

# On consensus through communication without a commonly known protocol\*

ELIAS TSAKAS<sup>†</sup>

*Department of Economics, Maastricht University, The Netherlands*

MARK VOORNEVELD<sup>‡</sup>

*Department of Econometrics and Operations Research, Tilburg University, The Netherlands*

*Department of Economics, Stockholm School of Economics, Sweden*

March, 2010

## Abstract

The present paper extends the standard model of pairwise communication among Bayesian agents to cases where the structure of the communication protocol is not commonly known. We show that, even under strict conditions on the structure of the protocols and the nature of the transmitted signals, a consensus may never be reached if very little asymmetric information about the protocol is introduced.

KEYWORDS: Communication protocol, common knowledge, consensus.

JEL CODES: D82, D89.

---

\*We would like to thank David Ahn, Alpaslan Akay, Geir Asheim, Martin Dufwenberg, Amanda Friedenberg, John Geanakoplos, Aviad Heifetz, Lucie Menager, Friederike Mengel, Rohit Parikh, Andrés Perea, Andrew Postlewaite, Jeff Steif, Fernando Vega-Redondo, Jörgen Weibull, and the audience at UC Berkeley, Northwestern (GAMES 2008), Arizona, UCL (RES), Maastricht, Göteborg, Cardiff and C.R.E.T.E. for fruitful discussions and comments. Financial support from the Netherlands Organization for Scientific Research (NWO), the Wallander/Hedelius Foundation, the Marie Curie Fellowship, and the Adlerbertska Forskningsstiftelsen Foundation is gratefully acknowledged. Tsakas thanks UC Berkeley and the Stockholm School of Economics for their hospitality while working on this paper.

<sup>†</sup>Tel: +31-43-38 83649; Fax: +31-43-38 84878; E-mail: e.tsakas@maastrichtuniversity.nl

<sup>‡</sup>Tel: +46-8-736 9217; Fax: +46-8-31 32 07; E-mail: mark.voorneveld@hhs.se

# 1 Introduction

Common knowledge is a very central component of game theory. The concept of common knowledge was formalized by Aumann (1976) who shows in his seminal paper that if two people have the same prior beliefs, and their posteriors for an event are common knowledge, then they necessarily agree on their posteriors beliefs. Many well-known results rely on common knowledge of some elements of the game, e.g., common knowledge of rationality leads to correlated equilibrium (Aumann, 1987; Brandenburger and Dekel, 1987), common knowledge of everybody's willingness to participate in a trade plan precludes trading (Milgrom and Stokey, 1982), common knowledge of rationality and the opponents' conjectures about everybody's strategy in a normal form game suffices for Nash equilibrium (Aumann and Brandenburger, 1995).

In most cases common knowledge was a priori taken for granted. Geanakoplos and Polemarchakis (1982) were the first ones to study how common knowledge emerges in a dynamic environment where individuals start with asymmetric information. They show that if two individuals communicate their probabilistic beliefs back and forth, they will eventually agree on a – commonly known – probability assessment. Their setting has been the stepping stone for further development of models of communication in populations with Bayesian agents. The main aim of this literature is to study the conditions for reaching a consensus in groups of people through different communication mechanisms.

All the previous models assume that communication takes place through public announcement of the signals, which is quite restrictive. Parikh and Krasucki (1990) relax this assumption by introducing a model of pairwise private communication: They show that under some connectedness assumption (roughly, everybody talks to everybody either directly or via others) on the structure of the communication protocol, a consensus will be reached. A number of subsequent papers study the possibility of agreeing in environments with pairwise communication, under different assumptions about the signal functions, the protocol structure and information structure (Krasucki, 1996; Heifetz, 1996; Koessler, 2001; Houy and Menager, 2008).

A usual assumption in the existing literature – and very crucial for the results – is that the structure of the communication protocol is commonly known, i.e., everybody knows who talks to whom at every period, everybody knows that everybody knows this, and so on. In this paper we relax this assumption by introducing uncertainty about conversations that took place between third parties. Consider for instance the following example. Three individuals – Ann, Bob and Carol – privately talk about the probability they assign to some event. Communication takes place as follows: Ann talks to Bob, who talks to Carol, who talks back to Bob, who talks to Ann, and so on. If this structure is commonly known, the three individuals will eventually agree on a

common assessment (Krasucki, 1996). However, it is not straightforward whether this would still be the case if Carol did not know whether Ann had already talked to Bob in the first period or not.

First, notice that in order to address the previous question we need to formally model interactive knowledge about the protocol. The existing models do not serve this purpose, as the structure of the protocol is not formally an event of the state space, and therefore we cannot formally refer to knowledge or common knowledge about it. Hence, our first step is to incorporate the structure of the protocol into every state. This process induces a generalized state space which allows us to express the structure of a protocol as an event, and therefore knowledge of the protocol is well-defined.

Then, we provide an example which illustrates that a consensus may not be reached, even if we impose very stringent restrictions on the structure of the possible protocols and the nature of the transmitted signals. Namely, in our example, the individuals fail to agree, even though it is common knowledge from the beginning and throughout the entire process that communication satisfies information exchange<sup>1</sup>, which is a very strong requirement (Krasucki, 1996). Furthermore, even though we do not assume a commonly known protocol, the corresponding graph<sup>2</sup> is: This is quite surprising as all existing results on consensus provide sufficient conditions on the structure of the graph, after having implicitly assumed that the protocol inducing this graph is commonly known. It turns out that this implicit assumption is rather crucial: As we show, common knowledge of the graph may not suffice for reaching an agreement in the absence of common knowledge of the protocol that induces this graph.

Finally, in our example it is not the nature of the signals that drives the result. Parikh and Krasucki (1990) introduce the notion of convex<sup>3</sup> signals. They show that if the transmitted signals fail to satisfy the convexity requirement a consensus may not be reached. However, in our example the failure to reach an agreement does not rely on the lack of convexity, as we assume convex signals.

Our main (negative) result provides an alternative explanation why people often fail to agree. The usual justification until now, for such a failure, has been that the individuals may have dif-

---

<sup>1</sup>A protocol satisfies information exchange whenever,  $i$  talks to  $j$  if and only if  $j$  talks to  $i$ .

<sup>2</sup>Every protocol induces a graph summarizing how information is transmitted: Each individual corresponds to a vertex, and there is a directed edge from  $i$  to  $j$  if  $i$  talks to  $j$  infinitely often.

<sup>3</sup>A signal function  $f$  maps every conditioning event  $E \subseteq \Omega$  to a real-valued signal. We say that the signal function is convex whenever there is some  $\alpha \in [0, 1]$  such that  $f(E_1 \cup E_2) = \alpha f(E_1) + (1 - \alpha)f(E_2)$  for non-empty disjoint observations  $E_1, E_2 \subseteq \Omega$ . That is, the signal transmitted by  $i$  after having observed  $E_1 \cup E_2$  lays between the signals the same individual would have transmitted after having observed  $E_1$ , or  $E_2$ .

ferent prior beliefs. However, as we show even little asymmetric information on the structure of the protocol may be very significant for reaching an agreement.

The paper is organized as follows: Section 2 introduces the (standard) notation and terminology used throughout the paper. Section 3 generalizes the state space by incorporating the protocol structure into the states. The main result is presented in Section 4, while Section 5 concludes.

## 2 Notation and preliminaries

### 2.1 Information and knowledge

Consider a finite state space  $\Omega$  and a finite population  $N = \{1, \dots, n\}$ . Every state  $\omega \in \Omega$  is a complete description of all the relevant characteristics of the environment, i.e., it determines all the natural facts that occur at every instance.

Every individual  $i \in N$  is endowed with an *information partition*  $P_i$  of  $\Omega$ , with  $P_i(\omega)$  being the element of the information partition that contains  $\omega \in \Omega$ : It is the set of states that  $i$  cannot distinguish from  $\omega$ .

We define *knowledge* as usual:  $i$  knows a natural event  $E \subseteq \Omega$  at some state  $\omega$  whenever  $P_i(\omega) \subseteq E$ . An event  $E \subseteq \Omega$  is *common knowledge* at  $\omega$ , whenever everybody knows it, everybody knows that everybody knows it, and so on. Aumann (1976) showed that  $E$  is commonly known if and only if  $(P_1 \wedge \dots \wedge P_n)(\omega) \subseteq E$ , where  $P_1 \wedge \dots \wedge P_n$  denotes the finest common coarsening of the partitions and  $(P_1 \wedge \dots \wedge P_n)(\omega)$  is its element that contains  $\omega$ .

### 2.2 Signals and consensus

Let  $A$  be a non-empty set of signals, which contains the values of some parameter, e.g., the subjective probability assessments assigned to some event. A *signal (action) function*  $f_i : \Omega \rightarrow A$  determines the signal that agent  $i$  transmits at every  $\omega \in \Omega$ . We assume that  $i$ 's signal is  $P_i$ -measurable implying that  $i$  knows her own signal, i.e.,  $f_i(\omega') = f_i(\omega)$  for every  $\omega \in \Omega$  and every  $\omega' \in P_i(\omega)$ . A *consensus* has been reached at some state  $\omega$  if all individuals transmit the same signal at  $\omega$ , i.e., if there is some  $a \in A$  such that  $f_i(\omega) = a$  for all  $i \in N$ .

Agents in the population are *like-minded* if there is a function  $f : 2^\Omega \rightarrow A$ , called the virtual signal function, such that  $f_i(\omega) = f(P_i(\omega))$  for every  $i \in N$  and  $\omega \in \Omega$ . Throughout the paper, we assume that the agents are like-minded.

The function  $f$  satisfies *union consistency*<sup>4</sup> (Cave, 1983) if for all non-empty, disjoint  $E_1, E_2 \subseteq$

---

<sup>4</sup>Bacharach (1985) used the term *sure-thing principle* for the same property.

$\Omega$  with  $f(E_1) = f(E_2)$ , it holds that  $f(E_1 \cup E_2) = f(E_1)$ . Henceforth, unless stated otherwise, we assume without loss of generality that  $f$  is real-valued.

### 2.3 Communication protocols

Suppose that every individual  $i \in N$  starts with a prior information partition  $P_i^0$  over  $\Omega$ , which is transparent among the individuals in  $N$ . The information partition of agent  $i \in N$  at time  $t \in \mathbb{N}$  is denoted by  $P_i^t$ . At every time  $t = 0, 1, \dots$ , a conversation between two individuals may take place: When  $i$  talks to  $j$  at  $t$ , we say that  $i$  is the sender and  $j$  the receiver, and we write  $s_t = i$  and  $r_t = j$ . In this case,  $j$  updates her information by refining her partition  $P_j^t$  to  $P_j^{t+1}$ .

Updating is carried out in the standard way: Let  $V_i^t$  denote  $i$ 's *working partition* of  $\Omega$  at time  $t$  (Krasucki, 1996; Heifetz, 1996), with  $V_i^t(\omega)$  being the signal equivalence class that contains  $\omega$ , i.e.,

$$V_i^t(\omega) := \{\omega' \in \Omega : f_i^t(\omega') = f_i^t(\omega)\},$$

where  $f_i^t(\omega) := f(P_i^t(\omega))$  is the signal transmitted by  $i$  at  $\omega$ . Then  $j$  refines her partition in the following standard way (Parikh and Krasucki, 1990; Krasucki, 1996; Heifetz, 1996): For all  $\omega \in \Omega$ ,

$$P_j^{t+1}(\omega) = \begin{cases} P_j^t(\omega) & \text{if } j \neq r_t, \\ P_j^t(\omega) \cap V_i^t(\omega) & \text{if } j = r_t, \text{ where } i = s_t. \end{cases} \quad (1)$$

That is, the receiver rules out all states which are not consistent with the sender's signal, i.e., the receiver conditions<sup>5</sup> with respect to the signal she would hear at every state.

The sequence of senders and receivers  $\{(s_t, r_t)\}_{t=0}^\infty$  is called a *protocol* and determines who talks<sup>6</sup> to whom at every time. The protocol induces a graph on  $N$ : There is a directed edge from  $i$  to  $j$ , if  $i$  talks to  $j$  infinitely often, i.e., if there are infinitely many  $t$  with  $(s_t, r_t) = (i, j)$ .

Parikh and Krasucki (1990) called a protocol *fair* if the graph of directed edges is strongly connected, i.e., if there is a path of directed edges which starts from some individual, passes from all the vertexes (individuals), returning to its origin. In other words, everybody communicates with everybody directly or indirectly.

<sup>5</sup>An equivalent way of writing the refining mechanism is the following

$$P_j^{t+1} = \begin{cases} P_j^t & \text{if } j \neq r_t, \\ P_j^t \vee V_i^t & \text{if } j = r_t, \text{ where } i = s_t, \end{cases}$$

where the operator  $\vee$  denotes the coarsest common refinement (join) of the two partitions (Krasucki, 1996; Heifetz, 1996).

<sup>6</sup>We adopt the following notational convention:  $s_t = r_t$  corresponds to "nobody talking to anybody at time  $t$ ".

A protocol satisfies *information exchange* if, for all distinct  $i, j \in N$  with a directed edge from  $i$  to  $j$ , there is a directed edge from  $j$  to  $i$ , i.e.,  $i$  talks to  $j$  infinitely often if and only if  $j$  talks back to  $i$  infinitely often. Krasucki (1996) showed that communicating union-consistent signals through a fair protocol which satisfies information exchange leads to a consensus.

An implicit – but crucial – assumption underlying all results in the literature, is that the protocol is commonly known among the individuals in  $N$ . What happens otherwise is not very clear. We address this question in the following sections.

### 3 The generalized state space

In this section, we formalize the notion of knowledge of the protocol, and also study the implications of relaxing the assumption about common knowledge of the protocol.

#### 3.1 An example

This example illustrates how the receiver interprets a signal in the existence of asymmetric information about the protocol, and how she consequently updates her information upon hearing this signal.

Let  $N = \{a, b, c\}$  and consider the state space  $\Omega = \{\omega_1, \dots, \omega_4\}$ , endowed with the following virtual signal function: For every  $E \subseteq \Omega$ , let

$$f(E) = \frac{1}{\#E} \sum_{\omega \in E} f(\{\omega\}), \quad (2)$$

where  $\#E$  denotes the cardinality of  $E$  and

$$f(\{\omega_1\}) = 1, f(\{\omega_2\}) = 3, f(\{\omega_3\}) = 4 \text{ and } f(\{\omega_4\}) = 0.$$

It is straightforward to show that  $f$  is union-consistent. The prior information partitions over  $\Omega$  are

$$\begin{aligned} P_a^0 &= \{\{\omega_1, \omega_2\}_2, \{\omega_3\}_4, \{\omega_4\}_0\}, \\ P_b^0 &= \{\{\omega_1, \omega_2, \omega_3, \omega_4\}_2\}, \\ P_c^0 &= \{\{\omega_1\}_1, \{\omega_2, \omega_3, \omega_4\}_{7/3}\}. \end{aligned}$$

with the indexes denoting the signals that the individuals would transmit given the corresponding cell of their information partition.

Suppose that according to the actual communication protocol  $a$  talks to  $b$  at  $t = 0$ , but  $c$  does not know whether this conversation has occurred or not. Then  $c$  does not know whether  $b$  has

refined her partition to  $P_b^1 = \{\{\omega_1, \omega_2\}_2, \{\omega_3\}_4, \{\omega_4\}_0\}$  (after having heard  $a$ 's signal), or still holds her prior partition (which is the case if  $a$  has not talked to  $b$ ). Therefore, if  $c$  in the next period receives signal "2" from  $b$  at  $\omega_1$ , she cannot rule out the states  $\omega_3$  and  $\omega_4$ : If  $a$  did not talk to  $b$ , then  $b$  would say "2" at every state in  $\Omega$ , whereas if  $a$  talked to  $b$ , then  $b$  would say "2" at  $\omega_1$  and  $\omega_2$ . Since  $c$  cannot rule out any of the two possibilities, she can neither rule out any state. Hence,  $c$  does not update her information at  $t = 1$ .

### 3.2 Knowledge of the protocol and the generalized state space

Let  $Z$  denote the a (finite) set of protocols, with typical element  $z$ . Let  $s_t(z)$  and  $r_t(z)$  denote the sender and the receiver at time  $t$ , given the protocol  $z$ . We endow  $Z$  with a *partition*  $I_i^0$  for every individual. For every  $z \in Z$ , let  $I_i^0(z)$  denote the element of  $I_i^0$  that contains  $z$ : It is the set of protocols that  $i$  cannot distinguish from  $z$  before the communication begins. *Knowledge of the protocols* is defined as usual: Individual  $i$  knows the event  $G \subseteq Z$  whenever  $I_i^0(z) \subseteq G$ . *Common knowledge of the protocols* is defined analogously.

It is intuitively straightforward that  $i$  can always distinguish between two protocols that induce different communication structure *at the times when she participates in the communication*. Formally, let  $S_i(z) := \{t \in \mathbb{N} : s_t(z) = i\}$  and  $R_i(z) := \{t \in \mathbb{N} : r_t(z) = i\}$  denote the times when  $i$  acts as a sender and as a receiver respectively, given the protocol  $z$ . The following condition is assumed throughout the paper.

**Assumption 1.** For all  $i \in N$  and  $z, z' \in Z$ : if  $z' \in I_i^0(z)$ , then  $(s_t(z'), r_t(z')) = (s_t(z), r_t(z))$  for every  $t \in S_i(z) \cup R_i(z)$ .

That is, each individual knows (i) when she is spoken to and by whom and, (ii) when she speaks and to whom.

Let  $\Theta := \Omega \times Z$  be the *generalized state space*, each element of which fully describes all natural facts and also determines the protocol structure. Every individual inherits a *generalized information partition*  $\Pi_i^0$  over  $\Theta$ , which is derived from  $P_i^0$  and  $I_i^0$ , as follows:

$$\Pi_i^0(\omega, z) := \{(\omega', z') \in \Theta : \omega' \in P_i^0(\omega) \text{ and } z' \in I_i^0(z)\}. \quad (3)$$

The partition  $\Pi_i^0$  obviously satisfies  $\text{proj}_\Omega \Pi_i^0(\omega, z) = P_i^0(\omega)$  and  $\text{proj}_Z \Pi_i^0(\omega, z) = I_i^0(z)$ , where  $\text{proj}$  denotes the projection of a subset of  $\Theta$  on one of the coordinates: the private information induced by  $\Pi_i^0$  is consistent with the (prior) information induced by both  $P_i^0$  and  $I_i^0$ .

The following example illustrates how the generalized state space, together with the generalized information partitions, are constructed.

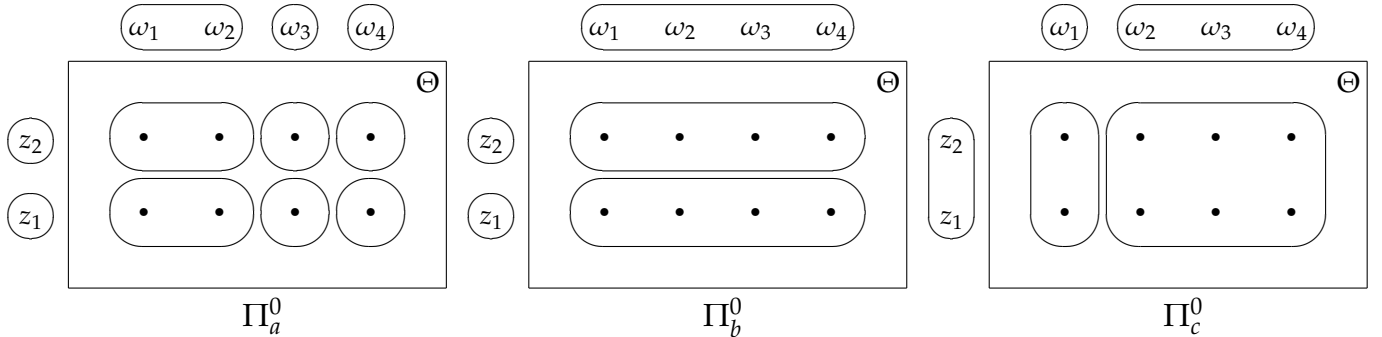
**Example 1.** Recall the Example from Section 3.1: There are three individuals  $N = \{a, b, c\}$  and the state space  $\Omega = \{\omega_1, \dots, \omega_4\}$ , together with the partitions

$$\begin{aligned} P_a^0 &= \{\{\omega_1, \omega_2\}, \{\omega_3\}, \{\omega_4\}\}, \\ P_b^0 &= \{\{\omega_1, \omega_2, \omega_3, \omega_4\}\}, \\ P_c^0 &= \{\{\omega_1\}, \{\omega_2, \omega_3, \omega_4\}\}. \end{aligned}$$

Consider also two possible protocols  $z = \{z_1, z_2\}$ , with  $z_1$  being the Round-Robin protocol ( $a$  talks to  $b$ , who talks to  $c$ , who talks to  $a$  and so on), and  $z_2$  being the same protocol with the difference that  $a$  does not talk to  $b$  at  $t = 0$ . Individuals  $a$  and  $b$  are able to distinguish between the two protocols, as  $z_1$  and  $z_2$  differ only in conversations they participate themselves. Individual  $c$  on the other hand, cannot distinguish between  $z_1$  and  $z_2$ . Note that  $c$ 's inability to distinguish between the two protocols is consistent with Assumption 1, as the two protocols differ only in conversations at which  $c$  does not participate either as a sender or as a receiver. Thus,

$$\begin{aligned} I_a^0 &= \{\{z_1\}, \{z_2\}\}, \\ I_b^0 &= \{\{z_1\}, \{z_2\}\}, \\ I_c^0 &= \{\{z_1, z_2\}\}. \end{aligned}$$

The following information partitions over  $\Theta$  respect  $P_i^0$  and  $I_i^0$  when projected to the corre-



sponding coordinate, as illustrated in the following figure. ◁

Obviously, when the protocol is commonly known, the  $Z$ -dimension becomes irrelevant, and therefore we are back to the standard setting, where  $\Theta$  is degenerated to  $\Omega$ , and  $\Pi_i^0$  is degenerated to  $P_i^0$ .

### 3.3 Signals and communication in the generalized state space

When the protocol is commonly known, signals are generated by the virtual signal function, as described in Section 2.2.

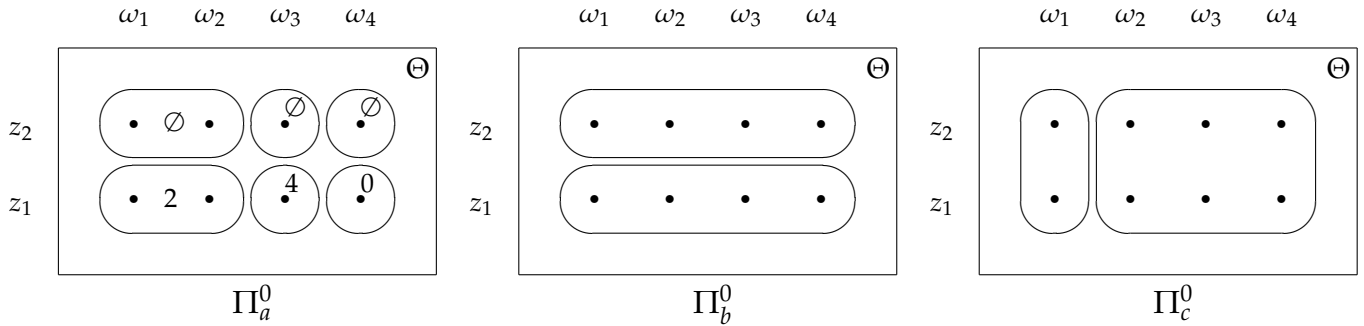


Now, suppose that there is incomplete information about the protocol. Recall that the signal function maps the individual's private information over  $\Omega$  to an element of  $A$ . That is, at  $(\omega, z)$  the sender  $s_t(z)$  transmits the signal that follows from her private information over  $\Omega$ . The following example illustrates how we generalize signals when the protocol is not commonly known.

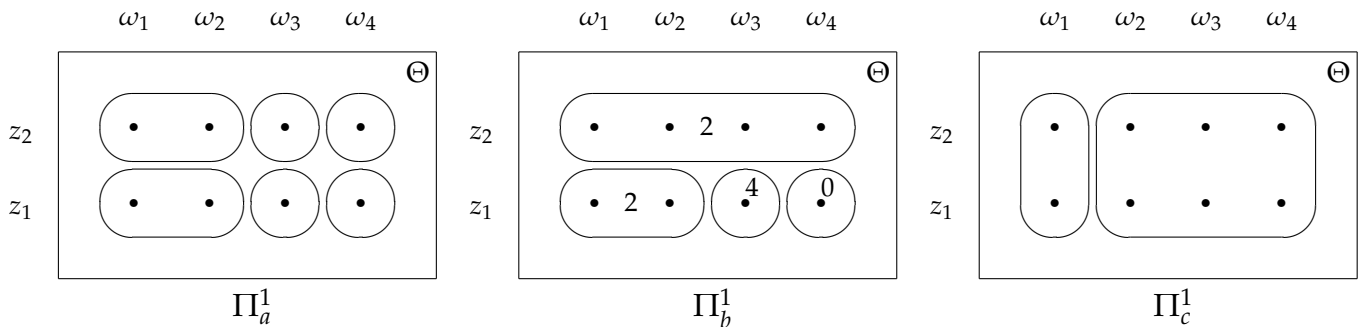
**Example 2.** Recall the generalized partitions from Example 1, and let the virtual signal function be the one defined in Equation (2), with

$$f(\{\omega_1\}) = 1, f(\{\omega_2\}) = 3, f(\{\omega_3\}) = 4 \text{ and } f(\{\omega_4\}) = 0.$$

If the actual protocol is  $z_2$  – according to which  $a$  never talks to  $b$  – all (generalized) states  $(\omega, z_2)$  yield the same signal  $\emptyset$ , which stands for “no signal”. If, on the other hand, the actual protocol is  $z_1$  – according to which  $a$  talks to  $b$  at the first period – the (generalized) states in  $\{(\omega_1, z_1), (\omega_2, z_1)\}$  yield signal “2”,  $(\omega_3, z_1)$  yields “4”, and  $(\omega_4, z_1)$  yields “0” (see the figure below).



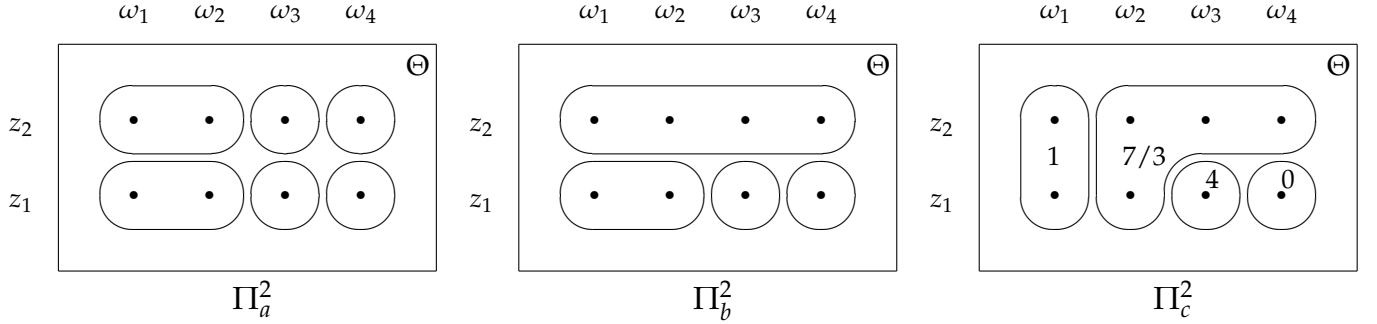
Since  $a$ 's generalized partition  $\Pi_a^0$  is transparent to everybody,  $b$  can associate every signal with a unique generalized state, and therefore refines his generalized partition (in the usual way) to  $\Pi_b^1$ , as depicted below.



In the next period, both  $z_1$  and  $z_2$  are such that  $b$  talks to  $c$ , who infers that, if the protocol is  $z_2$  – according to which  $a$  has not talked to  $b$  in the previous period – then  $b$  says “2” at all (natural)

states, i.e.,  $b$  says “2” at  $(\omega, z_2)$  for all  $\omega \in \Omega$ . If, on the other hand, the protocol is  $z_1$  – according to which  $a$  has talked to  $b$  at the first period – the (generalized) states in  $\{(\omega_1, z_1), (\omega_2, z_1)\}$  yield signal “2”,  $(\omega_3, z_1)$  yields “4”, and  $(\omega_4, z_1)$  yields “0”.

Since  $\Pi_b^1$  is transparent to everybody,  $c$  refines his generalized partition (in the standard way) to  $\Pi_c^2$ , as depicted below.



Now, at the third period,  $c$  talks to  $a$  according to both  $z_1$  and  $z_2$ . At the generalized state  $(\omega_2, z_1)$ , individual  $c$  has ruled out the natural states which do not occur at any generalized state in  $\Pi_c^2(\omega_2, z_1)$ . All other natural states may occur according to some protocol deemed as possible by  $c$  at  $(\omega_2, z_1)$ . Thus,  $c$  rules out only  $\omega_1$ , and therefore says “7/3”, as depicted above.  $\triangleleft$

Formally, at each time  $t$ , agent  $i \in N$  has the generalized information partition  $\Pi_i^t$ , with the initial partition  $\Pi_i^0$  being as defined in (3). The *generalized signal function*  $h_i^t : \Theta \rightarrow \mathbb{R} \cup \{\emptyset\}$  is such that at every generalized state  $(\omega, z)$  the value of the signal is

$$h_i^t(\omega, z) = \begin{cases} \emptyset & \text{if } i \neq s_t(z), \\ f(\text{proj}_\Omega \Pi_i^t(\omega, z)) & \text{if } i = s_t(z). \end{cases} \quad (4)$$

That is,  $i$  does not transmit any signal at  $(\omega, z)$  if she is not assigned (by the protocol  $z$ ) to be the sender at  $t$ . If, on the other hand,  $z$  assigns  $i$  to speak at time  $t$ , she sends the value that the virtual signal function assigns to the subset of natural states that  $i$  cannot rule out at  $(\omega, z)$ .

Let  $W_i^t$  denote  $i$ 's *generalized working partition* of  $\Theta$  at  $t$ , with  $W_i^t(\omega, z)$  being the element of the partition which contains  $(\omega, z)$ , i.e.,

$$W_i^t(\omega, z) = \{(\omega', z') \in \Theta : h_i^t(\omega', z') = h_i^t(\omega, z)\} \quad (5)$$

contains the generalized states that yield the same signal as  $(\omega, z)$  at  $t$ .

Information refining takes place in the standard way (Parikh and Krasucki, 1990; Krasucki, 1996; Heifetz, 1996): For every  $(\omega, z) \in \Theta$ ,

$$\Pi_j^{t+1}(\omega, z) = \begin{cases} \Pi_j^t(\omega, z) & \text{if } j \neq r_t(z), \\ \Pi_j^t(\omega, z) \cap W_i^t(\omega, z) & \text{if } j = r_t(z), \text{ where } i = s_t(z). \end{cases} \quad (6)$$

That is,  $j$  updates her information only at generalized states that assign her to be the receiver.

**Remark 1.** By Assumption 1 all individuals know at all states ( $i$ ) when they are engaged in communication, and ( $ii$ ) who the sender and receiver are at such occasions. Since, by construction, the (prior) generalized partition respects the information about the protocol, it follows from  $\Pi_i^{t+1}(\omega, z) \subseteq \Pi_i^t(\omega, z)$  that  $i$  keeps knowing this information. Thus, the updating mechanism is well-defined.  $\triangleleft$

## 4 Failing to agree when the protocol is not commonly known

Consider a commonly known fair protocol which satisfies information exchange, and let the virtual signal function satisfy union consistency. Under these assumptions, a consensus is necessarily reached (Krasucki, 1996). The following result shows that this is no longer true if the protocol is not commonly known, even if it remains common knowledge that the protocol satisfies information exchange.

**Negative result.** *If the protocol is not common knowledge, then a consensus may never be reached, even if (a) agents are like-minded, (b) signals are union-consistent, and (c) it is common knowledge that the protocol is fair and satisfies information exchange.*

The following example proves the result. Let  $N = \{a, b, c, d\}$  and  $\Omega = \{\omega_1, \dots, \omega_6\}$ . Let the virtual signal function be the one defined in Equation (2), with

$$f(\{\omega_1\}) = 2, f(\{\omega_2\}) = f(\{\omega_6\}) = 5, f(\{\omega_3\}) = f(\{\omega_5\}) = 8 \text{ and } f(\{\omega_4\}) = 4.$$

The two protocols differ only in the conversations that take place at  $t = 0$ :

	0	1	2	3	4	5	6	7	...
$z_1$	$c \rightarrow a$	$a \rightarrow b$	$b \rightarrow a$	$a \rightarrow c$	$c \rightarrow a$	$a \rightarrow d$	$d \rightarrow a$	$a \rightarrow b$	...
$z_2$	$d \rightarrow a$	$a \rightarrow b$	$b \rightarrow a$	$a \rightarrow c$	$c \rightarrow a$	$a \rightarrow d$	$d \rightarrow a$	$a \rightarrow b$	...

Note that the conversations that take place at every  $t \geq 1$  are common knowledge, whereas what happens at the first period is only known to  $a, c$  and  $d$ . Therefore – in line with Assumption 1 – we assume that  $a, c$  and  $d$  can distinguish the protocols, whereas  $b$  cannot:  $I_i^0 = \{\{z_1\}, \{z_2\}\}$  if  $i \in \{a, c, d\}$  and  $I_b^0 = \{\{z_1, z_2\}\}$ .

The prior generalized information partitions, together with the signals transmitted by the corresponding senders at  $t = 0$ , are depicted in Figure 1.

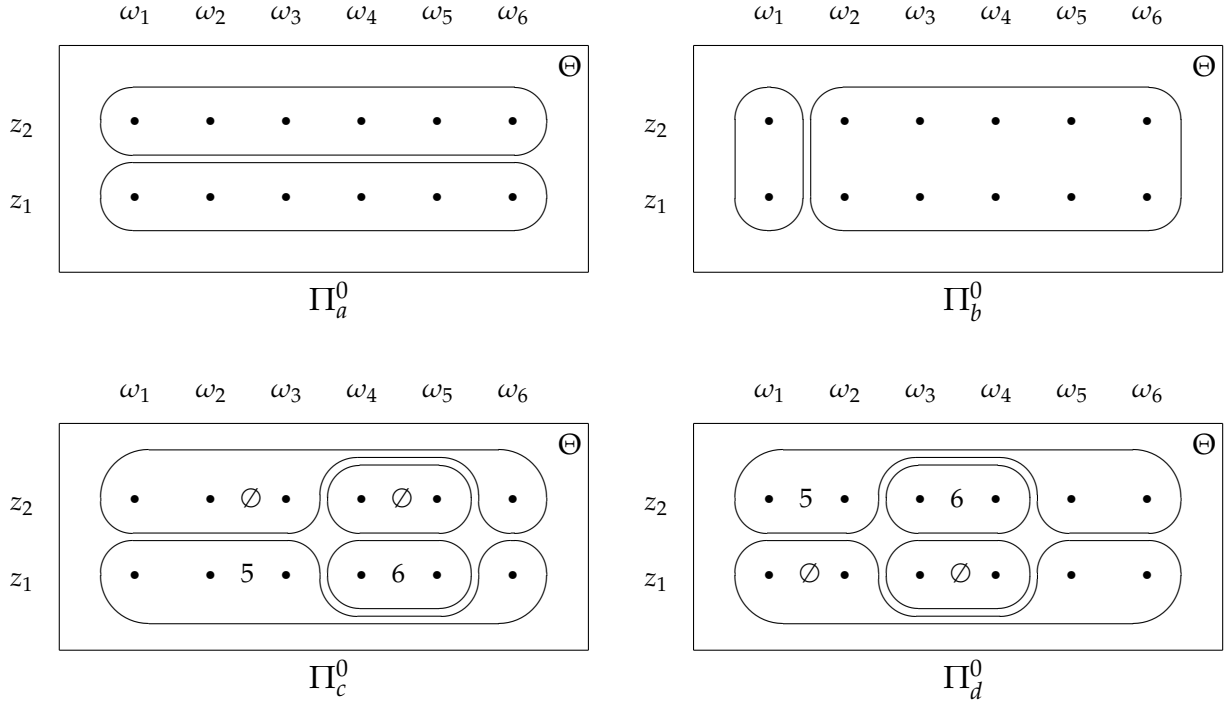


Figure 1: Prior generalized partitions and signals

One easily verifies that after  $t = 6$ , the generalized information partitions are the ones depicted in Figure 2, together with the corresponding signals. All the steps are presented in the Appendix. Notice that no refinement takes place after this time, implying that a consensus is never reached at  $(\omega_2, z_2)$  for instance.

The remarks below stress that the lack of consensus is ultimately due to the lack of common knowledge of the protocol, not due to violations of other conditions that are known causes for lack of agreement.

**Remark 2.** The previous result may be surprising, as information exchange is a very strong requirement, which always leads to consensus (Krasucki, 1996). This is the case because the parties involved in a repeated bilateral communication, could in principle disregard all communications in which they do not participate, and still agree with everybody they are connected. However, in the absence of common knowledge of the protocol structure, the receiver at time  $t$  fails to ignore the fact that the sender has already heard signals from other individuals at various  $t' < t$ . This is mainly because the actual signal aggregates the signals that would have been sent given every possible protocol. Thus, the receiver cannot disaggregate them, and therefore cannot ignore the fact that the sender has already received information by other individuals in the past.  $\triangleleft$

**Remark 3.** As the two possible protocols in the example differ in finitely many periods – in fact, the first period only – they correspond to the same graph. In other words, common knowledge

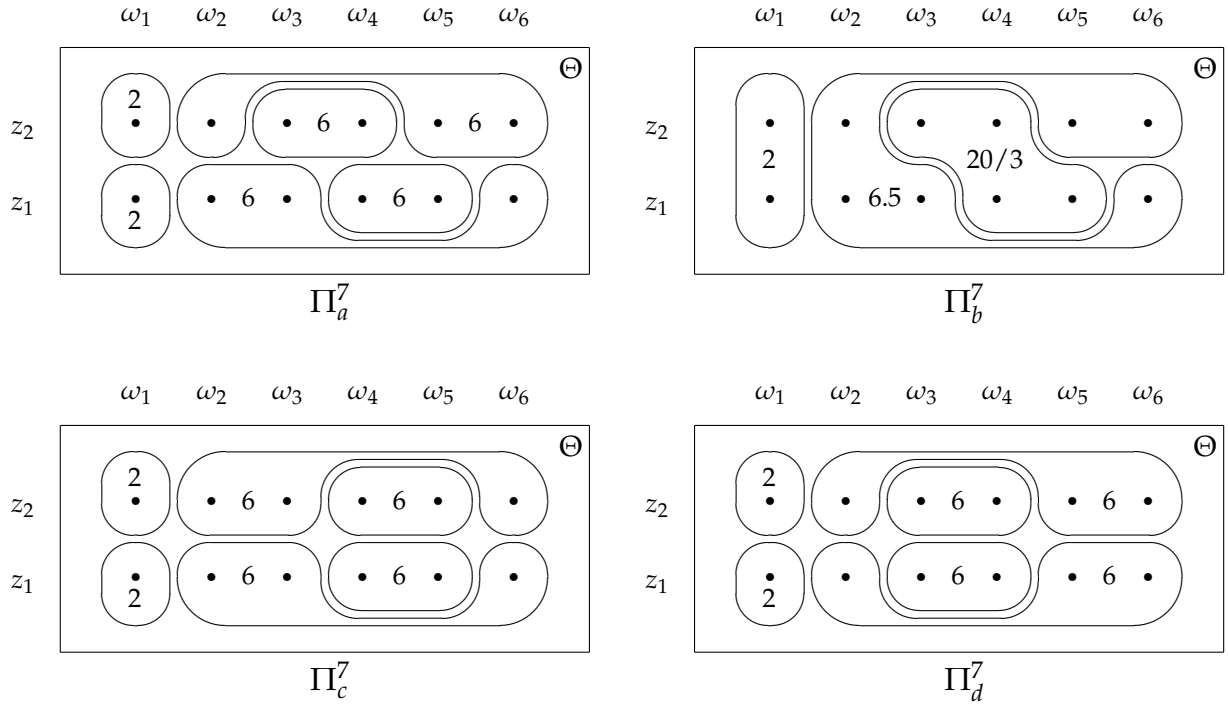


Figure 2: The final generalized information partitions

of the graph induced by the protocol does not suffice either for consensus: even little asymmetric information can lead to disagreement.  $\triangleleft$

**Remark 4.** Parikh and Krasucki (1990) show that consensus may not be reached through an arbitrary (commonly known) fair protocol, if the signals are union-consistent, but not convex (see Footnote 3). In our example above, the signal function is clearly convex. Hence, the lack of consensus is not due to a lack of convexity.  $\triangleleft$

## 5 Concluding discussion

### 5.1 Relationship to the existing literature

As we have already mentioned, the literature on communication and consensus, almost unanimously assumes common knowledge of the protocol. The only attempts to depart from such an environment are those of Heifetz (1996) and Koessler (2001), who study a particular form of asymmetric information about the protocol. Namely, they allow for the possibility that the message transmitted by the sender at some  $t \in \mathbb{N}$  fails to be delivered to the receiver with positive probability. This is a special case of our setting: Recall Example 1, where  $c$  does not know whether  $a$  has talked to  $b$  or not, at  $t = 0$ . Formally, this is equivalent to Heifetz's "closed eyes case", where

$c$  also considers the possibility that  $b$  has failed to receive  $a$ 's signal.

Though their research question is similar to ours, the model they employ differs from ours in that they enlarge the state space by adding time – instead of the protocols space – as the second dimension. In the previous example, their generalized state space would be such that  $c$  cannot distinguish between times  $t = 0$  and  $t = 1$ . That is, according to  $c$  the actual time is  $t = 1$  if  $a$  has already talked to  $b$ , and  $t = 0$  otherwise. Since, she does not know whether the conversation has taken place, she cannot distinguish between  $t = 0$  and  $t = 1$ .

Obviously, our framework is more general as it can capture any kind of asymmetric information about the protocol structure, e.g., recall the example in Section 4: At  $t = 0$  either  $c$  talks to  $a$  or  $d$  talks to  $a$ , implying that  $b$  knows at  $t = 1$  that this is the second round of communication, and therefore using generalizing the state space by simply incorporating time does not allow us to model  $b$ 's uncertainty about the protocol.

## 5.2 Towards a positive result

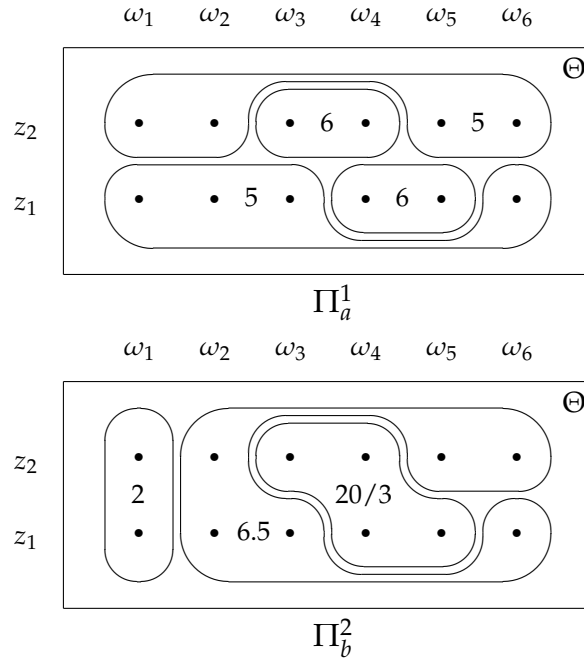
As we have already mentioned in Remark 4, the failure to reach an agreement does not rely on the signals not being convex. Therefore, for a positive result to be established, even stricter conditions on the virtual signal function are required. It is not hard to prove that by extending union consistency to non-disjoint events, a consensus is always reached. However, such an extension would rule out many classes of signal functions, such as conditional probabilities. Aumann et al. (2005) study the class of signals that satisfy this particular property, which they call *logical sure-thing principle*, and also clarify its relationship with union consistency. In any case, providing weaker (than the logical sure-thing principle) sufficient conditions for a consensus under asymmetric information about the protocols remains an open question for future research.

## A Appendix

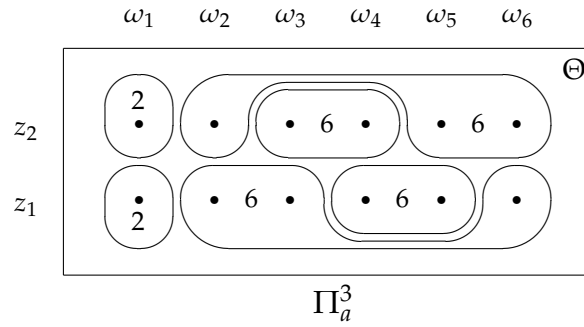
Below, we show how the generalized information partitions in the example of Section 4 evolve over time until  $t = 7$ . As we have already shown, after  $t = 7$  no further updating occurs, implying that the individuals never reach a consensus. For the sake of compactness, we present for every  $t = 1, \dots, 7$  only the generalized partitions of individuals who refine at that period.

At  $t = 0$ , individual  $c$  talks to  $a$  according to  $z_1$ , and  $d$  talks to  $a$  according to  $z_2$ . Therefore,  $a$  is the only one who refines her partition to  $\Pi_a^1$ .

At  $t = 1$ , individual  $a$  talks to  $b$  according to both protocols, and  $b$  refines her partition to  $\Pi_b^2$ , while everybody maintains their current generalized partition.



At  $t = 2$ , individual  $b$  talks back to  $a$  according to both protocols, and  $a$  refines her partition to  $\Pi_a^3$ , while everybody maintains their current generalized partition.



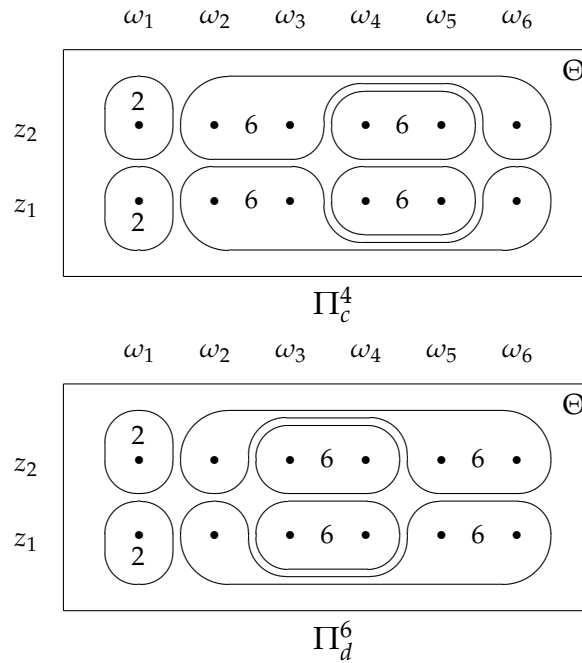
At  $t = 3$ , individual  $a$  talks to  $c$  according to both protocols, and  $c$  refines her partition to  $\Pi_c^4$ , while everybody maintains their current generalized partition.

At  $t = 4$ , individual  $c$  talks back to  $a$  according to both protocols. Since  $c$ 's generalized working partition is  $\Pi_a^4$ -measurable,  $a$  does not refine her partition, and therefore  $\Pi_a^5 = \Pi_a^4$ . Everybody maintains their current generalized partition.

At  $t = 5$ , individual  $a$  talks to  $d$  according to both protocols, and  $d$  refines her partition to  $\Pi_d^6$ , while everybody maintains their current generalized partition.

At  $t = 6$ , individual  $d$  talks back to  $a$  according to both protocols. Since  $d$ 's generalized working partition is  $\Pi_a^6$ -measurable,  $a$  does not refine her partition, and therefore  $\Pi_a^7 = \Pi_a^6$ . Everybody maintains their current generalized partition.

As we have already discussed in Section 4, nobody ever updates her generalized information partition after  $t = 7$ , implying that  $\Pi_i^t = \Pi_i^7$  for all  $t > 7$ , and for every  $i \in N$ , implying that a consensus will never be reached.



## References

- AUMANN, R.J. (1976). Agreeing to disagree. *Annals of Statistics* 4, 1236–1239.
- (1987). Correlated equilibrium as an expression of Bayesian rationality. *Econometrica* 55, 1–18.
- AUMANN, R.J., BRANDENBURGER, A. (1995). Epistemic conditions for Nash equilibrium. *Econometrica* 63, 1161–1180.
- AUMANN, R.J., HART, S., PERRY, M. (2005). Conditioning and the sure-thing principle. *The Hebrew University of Jerusalem, Center for Rationality Discussion Paper* 393.
- BACHARACH, M. (1985). Some extensions of a claim of Aumann in an axiomatic model of knowledge. *Journal of Economic Theory* 37, 167–190.
- BRANDENBURGER, A., DEKEL, E. (1987). Rationalizability and correlated equilibria. *Econometrica* 55, 1391–1402.
- CAVE, J.A.K. (1983). Learning to agree. *Economics Letters* 12, 147–152.
- GEANAKOPOLOS, J., POLEMARCHAKIS, H. (1982). We can't disagree forever. *Journal of Economic Theory* 28, 192–200.
- HEIFETZ, A. (1996). Comment on consensus without common knowledge. *Journal of Economic Theory* 70, 273–277.



- HOUY, N., MENAGER, L. (2008). Communication, consensus, and order: Who wants to speak first? *Journal of Economic Theory* 143, 140–152.
- KOESSLER, F. (2001). Common knowledge and consensus with noisy communication. *Mathematical Social Sciences* 42, 139–159.
- KRASUCKI, P. (1996). Protocols forcing consensus. *Journal of Economic Theory* 70, 266–272.
- MILGROM, P., STOKEY, N. (1982). Information, trade and common knowledge. *Journal of Economic Theory* 26, 17–27.
- PARIKH, R., KRASUCKI, P. (1990). Communication, consensus, and knowledge. *Journal of Economic Theory* 52, 178–189.