a^*)$.

Lemma 3 *For every $t \in \mathbb{N}$, we have*

$$(1 - \delta) \delta^{t-1} \max \{g_t^\mu, 0\} \leq \mathbb{E}[L_t(v_\infty - u_i(a^*))].$$

Proof. Since μ is an equilibrium outcome, for every $t \in \mathbb{N}$, we have

$$(1 - \delta) \delta^{t-1} \max \{g_t^\mu, 0\} \leq \mathbf{1} \{g_t^\mu \geq 0\} \times \sum_{h^t, a_t, y_t} \mu(h^t, a_t) (p(y_t|a_t) - p(y_t|a'_i, a_{-i,t})) \delta^t \mathbb{E}[w_{t+1}|h^t, a_t, y_t]. \quad (1)$$

If $g_t^\mu < 0$, this is trivial. If $g_t^\mu \geq 0$, this holds because, if she follows the equilibrium until period t and then takes a'_i , player i can guarantee an expected continuation payoff of $\sum_{h^t, a_t, y_t} \mu(h^t, a_t) p(y_t|a'_i, a_{-i,t}) \mathbb{E}[w_{t+1}|h^t, a_t, y_t]$ by following the mediator's recommendations from period $t+1$ onward, and this deviation must be unprofitable at equilibrium.

Next, note that, by definition, $v_{t+1} - v_t = \delta^t (\mathbb{E}[w_{t+1}|h^t, a_t, y_t] - \mathbb{E}[w_{t+1}|h^t])$. Therefore,

$$\begin{aligned} & \mathbf{1} \{g_t^\mu \geq 0\} \times \sum_{h^t, a_t, y_t} \mu(h^t, a_t) (p(y_t|a_t) - p(y_t|a'_i, a_{-i,t})) \delta^t \mathbb{E}[w_{t+1}|h^t, a_t, y_t] \\ &= \mathbf{1} \{g_t^\mu \geq 0\} \times \sum_{h^t, a_t, y_t} \mu(h^t, a_t) (p(y_t|a_t) - p(y_t|a'_i, a_{-i,t})) \delta^t (\mathbb{E}[w_{t+1}|h^t, a_t, y_t] - \mathbb{E}[w_{t+1}|h^t]) \\ &= \mathbf{1} \{g_t^\mu \geq 0\} \times \sum_{h^t, a_t, y_t} \mu(h^t, a_t) p(y_t|a_t) \frac{p(y_t|a_t) - p(y_t|a'_i, a_{-i,t})}{p(y_t|a_t)} \delta^t (v_{t+1} - v_t) \\ &= \mathbb{E}[L_t(v_{t+1} - v_t)] \\ &= \mathbb{E}[L_t(v_\infty - u_i(a^*))] + \mathbb{E}[L_t(v_{t+1} - \mathbb{E}[v_\infty|h^t, a_t, y_t])] - \mathbb{E}[\mathbb{E}[L_t|h^t, a_t] (v_{t+1} - u_i(a^*))] \\ &= \mathbb{E}[L_t(v_\infty - u_i(a^*))], \end{aligned} \quad (2)$$

where the second-to-last equality follows by iterated expectation, and the last equality follows from $v_{t+1} = \mathbb{E}[v_\infty|h^t, a_t, y_t]$ and $\mathbb{E}[L_t|h^t, a_t] = \mathbf{1} \{g_t^\mu \geq 0\} \times \sum_{y_t} (p(y_t|a_t) - p(y_t|a'_i, a_{-i,t})) = 0$. The lemma now follows from (1) and (2). ■

By Lemmas 1 and 2, we have

$$d^{\alpha^\mu} \geq \eta \sum_h \mu(h) |v(h) - u_i(a^*)|.$$

Hence, by Lemma 3, d^{α^μ} is lower-bounded by the value of the program

$$\begin{aligned} & \inf_{(v(h))_h} \eta \sum_h \mu(h) |v(h) - u_i(a^*)| \quad \text{subject to} \\ (1 - \delta) \delta^{t-1} \max\{g_t^\mu, 0\} & \leq \sum_h \mu(h) L_t(h) (v(h) - u_i(a^*)) \quad \text{for all } t, \\ |v(h) - u_i(a^*)| & \leq \bar{u} \quad \text{for all } h. \end{aligned}$$

By weak duality, we have

$$d^{\alpha^\mu} \geq \sup_{(\lambda_t)_{t \geq 0}} \inf_{(x(h))_h \in [-\bar{u}, \bar{u}]} \left(\sum_h \mu(h) (\eta |x(h)| - \sum_t \lambda_t L_t(h) x(h)) + \sum_t \lambda_t (1 - \delta) \delta^{t-1} \max\{g_t^\mu, 0\} \right). \quad (3)$$

Now fix any $\varepsilon' \in (0, \varepsilon)$, and fix

$$\lambda_t = \left(\rho + \sqrt{\frac{\chi^2}{1 - \delta}} \right)^{-(1 + \varepsilon')} \delta^{t-1} \quad \text{for all } t.$$

The remainder of the proof shows that, for this choice of $(\lambda_t)_t$, there exists $k > 0$ such that if $\rho + \sqrt{\chi^2 / (1 - \delta)} \geq k$ then the right-hand side of (3) is at least $\left(\rho + \sqrt{\chi^2 / (1 - \delta)} \right)^{-(1 + \varepsilon)}$. This suffices to establish the theorem.

We start with a simple inequality for geometric series.

Lemma 4 *For any sequence of numbers $\{z_t\}_t$ with $z_t \in [0, \bar{u}]$ for all t , we have*

$$\frac{1}{\bar{u}} \left((1 - \delta) \sum_t \delta^{t-1} z_t \right)^2 \leq (1 - \delta^2) \sum_t \delta^{2(t-1)} z_t.$$

Proof. For any $X > 0$, consider the program

$$\begin{aligned} & \min_{\{z_t\}_t} (1 - \delta) \sum_t \delta^{2(t-1)} z_t \quad \text{subject to} \\ (1 - \delta) \sum_t \delta^{t-1} z_t & = X \quad \text{and} \\ z_t & \in [0, \bar{u}] \quad \text{for all } t. \end{aligned}$$

Note that if $z_t > 0$ and $z_{t'} < \bar{u}$ for $t' > t$, then decreasing z_t by ε and increasing $z_{t'}$ by

$\varepsilon \delta^{t-1} / \delta^{t'-1}$ continues to satisfy the constraints (for small enough $\varepsilon > 0$), while decreasing the objective by $(1 - \delta) \varepsilon \delta^{t-1} (\delta^{t-1} - \delta^{t'-1}) > 0$. Thus, there exists T such that $z_t = 0$ for all $t < T$ and $z_t = \bar{u}$ for all $t > T$. The first constraint then implies that

$$\begin{aligned}
(1 - \delta) \delta^{T-1} z_T + \delta^T \bar{u} &= X && \Longleftrightarrow \\
(1 - \delta)^2 \delta^{2(T-1)} (z_T)^2 + 2\delta (1 - \delta) \delta^{2(T-1)} z_T \bar{u} + \delta^{2T} \bar{u}^2 &\geq X^2 && \implies \\
(1 - \delta) (1 + \delta) \delta^{2(T-1)} z_T \bar{u} + \delta^{2T} \bar{u}^2 &\geq X^2 && \Longleftrightarrow \\
(1 - \delta) \delta^{2(T-1)} z_T + \frac{\delta^{2T}}{1 + \delta} \bar{u} &\geq \frac{X^2}{(1 + \delta) \bar{u}},
\end{aligned}$$

and hence the value of the objective is at least

$$(1 - \delta) \delta^{2(T-1)} z_T + (1 - \delta) \sum_{t=T+1}^{\infty} \delta^{2(t-1)} \bar{u} = (1 - \delta) \delta^{2(T-1)} z_T + \frac{\delta^{2T}}{1 + \delta} \bar{u} \geq \frac{X^2}{(1 + \delta) \bar{u}}.$$

In total, we have

$$(1 - \delta) \sum_t \delta^{2(t-1)} z_t \geq \frac{1}{(1 + \delta) \bar{u}} \left((1 - \delta) \sum_t \delta^{t-1} z_t \right)^2.$$

Rearranging yields the desired result. ■

We now bound the second line in (3) (which, note, does not depend on $x(h)$).

Lemma 5 *We have*

$$\sum_t \lambda_t (1 - \delta) \delta^{t-1} \max \{g_t^\mu, 0\} \geq \frac{1}{(1 + \delta) \bar{u}} \left(\rho + \sqrt{\frac{\chi^2}{1 - \delta}} \right)^{-(1 + \varepsilon')} \left(\gamma - \frac{\gamma + \bar{u}}{\eta \bar{u}} d^{\alpha_\mu} \right)^2.$$

Proof. We have

$$\begin{aligned}
\sum_t \lambda_t (1 - \delta) \delta^{t-1} \max \{g_t^\mu, 0\} &= \left(\rho + \sqrt{\frac{\chi^2}{1 - \delta}} \right)^{-(1+\varepsilon')} (1 - \delta) \sum_t \delta^{2(t-1)} \max \{g_t^\mu, 0\} \\
&\geq \left(\rho + \sqrt{\frac{\chi^2}{1 - \delta}} \right)^{-(1+\varepsilon')} \frac{1}{(1 + \delta) \bar{u}} \left((1 - \delta) \sum_t \delta^{t-1} \max \{g_t^\mu, 0\} \right)^2 \\
&\geq \left(\rho + \sqrt{\frac{\chi^2}{1 - \delta}} \right)^{-(1+\varepsilon')} \frac{1}{(1 + \delta) \bar{u}} \left(\max \left\{ (1 - \delta) \sum_t \delta^{t-1} g_t^\mu, 0 \right\} \right)^2 \\
&= \left(\rho + \sqrt{\frac{\chi^2}{1 - \delta}} \right)^{-(1+\varepsilon')} \frac{1}{(1 + \delta) \bar{u}} (g^\mu)^2 \\
&\geq \left(\rho + \sqrt{\frac{\chi^2}{1 - \delta}} \right)^{-(1+\varepsilon')} \frac{1}{(1 + \delta) \bar{u}} \left(\gamma - \frac{\gamma + \bar{u}}{\eta \bar{u}} d^{\alpha_\mu} \right)^2,
\end{aligned}$$

where the first inequality follows by applying Lemma 4 to the sequence $\{\max \{g_t^\mu, 0\}\}_t$, the second inequality follows by Jensen (as the max operator is convex), and the third inequality follows by Lemma 1. ■

The next three lemmas (Lemmas 6–8) will be used to bound the first line in (3).

Lemma 6 *We have*

$$\begin{aligned}
&\inf_{(x(h))_{h \in [-\bar{u}, \bar{u}]}} \sum_h \mu(h) \left(\eta |x(h)| - \sum_t \lambda_t L_t(h) x(h) \right) \\
&= \left(\Pr \left(\left| \sum_t \lambda_t L_t(h) \right| \geq \eta \right) \eta - \mathbb{E} \left[\mathbf{1} \left\{ \left| \sum_t \lambda_t L_t(h) \right| \geq \eta \right\} \left| \sum_t \lambda_t L_t(h) \right| \right] \right) \bar{u}.
\end{aligned}$$

Proof. The left-hand side is minimized by taking

$$x(h) = \begin{cases} \bar{u} & \text{if } \sum_t \lambda_t L_t(h) \geq \eta, \\ 0 & \text{if } \sum_t \lambda_t L_t(h) \in [-\eta, \eta], \\ -\bar{u} & \text{if } \sum_t \lambda_t L_t(h) \leq -\eta. \end{cases}$$

This yields the desired equality. ■

Lemma 7 *Let X be a non-negative random variable, whose distribution is first-order stochastically dominated by a continuous distribution F satisfying $\lim_{x \rightarrow \infty} x(1 - F(x)) = 0$.*

Then, for all $\eta \geq 0$,

$$\Pr(X \geq \eta) \eta - \mathbb{E}[\mathbf{1}\{X \geq \eta\} X] \geq - \int_{x \geq \eta} (1 - P(x)) dx.$$

Proof. Let $G(x) = \Pr(X \geq x)$. By hypothesis, $G(x) \leq 1 - F(x)$ for all x , and hence $\lim_{x \rightarrow \infty} xG(x) \leq \lim_{x \rightarrow \infty} x(1 - F(x)) = 0$. By integration by parts,

$$\int_{x \geq \eta} x dG(x) = \lim_{x \rightarrow \infty} xG(x) - \eta G(\eta) - \int_{x \geq \eta} G(x) dx = -\eta G(\eta) - \int_{x \geq \eta} G(x) dx.$$

Hence,

$$\begin{aligned} \Pr(X \geq \eta) \eta - \mathbb{E}[\mathbf{1}\{X \geq \eta\} X] &= G(\eta) \eta + \int_{x \geq \eta} x dG(x) \\ &= G(\eta) \eta - \eta G(\eta) - \int_{x \geq \eta} G(x) dx \\ &= - \int_{x \geq \eta} G(x) dx \geq - \int_{x \geq \eta} (1 - F(x)) dx. \end{aligned}$$

■

The next lemma is a well-known martingale version of Bernstein's inequality, due to Freedman (1975).

Lemma 8 (Freedman's Inequality) *Let $(X_t)_{t \geq 1}$ be a sequence of martingale increments adapted to a filtration $(H_t)_{t \geq 1}$, so that $\mathbb{E}[X_t | H_{t-1}] = 0$. Assume that $|X_t| \leq R$, and let $S_T = \sum_{t=1}^T X_t$, $V_t = \text{Var}[X_t | H_{t-1}]$, and $W_T = \sum_{t=1}^T V_t$. Then, for all positive numbers x and y ,*

$$\Pr(S_T \geq x \text{ and } W_T \leq y) \leq \exp\left(-\frac{x^2}{2(Rx + y)}\right).$$

Now fix \bar{k} such that, for all $k \geq \bar{k}$, we have

$$\frac{4(1+\eta)^2}{\eta(2+\eta)} \exp\left(-k^{\varepsilon'} \frac{\eta^2}{2(1+\eta)}\right) \leq \frac{\gamma^2}{8(1+\delta)\bar{u}} k^{-1}, \quad (4)$$

$$\gamma - \frac{\gamma + \bar{u}}{\eta \bar{u}} k^{-(1+\varepsilon)} \geq \frac{\gamma}{2}, \quad \text{and} \quad (5)$$

$$k^{\varepsilon - \varepsilon'} \frac{\gamma^2}{8(1+\delta)\bar{u}} \geq 1. \quad (6)$$

We are now ready to bound the first line in (3).

Lemma 9 *If $\rho + \sqrt{\chi^2/(1-\delta)} \geq \bar{k}$, then we have*

$$\inf_{(x(h))_h \in [-\bar{u}, \bar{u}]} \sum_h \mu(h) \left(\eta |x(h)| - \sum_t \lambda_t L_t(h) x(h) \right) \geq -\frac{\gamma^2}{8(1+\delta)\bar{u}} \left(\rho + \sqrt{\frac{\chi^2}{1-\delta}} \right)^{-(1+\varepsilon')}$$

Proof. Note that

$$|\lambda_t L_t| \leq \left(\rho + \sqrt{\frac{\chi^2}{1-\delta}} \right)^{-(1+\varepsilon')} \quad \rho \leq \left(\rho + \sqrt{\frac{\chi^2}{1-\delta}} \right)^{-\varepsilon'},$$

and

$$\begin{aligned} \sum_t \mathbb{E} [(\lambda_t L_t)^2 | h^t] &\leq \sum_t \left(\rho + \sqrt{\frac{\chi^2}{1-\delta}} \right)^{-2(1+\varepsilon')} \delta^{2(t-1)} \chi^2 \\ &\leq \left(\rho + \sqrt{\frac{\chi^2}{1-\delta}} \right)^{-2(1+\varepsilon')} \frac{\chi^2}{1-\delta} \leq \left(\rho + \sqrt{\frac{\chi^2}{1-\delta}} \right)^{-\varepsilon'}. \end{aligned}$$

Thus, applying Freedman's inequality with

$$R = y = \left(\rho + \sqrt{\frac{\chi^2}{1-\delta}} \right)^{-\varepsilon'}$$

to the martingale increments $(\lambda_t L_t)_t$ and $(-\lambda_t L_t)_t$, and taking the union bound, we have, for all $x > 0$,

$$\Pr \left(\left| \sum_t \lambda_t L_t \right| \geq x \right) \leq 2 \exp \left(- \left(\rho + \sqrt{\frac{\chi^2}{1-\delta}} \right)^{\varepsilon'} \frac{x^2}{2(1+x)} \right).$$

By Lemmas 6 and 7, we have

$$\inf_{(x(h))_h \in [-\bar{u}, \bar{u}]} \sum_h \mu(h) \left(\eta |x(h)| - \sum_t \lambda_t L_t(h) x(h) \right) \geq - \int_{x \geq \eta} 2 \exp \left(- \left(\rho + \sqrt{\frac{\chi^2}{1-\delta}} \right)^{\varepsilon'} \frac{x^2}{2(1+x)} \right) dx.$$

Since

$$\begin{aligned}\frac{d}{dx} \frac{x^2}{2(1+x)} &= \frac{1}{2} \frac{x(x+2)}{(x+1)^2} \quad \text{and} \quad \frac{d}{dx} \frac{1}{2} \frac{x(x+2)}{(x+1)^2} = \frac{1}{(x+1)^3} \geq 0 \quad \text{for all } x \geq 0, \quad \text{we have} \\ \frac{d}{dx} \frac{x^2}{2(1+x)} &= \frac{x(x+2)}{2(x+1)^2} \geq \frac{\eta(2+\eta)}{2(1+\eta)^2} \quad \text{for all } x \geq \eta.\end{aligned}$$

Thus, by integration by substitution,

$$\begin{aligned}& - \int_{x \geq \eta} 2 \exp \left(- \left(\rho + \sqrt{\frac{\chi^2}{1-\delta}} \right)^{\varepsilon'} \frac{x^2}{2(1+x)} \right) dx \\ & \geq - \frac{4(1+\eta)^2}{\eta(2+\eta)} \left(\rho + \sqrt{\frac{\chi^2}{1-\delta}} \right)^{-\varepsilon'} \int_{y \geq \left(\rho + \sqrt{\frac{\chi^2}{1-\delta}} \right)^{\varepsilon'} \frac{\eta^2}{2(1+\eta)}} \exp(-y) dy \\ & = - \frac{4(1+\eta)^2}{\eta(2+\eta)} \left(\rho + \sqrt{\frac{\chi^2}{1-\delta}} \right)^{-\varepsilon'} \exp \left(- \left(\rho + \sqrt{\frac{\chi^2}{1-\delta}} \right)^{\varepsilon'} \frac{\eta^2}{2(1+\eta)} \right) \\ & \geq - \frac{\gamma^2}{8(1+\delta)\bar{u}} \left(\rho + \sqrt{\frac{\chi^2}{1-\delta}} \right)^{-(1+\varepsilon')},\end{aligned}$$

where the last inequality follows from $\rho + \sqrt{\chi^2/(1-\delta)} \geq \bar{k}$ and (4). ■

Finally, by (3) and Lemmas 5 and 9, we have

$$\begin{aligned}d^{\alpha^\mu} & \geq \frac{1}{(1+\delta)\bar{u}} \left(\rho + \sqrt{\frac{\chi^2}{1-\delta}} \right)^{-(1+\varepsilon')} \left(\gamma - \frac{\gamma + \bar{u}}{\eta\bar{u}} d^{\alpha^\mu} \right)^2 - \frac{\gamma^2}{8(1+\delta)\bar{u}} \left(\rho + \sqrt{\frac{\chi^2}{1-\delta}} \right)^{-(1+\varepsilon')} \\ & = \left(\rho + \sqrt{\frac{\chi^2}{1-\delta}} \right)^{-(1+\varepsilon)} \left(\rho + \sqrt{\frac{\chi^2}{1-\delta}} \right)^{\varepsilon-\varepsilon'} \left(\frac{1}{(1+\delta)\bar{u}} \left(\gamma - \frac{\gamma + \bar{u}}{\eta\bar{u}} d^{\alpha^\mu} \right)^2 - \frac{\gamma^2}{8(1+\delta)\bar{u}} \right).\end{aligned}$$

If $\gamma - \frac{\gamma + \bar{u}}{\eta\bar{u}} d^{\alpha^\mu} \leq \frac{\gamma}{2}$, then $d^{\alpha^{\bar{\mu}}} \geq \bar{k}^{-(1+\varepsilon)}$ by (5), and hence $d^{\alpha^{\bar{\mu}}} \geq \left(\rho + \sqrt{\frac{\chi^2}{1-\delta}} \right)^{-(1+\varepsilon)}$. Otherwise,

we have

$$\begin{aligned}
d^{\alpha\mu} &\geq \left(\rho + \sqrt{\frac{\chi^2}{1-\delta}}\right)^{-(1+\varepsilon)} \left(\rho + \sqrt{\frac{\chi^2}{1-\delta}}\right)^{\varepsilon-\varepsilon'} \left(\frac{\gamma^2}{4(1+\delta)\bar{u}} - \frac{\gamma^2}{8(1+\delta)\bar{u}}\right) \\
&= \left(\rho + \sqrt{\frac{\chi^2}{1-\delta}}\right)^{-(1+\varepsilon)} \left(\rho + \sqrt{\frac{\chi^2}{1-\delta}}\right)^{\varepsilon-\varepsilon'} \frac{\gamma^2}{8(1+\delta)\bar{u}} \\
&\geq \left(\rho + \sqrt{\frac{\chi^2}{1-\delta}}\right)^{-(1+\varepsilon)},
\end{aligned}$$

where the last inequality follows from (6).

References

- [1] Abreu, Dilip, Paul Milgrom, and David Pearce. “Information and Timing in Repeated Partnerships.” *Econometrica* 59.6 (1991): 1713-1733.
- [2] Abreu, Dilip, David Pearce, and Ennio Stacchetti. “Toward a Theory of Discounted Repeated Games with Imperfect Monitoring.” *Econometrica* 58.5 (1990): 1041-1063.
- [3] Fudenberg, Drew, David Levine, and Eric Maskin. “The Folk Theorem with Imperfect Public Information.” *Econometrica* 62.5 (1994): 997-1039.
- [4] Freedman, David A. “On Tail Probabilities for Martingales.” *Annals of Probability* 3.1 (1975): 100-118.
- [5] Hörner, Johannes, and Satoru Takahashi. “How Fast do Equilibrium Payoff Sets Converge in Repeated Games?” *Journal of Economic Theory* 165 (2016): 332-359.
- [6] Kandori, Michihiro. “Introduction to Repeated Games with Private Monitoring.” *Journal of Economic Theory* 102.1 (2002): 1-15.
- [7] Kandori, Michihiro, and Ichiro Obara. “Efficiency in Repeated Games Revisited: The Role of Private Strategies.” *Econometrica* 74.2 (2006): 499-519.
- [8] Sugaya, Takuo, and Alexander Wolitzky, “Informational Requirements for Cooperation.” *Working Paper* (2022).