THE FOLK THEOREM IN REPEATED GAMES WITH PRIVATE MONITORING

Takuo Sugaya

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF
ECONOMICS

Adviser: Stephen Morris

June 2012

Abstract

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

In Chapter 1, we introduce the model, states the assumptions and the main result, and offer the overview of the proof. In Chapter 2, we show the folk theorem in the two-player prisoners' dilemma, assuming special forms of communication. Given this chapter, we are left to extend the folk theorem to the general two-player game and the general N-player game with $N \geq 3$ and dispense with the special forms of communication. In Chapter 3, we summarize what new assumptions are sufficient for each extension. In the following chapters, we offer the proof: in Chapters 4 and 5, we extend the result to the general two-player game and the general N-player game, respectively, with the special forms of communication. In Chapters 6 and 7, we dispense with the special forms of communication in the two-player game and N-player game, respectively.

Acknowledgement

I thank Stephen Morris, my adviser, for invaluable advice, Yuliy Sannikov for helpful discussion, and Dilip Abreu, Eduardo Faingold, Drew Fudenberg, Edoardo Grillo, Johannes Hörner, Yuhta Ishii, Michihiro Kandori, Vijay Krishna, David Levine, George Mailath, Eric Maskin, Mihai Manea, Larry Samuelson, Tadashi Sekiguchi, Andrzej Skrzypacz, Satoru Takahashi, Yuichi Yamamoto and seminar participants at Harvard University, Penn State University, Princeton University, Stanford Graduate School of Business, University of California Berkeley, University of California Los Angeles, University of Pennsylvania, Washington University St. Louis, 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.

Contents

Chapter 1

Introduction

1.1 Introduction

One of the key results in the literature on infinitely repeated games is the folk theorem: any feasible and individually rational payoff can be sustained in equilibrium when players are sufficiently patient. Even if a stage game does not have an efficient Nash equilibrium, the repeated game does. Hence, the repeated game gives a formal framework to analyze a cooperative behavior. Fudenberg and Maskin (1986) establish the folk theorem under perfect monitoring, that is, when players can directly observe the action profile. Fudenberg, Levine, and Maskin (1994) extend the folk theorem to imperfect public monitoring, where players can observe only public noisy signals about the action profile.

The driving force of the folk theorem is reciprocity: if a player deviates today, she will be punished in future. For this mechanism to work, players need to infer what actions are expected by the other players. For example, in the trigger strategy equilibrium of the prisoners' dilemma with perfect monitoring, if a player takes defection before the other players take defection, then it is seen as a deviation of that player. On the other hand, if a player takes defection after some player takes defection, then it is seen as an equilibrium behavior. Hence, to know whether to cooperate or defect, each player needs to infer which action is expected by the other players. In other words, each player needs to coordinate her action with the other players' histories.

This coordination is straightforward if the players' strategies only depend on the public component of histories, such as action profiles in perfect monitoring or public signals in public monitoring. Since this public information is common knowledge, players can coordinate a punishment contingent on the public information (reciprocity), and thereby provide dynamic incentives to choose actions that are not static best responses. However, with private monitoring, since they do not share common information about histories, this coordination becomes complicated as periods proceed. Hence, "coordination failure" may arise.

Hörner and Olszewski (2006) and Hörner and Olszewski (2009) show the robustness of this coordination to private monitoring, where players can observe only private noisy signals about the action profile, if monitoring is almost perfect and almost public, respectively. If monitoring is almost perfect, then players can believe that every player observes the same signal corresponding to the action profile with a high probability. If monitoring is almost public, then players can believe that every player observes the same signal with a high probability.[1] Hence, almost common knowledge about relevant histories still exists.

However, with general private monitoring, almost common knowledge may not exist and coordination is difficult (we call this problem "coordination failure").[2] Hence, the robustness of the folk theorem to general private monitoring has been an open question. For example, Kandori (2002) states that "[t]his is probably one of the best known long-standing open questions in economic theory."[3]

This paper is the first to show that the folk theorem holds in repeated games with discounting and generic monitoring: in any N-player repeated game with private monitoring, we give sufficient conditions with which any feasible and individually rational payoff is sus-

[1] See also Mailath and Morris (2002) and Mailath and Samuelson (2006).

[2] Mailath and Morris (2002), Mailath and Samuelson (2006) and Sugaya and Takahashi (2011) offer the formal models of this argument.

[3] See Mailath and Samuelson (2006) for a survey.

tainable in a sequential equilibrium for a sufficiently large discount factor.[4] We also show that these sufficient conditions are generic if the cardinality of the support of each player's signal distribution is sufficiently large.

Repeated games with private monitoring are relevant for many traditional economic problems. For example, Stigler (1964) proposes a repeated price-setting oligopoly, where firms set their own prices in face-to-face negotiations and cannot directly observe their opponents' prices. Instead, a firm obtains some information about opponents' prices through its own sales. Since the level of sales depends on both opponents' prices and unobservable demand shocks, the sales level is an imperfect signal. Moreover, each firm's sales level is often private information since it is also determined in a face-to-face negotiation. Thus, the monitoring is imperfect and private. In principal-agent problems, if the principal evaluates the agent subjectively, then the monitoring by the principal about the agent becomes private. Despite the importance of these problems, only a limited number of papers successfully analyze repeated games with private monitoring.[5] Our result offers a benchmark to analyze these problems in a general private-monitoring setting.

To show the folk theorem under general monitoring, we unify and improve on three approaches in the literature on private monitoring that have been used to show the partial results so far: belief-free, belief-based and communication approaches.

The belief-free approach (and its generalizations) has been successful in showing the folk theorem in the prisoners' dilemma.[6] A strategy profile is belief-free if, for any history

[4]See Lehrer (1990) for the case of no discounting.

[5]Harrington and Skrzypacz (2011) show evidence of cooperative behavior (cartels) among firms in lysine and vitamin industries. After arguing that these industries fit Stigler's setup, they write a repeated-game model with private monitoring and solve a special case. See also Harrington Jr and Skrzypacz (2007).

Fuchs (2007) applies a repeated game with private monitoring to a contract between a principal and an agent with subjective evaluation.

[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 where a player can improve the precision of monitoring by paying cost.

profile, the continuation strategy of each player is optimal conditional on the histories of the opponents. Hence, coordination failure never happens. With almost perfect monitoring, Piccione (2002) and Ely and Välimäki (2002) show the folk theorem for the two-player prisoners' dilemma.[7] Without any assumption on the precision of monitoring but with conditionally independent monitoring, Matsushima (2004) obtains the folk theorem in the two-player prisoners' dilemma, which is extended by Yamamoto (2012) to the N-player prisoners' dilemma with conditionally independent monitoring.[8]

Previously, attempts to generalize Matsushima (2004) have shown only limited results without almost perfect or conditionally independent monitoring: for some restricted classes of the distributions of private signals, Fong, Gossner, Hörner, and Sannikov (2010) show that the payoff of the mutual cooperation is approximately attainable in the two-player prisoners' dilemma. Sugaya (2012) shows that the folk theorem holds with a general monitoring structure in the prisoners' dilemma if the number of players is no less than four.

Several papers construct belief-based equilibria, where players' strategies involve statistical inference about the opponents' past histories. That is, since common knowledge about relevant histories no longer exists, each player calculates the beliefs about the opponents' histories to calculate best responses. With almost perfect monitoring, Sekiguchi (1997) shows that the payoff of the mutual cooperation is approximately attainable and Bhaskar and Obara (2002) show the folk theorem in the two-player prisoners' dilemma.[9] Phelan and Skrzypacz (2012) characterize the set of possible beliefs about opponents' states in a finite-state automaton strategy and Kandori and Obara (2010) offer a way to verify if a finite-state automaton strategy is an equilibrium.

[7]See Yamamoto (2007) for the N-player prisoners' dilemma. Ely, Hörner, and Olszewski (2005) and Yamamoto (2009) characterize the set of belief-free equilibrium payoffs for a general game. Except for the prisoners' dilemma, this set is not so large as that of feasible and individually rational payoffs.

[8]The strategy used in Matsushima (2004) is called a "belief-free review strategy." See Yamamoto (2012) for the characterization of the set of belief-free review-strategy equilibrium payoffs for a general game with conditional independence. Again, except for the prisoners' dilemma, this set is not so large as that of feasible and individually rational payoffs.

[9]Bhaskar and Obara (2002) also derive a sufficient condition for the N-player prisoners' dilemma.

Another approach to analyze repeated games with private monitoring introduces public communication. Folk theorems have been proven by Compte (1998), Kandori and Matsushima (1998), Aoyagi (2002), Fudenberg and Levine (2007) and Obara (2009). Introducing a public element (the result of communication) and letting a strategy depend only on the public element allow these papers to sidestep the difficulty of coordination through private signals. However, the analyses are not applicable to settings where communication is not allowed: for example, in Stigler (1964)'s oligopoly example, anti-trust laws prohibit communication. Hörner and Olszewski (2006) also argue that "communication reintroduces an element of public information that is somewhat at odds with the motivation of private monitoring as a robustness test" to the lack of common knowledge.

This paper incorporates all three approaches. First, the equilibrium strategy to show the folk theorem is phase-belief-free. That is, we see the repeated game as the repetition of long review phases. Each player has two strategies for the review phase; one is generous to the opponent and the other is harsh to the opponent.[10] At the beginning of each review phase, for each player, both generous and harsh strategies are optimal conditional on any realization of the opponents' histories. Between review phases, each player can change the opponent's continuation payoff from the next review phase by changing the transition probability between the two strategies, without considering the other players' histories. This equilibrium is immune to coordination failure at the beginning of each phase and gives us freedom to control the continuation payoffs.

Second, however, the belief-free property does not hold except at the beginning of the phases. Hence, we consider each player's statistical inference about the opponents' past histories as in the belief-based approach within each phase.

Finally, in our equilibrium, to coordinate the play in the middle of the phase, the players do communicate. We offer sufficient conditions with which this message exchange can

[10]As will be seen in Section 1.6, for a game with more than two players, one of player i's strategies is generous to player $i+1$ and the other is harsh to player $i+1$. In addition, players $-(i, i+1)$'s payoffs are constant regardless of which strategy player i picks from the two.

be done with their actions, that is, without assuming any explicit communication device. The difficulty to communicate via actions is that, since the players need to infer the opponents' messages from their private histories, common knowledge about the past messages no longer exists. One of our methodological contributions is to offer a systematic way to replace the public communication with message exchange via actions in general monitoring by overcoming the lack of common knowledge.

The paper is organized as follows: in the rest of Chapter 1, we introduce the model, states the assumptions and the main result, and offer the overview of the proof. After that, we relate the infinitely repeated game to a finitely repeated game with a "reward function" and derives sufficient conditions on the finitely repeated game to show the folk theorem in the infinitely repeated game. The remaining parts of the paper are devoted to the proof of the sufficient conditions.

In Chapter 2, we show the sufficient conditions in the two-player prisoners' dilemma, assuming special forms of communication. Given this chapter, we are left to extend the folk theorem to the general two-player game and the general N-player game with $N \geq 3$ and dispense with the special forms of communication. In Chapter 3, we summarize what new assumptions are sufficient for each extension. In the following chapters, we offer the proof: in Chapters 4 and 5, we extend the result to the general two-player game and the general N-player game, respectively, with the special forms of communication. In Chapters 6 and 7, we dispense with the special forms of communication in the two-player game and N-player game, respectively.

1.2 Model

1.2.1 Stage Game

The stage game is given by $\left\{I, \{A_i, Y_i, U_i\}_{i\in I}, q\right\}$. $I = \{1, \ldots, N\}$ is the set of players, A_i with $|A_i| \geq 2$ is the finite set of player i's pure actions, Y_i is the finite set of player i's private signals, and U_i is the finite set of player i's ex-post utilities. Let $A \equiv \prod_{i \in I} A_i$, $Y \equiv \prod_{i \in I} Y_i$ and $U \equiv \prod_{i \in I} U_i$ be the set of action profiles, signal profiles and ex post utility profiles, respectively.

In every stage game, player i chooses an action $a_i \in A_i$, which induces an action profile $a \equiv (a_1, \ldots, a_N) \in A$. Then, a signal profile $y \equiv (y_1, \ldots, y_N) \in Y$ and an ex post utility profile $\tilde{u} \equiv (\tilde{u}_1, \ldots, \tilde{u}_N) \in U$ are realized according to a joint conditional probability function $q(y, \tilde{u} \mid a)$.

Following the convention in the literature, we assume that $\tilde{u}_i$ is a deterministic function of a_i and y_i so that observing the ex post utility does not give any further information than (a_i, y_i). If this were not the case, then we could see a pair of a signal and an ex post utility, $(y_i, \tilde{u}_i)$, as a new signal.

Player i's expected payoff from $a \in A$ is the ex ante value of $\tilde{u}_i$ given a and is denoted by $u_i(a)$. For each $a \in A$, let $u(a)$ represent the payoff vector $\{u_i(a)\}_{i \in I}$.

In this paper, we only consider independent mixture of actions, that is, $\Delta(A)$ is the set of independent mixed strategies $\Delta(A) = \prod_{i \in I} \Delta(A_i)$ in the stage game and by $\alpha \in \Delta(A)$, each player i takes $a_i \in A_i$ with probability $\alpha_i(a_i)$. $\Delta(A_{-i})$ and $\alpha_{-i} \in \Delta(A_{-i})$ are similarly defined.

1.2.2 Repeated Game

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

1.3 Assumptions

In this section, we state assumptions. First, we assume the full dimensionality condition. Let $F \equiv \text{co}(\{u(a)\}_{a\in A})$ be the set of feasible payoffs. The minimax payoff for player i is

$$v_i^* \equiv \min_{\alpha_{-i}\in\Delta(A_{-i})} \max_{a_i\in A_i} u_i(a_i, \alpha_{-i}).$$

In addition, let α_{-i}^* be a minimaxing strategy against player i in the stage game.

Then, the set of feasible and individually rational payoffs is given by $F^* \equiv \{v \in F : v_i \geq v_i^* \text{ for all } i\}$. We assume the full dimensionality of F^*.

Assumption 1 The stage game payoff structure satisfies the full dimensionality condition: $\dim(F^*) = N$.

Second, we state assumptions on the signal structure. In the proof of the folk theorem, we proceed in the following steps: first, we show the folk theorem, assuming the availability of special forms of cheap talk. Second, we dispense with the special forms of cheap talk so that all the communication is done by actions.

Hence, in Section 1.3.1, we state assumptions that we use whether or not the special forms of cheap talk are available. Then, in Section 1.3.2, we state what special forms of cheap talk we assume in the first step of the proof. Finally, Section 1.3.3 states a condition on the cardinality of each player's support of signals under which we can generically dispense with all forms of cheap talk assumed in Section 1.3.2. Exact conditions about the signal distributions to dispense with cheap talk are somewhat complicated and explained in Chapter 3.

In summary, the folk theorem holds if (i) the assumptions in Section 1.3.1 are satisfied and either (ii-a) the special forms of cheap talk in Section 1.3.2 are available or (ii-b) the assumptions in Chapter 3 are satisfied. Further, the assumptions in (ii-b) are generic if the cardinality of each player's support of signals satisfies Assumption 6 in Section 1.3.3.

1.3.1 Common Assumptions

First, we assume that full support for monitoring:

Assumption 2 For all $y \in Y$ and $a \in A$, $q(y \mid a) > 0$.

By Sekiguchi (1997), with this assumption, sequential equilibria are realization equivalent to Nash equilibria. Hence, for the rest of the paper, we concentrate on Nash equilibria.

Second, we assume that, for any pair of players (i, j), given any action profile $a \in A$, each player j can statistically identify player i's deviation. Let $q_j(\tilde{a}_i, a_{-i}) \equiv (q_j(y_j \mid \tilde{a}_i, a_{-i}))_{y_j}$ be the vector expression of the conditional distribution of player j's signals given $\tilde{a}_i, a_{-i}$. We assume that all the vectors $q_j(\tilde{a}_i, a_{-i})$ with $\tilde{a}_i \in A_i$ are linearly independent:

Assumption 3 For any $i, j \in I$ and $a \in A$, all the vectors $q_j(\tilde{a}_i, a_{-i})$ with $\tilde{a}_i \in A_i$ are linearly independent.

Third, with more than two players, for any trio (i, n, j), given any action profile $a \in A$, player j can statistically identify which of players i and n is more suspicious about deviations. We assume that all the vectors $q_j(\tilde{a}_i, a_{-i})$ with $\tilde{a}_i \in A_i$ and $q_j(\tilde{a}_n, a_{-n})$ with $\tilde{a}_n \in A_n, \tilde{a}_n \neq a_n$ are linearly independent:

Assumption 4 For $N \geq 3$, for any $i, n, j \in I$ and $a \in A$, $q_j(\tilde{a}_i, a_{-i})$ with $\tilde{a}_i \in A_i$ and $q_j(\tilde{a}_n, a_{-n})$ with $\tilde{a}_n \in A_n$ and $\tilde{a}_n \neq a_n$ are linearly independent.

Note that this is the same as pairwise identifiability condition for each a in Fudenberg, Levine, and Maskin (1994).

Fourth, we assume that, for each player i, for player j whose index is defined as

$$j = \begin{cases} i - 1 \text{ (right before player } i) & \text{if } i \neq 1, \\ 2 \text{ (right after player } i) & \text{if } i = 1, \end{cases}$$

there exists player j's mixed strategy $\hat{\alpha}_j$ in the stage game such that, given the other players' histories $a_{-(i,j)}, y_{-(i,j)}$, different (a_i, y_i) has different information about (a_j, y_j):

Assumption 5 For any $i \in I$, with

$$j = \begin{cases} i-1 & \text{if } i \neq 1, \\ 2 & \text{if } i = 1, \end{cases}$$

there exists $\hat{\alpha}_j \in \Delta(A_j)$ such that, for all $a_{-(i,j)} \in A_{-(i,j)}$, $y_{-(i,j)} \in Y_{-(i,j)}$, $a_i, a_i' \in A_i$ and $y_i, y_i' \in Y_i$, if $(a_i, y_i) \neq (a_i', y_i')$, then

$$\mathbb{E}\left[\mathbf{1}_{a_j,y_j} \mid y_i, y_{-(i,j)}, a_i, a_{-(i,j)}, \hat{\alpha}_j\right] \neq \mathbb{E}\left[\mathbf{1}_{a_j,y_j} \mid y_i', y_{-(i,j)}, a_i', a_{-(i,j)}, \hat{\alpha}_j\right]. \tag{1.1}$$

1.3.2 Assumptions about Cheap Talk

We assume the following two forms of cheap talk are available until Chapter 5.

Perfect Cheap Talk We first assume the availability of perfect cheap talk. When a player sends a message m via perfect cheap talk, the other players observe m directly and m becomes common knowledge.

This communication is (i) cheap (not directly payoff-relevant), (ii) instantaneous and (iii) public and perfect (it generates the same signal as the message to each player).

Error-Reporting Noisy Cheap Talk Between j and i with Precision $p \in (0,1)$ We second assume that, for each pair of players j and i with $j \neq i$, player j has an access to the following special form of cheap talk named "error-reporting noisy cheap talk with precision p" with precision $p \in (0,1)$ to send a binary message $m \in \{G, B\}$ to player i.[11]

Intuitively, this communication is (i) cheap (not directly payoff-relevant), (ii) instantaneous but (iii) private and noisy (it generates a private signal to player i that can be different from the original message of player j). In addition, as the name suggests, (iv) when player

[11]Except for Chapter 6, p is always equal to $\frac{1}{2}$.

i's signal is wrong, the "error is reported" to player $i-1 \pmod N$ with a high probability. As we will see in Section 1.6, player $i-1$ is a "controller" of player i's payoff.

Formally, when player j sends m to player i via error-reporting noisy cheap talk with precision $p \in (0,1)$, it generates player i's private signal $f[i](m) \in \{G, B\}$ with the following probability:

$$\Pr(\{f[i](m) = f\} \mid m) = \begin{cases} p & \text{for all } (m, f) \text{ with } f = m, \\ 1-p & \text{for all } (m, f) \text{ with } f \neq m. \end{cases}$$

That is, $f[i](m)$ is correct with probability p but incorrect with probability $1-p$.

Given the original message m and player i's signal $f[i](m)$, it generates player $(i-1)$'s private signal $g[i-1](m) \in \{m, E\}$. Intuitively, if player i's signal is an error, then $g[i-1](m) = E$ with probability p, that is, "the error is reported" to the controller of player i with probability p:

$$\Pr(\{g[i-1](m) = E\} \mid m, f[i](m)) = \begin{cases} p & \text{for } (m, f[i](m)) \text{ with } f[i](m) \neq m, \\ 1-p & \text{for } (m, f[i](m)) \text{ with } f[i](m) = m. \end{cases}$$

Finally, player $j-1$ (the controller of player j, the sender) observes a private signal $f_2[j-1](m) \in \{G, B\}$ and player $i-1$ (the controller of player i, the receiver) observes a private signal $g_2[i-1](m) \in \{G, B\}$. We assume that there exists $\eta > 0$ such that

- with arbitrarily fixed η, for sufficiently large p, even after observing any $f_2[j-1](m)$, player i still believes that if $f[i](m) \neq m$, then $g[i-1](m) = E$ with a high probability.[12]

  Formally, with arbitrarily fixed η, for sufficiently large p, $f_2[i](m)$ and $g_2[j](m)$ are very imprecise signals compared to $f[i](m)$ and $g[j](m)$: for all $m \in \{G, B\}$, $f[i](m) \in$

[12] If $j-1 \neq i$, then player i does not observe $f_2[j-1](m)$.

$\{G, B\}$, $g[i-1](m) \in \{G, B\}$, $f_2 \in \{G, B\}$, and $g_2 \in \{G, B\}$,

$$\Pr\left(\{f_2\left[j-1\right](m) = f_2, g_2\left[i-1\right](m) = g_2\} \mid m, f[i](m), g\left[i-1\right](m)\right) \geq \eta; \tag{1.2}$$

- a pair $(f[i](m), f_2[j-1](m))$ contains some information about the other players' signals.

  Formally, for any $m \in \{G, B\}$, $g[i-1](m) \in \{G, B\}$, $f[i](m), f[i](m)' \in \{G, B\}$ and $f_2[j-1](m), f_2[j-1](m)' \in \{G, B\}$, if $(f[i](m), f_2[j-1](m)) \neq (f[i](m)', f_2[j-1](m)')$, then

$$\left\| \begin{array}{c} \mathbb{E}\left[\mathbf{1}_{g_2[i-1](m)} \mid m, g[i-1](m), f[i](m), f_2[j-1](m)\right] \\ -\mathbb{E}\left[\mathbf{1}_{g_2[i-1](m)} \mid m, g[i-1](m), f[i](m)', f_2[j-1](m)'\right] \end{array} \right\| \geq \eta. \tag{1.3}$$

  In this paper, we use the Euclidean norm. In general, for a random variable $x \in X$, we define $\mathbf{1}_x$ as a $|X| \times 1$ vector such that, if $x = \hat{x}$, the element corresponding to $\hat{x}$ is equal to one and the other elements are zero. For example,

$$\mathbf{1}_{g_2[i-1](m)} = \begin{cases} \begin{bmatrix} 1 \\ 0 \end{bmatrix} & \text{if } g_2[i-1](m) = G, \\ \begin{bmatrix} 0 \\ 1 \end{bmatrix} & \text{if } g_2[i-1](m) = B; \end{cases} \tag{1.4}$$

  and

- a pair $(g[i-1](m), g_2[i-1](m))$ contains some information about the other players' signals.

  Formally, for any $m, m' \in \{G, B\}$, $f[i](m) \in \{G, B\}$, $g[i-1](m), g[i-1](m)' \in \{G, B\}$ and $g_2[i-1](m), g_2[i-1](m)' \in \{G, B\}$, if $(m, g[i-1](m), g_2[i-1](m)) \neq$

$(m', g[i-1](m)', g_2[i-1](m)')$, then

$$\left\| \begin{array}{c} \mathbb{E}\left[\mathbf{1}_{f_2[j-1](m)} \mid m, g[i-1](m), g_2[i-1](m), f[i](m)\right] \\ -\mathbb{E}\left[\mathbf{1}_{f_2[j-1](m)} \mid m', g[i-1](m)', g_2[i-1](m)', f[i](m)\right] \end{array} \right\| \geq \eta. \tag{1.5}$$

We assume that all the signals are private and so

- player j knows only m;
- player i knows only $f[i]\,(m)$;
- player $i-1$ knows only $g[i-1]\,(m)$ and $g_2[i-1]\,(m)$; and
- player $j-1$ knows only $f_2[j-1](m)$.[13]

1.3.3 Assumptions about Dispensability of Cheap Talk

In Chapters 6 and 7, we show that both perfect cheap talk and error-reporting cheap talk are dispensable. Exact sufficient conditions for dispensability are stated in Chapter 3 and here, we state the assumptions about the cardinality of each player's support of signals under which the sufficient conditions in Chapter 3 are generic:

Assumption 6 The cardinality of each player's support of signals is sufficiently large: for any $i \in I$, we have

$$|Y_i| \geq |A_j| \text{ with } j \neq i \text{ if } N = 2, \tag{1.6}$$

$$|Y_i| \geq \max_{\substack{j \neq n \\ j,n \neq i}} |A_j| + |A_n| \text{ if } N \geq 3. \tag{1.7}$$

Intuitively speaking, when player j sends a message to player i by taking actions, with two players, player i needs to statistically infer player j's actions, which is generically possible

[13]If there is a player whose index appears multiple times, then we assume that the player knows all the signals of the players with that index. For example, if player j and player $i-1$ are the same player, she knows m, $g[i-1]\,(m)$ and $g_2[i-1]\,(m)$.

if (1.6) is satisfied. With more than two players, player i needs to statistically infer player j's action in such a way that another player $n \in -(i,j)$ cannot manipulate player i's inference, which is generically possible if (1.7) is satisfied.

1.4 Result

First, with Assumptions 1, 2, 3, 4 and 5, we can generically construct an equilibrium to attain any point in $\text{int}(F^*)$ if perfect cheap talk and error-reporting noisy cheap talk are available.

Theorem 7 If Assumptions 1, 2, 3, 4 and 5 are satisfied and perfect cheap talk and error-reporting noisy cheap talk are available, then the folk theorem holds: for any $v \in \text{int}(F^*)$, there exists $\bar{\delta} < 1$ such that, for all $\delta > \bar{\delta}$, $v \in E(\delta)$.

In Chapter 3, we provide sufficient conditions to dispense with both perfect cheap talk and error-reporting noisy cheap talk. These sufficient conditions are generic if Assumption 6 is satisfied:

Theorem 8 If Assumption 6 is satisfied, then both perfect cheap talk and error-reporting noisy cheap talk are generically dispensable in Theorem 7: if Assumptions 1 and 6 are satisfied, then for a generic signal distribution $\{q(y \mid a)\}_{y,a}$, for any $v \in \text{int}(F^*)$, there exists $\bar{\delta} < 1$ such that, for all $\delta > \bar{\delta}$, $v \in E(\delta)$.

Assumptions 3, 4 and 5 are also generic under Assumption 6 and so omitted in the statement of Theorem 8.

From now on, we arbitrarily fix $v \in \text{int}(F^*)$ and construct an equilibrium to support v in a Nash equilibrium.

1.5 Overview of the Argument

This section provides an intuitive explanation for our construction. Following Hörner and Olszewski (2006), we see a repeated game as repetition of T_P-period review phases. T_P will be formally defined later. In Section 1.5.1, we explain that our equilibrium is "phase-belief-free" and how it makes our equilibrium immune to coordination failure at the beginning of each phase. Section 1.5.2 offers the basic structure of the review phase.

To explain the details of the review phase, we assume perfect and error-reporting noisy cheap talk is available as explained in Section 1.3.2. With these communication devices, in Sections 1.5.3, 1.5.4 and 1.5.5, we offer the detailed explanation of the review phase.

Finally, we explain how to dispense with the communication devices in Section 1.5.6.

1.5.1 Phase-Belief-Free

As Hörner and Olszewski (2006), the equilibrium is phase-belief-free. Each player i has two T_P-period-finitely-repeated-game strategies, denoted by $\sigma_i(G)$ and $\sigma_i(B)$. Since T_P-period-finitely-repeated-game strategies are not fully specified strategies in the infinitely repeated game, we call them "action plans" in the context of the infinitely repeated game. At the beginning of each review phase, for each player i, independently of her history, any continuation strategy that adheres to one of the two action plans $\sigma_i(G)$ and $\sigma_i(B)$ in the review phase is optimal. We say that player i taking $\sigma_i(x_i)$ with $x_i \in \{G, B\}$ in the review phase is "in state $x_i \in \{G, B\}$."

Intuitively speaking, $\sigma_i(G)$ is a "generous" action plan that gives a high payoff to player $i+1 \pmod N$ who takes either $\sigma_{i+1}(G)$ or $\sigma_{i+1}(B)$, regardless of the other players' state profile $x_{-(i,i+1)} \in \{G, B\}^{N-2}$. On the other hand, $\sigma_i(B)$ is a "harsh" action plan that gives a low payoff to player $i+1$ regardless of player $(i+1)$'s action plans (including those different from $\sigma_{i+1}(G)$ and $\sigma_{i+1}(B)$) and $x_{-(i,i+1)}$. Hence, player $(i-1)$'s state controls player i's value regardless of $x_{-(i-1)}$, replacing i with $i-1$ in the previous two sentences. Since these two action plans are optimal at the beginning of the next phase, it is up to player $i-1$ whether player $i-1$ will take $\sigma_{i-1}(G)$ or $\sigma_{i-1}(B)$ in the next phase. Therefore, player $i-1$ with $\sigma_{i-1}(G)$ in the current phase can freely reduce player i's continuation payoff from the next review phase by transiting to $\sigma_{i-1}(B)$ with a higher probability while player $i-1$ with $\sigma_{i-1}(B)$ can freely increase player i's continuation payoff by transiting to $\sigma_{i-1}(G)$ with a higher probability.[14] When we say "strategies" in the context of the infinitely repeated game, they include the specification of the state transition probability. In summary, we do not need to consider player $(i-1)$'s incentive to punish player i after a "bad history" in state G or to reward player i after a "good history" in state B.

1.5.2 Structure of the Review Phase

The basic structure of the review phase is summarized as follows. At the beginning of the review phase, the players communicate a state profile $x \in \{G, B\}^N$. This communication stage is named the "coordination block" since the players try to coordinate on x. The details will be explained in Section 1.5.3.

Based on the result of the coordination block, the players play the finitely repeated game for many periods. This step consists of multiple "review rounds." The details will be explained in Section 1.5.5.

Finally, at the end of the phase, the players communicate the histories in the coordination block and review rounds. This stage is named the "report block" since the players report the histories in the review rounds. The role of this communication will be explained in Section 1.5.4.

[14] Here, the changes in the continuation payoffs are measured by the differences between player i's ex ante value given x_{i-1} at the beginning of the review phase and the ex post value at the end of the review phase after player $i-1$ observes the history in the phase. See Section 1.6 for the formal definition.

For example, if player $i-1$ with $x_{i-1} = G$ does not reduce player i's continuation value, then it means that the state of player $i-1$ in the next review phase is G with probability one, so that the ex post value is the same as the ex ante value.

1.5.3 Coordination Block

The role of the coordination block is to coordinate on x as in Hörner and Olszewski (2006). With the perfect cheap talk, each player tells the truth about her own state x_i and the state profile $x \in \{G, B\}^N$ becomes common knowledge. In the review rounds, based on x, the players play $\alpha(x)$ with a high probability on the equilibrium path. Intuitively, $\alpha(x)$ is the mixed action profile taken in the "usual" histories when the state profile is x. See Section 1.6 for the formal definition of $\alpha(x)$.

1.5.4 Report Block

We introduce the report block where the players communicate the histories in the coordination block and review rounds. This communication enables us to concentrate on ε-equilibrium until the end of the last review round. Suppose that we have constructed a strategy profile which is ε-equilibrium at the end of the last review round if we neglect the report block. We explain how to attain the exact equilibrium by using the report block.

Suppose that the perfect cheap talk and public randomization are available. Each player i is picked by the public randomization with probability $\frac{1}{N}$.[15] The picked player i sends the whole history in the coordination block and review rounds (denoted by h_i^{main}) to player $i-1$. That is, h_i^{main} is player i's history from the beginning of the coordination block to the end of the last review round.

Assume that player i always tells the truth about h_i^{main}. To make ε-equilibrium exact equilibrium, player $i-1$ changes the continuation payoff of player i so that, for all t, after h_i^t, it is exactly optimal to take an action prescribed by the equilibrium strategy. Since the original strategy profile was ε-equilibrium with arbitrarily small ε, this can be done by slightly changing the continuation strategy based on h_{i-1}^{main} and h_i^{main}.[16]

[15] For $N \geq 3$, the formal procedure is slightly different. See Section 5.9.

[16] With more than two players, player $i-1$ also needs to know the histories of players $-(i-1, i)$. So that players $-(i-1, i)$ can send their histories to player $i-1$, we introduce another communication stage after the report block, named the "re-report block." Since this

The remaining task with the perfect cheap talk and public randomization is to show the incentive to tell the truth about h_i^{main}. Intuitively, with defining a linear space and norm properly for the history, player $i-1$ punishes player i proportionally to $\left\| h_{i-1}^{\text{main}} - \mathbb{E}\left[h_{i-1}^{\text{main}} \mid \hat{h}_i^{\text{main}} \right] \right\|^2$ with $\hat{h}_i^{\text{main}}$ being the reported history. The optimal report $\hat{h}_i^{\text{main}}$ to minimize the expected punishment $\mathbb{E}\left[\left\| h_{i-1}^{\text{main}} - \mathbb{E}\left[h_{i-1}^{\text{main}} \mid \hat{h}_i^{\text{main}} \right] \right\|^2 \mid h_i^{\text{main}} \right]$ is to tell the truth: $\hat{h}_i^{\text{main}} = h_i^{\text{main}}$.[17] Since the adjustment for exact equilibrium is small, the small punishment is enough to incentivize player i to tell the truth. Therefore, the total changes in the continuation payoff based on the report block do not affect the equilibrium payoff.

1.5.5 Review Rounds

Between the coordination block and the report block, the players play a T-period "review round" for L times. Here, $L \in \mathbb{N}$ is a fixed integer that will be determined in Section 2.6, and

$$T = (1-\delta)^{-\frac{1}{2}}$$

so that

$$T \to \infty \text{ and } \delta^{LT} \to 1 \text{ as } \delta \to 1. \tag{1.8}$$

Intuitively, if the discount factor is large, T is sufficiently long to aggregate information efficiently and, at the same time, the discounting over T periods is negligible since δ^T goes to unity. Throughout the paper, we neglect the integer problem since it is handled by replacing each variable s that should be an integer with $\min_{\substack{n \in \mathbb{N} \\ n \geq s}} n$.

The reason why we have T periods in each review round is to aggregate private signals for many periods to get precise information as in Matsushima (2004).[18] There are two reasons why we have L review rounds. The first reason is new: as we will explain, the signals of the

information sent by players $-(i-1, i)$ in the re-report block is used only to control player i's continuation payoff, the truthtelling incentive for players $-(i-1, i)$ is trivially satisfied. See Section 5.10.

[17] Note that this logic is the same as we show the consistency of generalized-method-of-moments estimators.

[18] See also Radner (1985) and Abreu, Milgrom, and Pearce (1991).

players can be correlated while Matsushima (2004) assumes that the signals are conditionally independent. To deal with correlation, we need multiple review rounds.

The second reason is the same as Hörner and Olszewski (2006). If we replace each period of Hörner and Olszewski (2006) with a T-period review round, then we need a sufficiently large number of review rounds so that a deviator should be punished sufficiently long to cancel out the gains in the instantaneous utility from deviation.

Below, we offer a more detailed explanation of the review rounds. In Section 1.5.5.1, we concentrate on the first role of the L rounds by considering the case where the block of Hörner and Olszewski (2006) has one period, that is, the stage game is the two-player prisoners' dilemma. We will explain the general two-player game and the general more-than-two-player game in Sections 1.5.5.2 and 1.5.5.3, respectively, where the second role of the L rounds is important.

Whenever we consider the two-player case and we say players i and j, we assume that player j is player i's (unique) opponent unless otherwise specified.

1.5.5.1 The Two-Player Prisoners' Dilemma

In the two-player prisoners' dilemma, we consider player i's incentive to take $\sigma_i(G)$ when player j takes $\sigma_j(G)$. The other combinations of (x_i, x_j) are symmetric. Remember that since x is communicated via perfect cheap talk, x is common knowledge.

So that $\sigma_i(G)$ is generous to player j, player i needs to take cooperation with ex ante high probability. On the other hand, player j can reduce player i's continuation payoff from the next review phase based on her history within the current review phase (see the explanation of the phase-belief-free in Section 1.5.1).

To incentivize player i to take C_i, player j needs to punish player i after observing a suspicious history. On the other hand, for efficiency, player j should not punish player i if player i cooperates. To satisfy these two requirements simultaneously in a noisy environment, player j needs to aggregate information over long T periods (review round).

Information Aggregation Let us explain how player j aggregates information. Since Assumption 3 implies that player j can statistically identify player i's action, player j can map her history in each period into a real number so that

$$\mathbb{E}\left[\pi_i[\alpha]\left(y_j\right) \mid \alpha_j, C_i\right] - \mathbb{E}\left[\pi_i[\alpha]\left(y_j\right) \mid \alpha_j, D_i\right] = u_i\left(\alpha_j, D_i\right) - u_i\left(\alpha_j, C_i\right), \tag{1.9}$$

where α_j is player j's equilibrium mixed action in that period and α_i is the mixed action that player j expects player i to take. Intuitively, conditional on α_j, after observing a "good" signal y_j which occurs more likely after player i's cooperation, player j gives a high point $\pi_i[\alpha]\left(y_j\right)$ while after observing a "bad" signal y_j which occurs more likely after player i's defection, player j gives a low point $\pi_i[\alpha]\left(y_j\right)$, so that the expected gain in points from cooperation cancels out the loss in instantaneous utilities. We normalize $\pi_i[\alpha]\left(y_j\right)$ by adding or subtracting a constant so that

$$\mathbb{E}\left[\pi_i[\alpha]\left(y_j\right) \mid \alpha_j, C_i\right] = 0. \tag{1.10}$$

Further, let $\bar{u}$ be the maximum absolute value of the points:

$$\bar{u} = \max_{j,\alpha,y_j} \left|\pi_i[\alpha]\left(y_j\right)\right| > 0.$$

As we will see in Lemma 12, this $\bar{u}$ is well defined.

Recall that we have L review rounds. For each lth review round, player j aggregates $\pi_i[\alpha(l)]\left(y_{j,t}\right)$ and creates player j's score about player i:

$$X_j(l) = \sum_{t:\ l\text{th review round}} \pi_i[\alpha\left(l\right)]\left(y_{j,t}\right). \tag{1.11}$$

Here $\alpha_j(l)$ is player j's equilibrium mixed action in the lth review round and $\alpha_i(l)$ is the mixed action that player j expects player i to take (as will be seen, the players take *i.i.d.* mixed actions in each review round).

Conditional Independence Following Matsushima (2004), assume that player i's signals were independent of player j's signals conditional on any action profile a.

With $x_i = G$, let $\alpha_i(x) = (1-2\rho)C_i + 2\rho D_i$ with a small $\rho > 0$ be an action plan that takes C_i with high probability $1-2\rho$. Intuitively, player j wants to incentivize player i to take $\alpha_i(x)$ by aggregating information over the review round and the punishment should be small if player i takes C_i frequently.

This is done as follows. Let the change in player i's continuation payoff be equal to

$$\left\{-2\bar{u}T + \sum_{l=1}^{L} X_j(l)\right\}_{-}, \tag{1.12}$$

where, in general, $\{X\}_{-}$ is equal to X if $X \leq 0$ and 0 otherwise. That is, player j adds the scores from all the review rounds. From (1.9), the expected score decreases when player i takes defection and this cancels out the gain in instantaneous utilities from defection. Hence, as long as $\sum_{l=1}^{L} X_j(l) \leq 2\bar{u}T$, player i is indifferent between cooperation and defection.

Hence, for ε-equilibrium, we are left to show that player i after any history believes that $\sum_{l=1}^{L} X_j(l) \leq 2\bar{u}T$ with a high probability and that efficiency is not destroyed.

From (1.9) and (1.10), the expected increase in the score in each period (that is, the expected point) is non-positive. Therefore, by the law of large numbers, with $\bar{u} > 0$ and $L \in \mathbb{N}$, for sufficiently large T, player i believes that $\sum_{l=1}^{L} X_j(l) \leq 2\bar{u}T$ with a high probability. Since player i's signals are independent of player j's signals, player i cannot update any information about player j's signals about player i from player i's history. Therefore, this statement is correct after any history of player i.

At the same time, since the expected value of $\pi_i[\alpha(l)](y_j)$ under cooperation is 0, for sufficiently small ρ, the ex ante value of $X_j(l)$ is close to 0 and so the ex ante reduction of the continuation payoff is close to $-2\bar{u}T$. Since there are L T-period review rounds, per-period efficiency loss is equal to $2\bar{u}/L$, which can be arbitrarily small for large L. Therefore, we are done.

Conditional Dependence Now, we dispense with conditional independence. That is, player i's signals and player j's signals can be correlated arbitrarily. Since the expected score is 0 under a constant cooperation, to prevent an inefficient punishment, player j cannot punish player i after the score is excessively high (in the above example, more than $2\bar{u}T$). On the other hand, if the signals are correlated, then it happens with a positive probability that player i believes that, judging from her own history and correlation, player j's score about player i has been excessively high already. Then, player i wants to start to defect.

More generally, it is impossible to create a punishment schedule that is approximately efficient and that at the same time incentivizes player i to cooperate after any history. Hence, we need to let player i's incentive to cooperate break down after some history. Symmetrically, player j also switches her own action after some history.

Intuitively, player i switches to a constant defection after player i's expectation of player j's score about player i is much higher than the ex ante mean. We want to specify exactly when each player i switches to a constant defection based on player i's expectation of player j's score about player i.

Reflective Learning Problem However, this creates the following problem: since player i switches her action based on player i's expectation of player j's score about player i, player i's action reveals player i's expectation of player j's score about player i. Since both "player i's expectation of player j's score about player i" and "player i's score about player j" are calculated from player i's history, player j may want to learn "player i's expectation of player j's score about player i" from "player j's signals about player i's action." If so, player j's decision of actions depends also on player j's expectation of player i's expectation of player j's score about player i. Proceeding one step further, player i's decision of actions depends on player i's expectation of player j's expectation of player i's expectation of player j's score about player i. This chain of "reflective learning" continues infinitely.

Error-Reporting Noisy Cheap Talk Cuts off the Reflecting Learning We want to construct an equilibrium that is not destroyed by the reflective learning. From the discussion of the report block, we can focus on ε-equilibrium. This means that, to verify an equilibrium, it is enough to show that each player believes that her action is optimal with a high probability (not probability one). To prevent the reflective learning, we take advantage of this "ε slack" in ε-equilibrium and the noise in the error-reporting noisy cheap talk explained in Section 1.3.2.

The basic structure is as follows. Recall that we have L T-period review rounds. At the beginning of each lth review round, in a normal history (we will define the "normal" history later), player j decides to take one of the following three mixed action plan

$$\alpha_j(l) = \begin{cases} \bar{\alpha}_j(x) \equiv (1-\rho) C_i + \rho D_i, \\ \alpha_j(x) \equiv (1-2\rho) C_i + 2\rho D_i, \\ \underline{\alpha}_j(x) \equiv (1-3\rho) C_i + 3\rho D_i, \end{cases} \tag{1.13}$$

with probability

$$\begin{cases} \eta/2, \\ 1-\eta, \\ \eta/2, \end{cases}$$

respectively with small $\rho, \eta > 0$. Once player j decides $\alpha_j(l)$, player j takes an action according to $\alpha_j(l)$ *i.i.d.* within the lth review phase.

Note that player j takes the same action plan $\alpha_j(x)$ as in the case with conditional independence with a high probability. If $\alpha_j(l) \neq \alpha_j(x)$, then player j makes player i indifferent between any action profile sequence from the lth review round. This can be done by changing the transition probability to $x_j = B$ at the beginning of the next review phase. See the definition of $\pi_i^{x_j}[\alpha_j(l)](y_j)$ in Section 2.3.

At the end of each lth review phase, player j's history is partitioned into the following two:

- player j's score about player i has been "not erroneously high": $X_j(\tilde{l}) \leq \frac{\bar{u}}{L}T$ for all $\tilde{l} = 1, ..., l$; or
- player j's score about player i has been erroneously high: there exists $\tilde{l} = 1, ..., l$ with $X_j(\tilde{l}) > \frac{\bar{u}}{L}T$.

In the former case, we say $\lambda_j(l+1) = G$ and in the latter case, we say $\lambda_j(l+1) = B$. Intuitively speaking, if $\lambda_j(l+1) = G$, then player i will be indifferent between C_i and D_i in the $(l+1)$th review round and so $\alpha_i(l)$ is optimal. On the other hand, if $\lambda_j(l+1) = B$, then player i should switch to a constant defection from the $(l+1)$th review round.

Player j informs player i of $\lambda_j(l+1)$ by sending $\lambda_j(l+1)$ to player i by the error-reporting noisy cheap talk with precision $p = 1 - \exp(-T^{\frac{1}{2}})$. Player i observes $f[i](\lambda_j(l+1))$. If there is an error, that is, if $f[i](\lambda_j(l+1)) \neq \lambda_j(l+1)$, then the error is reported to player j, that is, $g[j](\lambda_j(l+1)) = E$ with high probability $1 - \exp(-T^{\frac{1}{2}})$. If the error is reported, then player j will make player i indifferent between any action profile from the $(l+1)$th review round. Again, see the definition of $\pi_i^{x_j}[\alpha_j(l)](y_j)$ in Section 2.3 for how to do this. Note that this change of player j's continuation strategy does not affect player i's incentive since the probability of an error and the probability of error-reporting are independent of player j's message.

Consider player i's incentive at the end of the lth review round, calculating the optimal action in the $(l+1)$th review round. We partition player i's history into two classes: normal and abnormal.

We say player i's history at the end of the lth review round is normal if, for all the previous rounds $\tilde{l} = 1, ..., l$,

1. player i (symmetrically to player j) picks $\alpha_i(\tilde{l}) = \alpha_i(x)$;

2. the realized frequency of player i's action in the $\tilde{l}$th review round is actually close to $\alpha_i(x)$; and

3. player i's signal frequency during the periods when player i takes cooperation in the $\tilde{l}$th review round is close to the affine full of player i's signal frequency with respect to player j's actions:

$$\text{aff}\left(\{q_i\left(C_i, a_j\right)\}_{a_j \in A_j}\right).$$

As defined in Assumption 3, $q_i\left(\alpha\right) = \left(q_i(y_i \mid \alpha)\right)_{y_i}$ is player i's signal distribution under α.

Otherwise, we say player i's history is abnormal. If player i's history is abnormal, then player i will make player j indifferent between any action profile. See the definition of $\pi_j^{x_i}[\alpha_i(l)](y_i)$ in Section 2.3 for how to do this. Note that this change of player i's continuation strategy does not affect player j's incentive since whether player i's history is normal or not is not controllable for player j (Conditions 1 and 2 are determined by player i's mixture and Condition 3 takes the affine hull with respect to player j's action).

After the normal history, player i disregards player j's message about $\lambda_j(l+1)$ and keeps taking $\alpha_i(l+1)$ in the $(l+1)$th review round defined symmetrically to (1.13). The almost optimality of this action plan is explained as follows: roughly speaking, since player i has taken C_i very often in the previous rounds (see Conditions 1 and 2 of the normal history), player i can concentrate on periods when player i took C_i to infer player j's score about player i.

If player i's signal frequency is very close to the ex ante distribution under $\alpha_j(x)$, then player i's conditional expectation of player j's score about player i is also close to the ex ante mean of the score, that is, player i puts little belief on the event that $\lambda_j(l+1) = B$. Since the length of the review round is T, this believe is no more than $\exp(-\Theta(T))$.[19]

If player i's signal frequency is not close to the ex ante distribution under $\alpha_j(x)$, then it is likely that player j took $\bar{\alpha}_j(x)$ ($\underline{\alpha}_j(x)$, respectively) if the frequency is skewed toward $q_i(C_i, C_j)$ ($q_i(C_i, D_j)$, respectively) compared to $q_i(C_i, \alpha_j(x))$ since $\bar{\alpha}_j(x)$ ($\underline{\alpha}_j(x)$, respectively) takes C_j (D_j, respectively) more often than $\alpha_j(x)$. In this case, player j makes any action optimal for player i. Note that since player i's signal frequency is close to the affine hull of player i's signal distributions with respect to a_j, whenever player i's signal frequency is not close to the ex ante distribution under $\alpha_j(x)$, it should be skewed either toward $q_i(C_i, C_j)$ or $q_i(C_i, D_j)$.

Hence, in both cases, player i is almost indifferent between C_i and D_i. Note that the above discussion is only before player i learns about player j's history from player j's continuation action plan.

Before proceeding to the learning problem, let us specify player i's action plan after the abnormal history (that is, at least one of Conditions 1, 2 and 3 is violated). In this case, player i obeys player j's message: if $f[i](\lambda_j(l+1)) = G$, then player i keeps taking $\alpha_i(l+1)$ as prescribed by (1.13) in the $(l+1)$th review round. If $f[i](\lambda_j(l+1)) = B$, then player i switch to a constant defection from the $(l+1)$th review round.

The remaining questions are (i) how we can make sure that the reflective learning does not destroy an equilibrium, (ii) how we can incentivize player j to tell the truth about $\lambda_j(l+1)$, and (iii) whether efficiency is preserved.

Consider the first question. When player i obeys the message, the error is reported with a high probability (if any) and player j's action plan symmetrically defined to player i's action plan is independent of whether the error is reported or not. Hence, regardless of the learning about player j's continuation action plan, player i keeps a high belief on the event that, if there is an error in $f[i](\lambda_j(l+1))$, then the error is reported to player j.

[19] For a variable X_T which depends on T, we say $X_T = \exp(-\Theta(T))$ if and only if there exist $k_1, k_2 > 0$ such that $\exp(-k_1 T) \leq X_T \leq \exp(-k_2 T)$ for sufficiently large T.

When player i disregards the message, there are two channels for the learning about player j's score through player j's continuation action plan. The first one is player i's signals coming from player j's message about $\lambda_j(l+1)$ by the error-reporting noisy cheap talk. Since the order of noise, $\exp(-T^{\frac{1}{2}})$, is much larger than the original belief on $\lambda_j(l+1) = B$, $\exp(-\Theta(T))$, it is almost optimal for player i to disregard the message.

The second channel is through player j's reaction to player i's message. When player i learns that player j will play a constant defection, this means player j has obeyed player i's message, which means player j makes player i indifferent between any action profile sequence. Hence, learning this event is not a problem.

When player i learns that player j will play $\alpha_i(l+1)$ as prescribed by (1.13), the problematic history of player i is as follows: player i originally believes that player j's history should have been abnormal (for example, player i believes that player j took $\bar{\alpha}_j(x)$), player i has sent $\lambda_i(\tilde{l}+1) = B$ for some $\tilde{l} = 1, ..., l$ (that is, has told player j to defect), no error is reported to player i, but player j will play $\alpha_i(l+1)$ as prescribed by (1.13). If player j took $\bar{\alpha}_j(x)$ as player i believes, then player j should have obeyed the message. Hence, without an error, it would be inconsistent. There are two possibilities: player i's original belief about player j's history was wrong or although the error was not reported, player j's signal was wrong ($f[j](\lambda_i(\tilde{l}+1)) = G$) and player j has obeyed player i's message. The second event can happen with probability $\exp(-2T^{\frac{1}{2}})$, which is bigger than the original belief about the first event $\exp(-\Theta(T))$, and player i can attribute this inconsistency to the second case and adhere to player i's original belief about player j's history.

Therefore, due to the noise in the error-reporting noisy cheap talk and ε slack in ε-equilibrium, player i can neglect the learning from player j's continuation action plan.

Consider the second problem to incentivize player j to tell the truth about $\lambda_j(l+1)$. Remember that player i does not disregard the message only after the abnormal history. Hence, player i has made player j indifferent between any action profile sequence whenever player j's message matters, which implies player j is indifferent between player i cooperating and defecting. Hence, the truthtelling incentive is satisfied.

Finally, with high probability $1-2\rho$, player i takes $\alpha_i(l) = \alpha_i(x)$ for each lth review round. Then, by the law of large numbers, the realized frequency of player i's action is close to $\alpha_i(x)$. Again, by the law of large numbers, player i's signal frequency is close to the affine full of player i's signal frequency with respect to player j's actions. Hence, player i's history is normal and takes cooperation with a high probability. Further, an error and error-reporting do not happen with a high probability. Hence, efficiency is preserved.

Summary Let us intuitively summarize the equilibrium construction. Although the breakdown of cooperation after abnormal histories is inevitable, we need to verify that the reflective learning does not destroy the incentives.

To make it possible for player i to take an optimal action depending on player j's score about player i, player j informs player i of the optimal action via the error-reporting noisy cheap talk.

When player i calculates player i's optimal action, although player i's signal frequency is not close to the true distribution under $\alpha_j(x)$, as long as it is close to the affine hull of player i's signal distributions with respect to player j's action, player i believes that player j's action is not equal to $\alpha_j(x)$ and that player i is indifferent between any action. Therefore, player i disregards player j's message. Player i can neglect the learning from player j's continuation action plan since there is noise in the error-reporting noisy cheap talk and we can concentrate on almost optimality because of the report block.

When player i obeys player j's message, on the other hand, the error (if any) should be reported to player j and player j's continuation action plan is independent of whether the error is reported or not. Therefore, player i can obey the message without being worried about a mistake, neglecting the learning from player j's continuation action plan.

1.5.5.2 General Two-Player Game

Now, we consider the second role of L, that is, we consider the general two-player game where the block of Hörner and Olszewski (2006) has more than one period. We still concentrate on the two-player case.

Imagine that we replace each period in Hörner and Olszewski (2006) with a T-period review round. We need L review rounds so that, when player i uses the harsh strategy, regardless of player j's deviation, we can keep player j's value low enough. If player j deviates for a non-negligible part of a review round, then by the law of large numbers, player i can detect player j's deviation with a high probability. If player i minimaxes player j from the next review round after such an event, then player j can get a payoff higher than the targeted payoff only for one review round. With sufficiently long L, therefore, player j's average payoff from a review phase can be arbitrarily close to the minimax payoff.

A known problem to replace one period in Hörner and Olszewski (2006) with a review round is summarized in Remark 5 in their Section 5. Player i's optimal action in a round depends on player j's signals in the past rounds. Player i calculates the belief about player j's past signals at the beginning of the round and starts to take an action that is optimal from her belief. While player i observing signals in that round, since player j's actions depend on player j's signals in the past rounds, player i may realize that player j's actions are different from what player i expected from her belief about player j's signals. Then, player i needs to correct her belief about player j's past signals.

Realize that this is the same "reflective learning" problem as we have dealt with for $\lambda_j(l+1)$. Here, we will proceed as follows: first, as in (1.13), player j takes $\alpha_j(l) = \alpha_j(x)$ with a high probability. However, player j also takes action plans different from $\alpha_j(x)$ with a positive probability, which are comparable to $\bar{\alpha}_j(x)$ and $\underline{\alpha}_j(x)$ in the two-player prisoners' dilemma. In a general game, player j takes a minimaxing action plan with a positive probability in addition to those comparable to $\bar{\alpha}_j(x)$ and $\underline{\alpha}_j(x)$.

Second, at the end of the lth review round, when player i has a history such that

1. player i's history is normal as defined in Section 1.5.5.1 and

2. for some past $\tilde{l}$th review round with $\tilde{l} \leq 1, ..., l$, player i observes a signal frequency that is not close to the ex ante mean of player i's signal distribution under $\alpha_j(x)$,

then, player i will minimax player j from the $(l+1)$th review round with a high probability. However, there is a positive probability with which player i "forgives" player j and keeps taking $\alpha_i(l)$ defined in the first step. Player j's strategy is symmetrically defined. Importantly, since player j takes a minimaxing action plan in the first step, the support of player j's action plans in the next review round does not depend on whether player j's history satisfies Conditions 1 and 2.

Note that, whether or not player j deviates, player i's history is normal with a high probability as explained in Section 1.5.5.1. If Condition 2 is satisfied, then it is likely that player j took an action plan different from $\alpha_j(x)$. This is comparable to player i putting a high belief on $\bar{\alpha}_j(x)$ ($\underline{\alpha}_j(x)$, respectively) if the frequency is skewed toward $q_i(C_i, C_j)$ ($q_i(C_i, D_j)$, respectively) compared to $q_i(C_i, \alpha_j(x))$ in the prisoners' dilemma. If this belief is correct, then player j makes any action optimal for player i. Therefore, player i will punish player j by taking a minimax action plan from the next review round with a high probability.

We are left to verify that player i's learning about the optimal action from player j's continuation action plan does not change player i's incentive. The new learning in addition to the learning explained in Section 1.5.5.1 is whether player j will take a minimaxing action plan or not. However, as we have explained above, the support of player j's action plans in the next review round does not depend on whether player j will minimax player i or not as long as player j's history is normal. Hence, player i cannot update the belief so much.

1.5.5.3 General More-Than-Two-Player Game

Finally, we consider a general game with more than two players. There are two problems unique to a game with more than two players: first, if player i's state x_i is B, then player

$(i+1)$'s value should be low. Since player i is in the bad state, player i can only increase the continuation payoff of player $i+1$. That is, we cannot punish player $i+1$ by reducing the continuation payoff. Hence, players $-(i+1)$ need to minimax player $i+1$ if player $i+1$ seems to have deviated. With two players, player i is the only opponent of player $i+1$ ($i+1=j$ in the two-player game) and so it suffices for player i to unilaterally punish player $i+1$. Thus, the punishment explained in Section 1.5.5.2 works. On the other hand, with more than two players, we need to make sure that players $-(i+1)$ can coordinate on the punishment. This coordination can be done by communication among all the players about who will be punished at the end of each review round. See Chapter 5 for the details.

Second, there will be a new problem when we dispense with the perfect cheap talk in the coordination block. We will address this issue when we discuss dispensability of the perfect cheap talk in Section 1.5.6.2.

1.5.6 Dispensing with Special Communication Devices

We are left to dispense with the special communication devices introduced in Section 1.3.2. We first explain the dispensability in the two-player game and then proceed to the dispensability in the more-than-two-player game.

1.5.6.1 Two Players

Dispensing with the Perfect Cheap Talk for x We explain how to replace the perfect cheap talk for the coordination on x in the coordination block with messages via actions. We proceed in steps.

First, we replace the perfect cheap talk with the error-reporting noisy cheap talk. By exchanging messages by the error-reporting noisy cheap talk several times, each player i can construct the inference of x, denoted by $x(i)$. The important properties to establish are (i) $x(i)=x$ for all i with a high probability, (ii) the communication is incentive compatible, and (iii) after realizing that $x(i) \neq x(j)$, that is, after player i realizes that player i's inference is different from player j's inference, player i believes that player j should have realized that there was an error in the communication and that player j has made player i indifferent between any action profile sequence in all the review rounds with a high probability. This enables player i to stick to her own inference. See Chapter 6 for the details.

Dispensing with the Error-Reporting Noisy Cheap Talk Second, we replace the error-reporting noisy cheap talk with messages via actions. Given the discussion above, by doing so, we can dispense with the perfect cheap talk in the coordination block and the error-reporting noisy cheap talk in the review rounds.

Consider the situation where player j sends a binary noisy cheap talk message $m \in \{G, B\}$ to player i with precision $p = 1 - \exp(-T^k)$ with $k \in (0, 1)$ (in Section 1.5.5.1, $k = 1/2$). With two players, player $i - 1$ is equal to player j. Remember that the important properties that we use in Section 1.5.5.1 are (i) cheap, (ii) instantaneous, and (iii) precise with probability $1 - \exp(-\Theta(T^k))$: there exist $c_1, c_2, c_3 > 0$ such that, for sufficiently large T, (iii-a) $f[i](m) = m$ with probability no less than $1 - c_1$; (iii-b) if $f[i](m) \neq m$, then $g[j](m) = E$ with probability no less than $1 - \exp(-c_2 T^k)$; (iii-c) any signal pair can occur with probability no less than $\exp(-c_3 T^k)$.

Instead of the error-reporting noisy cheap, player j (sender) sends the message via actions: player j with message m determines $z_j(m) \in \{G, B, M\}$ such that

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

and player j takes

$$\alpha_j^{z_j(m)} \equiv \begin{cases} (1-\rho)\, a_j^G + \rho a_j^B & \text{if } z_j(m) = G, \\ (1-\rho)\, a_j^B + \rho a_j^G & \text{if } z_j(m) = B, \\ \frac{1}{2} a_j^G + \frac{1}{2} a_j^B & \text{if } z_j(m) = M \end{cases}$$

with $\rho < \frac{1}{2}$ for T^k period. That is, player j sends the "true" message $\alpha_j^{z_j(m)} = \alpha_j^m$ with high probability $1-\eta$. On the other hand, player j "tells a lie" with probability η: with probability $\eta/2$, player j sends the opposite message $z_j(m) = \{G, B\} \setminus \{m\}$ and with probability $\eta/2$, player j "mixes" two messages: $z_j(m) = M$ and $\alpha_j^M = \frac{1}{2}a_j^G + \frac{1}{2}a_j^B$. When player j tells a lie, player j makes player i indifferent between any action profile, which corresponds to the situation where the error is reported: $g[j](m) = E$.

Player i (receiver) takes some mixed action $\alpha_i^{\text{receive}}$. Player i needs to infer the message from her private history.

There are three difficulties: the message exchange is now (i) payoff-relevant, (ii) takes time and (iii) imprecise.

Since $T^k < T$ with $k \in (0, 1)$, the length of the communication is much shorter than that of the review rounds. Therefore, we can deal with the first difficulty by changing the continuation payoffs to cancel out the differences in instantaneous utilities. With $T^k < T$, this does not affect the equilibrium payoff, that is, the equilibrium payoff is mainly determined by instantaneous utilities and changes in the continuation payoff from the T-period review rounds. (ii) In addition, $T^k < T$ implies that the second difficulty does not affect the equilibrium payoff either.

(iii) We are left to consider the third difficulty. We want to create a mapping from player i's history to $f[i](m) \in \{G, B\}$ to preserve (iii-a), (iii-b) and (iii-c).

The basic intuition is as follows. Suppose that player i calculates the log likelihood between $z_j(m) = G$ and $z_j(m) = B$. If one of them is sufficient larger than the other, then player i infers that the one with the higher likelihood is the true message. If the log likelihoods for $z_j(m) = G$ and $z_j(m) = B$ are similar, then since the log likelihood is strictly concave, player i puts a high belief on the event that $z_j(m) = M$, which means $z_j(m) \neq m$ and player j makes player i indifferent. Therefore, pick $f[i](m) \in \{G, B\}$ arbitrarily.

Since player j takes all the possible $z_j(m)$ with probability at least $\eta/2$, the likelihood conditional on $m = G$ and that conditional on $m = B$ are close to each other. Therefore, the following inferences are well defined:

- if there exists $z_j \in \{G, B\}$ such that, for all $m \in \{G, B\}$, the likelihood of $z_j(m) = z_j$ is sufficiently higher than that of $z_j(m) = \{G, B\} \setminus \{z_j\}$ conditional on m, then player i infers $f[i](m) = z_j$; and
- otherwise, for all $m \in \{G, B\}$, conditional on m, player i puts a high belief on $z_j(m) = M$ and picks $f[i](m) \in \{G, B\}$ arbitrarily.

This satisfies (iii-a) and (iii-b) since (iii-a) $f[i](m) = z_j(m)$ with a high probability by the law of large numbers and (iii-b) conditional on true m, if $f[i](m) \neq m$, then it is very likely that $z_j(m) \neq m$ and that player j makes player i indifferent.

We are left to show that (iii-c) any signal pair can occur with probability no less than $\exp(-c_2 T^k)$ for some c_2. With full support of the distribution of signal profile y, this is true.

Dispensing with the Perfect Cheap Talk and Public Randomization in the Report Block We are left to dispense with the perfect cheap talk and public randomization in the report block about h_i^{main}.

First, we replace the perfect rich cheap talk to send h_i^{main} with perfect cheap talk that can send only a binary message. We attach a sequence of binary messages to h_i^{main}. To send h_i^{main}, player i sends the sequence of binary messages corresponding to h_i^{main}. Expecting that we will replace the perfect cheap talk with messages via actions, we make sure that the number of binary messages sent is sufficiently smaller than T. Otherwise, it would be impossible to replace the cheap and instantaneous talk with payoff-relevant and taking-time messages via actions. Since each period in each review round is *i.i.d.*, it suffices that player i reports how many times player i observes an action-signal pair (a_i, y_i) for each $(a_i, y_i) \in A_i \times Y_i$ for each review round. Hence, the cardinality of the relevant history is approximately $T^{L|A_i||Y_i|}$.

Since each message is binary, the number of binary messages necessary to send the relevant history is $\log_2 T^{L|A_i||Y_i|}$, which is much smaller than T.

Second, we dispense with the public randomization. Recall that we use the public randomization to determine who will report the history such that (i) ex ante (before the report block), every player has a positive probability to report the history, and that (ii) ex post (after the realization of the public randomization), there is only one player who reports the history.

To see why both (i) and (ii) are important, remember that the equilibrium strategy would be only ε-optimal without the adjustment based on the report block. Thus, to attain the exact optimality, it is important for each player in the review rounds to believe that the reward will be adjusted with a positive probability. Therefore, (i) is essential.

(ii) is important because, the logic to incentivize player i to tell the truth uses the fact that player i does not know h_j^{main} (again, with two players, player $i-1$ is player j). If player i could observe a part of player j's sequential messages which partially reveal h_j^{main} before finishing reporting h_i^{main}, then player i may want to tell a lie.

We show that the players use their actions and private signals to establish the properties (i) and (ii), without the public randomization.

Third, we replace the perfect binary cheap talk with noisy binary cheap talk. Before doing so, we explain what property of the communication is important in the report block. The role of the report block is for player j to adjust player i's continuation payoff so that $\sigma_i(G)$ and $\sigma_i(B)$ are both exactly optimal. Since this adjustment does not affect player j's payoff, while player i sends h_i^{main}, player j (the receiver) does not care about the precision of the message. On the other hand, if player i realizes that her past messages may not have transmitted correctly in the middle of sending a sequence of messages, then we cannot pin down player i's optimal action plan after that.

Therefore, we consider conditionally independent noisy cheap talk such that, when player i sends $m \in \{G, B\}$, player j receives a signal $f^{\text{ci}}[j](m) \in \{G, B\}$. The message transmits correctly, that is, $f^{\text{ci}}[j](m) = m$, with a high probability. Player i receives no information about $f^{\text{ci}}[j](m)$, so that player i can always believe that the message transmits correctly with a high probability. Then, the truthtelling is still optimal after any history.

Finally, we replace the conditionally independent noisy cheap talk with messages via repetition of actions. Although we do not assume conditional independence of signals *a priori* or do not assume that $2|Y_i| \leq |A_j||Y_j|$,[20] as long as the adjustment of the continuation payoff based on the messages is sufficiently small, we can construct a message exchange protocol such that the sender always believes that the message transmits correctly with a high probability. We defer the detailed explanation to Section 6.7.5.1 in Chapter 6.

1.5.6.2 More Than Two Players

With more than two players, we follow the same step as in the two-player case to dispense with the communication devices. Each step is the same as in the two-player case with player j replaced with player $i-1$ except for the following two differences: first, how to replace the perfect cheap talk in the coordination block with the noisy cheap talk and second, how to make sure that the players other than a sender and a receiver do not have an incentive to manipulate the communication by changing their actions.

Recall that player i informs the other players $-i$ of x_i in the coordination block. With two players, there is only one receiver of the message. On the other hand, with more than two players, there are more than one receivers of the message. If some players infer x_i is G while the others infer x_i is B, then the action that will be taken with a high probability in the review rounds may not be included in $\{a(x)\}_x$. Since we do not have any bound on player i's payoff in such a situation, it might be of player i's interest to induce this. Since we assume that the signals from the error-reporting noisy cheap talk when player i sends the message to player j are private, if we let player i inform each player j of x_i

[20]The latter implies that we cannot use the method that Fong, Gossner, Hörner, and Sannikov (2010) create $\lambda^j(y^j)$ in their Lemma 1 to preserve the conditional independence property.

separately, then player i may want to tell a lie to a subset of players. In Chapter 7, we create a message protocol so that, while the players exchange messages and infer the other players' messages from private signals in order to coordinate on x_i, there is no player who can induce a situation where some players infer x_i is G while the others infer x_i is B in order to increase her own equilibrium payoff. Yamamoto (2012) offers a procedure to achieve this goal with conditionally independent monitoring. Our contribution is a non-trivial extension of his procedure so that it is applicable to a general monitoring structure.

When we dispense with the error-reporting noisy cheap talk, in the two-player game, it suffices to verify that the sender has an incentive to tell the truth (take actions as prescribed by the strategy) and the receiver has an incentive to receive a message as prescribed by the strategy. With more than two players, we also need to make sure that players other than the sender and the receiver do not have an incentive to deviate in order to manipulate the receiver's signal distribution and inference. See Chapter 7 for the formal treatments of this incentive problem.

1.6 Finitely Repeated Game

In this section, we consider a T_P-period finitely repeated game with a "reward function." Intuitively, a finitely repeated game corresponds to a review phase in the infinitely repeated game and a reward function correspond to changes in the continuation payoff.

We derive sufficient conditions on strategies and reward functions in the finitely repeated game such that we can construct a strategy in the infinitely repeated game to support v. The sufficient conditions are summarized in Lemma 9.

Let $\sigma_i^{T_P} : H_i^{T_P} \to \Delta(A_i)$ be player i's strategy in the finitely repeated game. Let $\Sigma_i^{T_P}$ be the set of all strategies in the finitely repeated game. Each player i has a state $x_i \in \{G, B\}$. In state x_i, player i plays $\sigma_i(x_i) \in \Sigma_i^{T_P}$.

In addition, locate all the players on a circle clockwise. Each player i with x_i gives a "reward function" $\pi_{i+1}(x_i, \cdot : \delta) : H_i^{T_P+1} \to \mathbb{R}$ to the left-neighbor $i+1$ (identify player $N+1$ as player 1).[21] The reward functions are mapping from player i's histories in the finitely repeated game to the real numbers.

Our task is to find $\{\sigma_i(x_i)\}_{x_i,i}$ and $\{\pi_{i+1}(x_i, \cdot : \delta)\}_{x_i,i}$ such that, for each $i \in I$, there are two numbers $\underline{v}_i$ and $\bar{v}_i$ to contain v between them:

$$\underline{v}_i < v_i < \bar{v}_i \tag{1.14}$$

and such that there exists T_P with $\lim_{\delta\to 1} \delta^{T_P} = 1$ which satisfies the following conditions: for sufficiently large δ, for any $i \in I$,

1. for any combination of the other players' states $x_{-i} \equiv (x_n)_{n\neq i} \in \{G, B\}^{N-1}$, it is optimal to take $\sigma_i(G)$ and $\sigma_i(B)$:

$$\sigma_i(G), \sigma_i(B) \in \arg\max_{\sigma_i^{T_P}\in\Sigma_i^{T_P}} \mathbb{E}\left[\sum_{t=1}^{T_P} \delta^{t-1} u_i(a_t) + \pi_i(x_{i-1}, h_{i-1}^{T_P+1} : \delta) \mid \sigma_i^{T_P}, \sigma_{-i}(x_{-i})\right]; \tag{1.15}$$

2. regardless of $x_{-(i-1)}$, the discounted average of player i's instantaneous utilities and player $(i-1)$'s reward function on player i is equal to $\bar{v}_i$ if player $(i-1)$'s state is good ($x_{i-1} = G$) and equal to $\underline{v}_i$ if player $(i-1)$'s state is bad ($x_{i-1} = B$):

$$\frac{1-\delta}{1-\delta^{T_P}} \mathbb{E}\left[\sum_{t=1}^{T_P} \delta^{t-1} u_i(a_t) + \pi_i(x_{i-1}, h_{i-1}^{T_P+1} : \delta) \mid \sigma(x)\right] = \begin{cases} \bar{v}_i \text{ if } x_{i-1} = G, \\ \underline{v}_i \text{ if } x_{i-1} = B \end{cases} \tag{1.16}$$

for all $x_{-(i-1)} \in \{G, B\}^{N-1}$.

Intuitively, since $\lim_{\delta\to 1} \frac{1-\delta}{1-\delta^{T_P}} = \frac{1}{T_P}$, this requires that the time average of the expected sum of the instantaneous utilities and the reward function is close to the targeted payoffs $\underline{v}_i$ and $\bar{v}_i$; and

[21] The players are inward-looking.

3. $\frac{1-\delta}{\delta^{T^P}}$ converges to 0 faster than $\pi_i(x_{i-1}, h_{i-1}^{T_P+1} : \delta)$ diverges and the sign of $\pi_i(x_{i-1}, h_{i-1}^{T_P+1} : \delta)$ satisfies a proper condition:

$$\begin{cases} \lim_{\delta \to 1} \frac{1-\delta}{\delta^{T_P}} \sup_{x_{i-1}, h_{i-1}^{T_P+1}} \left|\pi_i(x_{i-1}, h_{i-1}^{T_P+1} : \delta)\right| = 0, \\ \pi_i(G, h_{i-1}^{T_P+1} : \delta) \leq 0, \\ \pi_i(B, h_{i-1}^{T_P+1} : \delta) \geq 0. \end{cases} \tag{1.17}$$

We call (1.17) the "feasibility constraint."

As seen in Section 1.5, (1.17) implies that player $i-1$ with $x_{i-1} = G$ can reduce player i's continuation payoff by transiting to $x_{i-1} = B$ with a higher probability while player $i-1$ with $x_{i-1} = B$ can increase player i's continuation payoff by transiting to $x_{i-1} = G$ with a higher probability.

We explain why these conditions are sufficient. As explained in Section 1.5, we see the infinitely repeated game as the repetition of T_P-period review phases.

In each review phase, each player i has two possible states $\{G, B\} \ni x_i$ and player i with state x_i takes $\sigma_i(x_i)$ in the phase. (1.15) implies that both $\sigma_i(G)$ and $\sigma_i(B)$ are optimal regardless of the other players' states. (1.16) implies that player i's ex ante value at the beginning of the phase is solely determined by player $(i-1)$'s state: $\sigma_{i-1}(G)$ gives a high value while $\sigma_{i-1}(B)$ gives a low value.

Here, $\pi_i(x_{i-1}, h_{i-1}^{T_P+1} : \delta)$ represents the differences between player i's ex ante value given x_{i-1} at the beginning of the phase and the ex post value at the end of the phase after player $i-1$ observes $h_{i-1}^{T_P+1}$. $\pi_i(x_{i-1}, h_{i-1}^{T_P+1} : \delta) = 0$ implies that the ex post value is the same as the ex ante value since player $i-1$ transits to the same state in the next phase with probability one. With $x_{i-1} = G$ (B, respectively), the smaller $\pi_i(G, h_{i-1}^{T_P+1} : \delta)$ (the larger $\pi_i(B, h_{i-1}^{T_P+1} : \delta)$, respectively), the more likely it is for player $i-1$ to transit to the opposite state B (G, respectively) in the next phase. The feasibility of this transition is guaranteed by (1.17).

The following lemma summarizes the discussion:

Lemma 9 For Theorem 7, it suffices to show that, for any $v \in \text{int}(F^*)$, for sufficiently large δ, there exist $\{\underline{v}_i, \bar{v}_i\}_{i \in I}$ with (1.14), T_P with $\lim_{\delta \to 1} \delta^{T_P} = 1$, $\{\{\sigma_i(x_i)\}_{x_i \in \{G,B\}}\}_{i \in I}$ and $\{\{\pi_i(x_{i-1}, \cdot : \delta)\}_{x_{i-1} \in \{G,B\}}\}_{i \in I}$ such that (1.15), (1.16) and (1.17) are satisfied in the T_P-period finitely repeated game.

Proof: See Section 1.9. ∎

From now on, when we say player i's action plan, it means player i's behavioral mixed strategy $\sigma_i(x_i)$ within the current review phase (or, the finitely repeated game). On the other hand, when we say player i's strategy, it contains both $\sigma_i(x_i)$ and $\pi_{i+1}(x_i, \cdot : \delta)$ which determines player i's continuation strategy from the next review phase.

Let us specify $\underline{v}_i$ and $\bar{v}_i$. This step is the same as Hörner and Olszewski (2006). Given $x \in \{G, B\}^N$, pick 2^N action profiles $\{a(x)\}_{x \in \{G,B\}^N}$. As we have mentioned, player $(i-1)$'s state x_{i-1} refers to player i's payoff and indicates whether this payoff is strictly above or below v_i no matter what the other players' states are. That is, player $(i-1)$'s state controls player i's payoff. Formally,

$$\max_{x:x_{i-1}=B} u_i(a(x)) < v_i < \min_{x:x_{i-1}=G} u_i(a(x)) \text{ for all } i \in I.$$

For example, in the two-player prisoners' dilemma, for $x = (G, G)$, $a_i(x) = C_i$.

Take $\underline{v}_i$ and $\overline{v}_i$ such that

$$\max\left\{v_i^*, \max_{x:x_{i-1}=B} u_i(a(x))\right\} < \underline{v}_i < v_i < \overline{v}_i < \min_{x:x_{i-1}=G} u_i(a(x)). \tag{1.18}$$

Remember that v_i^* is player i's minimax value. From now on, without loss, we assume that α_{-i}^* is a perfectly mixed action plan: $\alpha_{-i}^*(a_{-i}) > 0$ for all $a_{-i} \in A_{-i}$. Otherwise, perturb α_{-i}^* slightly so that each player $j \in -i$ takes all the actions in A_j with a positive probability and (1.18) still holds.

Action profiles that satisfy the desired inequalities may not exist. However, if Assumption 1 is satisfied, then there always exist an integer z and 2^z finite sequences $\{a_1(x), \ldots, a_z(x)\}$ $_{x\in\{G,B\}^N}$ such that each vector $w_i(x)$, the average discounted payoff vector over the sequence $\{a_1(x),$

$\ldots, a_z(x)\}_{x\in\{G,B\}^N}$, satisfies the appropriate inequalities provided δ is close enough to 1. 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 as in Hörner and Olszewski (2006).

Given $\rho > 0$ that will be determined later, for each i, given $a(x)$, we perturb $a_i(x)$ to $\alpha_i(x)$ so that player i takes all the actions in A_i with a positive probability no less than 2ρ: taking $a_i(x)$ with probability $1 - 2(|A_i| - 1)\rho$ and take $a_i \neq a_i(x)$ with probability 2ρ. For example, in the two-player prisoners' dilemma, for $x = (G, G)$, $\alpha(x) = (1 - 2\rho)\, C_i + 2\rho D_i$ as in (1.13).

Let $\{w(x)\}_{x\in\{G,B\}^N}$ be the corresponding payoff vectors under $\alpha(x)$:

$$w(x) \equiv u(\alpha(x)) \text{ with } x \in \{G, B\}^N. \tag{1.19}$$

As we will see in Section 2.6, with sufficiently small ρ, (1.18) implies

$$\max\left\{v_i^*, \max_{x:x_{i-1}=B} w_i(x)\right\} < \underline{v}_i < v_i < \overline{v}_i < \min_{x:x_{i-1}=G} w_i(x). \tag{1.20}$$

Below, we construct $\{\sigma_i(x_i)\}_{x_i,i}$ and $\{\pi_i(x_{i-1}, \cdot : \delta)\}_{x_{i-1},i}$ satisfying (1.15), (1.16) and (1.17) with $\bar{v}_i$ and $\underline{v}_i$ defined above in the finitely repeated game.

1.7 Coordination, Main and Report Blocks

In this section, we explain the basic structure of the T_P-period finitely repeated game. At the beginning of the finitely repeated game, there is the "coordination block." In the finitely repeated game, the players play the action-plan profile $\alpha(x)$ depending on the state profile $x = (x_n)_{n\in I} \in \{G, B\}^N$. Since x_i is player i's private state, player i informs the other players $-i$ of x_i by sending messages about x_i.

As seen in Section 1.5, we first assume that the players can communicate x via perfect cheap talk. The players take turns: player 1 tells x_1 first, player 2 tells x_2 second, and so on until player N tells x_N. With the perfect cheap talk, this block is instantaneous and x becomes common knowledge. Second, we replace the perfect cheap talk with the error-reporting noisy cheap talk. As we will see, with two players, this block is still instantaneous while with more than two players, this block now consists of many periods. More importantly, x is no longer common knowledge. Finally, we replace the error-reporting noisy cheap talk with messages via actions. Since the players repeat the messages to increase the precision, this block takes time.

After the coordination block, we have "main blocks." One main block consists of a review round and a few supplemental rounds. The review round lasts T periods with

$$T = (1-\delta)^{-\frac{1}{2}}$$

as seen in Section 1.5. After that, for each player i, each player $j \in -i$ sends messages about what is player i's optimal action in the next round. As explained in Section 1.5, we first assume that player j sends the messages via error-reporting noisy cheap talk. With the error-reporting noisy cheap talk, this message is sent instantaneously. Then, we replace the error-reporting noisy cheap talk with messages via actions. Since the players repeat the messages to increase the precision, sending the messages takes time.

Let h_i^{main} be a generic element of player i's history at the end of the last main block, that is, player i's history in the coordination block and all the main blocks.

After the last main block, we have the "report block" where each player reports h_i^{main}. We first assume that the players decide who will report the history by the public randomization device and that the picked player reports h_i^{main} by the perfect cheap talk. Then, this block is instantaneous. Second, we dispense with the public randomization. Third, we replace the perfect cheap talk with conditionally independent (noisy) cheap talk. Fourth, we dispense with the conditionally independent cheap talk.

When we say $h_i^{T_P+1}$, this denotes player i's history at the end of the report block, that is, $h_i^{T_P+1}$ contains both h_i^{main} and what information player i receives about $\left(h_n^{\text{main}}\right)_{n\in I}$ in the report block.

1.8 Almost Optimality

As seen in Section 1.5, we first show that player i's strategy is "almost optimal," or that the strategy profile is "ε-equilibrium" with $\varepsilon = \exp(-\Theta(T^{\frac{1}{2}}))$ until the end of the last main block if we neglect the report block. After that, based on the communication in the report block, player $i-1$ adjusts the reward function so that player i's strategy is exactly optimal after any history in any period of the review phase if we take the report block into account.

We divide the reward function into two parts:

$$\pi_i(x_{i-1}, h_{i-1}^{T_P+1} : \delta) = \pi_i^{\text{main}}(x_{i-1}, h_{i-1}^{\text{main}} : \delta) + \pi_i^{\text{report}}(x_{i-1}, h_{i-1}^{T_P+1} : \delta).$$

Note that $\pi_i^{\text{main}}(x_{i-1}, h_{i-1}^{\text{main}} : \delta)$ is the reward based on player $(i-1)$'s history except for the report block and that $\pi_i^{\text{report}}(x_{i-1}, h_{i-1}^{T_P+1} : \delta)$ is the reward based on player $(i-1)$'s whole history including the report block. As we will see, $\pi_i^{\text{report}}(x_{i-1}, h_{i-1}^{T_P+1} : \delta)$ is the adjustment that we mention above.

As a preparation to prove the existence of π_i with (1.15), (1.16) and (1.17), we first construct π_i^{main} such that

1. $\sigma_i(x_i)$ is "almost optimal with $\exp(-\Theta(T^{\frac{1}{2}}))$ if we ignore the report block": for all $i \in I$ and $x \in \{G, B\}^N$, for any τ and h_i^τ in the coordination and main blocks,

$$\begin{aligned} & \max_{\sigma_i \in \Sigma_i^{\text{main}}} \mathbb{E}\left[\sum_{t=1}^{T_P} \delta^{t-1} u_i(a_t) + \pi_i^{\text{main}}(x_{i-1}, h_{i-1}^{\text{main}} : \delta) \mid h_i^\tau, \sigma_i, \sigma_{-i}(x_{-i})\right] \\ & \quad -\mathbb{E}\left[\sum_{t=1}^{T_P} \delta^{t-1} u_i(a_t) + \pi_i^{\text{main}}(x_{i-1}, h_{i-1}^{\text{main}} : \delta) \mid h_i^\tau, \sigma(x)\right] \\ \leq & \ \exp(-\Theta(T^{\frac{1}{2}})). \end{aligned} \tag{1.21}$$

Here, Σ_i^{main} is the set of all possible strategies in the coordination and main blocks; and

2. (1.16) and (1.17) are satisfied with π_i replaced with π_i^{main} (neglecting π_i^{report}).

That is, our first objective is to construct $\sigma_i(x_i)$ and $\pi_i^{\text{main}}(x_{i-1}, h_{i-1}^{\text{main}} : \delta)$ satisfying (1.21), (1.16) and (1.17). After constructing such π_i^{main}, our second (and final) objective is to construct the action plan in the report block and the adjustment π_i^{report} such that $\sigma_i(x_i)$ and $\pi_i = \pi_i^{\text{main}} + \pi_i^{\text{report}}$ satisfy (1.15), (1.16) and (1.17).

In Chapter 2, we pursue these two objectives in the two-player prisoners' dilemma with the special forms of cheap talk. Given Chapter 2, we are left to extend the result to the general two-player game and the general N-player game with $N \geq 3$ and dispense with the special forms of communication. In Chapter 3, we summarize what new assumptions are sufficient for each extension. In the following chapters, we offer the proof: in Chapters 4 and 5, we extend the result to the general two-player game and the general N-player game, respectively, with special forms of communication. In Chapters 6 and 7, we dispense with the special forms of communication in the two-player game and N-player game, respectively.

1.9 Appendix of Chapter 1

1.9.1 Proof of Lemma 9

To see why this is enough for Theorems 7 and 8, define the strategy in the infinitely repeated game as follows: define

$$\begin{aligned} p(G, h_{i-1}^{T_P+1} &: \delta) \equiv 1 + \frac{1-\delta}{\delta^{T_P}} \frac{\pi_i(G, h_{i-1}^{T_P+1} : \delta)}{\bar{v}_i - \underline{v}_i}, \\ p(B, h_{i-1}^{T_P+1} &: \delta) \equiv \frac{1-\delta}{\delta^{T_P}} \frac{\pi_i(B, h_{i-1}^{T_P+1} : \delta)}{\bar{v}_i - \underline{v}_i}. \end{aligned} \tag{1.22}$$

If (1.17) is satisfied, then for sufficiently large δ, $p(G, h_{i-1}^{T_P+1} : \delta)$, $p(B, h_{i-1}^{T_P+1} : \delta) \in [0,1]$ for all $h_{i-1}^{T_P+1}$. We see the repeated game as the repetition of T_P-period "review phases." In each phase, player i has a state $x_i \in \{G, B\}$. Within the phase, player i with state x_i plays according to $\sigma_i(x_i)$ in the current phase. After observing $h_i^{T_P+1}$ in the current phase, the state in the next phase is equal to G with probability $p(x_i, h_i^{T_P+1} : \delta)$ and B with the remaining probability.

Player $(i-1)$'s initial state is equal to G with probability p_v^{i-1} and B with probability $1 - p_v^{i-1}$ such that

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

Then, since

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

and

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

(1.15) and (1.16) imply that, for sufficiently large discount factor δ,

1. conditional on the opponents' state, the above strategy in the infinitely repeated game is optimal;

2. regardless of $x_{-(i-1)}$, if player $i-1$ is in the state G, then player i's payoff from the infinitely repeated game is $\bar{v}_i$ and if player $i-1$ is in the state B, then player i's payoff is $\underline{v}_i$; and

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

Chapter 2

Two-Player Prisoners' Dilemma

2.1 Special Case

In this chapter, we illustrate the proof of the folk theorem by focusing on a special case where (i) public randomization is available, (ii) perfect cheap talk is available, (iii) error-reporting noisy cheap talk with precision $p \in (0,1)$ is available, (iv) there are two players ($N = 2$), and (v) the payoff structure is prisoners'-dilemma

$$u_i(D_i, C_j) > u_i(C_i, C_j) > u_i(D_i, D_j) > u_i(C_i, D_j), \tag{2.1}$$

and for all i,

$$v \in \text{int}([u_1(D_1, D_2), u_1(C_1, C_2)] \times [u_2(D_2, D_1), u_2(C_2, C_1)]). \tag{2.2}$$

We comment on each of these five simplifications.

Public Randomization As mentioned in Section 1.5, the players use public randomization in the report block to determine who will report the history h_i^{main} such that (i) ex ante (during the main blocks), every player has a positive probability to report the history, and that (ii) ex post (after the realization of the public randomization), there is only one player who reports the history.

Specifically, we assume that the players can draw a public random variable y^p from the uniform distribution on $[0,1]$ whenever they want.

In Chapters 6 and 7, we show that the public randomization is dispensable and that the players use their actions and private signals to establish the properties (i) and (ii).

Perfect Cheap Talk Perfect cheap talk will be used in the coordination block to coordinate on x and in the report block to report the whole history h_i^{main}.

In Chapters 6 and 7, we show that the perfect cheap talk is dispensable. As explained in Section 1.5, for the coordination block, we first replace the perfect cheap talk with the error-reporting noisy cheap talk and then replace the error-reporting noisy cheap talk with messages via actions. For the report block, we first replace the perfect cheap talk with the conditional independence (noisy) cheap talk and then replace the conditional independence cheap talk with messages via actions.

Error-Reporting Noisy Cheap Talk Between Players j and i with Precision $p \in (0,1)$ We assume that player j has an access to the error-reporting noisy cheap talk to send a binary message $m \in \{G, B\}$ to player i with precision p, as explained in Section 1.3.2.

As we will see below, in the two-player game, p will be either $p = 1 - \exp(-T^{\frac{1}{2}})$ or $p = 1 - \exp(-T^{\frac{2}{3}})$ while η for (1.2), (1.3) and (1.5) is a fixed number. Noting that $j = i - 1$ and $i = j - 1$ with two players, the properties that we will use are summarized in the following lemma:

Lemma 10 The signals by the error-reporting noisy cheap talk between j and i with precision $1 - \exp(-T^k)$ with $k = 1/2, 2/3$ satisfy the following conditions:

1. for any $m \in \{G, B\}$, player i's signal $f[i](m)$ is correct with a high probability:

$$\Pr(\{f[i](m) = m\} \mid m) \geq 1 - \exp(-T^k);$$

2. for any $m \in \{G, B\}$, $f[i](m) \in \{G, B\}$ and $f_2[i](m) \in \{G, B\}$, after knowing m, $f[i](m)$ and $f_2[i](m)$, player i puts a high belief on the events that either $f[i](m)$ is correct or $g[j](m) = E$. That is,

$$\begin{aligned} & \Pr(\{f[i](m) = m \text{ or } g[j](m) = E\} \mid m, f[i](m), f_2[i](m)) \\ = \ & 1 - \exp(-\Theta(T^k)); \end{aligned}$$

3. for any $m \in \{G, B\}$ and $g_2[j](m) \in \{G, B\}$, if $g[j](m) = m$ (the error is not reported), then player j puts a high belief on the event that player i's first signal is correct:

$$\Pr(\{f[i](m) = m\} \mid m, \{g[j](m) = m\}, g_2[j](m)) \geq 1 - \exp(-\Theta(T^k));$$

and

4. for any $m \in \{G, B\}$, any signal profile can happen with a positive probability:

$$\begin{aligned} & \Pr(\{(f[i](m), g[j](m), f_2[i](m), g_2[j](m)) = (f, g, f_2, g_2)\} \mid m) \\ \geq \ & \exp(-\Theta(T^k)) \end{aligned}$$

for all (f, g, f_2, g_2).

Proof: The only nontrivial statement is Condition 3. Since η is fixed, it suffices to show that

$$\Pr(\{f[i](m) = m\} \mid m, \{g[j](m) = m\}) \geq 1 - \exp(-\Theta(T^k)).$$

By Bayes' rule,

$$\begin{aligned}
& \Pr\left(\{f[i](m) = m\} \mid m, \{g[j](m) = m\}\right) \\
= & \frac{\Pr\left(\{f[i](m) = g[j](m) = m\} \mid m\right)}{\Pr\left(\{g[j](m) = m\} \mid m\right)} \\
\geq & \frac{1 - \exp(-2T^k)}{1 - \exp(-T^k)},
\end{aligned}$$

as desired. ∎

Condition 1 implies that the signal is correct with a high probability. Condition 2 implies that, even after player i realizes that her signal is not correct ($f[i](m) \neq m$), player i believes that player j realizes the mistake (that is, $g[j](m) = E$) with a high probability. On the other hand, Condition 3 implies that, after observing $g[j](m) = m$, player j believes that player i received the correct signal (since otherwise player i should have received $g[j](m) = E$) with a high probability.[1] Further, Condition 4 implies that all the players believe that any mistake happens with a positive probability. As seen in Section 1.5, this is important to solve the reflective learning problem.

On the other hand, with more than two players, p will be $p = 1 - \exp(-T^{\frac{1}{2}})$ while η for (1.2), (1.3) and (1.5) is a fixed number. As for Lemma 10, we can summarize the important features of the error-reporting noisy cheap talk in the following lemma:

Lemma 11 The signals by the error-reporting noisy cheap talk satisfies the following conditions:

1. for any $m \in \{G, B\}$, player i's signal $f[i](m)$ is correct with a high probability:

$$\Pr\left(\{f[i](m) = m\} \mid m\right) \geq 1 - \exp(-T^{\frac{1}{2}});$$

2. for any $m \in \{G, B\}$, $f[i](m) \in \{G, B\}$ and $f_2[j-1](m) \in \{G, B\}$, after knowing m, $f[i](m)$ and $f_2[j-1](m)$, player i puts a high belief on the events that either $f[i](m)$

[1] We use this property only in Section 6.2.

is correct or $g[i-1](m) = E$. That is,

$$\begin{aligned} & \Pr\left(\{f[i](m) = m \text{ or } g[i-1](m) = E\} \mid m, f[i](m), f_2[j-1](m)\right) \\ \geq \ & 1 - \exp(-\Theta(T^{\frac{1}{2}})); \end{aligned}$$

and

3. for any $m \in \{G, B\}$, any signal profile can happen with a positive probability:

$$\begin{aligned} & \Pr\left(\{(f[i](m), g[i-1](m), f_2[j-1](m), g_2[i-1](m)) = (f, g, f_2, g_2)\} \mid m\right) \\ \geq \ & \exp(-\Theta(T^{\frac{1}{2}})) \end{aligned}$$

for all (f, g, f_2, g_2).

We do not have a condition corresponding to Condition 3 of Lemma 10.

As seen in Lemmas 10 and 11, the learning from f_2 and g_2 is negligible for the almost optimality. Hence, while we are considering the almost optimality, we neglect f_2 and g_2, which will play roles when we consider the exact optimality in Section 2.8.

In Chapters 6 and 7, we show that we can replace the error-reporting noisy cheap talk with messages via actions, so that we can keep the important features summarized in Lemmas 10 and 11.

History with Cheap Talk and Public Randomization Since the players communicate via cheap talk, the players store the signals from the cheap talk in the history. When a sender sends a message m via perfect cheap talk (error-reporting noisy cheap talk, respectively) then the players observe a perfect signal m (private signals f, g, f_2 and g_2 depending on their indices, respectively). In addition, the sender observes the true message. With abuse of notation, when the communication is done before the players take actions in period t, we include these observations to the history in period t of player i, h_i^t.

In addition, since the players coordinate the future play via public randomization, the players store the realization of the public randomization in the history. Again, when a public randomization device is drawn before the players take actions in period t, we include the realization of the public randomization device to the history in period t of each player i, h_i^t.

Two Players As explained in Sections 1.5.5.3 and 1.5.6.2, the two-player case is special in the following three aspects: first, if player $(i-1)$'s state x_{i-1} is B, then with more than two players, players $-i$ need to coordinate on minimaxing player i if player i seems to have deviated. Second, when the players coordinate on x_i in the coordination block, we need to make sure that no player can induce the situation where some players infer x_i is G while the others infer x_i is B. Third, while player j sends a message to player n, no other players can manipulate player n's inference. See Chapter 5 for the first problem and Chapter 7 for the second and third problems.

Below, we concentrate on the two-player case. Since we assume two players, let player j be player i's unique opponent.

The Prisoners' Dilemma Remember that we take $a(x)$ such that (1.18) holds. If (2.2) is the case, then we can take

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

Then, as we mentioned in (1.20), we can take

$$\alpha_i(x) \equiv \begin{cases} (1-2\rho)\, C_i + 2\rho D_i & \text{if } x_i = G, \\ (1-2\rho)\, D_i + 2\rho C_i & \text{if } x_i = B \end{cases}$$

with sufficiently small $\rho > 0$. Then, it happens to be the case that $\alpha_i(x)$ with $x_i = B$ is very close to minimaxing player j at the same time of satisfying (1.20).

In a general game, $\alpha_i(x)$ with $x_i = B$ is not close to a minimaxing action plan. Since player i with $\sigma_i(B)$ needs to keep player j's payoff low with a non-negative reward for any action plan of player j, player i needs to switch to a minimaxing action plan if player i believes that player j has deviated with a high probability.

For this reason, Hörner and Olszewski (2006) have a block consisting of more than one period and in each period, if player i observes a signal indicating player j's deviation, then player i switches to a minimaxing action. As explained in Section 1.5, in our equilibrium, if player i observes signals indicating player j's deviation in a review round, then player i minimaxes player j from the next review round. See Chapter 4 for the formal treatment of the general two-player game.

Summary In summary, for the proof in Chapter 2, we assume that the perfect cheap talk, error-reporting noisy cheap talk and public randomization are available, all of which are shown to be dispensable in Chapter 6.

We focus on the two-player prisoners' dilemma: $I = 2$ and $A_i = \{C_i, D_i\}$ satisfying (2.1). Further, we focus on v with (2.2).

We can take $\alpha(x)$ such that

$$\alpha_i(x) \equiv \begin{cases} (1-2\rho)\, C_i + 2\rho D_i & \text{if } x_i = G, \\ (1-2\rho)\, D_i + 2\rho C_i & \text{if } x_i = B \end{cases} \tag{2.3}$$

with small $\rho > 0$ to be determined.

In addition, for notational convenience, whenever we say players i and j, unless otherwise specified, i and j are different.

For the rest of Chapter 2, we prove the folk theorem in this special case: we arbitrarily fix v with (2.2) and then construct a strategy profile (action plans and rewards) in the finitely repeated game with (1.15), (1.16) and (1.17).

2.2 Structure of the Phase

In this section, we formally define the structure of the T_P-period finitely repeated game (review phase), whose structure will be explained below. T_P depends on L and T. $L \in \mathbb{N}$ will be pinned down in Section 2.6 and $T = (1-\delta)^{-\frac{1}{2}}$.

As seen in Section 1.7, at the beginning of the phase, there is the coordination block. The players take turns to communicate x. First, player 1 sends x_1 via perfect cheap talk. Second, player 2 sends x_2 via perfect cheap talk. For notational convenience, let the round for x_i denote the moment that player i sends x_i.

After the coordination blocks, there are L "main blocks." Each of the first $(L-1)$ main blocks are further divided into three rounds. That is, for $l \in \{1, ..., L-1\}$, the lth main block consists of the following three rounds: first, the players play a T-period review round. Second, there is a supplemental round for $\lambda_1(l+1)$. Third, there is a supplemental round $\lambda_2(l+1)$. As seen in Section 1.5, $\lambda_i(l+1) \in \{G, B\}$ is an index of whether player i has observed an "erroneous score" in the review rounds $1, \ldots, l$. In the supplemental round for $\lambda_i(l+1)$, player i sends $\lambda_i(l+1)$ via error-reporting noisy cheap talk with precision $p = 1 - \exp(-T^{\frac{1}{2}})$.

The last Lth main block has only the T-period review round.

Let $T(l)$ be the set of T periods in the lth review round. In addition, in each lth review round, each player i randomly picks one period $t_i(l)$ from $T(l)$: $\Pr(\{t_i(l) = t\}) = \frac{1}{T}$ for all $t \in T(l)$. Let $T_i(l) \equiv T(l) \setminus \{t_i(l)\}$ be the other periods than $t_i(l)$ in the lth review round. As we will see, player i "excludes" period $t_i(l)$ when she determines her continuation action plan so that player j cannot learn player i's history in period $t_i(l)$ by observing player i's continuation action plan. As will be seen in Section 2.8, this is important to incentivize player i to tell the truth in the report block.

After the last main block, there is the report block, where player i who is picked by the public randomization reports the whole history h_i^{main}.

Given this structure, we show that, for sufficiently large δ, with $T_P = L(1-\delta)^{-\frac{1}{2}}$, there exist $\sigma_i(x_i)$ and $\pi_i(x_j, h_j^{T_P+1} : \delta)$ satisfying (1.15), (1.16) and (1.17).

2.3 Preparation

Before constructing an equilibrium, we define functions and statistics useful for the equilibrium construction.

First, we formally define the point $\pi_i[\alpha](y_j)$ as briefly explained in Section 1.5.5.1. Specifically, for each $\alpha \in \Delta(A)$, we want to create a statistics (point) $\pi_i[\alpha](y_j)$ that cancels out the differences in the instantaneous utilities for different a_i's:

$$u_i(a_i, \alpha_j) + \mathbb{E}[\pi_i[\alpha](y_j) \mid a_i, \alpha_j] \tag{2.4}$$

is independent of $a_i \in A_i$, as in (1.9).

Further, we want to make sure that if $\alpha = \alpha(x)$, then the expected sum of the instantaneous utility and $\pi_i[\alpha(x)](y_j)$ satisfies

$$u_i(a_i, \alpha_j(x)) + \mathbb{E}[\pi_i[\alpha(x)](y_j) \mid a_i, \alpha_j(x)] = w_i(x) \tag{2.5}$$

for all $a_i \in A_i$. This corresponds to (1.10) in Section 1.5. From (1.19), this implies

$$\mathbb{E}[\pi_i[\alpha(x)](y_j) \mid \alpha(x)] = \begin{cases} \leq 0 & \text{if } x_j = G, \\ \geq 0 & \text{if } x_j = B. \end{cases} \tag{2.6}$$

Since Assumption 3 implies that player j can statistically infer player i's action, the existence of such π_i is guaranteed.

Lemma 12 If Assumption 3 is satisfied, then there exists $\bar{u} > 0$ such that, for each $i \in I$, $\alpha \in \Delta(A)$ and $\{\alpha(x)\}_{x \in \{G,B\}^2}$, there exists $\pi_i[\alpha] : Y_j \to [-\bar{u}, \bar{u}]$ with (2.4) and (2.5).

Proof: See Section 2.9. ∎

Note that the bound for π_i, $\bar{u}$, is independent of x and α_j. Hence, by re-taking $\bar{u}$ if necessary, we can make sure that, for all $i \in I$ and $\alpha \in \Delta(A)$, we have

$$u_i(a_i, \alpha_j) + \mathbb{E}\left[\pi_i[\alpha](y_j) \mid a_i, \alpha_j\right] \in [-\bar{u}, \bar{u}] \text{ for all } a_i \in A_i. \tag{2.7}$$

Second, define

$$\pi_i^{x_j}[\alpha_j](y_j) \begin{cases} \leq 0 \text{ for all } y_j & \text{if } x_j = G, \\ \geq 0 \text{ for all } y_j & \text{if } x_j = B \end{cases} \tag{2.8}$$

such that, for all $i \in I$,

$$u_i(a_i, \alpha_j) + \mathbb{E}\left[\pi_i^{x_j}[\alpha_j](y_j) \mid a_i, \alpha_j\right] \tag{2.9}$$

is independent of $a_i \in A_i$ and $\alpha_j \in \Delta(A_j)$ and included in $[-2\bar{u}, 2\bar{u}]$.

The existence of such $\pi_i^{x_j}[\alpha_j](y_j)$ is guaranteed: arbitrarily fix $\bar{a}_i \in A_i$ and define

$$\pi_i^{x_j}[\alpha_j](y_j) = \begin{cases} \pi_i[\bar{a}_i, \alpha_j](y_j) - \bar{u} & \text{if } x_j = G, \\ \pi_i[\bar{a}_i, \alpha_j](y_j) + \bar{u} & \text{if } x_j = B. \end{cases}$$

Then, Lemma 12 and (2.7) imply (2.8). In addition, (2.9) holds for all $a_i \in A_i$. To make sure that (2.9) holds for all $\alpha_j \in \Delta(A_j)$, add or subtract the payoff difference with respect to $\alpha_j \in \Delta(A_j)$ to cancel out the difference, keeping (2.8). Since $\pi_i^{x_j}[\alpha_j](y_j)$ can depend on α_j, this is possible.

Third, since, again, Assumption 3 implies that player j can statistically infer player i's action, player j can give a reward that cancels out the effect of discounting:

Lemma 13 If Assumption 3 is satisfied, then for each $i \in I$, there exists $\pi_i^{\delta} : \mathbb{N} \times A_j \times Y_j \to \mathbb{R}$ such that

$$\delta^{t-1} u_i(a_t) + \mathbb{E}\left[\pi_i^{\delta}(t, a_{j,t}, y_{j,t}) \mid a_t\right] = u_i(a_t) \text{ for all } a_t \in A \text{ and } t \in \{1, ..., T_P\} \tag{2.10}$$

and

$$\lim_{\delta\to 1}\frac{1-\delta}{1-\delta^{T_P}}\sum_{t=1}^{T_P}\sup_{a_{j,t},y_{j,t}}\left|\pi_i^{\delta}\left(t,a_{j,t},y_{j,t}\right)\right|=0 \tag{2.11}$$

for $T_P=\Theta(T)$ with $T=(1-\delta)^{-\frac{1}{2}}$.

Proof: See Section 2.9. ∎

The intuition is straightforward. Since player j can identify player i's action, player j rewards player i if player i takes an action with a lower instantaneous utility in earlier periods rather than postponing it. Since the discount factor converges to unity, this adjustment is small.

2.4 Equilibrium Strategies

Given the preparation above, we define $\sigma_i\left(x_i\right)$ in the coordination and main blocks and $\pi_i^{\text{main}}(x_j,h_j^{\text{main}}:\delta)$. See Section 2.8 for the definition of $\sigma_i(x_i)$ in the report block and $\pi_i^{\text{report}}(x_j,h_j^{T_P+1}:\delta)$.

In Section 2.4.1, we define the state variables that will be used to define the action plans and rewards. Given the states, Section 2.4.2 defines the action plan $\sigma_i(x_i)$ and Section 2.4.3 defines the reward function $\pi_i^{\text{main}}(x_j,h_j^{\text{main}}:\delta)$. Finally, Section 2.4.4 determines the transition of the states defined in Section 2.4.1.

2.4.1 States x_i, $\lambda_j(l+1)$, $\hat{\lambda}_j(l+1)$ and $\theta_i(l)$

The state $x_i\in\{G,B\}$ is determined at the beginning of the review phase and fixed. With the perfect cheap talk, after player 2 sends x_2 in the coordination block, x becomes common knowledge. Hence, from now on, we use x_j for the definition of player i's strategy after the coordination block.

As seen in Section 1.5, $\lambda_j(l+1)\in\{G,B\}$ is player j's state. Intuitively, $\lambda_j(l+1)=B$ implies that player j has observed an erroneous score about player i in the lth round or before. As will be formally defined in Section 2.4.4, $\lambda_j(l+1)$ is determined at the end of the lth review round.

On the other hand, since player j's reward on player i in the $(l+1)$th review round depends on $\lambda_j(l+1)$ as seen in Section 1.5, it is natural to consider player i's belief about $\lambda_j(l+1) = G$. The space for player i's possible beliefs about $\lambda_j(l+1) = G$ in each period t in the $(l+1)$th review round is $[0,1]$ and it depends on the details of a history h_i^t. However, we classify the set of player i's histories into two partitions: the set of histories labeled as $\hat{\lambda}_j(l+1) = G$ and that labeled as $\hat{\lambda}_j(l+1) = B$. Intuitively, $\hat{\lambda}_j(l+1) = G$ ($\hat{\lambda}_j(l+1) = B$, respectively) implies that player i believes that $\lambda_j(l+1) = G$ ($\lambda_j(l+1) = B$, respectively) is likely.

To make the equilibrium tractable, $\hat{\lambda}_j(l+1)$ depends only on player i's history at the beginning of the $(l+1)$th review round and is fixed during the $(l+1)$th review block, as will be defined in Section 2.4.4. Given $x \in \{G,B\}^2$ and $\hat{\lambda}_j(l+1) \in \{G,B\}$, player i takes an *i.i.d.* action plan within the $(l+1)$th review round.

Further, as we have briefly mentioned in Sections 1.5, player j makes player i indifferent between any action profile after some history. If she does in the lth review round, then π_i^{main} will be $\sum_\tau \pi_i^{x_j}[\alpha_{j,\tau}](y_{j,\tau})$ for period τ in the lth review round and after with $\alpha_{j,\tau}$ being an action plan that player j takes in period τ. $\theta_j(l) \in \{G,B\}$ is an index of whether player j uses such a reward. See Section 2.4.3 for how the reward function depends on $\theta_j(l)$ and see Section 2.4.4 for the transition of the states.

2.4.2 Player i's Action Plan $\sigma_i(x_i)$

In this subsection, we define player i's action plan $\sigma_i(x_i)$. In the coordination block, player i tells the truth about x_i.

In the each lth review round, player i with $\sigma_i(x_i)$ takes an *i.i.d.* action plan $\alpha_i(l)$ as follows: given $\rho > 0$, remember that

$$\alpha_i(x) = \begin{cases} (1-2\rho)\, C_i + 2\rho D_i & \text{if } x_i = G, \\ (1-2\rho)\, D_i + 2\rho C_i & \text{if } x_i = B. \end{cases}$$

In addition to $\alpha_i(x)$, we define $\underline{\alpha}_i(x)$ and $\bar{\alpha}_i(x)$ as in Section 1.5.5.1:

$$\begin{aligned} \bar{\alpha}_i(x) &= \begin{cases} (1-\rho)\, C_i + \rho D_i & \text{if } x_i = G, \\ (1-3\rho)\, D_i + 3\rho C_i & \text{if } x_i = B. \end{cases} \\ \underline{\alpha}_i(x) &= \begin{cases} (1-3\rho)\, C_i + 3\rho D_i & \text{if } x_i = G, \\ (1-\rho)\, D_i + \rho C_i & \text{if } x_i = B. \end{cases} \end{aligned}$$

Intuitively, in equilibrium, with a high probability, player i takes $\alpha_i(x)$. However, with some small probability, player i takes $\bar{\alpha}_i(x)$ (taking C_i more often) or $\underline{\alpha}_i(x)$ (taking D_i more often): with $\eta > 0$,

- if $\hat{\lambda}_j(l) = G$, then player i believes that the score has not been erroneous and that player i is almost indifferent between any action. As in Section 1.5.5.1, player i takes $\alpha_i(x)$ with a high probability.
  - with probability $1-\eta$, player i takes $\alpha_i(l) = \alpha_i(x)$;
  - with probability $\frac{\eta}{2}$, player i takes $\alpha_i(l) = \bar{\alpha}_i(x)$; and
  - with probability $\frac{\eta}{2}$, player i takes $\alpha_i(l) = \underline{\alpha}_i(x)$; and
- if $\hat{\lambda}_j(l) = B$, then player i believes that the score has been erroneous and that player i should take a static best response to player j's action plan. Therefore, player i with $\hat{\lambda}_j(l) = B$ takes D_i with probability one.

In the supplemental round for $\lambda_i(l+1)$, player i sends the message $\lambda_i(l+1)$ truthfully via error-reporting noisy cheap talk with precision $p = 1 - \exp(-T^{\frac{1}{2}})$. We assume that the players cannot manipulate p.[2] That is, in the supplemental round for $\lambda_i(l+1)$, only the error-reporting noisy cheap talk with $p = 1 - \exp(-T^{\frac{1}{2}})$ is available.

2.4.3 Reward Function

In this subsection, we explain player j's reward function on player i, $\pi_i^{\text{main}}(x_j, h_j^{\text{main}} : \delta)$.

Reward Function The reward $\pi_i^{\text{main}}(x_j, h_j^{\text{main}} : \delta)$ is written as

$$\pi_i^{\text{main}}(x_j, h_j^{\text{main}} : \delta) = \sum_{l=1}^{L} \sum_{t \in T(l)} \pi_i^{\delta}(a_{j,t}, y_{j,t}) + \begin{cases} -2\bar{u}T + \sum_{l=1}^{L} \pi_i^{\text{main}}(x, h_j^{\text{main}}, l) & \text{if } x_j = G, \\ 2\bar{u}T + \sum_{l=1}^{L} \pi_i^{\text{main}}(x, h_j^{\text{main}}, l) & \text{if } x_j = B. \end{cases} \tag{2.12}$$

Note that the first term cancels out the effect of discounting. Intuitively, $\pi_i^{\text{main}}(x, h_j^{\text{main}}, l)$ is the reward for the lth review round.

Reward Function for the lth Review Round Next we define $\pi_i^{\text{main}}(x, h_j^{\text{main}}, l)$ for each $l = 1, ..., L$. There are following two cases: in the lth review round,

1. if $\theta_j(l) = B$, then player j makes player i indifferent between any action profile by

$$\pi_i^{\text{main}}(x, h_j^{\text{main}}, l) = \sum_{t \in T(l)} \pi_i^{x_j}[\alpha_j(l)](y_{j,t}). \tag{2.13}$$

   Remember that by (2.8),

$$\sum_{t \in T(l)} \pi_i^{x_j}[\alpha_j(l)](y_{j,t}) \begin{cases} \leq 0 & \text{if } x_j = G, \\ \geq 0 & \text{if } x_j = B. \end{cases}$$

   Intuitively, this means (1.17) is not an issue after $\theta_j(l) = B$; and

2. otherwise, that is, if $\theta_j(l) = G$, then consider the following two subcases:

[2] The same constraint is applicable whenever a player sends a message via error-reporting noisy cheap talk with precision $p \in (0, 1)$.

(a) if player j has observed an erroneous score, that is, if $\lambda_j(l) = B$, then $\pi_i^{\text{main}}(x, h_j^{\text{main}}, l)$ is a constant $\bar{\pi}_i(x, \lambda_j(l), l)$ to be determined; and

(b) if player j has not observed an erroneous score, that is, if $\lambda_j(l) = G$, then player j monitors player i by player j's score about player i in the lth review round, denoted by $X_j(l)$.

As we will see, if 2-(b) is the case, then player j takes $\alpha_j(x)$ in the lth review round.

We formally define $X_j(l)$. Intuitively, $X_j(l)$ is the summation of reward that makes any action optimal to player i: $\sum_{t \in T(l)} \pi_i[\alpha(x)](y_{j,t})$, except that player j keeps period $t_j(l)$ separated as mentioned in Section 2.2. That is, player j's score about player i is defined as

$$X_j(l) \equiv \sum_{t \in T_j(l)} \pi_i[\alpha(x)](y_{j,t}). \tag{2.14}$$

Then,

$$\pi_i^{\text{main}}(x, h_j^{\text{main}}, l) = \bar{\pi}_i(x, \lambda_j(l), l) + X_j(l) + \pi_i^{x_j}[\alpha_j(x)](y_{j,t_j(l)}). \tag{2.15}$$

Note that the reward in the separated period $t_j(l)$, $\pi_i^{x_j}[\alpha_j(x)](y_{j,t_j(l)})$, makes player i indifferent between any action in period $t_j(l)$.

Therefore, in total, if 2 is the case, then

$$\pi_i^{\text{main}}(x, h_j^{\text{main}}, l) = \begin{cases} \bar{\pi}_i(x, \lambda_j(l), l) + X_j(l) + \pi_i^{x_j}[\alpha_j(x)](y_{j,t_j(l)}) & \text{if } \lambda_j(l) = G, \\ \bar{\pi}_i(x, \lambda_j(l), l) & \text{if } \lambda_j(l) = B. \end{cases} \tag{2.16}$$

Here, $\bar{\pi}_i(x, \lambda_j(l), l)$ is a constant with

$$\bar{\pi}_i(x, \lambda_j(l), l) \begin{cases} \leq 0 & \text{if } x_j = G, \\ \geq 0 & \text{if } x_j = B \end{cases} \tag{2.17}$$

that will be determined in Section 2.7 so that (1.21), (1.16) and (1.17) are satisfied.

2.4.4 Transition of the States

In this subsection, we explain the transition of the players' states. Since x is fixed in the phase, we consider the following four states:

2.4.4.1 Transition of $\lambda_j(l+1) \in \{G, B\}$

As mentioned in Section 1.5, $\lambda_j(l+1) \in \{G, B\}$ is player j's index of the past erroneous score. The initial condition is $\lambda_j(1) = G$. Inductively, given $\lambda_j(l) \in \{G, B\}$, $\lambda_j(l+1)$ is determined as follows: if $\lambda_j(l) = B$, then $\lambda_j(l+1) = B$. That is, once $\lambda_j(l) = B$ happens, it lasts until the end of the phase. If $\lambda_j(l) = G$, then $\lambda_j(l+1) = G$ if and only if the score in the lth review round is not erroneous. That is,

1. if

$$X_j(l) \begin{cases} \leq \frac{\bar{u}}{L}T & \text{if } x_j = G, \\ \geq -\frac{\bar{u}}{L}T & \text{if } x_j = B. \end{cases}, \tag{2.18}$$

   then $\lambda_j(l+1) = G$; and

2. if

$$X_j(l) \begin{cases} > \frac{\bar{u}}{L}T & \text{if } x_j = G, \\ < -\frac{\bar{u}}{L}T & \text{if } x_j = B. \end{cases}, \tag{2.19}$$

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

Let us check feasibility (1.17). First, after $\lambda_i(l+1) = B$, (2.17) implies

$$\pi_i^{\text{main}}(x, h_j^{\text{main}}, l) \begin{cases} \leq 0 & \text{if } x_j = G, \\ \geq 0 & \text{if } x_j = B. \end{cases}$$

Therefore, to show (1.17), by (2.12) and (2.17), it suffices to show that, with $\bar{l}$ being the last review round with $\lambda_j(l) = G$ (if $\lambda_j(l) = G$ for all l, then $\bar{l} = L$),

$$\sum_{l=1}^{\bar{l}} X_j(l) \begin{cases} \leq 2\bar{u}T & \text{if } x_j = G, \\ \geq -2\bar{u}T & \text{if } x_j = B. \end{cases}$$

Note that, except for the $\bar{l}$th review round, (2.18) holds. In addition, by Lemma 12, for l with $\lambda_j(l) = G$,

$$X_j(l) \begin{cases} \leq \bar{u}T & \text{if } x_j = G, \\ \geq -\bar{u}T & \text{if } x_j = B. \end{cases}$$

In total,

$$\sum_{l=1}^{\bar{l}} X_j(l) \begin{cases} \leq \left(\bar{l}-1\right) \frac{\bar{u}}{L} T + \bar{u}T \leq 2\bar{u}T & \text{if } x_j = G, \\ \geq -\left(\bar{l}-1\right) \frac{\bar{u}}{L} T - \bar{u}T \geq -2\bar{u}T & \text{if } x_j = B, \end{cases}$$

as desired.

2.4.4.2 Transition of $\hat{\lambda}_j(l+1) \in \{G, B\}$

As we have mentioned in Section 2.4.1, $\hat{\lambda}_j(l+1) \in \{G, B\}$ is the partition of player i's histories. Intuitively, player i believes that $\lambda_j(l+1) = \hat{\lambda}_j(l+1)$ with a high probability.

Since $\lambda_j(1) = G$ is common knowledge, define $\hat{\lambda}_j(1) = G$. We define $\hat{\lambda}_j(l)$ inductively. If $\hat{\lambda}_j(l) = B$, then $\hat{\lambda}_j(l+1) = B$. Hence, once $\hat{\lambda}_j(l) = B$ happens, it lasts until the end of the phase. Hence, we concentrate on how $\hat{\lambda}_j(l+1) \in \{G, B\}$ is defined conditional on $\hat{\lambda}_j(l) = G$.

As we have explained in Section 2.4.2, player j sends $\lambda_j(l+1)$ via error-reporting noisy cheap talk and player i observes a signal $f[i](\lambda_j(l+1)) \in \{G, B\}$. Player i constructs $\hat{\lambda}_j(l+1)$ from her history in the lth review round and $f[i](\lambda_j(l+1))$.

Suppose player i's history in the lth review round satisfies the following three conditions:

1. player i takes $\alpha_i(l) = \alpha_i(x)$;

2. the empirical distribution of $a_{i,t}$'s is close to $\alpha_i(x)$; that is,

$$\left\| \frac{1}{T-1} \sum_{t \in T_i(l)} \mathbf{1}_{a_{i,t}} - \alpha_i(x) \right\| < \varepsilon \tag{2.20}$$

   with ε being a small number to be determined; and

3. player i's signal frequency in the periods where player i takes $a_i(x)$ in $T_i(l)$ is close to the affine hull of player i's signal distributions with respect to player j's action.

Before formally defining Condition 3, let us explain the intuitive meaning of these conditions. With these conditions, player i's history is normal as explained in Section 1.5.5.1.

First, from Condition 1 and 2, for sufficiently small ρ and ε, player i takes $a_i(x)$ for most of the time. Hence, we concentrate on the set of periods where player i takes $a_i(x)$ in $T_i(l)$, denoted by

$$T_i(l, x).$$

Second, if player i observes a signal frequency close to the ex ante mean under $\alpha_j(x)$, then player i believes that if player j takes $\alpha_j(l) = \alpha_j(x)$, then player j also observes the signal frequency close to the ex ante mean with a high probability by the law of large numbers. This means player j's score is not erroneous. In addition, if player j takes $\alpha_j(l) \neq \alpha_j(x)$, then player i will be indifferent between any action profile sequence as will be seen in Section 2.4.4.3. Therefore, when player i is told to defect from the error-reporting noisy cheap talk ($f[i](\lambda_j(l+1)) = B$), player i believes that this is an error with a high probability.

Third, if player i observes a signal frequency far from the ex ante mean under $\alpha_j(x)$, then since player i's signal frequency is close to the affine hull of player i's signal distributions with respect to player j's action, player i's signal frequency is skewed toward either C_j or D_j. If player i's signal frequency is skewed toward C_j, then player i believes that player j takes $\bar{\alpha}_j(x)$ rather than $\alpha_j(x)$ and if player i's signal frequency is skewed toward D_j, then player i believes that player j takes $\underline{\alpha}_j(x)$. Again, after player j takes $\alpha_j(l) \neq \alpha_j(x)$, as we will formally define in Section 2.4.4.3, player i will be indifferent between any action profile sequence. Therefore, when player i receives $f[i](\lambda_j(l+1))$, player i believes that this is irrelevant information.

Therefore, it is almost optimal for player i to disregard $f[i](\lambda_j(l+1))$ and to have $\hat{\lambda}_j(l+1) = G$.

In other cases, player i obeys the message: $\hat{\lambda}_j(l+1) = f[i](\lambda_j(l+1))$.

We are left to formally define Condition 3. Recall that $q_i(\alpha)$ is the vector of player i's signal distribution given α, $q_i(\alpha) \equiv (q_i(y_i \mid \alpha))_{y_i \in Y_i}$. In particular, we define

$$q_i(x) \equiv q_i(a_i(x), \alpha_j(x)).$$

In addition, let $\boldsymbol{Q}_i(a_i)$ be the affine hull of player i's signal distributions with respect to player j's action given a_i:

$$\boldsymbol{Q}_i(a_i) \equiv \text{aff}\left(\{q_i(a_i, a_j)\}_{a_j \in A_j}\right) \cap \mathbb{R}_+^{|Y_i|}.$$

Here, since the signal frequency should be non-negative, we restrict our attention to $\mathbb{R}_+^{|Y_i|}$. In particular, with $a_i = a_i(x)$, we define

$$\boldsymbol{Q}_i(x) \equiv \boldsymbol{Q}_i(a_i(x)).$$

We also consider the matrix representation of $\boldsymbol{Q}_i(a_i)$:

$$\boldsymbol{Q}_i(a_i) = \left\{ \mathbf{y}_i \in \mathbb{R}_+^{|Y_i|} : Q_i(a_i)\mathbf{y}_i = \mathbf{q}_i\left(a_i\right) \right\}.$$

Note that all the signal frequencies should be on the simplex over Y_i. Hence, by affine transformation, we can assume each element of $Q_i\left(a_i\right)$ and $\mathbf{q}_i\left(a_i\right)$ is included in $(0, 1)$:

Lemma 14 For each $i \in I$ and a_i, we can take $Q_i(a_i)$ such that all the elements are in $(0, 1)$.

Proof: See Section 2.9. ∎

With $T_i(l, x)$ being the set of periods when player i takes $a_i\left(x\right)$ in the lth review round, Condition 3 is equivalent to the condition that the distance between player i's signal frequency and $\boldsymbol{Q}_i(x)$ is small:

$$\left\| \boldsymbol{Q}_i(x) - \frac{1}{|T_i(l,x)|} \sum_{t \in T_i(l,x)} \mathbf{1}_{y_{i,t}} \right\| < \varepsilon. \tag{2.21}$$

Instead of using (2.21) directly, we consider the following procedure. First, with $L_{i,x}$ being the number of rows for $Q_i(x)$, we define an $L_{i,x} \times 1$ vector $\mathbf{1}_{Q_i(x)}$ as follows:[3] after observing y_i, player i calculates $Q_i(x)\mathbf{1}_{y_i}$. Then, player i draws $L_{i,x}$ random variables from the uniform $[0, 1]$ independently. If the lth realization of this random variable is no less than the lth element of $Q_i(x)\mathbf{1}_{y_i}$, we define the lth element of $\mathbf{1}_{Q_i(x)}$ equal to 1. Otherwise, the lth element of $\mathbf{1}_{Q_i(x)}$ is 0. By definition, the distribution of $\mathbf{1}_{Q_i(x)}$ is independent of player j's action.

Condition 3 is satisfied if

[3] Remember that, in Chapter 1, for a random variable $x \in X$, we define $\mathbf{1}_x$ as a $|X| \times 1$ vector such that, if $x = \hat{x}$, the element corresponding to $\hat{x}$ is equal to one and the other elements are zero. Here, $\mathbf{1}_{Q_i(x)}$ is a vector each of whose element can be either 0 or 1. These are not contradictory since $Q_i(x)$ is not a random variable.

- $Q_i(x)\left(\frac{1}{|T_i(l,x)|}\sum_{t\in T_i(l,x)}\mathbf{1}_{y_{i,t}}\right)$ and $\frac{1}{|T_i(l,x)|}\sum_{t\in T_i(l,x)}\mathbf{1}_{Q_i(x)}$ are close:

$$\left\| Q_i(x)\left(\frac{1}{|T_i(l,x)|}\sum_{t\in T_i(l,x)}\mathbf{1}_{y_{i,t}}\right) - \frac{1}{|T_i(l,x)|}\sum_{t\in T_i(l,x)}\mathbf{1}_{Q_i(x)} \right\| < \frac{\varepsilon}{K_1}; \tag{2.22}$$

  and

- $\frac{1}{|T_i(l,x)|}\sum_{t\in T_i(l,x)}\mathbf{1}_{Q_i(x)}$ and $\mathbf{q}_i(x)$ are close:

$$\left\| \frac{1}{|T_i(l,x)|}\sum_{t\in T_i(l,x)}\mathbf{1}_{Q_i(x)} - \mathbf{q}_i(x) \right\| < \frac{\varepsilon}{K_1}. \tag{2.23}$$

By Lipschitz continuity of $Q_i(x)$ and triangle inequality, for sufficiently large K_1, (2.22) and (2.23) implies (2.21). In addition, by definition, the probability of (2.23) is independent of player j's action.

Consider (2.22). For $\rho, \varepsilon < \frac{1}{4}$, if Conditions 1 and 2 are satisfied, the length of $T_i(l,x)$ is proportional to T:

$$|T_i(l,x)| = \Theta(T).$$

Hence, by the law of large numbers, conditional on player i's history $\{a_{i,t}, y_{i,t}\}$, (2.22) does not hold with probability

$$p[Q_i(x)](\{a_{i,t}, y_{i,t}\}_{t\in T_i(l)}) = \exp(-\Theta(T)).$$

We want to make this probability completely independent of player j's action plan. For that purpose, let $\bar{p}[Q_i]$ be the maximum of $p[Q_i(x)](\{a_{i,t}, y_{i,t}\}_{t\in T_i(l)})$ with respect to $\{a_{i,t}, y_{i,t}\}_{t\in T_i(l)}$ and x. For some $\{a_{i,t}, y_{i,t}\}_{t\in T_i(l)}$ and x, if $p[Q_i(x)](\{a_{i,t}, y_{i,t}\}_{t\in T_i(l)})$ is less than $\bar{p}[Q_i]$, then player i draws a random variable from the uniform $[0,1]$. If this realization is no less than $\bar{p}[Q_i] - p[Q_i(x)](\{a_{i,t}, y_{i,t}\}_{t\in T_i(l)})$, then player i behaves as if (2.22) were not satisfied.

Then, in total, player i behaves as if (2.22) is not satisfied with probability $\bar{p}[Q_i]$. From now on, when we say (2.22) is satisfied, it means that

$$\left\| Q_i(x) \left(\frac{1}{|T_i(l,x)|} \sum_{t \in T_i(l,x)} \mathbf{1}_{y_{i,t}} \right) - \frac{1}{|T_i(l,x)|} \sum_{t \in T_i(l,x)} \mathbf{1}_{Q_i(x)} \right\| < \frac{\varepsilon}{K_1}$$

and the realization of the random variable explained above is more than $\bar{p}[Q_i] - p[Q_i(x)]($ $\{a_{i,t}, y_{i,t}\}_{t \in T_i(l)})$. This implies that the probability of (2.22) is independent of player j's action plan.

In summary, we define the transition of $\hat{\lambda}_j(l+1)$ as follows: if player i's history in the lth review round satisfies the following two conditions, then player i disregards the message and $\hat{\lambda}_j(l+1) = G$:

1. player i takes $\alpha_i(l) = \alpha_i(x)$; and
2. (2.20), (2.22) and (2.23) are satisfied, which implies (2.21).

Otherwise, player i obeys the message and

$$\hat{\lambda}_j(l+1) = f[i](\lambda_j(l+1)). \tag{2.24}$$

2.4.4.3 Transition of $\theta_j(l+1) \in \{G, B\}$

As we have seen in Section 2.4.3, $\theta_j(l+1) = B$ implies that player j uses the reward (2.13) and player i is indifferent between any action profile in the $(l+1)$th review round by (2.9) (except for the incentives from π_i^{report}).

If $\theta_j(l) = B$, then $\theta_j(l+1) = B$. That is, once $\theta_j(l) = B$ happens, it lasts until the end of the phase. Hence, we concentrate on how $\theta_j(l+1) \in \{G, B\}$ is defined conditional on $\theta_j(l) = G$.

$\theta_j(l+1) = B$ if one of the following four conditions is satisfied:

1. when player j sends $\lambda_j(l+1)$ by the error-reporting noisy cheap talk, the error is reported:
$$g[j](\lambda_j(l+1)) = E;$$

2. at the beginning of the $(l+1)$th review round, player j with $\hat{\lambda}_j(l+1) = G$ takes $\alpha_j(l+1) \neq \alpha_j(x)$. With abuse of notation, for this condition, we include the case with $l+1 = 1$ (that is, $l = -1$); and

3. (2.20), (2.22) or (2.23) is not satisfied for player j (with indices i and j reversed).

Otherwise, $\theta_j(l+1) = G$.

We summarize the implications of the transitions of θ_j.

First, player j makes player i indifferent between any action profile after receiving $g[j](\lambda_j(l+1)) = E$. Since player i believes that, whenever her signal is wrong: $f[i](\lambda_j(l+1)) \neq \lambda_j(l+1)$, player j receives $g[j](\lambda_j(l+1)) = E$ and so any action will be optimal with a high probability. Therefore, (2.24) is an almost optimal inference.

Second, consider how player j constructs $\hat{\lambda}_i(l+1)$. Reversing the indices i and j in Sections 2.4.4.2, whenever player j uses the signal of the error-reporting noisy cheap talk $f[j](\lambda_i(l+1))$, $\theta_j(l+1) = B$ has been already determined. This implies that player j makes player i indifferent between any action profile (including player j's action plan), whenever player i's message has an impact on player j's continuation action plan. Hence, player i is indifferent between any message.

Third, consider all the cases where player j does not take $\alpha_j(x)$ in the $(l+1)$th review round:

- if $\alpha_j(l+1) \neq \alpha_j(x)$ with $\hat{\lambda}_i(l+1) = G$, then Condition 2 is the case and so $\theta_j(l+1) = B$; and

- if $\hat{\lambda}_i(l+1) = B$, then by the above discussion, player j has obeyed player i's message, which implies $\theta_j(l+1) = B$.

In total, if $\alpha_j(l+1) \neq \alpha_j(x)$, then player i is indifferent between any action profile.

Fourth, the distribution of $\theta_j(l+1)$ is independent of player i's action plan. To see why, consider each of the three conditions inducing $\theta_j(l+1) = B$:

1. the probability of error is $\exp(-T^{\frac{1}{2}})$ for all $\lambda_j(l+1)$;

2. with $\hat{\lambda}_i(l+1) = G$, $\alpha_j(l+1)$ is determined by player j's own randomization; and

3. (2.20) is determined by player j's own randomization and (2.22) and (2.23) are independent of player i's action plan (with indices i and j reversed).

2.5 Player i's Belief about Optimal Actions

Given the above strategy, we want to formally show that player i believes that player i's inference of $\lambda_j(l+1)$ is correct ($\hat{\lambda}_j(l+1) = \lambda_j(l+1)$) or any action is optimal ($\theta_j(l+1) = B$) with high probability $1-\exp(-\Theta(T^{\frac{1}{2}}))$, conditional on that player j has the state $\hat{\lambda}_i(l+1) = G$. Notice that if $\hat{\lambda}_i(l+1) = B$, then player j has obeyed player i's message and any action is optimal for player i ($\theta_j(l+1) = B$).

First, consider the case where player i obeys the signal $f[i](\lambda_j(l+1))$. Suppose that player i could know $\lambda_j(l+1)$ (she cannot in private monitoring). Consider the two possible realizations of the signals in the supplemental round for $\lambda_j(l+1)$. If $f[i](\lambda_j(l+1)) = \lambda_j(l+1)$, then player i receives a correct message. If $f[i](\lambda_j(l+1)) \neq \lambda_j(l+1)$, then with probability $1-\exp(-\Theta(T^{\frac{1}{2}}))$, player j should receive the signal telling that player i did not receive the correct signal, that is, $g[j](\lambda_j(l+1)) = E$. If $g[j](\lambda_j(l+1)) = E$, then as seen in Section 2.4.4.3, $\theta_j(l+1) = B$, as desired.

Notice that, symmetrically to player i's action plan $\sigma_i(x)$, player j's continuation action plan is independent of $g[j](\lambda_j(l+1))$ (except for the report block). Therefore, the learning from player j's continuation action plan does not change player i's belief.

On the other hand, suppose player i disregards the signal $f[i](\lambda_j(l+1))$. As seen in Section 2.4.4.2, player i believes either

1. player j's score is not erroneous with a high probability if player i observes a signal frequency close to the ex ante mean under $\alpha_j(x)$; or

2. player j takes either $\bar{\alpha}_j(x)$ or $\underline{\alpha}_j(x)$ rather than $\alpha_j(x)$ with a high probability if player i observes a signal frequency far from the ex ante mean under $\alpha_j(x)$.

Since the length of the review round is T, the belief on the above two events is no less than $1 - \exp(-\Theta(T))$, before learning from player j's continuation action plan.

The learning from player j's continuation action plan changes the belief in the following two ways. First, $\left\{f[i](\lambda_j(\tilde{l}))\right\}_{\tilde{l}=1}^{L}$ reveals $\left\{\lambda_j(\tilde{l})\right\}_{\tilde{l}=1}^{L}$. However, since the error occurs with a positive probability $\left(\exp(-\Theta(T^{\frac{1}{2}}))\right)^{L} = \exp(-\Theta(T^{\frac{1}{2}}))$, the update of the belief is sufficiently small compared to the original belief $1 - \exp(-\Theta(T))$.

Second, player i in the $\tilde{l}$th review round conditions that $\hat{\lambda}_j(\tilde{l}) = G$ (as mentioned, otherwise, $\theta_j(\tilde{l}) = B$ and player i is indifferent between any action profile). This conditioning changes player i's belief on player j's history in the following way: player i's belief on player j's history where player j obeys player i's message $f[j](\lambda_i(\tilde{l}))$ can be decreased. For example, if player i sends $\lambda_i(\tilde{l}) = B$ and the error is not reported, then the probability that $f[j](\lambda_i(\tilde{l})) = G$ is small. However, this decrease is at most by $\exp(-\Theta(T^{\frac{1}{2}}))$ since any signal profile can occur with positive probability $\exp(-\Theta(T^{\frac{1}{2}}))$ in the supplemental round for $\lambda_i(\tilde{l})$. On the other hand, player i's belief on player j's history where player j disregards player i's message $f[j](\lambda_i(\tilde{l}))$ remain unchanged. In total, the update of the belief is $\left(\exp(-\Theta(T^{\frac{1}{2}}))\right)^{L} = \exp(-\Theta(T^{\frac{1}{2}}))$, which is sufficiently small compared to the original belief $1 - \exp(-\Theta(T))$.

Formally, we can show the following lemma:

Lemma 15 For all $\bar{u}$ and L, there exists $\bar{\eta}$ such that, for all $\eta < \bar{\eta}$, there exist $\bar{\rho}$, $\bar{\varepsilon}$ such that, for all $\rho < \bar{\rho}$ and $\varepsilon < \bar{\varepsilon}$, for any history h_i^t with t being in the lth review round, conditional on $\hat{\lambda}_i(\tilde{l}) = G$ for all $\tilde{l} \leq l$ and $\alpha_j(l)$, player i after h_i^t believes that $\hat{\lambda}_j(l) = \lambda_j(l)$ or $\theta_j(l) = B$.

Proof: See Section 2.9. ∎

As we have mentioned, the statement is conditional on $\hat{\lambda}_i(\tilde{l}) = G$ for all $\tilde{l} \leq l$. We also condition on $\alpha_j(l)$ since given $\hat{\lambda}_i(l) = G$, $a_j(l)$ is independent of $\lambda_j(l)$.

2.6 Variables

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

First, $\bar{u}$ is determined in Lemma 12, independently of the other variables (note that $\bar{u}$ is independent of $\alpha_j(x)$).

Second, fix L so that

$$\max\left\{\max_{x:x_j=B} u_i(a(x)), v_i^*\right\} + 2\frac{\bar{u}}{L} < \underline{v}_i < \bar{v}_i < \min_{x:x_j=G} u_i(a(x)) - 2\frac{\bar{u}}{L}. \tag{2.25}$$

This is possible because of (1.18).

Third, given $\bar{u}$ and L, fix $\bar{\eta}$ so that Lemma 15 holds and for all $\eta < \bar{\eta}$,

$$\begin{aligned} &\max\left\{\max_{x:x_j=B} u_i(a(x)), v_i^*\right\} + 2\frac{\bar{u}}{L} + \eta L\left(2\bar{u} - \min\left\{\min_{i,x} u_i(a(x)), v_i^*\right\}\right) \\ < \quad & \underline{v}_i < \bar{v}_i < \min_{x:x_j=G} u_i(a(x)) - 2\frac{\bar{u}}{L} - \eta L\left(2\bar{u} + \max_{i,x} u_i(a(x))\right). \end{aligned} \tag{2.26}$$

This is possible because of (2.25). As explained in Section 2.4.2, a small η implies that the event that player i with $\hat{\lambda}_j(l) = G$ takes $\alpha_i(l) = \alpha_i(x)$ with a high probability.

Fourth, fix $\eta < \bar{\eta}$. Then, we can take $\bar{\rho}$ and $\bar{\varepsilon}$ so that Lemma 15 holds. Take $\varepsilon < \bar{\varepsilon}$ and $\rho < \bar{\rho}$ so that

$$\begin{aligned} &\max\left\{\max_{x:x_j=B} w_i(x), v_i^*(\rho)\right\} + 2\frac{\bar{u}}{L} + \eta L\left(2\bar{u} - \min\left\{\min_{x} w_i(x), v_i^*(\rho)\right\}\right) \\ < \quad & \underline{v}_i < \bar{v}_i < \min_{x:x_j=G} w_i(x) - 2\frac{\bar{u}}{L} - \eta L\left(2\bar{u} + \max_{i,x} w_i(x)\right). \end{aligned} \tag{2.27}$$

Here,

$$v_i^*(\rho) \equiv \max_{x:x_j=B,a_i} u_i(a_i, \alpha_j(x))$$

is the maximum payoff that player i can get if $x_j = B$ and player j takes $\alpha_j(x)$. Since $\alpha_j(x)$ with $x_j = B$ takes D_j with probability $1 - 2\rho$, for sufficiently small ρ, $v_i^*(\rho)$ is sufficiently close to the minimax value v_i^*. Together with (1.19), for sufficiently small ρ, (2.26) implies (2.27).

Since $T_P = LT$ and $T = (1-\delta)^{-\frac{1}{2}}$, we have

$$\lim_{\delta \to 1} \delta^{T_P} = 1.$$

Therefore, discounting for the payoffs in the next review phase goes to zero.

2.7 Almost Optimality of $\sigma_i(x_i)$

We have defined $\sigma_i(x_i)$ and π_i^{main} except for $\bar{\pi}_i(x, \lambda_j(l), l)$. In this section, based on Lemma 15, we show that if we properly define $\bar{\pi}_i(x, l)$, then $\sigma_i(x_i)$ and π_i^{main} satisfy (1.21), (1.16) and (1.17):

Proposition 16 For sufficiently large δ, there exists $\bar{\pi}_i(x, \lambda_j(l), l)$ such that

1. $\sigma_i(x_i)$ is almost optimal conditional on $\hat{\lambda}_i(l) = G$: for each $l \in \{1, ..., L\}$, conditional on $\hat{\lambda}_i(l) = G$,

    (a) for any period t in the lth review round, (1.21) holds; and

    (b) when player i sends the message about $\lambda_i(l+1)$ by the error-reporting noisy cheap talk, (1.21) holds;[4]

[4] With $l = L$, this is redundant.

2. (1.16) is satisfied with π_i replaced with π_i^{main}. Since each $x_i \in \{G, B\}$ gives the same value conditional on x_j, the action plan in the coordination block is optimal;[5] and

3. π_i^{main} satisfies (1.17).

Proof: See Section 2.9. ∎

Here, we offer the intuitive explanation. First, we construct $\bar{\pi}_i(x, \lambda_j(l), l)$, assuming that the players follow $\sigma_i(x_i)$. We want to make sure that

- for (1.16), player i's value from the lth review round

$$\frac{1}{T}\mathbb{E}\left[\sum_{t \in T(l)} u_i(a_t) + \pi_i^{\text{main}}(x, h_j^{\text{main}}, l) \mid x\right]$$

  is close to $\bar{v}_i$ ($\underline{v}_i$, respectively) if $x_j = G$ ($x_j = B$, respectively), $\hat{\lambda}_j(l) = \lambda_j(l)$ and $\theta_j(\tilde{l}) = G$ for all $\tilde{l} \leq l$. Note that the last condition implies that player j takes $\alpha_j(l) = \alpha_j(x)$; and

- for (1.17), $\bar{\pi}_i(x, \lambda_j(l), l) \leq 0$ ($\bar{\pi}_i(x, \lambda_j(l), l) \geq 0$, respectively) if $x_j = G$ ($x_j = B$, respectively), as seen in (2.19).

If $\hat{\lambda}_j(l) = \lambda_j(l) = G$, then the players take $\alpha(x)$ and from (2.5), player i's value from the lth review round is close to $w_i(x)$ except for $\bar{\pi}_i(x, \lambda_j(l), l)$. Hence, if we determine

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

properly, we can make sure that player i's value from the lth review round is close to $\bar{v}_i$ if $x_j = G$ and $\underline{v}_i$ if $x_j = B$.

[5] This is not precise since we will further adjust the reward function based on the report block. However, as we will see, even after the adjustment of the report block, any $x_i \in \{G, B\}$ still gives exactly the same value and so the strategy in the coordination block is exactly optimal.

If $x_j = G$ and $\hat{\lambda}_j(l) = \lambda_j(l) = B$, then player i takes a best response to player j's action $\alpha_j(x)$ and the average instantaneous utility during the lth review round is more than $w_i(x)$. The reward is 0 except for $\bar{\pi}_i(x, \lambda_j(l), l)$. Therefore, from (2.27), player i's value from the lth review round is close to $\bar{v}_i$ if we properly determine $\bar{\pi}_i(x, \lambda_j(l), l) \leq 0$.

If $x_j = B$ and $\hat{\lambda}_j(l) = \lambda_j(l) = B$, then player i takes a best response to player j's action $\alpha_j(x)$. Since $\alpha_j(x)$ is close to the minimaxing action D_j, the average instantaneous utility during the lth review round is $v_i^*(\rho) \leq \underline{v}_i$. The reward is 0 except for $\bar{\pi}_i(x, \lambda_j(l), l)$. Therefore, from (2.27), player i's value from the lth review round is close to $\underline{v}_i$ if we properly determine $\bar{\pi}_i(x, \lambda_j(l), l) \geq 0$.

Second, we verify 1-(a): in the lth review round, it is almost optimal for player i to follow $\sigma_i(x_i)$, conditional on $\hat{\lambda}_i(\tilde{l}) = G$ for all $\tilde{l}$. Lemma 15 guarantees that, for almost optimality, player i can assume $\lambda_j(l) = \hat{\lambda}_j(l)$ or $\theta_j(l) = B$. If the latter is the case, then any action is optimal. Hence, we concentrate on the case with $\lambda_j(l) = \hat{\lambda}_j(l)$ and $\theta_j(l) = G$, which implies that player j takes $\alpha_j(x)$.

For the last Lth review round, player i maximizes

$$\frac{1}{T}\mathbb{E}\left[\sum_{t \in T(l)} u_i(a_t) + \pi_i^{\text{main}}(x, h_j^{\text{main}}, l) \mid x\right] \tag{2.28}$$

with $l = L$. If $\lambda_j(L) = \hat{\lambda}_j(L) = G$, then $\pi_i^{\text{main}}(x, h_j^{\text{main}}, L)$ is the summation of $\pi_i[\alpha(x)](y_{j,t})$, which makes any action optimal for player i. If $\lambda_j(L) = \hat{\lambda}_j(L) = B$, then $\pi_i^{\text{main}}(x, h_j^{\text{main}}, L)$ is constant and so player i wants to take a static best response to $\alpha_j(x)$. Therefore, D_i is almost optimal in order to maximize (2.28). Therefore, $\sigma_i(x_i)$ is almost optimal for the Lth review round.

We proceed backward. Suppose that player i follows $\sigma_i(x_i)$ from the $(l+1)$th review round and consider player i's incentive in the lth review round. Note that we define $\bar{\pi}_i$ such that player i's value is almost independent of $\lambda_j(l+1)$ as long as player i follows $\sigma_i(x_i)$

from the $(l+1)$th review round and "$\lambda_j(l+1) = \hat{\lambda}_j(l+1)$ or $\theta_j(l+1) = B$."[6] In addition, Lemma 15 implies that player i in the main blocks believes that "$\lambda_j(l+1) = \hat{\lambda}_j(l+1)$ or $\theta_j(l+1) = B$" with a high probability. Further, Section 2.4.4.3 guarantees that the distribution of $\theta_j(l+1) = B$ is independent of player i's action plan. Therefore, for almost optimality, we can assume that player i in the lth review round maximizes (2.28), assuming that $\lambda_j(l) = \hat{\lambda}_j(l)$ and $\theta_j(l) = G$. Therefore, the same argument as for the Lth review round establishes that $\sigma_i(x_i)$ is almost optimal for the lth review round.

Third, 1-(b) is true since, as seen in Section 2.4.4.3, whenever player i's message affects player j's continuation action, player i has been indifferent between any action profile. In addition, although player i's message and signal observation affect player i's posterior about the optimality of $\hat{\lambda}_j(l+1)$, the effect is sufficiently small for almost optimality (see the discussion in Section 2.5).

Fourth, 2 is true since except for rare events $\hat{\lambda}_j(l) = B$, $\hat{\lambda}_i(l) = B$, $\theta_i(l) = B$ or $\theta_j(l) = B$, the players take $\alpha(x)$ in the lth review round and player j uses the reward (2.16). By construction of $\bar{\pi}_i(x, \lambda_j(l), l)$, the ex ante value is $\bar{v}_i$ ($\underline{v}_i$, respectively) if $x_j = G$ ($x_j = B$, respectively).

Finally, from Section 2.4.4.1, π_i^{main} satisfies 3 since we take $\bar{\pi}_i(x, \lambda_j(l), l) \leq 0$ ($\bar{\pi}_i(x, \lambda_j(l), l) \geq 0$, respectively) if $x_j = G$ (B, respectively).

Therefore, we are left to construct the action plan in the report block and π_i^{report} such that $\sigma_i(x_i)$ and $\pi_i^{\text{main}} + \pi_i^{\text{report}}$ satisfy (1.15), (1.16) and (1.17) based on Proposition 16.

2.8 Exact Optimality

In this section, we explain the action plan and the reward π_i^{report} in the report block.

[6] In the above discussion, we have verified that this claim is correct for the case with $\theta_j(l+1) = G$ (with L replaced with $l+1$).
For $\theta_j(l+1) = B$, player i is indifferent between any action profile sequence, as desired.

2.8.1 Preparation

We start with proving two Lemmas about incentivizing player i to tell the truth about h_i^{main}.

As we will see, player i conditions that player j believes that player i's score about player j is not erroneous: $\hat{\lambda}_j(l) = G$. In this case, player j takes a fully mixed action plan.

When player i reports her history $(a_{i,t}, y_{i,t})$ for some period t in the coordination or main blocks, player j punishes player i proportionally to

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

Here, $\mathbf{1}_{a_{j,t},y_{j,t}}$ is an $|A_j|\,|Y_j| \times 1$ vector whose element corresponding to $(a_{j,t}, y_{j,t})$ is one and other elements are zero. $(\hat{a}_{i,t}, \hat{y}_{i,t})$ is player i's message.

Intuitively,[7] player i wants to maximize

$$-\mathbb{E}\left[\left\| \mathbf{1}_{a_{j,t},y_{j,t}} - \mathbb{E}\left[\mathbf{1}_{a_{j,t},y_{j,t}} \mid \hat{y}_{i,t}, \hat{a}_{i,t}, \alpha_{j,t}\right] \right\|^2 \mid y_{i,t}, a_{i,t}, \alpha_{j,t}\right]. \tag{2.29}$$

We assume that player i knew player j's action plan $\alpha_{j,t}$. Since player j takes a fully mixed action plan conditional on $\hat{\lambda}_i(l) = G$, Assumption 5 implies that the truthtelling is uniquely optimal:

Lemma 17 If Assumption 5 is satisfied, then for each i, any fully mixed action plan $\alpha_{j,t} \in \Delta(A_j)$ and player i's history $(a_{i,t}, y_{i,t}) \in A_i \times Y_i$, $(\hat{a}_{i,t}, \hat{y}_{i,t}) = (a_{i,t}, y_{i,t})$ is a unique maximizer of (2.29).

Proof: By algebra. ∎

Take ex ante value of (2.29) before observing $y_{i,t}$ assuming the truthtelling $(\hat{a}_{i,t}, \hat{y}_{i,t}) = (a_{i,t}, y_{i,t})$:

$$-\mathbb{E}\left[\left\| \mathbf{1}_{a_{j,t},y_{j,t}} - \mathbb{E}\left[\mathbf{1}_{a_{j,t},y_{j,t}} \mid y_{i,t}, a_{i,t}, \alpha_{j,t}\right] \right\|^2 \mid a_{i,t}, \alpha_{j,t}\right]. \tag{2.30}$$

[7] That is, except that player i can learn about $(a_{j,t}, y_{j,t})$ from the continuation play between period t and the report block.

As in Lemma 12, we can show the existence of player j's reward on player i which cancels out the difference in (2.30) for different $a_{i,t}$'s:

Lemma 18 If Assumptions 3 is satisfied, then for any $j \in I$ and $\alpha_j \in \Delta(A_j)$, there exists $\Pi_i[\alpha_j] : Y_j \to \mathbb{R}$ such that

$$\mathbb{E}\left[\Pi_i[\alpha_j](y_j) \mid a_i, \alpha_j\right] - \mathbb{E}\left[\left\|\mathbf{1}_{a_j,y_j} - \mathbb{E}\left[\mathbf{1}_{a_j,y_j} \mid y_i, a_i, \alpha_j\right]\right\|^2 \mid a_i, \alpha_j\right] = 0$$

for all $a_i \in A_i$.

Proof: The same as Lemma 12. ∎

2.8.2 Report Block

Given these two lemmas, we explain the action plan and the reward π_i^{report} in the report block. As briefly mentioned in Sections 1.7 and 1.8, player i reports h_i^{main} to player j if player i is picked by the public randomization. Player j calculates π_i^{report} based on the reported history $\hat{h}_i^{\text{main}}$ so that $\sigma_i(x_i)$ is exactly optimal against $\sigma_j(x_j)$ and $\pi_i^{\text{main}} + \pi_i^{\text{report}}$.

With the perfect cheap talk, the players could report h_i^{main} simultaneously and instantaneously. However, as seen in Section 1.5, for the dispensability of the cheap talk, it is important to construct the report block so that only one player sends the message and that the cardinality of the messages is sufficiently small.

For the first purpose, the players use public randomization. Player 1 reports h_1^{main} if $y^p \leq \frac{1}{2}$ and player 2 reports h_2^{main} if $y^p > \frac{1}{2}$. Below, we consider the case where player i reports the history.

From Section 2.2, there is a chronological order for the rounds. Hence, we can number all the rounds serially. For example, the round for x_1 is round 1, the round for x_2 is round 2, the first review round is round 3, the supplemental round for $\lambda_1(l+1)$ is round 4, the supplemental round for $\lambda_2(l+1)$ is round 5, and so on.

Let h_i^{r+1} be player i's history at the beginning of the $(r+1)$th round. The reward from the report block is the summation of the rewards for each round:

$$\pi_i^{\text{report}}(x_j, h_j^{T_P+1} : \delta) = \sum_r \pi_i^{\text{report}}\left(h_j^{r+1}, \hat{h}_i^{r+1}, r\right).$$

Here, $\hat{h}_i^{r+1}$ is player i's report about h_i^{r+1}. Precisely, to reduce the cardinality of the messages, player i reports the summary of h_i^{r+1}. The details will be determined below. Note that the reward for round r, $\pi_i^{\text{report}}\left(h_j^{r+1}, \hat{h}_i^{r+1}, r\right)$, depends on the history until the end of round r.

We define $\pi_i^{\text{report}}\left(h_j^{r+1}, \hat{h}_i^{r+1}, r\right)$ such that

1. during the main blocks, for each period t and each h_i^t, player i believes that it will be optimal to tell the truth about h_i^{r+1}; and

2. based on the truthful report h_i^{r+1}, π_i^{report} will be adjusted so that $\sigma_i(x_i)$ is exactly optimal.

1 can be achieved by using the punishment similar to (2.29), which will be formally proven in Section 2.9.6. Here, suppose that we have shown the truthtelling incentive 1 and let us concentrate on the adjustment 2.

Since we need to keep the cardinality of the messages sufficiently small, we consider the summary statistics $\#_i^{\tilde{r}}$ for the history in each round $\tilde{r}$: for round 1, let $\#_i^1$ be x_1, the message sent by the perfect cheap talk in round 1. Similarly, for round 2, let $\#_i^2$ be x_2.

For round $\tilde{r}$ corresponding to a review round, for each $(a_i, y_i) \in A_i \times Y_i$, let $\#_i^{\tilde{r}}(a_i, y_i)$ be how many times player i observed an action-signal pair (a_i, y_i) in round $\tilde{r}$. Let $\#_i^{\tilde{r}}$ be a vector $\left(\#_i^{\tilde{r}}(a_i, y_i)\right)_{a_i, y_i}$.

For round $\tilde{r}$ where player i sends a message m via error-reporting noisy cheap talk, let $\#_i^{\tilde{r}}$ be player i's message and signals $(m, g[i](m), g_2[i](m))$.[8]

[8]Although we neglected the secondary signals $f_2[i](m)$ and $g_2[i](m)$ for the almost optimality, for the exact optimality, we need to take into account the belief updates from the secondary signals.

For round $\tilde{r}$ where player i receives a message m via error-reporting noisy cheap talk, let $\#_i^{\tilde{r}}$ be player i's signals $(f[i](m), f_2[i](m))$.

Let $\mathfrak{h}_i^{r+1} \equiv \left\{\#_i^{\tilde{r}}\right\}_{\tilde{r}\leq r}$ be a summary of player i's history at the beginning of round $r+1$.

By backward induction, we construct $\pi_i^{\text{report}}\left(h_j^{r+1}, \hat{h}_i^{r+1}, r\right)$. For round r corresponding to a review round, let $(T(r, a_i))_{a_i \in A_i} \in T^{|A_i|}$ be the set of strategies that take a_i for $T(r, a_i)$ times in round r.

In the last round (the Lth review round), since player j determines her continuation strategy treating each period within a past round identically, player i's belief about player j's continuation strategy at the beginning of the Lth review round conditional on $\alpha_j(L)$ and $\hat{\lambda}_i(L) = G$ is determined by $\mathfrak{h}_i^r$. Conditional on $\alpha_j(L)$, the learning from signal observations is redundant. Hence, conditional on $\alpha_j(L)$ and $\hat{\lambda}_i(L) = G$, player i's value of taking $(T(r, a_i))_{a_i}$ only depends on $\mathfrak{h}_i^r$ regardless of $\{a_{i,t}, y_{i,t}\}$ in the Lth review round. Since Proposition 16 holds conditional on $\alpha_j(L)$ and $\hat{\lambda}_i(L) = G$, for all $(T(r, a_i))_{a_i}$ consistent with the equilibrium strategy, player i's value of taking $(T(r, a_i))_{a_i}$ does not change by more than $\exp(-\Theta(T^{\frac{1}{2}}))$.

On the other hand, from player i's report, player j can know $\mathfrak{h}_i^r$ and how many times player i took a_i in round r. Let $(\#_i^r(a_i))_{a_i}$ be this number. With abuse of notation, we say player i reports that she took $(T(r, a_i))_{a_i}$ if $T(r, a_i)$ is equal to $\#_i^r(a_i)$ for all a_i.

Therefore, based on the report of $\mathfrak{h}_i^r$ and $(T(r, a_i))_{a_i}$, player j can construct $\pi_i^{\text{report}}(h_j^{r+1}, \hat{h}_i^{r+1}, r)$ such that, given $\mathfrak{h}_i^r$, $\alpha_j(L)$ and $\hat{\lambda}_i(L) = G$,

- between all $(T(r, a_i))_{a_i}$ that should be taken with a positive probability, player i is indifferent after any history in the Lth review round; and

- if player i reports $(T(r, a_i))_{a_i}$ that should not be taken on the equilibrium path, then player j punishes player i. We make sure that this punishment is sufficiently large to discourage any deviation after any history.

In addition, player j adjusts player i's learning at the beginning of the Lth review round. Let $V_i(\mathfrak{h}_i^r)$ be the expected increase of player i's continuation payoff at the beginning of round

r if player i could know $\lambda_j(l)$ and $\theta_j(l)$, where the lth review round is the first round to come after round r (if round r is the $\tilde{l}$th review round for some $\tilde{l}$, then $l = \tilde{l}$). By Proposition 16, $V_i(\mathfrak{h}_i^r)$ is very small. Player j adds $V_i(\mathfrak{h}_i^r)$ to the reward so that player i does not have an incentive to improve the learning until the beginning of round r.

Then, we proceed backwards. In the review rounds, the key differences from the last review round are

- player i's history in round r affects player i's expected continuation payoff from the next round through learning. This effect is canceled out by adding $V_i(\mathfrak{h}_i^{r+1})$; and
- the distribution of $\theta_j(\tilde{l})$ for the future rounds and whether player j obeys the message can affect player i's continuation payoff conditional on $\hat{\lambda}_i(\tilde{l}) = G$ for the future rounds, but the distribution of $\theta_j(\tilde{l})$ and whether player j obeys the message is independent of player i's action plan. Hence, this effect does not affect the optimality.

Hence, the same argument holds.

For round r where player i sends a message m, we replace $(T(r, a_i))_{a_i}$ with the set of possible messages m's in the above discussion.

For round r where player i receives a message m, player i does not take an action.

Therefore, by backward induction, we verify that $\sigma_i(x_i)$ is optimal, taking all the continuation strategies into account after a deviation given the truthtelling incentive. See Section 2.9.6 for how to incentivize the players to tell the truth.

2.9 Appendix of Chapter 2

2.9.1 Proof of Lemma 12

By linear independence of $(q_i(a_i, \alpha_j))_{a_i \in A_i}$ (Assumption 3), for all α, there exists $\pi_i[\alpha] : Y_j \to \mathbb{R}$ such that

$$u_i(a_i, \alpha_j) + \mathbb{E}\left[\pi_i[\alpha](y_j) \mid a_i, \alpha_j\right] = 0.$$

Without loss, we assume that $\pi_i[\alpha](y_j)$ is upper hemi-continuous with respect to α. Since $\Delta(A) \ni \alpha$ is compact, there exists $\bar{u}$ such that $\pi_i[\alpha] : Y_j \to [-\bar{u}, \bar{u}]$ for all $\alpha \in \Delta(A)$. We can add or subtract a constant so that (2.5) is satisfied. Re-take $\bar{u}$ if necessary.

2.9.2 Proof of Lemma 13

This follows from Assumption 3 as Lemma 12. Since $\left(1-\delta^{t-1}\right) u_i(a_t)$ converges to 0 as δ goes to unity for all $t \in \{1, ..., T_P\}$ with $T_P = \Theta((1-\delta)^{-\frac{1}{2}})$, we have

$$\lim_{\delta \to 1} \sup_{t \in \{1,\ldots,T_P\}, a_{j,t}, y_{j,t}} \left|\pi_i^{\delta}(t, a_{j,t}, y_{j,t})\right| = 0,$$

which implies (2.11).

2.9.3 Proof of Lemma 14

Let m be the minimum element of $Q_i(x)$ and M be the maximum element of $Q_i(x)$. Let $\tilde{Q}_i(x)$ be the matrix whose (l, n) element is $\frac{(Q_i(x))_{l,n}+|m|+1}{|M|+2|m|+2} \in (0, 1)$ and $\mathbf{q}_i(x)$ be the vector whose lth element is $\frac{(\mathbf{q}_i(x))_l+|m|+1}{|M|+2|m|+2} \in (0, 1)$.

Since aff $\left(\{q_i(a_i, a_j)\}_{a_j \in A_j}\right) \subset$ aff $(\{\mathbf{1}_{y_i}\}_{y_i \in Y_i})$, without loss of generality, we can assume that the first row of $Q_i(x)$ is parallel to $(1, ..., 1)$ and that the first element of $\mathbf{q}_i(x)$ is 1.

We will show

$$\left\{\mathbf{y}_i \in \mathbb{R}_+^{|Y_i|} : Q_i(x)\mathbf{y}_i = \mathbf{q}_i(x)\right\} = \boldsymbol{Q}_i(x) = \tilde{\boldsymbol{Q}}_i(x) \equiv \left\{\mathbf{y}_i \in \mathbb{R}_+^{|Y_i|} : \tilde{Q}_i(x)\mathbf{y}_i = \tilde{\mathbf{q}}_i(x)\right\}.$$

1. $\boldsymbol{Q}_i(x) \subset \tilde{\boldsymbol{Q}}_i(x)$

Suppose that $\mathbf{y}_i \in \boldsymbol{Q}_i(x)$. Then,

$$\begin{aligned}\left(\tilde{Q}_i(x)\mathbf{y}_i\right)_l &= \frac{(Q_i(x)\mathbf{y}_i)_l}{|M|+2|m|+2} + \frac{|m|+1}{|M|+2|m|+2} \underbrace{(1,...,1)}_{\text{first row of } Q_i(x)} \mathbf{y}_i \\ &= \frac{(\mathbf{q}_i(x))_l}{|M|+2|m|+2} + \frac{|m|+1}{|M|+2|m|+2} = (\tilde{\mathbf{q}}_i(x))_l\,,\end{aligned}$$

as desired.

2. $\boldsymbol{Q}_i(x) \supset \tilde{\boldsymbol{Q}}_i(x)$

Suppose that $\mathbf{y}_i \in \tilde{\boldsymbol{Q}}_i(x)$. Then, for all l, $\left(\tilde{Q}_i(x)\mathbf{y}_i\right)_l = (\tilde{\mathbf{q}}_i(x))_l$, that is,

$$\begin{aligned}&\frac{(Q_i(x)\mathbf{y}_i)_l}{|M|+2|m|+2} + \frac{|m|+1}{|M|+2|m|+2}(1,...,1)\mathbf{y}_i \\ =\;& \frac{(\mathbf{q}_i(x))_l}{|M|+2|m|+2} + \frac{|m|+1}{|M|+2|m|+2},\end{aligned} \tag{2.31}$$

If $\mathbf{y}_i \notin \text{aff}\left(\{\mathbf{1}_{y_i}\}_{y_i \in Y_i}\right)$, then

$$\begin{aligned}&\frac{(Q_i(x)\mathbf{y}_i)_1}{|M|+2|m|+2} + \frac{|m|+1}{|M|+2|m|+2}(1,...,1)\mathbf{y}_i \\ =\;& \frac{|m|+2}{|M|+2|m|+2}\sum_l y_{i,l} \\ \neq\;& \frac{|m|+2}{|M|+2|m|+2} = \frac{(\mathbf{q}_i(x))_1}{|M|+2|m|+2} + \frac{|m|+1}{|M|+2|m|+2}\end{aligned}$$

and so contradiction to (2.31) with $l = 1$.

If $\mathbf{y}_i \in \text{aff}\left(\{\mathbf{1}_{y_i}\}_{y_i \in Y_i}\right)$, then (2.31) implies

$$\frac{(Q_i(x)\mathbf{y}_i)_l}{|M|+2|m|+2} = \frac{(\mathbf{q}_i(x))_l}{|M|+2|m|+2},$$

or $(Q_i(x)\mathbf{y}_i)_l = (\mathbf{q}_i(x))_l$ and so $\mathbf{y}_i \in \boldsymbol{Q}_i(x)$.

2.9.4 Proof of Lemma 15

We consider the proof for a more general strategy of player j. Given some fixed set $A_j(x) \in 2^{\Delta(A_j)}$ of player j's mixed actions, player j determines $\alpha_j(l)$ as follows:

- $\alpha_j(l) = \alpha_j(x)$ with probability $1 - \eta$; and
- for each $\alpha_j \in A_j(x)$, $\alpha_j(l) = \alpha_j$ with probability $\frac{1}{|A_j(x)|}\eta$.

If the latter is the case, $\theta_j(l) = B$. That is, the latter corresponds to the case where $\alpha_j(l) \neq \alpha_j(x)$. The other conditions for $\lambda_j(l)$, $\hat{\lambda}_i(l)$ and $\theta_j(l)$ are the same.

Now, let us prove Lemma 15. Once $\lambda_j(\tilde{l}) = B$ is induced, then $\lambda_j(\tilde{l}') = B$ for all the following rounds. Hence, there exists a unique l^* such that $\lambda_j(\tilde{l}) = B$ is initially induced in the $(l^* + 1)$th review round: $\lambda_j(1) = \cdots = \lambda_j(l^*) = G$ and $\lambda_j(l^* + 1) = \cdots = \lambda_j(L) = B$. Similarly, there exists $\hat{l}^*$ with $\hat{\lambda}_j(1) = \cdots = \hat{\lambda}_j(\hat{l}^*) = G$ and $\hat{\lambda}_j(\hat{l}^* + 1) = \cdots = \hat{\lambda}_j(L) = B$. If $\lambda_j(L) = G$ ($\hat{\lambda}_j(L) = G$, respectively), then define $l^* = L$ ($\hat{l}^* = L$, respectively).

Then, there are following three cases:

2.9.4.1 $l^* = \hat{l}^*$

This means $\lambda_j(l) = \hat{\lambda}_j(l)$ for all l as desired.

2.9.4.2 $l^* > \hat{l}^*$

This means that player i obeys the message in the supplemental round for $\lambda_j(\hat{l}^* + 1)$:

$$\hat{\lambda}_j(\hat{l}^* + 1) = f[i](\lambda_j(\hat{l}^* + 1)).$$

By Lemma 10, player i believes that, conditional on $\lambda_j(\hat{l}^* + 1)$,

$$\hat{\lambda}_j(\hat{l}^* + 1) = f[i](\lambda_j(\hat{l}^* + 1)) = \lambda_j(\hat{l}^* + 1) = G$$

or $g[j](\lambda_j(\hat{l}^* + 1)) = E$ with probability no less than $1 - \exp(-\Theta(T^{\frac{1}{2}}))$. If the former is the case, then $l^* \leq \hat{l}^*$ (contradiction). If the latter is the case, then $\theta_j(\hat{l}^* + 1) = B$, as desired. Since player j's continuation action plan in the main blocks does not depend on $g[j](\lambda_j(\hat{l}^* + 1))$, we are done.

2.9.4.3 $l^* < \hat{l}^*$

There are following two cases. First, if player i obeys the message in the supplemental round for $\lambda_j(l^* + 1)$, then by the same reason as above, we are done.

Second, player i disregards $f[i](\hat{\lambda}_j(l^* + 1))$. As seen in Section 2.4.4.2, player i's history in the l^*th review round satisfies

1. (2.20): $\left\| \frac{1}{T-1} \sum_{t \in T_i(l^*)} \mathbf{1}_{a_{i,t}} - \alpha_i(x) \right\| < \varepsilon$; and
2. (2.21): $\left\| \frac{1}{|T_i(l^*,x)|} \sum_{t \in T_i(l^*,x)} \mathbf{1}_{y_{i,t}} - \boldsymbol{Q}_i(x) \right\| < \varepsilon$.

There are two cases: first, if player i's signal frequency in $T_i(l^*, x)$ is sufficiently close to the ex ante distribution:

$$\left\| \frac{1}{|T_i(l^*, x)|} \sum_{t \in T_i(l^*,x)} \mathbf{1}_{y_{i,t}} - q_i(x) \right\| < \eta. \tag{2.32}$$

Then, for sufficiently small ρ, ε and η, player i believes that

$$X_j(l^*) \begin{cases} \leq \frac{\bar{u}}{L} T & \text{if } x_j = G, \\ \geq -\frac{\bar{u}}{L} T & \text{if } x_j = B, \end{cases}$$

that is, $\lambda_j(l^* + 1) = G$, with probability $1 - \exp(-\Theta(T))$:

Lemma 19 Fix $\bar{u}$ and L. For any $i \in I$, $l \in \{1, ..., L\}$ and x, suppose player j takes an action as described above. There exist $\bar{\eta}, \bar{\rho}, \bar{\varepsilon}$ such that, for all $\eta < \bar{\eta}$, $\rho < \bar{\rho}$ and $\varepsilon < \bar{\varepsilon}$, if player i's history satisfies (2.20), (2.21) and (2.32), then given $\alpha_j(l^*) = \alpha_j(x)$, player i

believes that

$$X_j\left(l^*\right)\begin{cases}\leq \frac{\bar{u}}{L}T & \text{if } x_j=G,\\ \geq -\frac{\bar{u}}{L}T & \text{if } x_j=B\end{cases} \tag{2.33}$$

with probability $1-\exp(-\Theta(T))$.

Proof: For sufficiently small ρ and ε, $T_i(l,x)$ (the set of periods where player i takes $a_i\left(x\right)$) is sufficiently long and (2.33) is satisfied if the score in $T_i(l,x)$ is not erroneous:

$$\sum_{t\in T_i(l,x)}\pi_i[\alpha\left(x\right)](y_{j,t})\begin{cases}\leq \frac{2}{3}\frac{\bar{u}}{L}T & \text{if } x_j=G,\\ \geq -\frac{2}{3}\frac{\bar{u}}{L}T & \text{if } x_j=B.\end{cases}$$

For sufficiently small η, ρ and ε, the distribution of $\{y_{i,t}\}_{t\in T_i(l,x)}$ while player i takes $a_i\left(x\right)$ is close to the ex ante mean and so the conditional expectation of the score in $T_i(l,x)$ is close to the ex ante mean:

$$\mathbb{E}\left[\sum_{t\in T_i(l,x)}\pi_i[\alpha\left(x\right)](y_{j,t})\mid\{y_{i,t}\}_{t\in T_i(l,x)},a_i\left(x\right),\alpha_j(x)\right]\begin{cases}\leq \frac{1}{2}\frac{\bar{u}}{L}T & \text{if } x_j=G,\\ \geq -\frac{1}{2}\frac{\bar{u}}{L}T & \text{if } x_j=B.\end{cases}$$

Note that, for sufficiently small ρ, the ex ante mean of $\pi_i[\alpha\left(x\right)](y_{j,t})$ given $a_i\left(x\right),\alpha_j\left(x\right)$ is sufficiently close to that given $\alpha\left(x\right)$, which is zero.

In such a case, by Hoeffdling's inequality, player i believes that

$$\sum_{t\in T_i(l,x)}\pi_i[\alpha\left(x\right)](y_{j,t})\begin{cases}\leq \frac{2}{3}\frac{\bar{u}}{L}T & \text{if } x_j=G,\\ \geq -\frac{2}{3}\frac{\bar{u}}{L}T & \text{if } x_j=B\end{cases}$$

with probability $1-\exp(-\Theta(T))$, as desired. ∎

Second, if player i's signal frequency in $T_i(l^*,x)$ is not close to the ex ante distribution:

$$\left\|\frac{1}{|T_i(l^*,x)|}\sum_{t\in T_i(l^*,x)}\mathbf{1}_{y_{i,t}}-q_i(x)\right\|\geq\eta.$$

Then, player i believes that player j does not take $\alpha_j(l^*) = \alpha_j(x)$:

Lemma 20 For any $i \in I$, $l \in \{1, ..., L\}$ and x, suppose player j takes an action as described above. There exists $\bar{\eta} > 0$ such that, for all $\eta < \bar{\eta}$, there exists $\bar{\rho} > 0$ such that, for any $\rho < \bar{\rho}$, there exist $\bar{\varepsilon}$ and $\{A_j(x)\}_x$ such that for all $\varepsilon < \bar{\varepsilon}$, if player i's history satisfies

1. $\left\| \frac{1}{T-1} \sum_{t \in T_i(l)} \mathbf{1}_{a_{i,t}} - \alpha_i(x) \right\| < \varepsilon$;
2. $\left\| \frac{1}{|T_i(l,a_i)|} \sum_{t \in T_i(l,x)} \mathbf{1}_{y_{i,t}} - \boldsymbol{Q}_i(x) \right\| < \varepsilon$; and
3. $\left\| \frac{1}{|T_i(l,a_i)|} \sum_{t \in T_i(l,x)} \mathbf{1}_{y_{i,t}} - q_i(x) \right\| \geq \eta$,

then player i puts a belief at least $1 - \exp(-\Theta(T))$ on the events that $\alpha_j(l) \neq \alpha_j(x)$.

Proof: For notational simplicity, let f_i be player i's signal frequency in the periods where player i takes $a_i(x)$ in $T_i(l)$:

$$f_i = \frac{1}{|T_i(l,x)|} \sum_{t \in T_i(l,x)} \mathbf{1}_{y_{i,t}}.$$

Fix $\eta > 0$. Note that from 3,

$$\left\| \frac{1}{|T_i(l,x)|} \sum_{t \in T_i(l,x)} \mathbf{1}_{y_{i,t}} - q_i(x) \right\| \geq \eta.$$

Suppose player i's history in the lth review round is $\{a_{i,t}, y_{i,t}\}_{t \in T(l)}$. The likelihood between $\alpha_j \in A_j(x)$ against $\alpha_j(x)$ is given by

$$\begin{aligned}
& \frac{\Pr\left(\alpha_j \mid \{a_{i,t}, y_{i,t}\}_{t \in T(l)}\right)}{\Pr\left(\alpha_j(x) \mid \{a_{i,t}, y_{i,t}\}_{t \in T(l)}\right)} \\
= & \frac{\Pr\left(\{y_{i,t}\}_{t \in T_i(l,x)} \mid a_i(x), \alpha_j\right)}{\Pr\left(\{y_{i,t}\}_{t \in T_i(l,x)} \mid a_i(x), \alpha_j(x)\right)} \\
& \times \prod_{a_i \neq a_i(x)} \frac{\Pr\left(\{y_{i,t}\}_{t \in T_i(l,a_i)} \mid a_i, \alpha_j\right)}{\Pr\left(\{y_{i,t}\}_{t \in T_i(l,a_i)} \mid a_i, \alpha_j(x)\right)} \frac{\Pr\left(y_{i,t_i(l)} \mid a_{i,t_i(l)}, \alpha_j\right)}{\Pr\left(y_{i,t_i(l)} \mid a_{i,t_i(l)}, \alpha_j(x)\right)} \frac{\Pr(\alpha_j)}{\Pr(\alpha_j(x))},
\end{aligned}$$

where $T_i(l, a_i)$ is the set of periods in the lth review round with $a_{i,t} = a_i$. We want to show that if 1, 2 and 3 are satisfied, then this likelihood is $\exp(\Theta(T))$ for some $\alpha_j \in A_j(x)$ with $\alpha_j \neq \alpha_j(x)$.

For this purpose, we want to show that

$$\log \Pr\left(\{y_{i,t}\}_{t \in T_i(l,x)} \mid a_i(x), \alpha_j\right) = |T_i(l,x)| \sum_{y_i} f_i(y_i) \log q_i(y_i \mid a_i(x), \alpha_j)$$

is sufficiently bigger than $\log \Pr\left(\{y_{i,t}\}_{t \in T_i(l,x)} \mid a_i(x), \alpha_j(x)\right)$.

Instead of working on α_j, we work on the distribution directly. Define

$$t_i \equiv q_i(a_i(x), \alpha_j) - q_i(x).$$

By Taylor series, we have

$$\begin{aligned} & \sum_{y_i} f_i(y_i) \log q_i(y_i \mid a_i(x), \alpha_j) \\ = & \sum_{y_i} f_i(y_i) \log q_i(y_i \mid x) + \sum_{n=1}^{\infty} (-1)^{n-1} \sum_{y_i} f_i(y_i) \frac{1}{n\,(q_i(y_i \mid x))^n} (t_i(y_i))^n \end{aligned} \tag{2.34}$$

with $q_i(y_i \mid x) = q_i(y_i \mid a_i(x), \alpha_j(x))$. We want to show that the second term in the second line is sufficiently large.

First, we concentrate on the case where $f_i \in \boldsymbol{Q}_i(x)$ and consider the subspace of player i's signals, $\boldsymbol{Q}_i(x)$. There exists $K_1 > 0$ such that, for each $\rho > 0$, for any $e_\rho \leq K_1 \rho$, we have

$$B_{e_\rho}(q_i(x)) \subset \boldsymbol{Q}_i(x)$$

since $\alpha_j(x)$ mixes all $a_j \in A_j$ with probability at least ρ. Here, $B_{e_\rho}(q_i(x))$ is the closed ball with radius $e_\rho > 0$ and center $q_i(x)$ in the linear space $\mathrm{aff}\left(\{\mathbf{1}_{y_i}\}_{y_i \in Y_i}\right) \cap \boldsymbol{Q}_i(x)$.

Let $C_{e_\rho}(q_i(x))$ be the surface of $B_{e_\rho}(q_i(x))$. Intuitively, we will define $A_j(x)$ such that $\{q_i(a_i(x),\alpha_j)\}_{\alpha_j \in A_j(x)}$ is a discretization of $C_{e_\rho}(q_i(x))$.

Since $\|f_i - q_i(x)\| \geq \eta$ from 3, for sufficiently small ρ, for any $e_\rho \leq K_1\rho$, we have

$$f_i \notin B_{e_\rho}(q_i(x)).$$

In addition, since $q(y \mid a)$ has full support, for sufficiently small ρ, for any $e_\rho \leq K_1\rho$, we have

$$\bar{t} \equiv \max_{i,x,t_i,y_i \in Y_i, \alpha_j \in \Delta(A_j)} \frac{t_i(y_i)}{q_i(y_i \mid a_i(x), \alpha_j)} < \frac{1}{3}, \tag{2.35}$$

where the maximization with respect to t_i is subject to $\exists q_i \in B_{e_\rho}(q_i(x))$ with $t_i = q_i - q_i(x)$.

Given any $e_\rho \in (0, K_1\rho]$, given $f_i \in \boldsymbol{Q}_i(x)$ with $\|f_i - q_i(x)\| \geq \eta$, define t_i such that $q_i(x) + t_i$ is on $C_{e_\rho}(q_i(x))$. Since $\|t_i\| = e_\rho$, $f_i \notin B_{e_\rho}(q_i(x))$ and $\|f_i - q_i(x)\| \leq 1$, there exists $b \geq e_\rho$ such that

$$t_i \equiv b(f_i - q_i(x)).$$

Then,

$$\sum_{y_i} f_i(y_i) \frac{1}{q_i(y_i \mid x)} t_i(y_i) = b\left(\sum_{y_i} (f_i(y_i))^2 \frac{1}{q_i(y_i \mid x)} - 1\right).$$

Since $\sum_{y_i} (f_i(y_i))^2 \frac{1}{q_i(y_i|x)}$ is uniquely minimized at $f_i = q_i(x)$ and the minimized value is 1, we have

$$\min_{f_i:\|f_i - q_i(x)\| \geq \eta} \sum_{y_i} (f_i(y_i))^2 \frac{1}{q_i(y_i \mid x)} - 1 \equiv e_{\eta, a_i(x), \alpha_j(x)} > 0.$$

Since $\Delta(A_j)$ is compact, there exists e_η independent of ρ such that

$$\min_{\substack{i \in I \\ a_i(x) \in A_i, \alpha_j(x) \in \Delta(A_j)}} \min_{f_i:\|f_i - q_i(x)\| \geq \eta} \sum_{y_i} (f_i(y_i))^2 \frac{1}{q_i(y_i \mid x)} - 1 = 8e_\eta > 0.$$

Therefore, for any f_i with $\|f_i - q_i(x)\| \geq \eta$, there exists t_i such that

- $q_i(x) + t_i \in C_{e_\rho}(q_i(x))$; and

- we have

$$\sum_{y_i} f_i(y_i) \frac{1}{q_i\left(y_i \mid x\right)} t_i(y_i) \geq 8e_\eta e_\rho.$$

By (2.34), we have

$$\begin{aligned}
& \sum_{y_i} f_i(y_i) \log\left(q_i(y_i \mid x) + t_i(y_i)\right) \\
= & \sum_{y_i} f_i(y_i) \log q_i\left(y_i \mid x\right) + \sum_{n=1}^{\infty} (-1)^{n-1} \sum_{y_i} f_i(y_i) \frac{1}{n\left(q_i\left(y_i \mid x\right)\right)^n} \left(t_i(y_i)\right)^n \\
\geq & \sum_{y_i} f_i(y_i) \log q_i\left(y_i \mid x\right) + \left(1 - \sum_{n=1}^{\infty} (\bar{t})^n\right) 8e_\eta e_\rho \\
\geq & \sum_{y_i} f_i(y_i) \log q_i\left(y_i \mid x\right) + 4e_\eta e_\rho.
\end{aligned}$$

The last inequality uses the fact that $\bar{t} < \frac{1}{3}$ (See (2.35) for the definition of $\bar{t}$). Therefore,

$$\min_{\substack{i,x \\ f_i \in \boldsymbol{Q}_i(x)}} \max_{q_i \in C_{e_\rho}(q_i(x))} \left\{ \sum_{y_i} f_i(y_i) \log q_i(y_i) - \sum_{y_i} f_i(y_i) \log q_i(y_i \mid x) \right\} \geq 4e_\eta e_\rho.$$

By Assumption 2, there exists $\varepsilon_x > 0$ such that, for sufficiently small ρ,

$$C_{e_\rho}\left(q_i(x)\right) \subset \Delta_{\varepsilon_x} \equiv \left\{ q_i \in \Delta\left(\left\{\mathbf{1}_{y_i}\right\}_{y_i \in Y_i}\right) : q_i(y_i) \geq \varepsilon_x \text{ for all } y_i \in Y_i \right\}.$$

Since $\sum_{y_i} f_i\left(y_i\right) \log q_i(y_i)$ is Lipschitz continuous in $f_i \in \boldsymbol{Q}_i(x)$ and $q_i \in \Delta_{\varepsilon_x}$, there exists M such that there are M points in $C_{e_\rho}\left(q_i(x)\right)$, denoted by $C^M_{e_\rho}\left(q_i(x)\right)$, such that

$$\min_{\substack{i,x \\ f_i \in \boldsymbol{Q}_i(x)}} \max_{q_i \in C^M_{e_\rho}(q_i(x))} \left\{ \sum_{y_i} f_i(y_i) \log q_i(y_i) - \sum_{y_i} f_i(y_i) \log q_i(y_i \mid x) \right\} \geq 3e_\eta e_\rho.$$

Now we consider the case with $f_i \notin \boldsymbol{Q}_i(x)$. From 1 and 2, there exists $(s_i(y_i))_{y_i \in Y_i}$ such that

$$\begin{aligned} f_i + s_i &\in \boldsymbol{Q}_i(x), \\ \|s_i\| &\leq \varepsilon. \end{aligned}$$

Again, since $\sum_{y_i} f_i(y_i) \log q_i(y_i)$ is Lipschitz continuous in $f_i \in \boldsymbol{Q}_i(x)$ and $q_i \in \Delta_{\varepsilon_x}$, there exists $\varepsilon > 0$ such that, for all s_i with $\|s_i\| \leq \varepsilon$, we have

$$\begin{aligned} & \min_{\substack{i,x \\ f_i \in \boldsymbol{Q}_i(x)}} \max_{q_i \in C^M_{e_\rho}(q_i(x))} \left\{ \sum_{y_i} (f_i(y_i) + s_i(y)) \log q_i(y_i) - \sum_{y_i} (f_i(y_i) + s_i(y)) \log q_i(y_i \mid x) \right\} \\ \geq \;& e_\eta e_\rho. \end{aligned}$$

Therefore, in total, for sufficiently small ρ, for any $e_\rho \in (0, K_1\rho]$, there exists a finite set $C^M_{e_\rho}(q_i(x))$ such that, if Conditions 1, 2 and 3 are satisfied, then

$$\max_{q_i \in C^M_{e_\rho}(q_i(x))} \left\{ \sum_{y_i} f_i(y_i) \log q_i(y_i) - \sum_{y_i} f_i(y_i) \log q_i(y_i \mid x) \right\} \geq e_\eta e_\rho. \tag{2.36}$$

Fix $\eta > 0$ and then $\rho > 0$ and $e_\rho = K_1\rho$ so that the above statement is true. By definition of $\boldsymbol{Q}_i(x)$, there exists a set of M player j's strategies, denoted by $A_j(x)$, such that, for each point $q_i \in C^M_{e_\rho}(q_i(x))$, there exists a mixture $\alpha_j \in A_j(x)$ such that $q_i = q_i(y_i \mid a_i(x), \alpha_j)$. Then, (2.36) means that, if Conditions 1, 2 and 3 are satisfied, then there exists $\alpha_j \in A_j(x)$ such that

$$\begin{aligned} & \log \Pr\left(\{y_{i,t}\}_{t \in T_i(l,x)} \mid a_i(x), \alpha_j \right) - \log \Pr\left(\{y_{i,t}\}_{t \in T_i(l,x)} \mid a_i(x), \alpha_j(x) \right) \\ \geq \;& |T_i(l,x)|\, e_\eta e_\rho \geq (1 - 2|A_i|\rho - 2\varepsilon)\, e_\eta K_1 \rho T. \end{aligned}$$

The last inequality uses definition of $\alpha_i(x)$ and Condition 1.

On the other hand, consider the periods when player i takes $a_i \neq a_i(x)$. Note that, for all $\alpha_j \in A_j(x)$,

$$\begin{aligned} & |\log q_i(y_i \mid a_i, \alpha_j) - \log q_i(y_i \mid x)| \\ = \; & |\log (q_i(y_i \mid x) + t_i(y_i)) - \log q_i(y_i \mid x)| \leq \rho. \end{aligned}$$

Hence,

$$\begin{aligned} \prod_{a_i \neq a_i(x)} \frac{\Pr\left(\{y_{i,t}\}_{t \notin T_i(l,a_i)} \mid a_i, \alpha_j\right)}{\Pr\left(\{y_{i,t}\}_{t \in T_i(l,a_i)} \mid a_i, \alpha_j(x)\right)} & \geq \exp(-\rho \sum_{a_i \neq a_i(x)} |T_i(l,a_i)|) \\ & \geq \exp(-\rho 3 |A_i| \rho T) \end{aligned}$$

if we take $\bar{\varepsilon} \leq \rho$. The last inequality follows from $|T_i(l,a_i)| \leq 3|A_i| \rho T$. To see why $|T_i(l,a_i)| \leq 3|A_i| \rho T$, suppose not. Then, $|T_i(l,a_i)| > 3|A_i| \rho T$. Then, $\sum_{a_i \neq a_i(x)} |T_i(l,a_i)| > 3|A_i| \rho T$, which means $|T_i(l,x)| < (1 - 3|A_i| \rho) T \leq (1 - 2|A_i| \rho - \varepsilon) T$ for $\bar{\varepsilon} \leq \rho$, which contradicts to Condition 1.

Therefore, in total,

$$\begin{aligned} & \frac{\Pr\left(\alpha_j \mid \{a_{i,t}, y_{i,t}\}_{t \in T(l)}\right)}{\Pr\left(\alpha_j(x) \mid \{a_{i,t}, y_{i,t}\}_{t \in T(l)}\right)} \\ = \; & \frac{\Pr\left(\{y_{i,t}\}_{t \in T_i(l,x)} \mid a_i(x), \alpha_j\right)}{\Pr\left(\{y_{i,t}\}_{t \in T_i(l,x)} \mid a_i(x), \alpha_j(x)\right)} \\ & \times \prod_{a_i \neq a_i(x)} \frac{\Pr\left(\{y_{i,t}\}_{t \notin T_i(l,a_i)} \mid a_i, \alpha_j\right)}{\Pr\left(\{y_{i,t}\}_{t \in T_i(l,a_i)} \mid a_i, \alpha_j(x)\right)} \frac{\Pr\left(y_{i,t_i(l)} \mid a_{i,t_i(l)}, \alpha_j\right)}{\Pr\left(y_{i,t_i(l)} \mid a_{i,t_i(l)}, \alpha_j(x)\right)} \frac{\Pr(\alpha_j)}{\Pr(\alpha_j(x))} \\ \geq \; & \exp\left(\left(\left(1 - 2|A_i| \rho - 2\varepsilon\right) e_\eta K_1 - \rho 3 |A_i|\right) \rho T\right) \frac{1}{M}. \\ = \; & \exp(\Theta(T)) \end{aligned}$$

for sufficiently small ρ and ε, as desired. Here, we can neglect $\frac{\Pr\left(y_{i,t_i(l)}|a_{i,t_i(l)},\alpha_j\right)}{\Pr\left(y_{i,t_i(l)}|a_{i,t_i(l)},\alpha_j(x)\right)}$ by Assumption 2. ∎

In summary, we take η, ρ and ε so that Lemmas 19 and 20 are satisfied. Then, if player i disregards the message,

1. if
$$\left\|\frac{1}{|T_i(l^*,x)|}\sum_{t\in T_i(l^*,x)}\mathbf{1}_{y_{i,t}}-q_i(x)\right\|<\eta,$$
then player i believes that $\alpha_j(l^*)\neq\alpha_j(x)$ (which means $\theta_j(l^*)=B$) or
$$X_j(l^*)\begin{cases}\leq\frac{\bar{u}}{L}T & \text{if } x_j=G,\\ \geq-\frac{\bar{u}}{L}T & \text{if } x_j=B\end{cases}$$
with probability $1-\exp(-\Theta(T))$; and

2. if
$$\left\|\frac{1}{|T_i(l^*,x)|}\sum_{t\in T_i(l^*,x)}\mathbf{1}_{y_{i,t}}-q_i(x)\right\|\geq\eta,$$
then player i believes that $\alpha_j(l^*)\neq\alpha_j(x)$ (which means $\theta_j(l^*)=B$) with probability $1-\exp(-\Theta(T))$.

These bounds are before learning from player j's continuation action plan. The learning from player j's continuation action plan changes the belief in the following two ways. First, $\left\{f[i](\lambda_j(\tilde{l}))\right\}_{\tilde{l}=1}^{L}$ reveals $\left\{\lambda_j(\tilde{l})\right\}_{\tilde{l}=1}^{L}$. However, since the error occurs with positive probability $\left(\exp(-\Theta(T^{\frac{1}{2}}))\right)^L=\exp(-\Theta(T^{\frac{1}{2}}))$, the update of the belief is sufficiently small compared to the original belief $1-\exp(-\Theta(T))$.

Second, player i conditions that $\hat{\lambda}_j(\tilde{l})=G$ for all $\tilde{l}\leq l$. This conditioning changes player i's belief on player j's history in the following way: player i's belief on player j's history where player j obeys player i's message $f[j](\lambda_i(\tilde{l}))$ can be decreased at most by $\exp(-\Theta(T^{\frac{1}{2}}))$ for the supplemental round for $\lambda_i(\tilde{l})$ since any signal profile can occur with positive probability

$\exp(-\Theta(T^{\frac{1}{2}}))$ in the supplemental round for $\lambda_i(\tilde{l})$. On the other hand, player i's belief on player j's history where player j disregards player i's message $f[j](\lambda_i(\tilde{l}))$ remains unchanged. In total, the update of the belief is $\left(\exp(-\Theta(T^{\frac{1}{2}}))\right)^L = \exp(-\Theta(T^{\frac{1}{2}}))$, which is sufficiently small compared to the original belief $1 - \exp(-\Theta(T))$, as desired.

2.9.5 Proof of Proposition 16

For 3, it suffices to have

$$\bar{\pi}_i(x, \lambda_j(l), l) \begin{cases} \leq 0 & \text{if } x_j = G, \\ \geq 0 & \text{if } x_j = B, \end{cases} \tag{2.37}$$

$$|\bar{\pi}_i(x, \lambda_j(l), l)| \leq \max_{i,a} 2 \, |u_i(a)| \, T \tag{2.38}$$

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

To see why (2.37) and (2.38) are sufficient, notice the following: (2.38) with $T = (1-\delta)^{-\frac{1}{2}}$ implies

$$\lim_{\delta \to 1} \frac{1-\delta}{\delta^{T_P}} \sup_{x, h_j^{\text{main}}} \left| \sum_{l=1}^{L} \pi_i^{\text{main}}(x, h_j^{\text{main}}, l) \right| = 0.$$

See the discussion in Section 2.4.4.1 to see why (2.37) is enough for (1.17).

Now, we are left to prove 1 and 2. 1-(b) is true by the reasons that we have explained in Section 2.7.

We will verify 1-(a) by backward induction. Section 2.4.4.3 guarantees that the distribution of $\theta_j(l)$ is independent of player i's action plan and so we can neglect the effect of player i's action plan on $\theta_j(l)$. Further, for a moment, forget about the first term in π_i^{main}, $-2\bar{u}T$ ($2\bar{u}T$, respectively) for $x_j = G$ ($x_j = B$, respectively).

In the Lth review round, for almost optimality, we can assume that $\lambda_j(L) = \hat{\lambda}_j(L)$ and that player j uses (2.16) by the following reason: by Lemma 15, conditional on $\hat{\lambda}_i(L) = G$, player i has a posterior no less than $1 - \exp(-\Theta(T^{\frac{1}{2}}))$ on the event that $\lambda_j(L) = \hat{\lambda}_j(L)$ or any action is optimal. Since the per-period difference of the payoff from two different strategies is bounded by $\bar{U} \equiv 2\bar{u} + \max_{i,a} 2|u_i(a)|$, the expected loss from assuming $\lambda_j(L) = \hat{\lambda}_j(L)$ is no more than $\exp(-\Theta(T^{\frac{1}{2}}))\bar{U}T$. Therefore, for almost optimality, we can assume that $\lambda_j(L) = \hat{\lambda}_j(L)$. Further, if (2.13) is used, then any action is optimal. Therefore, we can assume that player j uses (2.16).

In addition, if player j does not play $\alpha_j(x)$, then it means that $\theta_j(L) = B$ and that any action is optimal for player i (that is, (2.13) is used). Hence, we can concentrate on the case where player j plays $\alpha_j(x)$.

If $\lambda_j(L) = \hat{\lambda}_j(L) = G$, then any action plan is almost optimal for sufficiently large T since Lemma 12 implies that the marginal expected increase in $X_j(L)$ cancels out the marginal decrease in instantaneous utilities. If $\lambda_j(L) = \hat{\lambda}_j(L) = B$, then D_i is strictly optimal since the reward (2.16) is constant. Therefore, $\sigma_i(x_i)$ is optimal. Remember that in the period $t_j(L)$ excluded from $X_j(L)$, $\pi_i^{x_j}[\alpha](y_{j,t_j(L)})$ makes any action optimal.

Further, if player j uses (2.16) and $\lambda_j(L) = \hat{\lambda}_j(L)$, then player i's average continuation payoff at the beginning of the Lth review round except for $\bar{\pi}_i(x, \lambda_j(L), L)$ is

$$\begin{array}{ll} w_i(x) & \text{if } \lambda_j(L) = \hat{\lambda}_j(L) = G, \\ u_i(D_i, (1-2\rho)\, C_j + 2\rho D_j) \geq w_i(x) & \text{if } x_j = G, \lambda_j(L) = \hat{\lambda}_j(L) = B, \\ u_i(D_i, (1-2\rho)\, D_j + 2\rho C_j) \leq \max\{\max\limits_{\substack{x: \\ x_j = B}} w_i(x), v_i^*(\rho)\} & \text{if } x_j = B, \lambda_j(L) = \hat{\lambda}_j(L) = B. \end{array} \tag{2.39}$$

Hence, there exists $\bar{\pi}_i(x, \lambda_j(L), L)$ with (2.37) and (2.38) such that player i's average continuation payoff is equal to $\min_{x:x_j=G} w_i(x)$ if $x_j = G$ and $\max\{\max_{x:x_j=B} w_i(x), v_i^*(\rho)\}$ if $x_j = B$.

Therefore, we define $\bar{\pi}_i(x, \lambda_j(L), L)$ such that player i's value from the Lth review round is independent of $\lambda_j(L)$ as long as $\lambda_j(L) = \hat{\lambda}_j(L)$.[9] In addition, Lemma 15 implies that player i in the main blocks does not put a belief more than $\exp(-\Theta(T^{\frac{1}{2}}))$ on the events

[9]In the above discussion, we have verified that this claim is correct for the case with $\theta_j(L) = G$.
Otherwise, player i is indifferent between any action profile sequence, as desired.

that $\lambda_j(L) \neq \hat{\lambda}_j(L)$ and player i's value depends on action profiles in the Lth review round. Further, again, we can neglect the effect of player i's action plan on θ_j. Therefore, for almost optimality, we can assume that player i in the $(L-1)$th review round maximizes

$$\mathbb{E}\left[\sum_{t \in T(L-1)} u_i(a_t) + \pi_i^{\text{main}}(x, h_j^{\text{main}}, L-1) \mid x\right], \tag{2.40}$$

assuming $\lambda_j(L-1) = \hat{\lambda}_j(L-1)$.

Therefore, the same argument as for the Lth review round establishes that $\sigma_i(x_i)$ is almost optimal in the $(L-1)$th review round.

Further, if player j uses (2.16) and $\lambda_j(L-1) = \hat{\lambda}_j(L-1)$, then player i's average payoff from the $(L-1)$th review round except for $\bar{\pi}_i(x, \lambda_j(L-1), L-1)$ is given by (2.39). The cases where (2.13) will be used in the Lth review round will happen with probability no more than η (player j takes $\alpha_j(L-1) = \alpha_j(x)$ with probability $1-\eta$) plus some negligible probabilities for not having (2.20), (2.22) or (2.23). When (2.13) is used, per period payoff is bounded by $[-2\bar{u}, 2\bar{u}]$ by (2.9).

Therefore, there exists $\bar{\pi}_i(x, \lambda_j(L-1), L-1)$ with (2.37) and (2.38) such that if player j uses (2.16) and $\lambda_j(L-1) = \hat{\lambda}_j(L-1)$, then player i's average continuation payoff from the $(L-1)$th and Lth review rounds is

$$\begin{cases} \min_{x:x_j=G} w_i(x) - \eta\left(2\bar{u} + \min_{x:x_j=G} w_i(x)\right) & \text{if } x_j = G, \\ \max\{\max_{x:x_j=B} w_i(x), v_i^*(\rho)\} + \eta\left(2\bar{u} - \max\{\max_{x:x_j=B} w_i(x), v_i^*(\rho)\}\right) & \text{if } x_j = B. \end{cases}$$

Recursively, for $l = 1$, 1-(a) of the proposition is satisfied and the average ex ante payoff of player i is

$$\begin{cases} \min_{x:x_j=G} w_i(x) - L\eta\left(2\bar{u} + \min_{x:x_j=G} w_i(x)\right) & \text{if } x_j = G, \\ \max\{\max_{x:x_j=B} w_i(x), v_i^*(\rho)\} + L\eta\left(2\bar{u} - \max\{\max_{x:x_j=B} w_i(x), v_i^*(\rho)\}\right) & \text{if } x_j = B. \end{cases}$$

Note that, in the first review round, player j uses (2.16) and $\hat{\lambda}_j(1) = \lambda_j(1) = G$ with probability one.

Taking the first term $-2\bar{u}T$ ($2\bar{u}T$, respectively) for $x_j = G$ ($x_j = B$, respectively) into account, the average ex ante payoff is

$$\begin{cases} \min_{x:x_j=G} w_i(x) - \frac{2\bar{u}}{L} - L\eta \left(2\bar{u} + \min_{x:x_j=G} w_i(x)\right) & \text{if } x_j = G, \\ \max\{\max_{x:x_j=B} w_i(x), v_i^*(\rho)\} \\ +\frac{2\bar{u}}{L} + L\eta\left(2\bar{u} - \max\{\max_{x:x_j=B} w_i(x), v_i^*(\rho)\}\right) & \text{if } x_j = B. \end{cases}$$

From (2.27), we can further modify $\bar{\pi}_i(x, G, 1)$ with (2.37) and (2.38) such that $\sigma_i(x_i)$ gives $\bar{v}_i$ ($\underline{v}_i$, respectively) if $x_j = G$ (B, respectively). Therefore, 2 of the proposition is satisfied.

2.9.6 Formal Construction of the Report Block

We formally construct the action plan in the report block and π_i^{report} so that player i tells the truth about h_i^{main} and that $\sigma_i(x_i)$ is exactly optimal. Here, we do not consider the feasibility constraint (1.17). As we will see, π_i^{report} is bounded by $[-T^{-1}, T^{-1}]$ and we can restore (1.17) by adding or subtracting a small constant depending on x_j without affecting efficiency or incentive.

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

- what state x_j player j is in;
- what action plan $\alpha_j(l)$ player j took in the lth review round if round r is the lth review round; and
- $\hat{\lambda}_i(l) = G$ in the lth review round if round r is the lth review round or after.

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

The following notations are useful: as defined in Section 2.8, $\#_i^r$ is the summary of player i's history within round r, $\mathfrak{h}_i^r$ is the summary of player i's history at the beginning of round r, and $(T(r,a_i))_{a_i}$ is the set of player i's strategies that take a_i for $T(r,a_i)$ times in round r ex ante (if round r corresponds to a review round). On the other hand, let $\hat{\#}_i^r$ be player i's report of $\#_i^r$ and $\left(\hat{T}(r,a_i)\right)_{a_i}$ be such that, according to player i's report $\hat{\#}_i^r$, player i takes each a_i for $\hat{T}(r,a_i)$ times in round r. In addition, let t_r be the first period of round r.

For round r corresponding to a review round, we divide a review round into $T^{\frac{3}{4}}$ review subrounds. Each kth subround is from period $t_r + (k-1)\,T^{\frac{1}{4}} + 1$ to period $t_r + kT^{\frac{1}{4}}$ with $k \in \{1,\ldots,T^{\frac{3}{4}}\}$. Let $T(r,k)$ be the set of periods in the kth subround of round r. Let $\#_i^r(k)(a_i,y_i) \in \{1,...,T^{\frac{1}{4}}\}$ be how many times player i observed an action-signal pair (a_i,y_i) in the kth subround of round r and $\#_i^r(k)$ be $(\#_i^r(k)(a_i,y_i))_{a_i,y_i}$.

When player i is picked by the public randomization device with probability $\frac{1}{2}$, player i sends the messages via perfect cheap talk: sequentially from round 1 to the last round, player i reports the history as follows:

- if round r corresponds to a review round, then
  - first, player i reports the summary $\#_i^r$;
  - second, for each subround k, player i reports the summary $\#_i^r(k)$;
  - third, public randomization is drawn such that each subround k is randomly picked with probability $T^{-\frac{3}{4}}$. Let $k(r)$ be the subround picked by the public randomization; and
  - fourth, for $k(r)$, player i reports the whole history $\{a_{i,t}, y_{i,t}\}_{t\in T(r,k(r))}$ in the $k(r)$th subround;
- if player i sends a message by the error-reporting noisy cheap talk in round r, then player i reports $\#_i^r$, which is her true message m and signals $g[i](m)$ and $g_2[i](m)$; and

- if player i receives a message by the error-reporting noisy cheap talk in round r, then player i reports $\#_i^r$, which is her signals $f[i](m)$ and $f_2[i](m)$.

Remember that we want to use binary perfect cheap talk as mentioned in Section 1.5. For round r corresponding to a review round, for each $\#_i^r \in \{1, ..., T\}^{|A_i||Y_i|}$, we can attach a sequence of binary messages $\{G, B\}$ so that the sequence uniquely identifies $\#_i^r$. The length of the sequence is $|A_i| |Y_i| \log_2 T$. Similarly, for each $\#_i^r(k)$, we can attach a sequence of binary messages $\{G, B\}$ with length $\frac{1}{4} |A_i| |Y_i| \log_2 T$. For each (a_i, y_i), we can attach a sequence of binary messages $\{G, B\}$ with length $\log_2 |A_i| |Y_i|$. Then, the number of binary messages to send the fourth message $\{a_{i,t}, y_{i,t}\}_{t \in T(r,k(r))}$ is $\log_2 |A_i| |Y_i| T^{\frac{1}{4}}$. Hence, in total, the length of the messages is $\Theta(T^{\frac{1}{4}})$. For the other rounds, the length of the necessary messages is at most 3. Therefore, in total, the number of messages we need is

$$\Theta(T^{\frac{1}{4}}). \tag{2.41}$$

As a preparation to show the incentive to tell the truth, we prove the following lemma:

Lemma 21 If Assumption 5 is satisfied, then given $\rho > 0$, there exists $\bar{\varepsilon} > 0$ such that

1. there exists $g_i(h_j^{\text{main}}, a_i, y_i)$ such that, for each $l \in \{1, ..., L\}$ and $t \in T(l)$, conditional on $\hat{\lambda}_i(l) = G$, it is better for player i to report $(a_{i,t}, y_{i,t})$ truthfully: for all h_i^{main} and h_j^{main} with $\hat{\lambda}_i(l) = G$, conditional on $\hat{\lambda}_i(l) = G$,

$$\begin{aligned} & \mathbb{E}\left[g_i(h_j^{\text{main}}, \hat{a}_{i,t}, \hat{y}_{i,t}) \mid h_i^{\text{main}}, (\hat{a}_{i,t}, \hat{y}_{i,t}) = (a_{i,t}, y_{i,t})\right] \\ > \ & \mathbb{E}\left[g_i(h_j^{\text{main}}, \hat{a}_{i,t}, \hat{y}_{i,t}) \mid h_i^{\text{main}}, (\hat{a}_{i,t}, \hat{y}_{i,t}) \neq (a_{i,t}, y_{i,t})\right] + \bar{\varepsilon} T^{-(N-1)}, \end{aligned} \tag{2.42}$$

where $(\hat{a}_{i,t}, \hat{y}_{i,t})$ is player i's report about $(a_{i,t}, y_{i,t})$ in the report block;

2. for the round where player i sends the message by the error-reporting noisy cheap talk, it is better to report $(m, g[i](m), g_2[i](m))$ truthfully:

$$\begin{aligned} & \mathbb{E}\left[\begin{array}{c} g_i(h_j^{\text{main}}, \hat{m}, \widehat{g[i](m)}, \widehat{g_2[i](m)}) \\ \mid h_i^{\text{main}}, \left(\hat{m}, \widehat{g[i](m)}, \widehat{g_2[i](m)}\right) = (m, g[i](m), g_2[i](m)) \end{array}\right] \\ > \ & \mathbb{E}\left[\begin{array}{c} g_i(h_j^{\text{main}}, \hat{m}, \widehat{g[i](m)}, \widehat{g_2[i](m)}) \\ \mid h_i^{\text{main}}, \left(\hat{m}, \widehat{g[i](m)}, \widehat{g_2[i](m)}\right) \neq (m, g[i](m), g_2[i](m)) \end{array}\right] + \bar{\varepsilon} T^{-1}, \end{aligned} \tag{2.43}$$

where $\left(\hat{m}, \widehat{g[i](m)}, \widehat{g_2[i](m)}\right)$ is player i's report in the report block about the message and signals of the supplemental round; and

3. for the round where player i receives the message from the error-reporting noisy cheap talk, it is better to report $(f[i](m), f_2[i](m))$ truthfully:

$$\begin{aligned} & \mathbb{E}\left[\begin{array}{c} g_i(h_j^{\text{main}}, \widehat{f[i](m)}, \widehat{f_2[i](m)}) \\ \mid h_i^{\text{main}}, \left(\widehat{f[i](m)}, \widehat{f_2[i](m)}\right) = (f[i](m), f_2[i](m)) \end{array}\right] \\ > \ & \mathbb{E}\left[\begin{array}{c} g_i(h_j^{\text{main}}, \widehat{f[i](m)}, \widehat{f_2[i](m)}) \\ \mid h_i^{\text{main}}, \left(\widehat{f[i](m)}, \widehat{f_2[i](m)}\right) \neq (f[i](m), f_2[i](m)) \end{array}\right] + \bar{\varepsilon} T^{-1}, \end{aligned} \tag{2.44}$$

where $\left(\widehat{f[i](m)}, \widehat{f_2[i](m)}\right)$ is player i's report in the report block about the signals of the supplemental round.

Proof: 1. Note that with $\hat{\lambda}_i(l) = G$, player j takes $\alpha_{j,t}$ that takes each action with probability no less than ρ. We show that

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

works.[10] To see this, consider the following two cases:

[10] Kandori and Matsushima (1998) use a similar reward to give a player the incentive to tell the truth about the history.

(a) if $t_j(l) \neq t$, then any report is optimal since $g_i(h_j^{\text{main}}, \hat{a}_{i,t}, \hat{y}_{i,t}) = 0$; and

(b) if $t_j(l) = t$, then period t is not used for the construction of player j's continuation action plan. Hence, player i cannot learn $a_{j,t}, y_{j,t}$ from h_i^{main}. Hence, player i, after knowing $t_j(l) = t$ and $\alpha_{j,t}$, wants to minimize

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

Lemma 21 implies that $(\hat{a}_{i,t}, \hat{y}_{i,t}) = (a_{i,t}, y_{i,t})$ is a unique minimizer.

Since $A_j(x) \ni \alpha_{j,t}$ is finite, we are left to show that there exists $\bar{\varepsilon} > 0$ such that, for any h_i^{main}, l and $t \in T(l)$, player i puts a belief at least $\bar{\varepsilon}T^{-1}$ on $t_j(l) = t$.

From now on, we condition on player j's history at the beginning of the lth review round with $\hat{\lambda}_i(l) = G$.

Let $\#_j^*(l)(a_j, y_j, \mathbf{1}_{Q_j(x)})$ denote how many times player j observed each pair $(a_j, y_j, \mathbf{1}_{Q_j(x)})$ in the lth review round and $\#_j^*(l)$ be a vector $(\#_j^*(l)(a_j, y_j, \mathbf{1}_{Q_j(x)}))_{(a_j,y_j,\mathbf{1}_{Q_j(x)})}$. Note that the distribution of player j's continuation action plan from the next round is fully determined by $\#_j^*(l)$ and $(a_{j,t_j(l)}, y_{j,t_j(l)}, \mathbf{1}_{Q_j(x),t_j(l)})$. Hence, it suffices to show that, if player i could know $\{a_{i,\tau}, y_{i,\tau}\}_{\tau\in T(l)}$, $\#_j^*(l)$ and $(a_{j,t_j(l)}, y_{j,t_j(l)}, \mathbf{1}_{Q_j(x),t_j(l)})$, player i puts a belief at least $\bar{\varepsilon}T^{-1}$ on $t_j(l) = t$ for each t.

For any t and $t' \in T(l)$, the likelihood ratio between $t_j(l) = t$ and $t_j(l) = t'$ is given by

$$\begin{aligned}
& \frac{\Pr\left(t_j(l) = t \mid \{a_{i,\tau}, y_{i,\tau}\}_{\tau\in T(l)}, \alpha_j(l), \#_j^*(l), (a_{j,t_j(l)}, y_{j,t_j(l)}, \mathbf{1}_{Q_j(x),t_j(l)})\right)}{\Pr\left(t_j(l) = t' \mid \{a_{i,\tau}, y_{i,\tau}\}_{\tau\in T(l)}, \alpha_j(l), \#_j^*(l), (a_{j,t_j(l)}, y_{j,t_j(l)}, \mathbf{1}_{Q_j(x),t_j(l)})\right)} \\
= & \frac{\Pr\left(\#_j^*(l), (a_{j,t_j(l)}, y_{j,t_j(l)}, \mathbf{1}_{Q_j(x),t_j(l)}) \mid \{a_{i,\tau}, y_{i,\tau}\}_{\tau\in T(l)}, \alpha_j(l), t_j(l) = t\right)}{\Pr\left(\#_j^*(l), (a_{j,t_j(l)}, y_{j,t_j(l)}, \mathbf{1}_{Q_j(x),t_j(l)}) \mid \{a_{i,\tau}, y_{i,\tau}\}_{\tau\in T(l)}, \alpha_j(l), t_j(l) = t'\right)} \\
\in & \left[\min q(a_j, y_j, \mathbf{1}_{Q_j(x)} \mid a_i, y_i, \alpha_j(l)), \frac{1}{\min q(a_j, y_j, \mathbf{1}_{Q_j(x)} \mid a_i, y_i, \alpha_j(l))}\right],
\end{aligned}$$

where the minimum is taken with respect to $a_j, y_j, \mathbf{1}_{Q_j(x)}$. By Assumption 2 and Lemma 14, there exists $\bar{\varepsilon} > 0$ such that

$$\begin{aligned} & \left[\min q(a_j, y_j, \mathbf{1}_{Q_j(x)} | a_i, y_i, \alpha_j(l)), \frac{1}{\min q(a_j, y_j, \mathbf{1}_{Q_j(x)} | a_i, y_i, \alpha_j(l))}\right] \\ \in & \left[\bar{\varepsilon}, \frac{1}{\bar{\varepsilon}}\right]. \end{aligned}$$

Since $A_j(x) \ni \alpha_j(l)$ is finite, we can take $\bar{\varepsilon}$ independently from $a_i, y_i, \alpha_j(l)$. Hence, we have

$$\begin{aligned} & \Pr\left(t_j(l) = t \mid \{a_{i,\tau}, y_{i,\tau}\}_{\tau \in T(l)}, \alpha_j(l), \#_j^*(l), (a_{j,t_j(l)}, y_{j,t_j(l)}, \mathbf{1}_{Q_j(x),t_j(l)})\right) \\ > & \ \bar{\varepsilon} \Pr\left(t_j(l) = t' \mid \{a_{i,\tau}, y_{i,\tau}\}_{\tau \in T(l)}, \alpha_j(l), \#_j^*(l), (a_{j,t_j(l)}, y_{j,t_j(l)}, \mathbf{1}_{Q_j(x),t_j(l)})\right). \end{aligned}$$

Since there exists at least one t with

$$\Pr\left(t_j(l) = t \mid \{a_{i,\tau}, y_{i,\tau}\}_{\tau \in T(l)}, \#_j^*(l), (a_{j,t_j(l)}, y_{j,t_j(l)}, \mathbf{1}_{Q_j(x),t_j(l)})\right) \geq T^{-(N-1)},$$

we are done.

2. From (1.5),

$$\begin{aligned} & g_i(h_j^{\text{main}}, \hat{m}, \widehat{g[i](m)}, \widehat{g_2[i](m)}) \\ = & -\left\| \mathbf{1}_{f_2[j](m)} - \mathbb{E}[\mathbf{1}_{f_2[j](m)} \mid f[j](m), \hat{m}, \widehat{g[i](m)}, \widehat{g_2[i](m)}] \right\|^2 \end{aligned}$$

works. Since player j's continuation action plan is independent of $f_2[j](m)$, conditioning on h_i^{main} does not change the optimality.

3. From (1.3),

$$\begin{aligned} & g_i(h_j^{\text{main}}, \widehat{f[i](m)}, \widehat{f_2[i](m)}) \\ = & -\left\| \mathbf{1}_{g_2[j](m)} - \mathbb{E}[\mathbf{1}_{g_2[j](m)} \mid m, g[j](m), \widehat{f[i](m)}, \widehat{f_2[i](m)}] \right\|^2 \end{aligned}$$

works. Since player j's continuation action plan is independent of $g_2[j](m)$, conditioning on h_i^{main} does not change the optimality. ∎

Since ρ is fixed in Section 2.6, $\bar{\varepsilon}$ is now fixed. Given this preparation, by backward induction, we construct $\pi_i^{\text{report}}\left(h_j^{r+1}, \hat{h}_i^{r+1}, r\right)$ for each r such that

$$\pi_i^{\text{report}}(x_j, h_j^{T_P+1} : \delta) = \sum_r \pi_i^{\text{report}}\left(h_j^{r+1}, \hat{h}_i^{r+1}, r\right)$$

makes it optimal to tell the truth in the report block and that $\sigma_i(x_i)$ is exactly optimal.

Formally, $\pi_i^{\text{report}}\left(h_j^{r+1}, \hat{h}_i^{r+1}, r\right)$ is determined as follows. If player i is not selected by the public randomization device or there exists l such that the lth review round is before or equal to round r and $\hat{\lambda}_i(l) = B$, then

$$\pi_i^{\text{report}}\left(h_j^{r+1}, \hat{h}_i^{r+1}, r\right) = 0. \tag{2.45}$$

Otherwise, $\pi_i^{\text{report}}\left(h_j^{r+1}, \hat{h}_i^{r+1}, r\right)$ is the summation of the following rewards and punishments.

Punishment for a Lie One component of $\pi_i^{\text{report}}\left(h_j^{r+1}, \hat{h}_i^{r+1}, r\right)$ is the punishment for telling a lie. For round r corresponding to a review round, the punishment is the summation of the following three:

- the number indicating player i's lie about $\{a_{i,t}, y_{i,t}\}_{t \in T(r,k(r))}$:

$$\sum_{t \in T(r,k(r))} T^{-3} g_i(h_j^{\text{main}}, \hat{a}_{i,t}, \hat{y}_{i,t}); \tag{2.46}$$

- the number indicating player i's lie about $\#_i^r(k)$:

$$T^{-3} \times T^{\frac{3}{4}} \times \mathbf{1}\left\{ \hat{\#}_i^r(k(r)) \neq \sum_{t \in T(r,k(r))} \mathbf{1}_{\hat{a}_{i,t},\hat{y}_{i,t}} \right\}, \tag{2.47}$$

  where $\mathbf{1}_{\hat{a}_{i,t},\hat{y}_{i,t}}$ is an $|A_i|\,|Y_i| \times 1$ vector such that the element corresponding to $(\hat{a}_{i,t}, \hat{y}_{i,t})$ is equal to 1 and the others are 0. Here, the term $T^{\frac{3}{4}}$ cancels out $T^{-\frac{3}{4}}$, the probability that each k is selected to be $k(r)$ by the public randomization; and

- the number indicating player i's lie about $\#_i^r$:

$$T^{-3} \times \mathbf{1}\left\{ \hat{\#}_i^r \neq \sum_k \hat{\#}_i^r(k) \right\}. \tag{2.48}$$

If player i sends a message by the error-reporting noisy cheap talk in round r, then player i reports $(m, g[i](m), g_2[i](m))$. Player j punishes player i if it is likely for player i to tell a lie by

$$T^{-3} g_i(h_j^{\text{main}}, \hat{m}, \widehat{g[i](m)}, \widehat{g_2[i](m)}). \tag{2.49}$$

If player i receives a message from the error-reporting noisy cheap talk in round r, then player i reports $(f[i](m), f_2[i](m))$. Player j punishes player i if it is likely for player i to tell a lie by

$$T^{-3} g_i(h_j^{\text{main}}, \widehat{f[i](m)}, \widehat{f_2[i](m)}). \tag{2.50}$$

Cancel Out the Expected Punishment by Telling the Truth Note that even if player i tells the truth, the expected punishment is positive for (2.46), (2.49) and (2.50) and the expectation of the punishment is different for different actions and messages.[11] We cancel out the differences in ex ante values of the punishment between different actions and messages:

[11] On the equilibrium, (2.47) and (2.48) are 0 after any history.

- if round r is a review round, then player j gives player i

$$\sum_{t \in T(r)} \mathbf{1}\{t_j(l) = t\} \mathbf{1}\{t \in T(r, k(r))\} T^{-3} \Pi_i[\alpha_j(l)](y_{j,t}); \tag{2.51}$$

- if player i sends the message in round r, then player j gives player i

$$T^{-3} \Pi_i(f[j](m)); \tag{2.52}$$

and

- if player i receives the message in round r, then player j gives player i

$$T^{-3} \Pi_i(m). \tag{2.53}$$

Here, $\Pi_i(f[j](m))$ ($\Pi_i(m)$, respectively) is defined so that the differences in (2.49) ((2.50), respectively) among messages are canceled out ex ante before sending (receiving, respectively) the message, as we define $\Pi_i[\alpha_j](y_{j,t})$ in Lemma 18. Assumption 3 guarantees the existence of such a reward.

Then, in each period of the main block, given truthtelling in the report block, before taking an action or sending a message, the ex ante punishments from (2.46), (2.47), (2.48), (2.49), (2.50), (2.51), (2.52) and (2.53) are zero.

Reward for the Optimal Action Another component of $\pi_i^{\text{report}}(h_j^{r+1}, \hat{h}_i^{r+1}, r)$ is the reward for taking an equilibrium action in round r (or, punishment for not taking an equilibrium action). From $\{\#_i^{\tilde{r}}\}_{\tilde{r} \leq r-1}$, we can calculate $\mathfrak{h}_i^r = \{\#_i^{\tilde{r}}\}_{\tilde{r} \leq r-1}$. Let $\hat{\mathfrak{h}}_i^r$ be player j's inference of $\mathfrak{h}_i^r$ based on player i's reports $\left\{\hat{\#}_i^{\tilde{r}}\right\}_{\tilde{r} \leq r-1}$.

If round r corresponds to a review round, then based on the reports $\hat{\mathfrak{h}}_i^r$ and $\hat{\#}_i^r$, player j gives the reward

$$f_i(\hat{\mathfrak{h}}_i^r, \hat{\#}_i^r, \alpha_j(r)), \tag{2.54}$$

which is to be determined. Here, $\alpha_j(r)$ is player j's action plan in round r.

If round r corresponds to a round where player i sends the message m, then based on the reports $\hat{\mathfrak{h}}_i^r$ and $\hat{m}$, player j gives the reward

$$f_i(\hat{\mathfrak{h}}_i^r, \hat{m}), \tag{2.55}$$

which is to be determined.

We will take f_i such that

$$f_i(\hat{\mathfrak{h}}_i^r, \hat{\#}_i^r, \alpha_j(r)), f_i(\hat{\mathfrak{h}}_i^r, \hat{m}) \in [-T^{-r-(N-1)-4}, T^{-r-(N-1)-4}] \tag{2.56}$$

for all $\hat{\mathfrak{h}}_i^r$, $\hat{\#}_i^r$, $\alpha_j(r)$ and $\hat{m}$.

Incentive to Tell the Truth Before specifying f_i, we establish player i's incentive to tell the truth. For the reports about the last round, all the reports about the previous rounds are sunk. Hence, what the reports affect is the punishment and f_i for the last round. Both are zero for $\hat{\lambda}_i(L) = B$. Hence, player i can condition $\hat{\lambda}_i(L) = G$. From (2.56), the effect on f_i is bounded by $\Theta(T^{-(N-1)-4})$ while the marginal effect on punishment from telling a lie is at least $\Theta(T^{-(N-1)-3})$ from (2.46), (2.47), (2.48), (2.49), (2.50) and Lemma 21. Hence, truthtelling is strictly optimal.

Given the incentive to tell the truth about the last round, the same argument holds for the second last round, and so on. By induction, we establish player i's incentive to tell the truth for all the rounds.

Ex Ante Expected Punishment Given the truthtelling incentive and (2.51), (2.52) and (2.53), in each period of each main block, before taking an action or sending a message, the ex ante punishments from (2.46), (2.47), (2.48), (2.49), (2.50), (2.51), (2.52) and (2.53) are zero.

Determination of f_i We determine f_i by backward induction.

In the last round (the Lth review round), since player j determines her continuation strategy treating each period within a past round identically, player i's belief about player j's continuation strategy at the beginning of the Lth review round conditional on $\alpha_j(L)$ and $\hat{\lambda}_i(L) = G$ is determined by $\mathfrak{h}_i^r$. Conditional on $\alpha_j(L)$, the learning from signal observations is redundant. In addition, given the truthtelling incentive, the expected punishments from (2.46), (2.47), (2.48), (2.49), (2.50), (2.51), (2.52) and (2.53) are zero and can be ignored.

Hence, conditional on $\alpha_j(L)$ and $\hat{\lambda}_i(L) = G$, player i's value of taking $(T(r, a_i))_{a_i}$ only depends on $\mathfrak{h}_i^r$ regardless of $\{a_{i,t}, y_{i,t}\}$ in the Lth review round. Let $v(\hat{\mathfrak{h}}_i^r, \alpha_j(r), (T(r, a_i))_{a_i})$ denote this value. On the other hand, from player i's report, player j can know $\mathfrak{h}_i^r$ and how many times player i took a_i in round r. Therefore, player j can calculate f_i such that, if $\hat{\mathfrak{h}}_i^r$ is an on-path history,

- if $(\hat{T}(r, a_i))_{a_i}$ is an equilibrium action plan given $\hat{\mathfrak{h}}_i^r$, then

$$f_i(\hat{\mathfrak{h}}_i^r, \hat{\#}_i^r, \alpha_j(r)) = \frac{v(\hat{\mathfrak{h}}_i^r, \alpha_j(r), (T^*(r, a_i))_{a_i}) - v(\hat{\mathfrak{h}}_i^r, \alpha_j(r), (\hat{T}(r, a_i))_{a_i})}{\Pr(\text{player } i \text{ is picked by the public randomization})} \tag{2.57}$$

  so that player i is indifferent between any equilibrium action plan. Here, $(T^*(r, a_i))_{a_i}$ is such $(T(r, a_i))_{a_i}$ that minimizes $v(\hat{\mathfrak{h}}_i^r, \alpha_j(r), (\hat{T}(r, a_i))_{a_i})$ among those taken with a positive probability in equilibrium after $\hat{\mathfrak{h}}_i^r$; and

- if $(\hat{T}(r, a_i))_{a_i}$ is not an equilibrium action plan given $\hat{\mathfrak{h}}_i^r$, then

$$f_i(\hat{\mathfrak{h}}_i^r, \hat{\#}_i^r, \alpha_j(r)) = -T^{-1-(N-1)-4} \tag{2.58}$$

On the other hand, if $\hat{\mathfrak{h}}_i^r$ is an off-path history, then f_i is defined to be 0.

We can take $f_i(\hat{\mathfrak{h}}_i^r, \hat{\#}_i^r, \alpha_j(r))$ satisfying (2.56) since (i) the original action plan is almost optimal by Proposition 16 conditional on $\alpha_j(L)$ and $\hat{\lambda}_i(L) = G$, (ii) we have established the incentive to tell the truth, and (iii) from (ii) and Π_i, the ex ante punishments from (2.46), (2.47), (2.48), (2.49), (2.50), (2.51), (2.52) and (2.53) are zero.

Consider player i's incentive on the equilibrium path. With $\hat{\lambda}_i(L) = B$, then $\pi_i^{\text{report}}\left(h_j^{r+1}, \hat{h}_i^{r+1}, r\right) = 0$ and $\theta_j(L) = B$. Hence, any action sequence is optimal. With $\hat{\lambda}_i(L) = G$, (2.57) makes player i indifferent between all the $(T(r,a_i))_{a_i}$'s that are taken with a positive probability in equilibrium after $\hat{\mathfrak{h}}_i^r$, regardless of $\{a_{i,t}, y_{i,t}\}$ in the Lth review round. In addition, (2.58) is sufficiently large to discourage any deviation after any history on the equilibrium path since the equilibrium action plan is almost optimal except for the report block.

Therefore, we are done with the last round. We proceed backward.

For the lth review round, given $\hat{\mathfrak{h}}_i^r$ and $\hat{\#}_i^r$, we can calculate the following effects of player i's history in the lth review round on the continuation payoff from the next round: given $\hat{\mathfrak{h}}_i^r$ and $\hat{\#}_i^r$, player i's optimal action plan from the next round is determined. Given this,

- the effect on f_i for round $\tilde{r} \geq r+1$ (this is well defined since we proceed backward);
- the effect on (2.46), (2.47), (2.48), (2.49), (2.50), (2.51), (2.52) and (2.53) for round $\tilde{r} \geq r+1$. Since all are zero, we neglect this effect;
- the effect on the distribution of $\lambda_j(l)$. By Proposition 16, as long as player i can learn $\lambda_j(l)$, this effect can be neglected. See below for the effect on learning;
- the effect on the distribution of $\theta_j(l)$. Since the distribution is independent of player i's action plan, we neglect this effect; and
- the effect on player i's learning about $\lambda_j(l)$ and $\theta_j(l)$.

Hence, we can concentrate on the effect on f_i and player i's learning about $\lambda_j(l)$ and $\theta_j(l)$. Since f_i for round $\tilde{r} \geq r+1$ is bounded by (2.56) (replace r with $r+1$), the former is bounded by $[-\Theta(T^{-r-(N-1)-5}), \Theta(T^{-r-(N-1)-5})]$. For the latter, let $V_i(\mathfrak{h}_i^r, \#_i^r)$ be the expected increase of player i's continuation payoff at the beginning of round $r+1$ if player i could know $\lambda_j(l)$

and $\theta_j(l)$ from round $r+1$. Note that $\lambda_j(l)$ and $\theta_j(l)$ fully determine player i's optimal action from the next round. Proposition 16 implies $V_i(\mathfrak{h}_i^r, \#_i^r) = \exp(-\Theta(T^{\frac{1}{2}}))$.

Player j constructs $f_i(\hat{\mathfrak{h}}_i^r, \hat{\#}_i^r, \alpha_j(r))$ in the following two steps: first, player j makes player i indifferent between all the $\#_i^r$ in terms of the above two effects so that player i does not have an incentive to affect f_i for round $\tilde{r} \geq r+1$ or improve the learning at the beginning of round $r+1$ (ex post). This is bounded by $[-\Theta(T^{-r-(N-1)-5}), \Theta(T^{-r-(N-1)-5})]$.

Second, given the first step, player j makes player i indifferent between all the $(T(r, a_i))_{a_i}$'s that are taken with a positive probability in equilibrium after $\hat{\mathfrak{h}}_i^r$ and discourages player i from any deviation after any history on the equilibrium, as in the Lth review round.

The total adjustment can be bounded by $[-T^{-r-(N-1)-4}, T^{-r-(N-1)-4}]$, as desired.

For round r corresponding to a round where player i sends a message m, we replace $(T(r, a_i))_{a_i}$ with the set of possible messages m's in the above discussion.

For round r corresponding to a round where player i receives a message m, player i does not take an action.

We can proceed until the first round and show the optimality of $\sigma_i(x_i)$ recursively.

Finally, without the reward in the report block, for all $x \in \{G, B\}^2$, $\sigma_i(x_i)$ gives a payoff $\underline{v}_i$ for $x_j = B$ and $\bar{v}_i$ for $x_j = G$. In this section, we have established the exact optimality of $\sigma_i(x_i)$ conditional on x_j. Since the summation of the reward in the report block is bounded by T^{-1}, for all $x \in \{G, B\}^2$, $\sigma_i(x_i)$ is optimal against $\sigma_j(x_j)$ and gives a payoff close to $\underline{v}_i$ for $x_j = B$ and $\bar{v}_i$ for $x_j = G$. By adjusting the reward based on x_j, we can make sure that $\sigma_i(x_i)$ is optimal against $\sigma_j(x_j)$ and the payoff is $\underline{v}_i$ for $x_j = B$ and $\bar{v}_i$ for $x_j = G$. Since $\sigma_i(x_i)$ is optimal conditional on x_j, it is optimal for both players to send x_i truthfully in the coordination block (although player 2, the second sender, knows x_1 when she sends x_2).

Chapter 3

Overview of the Extensions

In the following chapters, we prove the folk theorem for a general game without cheap talk or public randomization (Theorem 8) in steps. Remember that the arguments in Chapter 1 are valid for all the steps.

We offer an overview of the structure and summarize exactly what generic conditions we need to prove Theorems 7 and 8 in each step.

3.1 Structure

First, in Chapter 4, we show the folk theorem for the general two-player game with the perfect cheap talk, error-reporting noisy cheap talk and public randomization.

Second, in Chapter 5, we show the folk theorem for the general more-than-two-player game with the perfect cheap talk, error-reporting noisy cheap talk and public randomization.

Note that, by the end of the second step, we have shown Theorem 7. Hence, we are left to show Theorem 8, dispensing with the perfect cheap talk, error-reporting noisy cheap talk and public randomization.

Third, in Chapter 6, we prove the dispensability in the two-player game. We proceed in steps. In the coordination block, we replace the perfect cheap talk with the error-reporting noisy cheap talk. Then, we dispense with the error-reporting noisy cheap talk in the coordination and main blocks. On the other hand, in the report block, we dispense with public randomization, after which we replace the perfect cheap talk with conditionally independent noisy cheap talk. Then, we dispense with the conditionally independent cheap talk.

Fourth, in Chapter 7, we dispense with the perfect cheap talk, error-reporting noisy cheap talk and public randomization in the more-than-two-player game. The main difference from the two-player case is how to construct the coordination block without the perfect cheap talk but with the error-reporting noisy cheap talk.

3.2 Assumptions

Given the above structure, we mention what generic assumptions are sufficient to prove Theorems 7 and 8 in each step. Again, all the assumptions are generic under Assumption 6.

3.2.1 General Two-Player Game with Cheap Talk and Public Randomization

In the the general two-player game with the perfect cheap talk, error-reporting noisy cheap talk and public randomization, no additional assumption is necessary, that is, Assumptions 1, 2, 3 and 5 are sufficient.

3.2.2 General More-Than-Two-Player Game with Cheap Talk and Public Randomization

With more than two players, we also assume Assumption 4 so that player j can identify which of two players i and n is more suspicious in deviations since for each player j, there are more than one opponents.

Assumptions 1, 2, 3 and 5 are maintained as they are.

3.2.3 General Two-Player Game withOUT Cheap Talk

Now we consider how to dispense with the perfect cheap talk, error-reporting noisy cheap talk and public randomization device in the general two-player game.

3.2.3.1 Coordination Block

With cheap talk, each player communicates x_i via perfect cheap talk. We need to consider the protocol to coordinate on x_i by sending messages via actions.

Error-Reporting Noisy Cheap Talk First, we replace the perfect cheap talk in the coordination block with the error-reporting noisy cheap talk. To do so, we do not need any new assumptions.

Messages via Actions Second, we replace the error-reporting noisy cheap talk with messages via actions. Since we replace the perfect cheap talk in the coordination block with the error-reporting noisy cheap talk, this step enables us to dispense with the perfect cheap talk in the coordination block and the error-reporting noisy cheap talk in the main blocks. In this step, we make an assumption to make sure that we can create a message protocol to preserve the important features of the inferences which were guaranteed with the error-reporting noisy cheap talk (see Lemma 10). A sufficient condition will be Assumption 39 in Chapter 6.

3.2.3.2 Report Block

With public randomization, the players coordinate on who will report the history by the public randomization device. In addition, the picked player reports the history via perfect cheap talk.

Dispensing with Public Randomization We first dispense with the public randomization device. So that the players can coordination through their actions and private signals, Assumption 40 in Chapter 6 is sufficient.

Conditionally Independent Noisy Cheap Talk We second replace the perfect cheap talk with conditionally independent noisy cheap talk. For this step, no new assumption is necessary (except for the availability of the conditionally independent noisy cheap talk).

Messages via Actions We third replace the conditionally independent noisy cheap talk with messages via actions. As seen in Section 1.5.6, we need to create a statistics of a receiver to infer the messages from a sender so that the sender cannot get any information about the realization of the statistics through her private signals. See Assumption 41 in Chapter 6. Note that we do not assume that $2|Y_i| \leq |A_j||Y_j|$ for all i, j with $i \neq j$. Hence, we cannot use the method that Fong, Gossner, Hörner, and Sannikov (2010) create $\lambda^j(y^j)$ in their Lemma 1, which preserves the conditional independence property.

3.2.4 General More-Than-Two-Player Game withOUT Cheap Talk

Finally, we consider how to dispense with the perfect cheap talk, error-reporting noisy cheap talk and public randomization device in the general more-than-two-player game.

Before proceeding to the explanation of message exchange protocols, note that when the players exchange messages via actions, the message exchange becomes payoff relevant. Therefore, for each player i, player $i-1$ changes player i's continuation payoff to cancel out the differences in instantaneous utilities. Since the sender of a message may be different from players i and $i-1$, to control player i's continuation payoff properly, player $i-1$ needs to identify the sender's message and player i's action simultaneously. We will introduce Assumption 48 in Chapter 7 for this purpose.

3.2.4.1 Coordination Block

In Chapter 5, each player communicates x_i via perfect cheap talk in the coordination block.

Error-Reporting Noisy Cheap Talk We first replace the perfect cheap talk with the error-reporting noisy cheap talk. As we have explained in Section 1.5.6.2, with more than two players, it is important to create a message protocol so that, while the players exchange messages and infer the other players' messages in order to coordinate on x_i, there is no player who can induce a situation where some players infer x_i is G while the others infer x_i is B. Since the signals from the error-reporting noisy cheap talk are private as we will see in Section 1.3.2, we need a more sophisticated communication protocol than the case with two players. For that purpose, we will add Assumption 49 in Chapter 7.

Messages via Actions Then, we replace the error-reporting noisy cheap talk with messages via actions. So that we can create a message protocol to preserve the important features that were satisfied by the error-reporting noisy cheap talk, Assumption 51 is sufficient.

3.2.4.2 Report Block

We make the more-than-two-player-case counterparts of Assumptions 40 and 41 to dispense with the public randomization and perfect cheap talk in the report block. So that the players can coordinate through their actions and private signals, we add Assumption 53. In addition, to construct a statistics to preserve the conditional independence property, Assumption 54 is sufficient.

Chapter 4

General Two-Player Game with Cheap Talk

In this chapter, we prove Theorem 7 (folk theorem with cheap talk) for the general two-player game with the perfect cheap talk, error-reporting noisy cheap talk and public randomization devices.

Since there are only two players, when we say players i and j, unless otherwise specified, player i is different from player j.

4.1 Valid Lemmas

Since we maintain Assumptions 3 and 5, Lemmas 12, 13, 17 and 18 are still valid. Also, since the error-reporting noisy cheap talk is available, Lemma 10 holds.

4.2 Intuitive Explanation

The basic structure is the same as in the prisoners' dilemma: in each finitely repeated game, there are L review rounds and several supplemental rounds. In each review round, player j

monitors player i by making a score:

$$X_j(l) = \sum_{t \in T_j(l)} \pi_i[\alpha(x)](y_{j,t}).$$

If the realization of $X_j(l)$ is far from the ex ante mean, then player j will have $\lambda_j(l+1) = B$ and switch to a constant reward $\bar{\pi}_i(x, B, l+1)$.

Remember that player j with $x_j = B$ and $\lambda_j(l+1) = B$ needs to give a non-negative constant reward. On the other hand, player j with harsh strategy $\sigma_j(B)$ and π_i needs to ensure that player i's payoff is below $\underline{v}_i$ regardless of player i's action plan.

In the prisoners' dilemma, $a_j(x)$ with $x_j = B$ defined to satisfy (1.18) happens to be the minimaxing action. Hence, for sufficiently small ρ, player i's payoff is below $\underline{v}_i$ regardless of player i's action plan with a non-negative reward. However, in a general game, $a_j(x)$ with $x_j = B$ is not always a minimaxing action. Therefore, player j with $x_j = B$ and $\lambda_j(l+1) = B$ needs to switch to the minimaxing action if player j thinks that player i has deviated.

Reversing the indices i and j, we incentivize player i with $x_i = B$ and $\lambda_i(l+1) = B$ to switch to the minimaxing action if player i thinks that player j has deviated. The logic is based on Lemma 15.[1] That is, if

1. player i takes $\alpha_i(l) = \alpha_i(x)$;

2. the empirical distribution of $a_{i,t}$'s is close to $\alpha_i(x)$; and

3. player i's signal frequency in the periods where player i takes $a_i(x)$ in $T_i(l)$ is close to the affine hull of player i's signal distributions with respect to player j's action,

then whenever player i observes a signal frequency far from the ex ante mean under $\alpha_j(x)$, player i believes that player j takes $\alpha_j(l) \neq \alpha_j(x)$. After $\alpha_j(l) \neq \alpha_j(x)$, player j makes player i indifferent between any action profile, which means player i is willing to

[1] In Section 2.9, we have proven Lemma 15 for a general game.

take the minimaxing action. Intuitively, $d_i(l+1) \in \{G, B\}$ indicates whether or not player i minimaxes player j from the $(l+1)$th review round. Since the above three conditions are satisfied with a high probability regardless of player j's deviation, player j is punished sufficiently severely.

4.3 Structure of the Phase

As in the prisoners' dilemma, after the lth review round, we have the supplemental rounds for $\lambda_i(l+1)$.

In addition, player i with $\lambda_i(l+1) = B$ and $d_i(l+1) = B$ minimaxes player j in the $(l+1)$th review round. Since player i's reward function is constant after $\lambda_i(l+1) = B$, player j wants to take the static best response to player i's action. To best respond to player i, player j wants to know $d_i(l+1)$. Therefore, we also introduce the supplemental round for $d_i(l+1)$ so that player i can send $d_i(l+1)$ via error-reporting noisy cheap talk with precision $p = 1 - \exp(-T^{\frac{1}{2}})$. The truthtelling incentive will be verified later.

4.4 Preparation

Before constructing the equilibrium, we make three preparations.

4.4.1 Minimax and Values

Fix α_j^* be a player j's minimaxing strategy against player i. Given $\underline{v}_i$, take $\underline{u}$ sufficiently small so that

$$\underline{v}_i > v_i(B) \equiv \max\left\{\max_{x:x_j=B} u_i(a(x)), v_i^*\right\} + 2\underline{u}. \tag{4.1}$$

Given $\underline{u}$, for sufficiently small ρ,

$$v_i(B) \geq \max\left\{\max_{x:x_j=B} w_i(x), \max_{x:x_j=B} u_i(a_i(x), \alpha_j(x)), v_i^*\right\} + \underline{u}. \tag{4.2}$$

Re-take $\bar{u}$ if necessary to assume

$$\bar{u} \geq v_i(B) + \underline{u}, v_i(B) - v_i^* \tag{4.3}$$

4.4.2 Statistics to Monitor Player j

Instead of using player i's signal directly, player i constructs a statistics to infer whether player j has deviated. The following is the lemma to construct such a statistics:

Lemma 22 If Assumption 3 is satisfied, then there exist $q_2 > q_1$ such that, for all $i \in I$ and $a \in A$, there exists a function $\gamma_i^a : Y_i \to (0,1)$ such that player i can statistically infer whether player j takes a_j or not:

$$\mathbb{E}\left[\gamma_i^a(y_i) \mid \tilde{a}_j, a_i\right] = \begin{cases} q_2 & \text{if } \tilde{a}_j = a_j, \\ q_1 & \text{otherwise.} \end{cases}$$

Proof: Assumption 3 guarantees the existence of such γ_i^a. Further, we can assume that $\gamma_i^a : Y_i \to (0,1)$ and $q_1, q_2 \in (0,1)$ by applying an appropriate affine transformation of γ's. ∎

Intuitively, if $\frac{1}{|T_i(l,x)|}\sum_{t \in T_i(l,x)} \gamma_j^{a(x)}(y_{j,t})$ is far from q_2, player i infers that player j has deviated.

4.4.3 Perfect Monitoring

In this subsection, we consider a one-shot game with perfect monitoring parameterized with $l \in \mathbb{N}$ with the same sets of players and their possible actions. The result of this section is used when we consider how player j punishes player i in the T_P-period finitely repeated game with private monitoring.

In the game with parameter $l \in \{1, ..., L-1\}$, player j takes $\alpha_j(x)$. Take a pure strategy $a_i^{BR}(x) \in BR_i(\alpha_j(x))$.

Depending on player i's action, $d_j(l+1) \in \{G, B\}$ is determined as follows: if player i takes $a_i(x)$, then $d_j(l+1) = G$ with probability one. If player i takes $a_i \neq a_i(x)$, then $d_j(l+1) = B$ with probability $p_j^{l+1}(x)$ and $d_j(l+1) = G$ with the remaining probability $1 - p_j^{l+1}(x)$.

Player i's payoff will be the convex combination of the following two variables:

1. "instantaneous utility"

$$\tilde{u}_i(a_i) = \begin{cases} u_i(a_i, \alpha_j(x)) + \underline{u} & \text{if } a_i = a_i^{BR}(x), \\ u_i(a_i, \alpha_j(x)) & \text{if } a_i \neq a_i^{BR}(x), \end{cases}$$

   that is, if player i takes $a_i^{BR}(x)$, player i gets a small "bonus" $\underline{u} > 0$; and

2. "continuation payoff"

$$\mathbb{E}\left[W_i^{l+1}(d_j(l+1)) \mid a_i\right]$$

such that player i's payoff is equal to

$$V_i^l \equiv \max_{a_i} \frac{1}{L-l+1}\tilde{u}_i(a_i) + \frac{L-l}{L-l+1}\mathbb{E}\left[W_i^{l+1}(d_j(l+1)) \mid a_i\right]$$

with

$$W_i^{l+1}(G) = \frac{(L-l-1)\, v_i(B) + \bar{u}}{L-l} + \underline{u}, \tag{4.4}$$

$$W_i^{l+1}(B) = v_i^*. \tag{4.5}$$

In this game, we can show the following lemma:

Lemma 23 For any $L \geq 2$, $q_2 > q_1$ and $\underline{u}$, there exist $\bar{q} > 0$ and $\bar{\rho} > 0$ such that, for any $q < \bar{q}$ and $\rho < \bar{\rho}$, for any $i \in I$, there exist $\{\underline{p}_j^{l+1}(x), \bar{p}_j^{l+1}(x)\}_{l=1}^{L-1}$ with

$$\begin{aligned} \bar{p}_j^{l+1}(x) &\in [0,1] \\ \underline{p}_j^{l+1}(x) &= \frac{q_2 - q - (1 - 2(|A_i| - 1)\rho) q_1}{q_2 - q_1} \bar{p}_j^{l+1} \end{aligned} \tag{4.6}$$

for all $l = 1, ..., L-1$ such that, for $\{p_j^{l+1}(x)\}_{l=1}^{L-1}$,

1. if $p_j^{l+1}(x) \leq \bar{p}_j^{l+1}(x)$ for all l, then it is uniquely optimal for player i to take $a_i^{BR}(x)$; and
2. if $p_j^{l+1}(x) \in \left[\underline{p}_j^{l+1}(x), \bar{p}_j^{l+1}(x)\right]$ for all l, then

$$V_i^l \leq W_i^l(G) = \frac{(L-l) v_i(B) + \bar{u}}{L - l + 1} + \underline{u}.$$

Intuitively, $p_j^{l+1}(x)$ is the probability for player j to switch to minimaxing in the repeated game. $p_j^{l+1}(x)$ is sufficiently high so that player i is indifferent between taking $a_i(x)$ and $a_i^{BR}(x)$ except for a small "bonus" for $a_i^{BR}(x)$, which means player i should take $a_i^{BR}(x)$ as stated in Statement 1 and the punishment is sufficiently severe as stated in Statement 2.

Proof: Fix $\underline{u}$. Then, fix $\bar{\rho}_0$ so that (4.2) holds for all $\rho < \bar{\rho}_0$.

If $a_i(x) = a_i^{BR}(x)$, then take $\bar{p}_j^{l+1}(x) = 0$. The optimality of $a_i^{BR}(x)$ is obvious. In addition, with $u_i^*(x) = \tilde{u}_i(a_i^{BR}(x))$,

$$\begin{aligned} V_i^l &= \frac{1}{L-l+1} u_i^*(x) + \frac{L-l}{L-l+1} W_i^{l+1}(G) \\ &= \frac{1}{L-l+1} (u_i(a_i(x), \alpha_j(x)) + \underline{u}) + \frac{L-l}{L-l+1} \left(\frac{(L-l-1) v_i(B) + \bar{u}}{L-l} + \underline{u} \right) \\ &\leq \frac{1}{L-l+1} v_i(B) + \frac{(L-l-1) v_i(B) + \bar{u}}{L-l+1} + \underline{u} - \frac{1}{L-l+1} \underline{u} \\ &= W_i^l(G) - \frac{1}{L-l+1} \underline{u}, \end{aligned}$$

as desired. The inequality uses (4.2).

Otherwise, define $\bar{p}_j^{l+1}(x) \in (0,1)$ as the solution for

$$\bar{p}_j^{l+1}(x)\frac{(L-l-1)\,v_i\,(B)+\bar{u}}{L-l} + \left(1-\bar{p}_j^{l+1}(x)\right) v_i^* = v_i\,(B) + \underline{u}.$$

With $p_j^{l+1}(x) = \bar{p}_j^{l+1}(x)$,

- if player i takes $a_i^{BR}(x)$, then the payoff is

$$\frac{1}{L-l+1}u_i^*(x) + \frac{L-l}{L-l+1}v_i\,(B) + \frac{L-l}{L-l+1}\underline{u};$$

  and

- if player i takes $a_i(x)$, then the payoff is at most

$$\begin{aligned} & \frac{1}{L-l+1}u_i(a_i\,(x)) + \frac{L-l}{L-l+1}\left(\frac{(L-l-1)\,v_i\,(B)+\bar{u}}{L-l}+\underline{u}\right) \\ \leq \; & \frac{1}{L-l+1}\left(u_i^*(x)-\underline{u}\right) + \frac{L-l}{L-l+1}\left(\frac{(L-l-1)\,v_i\,(B)+\bar{u}}{L-l}+\underline{u}\right) \\ \leq \; & \frac{1}{L-l+1}u_i^*(x) + \frac{L-l}{L-l+1}v_i\,(B) + \frac{L-l}{L-l+1}\underline{u} - \frac{1}{L}\underline{u}. \end{aligned}$$

Hence, it is uniquely optimal for player i to take $a_i^{BR}(x)$.

Since $\frac{(L-l-1)v_i(B)+\bar{u}}{L-l} \geq v_i^*$, for all $p_j^{l+1}(x) \leq \bar{p}_j^{l+1}(x)$, we have the same incentive. In addition, with $\bar{p}_j^{l+1}\,(x)$,

$$\begin{aligned} V_i^l &= \frac{1}{L-l+1}u_i^*(x) + \frac{L-l}{L-l+1}v_i\,(B) + \frac{L-l}{L-l+1}\underline{u} \\ &\leq W_i^l\,(G) - \frac{1}{L-l+1}\underline{u} \leq W_i^l\,(G) - \frac{1}{L}\underline{u}. \end{aligned}$$

Hence, for sufficiently small ρ and q, we have $V_i^l \leq W_i^l\,(G)$ for all $p_j^{l+1}(x) \in [\underline{p}_j^{l+1}(x), \bar{p}_j^{l+1}(x)]$. ∎

4.5 Equilibrium Strategy

As in Section 2.4, we define $\sigma_i(x_i)$ and $\pi_i^{\text{main}}(x_j, h_j^{\text{main}} : \delta)$. In Section 4.5.1, we define the state variables that will be used to define the action plans and rewards. Given the states, Section 4.5.2 defines the action plan $\sigma_i(x_i)$ and Section 4.5.3 defines the reward function $\pi_i^{\text{main}}(x_j, h_j^{\text{main}} : \delta)$. Finally, Section 4.5.4 determines the transition of the states defined in Section 4.5.1.

4.5.1 States x_i, $\lambda_j(l+1)$, $\hat{\lambda}_j(l+1)$, $d_i(l+1)$, $\hat{d}_j(l+1)$ and $\theta_j(l+1)$

The state $x_i \in \{G, B\}$ is determined at the beginning of the phase and fixed. By the perfect cheap talk, x becomes common knowledge.

As in the prisoners' dilemma, $\lambda_j(l+1) \in \{G, B\}$ is player j's state, indicating whether player j's score about player i has been erroneous. On the other hand, $\hat{\lambda}_j(l+1) \in \{G, B\}$ indicates whether player i believes that $\lambda_j(l+1) = G$ or $\lambda_j(l+1) = B$ is likely.

As seen in Section 4.2, $d_i(l+1) \in \{G, B\}$ is player i's state, indicating whether or not player i minimaxes player j. On the other hand, $\hat{d}_j(l+1) \in \{G, B\}$ indicates whether player i believes that $d_j(l+1) = G$ or $d_j(l+1) = B$ (player j minimaxes player i) is likely.

Further, as in the prisoners' dilemma, player j makes player i indifferent between any action profile sequence after some history. If she does in the $(l+1)$th review round, then π_i^{main} will be $\sum_\tau \pi_i^{x_j}[\alpha_j(l)](y_{j,\tau})$ for period τ in the $(l+1)$th review round and after. $\theta_j(l+1) \in \{G, B\}$ is an index of whether player j uses such a reward. See Section 4.5.3 for how the reward function depends on these four states.

4.5.2 Player i's Action Plan $\sigma_i(x_i)$

In the coordination block, each player sends x_i truthfully via perfect cheap talk. Then, the state profile x becomes common knowledge.

In the lth review round, player i takes $\alpha_i(l) \in \Delta(A_i)$ *i.i.d.* within the round. $\alpha_i(l)$ depends on $\hat{\lambda}_j(l)$, $d_i(l)$ and $\hat{d}_j(l)$:

1. if $\hat{\lambda}_j(l) = G$, then

   (a) if $d_i(l) = G$, then, given some fixed set $A_i(x)$ of player i's mixed action plans (see Lemma 20 for the definition of $A_i(x)$);

      i. $\alpha_i(l) = \alpha_i(x)$ with probability $1 - \eta$;

      ii. $\alpha_i(l) = \alpha_i^*$ with probability $\frac{\eta}{2}$; and

      iii. for each $\alpha_i \in A_i(x)$, $\alpha_i(l) = \alpha_i$ with probability $\frac{1}{2|A_i(x)|}\eta$;

   (b) if $d_i(l) = B$;

      i. $\alpha_i(l) = \alpha_i^*$ with probability $1 - \eta$;

      ii. $\alpha_i(l) = \alpha_i(x)$ with probability $\frac{\eta}{2}$; and

      iii. for each $\alpha_i \in A_i(x)$, $\alpha_i(l) = \alpha_i$ with probability $\frac{1}{2|A_i(x)|}\eta$.

   Note that the support of $\alpha_i(l)$'s is constant regardless of $d_i(l)$; and

2. if $\hat{\lambda}_j(l) = B$, then

   (a) if $\hat{d}_j(l) = G$, then $\alpha_i(l) = a_i^{BR}(x)$; and

   (b) if $\hat{d}_j(l) = B$, then $BR_i(\alpha_j^*)$.

In the supplemental rounds for $\lambda_i(l+1)$ and $d_i(l+1)$, respectively, player i sends the message $\lambda_i(l+1)$ and $d_i(l+1)$, respectively, truthfully via error-reporting noisy cheap talk with precision $p = 1 - \exp(-T^{\frac{1}{2}})$.

4.5.3 Reward Function

In this subsection, we explain player j's reward function on player i, $\pi_i^{\text{main}}(x_j, h_j^{\text{main}} : \delta)$.

Reward Function As in the prisoners' dilemma, the reward $\pi_i^{\text{main}}(x_j, h_j^{\text{main}} : \delta)$ is written as

$$\pi_i^{\text{main}}(x_j, h_j^{\text{main}} : \delta) = \sum_{l=1}^{L} \sum_{t \in T(l)} \pi_i^{\delta}(t, a_{j,t}, y_{j,t}) \tag{4.7}$$
$$+ \begin{cases} -2\bar{u}T + \sum_{l=1}^{L} \pi_i^{\text{main}}(x, h_j^{\text{main}}, l) & \text{if } x_j = G, \\ 2\bar{u}T + \sum_{l=1}^{L} \pi_i^{\text{main}}(x, h_j^{\text{main}}, l) & \text{if } x_j = B. \end{cases}$$

Remember that $T(l)$ is the set of periods in the lth review round, that the first term cancels out the effect of discounting, and that $\pi_i^{\text{main}}(x, h_j^{\text{main}}, l)$ is the reward for the lth review round.

Reward Function for the lth Review Round Next we define $\pi_i^{\text{main}}(x, h_j^{\text{main}}, l)$ for each $l = 1, ..., L$. There are following two cases: in the lth review round,

1. if $\theta_j(l) = B$, then player j makes player i indifferent between any action profile by

$$\pi_i^{\text{main}}(x, h_j^{\text{main}}, l) = \sum_{t \in T(l)} \pi_i^{x_j}[\alpha_j(l)](y_{j,t}). \tag{4.8}$$

   Remember that by (2.8),

$$\sum_{t \in T(l)} \pi_i^{x_j}[\alpha_j(l)](y_{j,t}) \begin{cases} \leq 0 & \text{if } x_j = G, \\ \geq 0 & \text{if } x_j = B \end{cases}$$

   and (1.17) is not an issue after $\theta_j(l) = B$; and

2. otherwise, that is, if $\theta_j(l) = G$, then player j's reward on player i is based on the state profile x, the index of the past erroneous score $\lambda_j(l)$, index of minimaxing $d_j(l)$ and

player j's score about player i, $X_j(l)$:

$$\pi_i^{\text{main}}(x, h_j^{\text{main}}, l) = \begin{cases} \bar{\pi}_i(x, G, d_j(l), l) + X_j(l) + \pi_i^{x_j}[\alpha_j(x)](y_{j,t_j(l)}) & \text{if } \lambda_j(l) = G, \\ \bar{\pi}_i(x, B, d_j(l), l) & \text{if } \lambda_j(l) = B. \end{cases} \tag{4.9}$$

Here, $\bar{\pi}_i(x, \lambda_j(l), d_j(l), l)$ is a constant with

$$\bar{\pi}_i(x, \lambda_j(l), d_j(l), l) \begin{cases} \leq 0 & \text{if } x_j = G, \\ \geq 0 & \text{if } x_j = B \end{cases} \tag{4.10}$$

that will be determined later so that (1.15), (1.16) and (1.17) are satisfied.

Note that, as in the two-player prisoners' dilemma, one period $t_j(l)$ is randomly excluded from the construction of the score and we add $\pi_i^{x_j}[\alpha_j(x)](y_{j,t_j(l)})$ separately.

In addition, as we mentioned in Lemma 23, player j with $\lambda_j(l) = B$ and $d_j(l) = G$ gives a small bonus on $\underline{u}$ for taking $a_i^{BR}(x)$. By Assumption 3, there exists $\pi_i^{\underline{u}}[\alpha_j(l)](y_j)$ so that

$$\mathbb{E}\left[\pi_i^{\underline{u}}[\alpha_j(l)](y_j) \mid a_i, \alpha_j(l)\right] = \begin{cases} \underline{u} & \text{if } a_i = a_i^{BR}(x), \\ 0 & \text{if } a_i \neq a_i^{BR}(x). \end{cases}$$

Player j with $\lambda_j(l) = B$ and $d_j(l) = G$ adds

$$\sum_{t \in T(l)} \pi_i^{\underline{u}}[\alpha_j(l)](y_{j,t}) \tag{4.11}$$

to $\pi_i^{\text{main}}(x, h_j^{\text{main}}, l)$. For sufficiently small $\underline{u}$, $\pi_i^{\underline{u}}[\alpha_j(l)](y_j)$ is sufficiently small for all y_j. Hence, when we consider (1.17), we ignore (4.11).

4.5.4 Transition of the States

In this subsection, we explain the transition of player i's states. Since x_i is fixed in the phase, we consider the following five states:

4.5.4.1 Transition of $\lambda_j(l+1) \in \{G, B\}$

The transition of $\lambda_j(l+1) \in \{G, B\}$ and the proof that (4.10) is sufficient for (1.17) are the same as in Section 2.4.4.1.

4.5.4.2 Transition of $\hat{\lambda}_j(l+1) \in \{G, B\}$

The transition of $\hat{\lambda}_j(l+1) \in \{G, B\}$ is the same as in Section 2.4.4.2 except that, if player i has $\hat{\lambda}_j(l) = G$ and $d_i(l) = B$, then $\hat{\lambda}_j(l+1) = G$. As explained in Section 4.2, if $d_i(l) = B$, then player i believes that player j has made player i indifferent between any action profile. Hence, the belief about $\lambda_j(l)$ is irrelevant for almost optimality.

Specifically, $\hat{\lambda}_j(1) = G$. If $\hat{\lambda}_j(l) = B$, then $\hat{\lambda}_j(l+1) = B$. Below, we explain how $\hat{\lambda}_j(l+1) \in \{G, B\}$ is defined given $\hat{\lambda}_j(l) = G$.

Suppose player i's history in the lth review round satisfies the following two conditions, then player i disregards the message and

$$\hat{\lambda}_j(l+1) = G :$$

1. $d_i(l) = B$, that is, player i believes that player j has taken $\alpha_j(\tilde{l}) \neq \alpha_j(x)$ for $\tilde{l} \leq l-1$; or

2. $d_i(l) = G$ and the following two conditions are satisfied:

   (a) player i takes $\alpha_i(l) = \alpha_i(x)$; and

   (b) (2.20), (2.22) and (2.23) are satisfied. Remember the discussion in Section 2.4.4.2, which implies the probability of (2.22) is independent of player j's strategy.

Otherwise, player i obeys the message and

$$\hat{\lambda}_j(l+1) = f[i](\lambda_j(l+1)). \tag{4.12}$$

Note that (2.20), (2.21), (2.22) and (2.23) are well defined for the general game in Section 2.4.4.2.

4.5.4.3 Transition of $d_i(l+1) \in \{G, B\}$

Remember that $d_i(l+1) = B$ is the index of whether player i minimaxes player j in the $(l+1)$th review round. For the first round, $d_i(1) = G$ (no punishment in the initial review round). If $d_i(l) = B$, then $d_i(l+1) = B$. Below, we explain how $d_i(l+1) \in \{G, B\}$ is defined given $d_i(l) = G$.

Intuitively, by Lemma 22, for periods where player i takes $a_i(x)$, if player i observes a lot of low $\gamma_i^{a(x)}(y_{i,t})$, then it is more likely that player j takes $\alpha_j(l) \neq \alpha_j(x)$.

Hence, if player i's history in the lth review round satisfies the following four conditions, then player i minimaxes player j from the next review round:

$$d_i(l+1) = B:$$

1. player i uses the harsh strategy: $x_i = B$;

2. player i takes $\alpha_i(l) = \alpha_i(x)$;

3. (2.20), (2.22) and (2.23) are satisfied, which implies (2.21); and

4. $\frac{1}{|T_i(l,x)|} \sum_{t \in T_i(l,x)} \gamma_i^{a(x)}(y_{i,t})$ is low.

Otherwise, $d_i(l+1) = G$.

We are left to define Condition 4 formally. First, as $Q_i(x)\mathbf{1}_{y_i}$ and $\mathbf{1}_{Q_i(x)}$, we define $\Gamma_i^{a(x)} \in \{0, 1\}$ from $\gamma_i^{a(x)}(y_i)$: player i after taking $a_i(x)$ and observing y_i, player i draws a random variable from the uniform $[0, 1]$. If the realization of this random variable is no more than $\gamma_i^{a(x)}(y_i)$, then $\Gamma_i^{a(x)} = 1$. Otherwise, $\Gamma_i^{a(x)} = 0$. Lemma 22 implies

$$\Pr\left(\left\{\Gamma_i^{a(x)} = 1\right\} \mid \tilde{a}_j, a_i(x)\right) = \begin{cases} q_2 & \text{if } \tilde{a}_j = a_j(x), \\ q_1 & \text{otherwise.} \end{cases}$$

Second, let Condition 4 be satisfied if

- $\frac{1}{|T_i(l,x)|}\sum_{t\in T_i(l,x)} \gamma_i^{a(x)}(y_{i,t})$ and $\frac{1}{|T_i(l,x)|}\sum_{t\in T_i(l,x)} \Gamma_{i,t}^{a(x)}$ are close:

$$\left\| \frac{1}{|T_i(l,x)|} \sum_{t\in T_i(l,x)} \gamma_i^{a(x)}(y_{i,t}) - \frac{1}{|T_i(l,x)|} \sum_{t\in T_i(l,x)} \Gamma_{i,t}^{a(x)} \right\| < \frac{q}{3}; \tag{4.13}$$

  and

- player i draws a random variable from the uniform $[0,1]$ and the realization of this random variable is no less than

$$\bar{p}_i^{l+1}(x) \min\left\{1, \frac{\left\{q_2 T - qT - \sum_{t\in T_i(l,x)} \Gamma_{i,t}^{a(x)}\right\}_+}{q_2 T - q_1 T}\right\}. \tag{4.14}$$

As we have adjusted the probability of (2.22) in Section 2.4.4.2, we adjust the probability of (4.13) so that the probability of (4.13) is independent of $\{a_{i,t}, y_{i,t}\}_{t\in T(l)}$. When we say (4.13) is satisfied, we take this adjustment into account.

Notice that, if $d_i(l+1) = B$, then (4.13) and (4.14) imply

$$\sum_{t\in T_i(l,x)} \gamma_i^{a(x)}(y_{i,t}) \leq q_2 T - \frac{2}{3} qT.$$

Together with (2.20), we have

$$\begin{aligned}\frac{1}{|T_i(l,x)|}\sum_{t\in T_i(l,x)}\gamma_i^{a(x)}(y_{i,t}) &\leq q_2\frac{T}{|T_i(l,x)|}-\frac{2}{3}q\frac{T}{|T_i(l,x)|}\\ &\leq q_2\frac{1}{1-2(|A_i|-1)\rho}-\frac{2}{3}q\frac{1}{1-2(|A_i|-1)\rho}\\ &\leq q_2-\frac{q}{2} \qquad (4.15)\end{aligned}$$

for sufficiently small ρ. By Lemma 22, this implies that player i's signal frequency is far away from $q_i(x)$. Since (2.20), (2.22) and (2.23) imply that player i's signal frequency is in the affine hull of $q_i(a_i(x), a_j)$ with respect to $a_j \in A_j$, there exists $\alpha_j \in A_j(x)$ with $\alpha_j \neq \alpha_j(x)$ such that player i's signal frequency is skewed towards $q_i(a_i(x), \alpha_j)$ and player i believes that player j took $\alpha_j \in A_j(x)$ with $\alpha_j \neq \alpha_j(x)$ in the lth review round. As we will see in Section 4.5.4.5, this implies that player i believes that any action will be optimal from the $(l+1)$th review round and will have an incentive to minimax player j.

4.5.4.4 Transition of $\hat{d}_j(l+1) \in \{G, B\}$

Player j sends $d_j(l+1)$ via error-reporting noisy cheap talk in the supplemental round for $d_j(l+1)$. Player i always obeys the signal:

$$\hat{d}_j(l+1) = f[i](d_j(l+1)). \qquad (4.16)$$

4.5.4.5 Transition of $\theta_j(l+1) \in \{G, B\}$

As we have seen in Section 2.4.3, $\theta_j(l+1) = B$ implies that player i is indifferent between any action profile (except for the incentives from π_i^{report}).

As in the two-player prisoners' dilemma, if $\theta_j(l) = B$, then $\theta_j(l+1) = B$. Hence, we concentrate on how $\theta_j(l+1) \in \{G, B\}$ is defined conditional on $\theta_j(l) = G$. $\theta_j(l+1) = B$ if one of the following five conditions is satisfied:

1. when player j sends $\lambda_j(l)$ by the error-reporting noisy cheap talk, the error is reported: $g[j](\lambda_j(l)) = E$;

2. when player j sends $d_j(l)$ by the error-reporting noisy cheap talk, the error is reported: $g[j](d_j(l)) = E$;

3. at the beginning of the $(l+1)$th review round, player j with $\hat{\lambda}_j(l+1) = G$ and $d_j(l+1) = G$ takes $\alpha_j(l+1) \neq \alpha_j(x)$. With abuse of notation, we include the case with $l = -1$;

4. at the beginning of the $(l+1)$th review round, player j with $\hat{\lambda}_j(l+1) = G$ and $d_j(l+1) = B$ takes $\alpha_j(l+1) \neq \alpha_j^*$. With abuse of notation, we include the case with $l = -1$; or

5. (2.20), (2.22) or (2.23) is not satisfied for player j (that is, with indices i and j reversed).

Otherwise, $\theta_j(l+1) = G$.

There are following four implications. First, as in the two-player prisoners' dilemma, player j makes player i indifferent between any action profile after receiving $g[j](m) = E$ after sending some message m. Hence, it is almost optimal for player i to obey the message.

Second, if player j obeys player i's message about $\lambda_i(l+1)$ (this is the necessary condition for having $\hat{\lambda}_i(l+1) = B$) or player i's message about $d_i(l+1)$ affects player j's continuation strategy ($\hat{\lambda}_i(l+1) = B$ is the necessary condition for this), then $\theta_j(l+1) = B$. Hence, player i is indifferent between any message.

Third, if $\hat{\lambda}_i(l+1) = G$, then Conditions 3 and 4 imply that if "$d_i(l+1) = G$ and $\alpha_j(l+1) \neq \alpha_j(x)$" or "$d_i(l+1) = G$ and $\alpha_j(l+1) \neq \alpha_j^*$," then player i is indifferent between any action profile.

Fourth, the distribution of $\theta_j(l+1)$ is independent of player i's strategy.

4.6 Player i's Belief about Optimal Actions

As in the two-player prisoners' dilemma, we formally show that player i believes that $\hat{\lambda}_j(l+1) = \lambda_j(l+1)$ or $\theta_j(l+1) = B$ with high probability $1 - \exp(-\Theta(T^{\frac{1}{2}}))$, conditional on that player j has the state $\hat{\lambda}_i(l+1) = G$.

In addition, we show that if player i minimaxes player j ($\hat{\lambda}_j(l+1) = G$ and $d_i(l+1) = B$), then player i believes that $\theta_j(l+1) = B$ with high probability $1-\exp(-\Theta(T^{\frac{1}{2}}))$, conditional on that player j has the state $\hat{\lambda}_i(l+1) = G$.

Lemma 24 For all $\bar{u}$ and L, there exists $\bar{q}$ such that, for all $q < \bar{q}$, there exists $\bar{\eta}$ such that, for all $\eta < \bar{\eta}$, there exists $\bar{\rho}$ and $\bar{\varepsilon}$ such that, for all $\rho < \bar{\rho}$ and $\varepsilon < \bar{\varepsilon}$, for all $i \in I$ and $l \in \{1, ..., L\}$, for any history h_i^t with t being in the lth review round, conditional on $\hat{\lambda}_i(\tilde{l}) = G$ for all $\tilde{l} \leq l$ and $\alpha_j(l)$, player i after h_i^t believes that

1. $\hat{\lambda}_j(l) = \lambda_j(l)$ or $\theta_j(l) = B$;
2. $\hat{d}_j(l) = d_j(l)$ or $\theta_j(l) = B$; and
3. if $d_i(l) = B$, then $\theta_j(l) = B$.

with high probability $1 - \exp(-\Theta(T^{\frac{1}{2}}))$.

Proof: The proof of the first claim is the same as in Lemma 15 except that player i's learning from player j's continuation action plan is small since player j with $\hat{\lambda}_i(\tilde{l}) = G$ takes any action $A_j(x) \cup \{\alpha_j^*\}$ with a positive probability no less than $\frac{1}{2|A_j|}\eta$.

The second is proven by Lemma 10: player i believes that, conditional on $d_j(l)$,

$$\hat{d}_j(l) = f[i](d_j(l)) = d_j(l)$$

or $g[j](d_j(l)) = E$ (this implies $\theta_j(l) = B$) with probability no less than $1 - \exp(-\Theta(T^{\frac{1}{2}}))$. Since player j's continuation action plan in the main blocks does not depend on $g[j](d_j(l))$, we are done.

Let us prove the third claim. First, Take $\bar{q}$ and $\bar{\rho}_1$ such that Lemma 23 holds. Fix $q < \bar{q}$. Given q, take $\bar{\eta}_2$, $\bar{\rho}_2$ and $\bar{\varepsilon}_2$ such that for all $\eta < \bar{\eta}_2$, $\rho < \bar{\rho}_2$ and $\varepsilon < \bar{\varepsilon}_2$, (4.14) implies

$$\left\| \frac{1}{|T_i(l,x)|} \sum_{t \in T_i(l,x)} \mathbf{1}_{y_{i,t}} - q_i(x) \right\| \geq \eta.$$

Finally, fix $\bar{\eta} < \bar{\eta}_2$ such that Lemma 20 is satisfied. Then, for all $\eta < \bar{\eta}$, we can take take $\rho < \min\{\bar{\rho}_1, \bar{\rho}_2\}$ and $\varepsilon < \bar{\varepsilon}_2$ such that Lemma 20 holds for η, ρ and ε.

Then, take l^* such that $d_i(l^*) = G$ and $d_i(l^*+1) = B$. Then, for the l^*th review round, the premise of Lemma 20 is satisfied, which means player i believes $\theta_i(l^*) = B$ with probability $1 - \exp(-\Theta(T))$ conditional on $\hat{\lambda}_i(l^*) = G$. As for the first claim, player i's learning from player j's continuation action plan is small, as desired. ∎

4.7 Variables

In this section, we show that we can take all the variables necessary for the equilibrium construction appropriately: $\bar{u}$, q_2, q_1, $\underline{u}$, $\bar{q}$, L, η, ρ and ε.

First, Lemma 12 determines $\bar{u}$ and Lemma 22 determines q_1 and q_2.

Second, from (1.18), we have

$$\max\left\{v_i^*, \max_{x:x_j=B} u_i(a(x))\right\} < \underline{v}_i < v_i < \overline{v}_i < \min_{x:x_j=G} u_i(a(x)).$$

Take $\underline{u}$ sufficiently small so that

$$v_i(B) + \underline{u} < \underline{v}_i < v_i < \overline{v}_i < \min_{x:x_j=G} u_i(a(x))$$

with $v_i(B) = v \max\{v_i^*, \max_{x:x_j=B} u_i(a(x))\} + 2\underline{u}$ as defined in (4.1). Re-take $\bar{u}$ if necessary to have (4.3).

Third, take L sufficiently large such that

$$\frac{(L-1)\,v_i\,(B)+\bar{u}}{L}+\underline{u}+2\frac{\bar{u}}{L}<\underline{v}_i<v_i<\overline{v}_i<\min_{x:x_j=G}u_i(a(x))-2\frac{\bar{u}}{L}.$$

Fourth, given $L, q_2, q_1, \underline{u}$, fix $\bar{q}_1$ and $\bar{\rho}_1$ so that Lemma 23 holds.

Fifth, take $\bar{\rho}_2<\bar{\rho}_1$ sufficiently small so that for all $\rho<\bar{\rho}_2$, (4.2) holds and

$$\frac{(L-1)\,v_i\,(B)+\bar{u}}{L}+\underline{u}+2\frac{\bar{u}}{L}<\underline{v}_i<v_i<\overline{v}_i<\min_{x:x_j=G}w_i(x)-2\frac{\bar{u}}{L}.$$

Sixth, given $\bar{u}$ and L, take $\bar{\eta}$ so that Lemma 15 holds and for all $\eta<\bar{\eta}$, we have

$$\begin{aligned}&\frac{(L-1)\,v_i\,(B)+\bar{u}}{L}+\underline{u}+2\frac{\bar{u}}{L}+\eta L\left(2\bar{u}-\min_{i,x}\frac{(L-1)\,v_i\,(B)+\bar{u}}{L}\right)\\<\ &\underline{v}_i<\overline{v}_i<\min_{x:x_j=G}w_i\,(x)-2\frac{\bar{u}}{L}-\eta L\left(2\bar{u}+\max_{i,x}w_i\,(x)\right).\end{aligned}\tag{4.17}$$

Finally, fix $q<\bar{q}_1$ sufficiently small. Then, we can take $\eta<\bar{\eta}$, $\rho<\bar{\rho}_2$ and $\varepsilon<\bar{\varepsilon}_2$ so that (4.15) holds for fixed q and Lemma 24 holds.

4.8 Almost Optimality

Based on Lemma 24, we now show that if we properly define $\bar{\pi}_i(x,\lambda_j(l),d_j(l),l)$, then $\sigma_i(x_i)$ and π_i^{main} satisfy (1.21), (1.16) and (1.17).

We show that $\sigma_i(x_i)$ satisfies the following proposition by backward induction:

Proposition 25 For all $i\in I$, there exists $\bar{\pi}_i(x,\lambda_j(l),d_j(l),l)$ such that

1. $\sigma_i(x_i)$ is almost optimal conditional on $\hat{\lambda}_i(l)=G$: for each $l\in\{1,...,L\}$, conditional on $\hat{\lambda}_i(l)=G$,

   (a) for any period t in the lth review round, (1.21) holds; and

(b) when player i sends the message about $\lambda_i(l+1)$ or $d_i(l+1)$ by the error-reporting noisy cheap talk, (1.21) holds;[2]

2. (1.16) is satisfied with π_i replaced with π_i^{main}. Since each $x_i \in \{G, B\}$ gives the same value conditional on x_j, the strategy in the coordination block is optimal;[3] and

3. π_i^{main} satisfies (1.17).

Note that $\hat{\lambda}_i(l) = B$ implies $\theta_j(l) = B$ and that player i can condition on $\hat{\lambda}_i(l) = G$.

1-(b) follows from the following two facts: first, whenever player i's messages change player j's action, $\theta_j(l+1) = B$. Second, since there is an error of order $\exp(-\Theta(T^{\frac{1}{2}}))$ after any message, given Lemma 24, the effect on player i's learning about the optimal action is negligible.

To show 3, as in the prisoners' dilemma, it suffices to have

$$\bar{\pi}_i(x, \lambda_j(l), d_j(l), l) \begin{cases} \leq 0 & \text{if } x_j = G, \\ \geq 0 & \text{if } x_j = B, \end{cases} \tag{4.18}$$

$$|\bar{\pi}_i(x, \lambda_j(l), d_j(l), l)| \leq \max_{i,a} 2\,|u_i\,(a)|\,T \tag{4.19}$$

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

We are left to construct $\bar{\pi}_i$ so that 1-(a) and 2 are satisfied together with (4.18) and (4.19). Below, we consider the cases with $x_j = G$ and $x_j = B$ separately.

4.8.1 Case 1: $x_j = G$

Note that $d_j(l) = G$ for all l if $x_j = G$ (see Section 4.5.4.3). Hence, the logic is the same as in the prisoners' dilemma.

[2] With $l = L$, this is redundant.

[3] This is not precise since we will further adjust the reward function based on the report block. However, as we will see, even after the adjustment of the report block, any $x_i \in \{G, B\}$ still gives exactly the same value and so the strategy in the coordination block is exactly optimal.

We start backward induction from the Lth review round. Suppose that player j uses (4.9) in the Lth round. If $\lambda_j(L) = \hat{\lambda}_j(L)$ and $\hat{d}_j(L) = d_j(L)$, then from (4.9) and Section 4.5.2, if player i obeys $\sigma_i(x_i)$, then player i's average continuation payoff except for $\bar{\pi}_i$ is equal to $w_i(x)$ if $\lambda_j(L) = G$ and $u_i(a_i^{BR}(x), \alpha_j(x)) \geq w_i(x)$ if $\lambda_j(L) = B$.

Therefore, for $l = L$, there exists $\bar{\pi}_i(x, \lambda_j(l), d_j(l), l)$ with (4.18) and (4.19) such that player i's average continuation payoff is equal to $w_i(x)$ if (4.9) is used, $\lambda_j(L) = \hat{\lambda}_j(L)$, and $d_j(L) = \hat{d}_j(L)$.

Consider the almost optimality of $\sigma_i(x_i)$. For almost optimality, Lemma 24 guarantees that player i can always believe that (4.9) is used, that $\lambda_j(L) = \hat{\lambda}_j(L)$, and that $d_j(L) = \hat{d}_j(L)$. Therefore, if $\hat{\lambda}_j(L) = G$, then any action is almost optimal for player i and if $\hat{\lambda}_j(L) = B$, then "$a_i^{BR}(x)$ and $BR_i\left(\alpha_j^*\right)$, respectively, are optimal if $\hat{d}_j(L) = G$ and $\hat{d}_j(L) = B$, respectively," as desired.

Consider the $(L-1)$th review round. Since (i) we have define $\bar{\pi}_i(x, \lambda_j(L), d_j(L), L)$ such that player i's value from the Lth review round is independent of $\lambda_j(L)$ as long as $\lambda_j(L) = \hat{\lambda}_j(L)$, (ii) Lemma 24 implies that player i in the main blocks does not put a belief more than $\exp(-\Theta(T^{\frac{1}{2}}))$ on the events that $\lambda_j(L) \neq \hat{\lambda}_j(L)$ and $\theta_j(L) = G$, (iii) the distribution of $\theta_j(L)$ is independent of player i's strategy, and (iv) $d_j(L)$ is fixed, and so we can assume that player i in the $(L-1)$th review round maximizes

$$\frac{1}{T}\mathbb{E}\left[\sum_{t\in T(L-1)} u_i(a_t) + \pi_i^{\text{main}}(x, h_j^{\text{main}}, L-1) \mid x\right],$$

assuming $\lambda_j(L-1) = \hat{\lambda}_j(L-1)$ and $d_j(L-1) = \hat{d}_j(L-1)$.

Therefore, by the same reason as in the Lth review round, there exists $\bar{\pi}_i(x, \lambda_j(l), d_j(l), l)$ with (4.18) and (4.19) such that $\sigma_i(x_i)$ is almost optimal and player i's average payoff from the $(L-1)$th review round is equal to $w_i(x)$ if (4.9) is used, $\lambda_j(L-1) = \hat{\lambda}_j(L-1)$, $d_j(L-1) = \hat{d}_j(L-1)$, and player i obeys $\sigma_i(x_i)$.

If (4.9) is used in the $(L-1)$th review round, then (4.8) will be used in the Lth review round with probability no more than η (player j takes $\alpha_j(L) \neq \alpha_j(x)$ with $d_j(L) = G$ or $\alpha_j(L) \neq \alpha_j^*$ with $d_j(L) = B$) plus $\exp(-\Theta(T^{\frac{1}{2}}))$ (player j obeys the message or an error is reported in the supplemental rounds). When (4.8) is used, per period payoff is bounded by $[-2\bar{u}, 2\bar{u}]$ by (2.9).

Therefore, for $l = L-1$, re-taking $\bar{\pi}_i(x, \lambda_j(l), d_j(l), l)$ if necessary, player i's average continuation payoff for the next two review rounds is equal to $w_i(x) - \eta\left(2\bar{u} + \max_{i,x} w_i(x)\right)$ in the limit as $\delta \to 1$if player i obeys $\sigma_i(x_i)$.

Recursively, for $l = 1$, 1-(a) is satisfied and the average ex ante payoff of player i from the first review round is $w_i(x) - \eta L\left(2\bar{u} + \max_{i,x} w_i(x)\right)$ if $x_j = G$. Note that, in the first review round, $\lambda_j(1) = \hat{\lambda}_j(1) = G$ and $d_j(1) = \hat{d}_j(1) = G$.

Taking the first term $-2\bar{u}T$ in (4.7) into account, the average ex ante payoff is $w_i(x) - 2\frac{\bar{u}}{L} - \eta L\left(2\bar{u} + \max_{i,x} w_i(x)\right)$ if $x_j = G$.

From (4.17), we can further modify $\bar{\pi}_i(x, G, G, 1)$ with (4.18) and (4.19) such that $\sigma_i(x_i)$ gives $\bar{v}_i$ if $x_j = G$. Therefore, 2 is satisfied.

4.8.2 Case 2: $x_j = B$

The main difference from the case with $x_j = G$ is that, if $x_j = B$, then player i needs to take into account that player i's action in the lth review round will affect the probability of being minimaxed in the next review round ($d_j(l+1) = B$).

As in the case with $x_j = G$, $\hat{\lambda}_i(l) = B$ implies $\theta_i(l) = B$ and player i is indifferent between any action profile sequence. Hence, player i conditions $\hat{\lambda}_i(l) = G$. In addition, the distribution of $\theta_j(l+1)$ is independent of player i's strategy. Since $\theta_j(l) = B$ does not happen with probability more than $\eta + \exp(-\Theta(T^{\frac{1}{2}}))$ in each main block, we can deal with the effect of $\theta_j(l) = B$ as in the case with $x_j = G$. Therefore, we assume that $\theta_j(l) = G$ for each round.

Further, by Lemma 24, player i can neglect the possibility of mis-coordination for almost optimality. Hence, we assume that, for each l, $\hat{\lambda}_j(l) = \lambda_j(l)$, $\hat{d}_j(l) = d_j(l)$ and $\hat{\lambda}_i(l) = G$.

In the Lth review round, consider the following cases:

1. if player j uses (4.9) and $\lambda_j(L) = \hat{\lambda}_j(L) = G$, then any action is almost optimal. The average payoff of player i in the Lth review round except for $\bar{\pi}_i$ is $w_i(x)$;

2. if player j uses (4.9) and $\lambda_j(L) = \hat{\lambda}_j(L) = B$, then there are following two cases:

   (a) if $d_j(L) = \hat{d}_j(L) = G$, then $a_i^{BR}(x)$ is optimal and gives the average payoff $u_i^*(x)$. Note that we take (4.11) into account when we consider the payoff; and

   (b) if $d_j(L) = \hat{d}_j(L) = B$, then $BR_i(\alpha_j^*)$ is optimal and gives the average payoff v_i^*.

Therefore, for $l = L$, there exists $\bar{\pi}_i(x, \lambda_j(l), d_j(l), l)$ with (4.18) and (4.19) such that player i's average continuation payoff is equal to $\bar{u} + \underline{u}$ if 1 or 2-(a) is the case and v_i^* if 2-(b) is the case. Note that the former is higher than the latter by

$$\bar{u} + \underline{u} - v_i^*. \tag{4.20}$$

In the $(L-1)$th review round,

1. if player j uses (4.9) and $\lambda_j(L-1) = \hat{\lambda}_j(L-1) = G$, then any action is almost optimal since (i) the payoffs from $\lambda_j(L) = G$ and $\lambda_j(L) = B$ with $d_j(L) = G$ are the same and (ii) $\lambda_j(L-1) = G$ implies $d_j(L) = G$ (See Section 4.5.4.3) and player i can neglect the effect of player i's action in the $(L-1)$th review round on $d_j(L)$.

   The average payoff of player i from the $(L-1)$ and Lth review rounds except for $\bar{\pi}_i(x, L-1, \lambda_j(L-1), \hat{d}_i(L-1))$ is no more than

$$\frac{w_i(x) + \bar{u}}{2} + \frac{\underline{u}}{2} \leq \frac{v_i(B) + \bar{u}}{2} + \underline{u};$$

2. suppose that player j uses (4.9) and $\lambda_j(L-1) = \hat{\lambda}_j(L-1) = B$. Now, $\lambda_j(L)$ is fixed at B. Hence, the relevant cases are the following two:

   (a) if $d_j(L-1) = \hat{d}_j(L-1) = G$, then $a_i^{BR}(x)$ is optimal.

   To see why, remember that player j will have $d_j(L) = B$ with probability (4.14) plus negligible adjustment based on (4.13). The marginal decrease of this probability by not taking the static best response is bounded by

$$\bar{p}_j^L(x) \frac{\text{marginal decrease of } \Pr\left(\frac{1}{T}\left\{\Gamma_{i,t}^{a(x)} = 1\right\} \mid a_j(x)\right)}{q_2 - q_1} \leq \frac{\bar{p}_j^L(x)}{T}.$$

   On the other hand, the maximum gain of preventing $d_j(L) = B$ is equal to T times (4.20). Therefore, the expected gain is bounded by $\bar{p}_j^L(x)$ times (4.20).

   Since (4.20) corresponds to the gain of preventing $d_j(L) = B$ for $l = L-1$ in Lemma 23, player j should take $a_i^{BR}(x)$ for sufficiently large δ.[4]

   Note that player i does not have an incentive to take $a_i \neq a_j(x), a_i^{BR}(x)$ since $\Pr\left(\frac{1}{T}\left\{\Gamma_{i,t}^{a(x)} = 1\right\} \mid a_j(x)\right)$ is the same as $a_i^{BR}(x)$ and the instantaneous utility gain is smaller than $a_i^{BR}(x)$.

   Given player i's strategy, if $\lambda_j(L-1) = \hat{\lambda}_j(L-1) = B$ and $\hat{d}_i(L-1) = \hat{d}_i(L-1)(i) = G$, then, conditional on $\theta_j(L) = G$, $\hat{d}_i(L)$ happens with probability

$$\bar{p}_j^{l+1}(x) \min\left\{1, \frac{q_2 - q - (1 - 2(|A_i| - 1)\rho) q_1}{q_2 - q_1}\right\}$$

   from (4.14). Therefore, Lemma 23 guarantees that the average payoff from the $(L-1)$th and Lth main block is no more than

$$\frac{v_i(x) + \bar{u}}{2} + \underline{u};$$

[4] And so large T from (1.8).

(b) if $d_j(L-1) = \hat{d}_j(L-1) = B$, then $d_j(L)$ is fixed at B. Therefore, $BR_i(\alpha_j^*)$ is optimal and this gives the average payoff v_i^*.

Therefore, for $l = L-1$, there exists $\bar{\pi}_i(x, \lambda_j(l), d_j(l), l)$ with (4.18) and (4.19) such that player i's average continuation payoff is equal to $\frac{\max\left\{v_i(x), v_i^*\right\}+\bar{u}}{2} + \underline{u}$ if 1 or 2-(a) is the case and v_i^* if 2-(b) is the case.

Recursively, for $l = 1$, Proposition 16 is satisfied and the average ex ante payoff of player i at the first review round is $\frac{(L-1)v_i(x)+\bar{u}}{L} + \underline{u}$. Note that, in the first review round, $\lambda_j(1) = \hat{\lambda}_j(1) = G$ and $d_j(1) = \hat{d}_j(1) = G$.

Taking the probability of having $\theta_j(l) = B$ and the first term $2\bar{u}T$ in (4.7) into account, the average ex ante payoff is $\frac{(L-1)v_i(x)+\bar{u}}{L} + \underline{u} + 2\frac{\bar{u}}{L} + L\eta\left(2\bar{u} - \min_{i,x:x_j=B} v_i(x)\right)$.

From (4.17), we can further modify $\bar{\pi}_i(x, G, G, 1)$ with (4.18) and (4.19) such that $\sigma_i(x_i)$ gives $\underline{v}_i$ if $x_j = B$. Therefore, 2 is satisfied.

4.9 Exact Optimality

The report block is the same as in the prisoners' dilemma except that

1. for the proof for Lemma 21, player i can also learn $\Gamma_j^{a(x)}$. Since $\Gamma_j^{a(x)}$ satisfies full support $\gamma_j^{a(x)}(y_j) \in (0,1)$, the same proof works;

2. $V_i(\mathfrak{h}_i^r, \#_i^r)$ also includes the effect on learning $d_j(l)$; and

3. we have additional supplemental rounds.

Chapter 5

General N-Player Game with Cheap Talk

In this chapter, we prove Theorem 7 (folk theorem) for the general N-player game with $N \geq 3$, perfect cheap talk, error-reporting noisy cheap talk and public randomization. See Chapter 7 for the dispensability of special forms of cheap talk and public randomization.

In this chapter, all the error-reporting noisy cheap talk has precision $1 - \exp(-T^{\frac{1}{2}})$. In addition, when we say player i with $i \notin \{1, ..., N\}$, it means player $i \pmod N$. In particular, player 0 is player N and player $N+1$ is player 1.

Fix $v \in \text{int}(F^*)$ arbitrarily. We will find $\{\sigma_i(x_i)\}_{i,x_i}$ and $\left\{\pi_i(x_{i-1}, h_{i-1}^{T_P+1} : \delta)\right\}_{i,x_{i-1}}$ in the finitely repeated game with (1.15), (1.16) and (1.17).

As in Chapter 1, let

$$v_i^* \equiv \min_{\alpha_{-i} \in \Delta(A_{-i})} \max_{a_i \in A_i} u_i(a_i, \alpha_{-i})$$

be the minimax payoff (by independently mixed strategies). In addition, let $\alpha_{-i}^* \equiv (\alpha_{j,i}^*)_{j \neq i}$ be the solution for the above problem, that is, $\alpha_{j,i}^*$ is player j's stage game strategy when players $-i$ minimax player i.

Given $\underline{v}_i$, take $\underline{u}$ sufficiently small so that

$$\underline{v}_i > v_i(B) \equiv \max\left\{\max_{x:x_{i-1}=B} u_i(a(x)), v_i^*\right\} + 2\underline{u}. \tag{5.1}$$

Given $\underline{u}$, for sufficiently small ρ,

$$v_i(B) \geq \max\left\{\max_{x:x_{i-1}=B} w_i(x), \max_{x:x_{i-1}=B} u_i(a_i(x), \alpha_j(x)), v_i^*\right\} + \underline{u}. \tag{5.2}$$

Take $\bar{u}$ so that

$$\bar{u} \geq v_i(B) + \underline{u}, v_i(B) - v_i^* \tag{5.3}$$

5.1 Intuitive Explanation

Before proceeding to the proof, we offer an intuitive explanation. As in the two-player case, we have L main blocks, where L will be defined in Section 5.7. Player $i-1$ incentivizes player i by the reward function. Similarly to $\lambda_j(l)$ and $\hat{\lambda}_j(l)$ in the two-player case, player $i-1$ has $\lambda_{i-1}(l) \in \{G, B\}$ indicating whether player $i-1$ has observed an "erroneous score" and player i has $\lambda_{i-1}(l)(i) \in \{G, B\}$ indicating what is the optimal action for player i. With more than two players, player $i+1$ also has $\lambda_{i-1}(l)(i+1) \in \{G, B\}$, which is player $(i+1)$'s inference of $\lambda_{i-1}(l)$.

The game with more than two players is different from the one with two players in the following three aspects. First, as we mentioned in Section 1.5.5.3, the players need to coordinate on punishment.

Second, as seen in Section 1.5.6.2, player i informs the other players $-i$ of x_i in the coordination block. With two players, there is only one receiver of the message. On the other hand, with more than two players, there are more than one receivers of the message. If some players infer x_i is G while the others infer x_i is B, then the action that will be taken in the review rounds with a high probability may not be included in $\{a(x)\}_x$. Since we do not have any bound on the payoff of a player in such a situation, it might be of some player's interest to induce this.

In this chapter, since x_i is communicated by public cheap talk, this is not an issue. However, in Chapter 7, where the players communicate via actions, we need to make sure that no player can induce mis-coordination of x_i among the other players in order to increase her payoff.

Third, when player j sends a message to player n, there is another player i. In this chapter, player j sends the message by perfect or error-reporting noisy cheap talk. In each case, the distribution of player n's signal is exogenously given and there is no way for player i to manipulate the communication between players j and n. However, in Chapter 7, where player j sends the message via actions, we need to make sure that player i does not have an incentive to deviate in order to manipulate the signal distributions between players j and n and to "confuse" players j and n.

Let us go back to the first question: how players $-i$ coordinate on minimaxing player i after the histories where player i is likely to have deviated. To deal with this problem, we consider the following mechanism to coordinate on the punishment.

For each player i, there are two monitors, players $i-1$ and $i+1$. In other words, each player n monitors players $n-1$ and $n+1$.

After the lth review round, each player j constructs a variable $d_j(l+1) \in \{0, j-1, j+1\}$. $d_j(l+1) = 0$ implies that player j thinks that there was no deviator in players $j-1$ and $j+1$ in the lth review round. $d_j(l+1) = j-1$ implies that player j thinks that player $j-1$ deviated in the lth review round. $d_j(l+1) = j+1$ implies that player j thinks that player $j+1$ deviated. Intuitively speaking, since player j monitors player $j-1$, player j wants to know the realization of the score of player $j-2$ (the controller of player $j-1$). This is why $\lambda_{j-2}(l)(j)$ is also defined.

Player j sends the message $d_j(l+1)$ to each player $n \in -j$ via error-reporting noisy cheap talk.[1] Player n constructs the inference of $d_j(l+1)$, $d_j(l+1)(n) \in \{0, j-1, j+1\}$, from the private signal of the error-reporting noisy cheap talk.

Each player n minimaxes player i by $\alpha^*_{n,i}$ if and only if player n infers that the two monitors $i-1$ and $i+1$ think that player i has deviated:

$$d_{i-1}(l+1)(n) = d_{i+1}(l+1)(n) = i.$$

To incentivize the players to tell the truth about $d_j(l+1)$, whenever player j's message has an impact on the decision of whom to be minimaxed (we say "player j is pivotal"), we make player j indifferent between any action profile. Assumption 4 guarantees that player j cannot create a situation where player j is pivotal by deviation to increase her own payoff, as we will see in Lemma 33.

5.2 Almost Optimality

As seen in Section 1.8, we first show that player i's strategy is "almost optimal." We divide the reward function into two parts:

$$\pi_i(x_{i-1}, h_{i-1}^{T_P+1} : \delta) = \pi_i^{\text{main}}(x_{i-1}, h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}} : \delta) + \pi_i^{\text{report}}(x_{i-1}, h_{i-1}^{T_P+1}, h_{i-1}^{\text{rereport}} : \delta).$$

Contrary to the two-player case, we have $h_{i-1}^{\text{rereport}}$ in π_i^{main} and π_i^{report}. We will define h_{i-1}^{main}, $h_{i-1}^{T_P+1}$ and $h_{i-1}^{\text{rereport}}$ below.

With more than two players, player $i-1$ wants to use the information owned by players $-(i-1,i)$ to construct player $(i-1)$'s reward function on player i. Hence, as we will see in Section 5.4, after the report block where player i reports h_i^{main}, we have the "re-report block" where players $-(i-1,i)$ send their history to player $i-1$. This information is used

[1] Precisely, since $d_n(l+1)$ is ternary while the error-reporting noisy cheap talk can send a binary message, player n sends a sequence of binary messages. See Section 5.5.2.

only for π_i and does not affect the value of players $-(i-1,i)$. Therefore, we can assume that players $-(i-1,i)$ tell the truth. Further, since the information in the re-report block is used only for the reward (not for the action plan $\sigma_{i-1}(x_{i-1})$), it is sufficient for player $i-1$ to know the information by the end of the review phase.

Let h_{i-1}^{main}, $h_{i-1}^{T_P+1}$, and $h_{i-1}^{\text{rereport}}$, respectively, be the history of player $i-1$ in the main blocks, "in the coordination, main and report blocks," and in the re-report block, respectively.

We first construct $\sigma_i(x_i)$ and $\pi_i^{\text{main}}(x_{i-1}, h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}} : \delta)$ satisfying (1.21), (1.16) and (1.17) if we neglect the report block. After constructing such π_i^{main}, we construct the strategy in the report block such that $\pi_i = \pi_i^{\text{main}} + \pi_i^{\text{report}}$ satisfies (1.15), (1.16) and (1.17).

5.3 Preparations

Before proceeding to the equilibrium construction, we make the following preparations.

5.3.1 Functions and Statistics

We define functions and statistics useful for the equilibrium construction. First, since we assume the availability of the error-reporting noisy cheap talk, Lemma 11 holds.

Second, we define the point $\pi_i[\alpha(x)](y_{i-1})$, corresponding to point $\pi_i[\alpha(x)](y_j)$ in the two-player case. That is, $\pi_i[\alpha(x)](y_{i-1})$ cancels out the differences in the instantaneous utilities for different a_i's:

$$u_i(a_i, \alpha_{-i}(x)) + \mathbb{E}\left[\pi_i[\alpha(x)](y_{i-1}) \mid a_i, \alpha_{-i}(x)\right] \tag{5.4}$$

is independent of $a_i \in A_i$, as in (2.4).

Further, as in (2.5), we want to make sure that

$$u_i(a_i, \alpha_{-i}(x)) + \mathbb{E}\left[\pi_i[\alpha(x)](y_{i-1}) \mid a_i, \alpha_{-i}(x)\right] = w_i(x) \tag{5.5}$$

for all $a_i \in A_i$. From (1.19), this implies

$$\mathbb{E}\left[\pi_i[\alpha(x)](y_{i-1}) \mid \alpha(x)\right] = \begin{cases} \leq 0 & \text{if } x_{i-1} = G \\ \geq 0 & \text{if } x_{i-1} = B. \end{cases} \tag{5.6}$$

Since Assumption 3 implies that player $i-1$ can statistically infer player i's action, (5.4) and (5.5) can be satisfied simultaneously:

Lemma 26 If Assumption 3 is satisfied, then there exists $\bar{u} > 0$ such that, for each $i \in I$ and $\alpha(x) \in \Delta(A)$, there exists $\pi_i[\alpha(x)] : Y_{i-1} \to [-\bar{u}, \bar{u}]$ with (5.4) and (5.5).

Proof: The same as Lemma 12. ∎

Third, as for $\pi_i^{x_j}[\alpha_j](y_j)$, define

$$\pi_i^{x_{i-1}}[\alpha_{-i}](y_{i-1}) \begin{cases} \leq 0 \text{ for all } y_{i-1} & \text{if } x_{i-1} = G, \\ \geq 0 \text{ for all } y_{i-1} & \text{if } x_{i-1} = B \end{cases} \tag{5.7}$$

such that, for all $i \in I$,

$$u_i(a_i, \alpha_{-i}) + \mathbb{E}\left[\pi_i^{x_{i-1}}[\alpha_{-i}](y_{i-1}) \mid a_i, \alpha_{-i}\right] \tag{5.8}$$

is independent of $a_i \in A_i$ and $\alpha_{-i} \in \Delta(A_{-i})$ and included in $[-2\bar{u}, 2\bar{u}]$.

Fourth, as for $\pi_i^\delta(t, a_{j,t}, y_{j,t})$, we can construct the reward to cancel out discounting:

Lemma 27 If Assumption 3 is satisfied, then for each $i \in I$, there exists $\pi_i^\delta : \mathbb{N} \times \Delta(A_{-i}) \times Y_{i-1} \to [-\bar{u}, \bar{u}]$ such that, for all $a_{i,t} \in A_i$ and $\alpha_{-i,t}$, we have

$$\delta^{t-1} u_i(a_{i,t}, \alpha_{-i,t}) + \mathbb{E}\left[\pi_i^\delta(t, \alpha_{-i,t}, y_{i-1,t}) \mid a_{i,t}, \alpha_{-i,t}\right] = u_i(a_{i,t}, \alpha_{-i,t}) \text{ for all } t \in \{1, ..., T_P\}$$

and

$$\lim_{\delta \to 1} \frac{1-\delta}{1-\delta^{T_P}} \sum_{t=1}^{T_P} \sup_{\alpha_{-i,t}, y_{i-1,t}} \left|\pi_i^\delta(t, \alpha_{-i,t}, y_{i-1,t})\right| = 0 \tag{5.9}$$

for all $T_P = \Theta(T)$ with $T = (1-\delta)^{-\frac{1}{2}}$.

Proof: The same as Lemma 13. ∎

As we will see in Section 5.5.3, we add

$$\sum_{t=1}^{T_P} \pi_i^{\delta}\left(t, \alpha_{-i,t}, y_{i-1,t}\right) \tag{5.10}$$

to π_i^{main} so that we can ignore discounting.

Fifth, as $\gamma_i^a\left(y_i\right)$, player i constructs a statistics to infer whether player $i-1$ has deviated.

Lemma 28 If Assumption 3 is satisfied, then there exist $q_2 > q_1$ such that, for all $i \in I$, $a_i\left(x\right) \in A_i$, $a_{i-1}\left(x\right) \in A_{i-1}$ and $\alpha_{-(i-1,i)}\left(x\right) \in \Delta\left(A_{-(i-1,i)}\right)$, there exists a function $\gamma_i^x : Y_i \to (0,1)$ such that player i can statistically infer whether player $i-1$ takes $a_{i-1}(x)$ or not:

$$\mathbb{E}\left[\gamma_i^x\left(y_i\right) \mid a_i\left(x\right), \tilde{a}_{i-1}, \alpha_{-(i-1,i)}\left(x\right)\right] = \begin{cases} q_2 & \text{if } \tilde{a}_{i-1} = a_{i-1}\left(x\right), \\ q_1 & \text{otherwise.} \end{cases}$$

Proof: The same as Lemma 22. ∎

Sixth, as in the two-player case, let $q_i\left(\alpha\right)$ be the vector of player i's signal distribution given α:

$$q_i\left(\alpha\right) \equiv \left(q_i\left(y_i \mid \alpha\right)\right)_{y_i \in Y_i}.$$

In particular, with $\alpha = (a_i(x), \alpha_{-i}(x))$, we define

$$q_i\left(x\right) \equiv q_i\left(a_i(x), \alpha_{-i}(x)\right).$$

For each pair of players i and $j \neq i$, let $\boldsymbol{Q}_i^j(x)$ be the affine hull of player i's signal observations with respect to player j's actions given that player i takes $a_i\left(x\right)$ and players

$-(i,j)$ take $\alpha_{-(i,j)}(x)$:

$$\boldsymbol{Q}_i^j(x) \equiv \text{aff}\left(\left\{q_i(a_i(x), a_j, \alpha_{-(i,j)}(x))\right\}_{a_j \in A_j}\right) \cap \mathbb{R}_+^{|Y_i|}.$$

As in the two-player case, we consider the following representation:

$$\boldsymbol{Q}_i^j(x) = \left\{\mathbf{y}_i \in \mathbb{R}_+^{|Y_i|} : Q_i^j(x)\mathbf{y}_i = \mathbf{q}_i(x)\right\}.$$

As in Lemma 14, we can assume each element of $Q_i(x)$ and $\mathbf{q}_i(x)$ is included in $(0,1)$:

Lemma 29 For each $i, j \in I$, we can take $Q_i^j(x)$ such that all the elements of $Q_i(x)$ and $\mathbf{q}_i(x)$ are in $(0,1)$.

Proof: The same as Lemma 14. ∎

With Assumption 4, we make sure that $\boldsymbol{Q}_i^j(x)$ and $\boldsymbol{Q}_i^n(x)$ with $j \neq n$ does not intersect except for $q_i(x)$:

Lemma 30 If Assumption 4 is satisfied, then for all $i \in I$, we have

$$\boldsymbol{Q}_i^j(x) \cap \boldsymbol{Q}_i^n(x) = q_i(x) \text{ for all } j \neq n.$$

Proof: In this proof, we concentrate on the simplex on Y_i and omit the last component of $q_i(\alpha)$ since $\mathbf{1}q_i(\alpha) = 1$ for all $\alpha \in \Delta(A)$.

Since

$$\boldsymbol{Q}_i^j(x) = \left\{\mathbf{y}_i \in \mathbb{R}_+^{|Y_i|} : \begin{array}{c} \exists \{t_j(a_j)\}_{a_j \in A_j} \text{ with} \\ \mathbf{y}_i = q_i(x) + \sum_{a_j \in A_j} t_j(a_j) q_i(a_i(x), a_j, \alpha_{-(i,j)}(x)) \end{array}\right\},$$

it suffices to show that, for each j, n with $j \neq n$,

$$\sum_{a_j \in A_j} t_j(a_j) q_i(a_i(x), a_j, \alpha_{-(i,j)}(x)) = \sum_{a_n \in A_n} t_n(a_n) q_i(a_i(x), a_n, \alpha_{-(i,j)}(x))$$

implies

$$t_j(a_j) = t_n(a_n) = 0$$

for all a_j, a_n, which is guaranteed by Assumption 4. ∎

5.3.2 Perfect Monitoring

As in the two-player game, we consider a one-shot game with perfect monitoring parameterized with $l \in \mathbb{N}$. In the game with parameter $l \in \{1, ..., L-1\}$, players $-i$ takes $\alpha_{-i}(x)$. Depending on player i's action, $d_{i+1}(l+1) \in \{0, i\}$ is determined. If player i takes $a_i(x)$, then $d_{i+1}(l+1) = 0$ with probability one. If player i takes $a_i \neq a_i(x)$, then $d_{i+1}(l+1) = i$ with probability $p_{i+1}^{l+1}(x)$ and $d_{i+1}(l+1) = 0$ with the remaining probability $1 - p_{i+1}^{l+1}(x)$. The payoff of player i is determined as

$$V_i^l = \max_{a_i} \frac{1}{L-l+1} \tilde{u}_i(a_i) + \frac{L-l}{L-l+1} \mathbb{E}\left[W_i^{l+1}(d(l+1)) \mid a_i\right]$$

with

$$\begin{aligned}
\tilde{u}_i(a_i) &= \begin{cases} u_i\left(a_i, \alpha_{-i}(x)\right) + \underline{u} & \text{if } a_i = a_i^{BR}(x), \\ u_i\left(a_i, \alpha_{-i}(x)\right) & \text{if } a_i \neq a_i^{BR}(x), \end{cases} \\
W_i^{l+1}(G) &= \frac{(L-l-1)\, v_i(B) + \bar{u}}{L-l} + \underline{u}, \\
W_i^{l+1}(B) &= v_i^*.
\end{aligned}$$

Here, $a_i^{BR}(x) \in BR_i\left(\alpha_{-i}\left(x\right)\right)$. If $BR_i\left(\alpha_{-i}\left(x\right)\right)$ has multiple elements, then pick one arbitrarily.

As in the two-player case, we can show the following lemma:

Lemma 31 For any $L \geq 2$, $q_2 > q_1$ and $\underline{u}$, there exist $\bar{q} > 0$ and $\bar{\rho} > 0$ such that, for any $q < \bar{q}$ and $\rho < \bar{\rho}$, for all $i \in I$, there exist $\{\underline{p}_{i+1}^{l+1}(x), \bar{p}_{i+1}^{l+1}(x)\}_{l=1}^{L-1}$ with

$$\begin{aligned} \bar{p}_{i+1}^{l+1}(x) &\in [0,1] \\ \underline{p}_{i+1}^{l+1}(x) &= \frac{q_2 - q - (1 - 2(|A_{i+1}| - 1)\rho) q_1}{q_2 - q_1} \bar{p}_{i+1}^{l+1} \end{aligned} \tag{5.11}$$

for all $l = 1, ..., L-1$ such that, for $\{p_{i+1}^{l+1}(x)\}_{l=1}^{L-1}$,

1. if $p_{i+1}^{l+1}(x) \leq \bar{p}_{i+1}^{l+1}(x)$ for all l, then it is uniquely optimal for player i to take $a_i^{BR}(x)$; and
2. if $p_{i+1}^{l+1}(x) \in \left[\underline{p}_{i+1}^{l+1}(x), \bar{p}_{i+1}^{l+1}(x)\right]$ for all l, then

$$V_i^l \leq W_i^l(G) = \frac{(L-l)\, v_i(B) + \bar{u}}{L - l + 1} + \underline{u}.$$

Proof: The same as Lemma 23. ∎

5.4 Structure of the Phase

In this section, we explain the structure of the finitely repeated game. As in the two-player game, we have the coordination block at the beginning, where each player takes turns to send the cheap talk message $x_i \in \{G, B\}$: first, player 1 sends x_1, second, player 2 sends x_2, and so on until player N sends x_N. Note that x will become common knowledge for the rest of the game.

After the coordination block, we have L main blocks. The first $(L-1)$ blocks are further divided into $1 + 2N + N(N-1)$ rounds. That is, for $l \in \{1, ..., L-1\}$, the lth main block consists of the following rounds: first, the players play a T-period review round.

After that, as indicated in Section 5.1, each player $i-1$ sends $\lambda_{i-1}(l+1)$ to player $n \in \{i, i+1\}$ by the error-reporting noisy cheap talk between $i-1$ and n. The players take turns: player 1 sends $\lambda_1(l+1)$ to player 2, player 1 sends $\lambda_1(l+1)$ to player 3, player 2 sends $\lambda_2(l+1)$ to player 3, player 2 sends $\lambda_2(l+1)$ to player 4, and so on until player N sends $\lambda_N(l+1)$ to player 2. We call the instance where player $i-1$ sends $\lambda_{i-1}(l+1)$ to player $n \in \{i, i+1\}$ the "supplemental round for $\lambda_{i-1}(l+1)$ between $i-1$ and n."

After that, each player j sends $d_j(l+1)$ to each player $n \in -j$ by the error-reporting noisy cheap talk between j and n. The players take turns: player 1 sends $d_1(l+1)$ to player 2, player 1 sends $d_1(l+1)$ to player 3, and so on until player 1 sends $d_1(l+1)$ to player N. Then, player 2 sends $d_2(l+1)$ to player 1, player 2 sends $d_2(l+1)$ to player 3, and so on until player 2 sends $d_2(l+1)$ to player N. This step continues until player N sends $d_N(l+1)$ to player $N-1$. We call the instance where player j sends $d_j(l+1)$ to player n the "supplemental round for $d_j(l+1)$ between j and n."

The last Lth main block has only the T-period review round.

Let $T(l)$ be the set of T periods in the lth review round. As in the two-player case, player i randomly picks $t_i(l)$ from $T(l)$. The action plan is determined by the history in periods in $T_i(l) = T(l) \setminus \{t_i(l)\}$.

After the last main block, there is the report block, where each player i reports the whole history h_i^{main}.

Finally, after the report block, there is the re-report block, where each player i reports the whole history h_i^{main} again. This time, player i's message is used only for the reward $\pi_j^{\text{main}}(x_{j-1}, h_{j-1}^{\text{main}}, h_{j-1}^{\text{rereport}} : \delta)$ with $j \neq i$. That is, player i's message does not affect player i.

5.5 Equilibrium Strategy

In this section, we define $\sigma_i(x_i)$ and $\pi_i^{\text{main}}(x_{i-1}, h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}} : \delta)$. In Section 5.5.1, we define the state variables that will be used to define the action plans and rewards. Given the states, Section 5.5.2 defines player i's action plan $\sigma_i(x_i)$ and Section 5.5.3 defines player $(i-1)$'s reward function $\pi_i^{\text{main}}(x_{i-1}, h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}} : \delta)$ on player i. Finally, Section 5.5.4 determines the transition of the states defined in Section 5.5.1.

5.5.1 States x_i, $\lambda_{i-1}(l+1)$, $\lambda_{i-1}(l+1)(i)$, $\lambda_{i-2}(l+1)(i)$, $d_i(l+1)$, $d_j(l+1)(i)$, $c_i(l+1)$ and $\theta_{i-1}(l)$

The intuitive meaning of $x_i \in \{G, B\}$ and $\lambda_{i-1}(l+1) \in \{G, B\}$ is the same as in the two-player case with j replaced with $i-1$. Then, $\lambda_{i-1}(l+1)(n) \in \{G, B\}$ with $n \in \{i, i+1\}$ is player n's inference of $\lambda_{i-1}(l+1)$.

As seen in Section 5.1, $d_i(l+1) \in \{0, i-1, i+1\}$ indicates what player i thinks about a deviation by players $i-1$ and $i+1$.

Player $j \neq i$ sends $d_j(l+1)$ via error-reporting noisy cheap talk to player i in the supplemental round for $d_j(l+1)$ between j and i. Let $d_j(l+1)(i)$ be player i's inference of the message, which will be determined in Section 5.5.4.4.

Player i "confirms" player n's deviation if and only if player i infers that the two monitors $n-1$ and $n+1$ think that player n has deviated:

$$d_{n-1}(l+1)(i) = d_{n+1}(l+1)(i) = n.$$

In such a case, $n \in c_i(l+1)$. That is, $c_i(l+1) \subset \{0\} \cup I$ is the set of players whose deviation has been confirmed ($c_i(l+1) = \{0\}$ means that no player's deviation has been confirmed). If there is a unique player in $c_i(l+1)$ ($c_i(l+1) = \{n\}$ with $n \in I$), then player i minimaxes player n by taking $\alpha_{i,n}^*$.

The intuitive meaning of $\theta_{i-1}(l) \in \{G, B\}$ is the same as in the two-player case.

5.5.2 Player i's Action Plan $\sigma_i(x_i)$

In the coordination block, each player sends x_i truthfully. Then, the state profile x becomes common knowledge.

In the lth review round, player i's strategy depends on $\lambda_{i-1}(l)(i)$ and $c_i(l)$.

1. if $\lambda_{i-1}(l)(i) = G$, then

   (a) if there is no $n \in I$ with $c_i(l) = \{n\}$, then, given some fixed set $A_i(x)$ of player i's mixed action plans (see Lemma 35 for the definition of $A_i(x)$),

      i. $\alpha_i(l) = \alpha_i(x)$ with probability $1 - \eta$;

      ii. $\alpha_i(l) = \alpha^*_{i,n}$ for some player $n \in -i$ with probability $\frac{\eta}{2(N-1)}$; and

      iii. for each $\alpha_i \in A_i(x)$, $\alpha_i(l) = \alpha_i$ with probability $\frac{1}{2|A_i(x)|}\eta$;

   (b) if there is $n \in I$ with $c_i(l) = \{n\}$,

      i. $\alpha_i(l) = \alpha^*_{n,i}$ with probability $1 - \eta$;

      ii. $\alpha_i(l) = \alpha_i(x)$ with probability $\frac{\eta}{3}$;

      iii. $\alpha_i(l) = \alpha^*_{i,n'}$ for some player $n' \in -(i, n)$ with probability $\frac{\eta}{3(N-3)}$; and

      iv. for each $\alpha_i \in A_i(x)$, $\alpha_i(l) = \alpha_i$ with probability $\frac{1}{3|A_i(x)|}\eta$;

   Note that the support of $\alpha_i(l)$ is constant regardless of $c_i(l)$; and

2. if $\lambda_{i-1}(l)(i) = B$, then

   (a) if $c_i(l) = \{0\}$, then $\alpha_i(l) = a^{BR}_i(x)$; and

   (b) otherwise, $BR_i(\alpha^*_{-i})$.

In the supplemental round for $\lambda_i(l+1)$ between i and $n \in \{i+1, i+2\}$, player i sends $\lambda_i(l+1)$ truthfully via error-reporting noisy cheap talk to player n.

In the supplemental round for $d_i(l+1)$ between i and $n \in -i$, player i sends $d_i(l+1)$ truthfully via error-reporting noisy cheap talk to player n.

Since $d_i(l+1)$ is ternary while the error-reporting noisy cheap talk can send a binary message, we attach a sequence of binary messages to each $d_i(l+1)$. Specifically, given $d_i(l+1) \in \{0, i-1, i+1\}$, player i define a sequence $d_i(l+1)\{1\}, d_i(l+1)\{2\} \in \{G, B\}^2$:

if $d_i(l+1)=0$, then $d_i(l+1)\{1\}=G$ and $d_i(l+1)\{2\}=B$ with probability $\frac{1}{2}$ and $d_i(l+1)\{1\}=B$ and $d_i(l+1)\{2\}=G$ with probability $\frac{1}{2}$. If $d_i(l+1)=i-1$, then $d_i(l+1)\{1\}=G$ and $d_i(l+1)\{2\}=G$. If $d_i(l+1)=i+1$, then $d_i(l+1)\{1\}=B$ and $d_i(l+1)\{2\}=B$.

Player i with $d_i(l+1)$ sends $d_i(l+1)\{1\}$ and $d_i(l+1)\{2\}$ truthfully via error-reporting noisy cheap talk. We define

$$f[n](d_i(l+1)) = \begin{cases} 0 & \text{if } f[n](d_i(l+1)\{1\}) \neq f[n](d_i(l+1)\{2\}), \\ i-1 & \text{if } f[n](d_i(l+1)\{1\}) = f[n](d_i(l+1)\{2\}) = G, \\ i+1 & \text{if } f[n](d_i(l+1)\{1\}) = f[n](d_i(l+1)\{2\}) = B, \end{cases}$$

$$\begin{aligned} & g[n-1](d_i(l+1)) \\ = & \begin{cases} E & \text{if } g[n-1](d_i(l+1)\{1\}) = E \text{ or } g[n-1](d_i(l+1)\{2\}) = E, \\ d_i(l+1) & \text{otherwise,} \end{cases} \end{aligned}$$

$$g_2[n-1](d_i(l+1)) = (g_2[n-1](d_i(l+1)\{1\}), g_2[n-1](d_i(l+1)\{2\})),$$

and

$$f_2[i-1](d_i(l+1)) = (f_2[i-1](d_i(l+1)\{1\}), f_2[i-1](d_i(l+1)\{2\})).$$

Lemma 11 holds as if player i sent $d_i(l+1)$ via error-reporting noisy cheap talk since

- the message transmits correctly with probability $1-\exp(-\Theta(T^{\frac{1}{2}}))$;
- given $d_i(l+1)\{1\}, d_i(l+1)\{2\}$, player n puts a conditional belief $1-\exp(-\Theta(T^{\frac{1}{2}}))$ on the events that $f[n](d_i(l+1)) = d_i(l+1)$ or $g[n-1](d_i(l+1)) = E$; and
- given $d_i(l+1)\{1\}, d_i(l+1)\{2\}$, any signal sequences can happen with probability no less than $\exp(-\Theta(T^{\frac{1}{2}}))$.

5.5.3 Reward Function

In this subsection, we explain player $(i-1)$'s reward function on player i, $\pi_i^{\text{main}}(x_{i-1}, h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}} : \delta)$.

Score As in the two-player case, player $i-1$ picks $t_{i-1}(l)$ randomly from $T(l)$ and the score is the summation of the points $\pi_i[\alpha(x)](y_{i-1})$ over $T_{i-1}(l) \equiv T(l) \setminus \{t_{i-1}(l)\}$:

$$X_{i-1}(l) = \sum_{t \in T_{i-1}(l)} \pi_i[\alpha(x)](y_{i-1,t})$$

As we will see, $X_{i-1}(l)$ is used only if players $-i$ take $\alpha_{-i}(l) = \alpha_{-i}(x)$.

Reward Function As in the two-player case, the reward $\pi_i^{\text{main}}(x_{i-1}, h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}} : \delta)$ is written as

$$\begin{aligned}
&\pi_i^{\text{main}}\left(x_{i-1}, h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}} : \delta\right) = \sum_{l=1}^{L} \sum_{t \in T(l)} \pi_i^{\delta}\left(t, \alpha_{-i,t}, y_{i-1,t}\right) \\
&+ \begin{cases} -2\bar{u}T + \sum_{l=1}^{L} \pi_i^{\text{main}}(x_{i-1}, h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}}, l) & \text{if } x_{i-1} = G, \\ 2\bar{u}T + \sum_{l=1}^{L} \pi_i^{\text{main}}(x_{i-1}, h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}}, l) & \text{if } x_{i-1} = B. \end{cases}
\end{aligned} \tag{5.12}$$

Note that we add (5.10) to ignore discounting. Here, player $i-1$ uses information owned by players $-(i-1, i)$ such as $\alpha_{-i,t}$. In general, if player $i-1$ uses information owned by players $-(i-1, i)$, the information is sent by players $-(i-1, i)$ to player $i-1$ in the re-report block. As we will see in Section 5.5.4, since the information owned by players $-(i-1, i)$ does not affect player $(i-1)$'s action plan, player $(i-1)$'s equilibrium strategy is well defined.

Reward Function for the lth Review Round Next we define $\pi_i^{\text{main}}(x_{i-1}, h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}}, l)$ for each $l = 1, ..., L$. There are following two cases: in the lth review round,

1. if $\theta_{i-1}(l) = B$, then player $i-1$ makes player i indifferent between any action profile by

$$\pi_i^{\text{main}}(x_{i-1}, h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}}, l) = \sum_{t \in T(l)} \pi_i^{x_{i-1}}[\alpha_{-i}(l)](y_{i-1,t}). \tag{5.13}$$

(5.7) guarantees that (1.17) is not an issue after $\theta_{i-1}(l) = B$; and

2. otherwise, that is, if $\theta_{i-1}(l) = G$, then player $(i-1)$'s reward on player i is based on x, $\lambda_{i-1}(l)$ and $c_{i-1}(l)$. The formal description is given by

$$\begin{aligned} &\pi_i^{\text{main}}\left(x_{i-1}, h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}}, l\right) \qquad (5.14) \\ = &\begin{cases} \bar{\pi}_i(x, G, c_{i-1}(l), l) + X_{i-1}(l) + \pi_i^{x_{i-1}}[\alpha_{-i}(l)](y_{i-1,t_{i-1}(l)}) & \text{if } \lambda_{i-1}(l) = G, \\ \bar{\pi}_i(x, B, c_{i-1}(l), l) & \text{if } \lambda_{i-1}(l) = B, \end{cases} \end{aligned}$$

Here, $\bar{\pi}_i(x, \lambda_{i-1}(l), c_{i-1}(l), l)$ will be determined later so that (1.21), (1.16) and (1.17) are satisfied.

In addition, as in the two-player case, player $i-1$ with $\theta_{i-1}(l) = G$, $\lambda_{i-1}(l) = B$ and $d_{i-1}(l) = G$ gives a small bonus on $\underline{u}$ for taking $a_i^{BR}(x)$. By Assumption 3, there exists $\pi_i^{\underline{u}}[\alpha_{-i}(l)](y_{i-1})$ so that

$$\mathbb{E}\left[\pi_i^{\underline{u}}[\alpha_{-i}(l)](y_{i-1}) \mid a_i, \alpha_{-i}(l)\right] = \begin{cases} \underline{u} & \text{if } a_i = a_i^{BR}(x), \\ 0 & \text{if } a_i \neq a_i^{BR}(x). \end{cases}$$

Player $i-1$ with $\theta_{i-1}(l) = G$, $\lambda_{i-1}(l) = B$ and $d_{i-1}(l) = G$ adds

$$\sum_{t \in T(l)} \pi_i^{\underline{u}}[\alpha_{-i}(l)](y_{i-1,t}) \tag{5.15}$$

to $\pi_i^{\text{main}}\left(x_{i-1}, h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}}, l\right)$. For sufficiently small $\underline{u}$, $\pi_i^{\underline{u}}[\alpha_{-i}(l)](y_{i-1})$ is sufficiently small for all y_{i-1}. Hence, when we consider (1.17), we ignore (5.15).

5.5.4 Transition of the States

In this subsection, we explain the transition of player i's states. Since x_i is fixed in the phase, we consider the following six:

5.5.4.1 Transition of $\lambda_{i-1}(l+1) \in \{G, B\}$

The transition of $\lambda_{i-1}(l+1) \in \{G, B\}$ is the same as in Section 4.5.4.1 with j replaces with $i-1$, except that $\lambda_{i-1}(l+1) = G$ even if the score is erroneous if player $(i-1)$'s signal frequency is far away from the affine hull of player $(i-1)$'s signal distributions with respect to player i's action. As we will see, in such a case, $\theta_{i-1}(l+1) = B$ and $\lambda_{i-1}(l+1)$ is irrelevant for player i. In addition, this does not affect player i's incentive since player i cannot control whether player $(i-1)$'s signal frequency is far away from the affine hull of player $(i-1)$'s signal distributions with respect to player i's action.

Now, we define the transition of $\lambda_{i-1}(l)$: the initial condition is $\lambda_{i-1}(1) = G$. If $\lambda_{i-1}(l) = B$, then $\lambda_{i-1}(l+1) = B$. If $\lambda_{i-1}(l) = G$, then $\lambda_{i-1}(l+1) = G$ if at least one of the following three conditions is satisfied:

1. the score is regular:
$$X_{i-1}(l) \begin{cases} \leq \frac{\bar{u}}{L}T & \text{if } x_j = G, \\ \geq -\frac{\bar{u}}{L}T & \text{if } x_j = B; \end{cases}$$

2. some player's deviation has been confirmed: $c_{i-1}(l) \neq \{0\}$; or

3. player $i-1$ takes $\alpha_{i-1}(x)$, player $(i-1)$'s action frequency is close to $\alpha_{i-1}(x)$, and player $(i-1)$'s signal frequency while player $i-1$ is taking $a_{i-1}(x)$ is not close to $\boldsymbol{Q}^i_{i-1}(x)$. That is, the following three conditions are satisfied:

    (a) player $i-1$ takes $\alpha_{i-1}(x)$: $\alpha_{i-1}(l) = \alpha_{i-1}(x)$;

(b) player $(i-1)$'s action frequency is close to $\alpha_{i-1}(x)$:

$$\left\| \frac{1}{T-1} \sum_{t \in T_{i-1}(l)} \mathbf{1}_{a_{i-1,t}} - \alpha_{i-1}(x) \right\| < \varepsilon; \tag{5.16}$$

and

(c) player $(i-1)$'s signal frequency while player $i-1$ is taking $a_{i-1}(x)$ is not close to $\boldsymbol{Q}_{i-1}^{i}(x)$.

Otherwise, that is, if the score is erroneous

$$X_{i-1}(l) \begin{cases} > \frac{\bar{u}}{L} T & \text{if } x_j = G, \\ < -\frac{\bar{u}}{L} T & \text{if } x_j = B, \end{cases} \tag{5.17}$$

$c_{i-1}(l) = \{0\}$, and "not 3-(a), not 3-(b) or not 3-(c)," then $\lambda_{i-1}(l+1) = B$.

We are left to formally define Condition 3-(c). For notational convenience, we specify the conditions such that, if and only if these conditions are satisfied, Condition 3-(c) is not satisfied.

First, player $i-1$ creates $\mathbf{1}_{Q_{i-1}^{i}(x)}$ from $Q_{i-1}^{i}(x)\mathbf{1}_{y_{i-1}}$ as player j creates $\mathbf{1}_{Q_j(x)}$ from $Q_j(x)\mathbf{1}_{y_j}$ in the two-player case.

Condition 3-(c) is not satisfied if and only if

$$\left\| Q_{i-1}^{i}(x) \left(\frac{1}{|T_{i-1}(l,x)|} \sum_{t \in T_{i-1}(l,x)} \mathbf{1}_{y_{i-1,t}} \right) - \frac{1}{|T_{i-1}(l,x)|} \sum_{t \in T_{i-1}(l,x)} \mathbf{1}_{Q_{i-1}^{i}(x)} \right\| < \frac{\varepsilon}{K_1}, \tag{5.18}$$

and

$$\left\| \frac{1}{|T_{i-1}(l,x)|} \sum_{t \in T_{i-1}(l,x)} \mathbf{1}_{Q_{i-1}^{i}(x)} - \mathbf{q}_{i-1}^{i}(x) \right\| < \frac{\varepsilon}{K_1}. \tag{5.19}$$

Here, as for (2.22) in the two-player case, we adjust the probability of (5.18) so that the probability of (5.18) is independent of $a_{i-1,t}, y_{i-1,t}$ and so of players $-(i-1)$'s strategy. From now on, when we say (5.18) is satisfied, it takes this adjustment into account.

As in the two-player case, K_1 is sufficiently large so that (5.18) and (5.19) imply

$$\left\| \boldsymbol{Q}_i^j(x) - \frac{1}{|T_i(l,x)|} \sum_{t \in T_i(l,x)} \mathbf{1}_{y_{i,t}} \right\| < \varepsilon. \tag{5.20}$$

5.5.4.2 Transition of $\lambda_{i-1}(l+1)(i) \in \{G, B\}$ and $\lambda_{i+1}(l+1)(i) \in \{G, B\}$

The transition of $\lambda_{i-1}(l+1)(i) \in \{G, B\}$ is the same as in the two-player case with j replaced with $i-1$, except that if $c_i(l) \neq \{0\}$, then $\lambda_{i-1}(l+1)(i) = \lambda_{i-1}(l)(i)$. As will be seen in Section 5.6.2, player i with $c_i(l) \neq \{0\}$ and $\lambda_{i-1}(l)(i) = G$ believes that player i is indifferent between any action profile. This implies that player i does not need to infer $\lambda_{i-1}(l+1)$ seriously.

Specifically, $\lambda_{i-1}(1)(i) = G$. If $\lambda_{i-1}(l)(i) = B$, then $\lambda_{i-1}(l+1)(i) = B$. If $\lambda_{i-1}(l)(i) = G$, then $\lambda_{i-1}(l+1)(i) \in \{G, B\}$ is defined as follows.

Suppose player i's history in the lth review round satisfies the following two conditions, then player i disregards the message and

$$\lambda_{i-1}(l+1)(i) = G:$$

1. $c_i(l) \neq \{0\}$, that is, player i believes that some player's deviation has been confirmed; or

2. $c_i(l) = \{0\}$ and the following three conditions are satisfied:

   (a) player i takes $\alpha_i(l) = \alpha_i(x)$;

   (b) the actual frequency of player i's actions in the lth review round is close to $\alpha_i(x)$:

   $$\left\| \frac{1}{T-1} \sum_{t \in T_i(l)} \mathbf{1}_{a_{i,t}} - \alpha_i(x) \right\| < \varepsilon; \tag{5.21}$$

   and

(c) there exists some player $j \in -i$ such that player i's signal frequency while player i is taking $a_i(x)$ is close to $\boldsymbol{Q}_i^j(x)$.

Otherwise, player i obeys the message:

$$\lambda_{i-1}(l+1)(i) = f[i](\lambda_{i-1}(l+1)). \tag{5.22}$$

We are left to formally define 2-(c). As for (2.22) and (2.23), player i constructs the following random variables. First, player i creates $\mathbf{1}_{Q_i^j(x)}$ from $Q_i^j(x)\mathbf{1}_{y_i}$ as player i creates $\mathbf{1}_{Q_i(x)}$ from $Q_i(x)\mathbf{1}_{y_i}$.

Condition 2-(c) is satisfied if and only if there exists $j \in -i$ such that

$$\left\| Q_i^j(x) \left(\frac{1}{|T_i(l,x)|} \sum_{t \in T_i(l,x)} \mathbf{1}_{y_{i,t}} \right) - \frac{1}{|T_i(l,x)|} \sum_{t \in T_i(l,x)} \mathbf{1}_{Q_i^j(x)} \right\| < \frac{\varepsilon}{K_1}, \tag{5.23}$$

and

$$\left\| \frac{1}{|T_i(l,x)|} \sum_{t \in T_i(l,x)} \mathbf{1}_{Q_i^j(x)} - \mathbf{q}_i^j(x) \right\| < \frac{\varepsilon}{K_1}. \tag{5.24}$$

Here, as for (2.22) in the two-player case, we adjust the probability of (5.23) so that the probability of (5.23) is independent of $\{a_{i,t}, y_{i,t}\}_{t \in T(l)}$ and so of players $(-i)$'s strategy. From now on, when we say (5.23) is satisfied, it takes this adjustment into account.$\in$

As in the two-player case, K_1 is sufficiently large so that (5.23) and (5.24) implies

$$\left\| \boldsymbol{Q}_i^j(x) - \frac{1}{|T_i(l,x)|} \sum_{t \in T_i(l,x)} \mathbf{1}_{y_{i,t}} \right\| < \varepsilon. \tag{5.25}$$

For $\lambda_{i+1}(l+1)(i)$, player i always obeys the message

$$\lambda_{i+1}(l+1)(i) = f[i](\lambda_{i+1}(l+1)). \tag{5.26}$$

5.5.4.3 Transition of $d_i(l+1)$

We define the transition of $d_i(l+1) \in \{0, i-1, i+1\}$: $d_i(l+1) = 0$ implies that player i believes that both player $i-1$ and $i+1$ has taken $\alpha_{i-1}(x)$ and $\alpha_{i+1}(x)$ respectively in the lth review round; $d_i(l+1) = i-1$ implies that player i believes that player $i-1$ has taken $\alpha_{i-1}(l) \neq \alpha_{i-1}(x)$; $d_i(l+1) = i+1$ implies that player i believes that player $i+1$ has taken $\alpha_{i+1}(l) \neq \alpha_{i+1}(x)$.

Define $d_i(1) = 0$. For $l = 1, ..., L-1$, $d_i(l+1)$ is determined as follows:

1. if $d_i(l) \neq 0$ or $c_i(l) \neq \{0\}$, then $d_i(l+1) = d_i(l)$; and

2. if $d_i(l) = 0$ and $c_i(l) = \{0\}$,

   (a) if $\lambda_i(l) = G$ (player i has not observed an erroneous score), then player i monitors player $i-1$:

   $$d_i(l+1) = i-1$$

   if the following five conditions are satisfied:

   i. $\lambda_{i-2}(l)(i) = B$ (player i believes that player $i-2$ (the controller of player $i-1$) has observed an erroneous score);

   ii. $x_{i-2} = B$ (player $i-2$ takes the harsh strategy);

   iii. player i takes $\alpha_i(l) = \alpha_i(x)$;

   iv. (5.21), (5.23) and (5.24) are satisfied for $j = i-1$; and

   v. $\frac{1}{|T_i(l,x)|} \sum_{t \in T_i(l,x)} \gamma_i^x$ is low.

   Otherwise,

   $$d_i(l+1) = 0;$$

   and

(b) if $\lambda_i(l) = B$ (player i has not observed an erroneous score), then

$$\begin{cases} d_i(l+1) = 0 & \text{if } x_i = G, \\ d_i(l+1) = i+1 & \text{if } x_i = B. \end{cases}$$

Note that if $x_i = B$ and $\lambda_i(l) = B$, then player i always thinks player $i+1$ has deviated.

We are left to specify Condition 2-(a)-v. As player i constructs $\Gamma_i^{a(x)} \in \{0,1\}$ from $\gamma_i^{a(x)}(y_i)$ in the two-player case, player i constructs Γ_i^x from $\gamma_i^x(y_i)$.

As in the case with two players, Condition 2-(a)-v is satisfied if and only if

- $\frac{1}{|T_i(l,x)|}\sum_{t\in T_i(l,x)}\gamma_i^x(y_{i,t})$ and $\frac{1}{|T_i(l,x)|}\sum_{t\in T_i(l,x)}\Gamma_{i,t}^x$ are closed:

$$\left\| \frac{1}{|T_i(l,x)|}\sum_{t\in T_i(l,x)}\gamma_i^x(y_{i,t}) - \frac{1}{|T_i(l,x)|}\sum_{t\in T_i(l,x)}\Gamma_{i,t}^x \right\| < \frac{q}{3}; \tag{5.27}$$

  and

- player i draws a random variable from the uniform $[0,1]$ and the realization of this random variable is no less than

$$p_i^{l+1}(x)\min\left\{1, \frac{\left\{q_2T - qT - \sum_{t\in T_i(l,x)}\Gamma_{i,t}^x\right\}_+}{q_2T - q_1T}\right\}. \tag{5.28}$$

As in the two-player case, we adjust the probability of (5.27) so that the probability of (5.27) is independent of $\{a_{i,t}, y_{i,t}\}_{t\in T(l)}$. When we say (5.27) is satisfied, we take this adjustment into account.

Notice that, if $d_i(l+1) = i-1$, then

$$\frac{1}{|T_i(l,x)|}\sum_{t\in T_i(l,x)}\gamma_i^x(y_{i,t}) \le q_2 - \frac{q}{2} \tag{5.29}$$

for sufficiently small ρ, as for (4.15).

We postpone the intuitive explanation of this transition until Section 5.5.4.5.

5.5.4.4 Transition of $d_j(l+1)(i)$

If $j = i$, then player i knows $d_j(l+1)$. Hence, $d_j(l+1)(i) = d_j(l+1)$.

For each $j \neq i$, player i constructs $d_j(l+1)(i)$ by

$$d_j(l+1)(i) = f[i](d_j(l+1)) \tag{5.30}$$

using the signals that arrive when player j sends the message about $d_j(l+1)$ via error-reporting noisy cheap talk between players j and i.

5.5.4.5 Transition of $c_i(l+1) \subset \{0\} \cup I$

As seen in Section 5.5.2, $c_i(l+1) \ni n$ if and only if player n's deviation is "confirmed." Formally, $c_i(1) = \{0\}$ and for $l \geq 1$, if there exists $n \in I$ such that $c_i(l) \ni n$ (player n's deviation has been confirmed) or $d_{n-1}(l+1)(i) = d_{n+1}(l+1)(i) = n$ (the two monitors of player n agree that player n does not take $\alpha_n(l) = \alpha_n(x)$ except for the noise in communication), then $c_i(l+1) \ni n$. Otherwise, $c_i(l+1) = \{0\}$.

Let us explain the basic structure of the coordination on the punishment. For a simple explanation, assume that $d_j(l) = 0$ for all $j \in I$ and there is no noise in the communication: $\lambda_j(l) = \lambda_j(l)(n)$ and $d_j(l+1) = d_j(l+1)(n)$ for all j and n.

See Section 5.5.4.3. Since we want to consider player i's incentive, we consider the transition of $d_j(l+1)$ for the two monitors of player i, players $i-1$ and $i+1$. For $j \in \{i-1, i+1\}$,

1. if $\lambda_{i-1}(l) = G$, then

   (a) player $i-1$ monitors player $i-2 \neq i$ and so $d_{i-1}(l+1) \neq i$; and

(b) player $i+1$ monitors player i but since $\lambda_{i-1}(l)(i+1) = G$ (assuming no error), player $i+1$ always has $d_{i+1}(l+1) = 0$; and

2. if $\lambda_{i-1}(l) = B$, then

   (a) if $x_{i-1} = G$, then $d_{i-1}(l+1) = 0$; and

   (b) if $x_{i-1} = B$, then $d_{i-1}(l+1) = i$. Hence, it is up to player $i+1$ whether or not player i's deviation is confirmed. How player $i+1$ with $\lambda_{i+1}(l) = G$ and $\lambda_{i-1}(l)(i+1) = B$ monitors player i is the same as how player j with $\lambda_j(l) = B$ monitors player i in the two-player case.

From 1 and 2-(a), while $x_{i-1} = G$ or $\lambda_{i-1}(l) = G$, player i does not need to consider the possibility of $d_{i-1}(l+1) = i$. Consider the case with $x_{i-1} = \lambda_{i-1}(l) = B$. From 2-(b), as in the two-player case, we can show that player i should take the static best response to players $(-i)$'s action and player i's value is sufficiently small.

The remaining question is whether player i has an incentive to manipulate $\lambda_{i+1}(l)$ so that player $i+1$ monitors player $i+2$ rather than player i in Case 2-(b). Here is where Assumption 4 plays an important role: Lemma 32 below guarantees that player i cannot manipulate $\lambda_{i+1}(l)$ to increase her payoff.

5.5.4.6 Transition of $\theta_{i-1}(l) \in \{G, B\}$

As in the two-player case, if $\theta_{i-1}(l) = B$, then $\theta_{i-1}(l+1) = B$. Hence, we concentrate on how $\theta_{i-1}(l+1) \in \{G, B\}$ is defined conditional on $\theta_{i-1}(l) = G$. $\theta_{i-1}(l+1) = B$ if one of the following conditions are satisfied:

1. when player $i-1$ sends $\lambda_{i-1}(l+1)$ or when some player $j \in -i$ sends $d_j(l+1)$ to player i via error-reporting noisy cheap talk, the error is reported:

$$g[i-1](\lambda_{i-1}(l+1)) = E$$

or

$$g[i-1](d_j(l+1)) = E \text{ for some } j \in -i; \tag{5.31}$$

2. when player $j \in -i$ sends $\lambda_j(l+1)$ or $d_j(l+1)$ to player $n \in -i$ via error-reporting noisy cheap talk, the receiver (player n) makes a mistake:

   (a) $j \in -(i-1, i)$ and we have

   $$f[n](\lambda_j(l+1)) \neq \lambda_j(l+1) \text{ for some } n \in \{j+1, j+2\} \setminus \{i\};$$

   or

   (b) $j \in -i$ and

   $$f[n](d_j(l+1)) \neq d_j(l+1) \text{ for some } n \in -i. \tag{5.32}$$

   For $d_j(l+1)$, since we map a ternary signal into a sequence of binary signals in Section 5.5.2, the probability of (5.32) depends on $d_j(l+1)$. We adjust this probability so that the probability of (5.32) is independent of $d_j(l+1)$, as we do for (2.22). From now on, when we say (5.32) is satisfied, we take this adjustment into account;

3. for $j \in -i$, at the beginning of the $(l+1)$th review round, player j with $\lambda_{j-1}(l+1)(j) = G$ and $c_j(l+1) \neq \{n\} \subset I$ takes $\alpha_j(l+1) \neq \alpha_j(x)$. With abuse of notation, we include the case with $l = -1$;

4. for $j \in -i$, at the beginning of the $(l+1)$th review round, player j with $\lambda_{j-1}(l+1)(j) = G$ and $c_j(l+1) = \{n\} \subset I$ takes $\alpha_j(l+1) \neq \alpha^*_{j,n}$. With abuse of notation, we include the case with $l = -1$; or

5. for $j \in -i$, player j takes $\alpha_j(x)$ but (5.21), (5.23) or (5.24) is not satisfied (indices i and j are reversed).

Note that, in order to calculate the transition, player $i-1$ needs to know players $-i$'s history such as $\lambda_j(l+1)$ with $j \in -(i-1, i)$. These variables are sent in the re-report block and so included in $h_{i-1}^{\text{rereport}}$. Since $\theta_{i-1}(l+1)$ only affects the reward function (does not affect the action plan), it suffices that player $i-1$ knows the information by the end of the review phase.

The important property of $\theta_{i-1}(l+1)$ is that the distribution of $\theta_{i-1}(l+1)$ is independent of player i's strategy. Let us check all the conditions:

1. the probability for Condition 1 is exogenously given and independent of $\lambda_{i-1}(l+1)$ and $d_{j\in\in}(l+1)$;

2. the probability for Condition 2 is exogenously given and independent of player j's message;

3. Condition 3 is determined by the mixture by players $-i$ and the probability is fixed at η;

4. Condition 4 is determined by the mixture by players $-i$ and the probability is fixed at η; and

5. if Condition 1, 2, 3 or 4 is satisfied, then $\theta_{i-1}(l+1) = B$ is determined. Otherwise, as we will see in Section 5.6.2, players $-i$ take $\alpha_{-i}(x)$ and the probability of Conditions 5 is independent of player i's action plan.

5.6 Properties of the Equilibrium

5.6.1 Pairwise Distinguishability

Assumption 4 implies that each player i can distinguish which one of each pair $\{j,n\}$ is more likely to deviate, which derives the following two lemmas:

Lemma 32 For all $\bar{u}$ and L, there exists $\bar{\varepsilon}$ such that, for all $\varepsilon < \bar{\varepsilon}$, for each $i \in I$ and $l = 1,...,L-1$, if players $-i$ plays $\alpha_{-i}(l) = \alpha_{-i}(x)$, then $\lambda_{n-1}(l+1) = B$ with $n-1 \in -(i-1,i)$ implies $\theta_{i-1}(l+1) = B$.

Proof: Without loss, assume $\lambda_{n-1}(l) = G$ and $\lambda_{n-1}(l+1) = B$ with $n-1 \in -(i-1,i)$. Then,

- Condition 1 of Section 5.5.4.1 implies $\frac{1}{|T_{n-1}(l,x)|}\sum_{t\in T_{n-1}(l,x)} \mathbf{1}_{y_{n-1,t}}$ is far from $q_{n-1}(x)$; and
- Condition 3 of Section 5.5.4.1 implies $\frac{1}{|T_{n-1}(l,x)|}\sum_{t\in T_{n-1}(l,x)} \mathbf{1}_{y_{n-1,t}}$ is close to $\boldsymbol{Q}_{n-1}^{n}(x)$.

Together with Lemma 30, this implies $\frac{1}{|T_{n-1}(l,x)|}\sum_{t\in T_{n-1}(l,x)} \mathbf{1}_{y_{n-1,t}}$ is far from $\boldsymbol{Q}_{n-1}^{j}(x)$ for all $j \in -(n-1,n) \ni i$, which implies $\theta_{i-1}(l+1) = B$ from Condition 5 of Section 5.5.4.6. ∎

Lemma 33 For each q, there exists $\bar{\varepsilon}_q > 0$ such that, for all $\varepsilon < \bar{\varepsilon}_q$, for each $i \in I$ and $l = 1, ..., L-1$, if players $-i$ plays $\alpha_{-i}(l) = \alpha_{-i}(x)$, then $d_{n+1}(l+1) = n$ with $n \in -(i-1,i)$ implies $\theta_{i-1}(l+1) = B$.

Proof: The same as Lemma 32. ∎

5.6.2 $a_{-i}(l)$, $\lambda_j(l+1)$, $d_j(l+1)$ and $c_j(l+1)$

In this section, we assume each message m after each lth review round by player $j \in -i$ can transmit correctly to player $n \in -i$ since otherwise, Condition 2 of Section 5.5.4.6 implies $\theta_{i-1}(l+1) = B$ conditional on m.

Let l (if any) be the review round where player $i-1$ has $\lambda_{i-1}(l) = G$ and $\lambda_{i-1}(l+1) = B$. Without an error, $\lambda_{i-1}(l)(i+1) = G$.

First, consider the situation until the lth review round. Let $l^* \leq l$ (if any) be the first review round where (1) there exists $j \in -(i-1,i)$ with $\lambda_j(l^*+1) = B$ or (2) player $i+1$ decides $\lambda_i(l^*+1)(i+1)$ by $f[i+1](\lambda_i(l^*+1))$ (obeys player i's message).

Since $\lambda_j(l^*) = G$ for all $j \in -i$ and $l^* \leq l$, $d_j(\tilde{l}+1) \in \{0, j-1\}$ for all $j \in -i$ and $\tilde{l} \leq l^*$.

Let $\hat{l} \leq l^*$ (if any) be the first review round where (3) there exists $j \in -i$ with $d_j(\hat{l}+1) = j-1$. Then, players $-i$ take $\alpha_{-i}(\hat{l}) = \alpha_{-i}(x)$ or $\theta_{i-1}(\hat{l}) = B$ from Condition 3 of Section

5.5.4.6. Since $\lambda_{i-1}(l) = \lambda_{i-1}(l)(i+1) = G$ implies $d_{i+1}(\hat{l}+1) = 0$, (3) implies $d_j(\hat{l}+1) = j-1$ with $j \in -(i, i+1)$. Thus, Lemma 33 implies $\theta_{i-1}(\hat{l}+1) = B$.

Hence, players $-i$ take $\alpha_{-i}(l^*) = \alpha_{-i}(x)$ or $\theta_{i-1}(l^*) = B$. If the former is the case, then (1) $\lambda_j(l^*+1) = B$ with $j \in -(i-1, i)$ implies $\theta_{i-1}(l^*+1) = B$ from Lemma 32 and (2) player $i+1$ decides $\lambda_i(l^*+1)(i+1)$ by $f[i+1](\lambda_i(l^*+1))$ implies $\theta_{i-1}(l^*+1) = B$ from Condition 2 of Section 5.5.4.2 and Condition 5 of Section 5.5.4.6.

In total, for $\tilde{l} \le l$,

- players $-i$ play $\alpha_{i-1}(\tilde{l}) = \alpha_{-i}(x)$ or $\theta_{i-1}(\tilde{l}) = B$;
- "$\lambda_j(\tilde{l}+1) = G$ for all $j \in -(i-1, i)$, $\lambda_i(\tilde{l}+1)(i+1) = G$ regardless of player i's message, and $d_j(\tilde{l}+1) = 0$ for all $j \in -i$" or $\theta_{i-1}(\tilde{l}+1) = B$; and
- if $\alpha_{-i}(l+1) \neq \alpha_{-i}(x)$, then $\theta_{i-1}(l+1) = B$ (this follows from the second bullet point).

Second, consider the $\tilde{l}$th review rounds with $\tilde{l} \ge l+1$. Since $\lambda_{i-1}(l+1) = B$, $d_{i-1}(\tilde{l}+1) = 0$ for all $\tilde{l} \ge l+1$ if $x_{i-1} = G$ and $d_{i-1}(\tilde{l}+1) = i$ for all $\tilde{l} \ge l+1$ if $x_{i-1} = B$. From above, we can assume $d_j(l+1) = 0$ for all $j \in -i$, $\lambda_j(l+1) = G$ for all $j \in -(i-1, i)$ and $\lambda_i(l+1)(i+1) = G$ (otherwise, $\theta_{i-1}(l+1) = B$). Again, let $l^* \ge l+1$ (if any) be the first review round where (1) there exists $j \in -(i-1, i)$ with $\lambda_j(l^*+1) = B$ or (2) player $i+1$ decides $\lambda_i(l^*+1)(i+1)$ by $f[i+1](\lambda_i(l^*+1))$.

Since $\lambda_j(l^*) = G$ for $j \in -(i-1, i)$, $d_j(\tilde{l}+1) \in \{0, j-1\}$ for all $j \in -(i-1, i)$ and $\tilde{l} \le l^*$.

Let $\hat{l} \le l^*$ with $\hat{l} \ge l+1$ (if any) be the first review round where there exists $j \in -(i-1, i)$ with $d_j(\hat{l}+1) = j-1$. Note that players $-i$ take $\alpha_{-i}(\hat{l}) = \alpha_{-i}(x)$ or $\theta_{i-1}(\hat{l}) = B$ from Condition 3 of Section 5.5.4.6. Consider the following three cases:

1. if there exists $j \in -(i-1, i, i+1)$ with $d_j(\hat{l}+1) = j-1 \neq i$, then Lemma 33 implies $\theta_{i-1}(\hat{l}+1) = B$; and

2. if $d_j(\hat{l}+1) = 0$ for all $j \in -(i-1,i,i+1)$, then $d_{i+1}(\hat{l}+1) = i$. If $x_{i-1} = G$, then this is contradiction. Hence, consider the case with $x_{i-1} = B$. Since player $i-1$ with $x_{i-1} = \lambda_{i-1}(l+1) = B$ has $d_{i+1}(\tilde{l}+1) = i$ for all $\tilde{l} \geq l+1$, we have $c_j(\hat{l}+1) = \{i\}$ for all $j \in -i$ regardless of player i's message (note that the deviation by player $n \in \{i-1, i+1\}$ cannot be confirmed with $d_j(\hat{l}+1) = 0$ for all $j \in -(i-1,i,i+1)$). Hence, $d_j(\tilde{l}+1)$ with $\tilde{l} \geq \hat{l}+1$ will be fixed for all $j \in -i$. Further, since $c_j(\tilde{l}+1) \neq \{0\}$ for all $j \in -i$, $\lambda_j(\tilde{l}+1)$ will be fixed for all $j \in -i$ and $\tilde{l} \geq \hat{l}+1$ from Condition 2 of Section 5.5.4.1 and $\lambda_{j-1}(\tilde{l}+1)(j)$ will be fixed for all $j \in -i$ and $\tilde{l} \geq \hat{l}+1$ from Condition 2 of Section 5.5.4.2. Therefore,

   (a) if $\hat{l} < l^*$, then this contradicts $\lambda_j(l^*) = G$ and $\lambda_j(l^*+1) = B$ for some $j \in -(i-1,i)$; and

   (b) if $\hat{l} = l^*$, then players $-i$ take $\alpha_{-i}(l^*) = \alpha_{-i}(x)$ or $\theta_{i-1}(l^*) = B$. If the former is the case, then $\lambda_j(l^*+1) = B$ with $j \in -(i-1,i)$ implies $\theta_{i-1}(l^*+1) = B$ from Lemma 32.

In total, for $\tilde{l} \geq l+1$,

- unless $c_j(\tilde{l}) = \{i\}$ for all $j \in -i$, players $-i$ play $\alpha_{i-1}(\tilde{l}) = \alpha_{-i}(x)$ or $\theta_{i-1}(\tilde{l}) = B$; and
- "$\lambda_j(\tilde{l}+1) = G$ for all $j \in -(i-1,i)$, $\lambda_i(\tilde{l}+1)(i+1) = G$, $c_j(\tilde{l}+1) = \{i\}$ for all $j \in -i$ if $x_{i-1} = B$ and $d_{i+1}(\tilde{l}+1) = i$, and $c_j(\tilde{l}+1) = \{0\}$ if $x_{i-1} = G$ or $d_{i+1}(\tilde{l}+1) = 0$," regardless of player i's messages, or $\theta_{i-1}(\tilde{l}+1) = B$.

Therefore, we have the following lemma:

Lemma 34 For each $i \in I$ and $l = 1, ..., L$, in the lth review round, $\theta_{i-1}(l) = B$ or the following statement is true:

1. if $\lambda_{i-1}(l) = G$, then

(a) players $-i$ play $\alpha_{i-1}(l) = \alpha_{-i}(x)$ and $\lambda_j(l) = G$ for all $j \in -(i-1, i)$, $\lambda_i(l)(i+1) = G$, and $d_j(l) = 0$ for all $j \in -i$; and

(b) for the next review round,

i. $\alpha_{-i}(l+1) = \alpha_{-i}(x)$, $\lambda_j(l+1) = G$ for all $j \in -(i-1, i)$, $\lambda_i(l+1)(i+1) = G$ regardless of player i's message, and $d_j(l+1) = 0$ for all $j \in -i$; or

ii. $\theta_{i-1}(l+1) = B$; and

2. if $\lambda_{i-1}(l) = B$, then one of the following two is correct:

(a) $c_j(l) = \{i\}$ for all $j \in -i$ and players $-i$ takes $\alpha_{-i}(l) = \alpha^*_{-i}$. For the next round, player $-i$ take $\alpha_{-i}(l+1) = \alpha^*_{-i}$ or $\theta_{i-1}(l+1) = B$; or

(b) $c_j(l) = \{0\}$ for all $j \in -i$ and players $-i$ takes $\alpha_{-i}(l) = \alpha^*(x)$. For the next round,

i. $\lambda_j(l+1) = G$ for all $j \in -(i-1, i)$, $\lambda_i(l+1)(i+1) = G$, and "$c_j(l+1) = \{i\}$ for all $j \in -i$ if $x_{i-1} = B$ and $d_{i+1}(l+1) = i$" and "$c_j(l+1) = \{0\}$ for all $j \in -i$ if $x_{i-1} = G$ or $d_{i+1}(l+1) = 0$," regardless of player i's messages; or

ii. $\theta_{i-1}(l+1) = B$; and

3. $\lambda_{j-1}(l)(j) = G$ for all $j \in -i$.

Proof: Claims 1 and 2 follow the discussion above. Claim 3 requires the proof. Suppose $\lambda_{j-1}(l-1)(j) = B$ and $\lambda_{j-1}(l)(j) = G$. By Section 5.5.4.2, $c_j(l-1) \neq \{0\}$. From Claims 1 and 2 of the current lemma, we can concentrate on the case where players $-i$ play $\alpha_{i-1}(l) = \alpha_{-i}(x)$. Then, Condition 2 of Section 5.5.4.2 and Condition 5 of Section 5.5.4.6 implies $\theta_{i-1}(l) = B$. ∎

5.6.3 Player i's Belief

Consider player i's inference of $\lambda_{i-1}(l+1)$, $\lambda_{i-1}(l+1)(i) \in \{G, B\}$. As in the proof of Lemma 15, let l^* ($\hat{l}^*$, respectively) be such that $\lambda_{i-1}(l) = B$ ($\lambda_{i-1}(l)(i) = B$, respectively) is initially induced in the (l^*+1)th $((\hat{l}^*+1)$th, respectively) review round. If $\lambda_{i-1}(L) = G$ $(\lambda_{i-1}(L)(i) = G$, respectively), then define $l^* = L$ $(\hat{l}^* = L$, respectively). Then, there are following three cases:

$l^* = \hat{l}^*$ This means $\lambda_{i-1}(l) = \lambda_{i-1}(l)(i)$ for all l.

$l^* > \hat{l}^*$ This means that player i obeys the message in the supplemental round for $\lambda_{i-1}(\hat{l}^*+1)$ between $i-1$ and i:

$$\lambda_{i-1}(\hat{l}^*+1)(i) = f[i](\lambda_{i-1}(\hat{l}^*+1)).$$

By Lemma 11, player i believes that, conditional on $\lambda_{i-1}(\hat{l}^*+1)$,

$$\lambda_{i-1}(\hat{l}^*+1)(i) = f[i](\lambda_{i-1}(\hat{l}^*+1)) = \lambda_{i-1}(\hat{l}^*+1)$$

or $g[i-1](\lambda_{i-1}(\hat{l}^*+1)) = E$ with probability no less than $1-\exp(-\Theta(T^{\frac{1}{2}}))$.

If the former is the case, then $l^* \leq \hat{l}^*$ (contradiction). If the latter is the case, then $\theta_{i-1}(\hat{l}^*+1) = B$. Since players $(-i)$'s continuation actions plan in the main blocks do not depend on $g[i-1](\lambda_{i-1}(\hat{l}^*+1))$, this belief is valid after learning.

$l^* < \hat{l}^*$ If $\lambda_{i-1}(l^*+1)(i) = f[i](\lambda_{i-1}(l^*+1))$, then as in the case with $l^* > \hat{l}^*$, player i believes that $\theta_{i-1}(l) = B$ with probability $1-\exp(-\Theta(T^{\frac{1}{2}}))$ conditional on $\lambda_{i-1}(l^*+1)$, even after learning from the continuation action plan by players $-i$.

If player i disregards the message, then as in the case with two players, there are following two cases:

1. $c_i(l^*) \neq \{0\}$. Then, from Lemma 34, there should exist $j \in -i$ and $l < l^*$ such that

$$f[i](d_j(l)) \neq 0 = d_j(l),$$

or $\theta_{i-1}(l) = B$. If the former is the case, then by Lemma 11, player i believes that, conditional on $d_j(l)$, $g[i-1]\left(d_j(l)\right) = E$ with probability $1-\exp(-\Theta(T^{\frac{1}{2}}))$. Since players $(-i)$'s continuation action plans in the main blocks do not depend on $g[i-1](d_j(l))$, this belief is valid after learning; and

2. $c_i(l^*) = \{0\}$. From Lemma 34, unless $\theta_{i-1}(l^*) = B$, $c_j(l^*) = \{0\}$, $\lambda_j(l^*) = G$ with $j \in -i$ and $\lambda_i(l^*)(i+1) = G$. From $\lambda_j(l^*) = G$ with $j \in -i$, unless $\theta_{i-1}(l^*) = B$, $\lambda_{j-1}(l^*)(j) = B$ for all $j \in -(i, i+1)$. In total, unless $\theta_{i-1}(l^*) = B$, $c_j(l^*) = \{0\}$ and $\lambda_{j-1}(l^*)(j) = B$ for all $j \in -i$.

   Since player i disregards the message, from Condition 2 of Section 5.5.4.2, (5.21) is satisfied and there exists some player $j \in -i$ such that player i's signal frequency while player i is taking $a_i\left(x\right)$ is close to $\boldsymbol{Q}_i^j\left(x\right)$. Let $\mathcal{J}$ be the set of such player $j \in -i$.

   If $|\mathcal{J}| > 1$, then by Lemma 30 implies player i's signal frequency while player i is taking $a_i\left(x\right)$ is close to the ex ante mean $q_i\left(x\right)$ and player i believes that the score is not erroneous and $\lambda_{i-1}(l^*+1) = G$ with probability $1-\exp(-\Theta(T))$.

   If $|\mathcal{J}| = 1$, then the same proof as Lemma 15 shows that player i believes that the score is not erroneous or $\theta_{i-1}(l^*) = G$ with probability $1-\exp(-\Theta(T))$.

   These bounds are before learning from the continuation action plan by players $-i$. The learning from the continuation action plan changes the belief in the following two ways. First, when player $j \in -i$ sends a message m, $f[i](m)$ reveals the part of the histories m. However, since any sequence of signals occurs with positive probability $\left(\exp(-\Theta(T^{\frac{1}{2}}))\right)^{(2N+N(N-1))L} = \exp(-\Theta(T^{\frac{1}{2}}))$, the update of the belief is sufficiently small compared to the original belief $1-\exp(-\Theta(T))$.

   Second, player i conditions that $\lambda_{j-1}(\tilde{l})(j) = G$ with $j \in -i$. Since player j observes any $f[j](\lambda_{j-1}(\tilde{l}))$ with probability $\exp(-\Theta(T^{\frac{1}{2}}))$ regardless of a history of players $-j$, the update of the belief is $\left(\exp(-\Theta(T^{\frac{1}{2}}))\right)^{L} = \exp(-\Theta(T^{\frac{1}{2}}))$.

Third, since player j with $\lambda_{j-1}(\tilde{l})(j)$ takes all the actions with probability $\Theta(\eta)$, the update of the belief is $(\Theta(\eta))^L$.

Next, consider player i's inference of $c_j(l)$ with $j \in -i$. If there exist $j \in -i$ and $\tilde{l} \leq l$ with $c_j(\tilde{l}) \neq \{0\}, \{i\}$, then Lemma 34 implies $\theta_{i-1}(l) = B$. In addition, if there exist $j, j' \in -i$ and $\tilde{l} \leq l$ with $c_j(\tilde{l}) \neq c_{j'}(\tilde{l})$, then Lemma 34 implies $\theta_{i-1}(l) = B$. Further, if there exist $j \in -i$ and $\tilde{l} \leq l$ such that player i's message affects $c_j(\tilde{l})$, then $\theta_{i-1}(l) = B$.

Hence, we can concentrate on the case where $c_j(l) = c_{j'}(l) = \{0\}, \{i\}$ for $j, j' \in -i$ and $c_j(l) = c_{j'}(l)$ is independent of player i's message. If $c_i(\tilde{l}) \neq c_j(\tilde{l})$ for some $j \in -i$ and $\tilde{l} < l$, then there should exist $n \in -i$ and $\hat{l} \leq \tilde{l}$ such that

$$f[i](d_n(\hat{l})) \neq d_n(\hat{l})$$

or

$$f[j](d_n(\hat{l})) \neq d_n(\hat{l}).$$

If the former is the case, player i believes that, conditional on $d_n(\hat{l})$, $g[i-1](d_n(\hat{l})) = E$ with probability $1 - \exp(-\Theta(T^{\frac{1}{2}}))$. Since player $(i-1)$'s continuation action plan in the main blocks does not depend on $g[i-1](d_n(\hat{l}))$, this belief is valid after learning. If the latter is the case, then Condition 2 of Section 5.5.4.6 implies $\theta_{i-1}(l+1) = B$.

In total, we have shown the following lemma:

Lemma 35 For all $\bar{u}$ and L, there exists $\bar{\eta}$ such that, for all $\eta < \bar{\eta}$, there exists $\bar{\rho} > 0$ such that, for all $\rho < \bar{\rho}$, there exist $\bar{\varepsilon}$ and $\{A_i(x)\}_{i,x}$ such that, for all $\varepsilon < \bar{\varepsilon}$, for any history h_i^t with t being in the lth review round, conditional on $\lambda_{j-1}(\tilde{l})(j) = G$ for all $j \in -i$ and $\tilde{l} \leq l$ and $\alpha_{-i}(l)$, player i after h_i^t believes with probability $1 - \exp(-\Theta(T^{\frac{1}{2}}))$ that

1. $\hat{\lambda}_{i-1}(l)(i) = \lambda_{i-1}(l)$ and $c_i(l) = c_j(l)$ for all $j \in I$; or

2. $\theta_{i-1}(l) = B$.

5.7 Variables

In this section, we show that we can take all the variables necessary for the equilibrium construction appropriately: $\bar{u}$, q_2, q_1, $\underline{u}$, $\bar{q}$, L, η, ρ and ε.

First, Lemma 26 determines $\bar{u}$ and Lemma 28 determines q_1 and q_2.

Second, from (1.18), we have

$$\max\left\{v_i^*, \max_{x:x_{i-1}=B} u_i(a(x))\right\} < \underline{v}_i < v_i < \overline{v}_i < \min_{x:x_{i-1}=G} u_i(a(x)).$$

Take $\underline{u}$ sufficiently small so that

$$v_i\left(B\right) + \underline{u} < \underline{v}_i < v_i < \overline{v}_i < \min_{x:x_{i-1}=G} u_i(a(x)),$$

where $v_i\left(B\right) = \max\{v_i^*, \max_{x:x_{i-1}=B} u_i(a(x)) + 2\underline{u}$ as defined in (5.1). Re-take $\bar{u}$ if necessary to have (5.3).

Third, take L sufficiently large sufficiently small such that

$$\frac{(L-1)\,v_i\left(B\right)+\bar{u}}{L} + \underline{u} + 2\frac{\bar{u}}{L} < \underline{v}_i < v_i < \overline{v}_i < \min_{x:x_{i-1}=G} u_i(a(x)) - 2\frac{\bar{u}}{L}.$$

Fourth, given $L, q_2, q_1, \underline{u}$, fix $\bar{q}_1$ and $\bar{\rho}_1$ so that Lemma 31 holds.

Fifth, take $\bar{\rho}_2 < \bar{\rho}_1$ so that, for all $\rho < \bar{\rho}_2$, (5.2) holds and

$$\frac{(L-1)\,v_i\left(B\right)+\bar{u}}{L} + \underline{u} + 2\frac{\bar{u}}{L} < \underline{v}_i < v_i < \overline{v}_i < \min_{x:x_{i-1}=G} w_i(x) - 2\frac{\bar{u}}{L}.$$

Sixth, given $\bar{u}$ and L, take $\bar{\eta}_1$ so that for all $\eta < \bar{\eta}_1$, we have

$$\begin{aligned} &\frac{(L-1)\,v_i\left(B\right)+\bar{u}}{L} + \underline{u} + 2\frac{\bar{u}}{L} + \eta L\left(2\bar{u} - \min_{i,x}\frac{(L-1)\,v_i\left(B\right)+\bar{u}}{L}\right) \\ < \;\; &\underline{v}_i < \overline{v}_i < \min_{x:x_j=G} w_i\left(x\right) - 2\frac{\bar{u}}{L} - \eta L\left(2\bar{u} + \max_{i,x} w_i\left(x\right)\right). \end{aligned} \tag{5.33}$$

Finally, fix $q < \bar{q}_1$ sufficiently small. Then, we can take $\bar{\varepsilon}_1$ for Lemma 32 and $\bar{\varepsilon}_q$ for Lemma 33. Fix $\eta < \bar{\eta}_1$, ρ and $\varepsilon < \min\{\bar{\varepsilon}_1, \bar{\varepsilon}_q\}$ so that Lemma 35 hold.

5.8 Almost Optimality

Based on Lemmas 34 and 35, we now show that if we properly define $\bar{\pi}_i(x, \lambda_{i-1}(l), c_{i-1}(l), l)$, then $\sigma_i(x_i)$ and π_i^{main} satisfy (1.21), (1.16) and (1.17):

Proposition 36 For all $i \in I$, there exists $\bar{\pi}_i(x, \lambda_{i-1}(l), c_{i-1}(l), l)$ such that

1. $\sigma_i(x_i)$ is almost optimal conditional on $\lambda_{j-1}(l)(j) = G$ for all $j \in -i$ and $\tilde{l} \leq l$: for each $l \in \{1, ..., L\}$,
   (a) for any period t in the lth review round, (1.21) holds; and
   (b) when player i sends messages by the error-reporting noisy cheap talk in the supplemental rounds, (1.21) holds;[2]
2. (1.16) is satisfied with π_i replaced with π_i^{main}. Since each $x_i \in \{G, B\}$ gives the same value conditional on x_{-i}, the strategy in the coordination block is optimal;[3] and
3. π_i^{main} satisfies (1.17).

Lemma 34 shows that $\lambda_{j-1}(l)(j) = B$ for some $j \in -i$, then $\theta_{i-1}(l) = B$ and that player i can condition on $\lambda_{j-1}(l)(j) = G$ for all $j \in -i$.

Lemma 34 also shows that, whenever player i's message changes some player's action after the lth review round, $\theta_{i-1}(l+1) = B$ is predetermined. In addition, Lemma 35 implies that player i can infer $(c_n(l))_{n \in -i}$ and $\lambda_{i-1}(l)$ with probability $1 - \exp(-\Theta(T^{\frac{1}{2}}))$ (or any action is optimal) by sending messages as prescribed by the equilibrium action plan. Therefore, 1-(b) holds.

[2] If $l = L$, then this is redundant.

[3] As in the two-player case, even after the adjustment of the report block, any $x_i \in \{G, B\}$ still gives exactly the same value.

As in the two-player case, for 3, it suffices to have

$$\bar{\pi}_i(x, \lambda_{i-1}(l), c_{i-1}(l), l) \begin{cases} \leq 0 & \text{if } x_{i-1} = G, \\ \geq 0 & \text{if } x_{i-1} = B, \end{cases} \tag{5.34}$$

$$|\bar{\pi}_i(x, \lambda_{i-1}(l), c_{i-1}(l), l)| \leq \max_{i,a} 2 |u_i(a)| T \tag{5.35}$$

for all $x \in \{G, B\}^N$, $\lambda_{i-1}(l) \in \{G, B\}$, $c_{i-1}(l) \subset \{0\} \cup 2^I$ and $l \in \{1, ..., L\}$.

We are left to construct $\bar{\pi}_i$ so that 1-(a) and 2 are satisfied together with (5.34) and (5.35). Remember that, from Lemma 34 and 35, if $c_i(l) = \{n\} \in -i$, then player i believes $\theta_{i-1}(l) = B$ with probability $1 - \exp(-\Theta(T^{\frac{1}{2}}))$. Hence, it is almost optimal to take any $\alpha_i(l)$.

Further, from Lemmas 34 and 35, we can concentrate on the following five cases:

- $x_{i-1} = G$, $\lambda_{i-1}(l)(i) = \lambda_{i-1}(l) = G$, $c_j(l) = \{0\}$ for all $j \in I$, players $-i$ take $\alpha_{-i}(x)$, $\lambda_j(l) = G$ for all $j \in -(i-1, i)$, and $\lambda_{j-1}(l)(j) = G$ for all $j \in -i$.

  Further, "$\lambda_j(l+1) = G$ for all $j \in -(i-1, i)$, $\lambda_{j-1}(l+1)(j) = G$ for all $j \in -i$, and $c_j(l+1) = \{0\}$ for all $j \in -i$," or $\theta_{i-1}(l+1) = B$;

- $x_{i-1} = G$, $\lambda_{i-1}(l)(i) = \lambda_{i-1}(l) = B$, $c_j(l) = \{0\}$ for all $j \in I$, players $-i$ take $\alpha_{-i}(x)$, $\lambda_j(l) = G$ for all $j \in -(i-1, i)$, and $\lambda_{j-1}(l)(j) = G$ for all $j \in -i$.

  Further, "$\lambda_j(l+1) = G$ for all $j \in -(i-1, i)$, $\lambda_{j-1}(l+1)(j) = G$ for all $j \in -i$, and $c_j(l+1) = \{0\}$ for all $j \in -i$," or $\theta_{i-1}(l+1) = B$;

- $x_{i-1} = B$, $\lambda_{i-1}(l)(i) = \lambda_{i-1}(l) = G$, $c_j(l) = \{0\}$ for all $j \in I$, players $-i$ take $\alpha_{-i}(x)$, $\lambda_j(l) = G$ for all $j \in -(i-1, i)$, and $\lambda_{j-1}(l)(j) = G$ for all $j \in -i$.

  Further, "$\lambda_j(l+1) = G$ for all $j \in -(i-1, i)$, $\lambda_{j-1}(l+1)(j) = G$ for all $j \in -i$, and $c_j(l+1) = \{0\}$ for all $j \in -i$," or $\theta_{i-1}(l+1) = B$;

- $x_{i-1} = B$, $\lambda_{i-1}(l)(i) = \lambda_{i-1}(l) = B$, $c_j(l) = \{0\}$ for all $j \in I$, players $-i$ take $\alpha_{-i}(x)$, $\lambda_j(l) = G$ for all $j \in -(i-1, i)$, and $\lambda_{j-1}(l)(j) = G$ for all $j \in -i$.

Further, "$\lambda_j(l+1) = G$ for all $j \in -(i-1,i)$, $\lambda_{j-1}(l+1)(j) = G$ for all $j \in -i$, and $c_j(l+1) = \{i\}$ for all $j \in -i$ if $d_{i+1}(l+1) = i$ and $c_j(l+1) = \{0\}$ for all $j \in -i$ if $d_{i+1}(l+1) \neq i$," or $\theta_{i-1}(l+1) = B$; and

- $x_{i-1} = B$, $\lambda_{i-1}(l)(i) = \lambda_{i-1}(l) = \lambda_{i-1}(l)(i+1) = B$, $c_j(l) = \{i\}$ for all $j \in I$, players $-i$ take α^*_{-i}, $\lambda_j(l) = G$ for all $j \in -(i-1,i)$, and $\lambda_{j-1}(l)(j) = G$ for all $j \in -i$.

  Further, "$\lambda_j(l+1) = G$ for all $j \in -(i-1,i)$, $\lambda_{j-1}(l+1)(j) = G$ for all $j \in -i$, and $c_j(l+1) = \{i\}$ for all $j \in -i$," or $\theta_{i-1}(l+1) = B$.

Since the distribution of $\theta_{i-1}(l+1)$ is independent of player i's strategy, player i's incentive is the same as in the two-player case. Remember that in the fourth case, player $i+1$ determines $d_{i+1}(l+1) = i$ or not as player j determines $d_j(l+1) = B$ or not in the two-player case.

Therefore, almost optimality of $\sigma_i(x_i)$ and the existence of $\bar{\pi}_i$ with (5.34) and (5.35) can be shown as in the two-player case.

5.9 Report Block

We are left to construct the report and re-report blocks to attain the exact optimality of the equilibrium strategies. In this section, we explain the report block.

5.9.1 Preparation

As in the two-player case, we need to establish the truthtelling incentive in the report block. When player i reports her history $(a_{i,t}, y_{i,t})$ in some period t in the main blocks, intuitively, player $i-1$ punishes player i proportionally to

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

with

$$j=\begin{cases} i-1 & \text{if } i\neq 1,\\ 2 & \text{if } i=1.\end{cases} \tag{5.36}$$

Hence, player i wants to maximize

$$-\mathbb{E}\left[\begin{array}{c} \left\|\mathbf{1}_{a_{j,t},y_{j,t}}-\mathbb{E}\left[\mathbf{1}_{a_{j,t},y_{j,t}}\mid \hat{y}_{i,t},y_{-(i,j),t},\hat{a}_{i,t},a_{-(i,j),t},\alpha_{j,t}\right]\right\|^2 \\ \mid y_{i,t},y_{-(i,j),t},a_{i,t},a_{-(i,j),t},\alpha_{j,t}\end{array}\right]. \tag{5.37}$$

Compared to the two-player case, we assume that player i knows the history of players $-(i,j)$.

Since player j with $\lambda_{j-1}(l)(j)=G$ takes a fully mixed strategy, Assumption 5 implies that the truthtelling is uniquely optimal:

Lemma 37 If Assumption 5 is satisfied, then for each i, any fully mixed strategy $\alpha_{j,t}$ and any history of players $-j$, $(\hat{a}_{i,t},\hat{y}_{i,t})=(a_{i,t},y_{i,t})$ is a unique maximizer of (5.37).

5.9.2 Report Block

Given Lemma 37, we construct the report block.

5.9.2.1 Structure of the Report Block

The report block proceeds as follows:

1. player N sends the message about h_N^{main};
2. player $N-1$ sends the message about h_{N-1}^{main};

   ⋮

3. player 3 sends the message about h_3^{main};
4. then, public randomization y^p is drawn; and

5. player 1 reports h_1^{main} if $y^p \leq \frac{1}{2}$ and player 2 reports h_2^{main} if $y^p > \frac{1}{2}$.

We explain each step in the sequel.

5.9.2.2 Player i sends h_i^{main}

Since there is a chronological order for the rounds and r is a generic serial number of rounds, the notations $\#_i^r$, $\#_i^r(k)$, $T(r,k)$ and $\{a_{i,t}, y_{i,t}\}_{t \in T(r,k)}$ defined in Chapter 2 is still valid except that

- if player i sends m to player n via error-reporting noisy cheap talk in round r, then $\#_i^r$ contains m for sure. In addition, if and only if player $n-1$ is player i, $\#_i^r$ contains $g[n-1](m)$ and $g_2[n-1](m)$; and
- if player i receives m from player n via error-reporting noisy cheap talk in round r, then $\#_i^r$ contains $f[i](m)$ for sure. In addition, if and only if player $n-1$ is player i, $\#_i^r$ contains $f_2[n-1](m)$.

Player i sends the message about h_i^{main} in the same way as player i sends the message in Section 2.9.6. That is, for each round r,

- if round r corresponds to a review round, then
  - first, player i reports the summary $\#_i^r$;
  - second, for each subround k, player i reports the summary $\#_i^r(k)$;
  - third, public randomization is drawn such that each subround k is randomly picked with probability $T^{-\frac{3}{4}}$. Let $k(r)$ be the subround picked by the public randomization; and
  - fourth, for $k(r)$, player i reports the whole history $\{a_{i,t}, y_{i,t}\}_{t \in T(r,k(r))}$ in the $k(r)$th subround; and

- if player i sends or receives a message by error-reporting noisy cheap talk in round r, then player i reports $\#_i^r$.

Again, the necessary number of binary messages to send the information is

$$\Theta(T^{\frac{1}{4}}). \tag{5.38}$$

5.9.2.3 Reward Function π_i^{report}

We are left to define the reward function π_i^{report}. As a preparation, we prove the following lemma:

Lemma 38 Let h_i be player i's history right before player i sends the message about h_i^{main} in the report block. If Assumption 5 is satisfied, then there exists $\bar{\varepsilon} > 0$ such that

1. for each $l \in \{1, ..., L\}$, in the lth review round, there exists $g_i(h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}}, a_i, y_i)$ such that, for period $t \in T(l)$, it is better for player i to report $(a_{i,t}, y_{i,t})$ truthfully: if $\lambda_{j-1}(l)(j) = G$ for all $j \in -i$, then for all h_i,

$$\begin{aligned} & \mathbb{E}\left[g_i(h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}}, \hat{a}_{i,t}, \hat{y}_{i,t}) \mid h_i, (\hat{a}_{i,t}, \hat{y}_{i,t}) = (a_{i,t}, y_{i,t})\right] \\ > \; & \mathbb{E}\left[g_i(h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}}, \hat{a}_{i,t}, \hat{y}_{i,t}) \mid h_i, (\hat{a}_{i,t}, \hat{y}_{i,t}) \neq (a_{i,t}, y_{i,t})\right] + \bar{\varepsilon} T^{-(N-1)}, \end{aligned} \tag{5.39}$$

   where $(\hat{a}_{i,t}, \hat{y}_{i,t})$ is player i's message in the report block; and

2. for round r where player i sends or receives the message by the error-reporting noisy cheap talk, it is better for player i to report player i's history $\#_i^r$ truthfully:

$$\begin{aligned} & \mathbb{E}\left[g_i(h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}}, \hat{\#}_i^r) \mid h_i, \hat{\#}_i^r = \#_i^r\right] \\ > \; & \mathbb{E}\left[g_i(h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}}, \hat{\#}_i^r) \mid h_i, \hat{\#}_i^r \neq \#_i^r\right] + \bar{\varepsilon} T^{-1}, \end{aligned} \tag{5.40}$$

   where $\hat{\#}_i^r$ is player i's message about $\#_i^r$ in the report block.

Proof: 1. By the same proof as Lemma 21, we can show that

$$g_i(h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}}, \hat{a}_{i,t}, \hat{y}_{i,t}) = -\mathbf{1}\{t_j(r) = t \text{ for all } j \in -i\}$$
$$\times \left\| \mathbf{1}_{a_{j,t},y_{j,t}} - \mathbb{E}[\mathbf{1}_{a_{j,t},y_{j,t}} \mid \hat{y}_{i,t}, y_{-(i,j),t}, \hat{a}_{i,t}, a_{-(i,j),t}, \alpha_{j,t}] \right\|^2$$

with

$$j = \begin{cases} i-1 & \text{if } i \neq 1, \\ 2 & \text{if } i = 1 \end{cases} \tag{5.41}$$

works. Note that, compared to Lemma 21, we have

$$T^{-(N-1)} = \Pr\left(\mathbf{1}\{t_j(r) = t \text{ for all } j \in -i\} = 1\right)$$

instead of $T^{-1} = \Pr\left(\mathbf{1}\{t_j(r) = t\} = 1\right)$. Also, we assume player i could know how many times player $j \in -i$ observes each $(a_j, y_j, (\mathbf{1}_{Q_j^n(x)})_{n \in -j}, \Gamma_j^x)$ and player j's history in $t_j(l)$: $(a_{j,t_j(l)}, y_{j,t_j(l)}, (\mathbf{1}_{Q_j^n(x),t_j(l)})_{n \in -j}, \Gamma^x_{j,t_j(l)})_{j \in -i}$.

As we will see, players $-(i-1,i)$ sends the information $t_j(r), a_{-(i,j)}, y_{-(i,j)}$ to player $i-1$ in the re-report block and so they are in $h_{i-1}^{\text{rereport}}$.

2. If player i sends m to player n via error-reporting noisy cheap talk in round r, then

   (a) if $\#_i^r$ contains $g[n-1](m)$ and $g_2[n-1](m)$, then

$$g_i(h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}}, \hat{\#}_i^r) = -\left\| \mathbf{1}_{f_2[i-1](m)} - \mathbb{E}[\mathbf{1}_{f_2[i-1](m)} \mid f[n](m), \hat{\#}_i^r] \right\|^2.$$

   (b) if $\#_i^r$ does not contain $g[n-1](m)$ and $g_2[n-1](m)$, then

$$\begin{aligned} & g_i(h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}}, \hat{\#}_i^r) \\ = & -\left\| \mathbf{1}_{f_2[i-1](m)} - \mathbb{E}[\mathbf{1}_{f_2[i-1](m)} \mid f[n](m), \hat{\#}_i^r, g[n-1](m), g_2[n-1](m)] \right\|^2. \end{aligned}$$

As we will see, player $n \in -(i-1,i)$ sends $f[n](m)$ and player $n-1 \in -(i-1,i)$ sends $g[n-1](m)$ and $g_2[n-1](m)$ to player $i-1$ in the re-report block.

If player i receives m from player n via error-reporting noisy cheap talk in round r, then

(a) if $\#_i^r$ contains $f_2[n-1](m)$, then

$$g_i(h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}}, \hat{\#}_i^r) = -\left\| \mathbf{1}_{g_2[i-1](m)} - \mathbb{E}[\mathbf{1}_{g_2[i-1](m)} \mid m, g[i-1](m), \hat{\#}_i^r] \right\|^2.$$

(b) if $\#_i^r$ does not contain $f_2[n-1](m)$, then

$$\begin{aligned} & g_i(h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}}, \hat{\#}_i^r) \\ = & -\left\| \mathbf{1}_{g_2[i-1](m)} - \mathbb{E}[\mathbf{1}_{g_2[i-1](m)} \mid m, g[i-1](m), \hat{\#}_i^r, f_2[n-1](m)] \right\|^2. \end{aligned}$$

Then, (1.3) and (1.5) imply that truthtelling is optimal.

As we will see, player $n \in -(i-1,i)$ sends m and player $n-1 \in -(i-1,i)$ sends $f_2[n-1](m)$ to player $i-1$ in the re-report block. ∎

Given these preparations, by backward induction, we construct $\pi_i^{\text{report}}\left(h_{i-1}^{T_P+1}, h_{i-1}^{\text{rereport}}, \hat{h}_i^{r+1}, r\right)$ for each r such that

$$\pi_i^{\text{report}}(x_{i-1}, h_{i-1}^{T_P+1}, h_{i-1}^{\text{rereport}}) = \sum_r \pi_i^{\text{report}}\left(h_{i-1}^{r+1}, h_{i-1}^{\text{rereport}}, \hat{h}_i^{r+1}, r\right)$$

makes it optimal to tell the truth in the report block and $\sigma_i(x_i)$ is exactly optimal.

Formally, $\pi_i^{\text{report}}\left(h_{i-1}^{r+1}, h_{i-1}^{\text{rereport}}, \hat{h}_i^{r+1}, r\right)$ is determined as follows. If "$i \in \{1,2\}$ and player i is not selected by the public randomization" or there exists l such that the lth review round is before or equal to round r and $\lambda_{j-1}(l)(j) = B$ for some $j \in -i$, then

$$\pi_i^{\text{report}}\left(h_{i-1}^{r+1}, h_{i-1}^{\text{rereport}}, \hat{h}_i^{r+1}, r\right) = 0. \tag{5.42}$$

Otherwise, $\pi_i^{\text{report}}\left(h_{i-1}^{r+1}, h_{i-1}^{\text{rereport}}, \hat{h}_i^{r+1}, r\right)$ is the summation of the following rewards and punishments.

Punishment for a Lie As in the two-player case, we punish a lie. For round r corresponding to a review round, the punishment is the summation of the following three:

- the number indicating player i's lie about $\{a_{i,t}, y_{i,t}\}_{t \in T(r,k(r))}$:

$$\sum_{t \in T(r,k(r))} T^{-3} g_i(h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}}, \hat{a}_{i,t}, \hat{y}_{i,t}); \tag{5.43}$$

- the number indicating player i's lie about $\#_i^r(k)$:

$$T^{-3} \times T^{\frac{3}{4}} \times \mathbf{1}\left\{\hat{\#}_i^r(k(r)) \neq \sum_{t \in T(r,k(r))} \mathbf{1}_{\hat{a}_{i,t}, \hat{y}_{i,t}}\right\}; \tag{5.44}$$

  and

- the number indicating player i's lie about $\#_i^r$:

$$T^{-3} \times \mathbf{1}\left\{\hat{\#}_i^r \neq \sum_k \hat{\#}_i^r(k)\right\}. \tag{5.45}$$

For round r where player i sends or receives a message m, player $i-1$ punishes player i if it is likely for player i to tell a lie by

$$T^{-3} g_i(h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}}, \hat{\#}_i^r). \tag{5.46}$$

Cancel Out the Expected Punishment by Telling the Truth As in the two-player case, we cancel out the differences in ex ante value of the punishment between difference actions and messages: if player i reports the history (player $i \in \{1, 2\}$ needs to be picked by the public randomization to report the history), then we add the following variable to π_i^{main}:

- if round r is a review round, then

$$\sum_{t \in T(r)} \mathbf{1}\{t \in T\left(r, k(r)\right)\} \mathbf{1}\{t_j(l) = t \text{ for all } j \in -i\} T^{-3} \Pi_i[a_{-(i,j),t}, \alpha_{j,t}](y_{i-1,t});$$

- if player i sends the message in round r, then

$$T^{-3} \Pi_i(f[n]\,(m)),$$

  where player n is the receiver of the message; and

- if player i receives the message in round r, then

$$T^{-3} \Pi_i(n, m)$$

  where player n is the sender of the message m.

Here, $\Pi_i\left[a_{-(i,j)}, \alpha_j\right](y_{i-1})$ is defined so that the differences in (5.43) among action a_i's are canceled out ex ante before taking a_i. Since we assume that player $i-1$ knew $a_{-(i,j)}, \alpha_j$, Assumption 3 is sufficient to construct such $\Pi_i\left[a_{-(i,j)}, \alpha_j\right](y_{i-1})$. As we will see, player $i-1$ gets the information about $a_{-(i,j)}, \alpha_j$ from players $-(i-1, i)$ in the re-report block.

Similarly, $\Pi_i(f[n]\,(m))$ ($\Pi_i(n, m)$, respectively) is defined so that the differences in (5.46) among messages are canceled out ex ante before sending (receiving, respectively) the message. Again, player $n \in -(i-1, i)$ sends $f[n]\,(m)$ (m, respectively) to player $i-1$ in the re-report block. The identifiability to construct such Π_i is guaranteed by Lemma 11.

Reward for Optimal Action and Incentive to Tell the Truth This is the same as in the general two-player case. We construct the reward f_i so that, for each round r, for any period t in round r, for any history h_i^t, conditional on $\mathcal{A}_{-i}(r)$, $\sigma_i(x_i)$ is optimal. Here, $\mathcal{A}_{-i}(r)$ represents

- which state $x_{-i} \in \{G, B\}$ players $-i$ are in;

- for the review round l that is equal to round r, which action plan $\alpha_j(l)$ each player $j \in -i$ takes in the lth review rounds; and

- $\lambda_{j-1}(l)(j) = G$ for all $j \in -i$ in the lth review round if round r is the lth review round or after.

See Section 2.9.6 for the construction of f_i. With more than two players, the definition of $V_i(\mathfrak{h}_i^r, \#_i^r)$ includes the learning about $(\lambda_j(l+1), d_j(l+1))_{j \in -i}$.

5.10 Re-Report Block

This is the block for each player $i-1$ to collect the information owned by players $-(i-1, i)$ which is necessary to construct player$(i-1)$'s reward on player i, π_i.

In the re-report block, we have the following rounds in this chronological order:

- players $-(N-1, N)$ send the information to player $N-1$ to construct π_N;

- players $-(N-2, N-1)$ send the information to player $N-2$ to construct π_{N-1};

  $\vdots$

- players $-(1, 2)$ send the information to player 1 to construct π_2; and

- players $-(N, 1)$ send the information to player N to construct π_1.

We explain what information is sent for each step:

5.10.1 Information Sent by Players $-(i-1, i)$ to Player $i-1$

First, each player $n \in -(i-1, i)$ sends the information about their histories in the coordination and main blocks:

- which state x_n player n has;
- for each lth review round, what strategy $\alpha_n(l)$ player n took. Remember that player n's strategy within a round is $i.i.d$;
- for each lth review round, for each $\left(a_n, y_n, \left(\mathbf{1}_{Q_n^j(x)}\right)_{j\in -n}, \Gamma_n^x\right)$, how many times player n observed $\left(a_n, y_n, \left(\mathbf{1}_{Q_n^j(x)}\right)_{j\in -n}, \Gamma_n^x\right)$;
- for each lth review round, which period $t_n(l)$ is excluded from $T_n(l)$;
- what is the history $\left(a_n, y_n, \left(\mathbf{1}_{Q_n^j(x)}\right)_{j\in -n}, \Gamma_n^x\right)$ in the excluded period $t_n(l)$; and
- for each supplemental round, m, $f[n](m)$, $f_2[n](m)$, $g[n](m)$ and $g_2[n](m)$, whenever they are the pieces of information that player n sends or receives.

Second, the players communicate about the histories related to the report block:

- for each round r corresponding to a review round;
  - first, player $i-1$ sends to players $-(i-1,i)$ which $k(r)$ player $i-1$ and i coordinate about sending $(a_{i,t}, y_{i,t})$ for each r corresponding to the review round. With public randomization, $k(r)$ is public. Expecting that we replace public randomization with coordination through private signals, we let player $i-1$ speak $k(r)$ here; and
  - then, for each r, based on player $(i-1)$'s report $k(r)$, each player $n \in -(i-1,i)$ sends $(a_n, y_n, (\mathbf{1}_{Q_n^j(x)})_{j\in -n}, \Gamma_n^x)_{t\in T(r,k(r))}$; and
- for a supplemental round, each player $n \in -(i-1,i)$ sends player n's signals $f[n](m)$, $f_2[n](m)$, $g[n](m)$ and $g_2[n](m)$.[4]

Then, player $i-1$ collects all the information necessary to construct player $(i-1)$'s reward on player i, π_i. Further, the cardinality of the messages sent in the re-report block is

$$\Theta(T^{\frac{1}{4}}) \tag{5.47}$$

[4] If player n is not a receiver of a signal, then we exclude that signal.

by the same calculation as for (2.41).

Chapter 6

General Two-Player Game Without Cheap Talk

In this chapter, we prove the dispensability of the perfect cheap talk, error-reporting noisy cheap talk and public randomization in the proof of Theorem 7 and thereby prove Theorem 8.

After we summarize new notations and assumptions in Section 6.1, we show that the players can communicate and coordinate via actions and private signals. We take the following steps to dispense with the perfect and error-reporting noisy cheap talk and public randomization device.

Remember that the coordination block uses the perfect cheap talk to communicate x, that the supplemental rounds for $\lambda_i(l+1)$ and $d_i(l+1)$ use the error-reporting noisy cheap talk, and that the report block uses the public randomization and perfect cheap talk.

First, in Section 6.2, we replace the perfect cheap talk in the coordination block with the error-reporting noisy cheap talk. Although x_i is no longer common knowledge, by exchanging the messages via error-reporting noisy cheap talk several times, each player can construct an inference of x_i such that, given the opponent's inference, each player puts a belief no less than $1-\exp(-\Theta(T^{\frac{1}{2}}))$ on the event that if their inferences are different, then the opponent has made her indifferent between any action profile sequence in the main blocks. No additional assumption is necessary for this step.

Second, Section 6.3 explains the structure of the review phase (finitely repeated game) when we dispense with perfect cheap talk.

Third, in Section 6.4, we dispense with the error-reporting noisy cheap talk in the coordination block (given the first step above) and supplemental rounds. See Section 1.5 for the intuition and Section 6.1.1 for a new assumption sufficient for this step.

Third, in Section 6.5, we summarize the discussion and formally define the equilibrium strategy in the coordination and main blocks without perfect and error-reporting noisy cheap talk. The almost optimality of this strategy is verified in Section 6.6.

Fourth, in Section 6.7.3, we dispense with the public randomization in the report block, keeping the perfect cheap talk. Now, the players coordinate with their actions and private signals. Section 6.1.2 offers a sufficient condition for this step.

Fifth, in Section 6.7.4, we replace the perfect cheap talk in the report block with "conditionally independent noisy cheap talk," where each message transmits correctly with a high probability but there is a positive probability that an error happens. The sender does not receive any feedback about whether the message transmits correctly or not.

In the report block, the receiver does not have an incentive to infer the messages correctly since the messages are used only for the reward on the sender. Hence, we can disregard the incentives for the receiver.

We can show that the number of the messages that need to be sent to convey the information is sufficiently small compared to the precision of the conditionally independent noisy cheap talk, which means that all the messages transmit correctly with an ex ante high probability. Since the cheap talk is conditionally independent, the sender always believes that each message transmits correctly with a high probability. Given the strict incentive in the report block for the sender, this is enough to incentivize the sender to tell the truth and to construct π_i^{report} to make $\sigma_i(x_i)$ exactly optimal. This step does not require any assumption in addition to the availability of the conditionally independent cheap talk.

Then, we replace the conditionally independent noisy cheap talk with messages via actions. This step is novel since we do not assume anything about the differences in the cardinality of the support of each player's signals. See Section 6.1.3 for what generic assumption is sufficient.

When we say players i and j in this chapter, unless otherwise specified, it implies that $i \neq j$. In addition, without loss of generality, we assume that

$$|A_1|\,|Y_1| \geq |A_2|\,|Y_2|\,. \tag{6.1}$$

6.1 Notations and Assumptions

6.1.1 Assumption for Dispensing with the Error-Reporting Noisy Cheap Talk

As we will see, player j who wants to send a binary message $m \in \{G, B\}$ takes different actions for different messages. We assume that there exist a_j^G and a_j^B such that for any a_i, the different history of the receiver has different information about whether player j takes a_j^G or a_j^B:

Assumption 39 For each $j \in I$, there exist $a_j^G \in A_j$ and $a_j^B \in A_j$ such that

1. for each $a_i \in A_i$ and $y_i \in Y_i$, we have

$$q_i(y_i|a_i, a_j^G) \neq q_i(y_i|a_i, a_j^B);$$

and

2. for all $\alpha_j \in \Delta(A_j)$ with $\alpha_j(a_j^G) > 0$ and $\alpha_j(a_j^B) > 0$, for all (a_i, y_i) and (a_i', y_i') with $(a_i, y_i) \neq (a_i', y_i')$, we have

$$\mathbb{E}\left[\mathbf{1}_{a_j, y_j} \mid y_i, a_i, \alpha_j\right] \neq \mathbb{E}\left[\mathbf{1}_{a_j, y_j} \mid y_i', a_i', \alpha_j\right].$$

6.1.2 Assumption for Dispensing with the Public Randomization

When we dispense with the public randomization in the report block, the players use actions and private signals to coordinate. Fix $i \in I$ arbitrarily.

There are two different assumptions to achieve this goal, each one of which is sufficient. We will explain these two separately.

6.1.2.1 Conditionally Dependent Signals

When the players' signals are correlated given some action profile $a^{\text{p.r.}(i)} \in A$,[1] then we can proceed as follows: the players play $a^{\text{p.r.}(i)}$ and each player observes her own private signal.

Player j partitions the set of her signals into non-empty subsets $Y_{j,1}^i$ and $Y_{j,2}^i$ with $Y_j = Y_{j,1}^i \cup Y_{j,2}^i$.

Player i tries to infer which set player j's signal belongs to. With some $\bar{p}_i \in (0,1)$, player i classifies the set of her signals into two classes: the set of signals with which player i thinks that player j observes $y_j \in Y_{j,1}^i$ with probability more than $\bar{p}_i$ and the set of signals with which player i thinks that player j observes $y_j \in Y_{j,1}^i$ with probability less than $\bar{p}_i$. That is,

$$Y_{i,1}^i \equiv \left\{ y_i \in Y_i : \Pr(\{y_j \in Y_{j,1}^i\} \mid a^{\text{p.r.}(i)}, y_i) > \bar{p}_i \right\} \tag{6.2}$$

$$Y_{i,2}^i \equiv \left\{ y_i \in Y_i : \Pr(\{y_j \in Y_{j,1}^i\} \mid a^{\text{p.r.}(i)}, y_i) < \bar{p}_i \right\}. \tag{6.3}$$

[1] p.r. stands for "public randomization."

A sufficient condition to dispense with public randomization is that there exist $Y^i_{j,1}$ and $Y^i_{j,2}$ with $Y^i_{j,1} \cup Y^i_{j,2}$ and $\bar{p}_i \in (0,1)$ such that $Y^i_{i,1}$ and $Y^i_{i,2}$ are non-empty partitions of Y_i. Note that this is generic if the players' signals are correlated given some action profile $a^{\text{p.r.}(i)} \in A$.

6.1.2.2 Conditionally Independent Signals

We offer another sufficient condition that works with conditionally independent signals. In this assumption, each player takes a mixed strategy and then she needs to infer the action-signal pair observed by the other player. Even if the signals are conditionally independent, since the other player plays a mixed strategy, her signals contain non-trivial information about what action-signal pair the other player observes.

Specifically, take $i \in I$ arbitrarily. Then, player $n \in I$ (can be the same as player i) takes a mixed strategy $\alpha_n^{\text{p.r.}(i)}$.

Player $j \neq i$ has a function $\phi_j^{\text{p.r.}(i)}$ that maps her action-signal pairs into real numbers in $[0,1]$: $\phi_j^{\text{p.r.}(i)} : A_j \times Y_j \to (0,1)$. As player j constructs $\Gamma_j^{a(x)}$ from $\gamma_j^{a(x)}(y_j)$, player j constructs $\Phi_j^{\text{p.r.}(i)} \in \{0,1\}$ from $\phi_j^{\text{p.r.}(i)}(a_j, y_j)$.

Player i tries to infer whether player j has $\Phi_j^{\text{p.r.}(i)} = 1$ (corresponding to $Y^i_{j,1}$ in the previous assumption) or $\Phi_j^{\text{p.r.}(i)} = 0$ (corresponding to $Y^i_{j,2}$ in the previous assumption). With some $\bar{p}_i \in (0,1)$, player i classifies the set of her action-signal pairs into two classes: the set of action-signal pairs with which player i thinks that player j observes $\Phi_j^{\text{p.r.}(i)} = 1$ with probability more than $\bar{p}_i$ and the set of action-signal pairs with which player i thinks that player j observes $\Phi_j^{\text{p.r.}(i)} = 1$ with probability less than $\bar{p}_i$. That is,

$$H^i_{i,1} \equiv \left\{(a_i, y_i) \in A_i \times Y_i : \Pr(\{\Phi_j^{\text{p.r.}(i)} = 1\} \mid \alpha_j^{\text{p.r.}(i)}, a_i, y_i) > \bar{p}_i\right\} \tag{6.4}$$

$$H^i_{i,2} \equiv \left\{(a_i, y_i) \in A_i \times Y_i : \Pr(\{\Phi_j^{\text{p.r.}(i)} = 1\} \mid \alpha_j^{\text{p.r.}(i)}, a_i, y_i) < \bar{p}_i\right\}. \tag{6.5}$$

A sufficient condition to dispense with public randomization is that there exist $\alpha^{\text{p.r.}(i)}$ and $\phi_j^{\text{p.r.}(i)}$ such that

- $H^i_{i,1}$ and $H^i_{i,2}$ are non-empty partitions of $A_i \times Y_i$;

- given $\alpha_j^{\text{p.r.}(i)}$, the probability that player i observes $y_i \in Y_i$ with $(a_i, y_i) \in H^i_{i,1}$ is independent of $a_i \in A_i$; and

- given $\alpha_i^{\text{p.r.}(i)}$, the probability that player i observes $(a_i, y_i) \in A_i \times Y_i$ with $(a_i, y_i) \in H^i_{i,1}$ is independent of $a_j \in A_j$.

Note that the last two conditions are new, which implies that the probability that player i is willing to report her history in the report block is constant "for all a_i from player i's perspective" and "for all a_j from player j's perspective," as we will see in Section 6.7. This is important to incentivize them to take a mixed strategy.

This condition is generic under Assumption 6 since $\phi_j^{\text{p.r.}(i)}$ has $|A_j|\,|Y_j|$ degrees of freedom while we have $|A_i| + |A_j|$ constraints.

In summary, the following assumption is sufficient:

Assumption 40 For each $i \in I$, one of the following two conditions is satisfied:

1. there exists $a^{\text{p.r.}(i)} \in A$ such that there exist $Y^i_{j,1}$, $Y^i_{j,2}$, $\bar{p}_i$, $Y^i_{i,1}$ and $Y^i_{i,2}$ such that $Y^i_{i,1}$ and $Y^i_{i,2}$ satisfy (6.2), (6.3) and

$$Y^i_{i,1} \neq \emptyset, Y^i_{i,2} \neq \emptyset, Y_i = Y^i_{i,1} \cup Y^i_{i,2};$$

   or

2. there exists $\alpha^{\text{p.r.}(i)} \in \Delta(A)$ such that there exist $\phi_j^{\text{p.r.}(i)}$, $\bar{p}_i$, $H^i_{i,1}$ and $H^i_{i,2}$ such that

   (a) $H^i_{i,1}$ and $H^i_{i,2}$ satisfy (6.4), (6.5) and

$$H^i_{i,1} \neq \emptyset, H^i_{i,2} \neq \emptyset, A_i \times Y_i = H^i_{i,1} \cup H^i_{i,2};$$

(b) given $\alpha_j^{\text{p.r.}(i)}$, the probability that player i observes $y_i \in Y_i$ with $(a_i, y_i) \in H_{i,1}^i$ is independent of $a_i \in A_i$; and

(c) given $\alpha_i^{\text{p.r.}(i)}$, the probability that player i observes $(a_i, y_i) \in A_i \times Y_i$ with $(a_i, y_i) \in H_{i,1}^i$ is independent of $a_j \in A_j$.

6.1.3 Assumption for Dispensing with the Conditionally Independent Cheap Talk

As mentioned, after we replace the perfect cheap talk in the report block with the conditionally independent noisy cheap talk, we dispense with the conditionally independent cheap talk.

To do so, we want to construct a statistics that preserves the conditional independence property for player 2. Player 2 sends a binary message $m \in \{G, B\}$ by taking $a_2^m \in \{a_2^G, a_2^B\}$. Player 1 takes some mixed action $\alpha_1^{\text{report}} \in \Delta(A_1)$. Based on the realization of the mixture a_1 and signal observation y_1, player 1 calculates $\phi_1(a_1, y_1)$. We want to make sure that, regardless of player 2's signal observation, player 2 believes that player 1 statistically infers player 2's signal properly: there exist $q_2 > q_1$ such that

$$\mathbb{E}\left[\phi_1(a_1, y_1) \mid \alpha_1^{\text{report}}, a_2, y_2\right] = \begin{cases} q_2 & \text{if } a_2 = a_2^G, \\ q_1 & \text{if } a_2 \neq a_2^G \end{cases} \tag{6.6}$$

for all $y_2 \in Y_2$.

A sufficient condition for the existence of such ϕ is as follows:

Assumption 41 One of the following two assumptions is satisfied:

1. the monitoring is conditionally independent: $q(y_{-i} \mid a, y_i) = q(y_{-i} \mid a)$ for all $a \in A$ and $y \in Y$; or

2. there exists $\alpha_1^{\text{report}} \in \Delta(A_1)$ such that $\left(\Pr\left(a_1, y_1 \mid \alpha_1^{\text{report}}, a_2, y_2\right)\right)_{a_1,y_1}$ is linearly independent with respect to (a_2, y_2).

Note that Condition 2 is generic since we assume (6.1) and that we do not assume the counterpart for player 1 to send the message.

The following lemma shows that Assumption 41 is sufficient to have ϕ_1 with (6.6).

Lemma 42 If Assumption 41 is satisfied for $\alpha_1^{\text{report}} \in \Delta(A_1)$, then for all $a_2^G \in A_2$, there exist $q_2 > q_1$ and $\phi_1 : A_1 \times Y_1 \to (0, 1)$ such that (6.6) holds for all $y_2 \in Y_2$.

Proof: The same as Lemma 12. ∎

6.2 Replacing the Perfect Cheap Talk in the Coordination Block with the Error-Reporting Noisy Cheap Talk

Remember that, with the perfect cheap talk, the players communicate about x in the coordination block in the following way: first, player 1 tells x_1 and second, player 2 tells x_2.

We divide the step where player i sends the message about x_i into the following steps (remember that this step is called the "round for x_i" with the perfect cheap talk in Section 2.2):

1. first, player i sends x_i to player j via error-reporting noisy cheap talk with precision $p = 1 - \exp(-T^{\frac{1}{2}})$. Among other things,[2] $f[j](x_i, 1) \in \{G, B\}$ and $g[i](x_i, 1) \in \{x_i, E\}$ are generated. With abuse of notation, instead of x_i, we use $(x_i, 1)$ since, as we will see, player i will re-send the message x_i and we want to distinguish the results of the first message and those of the second message;

[2] Precisely, in addition to $f_2[j](x_i, 1)$ and $g_2[i](x_i, 1)$.

2. second, player i sends x_i to player j via error-reporting noisy cheap talk with precision $p = 1 - \exp(-T^{\frac{2}{3}})$. Among other things,[3] $f[j](x_i, 2) \in \{G, B\}$ and $g[i](x_i, 2) \in \{x_i, E\}$ are generated.

It is important to realize that the precision is higher in the second step. Given these two steps, player j constructs an inference of x_i, $x_i(j) \in \{G, B\}$; and

3. third, player j sends $x_i(j)$ to player i via error-reporting noisy cheap talk with precision $p = 1 - \exp(-T^{\frac{1}{2}})$. Among other things, $f[i](x_i(j)) \in \{G, B\}$ and $g[j](x_i(j)) \in \{x_i, E\}$ are generated.

Given these three steps, player i constructs an inference of x_i, $x_i(i) \in \{G, B\}$. Each player $n \in \{1, 2\}$ plays the continuation game as if x_i were $x_i(n)$ in Chapter 4.

In addition, after some events, player i (player j, respectively) makes player j (player i, respectively) indifferent between any action profile sequence in the main blocks by using $\pi_j^{x_i}[a_{i,t}](y_{i,t})$ $(\pi_i^{x_j}[a_{j,t}](y_{j,t})$, respectively) for π_j^{main} $(\pi_i^{\text{main}}$, respectively).

Intuitively, the coordination goes as follows: if player i observes $g[i](x_i, 2) = x_i$, then with a high probability, player i infers $x_i(i) = x_i$. With a small probability, however, player i uses the signal from player j's message: $x_i(i) = f[i](x_i(j))$. In addition, if the latter is the case, then player i makes player j indifferent between any action profile in the main blocks. If $g[i](x_i, 2) \neq x_i$, then player i uses the signal from player j's message: $x_i(i) = f[i](x_i(j))$.

On the other hand, player j uses the signals from the second message from player i and $x_i(j) = f[j](x_i, 2)$ with a high probability. With a small probability, however, player j uses the signal from the first message: $x_i(j) = f[j](x_i, 1)$. In addition, if the latter is the case, then player j makes player i indifferent between any action profile in the main blocks.

Consider player i's inference. If player i uses $f[i](x_i(j))$, then since player j's continuation action plan is independent of $g[j](x_i(j))$, 2 of Lemma 10 implies that player i can

[3]Precisely, in addition to $f_2[j](x_i, 2)$ and $g_2[i](x_i, 2)$.

always believe that player i's inference is correct or player j knows the mistake with a high probability.

If player i adheres to x_i after $g[i](x_i, 2) = x_i$, then player i before observing player j's continuation action plan believes that $f[j](x_i, 2) = g[i](x_i, 2)$ by 3 of Lemma 10 with a high probability. Hence, player i believes that $x_i(j) = f[j](x_i, 2) = g[i](x_i, 2)$ or any action profile is optimal. When player i realizes that $x_i(j) \neq g[i](x_i, 2)$ from learning, player i believes that player j uses $f[j](x_i, 1)$ rather than $f[j](x_i, 2)$. Here is where we use the assumption that the precision of the second message is higher than the first message. Since the precision of the second message is higher than the first message, player i after observing $x_i(j)$ contradictory to player i's expectation from the second message believes that player j uses the first message (this happens with a positive probability) and that there was an error in the first message. Remember that player j makes player i indifferent between any action profile in the main blocks if player j uses the first message. Therefore, after observing $x_i(j)$ contradictory to x_i, player i believes that any action is optimal with a high probability.

Therefore, player i is willing to obey the same strategy as in the case with the perfect cheap talk with x_i replaced with $x_i(i)$.

Consider player j's inference. If player j uses the first message: $x_i(j) = f[j](x_i, 1)$, then since player i's continuation action plan is independent of $g[i](x_i, 1)$, 2 of Lemma 10 implies that this is almost optimal.

Suppose that player j uses $f[j](x_i, 2)$ and realizes $x_i(i) \neq f[j](x_i, 2)$ from learning. There are two possibilities: $g[i](x_i, 2) = x_i(i) \in \{G, B\}$ or player i used player j's message and $f[i](x_i(j)) = x_i(i)$ (there is an error). Since the precision of player i's second message is higher than player j's message, player j believes that player i uses player j's message (this happens with a positive probability) and that there was an error in player j's message. Remember that player i makes player j indifferent between any action profile in the main blocks if player i uses player j's message. Therefore, after observing $x_i(i)$ contradictory to $f[j](x_i, 2)$, player j believes that any action is optimal with a high probability.

Verify that $g[i](x_i, 2) = x_i$ and that player i adheres to x_i with a high probability regardless of player j's message and so player j cannot manipulate $x_i(i)$.

The following lemma formalizes the argument:

Lemma 43 We can define $(x_i(1), x_i(2))_{i \in I}$ and the events that a player makes her opponent indifferent between any action profile sequence such that, conditional on $x \in \{G, B\}^2$, for each $i \in I$, the inferences in the coordination block satisfy the following:

1. given the true state x_i and player j's inference $x_i(j)$, player i puts a belief no less than $1 - \exp(-\Theta(T^{\frac{1}{2}}))$ on the events that $x_i(j) = x_i(i)$ or player i is indifferent between any action profile;

2. given the true state x_i and player i's inference $x_i(i)$, player j puts a belief no less than $1 - \exp(-\Theta(T^{\frac{1}{2}}))$ on the events that $x_i(i) = x_i(j)$ or player j is indifferent between any action profile; and

3. it is almost optimal for the players to send the messages truthfully.

We first define $x_i(i)$ and $x_i(j)$. Player i constructs $x_i(i)$ as follows:

1. if $g[i](x_i, 2) = x_i$ in the second step, then player i mixes the following two:

   (a) with probability $1-\eta$, $x_i(i) = x_i$. That is, with a high probability, player i adheres to her own state; and

   (b) with probability η, $x_i(i) = f[i](x_i(j))$. That is, with a low probability, player i uses the signal from player j's message; and

2. if $g[i](x_i) = E$ in the second step, then player i always uses the signal from player j's message: $x_i(i) = f[i](x_i(j))$.

For completeness, if player i deviates in the step 1 or 2 of the communication, then player i always uses the signal from player j's message: $x_i(i) = f[i](x_i(j))$.

Player j mixes the following two:

1. with probability $1-\eta$, $x_i(j) = f[j](x_i, 2)$. That is, with a high probability, player j uses the signal from player i's second message; and

2. with probability η, $x_i(j) = f[j](x_i, 1)$. That is, with a low probability, player j uses the signal from player i's first message.

Second, we identify after what history player i (player j, respectively) makes player j (player i, respectively) indifferent between any action profile sequence.

Player i makes player j indifferent between any action profile sequence if (and only if based on the round for x_i)[4] $g[i](x_i, 1) = E$, $g[i](x_i, 2) = E$, or "1-(b) or 2 is the case for the construction of $x_i(i)$."

Player j makes player i indifferent between any action profile sequence if (and only if based on the round for x_i) $g[j](x_i(j)) = E$ or 2 is the case for the construction of $x_i(j)$.

Given the above preparation, we prove the lemma:

Proof of 1 of Lemma 43: If 1-(b) or 2 of the construction of $x_i(i)$ is the case, then 2 of Lemma 10 guarantees the result. Note that player j's continuation action plan in the main blocks does not reveal $g[j](x_i(j))$.

If 1-(a) is the case, then without conditioning on $x_i(j)$, by 3 of Lemma 10, player i puts a belief no less than $1 - \exp(-\Theta(T^{\frac{2}{3}}))$ on the event that $f[j](x_i, 2) = x_i = x_i(i)$. Whenever player j uses $f[j](x_i, 1)$ for $x_i(j)$, player j makes player i indifferent between any action profile sequence. Hence, without conditioning on $x_i(j)$, player i puts the belief no less than $1 - \exp(-\Theta(T^{\frac{2}{3}}))$ on the event that $x_i(j) = x_i(i)$ or player i is indifferent between any action profile.

Suppose that player i learns $x_i(j) \neq x_i(i)$. Remember that with probability η, player j uses the signal of the first message $f[j](x_i, 1)$. Since the precision of the first message is $p = 1 - \exp(-T^{\frac{1}{2}})$, 4 of Lemma 10 implies that player i believes that any $f[j](x_i, 1)$ could

[4]With abuse of notation, for the multimple steps to coordinate on x_i, we use the same notation "the round for x_i" as in the case with the perfect cheap talk. In Section 6.3, we introduce a different notation from the case with the perfect cheap talk.

happen with probability $\exp(-\Theta(T^{\frac{1}{2}}))$ regardless of $g[i](x_i, 1)$ and $g_2[i](x_i, 1)$. Since player i's prior on $f[j](x_i, 2) = x_i(i)$ is $1-\exp(-\Theta(T^{\frac{2}{3}}))$, after learning $x_i(j) \neq x_i(i)$, player i puts a posterior no less than $1-\exp(-\Theta(T^{\frac{2}{3}}))/\exp(-\Theta(T^{\frac{1}{2}})) = 1-\exp(-\Theta(T^{\frac{2}{3}}))$ on the event that player j uses the signal of the first message $f[j](x_i, 1)$ and that $f[j](x_i, 1)$ was wrong. In that event, player j makes player i indifferent between any action profile sequence. Therefore, we are done.

Proof of 2 of Lemma 43: If 2 of the construction of $x_i(j)$ is the case, then 2 of Lemma 10 guarantees the result. Note that $x_i(i)$ never reveals $g[i](x_i, 1)$.

If 1 of the construction of $x_i(j)$ is the case, then 2 of Lemma 10 implies that, without conditioning on $x_i(i)$, player j puts the belief no less than $1 - \exp(-\Theta(T^{\frac{2}{3}}))$ on the events that $x_i(j) = x_i$ or player j is indifferent between any action profile since (i) if $g[i](x_i, 2) = E$, then player j is indifferent between any action profile and (ii) if $g[i](x_i, 2) = x_i$ and player i uses $f[i](x_i(j))$, then player j is indifferent between any action profile.

Suppose that player j learns that $x_i(i) \neq x_i(j)$ and $x_i(i) = x_i$.[5] If player j put a high belief on $x_i(i) = g[i](x_i, 2) = x_i$, then this lemma would not hold. However, with probability η, player i uses the signal from player j's message, $f[i](x_i(j))$. Since the precision of this message is $p = 1 - \exp(-T^{\frac{1}{2}})$, 4 of Lemma 10 implies that player j believes that any $f[i](x_i(j))$ could happen with probability $\exp(-\Theta(T^{\frac{1}{2}}))$ regardless of $g[j](x_i(j))$ and $g_2[j](x_i(j))$. Since player j's prior on the event "$x_i = f[j](x_i, 2)$ or $g[i](x_i, 2) = E$" is $1-\exp(-\Theta(T^{\frac{2}{3}}))$, after learning $x_i(i) \neq x_i(j)$ and $x_i(i) = x_i$, player i puts a posterior no less than $1 - \exp(-\Theta(T^{\frac{2}{3}}))/\exp(-\Theta(T^{\frac{1}{2}})) = 1 - \exp(-\Theta(T^{\frac{2}{3}}))$ on the event that player i uses the result of player j's message $f[i](x_i(j))$ and that $f[i](x_i(j))$ happened to be x_i. In that event, player i makes player j indifferent between any action profile sequence. Therefore, we are done.

[5] In the other cases, either $x_i(i) = x_i(j)$ or "$x_i(i) \neq x_i$ and so player j is indifferent between any action profile in the main blocks."

Proof of 3 of Lemma 43: Let us consider player i's incentive. First, the probability that player j makes player i indifferent is almost independent of player i's strategy: $g[j](x_i(j)) = E$ happens with probability no more than $\exp(-\Theta(T^{\frac{1}{2}}))$ regardless of $x_i(j)$. In addition, whether 1 or 2 is the case for the construction of $x_i(j)$ is determined by player j's own randomization.

Since x_i controls player j's value, not player i's value, player i does not have an incentive to deviate to coordinate on a different x_i. Since 1 of Lemma 43 guarantees that player i can infer player j's inference $x_i(j)$ correctly or player i is indifferent between any action profile, we are done.

Next, we consider player j's incentive. First, the probability that player i makes player j indifferent is independent of player j's strategy: the distribution of $g[i](x_i, 1)$ and $g[i](x_i, 2)$ is independent of player j's strategy. In addition, whether 1-(a) or 1-(b) is the case for the construction of $x_i(i)$ is determined by player i's own randomization.

Second, by 2 of Lemma 43, the equilibrium strategy enables player j to infer player i's inference $x_i(i)$ correctly or player j is indifferent between any action profile with probability no less than $1 - \exp(-\Theta(T^{\frac{1}{2}}))$.

Third, whenever player i uses the signal from player j's message, $f[i](x_i(j))$, 1-(b) or 2 is the case for the construction of $x_i(i)$ and player i makes player j indifferent.

Therefore, the truthtelling incentive for $x_i(j)$ is satisfied.

6.3 Structure of the Review Phase

Replacing the perfect cheap talk in the coordination block with the error-reporting noisy cheap talk, we have the following structure of the review phase. Now, the coordination block has six rounds with the following chronological order:

- the round for $(x_1, 1)$ where player 1 sends x_1 via error-reporting noisy cheap talk with precision $p = 1 - \exp(-T^{\frac{1}{2}})$;

- the round for $(x_1, 2)$ where player 1 sends x_1 via error-reporting noisy cheap talk with precision $p = 1 - \exp(-T^{\frac{2}{3}})$;

- the round for $(x_1, 3)$ where player 2 sends $x_1(2)$ via error-reporting noisy cheap talk with precision $p = 1 - \exp(-T^{\frac{1}{2}})$;

- the round for $(x_2, 1)$ where player 2 sends x_2 via error-reporting noisy cheap talk with precision $p = 1 - \exp(-T^{\frac{1}{2}})$;

- the round for $(x_2, 2)$ where player 2 sends x_2 via error-reporting noisy cheap talk with precision $p = 1 - \exp(-T^{\frac{2}{3}})$; and

- the round for $(x_2, 3)$ where player 1 sends $x_2(1)$ via error-reporting noisy cheap talk with precision $p = 1 - \exp(-T^{\frac{1}{2}})$.

After that, we have L review blocks. For each lth main block with $l = 1, ..., L-1$, there are following seven rounds in the following chronological order:

- the lth review round where the players play the stage game for T periods;

- the supplemental round for $\lambda_1(l+1)$ where player 1 sends $\lambda_1(l+1)$ via error-reporting noisy cheap talk with precision $p = 1 - \exp(-T^{\frac{1}{2}})$;

- the supplemental round for $\lambda_2(l+1)$ where player 2 sends $\lambda_2(l+1)$ via error-reporting noisy cheap talk with precision $p = 1 - \exp(-T^{\frac{1}{2}})$;

- the supplemental round for $d_1(l+1)$ where player 1 sends $d_1(l+1)$ via error-reporting noisy cheap talk with precision $p = 1 - \exp(-T^{\frac{1}{2}})$; and

- the supplemental round for $d_2(l+1)$ where player 2 sends $d_2(l+1)$ via error-reporting noisy cheap talk with precision $p = 1 - \exp(-T^{\frac{1}{2}})$.

The last main block only has the Lth review round where the players play the stage game for T periods.

After that, we have the report block, which will be explained fully in Section 6.7.

As we can see, there is a chronological order for the rounds. Hence, we can number all the rounds serially. For example, the round for $(x_1, 1)$ is round 1, the round for $(x_1, 2)$ is round 2, and so on.

In addition, if we replace the error-reporting noisy cheap talk with precision p with messages via actions, then round r where player i sends the message via error-reporting noisy cheap talk with precision $p = 1 - \exp(-T^k)$ consists of T^k periods. For example, in the round for $(x_1, 1)$, the players play the stage game for $T^{\frac{1}{2}}$ periods.

Finally, let $T(r)$ be the set of periods in round r. As in Chapter 2, player i randomly picks one period $t_i(r)$ from $T(r)$ and then use periods $T_i(r) = T(r) \setminus \{t_i(r)\}$ for the determination of action plans in the coordination and main blocks.

6.4 Dispensing with the Error-Reporting Noisy Cheap Talk

Consider round r where player j sends m via error-reporting noisy cheap talk with precision $p = 1 - \exp(-T^k)$ with $k \in \{1/2, 2/3\}$. We replace the error-reporting noisy cheap talk with precision $1 - \exp(-T^k)$ with messages via actions as follows. Now, round r consists of T^k periods.

With $\rho < \frac{1}{2}$, to send message m, player j takes

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

with

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

for T^k periods. That is, when $z_j(m)$ is not equal to m with probability η, player j tells a lie.

On the other hand, given some finite set $\mathcal{A}_i$ to be determined, player i (receiver) determines $\alpha_i(r)$ as follows:

- $\alpha_i(r) = \bar{\alpha}_i = (1 - 2(|A_i| - 1)\rho) a_i^G + \sum_{a_i \neq a_i^G} 2\rho a_i$ with probability $1-\eta$; and
- for each $\alpha_i \in \mathcal{A}_i$, $\alpha_i(r) = \alpha_i$ with probability $\frac{1}{|\mathcal{A}|}\eta$.

Player i takes $\alpha_i(r)$ for T^k periods. Note that $\bar{\alpha}_i$ is comparable to $\alpha_i(x)$ in the main rounds with $a_i(x)$ replaced with a_i^G.

Our task is to create a mapping from player j's history to $g[j](m) \in \{m, E\}$ and that from player i's history to $f[i](m) \in \{G, B\}$ such that important features of Lemma 10 are satisfied. The mapping from player i's history to $f[i](m)$ cannot depend on m since the receiver does not know the true message.[6]

6.4.1 Formal: $g[j](m) \in \{m, E\}$

Remember that player j makes player i indifferent between any action profile after $g[j](m) = E$. Player j has $g[j](m) = m$ if and only if the following three conditions are satisfied:

1. player j tells the truth: $z_j(m) = m$;

[6] As we will see, $y_{j,t_j(r)}$ is not revealed by player j's continuation play in the main block. Similarly to $g_2[j](m)$ for the error-reporting noisy cheap talk, $y_{j,t_j(r)}$ plays an important role to incentivize the players to tell the truth in the report block.

Symmetrically, $y_{i,t_i(r)}$ is not revealed by player i's continuation play in the main block, which plays a similar role to $f_2[i](m)$.

2. the empirical distribution of $a_{j,t}$'s is close to $\alpha_j(r) = \alpha_j^m$ (note that if $\alpha_j(r) \neq \alpha_j^m$, then Condition 1 is not the case):

$$\left\| \frac{1}{T^k - 1} \sum_{t \in T_j(r)} \mathbf{1}_{a_{j,t}} - \alpha_j^m \right\| < \varepsilon \tag{6.7}$$

with ε being a small number to be determined; and

3. player j's signal frequency in the periods where player j takes a_j^m in $T_j(r)$ is close to the affine hull of player j's signal distributions with respect to player i's action.

Otherwise, $g[j](m) = E$.

We are left to define Condition 3. First, define $T_j(l,m)$, $\boldsymbol{Q}_j(m)$, $Q_j(m)$, $\mathbf{q}_j(m)$ and $\mathbf{1}_{Q_j(m)}$ as $T_j(l,x)$, $\boldsymbol{Q}_j(x)$, $Q_j(x)$, $\mathbf{q}_j(x)$ and $\mathbf{1}_{Q_j(x)}$ with $a_j(x)$ replaced with a_j^m.

Condition 3 is satisfied if

- $Q_j(m)\left(\frac{1}{|T_j(l,m)|}\sum_{t\in T_j(l,m)} \mathbf{1}_{y_{j,t}}\right)$ and $\frac{1}{|T_j(l,m)|}\sum_{t\in T_j(l,m)} \mathbf{1}_{Q_j(m)}$ are close:

$$\left\| Q_j(m)\left(\frac{1}{|T_j(l,m)|}\sum_{t\in T_j(l,m)} \mathbf{1}_{y_{j,t}}\right) - \frac{1}{|T_j(l,m)|}\sum_{t\in T_j(l,m)} \mathbf{1}_{Q_j(m)} \right\| < \frac{\varepsilon}{K_1}. \tag{6.8}$$

  As we have adjusted the probability of (2.22) in Section 2.4.4.2, we adjust the probability of (6.8) so that the probability of (6.8) is independent of $\{a_{j,t}, y_{j,t}\}_{t\in T(l)}$. When we say (6.8) is satisfied, we take this adjustment into account; and

- $\frac{1}{|T_j(l,m)|}\sum_{t\in T_j(l,m)} \mathbf{1}_{Q_j(m)}$ and $\mathbf{q}_j(m)$ are close:

$$\left\| \frac{1}{|T_j(l,m)|}\sum_{t\in T_j(l,m)} \mathbf{1}_{Q_j(m)} - \mathbf{q}_j(m) \right\| < \frac{\varepsilon}{K_1}.$$

As for (2.22) and (2.23), we take K_1 so that Condition 3 implies

$$\left\| \boldsymbol{Q}_j(m) - \frac{1}{|T_j(l,m)|} \sum_{t \in T_j(l,m)} \mathbf{1}_{Q_j(m)} \right\| < \varepsilon.$$

6.4.2 Formal: $f[i](m) \in \{G, B\}$

Player i constructs $f[i](m)$ based only on $\{a_{i,t}, y_{i,t}\}_{t \in T_i(r)}$. Let $f_i(a_i, r)$ be the frequency of taking a_i in $T_i(r)$ and $\mathbf{y}_i(a_i, r)$ be the frequency of player i's signals in $T_i(r)$ while taking a_i. Since player j takes *i.i.d.* strategies,

$$h_i(r) = \left(f_i\left(a_i, r\right), \mathbf{y}_i\left(a_i, r\right)\right)_{a_i}$$

is the sufficient statistics.

Player i infers $z_j(m)$ (creates $f[i](m) \in \{G, B\}$) from the likelihood. Given $m \in \{G, B\}$ and $h_i(r)$, the conditional likelihood ratio between $z_j(m) = z_j \in \{G, B, M\}$ and $z_j(m) = z_j' \in \{G, B, M\}$ is

$$\begin{aligned} & \frac{\Pr\left(z_j(m) = z_j \mid m, h_i(r)\right)}{\Pr\left(z_j(m) = z_j' \mid m, h_i(r)\right)} \\ = \; & \frac{\Pr\left((\mathbf{y}_i(a_i, r))_{a_i} \mid z_j(m) = z_j, (f_i\left(a_i, r\right))_{a_i}\right)}{\Pr\left((\mathbf{y}_i(a_i, r))_{a_i} \mid z_j(m) = z_j', (f_i\left(a_i, r\right))_{a_i}\right)} \frac{\Pr\left(z_j(m) = z_j \mid m\right)}{\Pr\left(z_j(m) = z_j' \mid m\right)}. \end{aligned}$$

The log likelihood for

$$\frac{\Pr\left((\mathbf{y}_i(a_i, r))_{a_i} \mid z_j(m) = z_j, (f_i\left(a_i, r\right))_{a_i}\right)}{\Pr\left((\mathbf{y}_i(a_i, r))_{a_i} \mid z_j(m) = z_j', (f_i\left(a_i, r\right))_{a_i}\right)}$$

is expressed as

$$T^k \log\left(\mathcal{L}(h_i(r), G) - \mathcal{L}(h_i(r), B)\right)$$

with

$$\log \mathcal{L}(h_i(r), z_j) = \sum_{a_i} f_i(a_i, r) \mathcal{L}(\mathbf{y}_i(a_i, r), a_i, z_j)$$

and

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

From Assumption 39, $\mathcal{L}(h_i, z_j)$ is strictly concave with respect to the mixture of a_j^G and a_j^B for all $h_i(r)$ and there exists $\kappa > 0$ such that one of the following is true (this is formally shown in the proof of Lemma 44):

1. $z_j(m) = G$ is sufficiently more likely than $z_j(m) = B$: $\mathcal{L}(h_i, G) - \kappa \geq \mathcal{L}(h_i, B)$;
2. $z_j(m) = B$ is sufficiently more likely than $z_j(m) = G$: $\mathcal{L}(h_i, B) - \kappa \geq \mathcal{L}(h_i, G)$; or
3. if neither of Conditions 1 nor 2 is satisfied, then since $\mathcal{L}(h_i, z_j)$ is strictly concave, $z_j(m) = M$ is most likely: $\mathcal{L}(h_i, M) - \kappa \geq \mathcal{L}(h_i, G), \mathcal{L}(h_i, B)$.

Suppose that 1 is the case. This means that

$$\frac{\Pr\left(z_j(m) = G \mid m, (f_i(a_i, r), \mathbf{y}_i(a_i, r))_{a_i}\right)}{\Pr\left(z_j(m) = B \mid m, (f_i(a_i, r), \mathbf{y}_i(a_i, r))_{a_i}\right)} \geq \exp(\kappa T^k) \frac{\eta/2}{1-\eta}$$

for all $m \in \{G, B\}$. Remember that $z_j(m) = M$ implies that player j told a lie and that $g[j](m) = E$. Hence, given any $m \in \{G, B\}$, player i puts a conditional belief no less than $1 - \exp(\Theta(-T^k))$ on the event that $m = G$ or $g[j](m) = E$.

Similarly, if 2 is the case, then given any $m \in \{G, B\}$, player i puts a conditional belief no less than $1 - \exp(\Theta(-T^k))$ on the event that $m = B$ or $g[j](m) = E$.

Finally, if 3 is the case, then given any $m \in \{G, B\}$, player i puts a conditional belief no less than $1 - \exp(\Theta(-T^k))$ on the event that $g[j](m) = E$. In this case, player i can infer m arbitrarily for almost optimality.

Hence, using the likelihood, there exists a mapping from $\{a_{i,t}, y_{i,t}\}_{t \in T_i(r)}$ to $f[i](m) \in \{G, B\}$ such that, given any $m \in \{G, B\}$, player i puts a conditional belief more than $1 - \exp(\Theta(-T^k))$ on the event that $m = f[i](m)$ or $g[j](m) = E$.

We state the above discussion formally:

Lemma 44 If Assumption 39 is satisfied, then for any $j \in I$, there exists $\kappa > 0$ such that, for all $\rho \in \left[0, \frac{1}{3}\right]$, one of the following three is correct:

1. $z_j(m) = G$ is sufficiently more likely than $z_j(m) = B$ for all $m \in \{G, B\}$: $\mathcal{L}(h_i, G) - \kappa \geq \mathcal{L}(h_i, B)$;

2. $z_j(m) = B$ is sufficiently more likely than $z_j(m) = G$ for all $m \in \{G, B\}$: $\mathcal{L}(h_i, B) - \kappa \geq \mathcal{L}(h_i, G)$; or

3. if neither of Conditions 1 nor 2 is satisfied, then $z_j(m) = M$ is most likely for all $m \in \{G, B\}$:
$$\mathcal{L}(h_i, M) - \kappa \geq \mathcal{L}(h_i, G), \mathcal{L}(h_i, B).$$

Proof: From the above discussion, it suffices to show that there exists $\kappa > 0$ such that, with

$$\mathcal{L}(h_i(r), z_j, z_j') \equiv \mathcal{L}(h_i(r), z_j) - \mathcal{L}(h_i(r), z_j'),$$

for any $h_i(r)$, one of the following is true:

1. $\mathcal{L}(h_i(r), G, B) \geq \kappa$;

2. $\mathcal{L}(h_i(r), B, G) \geq \kappa$; or

3. $\mathcal{L}(h_i(r), M, G) \geq \kappa$ and $\mathcal{L}(h_i(r), M, B) \geq \kappa$.

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

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

Then,

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

for any $\mathbf{y}_i$ and a_i because of Assumption 39. Hence, $\mathcal{L}(\mathbf{y}_i, a_i, \lambda)$ is strictly concave. Therefore, since $\mathcal{L}(h_i(r), z_j, \tilde{z}_j)$ is the difference in

$$\sum_{a_i} f_i(a_i, r)\, \mathcal{L}(\mathbf{y}_i, a_i, \lambda),$$

we have

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

for all $\rho \in \left[0, \frac{1}{3}\right]$.

Since (i) LHS is continuous in $h_i(r)$ and ρ with Assumption 2 and (ii) $\Delta\left(\{\mathbf{1}_{y_i}\}_{y_i \in Y_i}\right) \ni \mathbf{y}_i$, $\Delta\left(\{\mathbf{1}_{a_i}\}_{a_i \in A_i}\right) \ni \left(f_i(a_i, r)\right)_{a_i}$ and $\left[0, \frac{1}{3}\right] \ni \rho$ are compact, there exists $\kappa > 0$ such that, for all $\rho \in \left[0, \frac{1}{3}\right]$ and $h_i(r)$,

$$\max\left\{\mathcal{L}(h_i(r), G, B), \mathcal{L}(h_i(r), B, G), \min\left\{\mathcal{L}(h_i(r), M, G), \mathcal{L}(h_i(r), M, B)\right\}\right\} > \kappa.$$

By Assumption 2 (full support), neglecting $\left(a_{i,t_i(r)}, y_{i,t_i(r)}\right)$ does not affect the posterior so much. ∎

6.4.3 Formal: $\theta_i(j \to_m i)$

In addition, player i constructs $\theta_i(j \to_m i) \in \{G, B\}$ for a round where player i receives a message m from player j. Intuitively, $\theta_i(j \to_m i) = B$ implies that player i makes player j indifferent between any action profile sequence in the subsequent rounds.

$\theta_i(j \to_m i) = G$ if and only if

1. player i takes $\alpha_i(r) = \bar{\alpha}_i$; and

2. the empirical distribution of $a_{i,t}$'s is close to $\alpha_i(r) = \bar{\alpha}_i$:

$$\left\| \frac{1}{T^k - 1} \sum_{t \in T_i(r)} \mathbf{1}_{a_{i,t}} - \alpha_i^m \right\| < \varepsilon$$

Consider player j who has $g[j](m) = m$. Then, from Section 6.4.1, player j takes α_j^m, the empirical distribution of $a_{j,t}$'s is close to $\alpha_j(r) = \alpha_j^m$, and player j's signal frequency in $T_j(m, r)$ is close to $\boldsymbol{Q}_j(m)$.

If player j's signal frequency is close to the ex ante mean under $\bar{\alpha}_i$, then player j believes that player i also should observe the signal close to the ex ante mean, which implies $f[i](m) = m$ by the consistency of the likelihood estimator.

On the other hand, if player j's signal frequency is far away from the ex ante mean under $\bar{\alpha}_i$, then player j believes that player i takes $\alpha_i(r) \neq \bar{\alpha}_i$ and $\theta_i(j \to_m i) = B$, as in Lemma 15.

In summary, we can show the following lemma:

Lemma 45 For all $\kappa > 0$ satisfying Lemma 44, there exists $\bar{\eta}$ such that, for all $\eta < \bar{\eta}$, there exist $\bar{\rho}$, $\bar{\varepsilon}$ and $\{\mathcal{A}_i\}_{i \in I}$ such that, for all $\rho < \bar{\rho}$ and $\varepsilon < \bar{\varepsilon}$, the above mappings satisfy the following: for all $i, j \in I$,

1. for any $m \in \{G, B\}$, $f[i](m) = m$ with probability $1 - \exp(-\Theta(T^k))$ and $g[j](m) = m$ with probability $1 - \eta - \exp(-\Theta(T^k))$;

2. for any $m \in \{G, B\}$, given m and any $h_i(r)$, player i puts a belief no less than $1 - \exp(-\Theta(T^k))$ on the event that $f[i](m) = m$ or $g[j](m) = E$;

3. for any $m \in \{G, B\}$, given m and any $h_j(r)$, player j with $g[j](m) = m$ puts a belief no less than $1 - \exp(-\Theta(T^k))$ on the event that $f[i](m) = m$ or $\theta_i(j \to_m i) = B$;

4. for any $m \in \{G, B\}$, any $(f[i](m), g[j](m))$ happens with probability at least $\exp(-\Theta(T^k))$;

5. the probability of $g[j](m)$ being E is independent of player i's strategy; and

6. the distribution of $\theta_i(j \to_m i)$ is independent of player j's strategy.

Note that, compared to the error-reporting noisy cheap talk, Condition 3 implies that player j with $g[j](m) = m$ believes that $f[i](m) = m$ or $\theta_i(j \to_m i) = B$ with probability no less than $1 - \exp(-\Theta(T^k))$, instead of believing $f[i](m) = m$. However, since $\theta_i(j \to_m i) = B$ implies that player j is indifferent between any action sequence, the inference defined in Section 6.2 is still almost optimal. In addition, Condition 6 guarantees that player j does not have an incentive to deviate to manipulate $\theta_i(j \to_m i)$. Further, as $f_2[i](m)$ is not revealed in the main blocks, $y_{i,t_i(r)}$ is not revealed by player i's continuation play in the main block. Similarly, as $g_2[j](m)$ is not revealed in the main blocks, $y_{j,t_j(r)}$ is not revealed by player j's continuation play in the main block. This fact will be important to incentivize the players to tell the truth in the report block.

Proof: 1. This follows from the law of large numbers.

2. This follows from Lemma 44.

3. Since the maximum likelihood estimator is consistent, for $h_i(r)$ which is equal to the ex ante frequency

$$\left(\Pr\left(a_i, y_i \mid \bar{\alpha}_i, \alpha_j^m\right)\right)_{a_i, y_i},$$

we should have

$$\log \mathcal{L}(h_i(r), m) - \log \mathcal{L}(h_i(r), m') > 0$$

with $m' = \{G, B\} \diagdown \{m\}$.

Since (i) $\mathcal{L}(h_i(r), m)$ is continuous in $h_i(r)$ and ρ with Assumption 2 and (ii) $\Delta\left(\{\mathbf{1}_{y_i}\}_{y_i \in Y_i}\right) \ni \mathbf{y}_i$, $\Delta\left(\{\mathbf{1}_{a_i}\}_{a_i \in A_i}\right) \ni \left(f_i\left(a_i, r\right)\right)_{a_i}$ and $\left[0, \frac{1}{3}\right] \ni \rho$ are compact, $\mathcal{L}(h_i(r), m)$ is uniform continuous in $h_i(r)$ and ρ. Since $(\Pr(a_i, y_i \mid \bar{\alpha}_i, \alpha_j^m))_{a_i, y_i}$ is also continuous in ρ, there exists $K_1 > 0$ such that, for all $\rho \in \left[0, \frac{1}{3}\right]$, for $h_i(r)$ with

$$h_i(r) = \left(\Pr\left(a_i, y_i \mid \bar{\alpha}_i, \alpha_j^m\right)\right)_{a_i, y_i},$$

$$\log \mathcal{L}(h_i(r), m) - \log \mathcal{L}(h_i(r), m') > K_1.$$

Since $\log \mathcal{L}(h_i(r), m)$ is uniform continuous with respect to $h_i(r)$ and ρ, there exists $\bar{\eta}$ such that, for all $\rho \in \left[0, \frac{1}{3}\right]$, for $h_i(r)$ with

$$\left\| h_i(r) - \left(\Pr\left(a_i, y_i \mid \bar{\alpha}_i, \alpha_j^m\right)\right)_{a_i, y_i} \right\| < \bar{\eta},$$

we have

$$\log \mathcal{L}(h_i(r), m) - \log \mathcal{L}(h_i(r), m') > \frac{K_1}{2}.$$

Fix $\kappa < \frac{K_1}{2}$ so that Lemma 44 holds. Then, player i infers m if

$$\left\| h_i(r) - \left(\Pr\left(a_i, y_i \mid \bar{\alpha}_i, \alpha_j^m\right)\right)_{a_i, y_i} \right\| < \bar{\eta}. \tag{6.9}$$

Given $\bar{\eta}$, there exist $\bar{\rho}_1$ and $\bar{\varepsilon}_1$ such that, for all $\rho < \bar{\rho}_1$ and $\varepsilon < \bar{\varepsilon}_1$, if Conditions 2 and 3 in Section 6.4.1 are satisfied, then if player i's observation of each pair of actions and signals in $T_i\colon (m, r)$ is close to the ex ante mean $\left(\Pr\left(a_i, y_i \mid \bar{\alpha}_i, a_j^m\right)\right)_{a_i, y_i}$, then (6.9) is satisfied: if

$$\left\| \frac{1}{|T_i\colon (m, r)|} \sum_{t \in T_i\colon (m,r)} \mathbf{1}_{a_{i,t}, y_{i,t}} - \left(\Pr\left(a_i, y_i \mid \bar{\alpha}_i, \alpha_j^m\right)\right)_{a_i, y_i} \right\| < \frac{\bar{\eta}}{2}, \tag{6.10}$$

then (6.9) is satisfied. For example, if we take $\bar{\rho}_1 + \bar{\varepsilon}_1 < \bar{\eta}/2$, then the number of periods not included in $T_i\colon (m, r)$ is small enough.

Further, since (i) small ρ implies $\Pr\left(y_i \mid \bar{\alpha}_i, \alpha_j^m\right)$ is almost equal to $\Pr\left(y_i \mid a_i^G, \alpha_j^m\right)$ and (ii) Conditions 1 and 2 in Section 6.4.3 with small ρ imply that player i takes a_i^G very often, (6.10) is satisfied without considering the frequency of actions: there exist $\bar{\rho}_2$ and $\bar{\varepsilon}_2$ such that, for all $\rho < \bar{\rho}_2$ and $\varepsilon < \bar{\varepsilon}_2$, if Conditions 1 and 2 in Section 6.4.3 are satisfied and

$$\left\| \frac{1}{|T_i\colon (m,r)|} \sum_{t \in T_i\colon (m,r)} \mathbf{1}_{y_{i,t}} - \left(\Pr\left(y_i \mid \bar{\alpha}_i, \alpha_j^m\right)\right)_{y_i} \right\| < \frac{\bar{\eta}}{4},$$

then (6.9) is satisfied. In addition, if one of Conditions 1 and 2 in Section 6.4.3 is not satisfied, then $\theta_i(j \to_m i) = B$.

In summary, if

$$\left\| \frac{1}{|T_i\colon (m,r)|} \sum_{t \in T_i\colon (m,r)} \mathbf{1}_{y_{i,t}} - \left(\Pr\left(y_i \mid \bar{\alpha}_i, \alpha_j^m\right)\right)_{y_i} \right\| < \frac{\bar{\eta}}{4},$$

then $f[i](m) = m$ or $\theta_i(j \to_m i) = B$. The rest of the proof is the same as in Lemma 15 with $\frac{\bar{u}}{4}$ replaced with $\frac{\bar{\eta}}{4}$.

4. Given m, any $(y_t)_{t \in T(r)}$ can occur with probability at least

$$\left\{ \min_{y,a} q(y \mid a) \right\}^{T^k}.$$

Assumption 2 (full support) implies that this probability is $\exp(-\Theta(T^k))$.

5. We adjusted the probability of (6.8) so that this probability is independent of $\{a_{j,t}, y_{j,t}\}_{t \in T(l)}$. Hence, the distribution of $g[j](m)$ is determined solely by player j's mixture.

6. The distribution of $\theta_i(j \to_m i)$ is determined solely by player i's mixture. ∎

6.5 Equilibrium Strategies

In this section, we define $\sigma_i(x_i)$ and π_i^{main}.

6.5.1 States

The states $\lambda_j(l+1)$, $\hat{\lambda}_j(l+1)$, $d_i(l+1)$ and $\hat{d}_j(l+1)$ are defined as in Chapter 4 except that x is replaced with $x(i) = (x_1(i), x_2(i))$ defined in Section 6.2.

If we replace the error-reporting noisy cheap talk with messages via actions, then we use $f[i](m)$ (when player i is a receiver) and $g[i](m)$ (when player i is a sender) defined in Section 6.4.

In addition, remember that each player i makes player j indifferent between any action profile sequence if the following events happen in the coordination block: $g[i](x_i, 1) = E$, $g[i](x_i, 2) = E$, $g[i](x_j(i)) = E$, "1-(b) or 2 is the case for the construction of $x_i(i)$," 2 is the case for the construction of $x_j(i)$, or "with m being the message sent by player j via actions, $\theta_i(j \to_m i) = B$ happens."

We create a new state $\theta_j(\mathrm{c}) \in \{G, B\}$ to summarize these events. For player i, if at least one of the events listed above happens, then we say $\theta_j(\mathrm{c}) = B$. Otherwise, $\theta_j(\mathrm{c}) = G$. Note that $\theta_j(\mathrm{c})$ is well defined for the coordination block with and without the error-reporting noisy cheap talk.

Then, we define $\theta_j(l+1) \in \{G, B\}$ as before except that

1. if $\theta_j(\mathrm{c}) = B$, then $\theta_j(1) = B$; and

2. if $\theta_j(i \to_m j) = B$ happens while player i sends the message in the supplemental rounds between the lth review round and the $(l+1)$th review round, then $\theta_j(l+1) = B$.

6.5.2 Player i's Action Plan $\sigma_i(x_i)$

6.5.2.1 With the Error-Reporting Noisy Cheap Talk

In the coordination block, the players play the game as explained in Section 6.2. For the other blocks, $\sigma_i(x_i)$ prescribes the same action with x replaced with $x(i)$ except for the report block. See Section 6.7 for the strategy in the report block.

6.5.2.2 Without the Error-Reporting Noisy Cheap Talk

In a round where player i would send a message m via error-reporting noisy cheap talk with precision p if it were available, the players' strategies are explained in Section 6.4. Here, since player i is sender, reverse i and j: player i (sender) takes $\alpha_i^{z_i(m)}$ and player j (receiver) takes $\bar{\alpha}_j$. $f[j](m) \in \{G, B\}$ and $g[i](m) \in \{m, E\}$ are determined as in Section 6.4.

6.5.3 Reward Function

In this subsection, we explain player j's reward function on player i, $\pi_i^{\text{main}}(x_j, h_j^{\text{main}} : \delta)$.

6.5.3.1 With the Error-Reporting Noisy Cheap Talk

The reward function is the same as in Chapter 4 except that x is replaced with $x(j)$.

6.5.3.2 Without the Error-Reporting Noisy Cheap Talk

Without the error-reporting noisy cheap talk, $\pi_i^{\text{main}}(x_j, h_j^{\text{main}} : \delta)$ is defined by

$$\pi_i^{\text{main}}(x_j, h_j^{\text{main}} : \delta) = \sum_{l=1}^{L} \sum_{t \in T(l)} \pi_i^{\delta}(t, a_{j,t}, y_{j,t}) + \sum_r \pi_i^{\text{main}}(x_j, h_j^{\text{main}}, r : \delta),$$

where $\pi_i^{\text{main}}(x_j, h_j^{\text{main}}, r : \delta)$ is the reward for each round r. Note that we add (2.10) only for the review rounds. For the other rounds where the players communicate, the reward for round r, $\pi_i^{\text{main}}(x_j, h_j^{\text{main}}, r : \delta)$, directly takes discounting into accounting as we will see in (6.11).

For round r corresponding to a review round, the reward function is the same as in the case with the error-reporting noisy cheap talk.

For round r where the players communicate, player j makes player i indifferent between any action profile sequence by

$$\pi_i^{\text{main}}(x_j, h_j^{\text{main}}, r) = \sum_{t \in T(r)} \delta^{t-1} \pi_i^{x_j}[\alpha_j(r)](y_{j,t}). \tag{6.11}$$

6.6 Almost Optimality of $\sigma_i(x_i)$

We first consider player i's incentive to receive a message m. When player i receives the message, Lemma 45 implies that $\bar{\alpha}_i$ gives player i the inference $f[i](m)$ satisfying $f[i](m) = m$ or $g[j](m) = E$ with $1 - \exp(-\Theta(T^{\frac{1}{2}}))$. In addition, the probability of $g[j](m)$ being E is independent of player i's strategy. Since (6.11) cancels out the difference in the instantaneous utilities, it is almost optimal to take $\bar{\alpha}_i$.

Second, we verify player i's incentive to send a message. Consider the rounds for $(x_i, 1)$ or $(x_i, 2)$. With the error-reporting noisy cheap talk, we are done with Lemma 43. Suppose that we replace the error-reporting noisy cheap talk with messages via actions. Remember that if player i deviates in these rounds, then player i uses the signal from player j's message in the round for $(x_i, 3)$. Therefore, from Lemmas 43 and 45, "$x(i) = x(j)$," "$\theta_j(i \to_{x_i} j) = B$ for the round $(x_i, 1)$ or $(x_i, 2)$" or "$\theta_j(\text{c}) = B$" with probability $1 - \exp(-\Theta(T^{\frac{1}{2}}))$ and the coordination on the same inference of x_i is achieved with a high probability, regardless of player i's strategy in the rounds for $(x_i, 1)$ or $(x_i, 2)$. In addition, the distribution of $\theta_j(i \to_{x_i} j) = B$ and $\theta_j(\text{c}) = B$ is independent of player i's strategy from Lemmas 43 and 45. Since x_i controls only player j's payoff, this implies that player i is indifferent between any action plan in the rounds for $(x_i, 1)$ or $(x_i, 2)$.

For the other rounds where player i sends a message, whenever player i's message affects player j's strategy (action or reward), $\theta_j(l) = B$ has been determined with l being the next review round.

Therefore, when player i sends a message, player i is almost indifferent between any action plan.

For the review rounds, from Lemmas 43 and 45, given $x(j)$, for any t in the main blocks and any h_i^t, player i puts a conditional belief no less than $1-\exp(-\Theta(T^{\frac{1}{2}}))$ on the event that $x(j) = x(i)$, $\theta_j(i \to_{x_i} j) = B$ or $\theta_j(\mathrm{c}) = B$. Therefore, the same proof for Proposition 25 works except that now player j makes player i indifferent between any action profile sequence with a higher probability:

1. $\theta_j(\mathrm{c}) = B$ with probability

$$\underbrace{\eta}_{\substack{z_j(x_j)\neq x_j \\ \text{in the round} \\ \text{for } (x_j,1)}} + \underbrace{\eta}_{\substack{z_j(x_j)\neq x_j \\ \text{in the round} \\ \text{for } (x_j,2)}} + \underbrace{\eta}_{\substack{\text{1-(b)} \\ \text{is the case} \\ \text{for } x_j(j)}} + \underbrace{\eta}_{\substack{z_j(x_i(j))\neq x_i(j) \\ \text{in the round} \\ \text{for } (x_i,3)}} + \underbrace{\eta}_{\substack{2 \\ \text{is the case} \\ \text{for } x_i(j)}} \leq 5\eta.$$

   plus negligible probability $\exp(-\Theta(T^{\frac{1}{2}}))$; and

2. for each supplemental rounds where player j sends a message m,

$$\underbrace{\eta}_{z_j(m)\neq m}$$

   plus negligible probability $\exp(-\Theta(T^{\frac{1}{2}}))$.

Therefore, instead of (4.17), we re-take η sufficiently small such that

$$\begin{aligned} & \frac{(L-1)\,v_i\,(B)+\bar{u}}{L} + \underline{u} + 2\frac{\bar{u}}{L} + (5+2L)\,\eta\left(2\bar{u} - \min_{i,x}\frac{(L-1)\,v_i\,(B)+\bar{u}}{L}\right) \\ < \ & \underline{v}_i < \bar{v}_i < \min_{x:x_j=G} w_i(x) - 2\frac{\bar{u}}{L} - (5+2L)\,\eta\left(2\bar{u} + \max_{i,x} w_i\,(x)\right). \end{aligned} \tag{6.12}$$

Finally, since the review round has T period while the other round has at most $T^{\frac{2}{3}}$ periods, the payoffs from the rounds other than the review rounds are negligible. Therefore, Proposition 25 holds.

6.7 Report Block

We are left to construct the report block. First, we explain the report block with the perfect cheap talk and public randomization. Although this is the same setup as in Chapter 2, since we replace the cheap talk in the coordination block and supplemental rounds with messages via actions, we need to change the structure accordingly.

Second, we construct the report block with the perfect cheap talk but without public randomization device.

Third, we replace the perfect cheap talk with conditionally independent noisy cheap talk.

Finally, we replace the conditionally independent noisy cheap talk with messages via actions.

6.7.1 Preparation

Before constructing an equilibrium, let us make three preparations.

First, consider the situation where player 2 sends a binary message to player 1 by taking actions. Suppose player 1 takes α_1^{report} defined in Lemma 42 and calculates $\phi_1(a_{1,t}, y_{1,t})$, defined in Lemma 42. Player 1 constructs $\Phi_{1,t}$ from $\phi_1(a_{1,t}, y_{1,t})$ as she constructs $\Gamma_{1,t}^a$ from $\gamma_1^a(y_{1,t})$. Lemma 42 implies that

$$\Pr\left(\{\Phi_{1,t} = 1\} \mid \alpha_{1,t}^{\text{report}}, a_{2,t}, y_{2,t}\right) = \begin{cases} q_2 & \text{if } a_{2,t} = a_2^G, \\ q_1 & \text{if } a_{2,t} \neq a_2^G \end{cases} \tag{6.13}$$

for all t and $y_{2,t}$. Intuitively, this is important to preserve the properties of the conditionally independent noisy cheap talk when we dispense with it in Section 6.7.5.

Second, consider the situation where player i sends a binary message to player j by taking actions. Suppose player j takes $a_j \in A_j$. Since Assumption 3 guarantees that player j can identify a_i, for each $i \in I$ and $a \in A$, there exist $\psi_j^a : Y_j \to (0,1)$ and $q_2 > q_1$ such that

$$\mathbb{E}\left[\psi_j^a(y_j) \mid \tilde{a}_i, a_j\right] = \begin{cases} q_2 & \text{if } \tilde{a}_i = a_i, \\ q_1 & \text{if } \tilde{a}_i \neq a_i. \end{cases} \tag{6.14}$$

Given such ψ_j^a, player j calculates $\psi_j^a(y_{j,t})$. Player j constructs $\Psi_{j,t}^a$ from $\psi_j^a(y_{j,t})$ as she constructs $\Gamma_{j,t}^a$ from $\gamma_j^a(y_{1,t})$. (6.14) implies that

$$\Pr\left(\{\Psi_{j,t} = 1\} \mid \tilde{a}_i, a_j\right) = \begin{cases} q_2 & \text{if } \tilde{a}_i = a_i, \\ q_1 & \text{if } \tilde{a}_i \neq a_i. \end{cases} \tag{6.15}$$

Note that we do not condition on y_i, that is, the conditional independence property does not necessarily hold.

Third, whenever the players play the stage game, we cancel out the difference in the instantaneous utilities and discounting by adding

$$\delta^{t-1}\pi_i^{x_j}[\alpha_{j,t}](y_{j,t}).$$

Since the report block lasts for $\Theta(T^{\frac{1}{3}})$ periods, this does not affect the equilibrium payoff. From now on, therefore, we ignore the instantaneous utilities.

6.7.2 Report Block with the Perfect Cheap Talk and Public Randomization

We formally construct π_i^{report} assuming that the players send messages via actions in the coordination block and supplemental rounds, keeping the perfect cheap talk and public randomization in the report block.

Remember that r is a serial number of the rounds. Let $\mathcal{A}_j(r)$ be the set of information up to and including round r consisting of

- what state x_j player j is in;
- $x(j)$ if round r is after the coordination block;
- for each lth review round, what action plan $\alpha_j(l)$ player j took in the lth review round if round r is the lth review round; and
- $\hat{\lambda}_i(l) = G$.

Remember that, for each round r, for any period t in round r and any history h_i^t, conditional on $\mathcal{A}_j(r)$, $\sigma_i(x_i)$ is almost optimal.

The reward π_i^{report} is the same as in the case with the cheap talk in the coordination block and supplemental rounds except for the following differences:

Subrounds As we divide a review round into review subrounds whose length is $T^{\frac{1}{4}}$ periods, we divide each round into subrounds whose length is $T^{\frac{1}{4}}$ periods.

Since the rounds for $(x_1, 2)$ and $(x_2, 2)$ have $T^{\frac{2}{3}}$ periods, there are $T^{\frac{2}{3}-\frac{1}{4}}$ subrounds. Since the other rounds for communication have $T^{\frac{1}{2}}$ periods, there are $T^{\frac{1}{4}}$ subrounds.

The coordination on $k(r)$ is analogously modified.

Truthtelling Incentive Since the players communicate via actions, we use (2.42) to give the incentive for player i to tell the truth instead of (2.43) and (2.44). Although the sender j only mixes a_j^G and a_j^B, Condition 2 of Assumption 39 guarantees the incentive. Note that player j's action plan is independent of the signal observation in period $t_j(r)$ and that player i cannot learn it from the history in the coordination and main blocks.

The Rounds where Player i Sends or Receives the Message Note that player i takes a mixed strategy in the round where player i sends or receives the message. Moreover, the history in this round affects the belief about the best responses at the beginning of the next round.

Therefore, as we incentivize player i to take a mixed strategy by using $V_i(\mathfrak{h}_i^r)$ and f_i in the review round, we cancel out the difference of the values coming from the learning at the beginning of the next round. Then, we cancel out the difference in the payoffs in the current round.

Since player i is almost indifferent between any action plan (see Section 6.6), the effect of this adjustment is sufficiently small.

6.7.3 Report Block withOUT the Public Randomization

In this subsection, we keep the availability of the perfect cheap talk and dispense with the public randomization device. We use the public randomization device to coordinate on the following two: first, who reports the history in the report block. Second, for each round r, the picked player sends the message about $(a_{i,t}, y_{i,t})$ for t included in $T(r, k(r))$ for some $k(r)$ determined by the public randomization. We explain how to dispense with the public randomization for each of them.

6.7.3.1 Coordination on Who will Report the History

Remember that the problem is that we want to require that (i) for each i, there is a positive probability that π_i^{report} adjusts the reward function and that (ii) each player should not learn about the opponent's history from the opponent's messages in the report block before player i sends the message in the report block.

Instead of using the public randomization device, we use actions and private signals to coordinate. The specific way of the coordination depends on whether Condition 1 or 2 of Assumption 40 is satisfied.

When Condition 1 of Assumption 40 is satisfied:

1. first, the players take the action profile $a^{\text{p.r.}(2)}$. Then, each player i observes $y_i \in Y_i$;

2. player 2 sends the message whether player 2 observed $y_2 \in Y_{2,1}^2$ or $y_2 \in Y_{2,2}^2$ in Step 1. See Assumption 40 to review the notation. Player 1 adjusts player 1's reward function on player 2 if and only if player 2 says that player 2 observed $y_2 \in Y_{2,1}^2$;

3. player 2 sends the messages first. Player 2 has the following two cases:

   (a) if player 2 observed $y_2 \in Y_{2,1}^2$, then player 2 sends the messages about her history h_2^{main} truthfully; and

   (b) if player 2 observed $y_2 \in Y_{2,2}^2$, then player 2 sends a meaningless message $\{\emptyset\}$; and

4. player 1 sends the messages about her history h_1^{main} truthfully.

When Condition 2 of Assumption 40 is satisfied:

1. first, the players take the action profile $\alpha^{\text{p.r.}(2)}$. Then, each player i observes $a_i, y_i \in A_i \times Y_i$;

2. player 2 sends the message whether player 2 observed $(a_2, y_2) \in H_{2,1}^2$ or $(a_2, y_2) \in H_{2,2}^2$ in Step 1. Player 1 adjusts player 1's reward function on player 2 if and only if player 2 says that player 2 observed $(a_2, y_2) \in H_{2,1}^2$;

3. player 2 sends the messages first. Player 2 has the following two cases:

   (a) if player 2 observed $(a_2, y_2) \in H_{2,1}^2$, then player 2 sends the messages about her history h_2^{main} truthfully; and

   (b) if player 2 observed $(a_2, y_2) \in H_{2,1}^2$, then player 2 sends a meaningless message $\{\emptyset\}$; and

4. player 1 sends the messages about her history h_1^{main} truthfully.

In Step 4, when Condition 1 (Condition 2, respectively) is satisfied in Assumption 40, player 2 adjusts player 2's reward function on player 1 if and only if player 2 observed

$y_2 \in Y^2_{2,2}$ ($(a_2, y_2) \in H^2_{2,1}$, respectively). Therefore, the probability that player j's reward on player i is adjusted is now $\Pr(y_2 \in Y^2_{2,2} \mid a^{\text{p.r.}(2)})$ ($\Pr(a_2, y_2 \in H^2_{2,2} \mid a_1, \alpha_2^{\text{p.r.}(2)})$, respectively) for $i = 1$ and $\Pr(y_2 \in Y^2_{2,1} \mid a^{\text{p.r.}(2)})$ ($\Pr(a_2, y_2 \in H^2_{2,1} \mid a_2, \alpha_1^{\text{p.r.}(2)})$, respectively) for $i = 2$. The term representing

$$\frac{1}{\Pr(\text{player } i \text{ is picked by the public randomization device})}$$

in π_i^{report} of Section 2.9.6 is analogously modified to

$$\begin{aligned} \Pr(y_2 &\in Y^2_{2,2} \mid a^{\text{p.r.}(2)}) \; (\Pr(a_2, y_2 \in H^2_{2,2} \mid a_1, \alpha_2^{\text{p.r.}(2)}), \text{ respectively) for } i = 1, \\ \Pr(y_2 &\in Y^2_{2,1} \mid a^{\text{p.r.}(2)}) \; (\Pr(a_2, y_2 \in H^2_{2,1} \mid a_2, \alpha_1^{\text{p.r.}(2)}), \text{ respectively) for } i = 2. \end{aligned}$$

Consider player 1's incentives. By Assumptions 2 and 40, there is a positive probability that $y_2 \in Y^2_{2,2}$ ($(a_2, y_2) \in H^2_{2,2}$, respectively) and player 2's reward on player 1 is adjusted. In addition, when player 1 sends the message in Step 4, player 1 conditions that player 2's message in Step 3 does not reveal h_2^{main}, as desired.

Therefore, we need to verify the incentives that the players take $a^{\text{p.r.}(2)}$ ($\alpha^{\text{p.r.}(2)}$, respectively) in Step 1 and player 2 tells the truth in Step 2 and 3. To establish the incentives, we add the following rewards:

- in Step 1,
  - when Condition 1 is satisfied in Assumption 40, to incentivize the players to take $a^{\text{p.r.}(2)}$, each player j gives a reward on $a_i^{\text{p.r.}(2)}$:

$$T^{-1}\Psi^{a^{\text{p.r.}(2)}}; \tag{6.16}$$

    and

– when Condition 2 is satisfied in Assumption 40, given the other player's action $\alpha_j^{\text{p.r.}(2)}$, each $a_i \in A_i$ gives the same probability for

$$\begin{aligned}\Pr(a_2, y_2 &\in H_{2,2}^2 \mid a_1, \alpha_2^{\text{p.r.}(2)}) \text{ for } i = 1, \\ \Pr(a_2, y_2 &\in H_{2,1}^2 \mid a_2, \alpha_1^{\text{p.r.}(2)}) \text{ for } i = 2.\end{aligned}$$

Hence, player i is indifferent between all the actions;

- in Step 2, to incentivize player 2 to tell the truth, player 1 punishes player 2 by $-T^{-2}$ if player 2 sends the message $y_2 \in Y_{2,1}^2$ ($(a_2, y_2) \in H_{2,1}^2$, respectively) while player 1 observed $y_1 \in Y_{1,2}^2$ ($\Phi_1^{\text{p.r.}(2)} = 0$, respectively) and punishes player 2 by $-\frac{1}{\bar{p}_2}T^{-2}$ if player 2 sends the message $y_2 \in Y_{2,2}^2$ while player 1 observed $y_1 \in Y_{1,1}^2$ ($\Phi_1^{\text{p.r.}(2)} = 1$, respectively).

  In addition, when Condition 2 is satisfied in Assumption 40, different a_2's can have different expected punishment in this step, given the truthtelling incentive. We cancel out that difference by changing the reward function so that player 2 before observing y_2 is indifferent between any action; and

- in Step 3,

  – if player 2 sent the message $y_2 \in Y_{2,1}^2$ ($(a_2, y_2) \in H_{2,1}^2$, respectively) in Step 2, then player 1's reward on player 2 is the same as π_2^{report} so that player 2 sends the messages about her history h_2^{main} truthfully; and

  – if player 2 sent the message $y_2 \in Y_{2,2}^2$ ($(a_2, y_2) \in H_{2,2}^2$, respectively) in Step 2, then player 1 changes π_2^{report} so that player 1 gives a small reward on $\{\emptyset\}$:

$$\pi_2^{\text{report}}(x_1, h_1^{T_P+1} : \delta) = T^{-3}\mathbf{1}\{\text{player 2 sends } \{\emptyset\}\} - T^{-3}.$$

Then, we can show the incentive compatibility of the above strategy by backward induction.

- in Step 3, since player 2 believes that the message in Step 2 was transmitted correctly, it is optimal to tell the truth by the same reason as in Section 2.9.6;

- in Step 2, since all the rewards affected by player i's current and future actions are bounded by $\Theta(T^{-3})$,[7] if

$$\Pr(\{y_1 \in Y_{1,1}^2\} \mid a_1^{\text{p.r.}(2)}, y_2) > \frac{\bar{L}T^{-2} + \Theta\left(T^{-3}\right)}{\frac{1}{\bar{p}_2}\bar{L}T^{-2} - \Theta\left(T^{-3}\right)} \to \bar{p}_2$$

or

$$\Pr(\{\Phi_1^{\text{p.r.}(2)} = 1\} \mid \alpha_1^{\text{p.r.}(2)}, a_2, y_2) > \frac{\bar{L}T^{-2} + \Theta\left(T^{-3}\right)}{\frac{1}{\bar{p}_2}\bar{L}T^{-2} - \Theta\left(T^{-3}\right)} \to \bar{p}_2,$$

then it is optimal to send $y_2 \in Y_{2,1}^2$ ($(a_2, y_2) \in H_{2,1}^2$, respectively) and if

$$\Pr(\{y_1 \in Y_{1,1}^2\} \mid a_1^{\text{p.r.}(2)}, y_2) < \frac{\bar{L}T^{-2} - \Theta\left(T^{-3}\right)}{\frac{1}{\bar{p}_2}\bar{L}T^{-2} + \Theta\left(T^{-3}\right)} \to \bar{p}_2$$

or

$$\Pr(\{\Phi_1^{\text{p.r.}(2)} = 1\} \mid \alpha_1^{\text{p.r.}(2)}, a_2, y_2) < \frac{\bar{L}T^{-2} + \Theta\left(T^{-3}\right)}{\frac{1}{\bar{p}_2}\bar{L}T^{-2} - \Theta\left(T^{-3}\right)} \to \bar{p}_2,$$

then it is optimal to send $y_2 \in Y_{2,2}^2$ ($(a_2, y_2) \in H_{2,2}^2$, respectively), as desired; and

- in Step 1, if Condition 1 of Assumption 40 is satisfied, since all the rewards affected by player i's action except for (6.16) are bounded by $\Theta(T^{-2})$, it is strictly optimal to take $a^{\text{p.r.}(2)}$. If Condition 2 of Assumption 40 is satisfied, then the future strategy and the reward will be the same for all the actions except for the punishment coming from Step 2. As we have seen, we cancel out the differences in the punishment in Step 2 so that each player is indifferent between any action.

[7] For the punishment in Step 5 in the coordination on $k(r)$ below, there is a punishment of order T^{-2}. However, as we will see, by backward induction, this punishment is not affected by Step 2 here.

Remember that player 1 rewards player 2 in Step 3 based on player 2's message about y_2 ((a_2, y_2), respectively), not depending on player 1's history (a_1, y_1). Therefore, after sending the message about y_2 ((a_2, y_2), respectively) in Step 2, it is optimal for player 2 to follow the equilibrium strategy.

6.7.3.2 Coordination on $k(r)$

For each player i, while player i sends the messages about h_i^{main}, for each round r, the players coordinate on $k(r) \in \{1, \ldots, K\}$ with $K \leq T^{\frac{3}{4}}$.

By abuse of language, in our equilibrium,

- in Step 3 in Section 6.7.3.1, even if player 2 sends $\{\emptyset\}$, the players play the following game for each round and player 1 punishes player 2; and
- in Step 4 in Section 6.7.3.1, even if player 2 sent h_2^{main}, the players play the following game for each round and player 2 punishes player 1.

We create a mapping between a sequence of $\{1, 2\}$, $\boldsymbol{i} \in \{1, 2\}^{\log_2 K}$, and $k(r)$ such that each $\boldsymbol{i}$ uniquely identifies $k(r)$ and that, for each $k(r)$, there is at least one $\boldsymbol{i}$ that is mapped into $k(r)$.

For each $n \in \{1, ..., \log_2 K\}$, the players coordinate on one element of $\{1, 2\}$ as in Steps 1, 2 and 3 in Section 6.7.3.1. That is, when Condition 1 (Condition 2, respectively) is satisfied in Assumption 40,

1. the players take $a^{\text{p.r.}(i)}$ ($\alpha^{\text{p.r.}(i)}$, respectively) for $\log_2 K$ times;
2. for each $n \in \{1, ..., \log_2 K\}$, if player j observes $y_j \in Y_{j,1}^i$ ($\Phi_j^{\text{p.r.}(i)} = 1$, respectively), then player j infers that the nth element of $\boldsymbol{i}$ is 1. Otherwise, that is, if player j observes $y_j \in Y_{j,2}^i$ ($\Phi_j^{\text{p.r.}(i)} = 0$, respectively), then player j infers that the nth element of $\boldsymbol{i}$ is 2. By doing so, player j infers $\boldsymbol{i}$. Let $\boldsymbol{i}(j)$ be player j's inference. Let $k(r, j)$ be player j's inference of $k(r)$ that corresponds to $\boldsymbol{i}(j)$;

3. on the other hand, for each $n \in \{1, ..., \log_2 K\}$, player i infers $\boldsymbol{i}(i)$ and $k(r,i)$ using the partitions $Y^i_{i,1}$ and $Y^i_{i,2}$ ($H^i_{i,1}$ and $H^i_{i,2}$, respectively);

4. player i sends the sequence of binary messages $\boldsymbol{i}(i) \in \{1,2\}^{\log_2 K}$;

5. for each $n \in \{1, ..., \log_2 K\}$, player j punishes player i if player i's message $\boldsymbol{i}_n(i)$ is different from $\boldsymbol{i}_n(j)$. Here, $\boldsymbol{i}_n(i)$ and $\boldsymbol{i}_n(j)$ are the nth element of $\boldsymbol{i}(i)$ and $\boldsymbol{i}(j)$, respectively.

   Specifically, player j punishes player i by $-T^{-2}$ if player i sends the message $\boldsymbol{i}_n(i) = 1$ while player j observed $\boldsymbol{i}_n(j) = 2$ and by $-\frac{1}{\bar{p}_i}T^{-2}$ if player i sends the message $\boldsymbol{i}_n(i) = 2$ while player j observed $\boldsymbol{i}_n(j) = 1$.

   Again, if Condition 2 is satisfied in Assumption 40, we adjust the reward function to cancel out the difference in the ex ante punishment, given the truthtelling incentive.

   Note that this is the same as in Step 2 in Section 6.7.3.1; and

6. from the message $\boldsymbol{i}(i)$, player j knows $k(r,i)$. Player j calculates the punishment (2.47) based on $k(r,i)$.

By backward induction, we can show that it is always optimal to follow the equilibrium strategy: consider the message for the last element of the sequence, $\boldsymbol{i}_K(i)$, for the last round. Since the punishment from the previous messages about $\boldsymbol{i}(i)$ is sunk and the reward or punishment affected by player i's continuation strategy except for the punishment coming from $\boldsymbol{i}_K(i) \neq \boldsymbol{i}_K(j)$ is $\Theta(T^{-3})$, if

$$\Pr(\{y_j \in Y^i_{j,1}\} \mid a^{\text{p.r.}(i)}, y_i) > \frac{\bar{L}T^{-2} + \Theta(T^{-3})}{\frac{1}{\bar{p}_i}\bar{L}T^{-2} - \Theta(T^{-3})} \to \bar{p}_i$$

or

$$\Pr(\{\Phi^{\text{p.r.}(i)}_i = 1\} \mid \alpha^{\text{p.r.}(i)}_j, a_i, y_i) > \frac{\bar{L}T^{-2} + \Theta(T^{-3})}{\frac{1}{\bar{p}_i}\bar{L}T^{-2} - \Theta(T^{-3})} \to \bar{p}_i,$$

then it is optimal to send $y_i \in Y_{i,1}^i$ ($(a_i, y_i) \in H_{i,1}^i$, respectively) and if

$$\Pr(\{y_j \in Y_{j,1}^i\} \mid a^{\text{p.r.}(i)}, y_i) < \frac{\bar{L}T^{-2} - \Theta(T^{-3})}{\frac{1}{\bar{p}_i}\bar{L}T^{-2} + \Theta(T^{-3})} \to \bar{p}_i,$$

or

$$\Pr(\{\Phi_i^{\text{p.r.}(i)} = 1\} \mid \alpha_j^{\text{p.r.}(i)}, a_i, y_i) < \frac{\bar{L}T^{-2} + \Theta(T^{-3})}{\frac{1}{\bar{p}_i}\bar{L}T^{-2} - \Theta(T^{-3})} \to \bar{p}_i,$$

then it is optimal to send $y_i \in Y_{i,2}^i$ ($(a_i, y_i) \in H_{i,2}^i$, respectively), as desired.

Given this, the players have an incentive to take $a^{\text{p.r.}(i)}$ since the reward on $a^{\text{p.r.}(i)}$ is sufficiently large if Condition 1 is satisfied in Assumption 40. If Condition 2 is satisfied in Assumption 40, then since (i) all a_i gives the same probability of $(a_i, y_i) \in H_{i,1}^i$, (ii) player i's continuation payoff only depends on whether $(a_i, y_i) \in H_{i,1}^i$ or not, and (iii) given the truthtelling incentive, we have adjusted the reward function to cancel out the difference in the ex ante punishment, it is optimal to take $\alpha^{\text{p.r.}(i)}$.

For the second $\boldsymbol{i}_{K-1}(i)$, since the expected punishment from the last message $\boldsymbol{i}_K(i)$ is fixed by the equilibrium strategy, the same argument holds. Recursively, we can show the optimality of the equilibrium strategy.

Although player j punishes player i for mis-coordination in Step 5, when player j calculates (2.47), player j uses player i's inference of $k(r)$, $k(r,i)$. Hence, once player i sends the messages about $\boldsymbol{i}(i)$, player i has the incentive to tell the truth about h_i^{main} based on her own inference $k(r,i)$.

Expected Punishment As we have mentioned, when the players coordinate on whether player 2 should send the message about h_2^{main}, player 1 rewards player 2 in Step 3 based on player 2's message about y_2 (a_2, y_2, respectively). In addition, when the players coordinate on $k(r)$, player j uses $k(r,i)$ to calculate (2.47) and the term $T^{\frac{3}{4}}$ in (2.47) is replaced by

$$\frac{1}{\Pr(k(r,i) \text{ is realized in the coordination explained in Section 6.7.3.2})}.$$

Therefore, given the truthtelling incentive during the coordination on who will report the history and which is $k(r)$ (which has been verified), the expected punishment from the coordination is independent of the players' history in $\left(h_1^{\text{main}}, h_2^{\text{main}}\right)$. Hence, this coordination in the report block does not affect any incentive in the coordination and main blocks.

In addition, this also implies that, as mentioned in footnote 7 of Chapter 6, when the players coordinate on who will report the history in Step 2 of Section 6.7.3.1, we can assume that the expected punishment from $\boldsymbol{i}_n(i) \neq \boldsymbol{i}_n(j)$ is fixed.

6.7.4 Report Block with Conditionally Independent Cheap Talk

In this subsection, we replace the perfect cheap talk with conditionally independent noisy cheap talk. That is, each player has the conditionally independent noisy cheap talk communication device to send a binary message $m \in \{G, B\}$. When player i sends the message m, the receiver (player j) observes the correct message m with high probability $1 - \exp(-T^{\frac{1}{3}})$ while player j observes the erroneous message $\{G, B\} \setminus \{m\}$ with low probability $\exp(-T^{\frac{1}{3}})$. Player i (sender) does not obtain any information about what message player j receives (conditional on m). Hence, the communication is conditionally independent.

When player j (receiver) constructs π_i^{report}, player j needs to take care of the possibility that player j receives an erroneous message.

First, the expected punishment for (2.47) and (2.48) together with Π_i now depends on what messages player i will send for $\{a_{i,t}, y_{i,t}\}_{t \in T(r,k(r,i))}$ since each $\{a_{i,t}, y_{i,t}\}_{t \in T(r,k(r,i))}$ has different probabilities of mis-transmission. However, the total expected punishment from a round could be calculated by $\{\#_i^r(k)\}_{k,r}$ if $\{\#_i^r(k)\}_{k,r}$ transmitted correctly. Since the expected punishment is sufficiently small under the truthtelling with conditionally independent cheap talk, player j can make all the $\{\#_i^r(k)\}_{k,r}$'s indifferent in terms of (2.47) and (2.48) together with Π_i. Then, the effect of the probability of making error for $\{a_{i,t}, y_{i,t}\}_{t \in T(r,k(r,i))}$ is canceled out.

Second, let us consider the other messages. Note that the number of binary messages sent in the report block is $\Theta(\log_2 T)$ since we exclude $\{a_{i,t}, y_{i,t}\}_{t\in T(r,k(r,i))}$. Therefore, all the messages transmit correctly with probability at least

$$1 - \Theta(\log_2 T)\exp(-T^{\frac{1}{3}}).$$

In addition, the cardinality of the information sent by all the messages is $\Theta(T)$. Let $\mathcal{M}_i$ be the set of information possibly sent by player i in the report block with $|\mathcal{M}_i| = \Theta(T)$. Let P_i be the $|\mathcal{M}_i| \times |\mathcal{M}_i|$ matrix whose (k, k') element represents

$$\Pr\left(\begin{array}{c}\text{player } j \text{ receives the message corresponding to the element } k' \text{ of } \mathcal{M}_i \\ \mid \text{player } i \text{ sends the message corresponding to the element } k \text{ of } \mathcal{M}_i\end{array}\right).$$

Since $|\mathcal{M}_i| = \Theta(T)$ and all the messages transmit correctly with probability no less than $1 - \Theta(\log_2 T)\exp(-T^{\frac{1}{3}})$,

$$\left(1 - \Theta(\log_2 T)\exp(-T^{\frac{1}{3}})\right)^{\Theta(T)} \geq 1 - \Theta(T)\Theta(\log_2 T)\exp(-T^{\frac{1}{3}}) \to 1$$

as T goes to infinity and so

$$\lim_{\delta\to 1} P_i^{-1} = \lim_{T\to\infty} P_i^{-1} = E \text{ (identity matrix)}.$$

Let

$$\pi_i^{\text{report}}(x_j, h_j^{T_P+1}, k : \delta)$$

be the reward function that player j with $h_j^{T_P+1}$ would construct after the history corresponding to the element k via perfect cheap talk, that is, if P_i were E.[8] In addition, let $\boldsymbol{\pi}_i^{\text{report}}(x_j, h_j^{T_P+1} : \delta)$ be the vector stacking all $\pi_i^{\text{report}}(x_j, h_j^{T_P+1}, k : \delta)$'s with respect to k.

[8]Here, we use $h_j^{T_P+1}$ instead of h_j^{main} since player j needs to use the signal observations while the players coordinate on who will report the history and $k(r)$.

If player j uses kth element of

$$P_i^{-1}\boldsymbol{\pi}_i^{\text{report}}(x_j, h_j^{T_P+1} : \delta)$$

when player j receives kth element of $\mathcal{M}_i$ from player i, then player i's incentive is the same as in the situation that the messages would always transmit correctly (as if P_i were E). Since the truthtelling incentive is strict, multiplying P_i^{-1} to $\boldsymbol{\pi}_i^{\text{report}}(x_j, h_j^{T_P+1} : \delta)$ does not affect the incentives in the report block if P_i^{-1} is sufficiently close to E.

6.7.5 Report Block withOUT the Conditionally Independent Cheap Talk

Finally, we dispense with the conditionally independent cheap talk. Consider Step 2 of Section 6.7.3.1. If player 2 wants to send $y_2 \in Y_{2,1}^2$ or $(a_2, y_2) \in H_{2,1}^2$, then player 2 takes a_2^G for $T^{\frac{1}{3}}$ periods. On the other hand, if player 2 wants to send $y_2 \in Y_{2,2}^2$ or $(a_2, y_2) \in H_{2,2}^2$, then player 2 takes a_2^B for $T^{\frac{1}{3}}$ periods.

Player 1 takes α_1^{report} for $T^{\frac{1}{3}}$ periods. Player 1 infers that player 2's message is G if

$$\frac{\sum_t \Phi_{1,t}}{T^{\frac{1}{3}}} > \frac{q_2 + q_1}{2} \tag{6.17}$$

and B otherwise. Here, the summation is taken over $T^{\frac{1}{3}}$ periods where player 2 sends the message.

Next, consider Step 3 of Section 6.7.3.1. If 3-(a) is the case, then player 2 would send binary messages about h_2^{main} with the conditionally independent noisy cheap talk. Since all the messages are binary, we can see player 2 sending a binary message $m \in \{G, B\}$. Without the conditionally independent cheap talk, for each message m, the players spend $T^{\frac{1}{3}}$ periods. Player 2 takes a_2^m and player 1 takes α_1^{report}. Player 1 infers player 2's message by (6.17).

On the other hand, if 3-(b) of Section 6.7.3.1 is the case, then the players spend the same number of periods as in 3-(a). For periods where player 2 would send the message about h_2^{main} if 3-(a) were the case, player 2 takes a_2^G. On the other hand, for periods where player 2 sends the message about $k(r)$, player 2 sends the same message as in 3-(a). Player 1 always takes α_1^{report}.[9]

Player 1's reward on player 2 is determined as follows: while player 2 sends the message m corresponding to Step 2 of Section 6.7.3.1, player 1 gives the following reward: let $t(m)$ be the first period when player 2 sends m.

1. at $t(m)$, both a_2^G and a_2^B are indifferent and are better than the other actions (if any). The reward is given by

$$\Psi_{1,t(m)}^{a_2^G,a_{1,t(m)}} + \Psi_{1,t(m)}^{a_2^B,a_{1,t(m)}} - (q_2 + q_1)\,. \tag{6.18}$$

   By algebra, we can verify that

   (a) the expected payoff of taking a_2^G or a_2^B in period $t(m)$ is 0; and

   (b) the expected payoff of taking another action is $-(q_2 - q_1)$; and

2. after that, the constant action is optimal:

$$\sum_{t=t(m)+1}^{t(m)+T^{\frac{1}{3}}-1}\left(c + T^{-1}\bar{L}\frac{1}{q_2\,(1-q_1)\,(q_1-q_2)^2}\left(\begin{array}{c}(1-q_1)\,\mathbf{1}\{\Phi_{1,t(m)}=1\}\Psi_{1,t}^{a_2^G,a_{1,t}}\\ +q_2\mathbf{1}\{\Phi_{1,t(m)}=0\}\Psi_{1,t}^{a_2^B,a_{1,t}}\end{array}\right)\right). \tag{6.19}$$

   Here, c is a constant such that the expected payoff of taking a_2^G after $a_{2,t(m)} = a_2^G$ is equal to 0.

   From (6.13) and (6.19), we can verify that

[9] Note that player 1 takes the same action between in 3-(a) and 3-(b). Therefore, player 1 does not need to know which is the case.

(a) the expected payoff of taking a_2^m in period t after taking a_2^m in period $t(m)$ is 0; and

(b) the expected payoff of taking another action in period t after taking a_2^m in period $t(m)$ is $-\Theta(T^{-1})$.

Next, let us consider the reward for the messages corresponding to Step 3 of Section 6.7.3.1.

If player 1 infers that player 2's message corresponding to Step 2 says that player 2 observed $y_2 \in Y_{2,1}^2$ or $(a_2, y_2) \in H_{2,1}^2$, then for each message m in Step 3 of Section 6.7.3.1, player 1 gives the same reward as (6.18) and (6.19).

If player 1 infers that player 2's message in Step 2 says that player 2 observed $y_2 \in Y_{2,2}^2$ or $(a_2, y_2) \in H_{2,2}^2$, then for periods where player 2 is supposed to take a_2^G, player 1 gives

$$\sum_t \left(\Psi_{1,t}^{a_2^G, a_{1,t}} - q_2 \right), \tag{6.20}$$

so that player 2 takes a_2^G. Note that the expected payoff from (6.20) by taking a_2^G is zero.

For periods where player 2 sends the message about $k(r)$, player 1 gives the same reward as (6.18) and (6.19).

By backward induction, we can show the following: suppose that player 2 constantly took a_2^G in Step 2 of Section 6.7.3.1. Then, with probability no less than $1 - \exp(-\Theta(T^{\frac{1}{3}}))$, player 1 uses the reward (6.18) and (6.19).[10] Consider the last message m. All the rewards from (6.18) and (6.19) determined by the previous messages are sunk. In addition, the punishment and reward in the report block (g_i and f_i) defined in Section 2.9.6 are bounded by $\Theta(T^{-3})$. Hence, the difference between 1-(a) and 1-(b) is sufficiently large that, in period $t(m)$, player 2 takes either a_2^G or a_2^B. In addition, the difference between 2-(a) and 2-(b) is sufficiently large that after taking a_2^G or a_2^B in period $t(m)$, player 2 should take the constant action.

[10] More precisely, given the truthtelling incentive, since player j takes into the account that P_i is not an identity matrix as seen in Section 6.7.4, player i believes that the messages transmit correctly with probability one.

Since this equilibrium strategy makes the expected payoff about the last message from (6.18) and (6.19) equal to 0, the same argument holds until the first message.

Symmetrically, if player 2 constantly took a_2^B in Step 2 of Section 6.7.3.1, then it is optimal for player 2 in Step 3 to take a_2^G when player 2 is supposed to take a_2^G and to take a_2^G or a_2^B constantly for periods where player 2 should send a message for $k(r)$.

Finally, when player 2 sends the message about y_2 in Step 2 of Section 6.7.3.1, it is strictly optimal to take a constant action since (i) there is a strict incentive for the constant action and (ii) it gives player 2 the better idea about whether player 2 should tell the truth about the history or take a_2^G constantly.

Therefore, (i) this replacement of the conditionally independent cheap talk with messages via actions does not affect the payoff and since player 2 repeats the message for $T^{\frac{1}{3}}$ periods, (ii) player 1 infers the correct message m with high probability $1 - \exp(-\Theta(T^{\frac{1}{3}}))$, (iii) player 2's private signal cannot update player 2's belief about player 1's inference of player 2's message (conditional independence) by (6.13), and (iv) since the number of necessary messages is $\Theta(T^{\frac{1}{4}})$, the number of necessary periods for player 2 to send all the messages is

$$\Theta(T^{\frac{1}{3}+\frac{1}{4}}) < \Theta(T),$$

as desired.

For player 1, since we cannot assume that $|A_1| |Y_1| \leq |A_2| |Y_2|$, we cannot generically find a function $\phi_2(a_2, y_2)$ with the conditional independence property symmetric to $\phi_1(a_1, y_1)$. Therefore, after Step 4 in Section 6.7.3.1, we add an additional round where player 1 sends the messages about player 1's histories in Step 4. Based on the information that player 2 obtains in this additional round, player 2 creates a statistics to infer player 1's messages in Step 4, so that while player 1 sends the messages about h_1^{main} in Step 4 (before observing the history in the additional round), player 1 cannot update player 2's inference of the messages from player 1's signal observations.

6.7.5.1 Recovery of Conditional Independence

In Step 4 of Section 6.7.3.1, without cheap talk, player 1 takes $a_1 \in \{a_1^G, a_1^B\}$ to send the message. To send each message m, player 1 repeats a_1^m for $T^{\frac{1}{3}}$ periods. Player 2 takes some mixed action $\alpha_2^{\text{c.i.}}$ with $\alpha_2^{\text{c.i.}}(a_2) > \rho$ for all $a_2 \in A_2$. As for player 2, this takes $\Theta(T^{\frac{1}{3}+\frac{1}{4}})$ periods.

After this step is over, we have the following round named the "round for conditional independence." The intuitive structure is as follows. For each period t in Step 4, player 1 reports the history in period t to player 2 in the round for conditional independence. Player 2 infers player 1's messages in Step 4 combining player 2's signals in Step 4 and player 1's reports about player 1's history.

Player 2 gives the following rewards to player 1: (i) the adjustment of player 1's reward so that $\sigma_1(x_1)$ is exactly optimal; (ii) a reward that makes an optimal strategy in the round for conditional independence given player 1's history in Step 4 unique; (iii) we make sure that (i) is much smaller than (ii), so that player 1's history in Step 4 and the strictness of player i's incentive in the round for conditional independence completely determines player 1's strategy in the round for conditional independence, independently of h_1^{main}. (iv) Given player 1's continuation strategy in the round for conditional independence, player 2 in Step 4 changes player 1's continuation payoff so that ex ante (before player 1 takes an action in each period of Step 4), the difference in the expected payoffs about the reward in (ii) from different actions in Step 4 is canceled out, taking (ii) into account. (iv) implies that the round for conditional independence does not affect player 1's incentive in Step 4.

Finally, since player 2 obtains rich information about player 1's history in Step 4 from the round for conditional independence, player 2 infers player 1's messages in Step 4 so that player 1 cannot update the belief about player 2's inference of player 1's message during Step 4. (iii) implies that the incentives in the round for conditional independence is not affected by this.

Now, we define the round of conditional independence formally. For each period t in Step 4, we attach $S \log_2 |A_1| |Y_1|$ periods in the round for conditional independence. S is a fixed number to be determined independently of T. Hence, this new round also takes $\Theta(T^{\frac{1}{3}+\frac{1}{4}})$ periods.

In these $S \log_2 |A_1| |Y_1|$ periods, player 1 sends the message about the history in each period t in Step 4, $(a_{1,t}, y_{1,t})$, as follows:

- we create a mapping between a sequence of $\{a_1^G, a_1^B\}$ (denoted by $\mathbf{a}_1(a_1, y_1) \in \{a_1^G, a_1^B\}^{\log_2|A_1||Y_1|}$), and $(a_1, y_1) \in |A_1| \times |Y_1|$, such that each $\mathbf{a}_1(a_1, y_1)$ uniquely identifies (a_1, y_1) and that, for each (a_1, y_1), there is at least one $\mathbf{a}_1(a_1, y_1)$ that is mapped into (a_1, y_1);

- player 1's strategy is to be determined. Intuitively, $S \log_2 |A_1| |Y_1|$ periods are separated into $\log_2 |A_1| |Y_1|$ sets of S periods. In each S periods, player 1 sends the message about the corresponding element of $\mathbf{a}_1(a_{1,t}, y_{1,t})$, depending on player 1's history $(a_{1,t}, y_{1,t})$ in Step 4.

- player 2 always takes a_2^G; and

- $S \log_2 |A_1| |Y_1|$ periods are separated into $\log_2 |A_1| |Y_1|$ sets of S periods. In each S periods, player 2 infers that player 1 sends the message a_1^G if
$$\frac{\sum_s \Psi_{2,s}^{a_1^G, a_2^G}}{S} > \frac{q_2 + q_1}{2} \tag{6.21}$$
and a_1^B otherwise. From these inferences and the correspondence $\mathbf{a}_1(a_1, y_1)$, player 2 infers player 1's message $(\hat{a}_{1,t}, \hat{y}_{1,t})$.

Let
$$-\left\|\mathbf{1}_{a_{2,t}, y_{2,t}} - \mathbb{E}\left[\mathbf{1}_{a_{2,t}, y_{2,t}} \mid \alpha_2^{\text{c.i.}}, \hat{a}_{1,t}, \hat{y}_{1,t}\right]\right\|^2 - f(h_2^{S \log_2 |A_1||Y_1|}), \tag{6.22}$$
be player 2's reward on player 1. Here, $h_2^{S \log_2 |A_1||Y_1|}$ is player 2's history in the $S \log_2 |A_1| |Y_1|$ periods where player 1 sends $(\hat{a}_{1,t}, \hat{y}_{1,t})$ and f will be determined in the following lemma:

Lemma 46 There exists e_1 such that, for each $S \in \mathbb{N}$ and $\varepsilon > 0$, there generically exist f and $e_2 > 0$ such that, suppose that the players play the following game:

1. Nature chooses $a_{1,t}$ (t is introduced to make the notations consistent) and $(y_{1,t}, y_{2,t})$ is distributed according to $q(y_t \mid a_{1,t}, \alpha_2^{\text{c.i.}})$. Player 1 can observe only $(a_{1,t}, y_{1,t})$;

2. the players play an $(S \log_2 |A_1| \, |Y_1|)$-period finitely repeated game where, in each period $s \in \{1, \ldots, S \log_2 |A_1| \, |Y_1|\}$, player 1 chooses $a_{1,s} \in A_1$, the signal profile $(y_{1,s}, y_{2,s})$ is generated by $q(y_s \mid a_{1,s}, a_2^G)$, player 1 can observe only $(a_{1,s}, y_{1,s})$, and there is no instantaneous utility;

3. player 2 infers $(\hat{a}_{1,t}, \hat{y}_{1,t})$ as explained above; and

4. player 1's utility is given by (6.22).

Then,

(a) the message transmits correctly with probability at least $1 - \varepsilon - e_1 \exp(-\Theta(S^{\frac{1}{2}}))$; and

(b) for any two pure strategies σ_1 and $\tilde{\sigma}_1$, if there exists h_1 where $\sigma_1 \mid h_1 \neq \tilde{\sigma}_1 \mid h_1$ on the path after h_1, then the continuation payoff from h_1 is different by at least e_2 (player 1's incentive is strict by e_2). Here, with abuse of notation, σ_1 and h_1 are player 1's strategy and history in the game just defined. Let σ_1^* be the (unique) optimal strategy.

Note that $e_2 > 0$ here corresponds to (ii) in the intuitive explanation above.

Proof: There exists $\bar{E}$ such that

$$\left\| \mathbf{1}_{a_{2,t}, y_{2,t}} - \mathbb{E}\left[\mathbf{1}_{a_{2,t}, y_{2,t}} \mid \alpha_2^{\text{c.i.}}, \hat{a}_{1,t}, \hat{y}_{1,t} \right] \right\|^2 < \bar{E}$$

for all $a_{2,t}$, $y_{2,t}$, $\hat{a}_{1,t}$ and $\hat{y}_{1,t}$. In addition, define

$$e = \min_{(a_{1,t},y_{1,t})} \left\{ \begin{array}{c} \min_{(\hat{a}_{1,t},\hat{y}_{1,t}) \neq (a_{1,t},y_{1,t})} \mathbb{E}\left[\begin{array}{c} \left\| \mathbf{1}_{a_{2,t},y_{2,t}} - \mathbb{E}\left[\mathbf{1}_{a_{2,t},y_{2,t}} \mid \alpha_2^{\text{c.i.}}, \hat{a}_{1,t}, \hat{y}_{1,t} \right] \right\|^2 \\ \mid \alpha_2^{\text{c.i.}}, a_{1,t}, y_{1,t} \end{array} \right] \\ -\mathbb{E}\left[\begin{array}{c} \left\| \mathbf{1}_{a_{2,t},y_{2,t}} - \mathbb{E}\left[\mathbf{1}_{a_{2,t},y_{2,t}} \mid \alpha_2^{\text{c.i.}}, \hat{a}_{1,t}, \hat{y}_{1,t} \right] \right\|^2 \\ \mid \alpha_2^{\text{c.i.}}, a_{1,t}, y_{1,t} \end{array} \right] \end{array} \right\}.$$

By Assumption 5, $e > 0$. Note that $\bar{E}$ and e are independent of S and f.

Fix S and $\varepsilon > 0$ arbitrarily. Lemma 12 implies that we can find and fix f and $e_2 > 0$ such that (b) is satisfied and that $f(h_2^{S \log_2 |A_1||Y_1|}) \in [-\frac{\varepsilon e}{2}, \frac{\varepsilon e}{2}]$ for all $h_2^{S \log_2 |A_1||Y_1|}$. Specifically, first, without loss of generality, we can make sure that player 1 has only two actions a_1^G and a_1^B since otherwise, Lemma 12 enables player 2 to give a very high punishment if player 1 takes an action other than a_1^G or a_1^B. Second, consider the last period of the game. If there is a history where a_1^G and a_1^B are indifferent, Lemma 12 guarantees that player 2 can break the ties by adding a small reward for a_1^G uniformly for all the histories. Hence, we establish the strictness in the last period. Third, we can proceed by backward induction. Whenever player 2 breaks a tie for some period, it does not affect the strict incentives in the later periods since the reward in a certain period will be sunk in the later periods.

Let $\bar{\sigma}_1$ be such that player 1 constantly takes a_1^G or a_1^B for each S periods that correspond to the proper counterpart of $\mathbf{a}_1(a_{1,t}, y_{1,t})$. In addition, define

$$R^*(a_{1,t}, y_{1,t}) = -\mathbb{E}\left[\left\| \mathbf{1}_{a_{2,t},y_{2,t}} - \mathbb{E}\left[\mathbf{1}_{a_{2,t},y_{2,t}} \mid \alpha_2^{\text{c.i.}}, a_{1,t}, y_{1,t} \right] \right\|^2 \mid \alpha_2^{\text{c.i.}}, a_{1,t}, y_{1,t} \right],$$

$$\bar{R}(a_{1,t}, y_{1,t}) = -\min_{(\hat{a}_{1,t},\hat{y}_{1,t}) \neq (a_{1,t},y_{1,t})} \mathbb{E}\left[\begin{array}{c} \left\| \mathbf{1}_{a_{2,t},y_{2,t}} - \mathbb{E}\left[\mathbf{1}_{a_{2,t},y_{2,t}} \mid \alpha_2^{\text{c.i.}}, \hat{a}_{1,t}, \hat{y}_{1,t} \right] \right\|^2 \\ \mid \alpha_2^{\text{c.i.}}, a_{1,t}, y_{1,t} \end{array} \right],$$

$$\underline{R} = -\max_{(a_{1,t},y_{1,t}),(\hat{a}_{1,t},\hat{y}_{1,t}) \neq (a_{1,t},y_{1,t})} \mathbb{E}\left[\begin{array}{c} \left\| \mathbf{1}_{a_{2,t},y_{2,t}} - \mathbb{E}\left[\mathbf{1}_{a_{2,t},y_{2,t}} \mid \alpha_2^{\text{c.i.}}, \hat{a}_{1,t}, \hat{y}_{1,t} \right] \right\|^2 \\ \mid \alpha_2^{\text{c.i.}}, a_{1,t}, y_{1,t} \end{array} \right].$$

Note that for all $(a_{1,t}, y_{1,t})$,

$$\begin{aligned} R^*(a_{1,t}, y_{1,t}) - \underline{R} &\leq 2E, \\ R^*(a_{1,t}, y_{1,t}) - \bar{R}(a_{1,t}, y_{1,t}) &\geq e. \end{aligned}$$

For (a), since the message transmits with ex ante probability $1 - \exp(-\Theta(S^{\frac{1}{2}}))$ with $\bar{\sigma}_1$, the optimal strategy σ_1^* should guarantee

$$\begin{aligned} & (1 - \Pr((\hat{a}_{1,t}, \hat{y}_{1,t}) \neq (a_{1,t}, y_{1,t}))) R^*(a_{1,t}, y_{1,t}) \\ & + \Pr((\hat{a}_{1,t}, \hat{y}_{1,t}) \neq (a_{1,t}, y_{1,t})) \bar{R}(a_{1,t}, y_{1,t}) \\ & - \mathbb{E}\left[f\left(h_2^{S \log_2 |A_1||Y_1|}\right) \mid \sigma_1^* \mid (a_{1,t}, y_{1,t})\right] \\ \geq & \left(1 - \exp(-\Theta(S^{\frac{1}{2}}))\right) R^*(a_{1,t}, y_{1,t}) + \exp(-\Theta(S^{\frac{1}{2}}))\underline{R} \\ & - \mathbb{E}\left[f\left(h_2^{S \log_2 |A_1||Y_1|}\right) \mid \bar{\sigma}_1 \mid (a_{1,t}, y_{1,t})\right] \end{aligned}$$

or

$$\begin{aligned} & \Pr((\hat{a}_{1,t}, \hat{y}_{1,t}) \neq (a_{1,t}, y_{1,t})) \\ \leq & \frac{R^*(a_{1,t}, y_{1,t}) - \underline{R}}{R^*(a_{1,t}, y_{1,t}) - \bar{R}(a_{1,t}, y_{1,t})} \exp(-\Theta(S^{\frac{1}{2}})) \\ & + \frac{\mathbb{E}\left[f\left(h_2^{S \log_2 |A_1||Y_1|}\right) \mid \bar{\sigma}_1 \mid (a_{1,t}, y_{1,t})\right] - \mathbb{E}\left[f\left(h_2^{S \log_2 |A_1||Y_1|}\right) \mid \sigma_1^* \mid (a_{1,t}, y_{1,t})\right]}{R^*(a_{1,t}, y_{1,t}) - \bar{R}(a_{1,t}, y_{1,t})} \\ \leq & \frac{2E}{e} \exp(-\Theta(S^{\frac{1}{2}})) + \varepsilon. \end{aligned}$$

Take

$$e_1 = \frac{2E}{e} > 0,$$

which is independent of S and f, then we are done. ∎

Given that, for each period t in Step 4 of Section 6.7.3.1, player 1 will take $\sigma_1^* \mid (a_{1,t}, y_{1,t})$ in the round for conditional independence and that player 2's reward on player 1 in the round for conditional independence is (6.22), the expected payoff from Nature's choice $a_{1,t}$ is determined. By Lemma 12, there exists $\bar{g}(a_1)$ such that

$$\sum_{a_1} \bar{g}\left(a_1\right) \Psi_{2,t}^{a_{2,t},a_1} \tag{6.23}$$

cancels out the difference from (6.22). Player 2 adds (6.25) to player 2's reward on player 1 in the report block. Then, seeing $a_{1,t}$ as player 1's action in Step 4 of Section 6.7.3.1, any action gives ex ante payoff 0 in terms of the payoffs in the round for conditional independence. Note that this corresponds to (iv) in the intuitive explanation above.

Then, based on the report σ_1^* in the round for conditional independence, player 2 can construct the statistics that indicates player 1's action with the conditional independence property from the perspective of Step 4 of Section 6.7.3.1.

Lemma 47 There exist $S \in \mathbb{N}$, $\varepsilon > 0$ and $\phi_2 : \hat{a}_1, \hat{y}_1, y_2 \to (0,1)$ such that, for all $(a_{1,t}, y_{1,t})$, if player 1 reports $(a_{1,t}, y_{1,t})$ by σ_1^*, then for all $y_{1,t}$,

$$\mathbb{E}\left[\phi_2(\hat{a}_{1,t}, \hat{y}_{1,t}, y_{2,t}) \mid \alpha_2^{\text{c.i.}}, a_{1,t}, y_{1,t}, \sigma_1^* \mid \alpha_2^{\text{c.i.}}, a_{1,t}, y_{1,t}\right] = \begin{cases} q_2 & \text{if } a_{1,t} = a_{1,t}^G, \\ q_1 & \text{if } a_{1,t} \neq a_{1,t}^G. \end{cases}$$

Proof: Since, from Condition (a) of Lemma 46, $(\hat{a}_{1,t}, \hat{y}_{1,t})$ transmits with probability no less than $1 - \varepsilon - e_1 \exp(-\Theta(S^{\frac{1}{2}}))$, for sufficiently large S and small ε, player 2 has enough information. Note that e_1 does not depend on S. ∎

Now, we are ready to construct Step 4 of Section 6.7.3.1. Fix $S \in \mathbb{N}$ and $\varepsilon > 0$ such that Lemma 47 holds. Then, fix f and e_2 such that Lemma 46 holds for those S and ε. Finally, take δ such that

$$e_2 > T^{-1} = (1-\delta)^{\frac{1}{2}}. \tag{6.24}$$

This implies that (iii) in the intuitive explanation is satisfied.

Step 4 of Section 6.7.3.1 For each $T^{\frac{1}{3}}$ periods when player 1 is supposed to take a constant action, player 2 infers that player 1's action is a_1^G if

$$\frac{\sum_t \Phi_{2,t}}{T^{\frac{1}{3}}} > \frac{q_2 + q_1}{2}$$

and a_2^B otherwise. Here $\Phi_{2,t} \in \{0, 1\}$ is calculated from $\phi_2(\hat{a}_{1,t}, \hat{y}_{1,t}, y_{2,t})$. Then, from Lemma 47, if player 1 uses $\sigma_1^* \mid a_{1,t}, y_{1,t}$ in the round for conditional independence, then $\Phi_{2,t}$ has the same property as $\Phi_{1,t}$ from Lemma 47.

If player 2 sent the message $y_2 \in Y_{2,2}^2$ or $(a_2, y_2) \in H_{2,2}^2$ in Step 2 of Section 6.7.3.1, then player 2's reward on player 1 in the Step 4 is symmetrically defined as player 1's reward on player 2 in Step 3 after player 1 infers that player 2's message in Step 2 says that player 2 observed $y_2 \in Y_{2,1}^2$ or $(a_2, y_2) \in H_{2,1}^2$. If $y_2 \in Y_{2,1}^2$ or $(a_2, y_2) \in H_{2,1}^2$, then player 2 only cancels out the differences in player 1's instantaneous utilities. Hence, player 1 can condition that $y_2 \in Y_{2,2}^2$ or $(a_2, y_2) \in H_{2,2}^2$. On the top of that, player 2 gives the reward

$$\sum_{a_1} \bar{g}(a_1)\, \Psi_{2,t}^{a_{2,t}, a_1} \tag{6.25}$$

to cancel out the effect of the round for conditional independence.

Round for Conditional Independence For $S \log_2 |A_1| |Y_1|$ periods that correspond to period t in Step 4 of Section 6.7.3.1, player 1 plays $\sigma_1^* \mid a_{1,t}, y_{1,t}$ and player 2 plays a_2^G. The reward is given by (6.22).

Optimality of Player 1's Strategy Note that all the rewards in Step 4 of Section 6.7.3.1 affected by the messages in the round for conditional independence are bounded by T^{-1}. Since we take T such that $e_2 > T^{-1}$ by (6.24), from Condition 2 of Lemma 46, regardless of the history in Step 4, the optimal strategy in the round for conditional independence is σ_1^* (note that (6.25) is sunk in the round for conditional independence).

Then, (6.25) together with the expected reward in the round for conditional independence makes player 1 indifferent between all the actions in terms of the expected reward in the round for conditional independence and yield 0 in expectation regardless of the history.

Therefore, the same argument as in Step 3 of Section 6.7.3.1 for player 2 establishes the incentive in Step 4 since, given that player 1 takes σ_1^*, the conditional independence property holds and player 1 conditions that $y_2 \in Y_{2,2}^2$ or $(a_2, y_2) \in H_{2,2}^2$, that is, player 1 cannot infer player 2's history from Step 3.

Chapter 7

General N-Player Game Without Cheap Talk

In this chapter, we prove the dispensability of the perfect and error-reporting noisy cheap talk and public randomization in the proof of Theorem 7 for the general N-player game with $N \geq 3$ (see Chapter 5 for the proof with the perfect and error-reporting noisy cheap talk and public randomization). Remember that in Chapter 5, the coordination block uses the perfect cheap talk, the supplemental rounds use the error-reporting noisy cheap talk, the report block uses the public randomization and perfect cheap talk, and the re-report block uses the perfect cheap talk.

First, in Section 7.2, we replace the perfect cheap talk in the coordination block with the error-reporting noisy cheap talk. As seen in Section 1.5.6.2, with more than two players, we need to make sure that while the players exchange messages and infer the other players' messages from private signals generated by the error-reporting noisy cheap talk in order to coordinate on x_i, there is no player who can induce a situation where some players infer x_i is G while the others infer x_i is B in order to increase her own equilibrium payoff. For this purpose, in Section 7.1.1, we introduce the new assumptions and explain why they are sufficient.

Second, Section 7.3 explains the structure of the review phase (finitely repeated game) without perfect cheap talk.

Third, in Section 7.4, we dispense with the error-reporting noisy cheap talk in the coordination block (given the first step above) and supplemental rounds. See Section 7.1.2 for what assumption is sufficient for this step.

Fourth, in Section 7.5, we define the equilibrium strategy in the coordination and main blocks. The almost optimality of this strategy is verified in Section 7.6.

Fifth, in Section 7.7, we dispense with the public randomization and the perfect cheap talk in the report and re-report blocks. Section 7.1.3 explains new assumptions for this step.

In this chapter, when we say player $i \notin \{1, ..., N\}$, without otherwise specified, it means player $i \pmod N$. In addition, without loss of generality, assume that

$$|A_1| |Y_1| \geq \cdots \geq |A_N| |Y_N| . \tag{7.1}$$

In addition, since we only use the error-reporting noisy cheap talk with precision $p = 1 - \exp(-T^{\frac{1}{2}})$, we omit the precision.

7.1 Notations and Assumptions

Before explaining the assumptions for dispensing with cheap talk, we will offer one assumption that is useful to construct a reward while players are exchanging the messages. Imagine the situation where player j takes two possible actions $\bar{A}_j$ with $\bar{A}_j \subset A_j$ and $|\bar{A}_j| = 2$ and players $-(i,j)$ take $\alpha_{-(i,j)}$ where each player takes each action with probability at least $\rho > 0$. We assume that player $i-1$ can identify actions by player $i \neq j$ from her action a_{i-1} and signal y_{i-1}:

Assumption 48 For all $\rho > 0$, $i, j \in I$, $\bar{A}_j \subset A_j$ with $|\bar{A}_j| = 2$ and $\{\alpha_n\}_{n \in -(i,j)}$ with $\alpha_n(a_n) \geq \rho$ for all n and a_n, $\left(\Pr\left(a_{i-1}, y_{i-1} \mid a_i, a_j, \alpha_{-(i,j)}\right)\right)_{a_{i-1}, y_{i-1}}$ is linearly independent with respect to $a_i \in A_i$ and $a_j \in \bar{A}_j$.

We argue that Assumption 48 is generic under Assumption 3. There are following two cases: if player j is player $i-1$, then player $i-1$ knows a_j and Assumption 48 is equivalent to linear independence of $\left(q\left(y_{i-1} \mid a_i, a_j, \alpha_{-(i,j)}\right)\right)_{y_{i-1}}$, which is equivalent to Assumption 3. If player j is not player $i-1$, then player $i-1$ mixes all the actions. Hence, $\left(\Pr\left(a_{i-1}, y_{i-1} \mid a_i, a_j, \alpha_{-(i,j)}\right)\right)_{a_{i-1}, y_{i-1}}$ is generically linearly independent if

$$|A_{i-1}| \, |Y_{i-1}| \geq \left|\bar{A}_j\right| |A_i| = 2 \, |A_i| \, ,$$

which is implied by Assumption 3.

7.1.1 Assumptions for Dispensing with the Perfect Cheap Talk in the Coordination Block

We explain how to replace the perfect cheap talk with the error-reporting noisy cheap talk in the coordination block. As explained in Section 1.3.2, the error-reporting noisy cheap talk is "private" in that when player j sends the message to player n via error-reporting noisy cheap talk, the main signal $f[n](m)$ is only observed by player n.

This creates the problem mentioned in Section 1.5.6.2: if player i sent the message x_i to each of the other players $-i$ via error-reporting noisy cheap talk separately, then player i could create a situation where some players infer x_i is G while the others infer x_i is B by telling a lie. Since the action that will be taken in the main blocks after such an event may not be included in $\{\alpha(x)\}_x$ and we do not have any bound on player i's payoff in such a situation, it might be of player i's interest to tell a lie.

To prevent this situation, we consider the following message protocol: let $N(i) \equiv \{i, i+1, i+2\}$ be the set of players whose index is in $\{i, i+1, i+2\}$. In addition, let

$$n^*(i) \in \arg\min_{j \in \{i, i+2\}} |A_j| |Y_j| \tag{7.2}$$

be the player whose $|A_j| |Y_j|$ is smaller between $\{i, i+2\}$. Let

$$n^{**}(i) = \{i, i+2\} \setminus \{n^*(i)\} \tag{7.3}$$

be the other player. Note that $N(i) = \{n^*(i), i+1, n^{**}(i)\}$. The players communicate as follows:

1. first, player i sends the message about x_i to player $n^*(i)$ via error-reporting noisy cheap talk. Let $w_i = f_1[n^*(i)](x_i)$ be player i's inference;
2. then, player $n^*(i)$ sends the message about w_i to players $N(i)$ via *actions*. This corresponds to "Phase 1" of Hörner and Olszewski (2006);
3. after that, each player j in $N(i)$ sends the message about her inference of x_i in Step 2 to each player $n \neq j$ via error-reporting noisy cheap talk; and
4. finally, each player n infers x_i based on the messages from $N(i)$. This corresponds to "Phase 2" of Hörner and Olszewski (2006).

As Hörner and Olszewski (2006), to incentive each player $j \in N(i)$ to tell the truth in Step 3, for each $j \in N(i)$, if there exists player $n \in -j$ such that player n's inference of player j's message changes player n's inference of x_i in Step 4 (that is, if player j is "pivotal"), then player $j-1$, who will know player j is pivotal in the re-report block, makes player j indifferent between any action profile sequence.

Given above, we will show that player $n^*(i)$ does not want to deviate in Step 2 in order to create a situation where player $n^*(i)$ herself will be pivotal with a high probability in

Step 3. Remember that we take $n^*(i)$ such that the set of player $n^*(i)$'s action-signal pairs is smaller than that of player $n^{**}(i)$ in (7.2). Heuristically, this guarantees that player $n^*(i)$ cannot infer player $n^{**}(i)$'s inference precisely, which prevents player $n^*(i)$ from creating the situation where player $n^*(i)$ is pivotal.

Given player $n^*(i)$'s truthtelling strategy in Step 2, the probability that player i is pivotal in Step 3 is almost independent of player i's strategy in Step 1. Since x_i controls player $(i+1)$'s payoff, players i and $n^*(i) \neq i+1$ do not have an incentive to manipulate the communication in Step 1.

Below, we explain what assumption is sufficient for what step.

The first step is straightforward if the error-reporting noisy cheap talk is available.

After player $n^*(i)$ infers w_i, player $n^*(i)$ sends the message about w_i to players $N(i) = \{n^*(i), i+1, n^{**}(i)\}$. While player $n^*(i)$ sends w_i, player $n^*(i)$ takes $a_{n^*(i)}^{w_i}$ and player $j \in -n^*(i)$ take $\alpha_j^{\text{receive}} \in \Delta(A_j)$ for $T^{\frac{1}{2}}$ periods. That is, in equilibrium, the players take $\alpha(i, w_i) \equiv \left(a_{n^*(i)}^{w_i}, \alpha_{-n^*(i)}^{\text{receive}}\right)$ for $T^{\frac{1}{2}}$ periods.

Take $n \in N(i) \setminus n^*(i)$. Suppose that player $j = N(i) \setminus \{n^*(i), n\}$ unilaterally deviates and takes $a_j \in A_j$. Then, the distribution of player n's action-signal pairs is

$$\mathbf{q}_n(a_j, \alpha_{-j}(i, w_i)) \equiv \left(q\left(a_n, y_n \mid a_j, \alpha_{-j}(i, w_i)\right)\right)_{a_n \in A_n, y_n \in Y_n}.$$

Consider the following linear equations: for any $a_j \in A_j$,

$$\boldsymbol{i}_n(i)\mathbf{q}_n(a_j, \alpha_{-j}(i, w_i)) = \begin{cases} q_2 & \text{if } w_i = G, \\ q_1 & \text{if } w_i = B. \end{cases} \tag{7.4}$$

Here, $\boldsymbol{i}_n(i)$ is a $1 \times |A_n| |Y_n|$ vector. Solve (7.4) for $\boldsymbol{i}_n(i)$. Suppose that there are $L_n(i)$ linearly independent solutions. Then, let

$$I_n(i) = \left(\boldsymbol{i}_n^l(i)\right)_{l=1}^{L_n(i)} \tag{7.5}$$

be the $L_n(i) \times |A_n| |Y_n|$ matrix collecting all the linearly independent $\boldsymbol{i}_n(i)$'s. Since all the expressions are linear, we can make sure that each element in $I_n(i)$ is in $(0,1)$ by a proper affine transformation as in Lemma 14.

Intuitively, if player n uses $\boldsymbol{i}_n(i)\mathbf{1}_{a_{n,t},y_{n,t}}$ after the history $(a_{n,t}, y_{n,t})$ to infer w_i, then player j cannot manipulate player n's inference by (7.4). That is,

1. if $\sum_t \boldsymbol{i}_n(i)\mathbf{1}_{a_{n,t},y_{n,t}}$ is close to $(q_2, ..., q_2)^\top$, then player n infers w_i is G;
2. if $\sum_t \boldsymbol{i}_n(i)\mathbf{1}_{a_{n,t},y_{n,t}}$ is close to $(q_1, ..., q_1)^\top$, then player n infers w_i is B; and
3. if $\sum_t \boldsymbol{i}_n(i)\mathbf{1}_{a_{n,t},y_{n,t}}$ is close to neither $(q_2, ..., q_2)^\top$ nor $(q_1, ..., q_1)^\top$, then player n infers w_i is M ("middle").

Note that (7.4) implies that player $j = N(i) \setminus \{n^*(i), n\}$ cannot manipulate this inference. Hence, we are left to show that player $n^*(i)$ cannot induce the situation that player $n^*(i)$ will be pivotal. Intuitively, it suffices to show that player $n^*(i)$ puts a negligible belief on the event that player $n^{**}(i)$ infers w_i is $\hat{w}_i \in \{G, B\}$ and player $i+1$ infers w_i is $\{G, B\} \setminus \{\hat{w}_i\}$ (we allow player $i+1$ to infer w_i is M) since, as we will see, these are only cases where player $n^*(i)$ will be pivotal with a non-negligible probability.

To calculate player $n^*(i)$'s belief, consider the matrix projecting player $n^*(i)$'s history on the conditional expectation of player n's history given an action profile by players $-n^*(i)$ being equal to $\alpha_{-n^*(i)}(i, w_i)$:

$$\underset{\left(|A_n||Y_n| \times \left|A_{n^*(i)}\right|\left|Y_{n^*(i)}\right|\right)}{Q_{n,n^*(i)}(i)},$$

where the element corresponding to $(a_n, y_n), \left(a_{n^*(i)}, y_{n^*(i)}\right)$ is the conditional probability that player n observes (a_n, y_n) given $\left(a_{n^*(i)}, y_{n^*(i)}\right)$ and $\alpha_{-n^*(i)}(i, w_i)$:

$$q(a_n, y_n | \alpha_{-n^*(i)}(i, w_i), a_{n^*(i)}, y_{n^*(i)}).$$

Since $\alpha_{-n^*(i)}(i, w_i) = \alpha_{-n^*(i)}^{\text{receive}}$ is independent of w_i, $Q_{n,n^*(i)}(i)$ is independent of w_i.

Given $Q_{n,n^*(i)}(i)$, the set of player $n^*(i)$'s histories such that player $n^*(i)$ believes that player n infers $\hat{w}_i \in \{G, B\}$ with a non-negligible probability should be included in

$$\boldsymbol{I}_{n,n^*(i)}[\varepsilon](i, \hat{w}_i) \equiv \left\{ \begin{array}{c} \mathbf{x} \in \mathbb{R}_+^{|A_{n^*(i)}||Y_{n^*(i)}|} : \exists \boldsymbol{\varepsilon} \in \mathbb{R}^{L_n(i)} \text{ such that} \\ \left\{ \begin{array}{c} \|\boldsymbol{\varepsilon}\| \leq \varepsilon, \\ I_n(i)\left(Q_{n,n^*(i)}(i)\mathbf{x} + \boldsymbol{\varepsilon}\right) = q(\hat{w}_i)\mathbf{1}. \end{array} \right. \end{array} \right\}$$

with $q(\hat{w}_i) = q_2$ if $\hat{w}_i = G$ and $q(\hat{w}_i) = q_1$ if $\hat{w}_i = B$.

So that player $n^*(i)$ puts a negligible belief on the event that player $n^{**}(i)$ infers $w_i = \hat{w}_i \in \{G, B\}$ and player $i+1$ infers $w_i = \{G, B\} \setminus \{\hat{w}_i\}$, we want to make sure that, for sufficiently small ε,

$$\boldsymbol{I}_{n^{**}(i),n^*(i)}[\varepsilon](i, G) \cap \boldsymbol{I}_{i+1,n^*(i)}[\varepsilon](i, B) = \emptyset$$

and

$$\boldsymbol{I}_{n^{**}(i),n^*(i)}[\varepsilon](i, B) \cap \boldsymbol{I}_{i+1,n^*(i)}[\varepsilon](i, G) = \emptyset.$$

We show that the following assumption is sufficient:

Assumption 49 There exist $\left\{a_i^G, a_i^B, \alpha_i^{\text{receive}}\right\}_{i \in I}$, q_2, q_1, and $\bar{\varepsilon} > 0$ such that $q_2, q_1 \in (0,1)$, $q_2 > q_1$, and for each $i \in I$,

1. there exists $\mathbf{x} \in \mathbb{R}^{L_{i+1}(i) + L_{n^{**}(i)}(i)}$ such that

$$\begin{bmatrix} I_{i+1}(i)Q_{i+1,n^*(i)}(i) \\ I_{n^{**}(i)}(i)Q_{n^{**}(i),n^*(i)}(i) \end{bmatrix}' \mathbf{x} \leq \mathbf{0}, \begin{bmatrix} q_2\mathbf{1} \\ q_1\mathbf{1} \end{bmatrix} \cdot \mathbf{x} > 0;$$

2. there exists $\mathbf{x} \in \mathbb{R}^{L_{i+1}(i) + L_{n^{**}(i)}(i)}$ such that

$$\begin{bmatrix} I_{i+1}(i)Q_{i+1,n^*(i)}(i) \\ I_{n^{**}(i)}(i)Q_{n^{**}(i),n^*(i)}(i) \end{bmatrix}' \mathbf{x} \leq \mathbf{0}, \begin{bmatrix} q_1\mathbf{1} \\ q_2\mathbf{1} \end{bmatrix} \cdot \mathbf{x} > 0;$$

and

3. for all $i \in I$,

$$j = \begin{cases} i-1 & \text{if } i \neq 1, \\ 2 & \text{if } i = 1, \end{cases}$$

$a_{-(i,j)} \in A_{-(i,j)}$ and $\alpha_j \in \Delta(A_j)$ with

(a) $\alpha_j(a_j^G) > 0$ and $\alpha_j(a_j^B) > 0$ if $j = n^*(i)$; and

(b) $\alpha_j = \alpha_j^{\text{receive}}$ for $j \neq n^*(i)$,

for all (a_i, y_i) and (a_i', y_i') with $(a_i, y_i) \neq (a_i', y_i')$, we have

$$\mathbb{E}\left[\mathbf{1}_{a_j,y_j} \mid y_i, a_i, \alpha_j, a_{-(i,j)}\right] \neq \mathbb{E}\left[\mathbf{1}_{a_j,y_j} \mid y_i', a_i', \alpha_j, a_{-(i,j)}\right].$$

As we will see in Section 7.7, Condition 3 will be used to show the truthtelling incentive in the report block.

Under Assumption 6, Assumption 49 is generic. Since Condition 3 is generic, we concentrate on Conditions 1 and 2. Note that (7.4) puts $2(|A_j| - 1)$ constraints while we have $|A_n||Y_n| - 1$ degrees of freedom for $\boldsymbol{i}_n(i)$ if $\mathbf{q}_n(a_j, \alpha_{-j}(i, w_i))$ is linearly independent for each w_i and a_j except for the constraint that "if we add all the elements up, then it should be one." Hence, generically $L_n(i)$ is equal to $|A_n||Y_n| - 2|A_j| + 1$. Therefore, for each of Conditions 1 and 2, we have $|A_{i+1}||Y_{i+1}| + |A_{n^{**}(i)}||Y_{n^{**}(i)}| - 2|A_{i+1}| - 2|A_{n^{**}(i)}| + 1$ degrees of freedom for $\mathbf{x}$,[1] while we have $|A_{n^*(i)}||Y_{n^*(i)}| + 1$ constraints. Since

$$\begin{aligned} & |A_{i+1}||Y_{i+1}| + |A_{n^{**}(i)}||Y_{n^{**}(i)}| - 2|A_{i+1}| - 2|A_{n^{**}(i)}| + 1 \\ \geq\ & |A_{i+1}||Y_{i+1}| + |A_{n^{*}(i)}||Y_{n^{*}(i)}| - 2|A_{i+1}| - 2|A_{n^{**}(i)}| + 1 \text{ by (7.2)} \\ \geq\ & |A_{i+1}|\left(|Y_{i+1}| - 2 - \frac{2}{|A_{i+1}|}|A_{n^{**}(i)}|\right) + |A_{n^{*}(i)}||Y_{n^{*}(i)}| + 1 \\ \geq\ & |A_{n^{*}(i)}||Y_{n^{*}(i)}| + 1 \text{ by Assumption 6,} \end{aligned}$$

[1]Note that two rows are parallel to $\mathbf{1}$.

we can generically find $\mathbf{x}$.

Lemma 50 If Assumption 49 is satisfied, then there exists $\bar{\varepsilon} > 0$ such that, for all $\varepsilon \leq \bar{\varepsilon}$,

$$\boldsymbol{I}_{n^{**}(i),n^*(i)}[\varepsilon](i,G) \cap \boldsymbol{I}_{i+1,n^*(i)}[\varepsilon](i,B) = \emptyset \tag{7.6}$$

and

$$\boldsymbol{I}_{n^{**}(i),n^*(i)}[\varepsilon](i,B) \cap \boldsymbol{I}_{i+1,n^*(i)}[\varepsilon](i,G) = \emptyset. \tag{7.7}$$

Proof: By Farkas Lemma,[2] for (7.6) and (7.7), it suffices to show that

1. there exists $\mathbf{x} \in \mathbb{R}^{L_{i+1}(i)+L_{n^{**}(i)}(i)}$ such that for all $\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}'$ with $\|\boldsymbol{\varepsilon}\| \leq \varepsilon$ and $\|\boldsymbol{\varepsilon}'\| \leq \varepsilon$,

$$\left\{ \begin{array}{l} \begin{bmatrix} I_{i+1}(i)Q_{i+1,n^*(i)}(i) \\ I_{n^{**}(i)}(i)Q_{n^{**}(i),n^*(i)}(i) \end{bmatrix}' \mathbf{x} \leq \mathbf{0}, \\ \left(\begin{bmatrix} q_2\mathbf{1} \\ q_1\mathbf{1} \end{bmatrix} + I_n(i) \begin{bmatrix} \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon}' \end{bmatrix} \right) \cdot \mathbf{x} > 0 \end{array} \right.$$

   and

2. there exists $\mathbf{x} \in \mathbb{R}^{L_{i+1}(i)+L_{n^{**}(i)}(i)}$ such that, for all $\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}'$ with $\|\boldsymbol{\varepsilon}\| \leq \varepsilon$ and $\|\boldsymbol{\varepsilon}'\| \leq \varepsilon$,

$$\left\{ \begin{array}{l} \begin{bmatrix} I_{i+1}(i)Q_{i+1,n^*(i)}(i) \\ I_{n^{**}(i)}(i)Q_{n^{**}(i),n^*(i)}(i) \end{bmatrix}' \mathbf{x} \leq \mathbf{0}, \\ \left(\begin{bmatrix} q_1\mathbf{1} \\ q_2\mathbf{1} \end{bmatrix} + I_n(i) \begin{bmatrix} \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon}' \end{bmatrix} \right) \cdot \mathbf{x} > 0 \end{array} \right.$$

With $\varepsilon = 0$, this is equivalent to

[2] Farkas Lemma has a constraint that each element of $\mathbf{x}$ should be non-negative. However, since we have equality constraints in the definition of $\boldsymbol{I}_{n,n^*(i)}[\varepsilon](i,\hat{w}_i)$, not inequality constraints, for each l with $x_l < 0$, we can multiply -1 to x_l, each element of lth row of the matrix in the LHS, and lth element of the vector in RHS. Therefore, non-negativity constraint is redundant.

- there exists $\mathbf{x} \in \mathbb{R}^{L_{i+1}(i)+L_{n^{**}(i)}(i)}$ such that

$$\begin{bmatrix} I_{i+1}(i)Q_{i+1,n^*(i)}(i) \\ I_{n^{**}(i)}(i)Q_{n^{**}(i),n^*(i)}(i) \end{bmatrix}' \mathbf{x} \leq \mathbf{0}, \begin{bmatrix} q_2\mathbf{1} \\ q_1\mathbf{1} \end{bmatrix} \cdot \mathbf{x} > 0$$

  and

- there exists $\mathbf{x} \in \mathbb{R}^{L_{i+1}(i)+L_{n^{**}(i)}(i)}$ such that

$$\begin{bmatrix} I_{i+1}(i)Q_{i+1,n^*(i)}(i) \\ I_{n^{**}(i)}(i)Q_{n^{**}(i),n^*(i)}(i) \end{bmatrix}' \mathbf{x} \leq \mathbf{0}, \begin{bmatrix} q_1\mathbf{1} \\ q_2\mathbf{1} \end{bmatrix} \cdot \mathbf{x} > 0.$$

By Assumption 49, we can find such $\mathbf{x}$. Fix $\mathbf{x}$. Then, now that $\mathbf{x}$ is fixed, by the continuity of the linear function, for sufficiently small ε, we have Conditions 1 and 2 as desired. ∎

7.1.2 Assumption for Dispensing with the Error-Reporting Noisy Cheap Talk

We explain how player j sends a binary message $m \in \{G, B\}$ to player n via actions instead of the error-reporting noisy cheap talk. Since we only use the error-reporting noisy cheap talk with precision $p = 1 - \exp(-T^{\frac{1}{2}})$, we concentrate on the case with $p = 1 - \exp(-T^{\frac{1}{2}})$.

As in the two-player case, with η being a small number to be defined, the sender (player j) determines

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

and player j takes

$$\alpha_j^{z_j(m)} = \begin{cases} (1-\rho)\, a_i^G + \rho a_i^B & \text{if } z_j(m) = G, \\ (1-\rho)\, a_i^B + \rho a_i^G & \text{if } z_j(m) = B, \\ \frac{1}{2} a_j^G + \frac{1}{2} a_j^B & \text{if } z_j(m) = M \end{cases}$$

for $T^{\frac{1}{2}}$ periods with $\rho \leq \frac{1}{3}$. Player n (receiver) takes

$$\bar{\alpha}_n = (1 - 2(|A_n| - 1)\rho) a_n^G + \sum_{a_n \neq a_n^G} 2\rho a_n$$

Player $i \in -(j,n)$ takes $\alpha_i^{\text{receive}}$ (compared to $\bar{\alpha}_n$, $\alpha_i^{\text{receive}}$ is independent of ρ).

For each $i \in I$, let $h_i(r)$ be the realized frequency of player i's signal-action pair in round r where player j sends m. We want to construct $f[n](m) \in \{G, B\}$ from $h_n(r)$ and $g[n-1](m) \in \{m, E\}$ from $h_{n-1}(r)$ such that

- player $n-1$ has $g[n-1](m) = m$ with a high probability;
- player n cannot manipulate $g[n-1](m)$;
- player n infers the message correctly ($f[n](m) = m$) with a high probability;
- given m, player n believes that $f[n](m) = m$ or $g[n-1](m) = E$ with a high probability; and
- player $i \in -(j,n)$ (other than the sender and receiver) cannot manipulate $f[n](m)$ to increase her own payoff.

Player $n-1$ has $g[n-1](m) = E$ if and only if at least one of the following three conditions are satisfied:

1. as in the two player case, $z_j(m) \neq m$;
2. player j's action frequency is not close to α_j^m; and
3. player j's signal frequency while player j takes a_j^m is not close to the affine hull of player j's signal distributions with respect to player n's action $\text{aff}(\{\mathbf{q}_j(a_j^m, a_n, \alpha_{-(j,n)}^{\text{receive}})\}_{a_n})$.

As for $\boldsymbol{Q}_j^n(x)$ in Chapter 6, let $\boldsymbol{Q}_j(j \to_m n) \equiv \text{aff}(\{\mathbf{q}_j(a_j^m, a_n, \alpha_{-(j,n)}^{\text{receive}})\}_{a_n}) \cap \mathbb{R}_+^{|Y_j|}$. Let

$$\boldsymbol{Q}_j(j \to_m n) = \left\{\mathbf{y}_j \in \mathbb{R}_+^{|Y_j|} : Q_j(j \to_m n)\mathbf{y}_j = \mathbf{q}_j(j \to_m n)\right\}. \tag{7.8}$$

As in Lemma 14, we can assume that each element in $Q_j(j \to_m n)$ is in $(0,1)$.

Perturb $\boldsymbol{Q}_j(j \to_m n)$ by ε so that

$$\boldsymbol{Q}_j[\varepsilon](j \to_m n) = \left\{ \begin{array}{c} \mathbf{y}_j \in \mathbb{R}_+^{|Y_j|} : \exists \boldsymbol{\varepsilon} \in \mathbb{R}^{|Y_j|} \text{ such that} \\ \left\{ \begin{array}{c} \|\boldsymbol{\varepsilon}\| \leq \varepsilon, \\ Q_j(j \to_m n)\left(\mathbf{y}_j + \boldsymbol{\varepsilon}\right) = \mathbf{q}_j(j \to_m n). \end{array} \right. \end{array} \right\} \tag{7.9}$$

By the law of large numbers, player $n-1$ has $g[n-1](m) = m$ with a high probability. In addition, since we take the affine hull with respect to player n's actions, player n cannot manipulate $g[n-1](m)$.

$f[n](m)$ is determined as follows:

1. if the following two conditions are satisfied, then player n infers $f[n](m) = \hat{m} \in \{G, B\}$:

   (a) player n's action frequency is close to $\bar{\alpha}_n$; and

   (b) there exists $i \in -(j,n)$ and $\hat{m} \in \{G, B\}$ such that player n's signal frequency while player n takes a_n^G is close to the affine hull of player n's signal distributions with respect to player i's action $\text{aff}(\{\mathbf{q}_n(a_j^{\hat{m}}, a_n^G, a_i, \alpha_{-(i,j,n)}^{\text{receive}})\}_{a_i})$.

   As we will see in Lemma 52, there is at most one $\hat{m} \in \{G, B\}$ such that $i \in -(j,n)$ and $\hat{m} \in \{G, B\}$ satisfy 1-(b); and

2. otherwise, player n infers $f[n](m)$ from the likelihood as in the two-player case

As for $\boldsymbol{Q}_j(j \to_m n)$, let $\boldsymbol{Q}_n^i(j \to_{\hat{m}} n) \equiv \text{aff}(\{\mathbf{q}_n(a_j^{\hat{m}}, a_n^G, a_i, \alpha_{-(i,j,n)}^{\text{receive}})\}_{a_i}) \cap \mathbb{R}_+^{|Y_n|}$ and

$$\boldsymbol{Q}_n^i[\varepsilon](j \to_{\hat{m}} n) \equiv \left\{ \begin{array}{c} \mathbf{y}_n \in \mathbb{R}_+^{|Y_n|} : \exists \boldsymbol{\varepsilon} \in \mathbb{R}^{|Y_n|} \text{ such that} \\ \left\{ \begin{array}{c} \|\boldsymbol{\varepsilon}\| \leq \varepsilon, \\ Q_n^i(j \to_{\hat{m}} n)\mathbf{y}_n + \boldsymbol{\varepsilon} = \mathbf{q}_n^i(j \to_{\hat{m}} n). \end{array} \right. \end{array} \right\} \tag{7.10}$$

As in Lemma 14, we can assume that each element in $Q_n(j \to_{\hat{m}} n)$ is in $(0,1)$.

Consider the properties of $f[n](m)$. Suppose Case 1 of $f[n](m)$ is the case. Then, since 1-(a) is the case, for sufficiently small ρ, player n takes a_n^G for most of the time. For simplicity, let us proceed as if player n took a_n^G all the time.

Below, we assume that there are no multiple $\hat{m}$'s in Case 1:

$$\left(\bigcup_{i \in -(j,n)} \boldsymbol{Q}_n^i[\varepsilon](j \to_G n) \right) \cap \left(\bigcup_{i \in -(j,n)} \boldsymbol{Q}_n^i[\varepsilon](j \to_B n) \right) = \emptyset$$

for sufficiently small ε. That is, $f[n](m)$ is uniquely determined in Case 1.

Suppose player j takes α_j^m (otherwise, $g[n-1](m) = E$). By Condition 2 of $g[n-1](m) = E$, for sufficiently small ρ, player n can believe that player j takes a_j^m for most of the time. For simplicity, let us proceed as if player j took a_j^m all the time.

Then, player n believes that, conditional on $\hat{m}$, $g[n-1](\hat{m}) = \hat{m}$ with a non-negligible probability only if player n's signal frequency is included in

$$\boldsymbol{H}_n[\varepsilon](j \to_{\hat{m}} n) = \left\{ \begin{array}{c} \mathbf{y}_n \in \mathbb{R}_+^{|Y_n|} : \exists \boldsymbol{\varepsilon} \in \mathbb{R}^{|Y_n|} \text{ such that} \\ \left\{ \begin{array}{c} \|\boldsymbol{\varepsilon}\| \leq \varepsilon, \\ Q_j(j \to_{\hat{m}} n)\left(M_{j,n}(\hat{m})\mathbf{y}_n + \boldsymbol{\varepsilon}\right) = \mathbf{q}_j(j \to_{\hat{m}} n). \end{array} \right. \end{array} \right\} \tag{7.11}$$

where

$$M_{j,n}(\hat{m}) = \left[\begin{array}{ccc} q(y_{j,1} \mid a_j^{\hat{m}}, a_n^G, \alpha_{-(j,n)}^{\text{receive}}, y_{n,1}) & \cdots & q(y_{j,1} \mid a_j^{\hat{m}}, a_n^G, \alpha_{-(j,n)}^{\text{receive}}, y_{n,|Y_n|}) \\ \vdots & & \vdots \\ q(y_{j,|Y_j|} \mid a_j^{\hat{m}}, a_n^G, \alpha_{-(j,n)}^{\text{receive}}, y_{n,1}) & \cdots & q(y_{j,|Y_j|} \mid a_j^{\hat{m}}, a_n^G, \alpha_{-(j,n)}^{\text{receive}}, y_{n,|Y_n|}) \end{array} \right]$$

is the matrix which projects player n's signal observation to the conditional distribution of player j's signal observation.

In other words, if

$$\begin{aligned}&\boldsymbol{H}_n[\varepsilon](j \to_G n) \cap \left(\bigcup_{i\in -(j,n)} \boldsymbol{Q}_n^i[\varepsilon](j \to_B n)\right)\\ = &\boldsymbol{H}_n[\varepsilon](j \to_B n) \cap \left(\bigcup_{i\in -(j,n)} \boldsymbol{Q}_n^i[\varepsilon](j \to_G n)\right) = \emptyset\end{aligned}$$

for sufficiently small ε, then player n inferring $f[n](m)$ according to Case 1 believes that, conditional on m, either $f[n](m) = m$ or $g[n-1](m) = E$ with a non-negligible probability, as desired.

As in the two-player case, player n inferring $f[n](m)$ according to Case 2 believes that, conditional on m, either $f[n](m) = m$ or $g[n-1](m) = E$ with a non-negligible probability, as desired.

Therefore, in total, given m, player n believes that $f[n](m) = m$ or $g[n-1](m) = E$ with a high probability. In addition, player n infers the message correctly with a high probability by the law of large numbers.

Hence, we are left to show player $i \in -(j,n)$ (other than the sender and receiver) cannot manipulate $f[n](m)$ to increase her payoff. For this purpose, define $\theta_{i-1}(j \to n) \in \{G, B\}$ as follows: $\theta_{i-1}(j \to n) = B$ if and only if $i \in -(j,i)$ and Case 1 is not the case, that is,

1. player n's action frequency is not close to $\bar{\alpha}_n$; or
2. player n's signal frequency while player n takes a_n^G is not close to the affine hull of player n's signal distributions with respect to player i's action $\text{aff}(\{\mathbf{q}_j(a_j^m, a_n^G, a_i, \alpha_{-(i,j,n)}^{\text{receive}})\}_{a_i})$.

If $\theta_{i-1}(j \to n) = B$ happens, then player $i-1$ makes player i indifferent between any action profile from the next review round.

Note that unless player n infers $f[n](m) = m$, $\theta_{i-1}(j \to n) = B$. Further, since we take the affine hull of player i, the distribution of $\theta_{i-1}(j \to n)$ is independent of player i's strategy.[3] Therefore, player i cannot manipulate $f[n](m)$ and increase her payoff.

In total, we want to assume that, for sufficiently small ε,

[3]Precisely speaking, we need to do a small adjustment as (2.22). See Section 7.4.2.

1.

$$\left(\bigcup_{i\in-(j,n)} \boldsymbol{Q}_n^i[\varepsilon](j\to_G n)\right)\cap\left(\bigcup_{i\in-(j,n)} \boldsymbol{Q}_n^i[\varepsilon](j\to_B n)\right)=\emptyset; \tag{7.12}$$

and

2.

$$\begin{aligned}\boldsymbol{H}_n[\varepsilon](j &\to {}_G n)\cap\left(\bigcup_{i\in-(j,n)} \boldsymbol{Q}_n^i[\varepsilon](j\to_B n)\right)\\ &= \boldsymbol{H}_n[\varepsilon](j\to_B n)\cap\left(\bigcup_{i\in-(j,n)} \boldsymbol{Q}_n^i[\varepsilon](j\to_G n)\right)=\emptyset. \end{aligned}\tag{7.13}$$

For (7.12) and (7.13), it suffices to assume that following:

Assumption 51 There exist $\left\{a_i^G, a_i^B, \bar{\alpha}_i, \alpha_i^{\text{receive}}\right\}_{i\in I}$ such that, for each $j\in I$ and $n\in -j$, the following five conditions are satisfied:

1. for all $i,i'\in -n$, there is $\mathbf{x}\in\mathbb{R}^{2|Y_n|-|A_i|-|A_{i'}|+2}$ such that

$$\begin{bmatrix} Q_n^i(j\to_G n)\\ Q_n^{i'}(j\to_B n)\end{bmatrix}'\mathbf{x}\le 0, \begin{bmatrix} \mathbf{q}_n^i(j\to_G n)\\ \mathbf{q}_n^{i'}(j\to_B n)\end{bmatrix}\cdot\mathbf{x}<0;$$

2. for all $i\in -n$, there is $\mathbf{x}\in\mathbb{R}^{|Y_j|+|Y_n|-|A_n|-|A_i|+2}$ such that

$$\begin{bmatrix} Q_j(j\to_G n)M_{j,n}(G)\\ Q_n^i(j\to_B n)\end{bmatrix}'\mathbf{x}\le 0, \begin{bmatrix} \mathbf{q}_j(j\to_G n)\\ \mathbf{q}_n^i(j\to_B n)\end{bmatrix}\cdot\mathbf{x}<0;$$

3. for all $i\in -n$, there is $\mathbf{x}\in\mathbb{R}^{|Y_j|+|Y_n|-|A_n|-|A_i|+2}$ such that

$$\begin{bmatrix} Q_j(j\to_G n)M_{j,n}(B)\\ Q_n^i(j\to_G n)\end{bmatrix}'\mathbf{x}\le 0, \begin{bmatrix} \mathbf{q}_j(j\to_B n)\\ \mathbf{q}_n^i(j\to_G n)\end{bmatrix}\cdot\mathbf{x}<0;$$

4. for each $a_n \in A_i$ and $y_n \in Y_i$, we have

$$q(y_n \mid a_n, a_j^G, \alpha_{-(n,j)}^{\text{receive}}) \neq q(y_n \mid a_n, a_j^B, \alpha_{-(n,j)}^{\text{receive}});$$

and

5. for all $i \in I$,

$$j' = \begin{cases} i-1 & \text{if } i \neq 1, \\ 2 & \text{if } i = 1, \end{cases}$$

$a_{-(i,j')} \in A_{-(i,j')}$ and $\alpha_{j'} \in \Delta(A_{j'})$ with

   (a) $\alpha_{j'}(a_{j'}^G) > 0$ and $\alpha_{j'}(a_{j'}^B) > 0$ if $j' = j$; and

   (b) $\alpha_{j'} = \bar{\alpha}_{j'}$ or $\alpha_{j'}^{\text{receive}}$ if $j' \neq j$,

for all (a_i, y_i) and (a_i', y_i') with $(a_i, y_i) \neq (a_i', y_i')$, we have

$$\mathbb{E}\left[\mathbf{1}_{a_{j'},y_{j'}} \mid y_i, a_i, \alpha_{j'}, a_{-(i,j')}\right] \neq \mathbb{E}\left[\mathbf{1}_{a_{j'},y_{j'}} \mid y_i', a_i', \alpha_{j'}, a_{-(i,j')}\right].$$

For notational simplicity, we assume that $\left(a_i^G, a_i^B, \alpha_i^{\text{receive}}\right)_{i\in I}$ in Assumption 49 satisfies Assumption 51.

Condition 4 is equivalent to Condition 1 of Assumption 39 in the two-player case. As we will see in Section 7.7, Condition 5 will be used to show the truthtelling incentive in the report block.

Under Assumption 6, Assumption 51 is generic. First, consider Condition 1. Note that there are $|Y_n| + 1$ constraints. On the other hand, after noting that the signal frequency is on the simplex over Y_n, $\mathbf{x}$ has $2|Y_n| - |A_i| - |A_{i'}| + 1$ degrees of freedom. By Assumption 6, since $2|Y_n| - |A_i| - |A_{i'}| + 1 \geq |Y_n| + 1$, $\mathbf{x}$ satisfying Condition 2 generically exists.

Second, consider Conditions 2. Note that there are $|Y_n|+1$ constraints while $\mathbf{x}$ has $|Y_j|+|Y_n|-|A_n|-|A_i|+1$ degrees of freedom. By Assumption 6, since $|Y_j|+|Y_n|-|A_n|-|A_i|+1 \geq |Y_n|+1$, $\mathbf{x}$ satisfying Condition 2 generically exists. Condition 3 is symmetric.

Finally, Conditions 4 and 5 are generic.

The next lemma shows that Assumption 49 is sufficient for (7.12) and (7.13):

Lemma 52 If Assumption 49 is satisfied, then there exists $\bar{\varepsilon} > 0$ such that, for all $\varepsilon < \bar{\varepsilon}$, (7.12) and (7.13) are satisfied.

Proof: It suffices to show that

1. for all $i, i' \in -n$, for sufficiently small ε,

$$\boldsymbol{Q}_n^i[\varepsilon](j \to_G n) \cap \boldsymbol{Q}_n^{i'}[\varepsilon](j \to_B n) = \emptyset;$$

2. for all $i \in -n$, for sufficiently small ε,

$$\boldsymbol{H}_n[\varepsilon](j \to_G n) \cap \boldsymbol{Q}_n^i[\varepsilon](j \to_B n) = \emptyset;$$

and

3. for all $i \in -n$, for sufficiently small ε,

$$\boldsymbol{H}_n[\varepsilon](j \to_B n) \cap \boldsymbol{Q}_n^i[\varepsilon](j \to_G n) = \emptyset.$$

They are equivalent to

1. for all $i, i' \in -n$, for sufficiently small ε, for all $\boldsymbol{\varepsilon} \in \mathbb{R}^{2|Y_n|-|A_i|-|A_{i'}|+2}$ with $\|\boldsymbol{\varepsilon}\| \leq \varepsilon$, there is no $\mathbf{y}_n \in \mathbb{R}_+^{|Y_n|}$ such that

$$\begin{bmatrix} Q_n^i(j \to_G n) \\ Q_n^{i'}(j \to_B n) \end{bmatrix} \mathbf{y}_n = \begin{bmatrix} \mathbf{q}_n^i(j \to_G n) \\ \mathbf{q}_n^{i'}(j \to_B n) \end{bmatrix} + \boldsymbol{\varepsilon};$$

2. for all $i \in -n$, for sufficiently small ε, for all $\boldsymbol{\varepsilon} \in \mathbb{R}^{|Y_j|+|Y_n|-|A_n|-|A_i|+2}$ with $\|\boldsymbol{\varepsilon}\| \leq \varepsilon$, there is no $\mathbf{y}_n \in \mathbb{R}_+^{|Y_n|}$ such that

$$\begin{bmatrix} Q_j(j \to_G n)M_{j,n}(G) \\ Q_n^i(j \to_B n) \end{bmatrix} \mathbf{y}_n = \begin{bmatrix} \mathbf{q}_j(j \to_G n) \\ \mathbf{q}_n^i(j \to_B n) \end{bmatrix} + \boldsymbol{\varepsilon};$$

and

3. for all $i \in -n$, for sufficiently small ε, for all $\boldsymbol{\varepsilon} \in \mathbb{R}^{|Y_j|+|Y_n|-|A_n|-|A_i|+2}$ with $\|\boldsymbol{\varepsilon}\| \leq \varepsilon$, there is no $\mathbf{y}_n \in \mathbb{R}_+^{|Y_n|}$ such that

$$\begin{bmatrix} Q_j(j \to_B n)M_{j,n}(B) \\ Q_n^i(j \to_G n) \end{bmatrix} \mathbf{y}_n = \begin{bmatrix} \mathbf{q}_j(j \to_B n) \\ \mathbf{q}_n^i(j \to_G n) \end{bmatrix} + \boldsymbol{\varepsilon}.$$

By Farkas Lemma, they are equivalent to

1. for all $i, i' \in -n$, for sufficiently small ε, for all $\boldsymbol{\varepsilon} \in \mathbb{R}^{2|Y_n|-|A_i|-|A_{i'}|+2}$ with $\|\boldsymbol{\varepsilon}\| \leq \varepsilon$, there is $\mathbf{x} \in \mathbb{R}^{2|Y_n|-|A_i|-|A_{i'}|+2}$ such that

$$\begin{bmatrix} Q_n^i(j \to_G n) \\ Q_n^{i'}(j \to_B n) \end{bmatrix}' \mathbf{x} \leq 0, \left(\begin{bmatrix} \mathbf{q}_n^i(j \to_G n) \\ \mathbf{q}_n^{i'}(j \to_B n) \end{bmatrix} + \boldsymbol{\varepsilon} \right) \cdot \mathbf{x} < 0;$$

2. for all $i \in -n$, for sufficiently small ε, for all $\boldsymbol{\varepsilon} \in \mathbb{R}^{|Y_j|+|Y_n|-|A_n|-|A_i|+2}$ with $\|\boldsymbol{\varepsilon}\| \leq \varepsilon$, there is $\mathbf{x} \in \mathbb{R}^{|Y_j|+|Y_n|-|A_n|-|A_i|+2}$ such that

$$\begin{bmatrix} Q_j(j \to_G n)M_{j,n}(G) \\ Q_n^i(j \to_B n) \end{bmatrix}' \mathbf{x} \leq 0, \left(\begin{bmatrix} \mathbf{q}_j(j \to_G n) \\ \mathbf{q}_n^i(j \to_B n) \end{bmatrix} + \boldsymbol{\varepsilon} \right) \cdot \mathbf{x} < 0;$$

and

3. for all $i \in -n$, for sufficiently small ε, for all $\boldsymbol{\varepsilon} \in \mathbb{R}^{|Y_j|+|Y_n|-|A_n|-|A_i|+2}$ with $\|\boldsymbol{\varepsilon}\| \leq \varepsilon$, there is $\mathbf{x} \in \mathbb{R}^{|Y_j|+|Y_n|-|A_n|-|A_i|+2}$ such that

$$\begin{bmatrix} Q_j(j \to_B n) M_{j,n}(B) \\ Q_n^i(j \to_G n) \end{bmatrix}' \mathbf{x} \leq 0, \left(\begin{bmatrix} \mathbf{q}_j(j \to_B n) \\ \mathbf{q}_n^i(j \to_G n) \end{bmatrix} + \boldsymbol{\varepsilon} \right) \cdot \mathbf{x} < 0.$$

The rest of the proof is the same as Lemma 50. ∎

7.1.3 Assumptions for Dispensing with the Public Randomization and Perfect Cheap Talk

First, to dispense with the public randomization, we want to make an assumption comparable to Assumption 40 in the two-player case. For each $i \in I$, with player j replaced with player $i-1$ (the controller of player i's payoff), all the definitions about $a^{\text{p.r.}(i)}$, $\alpha^{\text{p.r.}(i)}$, $\Phi_{i-1}^{\text{p.r.}(i)}$, (6.2), (6.3), (6.4), (6.5), $H_{i,1}^i$ and $H_{i,2}^i$ in Section 6.1.2 are valid with more than two players.

Now, we formally state the analogue of Assumption 40 for more than two players:

Assumption 53 For each $i \in I$, one of the following two conditions is satisfied:

1. there exists $a^{\text{p.r.}(i)} \in A$ such that there exist $Y_{i-1,1}^i$, $Y_{i-1,2}^i$, $\bar{p}_i$, $Y_{i,1}^i$ and $Y_{i,2}^i$ such that $Y_{i,1}^i$ and $Y_{i,2}^i$ satisfy (6.2) and (6.3) with j replaced with $i-1$ and

$$Y_{i,1}^i \neq \emptyset, Y_{i,2}^i \neq \emptyset, Y_i = Y_{i,1}^i \cup Y_{i,2}^i, Y_{i-1} = Y_{i-1,1}^i \cup Y_{i-1,2}^i;$$

and

2. there exists $\alpha^{\text{p.r.}(i)} \in \Delta(A)$ such that there exist $\phi_{i-1}^{\text{p.r.}(i)}$, $\bar{p}_i$, $H_{i,1}^i$ and $H_{i,2}^i$ such that

   (a) $H_{i,1}^i$ and $H_{i,2}^i$ satisfy (6.4), (6.5) with j replaced with $i-1$ and

$$H_{i,1}^i \neq \emptyset, H_{i,2}^i \neq \emptyset, A_i \times Y_i = H_{i,1}^i \cup H_{i,2}^i;$$

and

(b) for all $n \in I$, given $\alpha_{-n}^{\text{p.r.}(i)}$, the probability that player i observes $(a_i, y_i) \in A_i \times Y_i$ with $(a_i, y_i) \in H_{i,1}^i$ is independent of $a_n \in A_n$.

Second, when player i with $i \geq 2$ sends the message, player $i-1$ wants to construct a statistics $\phi_{i-1}(a_{i-1}, y_{i-1})$ such that player $i-1$ can infer player i's message statistically and that the conditional independence property holds for player i, as $\phi_j(a_j, y_j)$ in Lemma 42: for $a_i^G \in A_i$ and $\alpha_{-i}^{\text{report}}$, for all $(a_i, y_i) \in A_i \times Y_i$,

$$\mathbb{E}\left[\phi_{i-1}(a_{i-1}, y_{i-1}) \mid \alpha_{-i}^{\text{report}}, a_i, y_i\right] = \begin{cases} q_2 & \text{if } a_i = a_i^G, \\ q_1 & \text{if } a_i \neq a_i^G. \end{cases} \tag{7.14}$$

A sufficient condition for the existence of such $\phi_{i-1}(a_{i-1}, y_{i-1})$ is the linear independence of $(\Pr(a_{i-1}, y_{i-1} \mid \alpha_{-i}^{\text{report}}, a_i, y_i))_{a_{i-1}, y_{i-1}}$ with respect to $a_i \in A_i$ and $y_i \in Y_i$:

Assumption 54 There exists $\{\alpha_j^{\text{report}}\}_{j \in I}$ such that, for each $i \geq 2$, $(\Pr(a_{i-1}, y_{i-1} \mid \alpha_{-i}^{\text{report}}, a_i, y_i))_{a_{i-1}, y_{i-1}}$ is linearly independent with respect to $a_i \in A_i$ and $y_i \in Y_i$.

Note that this is generic since we assume (7.1) and that we do not assume the counterpart for player 1 to send the message.

The following lemma shows that Assumption 41 is sufficient to have ϕ_{i-1} with (7.14).

Lemma 55 If Assumption 54 is satisfied with $\left\{\alpha_j^{\text{report}}\right\}_{j \in I}$, then there exist $q_2 > q_1$ such that for all $i \in \{2, ..., N\}$, there exist $\phi_{i-1} : A_{i-1} \times Y_{i-1} \to (0, 1)$ such that (7.14) is satisfied.

Proof: The same as Lemma 42. ∎

7.2 Coordination Block with the Error-Reporting Noisy Cheap Talk

Now, we consider the coordination block without the perfect cheap talk but with the error-reporting noisy cheap talk with precision $p = 1 - \exp(-T^{\frac{1}{2}})$.

As mentioned in Section 7.1.1, we define

$$\begin{aligned} N(i) &= \{i, i+1, i+2\}, \\ n^*(i) &\in \arg\min_{j \in \{i, i+2\}} |A_j| \, |Y_j| \, , \\ n^{**}(i) &= \{i, i+2\} \setminus \{n^*(i)\}. \end{aligned}$$

First, player i sends the message about $x_i \in \{G, B\}$ to player $n^*(i)$ by the error-reporting noisy cheap talk and let $w_i = f[n^*(i)](x_i) \in \{G, B\}$ be player $n^*(i)$'s inference of this message. Second, player $n^*(i)$ sends the message about w_i to players $N(i)$ via actions. Each player $n \in N(i)$ constructs player n's inference of w_i, denoted by $w_i(n) \in \{G, M, B\}$. Here, the inference M ("middle") is introduced so that it prevents player $n^*(i)$ from creating a situation where player $n^*(i)$ is pivotal. See 7.1.1 for the definition of "pivotal."

7.2.1 Structure of the Coordination Block

Formally, the coordination block proceeds as follows:

- the periods where the players coordinate on x_1:
  - the coordination round 1 for x_1. Player 1 sends the message about x_1 to player $n^*(1)$ via error-reporting noisy cheap talk. If $n^*(1) = 1$, then this round does not exist;
  - the coordination round 2 for x_1. Player $n^*(1)$ sends the message about w_1 to players $N(1)$ via actions. Player $n \in N(1)$ creates the inference $w_1(n)$; and

  - for each $j \in N(1) = \{1, 2, 3\}$ and $n \in -j$, we have the coordination round 3 for x_1 between j and n, where player j sends the message $w_1(j)$ to player n via error-reporting noisy cheap talk. The players take turns: first, player 1 sends $w_1(1)$ to player 2, second, player 1 sends $w_1(1)$ to player 3, and so on until player 1 sends $w_1(1)$ to player N. Then, player 2 sends $w_1(2)$ to player 1, and so on until player 2 sends $w_1(2)$ to player N. After player 2, player 3 sends $w_1(3)$ to each of the opponents -3 sequentially;

⋮

- the periods where the players coordinate on x_i:

  - the coordination round 1 for x_i. Player i sends the message about x_i to player $n^*(i)$ via error-reporting noisy cheap talk. If $n^*(i) = i$, then this round does not exist;

  - the coordination round 2 for x_i. Player $n^*(i)$ sends the message about w_i to players $N(i)$ via actions. Player $n \in N(i)$ creates the inference $w_i(n)$; and

  - for each $j \in N(i)$ and $n \in -j$, we have the coordination round 3 for x_i between j and n, where player j sends the message $w_i(j)$ to player n via error-reporting noisy cheap talk. Again, the players take turns;

⋮

- the periods where the players coordinate on x_N:

  - the coordination round 1 for x_N. Player N sends the message about x_N to player $n^*(N)$ via error-reporting noisy cheap talk. If $n^*(N) = N$, then this round does not exist;

  - the coordination round 2 for x_N. Player $n^*(N)$ sends the message about w_N to players $N(N)$ via actions. Player $n \in N(N)$ creates the inference $w_N(n)$; and

– for each $j \in N(N)$ and $n \in -j$, we have the coordination round 3 for x_N between j and n, where player j sends the message $w_N(j)$ to player n via error-reporting noisy cheap talk. Again, the players take turns.

For notational convenience, let $T(n^*(i) \to_{w_i} N(i))$ be the set of periods in the coordination round 2 for x_i, where player $n^*(i)$ sends the message w_i to players $N(i)$ via actions. As in the review round, player n randomly picks $t_n(n^*(i) \to_{w_i} N(i))$ from $T(n^*(i) \to_{w_i} N(i))$ and does not use period $t_n(n^*(i) \to_{w_i} N(i))$ for inference. That is, player n uses the periods in

$$T_n(n^*(i) \to_{w_i} N(i)) \equiv T(n^*(i) \to_{w_i} N(i)) \setminus \{t_n(n^*(i) \to_{w_i} N(i))\}.$$

We explain each round in the sequel.

7.2.2 Coordination Round 1 for x_i

If player i is the same person as player $n^*(i)$, then this round does not exist. Let $w_i = x_i$ be player $n^*(i)$'s inference (player i's inference in other words).

Otherwise, player i sends x_i by the error-reporting noisy cheap talk and player $n^*(i)$ creates the inference of x_i denoted by w_i as $w_i = f[n^*(i)](x_i)$.

7.2.3 Coordination Round 2 for x_i

This is the round where player $n^*(i)$ sends w_i to players $N(i)$. Player $n^*(i)$ takes $a_{n^*(i)}^{w_i}$ and each player $j \in -i$ takes $\alpha_j^{\text{receive}}$ for $T^{\frac{1}{2}}$ periods, defined in Assumption 49. Remember that $T(n^*(i) \to_{w_i} N(i))$ be the set of periods in this round.

See (7.4) and (7.5) for the definition of the $L_n(i) \times |A_n| |Y_n|$ matrix $I_n(i)$. Based on $I_n(i)$, each player $n \in N(i) \setminus \{n^*(i)\}$ constructs a random variable $\mathbf{1}_{I_{n,t(i)}}$ ($L_n(i) \times 1$ vector) as follows: after taking a_n and observing y_n, player n calculates $I_n(i)\,\mathbf{1}_{a_n,y_n}$. Here, $\mathbf{1}_{a_n,y_n}$ is a $|A_n| |Y_n| \times 1$ vector such that the element corresponding to a_n, y_n is equal to 1 and the other elements are 0. Hence, $I_n(i)\,\mathbf{1}_{a_n,y_n}$ is a $L_n(i) \times 1$ vector. Then, player n draws $L_n(i)$ random variables independently from the uniform distribution on $[0,1]$. If the lth realization of these random variables is less than the lth element of $I_n(i)\,\mathbf{1}_{a_n,y_n}$, then the lth element of $\mathbf{1}_{I_n(i)}$ is equal to 1. Otherwise, the lth element of $\mathbf{1}_{I_n(i)}$ is equal to 0. We have

$$\Pr(\{(\mathbf{1}_{I_n(i)})_l = 1\} \mid a, y) = \boldsymbol{i}_n^l(i)\mathbf{1}_{a_n,y_n}. \tag{7.15}$$

Given $\{\mathbf{1}_{I_{n,t}(i)}\}_{t\in T_n(n^*(i)\to_{w_i} N(i))}$, player $n \in N(i)$ infers w_i as follows:

1. player $n^*(i)$ infers her own message straightforwardly: $w_i(n^*(i)) = w_i$; and
2. player $n \in N(i) \setminus \{n^*(i)\}$ infers as follows:
   (a) if
   $$\left\| \frac{1}{T^{\frac{1}{2}}-1} \sum_{t\in T_n(n^*(i)\to_{w_i} N(i))} \mathbf{1}_{I_{n,t}(i)} - q_2\mathbf{1} \right\| \le \varepsilon,$$
   then $w_i(n) = G$;
   (b) if
   $$\left\| \frac{1}{T^{\frac{1}{2}}-1} \sum_{t\in T_n(n^*(i)\to_{w_i} N(i))} \mathbf{1}_{I_{n,t}(i)} - q_1\mathbf{1} \right\| \le \varepsilon,$$
   then $w_i(n) = B$; and
   (c) otherwise, $w_i(n) = M$ (the posterior is not skewed enough for $w_i = G$ or $w_i = B$ and so player $n^*(i)$ infers that the message is "middle").

Assumption 49 implies the following Lemma:

Lemma 56 For any $\varepsilon < \bar{\varepsilon}$, for any $i \in I$ and $w_i \in \{G, B\}$,

1. for any $n \in N(i^*)$,
   (a) if players $n^*(i)$ and n follow the equilibrium strategy, then
   $$\Pr(\{w_i(n) = w_i\} \mid w_i) \ge 1 - \exp(-\Theta(T^{\frac{1}{2}}));$$

and

(b) the distribution of $w_i(n)$ given w_i is independent of player $j = N(i) \setminus \{n^*(i), n\}$'s unilateral deviation; and

2. for any history of player $n^*(i)$ at the end of the coordination round 2 for x_i, player $n^*(i)$ puts a belief $\exp(-\Theta(T^{\frac{1}{2}}))$ on the event

$$\{G, B\} \ni w_i(n^{**}(i)) \neq w_i(i+1) \in \{G, B\}.$$

Proof: 1. Follows from (7.4) and (7.15).

2. By Hoeffding's inequality, there exists K_1 such that, for each $k > 0$, for any frequency of action-signal pair $\mathbf{x}_{n^*(i)}$ for player $n^*(i)$, player $n^*(i)$ with $\mathbf{x}_{n^*(i)}$ believes that, for each $n \in \{n^{**}(i), i+1\}$, player n's frequency of action-signal pair $\mathbf{x}_n$ satisfies

$$\left\| Q_{n,n^*(i)}(i)\mathbf{x}_{n^*(i)} - \mathbf{x}_n \right\| < K_1 k \tag{7.16}$$

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

In addition, Hoeffding's inequality, there exists K_2 such that

$$\left\| \frac{1}{T^{\frac{1}{2}} - 1} \sum_{t \in T_n(n^*(i) \to_{w_i} N(i))} \mathbf{1}_{I_{n,t}(i)} - I_n(i)\mathbf{x}_n \right\| > K_2 k$$

with probability no more than $\exp(-kT^{\frac{1}{2}})$, conditional on $\mathbf{x}_n$.

Further, by Lipschitz continuity of the linear function, there exists K_3 such that (7.16) implies

$$\left\| I_n(i)Q_{n,n^*(i)}(i)\mathbf{x}_{n^*(i)} - I_n(i)\mathbf{x}_n \right\| < K_3 K_1 k.$$

In total, when player $n^*(i)$ believes player n has $w_i(n) \in \{G, B\}$ with probability no less than $2\exp(-kT^{\frac{1}{2}})$,

$$\left\| I_n(i) Q_{n,n^*(i)}(i) \mathbf{x}_{n^*(i)} - \mathbf{q}_n(w_i(n)) \right\| \leq (K_3 K_1 + K_2) k.$$

Hence, if we take k and ε sufficiently small compared to $\bar{\varepsilon}$ in Assumption 49, then player $n^*(i)$ cannot put a belief more than $4\exp(-kT^{\frac{1}{2}})$ on the event

$$\{G, B\} \ni w_i(n^{**}(i)) \neq w_i(i+1) \in \{G, B\},$$

as desired. ∎

As we will see, as long as the error-reporting noisy cheap talk by the other players transmits correctly in the coordination round 3 for x_i (this is true with an ex ante high probability at the end of the coordination round 2 for x_i), player $n^*(i)$ is pivotal for some player's inference of x_i if and only if $\{G, B\} \ni w_i(n^{**}(i)) \neq w_i(i+1) \in \{G, B\}$. 2 of Lemma 56 guarantees that, after any history (including those after player $n^*(i)$'s deviation), the probability that player $n^*(i)$ is pivotal is negligible for the almost optimality.

For each $n \in N(i) \setminus \{n^*(i)\}$, consider player $j = N(i) \setminus \{n^*(i), n\}$. As we will see, player j is not pivotal if players $n^*(i)$ and n infer the same state w_i. Therefore, 1-(b) of Lemma 56 guarantees that player j cannot manipulate player n's inference to create a situation where player j is pivotal.

For each $n \in -N(i)$, as long as the error-reporting noisy cheap talk transmits correctly in the coordination round 3 for x_i (this is true with an ex ante high probability at the end of the coordination round 2 for x_i), every player infers x_i in the same way. Since x_i controls the payoff of player $i+1 \in N(i)$, player $n \in -N(i)$ is indifferent for each $\{w_i(j)\}_{j \in N(i)}$.

7.2.4 Coordination Round 3 for x_i Between Players j and n

This is the round where player $j \in N(i)$ sends $w_i(j)$ to player $n \in I$. Let $w_i(j)(n) \in \{G, B, M\}$ be player n's inference of player j's message. Here, we assume that the error-reporting noisy cheap talk is available. See Section 7.4 for how to dispense with the error-reporting noisy cheap talk.

If player j is the same player as player n, then $w_i(j)(n) = w_i(j)$, that is, player j infers her own message straightforwardly.

Otherwise, player j sends messages as follows. From $w_i(j) \in \{G, M, B\}$, player j constructs a sequence of two binary messages $w_i(j)\{1\}, w_i(j)\{2\} \in \{G, B\}^2$: if $w_i(j) = G$, then $w_i(j)\{1\} = w_i(j)\{2\} = G$; If $w_i(j) = B$, then $w_i(j)\{1\} = w_i(j)\{2\} = B$; If $w_i(j) = M$, then $w_i(j)\{1\} = G$ and $w_i(j)\{2\} = B$ with probability $\frac{1}{2}$ and $w_i(j)\{1\} = B$ and $w_i(j)\{2\} = G$ with probability $\frac{1}{2}$.

Player j sends the two messages $w_i(j)\{1\}$ and $w_i(j)\{2\}$ sequentially via error-reporting noisy cheap talk.

With abuse of notation, we define $g[n-1](w_i(j)) \in \{w_i(j), E\}$ and $f[n](w_i(j)) \in \{G, M, B\}$ as follows: for $g[n-1](w_i(j))$,

1. $g[n-1](w_i(j)) = w_i(j)$ if and only if player $n-1$ thinks that there is no error for $f[n](w_i(j)\{1\})$ and $f[n](w_i(j)\{2\})$, that is, $g[n-1](w_i(j)\{1\}) = w_i(j)\{1\}$ and $g[n-1](w_i(j)\{2\}) = w_i(j)\{2\}$; and

2. $g[n-1](w_i(j)) = E$ otherwise.

For $f[n](w_i(j))$, player i infers $f[n](w_i(j))$ from $f[n](w_i(j)\{1\})$ and $f[n](w_i(j)\{2\})$, using the mapping between $w_i(j)$ and $w_i(j)\{1\}, w_i(j)\{2\}$.

1. $f[n](w_i(j)) = G$ if and only if $f[n](w_i(j)\{1\}) = f[n](w_i(j)\{2\}) = G$;

2. $f[n](w_i(j)) = B$ if and only if $f[n](w_i(j)\{1\}) = f[n](w_i(j)\{2\}) = B$; and

3. $f[n](w_i(j)) = M$ if and only if "$f[n](w_i(j)\{1\}) = G$ and $f[n](w_i(j)\{2\}) = B$" or "$f[n](w_i(j)\{1\}) = B$ and $f[n](w_i(j)\{2\}) = G$."

Finally, player n infers $w_i(j)$ as $w_i(j)(n) = f[n](w_i(j))$.

7.2.5 Player n's Inference of x_i

Based on these rounds, player n infers x_i as follows. Let $x_i(n) \in \{G, B\}$ be player n's inference of x_i. From $\{w_i(j)(n)\}_{j \in N(i)}$, player n constructs $x_i(n)$ such that

$$x_i(n) = \begin{cases} G & \text{if } \begin{cases} w_i(n^{**}(i))(n) = w_i(i+1)(n) = G, \\ w_i(n^{**}(i))(n) = M, w_i(i+1)(n) = G, \\ w_i(n^{**}(i))(n) = G, w_i(i+1)(n) = M, \\ w_i(n^{**}(i))(n) = B, w_i(i+1)(n) = G, w_i(n^{*}(i))(n) = G, \\ w_i(n^{**}(i))(n) = G, w_i(i+1)(n) = B, w_i(n^{*}(i))(n) = G, \end{cases} \\ B & \text{otherwise.} \end{cases} \tag{7.17}$$

Finally, let

$$x(n) = \{x_i(n)\}_{i \in I}$$

be the profile of the inferences.

7.2.6 Definition of $\theta_{i-1}(\mathrm{c}) \in \{G, B\}$

Based on the realization of the coordination block, if some events happen, then player $i-1$ makes player i indifferent between any action profile sequence in the main blocks. $\theta_{i-1}(\mathrm{c}) = B$ implies that such an event happens while $\theta_{i-1}(\mathrm{c}) = G$ implies that such an event does not happen.

We will define the events to induce $\theta_{i-1}(\mathrm{c}) = B$: for each $j \in I$, while the players coordinate on x_j,

1. there exists player $j' \in -i$ with $j' \in N(j)$ such that when player j' sends the message $w_j(j')$ to player i in the coordination round 3 for x_j between j' and i, player $i-1$ has $g[i-1](w_j(j')) = E$;

2. there exist players $j' \in -i$ and $n \in -i$ such that when player j' sends the message $w_j(j')$ to player n in the coordination round 3 for x_j between j' and n, player n has a wrong signal $f[n](w_j(j')) \neq w_j(j')$; and

3. player i is in $N(j)$ and consider the following inference:

$$x_j(n) = \begin{cases} G & \text{if } \begin{cases} w_j(n^{**}(j)) = w_j(j+1) = G, \\ w_j(n^{**}(j)) = M, w_j(j+1) = G, \\ w_j(n^{**}(j)) = G, w_j(j+1) = M, \\ w_j(n^{**}(j)) = B, w_j(j+1) = G, w_j(n^{*}(j)) = G, \\ w_j(n^{**}(j)) = G, w_j(j+1) = B, w_j(n^{*}(j)) = G, \end{cases} \\ B & \text{otherwise.} \end{cases} \tag{7.18}$$

Note that this is what we replace player n's inference of the messages in the coordination round 3 in (7.17) with the true messages. We have $\theta_{i-1}(\mathrm{c}) = B$ if there exist $n \in I$ and $j \in I$ such that player i's message $w_j(i)$ matters for $x_j(n)$ in (7.18). That is,

(a) if player i is $n^*(j)$, then

$$\{G, B\} \ni w_j(n^{**}(j)) \neq w_j(j+1) \in \{G, B\}; \tag{7.19}$$

and

(b) if player i is in $N(j) \setminus \{n^*(j)\}$, then

$$w_j \equiv w_j(n^*(j)) \neq w_j(i'). \tag{7.20}$$

with $i' = N(j) \setminus \{i, n^*(j)\}$ (player i' makes a mistake in the coordination round 3 for x_j between $n^*(j)$ and i').

Note that, although player n can be player i herself, whether or not $w_j(i)$ matters in (7.18) is determined by the other players' messages $\{w_j(i')\}_{i' \neq i}$.

In the definition of $\theta_{i-1}(\text{c})$, player $i-1$ uses the information owned by players $-(i-1, i)$. Section 7.8 explains how players $-(i-1, i)$ inform player $i-1$ of their history necessary to create $\theta_{i-1}(\text{c})$ in the re-report block. Since $\theta_{i-1}(\text{c})$ only affects the reward function (that is, does not affect action plan $\sigma_{i-1}(x_{i-1})$), it suffices that player $i-1$ knows the information by the end of the review phase.

We verify that the distribution of $\theta_{i-1}(\text{c})$ is almost independent of player i's strategy: for Cases 1 and 2, we need to verify that player i cannot manipulate $\theta_{i-1}(\text{c})$ by affecting some player's message m while coordinating on x_j in the coordination round 2 for x_j. The definition of the error-reporting noisy cheap talk implies that the probability of $g[i-1](m) = E$ when player i is a receiver and that of $f[n](m) \neq m$ when player $j \in -i$ is a sender and player $n \in -i$ is a receiver are almost independent of m.[4]

For Case 3-(a), 2 of Lemma 56 implies that player i puts a belief no more than $\exp(-\Theta(T^{\frac{1}{2}}))$ on (7.19) after *any history* (including those after player i's deviation) at the end of the coordination round 2 for x_j. Since $w_j(n^{**}(j))$ and $w_j(j+1)$ are fixed at the end of the coordination round 2 for x_j, whether (7.19) happens or not is almost independent of player i's strategy.

For Case 3-(b), note that if $w_j(n^{**}(j)) = w_j$, then (7.20) is not the case. In the coordination round 2 for x_j, the distribution of $w_j(n^{**}(j))$ is independent of player i's strategy because of (7.4). In addition, regardless of w_j, this event happens with probability no more than $\exp(-\Theta(T^{\frac{1}{2}}))$ from the perspective at the end of the coordination round 1 for x_j by

[4]Note that m can be affected by player i's strategy before the round where player j sends m to player n.

1 of Lemma 56.[5] Therefore, no player can change the distribution of $\theta_{i-1}(\text{c})$ by more than $\exp(-\Theta(T^{\frac{1}{2}}))$.

In summary, we have shown the following lemma:

Lemma 57 If

1. the probability of $g[i-1](m) = E$ when player i is a receiver of a message m is almost independent of m; and

2. the probability of $f[n](m) \neq m$ when player $j \in -i$ is a sender of a message m and player $n \in -i$ is a receiver is almost independent of m,

then, the distribution of $\theta_{i-1}(\text{c}) \in \{G, B\}$ is almost independent of player i's strategy.

The premise of lemma is stated to clarify what assumption about the error-reporting noisy cheap talk is used, expecting that we will dispense with it later.

7.2.7 Incentives in the Coordination Block

First, Lemma 57 implies that player i does not have an incentive to manipulate $\theta_{i-1}(\text{c})$.

Second, we consider player i's incentive to tell the truth about $w_n(i)$ with $i \in N(n)$ for the coordination round 3 for x_n between i and $i' \in -i$. If player i' with $i' \in -i$ received a wrong signal $f[i'](w_n(j))$ for some $n \in I$ and $j \in -i$, then Case 2 of $\theta_{i-1}(\text{c})$ implies $\theta_{i-1}(\text{c}) = B$. Hence, together with Case 3 of $\theta_{i-1}(\text{c})$, whenever player i's message matters for $x_n(i')$ for some $i' \in -i$, then $\theta_{i-1}(\text{c}) = B$ and player i is indifferent between any action profile sequence. Therefore, it is optimal for player i to tell the truth.

Third, we consider the incentive of player i in the coordination rounds 1 and 2 for x_n. If player i is player $n^*(n)$, then since x_n controls the value of player $n+1 \neq n^*(n)$, player $n^*(n)$ is indifferent between coordinating on $x_n(j) = G$ for all $j \in I$ or $x_n(j) = B$ for all

[5] Note that w_j can be affected by some player's strategy in the coordination round 1 for x_j.

$j \in I$. (7.17) and 2 of of Lemma 56 imply that, if the messages in the coordination round 3 transmit correctly if a sender is not player $n^*(n)$ (this is true with probability no less than $1 - \exp(-\Theta(T^{\frac{1}{2}}))$), then player $n^*(n)$ at the end of the coordination round 2 puts a conditional belief no less than $1 - \exp(-\Theta(T^{\frac{1}{2}}))$ on the event that $x_n(j) = G$ for all $j \in I$ or $x_n(j) = B$ for all $j \in I$ regardless of player $n^*(n)$'s history. Therefore, player $n^*(n)$ (player i) is almost indifferent between any strategy in the coordination rounds 1 and 2 for x_n.

If player i is player n (the initial holder of state x_n) but not player $n^*(n)$, then again, since x_n controls the value of player $n + 1 \neq n$, player n is indifferent between coordinating on $x_n(j) = G$ for all $j \in I$ or $x_n(j) = B$ for all $j \in I$. 1 of of Lemma 56 implies that regardless of player n's strategy in the coordination rounds 1 and 2 for x_n, players $n^*(n)$ and $n+1$ have $w_n(n^*(n)) = w_n(n+1) \in \{G, B\}$ with probability no less than $1 - \exp(-\Theta(T^{\frac{1}{2}}))$. Then, (7.17) implies that, if the messages in the coordination round 3 transmit correctly if a sender is not player n (again, this is true with probability no less than $1 - \exp(-\Theta(T^{\frac{1}{2}}))$), then $x_n(j) = G$ for all $j \in I$. Therefore, player n (player i) is almost indifferent between any strategy in the coordination rounds 1 and 2.

If player i is not player n or player $n^*(n)$, then 1 of of Lemma 56 implies that, regardless of player i's strategy in the coordination rounds 1 and 2 for x_n, players $n^*(n)$ and at least one player $i' \in N(n) \setminus \{i\}$ have $w_n(n^*(n)) = w_n(i') = x_n$ with probability no less than $1 - \exp(-\Theta(T^{\frac{1}{2}}))$. Then, (7.17) implies that, if the messages in the coordination round 3 transmit correctly if a sender is not player i (this is true with probability no less than $1 - \exp(-\Theta(T^{\frac{1}{2}}))$), then either $x_n(j) = G$ for all $j \in I$ or $x_n(j) = B$ for all $j \in I$. Therefore, player i is almost indifferent between any strategy in the coordination rounds 1 and 2.

Finally, we show that the definition of $\theta_{i-1}(\mathrm{c}) = B$ implies that, for any i, for any t in the main blocks, for any h_i^t, player i puts a belief no less than $1 - \exp(-\Theta(T^{\frac{1}{2}}))$ on the event that $x(j) = x(i)$ for all $j \in -i$ or $\theta_{i-1}(\mathrm{c}) = B$ by the following reasons: (i) if player i's signal $f[i](w_n(j))$ was wrong for some $n \in I$ and $j \in -i$, then, given $w_n(j)$, $g[i-1](w_n(j)) = E$ with probability no less than $1 - \exp(-\Theta(T^{\frac{1}{2}}))$. Since $g[i-1](w_n(j))$ is not revealed by players

$(-i)$'s continuation play in the main blocks, player i believes that $\theta_{i-1}(\mathrm{c}) = B$ because of Case 1. (ii) If player i' with $i' \in -i$ received a wrong signal $f[i'](w_n(j))$ for some $n \in I$ and $j \in -i$, then Case 2 of $\theta_{i-1}(\mathrm{c})$ implies $\theta_{i-1}(\mathrm{c}) = B$. From (i) and (ii), player i who considers almost optimality can condition that $f[i'](w_n(j)) = w_n(j)$ for all $i' \in I$, $n \in I$ and $j \in -i$. (iii) If player i is pivotal for player i''s inference of x_n with $i' \in -i$ and $n \in I$,[6] then $\theta_{i-1}(\mathrm{c}) = B$. Therefore, in total, $x(j) = x(i)$ for all $j \in -i$ or $\theta_{i-1}(\mathrm{c}) = B$.

The following lemma summarizes the above discussion:

Lemma 58 The following two statements are true:

1. if, for each player $i \in I$,

   (a) the probability of $g[i-1](m) = E$ when player i is a receiver of a message m is almost independent of m;

   (b) the probability of $f[n](m) \neq m$ when player $j \in -i$ is a sender of a message m and player $n \in -i$ is a receiver is almost independent of m and player i's strategy; and

   (c) for all n with $i \neq n+1$, player i's value is almost the same between $x_n(j) = G$ for all $j \in I$ and $x_n(j) = B$ for all $j \in I$ regardless of $\{x_{n'}(j)\}_{j \in I, n' \leq n-1}$ ($n' \leq n-1$ implies that the coordination rounds for $x_{n'}$ comes before those for x_n),

   then it is almost optimal for player i to follow the equilibrium strategy in the coordination block; and

2. for any $i \in I$, for any t in the main blocks, for any h_i^t, player i puts a belief no less than $1 - \exp(-\Theta(T^{\frac{1}{2}}))$ on the event that $x(j) = x(i)$ for all $j \in -i$ or $\theta_{i-1}(\mathrm{c}) = B$.

[6] (7.18) uses the true message $w_n(j)$ to define whether player i is pivotal or not. However, if there exists some player $i' \in -i$ receives a wrong signal from player $j \in -i$, then player i will be indifferent between any action profile sequence from (ii).

We mention 1-(a), 1-(b) and 1-(c) to clarify what properties of the error-reporting noisy cheap talk are used, expecting that we will dispense with it later. Note that, for the second statement, 1-(a), 1-(b) and 1-(c) are not necessary.

7.3 Structure of the Review Phase

Replacing the perfect cheap talk in the coordination block with the error-reporting noisy cheap talk, the structure of the coordination block is as explained in Section 7.2.1. Now, the coordination block has at most $N(1+1+3(N-1))$ rounds.[7] After the coordination block, the structure is the same as in Section 5.4 of Chapter 5. As in Chapter 5, let r be a generic serial number for a round.

If we replace the error-reporting noisy cheap talk with messages via actions, then as we will see in Section 7.4, we treat rounds where a player sends one message and rounds where a player send two messages separately. Each round where the sender sends one message has $T^{\frac{1}{2}}$ periods. Section 7.4 explains how the sender sends the message. On the other hand, each round where the sender would send two messages via error-reporting noisy cheap talk (for example, the coordination round 3 for x_i between j and n) is now divided into two rounds each of which has $T^{\frac{1}{2}}$ periods without the error-reporting noisy cheap talk. Using the first $T^{\frac{1}{2}}$-period round, the sender sends the first message as we will explain in Section 7.4. After that, using the second $T^{\frac{1}{2}}$-period round, the sender sends the second message. With abuse of notation, let r again be a generic serial number for a round and $T(r)$ be the set of periods in round r. As in the review round, player i randomly excludes period $t_i(r)$ from the periods used for the inferences. Let $T_i(r) \equiv T(r) \diagdown \{t_i(r)\}$ be the periods that player i uses for the inferences.

[7]The precise number depends on whether $n^*(i) = i$ or not for each i.

7.4 Dispensing with the Error-Reporting Noisy Cheap Talk

We consider how player j sends a binary message $m \in \{G, B\}$ to player n by taking actions rather than error-reporting noisy cheap talk in some round.

As mentioned in Section 7.1.2, with η being a small number to be defined, the sender (player j) determines

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

and player j takes

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

for $T^{\frac{1}{2}}$ periods. Player n (receiver) takes

$$\bar{\alpha}_n = (1 - 2\,(|A_n| - 1)\,\rho)\, a_n^G + \sum_{a_n \neq a_n^G} 2\rho a_n$$

Player $i \in -(j,n)$ takes $\alpha_i^{\text{receive}}$.

7.4.1 Formal: $g[n-1](m) \in \{m, E\}$

As in Section 7.1.2, player $n-1$ has $g[n-1](m) = m$ if and only if the following three conditions are satisfied:

1. as in the two player case, $z_j(m) = m$;

2. player j's action frequency is close to α_j^m:

$$\left\| \frac{1}{T^{\frac{1}{2}}-1} \sum_{t \in T_j(r)} \mathbf{1}_{a_{j,t}} - \alpha_j^m \right\| < \frac{\varepsilon}{3};$$

and

3. player j's signal frequency while player j takes a_j^m is close to $\boldsymbol{Q}_j(j \to_m n)$.

We are left to define Condition 3. Let $T_j(r,m)$ be the set of periods where player j takes a_j^m in round r, where player j sends m to player n. As player i creates $\mathbf{1}_{Q_i(x)}$ from $Q_i(x)$ and y_i, player j creates $\mathbf{1}_{Q_j(j \to_m n)}$ from $Q_j(j \to_m n)$ and y_j.

Condition 3 is satisfied if and only if the following two conditions are satisfied:

- $Q_j(j \to_m n)\left(\frac{1}{|T_j(r,m)|}\sum_{t \in T_j(r,m)} \mathbf{1}_{y_{j,t}}\right)$ and $\frac{1}{|T_j(r,m)|}\sum_{t \in T_j(r,m)} \mathbf{1}_{Q_j(j \to_m n)}$ are close:

$$\left\| Q_j(j \to_m n)\left(\frac{1}{|T_j(r,m)|}\sum_{t \in T_j(r,m)} \mathbf{1}_{y_{j,t}}\right) - \frac{1}{|T_j(r,m)|}\sum_{t \in T_j(r,m)} \mathbf{1}_{Q_j(j \to_m n)} \right\| < \frac{\varepsilon}{3}. \quad (7.21)$$

  As we have adjusted the probability of (2.22) in Section 2.4.4.2, we adjust the probability of (7.21) so that the probability of (7.21) is independent of $\{a_{j,t}, y_{j,t}\}_{t \in T(r)}$. When we say (7.21) is satisfied, we take this adjustment into account; and

- $\frac{1}{|T_j(r,m)|}\sum_{t \in T_j(r,m)} \mathbf{1}_{Q_j(j \to_m n)}$ and $\mathbf{q}_j(j \to_m n)$ are close:

$$\left\| \frac{1}{|T_j(r,m)|}\sum_{t \in T_j(r,m)} \mathbf{1}_{Q_j(j \to_m n)} - \mathbf{q}_j(j \to_m n) \right\| < \frac{\varepsilon}{3}.$$

Then, for sufficiently small ρ and ε compared to $\bar{\varepsilon}$ in Lemma 52, $g[n-1](m) = m$ only if $\mathbf{y}_j = \frac{1}{|T_j(r,m)|}\sum_{t \in T_j(r,m)} \mathbf{1}_{y_{j,t}}$ satisfies

$$\mathbf{y}_j \in \boldsymbol{Q}_j[\bar{\varepsilon}](j \to_m n).$$

7.4.2 Formal: $f[n](m) \in \{G, B\}$

On the other hand, as explained in Section 7.1.2, $f[n](m)$ is determined as follows:

1. if the following two conditions are satisfied, then player n infers $f[n](m) = \hat{m}$:

   (a) player n's action frequency is close to $\bar{\alpha}_n$:

$$\left\| \frac{1}{T^{\frac{1}{2}} - 1} \sum_{t \in T_n(r)} \mathbf{1}_{a_{n,t}} - \bar{\alpha}_n \right\| < \frac{\varepsilon}{3}; \tag{7.22}$$

   and

   (b) there exist $i \in -(j, n)$ and $\hat{m} \in \{G, B\}$ such that player n's signal frequency while player n takes a_n^G is close to $\boldsymbol{Q}_n^i(j \to_{\hat{m}} n)$.

   As we will see below, Lemma 52 guarantees that there is at most one $\hat{m} \in \{G, B\}$ such that $i \in -(j, n)$ and $\hat{m} \in \{G, B\}$ satisfy 1-(b); and

2. otherwise, player n infers $f[n](m)$ from the likelihood as in the two-player case.

We are left to define Condition 1-(b). Let $T_n(r, G)$ be the set of periods where player n takes a_n^G. As player j creates $\mathbf{1}_{Q_j(j \to_m n)}$ from $Q_j(j \to_m n)$ and y_j, player n creates $\mathbf{1}_{Q_n^i(j \to_{\hat{m}} n)}$ from $Q_n^i(j \to_{\hat{m}} n)$ and y_n.

We say player n's signal frequency while player n takes a_n^G is close to $\boldsymbol{Q}_n^i(j \to_{\hat{m}} n)$ if and only if the following two conditions are satisfied:

- $Q_n^i(j \to_{\hat{m}} n) \left(\frac{1}{|T_n(r,G)|} \sum_{t \in T_n(r,G)} \mathbf{1}_{y_{n,t}} \right)$ and $\frac{1}{|T_n(r,G)|} \sum_{t \in T_n(r,G)} \mathbf{1}_{Q_n^i(j \to_{\hat{m}} n)}$ are close:

$$\left\| Q_n^i(j \to_{\hat{m}} n) \left(\frac{1}{|T_n(r, G)|} \sum_{t \in T_n(r,G)} \mathbf{1}_{y_{n,t}} \right) - \frac{1}{|T_n(r, G)|} \sum_{t \in T_n(r,G)} \mathbf{1}_{Q_n^i(j \to_{\hat{m}} n)} \right\| < \frac{\varepsilon}{3}. \tag{7.23}$$

  Again, we adjust the probability of (7.23) so that this probability is independent of $\{a_{n,t}, y_{n,t}\}_{t \in T(r)}$; and

- $\frac{1}{|T_n(r,G)|}\sum_{t\in T_n(r,G)}\mathbf{1}_{Q_n^i(j\to_{\hat{m}} n)}$ and $\mathbf{q}_n^i(j\to_{\hat{m}} n)$ are close:

$$\left\|\frac{1}{|T_n(r,G)|}\sum_{t\in T_n(r,G)}\mathbf{1}_{Q_n^i(j\to_{\hat{m}} n)}-\mathbf{q}_n^i(j\to_{\hat{m}} n)\right\|<\frac{\varepsilon}{3}. \tag{7.24}$$

Then, for sufficiently small ρ and ε compared to $\bar{\varepsilon}$ in Lemma 52, Case 1 is the case only if $\mathbf{y}_n=\frac{1}{|T_n(r,G)|}\sum_{t\in T_n(r,G)}\mathbf{1}_{y_{n,t}}$ satisfies

$$\mathbf{y}_n\in\boldsymbol{Q}_n^i[\bar{\varepsilon}](j\to_m n)$$

for some $i\in-(j,n)$.

7.4.3 Definition of $\theta_{i-1}(j\to_m n)\in\{G,B\}$

While player $j\in I$ sends a message m to player $n\in -j$, for each $i\in I$, player $i-1$ creates $\theta_{i-1}(j\to_m n)\in\{G,B\}$. As for $\theta_{i-1}(\mathrm{c})$, $\theta_{i-1}(j\to_m n)=B$ implies that player $i-1$ makes player i indifferent between any action profile sequence in the subsequent rounds. We define $\theta_{i-1}(j\to_m n)\in\{G,B\}$ as follows:

1. for $i=j$ (sender), $\theta_{i-1}(j\to_m n)=G$ always;

2. for $i=n$ (receiver), $\theta_{i-1}(j\to_m n)=B$ if and only if $g[n-1](m)=E$; and

3. for $i\in -n$ (not a receiver), $\theta_{i-1}(j\to_m n)=B$ if and only if (7.22) or (7.23) is not satisfied for $\hat{m}=m$.

7.4.4 Summary of the Properties of $g[n-1](m)$, $f[n](m)$ and $\theta_{i-1}(j\to_m n)$

In summary, we can show the following lemma:

Lemma 59 For sufficiently large T, for any $j \in I$ and $n \in -j$, the above communication protocol satisfies the following:

1. $g[n-1](m) = E$ with probability $1 - \eta - \exp(-\Theta(T^{\frac{1}{2}}))$ for any $m \in \{G, B\}$;

2. given any $m \in \{G, B\}$ and any on-path history, player n puts a belief no less than $1 - \exp(-\Theta(T^{\frac{1}{2}}))$ on the event that $f[n](m) = m$ or $g[n-1](m) = E$;

3. for each $i \in -(j, n)$, if $f[n](m) \neq m$, then $\theta_{i-1}(j \to_m n) = B$;

4. given $m \in \{G, B\}$ and a history of players $-n$, any $f[n](m)$ happens with probability at least $\exp(-\Theta(T^{\frac{1}{2}}))$;

5. the distribution of $g[n-1](m)$ is independent of player n's strategy; and

6. the distribution of $\theta_{i-1}(j \to_m n)$ is independent of player i's strategy.

Proof: 1. This follows from the law of large numbers.

2. If Case 2 for $f[n](m)$ is the case, then as in the two-player case, we are done.

   Consider Case 1 is the case. Conditional on m, suppose player n infers $m' = \{G, B\} \setminus \{m\}$. Then, for sufficiently small ρ and ε,

$$\frac{1}{|T_n(r, G)|} \sum_{t \in T_n(r,G)} \mathbf{1}_{y_{n,t}} \in \boldsymbol{Q}_n^i[\bar{\varepsilon}](j \to_{m'} n). \tag{7.25}$$

   On the other hand, by Hoeffding's inequality, there exists K_1 such that, if player n believes $g[n-1](m) = m$ with probability $1 - \exp(-kT^{\frac{1}{2}})$, then the conditional expectation of $\frac{1}{|T_j(r,m)|} \sum_{t \in T_j(r,m)} \mathbf{1}_{y_{j,t}}$ given $\hat{m}$ should be in

$$\left\{ \begin{array}{c} \mathbf{y}_j \in \mathbb{R}_+^{|Y_n|} : \exists \boldsymbol{\varepsilon} \in \mathbb{R}^{|Y_n|} \text{ such that} \\ \left\{ \begin{array}{c} \|\boldsymbol{\varepsilon}\| \leq K_1 k, \\ Q_j(j \to_{\hat{m}} n)\,(\mathbf{y}_j + \boldsymbol{\varepsilon}) = \mathbf{q}_j(j \to_{\hat{m}} n). \end{array} \right. \end{array} \right\}.$$

Since $T_j(r,m)$ and $T_n(r,G)$ are different at most for $(2\rho(|A_n|-1)+\rho+2\varepsilon)T+2$ periods if $g[n-1](m)=m$ and Case 1 for $f[n](m)$ is the case, the above condition implies

$$\mathbf{y}_n = \frac{1}{|T_n(r,G)|}\sum_{t\in T_n(r,G)} \mathbf{1}_{y_{n,t}}$$

is included in

$$\left\{ \begin{array}{c} \mathbf{y}_j \in \mathbb{R}_+^{|Y_n|} : \exists \boldsymbol{\varepsilon} \in \mathbb{R}^{|Y_n|} \text{ such that} \\ \left\{ \begin{array}{c} \|\boldsymbol{\varepsilon}\| \leq K_1 k + (2\rho(|A_n|-1)+\rho+2\varepsilon) + \frac{2}{T}, \\ Q_j(j \to_{\hat{m}} n) M_{j,n}(\hat{m})(\mathbf{y}_n + \boldsymbol{\varepsilon}) = \mathbf{q}_j(j \to_{\hat{m}} n). \end{array} \right. \end{array} \right\}.$$

Take k, ρ and ε sufficiently small so that for $\bar{\varepsilon}$ in Lemma 52, we have

$$\bar{\varepsilon} < K_1 k + (2\rho(|A_n|-1)+\rho+2\varepsilon) + \frac{2}{T}$$

for large T. Then, the above condition implies

$$\frac{1}{|T_n(r,G)|}\sum_{t\in T_n(r,G)} \mathbf{1}_{y_{n,t}} \in \boldsymbol{H}_n[\bar{\varepsilon}](j \to_{m'} n). \tag{7.26}$$

By Lemma 52, (7.26) contradicts to (7.25).

Therefore, for sufficiently small k, ρ and ε, conditional on m, whenever player n infers $m' = \{G,B\} \setminus \{m\}$, player n believes $g[n-1](m') = E$ with probability no less than $1-\exp(-kT^{\frac{1}{2}})$, as desired.

3. By definition of $f[n](m)$, if $f[n](m) \neq m$, then either (7.22) or (7.23) is satisfied, which implies $\theta_{i-1}(j \to_m n) = B$.

4. Given $(a_t, y_{-n,t})_{t\in T(r)}$, any $(y_{n,t})_{t\in T(r)}$ can occur with probability at least

$$\left\{\min_{y,a} q(y_n \mid a, y_{-n})\right\}^{T^{\frac{1}{2}}}.$$

Assumption 2 (full support) implies that this probability is $\exp(-\Theta(T^{\frac{1}{2}}))$. The rest of the proof is the same as in Lemma 45.

5. We adjusted the probability of (7.21) so that this probability is independent of $\{a_{j,t}, y_{j,t}\}_{t\in T(l)}$. Hence, the distribution of $g[j](m)$ is determined solely by player j's mixture.

6. We adjusted the probability of (7.23) so that this probability is independent of $\{a_{n,t}, y_{n,t}\}_{t\in T(r)}$. Hence, the distribution of $\theta_{i-1}(j \to_m n)$ is determined solely by player n's mixture. ∎

7.5 Equilibrium Strategies

In this section, we define $\sigma_i(x_i)$ and π_i^{main}.

7.5.1 States

The states $\lambda_i(l+1)$, $\lambda_{i-1}(l+1)\,(i)$, $\lambda_{i-1}(l+1)\,(i+1)$, $d_i(l+1)$, $d_j(l+1)(i)$, $c_i(l+1)$ and $\theta_i(l)$ are defined as in Chapter 5 except that x is replaced with $x(i)$ defined in Section 7.2.5 and

- if $\theta_i(\text{c}) = B$ happens in the coordination block, then $\theta_i(l) = B$; and
- if $\theta_i(j \to_m n) = B$ happens in a round between the lth review round and $(l+1)$th review round where player $j \in I$ sends a message m to player $n \in -j$, then $\theta_i(l+1) = B$

If we replace the error-reporting noisy cheap talk with messages via actions, then we use $f[i](m)$ (when player i is a receiver) and $g[i](m)$ (when player $i+1$ is a receiver) defined in Section 7.4.

7.5.2 Player i's Action Plan $\sigma_i(x_i)$

7.5.2.1 With the Error-Reporting Noisy Cheap Talk

In the coordination block, the players play the game as explained in Section 7.2. For the other blocks, $\sigma_i(x_i)$ prescribes the same action with x replaced with $x(i)$ except for the report and re-report blocks. See Sections 7.7 and 7.8 for the strategy in the report and re-report blocks.

7.5.2.2 Without the Error-Reporting Noisy Cheap Talk

When player $j \in I$ sends a message m to player $n \in -j$, then the strategies are determined in Section 7.4.

7.5.3 Reward Function

In this subsection, we explain player $i-1$'s reward function on player i, $\pi_i^{\text{main}}(x_{i-1}, h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}} : \delta)$. In general, the total reward $\pi_i^{\text{main}}(x_{i-1}, h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}} : \delta)$ is the summation of rewards for each round r:

$$\begin{aligned}\pi_i^{\text{main}}(x_{i-1}, h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}} : \delta) &= \sum_{l=1}^{L}\sum_{t\in T(l)} \pi_i^{\delta}(t, \alpha_{-i,t}, y_{i-1,t}) \\ &\quad + \sum_r \pi_i^{\text{main}}(x_{i-1}, h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}}, r : \delta).\end{aligned}$$

Note that we add (5.10) to ignore discounting only for the review rounds. As we will see, for the round where the players communicate, we use reward function that take discounting into account directly.

We define $\pi_i^{\text{main}}(x, h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}}, r : \delta)$ for each r.

As a preparation, let us define the reward to cancel out the difference in the instantaneous utilities for player i. Note that while player j sends the message by taking actions, player j takes $a_j \in \{a_j^G, a_j^B\}$ and player $n \in -(i,j)$ takes either $\bar{\alpha}_n$ or $\alpha_n^{\text{receive}}$, both of which satisfy $\alpha_n(a_n) \geq \rho$ for all n and a_n. Then, Assumption 48 guarantees that player $i-1$ can statistically identify player i's action and construct the following reward function:

Lemma 60 For all $\rho > 0$, there exists $\bar{U} > 0$ such that, for all $i, j \in I$ and $\{\alpha_n\}_{n \in -(i,j)}$ with $\alpha_n(a_n) \geq \rho$ for all n and a_n, there exists $\pi_i^{x_{i-1}}[\alpha_{-(i,j)}] : A_{i-1} \times Y_{i-1} \to [-\bar{U}, \bar{U}]$ such that

1. for all $a_i \in A_i$, $a_j \in \{a_j^G, \alpha_j^B\}$ and $\alpha_{-(i,j)}$ satisfying the above condition,

$$u_i\left(a_i, a_j, \alpha_{-(i,j)}\right) + \mathbb{E}\left[\pi_i^{x_{i-1}}[\alpha_{-(i,j)}](a_{i-1}, y_{i-1}) \mid a_i, a_j, \alpha_{-(i,j)}\right] \tag{7.27}$$

   is independent of $a_i, a_j, \alpha_{-(i,j)}$; and

2. for all (a_{i-1}, y_{i-1}), we have

$$\pi_i^{x_{i-1}}[\alpha_{-(i,j)}](a_{i-1}, y_{i-1}) \begin{cases} \leq 0 & \text{if } x_{i-1} = G, \\ \geq 0 & \text{if } x_{i-1} = B. \end{cases}$$

Proof: For $\{\alpha_n\}_{n \in -(i,j)}$ with $\alpha_n(a_n) \geq \rho$, Assumption 48 guarantees the identifiability and the same proof as for Lemma 12 works. ∎

Note that $\bar{U}$ depends on ρ.

7.5.3.1 With the Error-Reporting Noisy Cheap Talk

In the coordination block, for round r where player j sends message m to player $n^*(i)$, player $i-1$ gives

$$\pi_i^{\text{main}}(x_{i-1}, h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}}, r : \delta) = \sum_{t \in T(r)} \delta^{t-1} \pi_i^{x_{i-1}}[\alpha_{-(i,j)}(r)](a_{i-1}, y_{i-1})$$

to make player i indifferent between any action profile sequence. Note that we take discounting into account.

In the main blocks, the reward function is the same as in Chapter 5 except that x replaced with $x(i-1)$.

7.5.3.2 Without the Error-Reporting Noisy Cheap Talk

For round r where player j sends a message, player $i-1$ gives

$$\pi_i^{\text{main}}(x_{i-1}, h_{i-1}^{\text{main}}, h_{i-1}^{\text{rereport}}, r : \delta) = \sum_{t \in T(r)} \delta^{t-1} \pi_i^{x_{i-1}} [\alpha_{-(i,j)}(r)](a_{i-1,t}, y_{i-1,t})$$

defined in Lemma 60. Again, we take discounting into account.

For round r corresponding to a review round, the reward function is the same as in the case with the error-reporting noisy cheap talk.

7.6 Almost Optimality of the Strategy

We want to verity (1.21), (1.16) and (1.17) are satisfied. First, by definition in Section 7.5.3, (1.17) is satisfied.

Second, since the length of the rounds other than the review rounds is $T^{\frac{1}{2}}$, the payoff from the review rounds approximately determines the payoff from the review phase for sufficiently large δ (and so sufficiently large T). Therefore, we neglect the payoffs from the rounds other than the review rounds.[8]

Third, we consider (1.21) and (1.16) in the case with the error-reporting noisy cheap talk. Suppose that $x(j) = x(i)$ for all $i, j \in I$ at the end of the coordination block. Then, (1.21) and (1.16) are shown as in the case with the perfect cheap talk.

This implies that the premises of Lemma 58 are satisfied. Therefore, (i) the incentive in the coordination block is satisfied and (ii) we can concentrate on the case with $x(j) = x(i)$ for all $i, j \in I$.

[8]Notice that $\bar{U}$ in Lemma 60 depends on ρ. We first fix ρ (and so fix $\bar{U}$) and then take T going to infinity.

(i) and (ii) imply (1.21). In addition, by the law of large numbers, $x(j) = x$ for all $j \in -i$ in the coordination block with probability no less than $1 - \exp(-\Theta(T^{\frac{1}{2}}))$. Therefore, (1.16) is satisfied at the beginning of the review phase.

Finally, we consider (1.21) and (1.16) in the case without the error-reporting noisy cheap talk. Again, suppose that $x(j) = x(i)$ for all $i, j \in I$ at the end of the coordination block. Then, (1.21) and (1.17) are verified as in the case with the error-reporting noisy cheap talk except for the following two points:

- player $i-1$ makes player i indifferent between any action profile sequence because of $g[i-1](m) = E$ or $\theta_{i-1}(j \to_m n) = B$ with a higher probability. However, the probability of $g[i-1](m) = E$ or $\theta_{i-1}(j \to_m n) = B$ is bounded by $\Theta(\eta)$. Hence, re-taking η sufficiently small as we do in (6.12), we can deal with this problem as in the two-player case; and

- when player $j \in -i$ sends a message m to player $n \in -(i,j)$, player i can manipulate the distribution of $f[n](m)$. However, Lemma 59 implies that player i cannot manipulate $\theta_{i-1}(j \to_m n)$. $f[n](m)$ matters for player i's continuation payoff if and only if $\theta_{i-1}(j \to_m n) = G$. Hence, the relevant events for player i are

  - $f[n](m) = m$ and $\theta_{i-1}(j \to_m n) = G$; or

  - $f[n](m) \neq m$ or $\theta_{i-1}(j \to_m n) = B$.

  Since $f[n](m) \neq m$ implies $\theta_{i-1}(j \to_m n) = B$, the relevant histories for player i are

  - $\theta_{i-1}(j \to_m n) = G$; or

  - $\theta_{i-1}(j \to_m n) = B$.

  Since player i cannot manipulate $\theta_{i-1}(j \to_m n)$ by Lemma 59, player i does not have an incentive to manipulate $f[n](m)$.

To verify the incentives in the coordination block, we consider the premises of Lemmas 57 and 58 in the case without the error-reporting noisy cheap talk.

The premise 1 of Lemmas 57 and premise 1-(a) of Lemma 58 are satisfied by Lemma 59.

As we have mentioned above, when player $j \in -i$ sends a message m to player $n \in -(i, j)$, player i does not have an incentive to manipulate $f[n](m)$. Therefore, the premise 2 of Lemmas 57 and premise 1-(b) of Lemma 58 are satisfied.

We are left to verify the premise 1-(c) of Lemma 58: player i's value is almost the same between $x_n(j) = G$ for all $j \in I$ and $x_n(j) = B$ for all $j \in I$ regardless of $\{x_{n'}(j)\}_{j \in I, n' \neq n}$. To formally show this, we proceed backward from player N's state. There are following two cases:

- suppose that $x_{n'}(j) \neq x_{n'}(j')$ happens for some $n' \in \{1, ..., N-1\}$, $j \in I$ and $j' \in -j$. Then, by definition of $\{\theta_{i-1}(j'' \to_m n'')\}_{j'', n''}$ and 2 of Lemma 58,[9] player i puts a belief no less than $1 - \exp(-\Theta(T^{\frac{1}{2}}))$ on the event that $\theta_{i-1}(\mathrm{c}) = B$ in the coordination rounds for $x_{n'}$ or that there exist $j'' \in I$ and $n'' \in -j''$ such that $\theta_{i-1}(j'' \to_m n'') = B$ happens when player $j'' \in -i$ sends a message m to player n'' in the coordination round 3 for $x_{n'}$. Therefore, if $x_{n'}(j) \neq x_{n'}(j')$ happens for some $n' \in \{1, ..., N-1\}$, $j \in I$ and $j' \in -j$, then player i is almost indifferent between any action profile sequence, which implies player i's value is almost constant; and

- suppose that $x_{n'}(j) = x_{n'}(j')$ for all $n \in \{1, ..., N-1\}$ and $j, j' \in I$. Then, if either $x_N(j) = G$ for all $j \in I$ or $x_N(j) = B$ for all $j \in I$ is the case, then we have verified that (1.16) holds with x replaced with $x(j)$. Since $i \neq N+1$, player i's value is almost the same between $x_N(j) = G$ for all $j \in I$ and $x_N(j) = B$ for all $j \in I$.

Therefore, 1-(c) of Lemma 58 holds for $n = N$. This implies that each player follows the equilibrium path in the coordination rounds for x_N. Hence, at the end of the coordination rounds for x_{N-1}, each player i expects that $x_N(j) = x_N$ for all $j \in I$ or $\theta_{i-1}(j \to_m n) = B$

[9] 2 of Lemma 58 does not use the premises 1-(a), 1-(b) and 1-(c).

in the coordination round 3 for x_N between some $j \in I$ and $n \in -j$ with probability no less than $1 - \exp(-\Theta(T^{\frac{1}{2}}))$. Hence, the same argument as for $n = N$ holds for $n = N - 1$. By induction, we are done.

Therefore, all the premises in Lemmas 57 and 58 are satisfied. This implies that

1. it is almost optimal for player i to follow the equilibrium strategy in the coordination block; and

2. for any i, for any t in the main blocks, for any h_i^t, player i puts a belief no less than $1 - \exp(-\Theta(T^{\frac{1}{2}}))$ on the event that $x(j) = x(i)$ for all $j \in -i$ or "$\theta_{i-1}(j \to_m n) = B$ or $\theta_{i-1}(\mathrm{c}) = B$ happens in the coordination block."

Note that 1 implies the almost optimality of $\sigma_i(x_i)$ in the coordination block and that 2 implies the almost optimality of $\sigma_i(x_i)$ in the main blocks. Hence, (1.21) is verified.

Since we have verified (1.16) for $x(j) = x(i)$ for all $i, j \in I$, we are left to show (1.16) at the beginning of the review phase. Compared to the case with the error-reporting noisy cheap talk, we need to deal with the fact that $g[n-1](m) = E$ and $\theta_{i-1}(j \to_m n) = B$ can happen when player j sends a message m to player n in the coordination block with a higher probability. However, since the ex ante probability of $g[n-1](m) = E$ or $\theta_{i-1}(j \to_m n) = B$ for some $j \in I$, $n \in -j$ and m is bounded by $\Theta(\eta)$, re-taking η smaller if necessary, we are done.

7.7 Report Block

7.7.1 Preparation

Before constructing an equilibrium, let us make three preparations. First, consider the situation where player $i \geq 2$ sends a binary message to player $i-1$ by taking actions. Suppose player $-i$ take $\alpha_{-i}^{\text{report}}$ defined in Lemma 55 and calculate $\phi_{i-1}(a_{i-1,t}, y_{i-1,t})$, defined in Lemma 55. Player $i-1$ constructs $\Phi_{i-1,t}$ as player 1 constructs $\Phi_{1,t}$ from $\phi_1(a_{1,t}, y_{1,t})$ in the two-player case. Lemma 55 implies that

$$\Pr\left(\{\Phi_{i-1,t}=1\} \mid \alpha_{-i,t}^{\text{report}}, a_{i,t}, y_{i,t}\right) = \begin{cases} q_2 & \text{if } a_{i,t} = a_i^G, \\ q_1 & \text{if } a_{i,t} \neq a_i^G \end{cases} \tag{7.28}$$

for all t and $y_{i,t}$. The role of Φ_{i-1} is the same as that of Φ_1 in the two-player case.

Second, consider the situation where player $j \in -(i-1,i)$ sends a binary message to player $i-1$ by taking actions by taking $a_j \in \{a_j^G, a_j^B\}$. Players $-(j,i)$ take $\alpha_{-(j,i)}^{\text{report}}$. Intuitively, player $i-1$ uses this information to monitor i and it is important to make sure that player i cannot manipulate player $(i-1)$'s inference. For this purpose, for some $q_2 > q_1$, for each $i-1$ and $j \in -(i-1,i)$, we want to construct $\psi_{j\to i-1}^i : A_{i-1} \times Y_{i-1} \to (0,1)$ such that, for each $a_i \in A_i$, we have

$$\mathbb{E}\left[\psi_{j\to i-1}^i(a_{i-1}, y_{i-1}) \mid a_i, a_j, \alpha_{-(j,i)}^{\text{report}}\right] = \begin{cases} q_2 & \text{if } a_j = a_j^G, \\ q_1 & \text{if } a_j = a_j^B. \end{cases}$$

Since Assumption 48 implies that $\left(\Pr\left(a_{i-1}, y_{i-1} \mid a_i, a_j, \alpha_{-(j,i)}\right)\right)_{a_{i-1},y_{i-1}}$ are linearly independent with respect to $a_i \in A_i$ and $a_j \in \{a_j^G, a_j^B\}$, such q_2, q_1 and $\psi_{j\to i-1}^i$ always exist.

Player $i-1$ constructs $\Psi_{j\to i-1}^i$ from $\psi_{j\to i-1}^i(a_{i-1}, y_{i-1})$ as player j constructs $\Psi_{j,t}$ from $\psi_j(a_j, y_j)$. Then,

$$\Pr\left(\left\{\Psi_{j\to i-1}^i = 1\right\} \mid a_i, a_j, \alpha_{-(j,i)}^{\text{report}}\right) = \begin{cases} q_2 & \text{if } a_j = a_j^G, \\ q_1 & \text{if } a_j = a_j^B \end{cases} \tag{7.29}$$

for all $a_i \in A_i$.

Third, as we will see, in the report block or re-report block, either "one player j takes $a_j \in \{a_j^G, a_j^B\}$ and the others are supposed to take $\alpha_{-j}^{\text{report}}$" or "all the players are supposed to take $\alpha^{\text{p.r.}(i)}$ or $a^{\text{p.r.}(i)}$, depending on whether Condition 1 or 2 is the case in Assumption 53."

In both cases, Assumptions 3 and 48 guarantee that player $i-1$ can statistically identify player i's action. Therefore, we can change the reward to cancel out the differences in the instantaneous utility in the report block. Since the report block lasts for $\Theta(T^{\frac{1}{3}})$ periods, this does not affect the equilibrium payoff.

We are left to construct the report and re-report blocks to attain the exact optimality of the equilibrium strategies. In this section, we explain the report block. Contrary to the two-player case, we directly construct the report block without public randomization or any cheap talk.

7.7.2 Structure of the Report Block

The report block proceeds as follows:

1. player N sends the messages about h_N^{main};
2. player $N-1$ sends the messages about h_{N-1}^{main};

   $\vdots$

3. player 3 sends the messages about h_3^{main};
4. as in the two-player case, players 1 and 2 coordinate on which of them will send messages: when Condition 1 (Condition 2, respectively) is satisfied in Assumption 53,
   (a) each player takes $a_i^{\text{p.r.}(2)}$ ($\alpha_i^{\text{p.r.}(2)}$, respectively) and each player i observes her private history (a_i, y_i); and
   (b) if player 2 observes $y_2 \in Y_{2,1}^2$ ($(a_2, y_2) \in H_{2,1}^2$, respectively), then player 2 sends the message that $y_2 \in Y_{2,1}^2$ ($(a_2, y_2) \in H_{2,1}^2$, respectively) to player 1. Otherwise, that is, if player 2 observes $y_2 \in Y_{2,2}^2$ ($(a_2, y_2) \in H_{2,2}^2$, respectively), then player 2 sends the message that $y_2 \in Y_{2,2}^2$ ($(a_2, y_2) \in H_{2,2}^2$, respectively) to player 1;

5. if player 2 has sent the message $y_2 \in Y^2_{2,1}$ ($(a_2, y_2) \in H^2_{2,1}$, respectively), then player 2 sends the meaningful messages about h_2^{main}. If player 2 has sent the message $y_2 \in Y^2_{2,2}$ ($(a_2, y_2) \in H^2_{2,2}$, respectively), then player 2 takes a_2^G for the periods where player 2 would send the messages about h_2^{main} otherwise;

6. player 1 sends the message about h_1^{main}; and

7. the players play the round for conditional independence.

We explain each step in the sequel.

7.7.3 Player $i \geq 3$ sends h_i^{main}

Since there is a chronological order for the rounds and r is a generic serial number of rounds, the notations $\#_i^r$, $\#_i^r(k)$, $T(r,k)$ and $\{a_{i,t}, y_{i,t}\}_{t \in T(r,k)}$ defined in Chapter 5 is still valid.

Player i sends the messages about h_i^{main} in the same way as player 2 sends the messages in Chapter 6 with two players.

That is, for each round r,

1. first, player i reports $\#_i^r$;

2. second, player i reports $\{\#_i^r(k)\}_{k \in \{1,\dots,K\}}$. See Section 2.9.6 for the definition of K;

3. third, players i and $i-1$ coordinate on $k(r)$ as players 2 and 1 coordinate on $k(r)$ in Section 2.9.6; and

4. fourth, player i sends $\{a_{i,t}, y_{i,t}\}_{t \in T(r,k(r,i))}$. $k(r,i)$ is the result of the coordination on $k(r)$ in Step 3.

In Steps 1, 2 and 4, player i sends a message as player 2 does in Chapter 6 and player $i-1$ interprets the message as player 1 does in Chapter 6: player i takes $a_i \in \{a_i^G, a_i^B\}$, players $-i$ take $\alpha_{-i}^{\text{report}}$, and player $i-1$ constructs $\Phi_{i-1} \in \{0,1\}$. From (7.28), player i cannot infer Φ_{i-1} from player i's signals.

In Step 3, the coordination between player i and $i-1$ is the same as in Section 6.7.3.2 with j replaced with $i-1$ (with the other players $-(i-1,i)$ taking $a_{-(i-1,i)}^{\text{p.r.}(i)}$ or $\alpha_{-(i-1,i)}^{\text{p.r.}(i)}$ depending on whether Condition 1 or 2 is satisfied in Assumption 53). Assumption 53 implies that this is a well defined procedure.

7.7.4 Player 2 sends h_2^{main}

Player 2 sends the messages about h_2^{main} as player $i \geq 3$ if and only if player 2 observed $y_2 \in Y_{2,1}^2$ (or $(a_2, y_2) \in H_{2,1}^2$ if Condition 2 is the case in Assumption 53) in Step 4 of Section 7.7.2. If player 2 observes $y_2 \in Y_{2,2}^2$ (or $(a_2, y_2) \in H_{2,2}^2$), then player 2 takes a_2^G for periods where player 2 would send $\#_2^r$, $\{\#_2^r(k)\}_{k\in\{1,\ldots,K\}}$ and $\{a_{2,t}, y_{2,t}\}_{t\in T(r,k(r,2))}$ otherwise. In addition, the coordination on $k(r)$ between players 2 and 1 is the same as in Chapter 6 (with the other players $-(1,2)$ taking $a_{-(1,2)}^{\text{p.r.}(2)}$ or $\alpha_{-(1,2)}^{\text{p.r.}(2)}$). Assumption 53 implies that this is a well defined procedure.

As for the case with $i \geq 3$, player 2 takes $a_2 \in \{a_2^G, a_2^B\}$, players -2 take $\alpha_{-2}^{\text{report}}$ and player 1 constructs $\Phi_1 \in \{0,1\}$. From (7.28), player 2 cannot infer Φ_1 from player 2's signals.

7.7.5 Player 1 sends h_1^{main}

Player 1 sends the messages about h_1^{main} to player N as player $i \geq 3$. As in the two-player case, player 1 takes $a_1 \in \{a_1^G, a_1^B\}$ and players -1 take $\alpha_{-1}^{\text{c.i.}}$ with $\alpha_j^{\text{c.i.}}(a_j) > \rho$ for all $j \in -1$ and $a_j \in A_j$.

After that, player 1 sends the histories in the report block to player N as player 1 does to player 2 in the round for conditional independence in Section 6.7.5.1. Again, this set of periods is called "the round for conditional independence." In this round, player 1 takes some action $a_1 \in A_1$ and players -1 take a_{-1}^G. Player N infers this message from y_N. By Assumption 3, player N can statistically identify player 1's action.

From the history in the round for conditional independence, player N constructs Φ_N. Compared to the two-player case, player 2 is replaced with player N.

7.7.6 Reward Function π_i^{report}

When Condition 1 is the case for Assumption 53, while the players should take $a^{\text{p.r.}(n)}$ to coordinate on $k(r)$ or whether player 2 reports the history, player $i-1$ incentivizes player i to take $a_i^{\text{p.r.}(n)}$. As in the two-player case, Assumption 3 is sufficient for player $i-1$ to construct a strict reward on $a_i^{\text{p.r.}(n)}$.

When Condition 2 is the case for Assumption 53, then as in the two-player case, for each $i \in I$, any $a_i \in A_i$ gives the same ex ante probability for the results of the coordination.

In the report block, when player i sends the message, no player $j \in -i$ has an incentive to manipulate player $(i-1)$'s inference of player i's message since player i's message only affects player $(i-1)$'s reward on player i and we construct the structure of the report block in Section 7.7.2 and the punishment for telling a lie, $g_j(h_{j-1}^{\text{main}}, h_{j-1}^{\text{rereport}}, \hat{a}_{j,t}, \hat{y}_{j,t})$, so that player j does not have an incentive to learn player i's history from the report block. Note that $g_j(h_{j-1}^{\text{main}}, h_{j-1}^{\text{rereport}}, \hat{a}_{j,t}, \hat{y}_{j,t})$ conditions players $-(i-1,i)$'s history.

Finally, we construct π_i^{report} that makes $\sigma_i(x_i)$ exactly optimal. This step is the same as in Section 5.9 except for the following:

1. at the end of the coordination round 2 for x_j,

   (a) player $i \notin N(j)$ is exactly indifferent between $\{w_j(n)\}_{n \in N(j)}$; and

   (b) player $i \in N(j)$ is exactly indifferent between $w_j(i) \in \{G, B\}$.

   This is possible since

   (a) player $i \notin N(j)$ is almost indifferent between $\{w_j(n)\}_{n \in N(j)}$ without adjustment since (i) players coordinate on the same $x_j(n)$ with a high probability for all $\{w_j(n)\}_{n \in N(j)}$, (ii) the probability that player $i-1$ makes player i indifferent in future does not depend on $\{w_j(n)\}_{n \in N(j)}$ by more than $\exp(-\Theta(T^{\frac{1}{2}}))$, and (iii) x_j controls the payoff of player $j+1 \neq i$.

Since $\{w_j(n)\}_{n \in N(j)}$ will be revealed in the re-report block, we can adjust player i's value so that player i's value is constant for all $\{w_j(n)\}_{n \in N(j)}$; and

(b) for player $i \in N(j)$, players will coordinate on the same $x_j(n)$ regardless of $w_j(i)$ or player $i-1$ makes player i indifferent between any action profile sequence. Since the probability of the latter case is independent of player i's continuation play after the coordination round 2 for x_j given $\{w_j(n)\}_{n \in N(j), n \neq i}$, we can make sure that, given $\{w_j(n)\}_{n \in N(j), n \neq i}$ which will be revealed in the re-report block, player i's value is determined solely by $\{w_j(n)\}_{n \in N(j), n \neq i}$;

2. given 1, only player $n^*(j)$ reports her history in the coordination round 2 for x_j:

   (a) for player $i = n^*(j)$, since the other players take $\alpha_{-i}^{\text{receive}}$, Condition 3 of Assumption 49 is sufficient to incentivize player $n^*(j)$ to tell the truth;

   (b) for player $i \in N(j) \setminus \{n^*(j)\}$, the distribution of $\{w_j(n)\}_{n \in N(j), n \neq i}$ is independent of player i's action plan and from 1-(b), player i's value is determined solely by $\{w_j(n)\}_{n \in N(j), n \neq i}$. Hence, player i is indifferent between any realization of the history in this round without adjustment; and

   (c) for player $i \in -N(j)$, from 1-(a), player i is indifferent between $\{w_j(n)\}_{n \in N(j)}$. Hence, player i is indifferent between any realization of the history in this round without additional adjustment;

3. in the coordination round 1 for x_j, players i and $n^*(i)$ report the history and we make players i and $n^*(i)$ indifferent between any realization of w_i;

4. when we construct a punishment g to incentivize player i as in Lemma 38, we assume player i could know the following variables for all $n \in -i$:

   (a) $\mathbf{1}_{I_{n,t}(j)}$ when player n infers the message by player $n^*(j)$ in the coordination round 2 for x_j;

(b) $\mathbf{1}_{Q_n^{i'}(j \to_G n)}$ and $\mathbf{1}_{Q_n^{i'}(j \to_B n)}$ for each $i' \in -(j,n)$ when player j sends a message to player n;

(c) $\Gamma_{n,t}^{a(x)}$ when player n tries to monitor player $n-1$ in the review round; and

(d) $\mathbf{1}_{Q_n^j(x)}$ for each $j \in -n$ when player n calculates (5.23) and (5.24).

Since all of these have full support, the same proof as Lemma 21 works. For notational convenience in Section 7.8, let $\varphi_{n,t}$ be the vector of variables in 3-(a), 3-(b), 3-(c) and 3-(d) that are valid in period t; and

5. for a round where the players communicate, we (i) first cancel out the effect of the history in the round on the learning about the best responses from the next rounds, and (ii) second make any action sequence is indifferent ex ante. The construction of f_i is the same as Section 2.8.

We are left to deal with the probability that the message does not transmit correctly with probability 1 without perfect cheap talk. We deal with this problem in Section 7.9 after we explain the re-report block.

7.8 Re-Report Block

As in Section 5.10, we introduce the re-report block so that, for each player i, player $i-1$ can collect the information necessary to construct π_i from players $-(i-1,i)$.

The basic structure of the re-report block is the same as in Section 5.10:

1. players $-(N-1,N)$ send the information to player $N-1$ to construct π_N;
2. players $-(N-2,N-1)$ send the information to player $N-2$ to construct π_{N-1};

   $\vdots$

N-1. players $-(1,2)$ send the information to player 1 to construct π_2; and

N. players $-(N,1)$ send the information to player N to construct π_1.

When players $-(i-1,i)$ send the information to player $i-1$, each player takes turns to send the information:

1. player 1 sends the information to player $i-1$ if $1 \in -(i-1,i)$. If $1 \notin -(i-1,i)$, then skip this step;

   $\vdots$

N player N sends the information to player $i-1$ if $N \in -(i-1,i)$. If $N \notin -(i-1,i)$, then skip this step.

When player $j \in -(i-1,i)$ sends the information about her history, she sends the following information chronologically:

- for each round r, what strategy α_j player j took in round r. Note that this contains the information about what message player j sent if player j sends a message in that round. The cardinality of this message is fixed and finite since the support of $\alpha_j(r)$ is fixed and finite, independently of T;

- for each round r, what was $t_n(r)$. The cardinality of this message is no more than T;

- for each round r, for each $\left(a_j, y_j, \varphi_j\right)$, how many times player j observed $\left(a_j, y_j, \varphi_j\right)$. The cardinality of this message is no more than $\Theta\left(T\right)$;

- for each round r, $\left(a_{j,t_j(r)}, y_{j,t_j(r)}, \varphi_{j,t_j(r)}\right)$. Note that the these two pieces of information are sufficient for player $i-1$ to know what was player j's inference of a message if player j receives a message in that round and to construct $\theta_{i-1}(j \to_m n)$ and $\theta_{i-1}(l)$;

- at the end of each lth review round, what was the realization of player n's randomization for the construction of some states. The cardinality of this message is a finite fixed number;

- so that player $i-1$ knows $\left(a_{-i,t}, y_{-i,t}, \varphi_{-i,t}\right)_{t\in T(r,k(r,i))}$,

  - first, for each r, player $i-1$ sends the message about $k(r,i)$ to players $-(i-1,i)$.[10] Each player $j \in -(i-1,i)$ infers $k(r,i)$ from their private signals. Let $k_j(r,i)$ be player j's inference. The cardinality of this message is no more than $T^{\frac{3}{4}}$; and
  - second, player j sends the messages about $\left(a_{j,t}, y_{j,t}, \varphi_{j,t}\right)_{t\in T(r,k_j(r,i))}$ to player $i-1$. The cardinality of this message is $\exp(\Theta(T^{\frac{1}{4}}))$; and

- if player j is player 1, then player 1 sends the message about player 1's history in the round for conditional independence: $(a_{1,t}, y_{1,t})$ for all t in the round for conditional independence as in the two-player case. The cardinality of this message is $\exp(\Theta(T^{\frac{1}{4}}))$.

Therefore, the cardinality of the whole message is $\exp(\Theta(T^{\frac{1}{4}}))$ and the length of the sequence of binary messages $\{G, B\}$ necessary to encode the information is $\Theta(T^{\frac{1}{4}})$. To send a binary message $m \in \{G, B\}$, player j repeats a_j^m for $T^{\frac{1}{3}}$ times to increase the precision. The other players $-j$ take $\alpha_{-j}^{\text{report}}$. Importantly, player $i-1$ can identify player j's message by $\Psi^i_{j\to i-1}$. (7.29) implies that player i cannot manipulate player $(i-1)$'s inference. The incentive to tell the truth is automatically satisfied since player j's message is used only for the reward on player i with $i \neq j$.

7.9 Probability of Errors in the Report and Re-Report Blocks

Note that the cardinality of the whole messages in the report and re-report blocks is $\exp(\Theta(T^{\frac{1}{4}}))$. Hence, the length of the sequence of binary messages $\{G, B\}$ that each player takes to send the messages in the report or re-report block is $\Theta(T^{\frac{1}{4}})$.

[10] We assume that player $i-1$ knew player i's inference $k(r,i)$. See Section 7.9 for how to deal with the small probability that player $i-1$ mis-interprets player i's message about $k(r,i)$ in the report block.

Since all the messages transmit correctly with probability at least

$$1 - \Theta(T^{\frac{1}{4}}) \exp(-\Theta(T^{\frac{1}{3}})),$$

by the same treatment as in Section 6.7.4, we can assume as if all the messages would transmit correctly. We do not apply this procedure for the messages in the round for conditional independence. As seen in Lemma 46, the incentive in the round for conditional independence is established taking into account the probability of mis-transmission.

Bibliography

Abreu, D., P. Milgrom, and D. Pearce (1991): "Information and timing in repeated partnerships," *Econometrica*, 56(2), 1713–1733.

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

Bhaskar, V., and I. Obara (2002): "Belief-based equilibria in the repeated prisoners' dilemma with private monitoring," *Journal of Economic Theory*, 102(1), 40–69.

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.

Fong, K., O. Gossner, J. Hörner, and Y. Sannikov (2010): "Efficiency in a repeated prisoners' dilemma with imperfect private monitoring," *mimeo*.

Fuchs, W. (2007): "Contracting with Repeated Moral Hazard and Private Evaluations," *American Economic Review*, 97(4), 1432–1448.

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.

HARRINGTON JR, J., AND A. SKRZYPACZ (2007): "Collusion under monitoring of sales," *The RAND Journal of Economics*, 38(2), 314–331.

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.

——— (2010): "Towards a belief-based theory of repeated games with private monitoring: An application of POMDP," *mimeo*.

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 S. Morris (2002): "Repeated games with almost-public monitoring," *Journal of Economic Theory*, 102(1), 189–228.

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.

Phelan, C., and A. Skrzypacz (2012): "Beliefs and private monitoring," *The Review of Economic Studies*.

Piccione, M. (2002): "The repeated prisoner's dilemma with imperfect private monitoring," *Journal of Economic Theory*, 102(1), 70–83.

Radner, R. (1985): "Repeated principal-agent games with discounting," *Econometrica*, 53(5), 1173–1198.

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): “belief-free review-strategy equilibrium without conditional independence,” *mimeo*.

Sugaya, T., and S. Takahashi (2011): “Coordination failure in repeated games with private monitoring,” *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*.