Folk Theorem in Repeated Games with Private Monitoring

Takuo Sugaya[*][†]
Stanford Graduate School of Business

April 29, 2015

Abstract

We show that the folk theorem generically holds for the repeated two-player game with private monitoring if the support of each player's signal distribution is sufficiently large. Neither cheap talk communication nor public randomization is necessary.

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

Keywords: repeated game, folk theorem, private monitoring

[*]tsugaya@stanford.edu

[†]I thank Stephen Morris for invaluable advice, Yuliy Sannikov for helpful discussion, Andrzej Skrzypacz for carefully reading the paper, Yuichi Yamamoto for his suggestions to improve the exposition of the paper, and Dilip Abreu, Eduardo Faingold, Drew Fudenberg, Edoardo Grillo, Johannes Hörner, Yuhta Ishii, Michihiro Kandori, Vijay Krishna, George Mailath, Mihai Manea, Larry Samuelson, Tadashi Sekiguchi, Satoru Takahashi, 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. I also thank Piotr Dworczak and Sergey Vorontsov for careful proofreading, and Jenna Abel for professional editorial service. The remaining errors are my own responsibility.

1 Introduction

A possibility of tacit collusion when a party observes only private signals about the other parties' actions has been a long-standing open problem. Stigler (1964) considers a case in which a firm can cut the price secretly and conjectures that, since the coordination on the punishment is difficult with private signals, policing an implicit collusive agreement (punishing a deviator from the agreement) does not work effectively. Harrington and Skrzypacz (2011) examine a recent cartel agreement, and construct a model to replicate how the firms police the agreement. Since their equilibrium requires cheat-talk communication and monetary transfer, they show the possibility of explicit collusion, rather than tacit collusion. A theoretical possibility of tacit collusion is, therefore, yet to be proven.[1]

The literature on infinitely repeated games offers a theoretical framework to study tacit collusion. One of the key findings is the folk theorem: Any feasible and individually rational payoff profile can be sustained in equilibrium when players are sufficiently patient. This implies that sufficiently patient firms can sustain tacit collusion to obtain any feasible and individually rational payoff. Fudenberg and Maskin (1986) establish the folk theorem under perfect monitoring, in which players can directly observe the action profile. Fudenberg, Levine, and Maskin (1994) extend the folk theorem to imperfect public monitoring, in which players can observe only public noisy signals about the action profile.

It has been an open question whether the folk theorem holds in general private monitoring, in which players observe private noisy signals about other players' actions. Matsushima (2004), Hörner and Olszewski (2006), Hörner and Olszewski (2009), and Yamamoto (2012) show the folk theorem in restricted classes of private monitoring (see below for details), but little is known about general private monitoring.

The objective of this paper is to solve this open problem: In general private monitoring with discounting,[2] the folk theorem holds in the two-player game. Specifically, we identify sufficient conditions with which the folk theorem holds, and show that these sufficient conditions generically hold in private monitoring. In our equilibrium, we use neither cheap-talk communication nor

[1] Another area where private signals are important is team-production. If each agent in a team subjectively evaluates how the other agents contribute to the production, then seeing subjective evaluation as a private signal, the situation is modeled as repeated games with private monitoring.

[2] See Lehrer (1990) for the case with no discounting.

public randomization.[3] Furthermore, in the companion paper Sugaya (2012), we generalize our equilibrium construction to a general game with more than two players.

We now discuss papers about the folk theorem[4] and then illustrate how we prove the folk theorem, built upon the pre-existing work. The driving force of the folk theorem is reciprocity: If a player deviates today, she will be punished in the future (Stigler (1964) call this reciprocity "policing the agreement"). For this mechanism to work, each player needs to coordinate her action with the other players' histories. Whether the players achieve coordination depends on the players' information and thus on the monitoring structure.

Fudenberg and Maskin (1986) establish the folk theorem under perfect monitoring. Because the histories are common knowledge in perfect monitoring, each player can coordinate her continuation strategy with the other players' histories. Fudenberg, Levine, and Maskin (1994) extend the folk theorem to imperfect public monitoring by focusing on the equilibrium in which each player's continuation strategy depends only on past public signals. Because the histories of public signals are common knowledge, players can coordinate their continuation play through public signals.

On the other hand, in general private monitoring, each player's actions and signals are her private information, so players do not share common information about the histories. Hence, the coordination becomes complicated as periods proceed. This is why Stigler (1964) conjectures that collusion is impossible with general private monitoring.

If monitoring is almost public, then players believe that every player observes the same signal with a high probability. Hence, almost common knowledge about finite-past histories exists. This almost common knowledge enables Hörner and Olszewski (2009) to show the folk theorem in almost public monitoring.

Without almost public monitoring, almost common knowledge may not exist, and coordination is difficult. To deal with this difficulty, Piccione (2002) and Ely and Välimäki (2002) focus on a tractable class of equilibria, a "belief-free equilibrium." A strategy profile is belief free if, after each history profile, the continuation strategy of each player is optimal, conditional on the histories of the opponents. Hence, coordination is not an issue. Piccione (2002) and Ely and Välimäki (2002) show that the belief-free equilibrium sustains the set of feasible and individually rational

[3] There are papers that prove folk theorems in private monitoring with cheap-talk or mediated communication (explicit collusion). See, for example, Compte (1998), Kandori and Matsushima (1998), Aoyagi (2002), Aoyagi (2005), Fudenberg and Levine (2007), Obara (2009), Rahman (2014), and Renault and Tomala (2004).

[4] See also Kandori (2002) and Mailath and Samuelson (2006) for a survey.

payoffs in the two-player prisoners' dilemma with almost perfect monitoring.[5,6] However, with general monitoring or in a general game beyond prisoners' dilemma, the belief-free equilibrium cannot sustain the folk theorem. See Ely, Hörner, and Olszewski (2005) and Yamamoto (2009) for the formal proof. Hence, there are two ways to generalize the belief-free equilibrium: One is to keep prisoners' dilemma payoff structure and to consider noisy monitoring; and the other is to keep almost perfect monitoring and to consider a more general stage game.

The former approach by Matsushima (2004) and Yamamoto (2012) recovers the precision of the monitoring, assuming that monitoring is not perfect but conditionally independent: Conditional on action profiles, the players observe statistically independent signals. The idea is to replace one period in the equilibrium of Ely and Välimäki (2002) with a long T-period review phase. When we aggregate information over a long phase and rely on the law of large numbers, we can recover the precision of the monitoring.

To see the obstacle to further generalize their result to general monitoring with correlated signals, let us explain their equilibrium construction in the two-player prisoners' dilemma in which each player has two signals, good and bad. Suppose that with player i's cooperation, player $j \neq i$ observes a good signal g_j with 60%, while with player i's defection, player j observes a bad signal b_j with 60%. To achieve approximately efficient equilibrium, player j should not punish player i if she observes g_j for more than $(0.6 + \varepsilon)T$ times during a review phase with a small ε. Note that by the central limit theorem, $0.6T$ is what player j can expect from player i's cooperation.

Suppose now we are close to the end of a review phase. If player i's signals are correlated with player j's, then after rare histories, player i may believe that, given her own history and correlation, player j has already observed g_j for more than $(0.6 + \varepsilon)T$ times. After such a history, player i may want to switch to defection. (If signals are conditionally independent, then player i after each history always believes that player j observes g_j at most $(0.6 + \varepsilon)T$ times with a very high probability. Hence, switching does not happen.)

Once player i switches her action based on her history, player j wants to learn player i's switch of actions via player j's private signals. Note that player i reviews player j's actions by player

[5] See Yamamoto (2007) for the N-player prisoners' dilemma.

[6] Kandori and Obara (2006) use a similar concept to analyze a private strategy in public monitoring. Kandori (2011) considers "weakly belief-free equilibria," which is a generalization of 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 in which a player can improve the precision of monitoring by paying a cost.

i's history and that player j wants to know how player i has reviewed player j so far. Player i's switch of actions is informative about her history and so about how she has reviewed player j. Recursively, if a player switches actions, then we have to deal with a high order belief about each player's history.

To deal with this problem, in our equilibrium, we divide a review phase into multiple review rounds. If the division is fine enough, we can make sure that the switch of actions only happens at the beginning of the review round. In addition, before each review round, player j tells player i whether player i should switch the actions. (Apart from the issue of player j's incentive to tell the truth,) if this communication is successful, then player i does not need to learn player j's switches. Since we do not assume cheap talk, this communication is done by player j taking different actions to send different messages. Since player i infers player j's actions (messages) from noisy private signals, player i may make a mistake and player j may not realize player i's mistake. In Section 10, we construct a module for a player to send the other player a message by taking actions with noisy private signals, so that, if player i suspects that she made a mistake, then she believes that player j has "realized" the mistake.

The latter approach is to generalize Ely and Välimäki (2002) to a general stage game (here we focus on two-player games), keeping monitoring almost perfect. In the belief-free equilibrium in Ely and Välimäki (2002), in each period, each player i picks a state x_i that can be G (good) or B (bad). In each period, given state $x_i \in \{G, B\}$, player i takes a mixed action $\sigma_i(x_i)$. Together with the state transition, they make sure that, for each state of the opponent $x_j \in \{G, B\}$, both $\sigma_i(G)$ and $\sigma_i(B)$ are optimal conditional on x_j.

A reason why the belief-free equilibrium in Ely and Välimäki (2002) cannot sustain the folk theorem in a general game is that it is hard for player j with $\sigma_j(B)$ to punish player i severely enough after a signal which statistically indicates her deviation, at the same time keeping both $\sigma_j(G)$ and $\sigma_j(B)$ optimal against both $x_i = G$ and $x_i = B$ and keeping the equilibrium payoff with $x_i = x_j = G$ sufficiently high.

Hörner and Olszewski (2006) overcome this difficulty as follows: They divide the repeated game into L-period phases.[7] In each phase, each player i picks a state $x_i \in \{G, B\}$. In each phase, given state $x_i \in \{G, B\}$, player i takes an L-period dynamic strategy $\sigma_i(x_i)$. Again, for each

[7]They use the term T-period blocks, but we use L and phases instead, in order to make the terminology consistent within the paper.

state of the opponent $x_j \in \{G, B\}$, both $\sigma_i(G)$ and $\sigma_i(B)$ are optimal conditional on x_j at the beginning of review phase. However, within a phase, since the players coordinate on x by taking actions at the beginning of the phase, it is optimal to adhere to $\sigma_i(x_i)$ once player i takes an action corresponding to $\sigma_i(x_i)$ at the beginning of the phase. Having L periods in the phase, they can create a severe punishment by letting player j with $\sigma_j(B)$ switch to a minimax strategy after a signal which statistically indicates the opponent's deviation. Moreover, since the switch happens only after a rare signal which indicates a deviation, we can keep the payoff of $\sigma_j(B)$ (and that of $\sigma_j(G)$ in order to keep belief-free property at the beginning of the phase) sufficiently high given $x_i = G$.

There are two difficulties to generalize their construction to general monitoring: With general monitoring, the coordination on x by taking actions becomes harder since the actions are less precisely observed. Second, as seen in the case with Matsushima (2004), since each player switches actions based on past signals, each player may start to learn the opponent's history by observing each other's switches via noisy signals. This learning becomes complicated with noisy signals. Again, by using the module to send a message by taking actions with noisy private signals, we make sure that the players can coordinate on x and learning does not change the player's optimal strategy.

In total, combining Matsushima (2004) and Hörner and Olszewski (2006), we construct a following equilibrium: We divide the repeated game into long review phases. At the beginning of the phase, the players coordinate on x by using the module to send messages via taking actions. Then, we have L review rounds. These review rounds serve two roles: One is to allow player i to switch actions when she believes that the opponent has observed a lot of good signals in order to deal with conditionally dependent signals. The other is to allow player j to switch to a minimaxing action when she observes a lot of signals statistically indicating the opponent's deviation as in Hörner and Olszewski (2006). Relatedly, player i also switches her own actions to take a best response against player j's minimaxing strategy, when player j switches to the minimaxing actions. To coordinate these switches, before each review rounds, the players communicate about the continuation play again by the module to send messages via taking actions. The module and communication are designed carefully so that learning from the opponent's actions in the review round does not change the optimal action within the review round.

The rest of the paper is organized as follows. In Section 2, we introduce the model, and in

Section 3, we state assumptions and result (folk theorem). The rest of the paper proves the folk theorem.

(After giving a short roadmap of the proof in Section 4), in Section 5, we derive the sufficient conditions in a finitely repeated game without discounting such that, once we prove the sufficient conditions, then it implies that the folk theorem holds in the infinitely repeated game with discounting. From there on, we focus on proving these sufficient conditions.

Section 6, we define multiple variables, with which Section 7 pins down the structure (of the strategy) of the finitely repeated game.

(After giving another short roadmap in Section 8,) since the rest of the proof is long and complicated, we first offer the overview in Section 9. Then, we prove two modules which may be of their own interests. Section 10 defines the module for player j to send a binary message $m \in \{G, B\}$ to player i. Section 11 defines the module that will be used for the equilibrium construction of the review round. Given these modules, we explain how to use these modules to prove the sufficient conditions in Section 12.

Sections 13-16 actually prove the sufficient conditions. (See Section 12 for the structure of these sections.) Since the entire proof is long, in Sections 13-16, we summarize each step as lemmas and offer the intuitive explanation of the proof, relegating the technical proof to Appendix A (online appendix). In addition, we offer the table of notation in Appendix B. Appendices A and B are provided as supplemental materials for the submission. In case the referees are not provided the supplemental materials, the online appendix is also available at https://sites.google.com/site/takuosugaya/home/research.

2 Model

2.1 Stage Game

We consider a two-player game with private monitoring. The stage game is given by $\{I, \{A_i, Y_i\}_{i \in I}, q\}$. Here, $I = \{1, 2\}$ is the set of players, A_i is the finite set of player i's pure actions, and Y_i is the finite set of player i's private signals.

In every stage game, player i chooses an action $a_i \in A_i$, which induces an action profile $a \equiv (a_1, a_2) \in A \equiv \prod_{i \in I} A_i$. Then, a signal profile $y \equiv (y_1, y_2) \in Y \equiv \prod_{i \in I} Y_i$ is realized according to a joint conditional probability function $q(y \mid a)$. Summing up these probabilities with respect to y_j, let $q_i(y_i \mid a) \equiv \sum_{y_j \in Y_j} q(y \mid a)$ denote the marginal distribution of player i's signals. Throughout the paper, when we say players i and j, players i and j are different: $i \neq j$.

Following the convention in the literature, we assume that player i's ex post utility is a deterministic function of player i's action a_i and her private signal y_i. This implies that observing her own ex post utility does not give player i any further information than (a_i, y_i).[8] Let $\tilde{u}_i(a_i, y_i)$ be player i's ex post payoff. Taking the expectation of the ex post payoff, we can derive player i's expected payoff from $a \in A$: $u_i(a) \equiv \sum_{y_i \in Y_i} q_i(y_i \mid a)\tilde{u}_i(a_i, y_i)$.

Let F^* be the set of feasible and individually rational payoffs:

$$F^* \equiv \left\{ v \in \text{co}(\{u(a)\}_{a \in A}) : v_i \geq \min_{\alpha_j \in \Delta(A_j)} \max_{a_i \in A_i} u_i(a_i, \alpha_j) \text{ for all } i \in I \right\}. \tag{1}$$

Let $\alpha_j^{\min}$ be the minimax strategy.

2.2 Repeated Game

Consider the infinitely repeated game with the (common) discount factor $\delta \in (0, 1)$. Let $h_i^t \equiv (a_{i,\tau}, y_{i,\tau})_{\tau=1}^{t-1}$ with $h_i^1 = \{\emptyset\}$ be player i's history in period t; and let H_i^t be the set of all the possible histories of player i. In each period t, player i takes an action $a_{i,t}$ according to her strategy $\sigma_i : \bigcup_{t=1}^{\infty} H_i^t \to \Delta(A_i)$. Let Σ_i be the set of all strategies of 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 our four assumptions and main result. First, we assume that the distribution of private signal profiles has full support:

Assumption 1 *For each $a \in A$ and $y \in Y$, we have $q(y \mid a) > 0$.*

Let us define

$$\varepsilon_{\text{support}} \equiv \min_{y \in Y, a \in A} q(y \mid a) > 0 \tag{2}$$

[8] Otherwise, we see the realization of player i's ex post utilities as a part of player i's private signals. Let U_i be the finite set of realizations of player i's ex post utilities, and let $\tilde{u}_i \in U_i$ be a generic element of U_i. We see a vector $(y_i, \tilde{u}_i)$ as player i's private signal. Hence, the set of player i's signals is now $Y_i \times U_i$.

be the lower bound of the probability. Note that this also implies the full support of the marginal distribution: $\min_{i\in I, y_i \in Y_i, a\in A} q_i(y_i \mid a) \geq \varepsilon_{\text{support}}$.

Assumption 1 excludes public monitoring, where $y_i = y_j$ with probability one. One may find allowing public monitoring of special interest: Our other Assumptions 2, 3, and 4 defined below generically hold if $|Y_i| \geq |A_j|$ for each i and j, while pairwise full rank in Fudenberg, Levine, and Maskin (1994) requires $|Y_i| \geq |A_1| + |A_2| - 1$. (See also Radner, Myerson, and Maskin (1986).) We can extend the result to allow public monitoring and interested readers are referred to the working paper.[9]

Second, we assume that player i's signal statistically identifies player j's action. Let

$$q_i(a_i, a_j) \equiv \left(q_i(y_i \mid a_i, a_j)\right)_{y_i \in Y_i} \tag{3}$$

be the vector expression of the marginal distribution of player i's signals.

Assumption 2 *For each $i \in I$ and $a_i \in A_i$, the collection of $|Y_i|$-dimensional vectors $(q_i(a_i, a_j))_{a_j \in A_j}$ is linearly independent.*

Suppose that player i takes a_i. When player j changes her actions, the change gives a different distribution of player i's signals, so that player i can statistically identify player j's deviation.

Third, we assume that each signal of player i happens with different probabilities after player j's different actions.

Assumption 3 *For each $i \in I$, $a_i \in A_i$, and $a_j, a_j' \in A_j$ satisfying $a_j \neq a_j'$, we have $q_i(y_i \mid a_i, a_j) \neq q_i(y_i \mid a_i, a_j')$ for all $y_i \in Y_i$.*

Fourth, suppose that in the repeated game, player j takes a mixed strategy $\alpha_j \in \Delta(A_j)$. If player i's history is (a_i, y_i), then player i believes that player j's history (a_j, y_j) is distributed according to $\Pr((a_j, y_j) \mid \alpha_j, a_i, y_i)$. Here, Pr is defined such that player j takes a_j according to α_j and signal profile y is drawn from $q(y \mid a)$. The following assumption about $\Pr((a_j, y_j) \mid \alpha_j, a_i, y_i)$ ensures that there exists a mixed strategy of player j such that different histories of player i give different beliefs about player j's history:

[9] We use *private strategies*. The possibility of private strategies to improve efficiency is first pointed out by Kandori and Obara (2006).

Assumption 4 *For each $i \in I$, there exists $\alpha_j \in \Delta(A_j)$ such that, for all $(a_i, y_i), (a_i', y_i') \in A_i \times Y_i$ with $(a_i, y_i) \neq (a_i', y_i')$, there exists $(a_j, y_j) \in A_j \times Y_j$ such that $\Pr((a_j, y_j) \mid \alpha_j, a_i, y_i) \neq \Pr((a_j, y_j) \mid \alpha_j, a_i', y_i')$.*[10]

With these four assumptions, we can show the following theorem.

Theorem 1 *If Assumptions 1, 2, 3, and 4 are satisfied, then in a two-player repeated game, for each payoff profile $v \in \mathrm{int}(F^*)$, there exists $\bar{\delta} < 1$ such that, for all $\delta > \bar{\delta}$, we have $v \in E(\delta)$.*

4 Short Road Map

Before giving the details of the proof, we first display a short road map to prove Theorem 1. First, in Section 5, we derive the sufficient conditions in a finitely repeated game without discounting such that, once we prove the sufficient conditions, then it implies that Theorem 1 holds in the infinitely repeated game with discounting. Second, in Section 6, we define multiple variables, with which Section 7 pins down the structure (of the strategy) of the finitely repeated game. The rest of the roadmap, which is easier to understand once we see the structure in Section 7, is postponed until Section 8.

5 Reduction to a Finitely Repeated Game without Discounting

To prove Theorems 1, we fix a payoff $v \in \mathrm{int}(F^*)$ arbitrarily. First, we derive sufficient conditions such that once we prove these sufficient conditions hold for given v, then $v \in E(\delta)$. These sufficient conditions are stated as conditions in a "finitely repeated game without discounting." This reduction to the finitely repeated game is standard (see Hörner and Olszewski (2006), for example), except that we successfully reduce the infinitely repeated game with discounting to the finitely repeated game without discounting. Since there is no discounting, we can treat each period identically. This identical treatment will simplify the equilibrium construction.

To this end, we see the infinitely repeated game as repetitions of T_P-period review phases. We make sure that each review phase is recursive. (Specifically, our equilibrium is a T_P-period block

[10] This conditional probability is well defined for each α_j and (a_i, y_i) given Assumption 1.

equilibrium in the language of Hörner and Olszewski (2006).) We decompose each player's payoff from the repeated game as the summation of instantaneous utilities from the current phase and continuation payoff from the future phases.

Instead of analyzing the infinitely repeated game, we concentrate on a T_P-period finitely repeated game, in which payoffs are augmented by a terminal reward function that depends on a history of the finitely repeated game. The reward function corresponds to the continuation payoff from the future phases.

Specifically, in the finitely repeated game, in each period $t = 1, ..., T_P$, each player i takes an action based on her private history $h_i^t \equiv (a_{i,\tau}, y_{i,\tau})_{\tau=1}^{t-1}$. Let Σ_i be the set of player i's strategies. In equilibrium, each player is in one of the two possible states $\{G, B\}$ and, given state $x_i \in \{G, B\}$, she takes a strategy $\sigma_i(x_i) \in \Sigma_i$. On the other hand, player i's reward function is determined by player j's state x_j and player j's history of the finitely repeated game $h_j^{T_P+1}$: $\pi_i(x_j, h_j^{T_P+1})$. Note that *player i's* reward function is the statistic that *player j* calculates.

We will show that if we find a natural number $T_P \in \mathbb{N}$, a strategy $\sigma_i(x_i)$, a reward function $\pi_i(x_j, h_j^{T_P+1})$, and values $\{v_i(x_j)\}_{x_j \in \{G,B\}}$ such that the following sufficient conditions hold, then we have $v \in E(\delta)$: For each $i \in I$,

1. [Incentive Compatibility] For all $x \in \{G, B\}^2$,

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

   In words, incentive compatibility requires that, given each state of the opponent, $x_j \in \{G, B\}$, both $\sigma_i(G)$ and $\sigma_i(B)$ are optimal for player i to maximize the summation of instantaneous utilities and reward function. Since $\pi_i(x_j, h_j^{T_P+1})$ is the movement of the continuation payoff in the context of the infinitely repeated game and $\sum_{t=1}^{T_P} \delta^{t-1} u_i(a_t) + \delta^{T_P} \pi_i(x_j, h_j^{T_P+1})$ is close to $\sum_{t=1}^{T_P} u_i(a_t) + \pi_i(x_j, h_j^{T_P+1})$ for a sufficiently large δ, this condition implies both $\sigma_i(G)$ and $\sigma_i(B)$ are optimal in each review phase regardless of the opponent's state.

2. [Promise Keeping] For all $x \in \{G, B\}^2$,

$$\frac{1}{T_P}\mathbb{E}\left[\sum_{t=1}^{T_P} u_i(a_t) + \pi_i(x_j, h_j^{T_P+1}) \mid \sigma(x)\right] = v_i(x_j). \tag{5}$$

Here, for notational convenience, we write $\sigma(x) \equiv (\sigma_1(x_1), \sigma_2(x_2))$. In words, (5) says that the time average of the instantaneous utilities and the change in continuation payoff, $\pi_i(x_j, h_j^{T_P+1})$, is equal to $v_i(x_j)$. This implies that, for a sufficiently high discount factor, $v_i(x_j)$ is approximately player i's value from the infinitely repeated game if player j's current state is x_j.

3. [Full Dimensionality] The values $v_i(B)$ and $v_i(G)$ contain v_i between them:

$$v_i(B) < v_i < v_i(G). \tag{6}$$

   Since this condition implies $v_i(B) < v_i(G)$, when player j switches from $x_j = G$ to $x_j = B$, the switch strictly decreases player i's payoff. Hence, player j can punish player i by history-contingent state transitions. In addition, since $v_i \in [v_i(B), v_i(G)]$, player j can mix the initial state properly so that player i's initial equilibrium payoff is v_i.

4. [Self Generation] The sign of $\pi_i(x_j, h_j^{T_P+1})$ satisfies a proper condition: For all $h_j^{T_P+1}$, $\pi_i(G, h_j^{T_P+1}) \leq 0$ and $\pi_i(B, h_j^{T_P+1}) \geq 0$. If we define

$$\text{sign}(x_j) \equiv \begin{cases} -1 & \text{if } x_j = G, \\ 1 & \text{if } x_j = B, \end{cases} \tag{7}$$

   then this condition is equivalent to the following:

$$\text{sign}(x_j)\pi_i(x_j, h_j^{T_P+1}) \geq 0. \tag{8}$$

   We call the condition (8) "self generation."

   This corresponds to the condition that Hörner and Olszewski (2006) impose: The reward function when the opponent's state is G (or B) is nonpositive (or nonnegative). As will be seen in the proof, in the infinitely repeated game, self generation ensures that player i's continuation payoff specified by the reward function at the end of a review phase is included in $[v_i(B), v_i(G)]$. Since $v_i(x_j)$ is player i's value when player j's state is x_j, this inclusion ensures that, by mixing $x_j = G$ and B properly in the next phase, player j implements the continuation payoff specified by the reward function.[11]

The following lemma proves that the above conditions are sufficient.

Lemma 1 *For any $v \in \mathbb{R}^2$, if there exist $T_P \in \mathbb{N}$, $\{\{\sigma_i(x_i)\}_{x_i \in \{G,B\}}\}_{i \in I}$, $\{\{\pi_i(x_j, h_j^{T_P+1})\}_{x_j \in \{G,B\}}\}_{i \in I}$, and $\{v_i(x_j)\}_{i \in I, x_j \in \{G,B\}}$ such that the conditions (5)–(8) are satisfied, then we have $v \in E(\delta)$ for a sufficiently large $\delta \in (0,1)$.*

Proof. See Appendix A.2. Appendices A and B are provided as supplemental materials for the submission. In case the referees are not provided the supplemental materials, the online appendix is also available at https://sites.google.com/site/takuosugaya/home/research. ■

Let us comment on why we can *ignore discounting*. Recall that Assumption 2 ensures that player j can statistically infer player i's action. Hence, if she infers that player i incurs a loss in earlier periods of the finitely repeated game rather than later, then player j can compensate player i slightly (of order $1-\delta$) so that player i is indifferent about when to incur the loss. For a sufficiently large δ, such compensation is arbitrarily small.

Therefore, we focus on, for any $v \in \text{int}(F^*)$, finding $T_P \in \mathbb{N}$, $\{\{\sigma_i(x_i)\}_{x_i \in \{G,B\}}\}_{i \in I}$, $\{\{\pi_i(x_j, h_j^{T_P+1})\}_{x_j \in \{G,B\}}\}_{i \in I}$, and $\{v_i(x_j)\}_{i \in I, x_j \in \{G,B\}}$ to satisfy (5)–(8).

6 Basic Variables

To construct the variables to satisfy (5)–(8), it will be useful to define some functions/variables. In Section 6.1, we fix $\pi_i[\alpha]$, $\pi_i^{x_j}$, $\bar{u} > 0$, and $\left(\bar{u}_i^{x_j}\right)_{i \in I, x_j \in \{G,B\}}$. Here, $\alpha \in \Delta(A)$ with $\Delta(A) \equiv \Delta(A_1) \times \Delta(A_2)$. Note that we do not allow correlation between players' actions; in Section 6.2, we fix $(a(x), \alpha_i^{\rho}(x), \alpha_i^{*,\rho}(x))_{x \in \{G,B\}^2}$ for each ρ, $\alpha_i^{\min,\rho}$ for each ρ, $\rho_{\text{payoff}} > 0$, $(v_i(x_j), u_i(x_j))_{i \in I, x_j \in \{G,B\}}$, $L \in \mathbb{N}$, $\eta > 0$, $a_i(G)$, $a_i(B)$, and α_i^{mix}; and in Section 6.3, we fix S, $\sigma_i^{\mathbb{S}(t)}$, ϕ_j, q_G, q_B, $\pi_i^{\text{c.i.}}$, and $\varepsilon_{\text{strict}}$.

[11] One may wonder whether we need the condition that $\left|\frac{(1-\delta)\pi_i(x_j, h_j^{T_P+1})}{\delta^{T_P}}\right|$ is sufficiently small. This condition is automatically satisfied for sufficiently large δ for the following reason: We first fix T_P. Then, $\left|\pi_i(x_j, h_j^{T_P+1})\right|$ is bounded since the number of histories given T_P is finite. Finally, by taking δ sufficiently close to 1, we can make $\left|\frac{(1-\delta)\pi_i(x_j, h_j^{T_P+1})}{\delta^{T_P}}\right|$ arbitrarily small.

6.1 Basic Reward Functions $\pi_i[\alpha]$ and $\pi_i^{x_j}$

First, for each $\alpha \in \Delta(A)$, we create $\pi_i[\alpha](a_j, y_j)$ that cancels out the differences in the instantaneous utilities for different a's. Since Assumption 2 implies that player j can statistically infer player i's action from her signals, for each $a_j \in A_j$, there exists $\pi_i(a_j, \cdot) : Y_j \to \mathbb{R}$ such that

$$u_i(a_i, a_j) + \mathbb{E}[\pi_i(a_j, y_j) \mid a_i, a_j] = 0 \tag{9}$$

for each $a_i \in A_i$. For each α, we define $\pi_i[\alpha](a_j, y_j) = \pi_i(a_j, y_j) + u_i(\alpha)$ so that

$$u_i(a) + \mathbb{E}[\pi_i[\alpha](a_j, y_j) \mid a] = u_i(\alpha) \tag{10}$$

for each $a \in A$. This equality also implies that the expected value of $\pi_i[\alpha](a_j, y_j)$ given α is zero: $\mathbb{E}[\pi_i[\alpha](a_j, y_j) \mid \alpha] = 0$. In addition, $\pi_i[\alpha](a_j, y_j)$ is continuous in α.

Second, by adding/subtracting a constant depending on x_j to/from $\pi_i(a_j, y_j)$, we create $\pi_i^{x_j}(a_j, y_j)$ that makes player i indifferent between any action profile and that satisfies self generation: $u_i(a) + \mathbb{E}\left[\pi_i^{x_j}(a_j, y_j) \mid a\right]$ does not depend on $a \in A$, and we have $\text{sign}(x_j)\pi_i^{x_j}(a_j, y_j) \geq 0$ for each $x_j \in \{G, B\}$ and (a_j, y_j). In summary,

Lemma 2 *There exist $\bar{u} > 0$ and $\left(\bar{u}_i^{x_j}\right)_{i \in I, x_j \in \{G,B\}}$ such that, for each $i \in I$, the following three properties hold:*

1. *For each $\alpha \in \Delta(A)$, there exists $\pi_i[\alpha] : A_j \times Y_j \to \left(-\frac{\bar{u}}{4}, \frac{\bar{u}}{4}\right)$ that makes any action optimal for player i: $u_i(a) + \mathbb{E}[\pi_i[\alpha](a_j, y_j) \mid a] = u_i(\alpha)$ for all $a \in A$. This implies $\mathbb{E}[\pi_i[\alpha](a_j, y_j) \mid \alpha] = 0$. Moreover, for each (a_j, y_j), $\pi_i[\alpha](a_j, y_j)$ is continuous in α.*

2. *For each $x_j \in \{G, B\}$, there exists $\pi_i^{x_j} : A_j \times Y_j \to \left(-\frac{\bar{u}}{4}, \frac{\bar{u}}{4}\right)$ satisfying $u_i(a) + \mathbb{E}\left[\pi_i^{x_j}(a_j, y_j) \mid a\right] = \bar{u}_i^{x_j}$ for all $a \in A$ and $\text{sign}(x_j)\pi_i^{x_j}(a_j, y_j) \geq 0$ for each $x_j \in \{G, B\}$ and $(a_j, y_j) \in A_j \times Y_j$.*

3. *$\bar{u}$ is sufficiently large: $\max_{i \in I, a \in A} |u_i(a)| + \max_{i \in I, x_j \in \{G,B\}} \left|\bar{u}_i^{x_j}\right| < \bar{u}$ and $\max_{i \in I, x_j \in \{G,B\}} \left|\bar{u}_i^{x_j}\right| < \frac{\bar{u}}{4}$.*

Proof. See Appendix A.3. ■

We fix $\pi_i[\alpha]$, $\pi_i^{x_j}$, $\bar{u} > 0$, and $\left(\bar{u}_i^{x_j}\right)_{i \in I, x_j \in \{G,B\}}$ so that Lemma 2 holds.

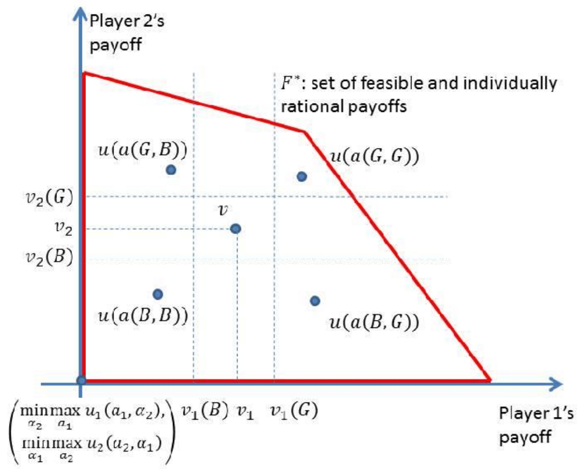

Figure 1: How to take $a(x)$

6.2 Basic Actions

We define an action profile $a(x)$ for each $x \in \{G, B\}^2$. As will be seen, $a(x)$ is the action profile which is taken with a high probability on equilibrium path when the players take $\sigma(x)$. Specifically, we fix $(a(x))_{x\in\{G,B\}^2}$ such that, for each $i \in I$ and $x_j \in \{G, B\}$, we have $\text{sign}(x_j)(v_i - u_i(a(x))) > 0$.[12] See Figure 1 for the illustration of $a(x)$. This definition of $a(x)$ is the same as Hörner and Olszewski (2006). Given $(a(x))_{x\in\{G,B\}^2}$, we perturb $a_i(x)$ to $\alpha_i^\rho(x)$ so that player i takes each action with probability no less than $\rho > 0$:

$$\alpha_i^\rho(x) \equiv (1 - (|A_i| - 1)\rho)\, a_i(x) + \rho \sum_{a_i \neq a_i(x)} a_i. \tag{11}$$

In addition, given player i's minimaxing strategy $\alpha_i^{\min}$, we also perturb $\alpha_i^{\min}$:

$$\alpha_i^{\min,\rho} = (1 - |A_i|\rho)\, \alpha_i^{\min} + \rho \sum_{a_i \in A_i} a_i. \tag{12}$$

[12] Action profiles that satisfy the desired inequalities may not exist. However, if $\dim(F^*) \geq 2$ (otherwise, $\text{int}(F^*) = \emptyset$ and Theorem 1 is vacuously true), then there always exist an integer z and 2^z finite sequences $\{a_1(x), \ldots, a_z(x)\}_{x\in\{G,B\}^2}$ such that each vector $w_i(x)$, the average payoff vector over the sequence $\{a_1(x), \ldots, a_z(x)\}_{x\in\{G,B\}^N}$, satisfies the appropriate inequalities. The construction that follows must then be modified by replacing each action profile $a(x)$ by the finite sequence of action profiles $\{a_1(x), \ldots, a_z(x)\}_{x\in\{G,B\}^N}$. Details are omitted since this is the same as Hörner and Olszewski (2006).

Let $v_i^{*,\rho} \equiv \max_{a_i \in A_i} u_i(a_i, \alpha_j^{\min,\rho})$ be the perturbed minimax value. Given $\alpha_i^{\min,\rho}$, we define

$$\alpha_i^{*,\rho}(x) = \begin{cases} \alpha_i^\rho(x) & \text{if } x_i = G, \\ \alpha_i^{\min,\rho} & \text{if } x_i = B. \end{cases} \tag{13}$$

By definition of $a(x)$ and $v \in \text{int}\,(F^*)$, we have

$$\max\left\{\max_{a_i \in A_i} u_i(a_i, \alpha_j^{\min}), \max_{x:x_j=B} u_i(a(x))\right\} < v_i < \min_{x:x_j=G} u_i(a(x)) \text{ for all } i \in I.$$

If we take ρ sufficiently small, then the targeted payoff v_i is in the interval of the payoffs induced by $\alpha_i^\rho(x)$ and $\alpha_i^{\min,\rho}$: For a sufficiently small $\rho_{\text{payoff}} > 0$, for each $\rho < \rho_{\text{payoff}}$, we have

$$\max\left\{v_i^{*,\rho}, \max_{x:x_j=B} u_i(\alpha^\rho(x))\right\} < v_i < \min_{x:x_j=G} u_i(\alpha^\rho(x)) \text{ for all } i \in I.$$

Hence, there exist $v_i(x_j)$ and $u_i(x_j)$ such that, for each $\rho < \rho_{\text{payoff}}$, we have

$$\max\left\{v_i^{*,\rho}, \max_{x:x_j=B} u_i(\alpha^\rho(x))\right\} < u_i(B) < v_i(B) < v_i < v_i(G) < u_i(G) < \min_{x:x_j=G} u_i(\alpha^\rho(x)) \text{ for all } i \in I.$$

Given $\bar{u}$ fixed in Section 6.1, there exists L such that, for each $\rho < \rho_{\text{payoff}}$, we have

$$\begin{cases} u_i(x_j) - u_i(\alpha^\rho(x)) \leq -\frac{\bar{u}}{L} & \text{if } x_j = G, \\ u_i(x_j) - \max\left\{u_i(\alpha^\rho(x)), \max_{a_i \in A_i} u_i(a_i, \alpha_i^{\min,\rho})\right\} \geq \frac{\bar{u}}{L} & \text{if } x_j = B. \end{cases}$$

Fix such L. Given $\bar{u}$ and L, we fix sufficiently small $\eta > 0$ such that, for each $x_j \in \{G, B\}$, we have

$$(15 + 8L)\,\eta\left\{|u_i(x_j)| + L\bar{u} + \left|\bar{u}_i^{x_j}\right|\right\} < |u_i(x_j) - v_i(x_j)|\,.$$

In summary, we have proven the following lemma:

Lemma 3 *Given $v \in \text{int}\,(F^*)$ and $\bar{u} \in \mathbb{R}$, there exist $(a(x))_{x \in \{G,B\}^2}$, $\rho_{\text{payoff}} > 0$, $(v_i(x_j), u_i(x_j))_{i \in I, x_j \in \{G,B\}}$, $L \in \mathbb{N}$, and $\eta > 0$ such that, for each $i \in I$ and $\rho < \rho_{\text{payoff}}$, we have*

$$\max\left\{v_i^{*,\rho}, \max_{x:x_j=B} u_i(\alpha^\rho(x))\right\} < u_i(B) < v_i(B) < v_i < v_i(G) < u_i(G) < \min_{x:x_j=G} u_i(\alpha^\rho(x)), \tag{14}$$

$$\begin{cases} u_i(x_j) - u_i(\alpha^\rho(x)) \le -\frac{\bar{u}}{L} & \text{if } x_j = G, \\ u_i(x_j) - \max\left\{u_i(\alpha^\rho(x)), \max_{a_i \in A_i} u_i(a_i, \alpha_i^{\min,\rho})\right\} \ge \frac{\bar{u}}{L} & \text{if } x_j = B, \end{cases} \tag{15}$$

and

$$(15 + 8L)\,\eta\left\{|u_i(x_j)| + L\bar{u} + \left|\bar{u}_i^{x_j}\right|\right\} < |u_i(x_j) - v_i(x_j)| \text{ for each } x_j \in \{G, B\}. \tag{16}$$

We fix $(a(x), \alpha_i^\rho(x), \alpha_i^{*,\rho}(x))_{x \in \{G,B\}^2}$ for each ρ, $\alpha_i^{\min,\rho}$ for each ρ, $\rho_{\text{payoff}} > 0$, $(v_i(x_j), u_i(x_j))_{i \in I, x_j \in \{G,B\}}$, $L \in \mathbb{N}$, and $\eta > 0$ so that Lemma 3 holds. In addition, we also fix two different actions $a_i(G), a_i(B) \in A_i$ arbitrarily with $a_i(G) \neq a_i(B)$. Moreover, let $\alpha_i^{\text{mix}} = \frac{1}{|A_i|}\sum_{a_i \in A_i} a_i$ be the random strategy of player i.

6.3 Conditional Independence

For the reason to be explained in Section 15, we want to construct S, $\sigma_i^{\mathbb{S}(t)}$, ϕ_j, q_G, q_B, $\pi_i^{\text{c.i.}}$, and $\varepsilon_{\text{strict}}$ with the following properties: In some period $t \in \mathbb{N}$, player j takes a random strategy α_j^{mix} and player i takes some action $a_{i,t}$ and observes $y_{i,t}$. After period t, there are S periods assigned to period t, denoted by $\mathbb{S}(t) \in \mathbb{N}^S$ with $|\mathbb{S}(t)| = S$. Player j takes α_j^{mix} in each period, while player i takes some pure strategy $\sigma_i^{\mathbb{S}(t)} : \bigcup_{s \in \mathbb{S}(t)} (a_{i,t}, y_{i,t}) \cup (a_{i,\tau}, y_{i,\tau})_{\tau \in \mathbb{S}(t), \tau \le s-1} \to A_i$ in $\mathbb{S}(t)$, which depends on her history in periods t and $\mathbb{S}(t)$.

Based on player j's history in t and $\mathbb{S}(t)$, player j calculates a function $\phi_i : (A_j \times Y_j)^{S+1} \to [0,1]$. The realization of ϕ_j statistically infers what action player i takes in period t: For some q_G and q_B with $q_G - \frac{1}{2} = \frac{1}{2} - q_B > 0$, we have

$$\mathbb{E}\left[\phi_j\left((a_{j,t}, y_{j,t}) \cup (a_{j,\tau}, y_{j,\tau})_{\tau \in \mathbb{S}(t)}\right) \mid \alpha_j^{\text{mix}}, a_{i,t}, y_{i,t}, \sigma_i^{\mathbb{S}(t)}\right] = \begin{cases} q_G & \text{if } a_{i,t} = a_i(G), \\ \frac{1}{2} & \text{if } a_{i,t} \neq a_i(G), a_i(B), \\ q_B & \text{if } a_{i,t} = a_i(B). \end{cases}$$

Importantly, the expected value of ϕ_j is *conditionally independent* of $y_{i,t}$: In period t, from player i's perspective, if she (rationally) expects that she will take $\sigma_i^{\mathbb{S}(t)}$, then the expected value of ϕ_j does not depend on $y_{i,t}$.

To incentivize player i to take $\sigma_i^{\mathbb{S}(t)}$, player j gives the reward function $\pi_i^{\text{c.i.}}\left((a_{j,t}, y_{j,t}) \cup (a_{j,\tau}, y_{j,\tau})_{\tau \in \mathbb{S}(t)}\right)$ (c.i. stands for conditional independence) such that, for each $(a_{i,t}, y_{i,t})$, if a pure strategy σ_i is different from $\sigma_i^{\mathbb{S}(t)}$ on equilibrium path, then

$$\begin{aligned}&\mathbb{E}\left[\pi_i^{\text{c.i.}}\left((a_{j,t},y_{j,t})\cup(a_{j,\tau},y_{j,\tau})_{\tau\in\mathbb{S}(t)}\right)\mid\alpha_j^{\text{mix}},a_{i,t},y_{i,t},\sigma_i^{\mathbb{S}(t)}\right]\\&-\mathbb{E}\left[\pi_i^{\text{c.i.}}\left((a_{j,t},y_{j,t})\cup(a_{j,\tau},y_{j,\tau})_{\tau\in\mathbb{S}(t)}\right)\mid\alpha_j^{\text{mix}},a_{i,t},y_{i,t},\sigma_i\right]\geq\varepsilon_{\text{strict}}\end{aligned}\tag{17}$$

for some $\varepsilon_{\text{strict}}>0$. (Here, we ignore the instantaneous utility.)

Moreover, we make sure that, given the equilibrium strategy $\sigma_i^{\mathbb{S}(t)}$, at the timing of taking $a_{i,t}$, the value does not depend on $a_{i,t}$: For each $a_{i,t}\in A_i$,

$$\mathbb{E}\left[\pi_i^{\text{c.i.}}\left((a_{j,t},y_{j,t})\cup(a_{j,\tau},y_{j,\tau})_{\tau\in\mathbb{S}(t)}\right)\mid\alpha_j^{\text{mix}},a_{i,t},\sigma_i^{\mathbb{S}(t)}\right]=0.$$

The following lemma ensures that we can find S, $\sigma_i^{\mathbb{S}(t)}$, ϕ_j, q_G, q_B, $\pi_i^{\text{c.i.}}$, and $\varepsilon_{\text{strict}}$ to satisfy the conditions above:

Lemma 4 *There exist $S\in\mathbb{N}$, $\varepsilon_{\text{strict}}>0$, $\bar{u}^{\text{c.i.}}>0$, and $1>q_G>q_B>0$ with $q_G-\frac{1}{2}=\frac{1}{2}-q_B>0$ such that, for each $i\in I$, $t\in\mathbb{N}$, and $\mathbb{S}(t)$, there exist $\sigma_i^{\mathbb{S}(t)}:\bigcup_{s\in\mathbb{S}(t)}(a_{i,t},y_{i,t})\cup(a_{i,\tau},y_{i,\tau})_{\tau\in\mathbb{S}(t),\tau\leq s-1}\to A_i$, $\phi_j:(A_j\times Y_j)^{S+1}\to[0,1]$, and $\pi_i^{\text{c.i.}}:(A_j\times Y_j)^{S+1}\to\left[-\bar{u}^{\text{c.i.}},\bar{u}^{\text{c.i.}}\right]$ such that the following properties hold:*

1. *Take any pure strategy of player i, denoted by $\sigma_i:\bigcup_{s\in\mathbb{S}(t)}(a_{i,t},y_{i,t})\cup(a_{i,\tau},y_{i,\tau})_{\tau\in\mathbb{S}(t),\tau\leq s-1}\to A_i$. For each $(a_{i,t},y_{i,t})$, if there exists $h_i=(a_{i,t},y_{i,t})\cup(a_{i,\tau},y_{i,\tau})_{\tau\in\mathbb{S}(t),\tau\leq s-1}$ for some $s\in\mathbb{S}(t)$ such that (a) h_i is reached by the equilibrium strategy $\sigma_i^{\mathbb{S}(t)}$ with a positive probability, and (b) $\sigma_i(h_i)\neq\sigma_i^{\mathbb{S}(t)}(h_i)$ (σ_i is an on-path deviation), then the continuation payoff from h_i is decreased by at least $\varepsilon_{\text{strict}}$:*

$$\begin{aligned}&\mathbb{E}\left[\pi_i^{\text{c.i.}}\left((a_{j,t},y_{j,t})\cup(a_{j,\tau},y_{j,\tau})_{\tau\in\mathbb{S}(t)}\right)\mid\alpha_j^{\text{mix}},a_{i,t},y_{i,t},\sigma_i^{\mathbb{S}(t)},h_i\right]\\&-\mathbb{E}\left[\pi_i^{\text{c.i.}}\left((a_{j,t},y_{j,t})\cup(a_{j,\tau},y_{j,\tau})_{\tau\in\mathbb{S}(t)}\right)\mid\alpha_j^{\text{mix}},a_{i,t},y_{i,t},\sigma_i,h_i\right]\geq\varepsilon_{\text{strict}}.\end{aligned}$$

2. *The conditional independence property holds: For each $a_{i,t}\in A$ and $y_{i,t}\in Y_i$,*

$$\mathbb{E}\left[\phi_j\left((a_{j,t},y_{j,t})\cup(a_{j,\tau},y_{j,\tau})_{\tau\in\mathbb{S}(t)}\right)\mid\alpha_j^{\text{mix}},a_{i,t},y_{i,t},\sigma_i^{\mathbb{S}(t)}\right]=\begin{cases}q_G & \text{if } a_{i,t}=a_i(G)\\ \frac{1}{2} & \text{if } a_{i,t}\neq a_i(G),a_i(B)\\ q_B & \text{if } a_{i,t}=a_i(B).\end{cases}$$

3. The expected value does not depend on $a_{i,t}$: For all $a_{i,t} \in A_i$.

$$\mathbb{E}\left[\pi_i^{\text{c.i.}}\left((a_{j,t}, y_{j,t}) \cup (a_{j,\tau}, y_{j,\tau})_{\tau \in \mathbb{S}(t)}\right) \mid \alpha_j^{\text{mix}}, a_{i,t}, \sigma_i^{\mathbb{S}(t)}\right] = 0.$$

Proof. See Appendix A.4. ■

Let us provide the intuition of the proof. Given player i's history in period t, $(a_{i,t}, y_{i,t})$, player i is asked to "report" $(a_{i,t}, y_{i,t})$ to player j in periods $\mathbb{S}(t)$. For a moment, imagine that player i sends the message via cheap talk. Given player i's message $(\hat{a}_{i,t}, \hat{y}_{i,t})$, suppose player j gives a reward

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

Here, $\mathbf{1}_{a_{j,t},y_{j,t}}$ is $|A_j|\,|Y_j|$-dimensional vector whose element corresponding to $(a_{j,t}, y_{j,t})$ is one and the other elements are zero. On the other hand, $\mathbb{E}\left[\mathbf{1}_{a_{j,t},y_{j,t}} \mid \alpha_j^{\text{mix}}, \hat{a}_{i,t}, \hat{y}_{i,t}\right]$ is the conditional expectation of $\mathbf{1}_{a_{j,t},y_{j,t}}$ given that player j takes α_j^{mix} and player i observes $(\hat{a}_{i,t}, \hat{y}_{i,t})$. Throughout the paper, we use Euclidean norm.

From player i's perspective, ignoring the instantaneous utility as in (17), she wants to minimize

$$\mathbb{E}\left[\left\|\mathbf{1}_{a_{j,t},y_{j,t}} - \mathbb{E}\left[\mathbf{1}_{a_{j,t},y_{j,t}} \mid \alpha_j^{\text{mix}}, \hat{a}_{i,t}, \hat{y}_{i,t}\right]\right\|^2 \mid \alpha_j^{\text{mix}}, a_{i,t}, y_{i,t}\right] \tag{18}$$

by taking $(\hat{a}_{i,t}, \hat{y}_{i,t})$ optimally. As will be seen in Lemma 16, given Assumption 4, we can show that $(\hat{a}_{i,t}, \hat{y}_{i,t}) = (a_{i,t}, y_{i,t})$ (telling the truth) is the unique optimal strategy.[13] Since $\hat{a}_{i,t} = a_{i,t}$, player j has enough information to create ϕ_j to satisfy Claim 2 of Lemma 4.

Without cheap talk, player i takes a message by taking an action sequence in $\mathbb{S}(t)$ and player j infers what action sequence/message player i sends from player j's history in $\mathbb{S}(t)$. Player i wants to minimize (18) with $(\hat{a}_{i,t}, \hat{y}_{i,t})$ replaced with player j's inference of the message. By taking $S = |\mathbb{S}(t)|$ sufficiently long, ex ante (after period t but before $\mathbb{S}(t)$), we can make sure that given player i's equilibrium strategy $\sigma_i^{\mathbb{S}(t)}$ to maximize the reward, player j infers $(a_{i,t}, y_{i,t})$ correctly with a high probability. Since Claim 2 of Lemma 4 requires conditional independence at the end of period t, this ex ante high probability is sufficient to create $\phi_j\left((a_{j,t}, y_{j,t}) \cup (a_{j,\tau}, y_{j,\tau})_{\tau \in \mathbb{S}(t)}\right)$ to satisfy Claim 2.

The strictness (Claim 1) can be achieved as follows: In the last period in $\mathbb{S}(t)$, if there is player

[13] The objective function is called "the scoring rule" in statistics.

i's history with which player i is indifferent, then player j can break the tie by giving a small reward for player i's specific action based on $(a_{j,\tau}, y_{j,\tau})$ with τ being the last period in $\mathbb{S}(t)$. We can make sure that this tie-breaking reward is small enough not to affect the strictness of the incentives for player i's histories after which player i had the strict incentive originally and not to affect player i's incentive to minimize (18) so much (that is, the ex ante probability of player j infering $(a_{j,t}, y_{j,t})$ is still high).

Then, we can proceed by backward induction. Whenever player j breaks a tie for some period τ, it does not affect the strict incentives in the later periods $\tau' > \tau$ since the reward to break a tie in a certain period τ based on $(a_{j,\tau}, y_{j,\tau})$ will be sunk in the later periods $\tau' > \tau$.

Finally, since player j can statistically identify player i's action $a_{i,t}$ from player j's signal $y_{j,t}$, by creating a reward function solely based on $y_{j,t}$ in order to reward (or punish) $a_{i,t}$ which gives a low value (or high value) in $\mathbb{S}(t)$, we can make sure that the expected value does not depend on $a_{i,t}$ to satisfy Claim 3. Since the reward based solely on $y_{j,t}$ is sunk in $\mathbb{S}(t)$, this reward does not affect player i's incentive in $\mathbb{S}(t)$.

We now fix S, $\sigma_i^{\mathbb{S}(t)}$, ϕ_j, q_G, q_B, $\pi_i^{\text{c.i.}}$, and $\varepsilon_{\text{strict}}$ so that Lemma 4 holds.

7 Structure

Given the variables fixed in Section 6, we pin down the structure of (the strategy in) each review phase/finitely repeated game with $T \in \mathbb{N}$ being a parameter. The review phase is divided into blocks, and the block is divided into rounds. See Figure 2 for illustration.

First, given each player i's state x_i, the players coordinate on x in order to take $a(x)$ depending on x with a high probability. To this end, at the beginning of the finitely repeated game, they play the coordination block.

In particular, players first coordinate on x_1. First, player 1 sends $x_1 \in \{G, B\}$ to player 2 by taking actions, spending $T^{\frac{1}{2}}$ periods with a large T. We call these $T^{\frac{1}{2}}$ periods "round $(x_1, 1)$," and let $\mathbb{T}(x_1, 1)$ be the set of periods in round $(x_1, 1)$. (In this step, we focus on the basic structure and postpone the explanation of the formal strategy in each round until Section 13.)

Throughout the paper, we ignore the integer problem since it is easily dealt with by replacing each variable n with the smallest integer no less than n.

Second, player 1 sends x_1 to player 2, spending $T^{\frac{2}{3}}$ periods. We call these $T^{\frac{2}{3}}$ periods "round

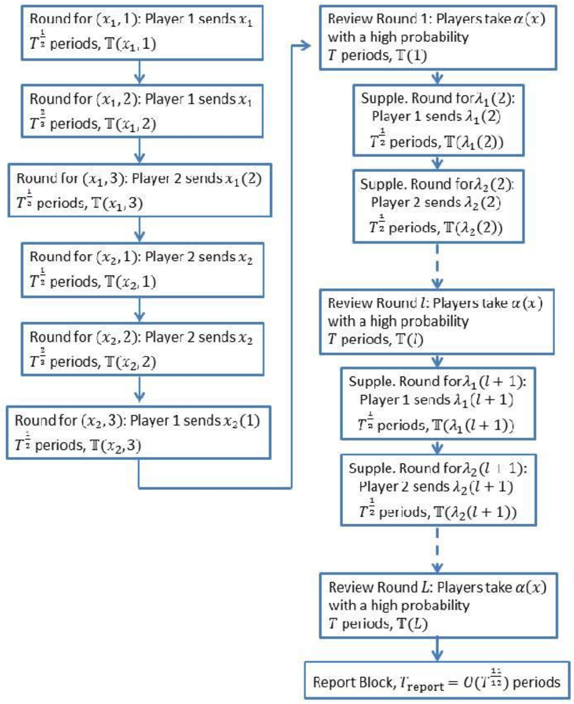

Figure 2: Structure of the review phase

$(x_1, 2)$," and let $\mathbb{T}(x_1, 2)$ be the set of periods in round $(x_1, 2)$. (We will explain in Section 13.2 why player 1 sends x_1 twice and why round $(x_1, 2)$ is longer.)

Based on rounds $(x_1, 1)$ and $(x_1, 2)$, player 2 creates the inference of x_1, denoted by $x_1(2) \in \{G, B\}$. (In general, we use subscript to denote the original owner of the variable and index in the parenthesis to denote a player who makes the inference of the variable. For example, variable$_j(i)$ means that player j knows variable$_j$, and that variable$_j(i)$ is player i's inference of variable$_j$.)

Third, player 2 sends $x_1(2)$ to player 1, spending $T^{\frac{1}{2}}$ periods. We call these $T^{\frac{1}{2}}$ periods "round $(x_1, 3)$," and let $\mathbb{T}(x_1, 3)$ be the set of periods in round $(x_1, 3)$. Based on rounds $(x_1, 1)$, $(x_1, 2)$, and $(x_1, 3)$, player 1 creates the inference of x_1, denoted by $x_1(1) \in \{G, B\}$. (This inference $x_1(1)$ may be different from her original state x_1.)

Once players are done with coordinating on x_1, they coordinate on x_2. The way to coordinate on x_2 is the same as the way to coordinate on x_1, with players' indices reversed.

After the coordination on x is done, the players play L "main blocks." In each main block, first, the players play a T-period review round. With a large T, we can recover the precision of the monitoring by the law of large numbers. After each review round, the players coordinate on what strategy they should take in the next review round, depending on the history (as explained in the

Introduction). This coordination is done by each player i sequentially sending a binary message $\lambda_i(l+1) \in \{G, B\}$, which summarizes the history at the end of the review round. (The precise meaning of $\lambda_i(l+1) \in \{G, B\}$ will be explained in Section 13.4.)

In total, for $l = 1, ..., L-1$, the players first play review round l for T periods. Let $\mathbb{T}(l)$ be the set of review round l. Then, player 1 sends $\lambda_1(l+1) \in \{G, B\}$, spending $T^{\frac{1}{2}}$ periods. We call these $T^{\frac{1}{2}}$ periods "the supplemental round for $\lambda_1(l+1)$," and let $\mathbb{T}(\lambda_1(l+1))$ be the set of periods in it. After player 1, player 2 sends $\lambda_2(l+1) \in \{G, B\}$, spending $T^{\frac{1}{2}}$ periods. "The supplemental round for $\lambda_2(l+1)$" and the set $\mathbb{T}(\lambda_2(l+1))$ are similarly defined. We call these three rounds "main block l." Once main block l is over, the players play review round $l+1$, recursively.

After the last review block L, since the players do not need to coordinate on the future strategy any more, main block L consists only of review round L.

Given this structure, we can chronologically order all the rounds in the coordination and main blocks, and name them round 1, round 2, ..., and round R. Here,

$$R \equiv 6 + 3(L-1) + 1 \tag{19}$$

is the total number of rounds in the coordination and main blocks. For example, round 1 is equivalent to round $(x_1, 1)$, round 2 is equivalent to round $(x_1, 2)$, and so on. Given such a chronological order, when we say $r \leq l$, this means that round r is review round l or a round chronologically before review round l. Similarly, $r < l$ means that round r is a round chronologically before (but not equal to) review round l. In addition, let $\mathbb{T}(r)$ be the set of periods in round r; for example, $\mathbb{T}(r) = \mathbb{T}(x_1, 1)$, and $t(r) + 1$ is the first period of round r.

Finally, the players play the report block, where the players send the summary statistics of the history in the coordination and main blocks. As will be seen in Section 15, we use this block to fine-tune the reward function. This round lasts for T_{report} periods with

$$\begin{aligned} T_{\text{report}} \quad \equiv \quad & 1+(S+1)\,T^{\frac{1}{2}}+(S+1)\,T^{\frac{1}{2}}\sum_{r=1}^{R}|\mathbb{T}(r)|^{\frac{1}{3}}\log_2\left(1+|\mathbb{T}(r)|^{\frac{2}{3}|A_1||Y_1|}\right) \\ & +\left((S+1)T^{\frac{1}{2}}+1\right)\sum_{r=1}^{R}\log_2|\mathbb{T}(r)|^{\frac{1}{3}}+(S+1)T^{\frac{1}{4}}\sum_{r=1}^{R}|\mathbb{T}(r)|^{\frac{2}{3}}\left(1+\log_2|A_1|\,|Y_1|\right) \\ & +(S+1)T^{\frac{1}{2}}\sum_{r=1}^{R}|\mathbb{T}(r)|^{\frac{1}{3}}\log_2|\mathbb{T}(r)|^{\frac{2}{3}|A_2||Y_2|}+\left((S+1)T^{\frac{1}{2}}+1\right)\sum_{r=1}^{R}\log_2|\mathbb{T}(r)|^{\frac{1}{3}} \\ & +(S+1)T^{\frac{1}{4}}\sum_{r=1}^{R}|\mathbb{T}(r)|^{\frac{2}{3}}\log_2|A_2|\,|Y_2|\,. \end{aligned} \tag{20}$$

Note that T_{report} is of order $T^{\frac{11}{12}}$.

We have now pinned down the structure of the review phase, with T being a parameter. For a sufficiently large T, the payoffs from the review round determine the equilibrium payoff from the review phase since the length of the review rounds is much longer than the other rounds/blocks:

Lemma 5 *Let*

$$T_P(T)=\underbrace{4T^{\frac{1}{2}}+2T^{\frac{2}{3}}}_{\textit{coordination block}}+\underbrace{LT}_{\textit{review round}}+\underbrace{(L-1)\,2T^{\frac{1}{2}}}_{\textit{supplemental round}}+\underbrace{T_{\text{report}}}_{\textit{report block, } O(T^{\frac{11}{12}})}$$

be the length of the review phase. We have $\lim_{T\to\infty}\frac{\textit{length of the review rounds}}{\textit{length of the review phase}}\equiv\lim_{T\to\infty}\frac{LT}{T_P(T)}=1.$

Given player i's history $h_i^{\mathbb{T}(r)}\equiv(a_{i,t},y_{i,t})_{t\in\mathbb{T}(r)}$ in round r in the coordination and main blocks, let

$$\begin{aligned} f_i[h_i^{\mathbb{T}(r)}] \quad & \equiv \quad \left(f_i[h_i^{\mathbb{T}(r)}](a_i,y_i)\right)_{(a_i,y_i)\in A_i\times Y_i} \\ & \equiv \quad \left(\frac{\#\left\{t\in\mathbb{T}\left(r\right):\left(a_{i,t},y_{i,t}\right)=\left(a_i,y_i\right)\right\}}{\left|\mathbb{T}\left(r\right)\right|}\right)_{(a_i,y_i)\in A_i\times Y_i} \end{aligned} \tag{21}$$

be the vector expression of the frequency of periods in round r where player i has $(a_{i,t},y_{i,t})=(a_i,y_i)$.

It will be useful to consider the following randomization of player i: For each round r, player i picks a period $t_i^{\text{exclude}}(r)\in\mathbb{T}\left(r\right)$ randomly: $t_i^{\text{exclude}}(r)=t$ with probability $\frac{1}{|\mathbb{T}(r)|}$ for each $t\in\mathbb{T}\left(r\right)$, independently of player i's history. Let $\mathbb{T}_i^{\text{include}}\left(r\right)\equiv\mathbb{T}\left(r\right)\setminus t_i^{\text{exclude}}(r)$ be the periods other than

$t_i^{\text{exclude}}(r)$, and let

$$f_i^{\text{include}}[h_i^{\mathbb{T}(r)}] \equiv \left(f_i^{\text{include}}[h_i^{\mathbb{T}(r)}](a_i, y_i)\right)_{(a_i,y_i)\in A_i\times Y_i} \equiv \left(\frac{\#\left\{t \in \mathbb{T}_i^{\text{include}}(r) : (a_{i,t}, y_{i,t}) = (a_i, y_i)\right\}}{|\mathbb{T}(r)| - 1}\right)_{(a_i,y_i)\in A_i\times Y_i} \tag{22}$$

be the frequency in $\mathbb{T}_i^{\text{include}}(r)$. As will be seen in Section 13, player i decides the continuation play in the subsequent rounds based only on $f_i^{\text{include}}[h_i^{\mathbb{T}(r)}]$. This implies that player j never learns $(a_{i,t_i^{\text{exclude}}(r)}, y_{i,t_i^{\text{exclude}}(r)})$ by observing her own private signals informative about player i's continuation play. This property will be important in Section 15.

Given $f_i[h_i^{\mathbb{T}(r)}]$, let

$$f_i[h_i^{\leq l}] \equiv \left(\begin{array}{c} f_i[h_i^{\mathbb{T}(x_1,1)}], f_i[h_i^{\mathbb{T}(x_1,2)}], f_i[h_i^{\mathbb{T}(x_1,3)}], f_i[h_i^{\mathbb{T}(x_2,1)}], f_i[h_i^{\mathbb{T}(x_2,2)}], f_i[h_i^{\mathbb{T}(x_2,3)}], \\ \left(f_i[h_i^{\mathbb{T}(\tilde{l})}], f_i[h_i^{\mathbb{T}\left(\lambda_1(\tilde{l}+1)\right)}], f_i[h_i^{\mathbb{T}\left(\lambda_2(\tilde{l}+1)\right)}]\right)_{\tilde{l}=1}^{l-1}, f_i[h_i^{\mathbb{T}(l)}] \end{array}\right) \tag{23}$$

be player i's frequency of the history at the end of review round l; and let $f_i[h_i^{<l}]$ be the frequency which exclude $f_i[h_i^{\mathbb{T}(l)}]$ from $f_i[h_i^{\leq l}]$ (that is, the frequency at the beginning of review round l). $f_i^{\text{include}}[h_i^{\leq l}]$ and $f_i^{\text{include}}[h_i^{<l}]$ are similarly defined. On the other hand, let $h_i^{\leq l}$ and $h_i^{<l}$ be player i's histories at the end and beginning of review round l, respectively. Similarly, let $h_i^{\leq r}$ and $h_i^{<r}$ be player i's histories at the end and beginning of round r, respectively.

Finally, let h_j^{report} be player j's history in the report block.

8 Road Map to Prove (5)–(8)

We have fixed $\pi_i[\alpha]$, $\pi_i^{x_j}$, $\bar{u} > 0$, $\left(\bar{u}_i^{x_j}\right)_{i\in I, x_j\in\{G,B\}}$, $(a(x), \alpha_i^{\rho}(x), \alpha_i^{\rho,*}(x))_{x\in\{G,B\}^2}$ for each ρ, $\alpha_i^{\min,\rho}$ for each ρ, $\rho_{\text{payoff}} > 0$, $(v_i(x_j), u_i(x_j))_{i\in I, x_j\in\{G,B\}}$, $L \in \mathbb{N}$, $\eta > 0$, $a_i(G)$, $a_i(B)$, α_i^{mix}, S, $\sigma_i^{\mathbb{S}(t)}$, ϕ_j, q_G, q_B, $\pi_i^{\text{c.i.}}$, and $\varepsilon_{\text{strict}}$ in Section 6, and then defined the structure in the finitely repeated game and $T_P(T)$ in Section 7. Given this structure, we are left to pin down $T \in \mathbb{N}$, $\{\{\sigma_i(x_i)\}_{x_i\in\{G,B\}}\}_{i\in I}$, and $\{\{\pi_i(x_j, h_j^{T_P+1})\}_{x_j\in\{G,B\}}\}_{i\in I}$ to satisfy (5)–(8) with fixed $\{v_i(x_j)\}_{i\in I, x_j\in\{G,B\}}$ and $T_P(T)$.

Since the proof is long and complicated, we first offer the overview in Section 9. Then, we prove two modules which may be of their own interests. Section 10 defines the module for player

j to send a binary message $m \in \{G, B\}$ to player i. This module will be used in the equilibrium construction so that the players can coordinate on the continuation play by sending messages via actions. Section 11 defines the module that will be used for the equilibrium construction of the review round. The proof for each module is relegated to the online appendix. The road map of how to use these modules in the equilibrium construction will be postponed after Section 11.

9 Overview of the Equilibrium Construction

As Lemma 5 ensures, the equilibrium payoff is determined by the payoffs in the review rounds. Hence, in this section, we focus on how to define the strategy and reward function in the review rounds.

9.1 Heuristic Reward Function

For a moment, suppose that the players can coordinate on the true x: $x(i) = x(j) = x$. Take $\rho < \rho_{\text{payoff}}$. Suppose that the players take $(\alpha_i^\rho(x(i)), \alpha_j^\rho(x(j))) = \alpha^\rho(x)$ in each review round and that player i's reward function is

$$LT\{v_i(x_j) - u_i(\alpha^\rho(x))\} + \sum_{l=1}^{L} \sum_{t \in T(l)} \pi_i[\alpha^\rho(x)](a_{j,t}, y_{j,t}). \tag{24}$$

Claim 1 of Lemma 2 ensures that player i's incentive is satisfied and that her equilibrium value is $v_i(x_j)$. Moreover, since $\mathbb{E}[\pi_i[\alpha^\rho(x)](a_{j,t}, y_{j,t}) \mid \alpha^\rho(x)] = 0$, the law of large numbers ensures that $\sum_{t=1}^{T_P} \pi_i[\alpha^\rho(x)](a_{j,t}, y_{j,t})$ is near zero with a high probability. Together with (14), we conclude that self generation is satisfied with a high probability. See Figure 3 for illustration.

However, self generation needs to be satisfied after each possible realization of player j's history, and after an erroneous realization of player j's history, self generation would be violated if we used (24). For example, consider prisoners' dilemma

	C	D
C	$2, 2$	$-1, 3$
D	$3, -1$	$0, 0$

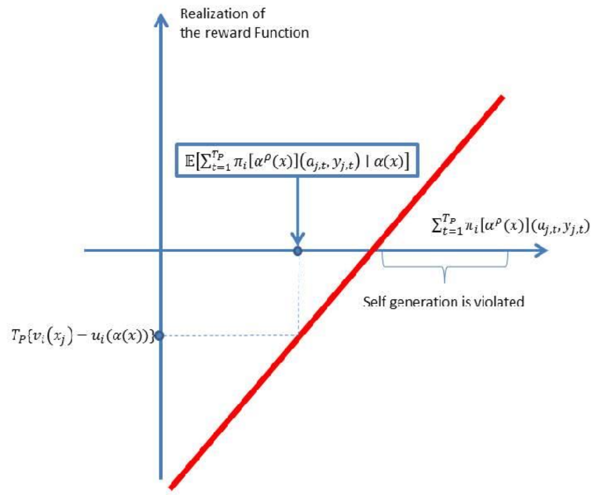

Figure 3: Illustration of the heuristic reward with $x_j = G$

where, for each $j \in I$, signal structure is $Y_j \in \{0_j, 1_j\}$, $q_j(1_j \mid C_i, a_j) = .6$, and $q_j(1_j \mid D_i, a_j) = .4$ for each $a_j \in \{C, D\}$. Then, $\pi_i[\alpha^\rho(x)](a_j, y_j) = 5 \times 1_{\{y_j = 1_j\}} - 5 + u_i(\alpha^\rho(x))$, where $1_{\{y_j=1_j\}} = 1$ if $y_j = 1_j$ and $1_{\{y_j=1_j\}} = 0$ otherwise. If $x = (G, G)$ and v is close to the mutual cooperation payoff $(2, 2)$, then both $v_i(x_j)$ and $u_i(\alpha^\rho(x))$ are near 2 by Figure 1. Hence, if player j observes $y_j = 1_j$ excessively often (say, almost T_P periods), then (24) is equal to

$$LT\{\underbrace{v_i(x_j) - u_i(\alpha^\rho(x))}_{\text{both of them are near 2}}\} + 5 \times \underbrace{\#\{t : y_{j,t} = 1_j\}}_{\text{near } T_P} - 5LT + \underbrace{u_i(\alpha^\rho(x))}_{\text{near 2}} LT > 0.$$

If signals were conditionally independent, then the fix suggested by Matsushima (2004) would work: If (24) violates self generation, then player j gives player i the reward of zero. Since signals are conditionally independent, player i after each history believes that such an event happens very rarely. (Since player i puts a small yet positive belief on (24) violating self generation, Matsushima uses pure strategy $a(x)$ and gives a strict incentive to follow $a_i(x)$, so that player i wants to take $a_i(x)$ even though the reward is zero with a small probability.)

However, with conditionally dependent signals, this fix does not work. For example, if player i's signals and player j's signals are positively correlated, after player i observes a lot of 1_i, she starts to believe that (24) violates self generation with a high probability. After such an event, incentive compatibility would not be satisfied if we gave the reward of zero.

9.2 Type-1, Type-2, and Type-3 Reward Functions

To deal with this problem, we define the overall reward is the summation of the reward for each review round l (we need more modification for the formal proof):

$$\pi_i(x_j, h_j^{T_P+1}) = LT\left\{u_i(x_j) - u_i(\alpha(x))\right\} + \sum_{l=1}^{L} \pi_i^{\text{review}}(x_j, h_j^{\mathbb{T}(l)}, l),$$

where $\pi_i^{\text{review}}(x_j, h_j^{\mathbb{T}(l)}, l)$ is player i's reward for review round l. We use $u_i(x_j)$ instead of $v_i(x_j)$ here to keep some slack in self generation for the later modifications (see (14) for why $u_i(x_j)$ gives us more slack). Now, player i's value without further adjustment is $u_i(x_j)$. The shape of $\pi_i^{\text{review}}(x_j, h_j^{\mathbb{T}(l)}, l)$ can be either type-1, type-2, or type-3.

The type-1 reward function is the same as heuristic reward: $\pi_i^{\text{review}}(x_j, h_j^{\mathbb{T}(l)}, l) = \sum_{t \in \mathbb{T}(l)} \pi_i[\alpha^\rho(x)]$ $(a_{j,t}, y_{j,t})$. By Lemma 2, the maximum realization of the absolute value of the type-1 reward per round is $\frac{\bar{u}}{4}T$. Hence, as long as the realized absolute value of the type-1 reward per round is no more than $\frac{\bar{u}}{4L}T$ until review round l, no matter what realization happens in review round $l+1$, the total absolute value at the end of review round $l+1$ is

$$\left| \sum_{\hat{l}=1}^{l+1} \sum_{t \in \mathbb{T}(\hat{l})} \pi_i[\alpha^\rho(x)](a_{j,t}, y_{j,t}) \right| \leq \frac{\bar{u}}{4L}T \times l + \frac{\bar{u}}{4}T \leq \frac{\bar{u}}{2}T.$$

Together with (15), this means that self generation is not an issue.

Therefore, as long as $\left| \sum_{t \in \mathbb{T}(\hat{l})} \pi_i[\alpha^\rho(x)](a_{j,t}, y_{j,t}) \right| \leq \frac{\bar{u}}{4L}T$ for each $\hat{l} = 1, ..., l$, player j uses the type-1 reward in review round $l+1$. Let $\lambda_j(l+1) \in \{G, B\}$ such that $\lambda_j(l+1) = G$ if and only if $\left| \sum_{t \in \mathbb{T}(\hat{l})} \pi_i[\alpha^\rho(x)](a_{j,t}, y_{j,t}) \right| \leq \frac{\bar{u}}{4L}T$ for each $\hat{l} = 1, ..., l$. Note that $\lambda_j(l+1)$ depends on player j's history $h_j^{\leq l}$. (Presicely speaking, we should write $\lambda_j(l+1)(h_j^{\leq l})$, but we omit $h_j^{\leq l}$ for simplicity.)

On the other hand, if $\lambda_j(l+1) = B$, then if player j used the type-1 reward function for review round $l+1$, then self generation might be violated. Hence, player j takes $\alpha_j^{*,\rho}(x(j))$ (which is equal to $\alpha_j^{*,\rho}(x)$ if $x(j) = x$) and gives a constant reward. (15) ensures that there exists a constant reward such that (i) if coordination goes well and player i takes $BR_i(\alpha_j^{*,\rho}(x(j)))$, then player i is indifferent between "player j taking $\alpha_j^\rho(x(j))$ and using the type-1 reward" and "player j taking $\alpha_j^{*,\rho}(x(j))$ and using the constant reward" (and so the switch of the reward function does not affect player i's incentive in the previous rounds) and that (ii) self generation is satisfied. We call this constant reward "the type-2 reward function."

In the equilibrium construction, player i creates an inference of $\lambda_j(l+1)$, $\lambda_j(l+1)(i) \in \{G, B\}$, which depends on player i's history $h_i^{<l+1}$ at the beginning of review round $l+1$. Intuitively, $\lambda_j(l+1)(i) = G$ means player i believes that $\lambda_j(l+1) = G$ and $\lambda_j(l+1)(i) = B$ means she believes $\lambda_j(l+1) = B$. Player i takes $\alpha_i^\rho(x(i))$ if $\lambda_j(l+1)(i) = G$ and takes $BR_i(\alpha_j^{*,\rho}(x(i)))$ if $\lambda_j(l+1)(i) = B$ since player j with $\lambda_j(l+1)$ takes $\alpha_j^{*,\rho}(x(j))$ (which is equal to $\alpha_j^{*,\rho}(x(i))$ if the coordination goes well: $x(i) = x(j)$) and the reward is type-2 and constant. (Here, we ignore player i's role to control player j's payoff. As player j takes an action based on $\lambda_j(l+1)$ in order to control player i's payoff, player i also takes her action based on not only $\lambda_j(l+1)(i)$ but also $\lambda_i(l+1)$ to control player j's payoff. We ignore this complication in this section.)

We say that the coordination goes well if $x(i) = x(j) = x$ and "$\lambda_j(l+1)(i) = B$ whenever $\lambda_j(l+1) = B$". The above discussion means that, if the coordination goes well, then player i's strategy is optimal. (One may wonder why her strategy is optimal if $\lambda_j(l+1)(i) = B$ and $\lambda_j(l+1) = G$. Note that player j with $\lambda_j(l+1) = G$ uses the type-1 reward, and Lemma 2 ensures that any strategy of player i maximizes the summation of the instantaneous utility and reward function. Hence, any strategy of player i is optimal as long as $\lambda_j(l+1) = G$.)

Finally, player j sometimes uses the following "type-3" reward:

$$\pi_i^{\text{review}}(x_j, h_j^{\mathbb{T}(l)}, l) = \sum_{t \in \mathbb{T}(l)} \pi_i^{x_j}(a_{j,t}, y_{j,t}). \tag{25}$$

By Claim 2 of Lemma 2, this reward always makes player i indifferent between any action profile (player i's incentive is irrelevant) and satisfies self generation, but the value $\bar{u}_i^{x_j}$ may be very different from the targeted value $v_i(x_j)$ (and so promise keeping may become an issue). We will make sure that the type-3 reward is used with a small probability so that we will not violate promise keeping.

Moreover, we will also make sure that player i cannot deviate to affect player j's decision of using the type-3 reward function (since otherwise, depending on whether $\bar{u}_i^{x_j}$ is larger or smaller than the value with the type-1 or type-2 reward, player i may want to deviate to affect the decision). That is, the distribution of the event that player j uses the type-3 reward does not depend on player i's strategy.

Further, whenever player j has $\lambda_i(l)(j)$ (that is, player j believes that her reward function is constant and takes a static best response to player i's action), player j has determined to use the type-3 reward for player i, so that player i can ignore the possibility of $\lambda_i(l)(j) = B$.

9.3 Miscoordination of the Continuation Strategy

Since players coordinate on the continuation strategy based on private signals, it is possible that coordination does not go well. There are following two possibilities. First, players coordinate on the same $x(i) = x(j)$, but this is different from true x. Recall that x_j controls player i's payoff. Hence, as long as players coordinate on the same $x(i) = x(j)$ and $x_j(i) = x_j(j) = x_j$, player i's payoff is equal to $u_i(x_j)$. We will make sure that player j with $x_j(j) \neq x_j$ uses the type-3 reward so that, as long as players coordinate on the same $x(i) = x(j)$, either player i's payoff is equal to $u_i(x_j)$ or player i's incentive is irrelevant. Since the type-3 reward is not used often, by adjusting the reward slightly, we can make sure that player i's ex ante equilibrium payoff is $v_i(x_j)$.

Second, player i has $x(i) \neq x(j)$ or "$\lambda_j(l+1)(i) = G$ if $\lambda_j(l+1) = B$". We will make sure that, player i with history $h_i^{\leq l+1}$ (at the end of review round $l+1$) believes

$$\Pr\left(\left\{\begin{array}{c} \{x(i) \neq x(j) \vee \{\lambda_j(l+1)(i) = G \wedge \lambda_j(l+1) = B\}\} \\ \wedge \left\{\pi_i^{\text{review}}(x_j, h_j^{\mathbb{T}(l+1)}, l+1) \text{ is not type-3 reward}\right\} \end{array}\right\} \mid x_j, h_i^{\leq l+1}\right) \leq \exp(-T^{\frac{1}{3}}). \quad (26)$$

Since the expected difference of player i's payoff for different actions is zero with the type-3 reward function, this means that player i's conditional expectation of the gain of changing her action to deal with the miscoordination is very small. (Note that (26) holds, conditioning on $h_i^{\leq l+1}$, which includes player i's history in review round $l+1$.) That is, without further adjustment, the equilibrium would be ε-*sequential equilibrium*, where deviation gain after each history is no more than ε being of order $\exp(-T^{\frac{1}{3}})$.

9.4 Adjustment and Report Block

To make the equilibrium strategy exactly optimal, we further modify the reward function as follows: The above discussion implies that, if we changed player i's reward function for review round l so that

$$
\begin{aligned}
& \pi_i^{\text{target}}(x_j, h_i^{\leq l}, h_j^{\leq l}, l) \\
= \; & 1_{\left\{\{x(i)\neq x(j) \vee \{\lambda_j(l)(i)=G \wedge \lambda_j(l)=B\}\} \wedge \{\text{type-3 reward is not used in } \pi_i^{\text{review}}(x_j, h_j^{\mathbb{T}(l)}, l)\}\right\}} \times \text{type-1 reward function} \\
& + \left(1 - 1_{\left\{\{x(i)\neq x(j) \vee \{\lambda_j(l)(i)=G \wedge \lambda_j(l)=B\}\} \wedge \{\text{type-3 reward is not used in } \pi_i^{\text{review}}(x_j, h_j^{\mathbb{T}(l)}, l)\}\right\}}\right) \times \pi_i^{\text{review}}(x_j, h_j^{\mathbb{T}(l)}, l)
\end{aligned}
$$

then player i's equilibrium strategy would be optimal. (Recall that Lemma 2 ensures that the type-1 reward function makes each strategy of player i optimal. Hence, player i does not need to worry about miscoordination.) Note that such a reward would depend on player i's history $h_i^{\leq l}$ as well, since $x(i)$ and $\lambda_j(l)(i)$ depends on $h_i^{\leq l}$. Moreover, (26) ensures that

$$
\mathbb{E}\left[\pi_i^{\text{target}}(x_j, h_i^{\leq l}, h_j^{\leq l}, l) \mid x_j, h_i^{\leq l}\right] - \mathbb{E}\left[\pi_i^{\text{review}}(x_j, h_j^{\mathbb{T}(l)}, l) \mid x_j, h_i^{\leq l}\right] \tag{27}
$$

is of order $\exp(-T^{\frac{1}{3}})$ with $l+1$ replaced with l.

If player j can construct a reward function $\pi_i^{\text{adjust}}(x_j, h_j^{\text{report}}, l)$, which depends on player j's history in the report block h_j^{report}, such that, from player i's perspective at the end of review round l, the expected value of $\pi_i^{\text{adjust}}(x_j, h_j^{\text{report}}, l)$ given $(x_j, h_i^{\leq l})$ is equal to (27), then by the law of iterated expectation, player i in review round l wants to maximize

$$
\mathbb{E}\left[\pi_i^{\text{review}}(x_j, h_j^{\mathbb{T}(l)}, l) + \pi_i^{\text{adjust}}(x_j, h_j^{\text{report}}, l) \mid x_j, h_i^{\leq l}\right] = \mathbb{E}\left[\pi_i^{\text{target}}(x_j, h_i^{\leq l}, h_j^{\leq l}, l) \mid x_j, h_i^{\leq l}\right]
$$

and so player i's equilibrium strategy is incentive compatible.

To this end, we define $\pi_i^{\text{adjust}}(x_j, h_j^{\leq l}, h_j^{\text{report}}, l)$ so that

$$
\begin{aligned}
& \mathbb{E}\left[\mathbb{E}\left[\pi_i^{\text{adjust}}(x_j, h_j^{\leq l}, h_j^{\text{report}}, l) \mid x_j, h_i^{\leq L}, \sigma_i^{\text{report}}|_{h_i^{\leq L}}\right] \mid x_j, h_i^{\leq l}\right] \\
= \; & \mathbb{E}\left[\pi_i^{\text{target}}(x_j, h_i^{\leq l}, h_j^{\leq l}, l) \mid x_j, h_i^{\leq l}\right] - \mathbb{E}\left[\pi_i^{\text{review}}(x_j, h_j^{\mathbb{T}(l)}, l) \mid x_j, h_i^{\leq l}\right].
\end{aligned} \tag{28}
$$

Here, $\sigma_i^{\text{report}}|_{h_i^{\leq L}}$ is player i's equilibrium strategy in the report block given her history at the end of the main block, $h_i^{\leq L}$. As will be seen in Section 15, we construct player i's reward function in the report block so that player i's optimal strategy $\sigma_i^{\text{report}}|_{h_i^{\leq L}}$ depends on her history in the coordination and main blocks, in particular, on $h_i^{\leq l}$. Then, player j's history in the report block is correlated with $h_i^{\leq l}$. Therefore, we can construct a function $\pi_i^{\text{adjust}}(x_j, h_j^{\leq l}, h_j^{\text{report}}, l)$ such that (28) holds.

From now on, before defining the strategy and reward function formally, we introduce two modules that will be useful for the equilibrium construction.

10 Module to Send a Binary Message

We explain how player j sends a binary message $m \in \{G, B\}$. Specifically, we fix $m \in \{G, B\}$ and define how player j sends m to player i in $\mathbb{T}$, where $\mathbb{T} \subset \mathbb{N}$ is the set of periods in which player j sends m. For example, if we take $\mathbb{T} = \mathbb{T}(x_j, 1)$ and $m = x_j$, then the following module is the one with which player j sends x_j in round $(x_j, 1)$.

Recall that, for each $i \in I$, $a_i(G), a_i(B) \in A_i$ with $a_i(G) \neq a_i(B)$, and $\eta > 0$ are fixed in Section 6. Given $m \in \{G, B\}$, player j (sender) takes actions as follows: With ρ^{send} to be determined in Lemma 6, player j picks one of the following mixtures of her actions at the beginning:

1. [Send: Regular] With probability $1-\eta$, player j picks $\bar{\alpha}_j^{\rho^{\text{send}}}(m) \equiv \left(1 - (|A_j| - 1)\rho^{\text{send}}\right) a_j(m) + \rho^{\text{send}} \sum_{a_j \neq a_j(m)} a_j$. That is, player i takes $a_j(m)$ (action corresponding to the message m) with a high probability.

2. [Send: Opposite] With probability $\frac{\eta}{2}$, player j picks $\bar{\alpha}_j^{\rho^{\text{send}}}(\hat{m}) \equiv \left(1 - (|A_j| - 1)\rho^{\text{send}}\right) a_j(\hat{m}) + \rho^{\text{send}} \sum_{a_j \neq a_j(\hat{m})} a_j$ with $\hat{m} = \{G, B\} \setminus \{m\}$. That is, player j takes an action as if the true message were $\hat{m} \neq m$.

3. [Send: Mixture] With probability $\frac{\eta}{2}$, player j picks $\bar{\alpha}_j^{\rho^{\text{send}}}(M) \equiv \frac{1}{2}\bar{\alpha}_j^{\rho^{\text{send}}}(G) + \frac{1}{2}\bar{\alpha}_j^{\rho^{\text{send}}}(B)$. That is, player j takes an action as if she mixed two messages G and B.

Let $\alpha_j(m) \in \{\bar{\alpha}_j^{\rho^{\text{send}}}(G), \bar{\alpha}_j^{\rho^{\text{send}}}(B), \bar{\alpha}_j^{\rho^{\text{send}}}(M)\}$ be the realization of the mixture. Given $\alpha_j(m)$, player j takes $a_{j,t}$ *i.i.d.* across periods according to $\alpha_j(m)$. (That is, the mixture over $\bar{\alpha}_j^{\rho^{\text{send}}}(G)$, $\bar{\alpha}_j^{\rho^{\text{send}}}(B)$, and $\bar{\alpha}_j^{\rho^{\text{send}}}(M)$ happens only once at the beginning. Given $\alpha_j(m)$, player j draws $a_{j,t}$ from $\alpha_j(m)$ every period.) On the other hand, player i (receiver) takes a_i according to α_i^{mix} (fully mixed strategy) *i.i.d.* across periods. See Figure 4. In general, thin lines in the figure represent the events that happen with small probabilities.

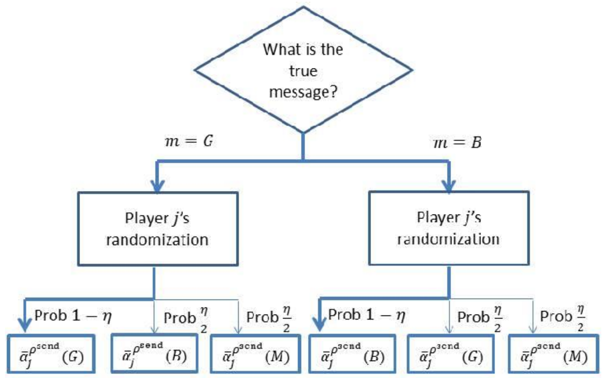

Figure 4: How to send message m

In periods $\mathbb{T}$, each player $n \in I$ observes a history $h_n^{\mathbb{T}} = \{a_{n,t}, y_{n,t}\}_{t\in\mathbb{T}}$. Given $h_n^{\mathbb{T}}$, let

$$f_n\left[h_n^{\mathbb{T}}\right](a_n, y_n) \equiv \frac{\#\{t \in \mathbb{T} : (a_{n,t}, y_{n,t}) = (a_n, y_n)\}}{|\mathbb{T}|}$$

be the frequency of periods with action a_n and signal y_n; and let $f_n\left[h_n^{\mathbb{T}}\right] \equiv \left(f_n\left[h_n^{\mathbb{T}}\right](a_n, y_n)\right)_{a_n, y_n}$ be the vector expression.

In addition, as in $t_i^{\text{exclude}}(r)$ in Section 7, each player n picks one period $t_n^{\text{exclude}} \in \mathbb{T}$ randomly: $t_n^{\text{exclude}} = t$ with probability $\frac{1}{|\mathbb{T}|}$ for each $t \in \mathbb{T}$, independently of player n's history. We define $\mathbb{T}_n^{\text{include}} \equiv \{t \in \mathbb{T} : t \neq t_n^{\text{exclude}}\}$ as the set of periods in $\mathbb{T}$ except for t_n^{exclude}. We define $f_n^{\text{include}}[h_n^{\mathbb{T}}](a_n, y_n)$ and $f_n^{\text{include}}[h_n^{\mathbb{T}}]$ as above, with $\mathbb{T}$ replaced with $\mathbb{T}_n^{\text{include}}$. That is, these are the frequencies in the periods except for t_n^{exclude}.

Player i creates an inference of m based on $f_i^{\text{include}}[h_i^{\mathbb{T}}]$, denoted by $m(f_i^{\text{include}}[h_i^{\mathbb{T}}]) \in \{G, B\}$. Since $f_i^{\text{include}}[h_i^{\mathbb{T}}]$ is in $\Delta(A_i \times Y_i)$, we can see $m(f_i^{\text{include}}[h_i^{\mathbb{T}}])$ as a function $m : \Delta(A_i \times Y_i) \to \{G, B\}$. On the other hand, player j creates a variable $\theta_j(m, f_j^{\text{include}}[h_j^{\mathbb{T}}]) \in \{R, E\}$ based on $f_j^{\text{include}}[h_j^{\mathbb{T}}]$, that is, $\theta_j(m, \cdot) : \Delta(A_j \times Y_j) \to \{R, E\}$.

As will be seen in the proof of Lemma 11, once $\theta_j(m, f_j^{\text{include}}[h_j^{\mathbb{T}}]) = E$ happens in a round where player j sends a message, then player j uses the type-3 reward for player i (that is, player j makes player i indifferent between any actions, as seen in (25)) in the subsequent review rounds. In order to satisfy (26) in the overview, we make sure that, given the true message m and player i's history $h_i^{\mathbb{T}}$, if $m(f_i^{\text{include}}[h_i^{\mathbb{T}}]) \neq m$ (if player i makes a mistake to infer the message), then the conditional probability of $\theta_j(m, f_j^{\text{include}}[h_j^{\mathbb{T}}]) = E$ is sufficiently high (see Claim 2 of Lemma 6 for the formal argument). That is, if player i realizes her mistake to infer player j's message, then player i believes that any strategy is optimal with a high probability.

In addition, as mentioned in Section 9.2, we want to make sure that the type-3 reward is used with a small probability so that we will not violate promise keeping. To this end, we define $\theta_j(m, f_j^{\text{include}}[h_j^{\mathbb{T}}])$ so that $\theta_j(m, f_j^{\text{include}}[h_j^{\mathbb{T}}]) = E$ with a small probability (see Claim 3 of Lemma 6). We use R and E for the realization of θ_j, meaning that R stands for "regular" and E stands for "erroneous."

Moreover, again as mentioned in Section 9.2, we want to make sure that the distribution of the event that player j uses the type-3 reward does not depend on player i's strategy. To this end, we make sure that the distribution of $\theta_j(m, f_j^{\text{include}}[h_j^{\mathbb{T}}])$ does not depend on player i's strategy (again see Claim 3 of Lemma 6).

In addition to $m(f_i^{\text{include}}[h_i^{\mathbb{T}}])$, player i also creates a variable $\theta_i(\text{receive}, f_i^{\text{include}}[h_i^{\mathbb{T}}]) \in \{R, E\}$, that is, $\theta_i(\text{receive}, \cdot) : \Delta(A_i \times Y_i) \to \{R, E\}$. As $\theta_j(m, f_j^{\text{include}}[h_j^{\mathbb{T}}])$, once $\theta_i(\text{receive}, f_i^{\text{include}}[h_i^{\mathbb{T}}]) = E$ happens in a round where player j sends a message, then player i uses the type-3 reward for player j in the subsequent review rounds. Again, we make sure that $\theta_i(\text{receive}, f_i^{\text{include}}[h_i^{\mathbb{T}}]) = E$ happens with a small probability and that the distribution of $\theta_i(\text{receive}, f_i^{\text{include}}[h_i^{\mathbb{T}}])$ does not depend on player j's strategy (see Claim 4 of Lemma 6).

Further, we make sure that player j with $\theta_j(m, f_j^{\text{include}}[h_j^{\mathbb{T}}]) = R$ (that is, if player j does not use the type-3 reward for player i) believes that $m_i(f_i^{\text{include}}[h_i^{\mathbb{T}}]) = m$ or $\theta_i(\text{receive}, f_i^{\text{include}}[h_i^{\mathbb{T}}]) = E$ (that is, player i infers the message correctly or player i uses the type-3 reward for player j) with a high probability (see Claim 5 of Lemma 6).

In particular, with $m = x_j$, this means that as long as player j is not using the type-3 reward for player i, she believes that player i received x_j correctly or player i uses the type-3 reward. Since player j is indifferent between any action in the latter case, in total, this belief incentivizes her to adhere to her own state x_j, as mentioned in Section 9.3 and will be shown in Claim 5 of Lemma 11.

Finally, since Assumption 1 assumes that signal profiles have full support, each player cannot figure out the other player's actions or inferences perfectly (see Claims 6 and 7 of Lemma 6).

In total, we can construct three functions m, θ_i, and θ_j, so that the following lemma holds:

Lemma 6 *There exist $\rho^{\text{send}} > 0$, $\varepsilon_{\text{message}} > 0$, $K_{\text{message}} < \infty$, $m : \Delta(A_i \times Y_i) \to \{G, B\}$, $\theta_i(\text{receive}, \cdot) : \Delta(A_i \times Y_i) \to \{R, E\}$, and $\theta_j(m, \cdot) : \Delta(A_j \times Y_j) \to \{R, E\}$ such that, for a sufficiently large $|\mathbb{T}|$, for each $i \in I$ and $m \in \{G, B\}$, the following claims hold:*

1. *Since t_n^{exclude} is random, the frequency is the sufficient statistic to infer the other players' variables: For each $h_j^{\mathbb{T}} \in (A_j \times Y_j)^{|\mathbb{T}|}$, $m, \hat{m} \in \{G, B\}$, and $\hat{\theta}_i \in \{R, E\}$, we have*

$$\begin{aligned} & \Pr_{i|j}\left(\left\{m(f_i^{\text{include}}[h_i^{\mathbb{T}}]) = \hat{m}_i, \theta_i(\text{receive}, f_i^{\text{include}}[h_i^{\mathbb{T}}]) = \hat{\theta}_i\right\} \mid m, h_j^{\mathbb{T}}\right) \\ = \; & \Pr_{i|j}\left(\left\{m(f_i^{\text{include}}[h_i^{\mathbb{T}}]) = \hat{m}_i, \theta_i(\text{receive}, f_i^{\text{include}}[h_i^{\mathbb{T}}]) = \hat{\theta}_i\right\} \mid m, f_j[h_j^{\mathbb{T}}]\right). \end{aligned}$$

In addition, for each $h_i^{\mathbb{T}} \in (A_i \times Y_i)^{|\mathbb{T}|}$ and $\hat{\theta}_j \in \{R, E\}$, we have

$$\begin{aligned} & \Pr_{j|i}\left(\left\{\theta_j(m, f_j^{\text{include}}[h_j^{\mathbb{T}}]) = \hat{\theta}_j\right\} \mid m, h_i^{\mathbb{T}}\right) \\ = \; & \Pr_{j|i}\left(\left\{\theta_j(m, f_j^{\text{include}}[h_j^{\mathbb{T}}]) = \hat{\theta}_j\right\} \mid m, f_i[h_i^{\mathbb{T}}]\right). \end{aligned}$$

Here, $\Pr_{i|j}\left(\cdot \mid m, h_j^{\mathbb{T}}\right)$ and $\Pr_{j|i}\left(\cdot \mid m, h_i^{\mathbb{T}}\right)$ are induced by the following assumptions: In $\Pr_{i|j}\left(\cdot \mid m, h_j^{\mathbb{T}}\right)$, player j's action sequence $\{a_{j,t}\}_{t\in\mathbb{T}}$ is given by $h_j^{\mathbb{T}}$. On the other hand, player i takes $a_{i,t}$ i.i.d. *across periods according to α_i^{mix}. Given $\{a_t\}_{t\in\mathbb{T}}$, the signal profile y_t is drawn from the conditional joint distribution function $q(y_t \mid a_t)$ for each t, and player j observes her signal $\{y_{j,t}\}_{t\in\mathbb{T}}$. Finally, player i draws t_i^{exclude} randomly. $\Pr_{j|i}\left(\cdot \mid m, h_i^{\mathbb{T}}\right)$ is defined in the same way, with i and j reversed, and α_i^{mix} replaced with $\alpha_j(m)$ given m.*

2. *For all $m \in \{G, B\}$ and $h_i^{\mathbb{T}} \in (A_i \times Y_i)^{|\mathbb{T}|}$, if $m(f_i^{\text{include}}[h_i^{\mathbb{T}}]) \neq m$ (player i misinfers the message), then we have*

$$\Pr_{j|i}\left(\left\{\theta_j(m, f_j^{\text{include}}[h_j^{\mathbb{T}}]) = E\right\} \mid m, h_i^{\mathbb{T}}\right) \geq 1 - \exp(-\varepsilon_{\text{message}} |\mathbb{T}|). \tag{29}$$

3. *For each m and for each strategy of player i denoted by $\sigma_i : \bigcup_{s=0}^{|\mathbb{T}|-1} (A_i \times Y_i)^s \to \Delta(A_i)$, the probability of $\theta_j = R$ does not depend on m or σ_i: $\Pr_{j|i}\left(\left\{\theta_j(m, f_j^{\text{include}}[h_j^{\mathbb{T}}]) = R\right\} \mid m, \sigma_i\right) = 1 - 2\eta$ for each $m \in \{G, B\}$ and σ_i.*

4. *For each strategy of player j denoted by $\sigma_j : \bigcup_{s=0}^{|\mathbb{T}|-1} (A_j \times Y_j)^s \to \Delta(A_j)$, the probability of $\theta_i = R$ does not depend on player j's strategy: $\Pr_{i|j}\left(\left\{\theta_i(\text{receive}, f_i^{\text{include}}[h_i^{\mathbb{T}}]) = R\right\} \mid \sigma_j\right) = 1 - 2\eta$ does not depend on σ_j.*[14]

[14] Here, we omit m from $\Pr_{i|j}\left(\cdot \mid m, h_j^{\mathbb{T}}\right)$ since player j's strategy σ_j and player i's equilibrium strategy α_i^{mix} fully determine the distribution of actions and signals.

5. *For all* $h_j^{\mathbb{T}} \in (A_j \times Y_j)^{|\mathbb{T}|}$, *if* $\theta_j(m, f_j^{\text{include}}[h_j^{\mathbb{T}}]) = R$, *then we have*

$$\begin{aligned} & \Pr_{i|j}\left(\left\{\theta_i(\text{receive}, f_i^{\text{include}}[h_i^{\mathbb{T}}]) = E \vee m(f_i^{\text{include}}[h_i^{\mathbb{T}}]) = m\right\} \mid m, h_j^{\mathbb{T}}\right) \\ \geq \; & 1 - \exp(-\varepsilon_{\text{message}} |\mathbb{T}|). \end{aligned}$$

6. *For all* $m, \hat{m} \in \{G, B\}$ *and* $h_j^{\mathbb{T}} \in (A_j \times Y_j)^{|\mathbb{T}|}$, *any inference of player* i *is possible:*

$$\Pr_{i|j}\left(\left\{m(f_i^{\text{include}}[h_i^{\mathbb{T}}]) = \hat{m}\right\} \mid m, h_j^{\mathbb{T}}\right) \geq \exp(-K_{\text{message}} |\mathbb{T}|).$$

7. *For all* $m \in \{G, B\}$ *and* $h_i^{\mathbb{T}} \in (A_i \times Y_i)^{|\mathbb{T}|}$, *any history of player* i *is possible:* $\Pr\left(h_i^{\mathbb{T}} \mid m\right) \geq \exp(-K_{\text{message}} |\mathbb{T}|)$.

Proof. See Appendix A.6. ■

Let us intuitively explain why Lemma 6 holds. Claim 1 holds once we define $m(f_i^{\text{include}}[h_i^{\mathbb{T}}])$ and $\theta_i(\text{receive}, f_i^{\text{include}}[h_i^{\mathbb{T}}])$ so that they depend only on $f_i^{\text{include}}[h_i^{\mathbb{T}}]$, and define $\theta_j(m, f_j^{\text{include}}[h_j^{\mathbb{T}}])$ so that it depends only on $f_j^{\text{include}}[h_j^{\mathbb{T}}]$. Hence, we concentrate on the other claims.

First, we define $\theta_j(m, f_j^{\text{include}}[h_j^{\mathbb{T}}])$ such that player j has $\theta_j(m, f_j^{\text{include}}[h_j^{\mathbb{T}}]) = R$ only if (and if, except for a small adjustment in the formal proof) all of the following conditions are satisfied:

1. [Regular Mixture Send] Player j picks $\alpha_j(m) = \bar{\alpha}_j^{\rho^{\text{send}}}(m)$.

2. [Regular Action Send] Let $f_j^{\text{include}}[h_j^{\mathbb{T}}](a_j) \equiv \sum_{y_j \in Y_j} f_j^{\text{include}}[h_j^{\mathbb{T}}](a_j, y_j)$ be the frequency of player j's actions. We say that player j's action frequency is regular if, for a small $\varepsilon_{\text{message}} > 0$, we have $\left|f_j^{\text{include}}[h_j^{\mathbb{T}}](a_j) - \alpha_j(m)(a_j)\right| \leq \varepsilon_{\text{message}}$ for each $a_j \in A_j$.

3. [Regular Signal Send] Let $f_j^{\text{include}}[h_j^{\mathbb{T}}](Y_j \mid a_j) \equiv \left(f_j^{\text{include}}[h_j^{\mathbb{T}}](y_j \mid a_j)\right)_{y_j \in Y_j}$ with

$$f_j^{\text{include}}[h_j^{\mathbb{T}}](y_j \mid a_j) \equiv \frac{f_j^{\text{include}}[h_j^{\mathbb{T}}](a_j, y_j)}{f_j^{\text{include}}[h_j^{\mathbb{T}}](a_j)}$$

be the vector-expression of the conditional frequency of player j's signals given player j's action a_j. (We define $f_j^{\text{include}}[h_j^{\mathbb{T}}](y_j \mid a_j) = 0$ if $f_j^{\text{include}}[h_j^{\mathbb{T}}](a_j) = 0$.) On the other hand, let $\text{aff}\left(\{q_j(a_j, a_i)\}_{a_i \in A_i}\right)$ be the affine hull of player j's signal distributions with respect to player

i's actions. We say that player j's signal frequency is regular if, for a small $\varepsilon_{\text{message}} > 0$,

$$d\left(f_j^{\text{include}}[h_j^{\mathbb{T}}](Y_j \mid a_j(m)), \text{aff}\left(\{q_j(a_j(m), a_i)\}_{a_i \in A_i}\right)\right) \leq \varepsilon_{\text{message}}. \tag{30}$$

Here and in what follows, we use Euclidean norm $\|\cdot\|$ and Hausdorff metric d.

Given this definition, we verify Claim 3 of Lemma 6. [Regular Mixture Send] and [Regular Action Send] are solely determined by player j's mixture, which player i cannot affect. Moreover, player i cannot affect the probability of [Regular Signal Send] since $\text{aff}\left(\{q_j(a_j(m), a_i)\}_{a_i \in A_i}\right)$ is the affine hull with respect to player i's actions. (Precisely speaking, taking affine hull ensures that player i cannot change the expected distance

$$\mathbb{E}\left[d\left(f_j^{\text{include}}[h_j^{\mathbb{T}}](Y_j \mid a_j), \text{aff}\left(\{q_j(a_j, a_i)\}_{a_i \in A_i}\right)\right)\right], \tag{31}$$

but does not guarantee that player i cannot change the distribution of the distance

$$\Pr\left(d\left(f_j^{\text{include}}[h_j^{\mathbb{T}}](Y_j \mid a_j), \text{aff}\left(\{q_j(a_j, a_i)\}_{a_i \in A_i}\right)\right)\right). \tag{32}$$

We take care of player i's incentive to change the distribution in the formal proof.) Hence, the probability of $\theta_j(m, f_j^{\text{include}}[h_j^{\mathbb{T}}]) = R$ does not depend on player i's strategy.

Moreover, $\alpha_j(m) = \bar{\alpha}_j^{\rho^{\text{send}}}(m)$ with a high probability, as seen in Figure 3, and by the law of large numbers, [Regular Action Send] and [Regular Signal Send] happen with a high probability given $\alpha_j(m) = \bar{\alpha}_j^{\rho^{\text{send}}}(m)$. Hence, $\theta_j(m, f_j^{\text{include}}[h_j^{\mathbb{T}}]) = R$ with a high probability. Hence, Claim 3 of Lemma 6 holds.

On the other hand, we define $\theta_i(\text{receive}, f_i^{\text{include}}[h_i^{\mathbb{T}}])$ as follows: Let $f_i^{\text{include}}[h_i^{\mathbb{T}}](a_i) \equiv \sum_{y_i \in Y_i} f_i^{\text{include}}[h_i^{\mathbb{T}}](a_i, y_i)$ be the frequency of player i's actions. We say that player i's action frequency is regular if, for a small $\varepsilon_{\text{message}} > 0$, we have $\left|f_i^{\text{include}}[h_i^{\mathbb{T}}](a_i) - \alpha_i^{\text{mix}}(a_i)\right| \leq \varepsilon_{\text{message}}$ for each $a_i \in A_i$. Let [Regular Action Receive] denote this event. Player i has $\theta_i(\text{receive}, f_i^{\text{include}}[h_i^{\mathbb{T}}]) = R$ only if (and if, except for a small adjustment in the formal proof) [Regular Action Receive] holds. Again, we can make sure that the probability of $\theta_i(\text{receive}, f_i^{\text{include}}[h_i^{\mathbb{T}}]) = R$ does not depend on player j's strategy and is high by the law of large numbers. Hence, Claim 4 of Lemma 6 holds.

We now define player i's inference $m(f_i^{\text{include}}[h_i^{\mathbb{T}}])$, so that Claim 2 of Lemma 6 holds. She calculates the log likelihood ratio between $\alpha_j(m) = \bar{\alpha}_j^{\rho^{\text{send}}}(G)$ and $\alpha_j(m) = \bar{\alpha}_j^{\rho^{\text{send}}}(B)$, ignoring the prior of $\alpha_j(m)$:

$$
\begin{aligned}
& \log \frac{\Pr\left(\{y_{i,t}\}_{t\in\mathbb{T}} \mid \left\{\alpha_j(m) = \bar{\alpha}_j^{\rho^{\text{send}}}(G)\right\}, \{a_{i,t}\}_{t\in\mathbb{T}}\right)}{\Pr\left(\{y_{i,t}\}_{t\in\mathbb{T}} \mid \left\{\alpha_j(m) = \bar{\alpha}_j^{\rho^{\text{send}}}(B)\right\}, \{a_{i,t}\}_{t\in\mathbb{T}}\right)} \qquad (33) \\
= \; & |\mathbb{T}| \left\{ \sum_{a_i,y_i} f_i[h_i^{\mathbb{T}}](a_i,y_i) \log q_i(y_i \mid a_i, \bar{\alpha}_j^{\rho^{\text{send}}}(G)) - \sum_{a_i,y_i} f_i[h_i^{\mathbb{T}}](a_i,y_i) \log q_i(y_i \mid a_i, \bar{\alpha}_j^{\rho^{\text{send}}}(B)) \right\}.
\end{aligned}
$$

Related to (33), let

$$
L_i(f_i[h_i^{\mathbb{T}}], G) \equiv \sum_{a_i,y_i} f_i[h_i^{\mathbb{T}}](a_i,y_i) \log q_i(y_i \mid a_i, \bar{\alpha}_j^{\rho^{\text{send}}}(G))
$$

be the log likelihood of $\alpha_j(m) = \bar{\alpha}_j^{\rho^{\text{send}}}(G)$. $L_i(f_i[h_i^{\mathbb{T}}], B)$ and $L_i(f_i[h_i^{\mathbb{T}}], M)$ are defined in the same way, with G replaced with B and M, respectively.

Given these likelihoods, player i creates $m(f_i^{\text{include}}[h_i^{\mathbb{T}}])$ as follows: For some $L_{\text{belief}}^{\text{receive}} > 0$,

1. [Case G] If $L_i(f_i^{\text{include}}[h_i^{\mathbb{T}}], G) > L_i(f_i^{\text{include}}[h_i^{\mathbb{T}}], B) + 2L_{\text{belief}}^{\text{receive}}$, then player i infers that the message is G: $m(f_i^{\text{include}}[h_i^{\mathbb{T}}]) = G$.

2. [Case B] If $L_i(f_i^{\text{include}}[h_i^{\mathbb{T}}], B) > L_i(f_i^{\text{include}}[h_i^{\mathbb{T}}], G) + 2L_{\text{belief}}^{\text{receive}}$, then player i infers that the message is B: $m(f_i^{\text{include}}[h_i^{\mathbb{T}}]) = B$.

3. [Case M] If neither of the above two conditions is satisfied, then player i infers the message randomly: $m(f_i^{\text{include}}[h_i^{\mathbb{T}}]) = G$ with probability $\frac{1}{2}$; and $m(f_i^{\text{include}}[h_i^{\mathbb{T}}]) = B$ with probability $\frac{1}{2}$.

See Figure 5 for the illustration.

Let us fix $L_{\text{belief}}^{\text{receive}} > 0$. Since the maximum likelihood estimator is consistent, there exists

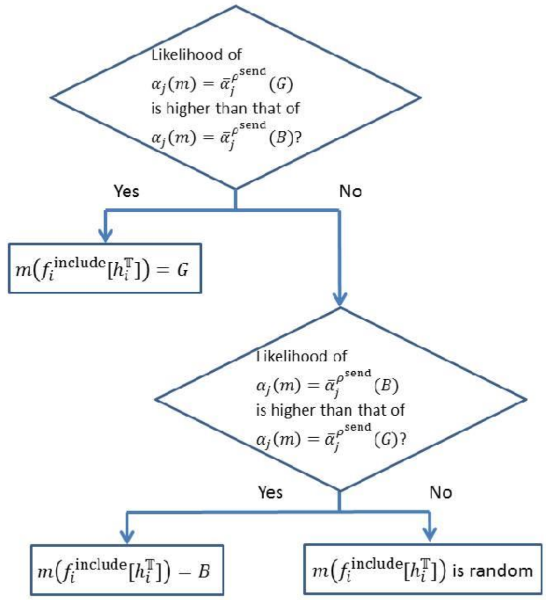

Figure 5: How to infer m

$\bar{\rho}^{\text{send}} > 0$ such that, for sufficiently small $L_{\text{belief}}^{\text{receive}} > 0$, for each $\rho^{\text{send}} < \bar{\rho}^{\text{send}}$, we have

$$\begin{cases} L_i(f_i^{\text{include}}[h_i^{\mathbb{T}}], G) \geq L_i(f_i^{\text{include}}[h_i^{\mathbb{T}}], B) + 3L_{\text{belief}}^{\text{receive}} \\ \text{if } f_i^{\text{include}}[h_i^{\mathbb{T}}](a_i, y_i) = \frac{1}{|A_i|} q_i\left(y_i \mid a_i, \bar{\alpha}_j^{\rho^{\text{send}}}(G)\right) \text{ for each } a_i, y_i \\ \text{(that is, if the frequency of player } i\text{'s history is close to the ex ante mean given } \bar{\alpha}_j^{\rho^{\text{send}}}(G), \\ \text{then the likelihood of } \bar{\alpha}_j^{\rho^{\text{send}}}(G) \text{ is higher than that of } \bar{\alpha}_j^{\rho^{\text{send}}}(B)), \\ L_i(f_i^{\text{include}}[h_i^{\mathbb{T}}], B) \geq L_i(f_i^{\text{include}}[h_i^{\mathbb{T}}], G) + 3L_{\text{belief}}^{\text{receive}} \\ \text{if } f_i^{\text{include}}[h_i^{\mathbb{T}}](a_i, y_i) = \frac{1}{|A_i|} q_i\left(y_i \mid a_i, \bar{\alpha}_j^{\rho^{\text{send}}}(B)\right) \text{ for each } a_i, y_i \\ \text{(that is, the frequency of player } i\text{'s history is close to the ex ante mean given } \bar{\alpha}_j^{\rho^{\text{send}}}(B), \\ \text{then the likelihood of } \bar{\alpha}_j^{\rho^{\text{send}}}(B) \text{ is higher than that of } \bar{\alpha}_j^{\rho^{\text{send}}}(G)). \end{cases} \tag{34}$$

Moreover, since the log-likelihood is strictly concave and $\bar{\alpha}_j^{\rho^{\text{send}}}(M) = \frac{1}{2}\bar{\alpha}_j^{\rho^{\text{send}}}(G) + \frac{1}{2}\bar{\alpha}_j^{\rho^{\text{send}}}(B)$, retaking $\bar{\rho}^{\text{send}} > 0$ sufficiently small if necessary, for sufficiently small $L_{\text{belief}}^{\text{receive}} > 0$, for each $\rho^{\text{send}} < \bar{\rho}^{\text{send}}$, if neither [Case G] nor [Case B] is the case, then $\alpha_j(m) = \bar{\alpha}_j^{\rho^{\text{send}}}(M)$ is more likely than both $\alpha_j(m) = \bar{\alpha}_j^{\rho^{\text{send}}}(G)$ and $\alpha_j(m) = \bar{\alpha}_j^{\rho^{\text{send}}}(B)$: If neither [Case G] or [Case B] holds, then

$$L_i(f_i^{\text{include}}[h_i^{\mathbb{T}}], M) \geq \max\left\{L_i(f_i^{\text{include}}[h_i^{\mathbb{T}}], G), L_i(f_i^{\text{include}}[h_i^{\mathbb{T}}], B)\right\} + 2L_{\text{belief}}^{\text{receive}}. \tag{35}$$

We fix $L_{\text{belief}}^{\text{receive}} > 0$ so that (34) and (35) hold.

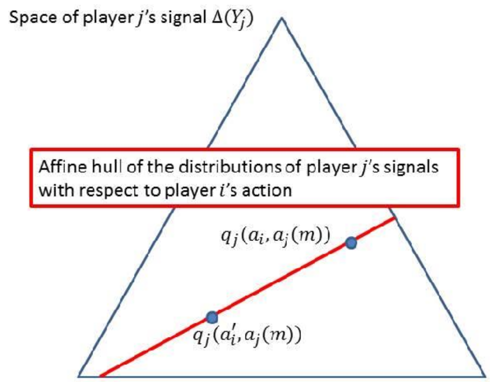

Figure 6: Affine hull of player j's signal observation

Note that with this inference, Claim 2 of Lemma 6 holds. To see why, suppose $m = G$ (the explanation with $m = B$ is the same and so is omitted). If $m(f_i^{\text{include}}[h_i^{\mathbb{T}}]) = B \neq G = m$, then [Case B] or [Case M] is the case. We consider these two cases in the sequel.

[Case B] implies that $\alpha_j(m) = \bar{\alpha}_j^{\rho^{\text{send}}}(B)$ is more likely than $\alpha_j(m) = \bar{\alpha}_j^{\rho^{\text{send}}}(G)$, except for the prior. Since player j takes each of $\bar{\alpha}_j^{\rho^{\text{send}}}(G)$, $\bar{\alpha}_j^{\rho^{\text{send}}}(B)$, and $\bar{\alpha}_j^{\rho^{\text{send}}}(M)$ with probability no less than $\frac{\eta}{2}$ by Figure 4, given $m = G$ and taking the prior into account, the law of large numbers ensures that $\alpha_j(m) \neq \bar{\alpha}_j^{\rho^{\text{send}}}(G)$ with probability of order $1 - \exp(-L_{\text{belief}}^{\text{receive}} |\mathbb{T}|)$. Since [Regular Mixture Send] ensures that $\alpha_j(m) \neq \bar{\alpha}_j^{\rho^{\text{send}}}(G)$ implies that $\theta_j(m, f_j^{\text{include}}[h_j^{\mathbb{T}}]) = E$, this high probability on $\alpha_j(m) \neq \bar{\alpha}_j^{\rho^{\text{send}}}(G)$ implies that player j has $\theta_j(m, f_j^{\text{include}}[h_j^{\mathbb{T}}]) = E$ with probability no less than $1 - \exp(-L_{\text{belief}}^{\text{receive}} |\mathbb{T}|)$. Taking $\varepsilon_{\text{message}} > 0$ sufficiently small, this is sufficient for (29).

If [Case M] is the case, then (35) implies that at least one of $\bar{\alpha}_j^{\rho^{\text{send}}}(B)$ and $\bar{\alpha}_j^{\rho^{\text{send}}}(M)$ is more likely than $\bar{\alpha}_j^{\rho^{\text{send}}}(G)$, except for the prior. The same proof as [Case B] establishes the result.

We now prove Claim 5 of Lemma 6. Recall that we have already fixed $L_{\text{belief}}^{\text{receive}}$ so that (34) and (35) hold. Note that we assume $\theta_j(m, f_j^{\text{include}}[h_j^{\mathbb{T}}]) = R$, which implies [Regular Mixture Send], [Regular Action Send], and [Regular Signal Send].

For a moment, suppose that $\rho^{\text{send}} = 0$ and $\varepsilon_{\text{message}} = 0$ for simplicity. Then, [Regular Mixture Send] and [Regular Action Send] imply that player j takes $\alpha_j(m) = \bar{\alpha}_j^{\rho^{\text{send}}}(m) = a_j(m)$ (the last equality holds only with $\rho^{\text{send}} = 0$). Since she takes $a_j(m)$ for sure, we can see $f_j^{\text{include}}[h_j^{\mathbb{T}}](Y_j \mid a_j(m))$ as her entire history. Moreover, [Regular Signal Send] implies that this frequency is equal to the affine hull aff $\left(\{q_j(a_j(m), a_i)\}_{a_i \in A_i}\right)$ (with $\varepsilon_{\text{message}} = 0$). See Figure 6.

If $f_j^{\text{include}}[h_j^{\mathbb{T}}]$ is close to the ex ante distribution given $\bar{\alpha}_j^{\rho^{\text{send}}}(m)$, then by the law of iterated expectation, player j believes that player i's history $f_i^{\text{include}}[h_i^{\mathbb{T}}]$ is close to the ex ante distribution given $\bar{\alpha}_j^{\rho^{\text{send}}}(m)$. (34) ensures that player j believes that player i has $L(m, f_i^{\text{include}}[h_i^{\mathbb{T}}]) \geq L(\hat{m}, f_i^{\text{include}}[h_i^{\mathbb{T}}]) + 2L_{\text{belief}}^{\text{receive}}$ with $\hat{m} \in \{G, B\} \setminus \{m\}$ and so $m(f_i^{\text{include}}[h_i^{\mathbb{T}}]) = m$ with a high probability.

In particular, there exists $\varepsilon_1 > 0$ such that, if

$$\begin{aligned} & \left\| f_j^{\text{include}}[h_j^{\mathbb{T}}](Y_j \mid a_j(m)) - q_j(a_j(m), \alpha_i^{\text{mix}}) \right\| \\ \leq \; & \varepsilon_1 \text{ (distance from the ex ante mean is small),} \end{aligned} \tag{36}$$

then by the law of large numbers, player j believes that $L(m, f_i^{\text{include}}[h_i^{\mathbb{T}}]) \geq L(\hat{m}, f_i^{\text{include}}[h_i^{\mathbb{T}}]) + 2L_{\text{belief}}^{\text{receive}}$ with $\hat{m} \in \{G, B\} \backslash \{m\}$ (and so $m(f_i^{\text{include}}[h_i^{\mathbb{T}}]) = m$) with probability $1 - \exp(-\varepsilon_{\varepsilon_1, L_{\text{belief}}^{\text{receive}}} |\mathbb{T}|)$, where the coefficient $\varepsilon_{\varepsilon_1, L_{\text{belief}}^{\text{receive}}} > 0$ depends on ε_1 and $L_{\text{belief}}^{\text{receive}}$. We fix such $\varepsilon_1 > 0$. In Figure 7, $\varepsilon_1 > 0$ determines the distance between point A and point D.

On the other hand, suppose (36) does not hold. This means that her signal frequency $f_j^{\text{include}}[h_j^{\mathbb{T}}](Y_j \mid a_j(m))$ is not close to $q_j(a_j(m), \alpha_i^{\text{mix}})$. Since $f_j^{\text{include}}[h_j^{\mathbb{T}}](Y_j \mid a_j(m))$ is in aff $\left(\{q_j(a_j(m), a_i)\}_{a_i \in A_i}\right)$, this means that her signal frequency is skewed toward $q_j(a_j(m), a_i)$ for some $a_i \in A_i$ compared to $q_j(a_j(m), \alpha_i^{\text{mix}})$. Player j believes that such a_i happens significantly more often than the ex ante probability $\alpha_i^{\text{mix}}(a_i) = \frac{1}{|A_i|}$.

Hence, given $\varepsilon_1 > 0$, there exists sufficiently small $\varepsilon_2 > 0$ such that, if (36) does not hold, then there exists $a_i \in A_i$ such that the conditional expectation of the frequency of a_i is not close to $\alpha_i^{\text{mix}}(a_i) = \frac{1}{|A_i|}$: For some $a_i \in A_i$,

$$\begin{aligned} & \left| \mathbb{E}\left[f_i^{\text{include}}[h_i^{\mathbb{T}}](a_i) \mid m, \bar{\alpha}_j^{\rho^{\text{send}}}(m), \alpha_i^{\text{mix}}, h_j^{\mathbb{T}} \right] - \frac{1}{|A_i|} \right| \\ > \; & \varepsilon_2 \text{ (the frequency of player } i\text{'s actions is irregular).} \end{aligned} \tag{37}$$

Fix such $\varepsilon_2 > 0$. In Figure 7, if we take $\varepsilon_2 > 0$ sufficiently smaller than $\varepsilon_1 > 0$, then points B and C will be included in the interval $[A, D]$.

The log likelihood is continuous in perturbation ρ^{send} and frequency $f_i^{\text{include}}[h_i^{\mathbb{T}}]$. In addition, the conditional expectation of $f_i^{\text{include}}[h_i^{\mathbb{T}}]$ is continuous in $f_j^{\text{include}}[h_j^{\mathbb{T}}]$. Hence, there exist sufficiently small $\rho^{\text{send}} > 0$ and $\varepsilon_{\text{message}} > 0$ such that the following claims hold: Given $\theta_j(m, f_j^{\text{include}}[h_j^{\mathbb{T}}]) = R$,

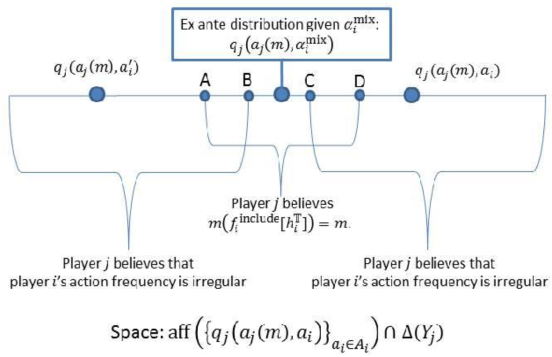

Figure 7: Player j's inference of player i's history

1. If

$$\begin{aligned}&\left\| f_j^{\text{include}}[h_j^{\mathbb{T}}](Y_j \mid a_j(m)) - q_j(a_j(m), \alpha_i^{\text{mix}}) \right\| \\ \leq\ &\varepsilon_1 \text{ (distance from the ex ante mean is small),}\end{aligned}$$

   then player j believes that $m(f_i^{\text{include}}[h_i^{\mathbb{T}}]) = m$ with probability $1 - \exp(-\frac{1}{2}\varepsilon_{\varepsilon_1, L_{\text{belief}}^{\text{receive}}} |\mathbb{T}|)$. For sufficiently small $\varepsilon_{\text{message}} > 0$ compared to $\varepsilon_{\varepsilon_1, L_{\text{belief}}^{\text{receive}}} > 0$, we can say that player j believes that $m(f_i^{\text{include}}[h_i^{\mathbb{T}}]) = m$ with probability $1 - \exp(-\varepsilon_{\text{message}} |\mathbb{T}|)$.

2. Otherwise, there exists $a_i \in A_i$ such that

$$\left| \mathbb{E}\left[f_i^{\text{include}}[h_i^{\mathbb{T}}](a_i) \mid m, \bar{\alpha}_j^{\rho^{\text{send}}}(m), \alpha_i^{\text{mix}}, h_j^{\mathbb{T}} \right] - \frac{1}{|A_i|} \right| > \frac{1}{2}\varepsilon_2.$$

   For sufficiently small $\varepsilon_{\text{message}} > 0$ compared to $\varepsilon_2 > 0$, by the law of large numbers, player j believes that $\left| f_i^{\text{include}}[h_i^{\mathbb{T}}](a_i) - \frac{1}{|A_i|} \right| > \varepsilon_{\text{message}}$ for some $a_i \in A_i$ and so $\theta_i(\text{receive}, f_i^{\text{include}}[h_i^{\mathbb{T}}]) = E$ with probability $1 - \exp(-\varepsilon_{\text{message}} |\mathbb{T}|)$.

In both cases, Claim 5 of Lemma 6 holds.

Finally, since each player takes a fully mixed strategy and Assumption 1 ensures that the distribution of the private signal profile has full support, we have Claims 6 and 7 of Lemma 6. In the working paper, we also consider public monitoring, where both players observe the same signal with probability one. There, we make sure that the same claims hold, only using the fact that players are taking private strategies.

11 Module for the Review Round

In this section, we consider the following $|\mathbb{T}|$-period finitely repeated game, in which $\mathbb{T} \subset \mathbb{N}$ is the (arbitrary) set of periods in this finitely repeated game. In our equilibrium construction later, $\mathbb{T}$ corresponds to the set of periods in the review round. In this section, we fix $x(i) \in \{G, B\}^2$ and $x(j) \in \{G, B\}^2$. As will be seen, $x(i)$ is player i's inference of state profile $x \in \{G, B\}^2$.

Recall that Section 6 defines $(\alpha_i^{\rho}(x))_{i \in I, x \in \{G,B\}^2}$ for each ρ. Given $x(i) \in \{G, B\}^2$, with ρ to be determined in Lemma 7, player i takes $a_{i,t}$ *i.i.d.* across periods according to $\alpha_i^{\rho}(x(i))$.

As in Section 10, let $h_i^{\mathbb{T}} = \{a_{i,t}, y_{i,t}\}_{t \in \mathbb{T}}$ be the history; $f_i\left[h_i^{\mathbb{T}}\right]$ be the frequency; t_i^{exclude} be the period excluded from $\mathbb{T}_i^{\text{include}} \equiv \{t \in \mathbb{T} : t \neq t_i^{\text{exclude}}\}$; and $f_i^{\text{include}}[h_i^{\mathbb{T}}]$ be the frequency in $\mathbb{T}_i^{\text{include}}$.

As in the case with the binary message protocol, each player i creates a variable $\theta_i(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}}]) \in \{R, E\}$ based on $f_i^{\text{include}}[h_i^{\mathbb{T}}]$, that is, $\theta_i(x(i), \cdot) : \Delta(A_i \times Y_i) \to \{R, E\}$. Again, once $\theta_i(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}}]) = E$ happens in a review round, player i uses the type-3 reward function for player j (makes player j indifferent between any action) in the subsequent review rounds. We make sure that $\theta_i(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}}]) = E$ with a small probability and the distribution of $\theta_i(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}}])$ does not depend on player j's strategy (see Claim 2 of Lemma 7 with indices i and j reversed for the formal argument).

In addition, related to the type-1 reward in Section 9.2, each player j calculates

$$\begin{aligned}
\pi_i(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}}]) &\equiv \sum_{t \in \mathbb{T}_j^{\text{include}}} \pi_i[\alpha^{\rho}(x(j))](a_{j,t}, y_{j,t}) \qquad (38) \\
&= \sum_{(a_j, y_j) \in A_j \times Y_j} \pi_i[\alpha^{\rho}(x(j))](a_j, y_j) \times f_j^{\text{include}}[h_j^{\mathbb{T}}](a_j, y_j),
\end{aligned}$$

using the reward function defined in Lemma 2. Except for the fact that it does not include the reward in period t_j^{exclude} and that it uses player j's inference $x(j)$ rather than true state profile x, this function is the same as the type-1 realization of $\pi_i^{\text{review}}(x_j, h_j^{\mathbb{T}(l)}, l)$ with $\mathbb{T} = \mathbb{T}(l)$. Note that this depends only on $f_j^{\text{include}}[h_j^{\mathbb{T}}]$ and $x(j)$ once we fix ρ, since we have fixed $(\pi_i[\alpha](a_j, y_j))_{(a_j, y_j) \in A_j \times Y_j}$ in Lemma 2 for each $\alpha \in \Delta(A)$.

As in Section 9.2, player j will have $\lambda_j(l + 1) = B$ only if $\left|\pi_i(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}}])\right| > \frac{\bar{u}}{4L} |\mathbb{T}|$ happens in review round l. We prove that if $x(i) = x(j)$ (coordination on x goes well) and $\theta_i(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}}]) = R$ (player i does not use the type-3 reward for player j), then player i believes that $\left|\pi_i(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}}])\right| \leq \frac{\bar{u}}{4L} |\mathbb{T}|$ (and so $\lambda_j(l+1) = G$) or $\theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}}]) = E$ (and so player

j uses the type-3 reward function for player i) with a high probability (see Claim 3 of Lemma 7).

This high probability implies the following: As will be seen in Lemma 8, if the coordination does not go well, then as a result of the message exchange about x, player j uses the type-3 reward with a high probability. Together with the argument in the previous paragraph, player i with $\theta_i(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}}]) = R$ believes that $\lambda_j(l+1) = G$ or player j uses the type-3 reward with a high probability. Hence, if we define player i's strategy such that she switches to $\lambda_j(l+1)(i) = B$ only after $\theta_i(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}}]) = E$, then (26) holds.

Given such a transition of $\lambda_j(l+1)$, whenever $\lambda_j(l+1)(i) = B$, player i uses the type-3 reward function for player j. With indices i and j reversed, as mentioned in Section 9.2, player i can condition that whenever $\lambda_i(l+1)(j) = B$, player j uses the type-3 reward.

Finally, we prove that if $x(i) = x(j)$, $\theta_i^{\text{review}}(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}}]) = R$, and $\left|\pi_j(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}}])\right| > \frac{\bar{u}}{4L}\left|\mathbb{T}\right|$, then player i believes that $\theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}}]) = E$ with a high probability (see Claim 4 of Lemma 7). This high probability implies the following: As will be seen in Section 13.4.1, if $x_i = B$ (player i wants to keep player j's equilibrium payoff low), $\theta_i^{\text{review}}(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}}]) = R$ (player i has not yet made player j indifferent), and $\left|\pi_j(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}}])\right| > \frac{\bar{u}}{4L}\left|\mathbb{T}\right|$ (self generation is an issue) in a review round, then player i will minimax player j to keep her payoff low in all the subsequent review rounds. In such a case, player i believes that player j uses the type-3 reward for player i and any strategy (including the minimaxing one) is optimal.

In total, we can construct $\theta_i(x(i), \cdot)$ so that the following lemma holds:

Lemma 7 *Given $\bar{\rho}_{\text{payoff}}$, $\bar{u}$, and L fixed in Section 6.2, there exist $\rho \in (0, \bar{\rho}_{\text{payoff}})$, $\theta_i(x(i), \cdot) : \Delta(A_i \times Y_i) \to \{R, E\}$, and $\varepsilon_{\text{review}} > 0$ such that, for a sufficiently large $|\mathbb{T}|$, the following four properties hold: For each $i \in I$,*

1. *Since t_j^{exclude} is random, the frequency is the sufficient statistic to infer the other players' variables: For each $\hat{\theta}_j \in \{R, E\}$ and $\hat{\pi}_i \in \mathbb{R}$, we have*

$$\begin{aligned}
& \Pr_{j|i}\left(\left\{\theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}}]) = \hat{\theta}_j, \pi_i(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}}]) = \hat{\pi}_i\right\} \mid x(j), h_i^{\mathbb{T}}\right) \\
= \; & \Pr_{j|i}\left(\left\{\theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}}]) = \hat{\theta}_j, \pi_i(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}}]) = \hat{\pi}_i\right\} \mid x(j), f_i[h_i^{\mathbb{T}}]\right).
\end{aligned}$$

Here, $\Pr_{j|i}\left(\cdot \mid x(j), h_i^{\mathbb{T}}\right)$ is induced by the following assumptions: Player i's action sequence $\{a_{i,t}\}_{t \in \mathbb{T}}$ is given by $h_i^{\mathbb{T}}$. On the other hand, player j takes $a_{j,t}$ i.i.d. *across periods according to*

$\alpha_j^{\rho}(x(j))$. Given $\{a_t\}_{t\in\mathbb{T}}$, the signal profile y_t is drawn from the conditional joint distribution function $q(y_t \mid a_t)$ for each t, and player i observes her signal $\{y_{i,t}\}_{t\in\mathbb{T}}$. Finally, player j draws t_j^{exclude} randomly.

2. *For each $x(j) \in \{G, B\}^2$ and for each strategy of player i denoted by $\sigma_i : \bigcup_{s=0}^{|\mathbb{T}|-1} (A_i \times Y_i)^s \to \Delta(A_i)$, the probability of $\theta_j = R$ does not depend on $x(j)$ or σ_i: $\Pr_{j|i}(\{\theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}}]) = R\} \mid x(j), \sigma_i) = 1 - 2\eta$ for each $x(j) \in \{G, B\}^2$ and σ_i.*

3. *For each $x(j) \in \{G, B\}^2$ and $h_i^{\mathbb{T}} \in (A_i \times Y_i)^{|\mathbb{T}|}$, conditional on $x(i) = x(j)$, if $\theta_i^{\text{review}}(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}}]) = R$, then player i believes that either $\pi_i(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}}])$ is near zero or $\theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}}]) = E$ with a high probability:*

$$\Pr_{j|i}\left(\left\{\begin{array}{c} \left|\pi_i(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}}])\right| \leq \frac{\bar{u}}{4L}|\mathbb{T}| \\ \vee \theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}}]) = E \end{array}\right\} \mid x(j), \{x(i) = x(j)\}, h_i^{\mathbb{T}}\right) \geq 1 - \exp(-\varepsilon_{\text{review}}|\mathbb{T}|).$$

4. *For each $x(i) \in \{G, B\}^2$ and $h_i^{\mathbb{T}} \in (A_i \times Y_i)^{|\mathbb{T}|}$, conditional on $x(i) = x(j)$, if $\theta_i^{\text{review}}(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}}]) = R$ and $\left|\pi_j(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}}])\right| > \frac{\bar{u}}{4L}|\mathbb{T}|$, then player i believes that $\theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}}]) = E$ with a high probability:*

$$\Pr_{j|i}\left(\{\theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}}]) = E\} \mid x(j), \{x(i) = x(j)\}, h_i^{\mathbb{T}}\right) \geq 1 - \exp(-\varepsilon_{\text{review}}|\mathbb{T}|).$$

Proof. See Appendix A.7. ■

Let us explain why Lemma 7 holds intuitively. First, we define $\theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}}])$ as follows: Let $f_j^{\text{include}}[h_j^{\mathbb{T}}](a_j)$ be the frequency of player j's actions; and let $f_j^{\text{include}}[h_j^{\mathbb{T}}](Y_j \mid a_j)$ be the conditional frequency of player j's signals given a_j, as in Section 10. Player j has $\theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}}]) = R$ only if (and if, except for a small adjustment in the formal proof) both of the following conditions are satisfied:

1. [Regular Action Review] We say that player j's action frequency is regular if, for a small $\varepsilon_{\text{review}} > 0$, we have $\left|f_j^{\text{include}}[h_j^{\mathbb{T}}](a_j) - \alpha_j^{\rho}(x(j))(a_j)\right| \leq \varepsilon_{\text{review}}$ for each $a_j \in A_j$.

2. [Regular Signal Review] We say that player j's signal frequency is regular if, for a small

$\varepsilon_{\text{review}} > 0$, we have

$$d\left(f_j^{\text{include}}[h_j^{\mathbb{T}}](Y_j \mid a_j(x(j))), \text{aff}\left(\{q_j(a_j(x(j)), a_i)\}_{a_i \in A_i}\right)\right) \leq \varepsilon_{\text{review}}.$$

Given this definition, Claim 1 holds since t_j^{exclude} is random and other variables depend only on $f_j^{\text{include}}[h_j^{\mathbb{T}}]$.

In addition, Claim 2 holds for the following reasons. [Regular Action Review] is solely determined by player j's mixture, which player i cannot affect. Moreover, player i cannot affect the probability of [Regular Signal Review] since aff $\left(\{q_j(a_j(x(j)), a_i)\}_{a_i \in A_i}\right)$ is the affine hull with respect to player i's actions. (The same caution as (31) and (32) is applicable here.)

By the law of large numbers, [Regular Action Review] and [Regular Signal Review] happen with a high probability. Hence, $\theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}}]) = R$ with a high probability.

Given this definition of $\theta_i(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}}])$ and $\theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}}])$, we prove Claim 3 of Lemma 7. Since we condition $x(j) = x(i)$, let $x(j) = x(i) = x$. As in Claim 5 of Lemma 6, with i and j reversed, $a_j(m)$ replaced with $a_i(x)$, ρ^{send} replaced with ρ, $\bar{\alpha}_j^{\rho^{\text{send}}}(m)$ replaced with $\alpha_i^{\rho}(x)$, and $\varepsilon_{\text{message}}$ replaced with $\varepsilon_{\text{review}}$, we have the following: Given $\theta_i(x, f_i^{\text{include}}[h_i^{\mathbb{T}}]) = R$, there are following two cases:

1. $f_i^{\text{include}}[h_i^{\mathbb{T}}](Y_i \mid a_i(x))$ is close to $q_i(a_i(x), \alpha_j^{\rho}(x))$. This means that $f_i^{\text{include}}[h_i^{\mathbb{T}}](Y_i \mid a_i(x))$ is close to the ex ante mean given $(a_i(x), \alpha_j^{\rho}(x))$. For a small perturbation ρ, this means that $f_i^{\text{include}}[h_i^{\mathbb{T}}]$ is close to the ex ante mean given by $(\alpha_i^{\rho}(x), \alpha_j^{\rho}(x))$. By the law of iterated expectation, the conditional expectation of $f_j^{\text{include}}[h_j^{\mathbb{T}}]$ is close to the ex ante mean given $(\alpha_i^{\rho}(x), \alpha_j^{\rho}(x))$. By the law of large numbers, player i believes that $\pi_i(x, f_j^{\text{include}}[h_j^{\mathbb{T}}])$ is close to the ex ante mean given $(\alpha_i^{\rho}(x), \alpha_j^{\rho}(x))$.

   By Claim 1 of Lemma 2, the ex ante mean of $\pi_i(x, f_j^{\text{include}}[h_j^{\mathbb{T}}]) = \sum_{t \in \mathbb{T}_j^{\text{include}}} \pi_i[\alpha^{\rho}(x)](a_{j,t}, y_{j,t})$ given $(\alpha_i^{\rho}(x), \alpha_j^{\rho}(x))$ is zero. Therefore, for sufficiently small $\varepsilon_{\text{review}} > 0$, player i believes that $\left|\pi_i(x, f_j^{\text{include}}[h_j^{\mathbb{T}}])\right| \leq \frac{\bar{u}}{4L} |\mathbb{T}|$ with probability $1 - \exp(-\varepsilon_{\text{review}} |\mathbb{T}|)$.

2. Otherwise, $f_i^{\text{include}}[h_i^{\mathbb{T}}](Y_i \mid a_i(x))$ is skewed toward $q_i(a_i(x), a_j)$ for some $a_j \in A_j$ compared to $q_i(a_i(x), \alpha_j^{\rho}(x))$. Again, for sufficiently small ρ, since player i takes $a_i(x)$ often and $a_i(x)$ and $\alpha_j^{\rho}(x)$ are close to each other, this implies that player i believes that $f_j^{\text{include}}[h_j^{\mathbb{T}}](a_j)$ is

not close to $\alpha_j^\rho(x(j))(a_j)$ for some $a_j \in A_j$ (and so $\theta_j(x, f_j^{\text{include}}[h_j^{\mathbb{T}}]) = E$) with probability $1 - \exp(-\varepsilon_{\text{review}} |\mathbb{T}|)$ for a sufficiently small $\varepsilon_{\text{review}}$.

In both cases, Claim 3 of Lemma 7 holds.

Similarly, we prove Claim 4 of Lemma 7. Since we assume $x(j) = x(i)$ and $\theta_i(x, f_i^{\text{include}}[h_i^{\mathbb{T}}]) = R$, again, there are following two cases: The first is that $f_i^{\text{include}}[h_i^{\mathbb{T}}]$ is close to the ex ante mean given by $(\alpha_i^\rho(x), \alpha_j^\rho(x))$ with $x = x(j) = x(i)$. Then, since the ex ante mean of $\pi_j(x, f_i^{\text{include}}[h_i^{\mathbb{T}}])$ is zero, $\left|\pi_j(x, f_i^{\text{include}}[h_i^{\mathbb{T}}])\right| > \frac{\bar{u}}{4L} |\mathbb{T}|$ is not the case. The remaining case is that player i believes that $\theta_j(x, f_j^{\text{include}}[h_j^{\mathbb{T}}]) = E$.

12 Road Map to Prove (5)–(8) given the Modules

Again, recall that we have fixed $\pi_i[\alpha]$, $\pi_i^{x_j}$, $\bar{u} > 0$, $(\bar{u}_i^{x_j})_{i \in I, x_j \in \{G,B\}}$, $(a(x), \alpha_i^\rho(x), \alpha_i^{*,\rho}(x))_{x \in \{G,B\}^2}$ for each ρ, $\alpha_i^{\min,\rho}$ for each ρ, $\rho_{\text{payoff}} > 0$, $(v_i(x_j), u_i(x_j))_{i \in I, x_j \in \{G,B\}}$, $L \in \mathbb{N}$, $\eta > 0$, $a_i(G)$, $a_i(B)$, α_i^{mix}, S, $\sigma_i^{\mathbb{S}(t)}$, ϕ_j, q_G, q_B, $\pi_i^{\text{c.i.}}$, and $\varepsilon_{\text{strict}}$ in Section 6. Then, we fix the structure of the finitely repeated game in Section 7. The length of the finitely repeated game is $T_P(T)$ with T being a parameter.

Given these variables, we fix $\bar{\alpha}_j^{\rho^{\text{send}}}(G)$, $\bar{\alpha}_j^{\rho^{\text{send}}}(B)$, $\bar{\alpha}_j^{\rho^{\text{send}}}(M)$, $\rho^{\text{send}} > 0$, $\varepsilon_{\text{message}} > 0$, $K_{\text{message}} < \infty$, $m : \Delta(A_i \times Y_i) \to \{G, B\}$, $\theta_i(\text{receive}, \cdot) : \Delta(A_i \times Y_i) \to \{R, E\}$, and $\theta_j(m, \cdot) : \Delta(A_j \times Y_j) \to \{R, E\}$ so that Lemma 6 holds; and we fix $\rho \in (0, \bar{\rho}_{\text{payoff}})$, $\theta_i(x(i), \cdot) : \Delta(A_i \times Y_i) \to \{R, E\}$, and $\varepsilon_{\text{review}} > 0$ such that Lemma 7 holds.

Given these variables/functions, in Sections 13-16, we define the strategy and reward, and then verify (5)–(8).

Specifically, we define player i's strategy in the coordination and main blocks for each T in Section 13. The definition of the strategy $\sigma_i^{\text{report}}|_{h_i^{\leq L}}$ in the report block will be postponed until Section 15.

In Section 14, we define player i's reward function $\pi_i(x_j, h_j^{T_P+1})$ for each T. As will be seen, $\pi_i(x_j, h_j^{T_P+1})$ is the summation of $\bar{\pi}_i(x_j)$, $\pi_i^{x_j}(a_{j,t}, y_{j,t})$, π_i^{review}, π_i^{adjust}, and π_i^{report}. We define the first three elements in Section 14 and will postpone the definition of π_i^{adjust} and π_i^{report} until Section 15.

In Section 15, Lemma 13 defines $\sigma_i^{\text{report}}|_{h_i^{\leq L}}$, π_i^{adjust}, and π_i^{report} for each T, which completes the definition of the strategy $\sigma_i(x_i)$ and the reward $\pi_i(x_j, h_j^{T_P+1})$ for each T.

Finally, in Section 16, we prove that, for a sufficiently large T, the defined strategy and reward satisfy (5)–(8).

The technical proofs for the claims in Sections 13-16 are further relegated to Appendix A (online appendix), and Appendix B contains the list of the frequetly used notation.

13 Strategy in the Coordination and Main Blocks

Sections 13.1 and 13.3 define the strategy in the coordination blocks for x_j and x_i, respectively. In particular, player i creates $x_j(i)$ and $x_i(i)$. Section 13.2 derives properties about $x_j(i)$.

Given $x(i) = (x_i(i), x_j(i))$, in Section 13.4, we define player i's strategy in main block l. As will be seen, her strategy in main block l depends on two variables $\lambda_i(l) \in \{G, B\}$ and $\lambda_j(l)(i) \in \{G, B\}$. Sections 13.4.1, 13.4.2, and 13.4.3 define her strategy in review round l, the supplemental round for $\lambda_j(l+1)$, and that for $\lambda_i(l+1)$, given $\lambda_i(l)$ and $\lambda_i(l)(j)$. Finally, we define the transition of $\lambda_i(l)$ and $\lambda_j(l)(i)$ in Section 13.4.4.

13.1 Strategy in the Coordination Block for x_j

In order to coordinate on x_j, player j sends $x_j \in \{G, B\}$ in rounds $(x_j, 1)$ and $(x_j, 2)$. Given these two rounds, player i creates her inference $x_j(i) \in \{G, B\}$. Given $x_j(i)$, player i sends $x_j(i)$ in round for $(x_j, 3)$.

In particular, in round $(x_j, 1)$, player j sends $x_j \in \{G, B\}$ spending $T^{\frac{1}{2}}$ periods as explained in Section 10 (with m replaced with x_j, and $\mathbb{T} = \mathbb{T}(x_j, 1)$ with $|\mathbb{T}| = T^{\frac{1}{2}}$). Player j constructs $\theta_j(x_j, f_j^{\text{include}}[h_j^{\mathbb{T}(x_j,1)}]) \in \{R, E\}$, and player i constructs $x_j(f_i^{\text{include}}[h_i^{\mathbb{T}(x_j,1)}]) \in \{G, B\}$ and $\theta_i(\text{receive}, f_i^{\text{include}}[h_i^{\mathbb{T}(x_j,1)}]) \in \{R, E\}$. In addition, for each $n \in \{i, j\}$, let $t_n^{\text{exclude}}(x_j, 1)$ be the period such that player n does not use her history in period $t_n^{\text{exclude}}(x_j, 1)$ when she determines the continuation play.

Then, in round $(x_j, 2)$, player j re-sends x_j spending $T^{\frac{2}{3}}$ periods as explained in Section 10 (with m replaced with x_j, and $\mathbb{T} = \mathbb{T}(x_j, 2)$ with $|\mathbb{T}| = T^{\frac{2}{3}}$). Player j constructs $\theta_j(x_j, f_j^{\text{include}}[h_j^{\mathbb{T}(x_j,2)}]) \in \{R, E\}$, and player i constructs $x_j(f_i^{\text{include}}[h_i^{\mathbb{T}(x_j,2)}]) \in \{G, B\}$ and $\theta_i(\text{receive}, f_i^{\text{include}}[h_i^{\mathbb{T}(x_j,2)}]) \in \{R, E\}$. For each $n \in \{i, j\}$, $t_n^{\text{exclude}}(x_j, 2)$ is randomly selected.

Given $x_j(f_i^{\text{include}}[h_i^{\mathbb{T}(x_j,1)}]) \in \{G, B\}$ and $x_j(f_i^{\text{include}}[h_i^{\mathbb{T}(x_j,2)}]) \in \{G, B\}$, player i creates $x_j(i) \in$

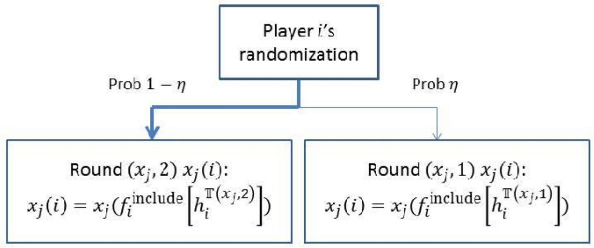

Figure 8: How to create $x_j(i)$

$\{G, B\}$ as follows:

1. [Round $(x_j, 2)$ $x_j(i)$] With probability $1-\eta$, player i uses $x_j(f_i^{\text{include}}[h_i^{\mathbb{T}(x_j,2)}])$ in round $(x_j, 2)$ to infer x_j: $x_j(i) = x_j(f_i^{\text{include}}[h_i^{\mathbb{T}(x_j,2)}])$;

2. [Round $(x_j, 1)$ $x_j(i)$] With probability η, player i uses $x_j(f_i^{\text{include}}[h_i^{\mathbb{T}(x_j,1)}])$ in round $(x_j, 1)$ to infer x_j: $x_j(i) = x_j(f_i^{\text{include}}[h_i^{\mathbb{T}(x_j,1)}])$.

We summarize player i's inference in Figure 8.

Finally, in round $(x_j, 3)$, player i sends $x_j(i) \in \{G, B\}$ spending $T^{\frac{1}{2}}$ periods as explained in Section 10 (with indices j and i reversed, m replaced with $x_j(i)$, and $\mathbb{T} = \mathbb{T}(x_j, 3)$ with $|\mathbb{T}| = T^{\frac{1}{2}}$). Player i constructs $\theta_i(x_j(i), f_i^{\text{include}}[h_i^{\mathbb{T}(x_j,3)}]) \in \{R, E\}$, and player j constructs $x_j(i)(f_j^{\text{include}}[h_j^{\mathbb{T}(x_j,3)}]) \in \{G, B\}$ and $\theta_j(\text{receive}, f_j^{\text{include}}[h_j^{\mathbb{T}(x_j,3)}]) \in \{R, E\}$. For each $n \in \{i, j\}$, $t_n^{\text{exclude}}(x_j, 3)$ is randomly selected.

Given $\theta_j(x_j, f_j^{\text{include}}[h_j^{\mathbb{T}(x_j,2)}]) \in \{R, E\}$ and $x_j(i)(f_j^{\text{include}}[h_j^{\mathbb{T}(x_j,3)}]) \in \{G, B\}$, player j creates $x_j(j) \in \{G, B\}$ as follows:

1. [Adhere $x_j(j)$] If player j has $\theta_j(x_j, f_j^{\text{include}}[h_j^{\mathbb{T}(x_j,2)}]) = R$ in round $(x_j, 2)$, then player j mixes the following cases:

   (a) [Round $(x_j, 2)$ $x_j(j)$] With probability $1-\eta$, player j adheres to her own state: $x_j(j) = x_j$;

   (b) [Round $(x_j, 3)$ $x_j(j)$] With probability η, player j listens to player i's message in round $(x_j, 3)$: $x_j(j) = x_j(i)(f_j^{\text{include}}[h_j^{\mathbb{T}(x_j,3)}])$.

2. [Not Adhere $x_j(j)$] If player j has $\theta_j(x_j, f_j^{\text{include}}[h_j^{\mathbb{T}(x_j,2)}]) = E$ in round $(x_j, 2)$, then player j always listens to player i's message in round $(x_j, 3)$: $x_j(j) = x_j(i)(f_j^{\text{include}}[h_j^{\mathbb{T}(x_j,3)}])$.

We summarize player j's inference in Figure 9.

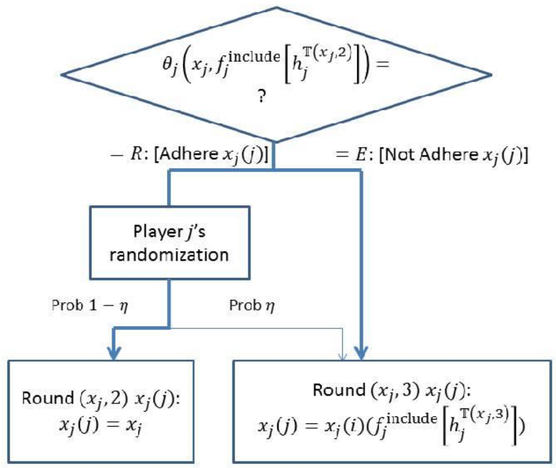

Figure 9: How to create $x_j(j)$

13.2 Properties of $x_j(i)$ and $x_j(j)$

It will be useful to derive the following properties about the inferences.

Lemma 8 *Given the above construction of $x_j(i)$ and $x_j(j)$, we have the following: For a sufficiently large T, for each $i, j \in I$ and $x_j, x_j(i), x_j(j) \in \{G, B\}$,*

1. *For each $h_i^{\mathbb{T}(x_j,1)}$, $h_i^{\mathbb{T}(x_j,2)}$, and $h_i^{\mathbb{T}(x_j,3)}$, if $x_j(i) \neq x_j(j)$, then player i believes that [Round $(x_j, 3)$ $x_j(j)$], $\theta_j(x_j, f_j^{\text{include}}[h_j^{\mathbb{T}(x_j,1)}]) = E$, or $\theta_j(x_j, f_j^{\text{include}}[h_j^{\mathbb{T}(x_j,2)}]) = E$ with a high probability:*

$$\Pr\left(\left\{\begin{array}{c}[\textit{Round } (x_j,3) \ x_j(j)] \\ \vee\theta_j(x_j, f_j^{\text{include}}[h_j^{\mathbb{T}(x_j,1)}]) = E \\ \vee\theta_j(x_j, f_j^{\text{include}}[h_j^{\mathbb{T}(x_j,2)}]) = E\end{array}\right\} \mid x_j, x_j(j), h_i^{\mathbb{T}(x_j,1)}, h_i^{\mathbb{T}(x_j,2)}, h_i^{\mathbb{T}(x_j,3)}\right) \geq 1-\exp(-2T^{\frac{1}{3}}). \tag{39}$$

2. *For each $h_j^{\mathbb{T}(x_j,1)}$, $h_j^{\mathbb{T}(x_j,2)}$, and $h_j^{\mathbb{T}(x_j,3)}$, if $x_j(j) \neq x_j(i)$, then player j believes that [Round $(x_j, 1)$ $x_j(i)$] or $\theta_i(\text{receive}, f_i^{\text{include}}[h_i^{\mathbb{T}(x_j,2)}]) = E$ with a high probability:*

$$\Pr\left(\left\{\begin{array}{c}[\textit{Round } (x_j,1) \ x_j(i)] \\ \vee\theta_i(\text{receive}, f_i^{\text{include}}[h_i^{\mathbb{T}(x_j,2)}]) = E \\ \vee\theta_i(x_j(i), f_i^{\text{include}}[h_i^{\mathbb{T}(x_j,3)}]) = E\end{array}\right\} \mid x_j, x_j(i), h_j^{\mathbb{T}(x_j,1)}, h_j^{\mathbb{T}(x_j,2)}, h_j^{\mathbb{T}(x_j,3)}\right) \geq 1-\exp(-2T^{\frac{1}{3}}). \tag{40}$$

Proof. See Appendix A.8. ■

As will be seen, if [Round $(x_j, 3)$ $x_j(j)$], $\theta_j(x_j, f_j^{\text{include}}[h_j^{\mathbb{T}(x_j,1)}]) = E$, or $\theta_j(x_j, f_j^{\text{include}}[h_j^{\mathbb{T}(x_j,2)}]) = E$ is the case, then player j uses the type-3 reward (makes any strategy of player i optimal). Hence, (39) ensures that, whenever player i realizes that player i's inference is miscoordinated with player j's inference, any strategy is optimal for player i with a high probability, as required by (26). Similarly, (40) ensures that, whenever player j realizes that player j's inference is miscoordinated with player i's inference, any strategy is optimal for player j with a high probability.

We explain why (39) holds. First, Figure 9 implies that if $x_j(j) \neq x_j$, then we have [Round $(x_j, 3)$ $x_j(j)$] or $\theta_j(x_j, f_j^{\text{include}}[h_j^{\mathbb{T}(x_j,2)}]) = E$. Hence, we concentrate on $x_j(j) = x_j$.

Look at Figure 8 for how player i infers $x_j(i)$. Suppose player i has $x_j(i) = x_j(f_i^{\text{include}}[h_i^{\mathbb{T}(x_j,1)}]) \neq x_j(j) = x_j$. Since the round $(x_j, 1)$ lasts for $T^{\frac{1}{2}}$ periods, Claim 2 in Lemma 6 with $m = x_j$ implies

$$\Pr\left(\left\{\theta_j(x_j, f_j^{\text{include}}[h_j^{\mathbb{T}(x_j,1)}]) = E\right\} \mid x_j, h_i^{\mathbb{T}(x_j,1)}\right) \geq 1 - \exp\left(-\varepsilon_{\text{message}} T^{\frac{1}{2}}\right).$$

Given x_j, by Figure 9, player j's strategy in rounds $(x_j, 2)$ and $(x_j, 3)$ and her inference $x_j(j)$ are independent of her history in $(x_j, 1)$. Hence, this inequality is sufficient for (39).

Suppose next player i has $x_j(i) = x_j(f_i^{\text{include}}[h_i^{\mathbb{T}(x_j,2)}]) \neq x_j(j) = x_j$. Since the round $(x_j, 2)$ lasts for $T^{\frac{2}{3}}$ periods, Claim 2 in Lemma 6 with $m = x_j$ implies

$$\Pr\left(\left\{\theta_j(x_j, f_j^{\text{include}}[h_j^{\mathbb{T}(x_j,2)}]) = E\right\} \mid x_j, h_i^{\mathbb{T}(x_j,2)}\right) \geq 1 - \exp\left(-\varepsilon_{\text{message}} T^{\frac{2}{3}}\right). \tag{41}$$

Since player j uses her history in the round $(x_j, 2)$ to determine $x_j(j)$ by Figure 9, this inequality does not immediately imply (39). Nonetheless, we can still prove (39) as follows.

Observe that by Figure 9, with probability at least η, regardless of player j's history in $(x_j, 2)$, player j uses the inference in $(x_j, 3)$: $x_j(j) = x_j(i)(f_j^{\text{include}}[h_j^{\mathbb{T}(x_j,3)}])$. Since the length of round $(x_j, 3)$ is $T^{\frac{1}{2}}$, Claim 6 of Lemma 6 (with i and j reversed and $m = x_j(i)$) implies that player i with $h_i^{\mathbb{T}(x_j,3)}$ believes that any $x_j(i)(f_j^{\text{include}}[h_j^{\mathbb{T}(x_j,3)}])$ happens with at least probability $\exp(-K_{\text{message}} T^{\frac{1}{2}})$. In total, player i believes that any $x_j(j)$ is possible with at least probability $\eta \exp(-K_{\text{message}} T^{\frac{1}{2}})$ given $(h_i^{\mathbb{T}(x_j,2)}, h_i^{\mathbb{T}(x_j,3)})$. Hence, the belief update about $\theta_j(x_j, f_j^{\text{include}}[h_j^{\mathbb{T}(x_j,2)}])$ from $h_i^{\mathbb{T}(x_j,3)}$ and $x_j(j)$ is of order $\exp(-T^{\frac{1}{2}})$.

Since $T^{\frac{2}{3}} \gg T^{\frac{1}{2}}$ (here is where we use the assumption that round $(x_j, 2)$ is much longer than round $(x_j, 3)$), (41) implies that

$$\Pr\left(\left\{\theta_j(x_j, f_j^{\text{include}}[h_j^{\mathbb{T}(x_j,2)}]) = E\right\} \mid x_j, x_j(j), h_i^{\mathbb{T}(x_j,2)}, h_i^{\mathbb{T}(x_j,3)}\right) \tag{42}$$

is of order $\exp(-T^{\frac{2}{3}})$. Conditional on x_j and $x_j(j)$, player j's strategy in round $(x_j, 1)$ is independent of her history in $(x_j, 2)$ and $(x_j, 3)$. Hence, (42) is sufficient for (39).

We second prove (40). Look at Figure 9 for how player j infers $x_j(j)$. Suppose player j has $x_j(j) = x_j(i)(f_j^{\text{include}}[h_j^{\mathbb{T}(x_j,3)}])$. Then, since round $(x_j, 3)$ lasts for $T^{\frac{1}{2}}$ periods, Claim 2 of Lemma 6 (with i and j reversed and $m = x_j(i)$) implies

$$\Pr\left(\left\{\theta_i(x_j(i), f_i^{\text{include}}[h_i^{\mathbb{T}(x_j,3)}]) = E\right\} \mid x_j(i), h_j^{\mathbb{T}(x_j,3)}\right) \geq 1 - \exp\left(-\varepsilon_{\text{message}} T^{\frac{1}{2}}\right).$$

Given $x_j(i)$, player i's strategy in rounds $(x_j, 1)$ and $(x_j, 2)$ is independent of her history in $(x_j, 3)$. Hence, this inequality is sufficient for (40).

Suppose next that player j adheres to x_j. In this case, by Figure 9, player j has $\theta_j(x_j, f_j^{\text{include}}[h_j^{\mathbb{T}(x_j,2)}]) = R$. Hence, by Claim 5 of Lemma 6, we have

$$\Pr\left(\left\{\begin{array}{c} \theta_i(\text{receive}, f_i^{\text{include}}[h_i^{\mathbb{T}(x_j,2)}]) = E \\ \vee x_j(f_i^{\text{include}}[h_i^{\mathbb{T}(x_j,2)}]) = x_j \end{array}\right\} \mid x_j, h_j^{\mathbb{T}(x_j,2)}\right) \geq 1 - \exp\left(-\varepsilon_{\text{message}} T^{\frac{2}{3}}\right). \tag{43}$$

To see why this is sufficient for (40), we deal with player j's learning from $x_j(i)$, $h_j^{\mathbb{T}(x_j,1)}$, and $h_j^{\mathbb{T}(x_j,3)}$ as follows. By Figure 8, with probability at least η, player i uses the inference in round $(x_j, 1)$: $x_j(i) = x_j(f_i^{\text{include}}[h_i^{\mathbb{T}(x_j,1)}])$. Since the length of round $(x_j, 1)$ is $T^{\frac{1}{2}}$, Claim 6 of Lemma 6 (with $m = x_j$) implies player j with $h_j^{\mathbb{T}(x_j,1)}$ believes that any inference $x_j(f_i^{\text{include}}[h_i^{\mathbb{T}(x_j,1)}])$ happens with probability at least $\exp(-K_{\text{message}} T^{\frac{1}{2}})$. In total, player j believes that any $x_j(i)$ is possible with probability at least $\eta \exp(-K_{\text{message}} T^{\frac{1}{2}})$ given $(h_j^{\mathbb{T}(x_j,1)}, h_j^{\mathbb{T}(x_j,2)})$.

Since $T^{\frac{2}{3}} \gg T^{\frac{1}{2}}$ (here is where we use the assumption that round $(x_j, 2)$ is much longer than round $(x_j, 1)$), (43) implies

$$\Pr\left(\left\{\begin{array}{c} \theta_i(\text{receive}, f_i^{\text{include}}[h_i^{\mathbb{T}(x_j,2)}]) = E \\ \vee x_j(f_i^{\text{include}}[h_i^{\mathbb{T}(x_j,2)}]) = x_j \end{array}\right\} \mid x_j, x_j(i), h_j^{\mathbb{T}(x_j,1)}, h_j^{\mathbb{T}(x_j,2)}\right)$$

is still of order $1 - \exp(-T^{\frac{2}{3}})$. Moreover, given $x_j(i)$, player i's strategy in round $(x_j, 3)$ is independent of her history in $(x_j, 1)$ and $(x_j, 2)$. Hence, (43) is sufficient for (40).

13.3 Strategy in the Coordination Block for x_i

The strategy in the coordination block for x_i is defined in the same way as Section 13.1, with indices i and j reversed. Player i creates $\theta_i(x_i, f_i^{\text{include}}[h_i^{\mathbb{T}(x_i,1)}]) \in \{R, E\}$, $\theta_i(x_i, f_i^{\text{include}}[h_i^{\mathbb{T}(x_i,2)}]) \in \{R, E\}$, $x_i(j)(f_i^{\text{include}}[h_i^{\mathbb{T}(x_i,3)}]) \in \{G, B\}$, $\theta_i(\text{receive}, x_i, f_i^{\text{include}}[h_i^{\mathbb{T}(x_i,3)}]) \in \{R, E\}$, and $x_i(i) \in \{G, B\}$. On the other hand, player j creates $x_i(f_j^{\text{include}}[h_j^{\mathbb{T}(x_i,1)}]) \in \{G, B\}$, $\theta_j(\text{receive}, f_j^{\text{include}}[h_j^{\mathbb{T}(x_i,1)}]) \in \{R, E\}$, $x_i(f_j^{\text{include}}[h_j^{\mathbb{T}(x_i,2)}]) \in \{G, B\}$, $\theta_j(\text{receive}, f_j^{\text{include}}[h_j^{\mathbb{T}(x_i,2)}]) \in \{R, E\}$, $x_i(j) \in \{G, B\}$, and $\theta_j(x_i(j), f_j^{\text{include}}[h_j^{\mathbb{T}(x_i,3)}]) \in \{R, E\}$.

13.4 Strategy in the Main Block

Given the coordination block, each player i has the inference of the state profile $x(i) = (x_i(i), x_j(i)) \in \{G, B\}^2$. Player i with $x(i)$ takes an action in main block l given $\lambda_i(l) \in \{G, B\}$ and $\lambda_j(l)(i) \in \{G, B\}$. We first define the strategy given $\lambda_i(l)$ and $\lambda_j(l)(i)$:

13.4.1 Review Round l given $\lambda_i(l)$ and $\lambda_j(l)(i)$

In review round l, player i takes actions, depending on (i) her inference of the state profile $x(i) \in \{G, B\}^2$, (ii) the summary statistic of the realization of player j's reward (calculated by player i) $\lambda_i(l) \in \{G, B\}$, and (iii) her inference of player j's statistic $\lambda_j(l)$, denoted by $\lambda_j(l)(i) \in \{G, B\}$. As will be seen, both $\lambda_i(l)$ and $\lambda_j(l)(i)$ are functions of the frequency of player i's history at the beginning of review round l, $f_i^{\text{include}}[h_i^{<l}]$. Hence, precisely speaking, we should write $\lambda_i(l)(f_i^{\text{include}}[h_i^{<l}])$ and $\lambda_j(l)(i)(f_i^{\text{include}}[h_i^{<l}])$. For notational convenience, we omit $f_i^{\text{include}}[h_i^{<l}]$.

Player i picks $\alpha_i(l) \in \Delta(A_i)$ based on $\lambda_i(l)$ and $\lambda_j(l)(i)$ at the beginning of review round l as follows (see Figure 10 for illustration). Given $\alpha_i(l)$, player i takes $a_{i,t}$ according to $\alpha_i(l)$ *i.i.d.* across periods for T periods in review round l. (That is, the mixture to decide $\alpha_i(l)$ is conducted only once at the beginning of review round l. Given $\alpha_i(l)$, player i draws $a_{i,t}$ from $\alpha_i(l)$ every period in the round.)

1. If player i has $\lambda_j(l)(i) = G$, then player i believes that with a high probability, any action is optimal. (If $\lambda_j(l) = G$ (coordination goes well), then the reward is not type-2. Whether the

reward is type-1 or type-3, both rewards make any strategy of player i optimal by Lemma 2.) In this case, she picks the strategy to control player j's payoffs.

(a) If $\lambda_i(l) = G$, then player j's reward function is not type-2. In this case, player i takes $\alpha_i(l) = \alpha_i^{\rho}(x(i))$ with probability $1-\eta$, and $\alpha_i(l) = \alpha_i^{*,\rho}(x(i))$ with probability η.

(b) If $\lambda_i(l) = B$, then player j's reward function is type-2. Then, player i takes $\alpha_i(l) = \alpha_i^{\rho}(x(i))$ with probability η, and $\alpha_i(l) = \alpha_i^{*,\rho}(x(i))$ with probability $1-\eta$.

That is, with a high probability, player i takes $\alpha_i^{\rho}(x(i))$ if $\lambda_i(l) = G$ and $\alpha_i^{*,\rho}(x(i))$ if $\lambda_i(l) = B$. As will be seen in Section 16, this strategy makes sure that player j's equilibrium payoff is $v_j(x_i)$ at the same time of satisfying self generation.

In addition, we make sure that the support of $\alpha_i(l)$ is the same regardless of $\lambda_i(l) \in \{G, B\}$ given $\lambda_j(l)(i) = G$. With indices i and j reversed, player i cannot learn $\lambda_j(l)$ by observing $\alpha_j(l)$ given $\lambda_i(l)(j) = G$. This will be important when we calculate player i's belief update about $\lambda_j(l)$ in Lemma 9.

2. If player i has $\lambda_j(l)(i) = B$, then player i believes that her reward function is type-2 and player j takes $\alpha_j^{*,\rho}(x(i))$ if $x(i) = x(j)$. (She believes that, if $x(i) \neq x(j)$, then player j uses the type-3 reward and any strategy is optimal with a high probability by Lemma 8.) Hence, player i takes the static best response to $\alpha_j^{*,\rho}(x(i))$: $\alpha_i(l) = BR_i(\alpha_j^{*,\rho}(x(i)))$.

For notational convenience, let $\text{action}_i(l) = R$ denote the event that player i takes $\alpha_i^{\rho}(x(i))$ if $\lambda_i(l) = G$ and takes $\alpha_i^{*,\rho}(x(i))$ if $\lambda_i(l) = B$; and let $\text{action}_i(l) = E$ denote the complementary event. If $\lambda_j(l)(i) = G$, then $\text{action}_i(l) = R$ denotes the event that player i takes the strategy which is taken with a high probability $1-\eta$. As will be seen in Claim 4 of Lemma 11 (with indices i and j reversed), if $\lambda_j(l)(i) = G$ and $\text{action}_i(l) = E$, then player i will use the type-3 reward for player j.

Let $h_i^{\mathbb{T}(l)} \equiv (a_{i,t}, y_{i,t})_{t \text{ in review round } l}$ be player i's history in review round l. Given $h_i^{\mathbb{T}(l)}$, if player i has $\lambda_j(l)(i) = \lambda_i(l) = G$ and takes $\alpha_i(l) = \alpha_i^{\rho}(x(i))$, then she creates the random variables $\theta_i^{\text{review}}(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}(l)}]) \in \{R, E\}$ as in Section 11 with $\mathbb{T} = \mathbb{T}(l)$ (T periods in review round l). Let $t_i^{\text{exclude}}(l)$ be the period randomly excluded by player i.

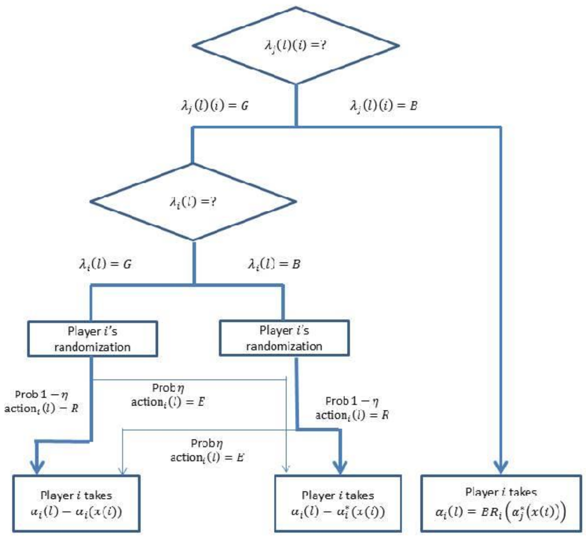

Figure 10: Determinaction of $\alpha_i(l)$

On the other hand, player i also calculates player j's type-1 reward

$$\pi_j(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}(l)}]) \equiv \sum_{t \in \mathbb{T}(l), t \neq t_i^{\text{exclude}}(l)} \pi_j[\alpha(x(i))](a_{i,t}, y_{i,t}). \tag{44}$$

The definitions are the same as (38) in Section 11 with indices i and j reversed and $\mathbb{T} = \mathbb{T}(l)$.[15]

Given $\lambda_i(l)$ and $\pi_j(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}(l)}])$, player i defines $\lambda_i(l+1)$ as follows: The initial condition is $\lambda_i(1) = G$. Given $\lambda_i(l)$, player i defines $\lambda_i(l+1)$ as follows (See Figure 11 for illustration):

1. [Case λ_i G] If player i has $\lambda_i(l) = G$, then player i has $\lambda_i(l+1) = B$ at the end of review round l if and only if player j's score in review round l is excessive:

   (a) [Case λ_i Not Excessive] If player i has $\left|\pi_j(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}(l)}])\right| \leq \frac{\bar{u}}{4L}T$, then $\lambda_i(l+1) = G$ at the end of review round l.

   (b) [Case λ_i Excessive] If player i has $\left|\pi_j(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}(l)}])\right| > \frac{\bar{u}}{4L}T$, then $\lambda_i(l+1) = B$ at the end of review round l.

2. [Case λ_i B] If player i has $\lambda_i(l) = B$, then player i has $\lambda_i(l+1) = B$ at the end of review round l. That is, $\lambda_i(l) = B$ is absorbing.

With indices i and j reversed, player j also creates $\lambda_j(l+1) \in \{G, B\}$.

[15] Note that the function $\pi_j(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}(l)}])$ is well defined even if player i takes $\alpha_i(l) \neq \alpha_i^{\rho}(x(i))$.

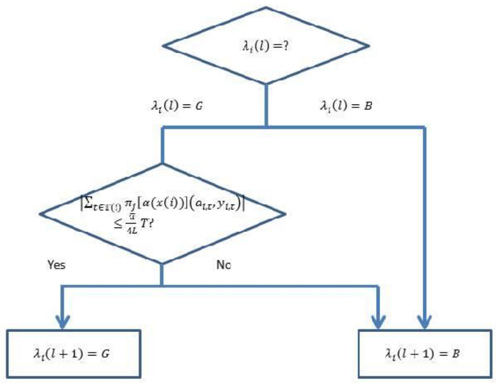

Figure 11: Transition of $\lambda_i(l)$

13.4.2 The Supplemental Round for $\lambda_j(l+1)$

Player j sends $\lambda_j(l+1) \in \{G, B\}$ in the supplemental round for $\lambda_j(l+1)$ spending $T^{\frac{1}{2}}$ periods as explained in Section 10 (with m replaced with $\lambda_j(l+1)$, and $\mathbb{T} = \mathbb{T}(\lambda_j(l+1))$ with $|\mathbb{T}| = T^{\frac{1}{2}}$). Player j creates $\theta_j(\lambda_j(l+1), f_j^{\text{include}}[h_j^{\mathbb{T}(\lambda_j(l+1))}]) \in \{R, E\}$ and player i creates $\lambda_j(l+1)(f_i^{\text{include}}[h_i^{\mathbb{T}(\lambda_j(l+1))}]) \in \{G, B\}$ and $\theta_i(\text{receive}, f_i^{\text{include}}[h_i^{\mathbb{T}(\lambda_j(l+1))}]) \in \{R, E\}$. For each $n \in \{i, j\}$, $t_n^{\text{exclude}}(\lambda_j(l+1))$ is randomly excluded.

13.4.3 The Supplemental Round for $\lambda_i(l+1)$

The strategy is the same as Section 13.4.2 with indices i and j reversed. Player i creates $\theta_i(\lambda_i(l+1), f_i^{\text{include}}[h_i^{\mathbb{T}(\lambda_i(l+1))}]) \in \{R, E\}$ and player j creates $\lambda_i(l+1)(f_j^{\text{include}}[h_j^{\mathbb{T}(\lambda_i(l+1))}]) \in \{G, B\}$ and $\theta_j(\text{receive}, f_j^{\text{include}}[h_j^{\mathbb{T}(\lambda_i(l+1))}]) \in \{R, E\}$. For each $n \in \{i, j\}$, $t_n^{\text{exclude}}(\lambda_i(l+1))$ is randomly excluded.

Now that we have defined player i's strategy given $\lambda_j(l)(i)$ and $\lambda_i(l)$ for all $l \in \{1, ..., L\}$ and that we have defined the transition of $\lambda_i(l)$. Hence, to complete the definition of the strategy in the main block, we are left to define the transition of $\lambda_j(l)(i)$.

13.4.4 Transitions of $\lambda_i(l)$ and $\lambda_j(l)(i)$

The initial condition is $\lambda_j(1)(i) = G$. Inductively, for $l \in \{1, ..., L-1\}$, given $\lambda_j(l)(i) \in \{G, B\}$, player i determines $\lambda_j(l+1)(i) \in \{G, B\}$ as follows (see Figure 12 for illustration):

1. [Case $\lambda_j(i)$ G] If player i has $\lambda_j(l)(i) = G$, then player i has $\lambda_j(l+1)(i) = B$ in the next block $l+1$ if and only if player i decides to "listen to" player j's message in the supplemental round for $\lambda_j(l+1)$ and "infers" that player j's message is B. Specifically,

(a) [Regular θ $\lambda_j(i)$] If player i has $\alpha_i(\hat{l}) = \alpha_i^{\rho}(x(i))$ and $\theta_i(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}(\hat{l})}]) = R$ for each $\hat{l} \leq l$ with $\lambda_i(\hat{l}) = G$, then player i mixes the following two cases:

i. [Case $\lambda_j(i)$ Ignore] With probability $1 - \eta$, player i "ignores" player j's message: Player i has $\lambda_j(l+1)(i) = G$, regardless of player i's inference of player j's message about $\lambda_j(l+1)$, denoted by $\lambda_j(l+1)(f_i^{\text{include}}[h_i^{\mathbb{T}(\lambda_j(l+1))}])$.

As will be seen in Lemma 9, if player i ignores player j's message, then player i believes that $\lambda_j(l+1) = G$ or player j uses the type-3 reward with a high probability. In other words, (26) holds when player i ignores the message.

To see why this is true with this transition, assume here that player j has $\lambda_j(l) = G$, $\lambda_i(l)(j) = G$, and $\alpha_j(l) = \alpha_j^{\rho}(x(j))$. The first condition assumes that player j had not yet switched to $\lambda_j(l) = B$ by the beginning of review round l. In addition, as will be seen in Lemma 11, if $\lambda_j(l) = G$ but either $\lambda_i(l)(j) \neq B$ or $\alpha_j(l) \neq \alpha_j^{\rho}(x(j))$, then player j uses the type-3 reward (and so (26) holds trivially). Given these conditions, if and only if $\left|\pi_i(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}(l)}])\right| > \frac{\bar{u}}{4L}$, player j has $\lambda_j(l+1) = B$ by Figure 11 (with indices i and j reversed).

Imagine first that player i has $\lambda_i(\hat{l}) = G$ for each $\hat{l} \leq l$. Then, [Regular θ $\lambda_j(i)$] implies $\theta_i(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}(l)}]) = R$. By Claim 3 of Lemma 7, player i with $\theta_i(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}(l)}]) = R$ believes that, if $x(i) = x(j)$, then $\left|\pi_i(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}(l)}])\right| \leq \frac{\bar{u}}{4L}$ or player j uses the type-3 reward. Moreover, if $x(i) \neq x(j)$, then player i believes that player j uses the type-3 reward by Lemma 8. In total, player i believes that $\lambda_j(l+1) = G$ or player j uses the type-3 reward.

Imagine second that player i switched $\lambda_i(\hat{l}) = G$ to $\lambda_i(\hat{l}+1) = B$ for some $\hat{l} \leq l-1$. This means $\left|\pi_j(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}(\hat{l})}])\right| > \frac{\bar{u}}{4L}$ by Figure 11. Claim 3 of Lemma 7 ensures that player i with $\theta_i(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}(\hat{l})}]) = R$ believes that player j uses the type-3 reward if $x(i) = x(j)$. Again, if $x(i) \neq x(j)$, then player i believes that player j uses the type-3 reward.

In both cases, player i believes that $\lambda_j(l+1) = G$ or player j uses the type-3 reward. Note that the above argument establishes player i believing "$\lambda_j(l+1) = G$ or player j using the type-3 reward" at the end of review round l (or in the second case, at the end of review round $\hat{l}$). Recall that (26) holds conditional on player i's history

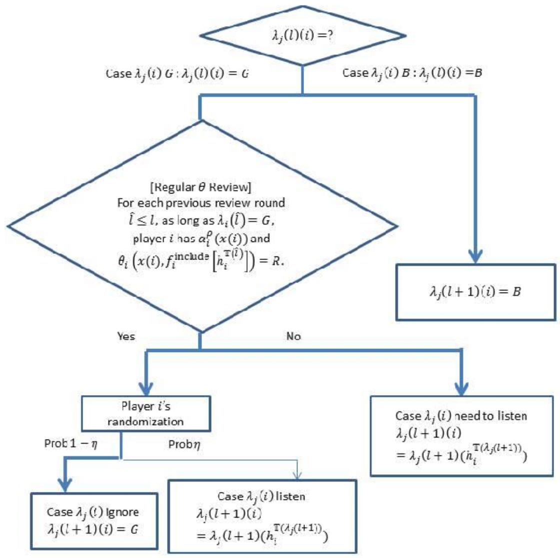

Figure 12: Transition of $\lambda_j(l)(i)$

at the end of review round $l+1$. Hence, we still have to deal with player i's learning about $\lambda_j(l+1)$ from her history between the end of review round l (or that of review round $\hat{l}$) and the end of review round $l+1$. See the explanation after Lemma 9 for the details.

ii. [Case $\lambda_j(i)$ Listen] With probability η, player i "listens to" player j's message: Player i has $\lambda_j(l+1)(i) = \lambda_j(l+1)\left(f_i^{\text{include}}[h_i^{\mathbb{T}(\lambda_j(l+1))}]\right)$.
Claim 2 of Lemma 6 ensures that player i believes that either $\lambda_j(l+1)(i) = \lambda_j(l+1)$ or player j uses the type-3 reward.

(b) [Irregular θ $\lambda_j(i)$] Otherwise, player i always listens to the message: Player i has $\lambda_j(l+1)(i) = \lambda_j(l+1)\left(f_i^{\text{include}}[h_i^{\mathbb{T}(\lambda_j(l+1))}]\right)$.

Note that, after each history (unless $\lambda_j(l)(i) = B$), player i listens to player j's message with probability at least η. See footnote 17 for why this lower bound of the probability is important.

2. [Case $\lambda_j(i)$ B] If player i has $\lambda_j(l)(i) = B$, then player i has $\lambda_j(l+1)(i) = B$ in the next block $l+1$. That is, $\lambda_j(l)(i) = B$ is absorbing.

13.4.5 Property of $\lambda_j(l)(i)$

Now that we have finished defining player i's strategy in the coordination and main blocks, it will be useful to summarize the property of the inference $\lambda_j(l)(i)$. Suppose that $x(i) = x(j)$ (otherwise, Lemma 8 ensures that player i believes that player j uses the type-3 reward function) and that player j has $\lambda_i(l)(j) = G$ (otherwise, as mentioned in the explanation of Claim 4 of Lemma 7, player j uses the type-3 reward).

The following lemma ensures that, if $x(i) = x(j)$, player j has $\lambda_i(l)(j) = G$, and player i has $\lambda_j(l)(i) = G$, then player i believes that, with a high probability, one of the following four events happens: $\lambda_j(l) = G$; $\theta_j(\lambda_j(l), f_j^{\text{include}}[h_j^{\mathbb{T}(\lambda_j(l))}]) = E$ happens when player j sends $\lambda_j(l)$ in the supplemental round; $\text{action}_j(\hat{l}) = E$ for some $\hat{l} \leq l-1$; or $\theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}(\hat{l})}]) = E$ happens for some review round $\hat{l} \leq l-1$.[16] As will be seen in Lemma 11, the cases except for $\lambda_j(l) = G$ imply that player j uses the type-3 reward. Hence, the following lemma is the basis for (26):

Lemma 9 *For a sufficiently large T, for each $i \in I$, $x_j \in \{G, B\}$, and each $h_i^{\leq l}$, if player i has $\lambda_j(l)(i) = G$, then we have*

$$\Pr\left(\begin{array}{c}\left\{\begin{array}{c}\lambda_j(l) = B \wedge \theta_j(\lambda_j(l), f_j^{\text{include}}[h_j^{\mathbb{T}(\lambda_j(l))}]) = R \\ \wedge \text{action}_j(\hat{l}) = \theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}(\hat{l})}]) = R \text{ for each } \hat{l} \leq l-1\end{array}\right\} \\ \mid x_j, \{x(j) = x(i)\}, \{\lambda_i(l)(j) = G\}, h_i^{\leq l}\end{array}\right) \leq \exp(-T^{\frac{1}{3}}). \qquad (45)$$

Proof. See Appendix A.9. ■

Let us give an intuitive explanation. Recall that, given $\lambda_i(l)(j) = G$, if player j has $\text{action}_j(\hat{l}) = R$ with $\hat{l} \leq l$, then we have

$$\begin{aligned} & \left\{\lambda_i(l)(j) = G \wedge \text{action}_j(\hat{l}) = R \text{ with } \hat{l} \leq l\right\} \\ \Rightarrow\ & \left\{\lambda_i(\hat{l})(j) = G \wedge \text{action}_j(\hat{l}) = R \text{ with } \hat{l} \leq l\right\} \text{ since } \lambda_i(l)(j) = B \text{ is absorbing by Figure 12} \\ \Rightarrow\ & \left\{\alpha_j(\hat{l}) = \alpha_j^{\rho}(x(j))\right\}. \end{aligned} \qquad (46)$$

See Figure 12 for how player i updates $\lambda_j(l)(i)$. If player i listens to $\lambda_j(l)$ in the supplemental

[16] $\theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}(\hat{l})}])$ is not defined if player j does not take $\alpha_j(\hat{l}) = \alpha_j^{\rho}(x(j))$. We define that if $\alpha_j(\hat{l}) \neq \alpha_j^{\rho}(x(j))$, then neither $\theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}(\hat{l})}]) = R$ nor $\theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}(\hat{l})}]) = E$ is the case.

round for $\lambda_j(l)$, then by Claim 2 of Lemma 6, if $\lambda_j(l)(i) \neq \lambda_j(l)$, then

$$\Pr\left(\left\{\theta_j(\lambda_j(l), f_j^{\text{include}}[h_j^{\mathbb{T}(\lambda_j(l))}]) = E\right\} \mid \lambda_j(l), h_i^{\mathbb{T}(\lambda_j(l))}\right) \geq 1 - \exp(-\varepsilon_{\text{message}} T^{\frac{1}{2}}).$$

Since player j's continuation strategy does not depend on $h_j^{\mathbb{T}(\lambda_j(l))}$ given $\lambda_j(l)$, this inequality is sufficient for (45) (for sufficiently large T).

Hence, we focus on the case where player i does not listen to $\lambda_j(l)$, and will prove

$$\Pr\left(\begin{array}{c}\left\{\begin{array}{c}\lambda_j(l) = B \\ \wedge \text{action}_j(\hat{l}) = \theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}(\hat{l})}]) = R \text{ for each } \hat{l} \leq l-1\end{array}\right\} \\ \mid x_j, \{x(j) = x(i)\}, \{\lambda_i(l)(j) = G\}, h_i^{\leq l}\end{array}\right) \leq \exp(-T^{\frac{1}{3}}).$$

Since player i does not listen to $\lambda_j(l)$, for each $\hat{l} \leq l-1$ with $\lambda_i(\hat{l}) = G$, we have $\alpha_i(\hat{l}) = \alpha_i^{\rho}(x(i))$ and $\theta_i(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}(\hat{l})}]) = R$.

First consider the case where player i has $\lambda_i(\hat{l}) = G$ for each $\hat{l} \leq l-1$. For each $\hat{l} \leq l-1$, suppose $\lambda_i(\hat{l})(j) = \lambda_j(\hat{l}) = G$. Unless $\text{action}_j(\hat{l}) = E$, player j takes $\alpha_j(\hat{l}) = \alpha_j^{\rho}(x(j))$. On the other hand, we have just verified that player i with $\lambda_i(\hat{l}) = G$ has $\alpha_i(\hat{l}) = \alpha_i^{\rho}(x(i))$ and $\theta_i(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}(\hat{l})}]) = R$ in the case we are focusing on. Hence, unless $\text{action}_j(\hat{l}) = E$, we have

$$\begin{aligned} & \Pr\left(\begin{array}{c}\left\{\begin{array}{c}\left|\pi_i(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}(\hat{l})}])\right| \leq \frac{\bar{u}}{4L}|\mathbb{T}| \\ \vee \theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}(\hat{l})}]) = E\end{array}\right\} \\ \mid x_j, \{x(j) = x(i)\}, \left\{\lambda_i(\hat{l})(j) = \lambda_j(\hat{l}) = G\right\}, \left\{\text{action}_j(\hat{l}) = R\right\}, f_i[h_i^{\mathbb{T}(\hat{l})}]\end{array}\right) \\ = & \Pr\left(\begin{array}{c}\left\{\begin{array}{c}\left|\pi_i(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}(\hat{l})}])\right| \leq \frac{\bar{u}}{4L}|\mathbb{T}| \\ \vee \theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}(\hat{l})}]) = E\end{array}\right\} \\ \mid x_j, \{x(j) = x(i)\}, \alpha_i^{\rho}(x(i)), \alpha_j^{\rho}(x(j)), f_i[h_i^{\mathbb{T}(\hat{l})}]\end{array}\right) \text{ by (46)} \\ \geq & 1 - \exp(-\varepsilon_{\text{review}} T) \text{ by Claim 3 of Lemma 7.} \end{aligned}$$

Since $\lambda_j(\hat{l}) = G$ and $\left|\pi_i(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}(\hat{l})}])\right| \leq \frac{\bar{u}}{4L}|\mathbb{T}|$ imply $\lambda_j(\hat{l}+1) = G$ by Figure 11, together with the case of $\text{action}_j(\hat{l}) = E$, the above inequality implies

$$\Pr\left(\begin{array}{c}\left\{\lambda_j(\hat{l}+1) = B \wedge \text{action}_j(\hat{l}) = \theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}(\hat{l})}]) = R\right\} \\ \mid x_j, \{x(j) = x(i)\}, \left\{\lambda_i(\hat{l})(j) = \lambda_j(\hat{l}) = G\right\}, h_i^{\leq \hat{l}}\end{array}\right) \leq \exp(-\varepsilon T) \qquad (47)$$

with a sufficiently small ε.

In order to prove (45), we are left to prove

$$\Pr\left(\begin{array}{c}\left\{\lambda_j(\hat{l}+1)=B\wedge\text{action}_j(\hat{l})=\theta_j(x(j),f_j^{\text{include}}[h_j^{\mathbb{T}(\hat{l})}])=R\right\}\\ \mid x_j,\{x(j)=x(i)\},\left\{\lambda_i(l)(j)=\lambda_j(\hat{l})=G\right\},h_i^{\leq l}\end{array}\right) \tag{48}$$

is of order $\exp(-T)$ from (47). (The difference from (47) is that, in (48), we condition on $\lambda_i(l)(j)=G$ and $h_i^{\leq l}$, rather than $\lambda_i(\hat{l})(j)=G$ and $h_i^{\leq \hat{l}}$.) To see why (48) is sufficient for (45), note that, since $\lambda_j(l)=B$ is absorbing, we have

$$\begin{aligned}&\Pr\left(\begin{array}{c}\left\{\begin{array}{c}\lambda_j(l)=B\\ \wedge\text{action}_j(\hat{l})=\theta_j(x(j),f_j^{\text{include}}[h_j^{\mathbb{T}(\hat{l})}])=R\text{ for each }\hat{l}\leq l-1\end{array}\right\}\\ \mid x_j,\{x(j)=x(i)\},\{\lambda_i(l)(j)=G\},h_i^{\leq l}\end{array}\right)\\ \leq\ &\sum_{\hat{l}=1}^{l-1}\Pr\left(\begin{array}{c}\left\{\begin{array}{c}\lambda_j(\hat{l}+1)=B\\ \wedge\text{action}_j(\hat{l})=\theta_j(x(j),f_j^{\text{include}}[h_j^{\mathbb{T}(\hat{l})}])=R\end{array}\right\}\\ \mid x_j,\{x(j)=x(i)\},\left\{\lambda_i(l)(j)=\lambda_j(\hat{l})=G\right\},h_i^{\leq l}\end{array}\right).\end{aligned}$$

If (48) holds for each $\hat{l}\leq l-1$, then this inequality implies (45), as desired.

We now prove (48). To this end, comparing (47) and (48), we need to deal with learning after review round $\hat{l}$ and conditioning on not only $\lambda_i(\hat{l})(j)=G$ but also $\lambda_i(l)(j)=G$. Given $\lambda_i(l)(j)=G$ (and so $\lambda_i(\tilde{l})(j)=G$ for each $\tilde{l}\leq l$), learning after review round $\hat{l}$ is small for the following reasons: Figure 10 implies that both $\alpha_j(\tilde{l})=\alpha_j^{\rho}(x(j))$ and $\alpha_j(\tilde{l})=\alpha_j^{*,\rho}(x(j))$ happen with probability at least η regardless of player j's history. Further, Figure 4 implies that, given $\lambda_j(\tilde{l})$, any $\alpha_j(\lambda_j(\tilde{l}))\in\{\bar{\alpha}_j^{\rho^{\text{send}}}(G),\bar{\alpha}_j^{\rho^{\text{send}}}(B),\bar{\alpha}_j^{\rho^{\text{send}}}(M)\}$ happens with probability at least $\frac{\eta}{2}$. Hence, player i cannot update the belief about $\lambda_j(\hat{l}+1)$ so much by observing $\alpha_j(\tilde{l})$ or $\alpha_j(\lambda_j(\tilde{l}))$ given $\lambda_i(l)(j)=G$.

Moreover, conditioning on $\lambda_i(l)(j)=G$ does not change player i's belief so much for the following reasons: Figure 12 with indices i and j reversed, as long as $\lambda_i(\tilde{l})(j)=G$, player j listens to player i's message about $\lambda_i(\tilde{l}+1)$ with probability at least η.[17] Claim 6 of Lemma 6 (with i and j reversed and $m=\lambda_i(\tilde{l}+1)$) implies that player i with $h_i^{\mathbb{T}(\lambda_i(\tilde{l}+1))}$ believes that any $\lambda_i(\tilde{l}+1)(f_j^{\text{include}}[h_j^{\mathbb{T}(\lambda_i(\tilde{l}+1))}])$

[17]This is where we use the mixture in the inference of $\lambda_j(l)$.

happens with probability $\exp(-K_{\text{message}}T^{\frac{1}{2}})$. Hence, player i believes that $\lambda_i(\tilde{l}+1)(j) = G$ is possible with probability at least $\eta \exp(-K_{\text{message}}T^{\frac{1}{2}})$.

Since $T^{\frac{1}{2}} \ll T$ for a sufficiently large T, (47) implies (48), as desired.

Second we consider the case where there exists $\hat{l} \le l-2$ such that player i switches from $\lambda_i(\hat{l}) = G$ to $\lambda_i(\hat{l}+1) = B$, that is, $\left|\pi_j(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}(\hat{l})}])\right| > \frac{\bar{u}}{4L}$ by Figure 11. Since player i has $\lambda_i(\tilde{l}) = G$ until review round $\hat{l}$, the same proof as above implies that she believes that, given $\lambda_i(\hat{l})(j) = G$, with a high probability, $\lambda_j(\hat{l}) = G$ unless $\text{action}_j(\tilde{l}) = E$ or $\theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}(\tilde{l})}]) = E$ for some $\tilde{l} \le \hat{l}$.[18]

Moreover, if $\lambda_j(\hat{l}) = G$, then since $\lambda_i(\hat{l}) = G$ and $\lambda_i(\hat{l}+1) = B$ imply $\left|\pi_j(x(i), f_i^{\text{include}}[h_i^{\mathbb{T}(\hat{l})}])\right| > \frac{\bar{u}}{4L}$ by Figure 11, unless $\text{action}_j(\hat{l}) = E$, we have

$$
\begin{aligned}
& \Pr\left(\begin{array}{c} \left\{\theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}(\hat{l})}]) = E\right\} \\ \mid x_j, \{x(j) = x(i)\}, \left\{\lambda_j(\hat{l}) = \lambda_i(\hat{l})(j) = G\right\}, \left\{\text{action}_j(\hat{l}) = R\right\}, f_i[h_i^{\mathbb{T}(\hat{l})}] \end{array}\right) \\
= & \Pr\left(\begin{array}{c} \left\{\theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}(\hat{l})}]) = E\right\} \\ \mid x_j, \{x(j) = x(i)\}, \alpha_i^{\rho}(x(i)), \alpha_j^{\rho}(x(j)), f_i[h_i^{\mathbb{T}(\hat{l})}] \end{array}\right) \\
& \text{since player } j \text{ takes } \alpha_j^{\rho}(x(j)) \text{ by (46)} \\
\ge & 1 - \exp(-\varepsilon_{\text{review}}T) \text{ by Claim 4 of Lemma 7.} \qquad (49)
\end{aligned}
$$

In total, player i believes that $\lambda_j(\hat{l}) = G$ unless $\text{action}_j(\tilde{l}) = E$ or $\theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}(\tilde{l})}]) = E$ for some $\tilde{l} \le \hat{l}$. Player i also believes that if $\lambda_j(\hat{l}) = G$, then $\text{action}_j(\hat{l}) = E$ or $\theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}(\hat{l})}]) = E$. Hence, player i at the end of review round $\hat{l}$ believes that

$$
\Pr\left(\left\{\text{action}_j(\tilde{l}) = \theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}(\tilde{l})}]) = R \text{ for each } \tilde{l} \le \hat{l}\right\} \mid x_j, \left\{\lambda_i(\hat{l})(j) = G\right\}, h_i^{\le \hat{l}}\right)
$$

is of order $\exp(-T)$. We can deal with learning after review round $\hat{l}$ and conditioning on $\lambda_i(l)(j) = G$ as above.

[18] The precise argument goes as follows: Since player i has $\lambda_i(\tilde{l}) = G$ for each $\tilde{l} \le \hat{l}$, she has $\lambda_i(\tilde{l}) = G$ for each $\tilde{l} \le \hat{l} - 1$. Replacing l with $\hat{l}$ in the above argument, player j believes that $\lambda_j(\hat{l}) = G$ unless $\text{action}_j(\tilde{l}) = E$ or $\theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}(\tilde{l})}]) = E$ for some $\tilde{l} \le \hat{l} - 1$. Since "$\text{action}_j(\tilde{l}) = E$ or $\theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}(\tilde{l})}]) = E$ for some $\tilde{l} \le \hat{l} - 1$" implies "$\text{action}_j(\tilde{l}) = E$ or $\theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}(\tilde{l})}]) = E$ for some $\tilde{l} \le \hat{l}$", the statement holds.

13.5 Full Support of $t_j^{\text{exclude}}(r)$

As will be seen, it will be useful (in the proof of Lemma 13) that player i believes that any $t_j^{\text{exclude}}(r)$ is possible in each round r, as long as player j's state $\lambda_i(l)(j)$ has not switched to $\lambda_i(l)(j) = B$:

Lemma 10 *For a sufficiently large T, for each $i \in I$, $x_j \in \{G, B\}$, round r, and player i's history $h_i^{\leq L}$, the following claim holds: Let l_i^* be the first review round with $\lambda_i(l)(j) = B$ and suppose round r satisfies $r < l_i^*$. (If $\lambda_i(l)(j) = G$ for each $l = 1, ..., L$, then we assume that each $r = 1, ..., R$ satisfies $r < l_i^*$.) For each $t \in \mathbb{T}(r)$, we have*

$$\Pr\left(\left\{t_j^{\text{exclude}}(r) = t\right\} \mid x_j, \left\{\lambda_i(l)(j) = G \text{ for all } l \leq r\right\}, h_i^{\leq L}\right) \geq T^{-2}.$$

Proof. Let $\alpha_j(r)$ be player j's strategy in round r. Since $r < l_i^*$, $\alpha_j(r)$ can be either $\bar{\alpha}_j^{\rho^{\text{send}}}(G)$, $\bar{\alpha}_j^{\rho^{\text{send}}}(B)$, $\bar{\alpha}_j^{\rho^{\text{send}}}(M)$, α_j^{mix}, $\alpha_j^{\rho}(x(j))$, or $\alpha_j^{*,\rho}(x(j))$. Let

$$\varepsilon_{\text{support}}^{\text{action}} = \min_{x(j) \in \{G,B\}^2, \alpha_j(r) \in \left\{\bar{\alpha}_j^{\rho^{\text{send}}}(G), \bar{\alpha}_j^{\rho^{\text{send}}}(B), \bar{\alpha}_j^{\rho^{\text{send}}}(M), \alpha_j^{\text{mix}}, \alpha_j^{\rho}(x(j)), \alpha_j^{*,\rho}(x(j))\right\}, a_j \in A_j} \alpha_j(r)(a_j) > 0$$

be the lower bound of the probability with which player j takes each action.

Since player j picks $t_j^{\text{exclude}}(r)$ from $\mathbb{T}(r)$, there exists one $s \in \mathbb{T}(r)$ with

$$\Pr\left(\left\{t_j^{\text{exclude}}(r) = s\right\} \mid \alpha_j(r), h_i^{\leq L}\right) \geq |\mathbb{T}(r)|^{-1} \geq T^{-1}.$$

By Assumption 1, the conditional probability is always well defined. Hence, we are left to show that the likelihood between $t_j^{\text{exclude}}(r) = t$ and $t_j^{\text{exclude}}(r) = s$ is bounded by T^{-1} for each $t \in \mathbb{T}(r)$:

$$\frac{\Pr\left(\left\{t_j^{\text{exclude}}(r) = t\right\} \mid \alpha_j(r), h_i^{\leq L}\right)}{\Pr\left(\left\{t_j^{\text{exclude}}(r) = s\right\} \mid \alpha_j(r), h_i^{\leq L}\right)} \geq T^{-1}. \tag{50}$$

Suppose that $t_j^{\text{exclude}}(r) = s$, $(a_{j,s}, y_{j,s}) = (\bar{a}_j, \bar{y}_j)$, and $(a_{j,t}, y_{j,t}) = (\tilde{a}_j, \tilde{y}_j)$ is the case. Imagine that player j changes $t_j^{\text{exclude}}(r)$ from s to t: $t_j^{\text{exclude}}(r) = t$. At the same time, suppose that Nature changes $(a_{j,s}, y_{i,s})$ from $(\bar{a}_j, \bar{y}_j)$ to $(\tilde{a}_j, \tilde{y}_j)$, and changes $(a_{j,t}, y_{j,t})$ from $(\tilde{a}_j, \tilde{y}_j)$ to $(\bar{a}_j, \bar{y}_j)$. Then, $f_j^{\text{include}}[h_j^{\mathbb{T}(r)}]$ is unchanged. Conditional on $\lambda_i(l)(j) = G$, player j takes fully mixed strategy in each round. Hence, the likelihood between "$(a_{j,s}, y_{j,s}) = (\bar{a}_j, \bar{y}_j)$ and $(a_{j,t}, y_{j,t}) = (\tilde{a}_j, \tilde{y}_j)$" and "$(a_{j,s}, y_{j,s}) = (\tilde{a}_j, \tilde{y}_j)$ and $(a_{j,t}, y_{j,t}) = (\bar{a}_j, \bar{y}_j)$" is uniformly bounded by $\left(\varepsilon_{\text{support}}^{\text{action}} \varepsilon_{\text{support}}\right)^2$. Hence, we have

$$\frac{\Pr\left(\left\{t_j^{\text{exclude}}(r)=t\right\}\mid \alpha_j(r),h_i^{\leq L}\right)}{\Pr\left(\left\{t_j^{\text{exclude}}(r)=s\right\}\mid \alpha_j(r),h_i^{\leq L}\right)}\geq\left(\varepsilon_{\text{support}}^{\text{action}}\varepsilon_{\text{support}}\right)^2,$$

which implies (50) for a sufficiently large T, as desired. ■

14 Reward Function in the Coordination and Main Blocks

Before defining the strategy in the report block, we define player i's reward function. In Section 14.1, we define the variable $\theta_j(l)\in\{G,B\}$, on which the reward function depends, so that $\theta_j(l)$ satisfies certain properties. Given the definition of $\theta_j(l)$, Section 14.2 defines the reward function $\pi_i(x_j,h_j^{T_P+1})$, which is the summation of $\bar{\pi}_i(x_j)$, $\pi_i^{x_j}$, π_i^{review}, π_i^{adjust}, and π_i^{report}. $\bar{\pi}_i(x_j)$ is defined in (52); and π_i^{review} is defined in (57). ($\pi_i^{x_j}$ has been defined in Lemma 2.) The remaining rewards π_i^{adjust} and π_i^{report} will be defined in Lemma 13.

14.1 Definition of $\theta_j(l)\in\{G,B\}$

As will be seen in (53) and (54), $\theta_j(l)=B$ means that player j uses the type-3 reward for player i in review round l. $\theta_j(l)$ depends on the frequency of player j's history at the beginning of review round l, $f_j^{\text{include}}[h_j^{<l}]$. Although we should write $\theta_j(l)(f_j^{\text{include}}[h_j^{<l}])$ precisely speaking, for notational convenience, we write $\theta_j(l)$.

The formal definition of $\theta_j(l)$ is relegated to Appendix A.10.1. In Appendix A.10.1, we will define $\theta_j(l)$ such that $\theta_j(l)=B$ if (and only if, except for a small adjustment) at least one of the following four cases happens: When player j sent a message m in a round $r<l$, $\theta_j(m,f_j^{\text{include}}[h_j^{\mathbb{T}(r)}])=E$ happened; when player j received a message in a round $r<l$, $\theta_j(\text{receive},f_j^{\text{include}}[h_j^{\mathbb{T}(r)}])=E$ happened; in some review round $\hat{l}\leq l-1$, player j had $\lambda_i(\hat{l})(j)=\lambda_j(\hat{l})=G$, took $\alpha_j(\hat{l})=\alpha_j^{\rho}(x(j))$, and $\theta_j^{\text{review}}(x(j),f_j^{\text{include}}[h_j^{\mathbb{T}(\hat{l})}])=E$ happened; or when player j inferred x_i, x_j, or $\lambda_i(\tilde{l})$ with $\tilde{l}\leq l-1$ or took $\alpha_j(\tilde{l})$ for $\tilde{l}\leq l$,[19] the result of player j's mixture was a rare event (for example, when she inferred x_j, [Round $(x_j,3)$ $x_j(j)$] happened. As seen in Figure 9, this event happens only with probability η (rare). After [Round $(x_j,3)$ $x_j(j)$], player j has $\theta_j(l)=B$).

[19] Precisely speaking, since the mixture to determine $\alpha_j(l)$ happens at the beginning of review round l, $\theta_j(l)$ depends on player j's history at the beginning of review round l, $h_j^{<l}$, and her own mixture at the beginning of review round l.

There are following six implications of this definition. First, $\theta_j(l) = B$ is absorbing: $\theta_j(l) = B$ implies $\theta_j(l+1) = B$.

Second, the distribution of $\theta_j(l)$ does not depend on player i's strategy. Since we will define that player j uses the type-3 reward if and only if $\theta_j(l) = B$, this ensures that player i cannot control whether player j uses the type-3 reward or not, as mentioned in Section 9.2.

To see why this property is true, recall that Claims 3 and 4 of Lemma 6 ensure that the distribution of $\theta_j(m, f_j^{\text{include}}[h_j^{\mathbb{T}(r)}])$ and $\theta_j(\text{receive}, f_j^{\text{include}}[h_j^{\mathbb{T}(r)}])$ does not depend on player i's strategy. Moreover, Claim 2 of Lemma 7 ensures that the distribution of $\theta_j^{\text{review}}(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}(\hat{l})}])$ does not depend on player i's strategy.[20] Finally, player j's own mixture is out of player i's control. Hence, in total, the distribution of $\theta_j(l)$ does not depend on player i's strategy.

Moreover, recalling that E stands for erroneous (and so it is rare that the realization of a random variable is E), it is rare to have $\theta_j(l) = B$. This ensures that player j does not use the type-3 reward often, as mentioned in Section 9.2.

Third, player j with $\lambda_i(l)(j) = B$ has $\theta_j(l) = B$. This ensures that player i can condition that player j has $\lambda_i(l)(j) = G$ since otherwise the type-3 reward makes all the strategies optimal.

To see why $\lambda_i(l)(j) = B$ implies $\theta_j(l) = B$, take $\tilde{l} \leq l - 1$ such that player j switched from $\lambda_i(\tilde{l})(j) = G$ to $\lambda_i(\tilde{l}+1)(j) = B$. As seen in Figure 12 with indices i and j reversed, this means that player j listened to player i's message $\lambda_i(\tilde{l}+1)$. If [Regular θ $\lambda_i(j)$] was the case (see Section 13.4.4 with indices i and j reversed), then listening to player i's message is the rare realization of player j's own mixture and so $\theta_j(l) = B$.

If [Regular θ $\lambda_i(j)$] is not the case, then player j with $\lambda_j(\hat{l}) = G$ did not take $\alpha_j(\hat{l}) = \alpha_j^{\rho}(x(j))$ or had $\theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}(\hat{l})}]) = E$ for some $\hat{l} \leq \tilde{l}$. Since player j had $\lambda_i(\tilde{l})(j) = G$, she also had $\lambda_i(\hat{l})(j) = G$. Hence, if she did not take $\alpha_j^{\rho}(x(j))$ in review round $\hat{l}$, then this is a rare realization of her own mixture for $\alpha_j(\hat{l})$. If $\theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}(\hat{l})}]) = E$, then player j has $\theta_j(l) = B$ by definition.

In total, player j has $\theta_j(l) = B$ whenever player j switched from $\lambda_i(\tilde{l})(j) = G$ to $\lambda_i(\tilde{l}+1)(j) = B$.

Fourth, given $\theta_j(l) = G$, player j takes $\alpha_j^{\rho}(x(j))$ if $\lambda_j(l) = G$ and takes $\alpha_j^{*,\rho}(x(j))$ if $\lambda_j(l) = B$. In other words, if $\text{action}_j(l) = E$, then $\theta_j(l) = B$. Moreover, since $\theta_j(l) = B$ is absorbing, if $\text{action}_j(\hat{l}) = E$ for some $\hat{l} \leq l$, then $\theta_j(l) = B$.

To see why, note that the third property implies that given $\theta_j(l) = G$, player j has $\lambda_i(l)(j) = G$.

[20] Since $\theta_j^{\text{review}}(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}(\hat{l})}])$ is defined only if player j takes $\alpha_j(\hat{l}) = \alpha_j^{\rho}(x(j))$, we need some adjustment so that player i does not want to control the distribution of $\theta_j(l)$ by changing the distribution of $\alpha_j(\hat{l})$.

Given $\lambda_i(l)(j) = G$, Figure 10 with indices i and j reversed ensures that the event that player j takes $\alpha_j^{*,\rho}(x(j))$ with $\lambda_j(l) = G$ or takes $\alpha_j^{\rho}(x(j))$ with $\lambda_j(l) = B$ is rare (happens with probability no more than η). Since we have $\theta_j(l) = B$ after a rare realization, the result follows.

Fifth, if player j has $x_j(j) \neq x_j$, then we have $\theta_j(l) = B$. Recall that player j's state x_j controls player i's payoff (see (5) for example). This property means that when player j does not adhere to her own state ("gives up" controlling player i's payoff), player i is indifferent between any action.

Figure 9 ensures that, if $x_j(j) \neq x_j$, then either $\theta_j(x_j, f_j^{\text{include}}[h_j^{\mathbb{T}(x_j,2)}]) = E$ or [Round $(x_j, 3)$ $x_j(j)$] (rare realization of player j's mixture) happens. Hence, $x_j(j) \neq x_j$ implies $\theta_j(l) = B$.

Finally, suppose that player i conditions on $\lambda_i(l)(j) = G$ (see the third property for why this conditioning is optimal). Player i believes that, if she has $x(i) \neq x(j)$ or $\lambda_j(l)(i) = G$, then $\lambda_j(l) = G$ or $\theta_j(l) = B$ with a high probability. That is, as seen in (26), player i believes that, if the coordination does not go well, then player j uses the type-3 reward with a high probability.

To see why player i holds such a belief, we first explain that, if player i has $x(i) \neq x(j)$, then she believes that $\theta_j(l) = B$ with a high probability. $x(i) \neq x(j)$ happens only if one of the following two cases happens: $x_j(i) \neq x_j(j)$ or $x_i(i) \neq x_i(j)$. We consider these two cases in the sequel.

Given (39) in Lemma 8, if player i has $x_j(i) \neq x_j(j)$, then given x_j and $x_j(j)$, player i believes that [Round $(x_j, 3)$ $x_j(j)$] (a rare realization of player j's mixture), $\theta_j(x_j, f_j^{\text{include}}[h_j^{\mathbb{T}(x_j,1)}]) = E$, or $\theta_j(x_j, f_j^{\text{include}}[h_j^{\mathbb{T}(x_j,2)}]) = E$ at the end of coordination block for x_j. Since player j's continuation strategy does not depend on her history in the coordination block for x_j conditional on $x_j(j)$, this belief means that, in review round l, player i believes that $\theta_j(l) = B$.

Again, given (40) in Lemma 8 with indices i and j reversed, if player i has $x_i(i) \neq x_i(j)$, then given $x_i(j)$, player i believes that [Round $(x_i, 1)$ $x_i(j)$] (a rare realization of player j's mixture), $\theta_j(\text{receive}, f_j^{\text{include}}[h_j^{\mathbb{T}(x_i,2)}]) = E$, or $\theta_j(x_i(j), f_j^{\text{include}}[h_j^{\mathbb{T}(x_i,3)}]) = E$ at the end of coordination block for x_i. Since player j's continuation strategy does not depend on her history in the coordination block for x_i given $x_i(j)$, this belief means that, in review round l, player i believes that $\theta_j(l) = B$.

Hence, we are left to show that, given $x(i) = x(j)$, if player i has $\lambda_j(l)(i) = G$, then she believes that $\lambda_j(l) = G$ or $\theta_j(l) = B$ with a high probability. Recall that Lemma 9 ensures that player i with $x(i) = x(j)$ and $\lambda_j(l)(i) = G$ believes that player j has $\lambda_j(l) = G$, $\theta_j(\lambda_j(l), f_j^{\text{include}}[h_j^{\mathbb{T}(\lambda_j(l))}]) = E$, $\text{action}_j(\hat{l}) = E$ for some $\hat{l} \leq l - 1$, or $\theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}(\hat{l})}]) = E$ for some $\hat{l} \leq l - 1$. As seen in the fourth property, $\text{action}_j(\hat{l}) = E$ for some $\hat{l} \leq l - 1$ means $\theta_j(l) = B$. In addition, $\theta_j(\lambda_j(l), f_j^{\text{include}}[h_j^{\mathbb{T}(\lambda_j(l))}]) = E$ and $\theta_j(x(j), f_j^{\text{include}}[h_j^{\mathbb{T}(\hat{l})}]) = E$ imply $\theta_j(l) = B$ by definition. Hence, Lemma 9 implies that player i believes that player j has $\lambda_j(l) = G$ or $\theta_j(l) = B$ with a high probability.

In total, we prove the following lemma:

Lemma 11 *For each $j \in I$ and $l \in \{1, ..., L\}$, we can define a random variable $\theta_j(l) \in \{G, B\}$ whose distribution is determined by player j's history $h_j^{\leq l}$ such that the following properties hold:*

1. *$\theta_j(l) = B$ is absorbing: $\theta_j(l) = B \Rightarrow \theta_j(l+1) = B$.*
2. *For each $x_j \in \{G, B\}$ and $\sigma_i \in \Sigma_i$, the distribution of $\theta_j(l)$, denoted by $\Pr(\theta_j(l) \mid \sigma_j(x_j), \sigma_i)$, does not depend on player i's strategy $\sigma_i \in \Sigma_i$. Moreover, $\theta_j(l) = B$ is rare: $\Pr(\{\theta_j(l) = B\} \mid \sigma_j(x_j), \sigma_i) \leq (15 + 8L)\eta$.*
3. *For each $h_j^{\leq l}$ with $\lambda_i(l)(j) = B$, we have $\theta_j(l) = B$.*
4. *For each $j \in I$ and $l \in \{1, ..., L\}$, for each $h_j^{\leq l}$ with $\theta_j(l) = G$, player j takes $\alpha_j^{\rho}(x(j))$ if $\lambda_j(l) = G$ and takes $\alpha_j^{*,\rho}(x(j))$ if $\lambda_j(l) = B$.*
5. *For each $j \in I$ and $l \in \{1, ..., L\}$, for each $h_j^{\leq l}$ with $x_j(j) \neq x_j$, we have $\theta_j(l) = B$.*
6. *For a sufficiently large T, for each $i \in I$, $x_j \in \{G, B\}$, $l \in \{1, ..., L\}$, and $h_i^{\leq l}$, if $x(i) \neq x(j)$ or $\lambda_j(l)(i) = G$ with a positive probability with $h_i^{\leq l}$, then player i believes that $\lambda_j(l) = G$ or $\theta_j(l) = B$ with a high probability conditional on x_j, $x(j)$, and $\lambda_i(l)(j) = G$: If $h_i^{\leq l}$ satisfies*

$$\Pr\left(\{x(i) \neq x(j) \vee \lambda_j(l)(i) = G\} \mid x_j, x(j), h_i^{\leq l}\right) > 0,$$

then we have

$$\Pr\left(\{\lambda_j(l) = G \vee \theta_j(l) = B\} \mid x_j, \{\lambda_i(l)(j) = G\}, h_i^{\leq l}\right) \geq 1 - \exp(-T^{\frac{1}{3}}).$$

Proof. See Appendix A.10. ■

14.2 Definition of the Reward Function

We are now ready to define player i's reward function, given the above definition of $\theta_j(l)$. The total reward is the summation of (i) a constant $\bar{\pi}_i(x_j)$, (ii) the reward for rounds other than review rounds,

$\sum_{r=1:\text{round } r \text{ is not a review round}}^{R}\sum_{t\in\mathbb{T}(r)}\pi_i^{x_j}(a_{j,t},y_{j,t})$, (iii) the reward for review round $l=1,...,L$, π_i^{review} and π_i^{adjust}, and (iv) the reward for the report block π_i^{report}:

$$\begin{aligned}\pi_i(x_j,h_j^{T_P+1}) &= \bar{\pi}_i(x_j)+\sum_{\substack{r=1\\ \text{round } r \text{ is not a review round}}}^{R}\sum_{t\in\mathbb{T}(r)}\pi_i^{x_j}(a_{j,t},y_{j,t}) \\ &\quad+\sum_{l=1}^{L}\left\{\pi_i^{\text{review}}(x_j,f_j^{\text{include}}[h_j^{\leq l}],f_j[h_j^{\mathbb{T}(l)}],l)+\pi_i^{\text{adjust}}(x_j,h_j^{\leq l},h_j^{\text{report}},l)\right\} \\ &\quad+\pi_i^{\text{report}}(x_j,h_j^{\leq L},h_j^{\text{report}}).\end{aligned}\tag{51}$$

We define and explain each term in the sequel.

First, we define the constant $\bar{\pi}_i(x_j)$ such that

$$\bar{\pi}_i(x_j)=T_P v_i(x_j)-\left\{\begin{array}{c}\mathbb{E}\left[\sum_{l=1}^{L}\left\{1_{\{\theta_j(l)=G\}}Tu_i(x_j)+1_{\{\theta_j(l)=B\}}\left(\text{sign}\left(x_j\right)LT\bar{u}+T\bar{u}_i^{x_j}\right)\right\}\mid x_j\right] \\ +\left\{\left(2+2L\right)T^{\frac{1}{2}}+2T^{\frac{2}{3}}\right\}\bar{u}_i^{x_j}\end{array}\right\}.\tag{52}$$

We define this constant so that player i's average value from the review phase is equal to $v_i(x_j)$ in order to satisfy promise keeping. As will be seen in (85), the value in the biggest parenthesis is the value (without being divided by T_P) from the coordination, main, and report block.

Second, the term

$$\sum_{\substack{r=1\\ \text{round } r \text{ is not a review round}}}^{R}\sum_{t\in\mathbb{T}(r)}\pi_i^{x_j}(a_{j,t},y_{j,t})$$

cancels out the effect of the instantaneous utilities on the incentives for the rounds which are not review rounds. This reward incentivizes the players to take actions in order to exchange messages in the coordination block and supplemental rounds.

Third, we define π_i^{review} such that π_i^{review} satisfies the properties mentioned in Section 9. Firstly, if $\lambda_i(l)(j)=B$ (recall that $\lambda_i(l)(j)=B$ implies $\theta_j(l)=B$), then player j uses the type-3 reward

$$\text{sign}\left(x_j\right)LT\bar{u}+\sum_{t\in\mathbb{T}(l)}\pi_i^{x_j}(a_{j,t},y_{j,t}).\tag{53}$$

Compared to (25) in Section 9, we add $\text{sign}(x_j)\,LT\bar{u}$. Since $\text{sign}(x_j)\times\text{sign}\left(x_j\right)LT\bar{u}=LT\bar{u}$, this term gives us enough slack in self generation once the type-3 reward is used.

Secondly, if $\lambda_i(l)(j) = G$, then if $\theta_j(l) = B$, then again player j uses the type-3 reward

$$\text{sign}(x_j)\, LT\bar{u} + \sum_{t \in \mathbb{T}(l)} \pi_i^{x_j}(a_{j,t}, y_{j,t}). \tag{54}$$

If $\lambda_i(l)(j) = \theta_j(l) = G$, then the reward function depends on $\lambda_j(l)$ (whether self generation is an issue or not). If $\lambda_j(l) = G$ (self generation is not an issue), then player j uses the type-1 reward

$$T\left\{u_i(x_j) - u_i(\alpha^\rho(x(j)))\right\} + \sum_{t \in \mathbb{T}(l)} \pi_i[\alpha^\rho(x(j))](a_{j,t}, y_{j,t}). \tag{55}$$

By Lemma 2, any strategy of player i is optimal. In addition, since the expected payoff from $u_i(a_t)+\pi_i[\alpha^\rho(x(j))](a_{j,t}, y_{j,t})$ is $u_i(\alpha^\rho(x(j)))$, the value from the review round (without being divided by T or T_P) is $Tu_i(x_j)$. Here, we ignore the effect of her strategy in review round l on the payoffs from the subsequent rounds. See Section 16 for the formal proof of why it is optimal for player i to ignore this effect.

On the other hand, if $\lambda_j(l) = B$ (self generation is an issue), then player j uses the type-2 reward

$$T\{u_i(x_j) - u_i(BR_i(\alpha_j^{*,\rho}(x(j))), \alpha_j^{*,\rho}(x(j)))\}. \tag{56}$$

By Claim 4 of Lemma 11, player j with $\theta_j(l) = G$ and $\lambda_j(l) = B$ takes $\alpha_j^{*,\rho}(x(j))$. Since the reward is constant, as long as the coodination goes well and player i has $x(i) = x(j)$ and $\lambda_j(l)(i) = B$, player i's equilibrium strategy, which takes a static best response to $\alpha_j^{*,\rho}(x(j))$, is optimal. In addition, since the expected payoff from $u_i(a_t)$ is $u_i(BR_i(\alpha_j^{*,\rho}(x(j))), \alpha_j^{*,\rho}(x(j)))$, the value from the review round (without being divided by T or T_P) is $Tu_i(x_j)$.

In total, we define the reward π_i^{review} as follows:

$$
\begin{aligned}
& \pi_i^{\text{review}}(x_j, f_j^{\text{include}}[h_j^{\leq l}], f_j[h_j^{\mathbb{T}(l)}], l) \qquad (57)\\
= \quad & 1_{\{\lambda_i(l)(j)=B\}} \underbrace{\left\{ \operatorname{sign}(x_j)\, LT\bar{u} + \sum_{t\in\mathbb{T}(l)} \pi_i^{x_j}(a_{j,t}, y_{j,t}) \right\}}_{(53)} \\
& +1_{\{\lambda_i(l)(j)=G\}} \left\{ \begin{array}{c} 1_{\{\theta_j(l)=G\}} \left\{ \begin{array}{c} 1_{\{\lambda_j(l)=G\}} \underbrace{\left\{ \begin{array}{c} T\{u_i(x_j) - u_i(\alpha^\rho(x(j)))\} \\ +\sum_{t\in\mathbb{T}(l)} \pi_i[\alpha^\rho(x(j))](a_{j,t}, y_{j,t}) \end{array} \right\}}_{(55)} \\ +1_{\{\lambda_j(l)=B\}} \underbrace{T\{u_i(x_j) - u_i(BR_i(\alpha_j^{*,\rho}(x(j))), \alpha_j^{*,\rho}(x(j)))\}}_{(56)} \end{array} \right\} \\ +1_{\{\theta_j(l)=B\}} \underbrace{\left\{ \operatorname{sign}(x_j)\, LT\bar{u} + \sum_{t\in\mathbb{T}(l)} \pi_i^{x_j}(a_{j,t}, y_{j,t}) \right\}}_{(54)} \end{array} \right\}.
\end{aligned}
$$

Note that $x(j)$, $\lambda_i(l)(j)$, $\lambda_j(l)$, and $\theta_j(l)$ are determined by x_j, $f_j^{\text{include}}[h_j^{\leq l}]$, and player j's own mixture, and $\sum_{t\in\mathbb{T}(l)} \pi_i^{x_j}(a_{j,t}, y_{j,t})$ and $\sum_{t\in\mathbb{T}(l)} \pi_i[\alpha(x(j))](a_{j,t}, y_{j,t})$ are determined by $f_j[h_j^{\mathbb{T}(l)}]$. Hence, π_i^{review} is a function of x_j, $f_j^{\text{include}}[h_j^{\leq l}]$, $f_j[h_j^{\mathbb{T}(l)}]$, and l.

Given this definition of π_i^{review}, player i's strategy is optimal unless the coordination goes wrong: Player i has $x(i) \neq x(j)$ or $\lambda_j(l)(i) = G$ even though player j has $\theta_j(l) = G$ (this implies $\lambda_i(l)(j) = G$ by Claim 3 of Lemma 11) and $\lambda_j(l) = B$. One may wonder what if player i has $x(i) \neq x(j)$ or $\lambda_j(l)(i) = B$ but player j has $\theta_j(l) = \lambda_j(l) = G$. This case is not a problem since any strategy of player i is optimal given $\lambda_j(l) = G$. To see why, note that Lemma 2 ensures that, with type-1 reward (55), any strategy is optimal.

Let $\Lambda_i(x(j), l)$ be the set of player i's histories $f_i[h_i^{<l}]$ such that player i with $f_i[h_i^{<l}]$ has $x(j) = x(i)$ and $\lambda_j(l)(i) = B$ with probability one:

$$
\Lambda_i(x(j), l) \equiv \left\{ f_i[h_i^{<l}] : \begin{array}{c} \text{For any realization of } t_i^{\text{exclude}}(r) \text{ for } r < l, \\ \text{player } i \text{ with } f_i[h_i^{<l}] \text{ has } x(j) = x(i) \text{ and } \lambda_j(l)(i) = B \end{array} \right\}.
$$

The discussion above means that, if and only if

$$\lambda_i(l)(j) = G \wedge \theta_j(l) = G \wedge \lambda_j(l) = B \wedge f_i[h_i^{<l}] \notin \Lambda_i(x(j), l), \tag{58}$$

player i's strategy would be suboptimal if the reward for review round l were $\pi_i^{\text{review}}(x_j, f_j^{\text{include}}[h_j^{\leq l}], f_j[h_j^{\mathbb{T}(l)}], l)$.

If player j used the following reward function π_i^{target} rather than π_i^{review}, then player i's strategy would be optimal after each history: Again, if $\lambda_i(l)(j) = B$ or $\theta_j(l) = B$, then player j uses the type-3 reward (53) or (54), respectively. If $\lambda_i(l)(j) = \theta_j(l) = G$ and $\lambda_j(l) = G$, player j uses the type-1 reward (55) as before. In addition, if $\lambda_i(l)(j) = \theta_j(l) = G$ and "$\lambda_j(l) = B$ and $f_i[h_i^{<l}] \notin \Lambda_i(x(j), l)$", then player j uses the type-1 reward as well. If $\lambda_i(l)(j) = \theta_j(l) = G$ and "$\lambda_j(l) = B$ and $f_i[h_i^{<l}] \in \Lambda_i(x(j), l)$", player j uses the type-2 reward (56).

Intuitively, whenever player i's history satisfies (58), player j uses the type-1 reward rather than type-2. Since Lemma 2 ensures that the type-1 reward makes any strategy of player i optimal, with such a reward function, player i's strategy would be optimal after each history.

In total, π_i^{target} is defined as

$$\begin{aligned}
&\pi_i^{\text{target}}(x_j, f_i[h_i^{\leq l}], f_j^{\text{include}}[h_j^{\leq l}], f_j[h_j^{\mathbb{T}(l)}], l) \qquad\qquad (59)\\
= \; & 1_{\{\lambda_i(l)(j)=B\}} \left\{ \text{sign}(x_j)\, LT\bar{u} + \sum_{t \in \mathbb{T}(l)} \pi_i^{x_j}(a_{j,t}, y_{j,t}) \right\} \\
&+ 1_{\{\lambda_i(l)(j)=G\}} \\
&\times \left\{ \begin{array}{c}
1_{\{\theta_j(l)=G\}} \left\{ \begin{array}{c}
1_{\{\lambda_j(l)=G \vee \underbrace{\{\lambda_j(l)=B \wedge f_i[h_i^{<l}] \notin \Lambda_i(x(j),l)\}}_{\substack{\text{the case with which} \\ \pi_i^{\text{review}} \text{ is suboptimal}}}\}} \\
\times \left\{ \begin{array}{c} T\{u_i(x_j) - u_i(\alpha^{\rho}(x(j)))\} \\ + \sum_{t \in \mathbb{T}(l)} \pi_i[\alpha^{\rho}(x(j))](a_{j,t}, y_{j,t}) \end{array} \right\} \\
+ 1_{\{\lambda_j(l)=B \wedge f_i[h_i^{<l}] \in \Lambda_i(x(j),l)\}} \{u_i(x_j) - u_i(BR_i(\alpha_j^{*,\rho}(x(j))), \alpha_j^{*,\rho}(x(j)))\}
\end{array} \right\} \\
+ 1_{\{\theta_j(l)=B\}} \left\{ \text{sign}(x_j)\, LT\bar{u} + \sum_{t \in \mathbb{T}(l)} \pi_i^{x_j}(a_{j,t}, y_{j,t}) \right\}
\end{array} \right\}.
\end{aligned}$$

Since π_i^{target} is a function of whether $f_i[h_i^{<l}]$ is included in $\Lambda_i(x(j), l)$ or not, this reward is a function of $f_i[h_i^{<l}]$ as well.

Instead of replacing π_i^{review} with π_i^{target}, we add the expected difference between π_i^{review} and π_i^{target} to π_i^{review}: The reward π_i^{adjust}, which is added to π_i^{review} in (51), is defined so that the expected value of π_i^{adjust} is equal to the expected difference between π_i^{review} and π_i^{target}. In particular, in Lemma 13, we will define $\pi_i^{\text{adjust}}(x_j, h_j^{\leq l}, h_j^{\text{report}}, l)$ and $\sigma_i^{\text{report}}|_{h_i^{\leq L}}$ so that

$$
\begin{aligned}
& \mathbb{E}\left[\pi_i^{\text{adjust}}(x_j, h_j^{\leq l}, h_j^{\text{report}}, l) \mid x_j, \{\lambda_i(l)(j) = G\}, \sigma_i^{\text{report}}|_{h_i^{\leq l}}, h_i^{\leq l}\right] \\
= \; & \mathbb{E}\left[\pi_i^{\text{target}}(x_j, f_i[h_i^{\leq l}], f_j^{\text{include}}[h_j^{\leq l}], f_j[h_j^{\mathbb{T}(l)}], l) \mid x_j, \{\lambda_i(l)(j) = G\}, h_i^{\leq l}\right] \\
& -\mathbb{E}\left[\pi_i^{\text{review}}(x_j, f_j^{\text{include}}[h_j^{\leq l}], f_j[h_j^{\mathbb{T}(l)}], l) \mid x_j, \{\lambda_i(l)(j) = G\}, h_i^{\leq l}\right] \qquad (60)
\end{aligned}
$$

if $\lambda_i(l)(j) = G$ and $\pi_i^{\text{adjust}}(x_j, h_j^{\leq l}, h_j^{\text{report}}, l)$ is zero if $\lambda_i(l)(j) = B$. (Note that the difference between π_i^{review} and π_i^{target} is identically equal to zero if $\lambda_i(l)(j) = B$.) Here, $\sigma_i^{\text{report}}|_{h_i^{\leq l}}$ is the conditional distribution of player i's strategy in the report block given $h_i^{\leq l}$. This property of π_i^{adjust} corresponds to (28) in Section 9.4.

We postpone the definition of π_i^{adjust} and π_i^{report} to Lemma 13.

14.3 Small Expected Difference between π_i^{target} and π_i^{review}

Recall that Claim 6 of Lemma 11 ensures that player i who has $x(i) \neq x(j)$ or $\lambda_j(l)(i) = G$ believes that it is rare for player j to have $\lambda_j(l) = B$ and $\theta_j(l) = G$ given $\lambda_i(l)(j) = G$. On the other hand, the difference between π_i^{target} and π_i^{review} is not zero only if player i has $x(i) \neq x(j)$ or $\lambda_j(l)(i) = G$ with a positive probability and player j has $\lambda_j(l) = B$ and $\theta_j(l) = G$. Hence, the expected difference (and so the expected value of π_i^{adjust}) is close to zero, conditional on $\lambda_i(l)(j) = G$. This closeness will play an important role when we define player i's strategy in the report block in Lemma 13.

Lemma 12 *For a sufficiently large T, for each $i \in I$, $l \in \{1, ..., L\}$, $x_j \in \{G, B\}$, and $h_i^{\leq l}$, the difference (60) depends only on the frequency of player i's history:*

$$
\begin{aligned}
& \mathbb{E}\left[\pi_i^{\text{target}}(x_j, f_i[h_i^{\leq l}], f_j^{\text{include}}[h_j^{\leq l}], f_j[h_j^{\mathbb{T}(l)}], l) \mid x_j, \{\lambda_i(l)(j) = G\}, h_i^{\leq l}\right] \\
& -\mathbb{E}\left[\pi_i^{\text{review}}(x_j, f_j^{\text{include}}[h_j^{\leq l}], f_j[h_j^{\mathbb{T}(l)}], l) \mid x_j, \{\lambda_i(l)(j) = G\}, h_i^{\leq l}\right] \\
= \; & \mathbb{E}\left[\pi_i^{\text{target}}(x_j, f_i[h_i^{\leq l}], f_j^{\text{include}}[h_j^{\leq l}], f_j[h_j^{\mathbb{T}(l)}], l) \mid x_j, \{\lambda_i(l)(j) = G\}, f_i[h_i^{\leq l}]\right] \\
& -\mathbb{E}\left[\pi_i^{\text{review}}(x_j, f_j^{\text{include}}[h_j^{\leq l}], f_j[h_j^{\mathbb{T}(l)}], l) \mid x_j, \{\lambda_i(l)(j) = G\}, f_i[h_i^{\leq l}]\right]
\end{aligned}
$$

Moreover, this difference is sufficiently small:

$$\left| \begin{array}{c} \mathbb{E}\left[\pi_i^{\text{target}}(x_j, f_i[h_i^{\leq l}], f_j^{\text{include}}[h_j^{\leq l}], f_j[h_j^{\mathbb{T}(l)}], l) \mid x_j, \{\lambda_i(l)(j) = G\}, f_i[h_i^{\leq l}]\right] \\ -\mathbb{E}\left[\pi_i^{\text{review}}(x_j, f_j^{\text{include}}[h_j^{\leq l}], f_j[h_j^{\mathbb{T}(l)}], l) \mid x_j, \{\lambda_i(l)(j) = G\}, f_i[h_i^{\leq l}]\right] \end{array} \right| \leq \exp(-T^{\frac{1}{4}}).$$

Proof. Both π_i^{review} and π_i^{target} depend only on x_j, $f_i[h_i^{\leq l}]$, $f_j^{\text{include}}[h_j^{\leq l}]$, and $f_j[h_j^{\mathbb{T}(l)}]$. In addition, $\lambda_i(l)(j)$ depends only on $f_j^{\text{include}}[h_j^{\leq l}]$. Since $t_j^{\text{exclude}}(r)$ is random, $f_i[h_i^{\leq l}]$ is a sufficient statistic.

Recall that $f_i[h_i^{<l}] \notin \Lambda_i(x(j), l)$ implies that player i has either $x(i) \neq x(j)$ or $\lambda_j(l)(i) = G$. Comparing (57) and (59), π_i^{target} and π_i^{review} differ only if $\lambda_i(l)(j) = G \wedge \theta_j(l) = G \wedge \lambda_j(l) = B \wedge f_i[h_i^{<l}] \notin \Lambda_i(x(j), l)$. Since both π_i^{review} and π_i^{target} are of order T, it suffices to show that, conditional on x_j, $x(j)$, $\lambda_i(l)(j) = G$, and $f_i[h_i^{\leq l}]$, if player i has $x(i) \neq x(j)$ or $\lambda_j(l)(i) = G$, then player i believes that $\theta_j(l) = B$ or $\lambda_j(l) = G$ with a high probability:

$$\Pr\left(\{\theta_j(l) = B \vee \lambda_j(l) = G\} \mid x_j, x(j), \{\lambda_i(l)(j) = G\}, h_i^{\leq l}\right) \geq 1 - \exp(-T^{\frac{1}{3}}).$$

Claim 6 of Lemma 11 establishes the result. ■

15 Report Block

We now define $\sigma_i^{\text{report}}|_{h_i^{\leq L}}$, π_i^{adjust}, and π_i^{report} in Lemma 13. This finishes defining the strategy $\sigma_i(x_i)$ and reward function $\pi_i(x_j, h_j^{T_P+1})$:

Lemma 13 *There exists $K_{\text{report}} \in \mathbb{N}$ such that there exist $\sigma_i^{\text{report}}|_{h_i^{\leq L}}$, $\pi_i^{\text{adjust}}(x_j, h_j^{\leq l}, h_j^{\text{report}}, l)$ with*

$$\pi_i^{\text{adjust}}(x_j, h_j^{\leq l}, h_j^{\text{report}}, l) \in \left[-\exp(-T^{\frac{1}{5}}), \exp(-T^{\frac{1}{5}})\right],$$

and

$$\pi_i^{\text{report}}(x_j, h_j^{\leq L}, h_j^{\text{report}}) \in [-K_{\text{report}} T^{\frac{11}{12}}, K_{\text{report}} T^{\frac{11}{12}}]$$

such that, for sufficiently large T, for each $x_j \in \{G, B\}$ and $h_i^{\leq L}$, *1.* $\sigma_i^{\text{report}}|_{h_i^{\leq L}}$ *maximizes*

$$\mathbb{E}\left[\sum_{t:\textit{report block}} u_i(a_t) + \pi_i^{\text{report}}(x_j, h_j^{\leq L}, h_j^{\text{report}}) + \sum_{l=1}^{L} \pi_i^{\text{adjust}}(x_j, h_j^{\leq l}, h_j^{\text{report}}, l) \mid x_j, h_i^{\leq L}\right]. \quad (61)$$

2. From player i's perspective at the end of review round l, the expected adjustment is equal to (60):

$$\begin{aligned} & \mathbb{E}\left[\pi_i^{\text{adjust}}(x_j, h_j^{\leq l}, h_j^{\text{report}}, l) \mid \{\lambda_i(l)(j) = G\}, \sigma_i^{\text{report}}|_{h_i^{\leq l}}, h_i^{\leq l}\right] \\ = \; & \mathbb{E}\left[\pi_i^{\text{target}}(x_j, f_i[h_i^{\leq l}], f_j^{\text{include}}[h_j^{\leq l}], f_j[h_j^{\mathbb{T}(l)}], l) \mid x_j, \{\lambda_i(l)(j) = G\}, h_i^{\leq l}\right] \\ & -\mathbb{E}\left[\pi_i^{\text{review}}(x_j, f_j^{\text{include}}[h_j^{\leq l}], f_j[h_j^{\mathbb{T}(l)}], l) \mid x_j, \{\lambda_i(l)(j) = G\}, h_i^{\leq l}\right] \end{aligned} \quad (62)$$

if $\lambda_i(l)(j) = G$*, and the adjustment is equal to zero if* $\lambda_i(l)(j) = B$*.*

3. The equilibrium value satisfies

$$\mathbb{E}\left[\sum_{t:\textit{report block}} u_i(a_t) + \pi_i^{\text{report}}(x_j, h_j^{\leq L}, h_j^{\text{report}}) \mid x_j, \sigma_i^{\text{report}}|_{h_i^{\leq L}}, h_i^{\leq L}\right] = 0. \quad (63)$$

Proof. See Appendix A.11. ■

Let us intuitively explain why this lemma is true. To this end, we first assume that the players have access to the public randomization device and can communicate via cheap talk in Section 15.1. Then, we explain how to dispense with cheap talk, keeping public randomization device in Section 15.2. Finally, we explain how to dispense with public randomization device in Section 15.3. The formal proof in Appendix A.11 does not use public randomization device or cheap talk.

15.1 Report Block with Cheap Talk

Intuitively, player i sends $f_i[h_i^{\leq L}] = (f_i[h_i^{\mathbb{T}(r)}])_{r=1}^R$ to player j so that player j can calculate π_i^{adjust} to satisfy (62). To this end, it will be useful to divide each round r into $|\mathbb{T}(r)|^{\frac{1}{3}}$ subrounds with equal length, that is, each subround lasts for $|\mathbb{T}(r)|^{\frac{2}{3}}$ periods: Let $t(r)+1$ be the first period of round r: $\mathbb{T}(r) = \{t(r)+1, ..., t(r)+|\mathbb{T}(r)|\}$. Subround $k(r)$ of round r consists of $\mathbb{T}(r, k(r)) \equiv \{t(r) + (k(r)-1)|\mathbb{T}(r)|^{\frac{2}{3}} + 1, ..., t(r) + k(r)|\mathbb{T}(r)|^{\frac{2}{3}}\}$. Let $f_i[h_i^{\mathbb{T}(r,k(r))}]$ be the frequency of player i's

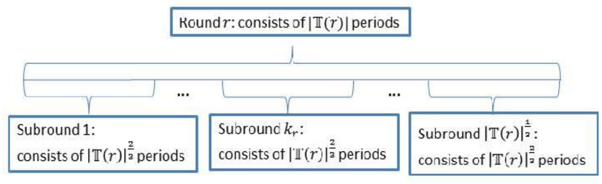

Figure 13: Rounds and subrounds

history in subround $k(r)$ of round r:

$$f_i[h_i^{\mathbb{T}(r,k(r))}](a_i, y_i) \equiv \frac{\#\{t \in \mathbb{T}(r, k(r)) : a_{i,t} = a_i, y_{i,t} = y_i\}}{|\mathbb{T}(r)|^{\frac{2}{3}}}.$$

See Figure 13 for illustration:

15.1.1 Player i's Strategy $\sigma_i^{\text{report}}|_{h_i^{\leq L}}$

At the beginning of the report block, the players draw a public randomization device, which determines whether player 1 or 2 sends the message $f_i[h_i^{\leq L}]$. Each player is picked with probability $\frac{1}{2}$ and only one player is picked at the same time (that is, with public randomization, only one player sends the history). Suppose player i is picked.

For each round r, player i sends the frequency of each subround $(f_i[h_i^{\mathbb{T}(r,k(r))}])_{k(r)=1}^{|\mathbb{T}(r)|^{\frac{1}{3}}}$. Then, for each round r, the players draw a public randomization device, which picks one subround $k(r)$ with probability $|\mathbb{T}(r)|^{-\frac{1}{3}}$ randomly. Suppose that $k_i(r) \in \{1, ..., |\mathbb{T}(r)|^{\frac{1}{3}}\}$ is picked for round r. Then, player i sends the entire history of the picked subround $k_i(r)$: $(a_{i,t}, y_{i,t})_{t \in \mathbb{T}(r,k_i(r))}$. (The reason why player i sends the message this way rather than simply sending $(a_{i,t}, y_{i,t})_{t \in \mathbb{T}(r)}$ is to reduce the cardinality of the message, looking ahead to dispensing with cheap talk.)

15.1.2 Reward Functions π_i^{report} and π_i^{adjust}

Player j incentivizes player i to tell the truth by π_i^{report} as follows: First, player j incentivizes player i to tell the truth about $(a_{i,t}, y_{i,t})_{t \in \mathbb{T}(r,k_i(r))}$ by giving the reward equal to

$$-T^{-11} \sum_{t \in \mathbb{T}(r,k_i(r))} 1_{\{t = t_j^{\text{exclude}}(r)\}} 1_{\{r < l_i^*\}} \left\| \mathbf{1}_{a_{j,t}, y_{j,t}} - \mathbb{E}\left[\mathbf{1}_{a_{j,t}, y_{j,t}} \mid \alpha_j(r), \hat{a}_{i,t}, \hat{y}_{i,t}\right] \right\|^2. \tag{64}$$

In general, $1_{\{X\}}$ is equal to one if statement X is true and zero otherwise. Here, $1_{\left\{t=t_j^{\text{exclude}}(r)\right\}}=1$ if and only if $t=t_j^{\text{exclude}}(r)$; and $1_{\left\{r<l_i^*\right\}}$ is equal to one if and only if round r is before (and not equal to) review round l_i^* (recall that l_i^* is the first review round with $\lambda_i(l)(j)=B$). We define $1_{\left\{r<l_i^*\right\}}=1$ for each $r=1,...,R$ if player j has $\lambda_i(l)(j)=G$ for each $l=1,...,L$. In addition, $\mathbf{1}_{a_{j,t},y_{j,t}}$ is $|A_i|\,|Y_i|$-dimensional vector whose element corresponding to $(a_{j,t},y_{j,t})$ is one and other elements are equal to zero. Moreover, $\alpha_j(r)$ is player j's strategy in round r and $(\hat{a}_{i,t},\hat{y}_{i,t})$ is player i's message.

$-\left\|\mathbf{1}_{a_{j,t},y_{j,t}}-\mathbb{E}\left[\mathbf{1}_{a_{j,t},y_{j,t}}\mid\alpha_j(r),\hat{a}_{i,t},\hat{y}_{i,t}\right]\right\|^2$ is called the scoring rule in statistics, and incentivizes player i to tell the truth about $(\hat{a}_{i,t},\hat{y}_{i,t})$ if player i believes that $(a_{j,t},y_{j,t})$ is distributed according to $\Pr(\cdot\mid\alpha_j(r),a_{i,t},y_{i,t})$. Moreover, given Assumption 4, the incentive is strict if $\alpha_j(r)$ is a fully mixing strategy. (See Lemma 46 for the formal proof.)

Here, player j punishes player i based on the period $t_j^{\text{exclude}}(r)$. Recall that player j does not use $(a_{j,t},y_{j,t})$ with $t=t_j^{\text{exclude}}(r)$ to determine her continuation strategy. Hence, player i cannot learn $(a_{j,t},y_{j,t})$ and believes that $(a_{j,t},y_{j,t})$ is distributed according to $\Pr(\cdot\mid\alpha_j(r),a_{i,t},y_{i,t})$. By Lemma 10, with Assumption 1, player i cannot learn what period is $t_j^{\text{exclude}}(r)$. Together with the fact that $\alpha_j(r)$ is fully mixing given $\lambda_i(l)(j)=G$, player i has the strict incentive to tell the truth about $(a_{i,t},y_{i,t})$ for each $t\in\mathbb{T}\left(r,k_i(r)\right)$ if $r<l_i^*$.

Second, we make sure that (64) does not affect player i's incentive in the coordination and main blocks. At the timing when player i takes $a_{i,t}$, the expected reward given that player i will tell the truth in the report block is

$$-T^{-11}1_{\left\{t=t_j^{\text{exclude}}(r)\right\}}1_{\left\{r<l_i^*\right\}}\mathbb{E}\left[\left\|\mathbf{1}_{a_{j,t},y_{j,t}}-\mathbb{E}\left[\mathbf{1}_{a_{j,t},y_{j,t}}\mid\alpha_j(r),a_{i,t},y_{i,t}\right]\right\|^2\mid\alpha_j(r),a_{i,t}\right],$$

given $t_j^{\text{exclude}}(r)$, l_i^*, and $\alpha_j(r)$. Note that the conditional expectation $\mathbb{E}\left[\mathbf{1}_{a_{j,t},y_{j,t}}\mid\alpha_j(r),a_{i,t},y_{i,t}\right]$ depends on $(a_{i,t},y_{i,t})$, and the distribution of $y_{i,t}$ depends on $(\alpha_j(r),a_{i,t})$. Hence, the conditional expectation given $(\alpha_j(r),a_{i,t})$ (but before observing $y_{i,t}$) depends on $(\alpha_j(r),a_{i,t})$. Since player j's signal $y_{j,t}$ statistically infers player i's action $a_{i,t}$, there exists $\pi_i^{\text{cancel}}[\alpha_j(r)](a_{j,t},y_{j,t})$ such that

$$\pi_i^{\text{cancel}}[\alpha_j(r)](a_{j,t},y_{j,t})=\mathbb{E}\left[\left\|\mathbf{1}_{a_{j,t},y_{j,t}}-\mathbb{E}\left[\mathbf{1}_{a_{j,t},y_{j,t}}\mid\alpha_j(r),a_{i,t},y_{i,t}\right]\right\|^2\mid\alpha_j(r),a_{i,t}\right].$$

To cancel out the effect of (64) on player i's incentive to take actions, player j gives the reward

$$T^{-11}1_{\left\{t=t_{j}^{\text{exclude}}(r)\right\}}1_{\left\{r<l_{i}^{*}\right\}}\pi_{i}^{\text{cancel}}[\alpha_{j}(r)](a_{j,t},y_{j,t}). \tag{65}$$

Third, player j punishes player i if player i's message about the frequency $f_i[h_i^{\mathbb{T}(r,k_i(r))}]$ about the picked subround is not compatible with player i's message about the history in this subround:

$$-T^{-15}1_{\left\{r<l_{i}^{*}\right\}}1_{\left\{f_{i}[\hat{h}_{i}^{\mathbb{T}(r,k_{i}(r))}]\neq|\mathbb{T}(r)|^{-\frac{2}{3}}\sum_{t\in\mathbb{T}(r,k_{i}(r))}\mathbf{1}_{\hat{a}_{i,t},\hat{y}_{i,t}}\right\}}. \tag{66}$$

The variable with hat denotes player i's message. Since $|\mathbb{T}(r)|^{\frac{2}{3}}$ is the length of the subround, $f_i[\hat{h}_i^{\mathbb{T}(r,k_i(r))}]$ and $|\mathbb{T}(r)|^{-\frac{2}{3}}\sum_{t\in\mathbb{T}(r,k_i(r))}\mathbf{1}_{\hat{a}_{i,t},\hat{y}_{i,t}}$ should be equal to each other if player i tells the truth.

In total, we have

$$\begin{aligned}
&\pi_{i}^{\text{report}}(x_{j},h_{j}^{\leq L},h_{j}^{\text{report}})\\
=\ &\sum_{r=1}^{R}1_{\left\{r<l_{i}^{*}\right\}}\left\{\begin{array}{c}
-T^{-11}\sum_{t\in\mathbb{T}(r,k_{i}(r))}1_{\left\{t=t_{j}^{\text{exclude}}(r)\right\}}\left\|\mathbf{1}_{a_{j,t},y_{j,t}}-\mathbb{E}\left[\mathbf{1}_{a_{j,t},y_{j,t}}\mid\alpha_{j}(r),\hat{a}_{i,t},\hat{y}_{i,t}\right]\right\|^{2}\\
+T^{-11}\sum_{t\in\mathbb{T}(r,k_{i}(r))}1_{\left\{t=t_{j}^{\text{exclude}}(r)\right\}}\pi_{i}^{\text{cancel}}[\alpha_{j}(r)](a_{j,t},y_{j,t})\\
-T^{-15}1_{\left\{f_{i}[\hat{h}_{i}^{\mathbb{T}(r,k_{i}(r))}]\neq|\mathbb{T}(r)|^{-\frac{2}{3}}\sum_{t\in\mathbb{T}(r,k_{i}(r))}\mathbf{1}_{\hat{a}_{i,t},\hat{y}_{i,t}}\right\}}
\end{array}\right\}.
\end{aligned}$$

Finally, given player i's message about the frequency of subrounds, $(f_i[\hat{h}_i^{\mathbb{T}(r,k(r))}])_{k(r)=1}^{|\mathbb{T}(r)|^{\frac{1}{3}}}$, player j calculates the frequency of the round:

$$f_{i}[\hat{h}_{i}^{\mathbb{T}(r)}]=|\mathbb{T}(r)|^{-\frac{1}{3}}\sum_{k(r)=1}^{|\mathbb{T}(r)|^{\frac{1}{3}}}f_{i}[\hat{h}_{i}^{\mathbb{T}(r,k(r))}]. \tag{67}$$

For each l, from $f_i[\hat{h}_i^{\mathbb{T}(1)}],...,f_i[\hat{h}_i^{\mathbb{T}(r)}]$ with round r being review round l, player j defines

$$\begin{aligned}
&\pi_{i}^{\text{adjust}}(x_{j},h_{j}^{\leq l},h_{j}^{\text{report}},l)\\
=\ &2\times1_{\{\lambda_{i}(l)(j)=G\}}\left\{\begin{array}{c}
\mathbb{E}\left[\pi_{i}^{\text{target}}(x_{j},f_{j}^{\text{include}}[h_{j}^{\leq l}],f_{j}[h_{j}^{l}],f_{i}[h_{i}^{\leq l}],l)\mid x_{j},\{\lambda_{i}(l)(j)=G\},f_{i}[\hat{h}_{i}^{\leq l}]\right]\\
-\mathbb{E}\left[\pi_{i}^{\text{review}}(x_{j},f_{j}^{\text{include}}[h_{j}^{\leq l}],f_{j}[h_{j}^{l}],l)\mid x_{j},\{\lambda_{i}(l)(j)=G\},f_{i}[\hat{h}_{i}^{\leq l}]\right]
\end{array}\right\},
\end{aligned}$$

so that (62) holds with truthtelling. Lemma 12 ensures that the expected difference between π_i^{target} and π_i^{review} depends only on the frequency of player i's history. Here, 2 cancels out the probability that player i is selected by the public randomization (we define that the adjustment is zero for player j who is not selected). Note that we have

$$\left|\pi_i^{\text{adjust}}(x_j, h_j^{\leq l}, h_j^{\text{report}}, l)\right| \leq \exp(-T^{\frac{1}{4}}) \tag{68}$$

for each x_j, $h_j^{\leq l}$, h_j^{report}, and l by Lemma 12.

15.1.3 Incentive Compatibility

Consider player i's incentive to tell the truth about the history in round r. If $1_{\{r<l_i^*\}} = 0$, that is, if round r is in review round l_i^* with $\lambda_i(l_i^*)(j) = B$ or after, then $\pi_i^{\text{report}}(x_j, h_j^{\leq L}, h_j^{\text{report}})$ does not depend on player i's message about round r. Moreover, the message about round r does not affect π_i^{adjust} either. To see why, note that, for each $l < l_i^*$,

$$\begin{aligned} & \pi_i^{\text{adjust}}(x_j, h_j^{\leq l}, h_j^{\text{report}}, l) \\ = \;\; & 2 \times 1_{\{\lambda_i(l)(j)=G\}} \left\{ \begin{array}{c} \mathbb{E}\left[\pi_i^{\text{target}}(x_j, f_j^{\text{include}}[h_j^{\leq l}], f_j[h_j^l], f_i[h_i^{\leq l}], l) \mid x_j, \{\lambda_i(l)(j) = G\}, f_i[\hat{h}_i^{\leq l}]\right] \\ -\mathbb{E}\left[\pi_i^{\text{review}}(x_j, f_j^{\text{include}}[h_j^{\leq l}], f_j[h_j^l], l) \mid x_j, \{\lambda_i(l)(j) = G\}, f_i[\hat{h}_i^{\leq l}]\right] \end{array} \right\} \end{aligned}$$

does not depend on player i's message about round r since round r is after review round l. For each $l \geq l_i^*$, we have $\pi_i^{\text{adjust}}(x_j, h_j^{\leq l}, h_j^{\text{report}}, l) = 0$ by definition since $\lambda_i(l)(j) = B$. Since the message does not affect π_i^{report} or π_i^{adjust}, any message is optimal about round r with $1_{\{r<l_i^*\}} = 0$. Therefore, we will concentrate on the case with $1_{\{r<l_i^*\}} = 1$ and $\lambda_i(l)(j) = G$ for each review round l which is equal to or before round r.

When player i sends $(a_{i,t}, y_{i,t})_{t\in\mathbb{T}(r,k_i(r))}$, player i believes that the expected reward from (64) given $\alpha_j(r)$ is

$$\begin{aligned} & -T^{-11}\Pr\left(\left\{t = t_j^{\text{exclude}}(r)\right\} \mid \alpha_j(r), h_i^{\leq L}\right) \\ & \times\mathbb{E}\left[\left\|\mathbf{1}_{a_{j,t},y_{j,t}} - \mathbb{E}\left[\mathbf{1}_{a_{j,t},y_{j,t}} \mid \alpha_j(r), \hat{a}_{i,t}, \hat{y}_{i,t}\right]\right\|^2 \mid \alpha_j(r), h_i^{\leq L}, \left\{t = t_j^{\text{exclude}}(r)\right\}\right]. \end{aligned} \tag{69}$$

First, $\Pr\left(\left\{t = t_j^{\text{exclude}}(r)\right\} \mid \alpha_j(r), h_i^{\leq L}\right) \geq T^{-2}$ for each $t \in \mathbb{T}(r)$ by Lemma 10. Second, since player j does not use information in period $t_j^{\text{exclude}}(r)$ to determine the continuation play, player i believes that given $t_j^{\text{exclude}}(r) = t$, player j's history $(a_{j,t}, y_{j,t})$ is distributed according to $\Pr(a_{j,t}, y_{j,t} \mid \alpha_j(r), a_{i,t}, y_{i,t})$.

Hence, (69) is equal to

$$-\varepsilon\mathbb{E}\left[\left\|\mathbf{1}_{a_{j,t},y_{j,t}} - \mathbb{E}\left[\mathbf{1}_{a_{j,t},y_{j,t}} \mid \alpha_j(r), \hat{a}_{i,t}, \hat{y}_{i,t}\right]\right\|^2 \mid \alpha_j(r), a_{i,t}, y_{i,t}\right] \tag{70}$$

for some $\varepsilon > 0$ of order T^{-13} for each $t \in \mathbb{T}(r, k_i(r))$.

Finally, since we are considering the case with $1_{\left\{r<l_i^*\right\}}$, player j's strategy $\alpha_j(r)$ has full support. Then, by Assumption 4, we can make sure that the expected loss in (70) from telling a lie about $(a_{i,t}, y_{i,t})$ is of order T^{-13}. (See Lemma 46 for the details.) Since T^{-13} is greater than the magnitude of the other reward in π_i^{report} (since $\pi_i^{\text{cancel}}[\alpha_j(r)](a_{j,t}, y_{j,t})$ is sunk in the report block, the other relevant reward in π_i^{report} is the one defined in (66)) and the adjustment $\pi_i^{\text{adjust}}(x_j, h_j^{\leq l}, h_j^{\text{report}}, l)$ is of order $\exp(-T^{\frac{1}{4}})$ by (68), it is strictly optimal to tell the truth about $(a_{i,t}, y_{i,t})_{t\in\mathbb{T}(r,k_i(r))}$ regardless of the past messages.

Given this truthtelling incentive about $(a_{i,t}, y_{i,t})_{t\in\mathbb{T}(r,k_i(r))}$, consider player i's incentive to tell the truth about the frequency in subrounds. While she sends the message $f_i[h_i^{\mathbb{T}(r,k(r))}]$, she believes that public randomization will pick any $k(r)$ with probability $|\mathbb{T}(r)|^{-\frac{1}{3}} \geq T^{-\frac{1}{3}}$. (Recall that the public randomization $k_i(r)$ is drawn after she finishes sending $(f_i[h_i^{\mathbb{T}(r,k(r))}])_{k(r)=1}^{|\mathbb{T}(r)|^{\frac{1}{3}}}$.) Hence, if player i tells a lie about subround $k(r)$, then the expected loss from (66) is $T^{-\frac{1}{3}} \times T^{-15}$, which is greater than the magnitude of the adjustment $\pi_i^{\text{adjust}}(x_j, h_j^{\leq l}, h_j^{\text{report}}, l)$. Hence, it is strictly optimal to tell the truth about $(f_i[h_i^{\mathbb{T}(r,k(r))}])_{k(r)=1}^{|\mathbb{T}(r)|^{\frac{1}{3}}}$ regardless of the past messages.

In summary, we construct an incentive compatible strategy and reward such that player i tells the truth about the history and from that report, player j calculates π_i^{adjust} to satisfy (62).

15.2 Dispensing with Cheap Talk

We now explain how player i sends the message by taking actions. Again, the players draw a public randomization device to decide who to report the history. Suppose that player i is selected. Player j takes α_j^{mix} *i.i.d.* across periods, and $\pi_j^{\text{adjust}}(x_i, h_i^{\leq l}, h_i^{\text{report}}, l) = 0$. Assumption 2 ensures that there exists player j's reward function π_j^{report} to incentivize her to take α_j^{mix} and to keep her equilibrium value in the report block equal to zero. Hence, we focus on player i's strategy and rewards.

15.2.1 Player i's Strategy $\sigma_i^{\text{report}}|_{h_i^{\leq L}}$

As in the case with cheap talk, player i sends $(f_i[h_i^{\mathbb{T}(r,k(r))}])_{k(r)=1}^{|\mathbb{T}(r)|^{\frac{1}{3}}}$ and $(a_{i,t}, y_{i,t})_{t\in\mathbb{T}(r,k_i(r))}$. Let m_i be a generic message that player i wants to send in the report block, that is, m_i can be $f_i[h_i^{\mathbb{T}(r,k(r))}]$ for some r and $k(r)$, or $(a_{i,t}, y_{i,t})$ for some r and $t \in \mathbb{T}(r, k_i(r))$; and let M_i be the set of possible messages. We have $|M_i| \leq |\mathbb{T}(r)|^{\frac{2}{3}|A_i||Y_i|} \leq T^{\frac{2}{3}|A_i||Y_i|}$ for $f_i[h_i^{\mathbb{T}(r,k(r))}]$ since the frequency of subround $f_i[h_i^{\mathbb{T}(r,k(r))}]$ can be expressed by "how many times out of $|\mathbb{T}(r)|^{\frac{2}{3}}$ periods player i observes each (a_i, y_i)." (Recall that $|\mathbb{T}(r)|^{\frac{2}{3}}$ is the length of the subround.) For $(a_{i,t}, y_{i,t})$, we have $|M_i| \leq |A_i| |Y_i|$ since (a_i, y_i) is included in $A_i \times Y_i$.

We now explain how player i sends $m_i \in M_i$. Given $a_i(G)$ and $a_i(B)$ fixed in Section 6.2, there exists a one-to-one mapping $\vec{a}_i : M_i \to \{a_i(G), a_i(B)\}^{\log_2|M_i|}$ between message m_i and sequence of binary actions since $\left|\{a_i(G), a_i(B)\}^{\log_2|M_i|}\right| = |M_i|$. (See Appendix A.4 for how to define $\vec{a}_i$ explicitly.) Note that the length of $\vec{a}_i(m_i)$ is bounded by $\log_2 T^{\frac{2}{3}|A_i||Y_i|}$ for $f_i[h_i^{\mathbb{T}(r,k(r))}]$ and by $\log_2|A_i||Y_i|$ for $(a_{i,t}, y_{i,t})$.

When player i sends the message m_i, player i takes an action sequence $\vec{a}_i(m_i)$. Moreover, player i repeats each element of $\vec{a}_i(m_i)$ for multiple periods, in order to increase the precision of the message. In particular, for m_i corresponding to $f_i[h_i^{\mathbb{T}(r,k(r))}]$, player i repeats the action for $T^{\frac{1}{2}}$ periods, and for m_i corresponding to $(a_{i,t}, y_{i,t})$, she repeats it for $T^{\frac{1}{4}}$ periods. Let $T(m_i)$ be the number of repetitions: $T(f_i[h_i^{\mathbb{T}(r,k(r))}]) = T^{\frac{1}{2}}$ and $T(a_{i,t}, y_{i,t}) = T^{\frac{1}{4}}$.

Given this strategy, the communication takes periods of order $T^{\frac{11}{12}}$:

$$\sum_{r=1}^{R}\left(\begin{array}{c} \underbrace{T^{\frac{1}{2}}}_{\text{repetition}} \times \underbrace{|\mathbb{T}(r)|^{\frac{1}{3}}}_{\text{number of subrounds}} \times \underbrace{\log_2|\mathbb{T}(r)|^{\frac{2}{3}|A_i||Y_i|}}_{\text{length of } \vec{a}_i(m_i) \text{ for } f_i[h_i^{\mathbb{T}(r,k(r))}]} \\ + \underbrace{T^{\frac{1}{4}}}_{\text{repetition}} \times \underbrace{|\mathbb{T}(r)|^{\frac{2}{3}}}_{\text{number of periods per subround}} \times \underbrace{\log_2|A_i||Y_i|}_{\text{length of } \vec{a}_i(m_i) \text{ for } (a_i,y_i)} \end{array}\right), \tag{71}$$

which is of order $T^{\frac{11}{12}} \ll T$, as seen in (20).

Let $t_{\text{report}} + 1$ be the first period of the report block; and let $t_{\text{report}} + 1, ..., t_{\text{report}} + T^{\frac{11}{12}}$ be the periods in which player i takes $\vec{a}_i(m_i)$ for some m_i. (Precisely speaking, $T^{\frac{11}{12}}$ should be of order $T^{\frac{11}{12}}$ by (71). For simple notation, we just write $T^{\frac{11}{12}}$ rather than of order $T^{\frac{11}{12}}$ in the text.) After

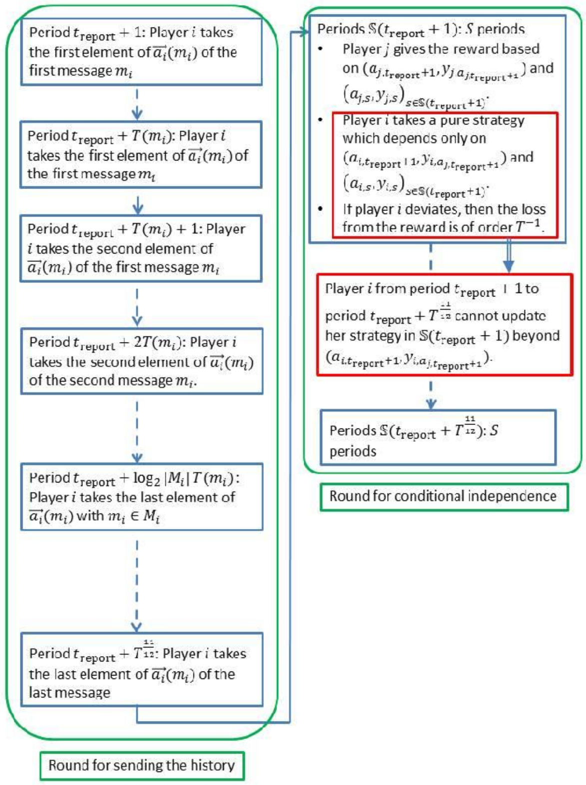

Figure 14: Structure of the report block with public randomization

player i finishes sending each $\vec{a}_i(m_i)$, the players sequentially assign S periods for each of periods $t_{\text{report}}+1,...,t_{\text{report}}+T^{\frac{11}{12}}$, where S is determined in Section 6.3. That is, periods $t_{\text{report}}+T^{\frac{11}{12}}+1,...,t_{\text{report}}+T^{\frac{11}{12}}+S$ are assigned to period $t_{\text{report}}+1$, periods $t_{\text{report}}+T^{\frac{11}{12}}+S+1,...,t_{\text{report}}+T^{\frac{11}{12}}+2S$ are assigned to period $t_{\text{report}}+2$, and so on. In general, let $\mathbb{S}(t)$ be the set of periods $t_{\text{report}}+T^{\frac{11}{12}}+(k-1)S+1,...,t_{\text{report}}+T^{\frac{11}{12}}+kS$ that are assigned to period t with $k=t-t_{\text{report}}$. In periods $\mathbb{S}(t)$, given player i's history in period t, player i takes $\sigma_i^{\mathbb{S}(t)}$ determined in Section 6.3.

The periods $t_{\text{report}}+1,...,t_{\text{report}}+T^{\frac{11}{12}}$ are called "the round for sending the history" and the other periods are called "the round for conditional independence." Figure 14 summarizes the entire structure.

15.2.2 Player j's Inference of Player i's Message

For each period t in which player i sends an element of an action sequence $\vec{a}_i(m_i)$ assigned to a message m_i, player j given her history in periods t and $\mathbb{S}(t)$ calculates a function $\phi_j((a_{j,t},y_{j,t})$ $\cup\,(a_{j,\tau},y_{j,\tau})_{\tau\in\mathbb{S}(t)})$ determined in Section 6.3. By Lemma 4, since the expected realization of ϕ_j is high (or low) if player i takes $a_i(G)$ (or $a_i(B)$), player j infers that the element of $\vec{a}_i(m_i)$ is $a_i(G)$

(or $a_i(B)$) if there are a lot of high (or low) realization of ϕ_j. Since player i repeats the element of $\vec{a}_i(m_i)$ for $T(m_i)$ periods, by the law of large numbers, player j can infer the element correctly with probability of order

$$1-\exp(-T(m_i)). \tag{72}$$

Importantly, Lemma 4 ensures that (if player i expects that she will take $\sigma_i^{\mathbb{S}(t)}$ in the round for conditional independence) player i in the round for sending the history cannot update player j's inference from player i's signal observation (*conditional independence property*).

Given this inference of each element of $\vec{a}_i(m_i)$, player j infers each message using the inverse of $\vec{a}_i(m_i)$. Let $(f_i[h_i^{\mathbb{T}(r,k(r))}](j))_{k(r)=1}^{|\mathbb{T}(r)|^{\frac{1}{3}}}$ and $(a_{i,t}(j), y_{i,t}(j))_{t\in\mathbb{T}(r,k_i(r))}$ be player j's inference. As in (67), player j infers $f_i[h_i^{\mathbb{T}(r)}](j)$ and $f_i[h_i^{\leq l}](j)$ from $((f_i[h_i^{\mathbb{T}(r,k(r))}](j))_{k(r)=1}^{|\mathbb{T}(r)|^{\frac{1}{3}}})_{r=1}^{R}$.

15.2.3 Reward Functions π_i^{adjust} and π_i^{report}

We modify π_i^{adjust} and π_i^{report} in order to deal with the possibility of errors. First, we consider an error when player j calculates

$$\begin{aligned}
& \pi_i^{\text{adjust}}(x_j, h_j^{\leq l}, h_j^{\text{report}}, l) \\
= \;\; & 2\times 1_{\{\lambda_i(l)(j)=G\}}\left\{\begin{array}{c} \mathbb{E}\left[\pi_i^{\text{target}}(x_j, f_j^{\text{include}}[h_j^{\leq l}], f_j[h_j^l], f_i[h_i^{\leq l}], l) \mid x_j, \{\lambda_i(l)(j)=G\}, f_i[h_i^{\leq l}](j)\right] \\ -\mathbb{E}\left[\pi_i^{\text{review}}(x_j, f_j^{\text{include}}[h_j^{\leq l}], f_j[h_j^l], l) \mid x_j, \{\lambda_i(l)(j)=G\}, f_i[h_i^{\leq l}](j)\right] \end{array}\right\}
\end{aligned} \tag{73}$$

from $f_i[h_i^{\leq l}](j)$. Since the cardinality of $f_i[h_i^{\mathbb{T}(r,k(r))}]$ is $|\mathbb{T}(r)|^{\frac{2}{3}|A_i||Y_i|}$, the length of $\vec{a}_i(m_i)$ to send $f_i[h_i^{\mathbb{T}(r,k(r))}]$ is $\log_2|\mathbb{T}(r)|^{\frac{2}{3}|A_i||Y_i|}\leq \log_2 T^{\frac{2}{3}|A_i||Y_i|}$. Since each round has $|\mathbb{T}(r)|^{\frac{1}{3}}\leq T^{\frac{1}{3}}$ subrounds and there are R rounds, the total length of $\vec{a}_i(m_i)$'s that are used to calculate $f_i[h_i^{\leq l}](j)$ is no more than $RT^{\frac{1}{3}}\log_2|\mathbb{T}(r)|^{\frac{2}{3}|A_i||Y_i|}$. By (72), recalling that $T(m_i)=T^{\frac{1}{2}}$ for m_i corresponding to $f_i[h_i^{\mathbb{T}(r,k(r))}]$, player j infers $f_i[h_i^{\leq l}](j)$ correctly with probability no less than

$$1-RT^{\frac{1}{3}}\log_2|\mathbb{T}(r)|^{\frac{2}{3}|A_i||Y_i|}\times\exp(-T^{\frac{1}{2}}). \tag{74}$$

On the other hand, the cardinality of the messages $((f_i[h_i^{\mathbb{T}(r,k(r))}](j))_{k(r)=1}^{|\mathbb{T}(r)|^{\frac{1}{3}}})_{r=1}^{R}$ used to calculate

$f_i[h_i^{\leq l}](j)$ is no more than

$$\prod_{r=1}^{R} \left(\text{cardinality of } f_i[h_i^{\mathbb{T}(r,k(r))}]\right)^{\text{number of subrounds in round } r} \leq \left(|\mathbb{T}\,(r)|^{\frac{2}{3}|A_i||Y_i|}\right)^{RT^{\frac{1}{3}}},$$

which is of order $\exp(T^{\frac{1}{3}})$.

As will be seen in Appendix A.11.3.2, if we have shown that the probability of the correct inference (74), to the power of the cardinality of the messages used in (73), converges to one, then we can create $\pi_i^{\text{adjust}}(x_j, h_j^{\leq l}, h_j^{\text{report}}, l)$ such that, given that player i follows the equilibrium strategy in the report block, (i) this adjustment is still small and (ii) (62) holds from the perspective of player i in review round l. Since we have $(|\mathbb{T}\,(r)|^{\frac{2}{3}|A_i||Y_i|})^{RT^{\frac{1}{3}}} \approx \exp(T^{\frac{1}{3}}) \ll \exp(T^{\frac{1}{2}})$ for a sufficiently large T, the probability of the correct inference, to the power of the cardinality of the messages, satisfies

$$\begin{aligned} & \left(1 - RT^{\frac{1}{3}} \log_2 |\mathbb{T}\,(r)|^{\frac{2}{3}|A_i||Y_i|} \times \exp(-T^{\frac{1}{2}})\right)^{\left(|\mathbb{T}(r)|^{\frac{2}{3}|A_i||Y_i|}\right)^{RT^{\frac{1}{3}}}} \\ \geq \ & 1 - RT^{\frac{1}{3}} \log_2 |\mathbb{T}\,(r)|^{\frac{2}{3}|A_i||Y_i|} \times \exp(-T^{\frac{1}{2}}) \times \left(|\mathbb{T}\,(r)|^{\frac{2}{3}|A_i||Y_i|}\right)^{RT^{\frac{1}{3}}} \to 1 \text{ as } T \to \infty, \end{aligned} \tag{75}$$

as desired.

We also modify $\pi_i^{\text{report}}(x_j, h_j^{\leq L}, h_j^{\text{report}})$ taking errors into account. Consider the reward to incentivize player i to tell the truth about $(a_{i,t}, y_{i,t})_{t \in \mathbb{T}(r,k_i(r))}$:

$$- 1_{\left\{t = t_j^{\text{exclude}}(r)\right\}} 1_{\left\{r < l_i^*\right\}} \left\| \mathbf{1}_{a_{j,t},y_{j,t}} - \mathbb{E}\left[\mathbf{1}_{a_{j,t},y_{j,t}} \mid \alpha_j(r), \hat{a}_{i,t}, \hat{y}_{i,t}\right] \right\|^2. \tag{76}$$

Without cheap talk, player j can infer each element of $\vec{a}_i(m_i)$ to send $(a_{i,t}, y_{i,t})$ correctly with probability of order $1 - \exp(-T^{\frac{1}{4}})$ since $T(m_i) = T^{\frac{1}{4}}$ for m_i corresponding to (a_i, y_i). Since the number of elements of $\vec{a}_i(m_i)$ to send $(a_{i,t}, y_{i,t})$ is $\log_2 |A_i|\,|Y_i|$, the probability that player j infers $(a_{i,t}, y_{i,t})$ correctly is

$$1 - \log_2 |A_i|\,|Y_i| \times \exp(-T^{\frac{1}{4}}).$$

On the other hand, the cardinality of the message $(a_{i,t}, y_{i,t})$ in (76) is $|A_i|\,|Y_i|$.

Hence, the probability of the correct inference, to the power of the cardinality of the messages in (76), converges to one as T goes to infinity:

$$\left(1-\log_2|A_i|\,|Y_i|\times\exp(-T^{\frac{1}{4}})\right)^{|A_i||Y_i|}\geq 1-|A_i|\,|Y_i|\log_2|A_i|\,|Y_i|\times\exp(-T^{\frac{1}{4}})\to_{T\to\infty}1.$$

As will be seen in Lemma 48 formally, this means that we can slightly modify the reward taking errors into account, so that it is strictly optimal for player i to tell the truth about $(a_{i,t},y_{i,t})$.

Similarly, we can modify $\pi_i^{\text{cancel}}[\alpha_j(r)](a_{j,t},y_{j,t})$ so that π_i^{cancel} cancels out the effect of $a_{i,t}$ in period t of the coordination or main block on the modified version of (76).

Next, consider the reward to incentivize player i to tell the truth about $f_i[h_i^{\mathbb{T}(r,k(r))}]$:

$$-\mathbf{1}_{\left\{f_i[h_i^{\mathbb{T}(r,k_i(r))}](j)\neq|\mathbb{T}(r)|^{-\frac{2}{3}}\sum_{t\in\mathbb{T}(r,k_i(r))}\mathbf{1}_{a_{i,t}(j),y_{i,t}(j)}\right\}}. \tag{77}$$

Without cheap talk, since the cardinality of $f_i[h_i^{\mathbb{T}(r,k_i(r))}]$ is $|\mathbb{T}(r)|^{\frac{2}{3}|A_i||Y_i|}$, the length of $\vec{a}_i(m_i)$ to send $f_i[h_i^{\mathbb{T}(r,k_i(r))}]$ is $\log_2|\mathbb{T}(r)|^{\frac{2}{3}|A_i||Y_i|}$. In addition, since the cardinality of $(a_{i,t},y_{i,t})$ is $|A_i|\,|Y_i|$, the length of $\vec{a}_i(m_i)$ to send $(a_{i,t},y_{i,t})$ is $\log_2|A_i|\,|Y_i|$. Since each subround has $|\mathbb{T}(r,k_i(r))|=|\mathbb{T}(r)|^{\frac{2}{3}}\leq T^{\frac{2}{3}}$ periods, the total length of $\vec{a}_i(m_i)$'s that are used to calculate $f_i[h_i^{\mathbb{T}(r,k_i(r))}](j)$ and $|\mathbb{T}(r)|^{-\frac{2}{3}}\sum_{t\in\mathbb{T}(r,k_i(r))}\mathbf{1}_{a_{i,t}(j),y_{i,t}(j)}$ is

$$\log_2|\mathbb{T}(r)|^{\frac{2}{3}|A_i||Y_i|}+|\mathbb{T}(r)|^{\frac{2}{3}}\log_2|A_i|\,|Y_i|\,.$$

Hence, all the messages transmit correctly with probability

$$1-(\log_2|\mathbb{T}(r)|^{\frac{2}{3}|A_i||Y_i|}\exp(-\underbrace{T^{\frac{1}{2}}}_{\#\text{ of repetitions for }f_i[h_i^{\mathbb{T}(r,k(r))}]})+|\mathbb{T}(r)|^{\frac{2}{3}}\log_2|A_i|\,|Y_i|\exp(-\underbrace{T^{\frac{1}{4}}}_{\#\text{ of repetitions for }(a_{i,t},y_{i,t})})). \tag{78}$$

On the other hand, the cardinality of the messages used in (77) is calculated as follows: The cardinality of $f_i[h_i^{\mathbb{T}(r,k_i(r))}](j)$ is $|\mathbb{T}(r)|^{\frac{2}{3}|A_i||Y_i|}$. As for $|\mathbb{T}(r)|^{-\frac{2}{3}}\sum_{t\in\mathbb{T}(r,k_i(r))}\mathbf{1}_{a_{i,t}(j),y_{i,t}(j)}$, when player i sends the message as if $(\tilde{a}_{i,t},\tilde{y}_{i,t})_{t\in\mathbb{T}(r,k_i(r))}$ were the true message, the distribution of this summation $\sum_{t\in\mathbb{T}(r,k_i(r))}\mathbf{1}_{a_{i,t}(j),y_{i,t}(j)}$ depends only on the frequency of each (a_i,y_i) in $(\tilde{a}_{i,t},\tilde{y}_{i,t})_{t\in\mathbb{T}(r,k_i(r))}$. Hence, the relevant cardinality of $(a_{i,t}(j),y_{i,t}(j))_{t\in\mathbb{T}(r,k_i(r))}$ for (77) is equal to that of the frequency, that is, $|\mathbb{T}(r)|^{\frac{2}{3}|A_i||Y_i|}$.

Hence, (78) to the power of the relevant cardinality is

$$\left(1-\left(\log_2 |\mathbb{T}(r)|^{\frac{2}{3}|A_i||Y_i|}\exp(-T^{\frac{1}{2}})+|\mathbb{T}(r)|^{\frac{2}{3}}\log_2|A_i|\,|Y_i|\exp(-T^{\frac{1}{4}})\right)\right)^{|\mathbb{T}(r)|^{\frac{2}{3}|A_i||Y_i|}\times|\mathbb{T}(r)|^{\frac{2}{3}|A_i||Y_i|}} \to_{T\to\infty} 1.$$

Therefore, we can slightly modify the reward function so that player i wants to tell the truth about $f_i[h_i^{\mathbb{T}(r,k(r))}]$.

In addition, we add $\pi_i^{\text{c.i.}}$ defined in Section 6.3 in order to incentivize player i to take $\sigma_i^{\mathbb{S}(t)}$ in the round for conditional independence.

Finally, we add the reward $\pi_i(x_j, y_j)$ defined in (9), so that it cancels out the effect of the instantaneous utilities.

In total, we have

$$\begin{aligned}
&\pi_i^{\text{report}}(x_j, h_j^{\leq L}, h_j^{\text{report}}) \\
= &\sum_{t:\text{report block}} \pi_i(x_{j,t}, y_{j,t}) \\
&+\sum_{r=1}^{R} 1_{\{r<l_i^*\}}\left\{\begin{array}{c}
-T^{-11}\sum_{t\in\mathbb{T}(r,k_i(r))} 1_{\{t=t_j^{\text{exclude}}(r)\}}\left\{\begin{array}{c}\text{modification of}\\ \left\|\mathbf{1}_{a_{j,t},y_{j,t}}-\mathbb{E}\left[\mathbf{1}_{a_{j,t},y_{j,t}}\mid \alpha_j(r), a_{i,t}(j), y_{i,t}(j)\right]\right\|^2\end{array}\right\} \\
+T^{-11}1_{\{t=t_j^{\text{exclude}}(r)\}}\left\{\text{modification of } \pi_i^{\text{cancel}}[\alpha_j(r)](a_{j,t}, y_{j,t})\right\} \\
-T^{-15}\left\{\text{modification of } 1_{\left\{f_i[h_i^{\mathbb{T}(r,k_i(r))}](j)\neq|\mathbb{T}(r)|^{-\frac{2}{3}}\sum_{t\in\mathbb{T}(r,k_i(r))}\mathbf{1}_{a_{i,t}(j),y_{i,t}(j)}\right\}}\right\}
\end{array}\right\} \\
&+\sum_{t:\text{round for sending the history}} T^{-1}\pi_i^{\text{c.i.}}\left((a_{j,t}, y_{j,t})\cup(a_{j,\tau}, y_{j,\tau})_{\tau\in\mathbb{S}(t)}\right).
\end{aligned}$$

15.2.4 Incentive Compatibility

Let us verify player i's incentive. By $\pi_i(x_{j,t}, y_{j,t})$, (9) ensures that we can neglect the instantaneous utility. In the round for conditional independence, Claim 1 of Lemma 4 implies that, once player i deviates, then she incurs a loss of $T^{-1}\varepsilon_{\text{strict}}$ from $T^{-1}\pi_i^{\text{c.i.}}$. Since the modification is small, the other terms in π_i^{report} are of order T^{-11}. Hence, it is optimal for player i to take $\sigma_i^{\mathbb{S}(t)}$ in the round for conditional independence.

Given this incentive, Claim 2 of Lemma 4 ensures that player i in the round for sending the history cannot update player j's inference of player i's message. Moreover, (9) and Claim 3 of Lemma 4 ensures that player i is indifferent between any action in the round for sending the history in terms of the expected value of

$$u_i(a_t) + \pi_i(a_{j,t}, y_{j,t}) + \sum_{\tau \in \mathbb{S}(t)} (u_i(a_\tau) + \pi_i(a_{j,\tau}, y_{j,\tau})) + \pi_i^{\text{c.i.}} \left((a_{j,t}, y_{j,t}) \cup (a_{j,\tau}, y_{j,\tau})_{\tau \in \mathbb{S}(t)} \right).$$

Since the incentive with cheap talk is strict, the message transmit correctly with a high probability, and the modification is small, it is incentive compatible for player i to follow the equilibrium strategy.

In summary, we construct an incentive compatible strategy and reward without cheap talk. The key idea is to use the round for conditional independence to keep conditional independence property in the round for sending the history.

15.3 Dispensing with Public Randomization

Now we explain how to dispense with the public randomization. Recall that the public randomization plays two roles in the report block. The first is to pick who reports the history. The second is to pick a subround $k(r)$.

Let us focus on the first one. (The second one is dispensed with by a similar procedure. See Appendix A.11 for the details.) The purpose of this first role is to establish the following two properties: (i) ex ante (before the report block), every player has a positive probability to report the history, and (ii) ex post (after the realization of the public randomization), there is only one player who reports the history.

The property (i) is important since, without adjusting the reward function by π_i^{adjust} based on player i's report in the report block, player i's equilibrium strategy would not be optimal. The property (ii) is important to incentivize player i to tell the truth. Remember that player j incentivizes player i to tell the truth by punishing player i according to (64). As seen in (70), the establishment of truthtelling incentive uses the fact that player i cannot update her belief about the realization of player j's history in period $t_j^{\text{exclude}}(r)$. If both players sent messages by taking actions and player i could observe a part of player j's messages before finishing reporting her own history, then player i might be able to learn player j's history in period $t_j^{\text{exclude}}(r)$ and would want to tell a lie.

In order to establish these two properties without public randomization, we consider the following procedure. For simplicity, let us assume that the signals are not conditionally independent (see the end of this subsection for the case with conditionally independent signals): There exists $a^{\text{p.r.}} \in A$ such that, given $a^{\text{p.r.}}$, there exist $y_1^{\text{p.r.}}, \bar{y}_1^{\text{p.r.}} \in Y_1$ and $y_2^{\text{p.r.}} \in Y_2$ such that

$$\Pr\left(y_2^{\text{p.r.}} \mid a^{\text{p.r.}}, y_1^{\text{p.r.}}\right) \neq \Pr\left(y_2^{\text{p.r.}} \mid a^{\text{p.r.}}, \bar{y}_1^{\text{p.r.}}\right).$$

Fix those $a^{\text{p.r.}}$, $y_1^{\text{p.r.}}, \bar{y}_1^{\text{p.r.}} \in Y_1$, and $y_2^{\text{p.r.}} \in Y_2$. Without loss, we assume that

$$\Pr\left(y_2^{\text{p.r.}} \mid a^{\text{p.r.}}, y_1^{\text{p.r.}}\right) > \Pr\left(y_2^{\text{p.r.}} \mid a^{\text{p.r.}}, \bar{y}_1^{\text{p.r.}}\right). \tag{79}$$

At the beginning of the report block, the players take a pure strategy $a^{\text{p.r.}}$.[21] Let $t_{\text{p.r.}}$ be the period when the players take $a^{\text{p.r.}}$. Each player i observes her own signal $y_{i,t_{\text{p.r.}}}$. By Assumption 2, since player j can statistically identify if player i takes $a_i^{\text{p.r.}}$, there exists $\pi_i^{\text{p.r.}} : A_j \times Y_j \to \mathbb{R}$ such that it is optimal for player i to take $a_i^{\text{p.r.}}$:

$$\mathbb{E}\left[u_i(a_i, a_j) + \pi_i^{\text{p.r.}}(a_j, y_j) \mid a_i, a_j^{\text{p.r.}}\right] = \begin{cases} 0 & \text{if } a_i = a_i^{\text{p.r.}}, \\ -1 & \text{if } a_i \neq a_i^{\text{p.r.}}. \end{cases} \tag{80}$$

Intuitively, player 2 asks player 1 to guess whether player 2 observed $y_{2,t_{\text{p.r.}}} = y_2^{\text{p.r.}}$ or $y_{2,t_{\text{p.r.}}} \neq y_2^{\text{p.r.}}$. On the other hand, since players' signals are conditionally dependent, player 1's conditional likelihood of $y_{2,t_{\text{p.r.}}} = y_2^{\text{p.r.}}$ against $y_{2,t_{\text{p.r.}}} \neq y_2^{\text{p.r.}}$ depends on $y_{1,t_{\text{p.r.}}} \in Y_1$. In particular, (79) ensures that there exists $\bar{p}_1^{\text{p.r.}}$ such that

$$\frac{\Pr\left(\left\{y_{2,t_{\text{p.r.}}} = y_2^{\text{p.r.}}\right\} \mid a^{\text{p.r.}}, y_1^{\text{p.r.}}\right)}{\Pr\left(\left\{y_{2,t_{\text{p.r.}}} \neq y_2^{\text{p.r.}}\right\} \mid a^{\text{p.r.}}, y_1^{\text{p.r.}}\right)} > \bar{p}_1^{\text{p.r.}} > \frac{\Pr\left(\left\{y_{2,t_{\text{p.r.}}} = y_2^{\text{p.r.}}\right\} \mid a^{\text{p.r.}}, \bar{y}_1^{\text{p.r.}}\right)}{\Pr\left(\left\{y_{2,t_{\text{p.r.}}} \neq y_2^{\text{p.r.}}\right\} \mid a^{\text{p.r.}}, \bar{y}_1^{\text{p.r.}}\right)}.$$

Perturbing $\bar{p}_1^{\text{p.r.}}$ if necessary, we make sure that there is no $y_1 \in Y_1$ such that

$$\frac{\Pr\left(\left\{y_{2,t_{\text{p.r.}}} = y_2^{\text{p.r.}}\right\} \mid a^{\text{p.r.}}, y_1\right)}{\Pr\left(\left\{y_{2,t_{\text{p.r.}}} \neq y_2^{\text{p.r.}}\right\} \mid a^{\text{p.r.}}, y_1\right)} = \bar{p}_1^{\text{p.r.}}.$$

[21] The superscript p.r. stands for "public randomization."

Given this definition, we partition the set of player 1's signals into Y_1^{report} and $Y_1^{\text{not-report}}$:

$$
\begin{aligned}
Y_1^{\text{report}} &\equiv \left\{ y_1 \in Y_1 : \frac{\Pr\left(\left\{y_{2,t_{\text{p.r.}}} = y_2^{\text{p.r.}}\right\} \mid a^{\text{p.r.}}, y_1\right)}{\Pr\left(\left\{y_{2,t_{\text{p.r.}}} \neq y_2^{\text{p.r.}}\right\} \mid a^{\text{p.r.}}, y_1\right)} > \bar{p}_1^{\text{p.r.}} \right\} \ni y_1^{\text{p.r.}}, \\
Y_1^{\text{not-report}} &\equiv \left\{ y_1 \in Y_1 : \frac{\Pr\left(\left\{y_{2,t_{\text{p.r.}}} = y_2^{\text{p.r.}}\right\} \mid a^{\text{p.r.}}, y_1\right)}{\Pr\left(\left\{y_{2,t_{\text{p.r.}}} \neq y_2^{\text{p.r.}}\right\} \mid a^{\text{p.r.}}, y_1\right)} < \bar{p}_1^{\text{p.r.}} \right\} \ni \bar{y}_1^{\text{p.r.}}
\end{aligned}
$$

with $Y_1^{\text{report}} \cup Y_1^{\text{not-report}} = Y_1$. Player 1 is asked to report whether "$y_{1,t_{\text{p.r.}}} \in Y_1^{\text{report}}$" or "$y_{1,t_{\text{p.r.}}} \in Y_1^{\text{not-report}}$." (In this explanation, we assume that cheap talk is available. Cheap talk is dispensable as in Section 15.2.) For notational convenience, let $\iota^{\text{p.r.}} = 1$ denote the event that $y_{1,t_{\text{p.r.}}} \in Y_1^{\text{report}}$; and $\iota^{\text{p.r.}} = 2$ denote the event that $y_{1,t_{\text{p.r.}}} \in Y_1^{\text{non-report}}$. Let $\hat{\iota}^{\text{p.r.}} \in \{1, 2\}$ be player 1's message about $\iota^{\text{p.r.}}$.

To incentivize player 1 to tell the truth, player 2 rewards player 1 if either "player 2 observes $y_2^{\text{p.r.}}$ and player 1 reports $\hat{\iota}^{\text{p.r.}} = 1$" or "player 2 does not observe $y_2^{\text{p.r.}}$ and player 1 reports $\hat{\iota}^{\text{p.r.}} = 2$":

$$
T^{-4}\left(1_{\left\{y_{2,t_{\text{p.r.}}} = y_2^{\text{p.r.}} \wedge \hat{\iota}^{\text{p.r.}} = 1\right\}} + \bar{p}_1^{\text{p.r.}} 1_{\left\{y_{2,t_{\text{p.r.}}} \neq y_2^{\text{p.r.}} \wedge \hat{\iota}^{\text{p.r.}} = 2\right\}}\right). \tag{81}
$$

After player 1 sends $\hat{\iota}^{\text{p.r.}}$, the players send the message about $h_i^{\leq L}$ sequentially: Player 1 sends the messages first, and then player 2 sends the messages. The players coordinate on how to send the messages based on $\hat{\iota}^{\text{p.r.}}$ as follows.

If player 1 reports $\hat{\iota}^{\text{p.r.}} = 1$, then player 2 adjusts player 1's reward function based on player 1's messages, and incentivizes player 1 to tell the truth according to (64), (65), and (66) with $i = 1$. On the other hand, if player 1 reports $\hat{\iota}^{\text{p.r.}} = 2$, then player 2 incentivizes player 1 to send an uninformative message (formally, we extend the message space to $\{\text{garbage}\} \cup M_1$ and player 2 gives the positive reward only if player 1 sends $\{\text{garbage}\}$ if $\hat{\iota}^{\text{p.r.}} = 2$. She does not adjust the reward function for player 1 if $\hat{\iota}^{\text{p.r.}} = 2$). As in the case with public randomization, given player 1's message $\hat{\iota}^{\text{p.r.}}$, player 1 wants to send the true message m_i after $\hat{\iota}^{\text{p.r.}} = 1$; and she wants to send $\{\text{garbage}\}$ after $\hat{\iota}^{\text{p.r.}} = 2$.

On the other hand, player 1 adjusts the reward and incentivizes player 2 to tell the truth according to (64), (65), and (66) with $i = 2$, only after player 1 sends $\{\text{garbage}\}$. (If $\hat{\iota}^{\text{p.r.}} = 1$, then player 1 makes player 2 indifferent between any messages and π_2^{adjust} is identically equal to zero.)

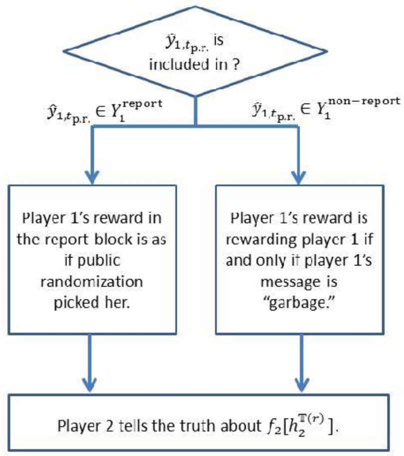

Figure 15: How to coordinate without public randomization

See Figures 15 and 16 for illustration.

Since Y_1^{report} and $Y_1^{\text{non-report}}$ are nonempty, with truthtelling, both $\hat{\imath}^{\text{p.r.}} = 1$ (that is, $y_{1,t_{\text{p.r.}}} \in Y_1^{\text{report}}$) and $\hat{\imath}^{\text{p.r.}} = 2$ (that is, $y_{1,t_{\text{p.r.}}} \in Y_1^{\text{not-report}}$) happen with a positive probability. Hence, each player has the positive probability to get her reward adjusted (that is, the property (i) is established). Moreover, the second sender (player 2) can condition that player 1's message was $\{\text{garbage}\}$, and so there is no learning about $h_1^{\leq L}$ (that is, the property (ii) is established).

We are left to verify that it is optimal for the players to take $a^{\text{p.r.}}$ and player 1 has the incentive to tell the truth about $\imath^{\text{p.r.}}$. Consider player i's incentive to take a pure strategy $a^{\text{p.r.}}$. $a^{\text{p.r.}}$ affects player i's payoff through (80), (81), adjustment of the reward π_i^{adjust}, the rewards for truthtelling (64), (65), and (66). Since the effects other than (80) are sufficiently small compared to (80), it is optimal to take $a_i^{\text{p.r.}}$ for each $i \in \{1, 2\}$.

Given that the players take $a^{\text{p.r.}}$, the report $\hat{\imath}^{\text{p.r.}}$ affects player 1's payoff through (81), adjustment of the reward π_1^{adjust}, the rewards for truthtelling (64), (65), and (66), since (80) is sunk at the point of sending $\hat{\imath}^{\text{p.r.}}$. Except for (81), the effect is of order T^{-5}. Hence, it is optimal to send $\hat{\imath}^{\text{p.r.}} = 1$ if

$$T^{-4}\mathbb{E}\left[1_{\left\{y_{2,t_{\text{p.r.}}}=y_2^{\text{p.r.}}\right\}} \mid a^{\text{p.r.}}, y_{1,t_{\text{p.r.}}}\right] - O(T^{-5}) > T^{-4}\bar{p}_1^{\text{p.r.}}\mathbb{E}\left[1_{\left\{y_{2,t_{\text{p.r.}}}\neq y_2^{\text{p.r.}}\right\}} \mid a^{\text{p.r.}}, y_{1,t_{\text{p.r.}}}\right] + O(T^{-5}),$$

where $O(T^{-5})$ is a random variable of order T^{-5}. That is,

$$\frac{\Pr\left(\left\{y_{2,t_{\text{p.r.}}} = y_2^{\text{p.r.}}\right\} \mid a^{\text{p.r.}}, y_{1,t_{\text{p.r.}}}\right)}{\Pr\left(\left\{y_{2,t_{\text{p.r.}}} \neq y_2^{\text{p.r.}}\right\} \mid a^{\text{p.r.}}, y_{1,t_{\text{p.r.}}}\right)} > \bar{p}_1^{\text{p.r.}} + O(T^{-1}).$$

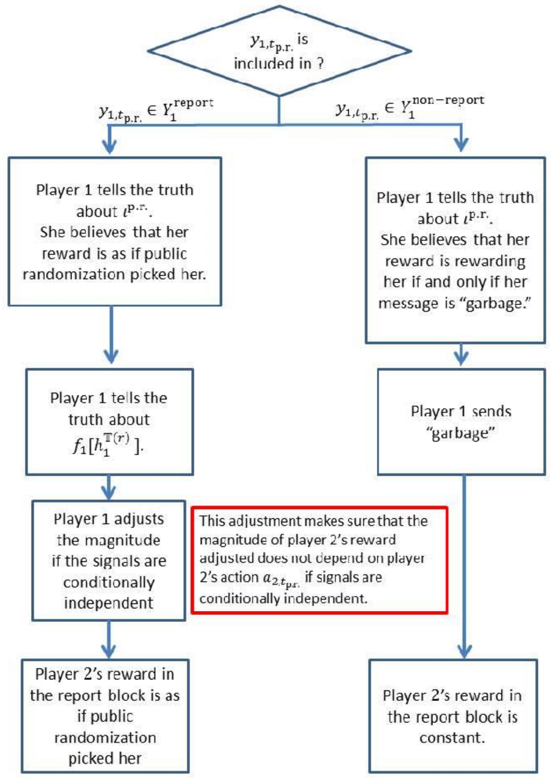

Figure 16: Player 2 can condition that she did not learn player 1's history

For a sufficiently large T, therefore, it is optimal to send $\hat{\iota}^{\text{p.r.}} = 1$ if $y_{1,t_{\text{p.r.}}} \in Y_1^{\text{report}}$. Similarly, we can show that it is optimal to send $\hat{\iota}^{\text{p.r.}} = 2$ if $y_{1,t_{\text{p.r.}}} \in Y_1^{\text{not-report}}$. That is, it is optimal to tell the truth about $\iota^{\text{p.r.}}$.

In summary, we construct an incentive compatible strategy and reward such that the players coordinate on whether player 1 sends the history truthfully, by asking player 1 to guess whether player 2 observed $y_{2,t_{\text{p.r.}}} = y_2^{\text{p.r.}}$ or $y_{2,t_{\text{p.r.}}} \neq y_2^{\text{p.r.}}$. Moreover, player 2's report matters only after player 1 sends the garbage, which incentivizes player 2 to tell the truth. Since the players' signals are correlated, player 1's guess differs after different observations of player 1's signals. Hence, both players have positive probabilities of getting their reward adjusted based on the report block.

Finally, if the players' signals are conditionally independent, player 2 takes a mixed strategy and asks player 1 to guess which action she takes, rather than asks player 1 to guess whether player 2 observed $y_{2,t_{\text{p.r.}}} = y_2^{\text{p.r.}}$ or $y_{2,t_{\text{p.r.}}} \neq y_2^{\text{p.r.}}$.

The proof is the same as above, except that, since player 2 takes a mixed strategy, player 2 has to be indifferent between multiple actions in period $t_{\text{p.r.}}$. If player 2 had different probabilities of getting her reward adjusted after different actions of hers, then depending on player 2's expectation of the adjustment, player 2 would have different incentives to take different actions. To avoid such complication, as will be seen in Appendix A.11.7.3 formally, we make sure that the magnitude of getting her reward adjusted does not depend on player 2's action. Then, player 2 becomes indifferent between actions.

16 Verification of (5)–(8)

Note that Section 7 defines the structure of the finitely repeated game with T being a parameter. For each T, Section 13 defines player i's strategy in the coordination and main blocks, and Lemma 13 defines her strategy in the report block $\sigma_i^{\text{report}}|_{h_i^{\leq L}}$. Hence, we have finished defining the strategy.

In addition, for each T, Section 14.2 defines the reward function as

$$
\begin{aligned}
\pi_i(x_j, h_j^{T_P+1}) &= \bar{\pi}_i(x_j) + \sum_{r=1:\text{round } r \text{ is not a review round}}^{R} \sum_{t\in\mathbb{T}(r)} \pi_i^{x_j}(a_{j,t}, y_{j,t}) \\
&\quad + \sum_{l=1}^{L} \left\{ \pi_i^{\text{review}}(x_j, f_j^{\text{include}}[h_j^{\leq l}], f_j[h_j^{\mathbb{T}(l)}], l) + \pi_i^{\text{adjust}}(x_j, h_j^{\leq l}, h_j^{\text{report}}, l) \right\} \\
&\quad + \pi_i^{\text{report}}(x_j, h_j^{\leq L}, h_j^{\text{report}}).
\end{aligned} \tag{82}
$$

Section 14.2 pins down $\bar{\pi}_i(x_j)$, $\pi_i^{x_j}$, and π_i^{review}, and Lemma 13 pins down π_i^{adjust} and π_i^{report}. Hence, we have finished defining the reward function. Therefore, we are left to verify (5)–(8).

It is useful to recall that Claim 2 of Lemma 13 defines $\pi_i^{\text{adjust}}(x_j, h_j^{\leq l}, h_j^{\text{report}}, l)$ so that, given player i's equilibrium strategy in the report block, we have

$$
\begin{aligned}
& \mathbb{E}\left[\pi_i^{\text{review}}(x_j, h_j^{\leq l}, l) + \pi_i^{\text{adjust}}(x_j, h_j^{\leq l}, h_j^{\text{report}}, l) \mid x_j, h_i^{\leq l}\right] \\
=\; & \mathbb{E}\left[\pi_i^{\text{target}}(x_j, f_i[h_i^{\leq l}], f_j^{\text{include}}[h_j^{\leq l}], f_j[h_j^{\mathbb{T}(l)}], l) \mid x_j, h_i^{\leq l}\right]
\end{aligned} \tag{83}
$$

since $\lambda_i(l)(j) = B$ implies $\pi_i^{\text{review}} = \pi_i^{\text{target}}$ by (57) and (59). Moroever, (59) ensures that

$$\begin{aligned}
& \pi_i^{\text{target}}(x_j, f_i[h_i^{\leq l}], f_j^{\text{include}}[h_j^{\leq l}], f_j[h_j^{\mathbb{T}(l)}], l) \qquad (84) \\
= \; & 1_{\{\lambda_i(l)(j)=B\}} \left\{ \text{sign}(x_j) LT\bar{u} + \sum_{t \in \mathbb{T}(l)} \pi_i^{x_j}(a_{j,t}, y_{j,t}) \right\} \\
& + 1_{\{\lambda_i(l)(j)=G\}} \\
& \times \left\{ \begin{array}{c} 1_{\{\theta_j(l)=G\}} \left\{ \begin{array}{c} 1_{\left\{\lambda_j(l)=G \vee \left\{\lambda_j(l)=B \wedge f_i[h_i^{<l}] \notin \Lambda_i(x(j),l)\right\}\right\}} \left\{ \begin{array}{c} T\left\{u_i(x_j) - u_i(\alpha^\rho(x(j)))\right\} \\ + \sum_{t \in \mathbb{T}(l)} \pi_i[\alpha^\rho(x(j))](a_{j,t}, y_{j,t}) \end{array} \right\} \\ + 1_{\left\{\lambda_j(l)=B \wedge f_i[h_i^{<l}] \in \Lambda_i(x(j),l)\right\}} \left\{u_i(x_j) - u_i(BR_i(\alpha_j^{*,\rho}(x(j))), \alpha_j^{*,\rho}(x(j)))\right\} \end{array} \right\} \\ + 1_{\{\theta_j(l)=B\}} \left\{ \text{sign}(x_j) LT\bar{u} + \sum_{t \in \mathbb{T}(l)} \pi_i^{x_j}(a_{j,t}, y_{j,t}) \right\} \end{array} \right\}.
\end{aligned}$$

16.1 Incentive Compatibility, Promise Keeping, and Full Dimensionality

We prove the optimality of player i's equilibrium strategy by backward induction: In the report block, player i maximizes

$$\mathbb{E}\left[\sum_{t:\text{report block}} u_i(a_{i,t}) + \sum_{l=1}^{L} \pi_i^{\text{adjust}}(x_j, h_j^{\leq l}, h_j^{\text{report}}, l) + \pi_i^{\text{report}}(x_j, h_j^{\leq L}, h_j^{\text{report}}) \mid x_j, h_i^{\leq L} \right]$$

since the other rewards in (82) are sunk in the report block. Hence, by Lemma 13, it is optimal to take $\sigma_i^{\text{report}}|_{h_i^{\leq L}}$ for each $x_j \in \{G, B\}$.

The equilibrium payoff (without taking the average) in the report block satisfies

$$\mathbb{E}\left[\sum_{t:\text{report block}} u_i(a_{i,t}) + \pi_i^{\text{report}}(x_j, h_j^{\leq L}, h_j^{\text{report}}) \mid x_j, \sigma_i^{\text{report}}|_{h_i^{\leq L}}, h_i^{\leq L} \right] = 0$$

by Claim 3 of Lemma 13 (we include $\pi_i^{\text{adjust}}(x_j, h_j^{\leq l}, h_j^{\text{report}}, l)$ to the equilibrium payoff in review round l, rather than in the report block).

In the last review round L, since the continuation payoff from the report block is zero, player i ignores $\sum_{t:\text{report block}} u_i(a_{i,t}) + \pi_i^{\text{report}}(x_j, h_j^{\leq L}, h_j^{\text{report}})$. In addition, by (62), the expected value of

$$\sum_{l \leq L-1} \pi_i^{\text{adjust}}(x_j, h_j^{\leq l}, h_j^{\text{report}}, l)$$

does not depend on player i's strategy after review round $L-1$. Since the rewards $\bar{\pi}_i(x_j)$, $\sum_{r=1:\text{round } r \text{ is not a review round}}^{R} \sum_{t\in\mathbb{T}(r)} \pi_i^{x_j}(a_{j,t}, y_{j,t})$, and $\pi_i^{\text{review}}(x_j, f_j^{\text{include}}[h_j^{\leq l}], f_j[h_j^{\mathbb{T}(l)}], l)$ for $l \leq L-1$ are sunk in review round L, player i wants to maximize, with $l = L$,

$$\mathbb{E}\left[\sum_{t\in\mathbb{T}(l)} u_i(a_t) + \pi_i^{\text{review}}(x_j, f_j^{\text{include}}[h_j^{\leq l}], f_j[h_j^{\mathbb{T}(l)}], l) + \pi_i^{\text{adjust}}(x_j, h_j^{\leq l}, h_j^{\text{report}}, l) \mid x_j, \sigma_i(x_i), h_i^{<l}\right].$$

Given (83), by the law of iterated expectation, player i wants to maximize, with $l = L$,

$$\mathbb{E}\left[\sum_{t\in\mathbb{T}(l)} u_i(a_t) + \pi_i^{\text{target}}(x_j, f_i[h_i^{\leq l}], f_j^{\text{include}}[h_j^{\leq l}], f_j[h_j^{\mathbb{T}(l)}], l) \mid x_j, \sigma_i(x_i), h_i^{<l}\right].$$

Hence, by Lemma 2 and (59), if $\theta_j(l) = B$, then player i wants to maximize

$$\mathbb{E}\left[\sum_{t\in\mathbb{T}(l)} u_i(a_t) + \text{sign}\,(x_j)\, LT\bar{u} + \sum_{t\in\mathbb{T}(l)} \pi_i^{x_j}(a_{j,t}, y_{j,t}) \mid x_j, \sigma_i(x_i), h_i^{<l}\right].$$

Claim 2 of Lemma 2 ensures that any strategy is optimal and the equilibrium value (without taking the average) is equal to $\text{sign}(x_j)\, LT\bar{u} + T\bar{u}_i^{x_j}$.

On the other hand, if $\theta_j(l) = G$, then we have $\lambda_i(l)(j) = G$ by Claim 3 of Lemma 11. Hence, if $\lambda_j(l) = G$ or "$\lambda_j(l) = B$ and $h_i^{<l} \notin \Lambda_i(x(j), l)$", then player i wants to maximize

$$\mathbb{E}\left[\sum_{t\in\mathbb{T}(l)} u_i(a_t) + T\left\{u_i(x_j) - u_i(\alpha^{\rho}(x(j)))\right\} + \sum_{t\in\mathbb{T}(l)} \pi_i[\alpha^{\rho}(x(j))](a_{j,t}, y_{j,t}) \mid x_j, \sigma_i(x_i), h_i^{<l}\right].$$

By Claim 1 of Lemma 2, any action is optimal and player i's equilibrium value is $Tu_i(x_j)$. In addition, $\lambda_j(l) = B$ and $h_i^{<l} \in \Lambda_i(x(j), l)$, then player i wants to maximize

$$\mathbb{E}\left[\sum_{t\in\mathbb{T}(l)} u_i(a_t) + u_i(x_j) - u_i(BR_i(\alpha_j^{*,\rho}(x(j))), \alpha_j^{*,\rho}(x(j))) \mid x_j, \sigma_i(x_i), h_i^{<l}\right].$$

Since the reward is constant, player i wants to take a static best response to player j's strategy. Moreover, Claim 4 of Lemma 11 ensures that player j with $\theta_j(l) = G$ and $\lambda_j(l) = B$ takes $\alpha_j^{*,\rho}(x(j))$. Hence, $BR_i\left(\alpha_j^{*,\rho}(x(j))\right)$ is optimal. Since $\Lambda_i(x(j), l)$ is the set of player i's history in which player

i has $x(i) = x(j)$ and $\lambda_j(l)(i) = B$ with probability one. Figure 10 ensures that player i with $\Lambda_i(x(j), l)$ takes $BR_i\left(\alpha_j^{*,\rho}(x(j))\right)$ with probability one. Hence, player i's strategy is optimal. Moreover, her equilibrium value is $Tu_i(x_j)$.

In total, with $l = L$, the equilibrium payoff is

$$\begin{aligned} &\mathbb{E}\left[\sum_{t\in\mathbb{T}(l)} u_i(a_t) + \pi_i^{\text{target}}(x_j, f_i[h_i^{\leq l}], f_j^{\text{include}}[h_j^{\leq l}], f_j[h_j^{\mathbb{T}(l)}], l) \mid x_j, h_i^{<l}\right] \\ = &\begin{cases} Tu_i(x_j) & \text{if } \theta_j(l) = G, \\ \text{sign}\,(x_j)\, LT\bar{u} + T\bar{u}_i^{x_j} & \text{if } \theta_j(l) = B. \end{cases} \end{aligned}$$

Since the distribution of $\theta_j(l)$ does not depend on player i's strategy by Claim 2 of Lemma 11, player i before review round L ignores the continuation payoff from review round L on. Again, (62) ensures that the expected value of

$$\sum_{l\leq L-1} \pi_i^{\text{adjust}}(x_j, h_j^{\leq l}, h_j^{\text{report}}, l)$$

does not depend on player i's strategy after review round $L-1$. Ignoring the rewards in (82) that have been sunk, player i in the supplemental rounds for $\lambda_1(l)$ and $\lambda_2(l)$ with $l = L$ maximizes

$$\mathbb{E}\left[\sum_{t\in\mathbb{T}(\lambda_1(l))\cup\mathbb{T}(\lambda_2(l))} u_i(a_t) + \sum_{t\in\mathbb{T}(\lambda_1(l))\cup\mathbb{T}(\lambda_2(l))} \pi_i^{x_j}(a_{j,t}, y_{j,t}) \mid x_j, h_i^{\leq l-1}\right].$$

Therefore, any strategy is optimal and the equilibrium value is equal to $2T^{\frac{1}{2}}\bar{u}_i^{x_j}$ since $|\mathbb{T}(\lambda_1(l))| = |\mathbb{T}(\lambda_2(l))| = T^{\frac{1}{2}}$.

Again, this value does not depend on the previous history. Hence, player i in review round $L-1$ ignores the continuation payoff from the supplemental round for $\lambda_1(L)$ on. Therefore, by the same proof as review round L, we can establish the optimality of the equilibrium strategy. Recursively, we can show that the equilibrium strategy is optimal in the main and report blocks, and the equilibrium payoff from the main and report blocks is equal to

$$2(L-1)\,T^{\frac{1}{2}}\bar{u}_i^{x_j} + \sum_{l=1}^{L}\left\{1_{\{\theta_j(l)=G\}} Tu_i(x_j) + 1_{\{\theta_j(l)=B\}}\left(\text{sign}\,(x_j)\, LT\bar{u} + T\bar{u}_i^{x_j}\right)\right\}.$$

Since this value does not depend on player i's strategy in the coordination block, player i in the coordination block maximizes

$$\mathbb{E}\left[\sum_{t:\text{coordination block}} \left\{u_i(a_t) + \pi_i^{x_j}(a_{j,t}, y_{j,t})\right\} \mid x_j\right].$$

Hence, any strategy is optimal and equilibrium value is equal to $\left(4T^{\frac{1}{2}} + 2T^{\frac{2}{3}}\right)\bar{u}_i^{x_j}$, recalling that $4T^{\frac{1}{2}} + 2T^{\frac{2}{3}}$ is the length of the coordination block.

Therefore, in total, player i's value from the review phase except for $\bar{\pi}_i(x_j)$ is equal to

$$\left(4T^{\frac{1}{2}} + 2T^{\frac{2}{3}} + 2(L-1)T^{\frac{1}{2}}\right)\bar{u}_i^{x_j} + \sum_{l=1}^{L}\left\{1_{\{\theta_j(l)=G\}}Tu_i(x_j) + 1_{\{\theta_j(l)=B\}}\left(\text{sign}(x_j)LT\bar{u} + T\bar{u}_i^{x_j}\right)\right\}. \tag{85}$$

Hence, by definition of $\bar{\pi}_i(x_j)$ in (52), the total equilibrium value (without taking the average) is equal to $T_P v_i(x_j)$, as desired. In total, we have verified incentive compatibility, promise keeping, and full dimensionality.

16.2 Self Generation

We first show that we can ignore the terms except for $\sum_{l=1}^{L}\pi_i^{\text{review}}(x_j, h_j^{\leq l}, l)$. To see this, define $\varepsilon > 0$ such that, for each $x_j \in \{G, B\}$, we have

$$(15+7L)\eta\left\{|u_i(x_j)| + L\bar{u} + \left|\bar{u}_i^{x_j}\right|\right\} + \varepsilon < |u_i(x_j) - v_i(x_j)|. \tag{86}$$

(16) ensures that there exists such $\varepsilon > 0$.

Note that the terms other than $\sum_{l=1}^{L}\pi_i^{\text{review}}(x_j, h_j^{\leq l}, l)$ in $\pi_i(x_j, h_j^{T_P+1})$ satisfy, for each $h_j^{\leq L}$ and h_j^{report},

$$\begin{cases} \bar{\pi}_i(x_j) = T_P v_i(x_j) - \mathbb{E}\left[\sum_{l=1}^{L}\left\{1_{\{\theta_j(l)=G\}}Tu_i(x_j) + 1_{\{\theta_j(l)=B\}}T\left(\text{sign}(x_j)LT\bar{u} + T\bar{u}_i^{x_j}\right)\right\}\right] \\ \qquad - \left\{(2+2L)T^{\frac{1}{2}} + 2T^{\frac{2}{3}}\right\}\bar{u}_i^{x_j}, \\ \left|\sum_{r=1:\text{round } r \text{ is not a review round}}^{R}\pi_i^{x_j}(a_{j,t}, y_{j,t})\right| \leq (T_P - LT)\bar{u} \text{ by Lemma 2,} \\ \left|\pi_i^{\text{adjust}}(x_j, h_j^{\leq l}, h_j^{\text{report}}, l)\right| \leq \exp\left(-T^{\frac{1}{5}}\right) \text{ by Lemma 13,} \\ \left|\pi_i^{\text{report}}(x_j, h_j^{\leq L}, h_j^{\text{report}})\right| \leq K_{\text{report}}T^{\frac{11}{12}} \text{ by Lemma 13.} \end{cases}$$

By Lemma 5, T_P and LT are similar to each other for a large T. Hence, for a sufficiently large T, for each $h_j^{\leq L}$ and h_j^{report}, we have

$$\frac{1}{T_P}\left\{\begin{array}{c}\bar{\pi}_i(x_j)+\sum_{r=1:\text{ round } r \text{ is not a review round}}^{R}\pi_i^{x_j}(a_{j,t},y_{j,t})\\ +\sum_{l=1}^{L}\pi_i^{\text{adjust}}(x_j,h_j^{\leq l},h_j^{\text{report}},l)+\pi_i^{\text{report}}(x_j,h_j^{\leq L},h_j^{\text{report}})\end{array}\right\}$$
$$\in\left[\begin{array}{c}v_i(x_j)-\mathbb{E}\left[\sum_{l=1}^{L}\left\{1_{\{\theta_j(l)=G\}}\frac{1}{L}u_i(x_j)+1_{\{\theta_j(l)=B\}}\frac{1}{L}\left(\text{sign}\left(x_j\right)L\bar{u}+\bar{u}_i^{x_j}\right)\right\}\right]-\varepsilon,\\ v_i(x_j)-\mathbb{E}\left[\sum_{l=1}^{L}\left\{1_{\{\theta_j(l)=G\}}\frac{1}{L}u_i(x_j)+1_{\{\theta_j(l)=B\}}\frac{1}{L}\left(\text{sign}\left(x_j\right)L\bar{u}+\bar{u}_i^{x_j}\right)\right\}\right]+\varepsilon\end{array}\right].$$

Moreover, by Claim 2 of Lemma 11, the probability of $\theta_j(l)=B$ is no more than $(15+8L)\,\eta$ for each l. Taking $l=L$, this means that $\theta_j(L)=G$ with probability no less than $1-(15+8L)\,\eta$. Since $\theta_j(l)=B$ is absorbing, this also means that $\theta_j(l)=G$ for each l with probability no less than $1-(15+8L)\,\eta$. Hence, if $x_j=G$, we have

$$\begin{aligned}&\frac{1}{T_P}\left\{\begin{array}{c}\bar{\pi}_i(x_j)+\sum_{r=1:\text{ round } r \text{ is not a review round}}^{R}\pi_i^{x_j}(a_{j,t},y_{j,t})\\ +\sum_{l=1}^{L}\pi_i^{\text{adjust}}(x_j,h_j^{\leq l},h_j^{\text{report}},l)+\pi_i^{\text{report}}(x_j,h_j^{\leq L},h_j^{\text{report}})\end{array}\right\}\\ \leq\ & v_i(x_j)-u_i(x_j)+(15+8L)\,\eta\left\{|u_i(x_j)|+L\bar{u}+\left|\bar{u}_i^{x_j}\right|\right\}+\varepsilon,\end{aligned}$$

which is non-positive by (86). Similarly, if $x_j=B$, we have

$$\begin{aligned}&\lim_{T\to\infty}\frac{1}{T_P}\left\{\begin{array}{c}\bar{\pi}_i(x_j)+\sum_{r=1:\text{ round } r \text{ is not a review round}}^{R}\pi_i^{x_j}(a_{j,t},y_{j,t})\\ +\sum_{l=1}^{L}\pi_i^{\text{adjust}}(x_j,h_j^{\leq l},h_j^{\text{report}},l)+\pi_i^{\text{report}}(x_j,h_j^{\leq L},h_j^{\text{report}})\end{array}\right\}\\ \geq\ & v_i(x_j)-u_i(x_j)-(15+8L)\,\eta\left\{|u_i(x_j)|+L\bar{u}+\left|\bar{u}_i^{x_j}\right|\right\}-\varepsilon\geq 0.\end{aligned}$$

Hence, for a sufficiently large T, we have $\text{sign}(x_j)\left\{\pi_i(x_j,h_j^{\leq T_P+1})-\sum_{l=1}^{L}\pi_i^{\text{review}}(x_j,h_j^{\leq l},l)\right\}\geq 0$. Therefore, it suffices to show that $\text{sign}(x_j)\sum_{l=1}^{L}\pi_i^{\text{review}}(x_j,h_j^{\leq l},l)\geq 0$.

Moreover, if $\theta_j(l)=B$ for some $l=1,...,L$, then we have $\text{sign}(x_j)\sum_{l=1}^{L}\pi_i^{\text{review}}(x_j,h_j^{\leq l},l)\geq 0$.

To see this, recall that, for each history of player j, we have

$$\begin{aligned}
& \pi_i^{\text{review}}(x_j, f_j^{\text{include}}[h_j^{\leq l}], f_j[h_j^{\mathbb{T}(l)}], l) \\
= \; & 1_{\{\lambda_i(l)(j)=B\}} \left\{ \text{sign}\,(x_j)\, LT\bar{u} + \sum_{t\in\mathbb{T}(l)} \pi_i^{x_j}(a_{j,t}, y_{j,t}) \right\} \\
& +1_{\{\lambda_i(l)(j)=G\}} \left\{ \begin{array}{c} 1_{\{\theta_j(l)=G\}} \left\{ \begin{array}{c} 1_{\{\lambda_j(l)=G\}} \left\{ \begin{array}{c} T\left\{u_i(x_j) - u_i(\alpha^{\rho}(x(j)))\right\} \\ +\sum_{t\in\mathbb{T}(l)} \pi_i[\alpha^{\rho}(x(j))](a_{j,t}, y_{j,t}) \end{array} \right\} \\ +1_{\{\lambda_j(l)=B\}} T\{u_i(x_j) - u_i(BR_i(\alpha_j^{*,\rho}(x(j))), \alpha_j^{*,\rho}(x(j)))\} \end{array} \right\} \\ +1_{\{\theta_j(l)=B\}} \left\{ \text{sign}\,(x_j)\, LT\bar{u} + \sum_{t\in\mathbb{T}(l)} \pi_i^{x_j}(a_{j,t}, y_{j,t}) \right\} \end{array} \right\}.
\end{aligned}$$

For each history of player j, by Lemma 2, we have $\text{sign}(x_j) \sum_{t\in\mathbb{T}(l)} \pi_i^{x_j}(a_{j,t}, y_{j,t}) \geq 0$. Moreover, by Claim 5 of Lemma 11, $x(j) = x_j$ if $\theta_j(l) = G$. Hence, (15) ensures that

$$\left\{ \begin{array}{c} \text{sign}\,(x_j)\, 1_{\{\theta_j(l)=G\}} T\left\{u_i(x_j) - u_i(\alpha^{\rho}(x(j)))\right\} \geq 0, \\ \text{sign}\,(x_j)\, 1_{\{\theta_j(l)=G\}} \{u_i(x_j) - u_i(BR_i(\alpha_j^{*,\rho}(x(j))), \alpha_j^{*,\rho}(x(j)))\} \geq 0. \end{array} \right.$$

Hence, together with Claim 3 of Lemma 11 ($\lambda_i(l)(j) = B \Rightarrow \theta_j(l) = B$), we have

$$\begin{aligned}
& \text{sign}\,(x_j)\, \pi_i^{\text{review}}(x_j, f_j^{\text{include}}[h_j^{\leq l}], f_j[h_j^{\mathbb{T}(l)}], l) \\
\geq \; & 1_{\{\theta_j(l)=G\}} \text{sign}\,(x_j) \sum_{t\in\mathbb{T}(l)} \pi_i[\alpha^{\rho}(x(j))](a_{j,t}, y_{j,t}) + 1_{\{\theta_j(l)=B\}} LT\bar{u} \\
\geq \; & -1_{\{\theta_j(l)=G\}} T\bar{u} + 1_{\{\theta_j(l)=B\}} LT\bar{u} \text{ by Lemma 2.}
\end{aligned}$$

Therefore, if $\theta_j(l) = B$ for some $l = 1, ..., L$, then we have

$$\sum_{l=1}^{L} \text{sign}\,(x_j)\, \pi_i^{\text{review}}(x_j, f_j^{\text{include}}[h_j^{\leq l}], f_j[h_j^{\mathbb{T}(l)}], l) \geq -\left(L-1\right) T\bar{u} + LT\bar{u} \geq 0,$$

as desired.

Therefore, we focus on the case with $\theta_j(l) = G$ for each $l = 1, ..., L$. By Claim 5 of Lemma 11, we have $x_j(j) = x_j$. Let $\bar{l}$ be the last review round with $\lambda_j(l) = G$ (define $\bar{l} = L$ if $\lambda_j(l) = G$ for each $l = 1, ..., L$). If $x_j = G$, then recalling that $\alpha_j^{*,\rho}(x(j)) = \alpha_j^{\rho}(x(j))$ with $x_j(j) = x_j = G$, we have

$$
\begin{aligned}
& \sum_{l=1}^{L} \pi_i^{\text{review}}(x_j, h_j^{\mathbb{T}(l)}, l) \\
= & \sum_{l=1}^{\bar{l}} \left\{ T\left\{u_i(x_j) - u_i(\alpha^{\rho}(x(j)))\right\} + \sum_{t\in\mathbb{T}(l)} \pi_i[\alpha^{\rho}(x(j))](a_{j,t}, y_{j,t}) \right\} \\
& + \sum_{l=\bar{l}+1}^{L} T\left\{u_i(x_j) - u_i(BR_i(\alpha_j^{*,\rho}(x(j))), \alpha_j^{*,\rho}(x(j)))\right\} \\
\leq & \bar{l}T \max_{x_i(j)\in\{G,B\}} \left\{u_i(x_j) - u_i(\alpha^{\rho}(x_j, x_i(j)))\right\} + (L-\bar{l})T\{ \underbrace{u_i(x_j) - \underbrace{u_i(BR_i(\alpha_j^{\rho}(x(j))), \alpha_j^{\rho}(x(j)))}_{\geq u_i(\alpha^{\rho}(x(j)))}}_{\leq \max_{x_i(j)\in\{G,B\}}\{u_i(x_j)-u_i(\alpha^{\rho}(x_j,x_i(j)))\} \text{ since } x_j(j)=x_j} \} \\
& + \underbrace{\sum_{l=1}^{\bar{l}-1} \frac{\bar{u}}{4L}T}_{\substack{\text{since } \lambda_j(\bar{l})=G, \\ \left|\sum_{t\in\mathbb{T}(l)} \pi_i[\alpha^{\rho}(x(j))](a_{j,t},y_{j,t})\right| \leq \frac{\bar{u}}{4L}T \\ \text{for } l=1,\dots,\bar{l}-1}} + \underbrace{\frac{\bar{u}}{4}T}_{\substack{\text{since } \pi_i[\alpha^{\rho}](a_j,y_j): A_j\times Y_j \to \left[-\frac{\bar{u}}{4},\frac{\bar{u}}{4}\right] \text{ by Lemma 2,} \\ \frac{\bar{u}}{4}T \text{ is the maximum realization} \\ \text{for } \sum_{t\in\mathbb{T}(\bar{l})} \pi_i[\alpha^{\rho}(x(j))](a_{j,t},y_{j,t})}} \\
\leq & LT \max_{x_i(j)\in\{G,B\}} \left\{u_i(x_j) - u_i(\alpha^{\rho}(x_j, x_i(j)))\right\} + \frac{\bar{u}}{4}T + \frac{\bar{u}}{4}T \leq 0
\end{aligned}
$$

by (15). Similarly, if $x_j = B$, we have

$$
\begin{aligned}
& \sum_{l=1}^{L} \pi_i^{\text{review}}(x_j, h_j^{\mathbb{T}(l)}, l) \\
= & \sum_{l=1}^{\bar{l}} \left\{ T\left\{u_i(x_j) - u_i(\alpha^{\rho}(x(j)))\right\} + \sum_{t\in\mathbb{T}(l)} \pi_i[\alpha^{\rho}(x(j))](a_{j,t}, y_{j,t}) \right\} \\
& + \sum_{l=\bar{l}+1}^{L} T\left\{u_i(x_j) - u_i(BR_i(\alpha_j^{*,\rho}(x(j))), \alpha_j^{*,\rho}(x(j)))\right\} \\
\geq & \bar{l}T \min_{x_i(j)\in\{G,B\}} \left\{u_i(x_j) - u_i(\alpha^{\rho}(x_j, x_i(j)))\right\} \\
& +(L-\bar{l})T\{ \underbrace{u_i(x_j) - u_i(BR_i(\alpha_j^{*,\rho}(x(j))), \alpha_j^{*,\rho}(x(j)))}_{\geq \min_{x_i(j)\in\{G,B\}}\left\{u_i(x_j)-u_i(BR_i(\alpha_j^{*,\rho}(x(j))),\alpha_j^{*,\rho}(x(j)))\right\} \text{ since } x_j(j)=x_j} \} \\
& -\sum_{l=1}^{\bar{l}-1} \frac{\bar{u}}{4L}T - \frac{\bar{u}}{4}T \\
\geq & LT \max_{x_i(j)\in\{G,B\}} \left\{u_i(x_j) - \max\left\{u_i(BR_i(\alpha_j^{*,\rho}(x(j))), \alpha_j^{*,\rho}(x(j))), u_i(\alpha^{\rho}(x_j, x_i(j)))\right\}\right\} + \frac{\bar{u}}{4}T + \frac{\bar{u}}{4}T \geq 0
\end{aligned}
$$

by (15). Therefore, self generation is satisfied.

Since we have proven (5)–(8), Lemma 1 ensures that Theorem 1 holds.

References

Aoyagi, M. (2002): "Collusion in dynamic Bertrand oligopoly with correlated private signals and communication," *Journal of Economic Theory*, 102(1), 229–248.

Aoyagi, M. (2005): "Collusion through Mediated Communication in Repeated Games with Imperfect Private Monitoring," *Economic Theory*, 25(2), pp. 455–475.

Boucheron, S., G. Lugosi, and P. Massart (2013): *Concentration Inequalities: A Nonasymptotic Theory of Independence*. OUP Oxford.

Compte, O. (1998): "Communication in repeated games with imperfect private monitoring," *Econometrica*, 66(3), 597–626.

Deb, J. (2011): "Cooperation and community responsibility: A folk theorem for repeated random matching games," *mimeo*.

Ely, J., J. Hörner, and W. Olszewski (2005): "Belief-free equilibria in repeated games," *Econometrica*, 73(2), 377–415.

Ely, J., and J. Välimäki (2002): "A robust folk theorem for the prisoner's dilemma," *Journal of Economic Theory*, 102(1), 84–105.

Fudenberg, D., and D. Levine (2007): "The Nash-threats folk theorem with communication and approximate common knowledge in two player games," *Journal of Economic Theory*, 132(1), 461–473.

Fudenberg, D., D. Levine, and E. Maskin (1994): "The folk theorem with imperfect public information," *Econometrica*, 62(5), 997–1039.

Fudenberg, D., and E. Maskin (1986): "The folk theorem in repeated games with discounting or with incomplete information," *Econometrica*, 53(3), 533–554.

Harrington, J., and A. Skrzypacz (2011): "Private monitoring and communication in cartels: Explaining recent collusive practices," *American Economic Review*, 101(6), 2425–2449.

Hörner, J., and W. Olszewski (2006): "The folk theorem for games with private almost-perfect monitoring," *Econometrica*, 74(6), 1499–1544.

——— (2009): "How robust is the folk theorem?," *The Quarterly Journal of Economics*, 124(4), 1773–1814.

Kandori, M. (2002): "Introduction to repeated games with private monitoring," *Journal of Economic Theory*, 102(1), 1–15.

——— (2011): "Weakly belief-free equilibria in repeated games with private monitoring," *Econometrica*, 79(3), 877–892.

Kandori, M., and H. Matsushima (1998): "Private observation, communication and collusion," *Econometrica*, 66(3), 627–652.

Kandori, M., and I. Obara (2006): "Efficiency in repeated games revisited: The role of private strategies," *Econometrica*, 74(2), 499–519.

Lehrer, E. (1990): "Nash equilibria of n-player repeated games with semi-standard information," *International Journal of Game Theory*, 19(2), 191–217.

Mailath, G., and L. Samuelson (2006): *Repeated games and reputations: long-run relationships.* Oxford University Press.

Matsushima, H. (2004): "Repeated games with private monitoring: Two players," *Econometrica*, 72(3), 823–852.

Miyagawa, E., Y. Miyahara, and T. Sekiguchi (2008): "The folk theorem for repeated games with observation costs," *Journal of Economic Theory*, 139(1), 192–221.

Obara, I. (2009): "Folk theorem with communication," *Journal of Economic Theory*, 144(1), 120–134.

Piccione, M. (2002): "The repeated prisoner's dilemma with imperfect private monitoring," *Journal of Economic Theory*, 102(1), 70–83.

Radner, R., R. Myerson, and E. Maskin (1986): "An example of a repeated partnership game with discounting and with uniformly inefficient equilibria," *Review of Economic Studies*, 53(1), 59–69.

Rahman, D. (2014): "The Power of Communication," *American Economic Review*, 104(11), 3737–51.

Renault, J., and T. Tomala (2004): "Communication equilibrium payoffs in repeated games with imperfect monitoring," *Games and Economic Behavior*, 49(2), 313–344.

Sekiguchi, T. (1997): "Efficiency in repeated prisoner's dilemma with private monitoring," *Journal of Economic Theory*, 76(2), 345–361.

Stigler, G. (1964): "A theory of oligopoly," *The Journal of Political Economy*, 72(1), 44–61.

Sugaya, T. (2012): "Folk theorem in repeated games with private monitoring: multiple players," *mimeo*.

Takahashi, S. (2010): "Community enforcement when players observe partners' past play," *Journal of Economic Theory*, 145(1), 42–62.

Yamamoto, Y. (2007): "Efficiency results in N player games with imperfect private monitoring," *Journal of Economic Theory*, 135(1), 382–413.

——— (2009): "A limit characterization of belief-free equilibrium payoffs in repeated games," *Journal of Economic Theory*, 144(2), 802–824.

——— (2012): "Characterizing belief-free review-strategy equilibrium payoffs under conditional independence," *mimeo*.