Folk Theorem in Repeated Games with Private Monitoring

Takuo Sugaya*†
Princeton University

September 2, 2011

Abstract

We show that the folk theorem with individually rational payoffs defined by pure strategies generically holds for N-player repeated games with private monitoring when each player's number of 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 invaluable advice, Yuliy Sannikov for helpful discussion, and Dilip Abreu, Edoardo Grillo, Johannes Hörner, Yuhta Ishii, Michihiro Kandori, Vijay Krishna, George Mailath, Tadashi Sekiguchi, Andrzej Skrzypacz, Satoru Takahashi, Yuichi Yamamoto and seminar participants at 10th SAET conference, 10th World Congress of the Econometric Society, 22nd Stony Brook Game Theory Festival and Summer Workshop on Economic Theory 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 future play based on common knowledge of relevant histories. Specifically, the public component of histories, such as action profiles in perfect monitoring or public signals in public monitoring, reveals past action profiles (at least statistically). Since this public information is common knowledge, players can coordinate a punishment contingent on the public information, and thereby provide dynamic incentives to choose actions that are not static best responses.

With *private monitoring*, players can observe only private noisy signals about the action profile. Common knowledge no longer exists and 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. For example, Kandori (2002) mentions that "[t]his is probably one of the best known long-standing open questions in economic theory."

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 and cannot observe their opponents' prices. Instead, a firm obtains some information about opponents' prices through its own sales. Since the sales level depends on both opponents' prices and unobservable shocks due to business cycles, the sales is an imperfect signal. In addition, each firm's sales is often private information. Thus, monitoring is imperfect and private. In addition, Fuchs (2007) applies a repeated game

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

with private monitoring to a contract between a principal and an agent, and Harrington and Skrzypacz (2007) analyze cartel behaviors using the framework of a repeated game with private monitoring.

This paper is the first to show that the folk theorem holds in repeated games with generic private monitoring. We unify and improve the three approaches in the literature on private monitoring that have been 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 the prisoners' dilemma.[2] A strategy profile is *belief-free* if, for any history profile, the continuation strategy of each player 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, only limited results have been shown without almost perfect or conditionally independent monitoring: Fong, Gossner, Hörner and Sannikov (2010) show the payoff of the mutual cooperation is approximately attainable 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 payoff of the mutual cooperation is approximately attainable 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 equilibria in public monitoring to almost public monitoring. Based on their insights, Hörner and Olszewski (2009) establish

[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) and Deb (2011) consider the community enforcement and Miyagawa, Miyahara and Sekiguchi (2008) consider the situation where a player can improve the monitoring by paying cost.

the folk theorem for almost public monitoring. Phelan and Skrzypacz (2009) characterize the set of possible beliefs about opponents' states in a finite-state automaton strategy and Kandori and Obara (2010) offer an easy way to verify if a finite-state automaton strategy is an equilibrium.

Another approach to analyzing repeated games with private monitoring is to introduce public communication. 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 coordination through private signals. However, the analyses are not applicable 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 although the robustness to a small private noise 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. At the beginning of each review phase, every on-path strategy of each player is optimal conditional on the histories of the opponents. Second, however, the belief-free property does not hold except at the beginning of the phases. Hence, we consider each player's statistical inference about the opponents' past histories as in belief-based approach within each phase. Finally, in our equilibrium, the players do communicate but the message exchange can be done with their actions. One of our methodological contributions is to offer a systematic way to dispense the cheap talk message protocol with message exchange via their actions.

The rest of the paper is organized as follows. Section 2 introduces the model and Section 3 states the assumptions and main result. The remaining parts of the paper are devoted to its proof. Since the complete proof is long and complicated, in the proof of the main text, we illustrate the main structure by focusing on the two-player prisoners' dilemma with cheap talk and public randomization and refer the complete proof for a general game without cheap talk and public randomization device to the Supplemental Materials. Section 4 relates the infinitely repeated game to a finitely repeated game with an auxiliary scenario (reward function) as Hörner and Olszewski (2006) and derives a sufficient condition on the finitely repeated game to show the folk theorem. In Section 5, we intuitively explain the equilibrium construction. In particular, after explaining Matsushima (2004) with conditionally independent signals, we explain how to extend his result to conditionally dependent signals. We will see that the players need to coordinate their future play through private signals. After we formally define the structure of the finitely repeated game in Section 6 and strategy in Section 8, Section 9 explains how the coordination works. Sections 11 and 12 verify the strategy satisfies the sufficient condition for the folk theorem derived in Section 4. All the proofs are given in the Appendix. In Sections 13 and 14, we comment on how we generalize the proof in the main text to the case for a general game without cheap talk and public randomization device in the Supplemental Materials.

2 Model

2.1 Stage Game

The stage game is given by $\left\{I, (A_i, Y_i, \tilde{u}_i)_{i \in I}, q\right\}$. $I = \{1, \ldots, N\}$ 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 $a \equiv (a_1, \ldots, a_N) \in A$. Then, a signal profile $y = (y_1, \ldots, y_N) \in Y$ is realized according to a joint conditional probability function $q(y \mid a)$. 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 (common) discount factor is $\delta \in (0,1)$. Let $a_{i,\tau}$ and $y_{i,\tau}$ denote respectively the action played and the private signal observed in period τ by player i. Player i's private history up to period $t \geq 1$ is given by $h_i^t \equiv (a_{i,\tau}, y_{i,\tau})_{\tau=1}^{t-1}$. With $h_i^1 = \{\emptyset\}$, for each $t \geq 1$, let H_i^t be the set of all h_i^t. 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 Assumptions and Result

In this section, we state two assumptions and the main result (folk theorem).

First, we state an assumption on the payoff structure. Let $F \equiv \text{co}(\{u(a)\}_{a \in A})$ 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 minimax. 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 *The stage game payoff structure satisfies the full dimensionality condition:* $\dim(F^*) = N$*.*

Second, we state an assumption on the signal structure.

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

$$|Y_i| \geq \max\left\{ \max_{j \neq i} |A_j| + 2 \sum_{n \neq i, i-1} |A_n|, \sum_{j \in I} |A_j|, \max_{j \in I} 2|A_j| \right\}.$$

Note that RHS is bounded by a lineaer function of $\sum_j |A_j|$. 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* $\delta > \bar{\delta}$*,* $v \in E(\delta)$*.*

Since the full support assumption $q(y \mid a) > 0$ for all $a \in A$ and $y \in Y$ is generic, we assume the monitoring is full support. Then, any sequential equilibrium is realization equivalent to a Nash equilibrium. Hence, for the rest of the paper, we consider a Nash equilibrium.

For the proof in the main text, we focus on the two-player prisoners' dilemma with perfect and public cheap talk and public randomization device: $I = 2$, $A_i = \{C_i, D_i\}$ 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. For notational convenience, whenever we say players i and j, unless otherwise mentioned, i and j are different. In the two-player prisoners' dilemma, Assumptions 1 and 2 are equivalent to $|Y_i| \geq 4$ for all i. Furthermore, we focus on v with

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

Section 13 briefly explains how to show the folk theorem for a general game and Section 14 explains how to dispense cheap talk and public randomization device. See the Supplemental Materials for the formal proofs.

We prove the theorem with the following steps. We arbitrarily fix v with (2) and 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 sufficient conditions on a strategy and an auxiliary scenario in the T_P-period finitely repeated game from which we can construct an equilibrium strategy to implement v in the infinitely repeated game. In Section 5, we intuitively explain the structure of the equilibrium. Given this, Section 6 explains the formal structure of the finitely repeated game and Section 8 explains the equilibrium strategy. In Sections 9, 11 and 12, we verify that the strategy satisfies the sufficient conditions in the finitely repeated game.

4 Finitely Repeated Game

In this section, we consider a T_P-period *finitely* repeated game with auxiliary scenarios. We derive sufficient conditions on strategies and auxiliary scenarios in the finitely repeated game such that we can construct a strategy in the infinitely repeated game to support v. The sufficient conditions are stated in Lemma 1.

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 an "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 histories in the finitely repeated game to the real numbers.

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 $\sigma_i(x_i)$ and $\pi_j(x_i, \cdot : \delta)$ such that, for sufficiently large δ, there exists T_P such that, for any $i \in I$,

1. For any $x_j \in \{G, B\}$, $\sigma_i(G)$ 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 ($x_j = B$):

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

Intuitively, for sufficiently large δ, since $\lim_{\delta\to 1}\frac{1-\delta}{1-\delta^{T_P}} = \frac{1}{T_P}$, this requires that the time average of the expected sum of the instantaneous utilities and the reward is close 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

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

We call (6) the "feasibility constraint."[3]

The following lemma gives us a sufficient condition about the finitely repeated game to show the folk theorem in the infinitely repeated game:

Lemma 1 *For Theorem 1, it suffices to show that there exist* $\{\{\sigma_i(x_i)\}_{x_i\in\{G,B\}}\}_{i\in I}$ *and* $\{\{\pi_i(x_j, \cdot : \delta)\}_{x_j\in\{G,B\}}\}_{i\in I}$ *satisfying (4), (5) and (6) in the* T_P*-period finitely repeated game.*

5 Intuitive Explanation

5.1 With Conditionally Independent Signals

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

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

[3] For notational convenience, when we say that (6) is satisfied, it also implies that $\pi_i(G, h_j^{T_P+1} : \delta)$ and $\pi_i(B, h_j^{T_P+1} : \delta)$ are uniformly bounded with respect to δ.

for all $a \in A$ and $y \in Y$. That is, conditional on an action profile a, player i's signal has no information about the opponent's signal. In this case, it is easy to construct $\sigma_i(x_i)$ and $\pi_i(x_j, \cdot : \delta)$ satisfying (4), (5) and (6).

With conditional independence, we can see a T_P-period finitely repeated game as a T-period "review round" with a sufficiently long $T_P = T$.

$\sigma_i(x_i)$ is simply defined as follows: At the beginning of the finitely repeated game, send the message $x_i \in \{G, B\}$ by cheap talk. Then, constantly take

$$a_i(x_i) \equiv \begin{cases} C_i & \text{if } x_i = G, \\ D_i & \text{if } x_i = B \end{cases} \tag{7}$$

for T periods. That is, player i with $\sigma_i(G)$ who wants to make player j's value high takes C_i and player i with $\sigma_i(B)$ who wants to make player j's value low takes D_i.

We are left to construct $\pi_i(x_j, \cdot : \delta)$. We concentrate on the case with $x_j = G$ since the case with $x_j = B$ is symmetric. 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:

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

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

$$\begin{aligned} & \lim_{T\to\infty} \lim_{\delta\to 1} \frac{1-\delta}{1-\delta^{T_P}} \mathbb{E}\left[\sum_{t=1}^{T_P} \delta^{t-1} u_{i,t}(a) + \pi_i(G, h_j^{T_P+1} : \delta) \mid \sigma_i(B), \sigma_j(G)\right] \\ = & \lim_{T\to\infty} \frac{1}{T}\left[\sum_{t=1}^{T} u_i(D_i, C_j) - \rho T\right] = u(D_i, C_j) - \rho > \bar{v}_i \text{ from (3).} \end{aligned} \tag{8}$$

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).

On the other hand, if player j receives the message $x_i = G$, then player j needs to incentivize player i to take C_i by the auxiliary scenario $\pi_i(x_j, \cdot : \delta)$. Suppose there exist statistics $\Psi_{j,t}^{C_i,a_j} \in \{0,1\}$ and $q_2 > q_1$ such that, for each $a_j \in A_j$, $\Psi_{j,t}^{C_i,a_j} = 1$ indicates that C_i is more likely to be played:

$$\mathbb{E}\left[\Psi_{j,t}^{C_i,a_j} \mid a_i, a_j\right] = \begin{cases} q_2 & \text{if } a_i = C_i, \\ q_1 & \text{if } a_i = D_i. \end{cases} \tag{9}$$

The existence of such $\Psi_{j,t}^{C_i,a_j}$, q_2 and q_1 will be proven in Lemma 3. Take $\bar{L}$ such that

$$\bar{L}\left(q_2 - q_1\right) > \max_{a,i} 2\left|u_i\left(a\right)\right|. \tag{10}$$

Since player j with $x_j = G$ takes C_j, player j rewards player i based on $\sum_{t=1}^{T} \Psi_{j,t}^{C_i,C_j}$, the summation of $\Psi_{j,t}^{C_i,C_j}$:

$$\pi_i(G, h_j^{T_P+1} : \delta) \equiv \bar{L}\left\{\sum_{t=1}^{T} \Psi_{j,t}^{C_i,C_j} - (q_2 T + 2\varepsilon T)\right\}_{-} - \rho T$$

with some small $\varepsilon > 0$. The reward is linearly increasing in $\Psi_{j,t}^{C_i,C_j}$ with slope $\bar{L}$ until $\sum_{t=1}^{T} \Psi_{j,t}^{C_i,C_j}$ hits the upper bound $q_2 T + 2\varepsilon T$. If $\sum_{t=1}^{T} \Psi_{j,t}^{C_i,C_j}$ does not hit the upper bound, since playing C_i instead of D_i increases the expectation of $\Psi_{j,t}^{C_i,C_j}$ by $(q_2 - q_1)$ from (9), the marginal gain of taking C_i is $\bar{L}\left(q_2 - q_1\right)$. Since (10) implies that this marginal gain dominates the difference in the instantaneous utilities, player i has the incentive to take C_i.

Hence, in order to show that player i takes C_i, it suffices to show that, *regardless of player i's signal observations*, player i believes that $\sum_{t=1}^{T} \Psi_{j,t}^{C_i,C_j}$ hits the upper bound with little probability: For all h_i,

$$\Pr\left(\left\{\sum_{t=1}^{T} \Psi_{j,t}^{C_i,C_j} > q_2 T + 2\varepsilon T\right\} \mid x_j = G, h_i\right) \leq \exp(-\left(q_2 - q_1\right) T).$$

This can be shown as follows: If the monitoring is conditionally independent, *regardless of player i's signal observations*, player i's beliefs on $\sum_{t=1}^{T}\Psi_{j,t}^{C_i,C_j} \mid x_j = G, h_i$ are approximately distributed according to the normal distribution with the mean

$$\mathbb{E}\left[\sum_{t=1}^{T}\Psi_{j,t}^{C_i,C_j} \mid C_i, C_j\right] = q_2 T$$

and the standard deviation $O(T^{\frac{1}{2}})$ by the central limit theorem.[4] Since $q_2T + 2\varepsilon T$ is greater than the mean 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 $\sum_{t=1}^{T}\Psi_{j,t}^{C_i,C_j}$ hits the upper bound is negligible.

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

$$\begin{aligned}
& \lim_{T\to\infty}\lim_{\delta\to 1}\frac{1-\delta}{1-\delta^{T_P}}\mathbb{E}\left[\sum_{t=1}^{T_P}\delta^{t-1}u_{i,t}\left(a\right) + \pi_i(G, h_j^{T_P+1} : \delta) \mid \sigma_i\left(G\right), \sigma_j\left(G\right)\right] \\
= & \lim_{T\to\infty}\frac{1}{T}\mathbb{E}\left[\sum_{t=1}^{T}u_{i,t}\left(a\right) + \bar{L}\left\{\sum_{t=1}^{T}\Psi_{j,t}^{C_i,C_j} - \left(q_2T + 2\varepsilon T\right)\right\}_{-} - \rho T \mid \sigma_i\left(G\right), \sigma_j\left(G\right)\right] \\
= & u\left(C_i, C_j\right) - 2\varepsilon\bar{L} - \rho,
\end{aligned}$$

which is larger than $\bar{v}_i$ for sufficiently small $\varepsilon > 0$ from (3). 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).

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.

[4] Here, we consider the incentive only on the equilibrium path. Since defection decreases the expected value of $\Psi_{j,t}^{C_i,C_j}$, together with conditional independence, defection reduces the probability that $\sum_{t=1}^{T}\Psi_{j,t}^{C_i,C_j}$ hits the upper bound even more.

5.2 With Conditionally Dependent Signals

5.2.1 Problem

Now we consider the two-player prisoners' dilemma with *conditionally dependent* monitoring and explain the difficulty. 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 $\sum_{t=1}^{T} \Psi_{j,t}^{C_i,C_j}$ from her own history very precisely. (i) Equation (5) requires that the time average of the expected sum of the instantaneous utilities 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 $\sum_{t=1}^{T} \Psi_{j,t}^{C_i,C_j}$ is close to its ex ante value $q_2 T$, then the time average of the reward should be close to 0. This means that player i cannot be rewarded if $\sum_{t=1}^{T} \Psi_{j,t}^{C_i,C_j}$ is unusually high. If, in period τ, player i infers from h_i^τ that $\sum_{t=1}^{\tau} \Psi_{j,t}^{C_i,C_j}$ is already unusually high (greater than $q_2 T + 2\varepsilon T$) with high probability, then player i stops cooperation after period τ.[5]

Specifically, since the reward after receiving $x_i = G$ is $\bar{L}\{\sum_{t=1}^{T} \Psi_{j,t}^{C_i,C_j} - (q_2 T + 2\varepsilon T)\}_- - \rho T$, if player i believes that $\sum_{t=1}^{T} \Psi_{j,t}^{C_i,C_j}$ has hit the upper bound $q_2 T + 2\varepsilon T$ with high probability, then player i wants to stop cooperation. The problem is that this stops increasing when $\sum_{t=1}^{\tau} \Psi_{j,t}^{C_i,C_j} > q_2 T + 2\varepsilon T$.

5.2.2 Modification

Structure of the Phase To deal with the problem above, we consider the following modification. First, for a moment,[6] let us see the T_P-period finitely repeated game as L

[5] This is also mentioned by Fong, Gossner, Hörner and Sannikov (2010).

[6] We will modify the structure further to deal with problems arising later. See Section 6 for the final structure.

repetition of the T-period review rounds with

$$\rho L > \bar{L}. \tag{11}$$

Now, $T_P = LT$.

[Insert Figure1].

For notational convenience, let $T(l)$ be the set of periods in the lth review round. When player j monitors player i, player j randomly drops one period $t_j(l)$ from $T(l)$. That is, $\Pr(\{t_j(l) = t\}) = \frac{1}{T}$ for all $t \in T(l)$. Let $T_j(l) \equiv T(l) \setminus \{t_j(l)\}$ be the remaining periods in the lth review round. Player j monitors player i during $T(l)$ by

$$X_j(l) \equiv \begin{cases} \sum_{t \in T_j(l)} \Psi_{j,t}^{C_i,C_j} + \mathbf{1}_{t_j(l)} & \text{if } x_j = G, \\ \sum_{t \in T_j(l)} \Psi_{j,t}^{C_i,D_j} + \mathbf{1}_{t_j(l)} & \text{if } x_j = B. \end{cases} \tag{12}$$

Here, $\mathbf{1}_{t_j(l)} \in \{0,1\}$ is a random variable with $\Pr(\{\mathbf{1}_{t_j(l)} = 1\}) = q_2$ conditional on $t_j(l)$. Hence, instead of monitoring by $\sum_{t \in T(l)} \Psi_{j,t}^{C_i,C_j}$, player j randomly picks $t_j(l)$ and replaces $\Psi_{j,t_j(l)}^{C_i,C_j}$ with the random variable $\mathbf{1}_{t_j(l)}$ that is independent of the players' action. The reason will be explained in Section 15.7.

Reward Function by Player j Second, we will heuristically define the reward function by player j.[7] There are following cases:

$x_i = B$**:** When player j receives the message $x_i = B$, then player i takes a dominant action and no incentive by the auxiliary scenario is necessary. Hence, it is straightforward to construct a constant reward as in Section 5.1, achieving (4), (5) and (6). In this case, we say $\lambda_j(l) = G$ for all $l = 1, ..., L$. See the case with $x_i = G$ for the meaning of $\lambda_j(l)$.

$x_i = G$**:** When player j receives the message $x_i = G$, then we need to deal with the problem mentioned in Section 5.2.1. See each of the L review rounds independent and

[7] See Section 8.3 for the formal definition of the reward function.

consider the following reward function for each lth round:

$$\pi_i(G, h_j^{T_P+1}, l) = \bar{L}\{X_j(l) - (q_2 T + 2\varepsilon T)\} - \rho T \quad (13)$$

$$\pi_i(B, h_j^{T_P+1}, l) = \bar{L}\{X_j(l) - (q_2 T - 2\varepsilon T)\} + \rho T \quad (14)$$

If we make

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

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 C_i.

Since $\pi_i(x_j, h_j^{T_P+1} : \delta) = \sum_{l=1}^{L} \pi_i(x_j, h_j^{T_P+1}, l)$ does not satisfy (6), we need further modification. Observe th following: For $x_j = G$, if $\pi_i > -\rho T$ (that is, $X_j(l) > q_2 T + 2\varepsilon T$) occurs at most for one review round, then (6) is still satisfied. To see why, the maximum reward in one review round 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 $X_j(l) \leq q_2 T + 2\varepsilon T$, then $\pi_i(G, h_j^{T_P+1}, l) \leq -\rho T$. Since we take $\bar{L} < L\rho$ in (11), the total reward in the finitely repeated game is still negative if $X_j(l) > q_2 T + 2\varepsilon T$ occurs only for one review round. Symmetrically, for $x_j = B$, if $\pi_i < \rho T$ (that is, $X_j(l) < q_2 T - 2\varepsilon T$) occurs at most for one review round, then (6) is still satisfied.

Therefore, once

$$X_j(l) \notin [q_2 T - 2\varepsilon T, q_2 T + 2\varepsilon T]$$

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

$$\pi_i(G, h_j^{T_P+1}, \tilde{l}) = -\rho T \quad (15)$$

$$\pi_i(B, h_j^{T_P+1}, \tilde{l}) = \rho T. \quad (16)$$

Importantly, (6) is recovered since $\pi_i(G, h_j^{T_P+1}, l) > -\rho T$ and $\pi_i(B, h_j^{T_P+1}, l) < \rho T$ happen at most for one review round by definition.

Let $\lambda_j(l) \in \{G, B\}$ denote which reward function player j is using in the lth review round:

- In the first review round, $\lambda_j(1) = G$ and the reward is (13) or (14).
- In the $(l+1)$th review round with $l \geq 1$,
  - If $\lambda_j(l) = B$, then $\lambda_j(l+1) = B$ and the reward is (15) or (16).
  - If $\lambda_j(l) = G$, then
    - If $X_j(l) \in [q_2T - 2\varepsilon T, q_2T + 2\varepsilon T]$, then $\lambda_j(l+1) = G$ and the reward is (13) or (14).
    - If $X_j(l) \notin [q_2T - 2\varepsilon T, q_2T + 2\varepsilon T]$, then $\lambda_j(l+1) = B$ and the reward is (15) or (16).

See Figure 2 for the automaton representation. Hence, $\lambda_j(l) = G$ implies that the reward in the lth review round is (13) or (14) while $\lambda_j(l) = B$ implies that the reward in the lth review round is (15) or (16).

Note that *ex ante*, if the players play $a_i(x_i), a_j(x_j)$, 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 that $X_j(l)$ is far away from its ex ante value by $2T^{\frac{1}{2}}$ times the order of standard deviation, $\lambda_j(l+1) = B$ happens only with small probability less than $\exp(-(q_2 - q_1)T)$. Hence, we call the observation is "erroneous" if $X_j(l) \notin [q_2T - 2\varepsilon T, q_2T + 2\varepsilon T]$.

[Insert Figure 2].

Optimal Action of Player i Third, given player j's reward function, we will derive the optimal action of player i. Since the shape of the reward function changes as $\lambda_j(l)$ changes, player i needs to infer $\lambda_j(l)$. 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.

After sending $x_i = B$ Since $\lambda_j(l) = G$ for any history of player j, $\hat{\lambda}_j(l) = G$. Since the reward function is constant, it is optimal to take $a_i(x_i) = D_i$. Therefore, $\sigma_i(B)$ such that player i always takes D_i is optimal.

After sending $x_i = G$ If $\hat{\lambda}_j(l) = B$, then since the reward is constant for the rest of the finitely repeated game, it is optimal to take D_i. We verify the incentive to take cooperation when $\hat{\lambda}_j(l) = G$ by the backward induction.

For the last review round $l = L$,

- If $\hat{\lambda}_j(L) = G$, then player i wants to take C_i since the reward is linearly increasing in $X_j(l)$ with slope $\bar{L}$ and there is no effect on $\lambda_j(L+1)$. Player i's average continuation payoff at the beginning of Lth review round is

  – If $x_j = G$, then

$$\begin{aligned}
&\lim_{T\to\infty}\lim_{\delta\to 1}\frac{1-\delta}{1-\delta^T}\mathbb{E}\left[\begin{array}{c}\sum_{t\in T(L)}\delta^{t-t_L}u_i(C_i,C_j)\\ -\delta^{-t_L+1}\left(\bar{L}\{X_j(L)+(q_2T+2\varepsilon T)\}+\rho T\right)\mid C_i,C_j\end{array}\right]\\
&= u(C_i,C_j)-2\varepsilon\bar{L}-\rho\geq\bar{v}_i
\end{aligned} \tag{17}$$

  with t_L being the first period of the Lth review round.

  – If $x_j = B$, then

$$\begin{aligned}
&\lim_{T\to\infty}\lim_{\delta\to 1}\frac{1-\delta}{1-\delta^T}\mathbb{E}\left[\begin{array}{c}\sum_{t\in T(L)}\delta^{t-t_L}u_i(C_i,D_j)\\ +\delta^{-t_L+1}\left(\bar{L}\{X_j(L)-(q_2T-2\varepsilon T)\}+\rho T\right)\mid C_i,D_j\end{array}\right]\\
&= u(C_i,D_j)+2\varepsilon\bar{L}+\rho\leq\underline{v}_i
\end{aligned} \tag{18}$$

  Note that replacing $\Psi_{j,t_j(l)}^{C_i,C_j}$ with $\mathbf{1}_{t_j(l)}$ in (12) does not change the incentive and value in the limit where T goes to infinity.

- If $\hat{\lambda}_j(L) = B$, then player i wants to take D_i since the reward is constant. Player i's average continuation payoff at the beginning of Lth review round is

– If $x_j = G$, then

$$\lim_{T\to\infty}\lim_{\delta\to 1}\frac{1-\delta}{1-\delta^T}\left[\sum_{t\in T(L)}\delta^{t-t_L}u_i(D_i,C_j)-\delta^{-t_L+1}\rho T\right]=u\left(D_i,C_j\right)-\rho\geq\bar{v}_i. \quad (19)$$

– If $x_j = B$, then

$$\lim_{T\to\infty}\lim_{\delta\to 1}\frac{1-\delta}{1-\delta^T}\left[\sum_{t\in T(L)}\delta^{t-t_L}u_i(D_i,D_i)+\delta^{-t_L+1}\rho T\right]=u\left(D_i,D_j\right)+\rho\leq\underline{v}_i. \quad (20)$$

We can subtract proper positive numbers depending only on x and $\lambda_j(L)$ from (17) and (19) such that the continuation payoff is exactly $\bar{v}_i$ if $x_j = G$. Similarly, we can add proper positive numbers depending only on x and $\lambda_j(L)$ to (18) and (20) such that the continuation payoff is exactly $\underline{v}_i$ if $x_j = B$. Note that we can keep (6).

Therefore, $\sigma_i(G)$ such that player i with $\hat{\lambda}_j(L) = G$ plays C_i and with $\hat{\lambda}_j(L) = B$ plays D_i is optimal and player i's continuation payoff is independent of $\lambda_j(L)$.

For the second last review round $l = L-1$, since the continuation payoff from the Lth review round is constant for $\lambda_j(L)$, player i do not need to consider the effect of her strategy in the $(L-1)$th review round on $\lambda_j(L)$. Therefore, the same proof shows that $\sigma_i(G)$ such that player i with $\hat{\lambda}_j(L-1) = G$ plays C_i and with $\hat{\lambda}_j(L-1) = B$ plays D_i is optimal. Further, we can make sure that player i's continuation payoff at the beginning of the $(L-1)$th review round is $\bar{v}_i$ or $\underline{v}_i$ depending on x_j but independently of $\lambda_j(L-1)$.

Recursively, we can show that $\sigma_i(G)$ such that player i with $\hat{\lambda}_j(l) = G$ plays C_i and with $\hat{\lambda}_j(l) = B$ plays D_i for all l is optimal and gives player i $\bar{v}_i$ or $\underline{v}_i$ depending on x_j.

In summary, player i's action in the lth review round is defined as in Figure 3:

[Insert Figure 3]

Reward Function of Player j Revisited Fourth, we modify player j's reward function further to incorporate the fact that player j with $\hat{\lambda}_i(l) = B$ (player j infers that player i

has observed erroneous histories) takes $D_j \neq a_j(x_j)$. As a preparation, we construct special reward functions $\pi_i^G(a_j, y_j)$ and $\pi_i^B(a_j, y_j)$ that make player i indifferent between any action profile.

Lemma 2 *Generically, the following statement is true: For any $\rho > 0$, there exists $\bar{u} > \rho$ such that, for each $i \in I$, there exist $\pi_i^G : A_j \times Y_j \to [-\bar{u}, -\rho]$ and $\pi_i^B : A_j \times Y_j \to [\rho, \bar{u}]$ such that*

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

When $\hat{\lambda}_i(l) = B$, player j uses the reward $\pi_i^{x_j}(a_{j,t}, y_{j,t})$ for each t in the $\tilde{l}$th review round with $\tilde{l} \geq l$ so that player i can always assume $\hat{\lambda}_i(l) = G$. Therefore, the modified reward function is explained in the following figure:

[Insert Figure 4]

The incentive compatibility that player i does not try to manipulate $\hat{\lambda}_j(l)$ will be proven in Proposition 1.

5.2.3 Summary from the Perspectives of Infinitely Repeated Games

Here, we offer the summary of our equilibrium construction, using the language of infinitely repeated games. As we can see from the proof of Lemma 1, we can see the finitely repeated game as the first T_P periods in the infinitely repeated game and a positive (negative, respectively) reward implies that in the continuation play from period $T_P + 1$ is higher (lower, respectively) than the value in the initial period.

Suppose $x_j = G$. Then, the value for player i in the initial period needs to be close to $u_i(C, C)$ to attain efficiency. Since the value in the initial period is very high, player j with $x_j = G$ cannot reward player i with high realization of auxiliary scenario ($\pi_i > -\rho T$) by going to a higher continuation payoff from period $T_P + 1$. Instead, player j "allows" player i to take defection and "rewards" player i by higher instantaneous utilities. While doing so, player j's incentive is given by the changes in the continuation payoff from period $T_P + 1$. That is, the changes in the continuation payoff for player j while player i with $\hat{\lambda}_j(l) = B$ defects correspond to the realization of the auxiliary scenario in Lemma 2 (the roles of players i and j are reversed).

5.2.4 Coordination Problem

We are left to show how player i infers $\lambda_j(l)$. The rest of the paper is mainly devoted to the coordination between $\lambda_j(l) \in \{G, B\}$ and $\hat{\lambda}_j(l) \in \{G, B\}$. In Section 6, we further modify the structure of the finitely repeated game by introducing the supplemental rounds where player j sends the message about $\lambda_j(l+1)$ by her actions after the lth review round. In Section 8, we formally define the equilibrium strategy (actions and rewards) except for how player i infers player j's message in the supplemental rounds. Then, in Section 9, we specify player i's inference of player j's message in the supplemental rounds. This fully pins down the strategy (actions and rewards).

While we define the equilibrium, we will introduce variables with various restrictions. In Section 10, we verify that we can take these variables consistently.

Then, in Sections 11 and 12, we verify the equilibrium strategy (actions and rewards) satisfies (4) (5) and (6).

6 Structure of the Phase

In this section, we explain the structure of the T_P-period finitely repeated game, which is summarized in Figure 5 below. T_P depends on L and T. $L \in \mathbb{N}$ will be pinned down in Section 10. $T \in \mathbb{N}$ is a parameter of the equilibrium.

At the beginning of the finitely repeated game, we insert the "coordination block" where each player i with $\sigma_i(x_i)$ sends x_i by cheap talk simultaneously. At the end of the finitely repeated game, we insert the "report block" where the players report the whole history in the finitely repeated game by cheap talk. The detail of the report block will be explained in Section 12. These blocks are instantaneous with cheap talk but will take multiple periods when we dispense cheap talk in the Supplemental Materials 3 and 4.

Between the coordination and report blocks, the players play T_P-periods finitely repeated game. We divide T_P periods into L "main blocks." The first $(L-1)$ blocks is further divided into the following five rounds: For $l \in \{1, ..., L-1\}$, the lth block consists of T-period review round, $T^{\frac{1}{2}}$-period supplemental round 1 for $\lambda_1(l+1)$, $T^{\frac{1}{2}}$-period supplemental round 2 for $\lambda_1(l+1)$, $T^{\frac{1}{2}}$-period supplemental round 1 for $\lambda_2(l+1)$ and $T^{\frac{1}{2}}$-period supplemental round 2 for $\lambda_2(l+1)$.[8] The last Lth block has only the T-period review round. In the lth block, the sets of periods in the review round, supplemental round 1 for $\lambda_i(l+1)$ and supplemental round 2 for $\lambda_i(l+1)$ respectively are denoted by $T(l)$, $T(l, \lambda_i, 1)$ and $T(l, \lambda_i, 2)$. For sufficiently large T, the length of the review round is much larger than that of the other four rounds and the payoffs from the review rounds approximately determine the equilibrium payoff. See Figure 5 for the illustration.

[Insert Figure 5]

We show that, for sufficiently large T, for sufficiently large δ, with $T_P = (L-1)\left\{T + 4T^{\frac{1}{2}}\right\} + T$, there exist $\sigma_i(x_i)$ and $\pi_i(x_j, \cdot : \delta)$ satisfying (4), (5) and (6).

7 Almost Optimality

Instead of proving (4), we only establish the "almost optimality with $\exp(-T^{\frac{1}{4}}) > 0$" until Section 12: For all $i \in I$ and $x \in \{G, B\}^2$, for any τ and h_i^τ, the loss of playing the

[8] Throughout the paper, we neglect the integer problem since it is handled by replacing each variable s that should be an integer with $\lfloor s \rfloor = \max_{\substack{n \in \mathbb{N} \\ n \leq s}} n$.

continuation strategy $\sigma_i(x_i) \mid h_i^\tau$ is smaller than $\exp(-T^{\frac{1}{4}})$:

$$\begin{array}{c} \max_{\sigma_i^{T_P} \in \Sigma_i^{T_P}} \mathbb{E}\left[\sum_{t=1}^{T_P} \delta^{t-1} u_i(a_t) + \pi_i(x_j, h_j^{T_P+1} : \delta) \mid h_i^\tau, \sigma_i^{T_P}, \sigma_j(x_j)\right] \\ -\mathbb{E}\left[\sum_{t=1}^{T_P} \delta^{t-1} u_i(a_t) + \pi_i(x_j, h_j^{T_P+1} : \delta) \mid h_i^\tau, \sigma(x)\right] \end{array} \leq \exp(-T^{\frac{1}{4}}). \quad (21)$$

To see why this is sufficient, remember that we insert the report block where the players report the whole history h_i at the end of the finitely repeated game (see Figure 5). Based on the reported history $\hat{h}_i$, player j adjusts the reward function $\pi_i(x_j, \cdot : \delta)$ so that the prescribed action is exactly optimal. This adjustment is very small for large T since the original strategy was optimal up to the loss of $\exp(-T^{\frac{1}{4}})$ by (21).

The remaining task is to show the incentive to tell the truth about h_i. Intuitively, with some norm, player j punishes player i proportionally to $\left\|h_j - \mathbb{E}\left[h_j \mid \hat{h}_i\right]\right\|^2$. For a properly chosen norm, the first order condition to minimize the expected punishment $\mathbb{E}\left[\left\|h_j - \mathbb{E}\left[h_j \mid \hat{h}_i\right]\right\|^2 \mid h_i\right]$ is to tell the truth: $\hat{h}_i = h_i$. Since the adjustment is small, this punishment can also be small and does not affect the equilibrium payoff. Section 12 formalizes the argument.

Therefore, until Section 12, our objective is to construct $\sigma_i(x_i)$ and $\pi_i(x_j, \cdot : \delta)$ satisfying (21), (5) and (6).

8 Equilibrium Strategies

In this section, we partially define $\sigma_i(x_i)$ and $\pi_i(x_j, \cdot : \delta)$. In Section 8.1, we define the equilibrium action given $\hat{\lambda}_j(l)$. Then, in Section 8.2, we define $\hat{\lambda}_j(l)$ given player i's inference of player j's messages in the supplemental rounds. The definition of player i's inference of player j's messages in the supplemental rounds will be deferred to Section 9. In addition, Section 8.3 defines the reward function $\pi_i(x_j, \cdot : \delta)$.

8.1 Actions Given $\hat{\lambda}_j(l)$

In this subsection, we explain the strategies $\sigma_i(x_i)$ given $\hat{\lambda}_j(l) \in \{G, B\}$. The transition of $\hat{\lambda}_j(l)$ will be explained in the next subsection.

In the each lth review round, player i with $\sigma_i(x_i)$ takes $a_i(x_i)$ with

$$a_i(x_i) \equiv \begin{cases} C_i & \text{if } x_i = G, \\ D_i & \text{if } x_i = B \end{cases} \tag{22}$$

if $\hat{\lambda}_j(l) = G$ and D_i if $\hat{\lambda}_j(l) = B$. See Figure 3 to check the equilibrium action.

As player j calculates $X_j(l)$ defined in (12) to monitor player i after $x_i = G$, player i calculates

$$X_i(l) \equiv \begin{cases} \sum_{t \in T_i(l)} \Psi_{i,t}^{C_j, C_i} + \mathbf{1}_{t_i(l)} & \text{if } x_i = G, \\ \sum_{t \in T_i(l)} \Psi_{i,t}^{C_j, D_i} + \mathbf{1}_{t_i(l)} & \text{if } x_i = B \end{cases} \tag{23}$$

if $x_j = G$. Then, $\lambda_i(l+1) = B$ is the record of an erroneous history

$$X_i(l) \notin [q_2 T - 2\varepsilon T, q_2 T + 2\varepsilon T]. \tag{24}$$

That is,

$$\lambda_i(l+1) = \begin{cases} G & \text{if } l = 0 \text{ or } X_i(\tilde{l}) \in [q_2 T - 2\varepsilon T, q_2 T + 2\varepsilon T] \text{ for all } \tilde{l} \leq l, \\ B & \text{otherwise} \end{cases} \tag{25}$$

if the cheap talk message is $x_j = G$ and $\lambda_i(l+1) = G$ if $x_j = B$. Review Figure 2 for the graphical explanation (note that the roles of i and j are reversed).

Then, in the supplemental rounds 1 and 2 for $\lambda_i(l+1)$, player i sends $\lambda_i(l+1)$. How to send the message will be defined in Section 9.2. On the other hand, in the supplemental rounds 1 and 2 for $\lambda_j(l+1)$, player j sends $\lambda_j(l+1)$ symmetrically defined. While player j sends the message, player i takes C_i. Let $\lambda_j(l+1)(i)$ be player i's inference of the message $\lambda_j(l+1)$, which will be defined in Section 9.2.

8.2 Inference of $\hat{\lambda}_j(l+1)$

In this subsection, we explain the transition of player i's inference of $\lambda_j(l+1)$, $\hat{\lambda}_j(l+1) \in \{G, B\}$. Here, we take player i's inference of player j's messages in the supplemental rounds for $\lambda_j(l+1)$ (denoted by $\lambda_j(l+1)(i) \in \{G, B\}$) as given. See Section 9 for the explanation of $\lambda_j(l+1)(i)$.

8.2.1 Statistics

Since $\lambda_j(l+1)$ depends on $X_j(l) = \sum_{t \in T_j(l)} \Psi_{j,t}^{a_i(x_i), a_j(x_j)} + \mathbf{1}_{t_j(l)}$ as we have mentioned in (23), (24) and (25), we formally define $\Psi_{j,t}^{a}$ first. Since $\Psi_{j,t}^{a}$ is i.i.d. within each review round, we omit the time subscript if it is not confusing.

We want Ψ_j^a to satisfy the following two conditions: First, player j needs to statistically monitor whether player i takes a_i or not. In particular, as we have mentioned in (9), we want to establish that

$$\mathbb{E}\left[\Psi_j^a \mid \tilde{a}_i, a_j\right] = \begin{cases} q_2 & \text{for } \tilde{a}_i = a_i, \\ q_1 & \text{for all } \tilde{a}_i \neq a_i \end{cases}$$

with $q_2 > q_1$.

Second, as we will see in Section 8.2.2, player i's continuation play depends on $\mathbb{E}\left[\Psi_j^a \mid a, y_i\right]$. We want to prevent player j from deviating from a_j to $\tilde{a}_j \neq a_j$ to manipulate $\mathbb{E}\left[\Psi_j^a \mid a, y_i\right]$. It is sufficient to have

$$\mathbb{E}_{y_i}\left[\mathbb{E}_{\Psi_j^a}\left[\Psi_j^a \mid a, y_i\right] \mid a_i, \tilde{a}_j\right]$$

is constant with respect to $\tilde{a}_j$. That is, player i calculates the conditional expectation of Ψ_j^a believing that a_j is taken. The ex ante value of this conditional expectation is constant even if player j secretly deviates to $\tilde{a}_j \neq a_j$.

In summary, we want to construct Ψ_j^a with the following conditions:

Condition 1 *We want to have*

1. Ψ_j^a *monitors* a_i*:*

$$\mathbb{E}\left[\Psi_j^a \mid \tilde{a}_i, a_j\right] = \begin{cases} q_2 & \text{for } \tilde{a}_i = a_i, \\ q_1 & \text{for all } \tilde{a}_i \neq a_i. \end{cases}$$

2. Player j *cannot manipulate the ex ante value of* $\mathbb{E}_{\Psi_j^a}\left[\Psi_j^a \mid a, y_i\right]$*: For all* $\tilde{a}_j \in A_j$*,*

$$\mathbb{E}_{y_i}\left[\mathbb{E}_{\Psi_j^a}\left[\Psi_j^a \mid a, y_i\right] \mid a_i, \tilde{a}_j\right] = q_2.$$

We construct Ψ_j^a in the following two steps. First, we define $\psi_j^a : Y_j \to (0,1)$. Second, player j constructs a random variable $\Psi_j^a \in \{0,1\}$ from $\psi_j^a(y_j)$ as follows: After taking a_j and 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 $y \in Y$, Condition 1 is satisfied for Ψ_j^a if and only if $\psi_j^a(y_j)$ satisfies

1.
$$\mathbb{E}\left[\psi_j^a(y_j) \mid \tilde{a}_i, a_j\right] \equiv \sum_{y_j} q(y_j \mid \tilde{a}_i, a_j)\psi_j^a(y_j) = \begin{cases} q_2 & \text{if } \tilde{a}_i = a_i, \\ q_1 & \text{if } \tilde{a}_i \neq a_i, \end{cases} \tag{26}$$

2. $\mathbb{E}_{y_i}\left[\mathbb{E}_{y_j}\left[\psi_j^a(y_j) \mid a, y_i\right] \mid a_i, \tilde{a}_j\right]$ is constant with respect to $\tilde{a}_j$. That is, for all $\tilde{a}_j \in A_j$,

$$\sum_{y_i}\left\{\sum_{y_j} \psi_j^a(y_j) q(y_j \mid a, y_i)\right\} q\left(y_i \mid a_i, \tilde{a}_j\right) = q_2. \tag{27}$$

Formally, we show the following lemma:

Lemma 3 *Generically, the following statement is true: 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 (26) and (27) are satisfied.*

8.2.2 Inference

Now we are ready to explain the transition of $\hat{\lambda}_j(l+1) \in \{G, B\}$. If $x_i = B$, then $\lambda_j(l+1) = G$ is common knowledge and so define $\hat{\lambda}_j(l+1) = G$.

We are left to consider the case with $x_i = G$. Since $\lambda_j(1) = G$ is common knowledge and so define $\hat{\lambda}_j(1) = G$. Further, since $\lambda_j(l+1) = B$ once $\lambda_j(\tilde{l}) = B$ has happened for some $\tilde{l} \le l$ from (25), define $\hat{\lambda}_j(l+1) = B$ once $\hat{\lambda}_j(\tilde{l}) = B$ has happened for some $\tilde{l} \le l$. Hence, we are left to specify, conditional on $x_i = G$ and $\hat{\lambda}_j(\tilde{l}) = G$ for all $\tilde{l} \le l$, how $\hat{\lambda}_j(l+1) \in \{G, B\}$ is determined.

Suppose $\hat{\lambda}_j(l) = G$ is a correct inference: $\lambda_j(l) = G$. Then, $\lambda_j(l+1)$ is determined as

$$\lambda_j(l+1) = \begin{cases} G & \text{if } X_j(l) \in [q_2 T - 2\varepsilon T, q_2 T + 2\varepsilon T] \\ B & \text{if } X_j(l) \notin [q_2 T - 2\varepsilon T, q_2 T + 2\varepsilon T] \end{cases} \tag{28}$$

with $X_j(l) = \sum_{t \in T_j(l)} \Psi_{j,t}^{a_i(x_i), a_j(x_j)} + \mathbf{1}_{t_j(l)}$. Therefore, it is natural to consider the conditional expectation of $X_j(l)$:

$$\mathbb{E}\left[X_j(l) \mid \{a_t, y_{i,t}\}_{t \in T(l)}\right]$$

Instead of using this, we consider

$$\sum_{t \in T_i(l)} \mathbb{E}\left[\Psi_{j,t}^{a(x)} \mid a(x), y_{i,t}\right] + q_2$$

with $a(x) = (a_i(x_i), a_j(x_j))$. Two reminders: First, player i calculates the expectation of the summation of $\Psi_{j,t}^{a(x)}$ over $T_i(l)$ (the set of periods when player i uses to monitor player j), not $T_j(l)$ (the set of periods when player j uses to monitor player i).[9] As we will see, since $T_i(l)$ and $T_j(l)$ are different at most for two periods, this difference is negligible for almost optimality (21). Second, player i conditions on $a(x)$ being taken. Player i with $\hat{\lambda}_j(l) = G$ takes $a_i(x_i) = C_i$ in the lth review round. In addition, as we have explained in Section 5.2.2, player j makes player i indifferent between any action profile for the rest of the finitely

[9] The term q_2 reflects the fact that the expected value of $\mathbf{1}_{t_j(l)}$ is q_2.

repeated game if $\hat{\lambda}_i(l) = B$. Therefore, player i always conditions $\hat{\lambda}_i(l) = G$ and player j takes $a_j(x_j)$ in the lth review round to calculate the conditional expectation.

Further, instead of using $\mathbb{E}\left[\Psi_{j,t}^{a(x)} \mid a(x), y_{i,t}\right]$ directly, player i constructs $(E_i\Psi_j^{a(x)})_t \in \{0,1\}$ as follows: After taking $a_i(x_i)$ and observing $y_{i,t}$, player i calculates $\mathbb{E}\left[\Psi_{j,t}^{a(x)} \mid a(x), y_{i,t}\right]$. After that, player i draws a random variable from the uniform distribution on $[0,1]$. If the realization of this random variable is less than $\mathbb{E}\left[\Psi_{j,t}^{a(x)} \mid a(x), y_{i,t}\right]$, then $(E_i\Psi_j^{a(x)})_t = 1$ and otherwise, $(E_i\Psi_j^{a(x)})_t = 0$. Let

$$E_iX_j(l) = \sum_{t\in T_i(l)} (E_i\Psi_j^{a(x)})_t + q_2.$$

Since

$$\Pr\left(\left\{(E_i\Psi_j^{a(x)})_t = 1\right\} \mid a_t, y_t\right) = \mathbb{E}\left[\Psi_{j,t}^{a(x)} \mid a(x), y_{i,t}\right]$$

for all a_t and y_t, *conditional on* $\{a_t, y_t\}_{t\in T(l)}$, the probability that

$$\left|\sum_{t\in T_i(l)} \mathbb{E}\left[\Psi_{j,t}^{a(x)} \mid a(x), y_{i,t}\right] + q_2 - E_iX_j(l)\right| \leq \frac{1}{4}\varepsilon T \tag{29}$$

is of order $\exp(-T)$ by the central limit theorem.

There are following cases:

1. (29) is not satisfied. Let $\zeta_i(l) = B$ denote this event. Player i will use the reward $\pi_j^{x_i}(a_{i,t}, y_{i,t})$ defined in Lemma 2 in the subsequent rounds and so player j is indifferent between any action profile. Hence, this case is excluded from player j's consideration. The inference of $\lambda_j(l+1)$ by player i will be equal to the inference of the messages about $\lambda_j(l+1)$ in the supplemental rounds 1 and 2 for $\lambda_j(l+1)$: $\hat{\lambda}_j(l+1) = \lambda_j(l+1)(i)$.

2. (29) is satisfied. Let $\zeta_i(l) = G$ denote this event. Consider the following subcases:

   (a) We have

   $$E_iX_j(l) \notin [q_2T - \frac{1}{2}\varepsilon T, q_2T + \frac{1}{2}\varepsilon T].$$

Let $\theta_i(l) = B$ denote this event. Player i will use the reward $\pi_j^{x_i}(a_{i,t}, y_{i,t})$ in the subsequent rounds. However, compared to Case 1, player j does not exclude this case from her consideration. Again, player i uses the inference of the messages in the supplemental rounds: $\hat{\lambda}_j(l+1) = \lambda_j(l+1)(i)$.

(b) We have

$$E_i X_j(l) \in [q_2 T - \frac{1}{2}\varepsilon T, q_2 T + \frac{1}{2}\varepsilon T]. \tag{30}$$

Player i randomly picks the following two procedures:

i. With small probability $\eta > 0$, player i will use the reward $\pi_j^{x_i}(a_{i,t}, y_{i,t})$ in the subsequent rounds. Again, player i uses the inference of the messages in the supplemental rounds: $\hat{\lambda}_j(l+1) = \lambda_j(l+1)(i)$. Let $\theta_i(l) = B$ denote this event (hence, $\theta_i(l) = B$ implies that either 2-(a) or 2-(b)-i happens).

ii. With large probability $1 - \eta$, player i believes $\hat{\lambda}_j(l+1) = G$ regardless of the history in the supplemental rounds 1 and 2 for $\lambda_j(l+1)$. Let $\theta_i(l) = G$ denote this event.

Then, we define $\theta_i(l) \in \{G, B\}$ after $\zeta_i(l) = B$ in the same way as after $\zeta_i(l) = G$: If (30) is not satisfied, then we have $\theta_i(l) = B$. If (30) is satisfied, then with probability η, we have $\theta_i(l) = B$, and with probability $1 - \eta$, we have $\theta_i(l) = G$.

For concreteness, we define that player i who has deviated before the supplemental round 1 for $\lambda_j(l+1)$ uses $\hat{\lambda}_j(l+1) = \lambda_j(l+1)(i)$.

If 2-(b)-ii is the case, then since $T_i(l)$ and $T_j(l)$ are different only for two periods, after knowing $T_i(l)$ and $T_j(l)$, (29) and (30) imply

$$\mathbb{E}\left[X_j(l) \mid a(x), \{y_{i,t}\}_{t\in T(l)}\right] \in [q_2 T - \varepsilon T, q_2 T + \varepsilon T]. \tag{31}$$

Conditional on $a(x)$ and $\{y_{i,t}\}_{t\in T(l)}$, conditional distribution of $X_j(l)$ is distributed approximately according to the normal distribution with mean $\mathbb{E}\left[X_j(l) \mid a(x), \{y_{i,t}\}_{t\in T(l)}\right]$ and the standard deviation $O(T^{\frac{1}{2}})$ by the central limit theorem. Since (31) implies that the conditional mean is inside of $[q_2T - 2\varepsilon T, q_2T + 2\varepsilon T]$ by $\varepsilon T^{\frac{1}{2}}$ times $T^{\frac{1}{2}}$, the order of the standard deviation of the conditional distribution $X_j(l) \mid a(x), \{y_{i,t}\}_{t \in T(l)}$, player i believes that $X_j(l) \in [q_2T - 2\varepsilon T, q_2T + 2\varepsilon T]$ with probability of order $1 - \exp(-T)$. Hence, player i believes $\lambda_j(l+1) = G$ with high probability and so $\hat{\lambda}_j(l+1) = G$.

If 1, 2-(a) or 2-(b)-i is the case, we say "player i is ready to listen (to player j's message in the supplemental rounds)."

See the following figure for all the classification:

[Insert Figure 6]

Boxes 3, 5, 7 and 10 respectively correspond to 1, 2-(b)-ii, 2-(a) and 2-(b)-i.

Note that as we have mentioned in Figure 4 (the roles of i and j are reversed), when player i has $\hat{\lambda}_j(l+1) = B$, one of Boxes 3, 7 and 10 must to be the case and player j will be indifferent between any action profile.

In addition, when player i uses the messages in the supplemental rounds to infer $\hat{\lambda}_j(l+1)$, player i has made player j indifferent between any action profile. Therefore, we do not need to consider the truthtelling incentive for player j in the supplemental rounds 1 and 2 for $\lambda_j(l+1)$.

Further, as we mentioned in (29), *conditional on* $\{a_t, y_t\}_{t \in T(l)}$, (29) is not satisfied with probability of order $\exp(-T)$ by the central limit theorem. And 2 of Lemma 3 implies that the distribution of $E_iX_j(l)$ is independent of player j's action. Therefore, player j cannot manipulate $\theta_i(l) \in \{G, B\}$ by changing her own action in the lth review round.

Therefore, we have proven the following lemma:

Lemma 4 *For sufficiently large T, for all $x \in \{G, B\}^2$, if $\lambda_j(l) = \hat{\lambda}_i(l) = G$, then:*

1. *Conditional on any $\{a_t, y_t\}_{t \in T(l)}$, $\zeta_i(l) = G$ with probability no less than $1 - \exp(-T^{\frac{4}{5}})$.*

2. *The distribution of $\theta_i(l) \in \{G, B\}$ is independent of player j's action.*

3. *If 2-(b)-ii is the case in the above explanation and* $\lambda_j(l) = \hat{\lambda}_j(l) = G$*, then player i's conditional belief on* $\lambda_j(l+1) = G$ *at the end of the* l*th review round is no less than* $1 - \exp(-T^{\frac{4}{5}})$.

4. *Whenever player i uses player i's inference of player j's messages in the supplemental rounds for* $\lambda_j(l+1)$*, player i has made player j indifferent between any action profile in the rounds after the* l*th review round.*

8.3 Reward Function

In this subsection, we explain player j's reward function on player i, $\pi_i(x_j, \cdot : \delta)$. The following notation is useful: Let $r \in \{1, ..., L\} \cup \{(l, \lambda_1, 1), (l, \lambda_1, 2)\}_l \cup \{(l, \lambda_2, 1) \cup (l, \lambda_2, 2)\}_l$ be the generic index for the rounds. From Figure 5, there is the chronological order of the rounds. Hence, with abuse of notation, we identify round $r+1$ as the round coming right after r. For example, with $r = l$ (lth review round), $r+1$ is the supplemental round 1 for $\lambda_1(l+1)$. In addition, let $r < l$ if and only if the round r is before the lth review round and $r \leq l$ if and only if $r < l$ or $r = l$.

The total reward is the summation of the following rewards in the review rounds and those in the supplemental rounds:

The lth review round If $\zeta_j(r) = B$, $\vartheta_j(r) = B$ or $\theta_j(r) = B$ happens for some $r < l$, then player j makes player i indifferent between any action profile sequence by

$$\pi_i(x_j, h_j^{T_P+1}, l) = \begin{cases} \sum_{t \in T(l)} \delta^{t-1} \pi_i^G(a_{j,t}, y_{j,t}) & \text{if } x_j = G, \\ \sum_{t \in T(l)} \delta^{t-1} \pi_i^B(a_{j,t}, y_{j,t}) & \text{if } x_j = B. \end{cases} \tag{32}$$

If $\zeta_j(r) = \vartheta_j(r) = \theta_j(r) = G$ for all $r < l$, then player j's reward based on x and $\lambda_j(l)$. Remember that the basic structure explained in Section 5.2.2 is summarized in the following figure:

[Insert Figure 7].

The formal description is given by

$$\pi_i(x_j, h_j^{T_P+1}, l) = \begin{cases} \bar{\pi}_i(x,l,G) - \rho T & \text{if } x_j = G \text{ and } x_i = B, \\ \bar{\pi}_i(x,l,G) + \rho T & \text{if } x_j = B \text{ and } x_i = B, \\ \bar{\pi}_i(x,l,G) + \bar{L}\{X_j(l) - (q_2 T + 2\varepsilon T)\} - \rho T & \text{if } x_j = G,\, x_i = G \text{ and } \lambda_j(l) = G, \\ \bar{\pi}_i(x,l,B) - \rho T & \text{if } x_j = G,\, x_i = G \text{ and } \lambda_j(l) = B, \\ \bar{\pi}_i(x,l,G) + \bar{L}\{X_j(l) - (q_2 T - 2\varepsilon T)\} - \rho T & \text{if } x_j = B,\, x_i = G \text{ and } \lambda_j(l) = G, \\ \bar{\pi}_i(x,l,B) + \rho T & \text{if } x_j = B,\, x_i = G \text{ and } \lambda_j(l) = B. \end{cases} \tag{33}$$

Here, $\bar{\pi}_i(x, l, \lambda_j(l))$ will be determined in Section 11 so that (21), (5) and (6) are satisfied.

The supplemental rounds 1 and 2 for $\lambda_1(l+1)$ and $\lambda_2(l+1)$ Player j makes player i indifferent between any action profile sequence by

$$\sum_t \delta^{t-1}\pi_i^{x_j}(a_{j,t}, y_{j,t}). \tag{34}$$

On the top of that, in the supplemental rounds 1 and 2 for $\lambda_j(l+1)$, where player i is the receiver of the message, player j adds

$$\begin{cases} -\bar{L}\sum_t (1 - \Psi_{j,t}^{C_i, a_{j,t}}) \leq 0 & \text{if } x_j = G, \\ \bar{L}\sum_t \Psi_{j,t}^{C_i, a_{j,t}} \geq 0 & \text{if } x_j = B \end{cases} \tag{35}$$

to make it strictly optimal to take C_i. Note that (35) is linearly increasing in $\Psi_{j,t}^{C_i, a_{j,t}}$.

The summary of rewards after $\zeta_j = B$, $\vartheta_j = B$ or $\theta_j = B$ In the lth review round, if $\zeta_j(r) = B$, $\vartheta_j(r) = B$ or $\theta_j = B$ with $r < l$ has happened, (32) implies that any action profile gives the same payoff. On the other hand, in the supplemental rounds where player i receives the message, (35) is valid even after $\zeta_j = B$, $\vartheta_j = B$, or $\theta_j = B$. Since (34) cancels out the difference in the instantaneous utilities w.r.t. an action profile, taking C_i is strictly optimal by (35) and gives a constant expected payoff.

With abuse of notation, we say "player i is indifferent between any action profile" if $\zeta_j = B$, $\vartheta_j = B$, or $\theta_j = B$ has happened, although player i strictly prefers C_i while she receives the message.

For concreteness, we define that, after player j deviates in the round r, player j has $\zeta_j(r) = B$ and makes any action optimal for player i.

8.3.1 Set of Almost Optimal Action

Given the above reward function, let $A_i(l)$ be the set of player i's optimal action in the lth review round. As we will show in Section 11, $A_i(l)$ is as follows: If $\zeta_j(r) = B$, $\vartheta_j(r) = B$ or $\theta_j(r) = B$ happens for some $r < l$ (and so (32) is being used), then $A_i(l) = A_i$. Otherwise (and so (33) is being used), $A_i(l)$ depends on $x_i \in \{G, B\}$ and $\lambda_j(l) \in \{G, B\}$. If $x_i = \lambda_j(l) = G$, since the reward is increasing in $X_j(l)$, $A_i(l) = \{C_i\}$. If $x_j = B$ or $\lambda_j(l) = B$, then since the reward is constant, $A_i(l) = \{D_i\}$.

9 Inference of the Messages

In this section, we define how player i infers player j's messages in the supplemental rounds for $\lambda_j(l+1)$. That is, how $\lambda_j(l+1)(i) \in \{G, B\}$ is determined.

9.1 Conditions on the Message Exchange

We derive conditions on $\lambda_j(l+1)(i)$ that makes the inferences explained in Figure 6 almost optimal. See Condition 2 below for the summary.

First, if (29) and (30) are satisfied, player i needs to believe $\lambda_j(l+1) = G$ is true with high probability regardless of player i's continuation history. From 3 of Lemma 4, player i before seeing player j's action in the continuation play believes that the probability of a mistake is no more than

$$\exp(-T^{\frac{4}{5}}). \tag{36}$$

As we will see in Sections 9.2.1 and 9.2.2, player j's strategy in the supplemental rounds is determined by $\lambda_j(l+1)$. In addition, as we have seen in Section 8.1, player j's strategy in the review rounds is determined by $\hat{\lambda}_i(l+1)$. Since the length of the supplemental rounds is $4LT^{\frac{1}{2}}$, this only updates player i's belief by of order $\exp(T^{\frac{1}{2}})$. Since (36) is much smaller than $1/\exp(T^{\frac{1}{2}})$, player i can keep high belief on $\lambda_j(l+1) = G$ regardless of player i's history in the supplemental rounds.

In addition, player j's actions in the review rounds indirectly reveal $\lambda_j(l+1)$ through $\hat{\lambda}_i(l+1)$. Suppose player i knows $\hat{\lambda}_i(l+1)$. See Figure 8, which summarizes the important features of $\hat{\lambda}_i$ and $\hat{\lambda}_j$ explained in Figure 6. Here, $a_i(l+1)$ is player i's equilibrium action in the $(l+1)$th review round prescribed in Section 8.1. Boxes 1 to 4 are about how player i infers $\lambda_j(l+1)$. Remember that player j infers $\lambda_i(l+1)$ symmetrically, which is expressed in Boxes 5 to 11. Whether player j is in Box 5 or 6, player j uses $\lambda_i(l+1)(j)$ to determine $\hat{\lambda}_i(l+1)$ with probability at least η. Again, since the length of the supplemental rounds is $4LT^{\frac{1}{2}}$, any $\hat{\lambda}_i(l+1)$ can occur with probability of order $\exp(-T^{\frac{1}{2}})$. Therefore, player i believes that any $\hat{\lambda}_i(l+1)$ occurs with probability at least of order $\eta \exp(-T^{\frac{1}{2}})$.[10] Since (36) is much smaller than $\eta \exp(-T^{\frac{1}{2}})$, player i can keep high belief on $\lambda_j(l+1) = G$ regardless of player i's history in the review rounds.

[Insert Figure 8]

In summary, if (29) and (30) are satisfied, player i believes that $\lambda_j(l+1) = G$ is true with high probability regardless of player i's continuation history, as desired.

Second, when player i uses $\lambda_j(l+1)(i)$ to determine $\hat{\lambda}_j(l+1)$, there are two cases: $\lambda_j(l+1)(i) = \lambda_j(l+1)$ and $\lambda_j(l+1)(i) \neq \lambda_j(l+1)$. If the former is the case, then the inference is optimal. If the latter is the case, for almost optimality, it suffices to show that, whenever $\lambda_j(l+1)(i) \neq \lambda_j(l+1)$, conditional on $\lambda_j(l+1)$ (that is, even after player i knows that her inference was not correct), player i believes that player j makes player i indifferent between any action with high probability: $A_i(l+1) = A_i$.

[10] As we will see in Lemma 5, excluding Box 12 does not affect the posterior so much.

Therefore, we want to have the following:

Condition 2 *For the supplemental rounds, we want to have the following: Conditional on* $\lambda_j(l+1)$ *and* $\lambda_j(l+1)(i)$*, if* $\lambda_j(l+1)(i) \neq \lambda_j(l+1)$*, then player* i *believes that* $A_i(l+1) = A_i$ *with high probability.*

9.2 Message Protocol

In this section, we explain how we establish Condition 2 in the supplemental rounds 1 and 2 for $\lambda_j(l+1)$, where player j sends $\lambda_j(l+1) \in \{G, B\}$ to player i.

The following notations are useful. Let $a_j^G = C_j$ and $a_j^B = D_j$.[11] Let $\mathbf{q}_j(a) = (q_j(y_j \mid a))_{y_j}$ be a vector of player j's signal distributions under a. In addition, let $\mathbf{1}_{y_{j,t}}$ be a $|Y_j| \times 1$ vector such that the element corresponding to $y_{j,t}$ is 1 and the other elements are 0. $\mathbf{q}_i(a) = (q_i(y_i \mid a))_{y_i}$ and $\mathbf{1}_{y_{i,t}}$ are symmetrically defined. $\varepsilon > 0$ is a small number that will be specified in Section 10.

9.2.1 Supplemental Round 1 for $\lambda_j(l+1)$

First, we intuitively explain the supplemental round 1 for $\lambda_j(l+1)$, which is summarized in Figure 9 below. Let $\bar{\lambda}_j \in \{G, B\}$ be a true message, that is, $\lambda_j(l+1) = \bar{\lambda}_j$. Player j (sender) takes $a_j^{\bar{\lambda}_j} \in \{a_j^G, a_j^B\}$ for $T^{\frac{1}{2}}$ periods while player i (receiver) takes a_i^G for $T^{\frac{1}{2}}$ periods.

Suppose player j requires player i to infer the message in the following way. From the set of periods in the current round $T(l, \lambda_j, 1)$ (see Figure 5), player j randomly picks $t_j(l, \lambda_j, 1)$. Player j calculates the frequency of player j's signal observation dropping period $t_j(l, \lambda_j, 1)$. With $T_j(l, \lambda_j, 1) \equiv T(l, \lambda_j, 1) \setminus \{t_j(l, \lambda_j, 1)\}$, let $\mathbf{y}_j(l, \lambda_j, 1) \equiv \frac{1}{T^{\frac{1}{2}}-1} \sum_{t \in T_j(l,\lambda_j,1)} \mathbf{1}_{y_{j,t}}$. (i) If the distance to $\mathbf{y}_j(l, \lambda_j, 1)$ from the affine hull of $\{\mathbf{q}_j(a_j^{\bar{\lambda}_j}, a_i)\}_{a_i}$ (aff $\left(\{\mathbf{q}_j(a_j^{\bar{\lambda}_j}, a_i)\}_{a_i}\right)$) is no more than ε, then player j requires player i to infer $\lambda_j(l+1)$ properly.[12] (ii) Otherwise, player j will make any action optimal to player i (let $\theta_j(l, \lambda_j, 1) = B$ denote this case. Then, from

[11] In a general game, as we will see in the Supplemental Materials, any a_j^G and a_j^B with $a_j^G \neq a_j^B$ generically works.

[12] As we will see in Section 9.2.2, there are subcases of (i) where player j makes player i indifferent between any action profile, depending on the realization of the supplemental round 2 for $\lambda_j(l+1)$.

(32), $A_i(\tilde{l}+1) = A_i$ for all $\tilde{l} \geq l$). Since we take the affine hull with respect to player i's action, whether (i) or (ii) is the case is out of player i's control. By the central limit theorem, (ii) occurs with small probability and does not affect the equilibrium payoff.

Player i (receiver) calculates the conditional expectation of $\mathbf{y}_j(l, \lambda_j, 1)$. For Condition 2, it suffices that player i infers (iii) $\lambda_j(l+1)(i) = G$ if the distance to $\mathbb{E}\left[\mathbf{y}_j(l, \lambda_j, 1) \mid \{y_{i,t}\}_{t \in T(l,\lambda_j,1)}, a_j^G, a_i^G\right]$ from aff $\left(\{\mathbf{q}_j(a_j^G, a_i)\}_{a_i}\right)$ is no more than 2ε and (iv) $\lambda_j(l+1)(i) = B$ if the distance to $\mathbb{E}\left[\mathbf{y}_j(l, \lambda_j, 1) \mid \{y_{i,t}\}_{t \in T(l,\lambda_j,1)}, a_j^B, a_i^G\right]$ from aff $\left(\{\mathbf{q}_j(a_j^B, a_i)\}_{a_i}\right)$ is no more than 2ε. To see why this is sufficient, suppose $\lambda_j(l+1) = G$ but player i infers $\lambda_j(l+1)(i) = B$. Moreover, assume that player i knows that $\lambda_j(l+1) = G$ and player j took a_j^G. Conditional on $\lambda_j(l+1) = G$, the distance to $\mathbb{E}\left[\mathbf{y}_j(l, \lambda_j, 1) \mid \{y_{i,t}\}_{t \in T(l,\lambda_j,1)}, a_j^G, a_i^G\right]$ from aff $\left(\{\mathbf{q}_j(a_j^G, a_i)\}_{a_i}\right)$ is more than 2ε (otherwise, (iii) should be the case and $\lambda_j(l+1)(i) = G$). Notice that the conditional expectation uses the true action a_j^G. Since the conditional variance of $\mathbf{y}_j(l, \lambda_j, 1)$ is of order $T^{\frac{1}{2}-1}$ by the central limit theorem, player i's belief on the event that the distance to $\mathbf{y}_j(l, \lambda_j, 1)$ from aff $\left(\{\mathbf{q}_j(a_j^G, a_i)\}_{a_i}\right)$ is no more than ε is at most of order $\exp(T^{\frac{1}{2}-1})$. Hence, player i believes that $A_i(\tilde{l}+1) = A_i$ for all $\tilde{l} \geq l$ with high probability. The symmetric argument holds for $\lambda_j(l+1) = B$.

For each $\hat{\lambda}_j \in \{G, B\}$, instead of using $\mathbb{E}\left[\mathbf{y}_j(l, \lambda_j, 1) \mid \{y_{i,t}\}_{t \in T(l,\lambda_j,1)}, a_j^{\hat{\lambda}_j}, a_i^G\right]$, player i calculates the following statistics: Player i randomly picks $t_i(l, \lambda_j, 1)$. With $T_i(l, \lambda_j, 1) \equiv T(l, \lambda_j, 1) \setminus \{t_i(l, \lambda_j, 1)\}$, player i calculates $\frac{1}{T^{\frac{1}{2}}-1} \sum_{t \in T_i(l,\lambda_j,1)} \mathbb{E}\left[\mathbf{1}_{y_{j,t}} \mid y_{i,t}, a_j^{\hat{\lambda}_j}, a_i^G\right]$ for each $\hat{\lambda}_j$. Here, player i drops her own $t_i(l, \lambda_j, 1)$, not $t_j(l, \lambda_j, 1)$ that player j drops. However, since $T_j(l, \lambda_j, 1)$ and $T_i(l, \lambda_j, 1)$ differs at most for two periods, the same argument as above asymptotically holds if we replace $\mathbb{E}\left[\mathbf{y}_j(l, \lambda_j, 1) \mid \{y_{i,t}\}_{t \in T(l,\lambda_j,1)}, a_j^{\hat{\lambda}_j}, a_i^G\right]$ with

$$
\begin{aligned}
& \frac{1}{T^{\frac{1}{2}}-1} \sum_{t \in T_i(l,\lambda_j,1)} \mathbb{E}\left[\mathbf{1}_{y_{j,t}} \mid y_{i,t}, a_j^{\hat{\lambda}_j}, a_i^G\right] \\
= \ & \mathbb{E}\left[\frac{1}{T^{\frac{1}{2}}-1} \sum_{t \in T_i(l,\lambda_j,1)} \mathbf{1}_{y_{j,t}} \mid \mathbf{y}_i(l, \lambda_j, 1), a_j^{\hat{\lambda}_j}, a_i^G\right].
\end{aligned}
$$

Here, $\mathbf{y}_i(l, \lambda_j, 1) \equiv \frac{1}{T^{\frac{1}{2}}-1} \sum_{t \in T_i(l,\lambda_j,1)} \mathbf{1}_{y_{i,t}}$.

The above inference is well defined if there is $\bar{\varepsilon} > 0$ such that, for all $\varepsilon < \bar{\varepsilon}$, there is no player i's signal observation $\{y_{i,t}\}_{t \in T_i(l,\lambda_j,1)}$ such that

$$\begin{aligned}
&\text{The distance to } \mathbb{E}\left[\frac{1}{T^{\frac{1}{2}}-1}\sum_{t\in T_i(l,\lambda_j,1)} \mathbf{1}_{y_{j,t}} \mid \mathbf{y}_i(l,\lambda_j,1), a_j^G, a_i^G\right] \\
&\text{from aff}\left(\{\mathbf{q}_j(a_j^G, a_i)\}_{a_i}\right) \text{ is no more than } 2\varepsilon, \qquad (37)\\
&\text{The distance to } \mathbb{E}\left[\frac{1}{T^{\frac{1}{2}}-1}\sum_{t\in T_i(l,\lambda_j,1)} \mathbf{1}_{y_{j,t}} \mid \mathbf{y}_i(l,\lambda_j,1), a_j^B, a_i^G\right] \\
&\text{from aff}\left(\{\mathbf{q}_j(a_j^B, a_i)\}_{a_i}\right) \text{ is no more than } 2\varepsilon. \qquad (38)
\end{aligned}$$

Since $\mathbf{y}_i(l,\lambda_j,1) \in \Delta\left(\{\mathbf{1}_{y_i}\}_{y_i \in Y_i}\right)$,[13] by continuity, a sufficient condition for the existence of such $\bar{\varepsilon}$ is that there is no $\mathbf{y}_i \in \Delta\left(\{\mathbf{1}_{y_i}\}_{y_i \in Y_i}\right)$ such that

$$\mathbb{E}\left[\mathbf{y}_j \mid \mathbf{y}_i, a_j^G, a_i^G\right] \in \text{aff}\left(\{\mathbf{q}_j(a_j^G, a_i)\}_{a_i}\right), \qquad (39)$$

$$\mathbb{E}\left[\mathbf{y}_j \mid \mathbf{y}_i, a_j^B, a_i^G\right] \in \text{aff}\left(\{\mathbf{q}_j(a_j^B, a_i)\}_{a_i}\right). \qquad (40)$$

Here, $\mathbb{E}[\mathbf{y}_j \mid \mathbf{y}_i, a]$ is the conditional expectation of the frequency of player j's signal observations given the frequency of player i's signal observation for given periods where the players take a. (39) imposes $|Y_j| - |A_i|$ conditions and (40) imposes $|Y_j| - |A_i|$ conditions. Hence, there are $2(|Y_j| - |A_i|)$ conditions in total. On the other hand, we have $|Y_i| - 1$ degrees of freedom for $\mathbf{y}_i$. Hence, if $2(|Y_j| - |A_i|) \leq |Y_j| - 1$, then there can be solutions for (39) and (40) and the inference above may not be well defined.

To deal with this problem, we consider the following procedure. (v) If player i observes $\mathbf{y}_i(l,\lambda_j,1)$ whose difference from the affine hull of $\{\mathbf{q}_i(a_j, a_i^G)\}_{a_j}$ is no more than ε, then player i infers $\lambda_j(l+1)(i)$ in the supplemental round 1 for $\lambda_j(l+1)$. (vi) Otherwise, player i infers $\lambda_j(l+1)(i)$ in the supplemental round 2 for $\lambda_j(l+1)$, which will be explained later. Since we take the affine hull with respect to player j's action, whether player i infers $\lambda_j(l+1)$ from the supplemental round 1 or 2 is out of player j's control (remember that the message

[13] Here, $\Delta\left(\{\mathbf{1}_{y_i}\}_{y_i \in Y_i}\right)$ is the set of distributions over Y_i.

receiver i takes a_i^G). By the central limit theorem, (vi) occurs with small probability and does not affect the equilibrium payoff.

If (v) is the case, then the inferences (iii) and (iv) are generically well defined. To see why, notice that this is well defined if there exists $\varepsilon > 0$ such that there is no $\mathbf{y}_i(l, \lambda_j, 1)$ such that the distance to $\mathbf{y}_i(l, \lambda_j, 1)$ from aff $\left(\{\mathbf{q}_i(a_j, a_i^G)\}_{a_j}\right)$ is no more than ε and (37) and (38) are satisfied. A sufficient condition is that there is no $\mathbf{y}_i \in \Delta\left(\{\mathbf{1}_{y_i}\}_{y_i \in Y_i}\right)$ such that

$$\mathbf{y}_i \in \text{aff}\left(\{\mathbf{q}_i(a_j, a_i^G)\}_{a_j}\right) \tag{41}$$

and such that (39) and (40) are satisfied. Since (41) imposes $|Y_i| - |A_j|$ constraints, there are $|Y_i| - |A_i| + 2(|Y_j| - |A_i|)$ constraints together with (39) and (40). Since the degree of freedom for $\mathbf{y}_i$ is $|Y_i| - 1$, if

$$|Y_i| - |A_j| + 2(|Y_j| - |A_i|) \geq |Y_i|$$

$\Leftrightarrow$

$$|Y_j| \geq |A_i| - \frac{1}{2}|A_i|,$$

then there is generically no $\mathbf{y}_i \in \Delta\left(\{\mathbf{1}_{y_i}\}_{y_i \in Y_i}\right)$ satisfying (41), (39) and (40), as desired.

If (vi) is the case, then player i infers $\lambda_j(l+1)(i)$ from the supplemental round 2. In this case, player i has $\zeta_i(l, \lambda_j, 1) = B$ or $\vartheta_i(l, \lambda_j, 1) = B$.[14] Then, from (32), player j is indifferent between any action profile. That is, player j does not care about how player i will play in the subsequent rounds. Therefore, player j excludes this case from the consideration. In particular, player j does not care about player i's inference in this case. That is, whenever player j's message in the supplemental round 2 matters, player j puts 0 belief on this event and player j is indifferent between $\lambda_j(l+1)(i) = G$ and $\lambda_j(l+1)(i) = B$.[15]

[14] See Section 15.4 for details.

[15] As we have seen in Figure 6, whenever player i uses $\lambda_j(l+1)(i)$ to determine $\hat{\lambda}_j(l+1)$, player j is indifferent between any action profile. Hence, player j is almost indifferent to any action profile in the supplemental round 1. However, player j's history (and so actions) in the supplemental round 1 affects player j's belief about player i's inference $\lambda_j(l+1)(i)$, which, in turn, affects player j's belief about the optimality of $\hat{\lambda}_i(l+1)$ since player i's continuation strategy is jointly determined by $\lambda_j(l+1)(i)$ and player

The summary of the cases is depicted in the following figure:

[Insert Figure 9].

9.2.2 Supplemental Round 2 for $\lambda_j(l+1)$

Second, we intuitively explain the supplemental round 2 for $\lambda_j(l+1)$, which is summarized in Figures 10 and 11 below. Again, this round matters only if player i is in Box 8 of Figure 9 as a result of the supplemental round 1 for $\lambda_j(l+1)$.

In this round, player j (sender) sends the message about $\lambda_j(l+1)$ again. Remember that $\bar{\lambda}_j$ is the true variable: $\lambda_j(l+1) = \bar{\lambda}_j$. This time, player j sends the message as follows: With $\eta > 0$ being a small number defined in Section 10, player j determines

$$z_j(\bar{\lambda}_j) = \begin{cases} \bar{\lambda}_j & \text{with probability } 1-2\eta, \\ \{G,B\} \setminus \{\bar{\lambda}_j\} & \text{with probability } \eta, \\ M & \text{with probability } \eta \end{cases}$$

and player j takes

$$\alpha_j^{z_j(\bar{\lambda}_j)} = \begin{cases} a_j^G & \text{if } z_j(\bar{\lambda}_j) = G, \\ a_j^B & \text{if } z_j(\bar{\lambda}_j) = B, \\ \frac{1}{2}a_j^G + \frac{1}{2}a_j^B & \text{if } z_j(\bar{\lambda}_j) = M \end{cases}$$

for $T^{\frac{1}{2}}$ periods. That is, player j sends the "true" message $\alpha_j^{z_j(\bar{\lambda}_j)} = a_j^{\bar{\lambda}_j}$ with high probability $1-2\eta$ while player j "tells a lie" with probability 2η. With probability η, player j sends the opposite message $z_j(\bar{\lambda}_j) = \{G,B\} \setminus \{\bar{\lambda}_j\}$. With probability η, player j "mixes" two messages: $z_j(\bar{\lambda}_j) = M$ and $\alpha_j^M = \frac{1}{2}a_j^G + \frac{1}{2}a_j^B$. When player j tells a lie, player j makes player

i's history in the lth review round on which $\lambda_i(l+1)$ is based, as we have seen in Figure 8 (with the roles of players i and j reversed). Hence, player j is almost indifferent, but not exactly indifferent in the supplemental round 1 for $\lambda_j(l+1)$.

On the other hand, in the supplemental round 2, since player j excludes the cases where player i's history in the supplemental round 2 affects player i's continuation strategy, player j is exactly indifferent to any strategy.

See Proposition 1 for the formal statement of the above discussion.

i indifferent between any action: $\theta_j(l, \lambda_j, 2) = B$ and $A_i(\tilde{l}+1) = A_i$ for all $\tilde{l} \geq l$. Since a lie occurs with small probability 2η, it does not affect the equilibrium payoff.

Player i (receiver) takes a_i^G. Player i infers $z_j(\bar{\lambda}_j)$ as follows. For the periods in the current round $T(l, \lambda_j, 2)$ (see Figure 5), let $\mathbf{y}_i(l, \lambda_j, 2) = (y_i(l, \lambda_j, 2))_{y_i}$ be the vector representing the frequency that player i observes y_i in $T(l, \lambda_j, 2)$. Conditional on $\bar{\lambda}_j \in \{G, B\}$, the likelihood ratio between $z_j(\bar{\lambda}_j) = z_j \in \{G, B, M\}$ and $z_j(\bar{\lambda}_j) = z_j' \in \{G, B, M\}$ is

$$\frac{\Pr\left(z_j(\bar{\lambda}_j) = z_j \mid \bar{\lambda}_j, \mathbf{y}_i(l, \lambda_j, 2)\right)}{\Pr\left(z_j(\bar{\lambda}_j) = z_j' \mid \bar{\lambda}_j, \mathbf{y}_i(l, \lambda_j, 2)\right)} = \frac{\Pr\left(\mathbf{y}_i(l, \lambda_j, 2) \mid z_j(\bar{\lambda}_j) = z_j\right)}{\Pr\left(\mathbf{y}_i(l, \lambda_j, 2) \mid z_j(\bar{\lambda}_j) = z_j'\right)} \frac{\Pr\left(z_j(\bar{\lambda}_j) = z_j \mid \bar{\lambda}_j\right)}{\Pr\left(z_j(\bar{\lambda}_j) = z_j' \mid \bar{\lambda}_j\right)}.$$

$\log \frac{\Pr\left(\mathbf{y}_i(l, \lambda_j, 2) | z_j(\bar{\lambda}_j) = z_j\right)}{\Pr\left(\mathbf{y}_i(l, \lambda_j, 2) | z_j(\bar{\lambda}_j) = z_j'\right)}$ is expressed as $T^{\frac{1}{2}} \left(\mathcal{L}(\mathbf{y}_i(l, \lambda_j, 2), z_j) - \mathcal{L}(\mathbf{y}_i(l, \lambda_j, 2), z_j')\right)$ with

$$\mathcal{L}(\mathbf{y}_i(l, \lambda_j, 2), z_j) = y_{i,1}(l, \lambda_j, 2) \log q(y_{i,1} | a_i^G, \alpha_j^{z_j}) + \cdots + y_{i,|Y_i|}(l, \lambda_j, 2) \log q(y_{i,|Y_i|} | a_i^G, \alpha_j^{z_j}).$$

Since $\mathcal{L}(\mathbf{y}_i(l, \lambda_j, 2), z_j)$ is strictly concave with respect to the mixture of a_j^G and a_j^B and $\Delta\left(\{\mathbf{1}_{y_i}\}_{y_i \in Y_i}\right) \ni \mathbf{y}_i(l, \lambda_j, 2)$ is compact, there exists $\kappa > 0$ such that one of the following is true:

1. $z_j(\bar{\lambda}_j) = G$ is sufficiently more likely than $z_j(\bar{\lambda}_j) = B$: $\mathcal{L}(\mathbf{y}_i(l, \lambda_j, 2), G) - \kappa \geq \mathcal{L}(\mathbf{y}_i(l, \lambda_j, 2), B)$.

2. $z_j(\bar{\lambda}_j) = B$ is sufficiently more likely than $z_j(\bar{\lambda}_j) = G$: $\mathcal{L}(\mathbf{y}_i(l, \lambda_j, 2), B) - \kappa \geq \mathcal{L}(\mathbf{y}_i(l, \lambda_j, 2), G)$.

3. If $z_j(\bar{\lambda}_j) = G$ and $z_j(\bar{\lambda}_j) = B$ are equally likely, then since $\mathcal{L}(\mathbf{y}_i(l, \lambda_j, 2), z_j)$ is strictly concave, $z_j(\bar{\lambda}_j) = M$ is most likely: $\mathcal{L}(\mathbf{y}_i(l, \lambda_j, 2), M) - \kappa \geq \mathcal{L}(\mathbf{y}_i(l, \lambda_j, 2), G), \mathcal{L}(\mathbf{y}_i(l, \lambda_j, 2), B)$.

If 1 is the case, then $\frac{\Pr\left(z_j(\bar{\lambda}_j) = G | \bar{\lambda}_j, \mathbf{y}_i(l, \lambda_j, 2)\right)}{\Pr\left(z_j(\bar{\lambda}_j) = B | \bar{\lambda}_j, \mathbf{y}_i(l, \lambda_j, 2)\right)} \geq \exp(\kappa T^{\frac{1}{2}}) \frac{\eta}{1 - 2\eta}$ for $\bar{\lambda}_j \in \{G, B\}$. Since $z_j(\bar{\lambda}_j) = M$ implies that player j told a lie and that player i is indifferent to any action profile, player i can believe that, conditional on $\lambda_j(l+1)$, inferring $\lambda_j(l+1)(i) = G$ is almost optimal. Similarly, if 2 is the case, then player i can believe that, conditional on $\lambda_j(l+1)$, inferring $\lambda_j(l+1)(i) = B$ is almost optimal. Finally, if 3 is the case, then player i believes conditional on $\lambda_j(l+1)$ that player j told a lie, that any action profile is optimal with high probability, and that inferring $\lambda_j(l+1)(i) = B$ is almost optimal.

In summary, there exists $\kappa^* > 0$ such that, for any history in the current round, conditional on the true $\lambda_j(l+1)$, one of the following three cases is true:

1. $z_j(\bar{\lambda}_j) = G$ is more likely than $z_j(\bar{\lambda}_j) = B$ by $\exp(\kappa^* T^{\frac{1}{2}})$.

2. $z_j(\bar{\lambda}_j) = B$ is more likely than $z_j(\bar{\lambda}_j) = G$ by $\exp(\kappa^* T^{\frac{1}{2}})$.

3. $z_j(\bar{\lambda}_j) = M$ is most likely by $\exp(\kappa^* T^{\frac{1}{2}})$.

See Figure 10 for the illustration of player j's requirement and Figure 11 for player i's inference. Consider Condition 2: Since almost optimality is established conditional on $\lambda_j(l+1)$, after realizing $\lambda_j(l+1)(i) \neq \lambda_j(l+1)$, player i believes that $\theta_j(l+1) = B$ and $A_i(\tilde{l}+1) = A_i$ for all $\tilde{l} \geq l$ with high probability.

[Insert Figures 10 and 11]

9.3 Formal Message Protocol

The formal definition of the supplemental rounds 1 and 2 for $\lambda_j(l+1)$ is given in the Appendix. The difference from the above intuitive explanation is that there are several events denoted by $\zeta_j = B$ which are excluded from player i's consideration. As we have mentioned in Section 9.2.1, when player j uses player i's message in the supplemental round 2 for $\lambda_i(l+1)$, $\zeta_j = B$ or $\vartheta_j = B$ has happened and player i always assumes that her message in the supplemental round 2 is irrelevant. In summary, we can prove the following Lemma.

Lemma 5 *There exists $\bar{\varepsilon} > 0$ such that, for any $\varepsilon \in (0, \bar{\varepsilon})$, for any $l = 1, ..., L-1$, for sufficiently large T, for any $i, j \in I$, there exists a message protocol in the supplemental rounds such that*

1. *In the supplemental round 1 for $\lambda_j(l+1)$,*

(a) *Player j takes $a_j^{\lambda_j(l+1)}$ and player i takes a_i^G for $T^{\frac{1}{2}}$ periods.*

(b) *Player j creates $\zeta_j(l,\lambda_j,1),\theta_j(l,\lambda_j,1)\in\{G,B\}$.*

(c) *Player i creates $\zeta_i(l,\lambda_j,1),\vartheta_i(l,\lambda_j,1)\in\{G,B\}$. If $\zeta_i(l,\lambda_j,1)=\vartheta_i(l,\lambda_j,1)=G$, then $\lambda_j(l+1)(i)$ is determined solely by player i's history in the supplemental round 1.*

2. *In the supplemental round 2 for $\lambda_j(l+1)$,*

(a) *Player j takes a mixed strategy and player i takes a_i^G for $T^{\frac{1}{2}}$ periods.*

(b) *Player j creates $\theta_j(l,\lambda_j,1)\in\{G,B\}$.*

(c) *If $\lambda_j(l+1)(i)$ is determined by player i's history in the supplemental round 2, then $\zeta_i(l,\lambda_j,1)=B$ or $\vartheta_i(l,\lambda_j,1)=B$ in the previous round.*

3. *Let ζ_j, ϑ_j and θ_j be $\{\zeta_j(\tilde{l}),\zeta_j(\tilde{l},\lambda_j,1),\zeta_j(\tilde{l},\lambda_i,1)\}_{\tilde{l}=1}^{L-1}$, $\{\vartheta_j(\tilde{l},\lambda_i,1)\}_{\tilde{l}=1}^{L-1}$ and $\{\theta_j(\tilde{l},\lambda_j,1),\theta_j(\tilde{l},\lambda_j,2)\}_{\tilde{l}=1}^{L-1}$, respectively. Then*

(a) *Conditional on $\{\lambda_j(\tilde{l}+1)\}_{\tilde{l}=1}^{L-1}$ and "ζ_j and ϑ_j being all G," player i believes that $\lambda_j(l+1)(i)=\lambda_j(l+1)$ or "$\theta_j(l,\lambda_j,1)=B$ or $\theta_j(l,\lambda_j,2)=B$" with probability no less than $1-\exp(-T^{\frac{1}{3}})$.*

(b) *Conditional on $\{\lambda_j(\tilde{l}+1)\}_{\tilde{l}=1}^{L-1}$ and "ζ_j and ϑ_j being all G," any history of player i in the supplemental rounds happens with probability no less than $\exp(-T^{\frac{1}{3}})$.*

(c) *Conditional on $\{\lambda_j(\tilde{l}+1)\}_{\tilde{l}=1}^{L-1}$ and "ζ_j and ϑ_j being all G," conditional on any history of player i in all the supplemental rounds, player i believes that any $\lambda_i(l+1)(j)$ is possible with probability no less than $\exp(-T^{\frac{1}{3}})$.*

(d) *ζ_j and ϑ_j are all G with probability no less than $1-\exp(-T^{\frac{1}{3}})$.*

(e) *The distribution of ζ_j is independent of player i's strategy with probability no less than $1-\exp(-T^{\frac{1}{3}})$.*

(f) *The distribution of $\{\theta_j,\vartheta_j\}$ is independent of player i's strategy.*

10 Variables

In this section, we show that all the variables can be taken consistently satisfying all the requirements that we have mentioned: ρ, $\bar{u}$, q_2, q_1, $\bar{\varepsilon}$, $\bar{L}$, L, η and ε.

First, ρ is determined in (3), $\bar{u}$ is determined in Lemma 2, q_1 and q_2 are determined in Lemma 3, and $\bar{\varepsilon} \in \left(0, \frac{1}{2}\right)$ is defined in Lemma 5, independently of the other variables.

Given q_1 and q_2, we define $\bar{L}$ to satisfy (10):

$$\bar{L}\left(q_2 - q_1\right) > \max_{a,i} 2\left|u_i\left(a\right)\right|. \tag{42}$$

Given ρ and $\bar{L}$, we define L sufficiently large so that (11) holds:

$$\bar{L} < L\rho. \tag{43}$$

We are left to pin down η and ε. Take $\varepsilon > 0$ and $\eta > 0$ sufficiently small such that

$$u_i\left(D_1, D_2\right) + \rho + 2\varepsilon\bar{L} + 3\bar{u}\eta < \underline{v}_i < \overline{v}_i < u_i\left(C_1, C_2\right) - \rho - 2\varepsilon\bar{L} - 3\bar{u}\eta, \tag{44}$$

and

$$\varepsilon < \min\left\{\bar{\varepsilon}, q_2 - q_1\right\}. \tag{45}$$

11 Almost Optimality of $\sigma_i(x_i)$

Since we have pinned down $\lambda_j(l+1)(i)$ in Section 9, Section 8.2.2 pins down $\hat{\lambda}_j(l+1)$. Then, Section 8.1 pins down the equilibrium action. In addition, since we have pinned down θ_j, ζ_j and ϑ_j, Section 9 pins down $\pi_i(x_j, \cdot, \delta)$ except for $\bar{\pi}_i(x, l, \lambda_j)$. Therefore, we want to show that there exists $\bar{\pi}_i(x, l, \lambda_j)$ such that these actions and reward functions satisfy (21), (5) and (6).

11.1 Almost Optimality of $\hat{\lambda}_j(l+1)$

First, we show the almost optimality of $\hat{\lambda}_j(l+1)$. Since we verified 2 of Condition 2 is satisfied in Lemma 5, this follows from Section 9.1.

For notational convenience, for each lth review round, let $\mathbf{a}_j(l)$ be the sequence of player j's equilibrium actions in each $\tilde{l}$th review round with $\tilde{l} \leq l$.

As we have mentioned in Section 9, in each lth review round ($r = l$), player i conditions that $\zeta_j(r) = \vartheta_j(r) = G$ for all $r < l$. Since Condition 2 is satisfied conditional on $\zeta_j = \vartheta_j = G$ in Lemma 5, the almost optimality still holds:

Lemma 6 *For any lth review round, conditional on $\mathbf{a}_j(l)$ and $\zeta_j(r) = \vartheta_j(r) = G$ for all $r < l$, for any history of player i where player i has not deviated in the supplemental rounds for $\lambda_j(\tilde{l})$ with $\tilde{l} \leq l$, player i can believe that $\hat{\lambda}_j(l) = \lambda_j(l)$ or there exists $r < l$ with $\theta_j(r) = B$ with probability no less than*

$$1 - \exp(-T^{\frac{1}{3}}). \tag{46}$$

11.2 Determination of $\bar{\pi}_i(x, l, \lambda_j)$

Second, based on Lemma 6, we determine $\bar{\pi}_i(x, l, \lambda_j)$ such that $\bar{\pi}_i(x, l, \lambda_j)$ satisfies (6) and show that $\sigma_i(x_i)$ is almost optimal:

Proposition 1 *There exists $\bar{\pi}_i(x, l, \lambda_j)$ that satisfies (6) and such that $\sigma_i(x_i)$ is almost optimal, that is, for sufficiently large T, for all round r,*

1. If $\zeta_j(\tilde{r}) = B$ or $\vartheta_j(\tilde{r}) = B$ with some $\tilde{r} < r$, then any strategy is exactly optimal.

2. If $\zeta_j(\tilde{r}) = \vartheta_j(\tilde{r}) = G$ for all $\tilde{r} < r$, then

   (a) For all $l \in \{1, ..., L-1\}$, for $r = (l, \lambda_i, 2)$ (supplemental round 2 for $\lambda_i(l+1)$) where player i takes a mixed strategy, any strategy is exactly optimal.

   (b) For all $l \in \{1, ..., L-1\}$, for $r = (l, \lambda_j, 1)$ or $(l, \lambda_j, 2)$ (supplemental round 1 or 2 for $\lambda_j(l+1)$) where player i receives the message, a_i^G is strictly optimal by at least $\frac{1}{2}\bar{L}(q_2 - q_1)$.

(c) For the other rounds, $\sigma_i(x_i)$ is optimal with loss up to $\exp(-T^{\frac{1}{4}})$.

1 follows from (32). 2-(a) is true since $\zeta_j(\tilde{r}) = \vartheta_j(\tilde{r}) = G$ for all $\tilde{r} < r$ imply that player j has inferred $\lambda_i(l+1)(j)$ from the supplemental round 1 (See Figure 13 reversing the roles of i and j). 2-(b) is true since taking a_i^G gives the almost optimal inference and (35) gives a high reward on a_i^G. 2-(c) is true intuitively because (i) in the supplemental round 1 for $\lambda_i(l+1)$, whenever player j is ready to listen, player i's continuation payoff is fixed regardless of player j's inference from 4 of Lemma 4 and (ii) Lemma 6 guarantees that the coordination goes well with high probability and that we can construct the reward functions in the review rounds as in Section 5.2.2.

12 Exact Optimality

We are left to show the truthtelling incentive for $h_i^{T_P+1}$ in the report block and to establish the exact optimality of $\sigma_i(x_i)$. Here, we offer the intuitive explanation. See Section 15.7 in the Appendix for the formal proof.

The report block proceeds 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 $h_1^{T_P+1}$ with probability $\frac{1}{2}$ and player 2 reports $h_2^{T_P+1}$ with probability $\frac{1}{2}$. Player i who is supposed to send $h_i^{T_P+1}$ according to the public randomization device sends two messages for each round r sequentially where player i does not take a mixed strategy (that is, except for the supplemental round 2 for $\lambda_i(l+1)$): (i-r-1) What action player i took and what signal player i observed in the first period of the round r, (a_{i,t_r}, y_{i,t_r}). Here, t_r is the first period of the round r. (i-r-2) The history in the round r, $\{a_{i,t}, y_{i,t}\}_{t\in T(r)}$.

Let h_i^r be the history in each round where player i does not take a mixed strategy before the round r. Notice that, before player i sends the message about (i-r-1), the messages about h_i^r have been sent.

By backward induction, for each r, player j who is not supposed to send $h_j^{T_P+1}$ by the realization of the public randomization device adjusts the reward function as follows.

Here, to verify the truthtelling incentive, we distinguish the true histories h_i^r, (a_{i,t_r}, y_{i,t_r}) and $\{a_{i,t}, y_{i,t}\}_{t\in T(r)}$ and the messages $\hat{h}_i^r$, $(\hat{a}_{i,t_r}, \hat{y}_{i,t_r})$ and $\{\hat{a}_{i,t}, \hat{y}_{i,t}\}_{t\in T(r)}$.

(j-r-1) Given $\hat{h}_i^r$, player j gives a reward based on $\left\{\Psi_{j,t_r}^{a_i,a_{j,t_r}}\right\}_{a_i\in A_i}$ such that any action that should be taken by $\sigma_i(x_i)$ after $\hat{h}_i^r$ is optimal in t_r (the first period of the round r). (j-r-2) Player j punishes player i if player i is likely to tell a lie about (a_{i,t_r}, y_{i,t_r}) in (i-r-1). (j-r-3) Player j gives a reward based on $\sum_{t\in T(r)} \Psi_{j,t}^{\hat{a}_{i,t_r},a_{j,t}}$ such that constantly taking the same action as $\hat{a}_{i,t_r}$ during the round r is optimal. (j-r-4) Player j punishes player i if player i is likely to tell a lie about $\{\hat{a}_{i,t}, \hat{y}_{i,t}\}_{t\in T(r)}$ in (i-r-2).

We take the punishment and reward satisfying the following requirements:

1. Within the round r, we take the reward in (j-r-1) much larger than the punishment or reward in (j-r-2), (j-r-3) and (j-r-4). Similarly, we take the punishment in (j-r-2) much larger than the punishment or reward in (j-r-3) and (j-r-4), and the reward in (j-r-3) much larger than the punishment in (j-r-4).

2. Between rounds, we take the punishment and reward from (j-r-1) to (j-r-4) much larger than those from (j-$r+1$-1) to (j-$r+1$-4).

By backward induction, we can verify the truthtelling incentive and make $\sigma_i(x_i)$ exactly optimal: In round r, we do not need to consider the punishment or reward for the round $\tilde{r} > r$ by Requirement 2 above. First, given that player i tells the truth in the report block, (j-r-1) is enough to make the equilibrium strategy optimal in t_r, taking into account the effect on (j-r-2), (j-r-3) and punishment and reward for the round $\tilde{r} > r$ by Requirement 1. Second, (j-r-2) is enough to give the truthtelling incentive for (a_{i,t_r}, y_{i,t_r}) since (i) the punishment and reward for (j-$\tilde{r}$-1), (j-$\tilde{r}$-2), (j-$\tilde{r}$-3) and (j-$\tilde{r}$-4) with $\tilde{r} < r$ are sunk, (ii) the reward (j-r-1) is sunk, and (iii) the punishment and reward for (j-r-3) and (j-r-4) are much smaller than the reward for (j-r-2). Third, given the truthtelling incentive for a_{i,t_r}, (j-r-3) is enough to incentivize player i to constantly take a_{i,t_r} within the round r since (i) the punishment and reward for (j-$\tilde{r}$-1), (j-$\tilde{r}$-2), (j-$\tilde{r}$-3) and (j-$\tilde{r}$-4) with $\tilde{r} < r$ are sunk, (ii) the punishment and reward for (j-r-1) and (j-r-2) are sunk, and (iii) the punishment for (j-r-4) is much smaller than the reward for (j-r-3). Finally, (j-r-4) is enough to incentivize player i to tell the truth about $\{a_{i,t}, y_{i,t}\}_{t \in T(r)}$ since (i) the punishment and reward for (j-$\tilde{r}$-1), (j-$\tilde{r}$-2), (j-$\tilde{r}$-3) and (j-$\tilde{r}$-4) with $\tilde{r} < r$ are sunk, and (ii) the punishment and reward for (j-r-1), (j-r-2) and (j-r-3) are sunk.

Note that we do not require player i to send the messages about the rounds where player i takes a mixed strategy. From Proposition 1, $\sigma_i(x_i)$ was exactly optimal without the adjustment in the report block. We cancel out the adjustment of the punishment and reward in the round r explained above once $\zeta_j(\tilde{r}) = B$ or $\vartheta_j(\tilde{r}) = B$ with $\tilde{r} < r$ happens. Then, from Proposition 1, $\sigma_i(x_i)$ is exactly optimal in all the rounds.

In addition, in the supplemental round 2 for $\lambda_i(l+1)$, if $\zeta_j(\tilde{r}) = \vartheta_j(\tilde{r}) = G$ until this round, then player i's strategy in this round does not affect the continuation strategy profile including the report block except for player j's messages in the report block. Since Proposition 1 guarantees that $\sigma_i(x_i)$ is exactly optimal in this round without adjustment, we can omit this round from the rounds whose history player i sends the messages in the report block about.

From Proposition 1, the last adjustment (j-R-3) can be very small. Hence, the total reward and punishment can be small enough not to affect the equilibrium payoff.

13 General Games

13.1 General Two-Player Games

The prisoners' dilemma is special in that, when player j' state is B, the equilibrium action D_j can (i) attain the targeted payoff $\bar{v}_j$ or $\underline{v}_j$ depending on player i's state and (ii) minimax player i at the same time. However, in a general game, there does not always exist such an action, which causes the following problem: Player j with $x_j = B$ and $\lambda_j(l) = B$ needs to give a *positive constant* reward as in (33) 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 minimax player i with high probability if player i

does not take a prescribed action although the minimaxing action can be different from the prescribed action to attain the targetted payoff $\bar{v}_j$ or $\underline{v}_j$.

For this purpose, player i constructs a statistics such that if its realization is low, then player i *allows player j to minimax player i*. Player i sends the message about whether she allows player j to minimax. Player j, on the other hand, calculates the conditional expectation of that statistics which *decreases if player i deviates*. Modifying the coordination protocol about $\lambda_i(l+1)$, we can make a protocol such that the players can coordinate on the punishment properly and that player j punishes player i with high probability if the realization of the conditional expectation is sufficiently low regardless of player i's message to keep player i's payoff below $\underline{v}_i$ regardless of player i's strategy. See the Supplemental Material 1 for the formal description.

The fact that there does not always exist an action satisfying both (i) and (ii) is the reason why the belief-free equilibrium payoff set is smaller than the feasible and individually rational payoff set except for the prisoners' dilemma. Hörner and Olszewski (2006) consider a way to coordinate on the punishment, which is different from ours. However, since their coordination uses almost perfect monitoring, as Hörner and Olszewski (2006) mention, it was not known how to obtain the coordination in not-almost-perfect monitoring.

13.2 General More-Than-Two-Player Games

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 (See Supplemental Material 2).

14 Equilibrium Construction without Cheap Talk

14.1 Two-Player Games

Throughout the main text, we use perfect and public cheap talk and public randomization devices.[16] In the Supplemental Material 3, we show that cheap talk and public randomization are completely *dispensable* and all the messages can be sent via actions without public randomization.

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 5 shows, *conditional on the opponent message*, the receiver can believe that her inference is correct or she is indifferent to any action profile sequence. Therefore, even after observing the continuation strategy that is not corresponding to her inference, the receiver can believe that she is indifferent to any action profile sequence with high probability and that following the equilibrium strategy is almost optimal. As we have seen in Section 12, we can attain the exact optimality afterwards.

There is one difference between $\lambda_i(l+1)$ and x_i. For $\lambda_i(l+1)$, whenever player j changes her strategy based on the inference of $\lambda_i(l+1)$, player j has made player i indifferent to any action profile sequence and the sender's inference of the receiver's inference is irrelevant (4 of Lemma 4). On the other hand, the receiver j constructs the reward function on the sender i based on x_i (remember that π_i depends on not only x_j but also x_i). Therefore, the sender needs to have the inference of the receiver's inference. For this purpose, player j sends back her inference of x_i to player i to inform what inference the reward function is based on. The incentive to tell the truth for player j can be provided.

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 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 $h_i^{T_P+1}$

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

since otherwise, while the opponent is reporting $h_j^{T_P+1}$, player i could update the belief over player j's signal observations and the incentive to tell the truth about $h_i^{T_P+1}$ would be destroyed. On the other hand, to have the exact optimality, there needs to be a positive probability for each player to report $h_i^{T_P+1}$ and the reward function will be adjusted. Third, there exists a positive probability that the messages do not transmit correctly. See the Supplemental Material 3 for the solutions to these problems.

14.2 More-Than-Two-Player Games

There is a problem in coordinating on x_i, unique to the case with more than two players: Each player takes an action based on the her own inference about x_i. It may be of some player's interest that some players take actions corresponding to $x_i = G$ and the others take actions corresponding to $x_i = B$. Then, that player may have an incentive to deviate from the equilibrium action in order to manipulate the signal distributions and to induce the miscoordination about x_i. See the Supplemental Material 4 for how to deal with this problem.

15 Appendix

15.1 Proof of Lemma 1

To see why this is enough for Theorem 1, define the strategy in the infinitely repeated game as follows: Define

$$\begin{aligned} p(G, h_j^{T_P+1} &: \delta) \equiv 1 + \frac{1-\delta}{\delta^{T_P}} \frac{\pi_i(G, h_j^{T_P+1} : \delta)}{\bar{v}_i - \underline{v}_i}, \\ p(B, h_j^{T_P+1} &: \delta) \equiv \frac{1-\delta}{\delta^{T_P}} \frac{\pi_i(B, h_j^{T_P+1} : \delta)}{\bar{v}_i - \underline{v}_i}. \end{aligned} \tag{47}$$

If (6) is satisfied, for sufficiently large δ, $p(G, h_j^{T_P+1} : \delta)$, $p(B, h_j^{T_P+1} : \delta) \in [0,1]$ for all $h_j^{T_P+1}$. 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 $\sigma_i(x_i)$ in the current phase. 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.

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

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

and

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

(4) and (5) imply, for sufficiently large discount factor δ,

1. Conditional on the opponent's state, the above strategy in the infinitely repeated game is optimal.

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 $\underline{v}_i$.

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.

15.2 Proof of Lemma 3

First, note that (27) is equivalent to

$$\sum_{y_i}\left\{\sum_{y_j}\psi_j^a(y_j)q(y_j \mid a, y_i)\right\}q\left(y_i \mid a_i, \tilde{a}_j\right) = \sum_{y_j}\left\{\sum_{y_i}q(y_j \mid a, y_i)q(y_i \mid a_i, \tilde{a}_j)\right\}\psi_j^a(y_j) = q_2. \tag{48}$$

Second, arbitrarily fix $q_2 > q_1$. Let $Q_1(\tilde{a}_i, a_j) \equiv \left(q\left(y_j \mid \tilde{a}_i, a_j\right)\right)_{y_j}$ be the vector of the distribution of player j's signal conditional on $\tilde{a}_i, a_j$. In addition, let $Q_2(a_i, \tilde{a}_j) \equiv \left(\sum_{y_i} q(y_j \mid a, y_i)q(y_i \mid a_i, \tilde{a}_j)\right)_{y_j}$. The vectors $Q_1(\tilde{a}_i, a_j)$ and $Q_2(a_i, \tilde{a}_j)$ 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 $\{\psi_i^a(y_j)\}_{y_j}$ for (26) and (27). 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 $\psi_j^a : Y_j \to (0,1)$.

15.3 Proof of Lemma 2

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

15.4 Proof of Lemma 5

15.4.1 Explanation of 1 of Lemma 5: Supplemental Round 1 for $\lambda_j(l+1)$

The following is the formal description of the message exchanges in the supplemental round 1 for $\lambda_j(l+1)$. Let us first introduce the following notations:

Notation 1 *For $\lambda_j \in \{G, B\}$, we define*

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

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

2. $(|Y_j| - |A_i| + 1) \times |Y_j|$ *matrix* $H_j(\bar{\lambda}_j)$ *and* $(|Y_j| - |A_i| + 1) \times 1$ *vector* $\mathbf{p}_j(\bar{\lambda}_j)$ *such that the affine hull of player j's signal distribution with respect to player i's action when player j takes* $a_j^{\bar{\lambda}_j}$ *is represented by*

$$\mathrm{aff}(\{\mathbf{q}_j(a_j^{\bar{\lambda}_j}, a_i)\}_{a_i}) \cap \mathbb{R}_+^{|Y_j|} = \left\{\mathbf{y}_j \in \mathbb{R}_+^{|Y_j|} : H_j(\bar{\lambda}_j)\mathbf{y}_j = \mathbf{p}_j(\bar{\lambda}_j)\right\}. \tag{49}$$

3. *The set of hyperplanes generated by perturbing RHS of the characterization of* $\mathrm{aff}(\{\mathbf{q}_j(a_j^{\bar{\lambda}_j}, a_i)\}_{a_i}) \cap \mathbb{R}_+^{|Y_j|}$*: for* $\varepsilon \geq 0$,[17]

$$\mathcal{H}_j[\varepsilon](\bar{\lambda}_j) \equiv \left\{\mathbf{y}_j \in \mathbb{R}_+^{|Y_j|} : \exists \boldsymbol{\varepsilon} \in \mathbb{R}_+^{|Y_j|-|A_i|+1} \text{ such that } \left\{\begin{array}{c} \|\boldsymbol{\varepsilon}\| \leq \varepsilon, \\ H_j(\bar{\lambda}_j)\mathbf{y}_j = \mathbf{p}_j(\bar{\lambda}_j)+\boldsymbol{\varepsilon} \end{array}\right\}\right\}.$$

4. *The set of frequencies of player i's signal observations such that the conditional expectation of the frequency of player j's signal observation is close to* $\mathcal{H}_j[\varepsilon](\hat{\lambda}_j)$ *when the players take* $a_j^{\hat{\lambda}_j}, a_i^G$*:*

$$\mathcal{H}_{j,i}[\varepsilon](\hat{\lambda}_j) = \left\{\begin{array}{c} \mathbf{y}_i \in \mathbb{R}_+^{|Y_i|} \text{ such that} \\ \text{there exist } \boldsymbol{\varepsilon}_1 \in \mathbb{R}_+^{|Y_j|},\ \boldsymbol{\varepsilon}_2 \in \mathbb{R}_+^{|Y_j|-|A_i|+1} \text{ and } \mathbf{y}_j \in \mathbb{R}_+^{|Y_j|} \text{ with} \\ \left\{\begin{array}{c} \mathbf{y}_j = Q_{j,i}(a_j^{\hat{\lambda}_j}, a_i^G)\mathbf{y}_i + \boldsymbol{\varepsilon}_1, \\ H_j(\hat{\lambda}_j)\mathbf{y}_j = \mathbf{p}_j(\hat{\lambda}_j) + \boldsymbol{\varepsilon}_2, \\ \|\boldsymbol{\varepsilon}_1\|, \|\boldsymbol{\varepsilon}_2\| \leq \varepsilon \end{array}\right. \end{array}\right\}.$$

Without loss of generality, we can make sure that all the elements in $H_j(G)$ and $H_j(B)$ are included in $(0,1)$ by the following lemma.

Lemma 7 *We can take* $H_j(G)$ *and* $H_j(B)$ *such that all the elements are in* $(0,1)$.

Proof. Let m_H be the minimum element of $H_j(\bar{\lambda}_j)$ and M_H be the maximum element of $H_j(\bar{\lambda}_j)$. Let $\tilde{H}_j(\bar{\lambda}_j)$ be the matrix whose (l,m) element is $\frac{\left(H_j(\bar{\lambda}_j)\right)_{l,m} + |m_H| + 1}{|M_H| + |m_H| + 2} \in (0,1)$ and $\tilde{\mathbf{p}}_j(\bar{\lambda}_j)$ be the vector whose lth element is $\frac{\left(\mathbf{p}_j(\bar{\lambda}_j)\right)_l + |m_H| + 1}{|M_H| + |m_H| + 2}$.

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

We will show

$$\left\{\mathbf{y}_j \in \mathbb{R}_+^{|Y_j|} : H_j(\bar{\lambda}_j)\mathbf{y}_j = \mathbf{p}_j(\bar{\lambda}_j)\right\} \equiv \mathcal{H}_j(\bar{\lambda}_j) = \tilde{\mathcal{H}}_j(\bar{\lambda}_j) \equiv \left\{\mathbf{y}_j \in \mathbb{R}_+^{|Y_j|} : \tilde{H}_j(\bar{\lambda}_j)\mathbf{y}_j = \tilde{\mathbf{p}}_j(\bar{\lambda}_j)\right\}.$$

1. $\mathcal{H}_j(\bar{\lambda}_j) \subset \tilde{\mathcal{H}}_j(\bar{\lambda}_j)$

   Suppose $\mathbf{y}_j \in \mathcal{H}_j(\bar{\lambda}_j)$. Since $\mathbf{y}_j \in \text{aff}(\{\mathbf{q}_j(a_j^{\bar{\lambda}_j}, a_i)\}_{a_i}) \subset \text{aff}\left(\{0,1\}^{|Y_j|}\right)$, $\tilde{H}_j(\bar{\lambda}_j)\mathbf{y}_j = \tilde{\mathbf{p}}_j(\bar{\lambda}_j)$ as desired.

2. $\mathcal{H}_j(\bar{\lambda}_j) \supset \tilde{\mathcal{H}}_j(\bar{\lambda}_j)$

   Suppose $\mathbf{y}_j \notin \mathcal{H}_j(\bar{\lambda}_j)$. Since $\text{aff}(\{\mathbf{q}_j(a_j^{\bar{\lambda}_j}, a_i)\}_{a_i}) \subset \text{aff}\left(\{0,1\}^{|Y_j|}\right)$, without loss of generality, we can assume that one row of $H_j(\bar{\lambda}_j)$ is parallel to $(1, ..., 1)$ and that the element of $\mathbf{p}_j(\bar{\lambda}_j)$ corresponding to that row is 1. If $(1, ..., 1)\mathbf{y}_j \neq 1$, then $\left(\frac{1+|m_H|+1}{|M_H|+|m_H|+2}, ..., \frac{1+|m_H|+1}{|M_H|+|m_H|+2}\right)\mathbf{y}_j = \frac{1+|m_H|+1}{|M_H|+|m_H|+2}(1, ..., 1)\mathbf{y}_j \neq \frac{1+|m_H|+1}{|M_H|+|m_H|+2}$ and $\mathbf{y}_j \notin \tilde{\mathcal{H}}_j(\bar{\lambda}_j)$ as desired. If $(1, ..., 1)\mathbf{y}_j = 1$, then there is another row $\mathbf{h}_j(\bar{\lambda}_j)$ and the corresponding element $p_j(\bar{\lambda}_j)$ of $\mathbf{p}_j(\bar{\lambda}_j)$ such that

   $$\mathbf{h}_j(\bar{\lambda}_j)\mathbf{y}_j \neq p_j(\bar{\lambda}_j).$$

   Let $\tilde{\mathbf{h}}_j(\bar{\lambda}_j)$ be the corresponding row of $\tilde{H}_j(\bar{\lambda}_j)$ and $\tilde{p}_j(\bar{\lambda}_j)$ is the corresponding element of $\tilde{\mathbf{p}}_j(\bar{\lambda}_j)$. Then,

   $$\begin{aligned}\tilde{\mathbf{h}}_j(\bar{\lambda}_j)\mathbf{y}_j &= \frac{1}{|M_H|+|m_H|+2}\left(\mathbf{h}_j(\bar{\lambda}_j) + (|m_H|+1)(1, ..., 1)\right)\mathbf{y}_j \\ &= \frac{1}{|M_H|+|m_H|+2}\left(\mathbf{h}_j(\bar{\lambda}_j)\mathbf{y}_j + |m_H|+1\right) \\ &\neq \frac{1}{|M_H|+|m_H|+2}\left(p_j(\bar{\lambda}_j) + |m_H|+1\right) = \tilde{p}_j(\bar{\lambda}_j)\end{aligned}$$

   and so $\mathbf{y}_j \notin \tilde{\mathcal{H}}_j(\bar{\lambda}_j)$.

■

Given the frequency of player j's signal observation $\mathbf{y}_j(l, \lambda_j, 1)$, $\mathbf{y}_j(l, \lambda_j, 1) \in \mathcal{H}_j[\varepsilon](\bar{\lambda}_j)$ corresponds to Box 2 of Figure 9. On the other hand, given the frequency of player i's signal observation $\mathbf{y}_i(l,\lambda_j,1)$, $\mathbf{y}_i(l,\lambda_j,1)\in\mathcal{H}_i[\varepsilon](G)$ corresponds to Boxes 5, 6 and 7 of Figure 9. In addition, since

$$\mathbb{E}\left[\frac{1}{T^{\frac{1}{2}}-1}\sum_{t\in T_i(l,\lambda_j,1)}\mathbf{1}_{y_{j,t}}\mid\{y_{i,t}\}_{t\in T_i(l,\lambda_j,1)},a_j^{\hat{\lambda}_j},a_i^G\right]=Q_{j,i}(a_j^{\hat{\lambda}_j},a_i^G)\mathbf{y}_i(l,\lambda_j,1),$$

$\mathbf{y}_i(l,\lambda_j,1)\in\mathcal{H}_{j,i}[\varepsilon](\hat{\lambda}_j)$ corresponds to Box 5 or 6 of Figure 9, depending on $\hat{\lambda}_j=G$ or B.

Strategies The strategy in the supplemental round 1 for $\lambda_j(l+1)$ is as follows: Player j (sender) with $\lambda_j(l+1)=\bar{\lambda}_j\in\{G,B\}$ takes $a_j^{\bar{\lambda}_j}$ and player i (receiver) takes a_i^G for $T^{\frac{1}{2}}$ periods in the supplemental round 1 for $\lambda_j(l+1)$, that is, for $t\in T(l,\lambda_j,1)$.

Requirement by Player j As we have mentioned in Section 9.2.1, player j requires player i to infer $\lambda_j(l+1)$ correctly only if $\mathbf{y}_j(l,\lambda_j,1)\in\mathcal{H}_j[\varepsilon](\bar{\lambda}_j)$.

Instead of using $\mathbf{y}_j(l,\lambda_j,1)$ directly, player j constructs random variables $\{\mathbf{\Omega}_{j,t}^H\}_{t\in T_j(l,\lambda_j,1)}$ as follows. After taking $a_j^{\bar{\lambda}_j}$ and observing $y_{j,t}$, player j calculates $H_j(\bar{\lambda}_j)\mathbf{1}_{y_{j,t}}$. Then, player j draws $(|Y_j|-|A_i|+1)$ 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 $H_j(\bar{\lambda}_j)\mathbf{1}_{y_{j,t}}$, then the lth element of $\mathbf{\Omega}_{j,t}^H$ is equal to 1. Otherwise, the lth element of $\mathbf{\Omega}_{j,t}^H$ is equal to 0. From Lemma 7,

$$\Pr\left(\left\{\left(\mathbf{\Omega}_{j,t}^H\right)_l=1\right\}\mid a,y\right)=\left(H_j(\bar{\lambda}_j)\mathbf{1}_{y_{j,t}}\right)_l\in(0,1)\tag{50}$$

for all a and y.

Given $\{\mathbf{\Omega}_{j,t}^H\}_{t\in T_j(l,\lambda_j,1)}$, let

$$\Omega_j^H(l,\lambda_j,1)=\frac{1}{T^{\frac{1}{2}}-1}\sum_{t\in T_j(l,\lambda_j,1)}\mathbf{\Omega}_{j,t}^H.\tag{51}$$

Player j requires player i to infer $\lambda_j(l+1)$ correctly only if

$$
\begin{aligned}
& \left\| \Omega_j^H(l,\lambda_j,1) - H_j(\bar{\lambda}_j)\mathbf{y}_j(l,\lambda_j,1) \right\| \\
= & \left\| \frac{1}{T^{\frac{1}{2}}-1} \sum_{t\in T_j(l,\lambda_j,1)} \mathbf{\Omega}_{j,t}^H - \frac{1}{T^{\frac{1}{2}}-1} \sum_{t\in T_j(l,\lambda_j,1)} H_j(\bar{\lambda}_j)\mathbf{1}_{y_{j,t}} \right\| \leq \frac{\varepsilon}{2}
\end{aligned} \tag{52}
$$

and

$$
\left\| \Omega_j^H(l,\lambda_j,1) - \mathbf{p}_j(\bar{\lambda}_j) \right\| \leq \frac{\varepsilon}{2}. \tag{53}
$$

In summary, there are following cases:

1. (52) is not satisfied. Let $\zeta_j(l,\lambda_j,1) = B$ denote this event. This case is excluded from player i's consideration. From (50) and (51), by the central limit theorem, conditional on $\{y_{j,t}\}_{t\in T(l,\lambda_j,1)}$, this occurs with probability no more than of order $\exp(-T^{\frac{1}{2}})$.

2. (52) is satisfied. Let $\zeta_j(l,\lambda_j,1) = G$ denote this event.

    (a) (53) is not satisfied. Let $\theta_j(l,\lambda_j,1) = B$ in this event. Then, player j makes player i indifferent between any action profile sequence in the subsequent rounds. This case is not excluded from player i's consideration. Since we take the affine hull with respect to player i's action in (49), together with (50), the central limit theorem implies that this occurs with probability no more than of order $\exp(-T^{\frac{1}{2}})$ ex ante at the beginning of the supplemental round 1 regardless of player i's strategy.

    (b) (53) is satisfied. Depending on the supplemental round 2, player j may require player i to infer $\lambda_j(l+1)$ correctly. Let $\theta_j(l,\lambda_j,1) = G$ denote this event. Note that (52) and (53) imply $\left\| H_j(\bar{\lambda}_j)\mathbf{y}_j(l,\lambda_j,1) - \mathbf{p}_j(\bar{\lambda}_j) \right\| \leq \varepsilon$ by triangle inequality and so $\mathbf{y}_j(l,\lambda_j,1) \in \mathcal{H}_j[\varepsilon](\bar{\lambda}_j)$. Therefore, player j requires player i to infer $\lambda_j(l+1)$ correctly only if $\mathbf{y}_j(l,\lambda_j,1) \in \mathcal{H}_j[\varepsilon](\bar{\lambda}_j)$ as mentioned.

In addition, if 1 is the case, we define $\theta_j(l,\lambda_j,1) \in \{G,B\}$ as in the case with

$\zeta_j(l, \lambda_j, 1) = G$: If (53) is not satisfied, then $\theta_j(l, \lambda_j, 1) = B$. If (53) is satisfied, then $\theta_j(l, \lambda_j, 1) = G$.

See the upper half of Figure 12 for the illustration. Figure 12 corresponds to Figure 9 in the intuitive explanation. Box 2 of Figure 9 corresponds to Box 5 in Figure 12 and it is still true that player j requires player i to infer $\lambda_j(l+1)$ correctly only if $\mathbf{y}_j(l, \lambda_j, 1) \in \mathcal{H}_j[\varepsilon](\bar{\lambda}_j)$. Therefore, the sufficient condition for the almost optimal inference by player i is still valid.

[Insert Figure 12]

Inference by Player i As mentioned in Section 9.2.1, if player i infers $\lambda_j(l+1)$ from the supplemental round 1, then player i infers $\lambda_j(l+1) = \hat{\lambda}_j \in \{G, B\}$ if

$$\mathbb{E}\left[\frac{1}{T^{\frac{1}{2}} - 1} \sum_{t \in T_i(l, \lambda_j, 1)} \mathbf{1}_{y_j} \mid \mathbf{y}_i(l, \lambda_j, 1), a_j^{\hat{\lambda}_j}, a_i^G\right]$$

is close to aff $\left(\{\mathbf{q}_j(a_j^{\hat{\lambda}_j}, a_i)\}_{a_i}\right)$. In other words, player i infers $\lambda_j(l+1) = G$ if $\mathbf{y}_i(l, \lambda_j, 1) \in \mathcal{H}_{j,i}[\varepsilon](G)$ and $\lambda_j(l+1) = B$ if $\mathbf{y}_i(l, \lambda_j, 1) \in \mathcal{H}_{j,i}[\varepsilon](B)$.

Instead of using $\mathbf{y}_i(l, \lambda_j, 1)$ directly, player i constructs random variables $\{\mathbf{\Omega}_{i,t}(G)\}_{t \in T_i(l, \lambda_j, 1)}$, $\{\mathbf{\Omega}_{i,t}(B)\}_{t \in T_i(l, \lambda_j, 1)}$ and $\{\mathbf{\Omega}_{i,t}^G\}_{t \in T_i(l, \lambda_j, 1)}$ as follows.

Construction of $\left\{\mathbf{\Omega}_{i,t}(\hat{\lambda}_j)\right\}_{t \in T_i(l, \lambda_j, 1)}$ with $\hat{\lambda}_j \in \{G, B\}$ After taking a_i^G and observing $y_{i,t}$, player i calculates $H_j(\hat{\lambda}_j) Q_{j,i}(a_j^{\hat{\lambda}_j}, a_i^G) \mathbf{1}_{y_{i,t}}$. Then, player i draws $|Y_j| - |A_i| + 1$ 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 $H_j(\hat{\lambda}_j) Q_{j,i}(a_j^{\hat{\lambda}_j}, a_i^G) \mathbf{1}_{y_{i,t}}$, then the lth element of $\mathbf{\Omega}_{i,t}(\hat{\lambda}_j)$ is equal to 1. Otherwise, the lth element of $\mathbf{\Omega}_{i,t}(\hat{\lambda}_j)$ is equal to 0. Since all the elements of $H_j(\hat{\lambda}_j)$ is in $(0, 1)$ from Lemma 7 and $Q_{j,i}(a_j^{\hat{\lambda}_j}, a_i^G) \mathbf{1}_{y_{i,t}}$ is a conditional probability distribution, we have $H_j(\hat{\lambda}_j) Q_{j,i}(a_j^{\hat{\lambda}_j}, a_i^G) \mathbf{1}_{y_{i,t}} \in (0, 1)$, and

$$\Pr\left(\left\{\left(\mathbf{\Omega}_{i,t}(\hat{\lambda}_j)\right)_l = 1\right\} \mid a, y\right) = \left(H_j(\hat{\lambda}_j) Q_{j,i}(a_j^{\hat{\lambda}_j}, a_i^G) \mathbf{1}_{y_{i,t}}\right)_l. \tag{54}$$

Given $\{\mathbf{\Omega}_{i,t}(\hat{\lambda}_j)\}_{t\in T_i(l,\lambda_j,1)}$, let

$$\Omega_i(\hat{\lambda}_j)(l,\lambda_j,1)=\frac{1}{T^{\frac{1}{2}}-1}\sum_{t\in T_i(l,\lambda_j,1)}\mathbf{\Omega}_{i,t}(\hat{\lambda}_j). \tag{55}$$

Construction of $\{\mathbf{\Omega}_{i,t}^G\}_{t\in T_i(l,\lambda_j,1)}$ After taking a_i^G and observing $y_{i,t}$, player i calculates $H_i(G)\mathbf{1}_{y_{i,t}}$. 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 $H_i(G)\mathbf{1}_{y_{i,t}}$, then the lth element of $\mathbf{\Omega}_{i,t}^G$ is equal to 1. Otherwise, the lth element of $\mathbf{\Omega}_{i,t}^G$ is equal to 0. By Lemma 7,

$$\Pr\left(\left\{\left(\mathbf{\Omega}_{i,t}^G\right)_l=1\right\}\mid a,y\right)=\left(H_i(G)\mathbf{1}_{y_{i,t}}\right)_l\in(0,1). \tag{56}$$

Given $\{\mathbf{\Omega}_{i,t}^G\}_{t\in T_i(l,\lambda_j,1)}$, let

$$\Omega_i^G(l,\lambda_j,1)=\frac{1}{T^{\frac{1}{2}}-1}\sum_{t\in T_i(l,\lambda_j,1)}\mathbf{\Omega}_{i,t}^G. \tag{57}$$

Player j infers $\lambda_j(l+1)(i)$ from the supplemental round 1 if and only if

$$\begin{aligned}&\left\|\Omega_i(G)(l,\lambda_j,1)-H_j(G)Q_{j,i}(a_j^G,a_i^G)\mathbf{y}_i(l,\lambda_j,1)\right\|\\=&\ \frac{1}{T^{\frac{1}{2}}-1}\left\|\sum_{t\in T_i(l,\lambda_j,1)}\mathbf{\Omega}_{i,t}(G)-\sum_{t\in T_i(l,\lambda_j,1)}H_j(G)Q_{j,i}(a_j^G,a_i^G)\mathbf{1}_{y_{i,t}}\right\|\leq\frac{\varepsilon}{2},\end{aligned} \tag{58}$$

$$\begin{aligned}&\left\|\Omega_i(B)(l,\lambda_j,1)-H_j(B)Q_{j,i}(a_j^B,a_i^G)\mathbf{y}_i(l,\lambda_j,1)\right\|\\=&\ \frac{1}{T^{\frac{1}{2}}-1}\left\|\sum_{t\in T_i(l,\lambda_j,1)}\mathbf{\Omega}_{i,t}(B)-\sum_{t\in T_i(l,\lambda_j,1)}H_j(B)Q_{j,i}(a_j^B,a_i^G)\mathbf{1}_{y_{i,t}}\right\|\leq\frac{\varepsilon}{2},\end{aligned} \tag{59}$$

$$
\begin{aligned}
& \left\|\Omega_i^G(l,\lambda_j,1) - H_i(G)\mathbf{y}_i(l,\lambda_j,1)\right\| \\
= \; & \frac{1}{T^{\frac{1}{2}}-1}\left\|\sum_{t\in T_i(l,\lambda_j,1)} \boldsymbol{\Omega}_{i,t}^G - \frac{1}{T^{\frac{1}{2}}}\sum_{t\in T_i(l,\lambda_j,1)} H_i(G)\mathbf{1}_{y_{i,t}}\right\| \leq \frac{\varepsilon}{2},
\end{aligned} \tag{60}
$$

and

$$
\left\|\Omega_i^G(l,\lambda_j,1) - p_i(G)\right\| \leq \frac{\varepsilon}{2}, \tag{61}
$$

There are following cases:

1. (58), (59) or (60) are not satisfied. Let $\zeta_i(l,\lambda_j,1) = B$ denote this event. This case is excluded from player j's consideration. From (54), (55), (56) and (57), by the central limit theorem, conditional on $\{y_{i,t}\}_{t\in T(l,\lambda_j,1)}$, this occurs with probability no more than of order $\exp(-T^{\frac{1}{2}})$.

2. (58), (59) and (60) are satisfied. Let $\zeta_i(l,\lambda_j,1) = G$ denote this event.

   (a) (61) is not satisfied. Let $\vartheta_i(l,\lambda_j,1) = B$ denote this event. This case is also excluded from player j's consideration. Since we take the affine hull with respect to player j's action in the definition of $\mathcal{H}_i(G)$ (this is symmetric to (49)), together with (56), the central limit theorem implies that this occurs with probability no more than of order $\exp(-T^{\frac{1}{2}})$ ex ante at the beginning of the supplemental round 1 regardless of player j's strategy.

   (b) (61) is satisfied. Let $\vartheta_i(l,\lambda_j,1) = G$ denote this event. In this case, player i uses the supplemental round 1 to infer $\lambda_j(l+1)$.

   For the following classification, define $\bar{K}$ as follows: Since a continuous linear transformation is Lipschitz continuous, we can take $\bar{K}$ such that, for any i and $\hat{\lambda}_j \in \{G,B\}$, if $\mathbf{y}_i \in \mathcal{H}_{j,i}[\varepsilon](\hat{\lambda}_j)$, then there exists $\boldsymbol{\varepsilon} \in \mathbb{R}_+^{|Y_j|-|A_i|+1}$ such that

$$
\begin{aligned}
H_j(\hat{\lambda}_j)Q_{j,i}(a_j^{\hat{\lambda}_j}, a_i^G)\mathbf{y}_i &= \mathbf{p}_j(\hat{\lambda}_j) + \boldsymbol{\varepsilon}, \\
\|\boldsymbol{\varepsilon}\| &\leq (\bar{K}-1)\varepsilon.
\end{aligned} \tag{62}
$$

Given this $\bar{K}$, we consider following subsubcases.

i. If we have

$$\|\Omega_i(G)(l,\lambda_j,1) - \mathbf{p}_j(G)\| \le \bar{K}\varepsilon, \tag{63}$$

then player i infers $\lambda_j(l+1)(i) = G$.

ii. If we have

$$\|\Omega_i(B)(l,\lambda_j,1) - \mathbf{p}_j(B)\| \le \bar{K}\varepsilon, \tag{64}$$

then player i infers $\lambda_j(l+1)(i) = B$.

iii. Otherwise, player i infers $\lambda_j(l+1)(i) = B$.

In addition, if 1 is the case, we define $\vartheta_i(l,\lambda_j,1) \in \{G,B\}$ as in the case with $\zeta_i(l,\lambda_j,1) = G$: If (61) is not satisfied, then $\vartheta_i(l,\lambda_j,1) = B$. If (61) is satisfied, then $\vartheta_i(l,\lambda_j,1) = G$.

The following lemma shows that the above inferences are well defined. The intuition is the same as one in Section 9.2.1.

Lemma 8 *Generically, the following statement is true: There exists $\bar{\varepsilon} > 0$ such that, for all $\varepsilon < \bar{\varepsilon}$, if Case 2-(a)-i above is the case, then Case 2-(a)-ii is not the case. Symmetrically, if Case 2-(a)-ii above is the case, then Case 2-(a)-i is not the case.*

Proof. (60) and (61) imply that there exists $\boldsymbol{\varepsilon}_1$ with

$$\begin{aligned} H_i(G)\mathbf{y}_i &= \mathbf{p}_i(G) + \boldsymbol{\varepsilon}_1, \\ \|\boldsymbol{\varepsilon}_1\| &\le \varepsilon. \end{aligned}$$

(58) and (63) imply that there exists $\boldsymbol{\varepsilon}_2$ with

$$\begin{aligned} H_j(G)Q_{j,i}(a_j^G, a_i^G)\mathbf{y}_i &= \mathbf{p}_j(G) + \boldsymbol{\varepsilon}_2, \\ \|\boldsymbol{\varepsilon}_2\| &\le (\bar{K}+1)\varepsilon. \end{aligned}$$

On the other hand, (59) and (64) imply that there exists $\boldsymbol{\varepsilon}_3$ with

$$\begin{aligned} H_j(B)Q_{j,i}(a_j^B, a_i^G)\mathbf{y}_i &= \mathbf{p}_j(B) + \boldsymbol{\varepsilon}_3, \\ \|\boldsymbol{\varepsilon}_3\| &\leq (\bar{K}+1)\varepsilon. \end{aligned}$$

Therefore, it suffices to show that, for sufficiently small e, for any $\|\boldsymbol{\varepsilon}\| \leq e$, there does not exist $\mathbf{y}_i \in \mathbb{R}_+^{|Y_i|}$ such that

$$\begin{bmatrix} H_i(G) \\ H_j(G)Q_{j,i}(a_j^G, a_i^G) \\ H_j(B)Q_{j,i}(a_j^B, a_i^G) \end{bmatrix} \mathbf{y}_i = \begin{bmatrix} \mathbf{p}_i(G) \\ \mathbf{p}_j(G) \\ \mathbf{p}_j(B) \end{bmatrix} + \boldsymbol{\varepsilon}. \tag{65}$$

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

By Farkas Lemma, this nonexistence of $\mathbf{y}_i$ for $\boldsymbol{\varepsilon} = \mathbf{0}$ is equivalent to the existence of $\mathbf{x} \in \mathbb{R}_+^{|Y_i|+2|Y_j|-|A_j|-2|A_i|+1}$ with

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

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

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

Again, by Farkas Lemma, (65) does not have a solution for sufficiently small e. ■

See the lower half of Figure 13. Case 1 above is Box 7 of Figure 13. Case 2-(a) is Box 12.

Case 2-(b)-i corresponds to Box 9, 2-(b)-ii corresponds to Box 10, and 2-(b)-iii corresponds to Box 11.

Figure 13 corresponds to Figure 11 in the intuitive explanation. Box 2 of Figure 11 corresponds to Box 9 in Figure 13, Box 3 in Figure 11 corresponds to Box 10 of Figure 13, and Box 4 of Figure 11 corresponds to Box 11 of Figure 13. The cases where player i answers "No" to the question in Box 1 in Figure 11 are included in Boxes 7 and 12 in Figure 13.

[Insert Figure 13]

Notice that whenever player i infers $\hat{\lambda}_j(l+1) = B$, then at least one of $\zeta_i(l) = B$, $\theta_i(l) = B$, $\zeta_i(l, \lambda_j, 1) = B$ and $\vartheta_i(l, \lambda_j, 1) = B$ happens. Reversing the roles of i and j, whenever player j takes $a_j(l+1) \neq a(x_j)$ (this implies $\hat{\lambda}_i(l+1) = B$), any action profile is optimal to player i.

Finally, we show that Condition 2 in Section 9.1 is satisfied.

Lemma 9 *For any $\varepsilon \in (0, \bar{\varepsilon})$, for sufficiently large T, for any $i \in I$ and $\lambda_j(l+1) \in \{G, B\}$,*

1. *Conditional on $\lambda_j(l+1)$ and $\zeta_j(l) = \zeta_j(l, \lambda_j, 1) = G$, player i believes that $\lambda_j(l+1)(i) = \lambda_j(l+1)$ or $\theta_j(l, \lambda_j, 1) = B$ with probability no less than $1 - \exp(-T^{\frac{2}{5}})$.*

2. *Conditional on $\lambda_j(l+1)$ and $\zeta_j(l) = \zeta_j(l, \lambda_j, 1) = G$, any history of player i can happen with probability at least $\exp(-T^{\frac{2}{5}})$.*

3. *Conditional on $\lambda_j(l+1)$ and $\zeta_i(l) = \zeta_i(l, \lambda_j, 1) = \vartheta_i(l, \lambda_j, 1) = G$, conditional on any history in the supplemental rounds, player j believes that any $\lambda_j(l+1)(i)$ is possible with probability no less than $\exp(-T^{\frac{2}{5}})$.*

4. *$\zeta_j(l, \lambda_j, 1) = \theta_j(l, \lambda_j, 1) = \zeta_i(l, \lambda_j, 1) = \vartheta_i(l, \lambda_j, 1) = G$ with probability no more than $\exp(-T^{\frac{2}{5}})$.*

5. *The distribution of $\zeta_j(l, \lambda_j, 1)$ is independent of player i's strategy with probability no less than $1 - \exp(-T^{\frac{2}{5}})$.*

6. *The distribution of* $\theta_j(l, \lambda_j, 1)$ *is independent of player i's strategy.*

7. *The distribution of* $\zeta_i(l, \lambda_j, 1)$ *is independent of player j's strategy with probability no less than* $1 - \exp(-T^{\frac{2}{5}})$.

8. *The distribution of* $\vartheta_i(l, \lambda_j, 1)$ *is independent of player j's strategy.*

Notice that, as we have mentioned, player i excludes the case with $\zeta_j = B$ in 1 and 2 of Lemma 9 and player j excludes the case with $\zeta_i = B$ or $\vartheta_i = B$ in 3 of Lemma 9.

Proof.

1. Suppose $\lambda_j(l+1) = \bar{\lambda}_j$. If $\lambda_j(l+1)(i) = \bar{\lambda}_j$, then we are done. So, let us concentrate on $\lambda_j(l+1)(i) \neq \bar{\lambda}_j$. Forget about the conditioning on $\zeta_j(l, \lambda_j, 1) = G$ for a while. If $\lambda_j(l+1)(i) \neq \bar{\lambda}_j$ but player i uses the supplemental round 1 to infer $\lambda_j(l+1)$, then

$$\left\|\Omega_i(\bar{\lambda}_j)(l, \lambda_j, 1) - \mathbf{p}_j(\bar{\lambda}_j)\right\| > \bar{K}\varepsilon$$

and player i has (58) and (59). Therefore, by triangle inequality,

$$\left\|H_j(\bar{\lambda}_j)Q_{j,i}(a_j^{\bar{\lambda}_j}, a_i^G)\mathbf{y}_i(l, \lambda_j, 1) - \mathbf{p}_j(\bar{\lambda}_j)\right\| > \left(\bar{K} - 1\right)\varepsilon.$$

From (62), this implies that $\mathbf{y}_i(l, \lambda_j, 1) \notin \mathcal{H}_{j,i}[\varepsilon](\bar{\lambda}_j)$, which means that any $\mathbf{y}_j \in \Delta\left(\{\mathbf{1}_{y_j}\}_{y_j \in Y_j}\right)$ within ε from the conditional expectation of $\frac{1}{T^{\frac{1}{2}}-1}\sum_{t \in T_i(l,\lambda_j,1)} \mathbf{1}_{y_j}$ conditional on the true message $\lambda_j(l+1) = \bar{\lambda}_j$ is not included in $\mathcal{H}_j[\varepsilon](\bar{\lambda}_j)$. Since $T_i(l, \lambda_j, 1)$ and $T_j(l, \lambda_j, 1)$ differ at most for two periods, this implies that any $\mathbf{y}_j \in \Delta\left(\{\mathbf{1}_{y_j}\}_{y_j \in Y_j}\right)$ within $\frac{\varepsilon}{2}$ from player i's conditional expectation of $\mathbf{y}_j(l, \lambda_j, 1)$ conditional on the true message $\lambda_j(l+1) = \bar{\lambda}_j$ is not included in $\mathcal{H}_j[\varepsilon](\bar{\lambda}_j)$. Since $\mathbf{y}_j(l, \lambda_j, 1) \in \mathcal{H}_j[\varepsilon](\bar{\lambda}_j)$ is a necessary condition for $\theta_j(l, \lambda_j, 1) = G$, player i puts the belief no less than $\exp(-T^{\frac{3}{7}})$ on $\theta_j(l, \lambda_j, 1) = G$ by Hoeffding's inequality.

Since $\Pr(\zeta_j(l, \lambda_j, 1) = G \mid \{y_{j,t}\}_t) \geq 1 - \exp(-T^{\frac{3}{7}})$ for all $\{y_{j,t}\}_t$, we have

$$\begin{aligned}
\Pr(\{y_{j,t}\}_t \quad &| \quad \zeta_j(l, \lambda_j, 1) = G, \{y_{i,t}\}_t) \\
&= \frac{\Pr(\zeta_j(l, \lambda_j, 1) = G \mid \{y_{i,t}\}_t, \{y_{j,t}\}_t) \Pr(\{y_{j,t}\}_t \mid \{y_{i,t}\}_t)}{\Pr(\zeta_j(l, \lambda_j, 1) = G \mid \{y_{i,t}\}_t)} \\
&= \frac{\Pr(\zeta_j(l, \lambda_j, 1) = G \mid \{y_{j,t}\}_t) \Pr(\{y_{j,t}\}_t \mid \{y_{i,t}\}_t)}{\sum_{\{y_{j,t}\}_t} \Pr(\zeta_j(l, \lambda_j, 1) = G \mid \{y_{j,t}\}_t) \Pr(\{y_{j,t}\}_t \mid \{y_{i,t}\}_t)} \\
&\in \begin{bmatrix} (1 - 2\exp(-T^{\frac{3}{7}})) \Pr(\{y_{j,t}\}_t \mid \{y_{i,t}\}_t), \\ (1 + 2\exp(-T^{\frac{3}{7}})) \Pr(\{y_{j,t}\}_t \mid \{y_{i,t}\}_t) \end{bmatrix}. \qquad (66)
\end{aligned}$$

Hence, even after conditioning on $\zeta_j(l, \lambda_j, 1) = G$, player i believes $\theta_j(l, \lambda_j, 1) = G$ with probability at most $\exp(-T^{\frac{2}{5}})$. In addition, conditional on $\lambda_j(l+1)$, $\zeta_j(l)$ is independent of the supplemental rounds. Therefore, we are done.

2. Suppose $\zeta_j(l, \lambda_j, 1) = B$ never happens. Conditional on $\lambda_j(l+1) = \bar{\lambda}_j$, any $(y_{i,t})_{t \in T(l, \lambda_j, 1)}$ can occur with probability at least

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

Hence, for

$$0 < e < \frac{1}{-\log\left\{\min_{y_j, y_i, a} q(y_i \mid a, y_j)\right\}},$$

any history of player i can happen.

By the symmetric proof to (66), conditioning on $\zeta_j(l, \lambda_j, 1) = G$ does not change the probability so much. Again, $\zeta_j(l)$ is independent of the supplemental rounds.

3. Suppose $\zeta_i(l, \lambda_j, 1) = B$ never happens. Conditional on $(a_t, y_{j,t})_{t \in T(l, \lambda_j, 1)}$, any $\mathbf{y}_i(l, \lambda_j, 1)$ can occur with probability at least

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

Further, for any $\hat{\lambda}_j \in \{G, B\}$, for $\mathbf{y}_i(l, \lambda_j, 1)$ sufficiently close to $\mathbf{q}(a_j^{\hat{\lambda}_j}, a_i^G)$, (61) and (63) are satisfied and $\lambda_j(l+1)(i) = \hat{\lambda}_j$. Hence, for

$$0 < e < \frac{1}{-\log\left\{\min_{y_j, y_i, a} q(y_i \mid a, y_j)\right\}},$$

$\lambda_j(l+1)(i) = \hat{\lambda}_j$ can happen together with (61) (and so $\vartheta_i(l, \lambda_j, 1) = G$) with probability no less than $\exp(-\frac{1}{e}T^{\frac{1}{2}})$.

By the symmetric proof to (66), conditioning on $\zeta_i(l, \lambda_j, 1) = G$ does not change the belief so much. Again, $\zeta_i(l)$ is independent of the supplemental rounds.

4 to 8 Follows from Hoeffding's inequality and the fact that we take the affine hull with respect to a_j for the definition of $\mathcal{H}_i(G)$.

■

15.4.2 Explanation of 2 of Lemma 5: Supplemental Round 2 for $\lambda_j(l+1)$

The following is the formal description of the message exchanges in the supplemental round 1 for $\lambda_j(l+1)$.

Strategies Player j (sender) with the message $\lambda_j(l+1) = \bar{\lambda}_j \in \{G, B\}$ determines $z_j(\bar{\lambda}_j) \in \{G, B, M\}$ as follows:

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

Here, η is determined in Section 10. Player j with $z_j(\bar{\lambda}_j)$ takes

$$\alpha_j^{z_j(\bar{\lambda}_j)} = \begin{cases} a_j^G & \text{if } z_j(\bar{\lambda}_j) = G, \\ a_j^B & \text{if } z_j(\bar{\lambda}_j) = B, \\ \frac{1}{2}a_j^G + \frac{1}{2}a_j^B & \text{if } z_j(\bar{\lambda}_j) = M \end{cases}$$

for $T^{\frac{1}{2}}$ periods in the supplemental round 2 for $\lambda_j(l+1)$, that is, for $t \in T(l, \lambda_j, 2)$. Player i (receiver) takes a_i^G.

Requirement of Player j Player j who has had $\zeta_j(l, \lambda_j, 1) = B$ or $\theta_j(l, \lambda_j, 1) = B$ in the supplemental round 1 has already made player i indifferent between any action profile in the subsequence rounds. Player j who has $\zeta_j(l, \lambda_j, 1) = \theta_j(l, \lambda_j, 1) = G$ requires player i to infer $\lambda_j(l+1)$ correctly if and only if $z_j(\bar{\lambda}_j) = \bar{\lambda}_j$. $\theta_j(l+1) = B$ if $z_j(\bar{\lambda}_j) \neq \bar{\lambda}_j$.

See Figure 12 again. Notice that Box 9 is the only place where player j requires player i to infer $\lambda_j(l+1)$ correctly and it happens only if $\zeta_j(l, \lambda_j, 1) = G$ and $z_j(\bar{\lambda}_j) = \bar{\lambda}_j$.

Inference of Player i If player i has used the supplemental round 1 to infer $\lambda_j(l+1)$, player i is stick to that inference. Otherwise, player i infers $\lambda_j(l+1)(i)$ using the supplemental round 2. Remember that player i uses the supplemental round 2 only if $\zeta_i(l, \lambda_j, 1) = B$ or $\vartheta_i(l, \lambda_j, 1) = B$ and player j (sender) excludes these cases from consideration (see Figure 13). Therefore, player j does not care about player i's inference in the supplemental round 2.

Based on the signal observations in the supplemental round 2, $\{y_{i,t}\}_{t \in T(l, \lambda_j, 2)}$, the belief of player i can be classified into the following three cases:

Lemma 10 *For each $\bar{\lambda}_j \in \{G, B\}$, conditional on $\lambda_j(l+1) = \bar{\lambda}_j$, one of the following three is correct:*

1. *the likelihood ratio of $z_j(\bar{\lambda}_j) = G$ compared to $z_j(\bar{\lambda}_j) = B$ is no less than* $\exp(T^{\frac{3}{7}})$.

2. *the likelihood ratio of $z_j(\bar{\lambda}_j) = B$ compared to $z_j(\bar{\lambda}_j) = G$ is no less than* $\exp(T^{\frac{3}{7}})$.

3. *the likelihood ratio of $z_j(\bar{\lambda}_j) = M$ compared to $z_j(\bar{\lambda}_j) = G, B$ is no less than* $\exp(T^{\frac{3}{7}})$.

If 1 of Lemma 10 is the case, then player i infers $\lambda_j(l+1)(i) = G$. This is almost optimal since we condition on $\lambda_j(l+1) = \bar{\lambda}_j \in \{G, B\}$ and player i is indifferent between any action if $z_j(\bar{\lambda}_j) \neq \bar{\lambda}_j$. If 2 of Lemma 10 is the case, then player i infers $\lambda_j(l+1)(i) = B$. This is almost optimal by the same reason. If 3 of Lemma 10 is the case, then player i infers $\lambda_j(l+1)(i) = B$. In this case, $z_j(\bar{\lambda}_j) = M \neq \bar{\lambda}_j$ is highly likely and player i is indifferent to any action. Therefore, this is also almost optimal.

Proof. Conditional on $\lambda_j(l+1) = \bar{\lambda}_j$,

$$\frac{\Pr\left(z_j(\bar{\lambda}_j) = z_j \mid \bar{\lambda}_j, \{y_{i,t}\}_{t \in T(l,\lambda_j,2)}\right)}{\Pr\left(z_j(\bar{\lambda}_j) = z_j' \mid \bar{\lambda}_j, \{y_{i,t}\}_{t \in T(l,\lambda_j,2)}\right)} = \frac{\Pr\left(\{y_{i,t}\}_{t \in T(l,\lambda_j,2)} \mid z_j(\bar{\lambda}_j) = z_j\right)}{\Pr\left(\{y_{i,t}\}_{t \in T(l,\lambda_j,2)} \mid z_j(\bar{\lambda}_j) = z_j'\right)} \frac{\Pr\left(z_j(\bar{\lambda}_j) = z_j \mid \bar{\lambda}_j\right)}{\Pr\left(z_j(\bar{\lambda}_j) = z_j' \mid \bar{\lambda}_j\right)}$$

and $\frac{\Pr\left(z_j(\bar{\lambda}_j) = z_j | \bar{\lambda}_j\right)}{\Pr\left(z_j(\bar{\lambda}_j) = z_j' | \bar{\lambda}_j\right)}$ is bounded by $\left[\frac{\eta}{1-2\eta}, \frac{1-2\eta}{\eta}\right]$. With $\mathbf{y}_i(l, \lambda_j, 2)$ being the frequency of each y for $\{y_{i,t}\}_{t \in T(l,\lambda_j,2)}$, $\log \Pr\left(\{y_{i,t}\}_{t \in T(l,\lambda_j,2)} \mid z_j(\bar{\lambda}_j) = z_j\right)$ is expressed as $T^{\frac{1}{2}} \mathcal{L}(\mathbf{y}_i(l, \lambda_j, 2), z_j)$ with

$$\mathcal{L}(\mathbf{y}_i(l, \lambda_j, 2), z_j) = y_{i,1}(l, \lambda_j, 2) \log q(y_{i,1}|a_i^G, \alpha_j^{z_j}) + \cdots + y_{i,|Y_i|}(l, \lambda_j, 2) \log q(y_{i,|Y_i|}|a_i^G, \alpha_j^{z_j}).$$

Hence, it suffices to show that there exists $\tilde{\kappa}$ such that, with

$$\mathcal{L}(\mathbf{f}_i, z_j, z_j') \equiv \mathcal{L}(\mathbf{f}_i, z_j) - \mathcal{L}(\mathbf{f}_i, z_j'),$$

for any $\mathbf{f}_i \in \Delta\left(\{\mathbf{1}_{y_i}\}_{y_i \in Y_i}\right)$, one of the following is true:

1. $z_j(\bar{\lambda}_j) = G$ is more likely than $z_j(\bar{\lambda}_j) = B$: $\mathcal{L}(\mathbf{f}_i, G, B) \geq \tilde{\kappa}$,
2. $z_j(\bar{\lambda}_j) = B$ is more likely than $z_j(\bar{\lambda}_j) = G$: $\mathcal{L}(\mathbf{f}_i, B, G) \geq \tilde{\kappa}$,
3. $z_j(\bar{\lambda}_j) = M$ is more likely than $z_j(\bar{\lambda}_j) = G, B$: $\mathcal{L}(\mathbf{f}_i, M, G) \geq \tilde{\kappa}$ and $\mathcal{L}(\mathbf{f}_i, M, B) \geq \tilde{\kappa}$.

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

$$q(y_{i,k}|a_i^G, \alpha_j^G) \neq q(y_{i,k}|a_i^G, \alpha_j^B). \tag{67}$$

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

$$g(\mathbf{f}_i, \lambda) = f_{i,1} \log q(y_{i,1}|a_i^G, \alpha_j^\lambda) + \cdots + f_{i,|Y_i|} \log q(y_{i,|Y_i|}|a_i^G, \alpha_j^\lambda).$$

Then,

$$\frac{d^2 g(\mathbf{f}_i,\lambda)}{d\lambda^2} = -\sum_{k=1}^{|Y_i|} f_{i,k} \left\{ \frac{q(y_{i,k}|a_i^G, \alpha_j^G) - q(y_{i,k}|a_i^G, \alpha_j^B)}{q(y_{i,k}|a_i^G, \alpha_j^\lambda)} \right\}^2 < 0$$

for any $f_{i,k}$ because of (67). Hence, $g(f_i, \lambda)$ is strictly concave. Therefore, since $\mathcal{L}(\mathbf{f}_i, z_j, \tilde{z}_j)$ is the difference in $g(\mathbf{f}_i, \lambda)$, one of the following is true:

1. $z_j(\bar{\lambda}_j) = G$ is more likely than $z_j(\bar{\lambda}_j) = B$: $\mathcal{L}(\mathbf{f}_i, G, B) > 0$,

2. $z_j(\bar{\lambda}_j) = B$ is more likely than $z_j(\bar{\lambda}_j) = G$: $\mathcal{L}(\mathbf{f}_i, B, G) > 0$,

3. $z_j(\bar{\lambda}_j) = M$ is more likely than $z_j(\bar{\lambda}_j) = G, B$: $\mathcal{L}(\mathbf{f}_i, M, G) > 0$ and $\mathcal{L}(\mathbf{f}_i, M, B) > 0$.

Hence,

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

Since LHS is continuous in $\mathbf{f}_i$ and $\Delta\left(\{\mathbf{1}_{y_i}\}_{y_i \in Y_i}\right)$ is compact, there exists $\tilde{\kappa} > 0$ such that

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

for all $\mathbf{f}_i \in \Delta\left(\{\mathbf{1}_{y_i}\}_{y_i \in Y_i}\right)$ as desired. ■

15.4.3 Proof of 3-(a) to 3-(f)

Follows from Lemma 9 and conditional independence of each round.

15.5 Proof of Lemma 6

Since $\hat{\lambda}_j(l) = \lambda_j(l)$ always holds for $\sigma_i(B)$, we concentrate on $\sigma_i(G)$.

Since once $\lambda_j(\tilde{l}) = B$ is induced, then $\lambda_j(\tilde{l}') = B$ for all the following rounds, there exists a unique l^* such that $\lambda_j(\tilde{l}) = B$ is initially induced in the $(l^* + 1)$th review round: $\lambda_j(1) = \cdots = \lambda_j(l^*) = G$ 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 that player i in the supplemental rounds 1 and 2 for $\lambda_j(\hat{l}^* + 1)$ inferred $\lambda_j(\hat{l}^* + 1)(i) = B$. Then, for any $l \geq \hat{l}^* + 1$, by 1 of Lemma 5, player i believes that conditional on $\zeta_j(\tilde{r}) = \vartheta_j(\tilde{r}) = G$ for all $r < l$, player i in the lth review round believes that $\lambda_j(\hat{l}^* + 1)(i) = \lambda_j(\hat{l}^* + 1)$ or "$\theta_j(\hat{l}^*, \lambda_j, 1) = B$ or $\theta_j(\hat{l}^*, \lambda_j, 2) = B$" with probability no less than $1 - \exp(-T^{\frac{1}{3}})$ as desired.
- $l^* < \hat{l}^*$: There are following two cases:
  - 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 that player i has not deviated until the end of the main round of the l^*th block). This means that 3 of Lemma 4 is applicable and player i at the end of the l^*th review round believes that $\lambda_j(l^* + 1) = G$ with probability no less than $1 - \exp(-T^{\frac{4}{5}})$.

    As we have mentioned in Section 9.1, player j's continuation strategy reveals $\lambda_j(l+1)$ through (i) the strategy in the supplemental rounds and (ii) $\hat{\lambda}_i(l + 1)$. From 2 of Lemma 5, conditional on $\lambda_j(l + 1)$ and $\zeta_j(l^*) = \zeta_j(l^*, \lambda_j, 1) = \zeta_j(l^*, \lambda_i, 1) = \vartheta_j(l^*, \lambda_i, 1) = G$, any history of player i can happen in the supplemental rounds with probability no less than $\exp(-T^{\frac{2}{5}})$. Hence, the update from (i) is bounded by $\exp(T^{\frac{2}{5}})$. In addition, for $\hat{\lambda}_i(l + 1)$, player j is ready to listen with probability at least η and 3 of Lemma 5 (with the roles of i and j being reversed) implies that any $\lambda_i(l+1)(j)$ happens with probability no less than $\exp(-T^{\frac{2}{5}})$. Hence, the update from (ii) is bounded by $\frac{1}{\eta}\exp(T^{\frac{2}{5}})$. Therefore, after observing player j's continuation strategy, player i believe $\lambda_j(l+1) = B$ with probability at most

$$\frac{\exp(-T^{\frac{4}{5}})\frac{1}{\eta}\exp(2T^{\frac{2}{5}})}{1-\exp(-T^{\frac{4}{5}})} < \exp(-T^{\frac{1}{3}}) \tag{68}$$

as desired.

15.6 Proof of Proposition 1

For (6), it suffices to have

$$\bar{\pi}_i(x,l,\lambda_j) \begin{cases} \leq 0 & \text{if } x_j = G, \\ \geq 0 & \text{if } x_j = B, \end{cases} \tag{69}$$

$$|\bar{\pi}_i(x,l,\lambda_j)| \leq \max_{i,a} 2|u_i(a)|T \tag{70}$$

for all $x \in \{G,B\}^2$, $l \in \{1,...,L\}$ and $\lambda_j \in \{G,B\}$.

To see why (69) and (70) are sufficient for (6), notice the following: (70) implies uniform boundedness. (69) implies that $\pi_i(x_j, h_j^{T_P+1}, l) > -\rho T$ with $x_j = G$ or $\pi_i(x_j, h_j^{T_P+1}, l) < \rho T$ with $x_j = B$ only if $\lambda_j(l) = G$ and $X_j(l) \notin [q_2T - 2\varepsilon T, q_2T + 2\varepsilon T]$. Since we have $\lambda_j(\tilde{l}) = B$ for $\tilde{l} > l$ after those events from (28), $\pi_i(x_j, h_j^{T_P+1}, l) \leq -\rho T$ with $x_j = G$ or $\pi_i(x_j, h_j^{T_P+1}, l) \geq \rho T$ with $x_j = B$ except for one round. In addition, (69) implies $-\bar{L}T - \rho T \leq \pi_i(x_j, h_j^{T_P+1}, l) \leq \bar{L}T + \rho T$ for all x_j. Hence, in total, $\sum_{l=1}^{L} \pi_i(x_j, h_j^{T_P+1}, l) \leq \bar{L}T - L\rho T < 0$ with $x_j = G$ and $\pi_i(x_j, h_j^{T_P+1}, l) \geq -\bar{L}T + L\rho T > 0$ with $x_j = B$. The strict inequalities follow from (11). Therefore, (6) is satisfied.

Therefore, we prove Proposition 1 with (6) replaced with (69) and (70).

1 follows from (32). 2-(a) is true since $\zeta_j(\tilde{r}) = \vartheta_j(\tilde{r}) = G$ for all $\tilde{r} < r$ imply that player j has inferred $\lambda_i(l+1)(j)$ from the supplemental round 1 (See Figure 13 reversing the roles of i and j). 2-(b) is true since taking a_i^G gives the almost optimal inference and (35) gives a high reward on a_i^G.

Therefore, we are left to show the existence of $\bar{\pi}_i(x, l, \lambda_j)$ with (69) and (70) and 2-(c) by backward induction. From Lemmas 4 and 5, the distribution of ζ_j, ϑ_j and θ_j is independent of player i's strategy with high probability and so we can neglect the effect of player i's strategy on ζ_j, ϑ_j and θ_j.

In the Lth review round, consider the case with $x_i = B$ first. If player j uses (33), then D_i is strictly optimal. If player j uses (32), then any action is optimal.

Consider the case with $x_i = G$ next. If player j uses (33) and $\hat{\lambda}_j(L) = \lambda_j(L) = G$, then C_i is strictly optimal for sufficiently large T since the reward (33) is increasing in $X_j(L)$ and (42) implies that the marginal expected increase in $\bar{L}X_j(L)$ is sufficiently large.[18] If player j uses (33) and $\hat{\lambda}_j(L) = \lambda_j(L) = B$, then D_i is strictly optimal since the reward (33) is constant. If player j uses (32), then any action is optimal. Hence, $\sigma_i(x_i)$ is optimal for these cases.

For the other cases, by Lemma 6, player i does not have posteriors more than $\exp(-T^{\frac{1}{3}})$. Since the per-period difference of the payoff from two different strategies is bounded by $\bar{U} \equiv \bar{L} + \max_{i,a} 2\,|u_i(a)|$, the expected loss from $\sigma_i(x_i)$ (taking C_i if $\hat{\lambda}_j(L) = G$ and D_i if $\hat{\lambda}_j(L) = B$) is no more than $\exp(-T^{\frac{1}{3}})\bar{U}T \leq \exp(-T^{\frac{1}{4}})$ for sufficiently large T.

Therefore, in total, $\sigma_i(x_i)$ is almost optimal in the Lth review round.

Further, if player j uses (33) and $\hat{\lambda}_j(L) = \lambda_j(L)$, then player i's average continuation payoff at the beginning of the Lth review round except for $\bar{\pi}_i(x, L, \lambda_j)$ is

$$
\begin{array}{ll}
u_i(D_i, C_j) - \rho & \text{if } x_i = B, x_j = G, \\
u_i(D_i, D_j) + \rho & \text{if } x_i = B, x_j = B, \\
u_i(C_i, C_j) - \rho - 2\varepsilon\bar{L} & \text{if } x_i = G, x_j = G, \hat{\lambda}_j(L) = \lambda_j(L) = G, \\
u_i(C_i, D_j) + \rho + 2\varepsilon\bar{L} & \text{if } x_i = G, x_j = B, \hat{\lambda}_j(L) = \lambda_j(L) = G, \\
u_i(D_i, C_j) - \rho & \text{if } x_i = G, x_j = G, \hat{\lambda}_j(L) = \lambda_j(L) = B, \\
u_i(D_i, D_i) + \rho & \text{if } x_i = G, x_j = B, \hat{\lambda}_j(L) = \lambda_j(L) = B.
\end{array}
\tag{71}
$$

Hence, there exists $\bar{\pi}_i(x, L, \lambda_j)$ with (69) and (70) such that player i's average continuation

[18] For large T, the effect of dropping one $t_j(l)$ is negligible.

payoff is equal to $u_i(C_i, C_j) - \rho - 2\varepsilon \bar{L}$ if $x_j = G$ and $u_i(D_i, D_j) + \rho + 2\varepsilon \bar{L}$ if $x_j = B$.

In the supplemental rounds for $\lambda_j(L)$ (where player i receives the message), as we have mentioned for 2-(b) of Proposition 1, $\sigma_i(x_i)$ is exactly optimal.

In the supplemental round 2 for $\lambda_i(L)$, as we have mentioned for 2-(a) of Proposition 1, $\sigma_i(x_i)$ is exactly optimal.

In the supplemental round 1 for $\lambda_i(L)$, since (i) (34) cancels out the difference in the instantaneous utilities, (ii) 4 of Lemma 4 guarantees that player i's continuation payoff is fixed regardless of player j's inference of player i's message if player j uses the inference $\lambda_i(L)(j)$ for $\hat{\lambda}_i(L)$, and (iii) the equilibrium strategy gives the almost optimal inferences,[19] $\sigma_i(x_i)$ is optimal up to the loss of $\exp(-T^{\frac{1}{3}})\{\bar{U}T + \max_{x,\lambda_j} \bar{\pi}_i(x, L, \lambda_j)\} \leq \exp(-T^{\frac{1}{4}})$ for sufficiently large T.

In the $(L-1)$th main round, for the almost optimality, only difference from the Lth review round is that, if $\lambda_j(L-1) = G$, then player i's action in the $(L-1)$th review round can affect the distribution of $\lambda_j(L)$. However, since (i) the miscoordination between $\lambda_j(L)$ and $\hat{\lambda}_j(L)$ will not occur with probability more than $\exp(-T^{\frac{1}{3}})$ by Lemma 6 and (ii) $\bar{\pi}_i(x, L, \lambda_j)$ is determined so that player i's continuation payoff is the same between $\lambda_j(L) = G$ and B if the coordination goes well, this difference is negligible for the almost optimality.

Further, if player j uses (33) and $\hat{\lambda}_j(L-1) = \lambda_j(L-1)$, then player i's average payoff from the $(L-1)$th review round except for $\bar{\pi}_i(x, L-1, \lambda_j)$ is given by (71). The cases where (32) will be used in the Lth review round will happen with probability no more than 3η (player j is ready to listen to the message $\lambda_i(L)$ and $z_j(\bar{\lambda}_j) \neq \lambda_j(L)$) plus some negligible probabilities for $\zeta_j = B$ or $\vartheta_j = B$. When (32) is used, per period payoff is bounded by $[-\bar{u}, \bar{u}]$ by Lemma 2. Therefore, there exists $\bar{\pi}_i(x, L, \lambda_j)$ with (69) and (70) such that player i's average continuation payoff from the $(L-1)$th and Lth review rounds is equal to $u_i(C_i, C_j) - \rho - 2\varepsilon \bar{L} - \frac{3\eta\bar{u}}{2}$ if $x_j = G$ and $u_i(D_i, D_j) + \rho + 2\varepsilon \bar{L} + \frac{3\eta\bar{u}}{2}$ if $x_j = B$. Since the length of the supplemental rounds is much smaller than that of the review rounds, the instantaneous utilities and rewards in the supplemental rounds do not affect the average

[19] See (68) to see how the history in this round can affect the optimality of player i's inferences slightly.

payoff.

Recursively, for $l = 1$, Proposition 1 is satisfied and the average ex ante payoff of player i is $u_i(C_i, C_j) - \rho - 2\varepsilon\bar{L} - \frac{3(L-1)\bar{u}\eta}{L}$ if $x_j = G$ and $u_i(D_i, D_j) + \rho + 2\varepsilon\bar{L} + \frac{3(L-1)\bar{u}\eta}{L}$ if $x_j = B$. From (44), we can further modify $\bar{\pi}_i(x, 1, G)$ with (69) and (70) such that $\sigma_i(x_i)$ gives $\bar{v}_i$ ($\underline{v}_i$, respectively) if $x_j = G$ (B, respectively).[20]

15.7 Formal Construction of the Report Block

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 $\hat{h}_i^{T_P+1}$. In general, when we write a variable in player i's history with "hat," it means player i's message (and with cheap talk, equivalently player j's inference) about that variable.

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

- What state x_j player j is in,
- What action $a_j(l)$ player j took in the lth review round with $l \leq r$, and
- What states $\zeta_j(\tilde{r}), \vartheta_j(\tilde{r}) \in \{G, B\}$ with $\tilde{r} < r$ player j had.

We want to show that $\sigma_i(x_i)$ is exactly optimal in the round r conditional on $\mathcal{A}_j(r)$. Note that $\mathcal{A}_j(r)$ contains x_j and so the equilibrium is belief-free at the beginning of the finitely repeated game.

We introduce the following variables. Let $R_i(r)$ be the set of rounds $\tilde{r} \leq r-1$ that are not a supplemental round 2 for $\lambda_i(l+1)$ (with $R_i(1) = \emptyset$), t_r be the initial period of the round r, and $\mathfrak{h}_i^r$ be the summary of player i's history at the beginning of the round r. $\mathfrak{h}_i^r$ is a collection of $|A_i||Y_i| \times 1$ vectors, one for each $\tilde{r} \in R_i(r)$. The element of the vector for $\tilde{r}$ corresponding to (a_i, y_i) represents how many times player i observed (a_i, y_i) in the round $\tilde{r}$.

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

[20] Remember that $\lambda_j(1) = G$.

- 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 $h_1^{T_P+1}$ with probability $\frac{1}{2}$ and player 2 reports $h_2^{T_P+1}$ with probability $\frac{1}{2}$. Suppose player i is picked by the public randomization device.

- For each round r that is not a supplemental round 2 for $\lambda_i(l+1)$,

  - Player i sends the history in the initial period of the round r: (a_{i,t_r}, y_{i,t_r}).

  - Player i sends the history of the round r: $\{a_{i,t}, y_{i,t}\}_{t \in T(r)}$.

With abuse of notation, we assume the players can send multiple messages sequentially. Note that $\hat{\mathfrak{h}}_i^r$ can be calculated by the messages sent before (a_{i,t_r}, y_{i,t_r}).

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 restore (6) by adding or subtracting a small constant depending on x_j without affecting the incentive.

As a preparation, we prove the following lemma:

Lemma 11 *Let h_i and h_j be player i's and player j's histories at the end of the main blocks, respectively. There generically exist $\bar{\varepsilon} > 0$ and $g_i(h_j, a_i, y_i)$ such that, for sufficiently large T, for any round $r \in R_i(R)$ and period t in $T(r)$, conditional on $\zeta_j(\tilde{r}) = \vartheta_j(\tilde{r}) = G$ with $\tilde{r} < r$, it is better for player i to report $a_{i,t}, y_{i,t}$ truthfully: For all h_i,*

$$\begin{aligned} & \mathbb{E}\left[g_i(h_j, \hat{a}_{i,t}, \hat{y}_{i,t}) \mid \zeta_j(\tilde{r}) = \vartheta_j(\tilde{r}) = G \text{ with } \tilde{r} < r, h_i, (\hat{a}_{i,t}, \hat{y}_{i,t}) = (a_{i,t}, y_{i,t})\right] \qquad (72) \\ > \ & \mathbb{E}\left[g_i(h_j, \hat{a}_{i,t}, \hat{y}_{i,t}) \mid \zeta_j(\tilde{r}) = \vartheta_j(\tilde{r}) = G \text{ with } \tilde{r} < r, h_i, (\hat{a}_{i,t}, \hat{y}_{i,t}) \neq (a_{i,t}, y_{i,t})\right] + \bar{\varepsilon} T^{-1}, \end{aligned}$$

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

Proof. We show

$$g_i(h_j, \hat{a}_{i,t}, \hat{y}_{i,t}) = -\mathbf{1}_{\{t_j(r)=t\}} \left\| \mathbf{1}_{y_{j,t}} - \mathbb{E}[\mathbf{1}_{y_{j,t}} \mid \hat{a}_{i,t}, \hat{y}_{i,t}, a_{j,t}] \right\|^2$$

works.[21,22] To see this, consider the following two cases:

1. If $t_j(r) \neq t$, any report is optimal since $g_i(h_j, \hat{a}_{i,t}, \hat{y}_{i,t}) = 0$.

2. If $t_j(r) = t$, then period t is not used for the construction of the continuation strategy. Hence, player i, after knowing $t_j(r) = t$ and $a_{j,t}$, wants to maximize

$$\max_{\hat{a}_{i,t}, \hat{y}_{i,t}} \mathbb{E}\left[- \left\| \mathbf{1}_{y_{j,t}} - \mathbb{E}[\mathbf{1}_{y_{j,t}} \mid \hat{a}_{i,t}, \hat{y}_{i,t}, a_{j,t}] \right\|^2 \mid a_{i,t}, y_{i,t}, a_{j,t}\right].$$

   The first order condition is

$$\mathbb{E}\left[\mathbf{1}_{y_{j,t}} \mid a_{i,t}, y_{i,t}, a_{j,t}\right] = \mathbb{E}[\mathbf{1}_{y_{j,t}} \mid \hat{a}_{i,t}, \hat{y}_{i,t}, a_{j,t}],$$

   which generically implies $(\hat{a}_{i,t}, \hat{y}_{i,t}) = (a_{i,t}, y_{i,t})$, and the second order condition is also satisfied.

We are left to show that there exists $\varepsilon > 0$ such that, for any h_i, $r \in R_i(R)$ and $t \in T(r)$, player i puts belief at least εT^{-1} on $t_j(r) = t$. Suppose player i knows $\{a_{j,\tau}\}_{\tau \in T(r)}$ and $\{y_{j,\tau}, \varphi_{j,\tau}\}_{\tau \in T_j(r)}$ in addition to h_i. Here, $\varphi_{j,t}$ is

- $\mathbf{\Omega}_{j,t}^H$ for t in the round r where player j is the sender of the message by a pure strategy,

- $\emptyset$ (no information) for t in the round r where player j is the sender of the message by a mixed strategy,

- $\mathbf{\Omega}_{j,t}^G$, $\mathbf{\Omega}_{j,t}(G)$ and $\mathbf{\Omega}_{j,t}(B)$ for t in the round r where player j is the receiver of player i's message set by a pure strategy, and

- $\varphi_{j,t} = (\Psi_{j,t}^{a(x)}, (E_j \Psi_i)_t)$ for t in the round r that is a review round.

[21] We use Euclidean norm in Section 15.7.

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

Since $\{a_{j,\tau}, y_{j,\tau}, \varphi_{j,\tau}\}_{\tau \in T_j(r)}$ determines player j's continuation strategy (including ζ_j and ϑ_j), it suffices to show that, conditional on $\{a_\tau, y_{i,\tau}\}_{\tau \in T(r)}$, $\{y_{j,\tau}, \varphi_{j,\tau}\}_{\tau \in T(r)}$ and $\left(y_{j,t_j(r)}, \varphi_{j,t_j(r)}\right) = (\bar{y}_j, \bar{\varphi}_j)$, player i puts belief at least εT^{-1} on $t_j(r) = t$. For any t and $t' \in T(r)$, the likelihood ratio between $t_j(r) = t$ and $t_j(r) = t'$ is given by

$$\begin{aligned}
& \frac{\Pr(t_j(r) = t \mid \{a_\tau, y_{i,\tau}\}_{\tau \in T(r)}, \{y_{j,\tau}, \varphi_{j,\tau}\}_{\tau \in T(r)}, \left(y_{j,t_j(r)}, \varphi_{j,t_j(r)}\right) = (\bar{y}_j, \bar{\varphi}_j))}{\Pr(t_j(r) = t' \mid \{a_\tau, y_{i,\tau}\}_{\tau \in T(r)}, \{y_{j,\tau}, \varphi_{j,\tau}\}_{\tau \in T(r)}, \left(y_{j,t_j(r)}, \varphi_{j,t_j(r)}\right) = (\bar{y}_j, \bar{\varphi}_j))} \\
= & \frac{\Pr(\{y_{j,\tau}, \varphi_{j,\tau}\}_{\tau \in T(r)}, \left(y_{j,t_j(r)}, \varphi_{j,t_j(r)}\right) = (\bar{y}_j, \bar{\varphi}_j) \mid \{a_\tau, y_{i,\tau}\}_{\tau \in T(r)}, t_j(r) = t)}{\Pr(\{y_{j,\tau}, \varphi_{j,\tau}\}_{\tau \in T(r)}, \left(y_{j,t_j(r)}, \varphi_{j,t_j(r)}\right) = (\bar{y}_j, \bar{\varphi}_j) \mid \{a_\tau, y_{i,\tau}\}_{\tau \in T(r)}, t_j(r) = t')} \\
\in & \left[\min_{a, y_i} q(\bar{y}_j, \bar{\varphi}_j \mid a, y_i), \frac{1}{\min_{a, y_i} q(\bar{y}_j, \bar{\varphi}_j \mid a, y_i)}\right] \\
\in & \left[\min_{a, y_i, y_j, \varphi_j} q(y_j, \varphi_j \mid a, y_i), \frac{1}{\min_{a, y_i, y_j, \varphi_j} q(y_j, \varphi_j \mid a, y_i)}\right].
\end{aligned}$$

Since $\min q(y_j, \varphi_j \mid a, y_i) \in (0, 1)$ from Lemma 7, there exists $\varepsilon > 0$ such that

$$\begin{aligned}
& \Pr\left(t_j(r) = t \mid \{a_\tau, y_{i,\tau}\}_{\tau \in T(r)}, \{y_{j,\tau}, \varphi_{j,\tau}\}_{\tau \in T(r)}, \left(y_{j,t_j(r)}, \varphi_{j,t_j(r)}\right) = (\bar{y}_j, \bar{\varphi}_j)\right) \\
> & \ \varepsilon \Pr(t_j(r) = t' \mid \{a_\tau, y_{i,\tau}\}_{\tau \in T(r)}, \{y_{j,\tau}, \varphi_{j,\tau}\}_{\tau \in T(r)}, \left(y_{j,t_j(r)}, \varphi_{j,t_j(r)}\right) = (\bar{y}_j, \bar{\varphi}_j))
\end{aligned}$$

for all t and t'. Since there exists at least one t' with $\Pr(t_j(r) = t' \mid \{a_\tau, y_{i,\tau}\}_{\tau \in T(r)}, \{y_{j,\tau}, \varphi_{j,\tau}\}_{\tau \in T(r)}, \left(y_{j,t_j(r)}, \varphi_{j,t_j(r)}\right) = (\bar{y}_j, \bar{\varphi}_j)) > T^{-1}$, we are done. ■

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

- If we come to the round r with $\zeta_j(\tilde{r}) = B$ or $\vartheta_j(\tilde{r}) = B$ with some $\tilde{r} < r$, then we cancel out all the rewards explained below about the following rounds $\hat{r} \geq r$. Then, since we have established the exact optimality of $\sigma_i(x_i)$ in the round r with $\zeta_j(\tilde{r}) = B$ or $\vartheta_j(\tilde{r}) \in B$ with some $\tilde{r} < r$ without adjustment, any action is exactly optimal after (and including) the round r.

- If the round r is a supplemental round 2 for $\lambda_i(l+1)$, player i does not report the history in that round and the rewards below are all 0 for that round. Since Proposition 1 establishes the exact optimality without adjustment, we can keep the exact optimality.

- Otherwise, we consider the following punishment and reward:
  - (r-j-1) Based on $\hat{\mathfrak{h}}_i^r$, player j gives

$$\sum_{a_i} f(a_i \mid \hat{\mathfrak{h}}_i^r, \mathcal{A}_j(r)) \Psi_{j,t_r}^{a_i,a_{j,t_r}} \tag{73}$$

    with

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

    such that, after $\mathfrak{h}_i^r$, it is optimal to take $a_i \in A_i(\mathfrak{h}_i^r)$. Note that if $\Psi_{j,t_r}^{a_i,a_{j,t_r}} = 1$, that is, if it is likely that player i took a_i at the beginning of the round r, player j rewards player i by $f(a_i \mid \hat{\mathfrak{h}}_i^r, \mathcal{A}_1(r))$ so that player i wants to follow the equilibrium strategy. The existence of such a function f will be verified below.
  - (r-j-2) Player j punishes player i if it is likely for player i to tell a lie about the history in the initial period of the round r by

$$T^{-10r+5} g_i(h_j, \hat{a}_{i,t_r}, \hat{y}_{i,t_r}). \tag{74}$$

  - (r-j-3) Player j makes it optimal to constantly take a_{i,t_r} within the round r by adding

$$T^{-10r+3} \sum_{t \in T(r)} \Psi_{j,t}^{\hat{a}_{i,t_r},a_{j,t}} \tag{75}$$

    for t included in the round r, that is, if player i tells the truth and $\hat{a}_{i,t_r} = a_{i,t_r}$, then player j rewards player i if it is likely that player i in the round r takes the same action as the one in the initial period a_{i,t_r}.
  - (r-j-4) Player j punishes player i if it is likely for player i to tell a lie about

$\{a_{i,t}, y_{i,t}\}_{t \in T(r)}$ by

$$\sum_{t \in T(r)} T^{-10r} g_i(h_j, \hat{a}_{i,t}, \hat{y}_{i,t}). \tag{76}$$

Based on the message $\hat{\mathfrak{h}}_i^r$, player j calculates $A_i(\hat{\mathfrak{h}}_i^r)$, the set of player i's action that should be taken with positive probability in the round r after history $\hat{\mathfrak{h}}_i^r$.

We show the truthtelling incentive about (a_{i,t_r}, y_{i,t_r}) and $\{a_{i,t}, y_{i,t}\}_{t \in T(r)}$ by backward induction. We start from $r = R$, the last round. If $\zeta_j(\tilde{r}) = B$ or $\vartheta_j(\tilde{r}) = B$ with some $\tilde{r} < r$, the messages about the history in the round r are irrelevant. If $\zeta_j(\tilde{r}) = \vartheta_j(\tilde{r}) = G$ with $\tilde{r} < r$, then regardless of the specification of f, since (73), (74) and (75) have been sunk, it is optimal for player i to tell the truth about $\{a_{i,t}, y_{i,t}\}_{t \in T(R)}$. Since (74) dominates (75), it is optimal to tell the truth about a_{i,t_r}, y_{i,t_r}.

For the round $(R-1)$, (73), (74), (75) and (76)for $r = R$ are dominated by the smallest loss in (76) and (74) for $r = R - 1$. Therefore, the same argument for the round R 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 $f(a_i \mid \hat{\mathfrak{h}}_i^r, \mathcal{A}_j(r))$ by backward induction.

At the beginning of the round R, if $\zeta_j(\tilde{r}) = B$ or $\vartheta_j(\tilde{r}) = B$ with $\tilde{r} \leq R - 1$, then any action is exactly optimal and the specification of f is irrelevant. Otherwise, player i's value of taking a constant action $a_i \in A_i$ conditional on $\mathcal{A}_j(R)$ only depends on $\mathfrak{h}_i^R$ given the truthtelling strategy in the report block. This is true even after player i's deviation since player j's strategy is i.i.d. within each round. Hence, we can write the value as $v_i(a_i \mid \mathfrak{h}_i^R, \mathcal{A}_j(R))$. Let $f(a_i \mid \mathfrak{h}_i^R, \mathcal{A}_j(R))$ be such that

$$\Pr(\text{player } i \text{ reports the history}) \sum_{\tilde{a}_i} f\left(\tilde{a}_i \mid \mathfrak{h}_i^R, \mathcal{A}_j(R)\right) \Pr\left(\left\{\Psi_{j,t}^{\tilde{a}_i, a_{j,t}} = 1\right\} \mid a_{i,t} = a_i\right)$$

$$= \begin{cases} \max_{\tilde{a}_i \in A_i} v_i(\tilde{a}_i \mid \mathfrak{h}_i^R, \mathcal{A}_j(R)) - v_i(a_i \mid \mathfrak{h}_i^R, \mathcal{A}_j(R)) \text{ if } a_i \in A_i(\mathfrak{h}_i^R) \\ 0 \text{ otherwise} \end{cases}$$

for all $\mathfrak{h}_i^R$. Remember that $\sigma_i(x_i)$ is almost optimal except for the adjustment of the reward in the report block, that all the messages about $\mathfrak{h}_i^R$ transmit correctly, and that the variance of the reward in the report block based on the histories in the round R is bounded by T^{-10R+5}. Hence, we can make sure that

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

This makes it exactly optimal to take $a_i \in A_i(\mathfrak{h}_i^R)$ at t_R. After that, (75) and truthtelling incentive imply that it is optimal to constantly take a_{i,t_R}.

We can proceed until the first round and show the optimality of $\sigma_i(x_i)$. The difference from the round R is that, when player i takes a_t, it affects the reward for the messages sent after a_{t_r}, y_{t_r} about the history in the following rounds. Since this effect is dominated by (75), it is optimal for player i to take the same action as a_{t_r} constantly.

Note that if the round r is a supplemental round 2 for $\lambda_i(l+1)$ and this round does not have an impact ($\zeta_j(\tilde{r}) = \vartheta_j(\tilde{r}) = G$ with $\tilde{r} < r$ guarantee this), then the expected rewards in the report block are not affected by the strategy in the round r. Therefore, $\sigma_i(x_i)$ is exactly optimal as stated in Proposition 1. In addition, if the round r is a supplemental rounds for $\lambda_j(l+1)$, since $\sigma_i(x_i)$ is strictly optimal by $\frac{1}{2}\bar{L}(q_2 - q_1)$ without f from 2-(b) of Proposition 1, the equilibrium strategy is optimal regardless of the specification of f. Hence, we make f constant at 0.

Finally, we consider the reward on the message x_i:

$$f(x_i \mid x_j).$$

Conditional on $x_j \in \{G, B\}$, $\sigma_i(x_i)$ gives $\bar{v}_i$ ($\underline{v}_i$, respectively) if $x_j = G$ (B, respectively) without the reward in the report block from Section 11. Since the reward in the report block so far is bounded by $[-T^{-2}, T^{-2}]$, we can take $f(x_i \mid x_j) \in [-T^{-1}, T^{-1}]$ for all $x_i, x_j \in \{G, B\}$ such that $\sigma_i(x_i)$ gives $\bar{v}_i$ ($\underline{v}_i$, respectively) if $x_j = G$ (B, respectively) without the reward in the report block.

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] Deb, D. (2011): "Cooperation and Community Responsibility: A Folk Theorem for Repeated Random Matching Games," *mimeo.*

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

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

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

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

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

[10] 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.

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

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

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

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

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

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

[17] Kandori, M. (2010): "Weakly Belief-Free Equilibria in Repeated Games with Private Monitoring," *Econometrica*, 79, 877-892.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

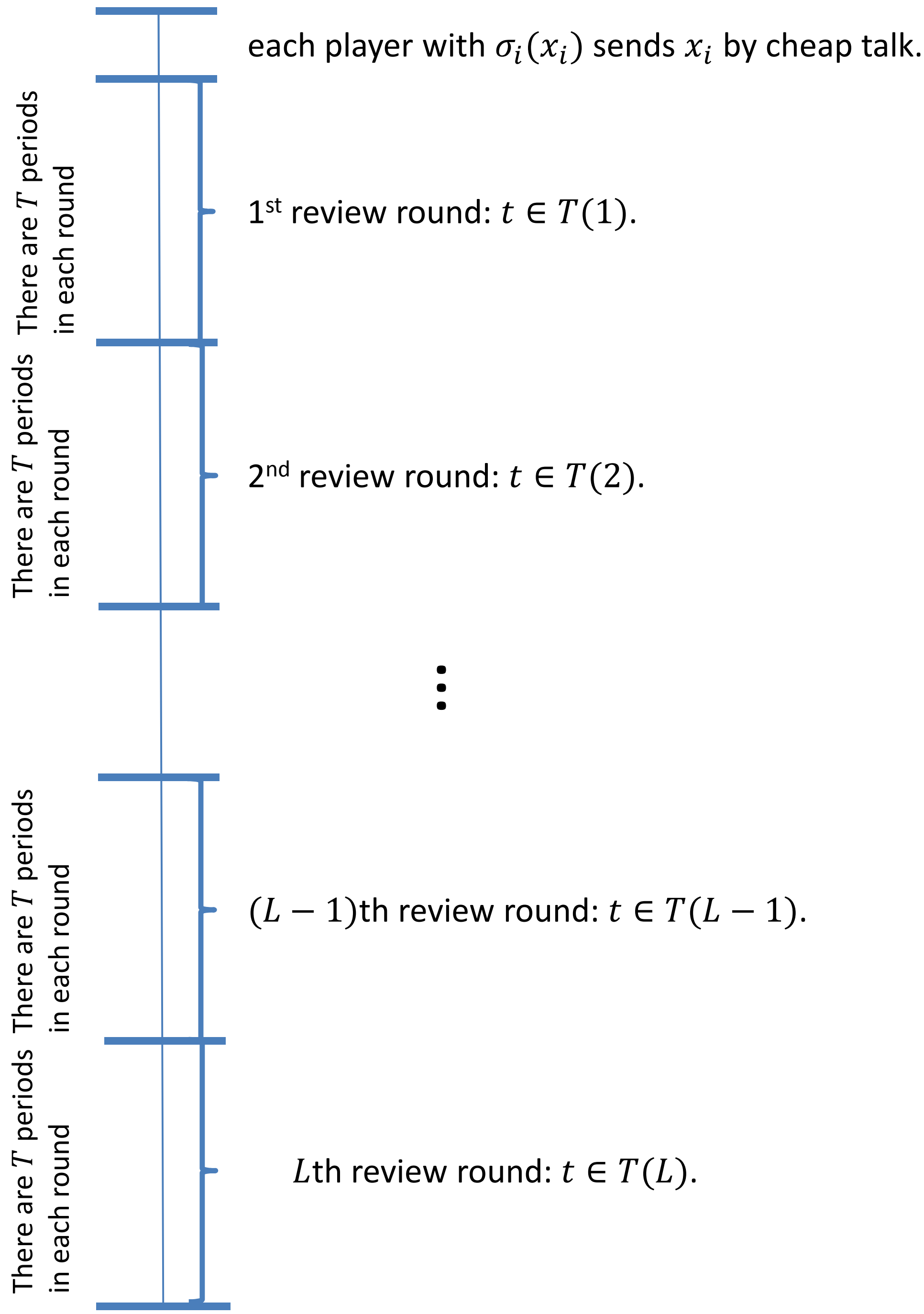

Figure 1:
The Informal Structure of the Phase

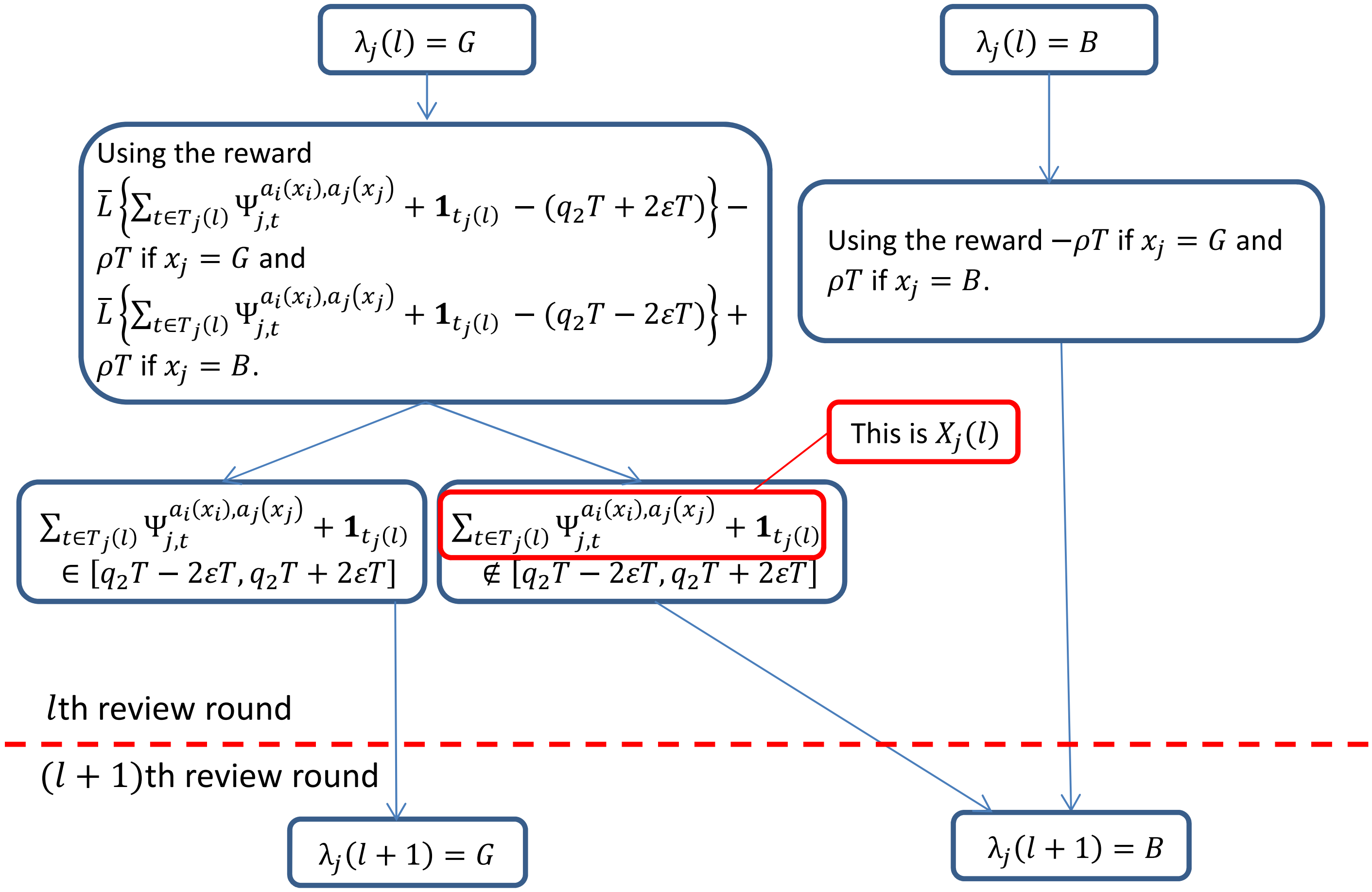

Figure 2:
Transition of λ_j

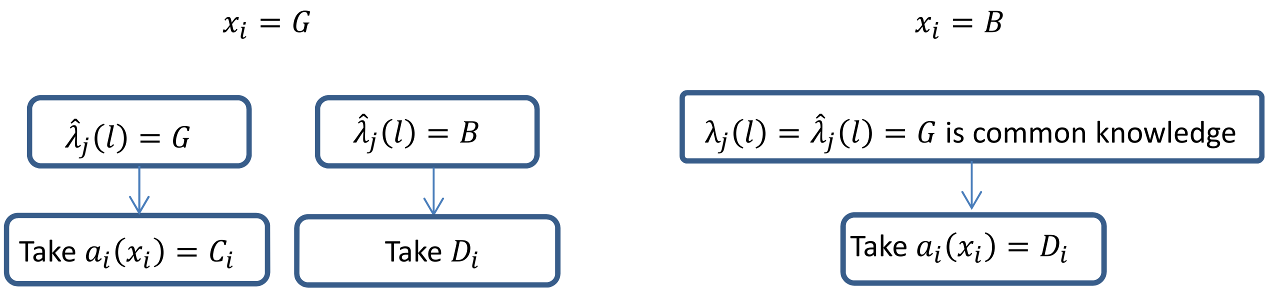

Figure 3:
Player i's Action

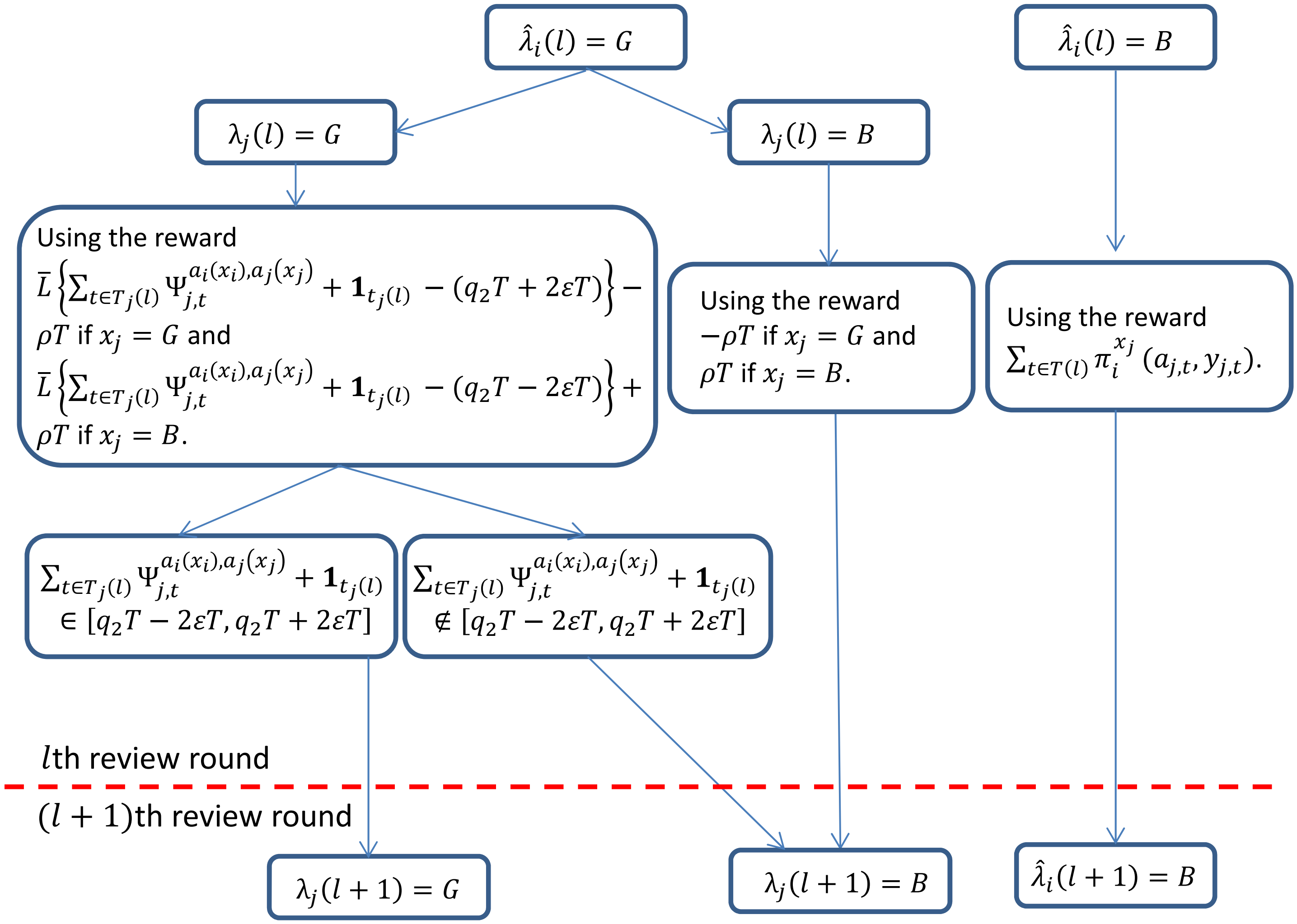

Figure 4:
Transition of $\lambda_j(l+1)$ and Player j's Reward After $\hat{\lambda}_i(l) = B$

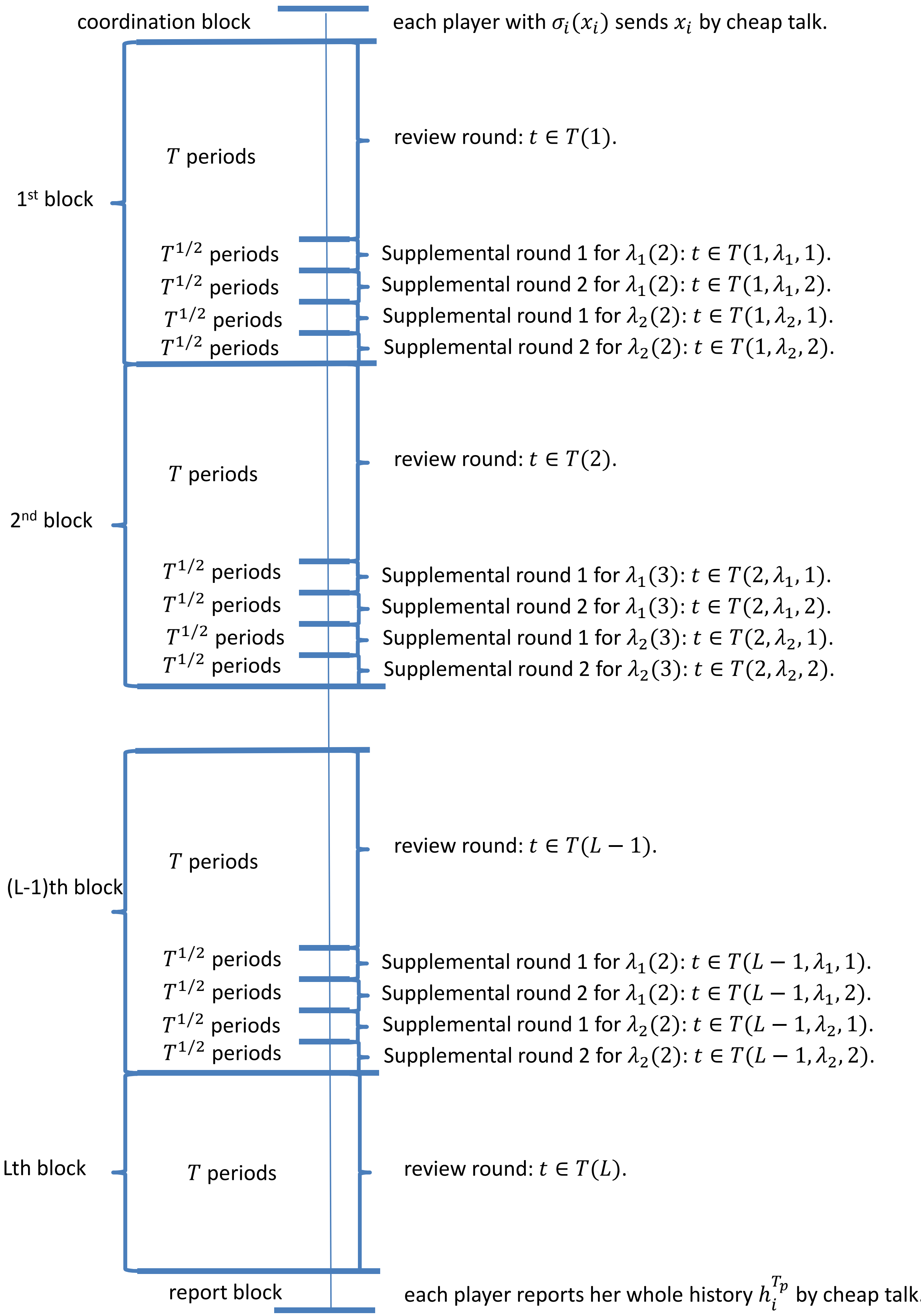

Figure 5: Formal Structure of the Phase

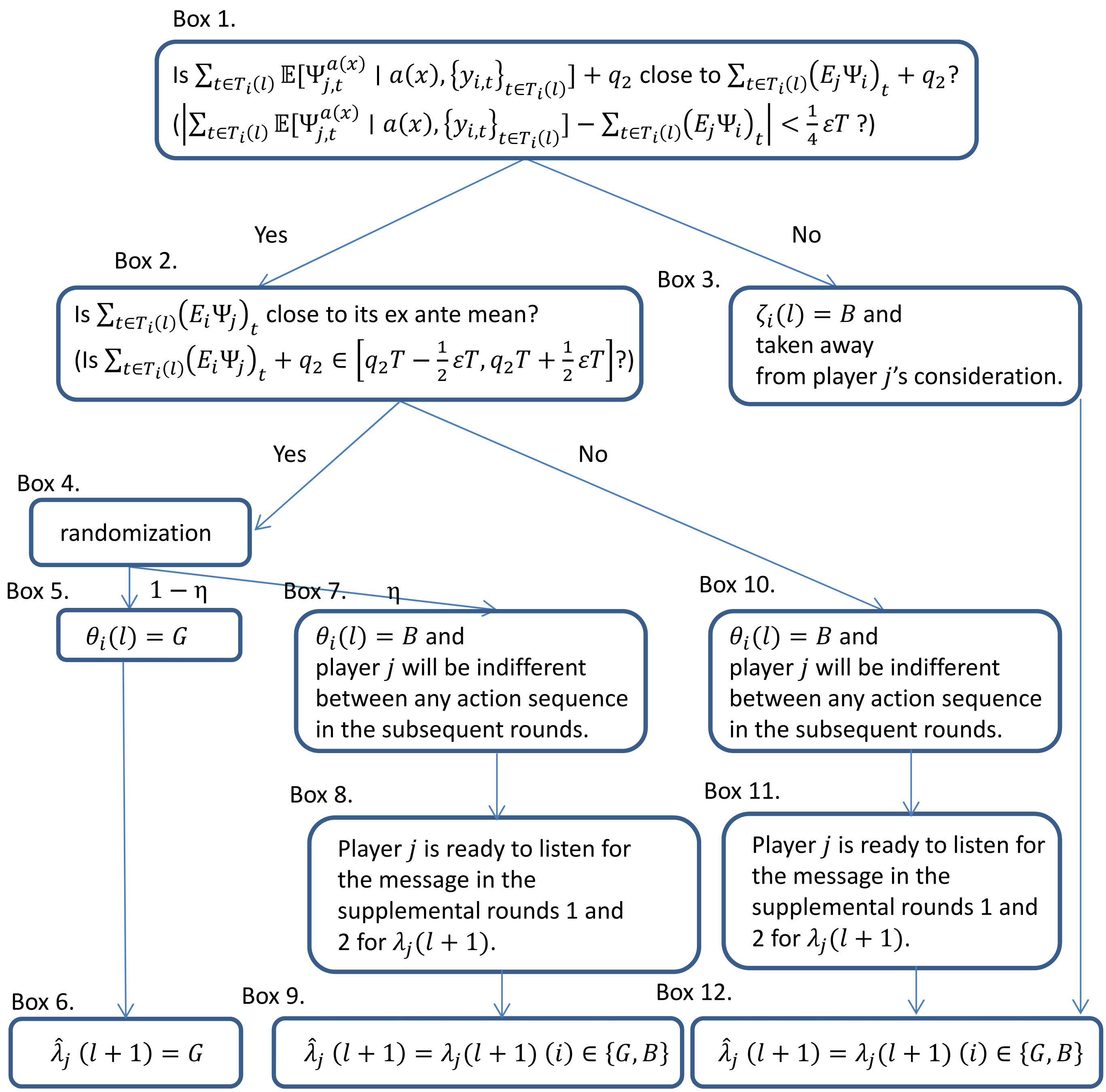

Figure 6:
Inference of λ_j

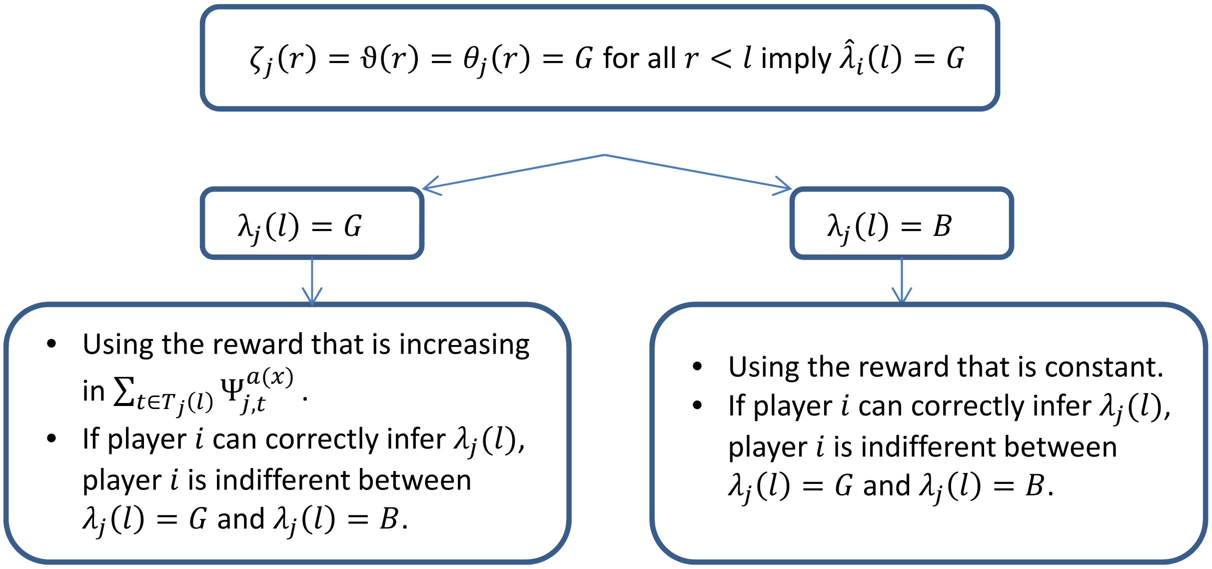

Figure 7:
Reward Function by Player j on Player i

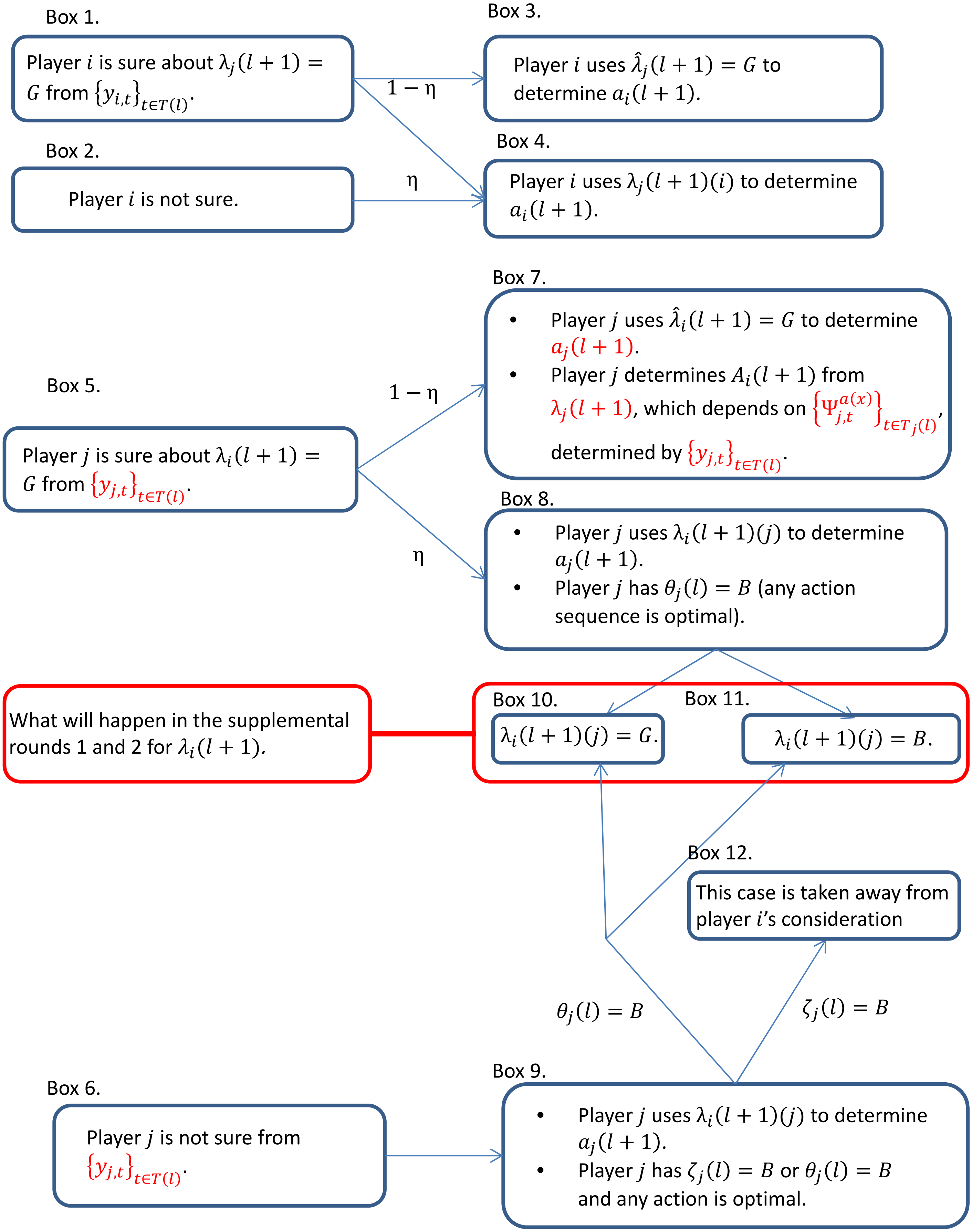

Figure 8: Player i's Inference of λ_j and Player j's Inference of λ_i

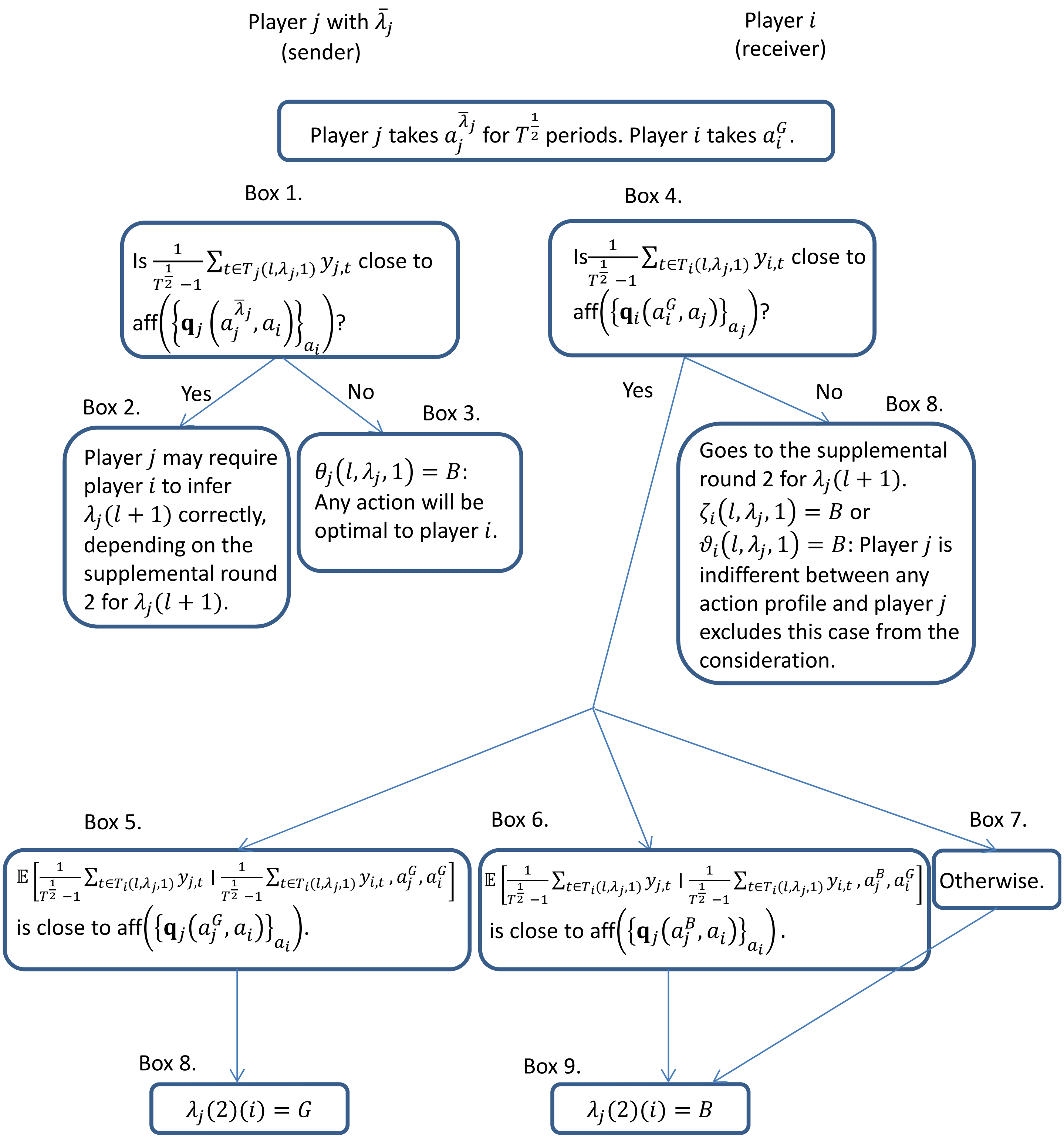

Figure 9:
Player i's Inference of Player j's Message

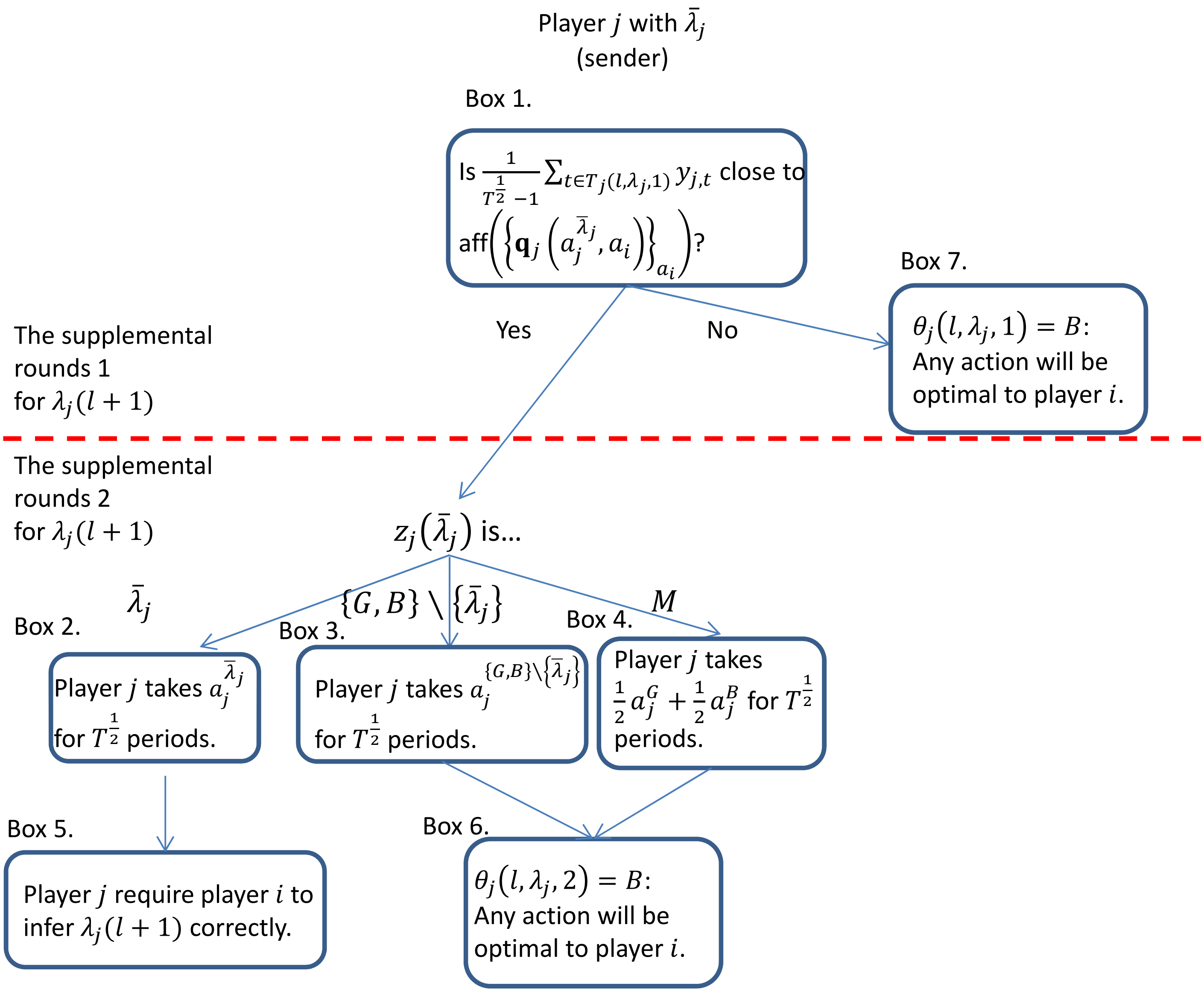

Figure 10:
Player j's Requirement on Player i's Inference

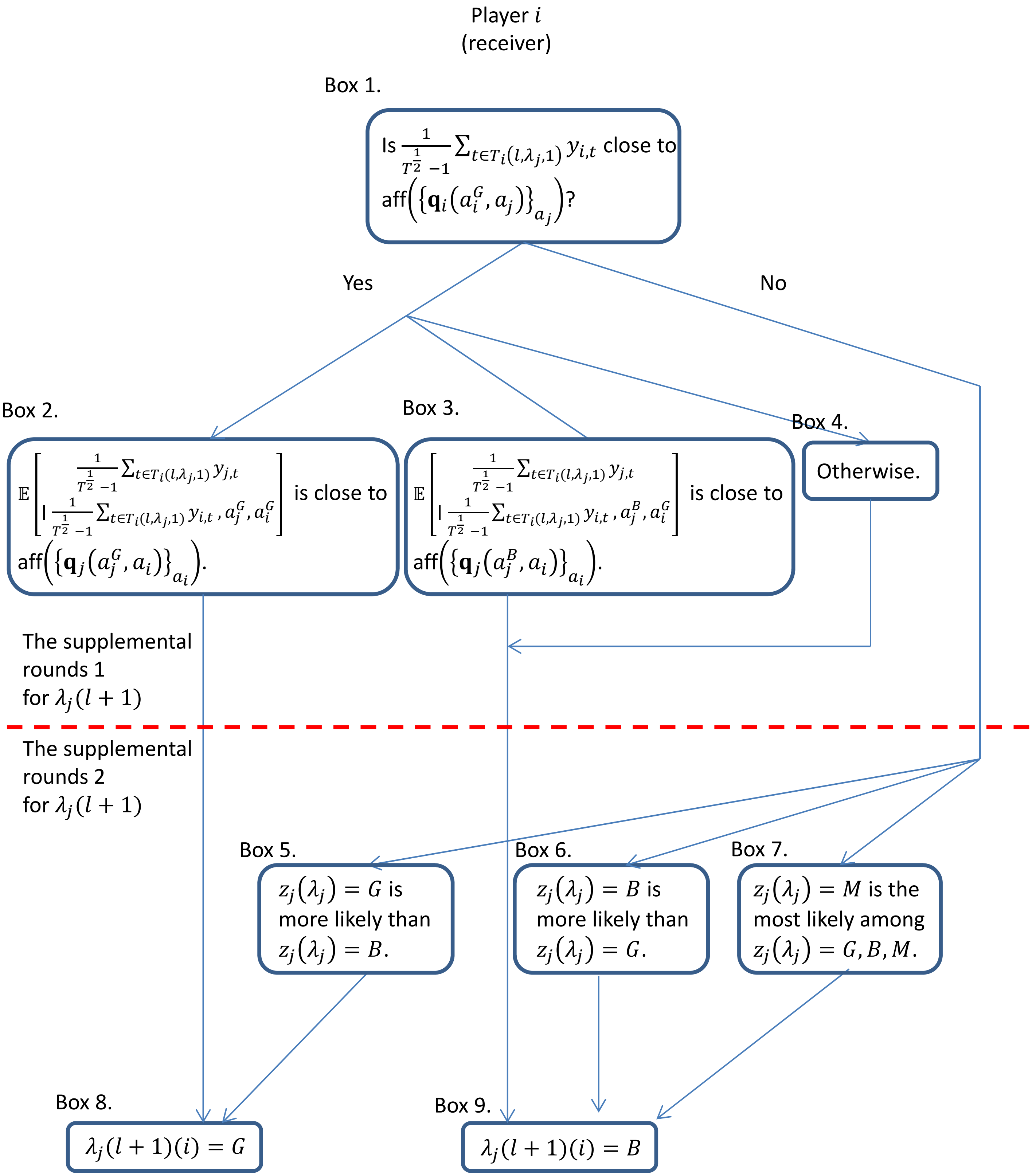

Figure 11:
Player i's Inference of Player j's Message

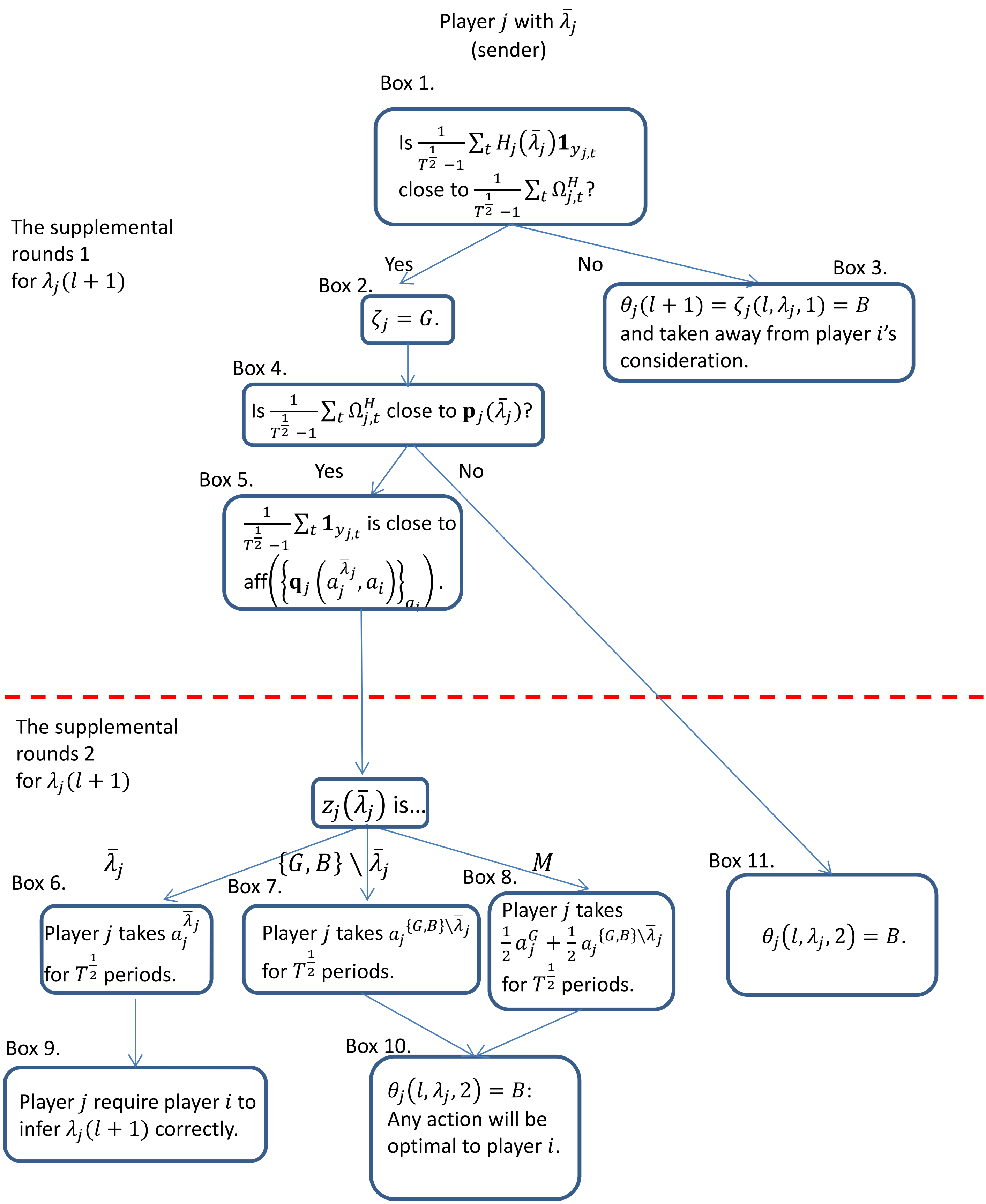

Figure 11:
Formal: Player j's Requirement on Player i's Inference

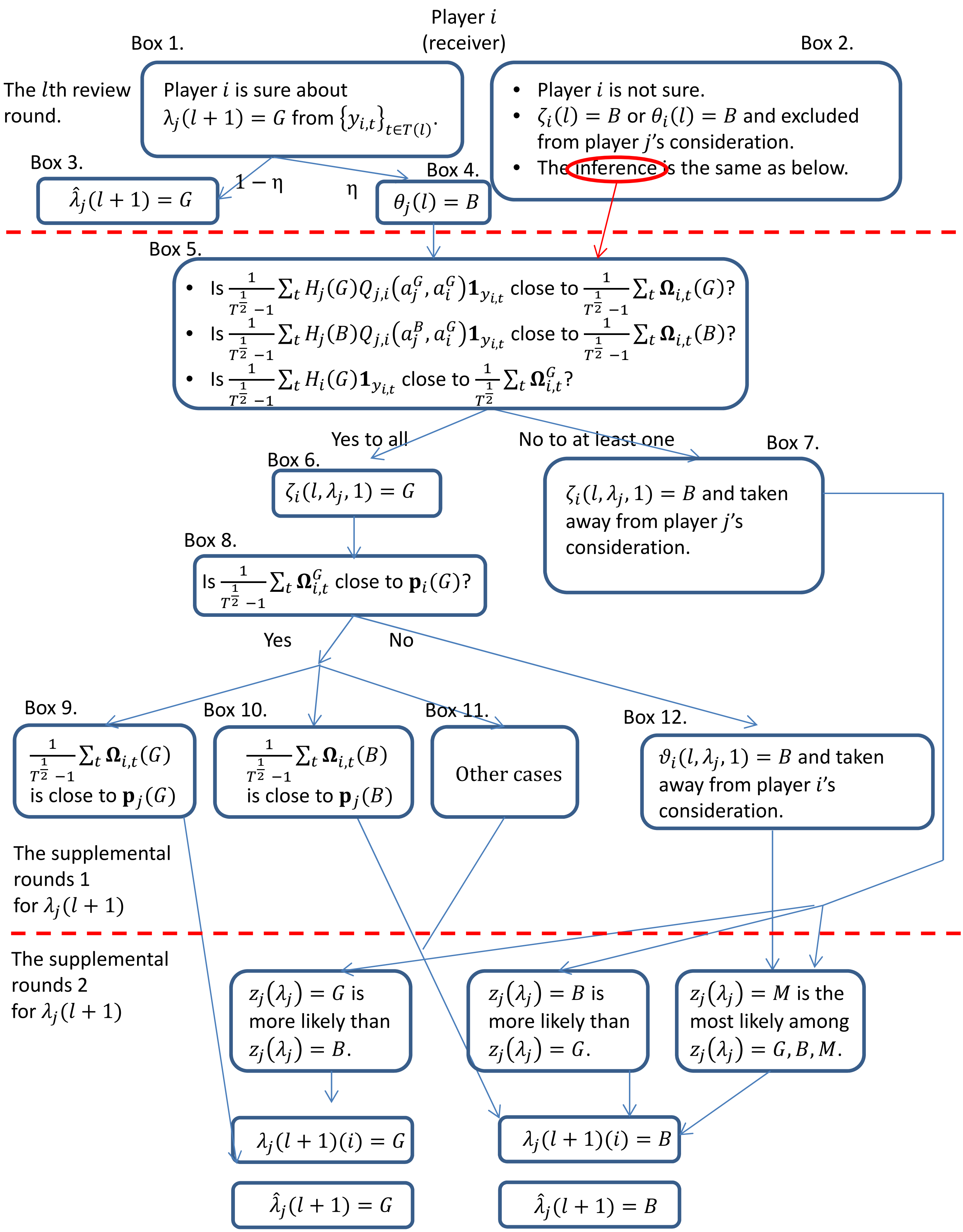

Figure 13: Formal: Player i's Inference of Player j's Message