UTMD Working Paper
The University of Tokyo
Market Design Center

UTMD-066

Dynamics of Common Beliefs and the Value of Communication

Konan Shimizu
Graduate School of Economics, the University of Tokyo

August 27, 2024

Dynamics of Common Beliefs and the Value of Communication

Konan Shimizu*
Graduate School of Economics, the University of Tokyo.

August 27, 2024

Abstract

We analyze the evolution of common beliefs among two agents, where one agent (the informed) knows the true state while the other (the uninformed) learns the true state over T periods via i.i.d. private signals. We examine how information sharing among them affects the emergence of common beliefs. Information sharing is assumed to be non-strategic and imperfect in the sense that it imperfectly reveals the signals received by the uninformed agent. We define *the value of information sharing* as (the probability that the true state is common p-believed under information sharing) - (the probability under no information sharing), and we show that the value of information sharing depends on T (how long the uninformed agent learns). Intuitively, information sharing appears to help achieve common beliefs, but this is only true when T is small. The value of information sharing is positive for small T, but it becomes negative when T is larger. Moreover, the value of information sharing can be equal to zero for all sufficiently large values of T, and we provide almost necessary and sufficient conditions for this to hold.

Keywords: common beliefs, higher-order beliefs, coordinated attack, comparison of signal structures

1 Introduction

1.1 Overview

Consider a situation in which there are two divisions of an army, each with a commanding general, general 1 and general 2. Each will decide simultaneously whether to attack the enemy or not after T days from now. The enemy is either weak or strong, and it can be defeated only when (i) it is weak *and* (ii) two divisions attack simultaneously. Suppose only general 1 knows that the enemy is weak, whereas general 2 receives a signal every day $t = 1, 2, ..., T$ to learn if the enemy is weak or not. We consider two cases. The first is without information sharing: general 1 does not receive any information about the signal received by general 2. The other is with (imperfect) information sharing: general 2 imperfectly communicates the

*shimizu-konan@g.ecc.u-tokyo.ac.jp
I would like to express my sincere gratitude to my advisor, Michihiro Kandori, for his continuous support. I am also particularly grateful to Satoru Takahashi for his constructive suggestions. A special gratitude I give to Drew Fudenberg, Ryota Iijima, and Yu Awaya for their grateful comments.

value of her signal to general 1. Does information sharing help to achieve a coordinated attack? A coordinated attack, or simultaneous attack, succeeds when the fact that the enemy is weak is close to common knowledge. This proximity to common knowledge is captured by the concept of common p-belief, whose precise definition is given shortly. Intuitively, looking at the definition of common knowledge, one might expect that information sharing would facilitate a coordinated attack. It turns out, however, that the answer depends on how many days before the choice to attack is made (i.e., how long general 2 learns). In the toy example presented in the next section, if $T < 5$, coordinated attacks are more likely to succeed with information sharing, but when $5 \leq T < 10$, information sharing backfires, making a coordinated attack less likely. Furthermore, when the time to attack is sufficiently long ($T \geq 10$), the presence or absence of information sharing does not affect the success probability of the coordinated attack at all.

In this paper, we will show that results similar to this particular example hold in quite general settings. Specifically, we consider the following situation. There are two agents (agent 1 (he), agent 2 (she)), and agent 1 knows the true state of the world θ in advance. Assume that the state space is finite. Time is discrete, and at the beginning of each period $t = 1, 2, ..., T$, agent 2 receives an i.i.d. signal x_{t2} about the true state. We studied two cases. In one, there is no information sharing; thus, only agent 2 receives the signal T times. In the other, there is information sharing[1]. In this case, after agent 2 receives a signal in each period, she shares the information about the value of the signal. This information may be subject to some noise, so mathematically, agents 1 and 2 will receive a signal (x_{t1}, x_{t2}) in each period from a joint distribution depending on θ.

We employ the notion of ***common*** ***p-belief*** introduced by Monderer and Samet (1989) to formalize the situation in which the actual state θ is approximately common knowledge. An event is said to be p-believed when each agent assigns at least probability p to that event. Also, an event is said to be common p-believed if it is p-believed, it is p-believed that it is p-believed, and so on ad infinitum. The formal definition is given in Section 2. Given the distribution of signals, we define the ***value of information sharing with respect to*** T ($V_p(T)$) as the difference in probability of a true state being common p-believed at the end of T periods with and without information sharing, i.e.,

$$V_p(T) = Prob(\textit{true state is common p-believed at T with information sharing}) \\ - Prob(\textit{true state is common p-believed at T without information sharing}).$$

Given the definition of common p-belief, it might seem intuitive that $V_p(T) \geq 0$ holds in all cases. This is certainly true when the information sharing is perfect (Lemma 2). In the imperfect case, however, we show in Theorem 1 that the sign of $V_p(T)$ depends on T. Specifically when T is small, $V_p(T) \geq 0$ holds, but when T is large, the naive intuition fails and $V_p(T) \leq 0$ holds. Under some regularity conditions, those inequalities are strict (Proposition 4 and 6). Furthermore, there is an almost necessary and sufficient condition that $V_p(T) = 0$ holds when T is sufficiently large (Theorem 2).

To understand this result and the intuitive reason why it holds, consider the following coordinated attack example. Figure 1(b) in the next subsection shows how $V_p(T)$ changes over time in this example.

[1] Here, information sharing is not strategic.

1.2 Example: Coordinated Attack With and Without Information Sharing

An army has two divisions, each with a commanding general: general 1 (he) and general 2 (she). They can potentially attack an enemy in T days. There are two possible states $\Theta = \{S, W\}$, and these correspond to the enemy being either strong (state S) or weak (state W). Only general 1 knows the strength of the enemy. Each general has to choose simultaneously whether to attack the enemy (action A) or not (action N) at the end of day T. Before the attack, on each day $t = 1, 2, ..., T$, general 2 receive a private i.i.d. signal about the true state. In particular, on each day, general 2 receives news that the enemy is weak with probability $1 - \varepsilon$ and no news with probability ε, where $0 < \varepsilon < 1$, if the state is W. She always receives no news if the true state is S. The enemy can be defeated only if the enemy is weak and the two divisions attack simultaneously. Each gains M if the enemy is defeated but otherwise receives $-L$, where $L > M > 0$.

The payoff matrix for each state is as follows.

1 \2	N	A
N	$0, 0$	$0, -L$
A	$-L, 0$	$-L, -L$

State: $\theta = S$

1 \2	N	A
N	$0, 0$	$0, -L$
A	$-L, 0$	M, M

State: $\theta = W$

In state W, the game is a symmetric coordination game, where the Pareto inferior equilibrium (N, N) is risk-dominant (because $L > M > 0$) In state S, N is a weakly dominant strategy for both generals.

We consider the following two cases.

Case 1: without information sharing: General 1 does not know if general 2 received news about the weakness of the enemy or not.

Case 2: with information sharing: When general 2 receives the news that the enemy is weak, a message is sent to general 1. However, this process is noisy, and the message fails to reach general 1 with probability δ, where $0 < \delta < 1$, in which case general 1 does not know whether general 2 received news or not, and general 2 knows that the message was not received by general 1. A message is sent each day when general 2 receives the news about the enemy's weakness.

In this game, it is a Bayesian Nash equilibrium (BNE) for both generals to always take action N regardless of the signals they receive. Since (A, A) is efficient in state W, however, we would like to examine whether there exists a BNE where a ***coordinated attack*** (both choose A in state W) occurs.

Define ***the best BNE*** as the BNE that achieves the highest probability of coordinated attack among the BNEs in this game. It turns out that the notion of common p-belief (Monderer and Samet, 1989) characterizes the probability of a coordinated attack in the best BNE, as the following Proposition shows.

Proposition 1. *Let $p = \frac{L}{L+M}$. In both cases, i.e., with or without information sharing, a best BNE exists, and the probability that a coordinated attack occurs under the best BNE is equal to the probability that state W is common p-believed at the end of day T.*

The proof can be found in Appendix A.

Figure 1(a) shows the relationship between the number of signals received and the probability of a coordinated attack under the best BNE for each case when

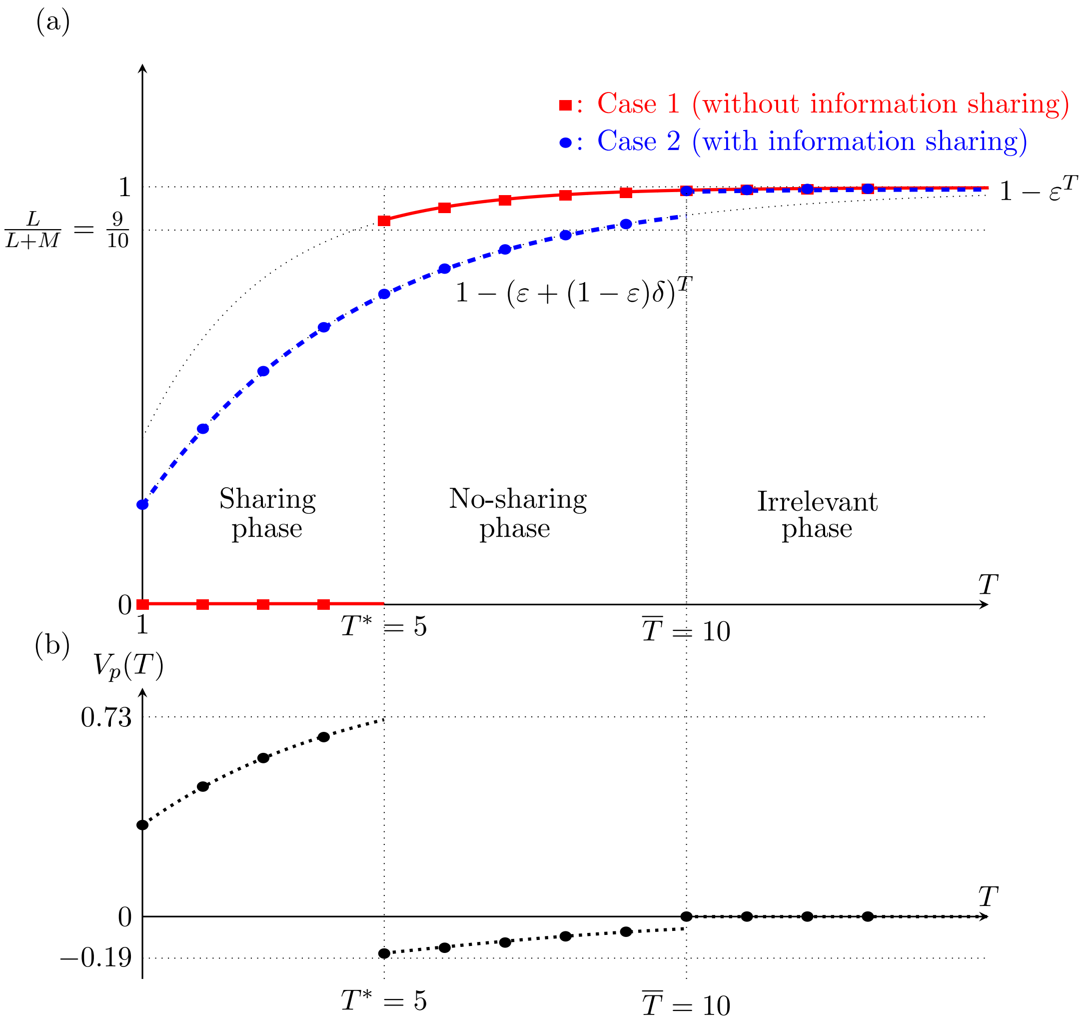

Figure 1: (a): Probability of coordinated attack under the best BNE when the state is W. (b): The value of $V_p(T)$ when the state is W.

$\frac{L}{M} = 9$, $\varepsilon = 0.6$, and $\delta = 0.4$[2]. The formal analysis used to derive this figure is found in Example 4 in the following section.

The vertical axis represents the probability of a coordinated attack in the best BNE when the state is W, which is the same as the probability that state W is common $\frac{L}{L+M}$-believed in each case. Intuitively, information sharing seems to facilitate better coordination, but this is true only when T is sufficiently small, and the opposite is true when T is larger. For instance, when general 2 has $T = 4$ days to learn the weakness of the enemy, the coordinated attack succeeds about 65% of the time when there is information sharing but never succeeds when there is no information sharing. When she can learn for $T = 5$ days, however, information sharing backfires; the coordinated attack succeeds about 73% of the time with information sharing, while it succeeds about 92% of the time without information sharing.

Figure 1(b) shows the associated value of information sharing $V_p(T)$ for $p = \frac{L}{L+M}$. As the figure shows, we first have a range of T where the value of information sharing is positive, and we call this the ***sharing phase***. Then comes the range of T where

[2]It is assumed that prior distribution $p_0 \in \Delta(\Theta)$ satisfies $p_0(S) > \frac{1}{16}$ to avoid the situation in which general 2 always $\frac{L}{L+M}$-believes state W regardless of the signals she received.

the value of information sharing is negative, which is called the ***no-sharing phase***. After that, we have the ***irrelevant phase***, where the value of information sharing is exactly equal to zero. We will show that this pattern in the dynamics of the value of information sharing happens under quite a general set of conditions.

1.3 Intuition

Using this example, we illustrate the intuitive reason why the sign of $V_p(T)$ depends on T. First, when T is small, in the absence of information sharing, general 1 considers the probability that general 2 p-believes the true state W to be small. Hence, the true state is never common p-believed in this case. On the other hand, if there is information sharing, W might be common p-believed if general 1 receives a signal that indicates that general 2 knows that the state is W. Therefore, $V_p(T) > 0$ holds, and this part matches our intuition. When T is of intermediate size, in contrast, $V_p(T) < 0$ holds for the following reason. If there is no information sharing, general 1 considers the probability that general 2 receives the news that the enemy is weak at least once (and thus she p-believes W) to be high because T is large. Then, it can be shown (Proposition 2) that W is common p-believed if and only if general 2 receives that news at least once. However, in the information sharing case, even when general 2 receives the news that the enemy is weak, there is a positive probability that general 1 does not receive any message at all during the entire T periods. If this happens, general 1 seriously doubts that general 2 knows the enemy is weak. Under such circumstances, W is not common p-believed even if general 2 p-believes W. Thus, $V_p(T) < 0$ holds. This adversary effect of information sharing never occurs when T is sufficiently large. If T is sufficiently large, even if general 1 never receives a message, general 1 believes that with a high probability general 2 must have received the news that the enemy is weak at least once. Then, W is common p-believed if and only if general 2 receives the news at least once for both cases, i.e., with and without information sharing. Thus, $V_p(T) = 0$ holds for sufficiently large T.

1.4 Comparative Statics

In the coordinated attack example, we can show that the lengths of the sharing phase and the no-sharing phase depend on the parameters of the model. First, T^*, the length of the sharing phase, is non-decreasing with respect to $\frac{L}{M}$ and ε. When $\frac{L}{M}$ is larger, "no attack" (N, N) is very strongly risk-dominant, and this means that state W must be closer to common knowledge for a coordinated attack to happen. A larger ε means that general 2 has a lower probability of receiving news. In both cases, a coordinated attack will be more difficult, and therefore, the sharing phase will be longer in cases where a coordinated attack is more difficult. Second, $\overline{T} - T^*$, the length of the no-sharing phase, is non-increasing with respect to δ, the probability that a message from general 2 does not reach general 1. When δ is small and general 1 has not received a message, he tends to think that general 2 has not received the news that the enemy is weak. Since the irrelevant phase begins when general 1 believes with a high probability that general 2 has received that news while general 1 has received no message at all, the no-sharing phase becomes longer when δ is smaller. In particular, it follows that $\overline{T} - T^* \to \infty$ as $\delta \to 0$ and $\overline{T} - T^* \to 0$ as

$\delta \to 1$. The length of the no-sharing phase is roughly an increasing function of $\frac{L}{M}$ [3]. In addition, $\overline{T} - T^*$ is roughly increasing with respect to ε if δ is relatively small, and decreasing if δ is relatively high. These comparative statics results for the example are generalized in Section 3.5. [4]

1.5 Related Literature

There are several prior studies about the dynamics of common p-beliefs, where agents receive signals about the true state independently over time. Within each period, agents' signals may be either correlated or independent given the true state, and these alternatives correspond to our information-sharing and no-sharing case. Cripps, Ely, Mailath, and Samuelson (2008, henceforth CEMS) consider the long-run outcomes and show that the true state is eventually common p-believed irrespective of information sharing, as $T \to \infty$ if the signal space is finite. Frick, Iijima, and Ishii (2022) examine the speed of convergence to those long-run outcomes and show that the speed of convergence does not depend on the presence or absence of information sharing. These papers together show that the degree of information sharing does not matter for the emergence of common p-belief in the limit $T \to \infty$. Awaya and Krishna (2022) consider the effects of information sharing when T is finite but large. In contrast, this paper provides a sharp characterization of the role of information sharing in the emergence of common p-belief for all T, including the case when T is small.

The following are the main differences between this paper and Awaya and Krishna. Awaya and Krishna show that the set of common p-belief events shrinks when there is information sharing in terms of set inclusion. On the other hand, this paper discusses how the probability of common p-belief events changes with and without information sharing. As Awaya and Krishna mention, the larger set of events does not necessarily have a larger probability. This is because the probability measure changes if we introduce information sharing. In this regard, our assessment of probabilities provides sharper conclusions about the value of information sharing. In addition, we do not assume that the signal space is finite, while this is assumed in all the literature mentioned so far. We are able to obtain stronger results because we assume that agent 1 knows the state of the world. As we discuss in Section 4, this assumption plays an important role.

2 Model

2.1 Setting

Let $I = \{1, 2\}$ be a set of agents and Θ be a finite set of states. Each period is denoted by $t = 1, 2, 3, ...$ Before $t = 1$, nature selects the state θ according to the common prior $p_0 \in \Delta(\Theta)$. Conditional on θ, in each period $t = 1, 2, 3, ...$, agent 1 and agent 2 receive a private signal x_1 and x_2. (x_1, x_2) is generated from the discrete probability distribution $\mu^\theta \in \Delta(X) = \Delta(X_1 \times X_2)$, where X_i is the set of

[3] Precisely, $T^* = \lceil \frac{log(1-p)}{log\varepsilon} \rceil$ and $\overline{T} = \lceil \frac{log(1-p)}{log\varepsilon - log[\varepsilon + (1-\varepsilon)\delta]} \rceil$ hold. Then, we can say that $\{ \frac{log(1-p)}{log\varepsilon - log[\varepsilon+(1-\varepsilon)]} \} - \{ \frac{log(1-p)}{log\varepsilon} \}$ is increasing with respect to ε.

[4] Precisely, there exists δ^* such that $\{ \frac{log(1-p)}{log\varepsilon - log[\varepsilon+(1-\varepsilon)]} \} - \{ \frac{log(1-p)}{log\varepsilon} \}$ is increasing with respect to ε if $\delta \leq \delta^*$, and decreasing if $\delta \geq \delta^*$

signals for agent i. Assume that X_i is at most a countable set[5]. In this paper, we assume that agent 1 knows the true state. Due to this assumption, we only consider the case that $X_1 \subset \Theta \times M_1$ and the signal $(\mu^\theta)_{\theta\in\Theta}$ such that the first component of the signal received by agent 1 is always θ, where M_1 is a set of messages for agent 1. We call an information structure that satisfies these conditions a signal structure. Precisely, a ***signal structure*** $\mathcal{S}$ is a tuple of X and $(\mu^\theta)_{\theta\in\Theta}$ that satisfies $X_1 \subset \Theta \times M_1$ for some set M_1, $\mu^\theta \in \Delta(X)$ and $\mu^\theta(x_1, x_2) = 0$ for all $\theta \in \Theta$, and $x_1 = (\theta', m_1) \in X_1$ with $\theta' \neq \theta$, and $x_2 \in X_2$. Note that since $\mathcal{S}$ only specifies the joint distribution of signals, it can represent a situation where, for example, agent 2 first receives the signal and then tries to convey the information to agent 1. Assume that in each period, the signals are generated independently and identically from μ^θ. Let $\mu_i^\theta \in \Delta(X_i)$ be the marginal distribution of μ^θ for each agent i. Denote as $\mu^t \in \Delta(\Theta \times X^t)$ the distribution over states and signal profiles induced by the prior p_0 and signal distribution μ^θ. Define its marginal distributions $\mu^{t,\theta} \in \Delta(X^t)$, $\mu_i^{t,\theta} \in \Delta(X_i^t)$, and $\mu_i^t \in \Delta(\Theta \times X_i^t)$.[6] By an abuse of notation, we use μ to mean μ^t when $t = 1$. In this paper, we assume, without loss of generality, that each agent's marginal distribution has full support. That is, for all $x_1 \in X_1$ and $x_2 \in X_2$,

$$\mu(\Theta \times \{x_1\} \times X_2) > 0 \textit{ and } \mu(\Theta \times X_1 \times \{x_2\}) > 0.$$

Now, denote the set of signals of agent 1 when the state is θ as $X_1(\theta)$. That is,

$$X_1(\theta) = X_1 \cap (\{\theta\} \times M_1).$$

Event E is a subset of $\Omega = \Theta \times X^t$. We assume that $\mu_2^\theta \neq \mu_2^{\theta'}$ for all θ, θ' with $\theta \neq \theta'$. Therefore, in this paper, we consider a situation where agent 1 knows the state and agent 2 does not know it but learns it by repeatedly receiving signals. It is important to note that although agent 1 knows the state, it may not be possible to determine exactly to what extent agent 2 has knowledge of the state. An example of this situation is the structure of the email game introduced by Rubinstein (1989). We will discuss this example later.

A ***basic game*** $\mathcal{G}$ consists of a finite set of actions A_i for each $i \in I$, and utility function $u_i : \Theta \times A_1 \times A_2 \to \mathbb{R}$. In the previous example, we considered the ***incomplete information game*** $\mathcal{G}(t, \mathcal{S})$. The game $\mathcal{G}(t, \mathcal{S})$ proceeds as follows. First, the state and signals are drawn from μ^t. Then, each agent observes a private signal and chooses its action according to the signal. $\sigma_i : X_i^t \to \Delta(A_i)$ is a strategy for agent i in this game. A strategy profile (σ_1, σ_2) is a Bayesian Nash equilibrium (BNE) of $\mathcal{G}(t, \mathcal{S})$ if

$$\sum_{\theta\in\Theta, x_j^t\in X_j^t} \mu^t(\theta, x_i^t, x_j^t | x_i^t) \sum_{a_j\in A_j} \sigma_j(a_j|x_j^t) u_i(\theta, a_i, a_j) \geq \sum_{\theta\in\Theta, x_j^t\in X_j^t} \mu^t(\theta, x_i^t, x_j^t | x_i^t) \sum_{a_j\in A_j} \sigma_j(a_j|x_j^t) u_i(\theta, a_i', a_j)$$

for all $i, j \in \{1, 2\}$, $j \neq i$, $x_i^t \in X_i^t$, $a_i' \in A_i$ and $a_i \in A_i$ with $\sigma_i(a_i|x_i^t) > 0$.

[5]In this paper,we assume that X_i is at most countable except for in Section 5.1. Although almost the same proof shows that all the results hold true even if X_i is uncountable, we make this assumption to simplify the notation.

[6]X^t and X_i^t mean $(X)^t$ and $(X_i)^t$ respectively.

2.2 Learning

In this paper, we analyze the extent to which the signal distribution influences state learning. To do so, we use the notation for ***common learning*** followed by CEMS (2008).

Let $\mathcal{S} = \langle X, (\mu^\theta)_{\theta\in\Theta}\rangle$ be a signal structure. Fix any $p \in (0,1)$, $t \in \mathbb{N}$, and $E \subseteq \Omega$. Let $B_{ti}^{p,\mathcal{S}}(E)$ be the event that agent i ***p-believes*** E at period t. Formally,

$$B_{ti}^{p,\mathcal{S}}(E) = \Theta \times \mathcal{B}_{ti}^{p,\mathcal{S}}(E) \times X_j^t,$$

where $j \neq i$ and $\mathcal{B}_{ti}^{p,\mathcal{S}}(E) = \{x_i^t \in X_i^t | \mu^t(E|x_i^t) \geq p\}$.
Thus, $B_{ti}^{p,\mathcal{S}}(E)$ is an event such that, when it occurs, agent i assigns at least probability p to event E. Denote $B_t^{p,\mathcal{S}}(E)$ as the event that E is p-believed by all agents at period t, i.e.,

$$\begin{aligned} B_t^{p,\mathcal{S}}(E) &= B_{t1}^{p,\mathcal{S}}(E) \cap B_{t2}^{p,\mathcal{S}}(E) \\ &= \Theta \times \{x_1^t \in X_1^t | \mu^t(E|x_1^t) \geq p\} \times \{x_2^t \in X_2^t | \mu^t(E|x_2^t) \geq p\}. \end{aligned}$$

In addition, let $C_t^{p,\mathcal{S}}(E)$ denote the event that agents ***common p-believe*** E at period t. Formally,

$$C_t^{p,\mathcal{S}}(E) = \bigcap_{k\in\mathbb{N}} \left(B_t^{p,\mathcal{S}}\right)^k (E).$$

Hence in $C_t^{p,\mathcal{S}}(E)$, each agent assigns at least probability p to event E, and also to event $B_t^{p,\mathcal{S}}(E)$, and also to event $B_t^{p,\mathcal{S}}(B_t^{p,\mathcal{S}}(E))$, and so on. By construction, this can be written as

$$C_t^{p,\mathcal{S}}(E) = \Theta \times C_{t1}^{p,\mathcal{S}}(E) \times C_{t2}^{p,\mathcal{S}}(E).$$

For a tuple (θ, x_1^t, x_2^t) of some true state θ and a signal profile (x_1^t, x_2^t) to be an element of $B_t^{p,\mathcal{S}}(E)$, each agent only has to assign at least probability p to event E when each receives signal (x_1^t, x_2^t), and thus only first-order belief matters. On the other hand, for (θ, x_1^t, x_2^t) to be an element of $B_t^{p,\mathcal{S}}(B_t^{p,\mathcal{S}}(E))$, agent 1 must not only assign at least probability p to event E but also assign at least probability p to the fact that the signal being received by agent 2 is an element of $\{x_2^t \in X_2^t | \mu^t(E|x_2^t) \geq p\}$ (and this is also true for agent 2). Thus, second-order beliefs are at issue here, and, in general, all higher-order beliefs are taken into account when considering $C_t^{p,\mathcal{S}}(E)$.

Hereafter, as an abuse of the notation, the event $\{\theta\} \times X_1^t \times X_2^t$ is sometimes simply denoted as θ in clear cases. From this, $B_t^{p,\mathcal{S}}(\theta)$, $C_t^{p,\mathcal{S}}(\theta)$, and so on are defined.

Agent i ***(individually) learns*** θ under $\mathcal{S}$ if for each $p \in (0,1)$,

$$\lim_{t\to\infty} \mu^t(B_{ti}^{p,\mathcal{S}}(\theta)|\theta) = 1.$$

Agent i learns Θ if i learns θ for all $\theta \in \Theta$. Also, the agents ***commonly learn*** θ under $\mathcal{S}$ if for each $p \in (0,1)$,

$$\lim_{t\to\infty} \mu^t(C_t^{p,\mathcal{S}}(\theta)|\theta) = 1.$$

The agents commonly learn Θ if they commonly learn θ for all $\theta \in \Theta$.

The common learning result of CEMS implies that, if X is finite, the agents commonly learn Θ. However, CEMS also showed that the agents possibly do not commonly learn Θ if X is an infinite countable set.

3 General Results

3.1 Without information sharing

First, we describe the properties that hold for signals without information sharing.

Definition 1. *A* ***no-sharing*** *signal structure is a signal structure* $\mathcal{S} = \langle X, (\mu^\theta)_{\theta\in\Theta}\rangle$ *that satisfies* $|X_1(\theta)| = 1$ *for all* $\theta \in \Theta$.

In a no-sharing signal structure, agent 1 cannot obtain any information other than what the true state is.

Hereafter, for signal structure $\mathcal{S} = \langle X, (\mu^\theta)_{\theta\in\Theta}\rangle$, let $q_t^{\mathcal{S}}(\theta) = \mu_2^{t,\theta}(\mathcal{B}_{t2}^{p,\mathcal{S}}(\theta))$. $q_t^{\mathcal{S}}(\theta)$ is the probability that agent 2 p-believes state θ as a result of receiving a signal t times when the true state is θ. We see that in this case, we can explicitly specify the set for which state θ is common p-believed.

Proposition 2. *Suppose that signal structure* $\mathcal{S} = \langle X, (\mu^\theta)_{\theta\in\Theta}\rangle$ *is no-sharing. Then,*

$$C_t^{p,\mathcal{S}}(\theta) = \begin{cases} \emptyset & \text{if } q_t^{\mathcal{S}}(\theta) < p \\ \Theta \times \{X_1(\theta)\}^t \times \mathcal{B}_{t2}^{p,\mathcal{S}}(\theta) & \text{if } q_t^{\mathcal{S}}(\theta) \geq p. \end{cases}$$

The intuitive statement of the result of this proposition is as follows. First, for state θ to be common p-believed, agent 2 must p-believe that the state is θ. This fact corresponds to the condition that $x_2^t \in \mathcal{B}_{t2}^{p,\mathcal{S}}(\theta)$. However, for θ to be common p-believed, this condition alone may not be sufficient. Agent 1 must believe that the other agent has some understanding of the state. Therefore, $C_t^{p,\mathcal{S}}(\theta)$ is an empty set when it is unlikely that agent 2 p-believes θ. On the other hand, $x_2^t \in \mathcal{B}_{t2}^{p,\mathcal{S}}(\theta)$ is a sufficient condition when it is likely that agent 2 p-believes θ. In this case, agent 1 p-believes that agent 2 p-believes the state, and agent 2 also p-believes "agent 1 p-believes that agent 2 p-believes the state", and so on because there is no information sharing.

Using Proposition 2, the next statement follows immediately. Here, $\mu^t(C_t^{p,\mathcal{S}}(\theta)|\theta)$ is the probability that the true state becomes common p-believed when the state is θ.

$$\mu^t(C_t^{p,\mathcal{S}}(\theta)|\theta) = \begin{cases} 0 & \text{if } q_t^{\mathcal{S}}(\theta) < p \\ q_t^{\mathcal{S}}(\theta) & \text{if } q_t^{\mathcal{S}}(\theta) \geq p. \end{cases}$$

3.2 Comparison of the with and without information sharing cases

When there is no information sharing, the condition $q_t^{\mathcal{S}}(\theta) \geq p$ is satisfied and state θ is common p-believed if agent 2 assigns at least probability p to state θ. In fact, $\mu^t(C_t^{p,\mathcal{S}}(\theta)|\theta) = q_t^{\mathcal{S}}(\theta)$ is the highest possible value of the probability that θ is common p-believed for any signal structure. A precise statement of this claim follows in Lemma 1 below.

Lemma 1. *Let* $\mathcal{S} = \langle X, (\mu^\theta)_{\theta\in\Theta}\rangle$ *be any signal structure. Then,*

$$\mu^t(C_t^{p,\mathcal{S}}(\theta)|\theta) \leq q_t^{\mathcal{S}}(\theta)$$

for all t.

The interpretation of Lemma 1 is quite simple. $q_t^{\mathcal{S}}(\theta)$ is the probability that agent 2 p-believes that the state is θ, when the true state is θ. Since it is a necessary condition for state θ to be common p-believed that agent 2 p-believes θ, the probability that state θ is common p-believed is less than or equal to this probability.

From Proposition 2 and Lemma 1, we have the following relation for any signal structure $\mathcal{S}$, which is no-sharing, and signal structure $\mathcal{S}'$, which adds information sharing to $\mathcal{S}$. This is our main result.

Definition 2. *Let $\mathcal{S}' = \langle X', (\nu^\theta)_{\theta\in\Theta}\rangle$ be any signal structure. Let $\mathcal{S} = \langle X, (\mu^\theta)_{\theta\in\Theta}\rangle$ be the no-sharing signal structure which satisfies $X_2 = X_2'$ and $\mu_2^{t,\theta} = \nu_2^{t,\theta}$ for all $\theta \in \Theta$. Given that the true state is θ, define the* ***value of information sharing with respect to*** *t ($V_p(t)$) as follows.*

$$V_p(t) = \nu^t(C_t^{p,\mathcal{S}'}(\theta)|\theta) - \mu^t(C_t^{p,\mathcal{S}}(\theta)|\theta).$$

Then, we say ***common p-belief of the true state is weakly more likely under no information sharing at*** *t if $V_p(t) \leq 0$. Similarly, we say* ***common p-belief of the true state is weakly less likely under no information sharing at*** *t if $V_p(t) \geq 0$. Analogously, we use the terms "strictly more likely" and "strictly less likely" in cases where these inequalities are held in the strict sense.*

Theorem 1. *Let $\mathcal{S}'$ and $\mathcal{S}$ be the signal structures in Definition 2. Fix $\theta \in \Theta$ and $p \in (0, 1)$. Then, $V_p(t) \geq 0$ holds if $q_t^{\mathcal{S}}(\theta) < p$ and $V_p(t) \leq 0$ holds if $q_t^{\mathcal{S}}(\theta) \geq p$.*

Note that $q_t^{\mathcal{S}}(\theta) \to 1$ as $t \to 1$ since agent 2 individually learn θ. Therefore, the following corollary follows immediately from this theorem.

Corollary 1. *Let $\mathcal{S}'$ and $\mathcal{S}$ be the signal structures in Definition 2. Fix $\theta \in \Theta$ and $p \in (0, 1)$. There exist the largest integers t^* such that for all $t < t^*$, $q_t^{\mathcal{S}}(\theta) < p$ and the smallest integers t^{**} such that for all $t \geq t^{**}$, $q_t^{\mathcal{S}}(\theta) \geq p$. Take such t^* and t^{**}. Then, $V_p(t) \geq 0$ for $t < t^*$, and $V_p(t) \leq 0$ for $t \geq t^{**}$.*

Corollary 1 says that when t is small, the state is more likely to be common p-believed if there is information sharing, and when t is large, the state is more likely to be common p-believed if there is no information sharing. Recall that in the no-sharing case $\mathcal{S} = \langle X, (\mu^\theta)_{\theta\in\Theta}\rangle$, we have

$$\mu^t(C_t^{p,\mathcal{S}}(\theta)|\theta) = \begin{cases} 0 & \text{if } q_t^{\mathcal{S}}(\theta) < p \\ q_t^{\mathcal{S}}(\theta) & \text{if } q_t^{\mathcal{S}}(\theta) \geq p. \end{cases}$$

Hence, if $\mu^t(C_t^{p,\mathcal{S}}(\theta)|\theta) > 0$, common p-belief of the true state is weakly more likely under no information sharing at t.

The range of t where $t < t^*$ is called the ***sharing phase***. If there exists $\bar{t} \geq t^{**}$ such that $V_p(t) = 0$ for all $t \geq \bar{t}$, then take the smallest one and we call the range of t where $t^{**} \leq t < \bar{t}$ the ***no-sharing phase*** and where $t \geq \bar{t}$ the ***irrelevant phase***. If there does not exist such $\bar{t}$, we call the range of t where $t \geq t^{**}$ the no-sharing phase. Note that t^* and t^{**} in Corollary 1 may not be the same. This is because $q_t^{\mathcal{S}}(\theta)$ converges to 1 but does not necessarily increase monotonically. In this case, we call the range of t where $t^* \leq t < t^{**}$ the ***irregular phase***. The following is an example of an irregular phase.

Example 1.

Let $\Theta = \{\alpha, \beta\}$ and $\mathcal{S} = \langle X, (\mu^\theta)_{\theta \in \Theta} \rangle$ be the signal structure which is no-sharing and satisfies $X_2 = \{0, 1\}$ and

$$\mu^\alpha(X_1 \times \{0\}) = 0.5$$
$$\mu^\alpha(X_1 \times \{1\}) = 0.5$$
$$\mu^\beta(X_1 \times \{0\}) = 0.9$$
$$\mu^\beta(X_1 \times \{1\}) = 0.1.$$

Suppose that $p_0(\beta) = p_0(\alpha) = 0.5$. Let $p = 0.875$. The value of $q_t^{\mathcal{S}}(\beta)$ and $\mu^t(C_t^{p,\mathcal{S}}(\beta)|\beta)$ is shown in Figure 2.
In this case, $q_t^{\mathcal{S}}(\beta)$ does not increase monotonically, and $t^* = 6$ and $t^{**} = 14$ hold. The reason that $q_t^{\mathcal{S}}(\beta)$ is not monotonically increasing is as follows. When $t = 2$, in order for agent 2 to p-believe state β, she must receive the signal 0 both times. When $t = 3$, receiving the signal $x_2 = 0$ twice is not enough for agent 2 to assign a sufficient probability to state β, so agent 2 needs to receive the signal 0 all three times to p-believe state β. Similarly, it can be calculated that all signals received by agent 2 must be 0 until $t = 5$. Hence, $q_t^{\mathcal{S}}(\beta)$ is decreasing until $t = 5$. On the other hand, when $t = 6, 7, 8,$ and 9, it can be calculated that agent 2 p-believes state β even if the agent receives the signal $x_2 = 1$ at most once. Thus $q_t^{\mathcal{S}}(\beta)$ jumps up at $t = 6$ and then decreases until $t = 9$. At $t = 10$ onward, $q_t^{\mathcal{S}}(\beta)$ behaves as shown in Figure 2 for the same reason.

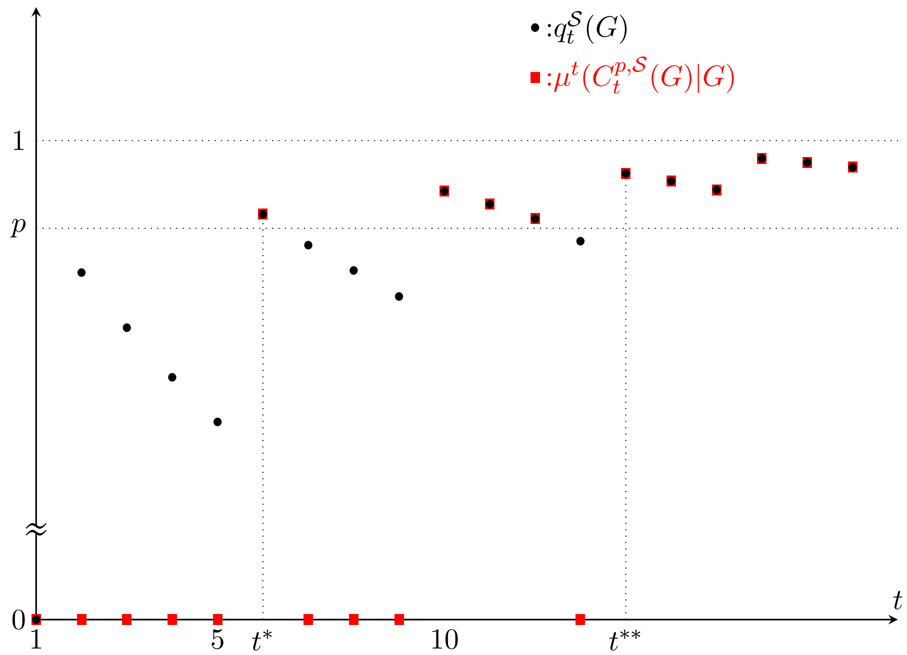

Figure 2: The case of $t^* \neq t^{**}$

As we saw in this example, t^* and t^{**} do not necessarily coincide, but the important thing is that such t^*, t^{**} always exist due to the property that $q_t^{\mathcal{S}}(\theta)$ converges to 1. Hence, the no-sharing phase or the irrelevant phase must exist. In addition, if the probability that agent 2 p-believes θ is non-decreasing function of t, $t^* = t^{**}$ must hold, and the irregular phase disappears.

The result that the state is weakly more likely to be common p-believed if there is no information sharing when t is large corresponds to that of Awaya and Krishna (2022). They show that for the case of $|\Theta| = 2$ and binary and conclusive full support signals, the more uncorrelated the signals are, the larger the set for which the state is common p-believed when t is sufficiently large in terms of set inclusion. We compare not the inclusion relationship of the common p-believed set, but the ***probability*** that a state is common p-believed. For this purpose, we assume that agent 1 knows the true state, but we do not make any additional assumptions about the distributional support, and we can, therefore, also consider the signal of perfect correlation.

There are cases in which $V_p(t) \leq 0$ for all t. That is, common p-belief of the true state is weakly more likely under no information sharing at t for all t. Specifically, we have the following Example 2. This is the well-known signal structure of the email game (Rubinstein, 1989) with repetitions.

Example 2. *Repeated email game.*

Let $\Theta = \{\alpha, \beta\}$ and $\mathcal{S}' = \langle X', (\nu_\theta)_{\theta \in \Theta} \rangle$ be the signal structure which satisfies $X_1'(\alpha) = \{(\alpha, 0)\}$, $X_1'(\beta) = \{\beta\} \times \mathbb{Z}_{\geq 1}$, $X_2' = \mathbb{Z}_{\geq 0}$, and

$$\begin{aligned}
&\nu^\alpha((\alpha, 0), 0) = 1 \\
&\nu^\beta((\beta, m), m-1) = (1-\varepsilon)^{2m-2}\varepsilon \\
&\nu^\beta((\beta, m), m) = (1-\varepsilon)^{2m-1}\varepsilon
\end{aligned}$$

for all $m \in \mathbb{Z}_{\geq 1}$, where $\varepsilon \in (0,1)$. Let $\mathcal{S} = \langle X, (\mu^\theta)_{\theta \in \Theta} \rangle$ be the signal structure which is no-sharing and induced by $\mathcal{S}'$, i.e., $X_2 = X_2'$ and μ and ν have he same marginal distribution for agent 2.

$\mathcal{S}'$ corresponds to the situation of Rubinstein's email game with reputation. That is, agent 1 knows the true state $\theta \in \{\alpha, \beta\}$. If the state is β, agent 1 tries to tell agent 2 that the state is β by sending an email to agent 2. Each agent replies to the other when it receives the email, but each email is lost with a small probability ε. If (m_1, x_2) is the number of times each agent has sent an email, then $((\beta, m_1), x_2)$ follows ν^β. Now suppose $\frac{p_0(\beta)\varepsilon}{p_0(\alpha)+p_0(\beta)\varepsilon} < p$, which is satisfied when ε is small enough. As is well known, when $p > \frac{1}{2}$ and $t = 1$, $C_t^{p,\mathcal{S}'}(\beta) = \emptyset$ (Rubinstein, 1989). However, increasing t creates the possibility that state β is common p-believed for relatively small t.

Proposition 3. *In this signal structure with* $\frac{p_0(\beta)\varepsilon}{p_0(\alpha)+p_0(\beta)\varepsilon} < p$,

$$C_t^{p,\mathcal{S}}(\beta) = \begin{cases} \emptyset & \text{if } p > 1 - \varepsilon^t \\ \Theta \times \{X_1(\beta)\}^t \times X_2^t \backslash \{(0,0,...,0)\} & \text{if } p \leq 1 - \varepsilon^t \end{cases}$$

$$C_t^{p,\mathcal{S}'}(\beta) = \begin{cases} \emptyset & \text{if } p > 1 - \left(\frac{1}{2-\varepsilon}\right)^t \\ \Theta \times \{X_1'(\beta)\}^t \times X_2^t \backslash \{(0,0,...,0)\} & \text{if } p \leq 1 - \left(\frac{1}{2-\varepsilon}\right)^t \end{cases}$$

Thus,

$$\mu^t(C_t^{p,\mathcal{S}}(\beta)|\beta) = \begin{cases} 0 & \textit{if } p > 1-\varepsilon^t \\ 1-\varepsilon^t & \textit{if } p \leq 1-\varepsilon^t \end{cases}$$

$$\nu^t(C_t^{p,\mathcal{S}'}(\beta)|\beta) = \begin{cases} 0 & \textit{if } p > 1-\left(\frac{1}{2-\varepsilon}\right)^t \\ 1-\varepsilon^t & \textit{if } p \leq 1-(\frac{1}{2-\varepsilon})^t \end{cases}$$

In particular, since $1-\left(\frac{1}{2-\varepsilon}\right)^t < 1-\varepsilon^t$,

$$V_p(t) \leq 0$$

for all t. Figure 3 is a graphical representation of this result. In this figure, t^* is the smallest integer t which satisfies $p \leq 1-\varepsilon^t$, which is the same value as t^{**}. Also, $\bar{t}$ is the smallest integer t which satisfies $p \leq 1-\left(\frac{1}{2-\varepsilon}\right)^t$. Hence, the smaller the value of ε, the smaller the values of t^* and $\bar{t}$.

Note that $\nu^t(C_t^{p,\mathcal{S}'}(\beta)|\beta) \to 1$ when $t \to \infty$, so agents commonly learn β. When $p > \frac{1}{2}$, we have seen that β is not common p-believed if the agents receive a signal only once in the email game setting, but there is a possibility that β is common p-believed if the agents receive signals repeatedly. However, the probability is even higher for the no-sharing case. In particular, when $t \in [t^*, \bar{t})$, common p-belief of the true state β is strictly more likely under no information sharing at t, i.e., $V_p(t) < 0$.

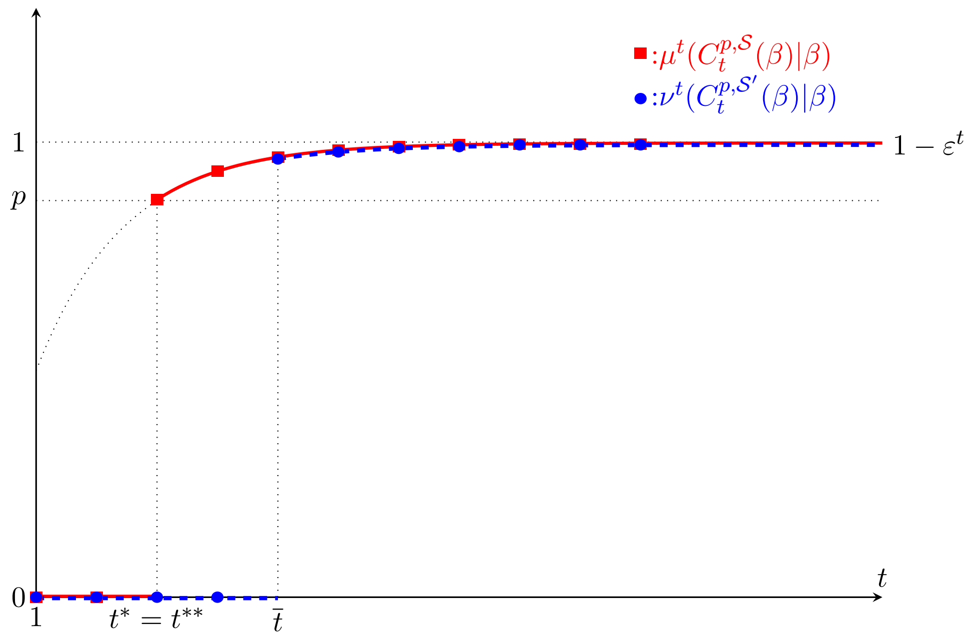

Figure 3: Probability of common p-belief in Example 2

On the other hand, there are cases where $V_p(t) \geq 0$ holds for all t. In such cases, when t is small and the condition $q_t^{\mathcal{S}}(\theta) \geq p$ is not satisfied, the state θ can be common p-believed with information sharing but cannot without it. The simplest case is that of perfect correlation.

Example 3. *Perfect correlation*

Let $\Theta = \{\alpha, \beta\}$ and $\mathcal{S}' = \langle X', (\nu_\theta)_{\theta\in\Theta}\rangle$ be the signal structure which satisfies $X_1(\alpha) = \{(\alpha, 0)\}$, $X_1(\beta) = \{(\beta, 0), (\beta, 1)\}$, $X_2 = \{0, 1\}$, and

$$\nu^\alpha((\alpha, 0), 0) = 1$$
$$\nu^\beta((\beta, 0), 0) = \varepsilon$$
$$\nu^\beta((\beta, 1), 1) = 1 - \varepsilon,$$

where $\varepsilon \in (0, 1)$. Let $\mathcal{S} = \langle X, (\mu^\theta)_{\theta\in\Theta}\rangle$ be the signal structure which is no-sharing and induced by $\mathcal{S}'$. Assume again that $\frac{p_0(\beta)\varepsilon}{p_0(\alpha)+p_0(\beta)\varepsilon} < p$.

Using almost the same argument as in Example 2, it follows that

$$C_t^{p,\mathcal{S}}(\beta) = \begin{cases} \emptyset & \textit{if } p > 1 - \varepsilon^t \\ \Theta \times \{X_1(\beta)\}^t \times X_2^t \backslash \{(0, 0, ..., 0)\} & \textit{if } p \leq 1 - \varepsilon^t, \end{cases}$$

and

$$\mu^t(C_t^{p,\mathcal{S}}(\beta)|\beta) = \begin{cases} 0 & \textit{if } p > 1 - \varepsilon^t \\ 1 - \varepsilon^t & \textit{if } p \leq 1 - \varepsilon^t \end{cases}$$

Next, consider $C_t^{p,\mathcal{S}'}(\beta)$. By the same argument as in Example 2, we have

$$B_t^{p,\mathcal{S}'}(\beta) = \Theta \times \{X_1'(\beta)\}^t \times X_2^t \backslash \{(0, 0, ..., 0)\}.$$

Since signals are perfectly correlated when $\theta = \beta$,

$$\begin{aligned} B_t^{p,\mathcal{S}'}(B_t^{p,\mathcal{S}'}(\beta)) &= \Theta \times \{X_1'(\beta)\}^t \backslash \{((\beta, 0), (\beta, 0), ..., (\beta, 0))\} \times X_2^t \backslash \{(0, 0, ..., 0)\} \\ &= B_t^{p,\mathcal{S}'}(B_t^{p,\mathcal{S}'}(B_t^{p,\mathcal{S}'}(\beta))). \end{aligned}$$

Hence, this is also equivalent to $C_t^{p,\mathcal{S}'}(\beta)$. Therefore, it follows that

$$C_t^{p,\mathcal{S}'}(\beta) = \Theta \times \{X_1'(\beta)\}^t \backslash \{((\beta, 0), (\beta, 0), ..., (\beta, 0))\} \times X_2^t \backslash \{(0, 0, ..., 0)\},$$

and

$$\nu^t(C_t^{p,\mathcal{S}'}(\beta)|\beta) = 1 - \varepsilon^t.$$

Figure 4 shows this result. In this figure, $t^* = t^{**}$ is the smallest integer t which satisfies $p \leq 1 - \varepsilon^t$.

Thus, in this case $V_p(t) \geq 0$ for any t. That is, common p-belief of the true state is weakly less likely under no information sharing at t for all t. In particular, it can be seen that β can be common p-believed in $\mathcal{S}'$ regardless of the value of t.
This result is also true for all the cases of perfect correlation.

Lemma 2. *Let $\mathcal{S}' = \langle X', (\nu_\theta)_{\theta\in\Theta}\rangle$ be a perfect correlation signal structure, i.e., $X_1' = \Theta \times X_2'$ and*

$$\nu^\theta((\theta, a), a | x_2 = a) = 1$$

for all $\theta \in \Theta$ and $a \in X_2$. Denote $\mathcal{S} = \langle X, (\mu^\theta)_{\theta\in\Theta}\rangle$ be the no-sharing signal structure induced by $\mathcal{S}'$. Then, $V_p(t) \geq 0$ for all $t = 1, 2, ...$

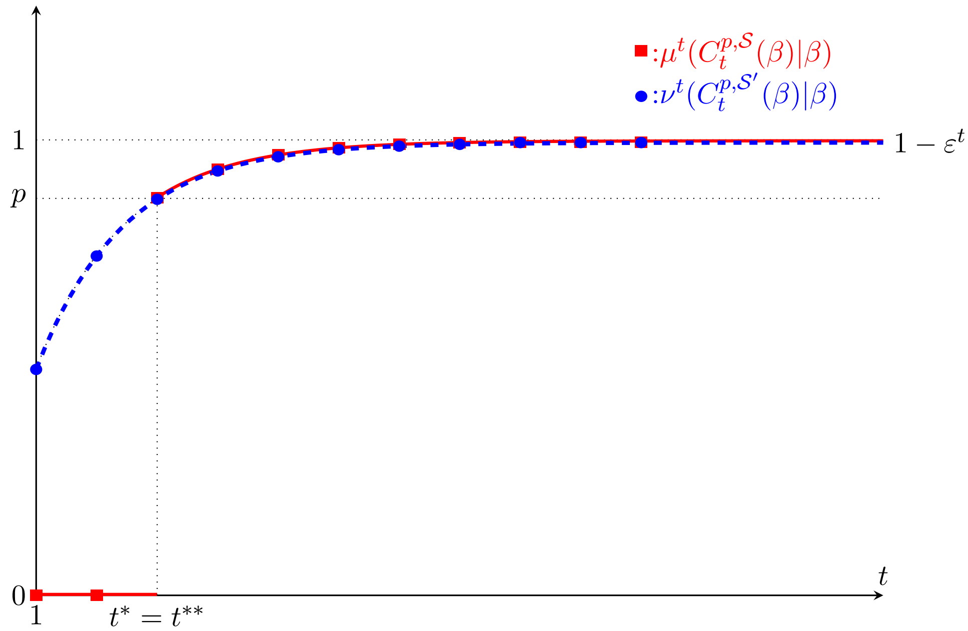

Figure 4: Probability of common p-belief in Example 3

Furthermore, as in the example we saw in the section on coordinated attacks, there is a case where both inequalities $V_p(t) < 0$ and $V_p(t') > 0$ hold strictly for some t and t'. Example 4 below is an exact expression of signal structures in the coordinated attack example.

Example 4.

Let $\Theta = \{\alpha, \beta\}$ and $\mathcal{S}' = \langle X', (\nu_\theta)_{\theta\in\Theta}\rangle$ be the signal structure which satisfies $X_1(\alpha) = \{(\alpha, 0)\}$, $X_1(\beta) = \{(\beta, 0), (\beta, 1)\}$, $X_2 = \{0, 1, 2\}$, and

$$\begin{aligned}
&\nu^\alpha((\alpha,0),0) = 1\\
&\nu^\beta((\beta,0),0) = \varepsilon\\
&\nu^\beta((\beta,0),1) = (1-\varepsilon)\delta\\
&\nu^\beta((\beta,1),2) = (1-\varepsilon)(1-\delta),
\end{aligned}$$

where $\varepsilon, \delta \in (0, 1)$. Let $\mathcal{S} = \langle X, (\mu^\theta)_{\theta\in\Theta}\rangle$ be the signal structure which is no-sharing and induced by $\mathcal{S}'$. Assume that $\frac{p_0(\beta)\varepsilon}{p_0(\alpha)+p_0(\beta)\varepsilon} < p$.

Again, it follows that

$$C_t^{p,\mathcal{S}}(\beta) = \begin{cases} \emptyset & \textit{if } p > 1-\varepsilon^t \\ \Theta \times \{X_1(\beta)\}^t \times X_2^t \backslash \{(0,0,...,0)\} & \textit{if } p \leq 1-\varepsilon^t, \end{cases}$$

and

$$\mu^t(C_t^{p,\mathcal{S}}(\beta)|\beta) = \begin{cases} 0 & \textit{if } p > 1-\varepsilon^t \\ 1-\varepsilon^t & \textit{if } p \leq 1-\varepsilon^t \end{cases}$$

Now, consider $C_t^{p,\mathcal{S}'}(\beta)$. By the same argument as in Example 2, we have

$$B_t^{p,\mathcal{S}'}(\beta) = \Theta \times \{X_1'(\beta)\}^t \times X_2^t \backslash \{(0,0,...,0)\}.$$

Recall that if agent 2 receives $x_2^t \in X_2^t \backslash \{(0,0,...,0)\}$, she assigns at least probability p to the event that the state is β, which is equivalent to the event that agent 1 receives $x_1^t \in \{X_1'(\beta)\}^t$. Conversely, if agent 1 receives $x_1^t \in \{X_1'(\beta)\}^t$, he knows that agent 2 receives $x_2^t \in X_2^t \backslash \{(0,0,...,0)\}$ unless $x_1 = ((\beta,0),(\beta,0),...,(\beta,0))$. Furthermore, if agent 1 receives $x_1^t = ((\beta,0),(\beta,0),...,(\beta,0))$, he assigns at least probability p to the event that agent 2 receives $x_2^t \in x_2^t \in X_2^t \backslash \{(0,0,...,0)\}$ if and only if $\nu^t(\Theta \times \{X_1'(\beta)\}^t \times \{(0,0,...,0)\}) \leq 1-p$, which is equivalent to $\left(\frac{\varepsilon}{\varepsilon+(1-\varepsilon)\delta}\right)^t \leq 1-p$. Therefore,

$$B_t^{p,\mathcal{S}'}(B_t^{p,\mathcal{S}'}(\beta)) = \begin{cases} \Theta \times \{X_1'(\beta)\}^t \times X_2^t \backslash \{(0,0,...,0)\} \\ \qquad\qquad\qquad\qquad \textit{if } 1 - \left(\frac{\varepsilon}{\varepsilon+(1-\varepsilon)\delta}\right)^t \geq p \\ \Theta \times \{X_1'(\beta)\}^t \backslash \{((\beta,0),(\beta,0),...,(\beta,0))\} \times X_2^t \backslash \{(0,0,...,0)\} \\ \qquad\qquad\qquad\qquad \textit{if } 1 - \left(\frac{\varepsilon}{\varepsilon+(1-\varepsilon)\delta}\right)^t < p \end{cases}$$

Hence, if $1 - \left(\frac{\varepsilon}{\varepsilon+(1-\varepsilon)\delta}\right)^t \geq p$,

$$C_t^{p,\mathcal{S}'}(\beta) = \Theta \times \{X_1'(\beta)\}^t \times X_2^t \backslash \{(0,0,...,0)\}.$$

Next, suppose that $1 - \left(\frac{\varepsilon}{\varepsilon+(1-\varepsilon)\delta}\right)^t < p$. In this case,

$$\begin{aligned} \{B_t^{p,\mathcal{S}'}\}^3(\beta) &= \Theta \times \{X_1'(\beta)\}^t \backslash \{((\beta,0),(\beta,0),...,(\beta,0))\} \times \{x_2^t \in X_2^t | \max_k x_{2k}^t = 2\} \\ &= \{B_t^{p,\mathcal{S}'}\}^4(\beta). \end{aligned}$$

Hence,

$$C_t^{p,\mathcal{S}'}(\beta) = \Theta \times \{X_1'(\beta)\}^t \backslash \{\{((\beta,0),(\beta,0),...,(\beta,0))\} \times \{x_2^t \in X_2^t | \max_k x_{2k}^t = 2\}.$$

Overall,

$$\nu^t(C_t^{p,\mathcal{S}'}(\beta)|\beta) = \begin{cases} 1 - (\varepsilon + (1-\varepsilon)\delta)^t & \textit{if } p > 1 - \left(\frac{\varepsilon}{\varepsilon+(1-\varepsilon)\delta}\right)^t \\ 1 - \varepsilon^t & \textit{if } p \leq 1 - \left(\frac{\varepsilon}{\varepsilon+(1-\varepsilon)\delta}\right)^t \end{cases}$$

Since $1 - \left(\frac{\varepsilon}{\varepsilon+(1-\varepsilon)\delta}\right)^t < 1 - \varepsilon^t$, $\nu^t(C_t^{p,\mathcal{S}'}(\beta)|\beta)$, we have following result.

$$\begin{cases} \nu^t(C_t^{p,\mathcal{S}'}(\beta)|\beta) > \mu^t(C_t^{p,\mathcal{S}}(\beta)|\beta) & \textit{if } 1 - \varepsilon^t < p \\ \nu^t(C_t^{p,\mathcal{S}'}(\beta)|\beta) < \mu^t(C_t^{p,\mathcal{S}}(\beta)|\beta) & \textit{if } 1 - \varepsilon^t \geq p > 1 - \left(\frac{\varepsilon}{\varepsilon+(1-\varepsilon)\delta}\right)^t \\ \nu^t(C_t^{p,\mathcal{S}'}(\beta)|\beta) = \mu^t(C_t^{p,\mathcal{S}}(\beta)|\beta) & \textit{if } 1 - \left(\frac{\varepsilon}{\varepsilon+(1-\varepsilon)\delta}\right)^t \geq p \end{cases}$$

Figure 5 shows this result. In this figure, $t^* = t^{**}$ is the smallest integer t which satisfies $p \leq 1 - \varepsilon^t$, and $\tilde{t}$ is the smallest integer t that satisfies $p \leq 1 - \left(\frac{\varepsilon}{\varepsilon+(1-\varepsilon)\delta}\right)^t$. If $p = 0.9$, $\varepsilon = 0.6$, and $\delta = 0.4$, we have the same plot as in Figure 1(a).

It has been observed that inequalities (1) and (2) in Definition 2 may be held both weakly and strictly. Thus, the next point of interest is to analyze when each inequality is satisfied in the strict sense.

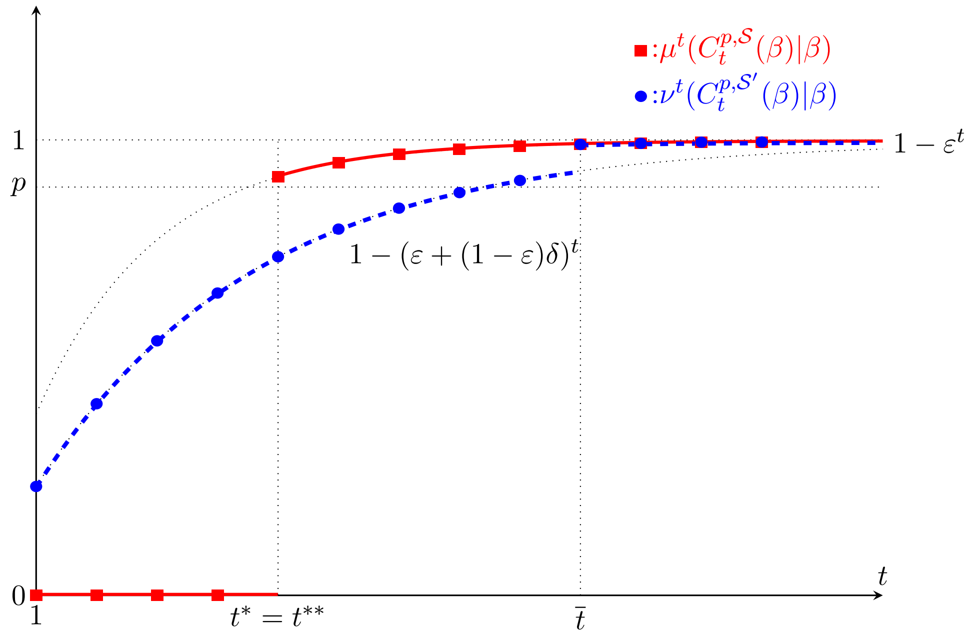

Figure 5: Probability of common p-belief in Example 4.

3.3 Conditions under which common p-belief of the true state is strictly less likely under no information sharing

First, consider the case of small t. The following proposition presents conditions, which we show to be equivalent, under which $V_p(t) > 0$ strictly holds during the sharing phase.

Proposition 4. *Let $\mathcal{S}'$ and $\mathcal{S}$ be the signal structures in Definition 2. Suppose also that $q_t^{\mathcal{S}}(\theta) < p$. Then, these conditions are equivalent.*

(1) $:\nu^t(C_t^{p,\mathcal{S}'}(\theta)) > 0$

(2) $:V_p(t) > 0$

(3) : *There exists non-empty set $E = \{\theta\} \times \tilde{X}_1^t \times \tilde{X}_2^t \subseteq \Theta \times \{X_1'(\theta)\}^t \times \mathcal{B}_{t2}^{p,\mathcal{S}'}(\theta)$ such that $\nu^t(E|x_2^t) \geq p$ and $\nu^t(E|x_1^t) \geq p$ for all $x_1^t \in \tilde{X}_1^t, x_2^t \in \tilde{X}_2^t$*

Proof. From Theorem 1, (1) and (2) are equivalent. Show that (1) and (3) are equivalent. Using the notion of p-evident event (Monderer and Samet(1989)), (3) says that there exists a nonempty p-evident event $E \subseteq \Theta \times \{X_1'(\theta)\}^t \times \mathcal{B}_{t2}^{p,\mathcal{S}'}(\theta)$ in $\mathcal{S}'$. According to Monderer and Samet, for each $\omega \in \Omega$, $\omega \in C_t^{p,\mathcal{S}'}(\theta)$ is equivalent to the condition that there exists p-evident event $E \subseteq \Omega$ such that $\omega \in E$ and $E \subseteq B_t^{p,\mathcal{S}'}(\theta)$.

Suppose $\nu^t(C_t^{p,\mathcal{S}'}(\theta)) > 0$. Since $C_t^{p,\mathcal{S}'}(\theta) \neq \emptyset$, there exists p-evident event $E' \subseteq \Omega$ which satisfies $E' = S \times \tilde{X}_1^t \times \tilde{X}_2^t \subseteq B_t^{p,\mathcal{S}'}(\theta) = \Theta \times \{X_1'(\theta)\}^t \times \mathcal{B}_{t2}^{p,\mathcal{S}'}(\theta)$. Let $E = \{\theta\} \times \tilde{X}_1^t \times \tilde{X}_2^t$. Then, E is also p-evident because $E' \subseteq B_t^{p,\mathcal{S}'}(\theta)$. Conversely, suppose there exists E that satisfies condition (3). Since $E \subseteq B_t^{p,\mathcal{S}'}(\theta)$, it follows that $E \subseteq C_t^{p,\mathcal{S}'}(\theta)$. Hence, $\nu^t(C_t^{p,\mathcal{S}'}(\theta)) > 0$. □

The next result is similar to Theorem 1. It says that if $q_t^{\mathcal{S}}(\theta) < p$, then the inequality in Theorem 1 holds in a strict sense if we take p appropriately for the true state.

Proposition 5. *Let $\mathcal{S}'$ and $\mathcal{S}$ be the signal structures in Definition 2. Fix $\theta \in \Theta$. Then there exists $p* \in (0,1)$ such that for all $p \in (0, p^*]$, $V_p(t) > 0$ holds if $q_t^{\mathcal{S}}(\theta) < p$.*

3.4 Conditions under which common p-belief of the true state is strictly more likely under no information sharing

Next, focus on the no-sharing phase and the irrelevant phase. The following is a condition that $V_p(t) < 0$ holds strictly.

Proposition 6. *Let $\mathcal{S}'$ and $\mathcal{S}$ be the signal structures in Definition 2. Fix the true state θ and suppose that $q_t^{\mathcal{S}}(\theta) \geq p$. Then, these conditions are equivalent.*

$$(1) : V_p(t) < 0$$
$$(2) : \text{There exists } x_1^t \in \{X_1'(\theta)\}^t \text{ such that } \nu^t(B_t^{p,\mathcal{S}'}(\theta)|x_1^t) \in (0,p)$$

Proof. Suppose $\mu^t(C_t^{p,\mathcal{S}}(\theta)|\theta) > \nu^t(C_t^{p,\mathcal{S}'}(\theta)|\theta)$. Prove by contradiction. Suppose that for all $x_1^t \in \{X_1'(\theta)\}^t, \nu^t(B_t^{p,\mathcal{S}'}(\theta)|x_1^t) \in \{0\} \cup [p,1]$. Let $E = \Theta \times \tilde{X_{1t}} \times \mathcal{B}_{t2}^{p,\mathcal{S}'}(\theta)$, where $\tilde{X_{1t}} = \{x_1^t \in \{X_1'(\theta)\}^t | \nu^t(B_t^{p,\mathcal{S}'}(\theta)|x_1^t) \geq p\}$. If $\tilde{X_{1t}} = \emptyset$, it follows that

$$\nu^t(B_t^{p,\mathcal{S}'}(\theta)|\theta) = 0.$$

Since $\nu^t(B_t^{p,\mathcal{S}'}(\theta)|\theta) = \nu_2^{t,\theta}(\mathcal{B}_{t2}^{p,\mathcal{S}'}(\theta)) = \mu_2^{t,\theta}(\mathcal{B}_{t2}^{p,\mathcal{S}}(\theta)) = q_t^{\mathcal{S}}(\theta)$, this contradicts the assumption that $q_t^{\mathcal{S}}(\theta) \geq p$. Hence, E is a nonempty set. Then $E \subseteq B_t^{p,\mathcal{S}'}(\theta)$ holds, and E is p-evident because

$$\begin{aligned}\nu^t(E|x_2^t) =& \nu^t(B_t^{p,\mathcal{S}'}(\theta)|x_2^t) - \nu^t(\Theta \times \{X_1'(\theta)\}^t \backslash \tilde{X_{1t}} \times \mathcal{B}_{t2}^{p,\mathcal{S}'}(\theta)|x_2^t)\\ =& \nu^t(B_t^{p,\mathcal{S}'}(\theta)|x_2^t)\\ \geq& p\end{aligned}$$

for all $x_2 \in \mathcal{B}_{t2}^{p,\mathcal{S}'}(\theta)$. The second equality holds because $\nu^{t,\theta}(x_1^t, x_2^t) = 0$ for all $x_1^t \in \{X_1'(\theta)\}^t \backslash \tilde{X_{1t}}, x_2 \in \mathcal{B}_{t2}^{p,\mathcal{S}'}(\theta)$, and $\theta \in \Theta$ by the definition of $\tilde{X_{t1}}$. Then it follows that

$$\nu^t(C_t^{p,\mathcal{S}'}(\theta)|\theta) \geq \nu^t(E|\theta) = q_t^{\mathcal{S}}(\theta) = \mu^t(C_t^{p,\mathcal{S}}(\theta)|\theta).$$

Using Theorem 1, we have

$$\nu^t(C_t^{p,\mathcal{S}'}(\theta)|\theta) = \mu^t(C_t^{p,\mathcal{S}}(\theta)|\theta).$$

Conversely, suppose $\mu^t(C_t^{p,\mathcal{S}}(\theta)|\theta) = \nu^t(C_t^{p,\mathcal{S}'}(\theta)|\theta)$. Then, $\nu^t(C_t^{p,\mathcal{S}'}(\theta)|\theta) = \mu_2^{t,\theta}(\mathcal{B}_{t2}^{p,\mathcal{S}}(\theta)) = \nu_2^{t,\theta}(\mathcal{B}_{t2}^{p,\mathcal{S}'}(\theta))$. Since $B_t^{p,\mathcal{S}'}(\theta) \subseteq C_t^{p,\mathcal{S}'}(\theta)$, it follows that

$$(\{\theta\}, x_1^t, x_2^t) \in C_t^{p,\mathcal{S}'}(\theta)$$

for all $x_2^t \in \mathcal{B}_{t2}^{p,\mathcal{S}'}(\theta)$ and $\nu^{t,\theta}(x_1^t, x_2^t) > 0$. Hence, for all $x_1^t \in \{X_1'(\theta)\}^t$, $\nu^t(B_t^{p,\mathcal{S}'}(\theta)|x_1^t) \in \{0\} \cup [p,1]$. □

CEMS (2008) showed that if X_1 and X_2 are both finite, agents commonly learn θ. This means that

$$\lim_{t\to\infty} \mu^t(C_t^{p,\mathcal{S}}(\theta)|\theta) = \lim_{t\to\infty} \nu^t(C_t^{p,\mathcal{S}'}(\theta)|\theta).$$

In fact, in Examples 2, 3, and 4, for sufficiently large t, we have $\mu^t(C_t^{p,\mathcal{S}}(\beta)|\beta) = \nu^t(C_t^{p,\mathcal{S}'}(\beta)|\beta)$. So it may seem that for any signal structure, if t is large enough, then $V_p(t) = 0$ will hold, but this is actually incorrect. By the law of large numbers, the probability that agent 2 p-believes the true state approaches 1 when $t \to \infty$. However, no matter how large t is, if there exists some "bad" signal for agent 1 such that, when he receives it, he thinks agent 2 has a low probability of p-believing the true state, then condition (2) of Proposition 6 is satisfied. Hence, $\mu^t(C_t^{p,\mathcal{S}}(\theta)|\theta) > \nu^t(C_t^{p,\mathcal{S}'}(\theta)|\theta)$ may hold for any large enough t. Example 5 below presents a case in which the irrelevant phase does not exist. In this example, $V_p(t) < 0$ holds for large enough t.

Example 5.

Fix $p \in \left(\frac{1}{2}, 1\right)$. Let $\Theta = \{\alpha, \beta\}$. Suppose that $p_0(\alpha) = p_0(\beta) = \frac{1}{2}$. Let $\mathcal{S}' = \langle X, (\nu_\theta)_{\theta\in\Theta}\rangle$ be the signal structure which satisfies $X_1'(\theta) = \{\theta\} \times \{0, 1, 2\}$ for each θ, $X_2 = \{0, 1\}$ and

$$\begin{aligned}
&\nu^\alpha((\alpha,0),0) = a\\
&\nu^\alpha((\alpha,1),0) = (1-a)\varepsilon\\
&\nu^\alpha((\alpha,1),1) = (1-a)(1-\varepsilon)\varepsilon\\
&\nu^\alpha((\alpha,2),1) = (1-a)(1-\varepsilon)^2\\
&\nu^\beta((\beta,0),0) = b\\
&\nu^\beta((\beta,1),0) = (1-b)\varepsilon\\
&\nu^\beta((\beta,1),1) = (1-b)(1-\varepsilon)\varepsilon\\
&\nu^\beta((\beta,2),1) = (1-b)(1-\varepsilon)^2,
\end{aligned}$$

where $\varepsilon \in (0, 2 - \frac{1}{p})$ and $0 < a < b < 1$.
Let $\mathcal{S} = \langle X, (\mu^\theta)_{\theta\in\Theta}\rangle$ be the signal structure which is no-sharing and induced by $\mathcal{S}'$.

$\mathcal{S}'$ corresponds to the situation of a slightly different version of the email game. That is, agent 1 knows the true state $\theta \in \{\alpha, \beta\}$. If the state is α, agent 1 sends an email to agent 2 with probability $1 - a$. If agent 2 receives the email, she replies to agent 1, but each email is lost with a small probability ε. Similarly, if the state is β, agent 1 sends an email to agent 2 with probability $1 - b$. If agent 2 receives the email, she replies to agent 1, but each email is lost with a small probability ε. Since $1 - b < 1 - a$, if agent 2 receives an email, she will assign a higher probability that the state is α.

Precisely,

$$\begin{aligned}
&\nu^t(\{\beta\} \times X_1^t \times X_2^t | x_2^t)\\
&\quad= \frac{(1-b)^m(1-\varepsilon)^m[b+(1-b)\varepsilon]^{t-m}}{(1-a)^m(1-\varepsilon)^m[a+(1-a)\varepsilon]^{t-m} + (1-b)^m(1-\varepsilon)^m[b+(1-b)\varepsilon]^{t-m}}\\
&\quad= \frac{(1-b)^m[b+(1-b)\varepsilon]^{t-m}}{(1-a)^m[a+(1-a)\varepsilon]^{t-m} + (1-b)^m[b+(1-b)\varepsilon]^{t-m}},
\end{aligned}$$

where m is the number of times that agent 2 receives the email, i.e., $m = |\{k | x^t_{2k} = 1\}|$. Hence, $x^t_2 \in \mathcal{B}^{p,\mathcal{S}'}_{t2}(\beta)$ is equivalent to

$$\frac{1}{p} - 1 \geq \left(\frac{1-a}{1-b}\right)^m [\frac{a+(1-a)\varepsilon}{b+(1-b)\varepsilon}]^{t-m}.$$

Let $f_t(m)$ be the value of the right-hand side of this inequality. Since $\frac{1-a}{1-b} > 1$ and $\frac{a+(1-a)\varepsilon}{b+(1-b)\varepsilon} < 1$ hold, there exists a large enough number T such that

$$\frac{1}{p} - 1 \geq f_t(0)$$
$$\frac{1}{p} - 1 < f_t(t)$$

for all $t \geq T$. Note that $f_t(m)$ is increasing with respect to m. Hence, for all $t \geq T$, there exists $n(t) \in \mathbb{Z} \cup [0, t-1]$ such that

$$\frac{1}{p} - 1 \geq f_t(m) \qquad \textit{for all } m \leq n(t)$$
$$\frac{1}{p} - 1 < f_t(m) \qquad \textit{for all } m \geq n(t)+1.$$

Now, consider agent 1's signal sequence $x^t_1 = ((\beta,2),(\beta,2),...,(\beta,2),(\beta,1),(\beta,1),...,(\beta,1))$ where $(\beta,2)$ appears $n(t)$ times. In this case, it follows that

$$0 < \nu^t(B^{p,\mathcal{S}'}_t(\beta)|x^t_1) = \left(\frac{1}{2-\varepsilon}\right)^{t-n(t)} \leq \frac{1}{2-\varepsilon} < p.$$

Therefore, by Proposition 6 and almost the same argument as for state B, for large enough t,

$$\mu^t(C^{p,\mathcal{S}}_t(\theta)|\theta) > \nu^t(C^{p,\mathcal{S}'}_t(\theta)|\theta)$$

holds for each $\theta \in \Theta$.

The next interesting question is that of when the irrelevant phase exists. In this regard, the following is an almost necessary and sufficient condition for $V_p(t) = 0$ to hold for large enough t in the case of finite signals.

Theorem 2. [7] *Let $\mathcal{S}'$ and $\mathcal{S}$ be the signal structures in Definition 2. Suppose $|\Theta| \geq 2$ and $|X'_1|, |X'_2| < \infty$. Fix true state θ. Let $X'_2(\theta) = \{x_2 \in X_2 | \nu^\theta_2(x_2) > 0\} = \{1, 2, ..., l\}$.*
For each $x_1 \in X'_1(\theta)$ and $s \in X'_2(\theta)$, define

$$p^{x_1}_s = \frac{\nu^\theta(x_1, s)}{\sum_{x_2 \in X'_2(\theta)} \nu^\theta(x_1, x_2)}.$$

Let $P = Conv(\{(p^{x_1}_1, p^{x_1}_2, ..., p^{x_1}_l)\}_{x_1 \in X'_1(\theta)})$. Then, the irrelevant phase exists if

$$\prod_{s=1}^{l} \left(\frac{\nu^{\theta'}_2(s)}{\nu^\theta_2(s)}\right)^{p_s} < 1$$

[7] In this proposition, define $0^0 = 1$.

for all $\theta' \in \Theta \backslash \{\theta\}$ *and* $(p_1, p_2, ..., p_l) \in P$*. In addition, the irrelevant phase does not exist if there exists* $\theta' \in \Theta \backslash \{\theta\}$ *and* $(p_1, p_2, ..., p_l) \in P \cap (\mathbb{R}_{>0})^l$ *such that*

$$\prod_{s=1}^{l} \left(\frac{\nu_2^{\theta'}(s)}{\nu_2^{\theta}(s)} \right)^{p_s} > 1$$

In this proposition, P is the set of the possible vectors of the proportion of each signal that is received by agent 2 from the perspective of agent 1 when the true state is θ. From Proposition 6, for an irrelevant phase to exist, whatever signal is received by agent 1, there must be a high probability that agent 2 p-believes θ if they receive signals a sufficiently large number of times. This represents the condition in the first half of the proposition. Conversely, when the latter condition is satisfied for some θ' and $(p_1, p_2, ..., p_l)$, there exists a situation where agent 1 receives a signal that makes the agent believe that the proportion of signals that agent 2 is receiving is almost $(p_1, p_2, ..., p_l)$ respectively. In this case, agent 2 is more likely to believe that it is θ' rather than θ, and thus the irrelevant phase disappears by Proposition 6.

If $\mathcal{S}'$ itself is a no-sharing signal structure, it follows that $P = \{\nu_2^{\theta}(1), \nu_2^{\theta}(2), ..., \nu_2^{\theta}(l)\}$. Since the Kullback-Leibler divergence of ν_2^{θ} from $\nu_2^{\theta'}$ is positive (because $\nu_2^{\theta} \neq \nu_2^{\theta'}$), we have

$$\sum_{s=1}^{l} \nu_2^{\theta}(s) log \left(\frac{\nu_2^{\theta}(s)}{\nu_2^{\theta'}(s)} \right) > 0.$$

Thus,

$$\sum_{s=1}^{l} \left(\frac{\nu_2^{\theta'}(s)}{\nu_2^{\theta}(s)} \right)^{\nu_2^{\theta}(s)} < 1.$$

Hence, it can be confirmed that there is indeed an irrelevant phase.

In Example 4, it can be checked that

$$\max \prod_{s \in X_2} \left(\frac{\nu_2^{\alpha}(s)}{\nu_2^{\beta}(s)} \right)^{p_s} = 0.$$

Hence, the irrelevant phase exists. However, in Example 5, it follows that for small enough $\delta > 0$, there exists $p \in P \cap (\mathbb{R}_{>0})^2$ such that

$$\prod_{s \in X_2} \left(\frac{\nu_2^{\alpha}(s)}{\nu_2^{\beta}(s)} \right)^{p_s} > \frac{1-a}{1-b} - \delta.$$

Since $\frac{1-a}{1-b} > 1$, we can see that there is no irrelevant phase.

Using this theorem, the following corollary can be shown immediately. This says that there is an irrelevant phase if there is a possibility that agent 2 receives a conclusive signal no matter what signal agent 1 receives.

Corollary 2. *Let* $\mathcal{S}'$ *and* $\mathcal{S}$ *be the signal structures in Definition 2. Suppose* $|X_1'| < \infty$ *and* $|X_2'| < \infty$*. Fix* θ*, and let* $E_2(\theta) = \{x_2 \in X_2 | \nu(\{\theta\} \times X_1 \times X_2 | x_2) = 1\}$*. Suppose that*

$$\inf_{x_1 \in \{X_1'(\theta)\}^t} \nu(\Theta \times X_1' \times E_2(\theta) | x_1) > 0$$

Then for all $p \in (0, 1)$*, the irrelevant phase exists.*

Corollary 2 is valid even if the signal is not finite (see Appendix for details).

3.5 Comparative Statistics

Here, we will look at the effect of signal accuracy and the value of p on the length of each phase. Let t^*,t^{**},$\bar{t}$ be the periods at the boundary of each phase as in Corollary 1. Recall that the sharing phase is the range of t where $1 \leq t < t^*$, the no-sharing phase is where $\bar{t} > t \geq t^{**}$, and the irrelevant phase is where $t \geq \bar{t}$.

Proposition 7. *Let $\mathcal{S}' = \langle X', (\nu^\theta)_{\theta\in\Theta}\rangle$ be a signal structure. Fix the true state θ. Then, t^*, which is the length of the sharing phase, and $\bar{t}$ are non-decreasing with respect to p.*

Proof. Since t^* is the largest integer such that $q_t^{\mathcal{S}}(\theta) < p$ for all $t < t^*$, the larger the value of p, the looser this condition becomes, and thus the larger the value of t^*. In the no-sharing phase and the irrelevant phase, Proposition 6 yields that $V_p(t) < 0$ implies $VC_{p'}(t) < 0$ for all $p' > p$. Hence, the starting period of the irrelevant phase $\bar{t}$, if it exists, is non-decreasing with respect to p. □

Fix p and θ. For two signal structures $\mathcal{S}$ and $\mathcal{S}'$, if $q_t^{\mathcal{S}}(\theta) \geq q_t^{\mathcal{S}'}(\theta')$ holds for all t, we say that $\mathcal{S}$ is a weakly more accurate signal structure for agent 2 than $\mathcal{S}'$. Intuitively, a more accurate signal structure means it is easier to p-believe the true state. From the definition of the sharing phase, if $\mathcal{S}$ is weakly more accurate for agent 2 than $\mathcal{S}'$, then the length of the sharing phase in $\mathcal{S}$ is equivalent to or shorter than the length in $\mathcal{S}'$.

Examples 2, 3, and 4, which include the coordinated attack example and the email game example, were examples in which agent 2 p-believes the true state only when she knows the true state. In such a signal structure, there is a relationship between the lower bound on the probability that agent 1 believes that agent 2 knows the state and $\bar{t} - t^{**}$, which is the length of the no-sharing phase. Fix $p \in (0, 1)$, $\theta \in \Theta$, and $q \in (0, 1]$. Define $\mathscr{S}(p, \theta, q)$ to be the class of signal structures such that the above conditions are satisfied and the probability that an agent 2 receives a conclusive signal for θ is q. Precisely, $\mathcal{S} \in \mathscr{S}(p, \theta, q)$ if $\mathcal{S} = \langle X, (\mu^\theta)_{\theta\in\Theta}\rangle$ is a signal structure,

$$\mu^\theta(X_1 \times E_2(\theta)) = q,$$

and

$$X_2^t \backslash \mathcal{B}_{t2}^{p,\mathcal{S}}(\theta) = \{X_2 \backslash E_2(\theta)\}^t$$

for each t, where $E_2(\theta) = \{x_2 \in X_2 | \mu(\{\theta\} \times X_1^t \times X_2^t | x_2) = 1\}$. For each $\mathcal{S} = \langle X, (\mu^\theta)_{\theta\in\Theta}\rangle \in \mathscr{S}(p, \theta, q)$, let $m(\mathcal{S})$ be

$$m(\mathcal{S}) = \inf_{x_1 \in X_1(\theta)} \mu(\Theta \times X_1 \times E_2(\theta) | x_1)$$

Proposition 8. *Take any p, θ, and q. In $\mathscr{S}(p, \theta, q)$, $\bar{t} - t^{**}$, which is the length of the no-sharing phase, is non-increasing with respect to $m(\mathcal{S})$ where $m(\mathcal{S}) > 0$. In particular, $\bar{t} - t^* \to \infty$ as $m(\mathcal{S}) \to 0+$ and $\bar{t} - t^* = 0$ if $m(\mathcal{S}) \geq p$.*

Proof. Take any $\mathcal{S} = \langle X, (\mu^\theta)_{\theta\in\Theta}\rangle \in \mathscr{S}(p, \theta, q)$. Then we have

$$q_t^{\mathcal{S}}(\theta) = \mu_2^{t,\theta}(\mathcal{B}_{t2}^{p,\mathcal{S}}(\theta)) = 1 - (1 - q)^t.$$

Hence, $q_t^{\mathcal{S}}$ does not depend on $\mathcal{S}$. Thus, Corollary 1 implies that t^* and t^{**} do not depend on $\mathcal{S}$. Proposition 6 yields that $V_p(t) = 0$ if and only if $1 - (1 - m(\mathcal{S}))^t \geq p$ when $m(\mathcal{S}) > 0$ at $t \geq t^{**}$. This condition is equivalent to $t \geq \frac{log(1-p)}{log(1-m(\mathcal{S}))}$, which is decreasing with respect to $m(\mathcal{S})$. □

Recall the example of a coordinated attack in the introduction. In this example, $m(\mathcal{S}) = \frac{(1-\varepsilon)\delta}{\varepsilon+(1-\varepsilon)\delta}$ holds. Note that this $m(\mathcal{S})$ is increasing with respect to δ. Since changing only δ does not affect the distribution of signals received by general 2, it follows from Proposition 8 that $\bar{t} - t^*$ is a weakly decreasing function of δ.

4 Discussion

4.1 Gaussian Signal

In this section, signal space X_i is assumed to be uncountable. Note that our main theorem (Theorem 1) and other results hold even in this case.
Let $\Theta \subset \mathbb{R}$, $2 \leq |\Theta| < \infty$, and $\mathcal{S} = \langle X, (\mu^\theta)_{\theta\in\Theta}\rangle$ be the signal structure which satisfies $X_2 = \mathbb{R}$, and $\mu_2^{t,\theta}$ follows normal distribution $N(\theta, \sigma^2)$, i.e.,

$$\mu_2^\theta(x_2) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(x_2-\theta)^2}{2\sigma^2}\right).$$

This represents the situation where agent 2 receives a signal with noise from a normal distribution for the true state.

Proposition 9. *Assume the above signal structure. Let* $m(\theta) = \min_{\theta'\in\Theta} \frac{p_0(\theta')}{p_0(\theta)}$. *Then, when the true state is* θ*, if* p *satisfies* $p \geq \frac{1}{1+m(\theta)}$*, the irregular phase disappears.*

Theorem 1 says that when t is small enough, the probability that the true state is common p-believed is higher with information sharing, and when t is large enough, the probability is higher without information sharing. In general, however, there may be an "irregular phase" in which we do not know which case is preferable between each phase. However, considering such a signal with Gaussian noise, it can be shown that the irregular phase disappears when p is sufficiently large.

Here is a sketch of the proof of Proposition 9. Since the signal is Gaussian, when multiple signals are received, only the average of the signals is an important value. The following Lemma 3 implies that the candidate state closest to the average of the signals received must be the true state in order for the true state to be common p-believed under the conditions of Proposition 9.

Lemma 3. *Assume the condition of Proposition 9. If agent 2 p-believes* θ *at* t, $|\theta - \overline{x_2^t}| \leq |\theta' - \overline{x_2^t}|$ *holds for all* $\theta' \in \Theta$*, where* $\overline{x_2^t} = \frac{1}{t}\sum_{k=1}^t x_{2k}^t$.

Using this Lemma 3, it can be shown that the set of signals for which agent 2 p-believes θ increases with respect to t.

Lemma 4. *Assume the condition of Proposition 9. If agent 2 p-believes* θ *at* t *when she receives the signal* x_2^t*, agent 2 also p-believes* θ *at* $t+1$ *when she receives the signal* z_2^{t+1} *that satisfies* $\overline{x_2^t} = \overline{z_2^{t+1}}$.

For the proof of Proposition 9, it is sufficient to show that the probability that agent 2 p-believes the true state is increasing with respect to t. However, Lemma 4 shows that it is monotonically increasing with respect to t in the sense of set inclusion, but it does not say this in the sense of probability. So we use the following Lemma 5 on the properties of signal sets such that agent 2 p-believes the true state.

Lemma 5. *Assume the condition of Proposition 9. Then, there exists interval $L(t) \subset \mathbb{R}$ (can be empty) such that agent 2 p-believes θ at t when she receives the signal x_2^t if and only if $\overline{x_2^t} \in L(t)$.*

From Lemma 4 and Lemma 5, it follows that $\theta \in L(t) \subset L(t+1)$ for all t that satisfies $q_t^{\mathcal{S}}(\theta) \geq p$ because $p \geq \frac{1}{1+m(\theta)} \geq \frac{1}{2}$. Since the distribution of the mean of the signals received by agent 2 is $N(\theta, \frac{\sigma^2}{t})$ and $N(\theta, \frac{\sigma^2}{t+1})$ when the number of signals she receives is t, $t+1$ respectively, the probability that agent 2 p-believes the true state is increasing with respect to t. Therefore, Proposition 9 holds.

4.2 Limitations of the Model

In this paper, the assumption that agent 1 knows the true state is important. We can construct a counterexample to Theorem 1 if we assume that both agents may not know the true state.

Example 6.

Let $\Theta = \{\alpha, \beta\}$ and $\mathcal{S}' = \langle X, (\nu_\theta)_{\theta\in\Theta}\rangle$, where $X_1 = \{0,1\}$, $X_2 = \{0,1\}$, and

$$\nu^\alpha(0,0) = 1$$
$$\nu^\beta(0,0) = \varepsilon$$
$$\nu^\beta(1,1) = 1-\varepsilon$$

where $\varepsilon \in (0,1)$. Let $\mathcal{S} = \langle X, (\mu^\theta)_{\theta\in\Theta}\rangle$ be the signal structure which is no-sharing and induced by $\mathcal{S}'$. Assume that $\frac{p_0(\beta)\varepsilon}{p_0(\alpha)+p_0(\beta)\varepsilon} < p$.
Then we have

$$C_t^{p,\mathcal{S}}(\beta) = \begin{cases} \emptyset & \textit{if } p > 1-\varepsilon^t \\ \Theta \times X_1^t \backslash \{(0,0,...,0)\} \times X_2^t \backslash \{(0,0,...,0)\} & \textit{if } p \leq 1-\varepsilon^t \end{cases}$$
$$C_t^{p,\mathcal{S}'}(\beta) = X_1^t \backslash \{(0,0,...,0)\} \times X_2^t \backslash \{(0,0,...,0)\}$$

Hence,

$$\mu^t(C_t^{p,\mathcal{S}}(\beta)|\beta) = \begin{cases} 0 & \textit{if } p > 1-\varepsilon^t \\ (1-\varepsilon^t)^2 & \textit{if } p \leq 1-\varepsilon^t \end{cases}$$
$$\nu^t(C_t^{p,\mathcal{S}'}(\beta)|\beta) = 1-\varepsilon^t.$$

Figure 6 shows this result. In this figure, t^* is the smallest integer t which satisfies $p \leq 1-\varepsilon^t$. In particular , for all p, t,

$$\nu^t(C_t^{p,\mathcal{S}'}(\beta)|\beta) > \mu^t(C_t^{p,\mathcal{S}}(\beta)|\beta),$$

and this is contrary to the results of Theorem 1.

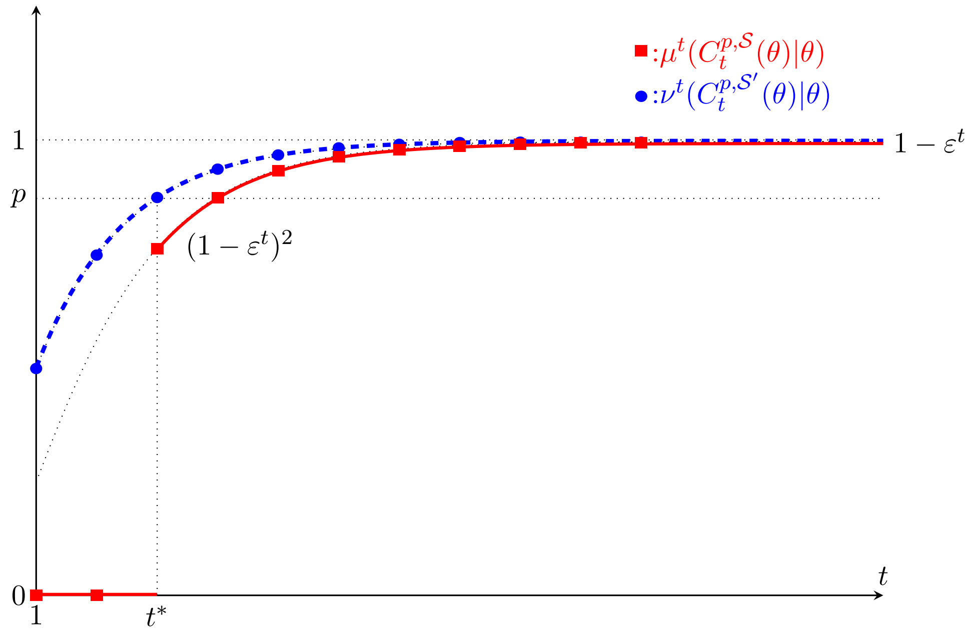

Figure 6: Probability of common p-belief in Example 6

References

[1] Awaya, Y. and V. Krishna (2022): "Commonality of Information and Commonality of Beliefs," Available at SSRN: `https://ssrn.com/abstract=4114760`.

[2] Blackwell, David (1951): "Comparison of Experiments," in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, Volume 1, University of California Press, 93-102.

[3] Cripps, M., J. Ely, G. Mailath and L. Samuelson (2008): "Common Learning," *Econometrica*, 76, 909-933.

[4] Cripps, M., J. Ely, G. Mailath and L. Samuelson (2013): "Common Learning with Intertemporal Dependence," *International Journal of Game Theory*, 42, 55-98.

[5] Frick, M., R. Iijima and Y. Ishii (2022): "Learning Efficiency of Multi-agent Information Structures," CEPR Discussion Paper No. DP16877, Available at SSRN: `https://ssrn.com/abstract=4026761`.

[6] Monderer, D. and D. Samet (1989): "Approximating Common Knowledge with Common Beliefs,". *Games and Economic Behavior*, 1, 170-190.

[7] Rubinstein, A. (1989): "The Electronic Mail Game: Strategic Behavior under Almost Common Knowledge," *American Economic Review*, 79, 385-291.

[8] Steiner, J. and C. Stewart (2011): "Communication, Timing and Common Learning," *Journal of Economic Theory*, 146, 230-247.

Appendix: Omitted Proofs

A: Proof of Proposition 1

Show not only for the two cases (case 1 and case 2) but also for the general signal structure $\mathcal{S} = \langle X, (\mu^\theta)_{\theta\in\Theta}\rangle$ treated in the text (i.e., agent 1 understands the state if it receives a signal even once).

First, show that the strategy whereby each agent chooses B if and only if $x_i^t \in C_{ti}^{p,\mathcal{S}}(\beta)$ is a BNE strategy. Suppose both agents follow this strategy. If $x_i^t \in C_{ti}^{p,\mathcal{S}}(\beta)$, agent i assigns at least probability p to the event that $C_t^{p,\mathcal{S}}(\beta)$ because $C_t^{p,\mathcal{S}}(\beta)$ is p-evident. Hence, agent i believes that the opponent chooses A with a probability of at least p. For agent 1, $x_1^t \in C_{t1}^{p,\mathcal{S}}(\beta)$ means that the true state is β, so the expected payoff when agent 1 chooses N is 0 and when he chooses A, at least $pM + (1-p)(-L) = 0$. For agent 2, since the state is β if agent 1 chooses A, the expected payoff when agent 2 chooses N is 0 and when she chooses A, at least $pM + (1-p)(-L) = 0$. Therefore, one of best replies for agent i is to chooses A if $x_i^t \in C_{i1}^{p,\mathcal{S}}(\beta)$ for each i.

Next, show that if $x_i^t \notin C_{ti}^{p,\mathcal{S}}(\beta)$, agent i must take N in any BNE. Note that agent 1 chooses N in a BNE if the true state is α since he can identify the state. If $x_1^t \notin C_{t1}^{p,\mathcal{S}}(\beta)$, it means that there exists the smallest non-negative integer n such that

$$x_1^t \notin \mathcal{B}_{t1}^{p,\mathcal{S}}(\{B_t^{p,\mathcal{S}}\}^n(\beta)).$$

Since agent 1 knows the true state, this n must be an odd number or 0. Similarly, if $x_2^t \notin C_{t2}^{p,\mathcal{S}}(\beta)$, it means that there exists the smallest non-negative even number m such that

$$x_2^t \notin \mathcal{B}_{t2}^{p,\mathcal{S}}(\{B_t^{p,\mathcal{S}}\}^m(\beta)).$$

If $n = 0$, the state is α, so agent 1 chooses A. If $m = 0$, the unique best reply for agent 2 is choosing N because the expected payoff of choosing A is less than or equal to $pM + (1-p)L = 0$. Then, if $n = 1$, the unique best reply for agent 1 is choosing N because he thinks his opponent chooses A with a probability less than p. In the same way, it can be shown by induction that when $m = 2, 4, 6, ...$, the only optimal action for agent 2 is to choose N, and when $n = 3, 5, 7, ...$, the only optimal action for agent 1 is to choose N. Therefore, the strategy whereby each agent chooses A if and only if $x_i^t \in C_{ti}^{p,\mathcal{S}}(\beta)$ is a BNE strategy and under any BNE strategy, agent i chooses A if $x_i^t \notin C_{ti}^{p,\mathcal{S}}(\beta)$. Thus, the probability that both agents simultaneously choose A under some BNE is at most the same as the probability that they common p-believe that the state is β.

B: Proof of Proposition 2

First, consider $B_t^{p,\mathcal{S}}(\theta)$, which is the event that both agents assign at least probability p to state θ. Since $(X_1(\theta))_{\theta\in\Theta}$ are disjoint, it follows that

$$\mathcal{B}_{t1}^{p,\mathcal{S}}(\theta) = \{X_1(\theta)\}^t.$$

Hence, we have

$$B_t^{p,\mathcal{S}}(\theta) = \Theta \times \{X_1(\theta)\}^t \times \mathcal{B}_{t2}^{p,\mathcal{S}}(\theta).$$

Now, consider $B_t^{p,\mathcal{S}}(B_t^{p,\mathcal{S}}(\theta))$, which corresponds to the second order belief. Note that

$$B_{t2}^{p,\mathcal{S}}(B_t^{p,\mathcal{S}}(\theta)) = \Theta \times X_1^t \times \mathcal{B}_{t2}^{p,\mathcal{S}}(\theta),$$

because $\mu^t(\Theta \times \{X_1(\theta)\}^t \times \mathcal{B}_{t2}^{p,\mathcal{S}}(\theta)|x_2^t) = \mu^t(\{\theta\} \times X_1^t \times X_2^t|x_2^t) \geq p$ for all $x_2^t \in \mathcal{B}_{t2}^{p,\mathcal{S}}(\theta)$.

Also, since $\mathcal{S}$ is no-sharing, it follows that $\mu^t(\Theta \times \{X_1(\theta)\}^t \times \mathcal{B}_{t2}^{p,\mathcal{S}}(\theta)|x_1^t) = \mu^t(\Theta \times X_1 \times \mathcal{B}_{t2}^{p,\mathcal{S}}(\theta)|\theta)$ for $x_1^t \in \{X_1(\theta)\}^t$. Therefore,

$$B_{t1}^{p,\mathcal{S}}(B_t^{p,\mathcal{S}}(\theta)) = \begin{cases} \emptyset & \textit{if } \mu^t(\Theta \times \{X_1(\theta)\}^t \times \mathcal{B}_{t2}^{p,\mathcal{S}}(\theta)|\theta) < p \\ \Theta \times \{X_1(\theta)\}^t \times X_2^t & \textit{if } \mu^t(\Theta \times M_\theta^t \times \mathcal{B}_{t2}^{p,\mathcal{S}}(\theta)|\theta) \geq p \end{cases}$$

From the above results, we have

$$B_t^{p,\mathcal{S}}(B_t^{p,\mathcal{S}}(\theta)) = \begin{cases} \emptyset & \textit{if } \mu^t(\Theta \times \{X_1(\theta)\}^t \times \mathcal{B}_{t2}^{p,\mathcal{S}}(\theta)|\theta) < p \\ \Theta \times \{X_1(\theta)\}^t \times \mathcal{B}_{t2}^{p,\mathcal{S}}(\theta) & \textit{if } \mu^t(\Theta \times M_\theta^t \times \mathcal{B}_{t2}^{p,\mathcal{S}}(\theta)|\theta) \geq p \end{cases}$$

Since $\mu^t(\Theta \times \{X_1(\theta)\}^t \times \mathcal{B}_{t2}^{p,\mathcal{S}}(\theta)|\theta) = \mu_2^{t,\theta}(\mathcal{B}_{t2}^{p,\mathcal{S}}(\theta)) = q_t^{\mathcal{S}}(\theta)$, it follows that

$$B_t^{p,\mathcal{S}}(B_t^{p,\mathcal{S}}(\theta)) = \begin{cases} \emptyset & \textit{if } q_t^{\mathcal{S}}(\theta) < p \\ \Theta \times \{X_1(\theta)\}^t \times \mathcal{B}_{t2}^{p,\mathcal{S}}(\theta) & \textit{if } q_t^{\mathcal{S}}(\theta) \geq p \end{cases}$$

Hence, if $\mu_2^{t,\theta}(\mathcal{B}_{t2}^{p,\mathcal{S}}(\theta)) \geq p$, it follows that $B_t^{p,\mathcal{S}}(B_t^{p,\mathcal{S}}(\theta)) = B_t^{p,\mathcal{S}}(\theta)$, so it is also equivalent to $C_t^{p,\mathcal{S}}(\theta)$. Otherwise, $C_t^{p,\mathcal{S}}(\theta) = \emptyset$ since $B_t^{p,\mathcal{S}}(\emptyset) = \emptyset$.

C: Proof of Lemma 1

First, note that

$$B_t^{p,\mathcal{S}}(\theta) = \Theta \times \{X_1(\theta)\}^t \times \mathcal{B}_{t2}^{p,\mathcal{S}}(\theta).$$

Since $C_t^{p,\mathcal{S}}(\theta) \subseteq B_t^{p,\mathcal{S}}(\theta)$, it follows that

$$C_t^{p,\mathcal{S}}(\theta) \subseteq \Theta \times \{X_1(\theta)\}^t \times \mathcal{B}_{t2}^{p,\mathcal{S}}(\theta).$$

Therefore, we have

$$\mu^t(C_t^{p,\mathcal{S}}(\theta)|\theta) \leq \mu^t(\Theta \times \{X_1(\theta)\}^t \times \mathcal{B}_{t2}^{p,\mathcal{S}}(\theta)) = \mu_2^{t,\theta}(\mathcal{B}_{t2}^{p,\mathcal{S}}(\theta)) = q_t^{\mathcal{S}}(\theta).$$

D: Proof of Theorem 1

From Proposition 2, it follows that

$$\mu^t(C_t^{p,\mathcal{S}'}(\theta)|\theta) \geq \mu^t(C_t^{p,\mathcal{S}}(\theta)|\theta) = 0 \qquad \textit{if } q_t^{\mathcal{S}}(\theta) < p.$$

Now, consider the case that $q_t^{\mathcal{S}}(\theta) \geq p$. By Lemma 1, we have

$$\mu^t(C_t^{p,\mathcal{S}'}(\theta)|\theta) \leq q_t^{\mathcal{S}}(\theta).$$

Note that $\mathcal{B}_{t2}^{p,\mathcal{S}'}(\theta) = \{x_2^t \in X_2^t | \nu^t(\{\theta\} \times X_1^t \times X_2^t | x_2^t) \geq p\}$. Then,

$$\begin{aligned}\mathcal{B}_{t2}^{p,\mathcal{S}'}(\theta) &= \{x_2^t \in X_2^t | \frac{p_0(\theta)\nu_2^{t,\theta}(x_2^t)}{\sum_{\theta' \in \Theta} p_0(\theta')\nu_2^{t,\theta'}(x_2^t)} \geq p\} \\ &= \{x_2^t \in X_2^t | \frac{p_0(\theta)\mu_2^{t,\theta}(x_2^t)}{\sum_{\theta' \in \Theta} p_0(\theta')\mu_2^{t,\theta'}(x_2^t)} \geq p\} \\ &= \mathcal{B}_{t2}^{p,\mathcal{S}}(\theta).\end{aligned}$$

Therefore, it follows that

$$\mu^t(C_t^{p,\mathcal{S}'}(\theta)|\theta) \leq \mu_2^{t,\theta}(\mathcal{B}_{t2}^{p,\mathcal{S}'}(\theta)) = q_t^{\mathcal{S}}(\theta) = \mu^t(C_t^{p,\mathcal{S}}(\theta)|\theta),$$

when $q_t^{\mathcal{S}}(\theta) \geq p$. The last equality follows from Proposition 2. This completes the proof of Theorem 1.

To prove corollary 1, fix $\theta \in \Theta$. Then, set t^* to be the smallest integer t which satisfies $q_t^{\mathcal{S}}(\theta) \geq p$. Note that $\lim_{t\to\infty} q_t^{\mathcal{S}}(\theta) = 1$, so t^* is well-defined. Again, since $\lim_{t\to\infty} q_t^{\mathcal{S}}(\theta) = 1$, there exists the smallest t^{**} such that $q_t^{\mathcal{S}}(\theta) \geq p$ for all $t \geq t^{**}$.

E: Proof of Proposition 3

First, consider $C_t^{p,\mathcal{S}}(\theta)$. If agent 2 receives a positive signal $x_2 > 0$ at least one time, she knows the state is β. On the other hand, if she receives $x_2^t = (0, 0, ..., 0)$, she assigns probability $\mu^t(\{\beta\} \times X_1^t \times X_2^t | x_2^t = (0, 0, ..., 0))$ to state β. Since

$$\mu^t(\{\beta\} \times X_1^t \times X_2^t | x_2^t = (0, 0, ..., 0)) = \frac{p_0(\beta)\varepsilon^t}{p_0(\alpha) + p_0(\beta)\varepsilon^t} < \frac{p_0(\beta)\varepsilon}{p_0(\alpha) + p_0(\beta)\varepsilon} < p,$$

state β is not p-believed by agent 2 in this case. Hence,

$$\mathcal{B}_{t2}^{p,\mathcal{S}}(\beta) = X_2^t \backslash \{(0, 0, ..., 0)\}.$$

Since $\mathcal{S}$ is no-sharing and $q_t^{\mathcal{S}}(\beta) = 1 - \varepsilon^t$, from Proposition 2, it follows that

$$C_t^{p,\mathcal{S}}(\beta) = \begin{cases} \emptyset & \textit{if } 1 - \varepsilon^t < p \\ \Theta \times \{X_1(\beta)\}^t \times X_2^t \backslash \{(0, 0, ..., 0)\} & \textit{if } 1 - \varepsilon^t \geq p \end{cases}$$

$$\mu^t(C_t^{p,\mathcal{S}}(\beta)|\beta) = \begin{cases} 0 & \textit{if } p > 1 - \varepsilon^t \\ 1 - \varepsilon^t & \textit{if } p \leq 1 - \varepsilon^t. \end{cases}$$

Next, consider $C_t^{p,\mathcal{S}'}(\beta)$. By the same argument,

$$B_t^{p,\mathcal{S}'}(\beta) = \Theta \times \{X_1'(\beta)\}^t \times X_2^t \backslash \{(0, 0, ..., 0)\}$$

holds. Again, if agent 2 receives a positive signal at least once, she knows that agent 1 receives $x_1^t \in \{X_1'(\beta)\}^t$. Now, if agent 1 receives at least one signal of $m_1 \geq 2$, agent 1 knows that agent 2 receives at least one email, i.e., he knows that $x_2^t \in$

$X_2^t\backslash\{(0,0,...,0)\}$. On the other hand, if he receives $x_1^t=((\beta,1),(\beta,1),...,(\beta,1))$, he assigns probability $\nu^t(\Theta\times X_1'^t\times X_2^t\backslash\{(0,0,...,0)\}|x_1^t=((\beta,1),(\beta,1),...,(\beta,1)))$ to the event that agent 2 receives $x_2^t\in X_2^t\backslash\{(0,0,...,0)\}$. Since

$$\nu^t(\Theta\times X_1^t\times X_2^t\backslash\{(0,0,...,0)\}|x_1^t=((\beta,1),(\beta,1),...,(\beta,1)))=1-\left(\frac{1}{2-\varepsilon^t}\right)^t,$$

it follows that

$$B_t^{p,\mathcal{S}'}(B_t^{p,\mathcal{S}'}(\beta))=\begin{cases}\Theta\times\{X_1'(\beta)\}^t\times X_2^t\backslash\{(0,0,...,0)\}=B_t^{p,\mathcal{S}}(\beta) & \\ \qquad\qquad\qquad\qquad \textit{if } 1-\left(\frac{1}{2-\varepsilon^t}\right)^t\geq p\\ \Theta\times\{X_1'(\beta)\}^t\backslash\{(1,1,...,1)\}\times X_2^t\backslash\{(0,0,...,0)\} & \\ \qquad\qquad\qquad\qquad \textit{if } 1-\left(\frac{1}{2-\varepsilon^t}\right)^t<p.\end{cases}$$

Hence, if $1-\left(\frac{1}{2-\varepsilon^t}\right)^t\geq p$, $C_t^{p,\mathcal{S}'}(\beta)=\Theta\times M_G^t\times X_2^t\backslash\{(0,0,...,0)\}$ holds.

Now, suppose that $1-\left(\frac{1}{2-\varepsilon^t}\right)^t<p$. By a similar argument as before, agent 1, who receives x_1^t, assigns at most probability $1-\left(\frac{1}{2-\varepsilon^t}\right)^t$ to the event that the maximum number of times agent 2 has sent an email is the same as him, i.e., $\max_k x_{2k}^t=\max_k m_{1k}^t$. In addition, agent 2, who receives x_2^t, assigns at most probability $1-\left(\frac{1}{2-\varepsilon^t}\right)^t$ to the event that the maximum number of times agent 1 has sent mail is one more than her, i.e., $\max_k m_{1k}^t=\max_k x_{2k}^t+1$. Therefore, for each $n=1,2,3,...$,

$$\{B_t^{p,\mathcal{S}'}\}^{2n-1}(\beta)=\Theta\times\{x_1^t\in X_1^t|\max_k m_{1k}^t\geq n\}\times\{x_2^t\in X_2^t|\max_k x_{2k}^t\geq n\}$$
$$\{B_t^{p,\mathcal{S}'}\}^{2n}(\beta)=\Theta\times\{x_1^t\in X_1^t|\max_k m_{1k}^t\geq n+1\}\times\{x_2^t\in X_2^t|\max_k x_{2k}^t\geq n\}.$$

Thus, $C_t^{p,\mathcal{S}'}(\beta)=\emptyset$ holds in this case.
Overall,

$$C_t^{p,\mathcal{S}'}(\beta)=\begin{cases}\emptyset & \textit{if } p>1-\left(\frac{1}{2-\varepsilon}\right)^t\\ \Theta\times\{X_1'(\beta)\}^t\times X_2^t\backslash\{(0,0,...,0)\} & \textit{if } p\leq 1-\left(\frac{1}{2-\varepsilon}\right)^t\end{cases}$$

$$\nu^t(C_t^{p,\mathcal{S}'}(\beta)|\beta)=\begin{cases}0 & \textit{if } p>1-\left(\frac{1}{2-\varepsilon}\right)^t\\ 1-\varepsilon^t & \textit{if } p\leq 1-\left(\frac{1}{2-\varepsilon}\right)^t\end{cases}$$

This completes the proof.

F: Proof of Proposition 5

Take $\overline{p}<1$ and let t^{**} be the integer which satisfies $q_t^{\mathcal{S}}(\theta)\geq p$ for all $t>t^{**}$. Show that there exist some p-evident events for all $t\leq t^{**}$ when p is small enough. For each t, let p_t be the supremum of p such that $\mathcal{B}_{t2}^{p,\mathcal{S}'}(\theta)$ is not an empty set. Since $p_t>0$, there exist $p'\in(0,\overline{p})$ such that $\mathcal{B}_{t2}^{p',\mathcal{S}'}(\theta)\neq\emptyset$ for all $t\leq t^{**}$. For each t, let $E_t = \{\theta\} \times \{x_1^t\} \times \{x_2^t\}$ where $x_2^t \in \mathcal{B}_{t2}^{p',\mathcal{S}'}(\theta)$ and x_1^t satisfies the condition that $\nu^{t,\theta}(x_1^t, x_2^t) > 0$. Then, it follows that

$$\nu^t(E_t|x_1^t) > 0$$
$$\nu^t(E_t|x_2^t) > 0.$$

Define $p_t^* = \min\{\nu^t(E_t|x_1^t), \nu^t(E_t|x_2^t)\}$, and $p^* = \min_{t \le t^{**}} p_t^*$. We have $p^* > 0$ and E_t is a p-evident event for all $t \le t^{**}$ and $p \in (0, p^*]$.
Therefore, by Proposition 4, it follows that $V_p(t) > 0$ if $q_t^{\mathcal{S}}(\theta) < p$ for all $p \in (0, p^*]$.

G: Proof of Theorem 2

Define $p^{x_1} = (p_1^{x_1}, p_2^{x_1}, ..., p_l^{x_1})$, $B_\varepsilon(p^{x_1}) = \{x \in \mathbb{R}^l | d(x, p^{x_1}) < \varepsilon\}$, and $P^\varepsilon = \{x \in \mathbb{R}^l | d(x, P) < \varepsilon\}$, where d is a usual Euclidean distance. Fix $\varepsilon \in (0, 1)$. First, show that there exists T such that $Prob\left(\left(\frac{n_1}{t}, \frac{n_2}{t}, ..., \frac{n_l}{t}\right) \in P^{2\varepsilon} | x_1^t\right) \ge p$ for all $t \ge T$ and $x_1^t \in \{X_1'(\theta)\}^t$, where n_s is the number of times agent 2 received signal $s \in X_2'(\theta)$.

From the law of large numbers, for each $x_1 \in X_1'(\theta)$, there exists T_{x_1} such that $Prob\left(\left(\frac{n_1}{t}, \frac{n_2}{t}, ..., \frac{n_l}{t}\right) \in B_\varepsilon(p^{x_1}) | x_1^t = (x_1, x_1, ..., x_1)\right) \ge p^{\frac{1}{|X_1'(\theta)|}}$ for all $t \ge T_{x_1}$. Define $T = \lceil \sqrt{l} \rceil \frac{|X_1'(\theta)|}{\varepsilon} \max_{x_1 \in X_1'(\theta)} T_{x_1}$. Show that this T satisfies the previous condition. Take any $t \ge T$ and $x_1^t \in X_1'(\theta)^t$. Let m_s be the number of times agent 1 received signal $s \in X_1'(\theta)$. Let $S = \{s \in X_1'(\theta) | m_s \ge \max_{x_1 \in X_1'(\theta)} T_{x_1}\}$. Note that this S is not empty because the definition of T implies $T \ge |X_1'(\theta)| \max_{x_1 \in X_1'(\theta)} T_{x_1}$. Define $t_S = \sum_{s \in S} m_s$ and consider the case where agent 1 receives only the signals that are in S t_S times, excluding those that are not in S from x_1^t (denote such a signal as $x_1^{t_S}$). Then, it follows that

$$Prob\left(\left(\frac{n_1}{t_S}, \frac{n_2}{t_S}, ..., \frac{n_l}{t_S}\right) \in P^\varepsilon | x_1^{t_S}\right) \ge p^{\frac{|S|}{|X_1'(\theta)|}} \ge p.$$

Since $t - t_S < |X_1'(\theta)| \max_{x_1 \in X_1'(\theta)} T_{x_1}$, we have $T > \frac{t - t_S}{\varepsilon}\sqrt{l}$, which yields

$$Prob\left(\left(\frac{n_1 + z_1}{t_S + (t - t_S)}, \frac{n_2 + z_2}{t_S + (t - t_S)}, ..., \frac{n_l + z_l}{t_S + (t - t_S)}\right) \in P^{2\varepsilon} | x_1^{t_S}\right) \ge p$$

for all $z_1, z_2, ..., z_l \in \mathbb{Z}_{\ge 0}$ such that $z_1 + z_2 + ... + z_l = t - t_S$, because

$$\begin{aligned}
& d(\left(\frac{n_1}{t_S}, \frac{n_2}{t_S}, ..., \frac{n_l}{t_S}\right), \left(\frac{n_1 + z_1}{t_S + (t - t_S)}, \frac{n_2 + z_2}{t_S + (t - t_S)}, ..., \frac{n_l + z_l}{t_S + (t - t_S)}\right)) \\
&= \sqrt{\sum_{k=1}^{l} \left(\frac{n_k}{t_s} - \frac{n_k + z_k}{t}\right)^2} \\
&\le \sqrt{\sum_{k=1}^{l} \max\{\left(\frac{n_k}{t_S} - \frac{n_k}{t}\right)^2, \left(\frac{n_k}{t_S} - \frac{n_k + t - t_S}{t}\right)^2\}} \\
&\le \sqrt{l}\left(1 - \frac{t_S}{t}\right) \\
&\le \sqrt{l}\frac{t - t_S}{T} < \varepsilon.
\end{aligned}$$

Therefore, it follows that

$$Prob\left(\left(\frac{n_1}{t}, \frac{n_2}{t}, ..., \frac{n_l}{t}\right) \in P^{2\varepsilon} | x_1^t\right) \ge p$$

for all $t \geq T$ and $x_1^t \in \{X_1'(\theta)\}^t$.

Next, prove the first half of the proposition. Suppose that

$$\prod_{s=1}^{l} \left(\frac{\nu_2^{\theta'}(s)}{\nu_2^{\theta}(s)} \right)^{p_s} < 1$$

for all $\theta' \in \Theta \backslash \{\theta\}$ and $(p_1, p_2, ..., p_l) \in P$. Since P is a compact set,

$$\max_{\theta' \in \Theta \backslash \{\theta\}, (p_1, p_2, ..., p_l) \in P} \prod_{s=1}^{l} \left(\frac{\nu_2^{\theta'}(s)}{\nu_2^{\theta}(s)} \right)^{p_s}$$

is well-defined. Let $1 - 2\delta$ be this maximum value. Then, there exists $\varepsilon > 0$ such that

$$\prod_{s=1}^{l} \left(\frac{\nu_2^{\theta'}(s)}{\nu_2^{\theta}(s)} \right)^{p_s} < 1 - \delta$$

for all $\theta' \in \Theta \backslash \{\theta\}$ and $(p_1, p_2, ..., p_l) \in P^{2\varepsilon}$. Take such ε. When the true state is θ, agent 2 p-believes θ if and only if

$$\sum_{\theta' \neq \theta} \frac{p_0(\theta')}{p_0(\theta)} [\prod_{s=1}^{l} \left(\frac{\nu_2^{\theta'}(s)}{\nu_2^{\theta}(s)} \right)^{\frac{n_s}{t}}]^t \leq \frac{1}{p} - 1. \tag{1}$$

Hence, if $t \geq T$, agent 1 always p-believes the event that

$$\prod_{s=1}^{l} \left(\frac{\nu_2^{\theta'}(s)}{\nu_2^{\theta}(s)} \right)^{\frac{n_s}{t}} < 1 - \delta.$$

Take $\bar{t}$ which satisfies $\bar{t} \geq T$ and $\sum_{\theta' \neq \theta} \frac{p_0(\theta')}{p_0(\theta)} [1 - \delta]^{\bar{t}} \leq \frac{1}{p} - 1$. Then, if $t \geq \bar{t}$, agent 1 always p-believes that inequality (1) holds. Hence, the irrelevant phase exists. In particular, $t \geq \bar{t}$ is a subset of the irrelevant phase.

Finally, prove the second half of the proposition. Take $\theta' \in \Theta \backslash \{\theta\}$ and $(p_1, p_2, ..., p_l) \in P \cap (\mathbb{R}_{>0})^l$ which satisfies

$$\prod_{s=1}^{l} \left(\frac{\nu_2^{\theta'}(s)}{\nu_2^{\theta}(s)} \right)^{p_s} > 1.$$

Let $1 + 2\delta$ be the value of the left-hand side. Note that $\nu_2^{\theta'}(s) > 0$ for all $s = 1, 2, ..., l$. Since rational numbers are dense in $\mathbb{R}$, there exists $\{q^{x_1}\}_{x_1 \in X_1'(\theta)} \subset \mathbb{Q}_{\geq 0}$ such that

$$\sum_{x_1 \in X_1'(\theta)} q^{x_1} = 1$$

$$\prod_{s=1}^{l} \left(\frac{\nu_2^{\theta'}(s)}{\nu_2^{\theta}(s)} \right)^{q_s} > 1 + \delta,$$

where $q_s = \sum_{x_1 \in X_1'(\theta)} q^{x_1} p_s^{x_1}$. Take $\varepsilon > 0$ such that

$$\prod_{s=1}^{l} \left(\frac{\nu_2^{\theta'}(s)}{\nu_2^{\theta}(s)} \right)^{r_s} > 1 + \delta,$$

for all $(r_1, r_2, ..., r_l) \in B_\varepsilon((q_1, q_2, ..., q_l))$. Fix any T, then, there exists $t \geq T$ such that $\frac{p_0(\theta')}{p_0(\theta)}[1+\delta]^t > \frac{1}{p} - 1$, $q^{x_1}t$ is an integer for all $x_1 \in X_1'(\theta)$, and $Prob((\frac{n_1}{t}, \frac{n_2}{t}, ..., \frac{n_l}{t}) \in B_\varepsilon(q_1, q_2, ..., q_l)|$ agent 1 receives each signal x_1 $q^{x_1}t$ times $) > 1 - p$. Hence, agent 1 assigns probability less than p to the event that agent 2 p-believes θ. Therefore, the irrelevant phase does not exist due to Proposition 6.

H: Proof of Corollary 2 (for the general case)

Let $m = \inf_{x_1 \in X_1'(\theta)} \nu(\Theta \times X_1 \times E_2(\theta)|x_1)$. Since $X_2^t \backslash \mathcal{B}_{t2}^{p,\mathcal{S}'}(\theta) \subseteq \{X_2 \backslash E_2(\theta)\}^t$, it follows that

$$\begin{aligned} 1 - \nu^t(\Theta \times \{X_1'(\theta)\}^t \times \mathcal{B}_{t2}^{p,\mathcal{S}'}(\theta)|x_1^t) &< \prod_{k=1}^{t} \nu(\Theta \times X_1'(\theta) \times X_2 \backslash E_2(\theta)|x_{1k}^t) \\ &< (1-m)^t \end{aligned}$$

for all $x_1^t \in \{X_1'(\theta)\}^t$. Hence, for all $p \in (0,1)$, there exists $\bar{t}$ such that $\nu^t(B_t^{p,\mathcal{S}'}(\theta)|x_1^t) > p$ for all $t \geq \bar{t}$ and for all $x_1^t \in \{X_1'(\theta)\}^t$.
Therefore, by Proposition 6, for all $p \in (0,1)$, there exists $\bar{t}$ such that $\mu^t(C_t^{p,\mathcal{S}}(\theta)|\theta) = \nu^t(C_t^{p,\mathcal{S}'}(\theta)|\theta)$ for all $t \geq \bar{t}$.

I: Proof of Lemma 3

First, note that

$$\begin{aligned} &\mu^t(\{\theta\} \times X_1^t \times X_2^t|x_2^t) \\ =&p_0(\theta)\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^t exp(-\frac{1}{2\sigma^2}(t(\theta - \overline{x_2^t})^2 + tS^2))[\sum_{\theta' \in \Theta}\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^t exp(-\frac{1}{2\sigma^2}(t(\theta' - \overline{x_2^t})^2 + tS^2))]^{-1}, \end{aligned}$$

where $\overline{x_2^t} = \frac{1}{t}\sum_{k=1}^t x_{2k}^t$ and $S^2 = \frac{1}{t}\sum_{k=1}^t (x_{2k}^t)^2 - \left(\frac{1}{t}\sum_{k=1}^t x_{2k}^t\right)^2$. Hence, for any $\theta, \theta' \in \Theta$, we have

$$\frac{\mu^t(\{\theta\} \times X_1^t \times X_2^t|x_2^t)}{\mu^t(\{\theta'\} \times X_1^t \times X_2^t|x_2^t)} = \frac{p_0(\theta)}{p_0(\theta')}\frac{exp(-\frac{1}{2\sigma^2}t(\theta - \overline{x_2^t})^2)}{exp(-\frac{1}{2\sigma^2}t(\theta' - \overline{x_2^t})^2)}.$$

Then, agent 2 p-believes θ at t if and only if

$$\frac{p_0(\theta)exp(-\frac{1}{2\sigma^2}t(\theta - \overline{x_2^t})^2)}{\sum_{\theta' \in \Theta} p_0(\theta')exp(-\frac{1}{2\sigma^2}t(\theta' - \overline{x_2^t})^2)} \geq p.$$

Since $p_0(\theta') > 0$ for all θ', this is equivalent to

$$\sum_{\theta' \neq \theta} \frac{p_0(\theta')}{p_0(\theta)} exp\left(\frac{t}{2\sigma^2}(\theta - \overline{x_2^t})^2 - \frac{t}{2\sigma^2}(\theta' - \overline{x_2^t})^2\right) \leq \frac{1-p}{p}.$$

Since it is assumed that $p \geq \frac{1}{1+m(\theta)}$,

$$\frac{p_0(\theta')}{p_0(\theta)} \geq m(\theta) \geq \frac{1-p}{p}$$

holds. Thus,

$$\begin{aligned}&\sum_{\theta'\neq\theta}\frac{p_0(\theta')}{p_0(\theta)}exp\left(\frac{t}{2\sigma^2}(\theta-\overline{x_2^t})^2-\frac{t}{2\sigma^2}(\theta'-\overline{x_2^t})^2\right)\\ \geq&\frac{1-p}{p}\sum_{\theta'\neq\theta}exp\left(\frac{t}{2\sigma^2}(\theta-\overline{x_2^t})^2-\frac{t}{2\sigma^2}(\theta'-\overline{x_2^t})^2\right).\end{aligned}$$

Therefore, for all $\theta'\neq\theta$, $\frac{t}{2\sigma^2}(\theta-\overline{x_2^t})^2-\frac{t}{2\sigma^2}(\theta'-\overline{x_2^t})^2\leq 0$ must hold. Hence, we have $|\theta'-\overline{x_2^t}|\leq|\theta-\overline{x_2^t}|$ for all θ'.

J: Proof of Lemma 4

Take any $\overline{x}\in\mathbb{R}$ that satisfies

$$\sum_{\theta'\neq\theta}\frac{p_0(\theta')}{p_0(\theta)}exp\left(\frac{t}{2\sigma^2}(\theta-\overline{x})^2-\frac{t}{2\sigma^2}(\theta'-\overline{x})^2\right)\leq\frac{1-p}{p}.$$

Then, it follows that

$$\begin{aligned}&\sum_{\theta'\neq\theta}\frac{p_0(\theta')}{p_0(\theta)}exp\left(\frac{t+1}{2\sigma^2}(\theta-\overline{x})^2-\frac{t+1}{2\sigma^2}(\theta'-\overline{x})^2\right)\\ =&\sum_{\theta'\neq\theta}\frac{p_0(\theta')}{p_0(\theta)}exp\left(\frac{t}{2\sigma^2}(\theta-\overline{x})^2-\frac{t}{2\sigma^2}(\theta'-\overline{x})^2\right)exp\left(\frac{1}{2\sigma^2}(\theta-\overline{x})^2-\frac{1}{2\sigma^2}(\theta'-\overline{x})^2\right)\\ \leq&\sum_{\theta'\neq\theta}\frac{p_0(\theta')}{p_0(\theta)}exp\left(\frac{t}{2\sigma^2}(\theta-\overline{x})^2-\frac{t}{2\sigma^2}(\theta'-\overline{x})^2\right)\\ \leq&\frac{1-p}{p}.\end{aligned}$$

The inequality in the third line follows from Lemma 3.

K: Proof of Lemma 5

Define $f:\Theta\times\mathbb{R}\to\mathbb{R}$ as

$$f(\theta,\overline{x})=p_0(\theta)exp\left(-\frac{t}{2\sigma^2}(\overline{x}-\theta)^2\right).$$

Let $g(\theta,\overline{x})=\frac{f(\theta)}{\sum_{\theta'\in\Theta}f(\theta')}$. Then, agent 2 p-believes θ at t if and only if

$$g(\theta,\overline{x_2^t})\geq p.$$

g is C^∞ with respect to $\overline{x}$ and

$$\frac{\partial g(\theta,\overline{x})}{\partial\overline{x}}=C(\theta,\overline{x})\sum_{\theta'\in\Theta}f(\theta',\overline{x})(\theta-\theta'),$$

where $C(\theta,\overline{x})=\frac{tf(\theta,\overline{x})}{\sigma^2\sum_{\theta'\in\Theta}f(\theta',\overline{x})}>0$.
Let $h(\theta,\overline{x})=\sum_{\theta'\in\Theta}f(\theta',\overline{x})(\theta-\theta')$. It follows that

$$\frac{\partial h(\theta,\overline{x})}{\partial\overline{x}}=-\frac{t}{\sigma^2}[\sum_{\theta'<\theta}(\theta-\theta')(\overline{x}-\theta')+\sum_{\theta'>\theta}(\theta-\theta')(\overline{x}-\theta')].$$

Hence, if $\overline{x}$ satisfies $|\theta - \overline{x}| \leq |\theta' - \overline{x}|$ for all θ', we have $\frac{\partial h(\theta,\overline{x})}{\partial \overline{x}} \leq 0$.
If agent 2 never p-believes θ regardless of which signals are received, this lemma holds, so suppose there exists $x_2^t \in X_2^t$ such that agent 2 p-believes θ if she receives x_2^t.
Let $I(\theta) \subset \mathbb{R}$ be the interval that

$$I(\theta) = \{\overline{x} \in \mathbb{R} || \theta - \overline{x}| \leq |\theta' - \overline{x}| \ \textit{for all}\ \theta'\}.$$

From Lemma 3, $\mathcal{B}_{t2}^{p,\mathcal{S}}(\theta) \subset I(\theta)$. Also, by the above argument, one of the following is true. (i) $g(\theta, \overline{x})$ is non-decreasing in $I(\theta)$, (ii) $g(\theta, \overline{x})$ is non-increasing in $I(\theta)$, or (iii) there exists $\alpha \in \mathbb{R}$ such that $g(\theta, \overline{x})$ is non-decreasing in $\{\overline{x} \in I(\theta) | x \leq \alpha\}$ and non-increasing in $\{\overline{x} \in I(\theta) | x \geq \alpha\}$. In all cases, $\mathcal{B}_{t2}^{p,\mathcal{S}}(\theta)$ must be an interval.