Folk Theorem in Repeated Games with Private Monitoring

Takuo Sugaya*†
Princeton University

February 27, 2011

Abstract

We show that the folk theorem with individually rational payoffs defined by pure strategies generically holds for a general N-player repeated game with private monitoring when the number of each player's signals is sufficiently large. No cheap talk communication device or public randomization device is necessary.

Journal of Economic Literature Classification Numbers: C72, C73, D82

Keywords: repeated game, folk theorem, private monitoring

*tsugaya@princeton.edu

†I thank Stephen Morris for the invaluable advice, Yuliy Sannikov for the helpful discussion, and Dilip Abreu, Edoardo Grillo, Johannes Hörner, Michihiro Kandori, George Mailath, Satoru Takahashi and Yuichi Yamamoto for useful comments. The usual disclaimer of course applies.

1 Introduction

One of the key results in the literature on infinitely repeated games is the folk theorem: any feasible and individually rational payoff can be sustained in equilibrium when players are sufficiently patient. Fudenberg and Maskin (1986) establish the folk theorem under perfect monitoring, that is, when players can directly observe the action profile. Fudenberg, Levine and Maskin (1994) extend the folk theorem to imperfect public monitoring, where players can observe only public noisy signals about the action profile.

The driving force of the folk theorem in perfect or public monitoring is the coordination of the future play based on common knowledge of the relevant histories. Specifically, the public component of the histories such as action profiles in perfect monitoring or public signals in public monitoring (at least statistically) reveals past action profiles. Since this public information is common knowledge, players can coordinate the punishment contingent on the public information to provide dynamic incentives to choose actions that are not static best responses.

With *private monitoring*, where players can observe only private noisy signals about the action profile, however, common knowledge no longer exists and the coordination is difficult (we call this problem "*coordination failure*").[1] Hence, the robustness of the folk theorem to a general private monitoring has been an open question. Kandori (2002), for example, mentions that "[t]his is probably one of the best known long-standing open questions in economic theory."

Moreover, many economic situations should be analyzed as repeated games with private monitoring. For example, Stigler (1964) proposes a repeated price-setting oligopoly, where firms set their own price in a face-to-face negotiation. Hence, they cannot observe their opponents' prices. Instead, a firm obtains some information about the opponents' prices through its own sales. Since the sales level depends on both the opponents' prices and unobservable shocks due to business cycles, it is an imperfect signal. In addition, the sales

[1] Mailath and Morris (2002 and 2006) and Sugaya and Takahashi (2010) offer the formal model of this argument.

of each firm are often private information. Thus, monitoring is imperfect and private. In addition, Fuchs (2007) applies a repeated game with private monitoring to a contract between a principal and an agent and Harrington and Skrzypacz (2007) analyze cartel behaviors using the frame work of a repeated game with private monitoring.

This paper is, to the best of our knowledge, the first to show that the folk theorem holds in repeated games with a generic private monitoring. We unify and improve the three approaches in the literature on private monitoring that are used to show partial results so far: belief-free, belief-based and communication approaches.

The belief-free approach (and its generalization) has been successful to show the folk theorem in prisoners' dilemma.[2] A strategy profile is *belief-free* if, for any history profile and for any player, the continuation strategy of that player after that history is optimal conditional on the opponent's history. With almost perfect monitoring, Piccione (2002) and Ely and Välimäki (2002) show the folk theorem for the two-player prisoners' dilemma. Without any assumption on the precision of monitoring but with conditionally independent monitoring, Matsushima (2004) obtains the folk theorem in the two-player prisoners' dilemma.

Unfortunately, however, without almost perfect or conditionally independent monitoring, only limited results have been shown: Fong, Gossner, Hörner and Sannikov (2010) show the efficiency and Sugaya (2010a) shows the folk theorem in the two-player prisoners' dilemma with some restricted classes of the distributions of the private signals.

Several papers construct belief-based equilibria, where players' strategies involve statistical inference about the opponents' past histories. With almost perfect monitoring, Sekiguchi (1997) shows the efficiency and Bhaskar and Obara (2002) show the folk theorem in the two-player prisoners' dilemma. Mailath and Morris (2002 and 2006) consider the robustness of the equilibrium in public monitoring to almost public monitoring. Based on their insights, Hörner and Olszewski (2009) establish the folk theorem for almost public monitoring. Phelan

[2]Kandori and Obara (2006) use a similar concept to analyze a private strategy in public monitoring. Kandori (2010) considers "weakly belief-free equilibria," which is a generalization of the belief-free equilibria. Apart from a typical repeated-game setting, Takahashi (2010) considers the community enforcement and Miyagawa, Miyahara and Sekiguchi consider the situation where a player can improve the monitoring by paying cost.

and Skrzypacz (2009) characterize the set of possible beliefs and Kandori and Obara (2010) offer an easy way to verify if a finite-state automaton strategy is an equilibrium.

Another approach to analyze repeated games with private monitoring is to introduce communication, whose results are publicly known. The folk theorems have been proven by Compte (1998), Kandori and Matsushima (1998), Aoyagi (2002), Fudenberg and Levine (2002) and Obara (2007). Introducing a public element and letting a strategy depend only on the public element allow these papers to sidestep the difficulty of the coordination through private signals. However, the analyses do not apply to settings where communication is not allowed; for example, in Stigler (1964)'s oligopoly example, anti-trust laws rule that communication is illegal. Furthermore, it is uncertain whether their equilibria are robust to the perturbation that the message transmission has small private noises, which is one of the main motivations in private monitoring.

This paper incorporates all the three approaches. First, the equilibrium strategy to show the folk theorem is *occasionally belief-free.* That is, we see the repeated game as the repetition of long review phases and at the beginning of each review phase, for any player, any strategy that can be taken on the equilibrium path with a positive probability is optimal conditional on the histories of the opponents. Second, however, since the belief-free property does not hold except for the beginning of the phase, we consider each player's statistical inference about the opponents' past histories seriously as in belief-based approach within each phase. Finally, in our equilibrium, the message exchange via taking actions plays an important role. Hence, one of our methodological contributions is to offer a systematic way to dispense the cheap talk message protocol with the message exchange via taking actions.

The rest of the paper is organized as follows. Section 2 introduces the model and Section 3 states the assumptions and result. The remaining parts of the paper are devoted to prove the result. In the proof, we focus on a two-player prisoners' dilemma. Section 4 relates the infinitely repeated game with the finitely repeated game with the auxiliary scenario and derives the sufficient condition on the finitely repeated game to show the folk theorem. Section 5 introduces the preliminary results important for the equilibrium construction.

Especially, in Section 5.2, we define the message protocol *via taking actions* to coordinate the future play based on the message exchange. Based on Section 5, we construct an equilibrium to attain the folk theorem in the two-player prisoners' dilemma in Section 6.

Throughout the main text of the paper, we use perfect and public cheap talk and public randomization devices for some of the message exchanges although we use the message protocol via taking action defined in Section 5.2 for the other to simplify the equilibrium construction.[3] In the Online Appendix, we show that cheap talk and public randomization are completely *dispensable* and all the messages can be sent via taking actions without public randomization. We comment on the dispensability of these devices in Section 7.

Let us finally comment on our focus on the prisoners' dilemma. It is well known that the prisoners' dilemma is exceptional: except for prisoners' dilemma, the belief-free equilibrium cannot obtain the folk theorem.[4] Hörner and Olszewski (2006) extends the belief-free equilibrium to attain the folk theorem for a general game if monitoring is almost perfect. However, as Hörner and Olszewski (2006) mention in their paper, it is not known how to generalize their result to not-almost-perfect monitoring. For our equilibrium construction in prisoners' dilemma, to deal with the fact that we have belief-free period occasionally, we establish the way to let the players coordinate their future play through private signals. Given this method, together with the insight of Hörner and Olszewski (2006), it is possible to attain the folk theorem for a general game. We come back to this point in Section 8.

[3] Even with cheap talk, however, the general folk theorem is new to the two-player case.

[4] Ely, Hörner and Olszewski (2005) characterize the set of belief-free equilibrium payoffs in a two-player game. Yamamoto (2009b), based on Yamamoto (2007 and 2009a), characterizes the set of belief-free review-strategy equilibrium payoffs in a general N-player game with conditionally independent monitoring. Sugaya (2010b) characterizes the set of belief-free review-strategy equilibrium payoffs with a generic monitoring in a general N-player game with $N \geq 4$. The characterized set is not large enough to show the folk theorem in a general game.

2 Model

2.1 Stage Game

The stage game is given by $\left\{I, (A_i, Y_i, \tilde{u}_i)_{i\in I}, q\right\}$. $I = \{1, \dots, N\}$ is the set of players, A_i with $|A_i| \geq 2$ is the finite set of player i's pure actions, Y_i is the finite set of player i's private signals and $\tilde{u}_i : A_i \times Y_i \to \mathbb{R}$ is player i's ex-post utility function. Let $A \equiv \prod_{i\in I} A_i$ and $Y \equiv \prod_{i\in I} Y_i$ be the set of action profiles and signal profiles, respectively.

In every stage game, player i chooses an action $a_i \in A_i$, which induces the action profile $a \equiv (a_1, \dots, a_N) \in A$. Then, a signal profile $y = (y_1, \dots, y_N) \in Y$ is jointly distributed according to the conditional probability function $q(\cdot \mid a)$. Given an action $a_i \in A_i$ and a private signal $y_i \in Y_i$, player i receives the ex-post utility $\tilde{u}_i(a_i, y_i)$. Thus, her expected payoff conditional on an action profile $a \in A$ is given by $u_i(a) \equiv \sum_{y\in Y} q(y \mid a)\, \tilde{u}_i(a_i, y_i)$. For each $a \in A$, let $u(a)$ represent the payoff vector $(u_i(a))_{i\in I}$.

2.2 Repeated Game

Consider the infinitely repeated game of the above stage game in which the discount factor is $\delta \in (0, 1)$. Let $a_{i,\tau}$ and $y_{i,\tau}$ denote the performed action and the observed private signal respectively in period τ by player i. Player i's private history up to period $t \geq 1$ is given by $h_i^t \equiv (a_{i,\tau}, y_{i,\tau})_{\tau=1}^{t-1}$. Let $h_i^1 = \emptyset$ and for each $t \geq 1$, let H_i^t be the set of all h_i^t. Then, a strategy for player i is defined to be a mapping $\sigma_i : \bigcup_{t=1}^{\infty} H_i^t \to \triangle(A_i)$. Let Σ_i be the set of all strategies for player i. Finally, let $E(\delta)$ be the set of sequential equilibrium payoffs with a common discount factor δ.

3 Assumption and Result

First, we state the assumption on the payoff structure. Let $F \equiv \mathrm{co}(\{u(a)\}_{a\in A})$ be the set of feasible payoffs. The individually rational payoff for player i is $v_i^* \equiv \min_{a_{-i}\in A_{-i}} \max_{a_i\in A_i} u_i(a_i, a_{-i})$. Note that we concentrate on the pure strategy minmax. Then, the set of feasible and individually rational payoffs is given by

$$F^* \equiv \{v \in F : v_i \geq v_i^* \text{ for all } i\}.$$

We assume the full dimensionality of F^*.

Assumption 1 *A stage game payoff structure satisfies the full dimensionality condition:*

$$\dim(F^*) = N.$$

Second, we state the assumption on the signal structure.

Assumption 2 *The number of each player's signals is sufficiently large: for any $i \in I$, we have*

$$|Y_i| \geq \max\left\{\max_{j \neq i} |A_j| + 2\sum_{\iota \neq i,j} |A_\iota|, \sum_{j \in I} |A_j|, \max_{j \in I} 2|A_j|\right\}.$$

Under these assumptions, we can generically construct an equilibrium to attain any point in $\text{int}(F^*)$.

Theorem 1 *If Assumptions 1 and 2 are satisfied, then the folk theorem generically holds: for generic $q(\cdot \mid \cdot)$, for any $v \in \text{int}(F^*)$, there exists $\bar{\delta} < 1$ such that, for all $\delta > \bar{\delta}$, $v \in E(\delta)$.*

For the proof, we focus on the two-player prisoners' dilemma: $I = 2$, $A_i = \{C_i, D_i\}$ and

$$u_i(D_i, C_j) > u_i(C_i, C_j) > u_i(D_i, D_j) > u_i(C_i, D_j) \tag{1}$$

for all i. In this case, Assumptions 1 and 2 are equivalent to having $|Y_i| \geq 4$ for all i.

In addition, since we concentrate on two-player case, whenever we say player i and j, unless otherwise mentioned, we assume $i \neq j$.

Further, since it is well known that, for the two-player prisoners' dilemma, if we can sustain any v with

$$v \in \text{int}([u_1(D_1, D_2), u_1(C_1, C_2)] \times [u_2(D_2, D_1), u_2(C_2, C_1)]). \tag{2}$$

as an equilibrium payoff, then it is easy to show that any $v \in \text{int}(F^*)$ is sustainable.[5] Moreover, as we will explain in Section 8, the equilibrium construction to support v with (2) is useful to show the folk theorem in a general game. Therefore, in this paper, we focus on the case with

We prove the theorem via following steps. We arbitrarily fix v so that (2) holds. We implement v by a strategy profile that is recursive in every T_P periods, where $T_P \in \mathbb{N}$ should be determined later. In Section 4, we relate the infinitely repeated game with a T_P-period finitely repeated game with an auxiliary scenario. Specifically, we derive conditions on a strategy and an auxiliary scenario in the T_P-period finitely repeated game from which we can construct a strategy to implement v in the infinitely repeated game. In Section 5.1, we construct random variables to construct the strategy and the auxiliary scenario. In Section 5.2, we construct message protocols for the players to send a binary message $\{G, B\}$. This is the key of our equilibrium construction to overcome the coordination failure. In Section 5.3, we finish defining all the necessary variables to specify the strategy and the auxiliary scenario. Given these preparations, Section 6 proves the theorem with perfect cheap talk. Section 7 offers an intuitive explanation of how to dispense cheap talk. See the Online Appendix for the formal proof without cheap talk. Finally, Section 8 explains how to extend the proof to a general game.

4 Finitely Repeated Game

Let us consider a T_P-period *finitely* repeated game. Let $\sigma_i^{T_P} : H_i^{T_P} \to \Delta(A_i)$ be player i's strategy in the finitely repeated game. Let $\Sigma_i^{T_P}$ be the set of all strategies in the finitely repeated game. Each player i has a state $x_i \in \{G, B\}$. In state x_i, player i plays $\sigma_i(x_i) \in \Sigma_i^{T_P}$. In addition, player i with x_i gives the "auxiliary scenario" (or "reward function") $\pi_j(x_i, \cdot : \delta)$ to player j. Here, $\pi_j(x_i, \cdot : \delta) : H_i^{T_P+1} \to \mathbb{R}$, that is, the auxiliary scenarios are functions from the history in the finitely repeated game to the real number.

[5] See Ely and Välimäki (2002) for the details.

For v with (2), we can take $\rho > 0$, $\underline{v}_i$ and $\bar{v}_i$ such that

$$u_i(D_1, D_2) + \rho < \underline{v}_i < v_i < \bar{v}_i < u_i(C_1, C_2) - \rho. \tag{3}$$

Our task is to find $\{\sigma_i(x_i)\}_i$ and $\{\pi_j(x_j, \cdot)\}_i$ such that, for sufficiently large δ, there exists T_P such that, for any i,

1. for any x_j, $\sigma_i(G)$ and $\sigma_i(B)$ are optimal in the finitely repeated game:

$$\sigma_i(G), \sigma_i(B) \in \arg\max_{\sigma_i^{T_P} \in \Sigma_i^{T_P}} \mathbb{E}\left[\sum_{t=1}^{T_P} \delta^{t-1} u_i(a_t) + \pi_i(x_j, h_j^{T_P+1} : \delta) \mid \sigma_i^{T_P}, \sigma_j(x_j)\right], \tag{4}$$

2. the discounted average of player i's instantaneous utilities and player j's auxiliary scenario on player i is equal to $\bar{v}_i$ if player j's state is good ($x_j = G$) and equal to $\underline{v}_i$ if player j's state is bad ($x_j = B$):

$$\frac{1-\delta}{1-\delta^{T_P}} \mathbb{E}\left[\sum_{t=1}^{T_P} \delta^t u_i(a_t) + \pi_i(x_j, h_j^{T_P+1} : \delta) \mid \sigma(x)\right] = \begin{cases} \bar{v}_i \text{ if } x_j = G, \\ \underline{v}_i \text{ if } x_j = B. \end{cases} \tag{5}$$

   Intuitively, for sufficiently large δ, since $\lim_{\delta \to 1} \frac{1-\delta}{1-\delta^{T_P}} = \frac{1}{T_P}$, this requires the time average of the expected sum of the instantaneous utilities and the reward is equal to the targeted payoffs $\underline{v}_i$ and $\bar{v}_i$.

3. $\pi_i(G, h_j^{T_P+1} : \delta)$ and $\pi_i(B, h_j^{T_P+1} : \delta)$ are uniformly bounded with respect to δ and

$$\begin{aligned} \pi_i(G, h_j^{T_P+1} &: \delta) \leq 0, \\ \pi_i(B, h_j^{T_P+1} &: \delta) \geq 0. \end{aligned} \tag{6}$$

This implies

$$
\begin{aligned}
p(G, h_j^{T_P+1} &: \delta) \equiv 1 + \frac{1-\delta}{\delta^{T_P}} \frac{\pi_i(G, h_j^{T_P+1} : \delta)}{\bar{v}_i - \underline{v}_i} \in [0,1], \\
p(B, h_j^{T_P+1} &: \delta) \equiv \frac{1-\delta}{\delta^{T_P}} \frac{\pi_i(B, h_j^{T_P+1} : \delta)}{\bar{v}_i - \underline{v}_i} \in [0,1] \qquad (7)
\end{aligned}
$$

for sufficiently large δ. The meaning of $p(x_j, h_j^{T_P+1} : \delta)$ will be specified later.

To see why this is enough for Theorem 1, define the strategy in the infinitely repeated game as follows: we see the repeated game as the repetition of T_P-period "review phases." In each phase, player i has a state $x_i \in \{G, B\}$. Within the phase, player i with state x_i plays according to $\sigma_i(x_i)$. After observing $h_i^{T_P+1}$ in the current phase, the state in the next phase is equal to G with probability $p(x_i, h_i^{T_P+1} : \delta)$ and B with the remaining probability. For this to be feasible, (7) is necessary, which is guaranteed by (6). Hence, we call (6) the "feasibility constraint."

Player i's initial state is equal to G with probability p_v^i and B with probability $1 - p_v^i$ with

$$
p_v^i \bar{v}_j + (1 - p_v^i)\underline{v}_j = v_j.
$$

Then, since

$$
\begin{aligned}
& (1-\delta)\sum_{t=1}^{T_P} \delta^{t-1} u_i(a_t) + \delta^{T_P}\left[p(G, h_j^{T_P+1})\bar{v}_i + \{1 - p(G, h_j^{T_P+1})\}\underline{v}_i\right] \\
= \quad & \left(1-\delta^{T_P}\right)\left[\frac{1-\delta}{1-\delta^{T_P}}\left\{\sum_{t=1}^{T_P} \delta^{t-1} u_i(a_t) + \pi_i(G, h_j^{T_P+1} : \delta)\right\}\right] + \delta^{T_P}\bar{v}_i
\end{aligned}
$$

and

$$
\begin{aligned}
& (1-\delta)\sum_{t=1}^{T_P} \delta^{t-1} u_i(a_t) + \delta^{T_P}\left[p(B, h_j^{T_P+1})\bar{v}_i + \{1 - p(B, h_j^{T_P+1})\}\underline{v}_i\right] \\
= \quad & \left(1-\delta^{T_P}\right)\left[\frac{1-\delta}{1-\delta^{T_P}}\left\{\sum_{t=1}^{T_P} \delta^{t-1} u_i(a_t) + \pi_i(B, h_j^{T_P+1} : \delta)\right\}\right] + \delta^{T_P}\underline{v}_i,
\end{aligned}
$$

(4) and (5) imply

1. conditional on the opponent's state, the above strategy in the infinitely repeated game is optimal for sufficiently large discount factor δ,

2. if player j is in the state G, then player i's payoff from the infinitely repeated game is $\bar{v}_i$ and if player j is in the state B, then player i's payoff is $\underline{v}_i$ for sufficiently large discount factor δ, and

3. the payoff in the initial period is $p_v^j \bar{v}_i + (1 - p_v^j)\underline{v}_i = v_i$ as desired.

Therefore, it suffices to show that there exist $\{\{\sigma_i(x_i)\}_{x_i \in \{G,B\}}\}_i$ and $\{\{\pi_i(x_j, \cdot : \delta)\}_{x_j \in \{G,B\}}\}_i$ satisfying (4) (5) and (6) with $\{\{\pi_i(x_j, \cdot : \delta)\}_{x_j \in \{G,B\}}\}_i$ being uniformly bounded with respect to δ in T_P-period finitely repeated game. Each review phase corresponds to the finitely repeated game and we use the terms "review phase" and "finitely repeated game" interchangeably.

5 Preliminaries

5.1 Statistics

In this section, we construct the statistics to be used for the equilibrium construction. First, we define $\psi_j^a(y_j)$ that allows player j to statistically distinguish player i's action. For our equilibrium to work, $\psi_j^a(y_j)$ should satisfy the following properties:

1. given that the other player takes a_j, the expected value of $\psi_j^a(y_j)$ is highest when player i takes a_i. Intuitively, we make the reward $\pi_i(x_j, h_j^{T_P+1} : \delta)$ linearly increasing in the summation of $\psi_j^a(y_{j,t})$ in the given review phase with a sufficiently high slope. Then, this incentivizes player i to take a_i.

2. suppose, on the equilibrium path, the action profile a is taken. In addition, player i calculates the conditional expectation of $\psi_j^a(y_j)$ conditional on a, y_i. The ex ante value

of this conditional expectation is

$$\sum_{y_i} \left\{ \sum_{y_j} \psi_j^a(y_j) q(y_j \mid a, y_i) \right\} q\left(y_i \mid a\right) = \sum_{y_j} \left\{ \sum_{y_i} q(y_j \mid a, y_i) q(y_i \mid a) \right\} \psi_j^a(y_j).$$

On the other hand, if player j secretly deviates to $\tilde{a}_j$ while player i still believes a is taken, then the ex ante value is

$$\sum_{y_i} \left\{ \sum_{y_j} \psi_j^a(y_j) q(y_j \mid a, y_i) \right\} q\left(y_i \mid a_i, \tilde{a}_j\right) = \sum_{y_j} \left\{ \sum_{y_i} q(y_j \mid a, y_i) q\left(y_i \mid a_i, \tilde{a}_j\right) \right\} \psi_j^a(y_j). \tag{8}$$

Note that player i's signal is distributed according to the true distribution $q(y_i \mid a_i, \tilde{a}_j)$ while player i calculates the conditional distribution $q(y_j \mid a, y_i)$ believing a is taken. As we will see, since player i's equilibrium strategy depends on the conditional expectation of $\psi_j^a(y_j)$, we want to exclude the possibility that player j can manipulate the ex ante value of the conditional expectation via changing her own action. That is, (8) is constant for all $\tilde{a}_j$.

The first condition requires that the distribution of player j's signal conditional on $\tilde{a}_i, a_j$, $Q_1(\tilde{a}_i, a_j) \equiv \left(q\left(y_j \mid \tilde{a}_i, a_j\right)\right)_{y_j}$, is linearly independent with respect to $\tilde{a}_i$. The second condition requires that the vector $Q_2(a_i, \tilde{a}_j) \equiv \left(\sum_{y_i} q(y_j \mid a, y_i) q(y_i \mid a_i, \tilde{a}_j)\right)_{y_j}$ is linearly independent with respect to $\tilde{a}_j$. Jointly, the vectors $Q_1(\tilde{a}_i, a_j)$ and $Q_2(a_i, \tilde{a}_j)$ are linearly independent except for $\tilde{a}_i = a_i$ and $\tilde{a}_j = a_j$.

Formally,

Lemma 1 *There exist $q_2 > q_1$ such that, for each $i \in I$ and $a \in A$, there exists a function $\psi_j^a : Y_j \to (0, 1)$ such that*

1. *ψ_j^a can distinguish whether $a_i \in A_i$ is taken or not:*

$$\mathbb{E}\left[\psi_j^a(y_j) \mid \tilde{a}_i, a_j\right] \equiv \sum_{y_j} q(y_j \mid \tilde{a}_i, a_j) \psi_j^a(y_j) = \begin{cases} q_2 & \text{if } \tilde{a}_i = a_i, \\ q_1 & \text{if } \tilde{a}_i \neq a_i, \end{cases} \tag{9}$$

and

2. *player j cannot change the ex ante expectation of player i's conditional expectation of $\psi_j^a(y_j)$: for any $\tilde{a}_j \in A_j$,*

$$\sum_{y_j} \left\{ \sum_{y_i} q(y_j \mid a, y_i) q(y_i \mid a_i, \tilde{a}_j) \right\} \psi_j^a(y_j) = \textit{constant}. \tag{10}$$

Proof. First, arbitrarily fix $q_2 > q_1$. The vectors $Q(\tilde{a}_i, a_j)$ and $Q(a_i, \tilde{a}_j)$ are generically linearly independent except for $\tilde{a}_i = a_i$ and $\tilde{a}_j = a_j$ since Assumption 2 guarantees that $|Y_j| \geq |A_i| + |A_j| - 1$. Therefore, there generically exists a solution $\{\psi_i^a(y_j)\}_{y_j}$ to satisfy (9) and (10). Note that if $\{\psi_i^a(y_j)\}_{y_j}$ solves the system for q_2 and q_1, then, for any $m, m' \in \mathbb{R}_{++}$, $\left\{ \frac{\psi_i^a(y_j)+m}{m'} \right\}_{y_j}$ solves the system for $\frac{q_2+m}{m'}$ and $\frac{q_1+m}{m'}$. Therefore, we can make sure that $\psi_j^a : Y_j \to (0,1)$ by adding a large number and dividing by a large number. ■

Second, for the equilibrium construction, it is convenient to use a variable taking only 0 or 1 instead of $\{\psi_j^a(y_j)\}$. Suppose player j constructs a random variable $\Psi_j^a \in \{0,1\}$ from $\{\psi_j^a(y_j)\}$ as follows: after observing y_j, player j calculates $\psi_j^a(y_j)$. After that, player j draws a random variable from the uniform distribution on $[0,1]$. If the realization of this random variable is less than $\psi_j^a(y_j)$, then $\Psi_j^a = 1$ and otherwise, $\Psi_j^a = 0$. Since $\Pr(\{\Psi_j^a = 1\} \mid \tilde{a}, y) = \psi_j^a(y_j)$ for all $\tilde{a} \in A$ and $y \in Y$, Ψ_j^a also satisfies

1. given that the other players take a_j, the expected value of Ψ_j^a is highest when player i takes a_i, and

2. the ex ante value of the conditional expectation of Ψ_j^a is equal to (8) and is constant for all $\tilde{a}_j$.

Further, player i, after believing that a being taken and observing y_i, constructs a random variable $E_i\Psi_j^a \in \{0,1\}$ from $\sum_{y_j} \psi_j^a(y_j) q(y_j \mid a, y_i)$ as player j constructs Ψ_j^a from $\psi_j^a(y_j)$. (10) implies that

$$\Pr(\{E_i\Psi_j^a = 1\} \mid a_i, a_j) = \Pr(\{E_i\Psi_j^a = 1\} \mid a_i, \tilde{a}_j) \tag{11}$$

for all $\tilde{a}_j \in A_j$.

Finally, we construct $\{\pi_i^- (a_j, y_j)\}$ and $\{\pi_i^+ (a_j, y_j)\}$ such that, if player i's payoff in the stage game is equal to the summation of the instantaneous utility $u_i (a)$ and $\pi_i^- (a_j, y_j)$ or $\pi_i^+ (a_j, y_j)$, then any action profile is indifferent to player i.

Lemma 2 *If Assumption 2 is satisfied, then there exists $\bar{u} > \rho$ such that, for each i, there exist $\pi_i^- : A_j \times Y_j \to [-\bar{u}, -\rho]$ and $\pi_i^+ : A_j \times Y_j \to [\rho, \bar{u}]$ such that*

$$\begin{aligned} u_i (a) + \mathbb{E}\left[\pi_i^- (a_j, y_j) \mid a\right] &= \text{constant for all } a \in A, \\ u_i (a) + \mathbb{E}\left[\pi_i^+ (a_j, y_j) \mid a\right] &= \text{constant for all } a \in A. \end{aligned}$$

Proof. This is generically satisfied if $|Y_j| \geq |A_i|$ as Yamamoto (2009b). ■

5.2 Message Protocol

As we will see, even with cheap talk, we need to let the players exchange messages via taking actions. We explain the message protocol to send a binary message $\{G, B\}$, which also gives the idea of how to dispense cheap talk in the Online Appendix. See Section 7 for the detailed discussion about the dispensability of cheap talk.

Let us introduce the following notations:

Notation 1 *Take $a_i^G, a_i^B \in A_i$ arbitrarily with $a_i^G \neq a_i^B$. For $m_i \in \{G, B\}$, we define*

1. *For $\iota \in I$, let $\mathbf{1}_{y_\iota}$ is a $|Y_\iota| \times 1$ vector such that the element corresponding to y_ι is equal to 1 and the other elements are 0.*

2. *a column vector of the signal distribution of player i conditional on a:*

$$\mathbf{q}_i(a) = (q(y_i \mid a))_{y_i \in Y_i},$$

3. *the matrix projecting player j's signal on the conditional distribution of player i's signals given an action profile a:*

$$M_{i,j}(a) = \begin{bmatrix} q(y_{i,1} \mid a, y_{j,1}) & \cdots & q(y_{i,1} \mid a, y_{j,|Y_j|}) \\ \vdots & & \vdots \\ q(y_{i,|Y_i|} \mid a, y_{j,1}) & \cdots & q(y_{i,|Y_i|} \mid a, y_{j,|Y_j|}) \end{bmatrix},$$

4. $(|Y_i| - |A_j| + 1) \times |Y_i|$ *matrix* $H_i(m_i)$ *and* $(|Y_i| - |A_j| + 1) \times 1$ *vector* $\mathbf{p}_i(m_i)$ *such that the affine hull of player i's signal distribution with respect to player j's action when player i takes* $a_i^{m_i}$ *is represented by*

$$\text{aff}(\{\mathbf{q}_i(a_i^{m_i}, a_j)\}_{a_j}) \cap \mathbb{R}_+^{|Y_i|} = \left\{\mathbf{x} \in \mathbb{R}_+^{|Y_i|} : H_i(m_i)\mathbf{x} = \mathbf{p}_i(m_i)\right\},$$

5. *the set of hyperplanes that are generated by perturbing RHS of the characterization of* $\text{aff}(\mathbf{q}_i(a_i^{m_i}, a_j)\}_{a_j} \cap \mathbb{R}_+^{|Y_i|}$*: for* $\varepsilon \geq 0$,[6]

$$\mathcal{H}_i[\varepsilon](m_i) \equiv \left\{\mathbf{x} \in \mathbb{R}_+^{|Y_i|} : \exists \boldsymbol{\varepsilon} \in \mathbb{R}_+^{|Y_i|-|A_j|+1} \text{ such that } \left\{ \begin{array}{c} \|\boldsymbol{\varepsilon}\| \leq \varepsilon \\ H_i(m_i)\mathbf{x} = \mathbf{p}_i(m_i) + \boldsymbol{\varepsilon} \end{array} \right\}\right\},$$

and

6. *the set of frequencies of player j's signal observations such that player j's conditional expectation of the frequency of player i's signal observation is in* $\mathcal{H}_i[\varepsilon](m_i)$ *when the players take* $a_i^{m_i}, a_j^G$*:*

$$\mathcal{H}_{i,j}[\varepsilon](m_i) = \left\{ \begin{array}{c} \mathbf{y} \in \mathbb{R}_+^{|Y_j|} : \exists \boldsymbol{\varepsilon} \in \mathbb{R}_+^{|Y_i|-|A_j|+1} \text{ such that} \\ \left\{ \begin{array}{c} \|\boldsymbol{\varepsilon}\| \leq \varepsilon \\ H_i(m_i) M_{i,j}(a_i^{m_i}, a_j^G)\mathbf{y} = \mathbf{p}_i(m_i) + \boldsymbol{\varepsilon} \end{array} \right. \end{array} \right\}.$$

Since a continuous linear transformation is Lipschitz continuous, we can take $\bar{K}$ such

[6] We use the sup norm unless otherwise notified: $\|\mathbf{x}\| = \max_i |x_i|$.

that, for any i, m_i and $\boldsymbol{\varepsilon} \in \mathbb{R}_+^{|Y_j|}$, $\|H_i(m_i)\boldsymbol{\varepsilon}\| \leq (\bar{K}-2)\|\boldsymbol{\varepsilon}\|$.

The following lemma is useful:

Lemma 3 *There generically exist* $\{a_i^G, a_i^B\}_{i\in I}$ *such that there exists* $\bar{\varepsilon} > 0$ *such that for any* $i, j \in I$ *and* $0 \leq \varepsilon \leq \bar{\varepsilon}$,

$$\mathcal{H}_j[\varepsilon](G) \cap \mathcal{H}_{i,j}[(\bar{K}+1)\varepsilon](G) \cap \mathcal{H}_{i,j}[(\bar{K}+1)\varepsilon](B) = \emptyset.$$

Proof. It suffices to show that, for sufficiently small ε, for any $\|\boldsymbol{\varepsilon}\| \leq \varepsilon$, there does not exist $\mathbf{x} \in \mathbb{R}_+^{|Y_j|}$ such that

$$\begin{bmatrix} H_j(G) \\ H_i(G)M_{i,j}(a_i^G, a_j^G) \\ H_i(B)M_{i,j}(a_i^B, a_j^G) \end{bmatrix} \mathbf{x} = \begin{bmatrix} \mathbf{p}_j(G) \\ \mathbf{p}_i(G) \\ \mathbf{p}_i(B) \end{bmatrix} + \boldsymbol{\varepsilon}. \tag{12}$$

For generic q, with $\boldsymbol{\varepsilon} = \mathbf{0}$, such $\mathbf{x}$ does not exist since we have $|Y_j|$ degrees of freedom and $|Y_j| + 2|Y_i| - |A_i| - 2|A_j| + 1$ constraints. Note that one row of each of $H_j(G)$, $H_i(G)M_{i,j}(a_i^G, a_j^G)$ and $H_i(B)M_{i,j}(a_i^B, a_j^G)$ is parallel to $\mathbf{1}$.

By Farkus Lemma, this is equivalent to the existence of $\mathbf{x} \in \mathbb{R}_+^{|Y_j|+2|Y_i|-|A_i|-2|A_j|+4}$ with

$$\begin{bmatrix} H_j(G) \\ H_i(G)M_{i,j}(a_i^G, a_j^G) \\ H_i(B)M_{i,j}(a_i^B, a_j^G) \end{bmatrix}' \mathbf{x} \leq \mathbf{0}, \begin{bmatrix} \mathbf{p}_j(G) \\ \mathbf{p}_i(G) \\ \mathbf{p}_i(B) \end{bmatrix} \cdot \mathbf{x} > 0.$$

Hence, for sufficiently small ε, for any $\|\boldsymbol{\varepsilon}\| \leq \varepsilon$,

$$\begin{bmatrix} H_j(G) \\ H_i(G)M_{i,j}(a_i^G, a_j^G) \\ H_i(B)M_{i,j}(a_i^B, a_j^G) \end{bmatrix}' \mathbf{x} \leq \mathbf{0}, \left(\begin{bmatrix} \mathbf{p}_j(G) \\ \mathbf{p}_i(G) \\ \mathbf{p}_i(B) \end{bmatrix} + \boldsymbol{\varepsilon} \right) \cdot \mathbf{x} > 0.$$

Again, by Farkus Lemma, (12) does not have a solution for sufficiently small ε. ■

Given above, the message protocol is determined as follows: fix $\{a_i^G, a_i^B\}_{i\in I}$ so that

Lemma 3 holds. Consider the following message protocol $(i \to_{m_i} j)$ where player i sends $m_i \in \{G, B\}$ to player j.

Player i sends the message using $T^{\frac{1-\varepsilon}{2}} + T^{1-\varepsilon}$ periods as follows.[7] As we will see, we use $\eta > 0$ in the definition of the protocol. In this section, we treat $\eta > 0$ as some small fixed number. In Section 5.3, we explain how to fix η.

1. player i decides the following two variables:

   (a) m_i, which is player i's message,

   (b) a new state variable that is equal to m_i with high probability, $z_i(m_i) \in \{G, B, M\}$, as follows:

$$z_i(m_i) = \begin{cases} m_i & \text{with probability } 1 - 2\eta, \\ \{G, B\} \setminus \{m_i\} & \text{with probability } \eta, \\ M & \text{with probability } \eta. \end{cases}$$

   In addition, if $z_i(m_i) \neq m_i$, then player i makes any action profile sequence indifferent to player j for the rest of the phase.

2. player i sends m_i by taking $a_i^{m_i}$ for $T^{\frac{1-\varepsilon}{2}}$ periods. Player j takes a_j^G. Let $S \subset \mathbb{N}$ be the set of $T^{\frac{1-\varepsilon}{2}}$ periods when player i constantly plays $a_i^{m_i}$.

3. player i constructs the random variable $\Omega_i(S)$ as follows: first, player i constructs a random variable $\{\mathbf{\Omega}_{i,t}\}_{t \in S}$ as follows. After taking $a_i^{m_i}$ and observing $y_{i,t}$, player i calculates $H_i(m_i)\mathbf{1}_{y_{i,t}}$. Then, player i draws $(|Y_i| - |A_j| + 1)$ random variables independently from the uniform distribution on $[0, 1]$. If the lth realization of these random variables is less than the lth element of $H_i(m_i)\mathbf{1}_{y_{i,t}}$, then the lth element of $\mathbf{\Omega}_{i,t}$ is equal to 1. Otherwise, the lth element of $\mathbf{\Omega}_{i,t}$ is equal to 0. Given $\{\mathbf{\Omega}_{i,t}\}_{t \in S}$, let

$$\Omega_i^H(S) = \frac{1}{|S|} \sum_{t \in S} \mathbf{\Omega}_{i,t}. \tag{13}$$

[7] Throughout the paper, we neglect the integer problem since it is handled easily.

By definition, the distribution of $\Omega_i^H(S)$ is independent of player j's strategy and $\Omega_i^H(S)$ is close to its mean $\mathbf{p}_i(m_i)$ with high probability for large $|S| = T^{\frac{1-\varepsilon}{2}}$. Further, if $|S| = T^{\frac{1-\varepsilon}{2}}$ is large, then, given $\mathbf{x} \in \mathbb{R}^{|Y_i|}$, the frequency of player i's signal observations in S, $H_i(m_i)\mathbf{x}$ and $\Omega_i^H(S)$ are sufficiently close with high probability. Hence, the following procedure does not change player j's incentive to take a_j^G while receiving the message:

(a) if

$$\left\|\Omega_i^H(S) - \mathbf{p}_i(m_i)\right\| \leq \frac{\varepsilon}{4} \wedge \left\|\Omega_i^H(S) - H_i(m_i)\mathbf{x}\right\| \leq \frac{\varepsilon}{8}, \tag{14}$$

then player i infers $m_i[1](i) = m_i$ and player i requires player j to infer m_i.

(b) otherwise, player i infers $m_i[1](i) = \emptyset$ and player i makes any action profile sequence indifferent to player j.

Since (14) implies $\|H_i(m_i)\mathbf{x} - \mathbf{p}_i(m_i)\| \leq \frac{\varepsilon}{2}\mathbf{x}$ by triangle inequality, which implies $\mathbf{x} \in \mathcal{H}_i\left[\frac{\varepsilon}{2}\right](m_i)$, for each m_i, it suffices for player j to infer the message is m_i only if $\mathbf{y} \in \mathcal{H}_{i,j}[(\bar{K}-1)\varepsilon](\hat{m}_i)$, where $\mathbf{y}$ is the frequency of player j's signal observations in S.

4. player j, on the other hand, constructs the random variable $\Omega_j^G(S)$ and $M_{i,j}^{m_i}(S)$ for each $m_i \in \{G, B\}$ as follows. Player j constructs $\mathbf{\Omega}_{j,t}^G$ and $\Omega_j^G(S)$ as player i constructs $\mathbf{\Omega}_{i,t}$ and $\Omega_j^H(S)$. Note that player j takes a_j^G to receive the message and so $m_j = G$.

   For $M_{i,j}^{m_i}(S)$, first, player j constructs a random variable $\{\mathbf{M}_{i,j,t}^{m_i}\}_{t\in S}$ as follows. After taking a_j^G and observing $y_{j,t}$, player j calculates $H_i(m_i)M_{i,j}(a_i^{m_i}, a_j^G)\mathbf{1}_{y_{j,t}}$. Then, player j draws $|Y_i| - |A_j| + 1$ random variables independently from the uniform distribution on $[0,1]$. If the lth realization of these random variables is less than the lth element of $M_{i,j}(a_i^{m_i}, a_j^G)\mathbf{1}_{y_{j,t}}$, then the lth element of $\mathbf{M}_{i,j,t}^{m_i}$ is equal to 1. Otherwise, the lth element of $\mathbf{M}_{i,j,t}^{m_i}$ is equal to 0. Given $\{\mathbf{M}_{i,j,t}^{m_i}\}_{t\in S}$, let

$$M_{i,j}^{m_i}(S) = \frac{1}{|S|}\sum_{t\in S}\mathbf{M}_{i,j,t}^{m_i}. \tag{15}$$

By definition, the distribution of $\Omega_j^G(S)$ is independent of player i's strategy. Further, if $|S|$ is large, then, given $\mathbf{y} \in \mathbb{R}^{|Y_j|}$, the frequency of player j's signal observations, "$H_j(G)\mathbf{y}$ and $\Omega_j^G(S)$," "$H_i(G)M_{i,j}(a_i^G, a_j^G)\mathbf{y}$ and $M_{i,j}^G(S)$" and "$H_i(B)M_{i,j}(a_i^B, a_j^G)\mathbf{y}$ and $M_{i,j}^B(S)$" are sufficiently close with high probability. Hence, the following procedure does not change player i's incentive to take $a_i^{m_i}$ to send m_i:

5. on the other hand, player j infers player i's state as follows: with S being the above $T^{\frac{1-\varepsilon}{2}}$ periods and $\mathbf{y}$ being the frequency of the signal observation in S,

   (a) if

$$\begin{aligned} \left\|\Omega_j^G(S) - \mathbf{p}_j(G)\right\| &\leq \frac{\varepsilon}{2} \wedge \left\|H_j(G)\mathbf{y} - \Omega_j^G(S)\right\| \leq \frac{\varepsilon}{8} \\ \wedge \left\|H_i(G)M_{i,j}(a_i^G, a_j^G)\mathbf{y} - M_{i,j}^G(S)\right\| &\leq \frac{\varepsilon}{8} \wedge \left\|H_i(B)M_{i,j}(a_i^B, a_j^G)\mathbf{y} - M_{i,j}^B(S)\right\| \leq \frac{\varepsilon}{8} \end{aligned} \tag{16}$$

   then

   i. if there exists $\hat{m}_i \in \{G, B\}$ with

$$\left\|M_{i,j}^{\hat{m}_i}(S) - \mathbf{p}_i(\hat{m}_i)\right\| \leq \bar{K}\varepsilon, \tag{17}$$

   then player j infers $m_i[2](j) = \hat{m}_i$,

   ii. otherwise, player j infers randomly: $m_i[2](j) = G$ with probability $\frac{1}{2}$ and B with $\frac{1}{2}$.

   i. and ii. are well defined since $\left\|\Omega_j^G(S) - \mathbf{p}_j(G)\right\| \leq \frac{\varepsilon}{2}$ and $\left\|H_j(G)\mathbf{y} - \Omega_j^G(S)\right\| \leq \frac{\varepsilon}{8}$ imply $\|H_j(G))\mathbf{y} - \mathbf{p}_j(G)\| \leq \varepsilon$ and $\left\|H_i(\hat{m}_i)M_{i,j}(a_i^{\hat{m}_i}, a_j^G)\mathbf{y} - M_{i,j}^{\hat{m}_i}(S)\right\| \leq \frac{\varepsilon}{8}$ and $\left\|M_{i,j}^{\hat{m}_i}(S) - \mathbf{p}_i(\hat{m}_i)\right\| \leq \bar{K}\varepsilon$ imply $\left\|H_i(\hat{m}_i)M_{i,j}(a_i^{\hat{m}_i}, a_j^G)\mathbf{y} - \mathbf{p}_i(\hat{m}_i)\right\| \leq (\bar{K}+1)\mathbf{x}$ by triangle inequality, which implies $\mathbf{y} \in \mathcal{H}_j[\varepsilon](G)(m_i) \cap \mathcal{H}_{i,j}[(\bar{K}+1)\varepsilon](\hat{m}_i)$. Hence, by Lemma 3, there is at most one $\hat{m}_i$ that satisfies ii.

   (b) otherwise, player j infers $m_i[2](j) = \emptyset$. In addition, if $m_i[2](j) = \emptyset$, then player j makes any action profile sequence indifferent to player i for the rest of the phase.

Further, player j will use $z_i(m_i)[2](j)$ below to infer m_i.

Suppose player j knows the message is m_i. Knowing m_i, $M_{i,j}(a_i^{m_i}, a_j^G)\mathbf{y}$ is the conditional expectation of $\mathbf{x} \in \mathbb{R}_+^{|Y_i|}$, the frequency of player i's signal observation. By the central limit theorem, player j puts nonnegligible posterior for $\mathbf{x}$ close to $M_{i,j}(a_i^{m_i}, a_j^G)\mathbf{y}$. Since $H_i(m_i)$ is Lipschitz continuous, it means player j puts nonnegligible posterior for $H_i(m_i)\mathbf{x}$ close to $H_i(m_i)M_{i,j}(a_i^{m_i}, a_j^G)\mathbf{y}$, that is, player j puts nonnegligible posterior on the event that $H_i(m_i)\mathbf{x}$ is close to $\mathbf{p}(m_i)$ (that is, $\mathbf{x} \in \mathcal{H}_i\left[\frac{\varepsilon}{2}\right](m_i)$) only if $H_i(m_i)M_{i,j}(a_i^{m_i}, a_j^G)\mathbf{y}$ is close to $\mathbf{p}(m_i)$ (that is, $\mathbf{y} \in \mathcal{H}_{i,j}[(\bar{K}-1)\varepsilon](m_i)$). Since $\mathbf{x} \notin \mathcal{H}_i\left[\frac{\varepsilon}{2}\right](m_i)$ implies any action profile indifferent from the previous paragraph, it suffices for player j to infer m_i only if $\mathbf{y} \in \mathcal{H}_{i,j}[(\bar{K}-1)\varepsilon](m_i)$. Even if player j does not know m_i, if suffices to infer m_i only if $\mathbf{y} \in \mathcal{H}_{i,j}[(\bar{K}-1)\varepsilon](m_i)$.

Therefore, it suffices player j to infer $\hat{m}_i$ if there is $\hat{m}_i \in \{G, B\}$ with $\mathbf{y} \in \mathcal{H}_{i,j}[(\bar{K}-1)\varepsilon](\hat{m}_i)$. This is true if 1. is the case by the following reason: if there exists $\hat{m}_i$ with $\mathbf{y} \in \mathcal{H}_{i,j}[(\bar{K}-1)\varepsilon](\hat{m}_i)$, then by definition, $\left\| H_i(\hat{m}_i)M_{i,j}(a_i^{\hat{m}_i}, a_j^G)\mathbf{y} - \mathbf{p}_i(\hat{m}_i) \right\| \leq (\bar{K}-1)\varepsilon$. Together with $\left\| H_i(\hat{m}_i)M_{i,j}(a_i^{\hat{m}_i}, a_j^G)\mathbf{y} - M_{i,j}^{\hat{m}_i}(S) \right\| \leq \frac{\varepsilon}{8}$ (coming from the fact that 1. is not the case), player j has $\left\| M_{i,j}^{\hat{m}_i}(S) - \mathbf{p}_i(\hat{m}_i) \right\| \leq \bar{K}\varepsilon$ and infers m_i.

If 2. is the case, then player j uses the result of the following message protocol. Note that this is used only after the history without (16) and that player i is indifferent to any action profile sequence. Therefore, we can let player i take any strategy.

6. player i takes either a_i^G, a_i^B, or $\frac{1}{2}a_i^G + \frac{1}{2}a_i^B$ for $T^{1-\varepsilon}$ periods. Player j takes a_j^G. Let

$$\alpha_i^{z_i(m_i)} = \begin{cases} a_i^G & \text{if } z_i(m_i) = G, \\ a_i^B & \text{if } z_i(m_i) = B, \\ \frac{1}{2}a_i^G + \frac{1}{2}a_i^B & \text{if } z_i(m_i) = M \end{cases}$$

be player i's possible strategy.

7. player j infers $z_i(m_i)$ as follows: for all $z_i \in \{G, B, M\}$, player j calculates, $T^{1-\varepsilon}\mathcal{L}(h_j, z_i)$,

the log likelihood of player j's history under z_i. With $f_{j,k}$ being a frequency of signal $y_{j,k}$ while receiving $z_i(m_i)$,

$$\mathcal{L}(h_j, z_i) = f_{j,1} \log q(y_{j,1}|a_j^G, \alpha_i^{z_i}) + \cdots + f_{j,|Y_j|} \log q(y_{j,|Y_j|}|a_j^G, \alpha_i^{z_i}).$$

Since $\mathcal{L}(h_j, z_i)$ is concave with the mixture of a_i^G and a_i^B and $\{f_{j,k}\}_{k=1}^{|Y_j|} \in \text{co}(\{0,1\}^{|Y_j|})$, which is compact, there exists $\kappa > 0$ such that one of the following three is true:

(a) the likelihood ratio of $z_i(m_i) = G$ compared to $z_i(m_i) = B$ is no less than $\exp(T^\kappa)$. Then, player j infers $z_i(m_i) = G$.

(b) the likelihood ratio of $z_i(m_i) = B$ compared to $z_i(m_i) = G$ is no less than $\exp(T^\kappa)$. Then, player j infers $z_i(m_i) = B$.

For (a) and (b), player j does not consider the case with $z_i(m_i) = M$ since if $z_i(m_i) = M$ is the case, then player j has been indifferent to any action profile sequence.

(c) the likelihood ratio of $z_i(m_i) = M$ compared to $z_i(m_i) = G, B$ is no less than $\exp(T^\kappa)$. Then, player j infers $z_i(m_i) = G$. Since player i makes any action profile sequence indifferent to player j if $z_i(m_i) = M \neq m_i \in \{G, B\}$ happens, if player j thinks that $z_i(m_i) = M$ is highly likely, then the loss from inferring G is very small.

8. based on above, player j infers m_i as follows:

$$m_i(j) = \begin{cases} m_i[2](j) & \text{if } m_i[2](j) \neq \emptyset, \\ z_i(m_i)[2](j) & \text{if } m_i[2](j) = \emptyset. \end{cases} \tag{18}$$

Further, based on the result up to 5., the players construct the following variables:

1. $\tilde{\zeta}_i(i \to_{m_i} j)$ and $\tilde{\zeta}_j(i \to_{m_i} j)$ that classify whether the summation (and so the average) of the variable constructed from a variable is close to the summation (and so the average) of the original variable. Specifically,

(a) $\tilde{\zeta}_i(i \to_{m_i} j) = B$ if $\left\|\Omega_i^H(S) - H_i(m_i)\mathbf{x}\right\| > \frac{\varepsilon}{8}$ happens. Otherwise, $\tilde{\zeta}_i(i \to_{m_i} j) = G$.

(b) $\tilde{\zeta}_j(i \to_{m_i} j) = B$ if $\left\|H_j(G)\mathbf{y} - \Omega_j^G(S)\right\| > \frac{\varepsilon}{8}$, $\left\|H_i(G)M_{i,j}(a_i^G, a_j^G)\mathbf{y} - M_{i,j}^G(S)\right\| \leq \frac{\varepsilon}{8}$ or $\left\|H_i(B)M_{i,j}(a_i^B, a_j^G)\mathbf{y} - M_{i,j}^B(S)\right\| \leq \frac{\varepsilon}{8}$ happens. Otherwise, $\tilde{\zeta}_j(i \to_{m_i} j) = G$.

As we have mentioned, conditional on $\{y_{i,t}\}_{t\in S}$, $\tilde{\zeta}_i(i \to_{m_i} j) = G$ with probability no less than $1 - \exp(-T^{\frac{1-2\varepsilon}{2}})$ and that, conditional on $\{y_{j,t}\}_{t\in S}$, $\tilde{\zeta}_j(i \to_{m_i} j) = G$ with probability no less than $1 - \exp(-T^{\frac{1-2\varepsilon}{2}})$.

2. further, let $\tilde{\vartheta}_j(i \to_{m_i} j) = G$ denote the situation where $\left\|\Omega_j^G(S) - \mathbf{p}_j(G)\right\| \leq \frac{\varepsilon}{2}$ and $\tilde{\vartheta}_j(i \to_{m_i} j) = B$ denote the situation where $\left\|\Omega_j^G(S) - \mathbf{p}_j(G)\right\| > \frac{\varepsilon}{2}$. Remember that the distribution of $\Omega_j^G(S)$ is independent of player i with mean $\mathbf{p}_j(G)$. Therefore, $\tilde{\vartheta}_j(i \to_{m_i} j) = G$ with high probability. In addition, Condition 5.(b). implies that

$$m_i[2](j) = \emptyset \Leftrightarrow \tilde{\zeta}_j(i \to_{m_i} j) = B \vee \tilde{\vartheta}_j(i \to_{m_i} j) = B \tag{19}$$

3. remember that player i makes player j indifferent to any action profile sequence if $m_i[1](i) = \emptyset$ or $z_i(m_i) \neq m_i$. We introduce $\tilde{\theta}_i(i \to_{m_i} j) = B$ to record these events. If these events do not happen, we say $\tilde{\theta}_i(i \to_{m_i} j) = G$.

4. on the other hand, player j makes player i indifferent to any action profile sequence if $m_i[2](j) = \emptyset$. We introduce $\tilde{\theta}_j(i \to_{m_i} j) = B$ to record these events. If these events do not happen, we say $\tilde{\theta}_j(i \to_{m_i} j) = G$.

The following Lemma summarizes the property of the inferences discussed above. Only the new result is 2, which follows from the fact that the length of the periods to send m_i (not $z_i(m_i)$) is $T^{\frac{1-\varepsilon}{2}}$.

Lemma 4 *There exists $\bar{\kappa}$ such that, for any $\varepsilon > 0$, for sufficiently large T,*

1. *(18) is well defined.*

2. *conditional on* m_i *and* $\tilde{\zeta}_i(S) = G$, *player* j *believes that* $m_i(j) = m_i$ *or* $\tilde{\theta}_i(i \rightarrow_{m_i} j) = B$ *with probability no less than* $1 - \exp(-T^{\bar{\kappa}})$,

3. *conditional on* m_i, *any history of player* i *and* $\tilde{\zeta}_j(S) = \tilde{\vartheta}_j(S) = G$, *any* $m_i[2](j) \in \{G, B\}$ *is possible with probability no less than* $\exp(-T^{\frac{1-\varepsilon}{2}+\varepsilon})$.

4. *the distribution of* $\tilde{\theta}_i(i \rightarrow_{m_i} j)$ *is independent of player* j*'s strategy with probability no less than* $1 - \exp(-T^{\frac{1-2\varepsilon}{2}})$.

5. *the distribution of* $\tilde{\theta}_j(i \rightarrow_{m_i} j)$ *is independent of player* i*'s strategy with probability no less than* $1 - \exp(-T^{\frac{1-2\varepsilon}{2}})$.

Proof. See Appendix. ■

There are two important features of this lemma. First, conditional on m_i, player j believes that $m_i(j) = m_i$ or any action profile sequence indifferent with high probability. This conditionality is important: player i's continuation strategy also depends on m_i. Hence, player j, after observing player j's future signals that statistically indicate player i's continuation strategy, can start to realize that $m_i(j) \neq m_i$. This lemma guarantees that even in this case, player j thinks that any action profile sequence are indifferent with high probability, which is the key to show that player j is willing to stick to the original inference.

Second, player i, after any history of player i, puts some positive probability on any inference of her own message by player j. As we will see, on the equilibrium path, the players play the game for T periods and then exchange the messages via protocol $(i \rightarrow_{m_i} j)$. After that, each player plays the continuation play based on the history in the first T periods and the inferences in the message exchange. Player i infers player j's past history in the fist T periods from her own history. Player i expects player j's continuation play from the inference of the first T periods and player i's message. Then, after player i observes player i's future signals that statistically indicate player j's continuation strategy, player i can realize that player j's continuation strategy is different from what player i expected, that is, player i's inference of player j's history in the first T periods was wrong or player j's inference of player i's message in the message exchange was wrong. Since $T^{\frac{1-\varepsilon}{2}}$ is sufficiently shorter than

T, player i can believe that, with high probability, player i's inference of player j's history in the first T periods was correct and player j's inference in the message exchange was wrong.

5.3 Variables

In this section, we finish defining the variables necessary for the equilibrium construction: $\bar{L}$, L, η, ε and κ^*.[8]

Given q_1 and q_2 in Lemma 1, we define $\bar{L}$ as follows: as we have mentioned, the reward $\pi_i(x_j, h_j^{T_P+1} : \delta)$ is usually linear in the summation of Ψ_j^a. When the slope of the reward is $\bar{L}$, the marginal increase of the expected reward by taking a_i instead of taking $\tilde{a}_i \neq a_i$ is given by $\bar{L}\,(q_2 - q_1)$ because of (9). We take $\bar{L}$ sufficiently large so that this increase is greater than the maximum gain in the instantaneous utility:

$$\bar{L}\,(q_2 - q_1) > \max_{a,\tilde{a},i} |u_i\,(a) - u_i\,(\tilde{a})|\,. \tag{20}$$

Given ρ and $\bar{L}$, we define L sufficiently large so that

$$\bar{L} < L\rho. \tag{21}$$

We are left to pin down η, ε and κ^*. Take $\varepsilon > 0$ and $\eta > 0$ sufficiently small such that Lemmas 3 and 4 are satisfied,

$$u_i\,(D_1, D_2) + \rho + 2\varepsilon\bar{L} + 2\bar{M}\bar{u}\eta < \underline{v}_i < \overline{v}_i < u_i\,(C_1, C_2) - \rho - 2\varepsilon\bar{L} - 2\bar{M}\bar{u}\eta, \tag{22}$$

and

$$\varepsilon < \frac{1}{2}, q_2 - q_1 \tag{23}$$

with

$$M \equiv 3\,(L-1) + 1. \tag{24}$$

[8] For the other variables, ρ is determined in (3) and $\bar{u}$ is determined in Lemma 2.

As we will explain in Section 6.1, we see each phase as the collection of several "rounds." (24) is the number of the total rounds.

Finally, take κ^* such that

$$\kappa^* < \bar{\kappa}, \frac{2}{3}\frac{1-\varepsilon}{2}. \tag{25}$$

6 Equilibrium Construction

Given T, we show that, for sufficiently large T, for sufficiently large δ, with

$$T_P = (L-1)\left\{T + 2\left(T^{\frac{1-\varepsilon}{2}} + T^{1-\varepsilon}\right)\right\} + T,$$

there exist $\{\{\sigma_i(x_i)\}_{x_i \in \{G,B\}}\}_i$ and $\{\{\pi_i(x_j, \cdot : \delta)\}_{x_{i+1} \in \{G,B\}}\}_i$ satisfying (4) (5) and (6).

In this section, we consider the game with cheap talk: after each stage game, each player i can send any message whose cardinality is equal to that of $\mathbb{N}$. Let $\mathfrak{m}_{i,t}$ be player i's message after period t and $\mathfrak{m}_t = (\mathfrak{m}_{i,t})_i$ be the message profile after period t. We assume that the message is perfect and public. Then, the history at the beginning of period $t \geq 1$ is given by

$$\mathfrak{m}_0, (a_{i,\tau}, y_{i,\tau}, \mathfrak{m}_\tau)_{\tau=1}^{t-1}.$$

Here, $\mathfrak{m}_0$ is the message sent at the beginning of the game.

With abuse of notation, we also let $H_i^{T_P+1}$ denote the set of all histories with cheap talk

$$\bigcup_{t=1}^{T_P+1} \left\{\mathfrak{m}_0, (a_{i,\tau}, y_{i,\tau}, \mathfrak{m}_\tau)_{\tau=1}^{t-1}\right\}.$$

See Section 7 for the dispensability of cheap talk.

6.1 Structure of the Finitely Repeated Game

Let

$$a_i(x) = \begin{cases} C_i & \text{if } x_i = G, \\ D_i & \text{if } x_i = B. \end{cases} \tag{26}$$

Intuitively, player j with $x_j = G$ who wants to make player i's value high takes $a_j(G) = C_j$ to reward player i and player j with $x_j = B$ who wants to make player i's value low takes $a_j(B) = D_j$ to punish player i.

Each phase has the following hierarchic structure: the phase is divided into L blocks and each block is divided into 5 rounds. The flow of the phase is as follows.

- at the beginning, each player i sends x_i by cheap talk. Note that x_i becomes public.
- the players play the lth normal block with $l = 1, \ldots, L$. For each l, the structure of the lth block is as follows:
  - the main round comes first. We use (l, μ) to represent this round. Each player i plays an equilibrium action $a_i(l, \mu)$ for T periods and constructs $\lambda_i(l+1) \in \{G, B\}$ at the end of the main round. The details will be explained in Section 6.3.2.
  - for $l \leq L-1$, the supplemental round 1 for $\lambda_1(l+1)$ comes next. We use $(l, \lambda_1, 1)$ to represent this round. This is the first $T^{\frac{1-\varepsilon}{2}}$ periods when player 1 sends $\lambda_1(l+1)$ via the protocol $(1 \rightarrow_{\lambda_1(l+1)} 2)$.
  - for $l \leq L-1$, the supplemental round 2 for $\lambda_1(l+1)$ comes next. We use $(l, \lambda_1, 2)$ to represent this round. This is the last $T^{1-\varepsilon}$ periods when player 1 sends $z_1(\lambda_1(l+1))$ via the protocol $(1 \rightarrow_{\lambda_1(l+1)} 2)$.
  - for $l \leq L-1$, the supplemental rounds 1 and 2 for $\lambda_2(l+1)$ comes next. $(l, \lambda_2, 1)$ and $(l, \lambda_2, 2)$ are symmetrically defined.
- after the Lth block, the report block comes, where each player sends the message about the history in all the normal blocks by cheap talk as we will explain in Section 6.4.

Even though this "block" is instantaneous with cheap talk, when we dispense cheap talk in the Online Appendix, this block becomes a collection of periods.

We attach a natural number m to each round such that m is a serial number counted from the beginning of the phase. For example, the first main round has $m = 1$ and the supplemental round 1 for $\lambda_1(2)$ has $m = 2$. In addition, we identify the name of the round with m. For example, we write $m = (1, \mu)$ if $m = 1$ and $m = (1, \lambda_1, 1)$ if $m = 2$. Moreover, if we say $(l, \mu) + 1$, then it represents the round next to (l, μ). If we say $(l, \mu) - 1$, then it represents the round before (l, μ). The similar notation is used for the other rounds. In addition, let $T(m)$ be the set of periods in the mth round.

6.2 With Conditionally Independent Signals

Following Matsushima's (2004), we first consider the case with *conditionally independent* monitoring:

$$q\left(y_j \mid a, y_i\right) = q\left(y_j \mid a\right)$$

for all a and y. That is, conditional on an action profile a, player i's signal has no information about the opponent's signal.

With conditional independence, $\lambda_1(l+1)$ and $\lambda_2(l+1)$ are irrelevant. Hence, we can omit the supplemental rounds for $\lambda_i(l+1)$. Therefore, the review phase is L collections of T-period main rounds.

Remember that we need to find $\{\{\sigma_i\left(x_i\right)\}_{x_i \in \{G,B\}}\}_i$ and $\{\{\pi_j(x_j, \cdot : \delta)\}\}_{x_j \in \{G,B\}}\}_j$ to satisfy (4), (5) and (6).

Let us specify the strategies: $\sigma_i(x_i)$ is to send the message x_i at the beginning of the phase and to always take $a_i(x_i)$ in (26).

By symmetry, we concentrate on the case with $x_j = G$. If player j receives the message $x_i = B$, then player i is supposed to take D_i. Since D_i is the dominant action, player i does not need to be incentivized by the auxiliary scenario. Consider the following constant reward function:

$$\pi_i(G, h_j^{T_P+1} : \delta) \equiv -\rho LT.$$

Since the reward is constant, player i plays D_i. Therefore, $\sigma_i(B)$ is optimal after the message $x_i = B$ and

$$\begin{aligned}
& \lim_{T\to\infty}\lim_{\delta\to 1}\frac{1-\delta}{1-\delta^{T_P}}\mathbb{E}\left[\sum_{t=1}^{T_P}\delta^t u_{i,t}\left(a\right)+\pi_i(G,h_j^{T_P+1}:\delta)\mid\sigma_i\left(B\right),\sigma_j\left(G\right)\right] \\
= & \lim_{T\to\infty}\frac{1}{T_P}\mathbb{E}\left[\sum_{t=1}^{T_P}u_i\left(D_i,C_j\right)-\rho LT\mid\sigma_i\left(B\right),\sigma_j\left(G\right)\right] \\
= & u\left(D_i,C_j\right)-\rho>\bar{v}_i \text{ from (3).}
\end{aligned} \tag{27}$$

By subtracting a proper fixed (depending only on x) positive number from the reward function, it is possible to attain $\bar{v}_i$ exactly for sufficiently large δ and T. Note that subtracting a positive number does not violate (6). Further, it is readily shown that the reward is uniformly bounded with respect to δ.

On the other hand, if player j receives the message $x_i = G$, then to incentivize player i to take C_i, player j rewards on player i in the lth block is based on

$$X_j\left(l\right) \equiv \sum_{t\in T_l}\Psi_{j,t}^{C_i,C_j},$$

the summation of $\Psi_{j,t}^{C_i,C_j}$ (remember that player j takes C_j with $x_j = G$) during the lth block. Here, T_l is the set of periods in the main round of the lth block. Consider the following reward:

$$\pi_i(G, h_j^{T_P+1} : \delta) \equiv \sum_{l=1}^{L}\pi_i(G, h_j^{T_P+1}, l)$$

with

$$\pi_i(G, h_j^{T_P+1}, l) \equiv \bar{L}\{X_j\left(l\right) - (q_2 T + 2\varepsilon T)\}_- - \rho T.$$

Intuitively, $\pi_i(G, h_j^{T_P+1}, l)$ is the reward for the lth block. Note that the reward is linearly increasing in $X_j\left(l\right)$ with slope $\bar{L}$ until $X_j\left(l\right)$ hits the upper bound $q_2 T + 2\varepsilon T$.

Since the monitoring is conditionally independent, *regardless of player i's signal observations*,

$$\Pr\left(\{X_j(l) > q_2 T + 2\varepsilon T\} \mid x_j = G, h_i\right) \leq \exp(-(q_2 - q_1) T) \leq \exp(-\varepsilon T)$$

by Hoeffding's inequality. The last inequality follows from (23). Intuitively, *regardless of player i's signal observations*, player i's belief $X_j(l) \mid x_j = G, h_i$ is approximately distributed according to $\mathcal{N}(q_2 T, O(T^{\frac{1}{2}}))$ by the central limit theorem if the monitoring is conditionally independent. Since $q_2 T + 2\varepsilon T$ is greater than the expectation of $X_j(l)$ by $2\varepsilon T^{\frac{1}{2}}$ times $T^{\frac{1}{2}}$, the order of the standard deviation, player i believes that the probability that $X_j(l)$ hits the upper bound is negligible and that the reward is linearly increasing in $X_j(l)$ with slope $\bar{L}$ with probability almost equal to one. Therefore, since (20) implies that the marginal increase of the expected reward by taking C_i instead of D_i is greater than the loss of the instantaneous utility, it is optimal for player i to take C_i always.

Therefore, $\sigma_i(G)$ is optimal after the message $x_i = G$ and

$$\begin{aligned}
& \lim_{T\to\infty}\lim_{\delta\to 1}\frac{1-\delta}{1-\delta^{T_P}}\mathbb{E}\left[\sum_{t=1}^{T_P}\delta^t u_{i,t}(a) + \pi_i(G, h_j^{T_P+1} : \delta) \mid \sigma_i(G), \sigma_j(G)\right] \\
= & \lim_{T\to\infty}\frac{1}{T_P}\mathbb{E}\left[\sum_{t=1}^{T_P} u_{i,t}(a) + \sum_{l=1}^{L}\pi_i(G, h_j^{T_P+1}, l) \mid \sigma_i(G), \sigma_j(G)\right] \\
= & u(C_i, C_j) - 2\varepsilon\bar{L} - \rho > \bar{v}_i \text{ from (3).}
\end{aligned}$$

By subtracting a proper fixed number from the reward function, it is possible to attain $\bar{v}_i$ exactly for sufficiently large δ and T. Again, we can keep (6) and uniform boundedness.

Since both $\sigma_i(G)$ and $\sigma_i(B)$ yield the same expected value $\bar{v}_i$ and are subgame perfect once message x_i is sent, both $\sigma_i(G)$ and $\sigma_i(B)$ are optimal. Therefore, we are done.

6.3 With Conditionally Dependent Signals

6.3.1 Intuitive Explanation

Now we consider the two-player prisoners' dilemma with *conditionally dependent* monitoring. To illustrate the problem, consider the case with $x_j = G$ and suppose we use the same $\sigma_i(G)$ and $\sigma_i(B)$ and the same reward as in the case with conditionally independent monitoring. A problem occurs when $x_i = G$ and player j needs to incentivize player i to take C_i. Suppose player i's signal and player j's signal are highly correlated and player i can infer $X_j(l)$ from her own history very precisely. (i) Equation (5) requires that the time average of the expected sum of the instantaneous utility and the reward function should be close to $u_i(C_i, C_j)$. (ii) Inequality (6) requires that the reward should be negative. Since (iii) the time average of the instantaneous utilities is close to $u_i(C_i, C_i)$ if player i takes C_i, (i), (ii) and (iii) together require that if $X_j(l)$ is close to its ex ante value $q_2 T$, then the time average of the reward should be close to 0, which means player i cannot be rewarded if $X_j(l)$ is unusually high. If player i infers from h_i that $X_j(l)$ is unusually high with high probability, then player i stops cooperation.

Specifically, since the reward in each lth block after receiving $x_i = G$ is $\pi_i(G, h_j^{T_P+1}, l) = \bar{L}\{X_j(l) - (q_2 T + 2\varepsilon T)\}_- - \rho T$, if player i believes that $X_j(l)$ has hit $q_2 T + 2\varepsilon T$ with high probability, then player i wants to stop cooperation. The problem is that the reward function stops increasing in $X_j(l)$ when $X_j(l) > q_2 T + 2\varepsilon T$.

Symmetrically, when we define the reward with $x_j = B$ after receiving $x_i = G$ as $\pi_i(B, h_j^{T_P+1}, l) = \bar{L}\left\{X_j(l) - (q_2 T - 2\varepsilon T)\right\}_+ + \rho T$, the problem is that this does not start to increase when $X_j(l) < q_2 T - 2\varepsilon T$.

Intuitively, we consider the following modification. For the intuitive explanation, we concentrate on the case with $x_i = G$.

Let us change the reward function to

$$\pi_i(G, h_j^{T_P+1}, l) = \bar{L}\{X_j(l) - (q_2 T + 2\varepsilon T)\} - \rho T \tag{28}$$

$$\pi_i(B, h_j^{T_P+1}, l) = \bar{L}\{X_j(l) - (q_2 T - 2\varepsilon T)\} + \rho T \tag{29}$$

for any l. Then, since the reward function is always increasing in $X_j(l)$ with slope $\bar{L}$, it is always optimal for player i to take cooperation.

Consider (6). For $x_j = G$, if $X_j(l) > q_2T + 2\varepsilon T$ occurs only for one block, then (6) is still satisfied by the following reasons. The maximum reward in one block is attained at $X_j(l) = T$ and $\pi_i(G, h_j^{T_P+1}, l) < \bar{L}T - \rho T$. On the other hand, if $X_j(l) \leq q_2T + 2\varepsilon T$, then $\pi_i(G, h_j^{T_P+1}, l) \leq -\rho T$. Since we take $\bar{L} < L\rho$ in (21), the total reward in the phase is still negative if $X_j(l) > q_2T + 2\varepsilon T$ occurs only for one block. Symmetrically, for $x_j = B$, if $X_j(l) < q_2T - 2\varepsilon T$ occurs only for one block, then (6) is still satisfied.

However, if $X_j(l) > q_2T + 2\varepsilon T$ or $X_j(l) < q_2T - 2\varepsilon T$ happens more than once, then the condition (6) can be violated. Therefore, once

$$X_j(l) \notin [q_2T - 2\varepsilon T, q_2T + 2\varepsilon T]$$

happens in the lth block, we make player j's reward function for the following blocks *will be a negative constant for* $x_j = G$ *and a positive constant for* $x_j = B$. Specifically, for the following blocks $\tilde{l}$ with $\tilde{l} \in \{l+1, \ldots, L\}$, suppose we modify the reward function to

$$\pi_i(G, h_j^{T_P+1}, \tilde{l}) = -\rho T \tag{30}$$

$$\pi_i(B, h_j^{T_P+1}, \tilde{l}) = \rho T. \tag{31}$$

Importantly, (6) is recovered since $\pi_i(G, h_j^{T_P+1}, l) \leq -\rho T$ and $\pi_i(B, h_j^{T_P+1}, l) \geq \rho T$ except for one block.

Let $\lambda_j(l) \in \{G, B\}$ denote which reward function player j is using in the lth block: if player j is using (28) or (29), then $\lambda_j(l) = G$ and if using (30) or (31), then $\lambda_j(l) = B$. Note that ex ante, if the players play $a(x)$, then $X_j(l)$ is approximately distributed according to the normal distribution with expectation q_2T and standard deviation $O(T^{\frac{1}{2}})$. Since $X_j(l) \notin [q_2T - 2\varepsilon T, q_2T + 2\varepsilon T]$ implies $X_j(l)$ is far way from its ex ante value by $2T^{\frac{1}{2}}$ times the order of standard deviation, $\lambda_j(l) = B$ happens only with small probability less than $\exp(-\varepsilon T)$. Hence, we call the observation is "erroneous" if $X_j(l) \notin [q_2T - 2\varepsilon T, q_2T + 2\varepsilon T]$.

On the other hand, let $\hat{\lambda}_j(l) \in \{G, B\}$ be player i's inference of $\lambda_j(l)$. Forgetting the question of how player i infers $\lambda_j(l)$ for a while, assume $\hat{\lambda}_j(l)$ is always correct: $\hat{\lambda}_j(l) = \lambda_j(l)$ for all l.

If $\hat{\lambda}_j(l) = B$, then since the reward is constant for the rest of the phase, it is optimal to take D_i. If $\hat{\lambda}_j(l) = G$, then the reward is linearly increasing in $X_j(l)$ with slope $\bar{L}$. Hence, if player i's continuation payoff is not changed when $\hat{\lambda}_j(l+1) = B$ is induced, then player i is willing to take C_i always in the lth block. To verify the incentive to take cooperation when $\hat{\lambda}_j(l) = G$, we do the backward induction. For $l = L$ and $\hat{\lambda}_j(L) = G$, player i wants to take C_i since there is no effect on $\lambda_j(L+1)$.

Suppose that player i takes C_i when $\hat{\lambda}_j(l+1) = G$ and let us verity player i takes C_i when $\hat{\lambda}_j(l) = G$. In the intuitive explanation here, we assume player j with x_j always takes $a_j(x_j)$ although, symmetrically to player i, player j also changes her own action depending on $\hat{\lambda}_i(l)$, that is, depending on whether player j believes player i has observed an erroneous history.

For $x_j = G$, when $\hat{\lambda}_j(l+1) = G$, the limit of the discounted average payoff from the main round of the $(l+1)$th block satisfies

$$\begin{aligned} & \lim_{T\to\infty}\lim_{\delta\to 1}\frac{1-\delta}{1-\delta^{T_P}}\mathbb{E}\left[\sum_{t\in T_{l+1}}\delta^{t-1}u_i(C_i,C_j) - \bar{L}\{X_j(l+1) - (q_2T + 2\varepsilon T)\} - \rho T \mid C_i, C_j\right] \\ = & \frac{1}{L}\left\{u(C_i,C_j) - 2\varepsilon\bar{L} - \rho\right\} \geq \frac{1}{L}\bar{v}_i \end{aligned} \tag{32}$$

since the expected value of $X_j(l)$ is q_2T. On the other hand, when $\hat{\lambda}_j(l+1) = B$, it satisfies

$$\begin{aligned} & \lim_{T\to\infty}\lim_{\delta\to 1}\frac{1-\delta}{1-\delta^{T_P}}\left[\sum_{t\in T_{l+1}}\delta^{t-1}u_i(D_i,C_j) - \rho T\right] \\ = & \frac{1}{L}\{u(D_i,C_j) - \rho\} \geq \frac{1}{L}\bar{v}_i. \end{aligned} \tag{33}$$

We can subtract proper positive numbers depending only on x and $\lambda_j(l+1)$ from (32) and (33) such that both attain $\frac{1}{L}\bar{v}_i$. Note that we can keep (6) and uniform boundedness.

For $x_j = B$, when $\hat{\lambda}_j(l+1) = G$, the limit of the discounted average payoff from the main round of the $(l+1)$th block satisfies

$$\begin{aligned} & \lim_{T\to\infty}\lim_{\delta\to 1}\frac{1-\delta}{1-\delta^{T_P}}\mathbb{E}\left[\sum_{t\in T_{l+1}}\delta^{t-1}u_i(C_i,D_j)+\bar{L}\{X_j\left(l+1\right)-\left(q_2T-2\varepsilon T\right)\}+\rho T\mid C_i,D_j\right] \\ = & \ \frac{1}{L}\left\{u\left(C_i,D_j\right)+2\varepsilon\bar{L}+\rho\right\}\leq\frac{1}{L}\underline{v}_i \end{aligned} \quad (34)$$

since the expected value of $X_j\left(l\right)$ is q_2T. On the other hand, when $\hat{\lambda}_j(l+1) = B$, the limit of the discounted average payoff from the main round of the $(l+1)$th block satisfies

$$\begin{aligned} & \lim_{T\to\infty}\lim_{\delta\to 1}\frac{1-\delta}{1-\delta^{T_P}}\left[\sum_{t\in T_{l+1}}\delta^{t-1}u_i(D_i,D_i)+\rho T\right] \\ = & \ \frac{1}{L}\left\{u\left(D_i,D_j\right)+\rho\right\}\leq\frac{1}{L}\underline{v}_i. \end{aligned} \quad (35)$$

We can add proper positive numbers depending only on x and $\lambda_j(l+1)$ to (34) and (35) such that both attain $\frac{1}{L}\underline{v}_i$. Note again that we can keep (6) and uniform boundedness.

Therefore, player i currently in block l with $\hat{\lambda}_j(l) = G$ is willing to take C_i always in block l since (i) the reward is linearly increasing in $X_j(l)$ with slope $\bar{L}$ and (ii) player i's continuation payoff is the same between $\lambda_j(l+1) = G$ and $\lambda_j(l+1) = B$.

In summary, $\sigma_i(G)$ is such that player i with $\hat{\lambda}_j(l) = G$ plays C_i and with $\hat{\lambda}_j(l) = B$ plays D_i.

6.3.2 Formal Description of $\sigma_i(x_i)$

The formal description of $\sigma_i(x_i)$ is more complicated since player i needs to infer $\lambda_j(l)$ and player j changes her action based on $\hat{\lambda}_i(l)$. For the structure of the phase, see Section 6.1. Each player has the following state variables: the state of the phase $x_i \in \{G, B\}$, the record of erroneous histories $\lambda_i(l) \in \{G, B\}$ and the inference of $\lambda_j(l)$: $\hat{\lambda}_j(l) \in \{G, B\}$.

The state of the phase x_i is determined once and for all at the beginning of the phase. In the main round of the lth block, each player i plays an equilibrium action $a_i(l, \mu)$ for T

periods, where

$$a_i(l,\mu) = \begin{cases} a_i(x) & \text{if } \hat{\lambda}_j(l) = G, \\ D_i & \text{if } \hat{\lambda}_j(l) = B. \end{cases}$$

Note that, as in the intuitive explanation above, player i switches to D_i if $\hat{\lambda}_j(l) = B$.

After the main round of each lth block, player j updates $\lambda_j(l)$ as follows: if $x_i = B$, then $\lambda_j(l+1) = G$ always. If $x_i = G$, then the initial condition is $\lambda_j(1) = G$. For $l \geq 1$, given $X_j(l) = \sum_{t \in T_l} \Psi_{j,t}^{C_i, a_j(x_j)}$, $\lambda_j(l+1)$ is determined as

$$\lambda_j(l+1) = \begin{cases} \lambda_j(l) & \begin{array}{c} \text{with probability } 1-\eta \text{ if } X_j(l) \in [q_2 T - 2\varepsilon T, q_2 T + 2\varepsilon T], \\ \text{with probability } \eta \text{ if } X_j(l) \notin [q_2 T - 2\varepsilon T, q_2 T + 2\varepsilon T], \end{array} \\ B & \begin{array}{c} \text{with probability } \eta \text{ if } X_j(l) \in [q_2 T - 2\varepsilon T, q_2 T + 2\varepsilon T], \\ \text{with probability } 1-\eta \text{ if } X_j(l) \notin [q_2 T - 2\varepsilon T, q_2 T + 2\varepsilon T]. \end{array} \end{cases} \tag{36}$$

Here, with abuse of notation, even if $\lambda_j(l) = B$, we see the event $\lambda_j(l+1) = \lambda_j(l)$ is different from $\lambda_j(l+1) = B$.

The basic structure is as follows: with high probability $1-\eta$, $\lambda_j(l+1) = B$ is the record of the events that player j observes an erroneous history as in the intuitive explanation. However, there exists a positive probability $\eta > 0$ with which $\lambda_j(l+1)$ "transits to the opposite state." As we will see, if $\lambda_j(l+1)$ transits to the opposite state, then player j makes any action profile sequence indifferent to player i from the $(l+1)$th block.

Then, in the supplemental rounds for $\lambda_j(l+1)$, player j sends $\lambda_j(l+1)$ to player i via the protocol $(i \rightarrow_{\lambda_j(l+1)} j)$ as explained in Section 5.2 with $m_i = \lambda_i(l+1)$. Remember $\lambda_j(l+1)(i)$ is player i's inference, which is, from Lemma 4, *conditional on* $\lambda_j(l+1)$, player i believes that $\lambda_j(l+1)(i) = \lambda_j(l+1)$ or any action profile sequence will be indifferent with high probability. In particular, this means that after player i realizes that $\lambda_j(l+1)(i) \neq \lambda_j(l+1)$, player i believes that any action profile sequence is indifferent with high probability.

Each player i, then, constructs $\hat{\lambda}_j(l+1)$, the inference of $\lambda_j(l+1)$, based on her own history in the lth main round and $\lambda_j(l+1)(i)$. If $x_i = B$, then $\hat{\lambda}_j(l+1) = G$ always. Since $\lambda_j(l+1) = G$ always for $x_i = B$, this is the correct inference of $\lambda_j(l+1)$.

If $x_i = G$, then the initial condition is $\hat{\lambda}_j(1) = G = \lambda_j(1)$. For $l \geq 1$, player i calculates the conditional expectation of $X_j(l)$, on which $\lambda_j(l+1)$ depends:

$$\mathbb{E}\left[X_j(l) \mid a(x), \{y_{i,t}\}_{t \in T_l}\right] = \sum_{t \in T_l} \sum_{y_{j,t}} \psi_j^{a(x)}(y_{j,t}) q(y_{j,t} \mid a(x), y_{i,t}). \tag{37}$$

In addition, player i constructs the two variables related to (37):

$$E_i X_j(l) = \sum_{t \in T_l} (E_i \Psi_j^{a(x)})_t \text{ and } \sum_{t \in T_l} \mathbb{E}\left[\Psi_{j,t}^{a(x)} \mid a(x), y_{i,t}\right].$$

Remember that, player i constructs $(E_i \Psi_j^{a(x)})_t$ from $\sum_{y_{j,t}} \psi_j^a(y_{j,t}) q(y_{j,t} \mid a(x), y_{i,t})$. For large T, with high probability, $E_i X_j(l)$ and $\mathbb{E}\left[X_j(l) \mid a(x), \{y_{i,t}\}_{t \in T_l}\right]$ are close to each other by the law of large numbers:

$$\left|E_i X_j(l) - \mathbb{E}\left[X_j(l) \mid a(x), \{y_{i,t}\}_{t \in T_l}\right]\right| \leq \frac{1}{8} \varepsilon T. \tag{38}$$

Let $\tilde{\zeta}_i(l, \mu) = G$ denote the case with (38) and $\tilde{\zeta}_i(l, \mu) = B$ denote the case without (38).

Suppose the players play $a(x)$ and player i has

$$E_i X_j(l) \in [q_2 T - \frac{1}{2} \varepsilon T, q_2 T + \frac{1}{2} \varepsilon T] \tag{39}$$

and $\tilde{\zeta}_i(S) = G$. Then, player i's posterior on

$$X_j(l) \in [q_2 T - 2\varepsilon T, q_2 T + 2\varepsilon T] \tag{40}$$

is less than $\exp(-T^{1-\varepsilon})$ by the following reason: by triangle inequality, we have

$$\mathbb{E}\left[X_j(l) \mid a(x), \{y_{i,t}\}_{t \in T_l}\right] \in [q_2 T - \varepsilon T, q_2 T + \varepsilon T].$$

Since the expectation of $X_j(l)$ is inside of $[q_2 T - 2\varepsilon T, q_2 T + 2\varepsilon T]$ by $\varepsilon T^{\frac{1}{2}}$ times $T^{\frac{1}{2}}$, the order of the standard variation of the conditional distribution $X_j(l) \mid a(x), \{y_{i,t}\}_{t \in T_l}$, the result follows from Hoeffding's inequality.

In summary, we have the following lemma:

Lemma 5 *For any $\varepsilon > 0$, for sufficiently large T, for all $x \in \{G, B\}^2$,*

1. *if player j plays $a_j(x)$ in (l, μ), then the distribution of $E_j X_i(l)$ is independent of $\{a_{i,t}\}_{t \in S}$ by (11).*

2. *conditional on any $\{a_{j,t}, y_{j,t}\}_{t \in T(l,\mu)}$, $\tilde{\zeta}_j(S) = G$ with probability no less than $1 - \exp(-T^{1-\frac{\varepsilon}{2}})$ by the central limit theorem.*

3. *if the players play $a(x)$ and player i has (39) and $\tilde{\zeta}_i(l, \mu) = G$, then player i's posterior on (40) is less than $\exp(-T^{1-\varepsilon})$.*

Based on above, on the equilibrium path,

- if (39) and $\tilde{\zeta}_i(l, \mu) = G$, then player i is "ready to listen to $\lambda_j(l+1)$" with probability η,

- otherwise, player i is "ready to listen to $\lambda_j(l+1)$" with probability 1.

If player i has deviated from $\sigma_i(x_i)$, then player i is ready to listen always. In addition, if player i is ready to listen, then player i makes any action profile sequence indifferent to player j.

$\hat{\lambda}_j(l+1)$ is determined as

$$\hat{\lambda}_j(l+1) = \begin{cases} \lambda_j(l+1)(i) & \text{when player } i \text{ is ready to listen,} \\ \hat{\lambda}_j(l) & \text{otherwise.} \end{cases}$$

That is, player i uses the inference of the message about $\lambda_j(l+1)$ to infer $\lambda_j(l+1)$ if player i is ready to listen. Otherwise, player i neglects the message.

The basic structure of the above equilibrium construction is summarized as follows. Since player i does not know about $\lambda_j(l+1)$, we let player j tell $\lambda_j(l+1)$. To incentivize player

j to tell the truth, we make sure that whenever the message about $\lambda_j(l+1)$ has an impact on player i's equilibrium action (that is, whenever player i is ready to listen), player i makes any action profile sequence indifferent to player j. This can be done without affecting the equilibrium payoff only if player i neglects player j's message with high probability on the equilibrium path. To incentivize player i to neglect the message, we let player j transit to "the opposite state" with small probability η in (36). If player j transits to the opposite state in the lth block, then player j makes any action profile sequence indifferent to player i. Then, player i with $\tilde{\zeta}_i(l,\mu) = G$ and (39) can believe that, even after player i knows the message is $\lambda_j(l+1) = B$, because of Condition 3 of Lemma 5, player j has transit to the opposite state with high probability and player j makes any action profile sequence indifferent to player i with high probability. Hence, with high probability, neglecting player j's message and taking $a_i(x_i)$ is optimal. If player i does not have $\tilde{\zeta}_i(l,\mu) = G$ and (39), on the other hand, player i is ready to listen. From Lemma 4, the message protocol gives player i almost correct inference of the message $\lambda_j(l+1)$, we are done.

We also introduce three state variables ζ_j, ϑ_j and θ_j which change at the end of each round. In Section 5 and this subsubsection, we define $\tilde{\zeta}_i$ and $\tilde{\zeta}_j$ in the various ways. With abuse of notation, we define $\tilde{\zeta}_j(m) = B$ if one of the following conditions is satisfied:

1. $m = (l,\mu)$ (main round) and $\tilde{\zeta}_j(l,\mu) = B$ happens as we just described,

2. $m = (l,\lambda_j,1)$ (when player j is a message sender) and $\tilde{\zeta}_j(j \rightarrow_{\lambda_j(l+1)} i)$, or

3. $m = (l,\lambda_i,1)$ (when player j is a message receiver) and $\tilde{\zeta}_j(i \rightarrow_{\lambda_i(l+1)} j)$.

Otherwise, $\tilde{\zeta}_j(m) = G$. Given $\tilde{\zeta}_j$, let $\zeta_j(m+1) = G$ denote $\tilde{\zeta}_j(\tilde{m}) = G$ for all $\tilde{m} \leq m$ and $\zeta_j(m+1) = B$ denote all the other cases.

In addition, we define $\tilde{\vartheta}_j(m) = B$ if $m = (l,\lambda_i,1)$ (when player j is a message receiver) and $\tilde{\vartheta}_j(i \rightarrow_{\lambda_i(l+1)} j) = B$. Otherwise, $\tilde{\vartheta}_j(m) = G$. Similarly to ζ_j, let $\vartheta_j(m+1) = G$ denote $\tilde{\vartheta}_j(\tilde{m}) = G$ for all $\tilde{m} \leq m$ and $\vartheta_j(m+1) = B$ denote all the other cases.

The following lemma is useful. The intuition is that, for each m, since conditional on $\{y_{j,t}\}_{t \in T(m)}$, $\tilde{\zeta}_j(m) = B$ happens only with a small probability, not conditioning on

$\tilde{\zeta}_j(m) = G$ does not change player i's posterior so much.

Lemma 6 *For any m, suppose the players take a for all $t \in T(m)$. Take $A \subset Y_j^{|T(m)|}$ that does not depend on $\tilde{\zeta}_j(m)$. Player i calculates the belief of $\{y_{j,t}\}_{t \in T(m)} \in A$ from player i's history in the mth round, $\{y_{i,t}\}_{t \in T(m)}$, conditional on $a_t = a$ for all $t \in T(m)$: $\Pr(\{\{y_{j,t}\}_{t \in T(m)} \in A\} \mid a, \{y_{i,t}\}_{t \in T(m)})$. This is close to the the true posterior conditional on $\tilde{\zeta}_j(m) = G$:*

$$\begin{aligned} & \left| \Pr(\{\{y_{j,t}\}_{t \in T(m)} \in A\} \mid a, \{y_{i,t}\}_{t \in T(m)}) - \Pr(\{\{y_{j,t}\}_{t \in T(m)} \in A\} \mid \tilde{\zeta}_j(m) = G, a, \{y_{i,t}\}_{t \in T(m)}) \right| \\ < \ & 2 \exp(-|T(m)|^{1-\frac{\varepsilon}{2}}). \end{aligned}$$

Proof. See Appendix. ■

As we have seen, the players make the opponent indifferent to any action profile sequence after some histories. $\theta_i(m) \in \{G, B\}$ is used for encoding the events that player i makes player j indifferent to any action profile from (and including) the mth round.

We now define the transition of $\theta_j(m) \in \{G, B\}$. We let player j make any action profile sequence indifferent to player i from (and including) the mth round if $\theta_j(m) = B$. The initial condition is that $\theta_j(1) = G$. The transition is determined as follows. We have $\theta_j(m+1) = B$ if one of the following conditions is satisfied:

1. $\theta_j(m) = B$. That is, once $\theta_j(m) = B$ is induced, it lasts until the end of the phase.

2. $m = (l, \mu)$ (the main round of the lth block) and $\lambda_j(m+1)$ transits to the opposite state as we have seen in (36). Note that this happens randomly with probability η regardless of player j's history.

3. $m = (l, \mu)$ (the main round of the lth block) and player j is ready to listen.

4. $m = (l, \lambda_j, 1)$ (the supplemental round 1 for $\lambda_j(l+1)$) and $\tilde{\theta}_j(j \to_{\lambda_j(l+1)} i) = B$ happens.

5. $\tilde{\zeta}_j(m) = B$ or $\tilde{\vartheta}_j(m) = G$ happens. Especially, (19) shows that, for $m = (l, \lambda_i, 2)$ (the supplemental round 1 for $\lambda_i(l+1)$), if player j has $\lambda_i(l+1)[2](j) = \emptyset$, then $\theta_j(m) = B$.

Note that if $\hat{\lambda}_i(l) = B$ or $a_j(l, \mu) \neq a_j(x_j)$, then there exists $\tilde{l} \leq l-1$ such that $\hat{\lambda}_i(\tilde{l}+1) = B$ is newly induced at the $\tilde{l}$th block (note that player j with $\hat{\lambda}_i(l) = G$ takes $a_j(x_j)$). Since player j needs to be ready to listen to $\lambda_i(\tilde{l}+1)$, Condition 3. of θ_j is satisfied. Hence,

$$\theta_j(l, \mu) = B. \tag{41}$$

From Lemmas 8, 4 and 5, the distribution of $\{\theta_j(\tilde{m}+1)\}_{\tilde{m} \geq m}$ is independent of player i's strategy in the mth round with probability at least $1 - \exp(-T^{\frac{1-2\varepsilon}{2}})$. To show this for Condition 3. of θ_j, conditional on $\theta_j(l, \mu) = G$ (this is the only case when the distribution of $\theta_j((l, \mu)+1)$ is nontrivial), player j takes $a_j(x)$ and Lemma 5 guarantees that $\tilde{\zeta}_j(l, \mu) = G$ and (39) with probability no less than $1 - \exp(-T^{1-\varepsilon})$.

Among all the events that induce $\theta_j(m) = B$, we let player i exclude $\zeta_j(m) = B$ or $\vartheta_j(m) = B$ from the consideration and condition that $\zeta_j(m) = \vartheta_j(m) = G$. This is rational since if the current state is $\theta_j(m) = B$, then any action profile sequence is indifferent. From (19), $\zeta_j(m) = \vartheta_j(m) = G$ implies player j does not use $z_i(\lambda_i(\tilde{l}+1))[2](j)$ to infer $\lambda_i(\tilde{l}+1)(j)$.

Let us formally show the almost optimality of the inference on the equilibrium path.[9] With $\sigma_i(B)$, $\hat{\lambda}_j(l) = \lambda_j(l)$ always. Hence, we concentrate on $\sigma_i(G)$. Specifically, we show that, if player i has not deviated in the supplemental rounds for λ_j, for each mth round included in the lth block, conditional on $\zeta_j(m) = \vartheta_j(m) = G$ and player j's action up to (and including) the mth round but excluding the mixed strategy[10] that player j has taken, player i can believe that $\hat{\lambda}_j(l) = \lambda_j(l)$ or $\theta_j(m) = B$ with probability no less than

$$1 - M \exp(-T^{\kappa^*}) \tag{42}$$

with M being the total number of all the rounds by the end of the normal blocks.

[9] See (42) below for the formal definition of almost optimality of the inference.

[10] When we say the mixed strategy, it refers to the mixed strategy within a round and does not refer to the randomization at the beginning of a round.

From now on, we condition $\zeta_j(m) = \vartheta_j(m) = G$ and that player i has not deviated in the supplemental rounds for λ_j.

Since once $\lambda_j(\tilde{l}) = B$ is induced, then $\lambda_j(\tilde{l}') = B$ for all the following blocks, there exists a unique l^* such that $\lambda_j(\tilde{l}) = B$ is initially induced in the (l^*+1)th block: $\lambda_j(1) = \cdots = \lambda_j(l^*) = G$ and $\lambda_j(l^*+1) = \cdots = \lambda_j(L) = B$. Similarly, there exists $\hat{l}^*$ with $\hat{\lambda}_j(1) = \cdots = \hat{\lambda}_j(\hat{l}^*) = G$ and $\hat{\lambda}_j(\hat{l}^*+1) = \cdots = \hat{\lambda}_j(L) = B$. If $\lambda_j(L) = G$ ($\hat{\lambda}_j(L) = G$, respectively), then define $l^* = L$ ($\hat{l}^* = L$, respectively).

Then, there are following three cases:

- $l^* = \hat{l}^*$: this means $\lambda_j(l) = \hat{\lambda}_j(l)$ for all l as desired.

- $l^* > \hat{l}^*$: this means $\lambda_j(l) = \hat{\lambda}_j(l)$ for all $l \leq \hat{l}^*$. Hence, it suffices to show that player i believes $\theta_j(m) = B$ for $m \geq (\hat{l}^*+1, \mu)$ with probability no less than $1 - \exp(-T^{\kappa^*})$.

  Conditional on $\zeta_j(m) = \vartheta_j(m) = G$, we have $\zeta_j(\hat{l}^*, \lambda_j, 1) = \vartheta_j(\hat{l}^*, \lambda_j, 1) = G$. In $(\hat{l}^*, \lambda_j, 1)$ and $(\hat{l}^*, \lambda_j, 2)$ (supplemental rounds for λ_j), for $\hat{\lambda}_j(\hat{l}^*+1) = B$ to be induced, player i's inference should be $\lambda_j(\hat{l}^*+1)(i) = B$. Further, for $m = (\hat{l}^*, \lambda_j, 1)$ and $(\hat{l}^*, \lambda_j, 2)$, $\zeta_j(m+1) = B$ or $\vartheta_j(m+1) = B$ happens if and only if $\tilde{\zeta}_j(m) = B$. Hence, from Lemma 4, conditional on $\lambda_j(\hat{l}^*+1)$ and $\zeta_j((\hat{l}^*, \lambda_j, 2)+1) = \vartheta_j((\hat{l}^*, \lambda_j, 2)+1) = G$, player i believes $\lambda_j(\hat{l}^*+1)(i) = \lambda_j(\hat{l}^*+1)$ or $\theta_j((\hat{l}^*, \lambda_j, 1)+1) = B$ with probability no less than $1 - \exp(-T^{\kappa^*})$. Since $l^* > \hat{l}^*$ implies $\lambda_j(\hat{l}^*+1) = G$, only the possible case is $\theta_j((\hat{l}^*, \lambda_j, 1)+1) = B$ as desired. Since we condition $\lambda_j(\hat{l}^*+1)$ and $\zeta_j((\hat{l}^*, \lambda_j, 2)+1) = \vartheta_j((\hat{l}^*, \lambda_j, 2)+1) = G$, learning and conditioning on $\tilde{\zeta}_j(m) = \tilde{\vartheta}_j(m) = G$ in the other rounds do not affect the posterior.

- $l^* < \hat{l}^*$: this means $\lambda_j(l) = \hat{\lambda}_j(l)$ for all $l \leq l^*$. Hence, it suffices to show that player i believes $\theta_j(m) = B$ for $m \geq (l^*+1, \mu)$ with probability no less than $1 - \exp(-T^{\kappa^*})$.

  We can condition that player j has $\hat{\lambda}_i(l^*) = G$ and takes $a_j(x)$ in the l^*th block (otherwise, $\theta_j(l^*, \mu) = B$ from (41) and we are done). In addition, since $\hat{\lambda}_j(l^*) = G$, player i takes $a_i(x)$ in the l^*th block. Hence, the players play $a(x)$ in the l^*th block. There are following two subcases:

– player i was ready to listen in the l^*th block: by the same reason as above, we are done.

– player i was not ready to listen in the l^*th block (this means player i has not deviated at the end of the main round of the l^*th block).

We condition $\zeta_j(l^*, \mu) = \vartheta_j(l^*, \mu) = G$. Consider $m = (l^*, \mu)$. Suppose, for a while, forget about conditioning on $\tilde{\zeta}_j(l^*, \mu) = G$. $\tilde{\vartheta}_j(l^*, \mu) = G$ for the main round by definition.

For player i not to be ready to listen, $\mathbb{E}\left[X_j(l^*) \mid a(x), \{y_{i,t}\}_{t \in T_{l^*}}\right]$ needs to be within $[q_2 T - \varepsilon T, q_2 T + \varepsilon T]$. Note that $\mathbb{E}[X_j(l^*) \mid a(x), \{y_{i,t}\}_{t \in T_{l^*}}] \in [q_2 T - \varepsilon T, q_2 T + \varepsilon T]$ implies that the conditional expectation is inside of $[q_2 T - 2\varepsilon T, q_2 T + 2\varepsilon T]$ by more than εT. Since the order of the standard deviation of the conditional distribution of $X_j(l^*) \mid a, h_i$ is $T^{\frac{1}{2}}$, this implies that player i' belief on $X_j(l^*) \notin [q_2 T - 2\varepsilon T, q_2 T + 2\varepsilon T]$ is no more than $\exp(-\varepsilon T)$.

Remember that player j sends the message $\lambda_j(l^* + 1) = B$ with a fixed probability $\eta > 0$ even if $X_j(l^*) \in [q_2 T - 2\varepsilon T, q_2 T + 2\varepsilon T]$. Therefore, player i knowing $\lambda_j(l^* + 1) = B$ believes with probability no less than $1 - \frac{1-\eta}{\eta} \exp(-\varepsilon T)$, player j transits to $\lambda_j(l^* + 1) = B$ although $X_j(l^*) \in [q_2 T - 2\varepsilon T, q_2 T + 2\varepsilon T]$, which means $\theta_j((l^*, \mu) + 1) = B$ and any action profile sequence is indifferent. Hence, having $\hat{\lambda}_j(l^* + 1) = G$ and taking C_i is also optimal with probability no less than $1 - \frac{1-\eta}{\eta} \exp(-\varepsilon T)$.

From Lemma 6, after conditioning on $\tilde{\zeta}_j(l^*, \mu) = \tilde{\vartheta}_j(l^*, \mu) = G$, player i believes $\theta_j((l^*, \mu) + 1) = B$ with probability no less than $1 - 2\exp(-T^{1-\varepsilon})$.

Player i can infer $X_j(l^*)$ further from player j's continuation strategy since player j's future action depends on $\hat{\lambda}_i(l^* + 1)$, which depends on player j's history in the main round of the l^*th block, which, in turn, constructs $X_j(l^*)$.

Suppose player i knows $\hat{\lambda}_i(l^* + 1) = \hat{\lambda}_i \in \{G, B\}$. Conditional on $\tilde{\zeta}_j(l^*, \mu) = \tilde{\vartheta}_j(l^*, \mu) = G$, player j is ready to listen to player i's message $\lambda_i(l^* + 1)$ with probability η and Lemma 4 implies that any inference $\lambda_i(l^* + 1)[2](j)$ is possible with probability at least $\exp(-T^{\frac{1-\varepsilon}{2}+\varepsilon})$ conditional on $\tilde{\zeta}_j(l^*, \lambda_i, 1) = \tilde{\vartheta}_j(l^*, \lambda_i, 1) = G$. Therefore, the minimum likelihood for $\lambda_j(l^*+1) = B$ with $\theta_j((l^*, \mu)+1) = B$ against $\lambda_j(l^*+1) = B$ with $\theta_j((l^*, \mu)+1) = G$ conditional on $\tilde{\zeta}_j(l^*, \mu) = \tilde{\vartheta}_j(l^*, \mu) = G$, $\tilde{\zeta}_j(l^*, \lambda_i, 1) = \tilde{\vartheta}_j(l^*, \lambda_i, 1) = G$ and $\hat{\lambda}_i(l^*+1) = \hat{\lambda}_i$ is no less than

$$\frac{(1 - 2\exp(-T^{1-\varepsilon}))\,\eta\exp(-T^{\frac{1-\varepsilon}{2}+\varepsilon})}{2\exp(-T^{1-\varepsilon})} > \frac{1}{2}\eta\exp(T^{\frac{1-3\varepsilon}{2}}) > \exp(T^{\kappa^*}) \qquad (43)$$

for sufficiently large T from (25) as desired.

After conditioning on $\lambda_j(l^*+1)$, $\hat{\lambda}_i(l^*+1)$, $\tilde{\zeta}_j(l^*, \mu) = \tilde{\vartheta}_j(l^*, \mu) = G$ and $\tilde{\zeta}_j(l^*, \lambda_j, 1) = \tilde{\vartheta}_j(l^*, \lambda_j, 1) = G$, learning and conditioning on $\tilde{\zeta}_j(m) = \tilde{\vartheta}_j(m) = G$ in the other rounds do not affect the posterior.

6.3.3 Reward Function

Now we formally show that there exists $\pi_i(x_j, h_j^{T_P+1} : \delta)$ that satisfies (4), (5) and (6) for $\sigma(x)$ with $x \in \{G, B\}^2$ with $\pi_i(x_j, h_j^{T_P+1} : \delta)$ being uniformly bounded with respect to δ. We see $\pi_i(x_j, h_j^{T_P+1} : \delta)$ as the summation of $\pi_i(x_j, h_j^{T_P+1}, m : \delta)$, which is the reward on the mth round.

For any m, if $\theta_j(m) = B$, then player j uses

$$\pi_i(x_j, h_j^{T_P+1}, (l, \mu)) = \begin{cases} \sum_{t \in T(m)} \delta^{t-1} \pi_i^-(a_{j,t}, y_{j,t}) & \text{if } x_j = G, \\ \sum_{t \in T(m)} \delta^{t-1} \pi_i^+(a_{j,t}, y_{j,t}) & \text{if } x_j = B \end{cases}$$

to make player i indifferent to any action profile sequence. Below, we concentrate on $\theta_j(m) = G$. Since the distribution of $\{\theta_j(\tilde{m}+1)\}_{\tilde{m} \geq m}$ is independent of player i's strategy in the mth round with probability at least $1 - \exp(-T^{\frac{1-2\varepsilon}{2}})$, player i neglects the effect of her strategy in the mth round on $\theta_j(m+1)$.

For $m = (l, \mu)$ (the main round), if player j receives the message $x_i = B$, then since D_i is a stage-game dominant strategy, no reward is necessary to incentivize player i. Therefore, the same argument as in Section 6.2 works. If player j receives the message $x_i = G$, then player j uses the reward function increasing in $X_j(l)$ while $\lambda_j(l) = G$ and the constant reward function while $\lambda_j(l) = B$. Specifically,

$$\pi_i(x_j, h_j^{T_P+1}, (l,\mu)) = \begin{cases} \bar{\pi}_i(x,l,G) - \rho T & \text{if } x_j = G \text{ and } x_i = B, \\ \bar{\pi}_i(x,l,G) + \rho T & \text{if } x_j = B \text{ and } x_i = B, \\ \bar{\pi}_i(x,l,G) + \bar{L}\{X_j(l) - (q_2 T + 2\varepsilon T)\} - \rho T & \text{if } x_j = G,\ x_i = G \text{ and } \lambda_j(l) = G, \\ \bar{\pi}_i(x,l,B) - \rho T & \text{if } x_j = G,\ x_i = G \text{ and } \lambda_j(l) = B, \\ \bar{\pi}_i(x,l,G) + \bar{L}\{X_j(l) - (q_2 T + 2\varepsilon T)\} - \rho T & \text{if } x_j = B,\ x_i = G \text{ and } \lambda_j(l) = G, \\ \bar{\pi}_i(x,l,B) + \rho T & \text{if } x_j = B,\ x_i = G \text{ and } \lambda_j(l) = B, \end{cases} \tag{44}$$

Here, $\bar{\pi}_i(x, l, \lambda_j(l)) < 0$ is determined so that player i is almost indifferent between $\lambda_j(l) = G$ and $\lambda_j(l) = B$ and that player i's value at the beginning of the phase is to $\bar{v}_i$ ($\underline{v}_i$, respectively) regardless of x_i if $x_j = G$ (B, respectively) as we have done for (32) and (33). The existence will be explained in Section 6.3.4.

For $m = (l, \lambda_i, 1)$, $(l, \lambda_i, 2)$, $(l, \lambda_j, 1)$ and $(l, \lambda_j, 2)$ (supplemental rounds), player j cancels out the difference of the instantaneous utility:

$$\begin{cases} \sum_{t \in T(m)} \delta^{t-1} \pi_i^-(a_{j,t}, y_{j,t}) & \text{if } x_j = G, \\ \sum_{t \in T(m)} \delta^{t-1} \pi_i^+(a_{j,t}, y_{j,t}) & \text{if } x_j = B. \end{cases} \tag{45}$$

On the top of that, if $m = (l, \lambda_j, 1)$ and $(l, \lambda_j, 2)$ (when player i is the receiver of the message), player j adds

$$\begin{cases} -\sum_{t \in T(m)} \delta^{t-1} (1 - \Psi_{j,t}^{a_i^G, a_{j,t}}) & \text{if } x_j = G, \\ \sum_{t \in T(m)} \delta^{t-1} \Psi_{j,t}^{a_i^G, a_{j,t}} & \text{if } x_j = B \end{cases} \tag{46}$$

to make it strictly optimal to take a_i^G as prescribed in Section 5.2.

6.3.4 Almost Optimality of $\sigma_i(x_i)$

We show $\sigma_i(x_i)$ is optimal by backward induction. The truthtelling incentive in the report round will be verified in 6.4.

For each mth round in the normal blocks, we want to establish the following "almost optimality":

Proposition 1 *$\sigma_i(x_i)$ is almost optimal, that is,*

1. *if $\zeta_j(m) = B$ or $\vartheta_j(m) = B$, then any strategy is exactly optimal.*

2. *if $\zeta_j(m) = \vartheta_j(m) = G$, then*

   (a) *for $m = (l, \lambda_i, 2)$, where player i takes a mixed strategy to send $\lambda_i(l+1)$, then any strategy is indifferent.*

   (b) *for the other rounds, $\sigma_i(x_i)$ is optimal with loss up to $M\exp(-T^{\kappa^*})2\max_{i,a} u_i(a)T_P$.*

1. is obvious since $\zeta_j(m) = B$ or $\vartheta_j(m) = B$ implies $\theta_j(m) = B$. 2.(a). is true since $\zeta_j(m) = \vartheta_j(m) = G$ with $m = (l, \lambda_i, 2)$ imply $\lambda_i(l+1)[2](j) \neq \emptyset$ from (19) and player j does not use the inference in the mth round. Therefore, we show 2.(b) by backward induction. For this purpose, since the distribution of $\{\theta_j(\tilde{m}+1)\}_{\tilde{m}\geq m}$ does not react to player i's strategy in the mth round by more than $M\exp(-T^{\frac{1-2\varepsilon}{2}}) < M\exp(-T^{\kappa^*})$, we can neglect the effect of player i's strategy in the mth round on $\{\theta_j(\tilde{m}+1)\}_{\tilde{m}\geq m}$.

In the Lth main round, consider the case with $x_i = B$ first. If $\theta_j(L, \mu) = G$, then D_i is strictly optimal since it maximizes the summation of the instantaneous utility and (44). If $\theta_j(L, \mu) = B$, then any action is indifferent.

Consider the case with $x_i = G$ next. There are following four cases: (i) if $\hat{\lambda}_j(L) = \lambda_j(L) = G$ and $\theta_j(L, \mu) = G$, then C_i is strictly optimal since the reward (44) is increasing in $X_j(L)$. (ii) If $\hat{\lambda}_j(L) = \lambda_j(L) = B$ and $\theta_j(L, \mu) = G$, then D_i is strictly optimal since the reward (44) is constant. (iii) If $\theta_j(L, \mu) = B$, then any action is optimal. (iv) For the other cases, if player i has not deviated in the supplemental rounds for λ_j, then player i does not have a posterior on the other events more than $M\exp(-T^{\kappa^*})$. Therefore, in total, $\sigma_i(x_i)$ is optimal up to the loss of $M\exp(-T^{\kappa^*})2\max_{i,a} u_i(a)T_P$, where $2\max_{i,a} u_i(a)$ is the maximum of the loss in the per-period utility due to miscoordination. Since player j's continuation play does not depend on the signal observations in the supplemental rounds for λ_j except for θ_j, no matter how we specify the strategy after the deviation in the supplemental rounds for λ_j, the payoff does not increase by more than $M \exp(-T^{\kappa^*}) 2 \max_{i,a} u_i(a) T_P$.

In the supplemental rounds for $\lambda_j(L)$ when player i receives the message, it is optimal to take a_i^G since (46) gives a high reward on a_i^G and $\hat{\lambda}_j(L)(i)$ is almost correct inference. Since player j's continuation strategy does not respond to her history in these rounds except for $\theta_j(m+1)$, no matter how we specify $\sigma_i(x_i)$ after deviation, the deviation in this round will not increase the payoff by more than $M \exp(-T^{\kappa^*}) 2 \max_{i,a} u_i(a) T_P$.

In the supplemental round 2 for $\lambda_i(L)$, as we have verified, $\sigma_i(x_i)$ is exactly optimal.

In the supplemental round 1 for $\lambda_i(L)$, any message is almost indifferent since (45) cancels out the difference in the instantaneous utilities and whenever player i's message has the impact on player j's continuation strategy, player j needs to be ready to listen, which means $\theta_j(m) = B$. Recall that player i's message $\lambda_i(L)$ affects player i's posterior through (43). However, regardless of player i's strategy in this round, the bound (42) is valid. Hence, any message is "almost indifferent," that is, it does not affect the equilibrium payoff more than $M \exp(-T^{\kappa^*}) 2 \max_{i,a} u_i(a) T_P$.

In the $(L-1)$th main round, only the difference from the main round of the Lth block is that, if $\lambda_j(L-1) = \hat{\lambda}_j(L-1) = G$ and $\theta_j(L-1,\mu) = G$, then player i's action in $(L-1,\mu)$ can affect the distribution of $\lambda_j(L)$. As we have explained in the intuitive explanation, we can make sure that player i's continuation payoff is the same between $\lambda_j(L) = G$ and $\lambda_j(L) = B$ if the coordination goes well. (42) implies that regardless of the deviation in this round, coordination goes well with probability no less than $1 - M \exp(-T^{\kappa^*})$ and this effect is negligible. We can proceed toward the first round to establish Proposition 1.

Let us verify the existence of $\bar{\pi}_i(x, L, \lambda_j(l))$ with

$$\bar{\pi}_i(x, l, \lambda_j(l)) \begin{cases} \leq 0 & \text{if } x_j = G, \\ \geq 0 & \text{if } x_j = B \end{cases} \tag{47}$$

for all x, l and $\lambda_j(l)$ by backward induction. Note that this is enough for (6).

For the main round of the Lth block, if $\theta_j((L,\mu)+1) = G$ and $\lambda_j(L) = \hat{\lambda}_j(L) = G$, then the players play $a(x)$ and the ex ante average payoff in the main round of the Lth block for player i is more than $u_i(C_i, C_j) - \rho - 2\varepsilon \bar{L}$ if $x_j = G$ and less than $u_i(D_i, D_j) + \rho + 2\varepsilon \bar{L}$ if $x_j = B$ except for $\bar{\pi}_i$. Note that $\theta_j((L, \mu) + 1) = B$ will not happen with probability more than 2η. Therefore, in total, the ex ante value is more than $u_i(C_i, C_j) - \rho - 2\varepsilon \bar{L} - 2\bar{u}\eta$ if $x_j = G$ and less than $u_i(D_i, D_j) + \rho + 2\varepsilon \bar{L} + 2\bar{u}\eta$ if $x_j = B$. On the other hand, if $\theta_j((L, \mu) + 1) = G$ and $\lambda_j(L) = \hat{\lambda}_j(L) = B$, then the average payoff is $u_i(D_i, C_j) - \rho \geq u_i(C_i, C_j) - \rho$ if $x_j = G$ and $u_i(D_i, D_i) + \rho$ if $x_j = B$ except for $\bar{\pi}_i$.

Therefore, there exists $\bar{\pi}_i(x, L, \lambda_j(L))$ such that (i) (47) is satisfied, that (ii) $\bar{\pi}_i(x, L, \lambda_j(L))$ is uniformly bounded with respect to δ, and that (iii) at the beginning of the main Lth block, conditional on $\theta_j(L, \mu) = G$ and $\lambda_j(L) = \hat{\lambda}_j(L)$, player i is indifferent between $\lambda_j(L) = G$ and $\lambda_j(L) = B$ and the value is $u_i(C_i, C_j) - \rho - 2\varepsilon \bar{L} - 2\bar{u}\eta$ if $x_j = G$ and $u_i(D_i, D_j) + \rho + 2\varepsilon \bar{L} + 2\bar{u}\eta$ if $x_j = B$. Since player i believes that $\theta_j(L, \mu) = G$ and $\lambda_j(L) \neq \hat{\lambda}_j(L)$ cannot happen with probability more than $M \exp(-T^{\kappa^*})$, player i is almost indifferent between $\lambda_j(L) = G$ and $\lambda_j(L) = B$.

Recursively, for $l = 1$, the average ex ante payoff of player i is more than $u_i(C_i, C_j) - \rho - 2\varepsilon \bar{L} - 2M\bar{u}\eta$ ($\geq \bar{v}_i$ from (22)) if $x_j = G$ and less than $u_i(D_i, D_j) + \rho + 2\varepsilon \bar{L} + 2M\bar{u}\eta$ ($\leq \underline{v}_i$ from (22)) if $x_j = B$. Recall that the length of the supplemental rounds is of order $O(T^{1-\varepsilon})$, which is smaller than $O(T)$, the length of the main rounds, and that we can neglect the effect of the supplemental rounds on the average payoff. Therefore, since $\lambda_j(1) = \hat{\lambda}_j(1)$ by definition, there exists $\bar{\pi}_i(x, 1, \lambda_j(1))$ such that $\sigma_i(x_i)$ gives $\bar{v}_i$ ($\underline{v}_i$, respectively) to player i regardless of x_i and $\lambda_j(1)$ if $x_j = G$ (B, respectively).

6.4 Exact Optimality

We are left to show the truthtelling incentive for $h_i^{T_P+1}$ and to establish the exact optimality of $\sigma_i(x_i)$. To verify the truthtelling incentive, we distinguish the true history $h_i^{T_P+1}$ and the message $\tilde{h}_i^{T_P+1}$. In general, when we write a variable that player i owns with "tilde", it means player j's inference of player i's message about that variable.

Let $\mathcal{A}_j(m)$ be the set of information up to and including the mth round consisting of

- what state x_j player j is in,

- what action $a_j(l, \mu)$ player j took in the main rounds with $(l, \mu) \leq m$,

- what message $\lambda_j(l)$ player j had with $(l, \lambda_j, 1) \leq m$, and

- what state $\vartheta_j(m)$ and $\zeta_j(m)$ player j had.

We want to show that $\sigma_i(x_i)$ is exactly optimal in the mth round conditional on $\mathcal{A}_j(m)$. Note that $\mathcal{A}_j(m)$ contains x_j and so the equilibrium is belief-free at the beginning of the phase. Note also that $\mathcal{A}_j(m)$ does not include player j's messages in the supplemental round 2 for $\lambda_j(l)$ since, for these messages, player i needs to infer by the likelihood ratio and we cannot condition them.

We introduce the following variables. Let $R_i(m)$ be the set of rounds $\tilde{m} \leq m$ that is not a supplemental round 2 for $\lambda_i(l+1)$, t_m be the initial period of the mth round, $\#_i^m$ be the summary of the history in the mth round, where $\#_i^m$ is a $|A_i| |Y_i| \times 1$ vector whose element corresponding to (a_i, y_i) represents how many times player i observed (a_i, y_i) in the mth round, and $\mathfrak{h}_i^m \equiv \{\#_i^{\tilde{m}}\}_{\tilde{m} \in R_i(m)}$ be the summary of player i's history at the beginning of the $(m+1)$th round.

We construct the message protocol for $\hat{h}_i^{T_P+1}$ as follows.

- by public randomization device, the players coordinate on who will report $h_i^{T_P+1}$. Only one player reports the history. Player 1 reports $h_1^{T_P}$ with probability $\frac{1}{2}$ and player 2 reports $h_2^{T_P}$ with probability $\frac{1}{2}$.

- for each $m \in R_i(M)$, player i who is supposed to send $h_i^{T_P+1}$ according to the public randomization device sends the history in the mth round $\{a_{i,t}, y_{i,t}\}_{t \in T(m)}$.

- based on $\{\tilde{a}_{i,t}, \tilde{y}_{i,t}\}_{t \in T(m)}$, player j calculates $\tilde{\#}_i^m$.

With abuse of notation, we assume the players can send cheap talk messages several times and the public randomization is available every time when the players send the messages.

How to dispense the public randomization device without cheap talk is explained in the Online Appendix.

Player j gives a reward on player i as follows. Here, we do not consider the feasibility constraint (6). As we will see, the total reward in the report block is bounded by T^{-1} and we can satisfy (6) by adding or subtracting a small constant without affecting the incentive if necessary.

As a preparation, we state the following lemma, whose proof is offered in the Appendix:

Lemma 7 *Let h_j^m be player j's history in the mth round. There generically exist $\bar{\varepsilon} > 0$ and $g_i(h_j^m, a_i, y_i)$ such that, for sufficiently large T, for any mth round and period t in the mth round, conditional on $\zeta_j(m) = \zeta_j(m+1) = G$, player i's history at the end of the normal block h_i, it is better for player i to report $a_{i,t}, y_{i,t}$ truthfully: for any h_i and t,*

$$\begin{aligned} & \mathbb{E}\left[g_i(h_j^m, \tilde{a}_{i,t}, \tilde{y}_{i,t}) \mid \zeta_j(m) = \zeta_j(m+1) = G, h_i, (\hat{a}_{i,t}, \hat{y}_{i,t}) = (a_{i,t}, y_{i,t})\right] \\ > \; & \mathbb{E}\left[g_i(h_j^m, \tilde{a}_{i,t}, \tilde{y}_{i,t}) \mid \zeta_j(m) = \zeta_j(m+1) = G, h_i, (\hat{a}_{i,t}, \hat{y}_{i,t}) \neq (a_{j,t}, y_{j,t})\right] + \bar{\varepsilon} T^{-1}, \end{aligned}$$

where $(\hat{a}_{i,t}, \hat{y}_{i,t})$ is player i's message.

Intuitively, $\|\mathbf{y}_{j,t} - E[\mathbf{y}_{j,t} \mid \tilde{a}_{i,t}, \tilde{y}_{i,t}, a_{j,t}]\|^2$ works for $g_i(h_j^m, \tilde{a}_{i,t}, \tilde{y}_{i,t})$, where $\mathbf{y}_{j,t}$ is a $|Y_j| \times 1$ vector whose element corresponding to $y_{j,t}$ is equal to 1 and that not corresponding to $y_{j,t}$ is equal to 0.[11,12] However, player i can learn about $y_{j,t}$ by player i's signals in the rounds $\tilde{m} \geq m$, which contains the information about player j's strategy in the rounds $\tilde{m} \geq m$, which depends on $\{y_{j,t}\}$ in the mth round. See the Appendix for the formal adjustment.

By backward induction, for each m, we will construct the following rewards based on player i's messages:

- based on the past messages $\tilde{\mathfrak{h}}_i^{m-1} = \{\tilde{\#}_i^{\tilde{m}}\}_{\tilde{m} \leq R_i(m-1)}$, player j calculates $A_i(\tilde{\mathfrak{h}}_i^{m-1})$, the set of player i's action that is taken with positive probability in the mth round after history $\tilde{\mathfrak{h}}_i^{m-1}$.

[11] Here, we use Euclidean norm.

[12] Kandori and Matsushima (1994) uses the similar reward to give a player the incentive to tell the truth abouth the history.

- if the mth round is a supplemental round 2 for $\lambda_i(l+1)$, player i does not report the history. Hence, the reward below for the mth round is 0.

- otherwise, player j gives the following reward: first, based on $\tilde{\mathfrak{h}}_i^{m-1}$, player j gives

$$\sum_{a_i} f(a_i \mid \tilde{\mathfrak{h}}_i^{m-1}, \mathcal{A}_j(m)) \Psi_{j,t_m}^{a_i, a_{j,t_m}} \tag{48}$$

  with

$$f(a_i \mid \tilde{\mathfrak{h}}_i^{m-1}, \mathcal{A}_1(m)) \in [-T^{-10m+6}, T^{-10m+6}] \text{ for all } a_i \in A_i$$

  such that, after $\mathfrak{h}_i^{m-1}$, it is optimal to take $a_i \in A_i(\mathfrak{h}_i^{m-1})$. The existence will be verified below.

  Second, player j punishes player i based on

$$T^{-10m+5} g_i(h_j^m, \tilde{a}_{i,t_m}, \tilde{y}_{i,t_m}). \tag{49}$$

  Further, player j makes it optimal to constantly take $a_i(m)$ within the mth round by adding

$$T^{-10m+3} \sum_{t \in T(m)} \Psi_{j,t}^{\tilde{a}_{i,t_m}, a_{j,t}} \tag{50}$$

  for t included in the mth round.

  Finally, player j punishes player i by

$$\sum_{t \in T(m)} T^{-10m} T g_i(h_j^m, \tilde{a}_{i,t}, \tilde{y}_{i,t}). \tag{51}$$

In addition, if we come to the mth round with $\zeta_j(m) = B$ or $\vartheta_j(m) = B$, then we cancel out all the rewards about the following rounds $\tilde{m} \geq m$. Then, since we have established the exact optimality of $\sigma_i(x_i)$ without adjustment after $\zeta_j(m) = B$ or $\vartheta_j(m) = B$, any action profile sequence is exactly optimal after (and including) the mth round.

Further, if we come to the $(m-1)$th round with $\zeta_j(m) = B$, then we cancel out all the rewards about the following rounds $\tilde{m} \geq m-1$. We will verity this does not affect player i's incentive below.

We show the truthtelling incentive about $\{a_{i,t}, y_{i,t}\}_{t \in T(m)}$ by backward induction. We start from M, the last round. Regardless of the specification of f, (51) incentivizes player i to tell the truth about $a_{i,t}, y_{i,t}$ for $t \neq t_m$. Since (49) dominates (50), it is optimal to tell the truth about a_{i,t_m}, y_{i,t_m}. Since Lemma 7 conditions that $\zeta_j(M+1) = G$, all the discussion above holds after conditioning on $\zeta_j(M+1) = G$. Hence, conditioning on $\zeta_j(M+1) = G$ does not affect player i's incentive.

For the $(M-1)$th round, (50) and (51) for $m = M$ are independent of the messages about the $(M-1)$th round. (48) with $m = M$ is dominated by the smallest loss in (49) and (51) for $m = M-1$. Therefore, the same argument for the Mth round works. We can proceed until the first round.

Recursively, therefore, regardless of the specification of f, we have established the optimality of the truthtelling incentive in the report block. Now, we construct $f(a_i \mid \tilde{\mathfrak{h}}_i^{m-1}, \mathcal{A}_j(m))$ by backward induction.

At the beginning of the Mth round, player i's value of taking a constant action $a_i \in A_i$ conditional on $\mathcal{A}_j(M)$ only depends on $\mathfrak{h}_i^{M-1}$. This is true even after player i's deviation since this is enough to infer player j's strategy that is i.i.d. within all the rounds. Note that (49), (50) and (51) with $m \leq M-1$ are already sank given the truthtelling strategy in the report block. Hence, we can write the value as $v_i(a_i \mid \mathfrak{h}_i^{M-1}, \mathcal{A}_1(M))$. Let $f(a_i \mid \tilde{\mathfrak{h}}_i^{M-1}, \mathcal{A}_1(M))$ be such that

$$\Pr(\text{player } i \text{ reports the history}) \sum_{\tilde{a}_i} f\left(\tilde{a}_i \mid \tilde{\mathfrak{h}}_i^{M-1}, \mathcal{A}_j(M)\right) \Pr\left(\left\{\Psi_{j,t}^{\tilde{a}_i, a_{j,t}} = 1\right\} \mid a_{i,t} = a_i\right)$$
$$= \begin{cases} \max_{\tilde{a}_i \in A_i} v_i(\tilde{a}_i \mid \mathfrak{h}_i^{M-1}, \mathcal{A}_j(M)) - v_i(a_i \mid \mathfrak{h}_i^{M-1}, \mathcal{A}_j(M)) \text{ if } a_i \in A_i(\mathfrak{h}_i^{M-1}) \\ 0 \text{ otherwise} \end{cases}$$

for all $\mathfrak{h}_i^{M-1}$. Since $\Pr(\zeta_j(M+1) = G \mid \zeta_j(M) = G, \#_j^M) \geq 1 - \exp(-T^{\frac{1-2\varepsilon}{2}})$ for all $\#_j^M$, all the messages about $\mathfrak{h}_i^{M-1}$ transmit correctly, $\sigma_i(x_i)$ is almost optimal except for the adjustment of the reward in the report block, and the variance of the reward in the report block based on the histories in the Mth round is bounded by T^{-10M+5}, we can make sure that

$$f(a_i \mid \tilde{\mathfrak{h}}_i^{M-1}, \mathcal{A}_j(M)) \in [-T^{-10M+6}, T^{-10M+6}].$$

This makes it strictly optimal to take $a_i \in A_i(\mathfrak{h}_i^{M-1})$. After that, (50) and truthtelling incentive imply that it is optimal to constantly take $a_i^{a_i(M)}$ since $\Pr(\zeta_j(M+1) = G \mid \zeta_j(M) = G, \#_j^M) \geq 1 - \exp(-T^{\frac{1-2\varepsilon}{2}})$ for all $\#_j^M$ and the the rewards that the following messages about the history in the Mth round affect is dominated by (50).

We can proceed until the first round and show the optimality of $\sigma_i(x_i)$. Note that if the mth round is a supplemental round 2 for $\lambda_i(l+1)$ and this round does not have an impact ($\zeta_j(m) = \vartheta_j(m) = G$ guarantees this), then the expected rewards in the report block are not affected by the strategy in the mth round.

7 Equilibrium Construction without Cheap Talk

Note that we use the cheap talk messages for x_i and $h_i^{T_P+1}$. The basic idea to send x_i is the same as in the supplemental rounds for $\lambda_i(l+1)$. As Lemma 4 shows, *conditional on the opponent message*, the receiver can believe that her inference is correct or any action profile sequence is indifferent to her. Therefore, even after observing the continuation strategy that is not corresponding to her inference, the receiver can believe that any action profile sequence is indifferent with high probability and that following the equilibrium strategy is almost optimal. As we have seen in 6.4, we can attain the exact optimality afterwards.

For x_i, however, there is one problem: the receiver j constructs the reward function on the sender i based on x_i (remember that for $\lambda_i(l+1)$, whenever player j changes her strategy based on the inference of $\lambda_i(l+1)$, player j makes any action profile sequence indifferent to player i and the sender's inference of the receiver's inference is irrelevant). Therefore, the sender need to have the inference of the receiver's inference. Further, even after observing the receiver's continuation play, the sender needs to think it is almost optimal to follow the equilibrium strategy. This complicates the problem.

As for $h_i^{T_P+1}$ in the report block, there are three problems. First, the cardinality of the messages for our equilibrium construction is of order $(|A_i|\,|Y_i|)^T$. If we replace cheap talk with the message exchange via taking actions straightforwardly, it takes too long to send all the messages and affects the equilibrium payoff. The second problem is that we need to dispense the public randomization device. It is important to have only one player who reports $h_i^{T_P+1}$ since otherwise, while the opponent is reporting $h_j^{T_P}$, player i can update the belief over player j's signal observations and the incentive to tell the truth will be destroyed. On the other hand, to have the exact optimality, there needs to be a positive probability during the main blocks for each player to report $h_i^{T_P+1}$ and the reward function on her will be adjusted. Third, there exists a positive probability that the messages do not transmit correctly. While player i sends the message about $h_i^{T_P+1}$, if player i's signal observations contain the information about the receiver's inference of the message, player i may start to think that the probability that her message transmits correctly will be small if player i's signal observation suggests that player j's signal observation is "erroneous." To deal with this problem, we need to ensure that conditional on player i's action, player i's signal observations and player j's inference is independent. That is, we need to attain the "conditionally independent inference" without further assumptions on the monitoring. See the Online Appendix for the solutions to these problems.

8 General Game

8.1 General Two-Player Game

The prisoners' dilemma is special in that, when player j' state is B, the equilibrium action D_j can attain the targeted payoff $\bar{v}_j$ or $\underline{v}_j$ depending on player i's state at the same time minmaxing player i. The second point is important. Player j with $x_j = B$ and $\lambda_j(l) = B$ needs to give a *positive constant* reward to sustain (6). On the other hand, player j needs to ensure that player i's payoff is below $\underline{v}_i$ regardless of player i's strategy. Therefore, player j

with $x_j = B$ and $\lambda_j(l) = B$ needs to minmax player i if player i does not take a prescribed action.

For this purpose, we let player i construct the statistics depending on whose realization player i *allows player j to minmax player i*. Player i sends the message about whether she allows player j to minmax. Player j, on the other hand, calculates the conditional expectation of that statistics which *decreases if player i deviates*. Modifying the coordination protocol about $\lambda_i(l)$, we can make the protocol such that the players can coordinate on the punishment properly and that player j punishes player j with high probability if the realization of the conditional expectation is sufficiently low regardless of the message to keep player i's payoff below $\underline{v}_i$ regardless of player i's strategy. See Sugaya (2010c) for the formal description.

8.2 General More-Than-Two-Player Game

If there are more than two players, as Hörner and Olszewski (2006), we let player $(i-1)$'s state determine whether player i's value is $\bar{v}_i$ or $\underline{v}_i$ independently of players $-(i-1)$'s state. With cheap talk, there is no conceptual difficulty to extend our result. Without cheap talk, however, there is a new problem to coordinate on x_i. Since each player takes an action based on the her own inference about x_i, if there exists a player who induces the different inferences for the different players for the same x_i, then it may be of that player's interest to induce the miscoordination. See Sugaya (2010c) for how to deal with this problem.

9 Appendix

9.1 Proof of Lemma 4

As a preparation, we prove the following three lemmas. In this subsection, let $S \subset \mathbb{N}$ be the set of periods when the players are supposed to constantly take $(a_i^{m_i}, a_j^G)$.

The first lemma is about the relationship between the history of the message sender i and $\Omega_i^H(S)$:

Lemma 8 *1. for any t and $\{a_{j,\tau}, y_{j,\tau}\}_{\tau\in S,\tau\neq t}$, the distribution of $\Omega_i^H(S)$ is independent of $a_{j,t}$.*

2. for any $\{a_{i,t}, y_{i,t}\}_{t\in S}$, the probability that

$$\left\| \Omega_i^H(S) - \frac{1}{|S|}\sum_{t\in S} H_i(m_i)\mathbf{1}_{y_{i,t}} \right\| \leq \frac{\varepsilon}{8}$$

with probability no less than $1 - \exp(|S|^{1-\frac{\varepsilon}{2}})$.

3. conditional on

$$\left\| \Omega_i^H(S) - \frac{1}{|S|}\sum_{t\in S} H_i(m_i)\mathbf{1}_{y_{i,t}} \right\| \leq \frac{\varepsilon}{8},$$

if the frequency of player i's signal observations in S, $\frac{1}{|S|}\sum_{t\in S}\mathbf{1}_{y_{i,t}}$, is not included in $H_i\left[\frac{\varepsilon}{2}\right](m_i)$, then

$$\left\| \Omega_i^H(S) - \mathbf{p}_i(m_i) \right\| > \frac{1}{4}\varepsilon.$$

Proof.

1. Follows from the definition.

2. Follows from Hoeffding's inequality.

3. Under the condition

$$\left\| \Omega_i^H(S) - \frac{1}{|S|}\sum_{t\in S} H_i(m_i)\mathbf{1}_{y_{i,t}} \right\| \leq \frac{\varepsilon}{8},$$

$$\left\| \Omega_i^H(S) - \mathbf{p}_i(m_i) \right\| \leq \frac{1}{4}\varepsilon$$

implies, by the triangle inequality,

$$\left\| \frac{1}{|S|}\sum_{t\in S} H_i(m_i)\mathbf{1}_{y_{i,t}} - \mathbf{p}_i(m_i) \right\| \leq \frac{1}{2}\varepsilon$$

and so $\frac{1}{|S|}\sum_{t\in S}\mathbf{1}_{y_{i,t}} \in \mathrm{H}_i\left[\frac{\varepsilon}{2}\right](m_i)$, which is the contradiction.

■

The second lemma is, on the other hand, about the relationship between the history of the message receiver j and $\Omega_j^G(S)$, $M_{i,j}^G(S)$ and $M_{i,j}^B(S)$:

Lemma 9 *1. for any t and $\{a_{i,\tau}, y_{i,\tau}\}_{\tau\in S,\tau\neq t}$, the distribution of $\Omega_j^G(S)$ is independent of $a_{i,t}$.*

2. if player j takes a_j, then

$$\left\|\Omega_j^G(S) - \mathbf{p}_j(G)\right\| \leq \frac{\varepsilon}{2}$$

with probability no less than $1 - \frac{1}{2}\exp(|S|^{1-\frac{\varepsilon}{2}})$.

Further, for any $\{a_{j,t}, y_{j,t}\}_{t\in S}$,

$$\left\|\Omega_j^G(S) - \frac{1}{|S|}\sum_{t\in S} H_j(G)\mathbf{1}_{y_{j,t}}\right\| \leq \frac{\varepsilon}{8},$$

$$\left\|M_{i,j}^G(S) - H_i(G)M_{i,j}(a_i^G, a_j^G)\frac{1}{|S|}\sum_{t\in S}\mathbf{1}_{y_{j,t}}\right\| \leq \frac{\varepsilon}{8}$$

and

$$\left\|M_{i,j}^B(S) - H_i(B)M_{i,j}(a_i^B, a_j^G)\frac{1}{|S|}\sum_{t\in S}\mathbf{1}_{y_{j,t}}\right\| \leq \frac{\varepsilon}{8}$$

with probability no less than $1 - \exp(|S|^{1-\frac{\varepsilon}{2}})$.

3. if, for some $\hat{m}_i \in \{G, B\}$,

$$\left\|\Omega_j^G(S) - \mathbf{p}_j(G)\right\| \leq \frac{\varepsilon}{2},$$

$$\left\|\Omega_j^G(S) - \frac{1}{|S|}\sum_{t\in S} H_j(G)\mathbf{1}_{y_{j,t}}\right\| \leq \frac{\varepsilon}{8},$$

$$\left\|M_{i,j}^G(S) - H_i(G)M_{i,j}(a_i^G, a_j^G)\frac{1}{|S|}\sum_{t\in S}\mathbf{1}_{y_{j,t}}\right\| \leq \frac{\varepsilon}{8},$$

$$\left\|M_{i,j}^B(S) - H_i(B)M_{i,j}(a_i^B, a_j^G)\frac{1}{|S|}\sum_{t\in S}\mathbf{1}_{y_{j,t}}\right\| \leq \frac{\varepsilon}{8},$$

and

$$\left\|M_{i,j}^{\hat{m}_i}(S) - \mathbf{p}_i(\hat{m}_i)\right\| \leq \bar{K}\varepsilon$$

then, for $\tilde{m}_i \in \{G, B\} \setminus \{\hat{m}_i\}$*, we have*

$$\left\|M_{i,j}^{\tilde{m}_i}(S) - \mathbf{p}_i(\tilde{m}_i)\right\| > \bar{K}\varepsilon.$$

4. *for any* $m_i \in \{G, B\}$*, if*

$$\left\|M_{i,j}^{m_i}(S) - H_i(m_i)M_{i,j}(a_i^{m_i}, a_j^G)\frac{1}{|S|}\sum_{t\in S}\mathbf{1}_{y_{j,t}}\right\| \leq \frac{\varepsilon}{8} \text{ and } \left\|M_{i,j}^{m_i}(S) - \mathbf{p}_i(m_i)\right\| > \bar{K}\varepsilon,$$

then there do not exist $\mathbf{x} \in \mathbb{R}_+^{|Y_i|}$ *and* $\boldsymbol{\varepsilon} \in \mathbb{R}_+^{|Y_i|}$ *with* $\|\boldsymbol{\varepsilon}\| \leq \varepsilon$ *such that*

$$\begin{aligned}\mathbf{x} &= M_{i,j}(a_i^{m_i}, a_j^G)\frac{1}{|S|}\sum_{t\in S}\mathbf{1}_{y_{j,t}} + \boldsymbol{\varepsilon} \\ \mathbf{x} &\in \mathcal{H}_i\left[\frac{1}{2}\varepsilon\right](m_i).\end{aligned}$$

5. *conditional on*

$$\left\|\Omega_j^G(S) - \frac{1}{|S|}\sum_{t\in S}H_j(G)\mathbf{1}_{y_{j,t}}\right\| \leq \frac{\varepsilon}{8},$$

$$\left\|M_{i,j}^G(S) - H_i(G)M_{i,j}(a_i^G, a_j^G)\frac{1}{|S|}\sum_{t\in S}\mathbf{1}_{y_{j,t}}\right\| \leq \frac{\varepsilon}{8},$$

$$\left\|M_{i,j}^B(S) - H_i(B)M_{i,j}(a_i^B, a_j^G)\frac{1}{|S|}\sum_{t\in S}\mathbf{1}_{y_{j,t}}\right\| \leq \frac{\varepsilon}{8},$$

if the frequency of player i's signal observations in S, $\frac{1}{|S|}\sum_{t\in S}\mathbf{1}_{y_{i,t}}$*, is included in* $H_j\left[\frac{\varepsilon}{4}\right](G) \cap H_{i,j}\left[(\bar{K}-1)\varepsilon\right](\hat{m}_i)$*, then*

$$\begin{aligned}\left\|\Omega_j^G(S) - \mathbf{p}_j(G)\right\| &\leq \frac{\varepsilon}{2} \\ \left\|M_{i,j}^{m_i}(S) - \mathbf{p}_i(m_i)\right\| &\leq \bar{K}\varepsilon.\end{aligned}$$

Proof.

1. Follows from Hoeffding's inequality.

2. Follows from Hoeffding's inequality.

3. From the triangle inequality, the first two conditions imply

$$\frac{1}{|S|}\sum_{t\in S}\mathbf{1}_{y_{j,t}} \in \mathcal{H}_j[\varepsilon](G).$$

   On the other hand, the last three conditions imply

$$\frac{1}{|S|}\sum_{t\in S}\mathbf{1}_{y_{j,t}} \in \mathcal{H}_{i,j}[(\bar{K}+1)\varepsilon](\hat{m}_i).$$

   If $\left\|M_{i,j}^{\tilde{m}_i}(S) - \mathbf{p}_i(\tilde{m}_i)\right\| \leq \bar{K}\varepsilon$ is the case, then we also have

$$\frac{1}{|S|}\sum_{t\in S}\mathbf{1}_{y_{j,t}} \in \mathcal{H}_{i,j}[(\bar{K}+1)\varepsilon](\tilde{m}_i),$$

   which contradicts to Lemma 3.

4. Suppose there exists such $\mathbf{x} \in \mathbb{R}_+^{|Y_i|}$ and $\boldsymbol{\varepsilon} \in \mathbb{R}_+^{|Y_i|}$. Then, for $\mathbf{y} = \left(\sum_{t\in S}\mathbf{1}_{y_{j,t}}\right)_{y_j}$,

$$\begin{aligned} H_i(m_i)\mathbf{x} &= H_i(m_i)\left(M_{i,j}(a_i^{m_i}, a_j^G)\mathbf{y} + \boldsymbol{\varepsilon}\right) \\ &= H_i(m_i)M_{i,j}(a_i^{m_i}, a_j^G)\mathbf{y} + H_i(m_i)\boldsymbol{\varepsilon} \\ &= H_i(m_i)M_{i,j}(a_i^{m_i}, a_j^G)\mathbf{y} + \tilde{\boldsymbol{\varepsilon}} \end{aligned}$$

   with $\|\tilde{\boldsymbol{\varepsilon}}\| = \|H(m_i)\boldsymbol{\varepsilon}\| \leq (\bar{K}-2)\boldsymbol{\varepsilon}$. Since $\mathbf{x} \in \mathcal{H}_i\left[\frac{1}{2}\varepsilon\right](m_i)$, there exists $\boldsymbol{\varepsilon}_i$ with $\|\boldsymbol{\varepsilon}_i\| \leq \frac{\varepsilon}{2}$ with

$$H_i(m_i)\mathbf{x} - \boldsymbol{\varepsilon}_i = \mathbf{p}_i(m_i),$$

   which implies

$$H_i(m_i)M_{i,j}(a_i^{m_i}, a_j^G)\mathbf{y} + \tilde{\boldsymbol{\varepsilon}} - \boldsymbol{\varepsilon}_i = \mathbf{p}_i(m_i)$$

and

$$\|\tilde{\boldsymbol{\varepsilon}} - \boldsymbol{\varepsilon}_i\| \le (\bar{K} - 1)\varepsilon,$$

which implies

$$\left\| H_i(m_i) M_{i,j}(a_i^{m_i}, a_j^G) \frac{1}{|S|} \sum_{t \in S} \mathbf{1}_{y_{j,t}} - \mathbf{p}_i(m_i) \right\| \le (\bar{K} - 1)\varepsilon.$$

By the triangle inequality,

$$\left\| M_{i,j}^{m_i}(S) - \mathbf{p}_i(m_i) \right\| \le \bar{K}\varepsilon,$$

which is the contradiction.

5. Follows from the triangle inequality.

■

The third one is about the inference $z_i(m_i)[2](j)$:

Lemma 10 *For generic q, there exists $\kappa > 0$ such that for any j and $f_j \in \Delta\left(\{0,1\}^{|Y_j|}\right)$, one of the following is true:*

1. *the likelihood ratio of $z_i(m_i) = G$ compared to $z_i(m_i) = B$ is no less than $\exp(T^\kappa)$ for any m_i,*

2. *the likelihood ratio of $z_i(m_i) = B$ compared to $z_i(m_i) = G$ is no less than $\exp(T^\kappa)$ for any m_i, or*

3. *the likelihood ratio of $z_i(m_i) = M$ compared to $z_i(m_i) = G, B$ is no less than $\exp(T^\kappa)$ for any m_i.*

Proof. It suffices to show that there exists $\tilde{\kappa}$ such that one of the following is true:

1. $z_i(m_i) = G$ is more likely than $z_i(m_i) = B$: $\mathcal{L}(f_j, G, B) \ge \tilde{\kappa}$,

2. $z_i(m_i) = B$ is more likely than $z_i(m_i) = G$: $\mathcal{L}(f_j, B, G) \geq \tilde{\kappa}$,

3. $z_i(m_i) = M$ is more likely than $z_i(m_i) = G, B$: $\mathcal{L}(f_j, M, G) \geq \tilde{\kappa}$ and $\mathcal{L}(f_j, M, B) \geq \tilde{\kappa}$.

Generically, we can assume that for each $k \in \{1, \ldots, |Y_j|\}$,

$$q(y_{j,k}|a_j^G, \alpha_i^G) \neq q(y_{j,k}|a_j^G, \alpha_i^B). \tag{52}$$

Let $\alpha_i^\lambda = \lambda a_i^G + (1-\lambda)\, a_i^B$ for $\lambda \in [0,1]$ and consider

$$g(f_j, \lambda) = f_{j,1} \log q(y_{j,1}|a_j^G, \alpha_i^\lambda) + \cdots + f_{j,|Y_j|} \log q(y_{j,|Y_j|}|a_j^G, \alpha_i^\lambda).$$

Then,

$$\frac{d^2 g(f_j, \lambda)}{d\lambda^2} = -\sum_{k=1}^{|Y_j|} f_{j,k} \left\{ \frac{q(y_{j,k}|a_j^G, \alpha_i^G) - q(y_{j,k}|a_j^G, \alpha_i^B)}{q(y_{j,k}|a_j^G, \alpha_i^\lambda)} \right\}^2 < 0$$

for any $f_{j,k}$ because of (52). Hence, $g(f_j, \lambda)$ is strictly concave. Therefore, since $\mathcal{L}(f_{j,k}, z_i, \tilde{z}_i)$ is the difference in $g(f_j, \lambda)$, one of the following is true:

1. $z_i(m_i) = G$ is more likely than $z_i(m_i) = B$: $\mathcal{L}(f_j, G, B) > 0$,

2. $z_i(m_i) = B$ is more likely than $z_i(m_i) = G$: $\mathcal{L}(f_j, B, G) > 0$,

3. $z_i(m_i) = M$ is more likely than $z_i(m_i) = G, B$: $\mathcal{L}(f_j, M, G) > 0$ and $\mathcal{L}(f_j, M, B) > 0$.

Hence,

$$\max\left\{\mathcal{L}(f_j, G, B), \mathcal{L}(f_j, B, G), \min\left\{\mathcal{L}(f_j, M, G), \mathcal{L}(f_j, M, B)\right\}\right\} > 0.$$

Since LHS is continuous in f_j and $\Delta\left(\{0,1\}^{|Y_j|}\right)$ is compact, there exists $\tilde{\kappa} > 0$ such that

$$\max\left\{\mathcal{L}(f_j, G, B), \mathcal{L}(f_j, B, G), \min\left\{\mathcal{L}(f_j, M, G), \mathcal{L}(f_j, M, B)\right\}\right\} > \tilde{\kappa}$$

for all $f_j \in \Delta\left(\{0,1\}^{|Y_j|}\right)$ as desired. ■

Given the above three lemmas, we can prove Lemma 4:

1. Follows from Condition 3 of Lemma 9.

2. If $m_i[2](j) = \emptyset$, then it follows from Lemma 10. The conditioning on $\tilde{\zeta}_i(S)$ for S being not the periods when player i sends $z_i(m_i)$ is irrelevant.

   Let us concentrate on $m_i[2](j) = \{G, B\} \setminus \{m_i\}$. Player i has $m_i[1](i) = m_i$ only if $\mathbf{x} \in \mathcal{H}_i \left[\frac{\varepsilon}{2}\right](m_i)$ from Condition 3 of Lemma 8.

   Forget about the condition on $\tilde{\zeta}_i(S) = G$ for a while. If $m_i[2](j) = \{G, B\} \setminus \{m_i\}$, then from Condition 4 of Lemma 9, player j's conditional expectation $M_{i,j}(a_i^{m_i}, a_j^G)\mathbf{y}$ is away from $\mathcal{H}_i \left[\frac{\varepsilon}{2}\right](m_i)$ by at least ε, by Hoeffding's inequality, player i puts probability no more than $\exp(-T^{\frac{1-2\varepsilon}{2}})$ on $\mathbf{x} \in \mathcal{H}_i \left[\frac{\varepsilon}{2}\right](m_i)$.

   Since

$$
\begin{aligned}
\Pr(\{y_{i,t}\}_{t\in S} \quad &| \quad \tilde{\zeta}_i(S) = G, \{y_{j,t}\}_{t\in S}) \\
&= \frac{\Pr(\tilde{\zeta}_i(S) = G \mid \{y_{i,t}\}_{t\in S}, \{y_{j,t}\}_{t\in S}) \Pr(\{y_{i,t}\}_{t\in S} \mid \{y_{j,t}\}_{t\in S})}{\Pr(\tilde{\zeta}_i(S) = G \mid \{y_{j,t}\}_{t\in S})} \\
&= \frac{\Pr(\tilde{\zeta}_i(S) = G \mid \{y_{j,t}\}_{t\in S}) \Pr(\{y_{i,t}\}_{t\in S} \mid \{y_{j,t}\}_{t\in S})}{\sum_{\{y_{i,t}\}_{t\in S}} \Pr(\tilde{\zeta}_i(S) = G \mid \{y_{i,t}\}_{t\in S}) \Pr(\{y_{i,t}\}_{t\in S} \mid \{y_{j,t}\}_{t\in S})} \\
&\in \left[\begin{array}{c} (1 - 2\exp(-T^{\frac{1-2\varepsilon}{2}})) \Pr(\{y_{i,t}\}_{t\in S} \mid \{y_{j,t}\}_{t\in S}), \\ (1 + 2\exp(-T^{\frac{1-2\varepsilon}{2}})) \Pr(\{y_{i,t}\}_{t\in S} \mid \{y_{j,t}\}_{t\in S}) \end{array} \right],
\end{aligned}
$$

   after conditioning $\tilde{\zeta}_i(S) = G$, we can say player j puts probability no more than $4\exp(-T^{\frac{1-2\varepsilon}{2}})$ on $\mathbf{x} \in \mathcal{H}_i \left[\frac{\varepsilon}{2}\right](m_i)$.

3. For a while, suppose $\tilde{\zeta}_j(S) = B$ never happens. Conditional on any m_i and $\mathbf{x}$, any $\mathbf{y}$ can occur with probability at least

$$
\left\{ \min_{y_i, y_j, a} q(y_j \mid a, y_i) \right\}^{T^{\frac{1-\varepsilon}{2}}}.
$$

   Hence, for

$$
e < \frac{1}{-\log \left\{ \min_{y_i, y_j, a} q(y_j \mid a, y_i) \right\}},
$$

conditional on m_i and $\mathbf{x}$, for any $\hat{m}_i \in \{G, B\}$,

$$\mathbf{y} \in \mathcal{H}_{i,j}\left[(\bar{K}-1)\varepsilon\right](\hat{m}_i) \cap \mathcal{H}_j\left[\frac{\varepsilon}{4}\right](G)$$

can happen with probability no less than $\exp(-\frac{1}{e}T^{\frac{1-\varepsilon}{2}})$. If $\mathbf{y} \in \mathcal{H}_{i,j}\left[(\bar{K}-1)\varepsilon\right](\hat{m}_i) \cap \mathcal{H}_j\left[\frac{\varepsilon}{4}\right](G)$, then player j has $m_i[2](j) = \hat{m}_i$ from Condition 5. of Lemma 9. Hence, player j believes any $\hat{m}_i$ is possible with probability no less than $\exp(-\frac{1}{e}T^{\frac{1-\varepsilon}{2}})$. Conditioning on $\tilde{\vartheta}_j(S) = G$ only increases the probability.

Since

$$\begin{aligned}
\Pr(\{y_{j,t}\}_{t\in S} \quad | \quad & \tilde{\zeta}_j(S) = G, \{y_{i,t}\}_{t\in S}) \\
= & \frac{\Pr(\tilde{\zeta}_j(S) = G \mid \{y_{j,t}\}_{t\in S}, \{y_{i,t}\}_{t\in S}) \Pr(\{y_{j,t}\}_{t\in S} \mid \{y_{i,t}\}_{t\in S})}{\Pr(\tilde{\zeta}_j(S) = G \mid \{y_{i,t}\}_{t\in S})} \\
= & \frac{\Pr(\tilde{\zeta}_j(S) = G \mid \{y_{i,t}\}_{t\in S}) \Pr(\{y_{j,t}\}_{t\in S} \mid \{y_{i,t}\}_{t\in S})}{\sum_{\{y_{i,t}\}_{t\in S}} \Pr(\tilde{\zeta}_j(S) = G \mid \{y_{i,t}\}_{t\in S}) \Pr(\{y_{j,t}\}_{t\in S} \mid \{y_{i,t}\}_{t\in S})} \\
\in & \left[\begin{array}{c} (1 - 2\exp(-T^{\frac{1-2\varepsilon}{2}})) \Pr(\{y_{j,t}\}_{t\in S} \mid \{y_{i,t}\}_{t\in S}), \\ (1 + 2\exp(-T^{\frac{1-2\varepsilon}{2}})) \Pr(\{y_{j,t}\}_{t\in S} \mid \{y_{i,t}\}_{t\in S}) \end{array}\right],
\end{aligned}$$

player i believes player j has $m_i[2](j) = \hat{m}_i$ with probability at least $\frac{1}{2}\exp(-\frac{1}{e}T^{\frac{1-\varepsilon}{2}})$.

4. Follows from Lemma 8.

5. Follows from Lemma 9.

9.2 Proof of Lemma 6

We have

$$\begin{aligned}
& \Pr(\{\{y_{j,t}\}_{t\in T(m)} \in A\} \mid \tilde{\zeta}_j(m) = G, a, \{y_{i,t}\}_{t\in T(m)}) \\
= \frac{\Pr(\tilde{\zeta}_j(m) = G \mid a, \{y_{j,t}\}_{t\in T(m)} \in A, \{y_{i,t}\}_{t\in T(m)})}{\Pr(\tilde{\zeta}_j(m) = G \mid a, \{y_{i,t}\}_{t\in T(m)})} & \\
& \times \Pr(\{\{y_{j,t}\}_{t\in T(m)} \in A\} \mid a, \{y_{i,t}\}_{t\in T(m)}).
\end{aligned}$$

Since

$$\begin{aligned}
\Pr(\tilde{\zeta}_j(m) &= G \mid a, \{y_{j,t}\}_{t\in T(m)}, \{y_{i,t}\}_{t\in T(m)}) \geq 1 - \exp(-|T(m)|^{1-\frac{\varepsilon}{2}}) \\
\Pr(\tilde{\zeta}_j(m) &= G \mid a, \{y_{i,t}\}_{t\in T(m)}) \leq 1
\end{aligned}$$

for all $\{y_{j,t}\}$, we have

$$\begin{aligned}
& \left| \left\{ \frac{\Pr(\tilde{\zeta}_j(m) = G \mid a, \{y_{j,t}\}_{t\in T(m)} \in A, \{y_{i,t}\}_{t\in T(m)})}{\Pr(\tilde{\zeta}_j(m) = G \mid a, \{y_{i,t}\}_{t\in T(m)})} - 1 \right\} \Pr(\{\{y_{j,t}\}_{t\in T(m)} \in A\} \mid a, \{y_{i,t}\}_{t\in T(m)}) \right| \\
\leq \;& \exp(-|T(m)|^{1-\frac{\varepsilon}{2}},
\end{aligned}$$

which derives the result.

9.3 Proof of Lemma 7

Let $\varphi_{j,t}$ be

- for m where player j is the sender of the message, $\mathbf{\Omega}_{j,t}$,
- for m where player j is the receiver of the message, $\mathbf{\Omega}_{j,t}^G$, $\mathbf{M}_{i,j,t}^G$ and $\mathbf{M}_{i,j,t}^B$, and
- for $m = (l, k, \mu)$, $\varphi_{j,t} = (\Psi_{j,t}^{a(x)}, E_j \Psi_{i,t}^{a(x)}, \Gamma_j^{a(x)}, E_j \Gamma_i^{a(x)})$.

Conditional on $\zeta_j(m+1) = G$, player j's history in the $\tilde{m}$th round with $\tilde{m} \geq m+1$ does not reveal $\{y_{j,t}\}_{t\in T(m)}$ conditional on $\{\varphi_{j,t}\}_{t\in T(m)}$.

We want to show that

$$g_i(h_j^m, \tilde{a}_{i,t}, \tilde{y}_{i,t}) = -\left\| \mathbf{y}_{j,t} - \mathbb{E}\left[\mathbf{y}_{j,t} \mid \zeta_j(m) = \zeta_j(m+1) = G, \{a_{j,\tau}, \varphi_{j,\tau}\}_{\tau\in T(m)}, \{y_{j,\tau}\}_{\tau\in T(m)\setminus\{t\}}, \tilde{a}_{i,t}, \tilde{y}_{i,t} \right] \right\|^2$$

works.[13] Let $\mathcal{Y}_{j,t}$ be the set of $y_{j,t}$ that is feasible under $\{\varphi_{j,\tau}\}_{\tau\in T(m)}, \{y_{j,\tau}\}_{\tau\in T(m)\setminus\{t\}}, \zeta_j(m) = \zeta_j(m+1) = G$. If $\mathcal{Y}_{j,t} = \emptyset$, then we set $g_i(h_j, \tilde{a}_{i,t}, \tilde{y}_{i,t}) = 0$.

[13] Here, we use Euclidean norm.

For notational simplicity, let $\Omega_t = \{a_{j,\tau}, \varphi_{j,\tau}\}_{\tau \in T(m)}, \{y_{j,\tau}\}_{\tau \in T(m) \setminus \{t\}}, a_{i,t}, y_{i,t}, \zeta_j(m) = \zeta_j(m+1) = G$, $\tilde{\Omega}_t = \{a_{j,\tau}, \varphi_{j,\tau}\}_{\tau \in T(m)}, \{y_{j,\tau}\}_{\tau \in T(m) \setminus \{t\}}, \tilde{a}_{i,t}, \tilde{y}_{i,t}, \zeta_j(m) = \zeta_j(m+1) = G$, and $\hat{\Omega}_t = \{a_{j,\tau}, \varphi_{j,\tau}\}_{\tau \in T(m)}, \{y_{j,\tau}\}_{\tau \in T(m) \setminus \{t\}}, \hat{a}_{i,t}, \hat{y}_{i,t}, \zeta_j(m) = \zeta_j(m+1) = G$. Further, let $h_i^G = (\zeta_j(m) = \zeta_j(m+1) = G, h_i)$.

Since

$$
\begin{aligned}
& \mathbb{E}\left[g_i(h_j, \tilde{a}_{i,t}, \tilde{y}_{i,t}) \mid h_i^G, (\hat{a}_{i,t}, \hat{y}_{i,t})\right] \\
= & \mathbb{E}\left[-\left\|\mathbf{y}_{j,t} - \mathbb{E}\left[\mathbf{y}_{j,t} \mid \tilde{\Omega}_t\right]\right\|^2 \mid h_i^G, h_i, (\hat{a}_{i,t}, \hat{y}_{i,t})\right] \\
= & \mathbb{E}\left[-\left\|\mathbf{y}_{j,t} - \mathbb{E}\left[\mathbf{y}_{j,t} \mid \hat{\Omega}_t\right]\right\|^2 \mid h_i^G, h_i\right] \\
= & -\left\{\mathbb{E}\left[\|\mathbf{y}_{j,t}\|^2 \mid h_i^G, h_i\right] - 2\mathbb{E}\left[\mathbf{y}_{j,t} \mid \hat{\Omega}_t\right] \cdot \mathbb{E}\left[\mathbf{y}_{j,t} \mid h_i^G, h_i\right] + \left\|\mathbb{E}\left[\mathbf{y}_{j,t} \mid \hat{\Omega}_t\right]\right\|^2\right\},
\end{aligned}
$$

we have

$$
\begin{aligned}
& \mathbb{E}\left[g_i(h_j, \tilde{a}_{i,t}, \tilde{y}_{i,t}) \mid h_i^G, (\hat{a}_{i,t}, \hat{y}_{i,t}) = (a_{i,t}, y_{i,t})\right] - \mathbb{E}\left[g_i(h_j, \tilde{a}_{i,t}, \tilde{y}_{i,t}) \mid h_i^G, h_i, (\hat{a}_{i,t}, \hat{y}_{i,t})\right] \\
= & 2\mathbb{E}\left[\mathbf{y}_{j,t} \mid \Omega_i\right] \cdot \mathbb{E}\left[\mathbf{y}_{j,t} \mid h_i^G\right] - \|\mathbb{E}\left[\mathbf{y}_{j,t} \mid \Omega_i\right]\|^2 \\
& -2\mathbb{E}\left[\mathbf{y}_{j,t} \mid \hat{\Omega}_i\right] \cdot \mathbb{E}\left[\mathbf{y}_{j,t} \mid h_i^G\right] - \left\|\mathbb{E}\left[\mathbf{y}_{j,t} \mid \hat{\Omega}_i\right]\right\|^2 \\
= & \left\{\mathbb{E}\left[\mathbf{y}_{j,t} \mid \Omega_i\right] - \mathbb{E}\left[\mathbf{y}_{j,t} \mid \hat{\Omega}_i\right]\right\} \cdot 2\mathbb{E}\left[\mathbf{y}_{j,t} \mid h_i^G\right] \\
& - \left\{\mathbb{E}\left[\mathbf{y}_{j,t} \mid \Omega_i\right] - \mathbb{E}\left[\mathbf{y}_{j,t} \mid \hat{\Omega}_i\right]\right\} \cdot \left\{\mathbb{E}\left[\mathbf{y}_{j,t} \mid \Omega_i\right] + \mathbb{E}\left[\mathbf{y}_{j,t} \mid \hat{\Omega}_i\right]\right\} \\
= & \left\{\mathbb{E}\left[\mathbf{y}_{j,t} \mid \Omega_i\right] - \mathbb{E}\left[\mathbf{y}_{j,t} \mid \hat{\Omega}_i\right]\right\} \\
& \cdot \left\{2\mathbb{E}\left[\mathbf{y}_{j,t} \mid h_i^G\right] - \mathbb{E}\left[\mathbf{y}_{j,t} \mid \Omega_i\right] - \mathbb{E}\left[\mathbf{y}_{j,t} \mid \hat{\Omega}_i\right]\right\}
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{E}\left[\mathbf{y}_{j,t} \mid \Omega_i\right] &= \left(\frac{q(y_j \mid a_{j,t}, \varphi_{j,t}, a_{i,t}, y_{i,t})}{\sum_{\hat{y}_j \in \mathcal{Y}_{j,t}} q(\hat{y}_j \mid a_{j,t}, \varphi_{j,t}, a_{i,t}, y_{i,t})}\right)_{y_j}, \\
\mathbb{E}\left[\mathbf{y}_{j,t} \mid \hat{\Omega}_i\right] &= \left(\frac{q(y_j \mid a_{j,t}, \varphi_{j,t}, \hat{a}_{i,t}, \hat{y}_{i,t})}{\sum_{\hat{y}_j \in \mathcal{Y}_{j,t}} q(\hat{y}_j \mid a_{j,t}, \varphi_{j,t}, \hat{a}_{i,t}, \hat{y}_{i,t})}\right)_{y_j}
\end{aligned}
$$

for $\mathcal{Y}_{j,t}$ with $|\mathcal{Y}_{j,t}| \geq 1$.

Generically, there exists $\bar{\varepsilon}$ such that, for any $\mathcal{Y}_j$ with $|\mathcal{Y}_j| \geq 2$ and φ_j,

$$\left| \frac{q(y_j \mid a_j, \varphi_j, a_i, y_i)}{\sum_{\hat{y}_j \in \mathcal{Y}_j} q(\hat{y}_j \mid a_j, \varphi_j, a_i, y_i)} - \frac{q(y_j \mid a_j, \varphi_j, \hat{a}_i, \hat{y}_i)}{\sum_{\hat{y}_j \in \mathcal{Y}_j} q(\hat{y}_j \mid a_j, \varphi_j, \hat{a}_i, \hat{y}_i)} \right| > \bar{\varepsilon} \tag{53}$$

for all (a_i, y_i) and $(\hat{a}_i, \hat{y}_i)$ with $(\hat{a}_i, \hat{y}_i) \neq (a_i, y_i)$.

For a moment, let us condition that $\mathcal{Y}_{j,t}$ is not empty or a singleton. Then, (53) implies that

$$\left\| \mathbb{E}\left[\mathbf{y}_{j,t} \mid \Omega_i\right] - \mathbb{E}\left[\mathbf{y}_{j,t} \mid \hat{\Omega}_i\right] \right\| > \bar{\varepsilon}$$

for all (a_i, y_i) and $(\hat{a}_i, \hat{y}_i)$ with $(\hat{a}_i, \hat{y}_i) \neq (a_i, y_i)$. Therefore, it suffices to show that, for all h_i^G,

$$\mathbb{E}\left[\mathbf{y}_{j,t} \mid h_i^G\right] - \mathbb{E}\left[\mathbf{y}_{j,t} \mid \Omega_i\right] = 0.$$

Since $\{a_{j,\tau}, \varphi_{j,\tau}\}_{\tau \in T(m)}, \{y_{j,\tau}\}_{\tau \in T(m) \setminus \{t\}}, \{a_{i,\tau}, y_{i,\tau}\}_{\tau \in T(m)}, \zeta_j(m) = \zeta_j(m+1) = G$ contain all the information in h_i^G to infer $y_{j,t}$, it suffices to show that

$$\left\| \begin{array}{c} \mathbb{E}\left[\mathbf{y}_{j,t} \mid \{a_{j,\tau}, \varphi_{j,\tau}\}_{\tau \in T(m)}, \{y_{j,\tau}\}_{\tau \in T(m) \setminus \{t\}}, \{a_{i,\tau}, y_{i,\tau}\}_{\tau \in T(m)}, \zeta_j(m) = \zeta_j(m+1) = G\right] \\ -\mathbb{E}\left[\mathbf{y}_{j,t} \mid \Omega_i\right] \end{array} \right\| = 0$$

for all $\{a_{j,\tau}, \varphi_{j,\tau}\}_{\tau \in T(m)}, \{y_{j,\tau}\}_{\tau \in T(m) \setminus \{t\}}, \{a_{i,\tau}, y_{i,\tau}\}_{\tau \in T(m)}$, which follows from

$$\begin{aligned} & \mathbb{E}\left[\mathbf{y}_{j,t} \mid \{a_{j,\tau}, \varphi_{j,\tau}\}_{\tau \in T(m)}, \{y_{j,\tau}\}_{\tau \in T(m) \setminus \{t\}}, \{a_{i,\tau}, y_{i,\tau}\}_{\tau \in T(m)}, \zeta_j(m) = \zeta_j(m+1) = G\right] \\ = & \frac{q(y_j \mid a_{j,t}, \varphi_{j,t}, a_{i,t}, y_{i,t})}{\sum_{\hat{y}_j \in \mathcal{Y}_{j,t}} q(\hat{y}_j \mid a_{j,t}, \varphi_{j,t}, a_{i,t}, y_{i,t})}. \end{aligned}$$

Therefore, the above discussion implies that, even after knowing $\{a_{j,\tau}, \varphi_{j,\tau}\}_{\tau \in T(m)}$, $\{y_{j,\tau}\}_{\tau \in T(m) \setminus \{t\}}$, $\{a_{i,\tau}, y_{i,\tau}\}_{\tau \in T(m)}$ and $\zeta_j(m) = \zeta_j(m+1) = G$,

- if $|\mathcal{Y}_{j,t}| \geq 2$, then the truthtelling is strictly optimal with incentive $\bar{\varepsilon}$, and
- if $|\mathcal{Y}_{j,t}| \leq 1$, then the truthtelling is weakly optimal.

After any history, conditional on $\zeta_j(m) = \zeta_j(m+1) = G$, player i puts probability at least $\frac{\min_{a,y,\varphi_j} q(y_j|a,y_i,\varphi_j)}{T}$ on $|\mathcal{Y}_{j,t}| \geq 2$. Therefore, after retaking $\bar{\varepsilon} = \min_{a,y,\varphi_j} q(y_j \mid a, y_i, \varphi_j)\bar{\varepsilon}$, we are done.

References

[1] Aoyagi, M. (2002): "Collusion in Dynamic Bertrand Oligopoly with Correlated Private Signals and Communication," *Journal of Economic Theory*, 102, 229-248.

[2] Bhaskar, V., and I. Obara (2002): "Belief-Based Equilibria in the Repeated Prisoners' Dilemma with Private Monitoring," *Journal of Economic Theory*, 102, 49-69.

[3] Compte, O. (1998): "Communication in Repeated Games with Imperfect Private Monitoring," *Econometrica*, 66, 597-626.

[4] Ely, J., J. Hörner, and W. Olszewski (2004): "Dispensability of Public Randomization Device," *mimeo.*

[5] Ely, J., J. Hörner, and W. Olszewski (2005): "Belief-Free Equilibria in Repeated Games," *Econometrica*, 73, 377-415.

[6] Ely, J., and J. Välimäki (2002): "A Robust Folk Theorem for the Prisoners' Dilemma," *Journal of Economic Theory*, 102, 84-105.

[7] Fong, K., O. Gossner, J. Hörner, and Y. Sannikov (2007): "Efficiency in a Repeated Prisoners' Dilemma with Imperfect Private Monitoring," *mimeo.*

[8] Fuchs, W. (2007): "Contracting with Repeated Moral Hazard and Private Evaluations," *American Economic Review*, 97, 1432-1448.

[9] Fudenberg, D., and D. Levine (2002): "The Nash-Threats Folk Theorem with Communication and Approximate Common Knowledge in Two Player Games," *Journal of Economic Theory*, 132, 461-473.

[10] Fudenberg, D., and D. Levine, and E. Maskin (1994): “The Folk Theorem with Imperfect Public Information,” *Econometrica*, 62, 997-1040.

[11] Fudenberg, D., and E. Maskin (1986): “The Folk Theorem in Repeated Games with Discounting and with Incomplete Information,” *Econometrica*, 54, 533-554.

[12] Harrington, J. E. Jr. and A. Skrzypacz (2007): “Collusion under Monitoring of Sales,” *RAND Journal of Economics*, 38, 314-331.

[13] Hörner, J. and W. Olszewski (2006): “The Folk Theorem for Games with Private Almost-Perfect Monitoring,” *Econometrica*, 74, 1499-1544.

[14] Hörner, J. and W. Olszewski (2009): “How Robust is the Folk Theorem with Imperfect Public Monitoring?,” *Quarterly Journal of Economics*, 124, 1773-1814.

[15] Kandori, M. (2002): “Introduction to Repeated Games with Private Monitoring,” *Journal of Economic Theory*, 102, 1-15.

[16] Kandori, M. (2008): “Weakly Belief-Free Equilibria in Repeated Games with Private Monitoring,” *mimeo.*

[17] Kandori, M., and H. Matsushima (1998): “Private Observation, Communication and Collusion,” *Econometrica*, 66, 627-652.

[18] Kandori, M., and I. Obara (2006): “Efficiency in Repeated Games Revisited: the Role of Private Strategies,” *Econometrica*, 72, 499-519.

[19] Kandori, M., and I. Obara (2010): “Towards a Belief-Based Theory of Repeated Games with Private Monitoring: An Application of POMDP,” *mimeo.*

[20] Mailath, G. and S. Morris (2002): “Repeated Games with Almost-Public Monitoring,” *Journal of Economic Theory* 102, 189-228.

[21] Mailath, G., and S. Morris (2006): “Coordination Failure in Repeated Games with Almost-Public Monitoring,” *Theoretical Economics*, 1, 311-340.

[22] Mailath, G., and L. Samuelson (2006): *Repeated Games and Reputations: Long-Run Relationships.* Oxford University Press, New York, NY.

[23] Matsushima, H. (2004): "Repeated Games with Private Monitoring: Two Players," *Econometrica*, 72, 823-852.

[24] Miyagawa, E., Y. Miyahara, and T. Sekiguchi (2008): "The Folk Theorem for Repeated Games with Observation Costs," *Journal of Economic Theory*, 139, 192-221.

[25] Obara, I. (2009): "Folk Theorem with Communication," *Journal of Economic Theory*, 144, 120-134.

[26] Phelan, C. and A. Skrzypacz: "Beliefs and Private Monitoring," *mimeo*

[27] Piccione, M. (2002): "The Repeated prisoners' Dilemma with Imperfect Private Monitoring," *Journal of Economic Theory*, 102, 70-83.

[28] Radner, R. (1985): "Repeated Principal-Agent Games with Discounting," *Econometrica*, 53, 1173-1198.

[29] Sekiguchi, T. (1997): "Efficiency in Repeated prisoners' Dilemma with Private Monitoring," *Journal of Economic Theory*, 76, 345-361.

[30] Stigler, G. (1964): "A Theory of Oligopoly," Journal of Political Economy, 72, 44-61.

[31] Sugaya, T. (2010a): "Generic Characterization of the Set of Belief-Free Review-Strategy Equilibrium Payoffs," *mimeo.*

[32] Sugaya, T. (2010b): "Folk Theorem in a Repeated Prisoners' Dilemma without Conditional Independence," *mimeo.*

[33] Sugaya, T. (2010c): "Folk Theorem in a Repeated Prisoners' Dilemma with a Generic Private Monitoring," *mimeo.*

[34] Sugaya, T. and S. Takahashi (2010): "Coordination Failure in Repeated Games with Private Monitoring," *mimeo.*

[35] Takahashi, S. (2007): "Community Enforcement when Players Observe Partners' Past Play," *mimeo.*

[36] Yamamoto, Y. (2007): "Efficiency Results in N Player Games with Imperfect Private Monitoring," *Journal of Economic Theory*, 135, 382–413.

[37] Yamamoto, Y. (2009a): "A Limit Characterization of Belief-Free Equilibrium Payoffs in Repeated Games," forthcoming in *Journal of Economic Theory.*

[38] Yamamoto, Y. (2009b): "Repeated Games with Private and Independent Monitoring," *mimeo.*