

# Agreeing to Disagree and Dilation

Jiji Zhang<sup>a</sup>, Hailin Liu<sup>b</sup>, Teddy Seidenfeld<sup>c</sup>

<sup>a</sup>*Department of Philosophy, Lingnan University, Hong Kong, China*

<sup>b</sup>*Institute of Logic and Cognition, Department of Philosophy, Sun Yat-sen University, Guangzhou, China*

<sup>c</sup>*Department of Philosophy, Carnegie Mellon University, Pittsburgh, USA*

---

## Abstract

We consider Geanakoplos and Polemarchakis’s generalization of Aumann’s famous result on “agreeing to disagree”, in the context of imprecise probability. The main purpose is to reveal a connection between the possibility of agreeing to disagree and the interesting and anomalous phenomenon known as dilation. We show that for two agents who share the same set of priors and update by conditioning on every prior, it is impossible to agree to disagree on the lower or upper probability of a hypothesis unless a certain dilation occurs. With some common topological assumptions, the result entails that it is impossible to agree not to have the same set of posterior probability values unless dilation is present. This result may be used to generate sufficient conditions for guaranteed full agreement in the generalized Aumann-setting for some important models of imprecise priors, and we illustrate the potential with an agreement result involving the density ratio classes. We also provide a formulation of our results in terms of “dilation-averse” agents who ignore information about the value of a dilating partition but otherwise update by full Bayesian conditioning.

### *Keywords:*

Agreeing to disagree, Common knowledge, Dilation, Imprecise probability.

---

## 1. Introduction

In a simple but insightful paper, Aumann (1976) famously showed that two (Bayesian) agents who start with the same (precise) prior cannot agree to disagree on their posteriors of a hypothesis, in the sense that if the posteriors of the hypothesis (as well as the structures of their respective information partitions) are common knowledge, then the posteriors must be equal. This result has been generalized in at least two ways. First, Aumann’s result applies only to those events whose posteriors happen to be common knowledge. Geanakoplos and Polemarchakis (1982) generalized the framework to a communication setting where the agents are invited to repeatedly make their credences common knowledge by announcements and update by conditioning on the announced credences, until no new information is conveyed. They showed that for any hypothesis/event, this communication procedure is guaranteed to lead to an agreement on the probability

---

*Email addresses:* [jjizhang@ln.edu.hk](mailto:jjizhang@ln.edu.hk) (Jiji Zhang), [liuhlin3@mail.sysu.edu.cn](mailto:liuhlin3@mail.sysu.edu.cn) (Hailin Liu), [teddy@stat.cmu.edu](mailto:teddy@stat.cmu.edu) (Teddy Seidenfeld)

of the hypothesis, if the agents start with the same (precise) prior (and each agent’s information partition is finite).

Second, Kajii and Ui (2005, 2009) and Carvajal and Correia-da-Silva (2010) generalized Aumman’s result in the setting of multiple priors. In this line of work, “agreement” is taken to mean “partial agreement”, in the sense that two sets of probabilities agree if they have a non-empty intersection. These authors established several sufficient conditions under which two agents who (partially) agree on their priors are guaranteed to (partially) agree on their posteriors of a hypothesis if these posteriors are common knowledge.

In this paper, we combine the two more general settings and establish a connection between the possibility of agreeing to disagree and the interesting and anomalous phenomenon known as dilation (Good, 1974; Seidenfeld, 1981; Walley, 1991; Seidenfeld and Wasserman, 1993; Herron et al., 1997). Dilation occurs when conditioning on each element of a partition, the lower and upper probabilities of a hypothesis become more divergent than the unconditional ones. In such a case, for agents who use conditioning as their updating rule, their credences on a hypothesis become less precise or determinate after learning the value of the dilating partition, no matter which value they learn! This counterintuitive phenomenon is often interpreted as a distinctive challenge to the orthodox Bayesian doctrine on the value of information and to the Bayesian merging of opinions, but as far as we know, it has never been discussed in connection to Aumman’s result. We shall show that it is the key obstacle to reaching agreements via communicating posteriors by Bayesian agents with imprecise priors.

We will establish the following. After introducing the setting and reviewing the special case of precise probabilities in Section 2, we show in Section 3 that dilation is the *only* obstacle for Bayesian agents with the same prior to reaching agreements on lower and upper probabilities of a hypothesis by communicating their posteriors on the hypothesis. Without dilation, the two agents in our setting are guaranteed to end up agreeing on lower and upper probabilities of the hypothesis of interest. An immediate consequence of this result, as we note in Section 4, is that under common topological assumptions, dilation is the *only* obstacle for Bayesian agents to reaching a full agreement, *full* in the sense that the sets of probability values representing their credences on the hypothesis of interest are identical. This result opens the door to generating sufficient conditions for reaching full consensus in the generalized Aumman-setting by plugging in sufficient conditions for the absence of dilation in common and important models of imprecise probabilities. As an example, we include a corollary about density ratio classes, which are shown to be dilation-immune by Seidenfeld and Wasserman (1993). This to our knowledge is the first attempt to generalize Aumman’s result to the context of imprecise probability that is concerned with “full agreement”. In Section 5, we provide another perspective on our results and reformulate the theorems in terms of “dilation-averse” agents, who update by full Bayesian conditioning unless the information is about the value of a dilating partition (in which case they ignore the information). For such agents, they are guaranteed to end up agreeing on lower and upper probabilities, and, when lower and upper probabilities are sufficient to identify the set of probabilities, end up fully agreeing.

## 2. A Procedure of Communicating Posteriors

In Geanakoplos and Polemarchakis (1982)’s setup, two agents share a common measurable space  $(\Omega, \mathcal{A})$  and have possibly different information partitions of  $\Omega$ ,  $\mathcal{P}^1$  and  $\mathcal{P}^2$ , which are assumed to be finite. Henceforth we use  $i \in \{1, 2\}$  to index the two agents, and when  $i$  is used in a statement we always intend that the statement is true for both  $i = 1$  and  $i = 2$ . For any

$w \in \Omega$ , let  $\mathcal{P}^i(w)$  denote the member of  $\mathcal{P}^i$  that contains  $w$ ; intuitively,  $\mathcal{P}^i(w)$  represents agent  $i$ 's initial information at state  $w$ . Both the space and the partitions are assumed to be common knowledge, in the standard sense of the term used in game theory: some proposition is common knowledge just in case agent  $i$  knows it, agent  $j$  (where  $j = 3 - i$ ) knows that agent  $i$  knows it, agent  $i$  knows that agent  $j$  knows that agent  $i$  knows it, ... and so on. Let  $\mathcal{P} = \mathcal{P}^1 \wedge \mathcal{P}^2$  be the meet of the two partitions (i.e., the finest common coarsening of  $\mathcal{P}^1$  and  $\mathcal{P}^2$ ). As Aumann (1976, p. 1237) explained, at state  $w$ ,  $\mathcal{P}(w)$  — the member of  $\mathcal{P}$  that contains  $w$  — is the finest event in  $\mathcal{A}$  that is common knowledge: any event that is common knowledge is a superset of  $\mathcal{P}(w)$ . In Geanakoplos and Polemarchakis's setting, common knowledge may grow as the agents communicate their posteriors of a hypothesis. So we call  $\mathcal{P}(w)$  the initial common knowledge and denote it by  $\mathcal{C}_0$ .

Instead of a common precise prior, we assume that the two agents have a common, (possibly) imprecise prior, i.e., a common, non-empty set of priors, denoted by  $\mathbf{Q}$ . Let  $\mathcal{P}^1 \vee \mathcal{P}^2$  denote the join (i.e., the coarsest common refinement) of  $\mathcal{P}^1$  and  $\mathcal{P}^2$ . We assume that every member of  $\mathcal{P}^1 \vee \mathcal{P}^2$  is non-null — in the sense that it receives a positive probability under every probability measure in  $\mathbf{Q}$  — so that all the relevant conditional probabilities are well defined as ratios of unconditional probabilities. Let  $H \in \mathcal{A}$  be a hypothesis of interest. Henceforth by credences or posteriors we mean the agents' credences or posteriors of  $H$  or on the simple algebra generated by the partition  $\{H, \neg H\}$ . We write  $\mathbf{Q}(H)$  to denote the set of prior probabilities of  $H$ , that is,  $\mathbf{Q}(H) = \{p(H) \mid p \in \mathbf{Q}\}$ . Similarly, for any  $E \in \mathcal{A}$  such that  $p(E) > 0$  for every  $p \in \mathbf{Q}$ , let  $\mathbf{Q}(H|E) = \{p(H|E) = p(H \cap E)/p(E) \mid p \in \mathbf{Q}\}$ . Unless otherwise noted (in Section 5), we assume that the agents update their credences by full Bayesian conditioning, where each and every prior in  $\mathbf{Q}$  is updated by conditioning.

Suppose the true state is  $w$ . At step 0, agent  $i$ 's information is  $\mathcal{P}^i(w) \cap \mathcal{C}_0 = \mathcal{P}^i(w)$ . Thus agent  $i$  updates her credence of  $H$  to  $\mathbf{Q}_0^i(H) = \mathbf{Q}(H|\mathcal{P}^i(w))$ . Let  $\mathcal{P}_0^i = \{E \in \mathcal{P}^i \mid E \cap \mathcal{C}_0 \neq \emptyset\}$ , which is the set of those members of  $\mathcal{P}^i$  that are not ruled out by the common knowledge at step 0.

At step 1, the agents announce  $\mathbf{Q}_0^1(H)$  and  $\mathbf{Q}_0^2(H)$ , respectively.<sup>1</sup> Consider  $\mathcal{N}_1^i = \{E \in \mathcal{P}_0^i \mid \mathbf{Q}(H|E) = \mathbf{Q}_0^i(H)\}$ . Intuitively,  $\mathcal{N}_1^i$  is the set of those members of  $\mathcal{P}_0^i$  that are compatible with  $\mathbf{Q}_0^i(H)$ , and the effect of agent  $i$ 's announcement of  $\mathbf{Q}_0^i(H)$  is that it becomes common knowledge that  $\mathcal{P}^i(w) \in \mathcal{N}_1^i$ , or that  $w \in \bigcup \mathcal{N}_1^i$  (where  $\bigcup \mathcal{N}_1^i$  denotes the union of all the sets in  $\mathcal{N}_1^i$ ). Therefore, after the announcements at this step,  $\mathcal{C}_1 = \bigcup \mathcal{N}_1^1 \cap \bigcup \mathcal{N}_1^2$  becomes common knowledge. Let  $\mathcal{P}_1^i = \{E \in \mathcal{N}_1^i \mid E \cap \mathcal{C}_1 \neq \emptyset\}$ , which is the set of those members of  $\mathcal{N}_1^i$  that are not ruled out by the common knowledge at this step. Clearly  $\mathcal{P}_1^i \subseteq \mathcal{N}_1^i \subseteq \mathcal{P}_0^i$  and  $\mathcal{C}_1 = \bigcup \mathcal{P}_1^1 \cap \bigcup \mathcal{P}_1^2$ . Now, if  $\mathcal{P}_1^i = \mathcal{P}_0^i$ , or equivalently, if  $\mathcal{C}_1 = \mathcal{C}_0$ , neither agent learns new information and their credences will stay the same no matter how many more exchanges take place; so the procedure stops. Otherwise, agent  $i$  updates credence of  $H$  to  $\mathbf{Q}_1^i(H) = \mathbf{Q}(H|\mathcal{P}^i(w) \cap \mathcal{C}_1)$ , and enters the next step.

In general, at step  $n + 1$ , the agents announce  $\mathbf{Q}_n^1(H)$  and  $\mathbf{Q}_n^2(H)$ , respectively. Let

$$\begin{aligned} \mathcal{N}_{n+1}^i &= \{E \in \mathcal{P}_n^i \mid \mathbf{Q}(H|E \cap \mathcal{C}_n) = \mathbf{Q}_n^i(H)\} \\ \mathcal{C}_{n+1} &= \bigcup \mathcal{N}_{n+1}^1 \cap \bigcup \mathcal{N}_{n+1}^2 \\ \mathcal{P}_{n+1}^i &= \{E \in \mathcal{N}_{n+1}^i \mid E \cap \mathcal{C}_{n+1} \neq \emptyset\} \end{aligned}$$

<sup>1</sup>In Geanakoplos and Polemarchakis's design, at each step, agent 2 announces her prior *after* agent 1's announcement, already taking into account whatever information is conveyed in agent 1's announcement. This feature is immaterial.

Again,  $\mathcal{N}_{n+1}^i$  is the set of those members of  $\mathcal{P}_n^i$  that are compatible with  $\mathbf{Q}_n^i(H)$ .<sup>2</sup> Hence, after the announcements at this step,  $C_{n+1}$  becomes common knowledge, and  $\mathcal{P}_{n+1}^i$  is the set of those members of  $\mathcal{N}_{n+1}^i$  that are not ruled out by  $C_{n+1}$ . Clearly,  $\mathcal{P}_{n+1}^i \subseteq \mathcal{N}_{n+1}^i \subseteq \mathcal{P}_n^i$  and  $C_{n+1} = \bigcup \mathcal{P}_{n+1}^1 \cap \bigcup \mathcal{P}_{n+1}^2$ . If  $\mathcal{P}_{n+1}^i = \mathcal{P}_n^i$ , or equivalently, if  $C_{n+1} = C_n$ , neither agent learns new information and the procedure stops; otherwise, agent  $i$  updates credence of  $H$  to  $\mathbf{Q}_{n+1}^i(H) = \mathbf{Q}(H|\mathcal{P}^i(w) \cap C_{n+1})$ , and enters the next step.

We will refer to this procedure as the (*Bayesian*) *procedure of communicating posteriors* (of  $H$ ). Obviously, since  $\mathcal{P}^1$  and  $\mathcal{P}^2$  are assumed to be finite, the procedure is guaranteed to stop at step  $m+1$  for some  $m \geq 0$ . Aumann's original setting — where  $\mathbf{Q}_0^1(H)$  and  $\mathbf{Q}_0^2(H)$  are assumed to be common knowledge at step 0 (i.e., it is assumed that  $\mathcal{N}_1^i = \mathcal{P}_0^i$ ) — is a special case in which the procedure stops at step 1. In general, the procedure stops at step  $m+1$  if and only if both  $\mathbf{Q}_m^1(H)$  and  $\mathbf{Q}_m^2(H)$  are already common knowledge at step  $m$  (i.e., before they are announced).

We adapt an example from Geanakoplos and Polemarchakis (1982) to illustrate this procedure.

**Example 2.1.** Let  $\Omega = \{w_1, w_2, w_3, w_4, w_5, w_6, w_7, w_8, w_9\}$  and  $\mathcal{A}$  the power set of  $\Omega$ . Let  $\mathcal{P}^1 = \{\{w_1, w_2, w_3\}, \{w_4, w_5, w_6\}, \{w_7, w_8, w_9\}\}$  and  $\mathcal{P}^2 = \{\{w_1, w_2, w_3, w_4\}, \{w_5, w_6, w_7, w_8\}, \{w_9\}\}$ . Let  $H = \{w_3, w_4\}$ , and suppose the true state of the world is  $w_1$ . For the common set of priors, suppose  $\mathbf{Q}$  is a density ratio class (Seidenfeld and Wasserman, 1993; see also Section 4):

$$\mathbf{Q} = \{(q_1, q_2, q_3, q_4, q_5, q_6, q_7, q_8, q_9) \mid \sum_{1 \leq j \leq 9} q_j = 1, \text{ and } \frac{1}{2} \leq \frac{q_k}{q_l} \leq 2, 1 \leq k, l \leq 9.\}$$

It is easy to calculate that the lower probability of  $H$  is:  $\underline{\mathbf{Q}}(H) = \inf_{p \in \mathbf{Q}} p(H) = 1/8$  (obtained at  $(1/8, 1/8, 1/16, 1/16, 1/8, 1/8, 1/8, 1/8, 1/8)$ ), and the upper probability of  $H$  is:  $\overline{\mathbf{Q}}(H) = \sup_{p \in \mathbf{Q}} p(H) = 4/11$  (obtained at  $(1/11, 1/11, 2/11, 2/11, 1/11, 1/11, 1/11, 1/11, 1/11)$ ). Since  $\mathbf{Q}$  is closed and connected,  $\mathbf{Q}(H) = [1/8, 4/11]$ .

Suppose the two agents in this example carry out the procedure of communicating posteriors. Here is a summary of the execution:

- Step 0  $C_0 = (\mathcal{P}^1 \wedge \mathcal{P}^2)(w_1) = \Omega$ . Agent 1 learns the private information  $\mathcal{P}^1(w_1) = \{w_1, w_2, w_3\}$  and updates her credence on  $H$  to  $\mathbf{Q}_0^1(H) = \mathbf{Q}(H|\mathcal{P}^1(w_1)) = [1/5, 1/2]$ ; agent 2 learns the private information  $\mathcal{P}^2(w_1) = \{w_1, w_2, w_3, w_4\}$ , and updates her credence on  $H$  to  $\mathbf{Q}_0^2(H) = \mathbf{Q}(H|\mathcal{P}^2(w_1)) = [1/3, 2/3]$ .
- Step 1 agent 1 announces  $\mathbf{Q}_0^1(H)$ , and so it becomes common knowledge that  $\mathcal{P}^1(w_1) \in \mathcal{N}_1^1 = \{\{w_1, w_2, w_3\}, \{w_4, w_5, w_6\}\}$  (for  $\mathbf{Q}(H|\{w_7, w_8, w_9\}) = \{0\} \neq \mathbf{Q}_0^1(H)$ .) And agent 2 announces  $\mathbf{Q}_0^2(H)$ , and so it becomes common knowledge that  $\mathcal{P}^2(w_1) \in \mathcal{N}_1^2 = \{\{w_1, w_2, w_3, w_4\}\}$  (for  $\mathbf{Q}(H|\{w_5, w_6, w_7, w_8\}) = \mathbf{Q}(H|\{w_9\}) = \{0\} \neq \mathbf{Q}_0^2(H)$ .) Thus  $C_1 = \bigcup \mathcal{N}_1^1 \cap \bigcup \mathcal{N}_1^2 = \{w_1, w_2, w_3, w_4\}$ , and  $\mathcal{P}_1^i = \mathcal{N}_1^i$ . Since  $C_1 \neq C_0$ , agent 1 updates her credence to  $\mathbf{Q}_1^1(H) = \mathbf{Q}(H|\mathcal{P}^1(w_1) \cap C_1) = [1/5, 1/2]$ , and agent 2 updates her credence to  $\mathbf{Q}_1^2(H) = \mathbf{Q}(H|\mathcal{P}^2(w_1) \cap C_1) = [1/3, 2/3]$ .<sup>3</sup>
- Step 2 agent 1 announces  $\mathbf{Q}_1^1(H)$ , and so it becomes common knowledge that  $\mathcal{P}^1(w_1) \in \mathcal{N}_2^1 = \{\{w_1, w_2, w_3\}\}$  (for  $\mathbf{Q}(H|\{w_4, w_5, w_6\} \cap C_1) = \{1\} \neq \mathbf{Q}_1^1(H)$ .) And agent 2 announces

<sup>2</sup>Note that the definition of  $\mathcal{N}_{n+1}^i$  also applies to  $n = 0$ , as for every  $E \in \mathcal{P}_0^i$ ,  $E \cap C_0 = E$ .

<sup>3</sup>Although  $\mathbf{Q}_1^i(H) = \mathbf{Q}_0^i(H)$ , the procedure goes on, because some agent still acquires new information in this step.

$\mathbf{Q}_1^2(H)$ , but that obviously does not change anything, and  $\mathcal{N}_2^2 = \mathcal{P}_1^2$ . So  $C_2 = \{w_1, w_2, w_3\}$ , and  $\mathcal{P}_2^i = \mathcal{N}_2^i$ . Since  $C_2 \neq C_1$ , agent 1 updates credence to  $\mathbf{Q}_2^1(H) = \mathbf{Q}(H|\mathcal{P}^1(w_1) \cap C_2) = [1/5, 1/2]$ , and agent 2 updates credence to  $\mathbf{Q}_2^2(H) = \mathbf{Q}(H|\mathcal{P}^2(w_1) \cap C_2) = [1/5, 1/2]$ .

Step 3 the agents announce  $\mathbf{Q}_2^1(H)$  and  $\mathbf{Q}_2^2(H)$ , but they do not convey any new information:  $C_3 = C_2$  (or equivalently,  $\mathcal{P}_3^i = \mathcal{P}_2^i$ ). The procedure stops.

In this example, the communication ends up making each agent's private information public. This is not always the case, as later examples will illustrate. When (at least one agent's) private information remains private, it is in general possible to agree to disagree. However, in the case of a precise prior, that is, if  $\mathbf{Q} = \{\tilde{p}\}$  is a singleton, Geanakoplos and Polemarchakis (1982, Proposition 1) showed that when the procedure stops at step  $m + 1$ , it is necessarily the case that  $\mathbf{Q}_m^1(H) = \mathbf{Q}_m^2(H)$ . We present a version of the argument here that will facilitate our subsequent presentation. Suppose the procedure stops at step  $m + 1$ . It means that  $\mathcal{P}_{m+1}^i = \mathcal{P}_m^i$  (for both  $i = 1, 2$ , as we always intend). This entails, by the definition of  $\mathcal{P}_{m+1}^i$ , that

$$\forall E \in \mathcal{P}_m^i, \mathbf{Q}(H|E \cap C_m) = \mathbf{Q}_m^i(H) = \mathbf{Q}(H|\mathcal{P}^i(w) \cap C_m) \quad (1)$$

Since  $\mathbf{Q} = \{\tilde{p}\}$ ,  $\mathbf{Q}(H|E \cap C_m) = \{\tilde{p}(H|E \cap C_m)\}$  and  $\mathbf{Q}_m^i(H) = \mathbf{Q}(H|\mathcal{P}^i(w) \cap C_m) = \{\tilde{p}(H|\mathcal{P}^i(w) \cap C_m)\}$ . It follows that

$$\forall E \in \mathcal{P}_m^i, \tilde{p}(H|E \cap C_m) = \tilde{p}(H|\mathcal{P}^i(w) \cap C_m) \quad (2)$$

Since all members of  $\mathcal{P}_m^i$  are mutually disjoint, (2) entails that

$$\tilde{p}(H|\bigcup \mathcal{P}_m^i \cap C_m) = \tilde{p}(H|\mathcal{P}^i(w) \cap C_m) \quad (3)$$

Recall that  $C_m = \bigcup \mathcal{P}_m^1 \cap \bigcup \mathcal{P}_m^2$ . Hence  $\bigcup \mathcal{P}_m^i \cap C_m = C_m$ . It then follows from (3) that

$$\tilde{p}(H|\mathcal{P}^i(w) \cap C_m) = \tilde{p}(H|C_m) \quad (4)$$

Therefore,  $\tilde{p}(H|\mathcal{P}^1(w) \cap C_m) = \tilde{p}(H|\mathcal{P}^2(w) \cap C_m)$ ; that is, the two agents end up agreeing.

Two comments are in order. First, equation (4) shows that the two agents are driven to the same posterior because when the communication stops, the resulting common knowledge ( $C_m$ ) renders each agent's private information ( $\mathcal{P}^i(w)$ ) irrelevant to  $H$  (even if  $C_m \neq \mathcal{P}^1(w) \cap \mathcal{P}^2(w)$ ). However, it does not follow that  $\mathcal{P}^1(w)$  and  $\mathcal{P}^2(w)$  are jointly irrelevant to  $H$  given  $C_m$ . As Geanakoplos and Polemarchakis (1982, Proposition 3) observed, the consensus reached via the procedure of communicating posteriors can be different from the consensus that would result from directly exchanging private information. Clearly, they are different if and only if  $\mathcal{P}^1(w)$  and  $\mathcal{P}^2(w)$  are jointly relevant to  $H$  given  $C_m$ , even though each on its own is guaranteed to be irrelevant given  $C_m$  (see Example 6.1 in Section 6).

Second, and more importantly for the purpose of this paper, a crucial step in the above argument is the move from (2) to (3), where what is needed is the following fact: if all members of a (finite) set of events  $\mathcal{E}$  are mutually disjoint, and for every  $E \in \mathcal{E}$ ,  $\tilde{p}(H|E) = q$  (where  $q$  is a real number), then  $\tilde{p}(H|\bigcup \mathcal{E}) = q$ . An analogous condition for imprecise probabilities would be the following: if all members of a (finite) set of events  $\mathcal{E}$  are mutually disjoint, and for every  $E \in \mathcal{E}$ ,  $\mathbf{Q}(H|E) = Q$  (where  $Q$  is a set of real numbers), then  $\mathbf{Q}(H|\bigcup \mathcal{E}) = Q$ . This condition does not hold in general for sets of probabilities.<sup>4</sup>

<sup>4</sup>Even in the case of precise probability, it is well known that this condition, as a special case of conglomerability,

### 3. Dilation and Agreeing to Disagree on Lower and Upper Probabilities

We borrow a simple example from Carvajal and Correia-da-Silva (2010) to illustrate the failure of the said condition for sets of probabilities.

**Example 3.1.** Let  $\Omega = \{w_1, w_2, w_3, w_4\}$  and  $\mathcal{A}$  the power set of  $\Omega$ . Let  $\mathcal{P}^1 = \{\{w_1, w_2\}, \{w_3, w_4\}\}$  and  $\mathcal{P}^2 = \{\Omega\}$ . Suppose  $\mathbf{Q} = \{(1/2, 0, 1/2, 0), (0, 1/2, 0, 1/2)\}$ ; that is, the common set of priors consists of just two probability measures, represented by the two probability vectors.<sup>5</sup> Let  $H = \{w_2, w_3\}$ , and suppose the true state of the world is  $w_1$ .

This is an Aumann case in that the agents' posteriors on  $H$  are already common knowledge at the beginning; the procedure of communicating posteriors stops at step 1, for  $C_1 = C_0 = \Omega$ . However,  $\mathbf{Q}_0^1(H) = \{0, 1\}$  and  $\mathbf{Q}_0^2(H) = \{1/2\}$ . Not only are the sets non-identical, they are in full disagreement in the sense that they do not even intersect and have different lower and upper probabilities. The agents agree to fully disagree. The condition we highlighted at the end of Section 2 fails dramatically in this case for agent 1's partition  $\mathcal{P}_1^1$  (which is identical to  $\mathcal{P}^1$  in this case):  $\mathbf{Q}(H|\{w_1, w_2\}) = \mathbf{Q}(H|\{w_3, w_4\}) = \{0, 1\}$ , while  $\mathbf{Q}(H|\{w_1, w_2, w_3, w_4\}) = \{1/2\}$ .

This dramatic failure of the condition is known as dilation (Good, 1974; Seidenfeld, 1981; Walley, 1991; Seidenfeld and Wasserman, 1993; Herron et al., 1997). No matter which member of  $\mathcal{P}_1^1$  is the case, the resulting conditional probability is less precise than the probability conditional on  $\bigcup \mathcal{P}_1^1$ . Given a non-empty set of probabilities  $\mathbf{R}$ , let  $\underline{\mathbf{R}}(A|E) = \inf_{p \in \mathbf{R}} p(A|E)$  denote the lower probability of  $A$  conditional on  $E$ , and  $\overline{\mathbf{R}}(A|E) = \sup_{p \in \mathbf{R}} p(A|E)$  denote the upper probability of  $A$  conditional on  $E$ . Here is a definition of dilation that suits the present purpose.

**Definition 3.2 (Dilation).** Let  $\mathbf{R}$  be a non-empty set of probability measures on  $(\Omega, \mathcal{A})$ . Let  $\mathcal{E}$  be a finite, non-empty set of mutually disjoint events.  $\mathcal{E}$  is said to dilate an event  $A$  with respect to  $\mathbf{R}$  (or  $\mathbf{R}(\bullet|\bigcup \mathcal{E})$ ) if for every  $E \in \mathcal{E}$ , the interval  $[\underline{\mathbf{R}}(A|E), \overline{\mathbf{R}}(A|E)]$  strictly contains the interval  $[\underline{\mathbf{R}}(A|\bigcup \mathcal{E}), \overline{\mathbf{R}}(A|\bigcup \mathcal{E})]$ .

This is a slight generalization of the standard definition of dilation (Seidenfeld and Wasserman, 1993, p. 1141)<sup>6</sup>, for it considers dilation in a subspace  $\bigcup \mathcal{E}$  (the definition reduces to the standard one when  $\bigcup \mathcal{E} = \Omega$ ), but the idea and the anomalous feature are exactly the same. Again, in example 3.1,  $\mathcal{P}_1^1$ , which happens to be the same as  $\{E \cap C_1 \mid E \in \mathcal{P}_1^1\}$ , dilates the hypothesis of interest with respect to the given prior. This is not a coincidence, as Theorem 3.4 below shows. It is a simple consequence of the following lemma, which is a straightforward generalization of Lemma 1 in Carvajal and Correia-da-Silva (2010; also see Kajii and Ui, 2005, Proposition 3).

---

can fail for finitely but not countably additive probability measures (de Finetti, 1972; Schervish et al., 1984; Hill and Lane, 1985). This does not matter in the setup we are considering, for the partitions are assumed to be finite. However, the original setup in Aumann (1976) seems to allow denumerable infinite partitions, in which case Aumann's result does not necessarily hold for merely finitely additive probabilities. More generally, Schervish et al. (2016) showed that conglomerability can fail in a partition of cardinality  $\kappa$  for a probability measure that is not  $\kappa$ -additive. Thus, if uncountable partitions are allowed, Aumann's result may fail even for countably additive measures.

<sup>5</sup>In case readers are concerned that the two probabilities are not positive and are mutually singular, these special features are not essential. We can also use  $\mathbf{Q} = \{(1/2 - \epsilon, \epsilon, 1/2 - \epsilon, \epsilon), (\epsilon, 1/2 - \epsilon, \epsilon, 1/2 - \epsilon)\}$ ,  $0 < \epsilon < 1/4$ , to make the same point.

<sup>6</sup>Herron et al. (1997, p. 412) gave a weaker definition of dilation, requiring only that all conditional intervals contain and some of them *strictly* contain the unconditional interval. This definition would work equally well for our purpose. We thank an anonymous referee for this point.

**Lemma 3.3.** *Suppose the procedure of communicating posteriors stops at step  $m + 1$ . Then*

$$\mathbf{Q}(H|C_m) \subseteq [\underline{\mathbf{Q}}(H|\mathcal{P}^i(w) \cap C_m), \overline{\mathbf{Q}}(H|\mathcal{P}^i(w) \cap C_m)]$$

for both  $i = 1, 2$ .

*Proof.* As already mentioned, when the procedure stops at step  $m + 1$ , we have equation (1), namely,

$$\forall E \in \mathcal{P}_m^i, \mathbf{Q}(H|E \cap C_m) = \mathbf{Q}(H|\mathcal{P}^i(w) \cap C_m)$$

Consider  $i = 1$  first. Let  $\mathcal{P}_m^1 = \{E_1, \dots, E_k\}$ . Notice that  $\{E_1 \cap C_m, \dots, E_k \cap C_m\}$  forms a partition of  $\bigcup \mathcal{P}_m^1 \cap C_m$ . Hence, for every  $p \in \mathbf{Q}$ , by the law of total probability

$$p(H|C_m) = p(H|\bigcup \mathcal{P}_m^1 \cap C_m) = \sum_{1 \leq j \leq k} p(H|E_j \cap C_m)p(E_j \cap C_m|\bigcup \mathcal{P}_m^1 \cap C_m) \quad (5)$$

Given equation (1), we have that for every  $1 \leq j \leq k$ ,  $p(H|E_j \cap C_m) \in \mathbf{Q}(H|E_j \cap C_m) = \mathbf{Q}(H|\mathcal{P}^1(w) \cap C_m)$ . It follows that for every  $1 \leq j \leq k$ ,

$$\underline{\mathbf{Q}}(H|\mathcal{P}^1(w) \cap C_m) \leq p(H|E_j \cap C_m) \leq \overline{\mathbf{Q}}(H|\mathcal{P}^1(w) \cap C_m) \quad (6)$$

(5) and (6) together entail that

$$p(H|C_m) \geq \underline{\mathbf{Q}}(H|\mathcal{P}^1(w) \cap C_m) \sum_{1 \leq j \leq k} p(E_j \cap C_m|\bigcup \mathcal{P}_m^1 \cap C_m) = \underline{\mathbf{Q}}(H|\mathcal{P}^1(w) \cap C_m) \quad (7)$$

and

$$p(H|C_m) \leq \overline{\mathbf{Q}}(H|\mathcal{P}^1(w) \cap C_m) \sum_{1 \leq j \leq k} p(E_j \cap C_m|\bigcup \mathcal{P}_m^1 \cap C_m) = \overline{\mathbf{Q}}(H|\mathcal{P}^1(w) \cap C_m) \quad (8)$$

Since (7) and (8) hold for every  $p \in \mathbf{Q}$ , the desired conclusion is established for  $i = 1$ . The case of  $i = 2$  is of course entirely parallel.  $\square$

Lemma 3.3 shows that although equation (1) does not entail that  $\mathbf{Q}(H|C_m) = \mathbf{Q}(H|\mathcal{P}^i(w) \cap C_m)$ , it does entail that  $\mathbf{Q}(H|C_m)$  is bounded by the infimum and supremum  $\mathbf{Q}(H|\mathcal{P}^i(w) \cap C_m)$ . The following theorem is then immediate.

**Theorem 3.4.** *Suppose the procedure of communicating posteriors stops at step  $m + 1$ . If for both  $i = 1, 2$ ,  $\{E \cap C_m \mid E \in \mathcal{P}_m^i\}$  does not dilate  $H$ , then  $\underline{\mathbf{Q}}(H|\mathcal{P}^1(w) \cap C_m) = \underline{\mathbf{Q}}(H|\mathcal{P}^2(w) \cap C_m)$  and  $\overline{\mathbf{Q}}(H|\mathcal{P}^1(w) \cap C_m) = \overline{\mathbf{Q}}(H|\mathcal{P}^2(w) \cap C_m)$ .*

*Proof.* Lemma 3.3 entails that for both  $i = 1, 2$ ,

$$\underline{\mathbf{Q}}(H|\mathcal{P}^i(w) \cap C_m) \leq \underline{\mathbf{Q}}(H|C_m), \overline{\mathbf{Q}}(H|C_m) \leq \overline{\mathbf{Q}}(H|\mathcal{P}^i(w) \cap C_m).$$

Since  $C_m = \bigcup \mathcal{P}_m^i \cap C_m$ , if either of the inequality is strict, then  $\{E \cap C_m \mid E \in \mathcal{P}_m^i\}$  dilates  $H$ , because of equation (1). Therefore, if  $\{E \cap C_m \mid E \in \mathcal{P}_m^i\}$  does not dilate  $H$ , then

$$\underline{\mathbf{Q}}(H|\mathcal{P}^i(w) \cap C_m) = \underline{\mathbf{Q}}(H|C_m), \overline{\mathbf{Q}}(H|C_m) = \overline{\mathbf{Q}}(H|\mathcal{P}^i(w) \cap C_m).$$

The desired conclusion follows.  $\square$

Thus, the two agents can agree to disagree on the lower or upper probability of a hypothesis only if a certain dilation takes place. Without dilation, the two agents are guaranteed to reach consensus on lower and upper probabilities by communicating posteriors.

It is worth noting that for the argument for Theorem 3.4 to go through, it is not necessary to require the agents to communicate their sets of posteriors. It is sufficient to ask them to communicate lower and upper probabilities at each step. Consider the procedure of communicating lower and upper posteriors: at step  $n + 1$ , agent  $i$  announces  $\underline{\mathbf{Q}}_n^i(H)$  and  $\overline{\mathbf{Q}}_n^i(H)$ . Let

$$\begin{aligned}\mathcal{N}_{n+1}^{i\dagger} &= \{E \in \mathcal{P}_n^{i\dagger} \mid \underline{\mathbf{Q}}(H|E \cap C_n) = \underline{\mathbf{Q}}_n^i(H) \text{ and } \overline{\mathbf{Q}}(H|E \cap C_n) = \overline{\mathbf{Q}}_n^i(H)\} \\ C_{n+1}^\dagger &= \bigcup \mathcal{N}_{n+1}^{1\dagger} \cap \bigcup \mathcal{N}_{n+1}^{2\dagger} \\ \mathcal{P}_{n+1}^{i\dagger} &= \{E \in \mathcal{N}_{n+1}^{i\dagger} \mid E \cap C_{n+1}^\dagger \neq \emptyset\}\end{aligned}$$

If  $\mathcal{P}_{n+1}^{i\dagger} = \mathcal{P}_n^{i\dagger}$ , or equivalently, if  $C_{n+1}^\dagger = C_n^\dagger$ , the procedure stops; otherwise, agent  $i$  updates credence to  $\mathbf{Q}_{n+1}^i(H) = \mathbf{Q}(H|\mathcal{P}^i(w) \cap C_{n+1}^\dagger)$ , and enters the next step.

As before, this modified procedure is guaranteed to stop at step  $m+1$  for some  $m \geq 0$ , because  $\mathcal{P}^1$  and  $\mathcal{P}^2$  are assumed to be finite. The version of Lemma 3.3 on this procedure remains valid, for equation (1) is not necessary for the argument. All that is needed is the weaker condition that

$$\forall E \in \mathcal{P}_m^i, \underline{\mathbf{Q}}(H|E \cap C_m) = \underline{\mathbf{Q}}(H|\mathcal{P}^i(w) \cap C_m) \text{ and } \overline{\mathbf{Q}}(H|E \cap C_m) = \overline{\mathbf{Q}}(H|\mathcal{P}^i(w) \cap C_m)$$

This weaker condition obviously remains true when  $\mathcal{P}_m^i$  is replaced by  $\mathcal{P}_m^{i\dagger}$  and  $C_m$  by  $C_m^\dagger$ . So Lemma 3.3 still holds for the procedure of communicating lower and upper posteriors. Hence we also have the following variant of Theorem 3.4.

**Theorem 3.5.** *Suppose the procedure of communicating lower and upper posteriors stops at step  $m + 1$ . If for both  $i = 1, 2$ ,  $\{E \cap C_m^\dagger \mid E \in \mathcal{P}_m^{i\dagger}\}$  does not dilate  $H$ , then  $\underline{\mathbf{Q}}(H|\mathcal{P}^1(w) \cap C_m^\dagger) = \underline{\mathbf{Q}}(H|\mathcal{P}^2(w) \cap C_m^\dagger)$  and  $\overline{\mathbf{Q}}(H|\mathcal{P}^1(w) \cap C_m^\dagger) = \overline{\mathbf{Q}}(H|\mathcal{P}^2(w) \cap C_m^\dagger)$ .*

*Proof.* Extremely similar to that of Theorem 3.4.  $\square$

Although both procedures result in consensus on lower and upper probabilities in the absence of dilation, in general the agreements they lead to may well be different, for in general communicating lower and upper posteriors conveys less information than communicating the full sets of posteriors.

#### 4. More Agreement Results

Under some common assumptions, however, lower and upper probabilities are sufficient to identify the full set, in which case the two procedures are equivalent and, more importantly, the consensus reached in the absence of dilation will be full consensus. For example, if we follow Carvajal and Correia-da-Silva (2010) to assume that the set of priors is closed and connected (or follow Kajii and Ui (2005) to assume that the set of posteriors is a closed interval), we get the following result.

**Theorem 4.1.** *Suppose  $\mathbf{Q}$  is closed and connected (with respect to the total variation topology), and suppose the procedure of communicating posteriors stops at step  $m + 1$ . If for both  $i = 1, 2$ ,  $\{E \cap C_m \mid E \in \mathcal{P}_m^i\}$  does not dilate  $H$ , then  $\mathbf{Q}(H|\mathcal{P}^1(w) \cap C_m) = \mathbf{Q}(H|\mathcal{P}^2(w) \cap C_m)$ .*

*Proof.* Given the assumption that all the relevant conditional probabilities are well defined as ratios of unconditional probabilities, the mapping from  $\mathbf{Q}$  to  $\mathbf{Q}(H|\mathcal{P}^i(w) \cap C_j)$  is continuous for every  $0 \leq j \leq m$  and for both  $i = 1, 2$ . Hence, since  $\mathbf{Q}$  is assumed to be closed and connected,  $\mathbf{Q}(H|\mathcal{P}^i(w) \cap C_m)$  is a closed interval. Thus  $\mathbf{Q}(H|\mathcal{P}^i(w) \cap C_m)$  is identified by  $\overline{\mathbf{Q}(H|\mathcal{P}^i(w) \cap C_m)}$  and  $\underline{\mathbf{Q}(H|\mathcal{P}^i(w) \cap C_m)}$ . Then Theorem 3.4 entails that  $\mathbf{Q}(H|\mathcal{P}^1(w) \cap C_m) = \mathbf{Q}(H|\mathcal{P}^2(w) \cap C_m)$ .  $\square$

To our knowledge, Theorem 4.1 is the first attempt to formulate a generalization of Aumann's agreement theorem in the context of imprecise probability that takes agreement to mean full agreement (identical set of posteriors) rather than partial agreement (intersecting sets of posteriors). In addition to revealing a connection to the important phenomenon of dilation, Theorem 4.1 may also be used to generate sufficient conditions for guaranteed full agreement via communicating posteriors, for important models of imprecise probability, if sufficient conditions for the absence of dilation in those models are known. As a simple example, consider the density ratio classes for finite spaces (Wasserman, 1992; Seidenfeld and Wasserman, 1993).

**Definition 4.2** (Density Ratio Prior). Let  $\Omega = \{w_1, \dots, w_n\}$  and  $\mathcal{A}$  the power set of  $\Omega$ . A density ratio prior is defined by

$$\mathbf{D}_{p,k} = \{(q_1, \dots, q_n) \mid \sum_{1 \leq j \leq n} q_j = 1 \text{ and } \frac{q_h}{q_j} \leq k \frac{p_h}{p_j}, \forall 1 \leq h, j \leq n\}$$

where  $k \geq 1$  and  $(p_1, \dots, p_n)$  is a probability vector such that  $p_j > 0$  for all  $1 \leq j \leq n$ .

For instance, Example 2.1 in Section 2 employs a density ratio prior, where  $p$  is the uniform distribution over the 9-atom algebra and  $k = 2$ .

**Corollary 4.3.** *If two agents start with a common density ratio prior and carry out the procedure of communicating posteriors, they are guaranteed to reach the same set of posteriors.*

*Proof.* Seidenfeld and Wasserman (1993, Theorem 4.1) showed that the density ratio priors are dilation-immune in the sense that no finite partition of the sample space dilates any hypothesis. Note also that if  $\mathbf{D}$  is a density ratio prior on  $(\Omega, \mathcal{A})$ , then for every  $E \in \mathcal{A}$ ,  $\mathbf{D}(\bullet|E)$  remains a density ratio prior on the space restricted to  $E$ , which follows easily from Definition 4.2. Moreover, a density ratio prior is obviously closed and connected. Then Theorem 4.1 entails the desired conclusion.  $\square$

Another simple consequence of Theorem 3.4 is that if we just consider partial agreement, in the sense of a non-empty intersection of sets of posteriors (Kajii and Ui, 2005, 2009; Carvajal and Correia-da-Silva, 2010), we can drop the assumption of connectedness in Theorem 4.1.

**Theorem 4.4.** *Suppose  $\mathbf{Q}$  is closed, and suppose the procedure of communicating posteriors stops at step  $m+1$ . If for both  $i = 1, 2$ ,  $\{E \cap C_m \mid E \in \mathcal{P}_m^i\}$  does not dilate  $H$ , then  $\mathbf{Q}(H|\mathcal{P}^1(w) \cap C_m) \cap \mathbf{Q}(H|\mathcal{P}^2(w) \cap C_m) \neq \emptyset$ .*

*Proof.* Since  $\mathbf{Q}$  is closed,  $\mathbf{Q}(H|\mathcal{P}^i(w) \cap C_m)$  is also closed, for the mapping from  $\mathbf{Q}$  to  $\mathbf{Q}(H|\mathcal{P}^i(w) \cap C_m)$  is continuous. Thus,  $\mathbf{Q}(H|\mathcal{P}^i(w) \cap C_m)$  contains its infimum and supremum. It then follows from Theorem 3.4 that  $\mathbf{Q}(H|\mathcal{P}^1(w) \cap C_m) \cap \mathbf{Q}(H|\mathcal{P}^2(w) \cap C_m) \neq \emptyset$ .  $\square$

In general the topological assumptions in Theorem 4.1 and Theorem 4.4 are not redundant. Consider the following variation of Example 3.1:

**Example 4.5.** The setup is the same as Example 3.1, except for the common prior  $\mathbf{Q}$ . Let  $\mathbb{Z}$  denote the set of all integers. Let  $\langle \alpha_z \mid 0 < \alpha_z < 1, z \in \mathbb{Z} \rangle$  be an increasing sequence of rational numbers (in the open interval  $(0, 1)$ ), such that  $\lim_{z \rightarrow -\infty} \alpha_z = 0$  and  $\lim_{z \rightarrow \infty} \alpha_z = 1$ . Let  $\langle \gamma_z \rangle$  be a sequence of positive irrational numbers such that  $\gamma_z < \min(\frac{\alpha_z}{2}, \frac{1-\alpha_z}{2})$ . Define  $\mathbf{Q} = \mathbf{Q}_1 \cup \mathbf{Q}_2$ , where

$$\begin{aligned}\mathbf{Q}_1 &= \left\{ \left( \frac{1-\alpha_z}{2} - \gamma_z, \frac{\alpha_z}{2} + \gamma_z, \frac{\alpha_z}{2} - \gamma_z, \frac{1-\alpha_z}{2} + \gamma_z \right) \mid z \in \mathbb{Z} \right\} \\ \mathbf{Q}_2 &= \left\{ \left( \frac{1-\alpha_z}{2} + \gamma_z, \frac{\alpha_z}{2} - \gamma_z, \frac{\alpha_z}{2} + \gamma_z, \frac{1-\alpha_z}{2} - \gamma_z \right) \mid z \in \mathbb{Z} \right\}\end{aligned}$$

In this example,  $\mathbf{Q}$  is neither connected nor closed. It is easy to see that

$$\mathbf{Q}(H|\{w_1, w_2\}) = \mathbf{Q}(H|\{w_3, w_4\}) = \{\alpha_z + 2\gamma_z \mid z \in \mathbb{Z}\} \cup \{\alpha_z - 2\gamma_z \mid z \in \mathbb{Z}\}$$

Thus it remains an Aumann case in which both agents' posteriors are already common knowledge before they are announced. Notice, however, that dilation does not occur here, for

$$\mathbf{Q}(H|\Omega) = \{\alpha_z \mid z \in \mathbb{Z}\}$$

which has the same infimum (i.e., 0) and supremum (i.e., 1) as  $\mathbf{Q}(H|\{w_1, w_2\})$  and  $\mathbf{Q}(H|\{w_3, w_4\})$ . As a result, the two agents agree on the posterior lower probability and upper probability. However,  $\mathbf{Q}_0^1 (= \mathbf{Q}(H|\{w_1, w_2\}))$  and  $\mathbf{Q}_0^2 (= \mathbf{Q}(H|\Omega))$  are clearly disjoint: the former is a set of irrational numbers and the latter is a set of rational numbers. Hence, without the condition of closure, Theorem 4.4 does not hold in general. If we make  $\mathbf{Q}$  in this example closed by adding the two limit points  $(0, 1/2, 1/2, 0)$  and  $(1/2, 0, 0, 1/2)$ , then the two sets of posteriors intersect (at the limit points), but of course they are far from identical. Therefore, in Theorem 4.1, the condition of connectedness is also needed.

## 5. Dilation-Averse Updating

The presence of dilation may alarm some agents, who may be inclined to think that they are permitted or even rationally required to ignore information that would increase the ambiguity of their credences. One of the best known examples of this kind is Kyburg's (1961, 1974) interval-valued theory of *Epistemological Probability*, which precludes the possibility of dilation, primarily because of its *Strength rule* for determining which is the relevant reference set for using information about an interval of frequencies also as the probability interval. Informally, the Strength rule requires that we ignore added information whenever that added information generates a narrower reference set with strictly less informative statistical, frequency information. More specifically, suppose that the knowledge base for Epistemological Probability (EP) includes the following information as the most informative frequency information about the event  $\mathbf{H}(a)$ : individual  $a$  has the property  $\mathbf{H}$ .

- Reference set  $R_1$  is a proper subset of reference set  $R_2$ , i.e.,  $R_1 \subset R_2$ .
- Individual  $a$  belongs to  $R_1$ ; hence,  $R_1$  is a strictly narrower, i.e. logically stronger reference set for  $a$  than is  $R_2$ .
- The relative frequency of  $\mathbf{H}$  in  $R_1$  falls in the interval  $[l_1, u_1]$ . The relative frequency of  $\mathbf{H}$  in  $R_2$  falls in the interval  $[l_2, u_2]$ . The interval  $[l_2, u_2]$  is strictly contained in the interval  $[l_1, u_1]$ .

Then with respect to this knowledge base, the Strength rule requires that  $EP(\mathbf{H}(a)) = [l_2, u_2]$ . Whatever additional information responsible for the difference between  $R_1$  and  $R_2$  is ignored.

The Strength rule is clearly analogous to an updating rule that would avoid dilation by sometimes ignoring information. However, Levi (1977) noted that it conflicts with basic Bayesian tenets even when the theory delivers precise values. Here is Levi's example.

**Example 5.1.** Suppose we know that Petersen (denoted  $a$ ) is a Swedish resident of Malmo. Let  $\mathbf{H}$  be the property that a person is Protestant. The corpus of knowledge  $K$  includes the following frequency facts about the two competing reference sets: Swedes, and residents of Malmo.

- 90% of Swedes are Protestants.
- However, regarding residents of Malmo, the frequency information is not precise. Either  $A_1$ : 85% of Malmo's residents are Protestants; or  $A_2$ : 91% of Malmo's residents are Protestants; or  $A_3$ : 95% of Malmo's residents are Protestants.

In this example, with respect to corpus  $K$ , Kyburg's theory entails that  $EP(\mathbf{H}(a)) = [0.9, 0.9]$ , i.e., a precise value, for the Strength rule fixes the logically weaker but statistically more informative reference set (Swedes) over the rival reference set of Malmo residents for delivering EP. Similarly, Kyburg's theory also entails each of the following *precise* probability statements with respect to expansions of  $K$  that result by adding some Boolean combinations of the three hypotheses:

$$\begin{aligned}
 EP(\mathbf{H}(a); K + A_1) &= [0.85, 0.85] \\
 EP(\mathbf{H}(a); K + A_2) &= [0.91, 0.91] \\
 EP(\mathbf{H}(a); K + A_3) &= [0.95, 0.95] \\
 EP(\mathbf{H}(a); K + (A_1 \vee A_3)) &= [0.9, 0.9] \\
 EP(\mathbf{H}(a); K + (A_1 \vee A_2)) &= [0.9, 0.9] \\
 EP(\mathbf{H}(a); K + (A_1 \vee A_2 \vee A_3)) &= [0.9, 0.9]
 \end{aligned}$$

Each of the last three of these six EP statements results by an application of the Strength rule, which picks the statistically more informative and logically weaker reference set (Swedes) for determining the EP that Petersen is a Protestant.

However, these six precise EP values are not consistent with a precise probability function in the sense that there is no "prior" distribution over these three simple statistical hypotheses  $\{A_1, A_2, A_3\}$  that satisfies all six precise EP values. In other words, the EP theory does not follow the Bayesian law that there exists a "prior" against which one may average the "likelihood" function, even when it delivers precise values!

Evidently, Levi's construction can be generalized to any inductive theory of interval-value probability that avoids dilation by appeal to a Strength-like rule — where priority is assigned to informativeness of the probability interval over the requirement of total evidence (see also Seidenfeld, 2007). Doing so conflicts with basic Bayesian theory.

On the other hand, Grünwald and Halpern (2004) showed that once non-Bayesian criteria such as minimax of loss are used, optimal decision rules often require ignoring information about the value of a dilating partition. It is fair to say that whether dilating information can or should be ignored remains a matter of debate.

We do not pretend to resolve the debate here, but we would like to reformulate the main ideas of this paper in terms of dilation-averse agents, which may provide a useful perspective to think

about our results. An agent participating in the communication procedure is said to be dilation-averse if she does not condition on her information about the value of a partition that dilates the hypothesis of interest, but is otherwise happy to update by conditioning.<sup>7</sup> Suppose it is common knowledge that the two agents who communicate posteriors are dilation-averse. To model this situation, the procedure of communicating posteriors should be modified as follows.

At step 0, for each agent  $i$ , if  $\mathcal{P}_0^i = \{E \in \mathcal{P}^i \mid E \cap C_0 \neq \emptyset\}$  dilates  $H$ , she updates her credence by conditioning on the common knowledge  $C_0$ :  $\mathbf{Q}_0^i(H) = \mathbf{Q}(H|C_0)$ ; otherwise, she updates her credence in the standard way:  $\mathbf{Q}_0^i(H) = \mathbf{Q}(H|\mathcal{P}^i(w))$ .

At step  $n+1$ , the agents announce  $\mathbf{Q}_{n+1}^1(H)$  and  $\mathbf{Q}_{n+1}^2(H)$ , respectively. Consider the set  $\tilde{\mathcal{N}}_{n+1}^i = \{E \in \mathcal{P}_n^i \mid \mathbf{Q}(H|E \cap C_n) = \mathbf{Q}_n^i(H)\}$ . It is easy to see that  $\tilde{\mathcal{N}}_{n+1}^i = \emptyset$  if and only if there was dilation at step  $n$ . Let

$$\mathcal{N}_{n+1}^i = \begin{cases} \mathcal{P}_n^i & \text{if } \tilde{\mathcal{N}}_{n+1}^i = \emptyset, \\ \tilde{\mathcal{N}}_{n+1}^i & \text{otherwise.} \end{cases}$$

That is, when  $\tilde{\mathcal{N}}_{n+1}^i = \emptyset$ , no new information is conveyed by the announcement of  $\mathbf{Q}_n^i(H)$ . As before, let  $C_{n+1} = \bigcup \mathcal{N}_{n+1}^1 \cap \bigcup \mathcal{N}_{n+1}^2$ , and  $\mathcal{P}_{n+1}^i = \{E \in \mathcal{N}_{n+1}^i \mid E \cap C_{n+1} \neq \emptyset\}$ . Clearly, it remains true that  $\mathcal{P}_{n+1}^i \subseteq \mathcal{N}_{n+1}^i \subseteq \mathcal{P}_n^i$  and  $C_{n+1} = \bigcup \mathcal{P}_{n+1}^1 \cap \bigcup \mathcal{P}_{n+1}^2$ . If  $\mathcal{P}_{n+1}^i = \mathcal{P}_n^i$ , or equivalently, if  $C_{n+1} = C_n$ , the procedure stops; otherwise, agent  $i$  updates credence of  $H$  according to whether  $\{E \in \mathcal{P}_{n+1}^i \mid E \in \mathcal{P}_{n+1}^i\}$  dilates  $H$ . If it does not dilate  $H$ , the credence is updated to  $\mathbf{Q}_{n+1}^i(H) = \mathbf{Q}(H|\mathcal{P}_{n+1}^i(w) \cap C_{n+1})$ ; otherwise, the credence is updated to  $\mathbf{Q}_{n+1}^i(H) = \mathbf{Q}(H|C_{n+1})$ .

For instance, if the agents in Example 3.1 are commonly known to be dilation-averse and follow the above procedure, then at step 0, seeing that her partition  $\{\{w_1, w_2\}, \{w_3, w_4\}\}$  dilates  $H$ , agent 1 will ignore her private information (i.e.,  $\{w_1, w_2\}$ ) and go with  $\mathbf{Q}_0^1(H) = \mathbf{Q}(H|C_0) = \{1/2\}$ . Then at step 1,  $\mathcal{P}_1^1 = \mathcal{N}_1^1 = \mathcal{P}_0^1$ , and the procedure stops.

As the original, Bayesian procedure of communicating posteriors, this dilation-averse procedure will surely stop at step  $m+1$  for some  $m \geq 0$ . It is then very easy to adapt the arguments for Theorems 3.4, 4.1, and 4.4 to show the following:

**Theorem 5.2.** *Suppose the dilation-averse procedure of communicating posteriors stops at step  $m+1$ . Then*

- 1)  $\underline{\mathbf{Q}}_m^1(H) = \underline{\mathbf{Q}}_m^2(H)$  and  $\overline{\mathbf{Q}}_m^1(H) = \overline{\mathbf{Q}}_m^2(H)$ ;
- 2) If  $\mathbf{Q}$  is closed, then  $\mathbf{Q}_m^1(H) \cap \mathbf{Q}_m^2(H) \neq \emptyset$ ; and
- 3) If  $\mathbf{Q}$  is closed and connected, then  $\mathbf{Q}_m^1(H) = \mathbf{Q}_m^2(H)$ .

*Proof.* For each  $i$ , either  $\{E \in \mathcal{P}_m^i \mid E \in \mathcal{P}_m^i\}$  dilates  $H$ , in which case  $\mathbf{Q}_m^i(H) = \mathbf{Q}(H|C_m)$  by the design of the procedure, or  $\{E \in \mathcal{P}_m^i \mid E \in \mathcal{P}_m^i\}$  does not dilate  $H$ , in which case the argument for Theorem 3.4 is applicable to derive that  $\underline{\mathbf{Q}}_m^i(H) = \underline{\mathbf{Q}}(H|C_m)$  and  $\overline{\mathbf{Q}}_m^i(H) = \overline{\mathbf{Q}}(H|C_m)$ . Either way we have 1). The derivations of 2) and 3) from 1) are the same as those of Theorems 4.4 and 4.1.  $\square$

Therefore, two agents who are commonly known to be dilation-averse cannot agree to disagree on lower or upper probabilities, and, under common assumptions, cannot agree not to fully agree.

<sup>7</sup>This dilation-averse updating rule is in a way weaker than Kyburg's Strength rule, in that it only ignores information when the information is about the value of a *dilating partition*. Still, Levi's example (Example 5.1) can be adapted to reveal the conflict of this weaker rule with basic Bayesian theory.

## 6. Discussion and Conclusion

Like Aumann's original result, the results in this paper are mathematically simple once the framework is set up, but they highlight an interesting connection between the possibility of agreeing to disagree and the phenomenon of dilation. We offered two perspectives to view this connection. For Bayesian agents with a common set of priors, agreeing to disagree on lower or upper posteriors entails the presence of dilation for at least one of them. For dilation-averse (but otherwise Bayesian) agents with a common set of priors, it is impossible to agree to disagree on lower or upper posteriors.

Although the absence of dilation is sufficient for Bayesian agents to reach agreements by communicating posteriors, it is certainly not necessary. Here is a simple example to show this.

**Example 6.1.** Let  $\Omega = \{w_1, w_2, w_3, w_4\}$  and  $\mathcal{A}$  its power set. Suppose  $\mathcal{P}^1 = \{\{w_1, w_2\}, \{w_3, w_4\}\}$  and  $\mathcal{P}^2 = \{\{w_1, w_3\}, \{w_2, w_4\}\}$ . Let  $H = \{w_1, w_4\}$ , and suppose the true state of the world is  $w_1$ . Let  $\tilde{p}$  be the uniform distribution over the 4-atom algebra, and  $\Lambda$  be the set of all distributions over the 4-atom algebra, and let  $\mathbf{Q} = \{(0.8\tilde{p} + 0.2q \mid q \in \Lambda)\}$ .<sup>8</sup>

Like Example 3.1, this is an Aumann case, where the posteriors of  $H$  are common knowledge without announcements, because  $\mathbf{Q}(H|\{w_1, w_2\}) = \mathbf{Q}(H|\{w_3, w_4\}) = [1/3, 2/3]$ , and  $\mathbf{Q}(H|\{w_1, w_3\}) = \mathbf{Q}(H|\{w_2, w_4\}) = [1/3, 2/3]$ . So the procedure of communicating posteriors stops at step 1, and  $C_1 = C_0 = \Omega$ . Dilations do occur, for both agents, because  $\mathbf{Q}(H|C_0) = \mathbf{Q}(H) = [0.4, 0.6]$ , which is strictly contained in  $[1/3, 2/3]$ . Despite the presence of dilations, the two agents will still reach an agreement even if they are not dilation-averse, though of course the agreement is different from the one dilation-averse agents would reach.

It is also worth noting that this example is a generalization of an example from Geanakoplos and Polemarchakis (1982), which was used to illustrate the fact we mentioned in Section 2, that the consensus resulting from communicating posteriors can be different from the consensus resulting from directly exchanging private information. If both pieces of private information in the example become public, the two agents will converge on a precise, extreme probability.

There are several ways in which our results may be expanded. First, a potentially interesting route is to take advantage of the connection between Aumann-style agreement theorems and the so-called no-trade theorems. Instead of emphasizing the agreement of posteriors when they are common knowledge, the no-trade theorems establish conditions under which it is impossible to have a trade that is commonly known to be acceptable to both parties. For example, Milgrom and Stokey (1982) showed that given an *ex ante* Pareto-efficient arrangement (and some mild assumptions), even after the agents receive different private information, there is no trade that can be commonly known to be mutually acceptable. This remarkable result implies that significant volumes of trade in real-world markets cannot be explained solely on the basis of informational differences. Sebenius and Geanakoplos (1983), Holmström and Myerson (1983), and Rubinstein and Wolinsky (1990), among others, continued this line of research and established several no-trade type results.

Here is a very simple example to illustrate the connection between no-trade results and Aumann's agreement result.

---

<sup>8</sup>This  $\epsilon$ -contamination model ( $\epsilon = 0.2$ ) can be equivalently specified as the largest set of probability measures on the 4-atom algebra that satisfy the constraint that every atom receives a lower probability of 0.2.

**Example 6.2.** Let  $\Omega = \{w_1, w_2, w_3, w_4\}$  and suppose two agents have the same precise prior:  $p(w_i) = \frac{1}{4}$ . The information partitions are  $\mathcal{P}^1 = \{\{w_1, w_2\}, \{w_3, w_4\}\}$  and  $\mathcal{P}^2 = \{\{w_1, w_4\}, \{w_2, w_3\}\}$ , respectively. For simplicity, consider just one possible transaction such that if it is carried out, the agents' state-dependent payoffs are given as follows:  $u_1 = 1.9 - i$  and  $u_2 = i - 1.9$ , where  $i$  is the index of the state  $w_i$ ; otherwise, each agent receives 0. Suppose the true state of the world is  $w_2$ .

In this simple situation, the arrangement is *ex ante* pareto-efficient with respect to the transaction in question, in the sense that the transaction is not acceptable to both agents before they receive private information. But after they receive private information at state  $w_2$ , agent 1's expected payoff from the transaction becomes 0.4, and agent 2's expected payoff from the transaction becomes 0.6. Both are positive and thus both parties are willing to trade. However, once it becomes common knowledge that both are willing to trade, further information is conveyed to the agents. In this case, once agent 2 knows that agent 1 is willing to trade, then agent 2 learns that agent 1 has received the information  $\{w_1, w_2\}$  rather than  $\{w_3, w_4\}$ , for the latter would have made the transaction unacceptable to agent 1. Combining this with her own private information, agent 2 gets to know that the true state is  $w_2$ . Similarly, agent 1 can reason that 2 can reason this way and distinguish between  $w_1$  and  $w_2$ . Since agent 2 is still willing to trade, agent 1 knows that agent 2 knows that the true state is  $w_2$ . Then the transaction is not acceptable to agent 1 after all.

The key idea is that once the willingness to trade becomes common knowledge, this new information changes the agents' beliefs in a way that destroys the incentives to trade. This should be reminiscent of what gives rise to Aumann-style agreement results. In this simple case, the two agents converge to the same posterior after it becomes common knowledge that both parties are willing to carry out the transaction. Consequently, the rational expectations based on the new, agreed posterior would make the transaction unacceptable for at least one party.

In view of the connection illustrated above and the generalization of Aumann's agreement theorem in the context of imprecise probability, it is natural to ask whether no-trade theorems can be similarly generalized. There have been many attempts to extend no-trade theorems beyond the context of standard expected utility preference (e.g., Dow et al. (1990), Ma (2001), Halvey (2004), Wakai (2002), Kajii and Ui (2009)). In particular, Kajii and Ui (2009) gave an elegant characterization of *interim efficiency* under Bewley's (2002) incomplete preferences derived from multiple priors (and concave utility functions). Using this characterization, they showed that there is a no-trade theorem (which essentially says that *ex ante efficiency* entails *interim efficiency*) in their setting if agents employ the full Bayesian updating rule, and they demonstrated how the result is intimately related to an agreement result on "partial" agreement between sets of probabilities. We suspect that our results on "full" agreement in this paper may be adapted to develop new no-trade results that apply to a wider range of preference models, for intuitively a full agreement between posteriors may ensure interim efficiency under a wider range of preference models than a merely partial agreement does. This strikes us as an interesting question to explore further.

Second, speaking of partial agreement, it is worth noting that when "agreement" is interpreted as partial agreement, the common prior assumption may also be relaxed to the assumption that priors (partially) agree, i.e., that the two sets of priors have a non-empty intersection. This is, for example, what Carvajal and Correia-da-Silva (2010) assume in their results. Their main agreement result about Bayesian agents (Proposition 1) is that if two Bayesian agents have closed and connected sets of priors that have a non-empty intersection, and both sets of posteriors on a hypothesis are common knowledge, then the sets of posteriors also have a non-empty inter-

section. This result, just like Aumann’s original result, is straightforwardly generalizable to the setting of communicating posteriors. An open question, in light of our results here, is what purchase the condition of no dilation has in the context of priors that do not fully agree, or to put it differently, whether stronger agreement results are available in this context for dilation-averse agents.

Third, we have only considered the full Bayesian updating rule (and the dilation-averse variant). Other updating rules may be examined in our setting, especially the Dempster-Shafer updating rule, which was considered by Kajii and Ui (2005, 2009) and Carvajal and Correia-da-Silva (2010). For Dempster-Shafer updating, Carvajal and Correia-da-Silva’s main agreement result requires each agent’s set of likelihood maximizers as well as their sets of posteriors to be common knowledge, which suggests that communication of posteriors alone is in general not enough to guarantee agreement, even if dilation — which can also occur with Dempster-Shafer updating (Seidenfeld, 1997) — is assumed away. One natural idea is to allow also the communication of likelihood maximizers. On the other hand, Seidenfeld (1997) showed that for  $\epsilon$ -contamination models (Huber, 1973; Berger, 1984) Dempster-Shafer updating is equivalent to full Bayesian updating. Therefore, if we can derive a corollary about  $\epsilon$ -contamination models (in the spirit of Corollary 4.3) from Theorem 4.1 and results on dilation in  $\epsilon$ -contamination models, that will also apply to Dempster-Shafer updating.

## Acknowledgements

This research was supported in part by the Research Grants Council of Hong Kong under the General Research Fund LU13600715, and by a Faculty Research Grant from Lingnan University. The second author was also supported by Youth Program of National Social Science Foundation of China (No. 17CZX073), and by the Fundamental Research Funds for the Central Universities (No. 16wkpy30).

## References

- Aumann, R. (1976). Agreeing to disagree. *Annals of Statistics* 4: 1236-1239.
- Berger, J. (1984). The robust Bayesian viewpoint (with discussion). In *Robustness in Bayesian Statistics* (J. Kadane, ed.), pp. 63-124. North-Holland, Amsterdam.
- Bewley, T. F. (2002). Knightian decision theory: Part I. *Decisions in Economics and Finance* 25: 79-110.
- Carvajal, A. and Correia-da-Silva, J. (2010). *Agreeing to disagree with multiple priors* (No. 368). Universidade do Porto, Faculdade de Economia do Porto.
- De Finetti, B. (1972). *Probability, Induction, and Statistics*. John Wiley, New York.
- Dow, J., Madrigal, V. and Werlang S. (1990) Preference, common knowledge, and speculative trade. Working paper, 1990.
- Geanakoplos, J. and Polemarchakis, M. (1982). We can’t disagree forever. *Journal of Economic Theory* 26: 363-390.
- Good, I. J. (1974). A little learning can be dangerous. *British Journal of Philosophy Science* 25: 340-342.
- Grünwald, P. D. and Halpern, J. Y. (2004). When ignorance is bliss. In *Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence*, pp. 226-234. AUA Press.
- Halevy, Y. (2004). The possibility of speculative trade between dynamically consistent agents. *Games and Economic Behavior* 46: 189-198.
- Herron, T., Seidenfeld, T. and Wasserman, L. (1997). Divisive conditioning: Further results on dilation. *Philosophy of Science* 64: 411-444.
- Hill, B. M. and Lane, D. (1985). Conglomerability and countable additivity. *Sankhyā: The Indian Journal of Statistics, Series A*: 366-379.
- Holmström, B. and Myerson, R. B. (1983). Efficient and durable decision rules with incomplete information. *Econometrica* 51: 1799-1820.

- Huber, P. J. (1973). The use of Choquet capacities in statistics. *Bulletin of the International Statistical Institute* 45: 181-191.
- Levi, I. (1977). Direct inference. *The Journal of Philosophy* 74(1): 5-29.
- Kajii, A. and Ui, T. (2005). Incomplete information games with multiple priors. *Japanese Economic Review* 56: 332-351.
- Kajii, A., and Ui, T. (2009). Interim efficient allocations under uncertainty. *Journal of Economic Theory* 144: 337-353.
- Kyburg H. E. (1961). *Probability and the Logic of Rational Belief*. Wesleyan University Press, Middleton.
- Kyburg H. E. (1974). *The Logical Foundations of Statistical Inference*. Reidel, Dordrecht.
- Ma, C. (2001). A no-trade theorem under Knightian uncertainty with general preferences. *Theory and Decision* 51: 173-181.
- Milgrom, P. R. and Stokey, N. L. (1982). Information, trade and common knowledge. *Journal of Economic Theory* 26: 17-27.
- Rubinstein, A. and Wolinsky A. (1990). On the logic of "Agreeing to disagree" type results. *Journal of Economic Theory* 51: 184-193.
- Schervish, M. J., Seidenfeld, T. and Kadane, J. B. (1984). The extent of non-conglomerability of finitely additive probabilities. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 66(2): 205-226.
- Schervish, M. J., Seidenfeld, T. and Kadane, J. B. (2016). Non-conglomerability for countably additive measures that are not  $\kappa$ -additive. *The Review of Symbolic Logic*, forthcoming.
- Sebenius, J. and Geanakoplos J. D. (1983). Don't bet on it: Contingent agreements with asymmetric information. *Journal of the American Statistical Association* 78: 424-426.
- Seidenfeld, T. (1981). Levi on the dogma of randomization. In *Henry E. Kyburg and Isaac Levi* (R. Bogdan, ed.), pp. 263-291. Reidel, Dordrecht.
- Seidenfeld, T. (1997). Some static and dynamic aspects of robust Bayesian theory. In *Random Sets*, pp. 385-406. Springer, New York.
- Seidenfeld, T. (2007). Forbidden fruit: When Epistemological Probability may not take a bite of the Bayesian apple. *Probability and inference: Essays in honor of Henry E. Kyburg, Jr.*, William Harper and Gregory Wheeler (eds.). London: Kings College Publications.
- Seidenfeld, T. and Wasserman, L. (1993). Dilation for sets of probabilities. *Annals of Statistics* 21: 1139-1154.
- Wakai, K. (2002). Linking Behavioral Economics, Axiomatic Decision Theory and General Equilibrium Theory. Ph.D. Dissertation, Yale University.
- Wasserman, L. (1992). Invariance properties of density ratio priors. *Annals of Statistics* 20: 2177-2182.
- Walley, P. (1991). *Statistical reasoning with imprecise probabilities*. New York: Chapman and Hall.