Monologues, Dialogues, and Common Priors*

A. Di Tillio † E. Lehrer ‡ D. Samet §

May 19, 2021

Abstract

The main purpose of this paper is to provide a simple criterion enabling to conclude that two agents do not share a common prior. The criterion is simple, as it does not require information about the agents' knowledge and beliefs, but rather only the record of a dialogue between the agents. In each stage of the dialogue the agents tell each other the probability they ascribe to a fixed event and update their beliefs about the event. To characterize dialogues consistent with a common prior, we first study monologues, which are sequences of probabilities assigned by a single agent to a given event in an exogenous learning process. A dialogue is consistent with a common prior if and only if each selection sequence from the two monologues comprising the dialogue is itself a monologue.

Keywords: Learning processes; Bayesian dialogue; Bayesian monologue; Ratio variation; Joint fluctuation; Agreement.

JEL classification: D83.

---

*Lehrer acknowledges the support of NSFC-ISF Grant #2510/17. Samet acknowledges the support of ISF Grant #722/18. The authors acknowledge an inspiring talk with Moshe Hoffman.

†Bocconi University; alfredo.ditillio@unibocconi.it

‡Tel Aviv University; lehrer@post.tau.ac.il

§Coller School of Management, Tel Aviv University; samet@gmail.com

Two monologues do not make a dialogue.

De Nevers' Law of Debate

1 Introduction

Theoretical arguments against the common prior assumption were raised most notably by Morris [1995] and Gul [1998]. This paper offers a simple criterion for showing that two agents do not have a common prior (CP).

We consider a learning process in which two agents exchange information about the probability they ascribe to a given event E. In the first stage of this process the agents truthfully and simultaneously report to each other their initial probabilities. This means that these probabilities become common knowledge. Acquiring this information each of the agents updates the probability she ascribes to E. In the second stage they again make their updated probabilities common knowledge. And so on.[1] We assume that in each stage the conditional probability of E given the new information is well defined. That is, the event on which the probability is conditioned has a positive probability. If, moreover, the cumulative information when the process is completed has a positive probability, we say that the process is positive. To easily meet the positivity conditions we assume that the state space is countable.[2]

A pair of sequences of probability numbers generated in such a learning process, one for each agent, is called a *dialogue*. We provide a necessary and sufficient condition for a pair of sequences to be a dialogue in a positive learning process, where the agents have a CP. This condition does not prove the existence of a CP, it only guarantees that there is a knowledge-belief structure with a CP in which the dialogue can be realized. However, failure of this condition proves that the agents do not have a CP.

To describe our condition for dialogues we first study learning processes of a single agent who sequentially acquires new information. In such a process the information acquired by the

[1]We assume that the agents report truthfully the probability of E in each stage. This would be the case if each is facing a decision problem whose result depends only on whether E is the case or not. Since the agents do not compete, it is the interest of both to truthfully share the information about E. For this reason they exchange their views of the probability of E. This leads them of course to generate a dialogue. In each stage, it is the interest of each to report her true probability, because the rest of the process depends on her report. If she reports a different probability, the information she receives in later stages will be also flawed.

[2]Our results hold for finite spaces and also for uncountable measurable spaces with countable or finite partitions.

agent is *exogenously*, given as opposed to the information in a dialogue which is *endogenously* generated by the agents. The sequence of the probabilities ascribed by the agent to a given event E along such a process is called a *monologue*. Obviously, a dialogue consists in particular of two monologues.

We first characterize sequences of numbers that are monologues. This characterization is made in terms of the fluctuation of the sequence. It is known that a positive monologue must have bounded variation. This condition, however, is not sufficient. We introduce a stronger notion of fluctuation, named *bounded ratio variation*. This condition is necessary and sufficient for a sequence to be a positive monologue.

The condition for dialogues reflects an intuitive understanding of priors and common prior. The prior characterizes an agent. Information can change, and with it the posterior beliefs. But it is the prior that remains constant. Now, if agents have a common prior it means that in a sense they are the same, but for informational differences. Therefore, we could possibly think of a dialogue of agents with a CP to be a monologue of one agent which is characterized by the CP. How can this be formalized?

Imagine that we listen to a dialogue as follows. At some points in time we hear only the report of agent 1, and in the rest of the times we hear only agent 2. For example, we may listen to 1 at odd periods and to 2 at even periods. This way we observe only one sequence of probability numbers which is a *selection* from the two monologues that comprise the dialogue. If two agents with a CP are essentially two faces of one agent, we can expect this selection to be a monologue of this single agent. We show indeed that a necessary and sufficient condition for two sequences to be a positive dialogue of agents that have a CP is that any selection of the two sequences is a positive monologue. The requirement that *any* selection of the two sequences should be a positive monologue seems at first glance highly demanding, as it involves a continuum of selection sequences. We show, however, that it suffices to check only the boundedness of three sequences.

Our result seems to be a formal rendering of the claim of the motto, known as De Nevers' Law of Debate. If agents share a CP, then not every pair of monologues makes a dialogue. However, we show that every pair of monologues is a dialogue if we do not require that the agents share a CP. In that case the dialogue sounds much like dialogues in the theater of the absurd.

Literature Contribution. We are bringing here together dialogues, monologues, and a necessary and sufficient condition for the existence of a common prior. Each of these three topics is discussed in the literature. We compare this literature to the results in this work.

Necessary and sufficient condition for the existence of a CP. There are several works that provide, like this work, a necessary and sufficient condition for the existence of a CP. The most conspicuous ones are no-trade theorems [Morris, 1994, Feinberg, 2000, Samet, 1998a, Lehrer and Samet, 2014]. In such theorems CP does not exist if and only if there is a state-contingent zero-sum trade which the agents commonly know to yield each of them positive gains. Heifetz [2006] provided a condition analogous to the no-trade condition in syntactic terms. Samet [1998b] provided yet another condition in terms of iterated expectations. Common to all these conditions is their dependence on the state space. More specifically, to refute the existence of a CP one needs to know the knowledge-belief space, or equivalently, know everything about the knowledge and belief of the agents. In contrast, in this paper all that one needs to know in order to refute the existence of a CP is a pair of sequences of probability numbers.

Dialogues. Dialogues of the type studied here were first delineated in the last paragraph of Aumann [1976]. He describes a simultaneous dialogue concerning the probability of a coin falling on H after each of the individuals made a number of observations known only to her. A dialogue is *simultaneous* when at each stage *both* posteriors become common knowledge, as in our paper. In light of the agreement theorem proved in Aumann [1976], common knowledge of the posteriors of an event implies that the two posteriors coincide. Aumann therefore concluded that the dialogue must end with the same posterior.

Geanakoplos and Polemarchakis [1982] proved formally that any serial dialogue in a finite model must end with the same probability ascribed by both individuals to the given event. A dialogue is *serial* when in each period only one of the individuals informs the other of his posterior. They showed, moreover, that in all but the last period, the individuals can repeat each the same probability, and only in the last period an agreement is reached which is commonly known. Polemarchakis [2016], which inspired our paper, showed that any two finite, internal sequences can be obtained as a serial dialogue in a finite model with a common prior. Hart and Taumann [2004] showed, in a similar model, but with communication replaced by observation of the market, that behavior in the market can remain constant for several periods, and then crash. In contrast to our work, the analysis in these papers is made locally. That is, a state is fixed and the updating of the knowledge of the players is followed in this state. All these papers assumed finite partitions, which guarantees that common knowledge of the posterior probability of the event is reached in finite time.

Nielsen [1984] extended both Aumann [1976] and Geanakoplos and Polemarchakis [1982] by allowing knowledge structures given by sigma algebras rather than finite partitions. He formulated and proved Aumann's agreement theorem for such knowledge structures and showed that dialogues, simultaneous and sequential, that may be infinite, converge almost surely to the same probability. His analysis, like ours, is global: in each period the knowledge of the individuals is described in *all* states by specifying a knowledge structure in each period.[3]

To show that a dialogue is inconsistent with a CP, we need to examine the infinite dialogue. Any finite part of the dialogue is consistent with a CP. An analogous result was presented by Lipman [2003], who showed that any finite set of descriptions of a player's beliefs is consistent with a CP.

Monologues. The sequence of probabilities of one individual in a dialogue is a monologue, which is simply the result of a learning process of one agent. The literature on individual learning processes dealt with such sequences. Burkholder [1966] showed that an L^1-bounded martingale sequence is of bounded variation almost surely on every atom of the basic probability space. A simpler proof was given in Tsuchikura and Yamasaki [1976]. We prove a stronger result: for our martingales, the sequence must be of bounded *ratio* variation. Moreover, we show that every sequence can be realized when the prior of a state is 0. Recently, Shaiderman [2018] has shown that any L^2-bounded martingale, when conditioned on a positive probability event, has bounded variation. This is typically false when the martingale is only L^1-bounded.

2 Monologues

A monologue is the sequence of conditional probabilities assigned to a fixed event along a learning process. Formally, a *learning process* is a tuple $(\Omega,\mu,E,(\pi_k)_{k\geqslant 1})$ where (Ω,μ) is a countable or finite probability space, $E\subseteq\Omega$ is an event, and $(\pi_k)_{k\geqslant 1}$ is a a sequence of partitions of Ω which is a *filtration*, that is, π_{k+1} refines π_k.

Let P_k be the set of ω's such that $\mu(\pi_k(\omega))>0$, where $\pi_k(\omega)$ is the element of π_k containing ω. By the countability of Ω it follows that $\mu(P_k)=1$. Define $P=\cap_k P_k$, then $\mu(P)=1$. For

[3] Bacharach [1979] looked at dialogues when information is normally distributed.

$\omega \in P$ we call the sequence $\big(p_k(\omega)\big)_k = \big(\mu(E|\pi_k(\omega))\big)_k$, the *monologue* at ω. We say that the monologue at ω is *positive* if $\mu(\cap_k \pi_k(\omega)) > 0$. Clearly, if the monologue at a state is not positive, then the probability of that state is zero. Thus, monologues are positive with probability one. A sequence of numbers (p_k) is a (*positive*) *monologue* if it is a (positive) monologue at some state in a learning process.

Example 1. Independent trials are conducted sequentially to find out whether a machine can fail. The probability of success in the k-th trial is $q_k > 0$, so the probability that all trials are successful is $\prod_{k=1}^{\infty} q_k$. After each trial, the probability that the machine is infallible is announced. The states of the world are $1, 2, \ldots, n, \ldots, \infty$, where state n means that trials $1, \ldots, n-1$ were successful and trial n failed. Obviously, for $n < \infty$, $\mu(n) = (\prod_{k=1}^{n-1} q_k)(1-q_n)$, and $\mu(\infty) = \prod_{k=1}^{\infty} q_k$. The first partition, π_1, reflects the knowledge before the trials, and it is the trivial partition. At that point it is not clear if the machine will ever fail, and if it fails at what time it happens. At time $k+1$, it is known if the machine failed at any time before $k+1$, but if it did not fail, it is not known if it ever fails or at what time after k it will fail. Thus the partition is $\pi_{k+1} = \{\{1\}, \{2\}, \ldots, \{k\}, \{k+1, \ldots, \infty\}\}$. Let $E = \{\infty\}$ be the event that the machine is infallible. The announcement sequence $p_k(\infty) = \mu(E|\pi_k(\infty)) = \prod_{n=k}^{\infty} q_n$, $k = 1, 2, \ldots$, is the monologue at state ∞.

Not every sequence in the interval $[0,1]$ is a monologue. For example, the boundaries of the unit interval are absorbing for monologues, that is, if $p_k = 0$ or $p_k = 1$ for some k, then $p_n = p_k$ for all $n > k$. Thus, any sequence that hits a boundary and is not absorbed there is not a monologue. Using the techniques presented below, it is easy to show that every sequence that is absorbed in one of the boundaries is a monologue in some learning process. This is why in what follows we consider only sequences lying in the open interval $(0,1)$, which we call *internal*. In particular, we are interested in characterizing internal sequences that are positive monologues, which are the monologues observed with probability one.

Our characterization involves a condition restricting the fluctuation of the sequences. Let (p_k) be an internal sequence and define $\bar{p}_k = 1 - p_k$ for every k. We define the *ratio variation* of the sequence as

$$\sum_k \max\left\{\frac{p_{k+1}}{p_k} - 1, \frac{\bar{p}_{k+1}}{\bar{p}_k} - 1\right\} \tag{1}$$

and say that the ratio variation is *bounded* if the sum is finite.

The logic behind the definition of ratio variation is as follows. If the sequence (p_k) is the monologue concerning an event E, then the sequence $(\bar{p}_k)$ is the monologue concerning the complement $\bar{E}$ of E. The ratios p_{k+1}/p_k and $\bar{p}_{k+1}/\bar{p}_k$ measure the change in the agent's beliefs at stage k. The closer they are to 1, the smaller is the change. Thus, the sums of $|p_{k+1}/p_k - 1|$ or $|\bar{p}_{k+1}/\bar{p}_k - 1|$ measure the total change in the agent's belief along the learning process. The ratio variation picks at each k one of $|p_{k+1}/p_k - 1|$ and $|\bar{p}_{k+1}/\bar{p}_k - 1|$ according to the following reasoning. When $p_{k+1} > p_k$ the information at stage k confirms E, that is, it increases the probability of E. In this case $\bar{p}_{k+1} < \bar{p}_k$, which means that $\bar{E}$ is disconfirmed. The ratio variation picks up the change in the probability of the confirmed event E, namely, $|p_{k+1}/p_k - 1| = p_{k+1}/p_k - 1$. When $\bar{p}_{k+1} > \bar{p}_k$ the event $\bar{E}$ is confirmed, while E is disconfirmed. In this case $|\bar{p}_{k+1}/\bar{p}_k - 1| = \bar{p}_{k+1}/\bar{p}_k - 1$ measures the change. Note that $p_{k+1}/p_k - 1$ and $\bar{p}_{k+1}/\bar{p}_k - 1$ are of opposite sign, and the maximum in the definition of ratio variation yields the choice we have just described.

Theorem 1. *An internal sequence is a positive monologue if and only if it has bounded ratio variation.*

We discuss later (after Corollary 1 below) the connection between positiveness of the process and boundedness of the ratio variation.

Variation and Ratio Variation. Our novel notion of ratio variation measures fluctuation of a sequence (p_k) by comparing the *ratios* p_{k+1}/p_k and $\bar{p}_{k+1}/\bar{p}_k$ to 1. A more standard measure of fluctuation, which compares the *differences* $p_{k+1} - p_k$ to 0, is the *variation* of the sequence, $\sum_k |p_{k+1} - p_k|$. In this case it does not matter if we use p_k or $\bar{p}_k$ to measure fluctuation, as $|p_{k+1} - p_k| = |\bar{p}_{k+1} - \bar{p}_k|$. Given this equality, we can rewrite the ratio variation in terms of differences as follows:

$$\sum_k |p_{k+1} - p_k|/r_k, \tag{2}$$

where $r_k = p_k$ when $p_{k+1} \geqslant p_k$ and $r_k = \bar{p}_k$ when $p_k \geqslant p_{k+1}$.[4] Since $r_k \leqslant 1$ for every k, we immediately obtain:

Observation 1. *Bounded ratio variation implies bounded variation.*

Thus, bounded variation is a necessary condition for a sequence to be a positive monologue,

[4]We can equally define ratio variation in terms of the ratios p_k/p_{k+1} rather than p_{k+1}/p_k, by $\max\{1 - p_k/p_{k+1}, 1 - \bar{p}_k/\bar{p}_{k+1}\}$. In this case, r_k in equation (2) is defined as p_{k+1} when $p_{k+1} \geqslant p_k$ and $\bar{p}_{k+1}$ when $p_{k+1} \leqslant p_k$. Our results hold also for this definition of ratio variation.

but it is the stronger notion of bounded ratio variation that turns out to be the necessary and sufficient condition. The next example shows that the two notions are not equivalent.

Example 2. Consider the sequence $x, y, x/2, y/2, \ldots, x/2^n, y/2^n, x/2^{n+1} \ldots$, where $y > x > 0$. The variation of this sequence is $\sum_{n=1}^{\infty} |x/2^n - y/2^n| + \sum_{n=1}^{\infty} |y/2^n - x/2^{n+1}|$. Each of the two sums is a geometric series with quotient $1/2$, hence it converges. Thus, the variation is bounded. But, for each n, $y/2^n > x/2^n$, hence the ratio variation contains the sum $\sum_n |x/2^n - y/2^n|/(x/2^n)$. Since each term in this sum is $|1 - y/x| > 0$, the ratio variation is unbounded.

We note that bounded variation of a sequence implies that the sequence is Cauchy and hence converges.[5] Thus, by Observation 1 and Theorem 1 we obtain the following:

Corollary 1. *Positive monologues are converging sequences.*

This claim can be easily verified also directly. Given a learning process $(\Omega, \mu, E, (\pi_k)_{k\geqslant 1})$, for every state ω the sequence of events $(\pi_k(\omega))_{k\geqslant 1}$ is decreasing and converges to $\bigcap_k \pi_k(\omega)$. Thus, if the learning process is positive at ω, then the sequence $(p_k(\omega))_{k\geqslant 1}$ converges to $\mu(E \mid \bigcap_k \pi_k(\omega))$.

Corollary 1 and the discussion thereafter gives an intuitive appeal to the connection between positiveness of the process and the boundedness of the ratio variation of the monologue. The basic intuition is that the steps in a journey cover a finite distance if and only if the journey reaches a final destination. Analogously, and more abstractly, it stands to reason that an incremental process reaches a terminal point if and only if the sum of the increments is finite. Positiveness of the learning process means that it reaches a terminal point. That is, it converges to the probability of E given everything that is learned in the process. The increments of the learning process are given by the terms in Equation (1), as argued before. Theorem 1 states that the sum of the increments is finite, that is, the ratio variation is bounded, if and only if the process reaches a terminal point, i.e., it is positive.

While bounded ratio variation is strictly stronger than bounded variation, there are cases in which the two notions are equivalent. Call a sequence *strictly internal* if for some $0 < \varepsilon < 1$ the sequence lies in the interval $(\varepsilon, 1 - \varepsilon)$.

[5] The converse of this claim is not true. For example, the internal sequence $p_k = \sum_{n=1}^{k} (-1)^{n+1}/(n+1)$ converges to $1 - \ln 2$, but its variation is $\sum_{n=3}^{\infty} 1/n$, which is unbounded.

Observation 2. *In the following two cases, bounded variation is equivalent to bounded ratio variation:*

(i) *the sequence is strictly internal;*

(ii) *the sequence is internal and converges monotonically to zero or to one.*

Thus, for the two cases listed in Observation 2, the equivalence in Theorem 1 can be stated in terms of bounded variation.

Corollary 2. *Fix a sequence that is either strictly internal, or internal and monotonically converging to zero or one. The sequence is a positive monologue if and only if it has bounded variation.*

The following example exhibits a family of sequences guaranteed by Theorem 1 to be positive monologues, and constructs a learning process in which this is satisfied.

Example 3. Consider a decreasing (and hence converging) internal sequence (p_k). The sequence converges and its variation, $p_1 - \lim p_k$, is bounded. The sequence also has bounded ratio variation. This follows from Observation 2(i) if the limit is positive (and hence the sequence is strictly internal) and from Observation 2(ii) if the limit is 0. Thus, by Theorem 1, there exists a learning process in which the sequence is a positive monologue. To construct such a process, let $\Omega = \{\omega_1, \omega_2, \omega_3, \ldots, \omega\}$ and $E = \Omega \setminus \{\omega\}$. Also let $(\pi_k)_{k \geqslant 1}$ be a filtration such that for every k, $\pi_k(\omega) = \{\omega_k, \omega_{k+1}, \ldots, \omega\}$. We construct the unique probability μ that makes (p_k) a monologue at ω, and show that the monologue is positive. Since for each k, $\mu(\omega_k) = \mu(\pi_k(\omega)) - \mu(\pi_{k+1}(\omega))$, and $\mu(\omega) = \lim_k \mu(\pi_k(\omega))$, it is enough that we define $\mu(\pi_k(\omega))$ for each k. Let $a_k = \mu(\pi_k(\omega))$, and $b_k = \mu(E \cap \pi_k(\omega))$. Then $b_k - b_{k+1} = a_k - a_{k+1} = \mu(\omega_k)$. If (p_k) is the monologue at ω, then for each k, $b_k = p_k a_k$. Subtracting from this equality the equality $b_{k+1} = p_{k+1} a_{k+1}$ yields $p_k a_k - p_{k+1} a_{k+1} = a_k - a_{k+1}$, or equivalently, $a_{k+1} = a_k \bar{p}_k / \bar{p}_{k+1}$. Given that $a_1 = 1$, we have $a_{k+1} = \prod_{n=2}^{k} \bar{p}_n / \bar{p}_{n+1} = (1 - p_2)/(1 - p_{k+1})$. Thus, $\mu(\cap \pi_k(\omega)) = \lim a_k = (1 - p_2)/(1 - \lim p_k) > 0$ and the monologue is positive.

Giving Up Positivity. We conclude this section by showing that if we do not require positivity, then there is nothing that prevents a sequence from being a monologue.

Proposition 1. *Every internal sequence is a monologue.*

To illustrate Proposition 1, we consider in the following example an internal sequence with unbounded variation. By Theorem 1 and Observation 1, the sequence is not a positive monologue. Yet, we construct a learning process in which the sequence is a monologue, the existence of which is guaranteed by Proposition 1.

Example 4. Consider the alternating sequence $2/3, 1/3, 2/3, 1/3, \ldots$. Obviously, this sequence does not converge and hence does not have bounded variation. We construct a learning process in which the sequence is a monologue at some state. Let Ω and (π_k) be as in Example 3, but now let $E = \{\omega_1, \omega_3, \omega_5, \ldots\}$. Also, let $\mu(\omega_k) = 2^{-k}$ and $\mu(\omega) = 0$, so that the monologue at ω is not positive. Then, for odd k, $\mu(E \cap \pi_k(\omega)) = 2^{-k+2}/3$, while for even k, $\mu(E \cap \pi_k(\omega)) = 2^{-k+1}/3$, and for each k, $\mu(\pi_k(\omega)) = 2^{-k+1}$. Thus for odd k, $\mu(E|\pi_k(\omega)) = 2/3$, while for even k, $\mu(E|\pi_k(\omega)) = 1/3$.

3 Dialogues and Common Priors

A dialogue is the pair of monologues generated by a joint learning processes where in each stage two agents simultaneously tell each other the probability they assign to a fixed event. By telling each other these probabilities, the agents make them not only known to both, but also commonly known.

Formally, a *joint learning process* is a tuple

$$(\Omega, \mu^1, \mu^2, E, (\pi_k^1)_{k\geqslant 1}, (\pi_k^2)_{k\geqslant 1}),$$

such that $(\Omega, \mu^1, E, (\pi_k^1)_{k\geqslant 1})$ and $(\Omega, \mu^2, E, (\pi_k^2)_{k\geqslant 1})$ are learning processes in the same countable space Ω, and the probabilities μ^1 and μ^2 have the same support.

Starting with the partitions π_1^1 and π_1^2 that are exogenously given, the remaining partitions are defined endogenously by induction. Let $k \geqslant 1$ and suppose that π_k^1 and π_k^2 are defined. The partitions π_{k+1}^1 and π_{k+1}^2 should describe the agents' knowledge after the pair of posteriors

$$p_k(\omega) = (p_k^1(\omega), p_k^2(\omega)) = (\mu(E \mid \pi_k^1(\omega)), \mu(E \mid \pi_k^2(\omega)))$$

becomes commonly known. First we describe the set of states in which this pair is well defined. For $i = 1, 2$, let P_k^i be the set of ω's such that $\mu(\pi_k^i(\omega)) > 0$, and let $P_k = P_k^1 \cap P_k^2$. Thus, $p_k(\omega)$

is well defined in all the states of P_k. Now, for each i, $\mu^i(P_k^i) = 1$, and since μ^1 and μ^2 have the same support, it follows that $\mu^1(P_k) = \mu^2(P_k) = 1$. Let π_k' be the partition of P_k induced by p_k, that is, $\pi_k'(\omega)$ consists of all states ω' such that $p_k(\omega') = p_k(\omega)$. We extend π_k' to a partition $\hat{\pi}_k$ of Ω by adding the complement of P_k, that is, $\hat{\pi}_k = \pi_k' \cup \{\bar{P}_k\}$. For the agents to commonly know $p_k(\omega)$ in stage $k+1$ means that for $\omega \in P_k$, the event $\hat{\pi}_k(\omega)$ is commonly known at ω. For this , the partition π_{k+1}^i is defined as the common refinement of π_k^i and $\hat{\pi}_k$, for each i. Note that for $\omega \in P_k$, $\hat{\pi}_k(\omega)$ is a union of elements of π_{k+1}^1 and also a union of elements of π_{k+1}^2. Since $p_k(\omega)$ is the pair of posteriors at all the states in $\hat{\pi}_k(\omega)$, it is commonly known at ω. This completes the definition of the filtrations (π_k^1) and (π_k^2).

Let $P = \bigcap P_k$. Then for $i = 1,2$, $\mu^i(P) = 1$, $(\Omega, \mu^i, E, (\pi_k^i)_{k\geqslant 1})$ is a learning process, and for each $\omega \in P$, the sequence $\left(p_k^i(\omega)\right)_k$ is the monologue at ω. For each state $\omega \in P$, we call the pair of monologues $\left((p_k^1(\omega)), (p_k^2(\omega))\right)$, the *dialogue* at ω. We say that the dialogue is *positive* if both monologues are positive. When $\mu^1 = \mu^2 = \mu$, we say that the dialogue has a *common prior*. A pair of sequences $\left((p_k^1), (p_k^2)\right)$ is a (*positive*) *dialogue* (*with a common prior*) if it is a (positive) dialogue (with a common prior) at some state in a joint learning process.

The next theorem characterizes pairs of sequences that are positive dialogues with a common prior. The characterization involves sequences obtained as selections from the two monologues. Formally, we say that a sequence (p_k) is a *selection* from two sequences (p_k^1) and (p_k^2) if $p_k \in \{p_k^1, p_k^2\}$ for each k. Note that, in particular, each of the sequences (p_k^1) and (p_k^2) is a selection from the two sequences.

Theorem 2.

(i) *If a pair of internal sequences is a positive dialogue with a common prior, then every selection from the two sequences is a positive monologue.*

(ii) *If every selection from a pair of strictly internal sequences is a positive monologue, then the pair of sequences is a positive dialogue with a common prior.*

By Theorem 1 and Observations 1 and 2(i), Theorem 2 can be stated equivalently as follows:

Theorem 2*.

(i) *If a pair of internal sequences is a positive dialogue with a common prior, then every selection from the two sequences has bounded ratio variation.*

(ii) *If every selection from a pair of strictly internal sequences has bounded variation, then the pair of sequences is a positive dialogue with a common prior.*

Of the two parts of the theorem, part (i) delivers the main purpose of this paper, which is to provide a criterion to falsify the existence of a common prior. If the condition in this part fails, then we can answer with a definite No the question whether agents have a CP. In contrast, the condition in part (ii) does not enable us to answer with a definite Yes. The condition only asserts that *some* joint learning process with a common prior yields the positive dialogue. But the dialogue might also arise in a joint learning process without a CP.

Unlike part (i), part (ii) of Theorem 2 assumes *strictly* internal sequences. We conjecture that part (ii) holds also for just internal sequences, but our construction of the learning process required the assumption of strictness. Obviously, part (i) holds in particular when the sequence is strictly internal. Thus, Theorem 2 provides a necessary and sufficient condition for *strictly* internal sequences to be a positive dialogue with a common prior.

The following example presents a pair of strictly internal sequences that satisfy the condition on selections and therefore by Theorem 2(ii), is a positive monologue with a common prior.

Example 5. Let (p_k) be a sequence that is a strictly internal positive monologue. Consider the pair of sequences $(p^1_k) = (p^2_k) = (p_k)$. All selections from these two sequences are (p_k) which is a positive monologue. Thus, by Theorem 2(ii), these two sequences form a positive dialogue with a common prior. In this example the agents always agree with each other, and yet, due to the dialogue they conduct, they learn from each other: along the process they change their beliefs about E.

The next example presents a pair of internal sequences that does not satisfy the condition on selections and therefore, by Theorem 2(i), is not a positive dialogue with a common prior.

Example 6. The following two internal sequences do not satisfy the condition on selections. Let $x_k = \sum_{n=1}^{k}(-1)^{n+1}/(n+1)$. Define

$$(p^1_k) = (x_1, x_1, x_3, x_3, x_5, x_5, \ldots),$$

and

$$(p^2_k) = (x_0, x_2, x_2, x_4, x_4, \ldots).$$

where x_0 is an arbitrary number in $(0,x_1)$. Now, $(x_k) = (p_1^1, p_2^2, p_3^1, p_4^2, \ldots)$, and hence it is a selection from (p_k^1) and (p_k^2). But the variation of (x_k) is $1/3+1/4+1/5+\cdots$, which is unbounded, and thus, by Observation 1 (x_k) does not have bounded ratio variation. Thus, the two sequences violate the condition on selections in Theorem 2(i), and therefore the pair of monologues $\big((p_k^1),(p_k^2)\big)$ is not a positive dialogue with a CP.

Giving Up Positivity or Common Prior. Theorem 2 concerns positive dialogues with a common prior. The proposition below characterizes dialogues where either positivity or the common prior property is omitted.

Proposition 2.

(i) *Any pair of internal sequences is a dialogue with a common prior.*

(ii) *Any pair of strictly internal positive monologues is a positive dialogue.*

Similarly to Proposition 1, part (i) of Proposition 2 shows that omitting positivity makes any pair of internal sequences a dialogue with a common prior. As part (ii) shows, omitting the common prior property while keeping positivity makes positivity of a pair of monologues—which holds by definition when the two monologues form a positive dialogue—a sufficient condition for them to form a positive dialogue.

When Two Monologues Make a Dialogue. Theorem 2 provides a necessary and sufficient condition for a pair of strictly internal *sequences* to be a positive dialogue with a common prior. Our third main result, Theorem 3 below, gives a simple necessary and sufficient condition for a pair of strictly internal *monologues* to be a positive dialogue with a common prior—addressing directly De Nevers' Law of Debate, the motto of this paper.

On our way to the result, we first strengthen Theorem 2(ii) by weakening the assumption that *all* selection sequences have bounded variation. As Proposition 3 below shows, to establish bounded variation of all selection sequences it suffices to check the bounded variation of only three sequences. Given two sequences (p_k^1) and (p_k^2) in the interval $(0,1)$, consider the following three sequences:

(a) $(p_1^1, p_2^2, p_3^1, p_4^2, \ldots)$,

(b) $(p_1^2, p_2^1, p_3^2, p_4^1, \ldots)$,

(c) $(p_k^1 - p_k^2)$

The sequence (a) is the selection sequence whose elements are selected alternately from the two sequences starting from (p_k^1). The sequence (b) is an alternating selection starting with (p_k^2). The sequence (c) is not a selection, but the difference of the two sequences.

Proposition 3. *Let (p_k^1) and (p_k^2) be strictly internal sequences. Then, the following four conditions are equivalent:*

(i) *All the selections from (p_k^1) and (p_k^2) have bounded variation;*

(ii) *The three sequences (p_k^1), (p_k^2), and (a) have bounded variation;*

(iii) *The three sequences (p_k^1), (p_k^2), and (b) have bounded variation;*

(iv) *The sequences (p_k^1), (p_k^2) have bounded variation, and the series defined by (c) absolutely converges, that is, $\sum_k \mid p_k^1 - p_k^2 \mid < \infty$.*

Suppose now that each of two strictly internal sequences (p_k^1) and (p_k^2) is a positive monologue. Then, by Theorem 1 and Observation 2(ii), these two sequences have bounded variation. By Theorem 2 and Proposition 3, to guarantee that the pair of sequences is a positive dialogue with a common prior it is sufficient that one of the sequences (a), (b) has bounded variation or that the series defined by (c) absolutely converges. Thus, the following theorem, which explicitly assumes that (p_k^1) and (p_k^2) are positive monologues, follows directly from Theorem 2 and Proposition 3.

Theorem 3. *A pair of strictly internal positive monologues (p_k^1) and (p_k^2) is a positive dialogue with a common prior if and only if $\sum_k |p_k^1 - p_k^2| < \infty$.*

Example 7. Consider the pair of sequences in Example 6. We claim that each sequence is a strictly internal positive monologue. Both sequences are strictly internal, as they lie in (x_0, x_1). The sequence (p_k^1) is weakly decreasing, while (p_k^2) is weakly increasing, hence they both have bounded variation. Being strictly internal, by Observation 2(i) they also have bounded ratio variation. Thus, by Theorem 1, each of the two sequences is a positive monologue. But, as we have shown, they do not form a positive dialogue with a common prior. Theorem 3 implies that the difference between the sequences does not have bounded variation. Of course, this can be also checked directly.

Eventual Agreement. Consider two internal sequences forming a positive dialogue with a common prior. By Theorem 2(i), every selection from the two sequences is a positive monologue and hence, by Theorem 1, it must converge. This implies that the two sequences converge to the same limit. Thus, we obtain the following:

Corollary 3. *If two internal monologues form a positive dialogue with a common prior, then they converge to the same limit.*

We remark that if for each k and i, $\pi^i_k = \pi^i_{k+1}$, then the claim in Corollary 3 is the agreement theorem of Aumann [1976]. In fact, we can prove Corollary 3 without using Theorem 2, in a way that resembles the proof in Aumann [1976]. Let π_k be the meet of the partitions π^1_k and π^2_k. Since the sequences π^i_k are ordered by refinement, so is the sequence π_k. In particular, $\pi_{k+1}(\omega) \subseteq \pi_k(\omega)$. Let $\omega \in P$, and $Q^i_k(\omega)$ be the set of states ω' in $\pi_k(\omega)$ such that $\mu(E \mid \pi^i_k(\omega')) = \mu(E \mid \pi^i_k(\omega)) = p^i_k$. Since $\pi_k(\omega)$ is a union of elements of π^i_k, it follows that $\mu(E \mid Q^i_k(\omega)) = p^i_k$. By definition, $Q^i_k(\omega) \subseteq \pi^i_k(\omega)$. As $\pi_{k+1}(\omega)$ is a union of elements of π^i_{k+1}, and since $\pi_{k+1}(\omega) \subseteq \pi_k(\omega)$, it follows by the definition of the partition π_{k+1} that $\pi^i_{k+1}(\omega) \subseteq Q^i_k(\omega)$. Therefore $\bigcap_k Q^i_k(\omega) = \bigcap_k \pi_k(\omega)$. Hence, $p^i_k \to \mu(E \mid \bigcap_k Q^i_k(\omega)) = \mu(E \mid \bigcap_k \pi_k(\omega))$. This shows that the two sequences p^i_k converge to the same limit.

4 Proofs – monologues

The condition of bounded ratio variation of a sequence (p_k) is given in terms of *sums* of ratios of the p_k's. In Claim 1, we characterize this condition in terms of *products* of ratios of the p_k's. We use the following lemma to establish a connection between sums and products.

Lemma 1. *Let (ε_k) be a non-negative sequence. Then $\lim_n \prod_{k=1}^n (1+\varepsilon_k) < \infty$ if and only if $\lim_n \sum_{k=1}^n \varepsilon_k < \infty$.*

Proof. For the "only if" direction, observe that $\prod_{k=1}^n (1+\varepsilon_k) \geqslant 1 + \sum_{k=1}^n \varepsilon_k$. For the "if" direction, note that $\lim_n \sum_{k=1}^n \varepsilon_k \geqslant \lim_n \sum_{k=1}^n \ln(1+\varepsilon_k)$, and $\lim_n \sum_{k=1}^n \ln(1+\varepsilon_k) < \infty$, if and only if $\lim_n \prod_{k=1}^n (1+\varepsilon_k) < \infty$. □

Using Lemma 1, we describe bounded ratio variation in terms of products rather than sums.

Claim 1. *An internal sequence (p_k) has bounded ratio variation if and only if*

$$\lim_n \prod_{k=1}^{n} \max\left\{\frac{p_{k+1}}{p_k}, \frac{\bar{p}_{k+1}}{\bar{p}_k}\right\} < \infty \tag{3}$$

Proof. Note that

$$\max\left\{\frac{p_{k+1}}{p_k}, \frac{\bar{p}_{k+1}}{\bar{p}_k}\right\} = 1 + \max\left\{\frac{p_{k+1}}{p_k} - 1, \frac{\bar{p}_{k+1}}{\bar{p}_k} - 1\right\} \tag{4}$$

Setting $\varepsilon_k = \max\{p_{k+1}/p_k - 1, \bar{p}_{k+1}/\bar{p}_k - 1\}$, it follows from Lemma 1 that (3) holds if and only if $\lim_n \sum_{k=1}^{n} \max\{p_{k+1}/p_k - 1, \bar{p}_{k+1}/\bar{p}_k - 1\} < \infty$, namely that (p_k) has bounded ratio variation. □

Let $(\Omega, \mu, (\pi_k)_{k\geqslant 1})$ be a learning process, $E \subseteq \Omega$ and $\omega \in \Omega$. Denote for brevity $Q_k = \pi_k(\omega)$ and $a_k = \mu(Q_k)$. The sequence $p_k = \mu(E|Q_k)$ is a monologue.

The following claim characterizes positivity of a monologue in terms of the product of ratios of the a_k's.

Claim 2. *The following statements are equivalent:*

(i) *(p_k) is a positive monologue;*

(ii) *$\lim_n a_n > 0$;*

(iii)

$$\lim_n \prod_{k=1}^{n} \frac{a_k}{a_{k+1}} < \infty. \tag{5}$$

Proof. The finite products (up to n) in (5) are equal to a_1/a_{n+1}. They converge if and only if $\lim_n a_n > 0$. This is equivalent to saying that the monologue is positive as $\mu(\bigcap Q_k) = \lim_n a_n$. □

The next claim relates ratios of p_k's with ratios of a_k's, which enables us, using Claim 1 and Claim 2, to tie together bounded ratio variation with positivity of a monologue.

Claim 3. *For each k,*

$$\frac{a_k}{a_{k+1}} \geqslant \max\left\{\frac{p_{k+1}}{p_k}, \frac{\bar{p}_{k+1}}{\bar{p}_k}\right\}. \tag{6}$$

Proof. Observe that inequality (6) holds if and only if $a_k/a_{k+1} \geqslant p_{k+1}/p_k$ and $a_k/a_{k+1} \geqslant \bar{p}_{k+1}/\bar{p}_k$. These two inequalities hold if and only if

$$0 \leqslant p_k a_k - p_{k+1} a_{k+1} \leqslant a_k - a_{k+1}. \tag{7}$$

But (7) holds because $\mu(E \cap (Q_k \setminus Q_{k+1})) = \mu(E \cap Q_k) - \mu(E \cap Q_{k+1}) = p_k a_k - p_{k+1} a_{k+1}$, and $0 \leqslant \mu(E \cap (Q_k \setminus Q_{k+1})) \leqslant \mu(Q_k \setminus Q_{k+1}) = a_k - a_{k+1}$. □

For the proof of Theorem 1 and Proposition 1 we use the same learning process $\big(\Omega, \mu, E, (\pi_k)_{k \geqslant 1}\big)$, which we call the *basic learning process*. In the basic learning process, $\Omega = \{\omega_1, \eta_1, \omega_2, \eta_2, \dots, \eta, \omega\}$. Thus, Ω consists of infinitely many states ω_k and η_k and a pair of states η and ω. We set $E = \{\omega_1, \omega_2, \dots, \omega\}$. The partitions are defined by $\pi_1 = \{\Omega\}$ and for $k > 1$,

$$\pi_k = \big\{\{\omega_1\}, \{\eta_1\}, \dots, \{\omega_{k-1}\}, \{\eta_{k-1}\}, \{\omega_k, \eta_k, \dots, \eta, \omega\}\big\}.$$

We are interested in the monologue at ω. As before, we define $Q_k = \pi_k(\omega) = \{\omega_k, \eta_k, \dots, \eta, \omega\}$ for brevity.

Proof of Theorem 1. We start by showing that every internal positive monologue has bounded ratio variation. Using the notation above, assume that $p_k = \mu(E \mid Q_k)$ is a positive monologue. Then, by Claim 2, (5) holds. Therefore, the inequality (6) implies the inequality (3). Thus, by Claim 1, (p_k) has bounded ratio variation.

Next, we prove that every internal sequence (p_k) with bounded ratio variation is a positive monologue in the basic learning process at ω. For this we need to define the probability measure μ such that $p_k = \mu(E \mid Q_k)$, and $\mu(\bigcap_k Q_k) > 0$, which means that the process is positive. Note that $\mu(Q_k) - \mu(Q_{k+1}) = \mu(\{\omega_k, \eta_k\})$, and $\lim_k \mu(Q_k) = \mu(\{\omega, \eta\})$. Thus, it is enough to define for each $k \geqslant 1$, $\mu(Q_k)$, and $\mu(\omega_k)$ and also $\mu(\omega)$.

Let $a_1 = 1$ and define by induction a sequence a_k that satisfies for every k,

$$\frac{a_k}{a_{k+1}} = \max\left\{\frac{p_{k+1}}{p_k}, \frac{\bar{p}_{k+1}}{\bar{p}_k}\right\}. \tag{8}$$

Note that (8) is obtained by replacing the inequality in (6) by an equality. The right-hand side of (8) is at least 1, hence the sequence (a_k) is weakly decreasing. Therefore, for the decreasing sequence of events Q_k, we can define $\mu(Q_k) = a_k$, and thus $\mu(Q_k \setminus Q_{k+1}) = \mu(\{\omega_k, \eta_k\}) =$

$a_k - a_{k+1}$. Since (6) is equivalent to (7), it follows that we can define $\mu(\omega_k) = p_k a_k - p_{k+1} a_{k+1}$, which is non-negative and does not exceed $\mu(Q_k \setminus Q_{k+1})$.

Let $a = \lim a_k$. By Observation 1, (p_k) has bounded variation and hence it converges. Let $p = \lim p_k$. We define $\mu(\omega) = pa$. This completes the definition of μ. Now, $\mu(E \cap Q_k) = (\sum_{i \geqslant k} p_i a_i - p_{i+1} a_{i+1}) + pa$. As $p_i a_i$ converges to pa, this sum is $p_k a_k$, implying that $\mu(E \mid Q_k) = p_k$, which means that (p_k) is the monologue at ω. To show that the monologue is positive, we note that, by Claim 1, (3) holds since (p_k) has bounded ratio variation. In virtue of (8), Claim 2 implies that $a > 0$ and as $\mu(\cap_k Q_k) = a$, this means that the monologue is positive. $\square$

Note that in the proof of Theorem 1, the sequence $a_k = \mu(Q_k)$ can be weakly decreasing. Indeed, when $p_k = p_{k+1}$, the right hand side of (8) is 1, and thus, $a_k = a_{k+1}$. In the proof of Theorem 2 we need to define a positive monologue on the basic learning process with a strictly decreasing sequence a_k. We show here that this is possible.

Lemma 2. *An internal sequence with bounded ratio variation is a positive monologue in the basic learning process at ω with a strictly decreasing sequence $\big(\mu(\pi_k(\omega))\big)$.*

Proof. Let β_k be a sequence such that for each k, $\beta_k > 1$ and $\lim_n \prod_{k=1}^{n} \beta_k < \infty$. In the proof of Theorem 1 define the sequence a_k by

$$\frac{a_k}{a_{k+1}} = \max\left\{\frac{p_{k+1}}{p_k}, \frac{\bar{p}_{k+1}}{\bar{p}_k}\right\}\beta_k. \tag{9}$$

rather than by (8). Obviously, a_k is strictly decreasing. Since (9) implies the inequality (6), the proof holds verbatim up to the point where we need to show that (5) holds in order to prove that the monologue is positive. This follows from (9), (3) and the boundedness of the product of the β's. $\square$

Proof of Observation 2. For (i) we note that if (p_k) is in $(\xi, 1-\xi)$ for some $\xi > 0$, then $|p_{k+1} - p_k|/r_k < |p_{k+1} - p_k|/\xi$. For (ii) assume that (p_k) converges monotonically to 0. We can assume that $p_k \leqslant 1/2$ for all k. The ratio variation of p_k is $\sum_k (\bar{p}_{k+1} - \bar{p}_k)/\bar{p}_k \leqslant 2\sum_k (\bar{p}_{k+1} - \bar{p}_k)$. If (p_k) converges monotonically to 1, we can assume that $p_k \geqslant 1/2$ for all k. The ratio variation of p_k is $\sum_k (p_{k+1} - p_k)/p_k \leqslant 2\sum_k (p_{k+1} - p_k)$. $\square$

Proof of Proposition 1. Let (p_k) be an internal sequence. If (p_k) has bounded ratio variation, the result follows from Theorem 1. So we assume that (p_k) does not have bounded ratio variation. We show that (p_k) is the monologue in the basic learning process at ω. We define μ for all Q_k and ω_k as in the proof of Theorem 1. Since (p_k) does not have bounded ratio variation, it follows from (8) and Claims 1 and 2 that $\lim a_k = 0$. Therefore, $\mu(\{\eta,\omega\}) = 0$. Thus, $\mu(E \cap Q_k) = \sum_{i \geqslant k}(p_i a_i - p_{i+1}a_{i+1}) = p_k a_k$, implying that $\mu(E \mid Q_k) = p_k$. $\square$

5 Proofs – dialogues

Before proving Theorem 2 and Proposition 2 we consider three auxiliary subjects:

- Proposition 3 which belongs to the theory of variation of sequences;
- Some basic lemmas concerning the probabilities of disjoint events and their intersection with another event;
- A basic joint learning process which is used in both Theorem 2 and Proposition 2.

Selection sequences and bounded variation

In the proof of part (ii) of Theorem 2 we use the equivalence of (i) and (iv) in Proposition 3. Therefore we first prove this proposition.

Proof of Proposition 3. Suppose that all the selections from the sequences (p^1_k) and (p^2_k) have bounded variation. Then in particular the two sequences (p^1_k) and (p^2_k) and the sequences (a) and (b) have bounded variation in virtue of being selections from (p^1_k) and (p^2_k). Since $|p^1_k - p^2_k| \leqslant |p^1_k - p^1_{k+1}| + |p^1_{k+1} - p^2_k|$, it follows that the series defined by (c) absolutely converges, because (p^1_k) and (b) have bounded variation. Thus (i) implies each of (ii), (iii), and (iv).

We now show that (ii), (iii) and (iv) are equivalent. Suppose first that (p^1_k), (p^2_k) and (a) have each bounded variation. Since $|p^2_k - p^1_k| \leqslant |p^2_k - p^2_{k+1}| + |p^2_{k+1} - p^1_k|$ it follows that the series defined by (c) absolutely converges. Thus, (ii) implies (iv) and similarly (iii) implies (iv). If (iv) holds, then, since $|p^2_{k+1} - p^1_k| \leqslant |p^2_{k+1} - p^2_k| + |p^2_k - p^1_k|$ it follows that (a) has bounded variation and in a similar way also (b). Thus, (iv) implies (ii) and (iii).

Finally, it is enough to show that (ii) and (iii) imply (i). Observe that any summand in a selection from (p^1_k), (p^2_k) appears in the variation of either (p^1_k), (p^2_k), (a) or (b). This shows that if these four sequences have bounded variation, then all selections have bounded variation. $\square$

Three lemmas concerning the probabilities of three events.

In the first two lemmas we study a function that we have already met in the definition of ratio variation. For every $x, y \in (0,1)$ we denote

$$\varphi(x,y) = \max\left\{\frac{y}{x} - 1, \frac{\bar{y}}{\bar{x}} - 1\right\}.$$

Note that whenever $\xi \leqslant x, y \leqslant 1-\xi$, for some $\xi > 0$,

$$\varphi(x,y) \leqslant |x-y|/\xi. \tag{10}$$

Lemma 3. *Let* $x, y \in (0,1)$*,* (Ω, μ) *be a measurable space,* $A, B \subseteq \Omega$ *be two disjoint events, and* $E \subseteq \Omega$ *be an event such that* (a) $\mu(A) > 0$, (b) $\mu(B) = 2\varphi(x,y)\mu(A)$; (c) $\mu(E \mid A) = y$; and (d) $\mu(E \cap B) = (x + \mathbb{1}_{x>y})\varphi(x,y)\mu(A)$. Then, $\mu(E \mid A \cup B) = x$.

Proof. Suppose first that $x > y$. Then,

$$\begin{aligned}\mu(E \mid A \cup B) &= \frac{\mu(A)(x\varphi(x,y) + \varphi(x,y) + y)}{\mu(A)(2\varphi(x,y)+1)} \\ &= \frac{(x-y)(1+x) + y(1-x)}{2(x-y)+1-x} = \frac{x(x-2y+1)}{x-2y+1} = x.\end{aligned}$$

Now suppose that $x \leqslant y$. Then,

$$\mu(E \mid A \cup B) = \frac{\mu(A)(x\varphi(x,y)+y)}{\mu(A)(2\varphi(x,y)+1)} = \frac{(y-x+y)x}{2(y-x)+x} = x.$$

□

Lemma 3 is of great importance for the constructive proofs that follow. Suppose that the probability of an event A, say $\mu(A)$, and the conditional probability $\mu(E \mid A)$ have already been defined. We would like to add another event B, disjoint from A, such that (i) the probability of B is $\mu(B) = 2\varphi(x,y)\mu(A)$, and (ii) the conditional probability of E given the union $A \cup B$ is equal to x. Is this possible, and if so, what should be the probability of E within B? Lemma 3 provides sufficient conditions for when this is possible. Furthermore, if

$$\mu(E \cap B) = (x + \mathbb{1}_{x>y}) 2\varphi(x,y)\mu(A), \tag{11}$$

then $\mu(E \mid A \cup B) = x$. The case where $x = y$ requires a special treatment. When $x = y$, $\varphi(x,y) = 0$, which would make $\mu(B) = 0$. In this case $\varphi(x,y)$ is replaced by $\varepsilon > 0$. We set $\mu(B) = 2\varepsilon\mu(A)$ and $\mu(E \cap B) = x \cdot 2\varepsilon\mu(A)$. In this case, the conditional probability of E given the union $A \cup B$ is equal to x. We give a name for the fixing of B and E in a way that gives rise to the result of Lemma 3:

> When we set $\mu(B) = 2\varphi(x,y)\mu(A)$ and $\mu(E \cap B) = (x + \mathbb{1}_{x>y})\varphi(x,y)\mu(A)$, we say that we apply the $(x;\mu(A),y)$*-scheme* on B.

Suppose that $\mu(A) > 0$ and $\mu(E|A) = y$. If we apply the $(x;\mu(A),y)$-scheme on B, then Lemma 3 states that $\mu(E \mid B \cup A) = x$. In the construction below we apply $(x;\mu(A),y)$-schemes only to events A whose probability is positive.

Lemma 4. *Let $x,y \in (0,1)$, (Ω,μ) be a measurable space, $A,B \subseteq \Omega$ two disjoint events, and $E \subseteq \Omega$ an event such that* (a) $\mu(A) > 0$, (b) $\mu(B) \geqslant 2\varphi(x,y)\mu(A)$; (c) $\mu(E \mid A) = y$. There exists a number z such that if $\mu(E \cap B) = z$, then $\mu(E \mid A \cup B) = x$.

Proof. In case $\mu(B) = 0$, the result is trivial. Assume then that $\mu(B) > 0$ and let

$$z = \frac{\mu(B)x + \mu(A)(x-y)}{\mu(B)}.$$

We show first that $z \geqslant 0$. It is clear when $x \geqslant y$. If $x < y$, then $\varphi(x,y) = (y-x)/x$ and $\mu(B)x + \mu(A)(x-y) \geqslant ((y-x)/x)x + \mu(A)(x-y) = (1-\mu(A))(y-x) \geqslant 0$. Thus, $z \geqslant 0$. We now show that $z \leqslant 1$. In case $x \leqslant y$, this is clear. When $x > y$, $\mu(B) \geqslant 2\varphi(x,y)\mu(A) \geqslant \mu(A)(x-y)/(1-x)$. Therefore, $z = x + \mu(A)(x-y)/\mu(B) \leqslant x + 1 - x = 1$. To complete the proof assume that $\mu(E \mid B) = z$. Then,

$$\mu(E \mid A \cup B) = \frac{\mu(B)\mu(E|B) + \mu(A)\mu(E|A)}{\mu(B) + \mu(A)} = \frac{\mu(B)x + \mu(A)(x-y) + \mu(A)y}{\mu(B) + \mu(A)} = x.$$

□

The next lemma will be used for estimating the growth in the total weight during the induction process.

Lemma 5. *Let $D = D_1 \cup D_2$ and E be events with $D_1 \cap D_2 = \emptyset$ and let μ be a measure. Suppose that $\mu(E \mid D_1) = p$, $\mu(E \mid D_2) = z$ and $\mu(E \mid D) = q$. Then, for every p',*

$$\mu(D_1 \mid D)|p - p'| + \mu(D_2 \mid D)|z - p'| \leqslant |p - p'| + |q - p|.$$

Proof. By assumption, $q = \mu(D_1 \mid D)p + \mu(D_2 \mid D)z = (1 - \mu(D_2 \mid D))p + \mu(D_2 \mid D)z$. Thus, $\mu(D_2 \mid D)z = q + \mu(D_2 \mid D)p - p$. Hence,

$$\begin{aligned} & \mu(D_1 \mid D)|p - p'| + \mu(D_2 \mid D)|z - p'| \\ = \ & \mu(D_1 \mid D)|p - p'| + |q + \mu(D_2 \mid D)p - p - \mu(D_2 \mid D)p'| \\ \leqslant \ & \mu(D_1 \mid D)|p - p'| + |q - p| + \mu(D_2 \mid D)|p - p'| = |p - p'| + |q - p|. \end{aligned}$$

□

The basic joint learning process.

We consider a joint learning process $(\Omega, \mu^1, \mu^2, E, (\pi^1_k)_{k\geqslant 1}, (\pi^2_k)_{k\geqslant 1})$ defined as follows. The state space is $\Omega = \{\omega_{i,j}, \eta_{i,j} \mid i, j \in [1,\ldots,\infty]\}$. We call the events $C_{i,j} = \{\omega_{i,j}, \eta_{i,j}\}$ *cells*. Define, $E_{i,j} = \{\eta_{i,j}\}$ and set $E = \bigcup_{i,j} E_{i,j}$. Thus, $\eta_{i,j}$ is the only common state to E and $C_{i,j}$. We denote $\text{Row}(m, \vec{k}) = \bigcup_{j\geqslant k} C_{m,j}$ and call it the *k-truncated m-row*. Similarly, the *k-truncated m-column* is $\text{Col}(\vec{k}, m) = \bigcup_{i\geqslant k} C_{i,m}$. We set $\omega = \omega_{\infty,\infty}$, and study the dialogue at ω. We assume that the agents have a CP, that is, $\mu^1 = \mu^2 = \mu$.

We assume that for two sequences (p^1_k) and (p^1_k) the CP μ satisfies the following conditions. For every $k = 1, 2, ...$ and for every $i, j \in [k+1, \infty]$,

$$\mu\big(E \mid \text{Row}(i, \vec{k})\big) = p^1_k \quad \text{and} \quad \mu\big(E \mid \text{Col}(\vec{k}, j)\big) = p^2_k, \tag{12}$$

and

$$\mu\big(E \mid \text{Row}(k, \vec{k})\big) \neq p^1_k \quad \text{and} \quad \mu\big(E \mid \text{Col}(\vec{k}, k)\big) \neq p^2_k. \tag{13}$$

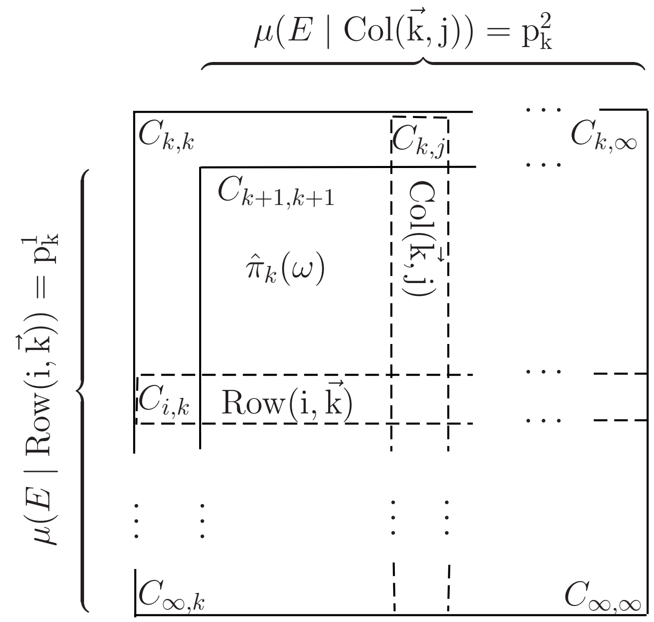
Note that (12) can be equivalently written in a different order of quantification: for $i = 1, 2, \ldots$ and for every $k \leqslant i - 1$, $\mu(E \mid \text{Row}(i, \vec{k})) = p^1_k$, and similarly for agent 2.

The initial partition of player 1, π^1_1, consists of all the 1-truncated rows $\text{Row}(i, \vec{1})$ for $i \in [1,\ldots\infty]$. For player 2, π^2_1 consists of all the 1-truncated columns $\text{Col}(\vec{1}, j)$ for $j \in [1,\ldots\infty]$. The rest of the partitions are defined endogenously as described in Section 3. In stage $k \geqslant 1$,

$\mathrm{Row}(i,\vec{k}))$ for $i \geqslant k$ are elements of the partition π_k^1, and $\mathrm{Col}(\vec{k},j)$ for $j \geqslant k$ are elements of π_k^2. By equations (12) and (13), the event that the pair of posteriors is $(p_k^1(\omega), p_k^2(\omega))$ is

$$
\begin{aligned}
\hat{\pi}_k(\omega) &= \Big(\bigcup_{i\geqslant k+1} \mathrm{Row}(i,\vec{k})\Big) \bigcap \Big(\bigcup_{j\geqslant k+1} \mathrm{Col}(\vec{k},j)\Big) \\
&= \bigcup_{i\geqslant k+1} \mathrm{Row}(i,\overrightarrow{k+1}) \\
&= \bigcup_{j\geqslant k+1} \mathrm{Col}(\overrightarrow{k+1},j).
\end{aligned}
$$

Thus, $\mathrm{Row}(i,\overrightarrow{k+1})$ for $i \geqslant k+1$ are elements of the partition π_{k+1}^1 and $\mathrm{Col}(\overrightarrow{k+1},j)$ for $j \geqslant k+1$ are elements of the partition π_{k+1}^2. It is common knowledge in $\hat{\pi}_k(\omega)$ that the posteriors at stage k are $(p_k^1(\omega), p_k^2(\omega))$. Stage k of the process is depicted in Figure 1.



The truncated rows $\mathrm{Row}(i,\vec{k})$ and columns $\mathrm{Col}(\vec{k},j)$, for $i,j \geqslant k$, are elements of the partitions π_k^1 and π_k^2, respectively. The conditional probability of E is p_k^1 in all these truncated rows, but the first, and p_k^2 in all these truncated columns, but the first. Thus the posterior probabilities of E are (p_k^1, p_k^2) in all the states of $\hat{\pi}_k(\omega)$, which consists of the cells $C_{i,j}$ with $i,j \geqslant k+1$. When $\hat{\pi}_k(\omega)$ becomes common knowledge, the truncated rows $\mathrm{Row}(i,\overrightarrow{k+1})$ and columns $\mathrm{Col}(\overrightarrow{k+1},j)$, for $i,j \geqslant k+1$, become elements of π_{k+1}^1 and π_{k+1}^2, respectively. The sequence (p_k^1, p_k^2) is the dialogue at $\omega = \omega_{\infty,\infty}$.

Figure 1: The basic joint learning process in stage k

Proof of Theorem 2.

We start with the short proof of part (i) of the theorem.

Proof of part (i) of Theorem 2. Suppose that the pair (p_k^1) and (p_k^2) forms a dialogue at ω with a common prior μ. We show that every selection of these sequences has bounded ratio variation. Consider a sequence $i(k)$ of names in $\{1,2\}$. We define a decreasing sequence of events Q_k such that $Q_k \subseteq \pi_k(\omega)$, where $\pi_k(\omega)$ is the element of the meet of π_k^1 and π_k^2 that contains ω. We define Q_k as follows:

$$Q_k = \{\omega \mid \pi_k^{i(k)}(\omega) \subseteq \pi_k(\omega) \text{ and } \mu(E \mid \pi_k^{i(k)}) = p_k^{i(k)}\}.$$

In other words, the event Q_k is the union of all the elements of the partition $\pi_k^{i(k)}$ contained in $\pi_k(\omega)$ in which agent $i(k)$ assigns probability $p_k^{i(k)}$ to E. Since $Q_{k+1} \subseteq \pi_{k+1}(\omega) \subseteq Q_k$, it follows that the sequence Q_k is decreasing. Thus, the sequence $p_k^{i(k)}$ is the monologue at ω generated by the sequence Q_k and the common prior μ. By construction, $\bigcap_k \pi_{k+1}(\omega) \subseteq \bigcap_k Q_k$. Since the dialogue is positive, $\mu(\bigcap_k \pi_{k+1}(\omega)) > 0$. Thus, the monologue $p_k^{i(k)}$ is positive as well. By Theorem 1 and Observation 2, the sequence $p_k^{i(k)}$ has bounded ratio variation. $\square$

Proof of part (ii) of Theorem 2. Assume that two strictly internal sequences (p_k^1) and (p_k^2) satisfy the condition in part (ii) of the theorem. The sequences are strictly internal and thus one can find $\xi > 0$ such that $\xi < p_k^i < 1-\xi$ for every k and $i = 1,2$. We use the basic joint learning process and construct the CP μ such that the two sequences form a positive dialogue at ω. In view of Proposition 3, we can assume that $\sum_{k=0}^{\infty} |p_k^1 - p_k^2| < \infty$.

A sketch of the construction of μ.

We construct the probabilities backward. Define, $H_{i,j} = \bigcup \mathrm{Row}_{k \geqslant i}(k, \overrightarrow{j})$. This is the bottom-right corner whose top-left cell is $C_{i,j}$. Suppose that all the probabilities of $C_{i,j}$ and $E_{i,j}$ in $H_{k+1,k+1}$ have been defined. We want to extend the definition to $H_{k,k}$. This is done is Steps 4–6 below. We first define the probabilities on the k-th column, $\mathrm{Col}(\overrightarrow{k+1},k)$. By doing it, we ensure that the conditional probability of E on every row $\mathrm{Row}(i,\overrightarrow{k})$ is p_k^1, $i \geqslant k+1$. We then define the probability of the cells on $\mathrm{Row}(k,\overrightarrow{k+1})$. This is done in a way that makes the probability of the event E conditional on every column starting at $C_{k,j}, j \geqslant k+1$ equal to p_k^2.

Finally, the probability on the $C_{k,k}$ is defined. The objective here is to make the probability of E conditional on the row $\mathrm{Row}(k,\overrightarrow{k})$ different from p_k^1, and at the same time the probability of E conditional on the column $\mathrm{Col}(\overrightarrow{k},k)$ different from p_k^2. This point in the construction guarantees, for instance, that when agent 1 announces p_k^1 at time k, agent 2 knows that the event related to the row $\mathrm{Row}(k,\overrightarrow{k})$ did not occur and he updates his belief accordingly.

An important issue in the construction is to control the size of $\mathrm{Col}(\overrightarrow{k+1},k)$, $\mathrm{Row}(k,\overrightarrow{k})$ and $C_{k,k}$ added in the induction process. It should not grow too fast, the reason being that when we define the measure on these events, we add a weight to each and by the end of the process, we normalize the measure obtained. If the added weight is too large, the normalized probabilities at the end of the process might be very small and eventually vanish.

The bounded variation of the sequences involved is the property that ensures that the normalizing factors are uniformly bounded. The probabilities in the limit, including the conditional ones, are therefore well defined.

The construction of μ.

We proceed in two stages. In the first stage we define for every integer ℓ a measure μ^ℓ on Ω. It will not necessarily be a probability measure. The measure μ^ℓ will be defined inductively in Steps 0–6 below. At each stage, the measure of more cells will be introduced. The added measure will be called also the weight or size added.

The idea of the construction is to add inductively weight to more and more cells without taking care of the total weight. Only at the end of the inductive process μ^ℓ is normalized in order to obtain a probability measure $\overline{\mu}^\ell$. Note that the conditional measures do not change after normalization. In the second stage a measure μ will be defined as a limit of the sequence $(\overline{\mu}^\ell)$. One of the objectives of the construction is to make sure that the the sequence $(\overline{\mu}^\ell)$ is not vanishing in the limit. That is, μ is indeed a measure.

Stage 1: Defining the measure μ^ℓ. Fix an integer ℓ. During the construction we are going to define a few arrays of weights, not necessarily probabilities, and conditional probabilities:

$$c_{i,j}^\ell := \mu^\ell(C_{i,j}) = \mu^\ell(\omega_{i,j},\eta_{i,j});$$

$$\alpha_{i,j}^\ell := \mu^\ell(E|C_{i,j}) = \mu^\ell(\eta_{i,j})/\mu^\ell(\omega_{i,j},\eta_{i,j});$$

$$d_{i,j}^\ell := \mu^\ell(\mathrm{Col}(\overrightarrow{i},j));$$

$$\gamma_{i,j}^\ell := \mu^\ell(E|\mathrm{Col}(\overrightarrow{i},j));$$

$$r_{i,j}^\ell := \mu^\ell(\mathrm{Row}(i,\overrightarrow{j}));$$

$\rho^{\ell}_{i,j} := \mu^{\ell}(E|\mathrm{Row}(i, \overrightarrow{j}));$

During the construction, for every $i, j < \infty$, we will take care to keep $c^{\ell}_{i,j}$ (across ℓ) away from 0. The reason is that in the second stage a converging subsequence (as ℓ goes to infinity) will define μ, and we want to make sure that limit $\lim_{\ell\to\infty} c^{\ell}_{i,j}$ does not vanish. This, in turn, will guarantee that the conditional probabilities are well defined.

Step 0: Defining the probabilities on the margins. We start with the weights on the right margin, $\mathrm{Col}(\overrightarrow{1}, \infty)$: $(\alpha^{\ell}_{i,\infty})$ and $(c^{\ell}_{i,\infty})$, $i = 1, \ldots, \infty$. By assumption, (p^2_k) is a positive monologue. That is, one can find a sequence of decreasing events (Q_k), an event E and a measure ν_2 such that $\nu_2(\bigcap_k Q_k) > 0$ and $(p^2_k) = (\nu_2(E|Q_k))$.

Define $c_{\infty,\infty} = \nu_2(\bigcap_k Q_k)$ and $\alpha^{\ell}(\eta_{\infty,\infty}) = (\mu^{\ell}(\eta_{\infty,\infty}))/(\mu^{\ell}(\eta_{\infty,\infty}, \omega_{\infty,\infty})) = \lim_k p^2_k$. This takes care of the limit cell in which the conditional probability is $\lim_k p^2_k$. Next define, $c_{k,\infty} = \nu_2(Q_k \setminus Q_{k+1})$ and $\mu^{\ell}(\eta_{k,\infty})) = \nu_2(E \cap (Q_k \setminus Q_{k+1}))$ (i.e., $\alpha^{\ell}_{k,\infty} = \nu_2(E \cap (Q_k \setminus Q_{k+1})) \cdot c_{k,\infty}$). By Lemma 2, $c_{k,\infty} > 0$ for every $k = 1, 2, \ldots$.

We turn to the probabilities defined on the bottom margin, $\mathrm{Row}(\infty, \overrightarrow{1})$: $(\alpha^{\ell}_{\infty,j})$ and $(c^{\ell}_{\infty,j})$. As (p^1_k) is a positive monologue, we can find a sequence of decreasing events (Q'_k), an event E' and a measure ν_1 such that $\nu_1(\bigcap_k Q'_k) > 0$ and $(p^1_k) = (\nu_1(E|Q'_k))$ is a monologue concerning E' with respect to (Q'_k). Without loss of generality[6] $\nu_1(\bigcap_k Q'_k) = \nu_2(\bigcap_k Q_k)$. We now define probabilities on the bottom margin in a manner similar to that of the right margin. These definitions guarantee that (a) $c_{\infty,k} > 0$ for every $k = 1, 2, \ldots$ and (b) (12) holds for $i, j = \infty$.

Note that the weights on the last row and column do not depend on ℓ. Denote by M their total size, i.e., $M := r^{\ell}_{\infty,1} + d^{\ell}_{1,\infty} - c^{\ell}_{\infty,\infty}$.

We proceed with the other cells. For any $\ell + 1 \leqslant i, j < \infty$, set $c^{\ell}_{i,j} = \alpha^{\ell}_{i,j} = 0$. That is, the bottom-right corner $H_{\ell+1,\ell+1}$, excluding the right and bottom margins, gets total weight 0. Steps 1–3 define the probabilities over $H_{\ell,\ell}$. For the following definitions recall the definition of the scheme introduced above.

Step 1: Defining the measure on the truncated column $\mathrm{Col}(\overrightarrow{\ell+1}, \ell)$**.** Fix $i \geqslant \ell + 1$. Apply the $(p^1_{\ell}; r^{\ell}_{i,\ell+1}, \rho^{\ell}_{i,\ell+1})$-scheme on $C_{i,\ell}$.

What is the total size of the ℓ-th column just added? From Step 0 we deduce that $r^{\ell}_{i,\ell+1} = c^{\ell}_{i,\infty}$

[6] This is so because otherwise we can assume, without loss of generality, that $\nu_1(\bigcap_k Q'_k) > \nu_2(\bigcap_k Q_k)$. By redefining the conditional probabilities $(\nu_1(E|Q'_k))$, actually by multiplying them all by $\nu_2(\bigcap_k Q_k)/\nu_1(\bigcap_k Q'_k)$, one can make their limit equal to $\nu_2(\bigcap_k Q_k)$.

and $\rho^\ell_{i,\ell+1} = \alpha^\ell_{i,\infty}$. The scheme dictates that the size of $C_{i,\ell}$ is $2\varphi(p^1_\ell, \alpha^\ell_{i,\infty})r^\ell_{i,\ell+1}$. This is bounded by $2r^\ell_{i+1,\ell}/\xi$ (recall that $\xi < p^1_\ell < 1-\xi$). Thus, the total weight added (due to all the cells $C_{i,\ell}, i \geqslant \ell+1$) is bounded by $2M/\xi$.

Step 2: Defining the measure on the truncated row $\mathrm{Row}(\ell, \overrightarrow{\ell+1})$**.** Fix $j \geqslant \ell+1$ and apply the $(p^2_\ell; d^\ell_{\ell+1,j}, \gamma^\ell_{\ell+1,j})$-scheme on the cell $C_{\ell,j}$. Again, the total added weight is bounded by $2M/\xi$.

Step 3: Defining the measure on the diagonal cell $C_{\ell,\ell}$**.** The diagonal cell requires a special treatment. We have to define the probabilities $\mu^\ell(C_{\ell,\ell})$ and $\mu^\ell(E \cap C_{\ell,\ell})$ in a way that (a) $\rho^\ell_{i,\ell+1}$ is close to p^1_ℓ; and (b) (13) holds for $k=\ell$. That is, $\rho^\ell_{\ell,\ell}$ should be different from p^1_ℓ and $\gamma^\ell_{\ell,\ell}$ should be different from p^2_ℓ. Here, we choose $\rho^\ell_{\ell,\ell}$ and $\gamma^\ell_{\ell,\ell}$ to be bounded away (across different ℓ's), respectively, from p^1_ℓ and p^2_ℓ.[7]

Let (ε_ℓ) be a sequence of positive numbers such that $\sum_\ell \varepsilon_\ell < \infty$ and $\varepsilon_\ell < \xi/2$ for every ℓ. Let $\tilde{q}^\ell_\ell$ be in the interval $[p^1_\ell + \varepsilon_\ell, p^1_\ell + 2\varepsilon_\ell]$. Its precise value will be determined shortly. Note that since $\varepsilon_\ell < \xi/2$, we have $0 < \tilde{q}^\ell_\ell < 1$). Let

$$\mu^\ell(C_{\ell,\ell}) = \max\left\{2\varphi(\tilde{q}^\ell_\ell; r^\ell_{\ell,\ell+1}, \rho^\ell_{\ell,\ell+1}), \varepsilon_\ell\right\}.$$

We now use Lemma 4 with $A = \mathrm{Row}(\ell, \overrightarrow{\ell+1})$, $y = \rho^\ell_{\ell,\ell+1}$ and $B = C_{\ell,\ell}$. The lemma states that there is z such that if $\mu^\ell(\varepsilon_{\ell,\ell})/\mu^\ell(C_{\ell,\ell}) = z$, then $\rho^\ell_{\ell,\ell} = \tilde{q}^\ell_\ell$.

The probability $\mu^\ell(C_{\ell,\ell})$ induces also the conditional probability on $\mathrm{Col}(\overrightarrow{\ell}, \ell)$. Our goal is to have it different from p^2_k by at least ε^2_ℓ. Since, $\mu^\ell(C_{\ell,\ell})$ is at least $\varepsilon_\ell = \tilde{q}^\ell_\ell$ when we choose $\tilde{q}^\ell_\ell$ in the interval $[p^2_\ell + \varepsilon_\ell, p^2_\ell + 2\varepsilon_\ell]$ and move from one end of the interval to the other, the probability of E conditional on $\mathrm{Col}(\overrightarrow{\ell}, \ell)$ is changing by at least ε^2_ℓ (one ε_ℓ for the size of $\mu^\ell(C_{\ell,\ell})$ and the other because the size of the interval is ε_ℓ). We conclude that by a proper choice of $\tilde{q}^\ell_\ell$ we have that $|\rho_{\ell,\ell} - p^1_\ell| \geqslant \varepsilon_\ell$ and $|\gamma_{\ell,\ell} - p^2_\ell| \geqslant \varepsilon^2_\ell/2$.

The added weight in this step is $\mu^\ell(C_{\ell,\ell})$, which is bounded by $\max\left\{2\varphi(\tilde{q}^\ell_\ell, \rho^\ell_{\ell,\ell+1})r^\ell_{\ell,\ell+1}, \varepsilon_\ell\right\}$. By Step 1, $r^\ell_{\ell,\ell} \leqslant 2M/\xi$ and therefore $2\varphi(\tilde{q}^\ell_\ell; \rho^\ell_{\ell,\ell+1}, r^\ell_{\ell,\ell+1}) \leqslant 4M/(\xi^2)$. Thus, the added weight in Step 3 is bounded by $\max\left\{4M/\xi^2, \varepsilon_\ell\right\}$. Since the sequence (ε_k) is bounded, we obtain that

[7] The reason is that in stage 2, at the end of this proof, we let ℓ go to infinity. This makes sure that the probabilities in the limit are still different.

the total weight added in Steps 0–3 is bounded by M_1 (which does not depend on ℓ). That is,

$$M + 4M/\xi + \max\left\{4M/\xi^2, \varepsilon_\ell\right\} \leqslant M_1. \tag{14}$$

To summarize, on $H_{\ell,\ell}$ the conditional probabilities satisfy $\rho^\ell_{i,\ell} = p^1_\ell$ for $i \geqslant \ell+1$, $\gamma^\ell_{\ell,j} = p^2_\ell$ for $j \geqslant \ell+1$, $\rho^\ell(\ell,\ell) = \tilde{q}^\ell_\ell \geqslant p^1_\ell + \varepsilon_\ell$ and $|\gamma^\ell_{\ell,\ell} - p^1_\ell| \geqslant (\varepsilon_\ell)^2$. The conditional probability of E given $H_{\ell,\ell}$ is a convex combination of p^2_ℓ and $\tilde{q}^\ell_\ell$ and we denote it by q^ℓ_ℓ.

We now continue the definitions of μ^ℓ on all other cells by a backward induction. Suppose that all $c^\ell_{i,j}$ and $\alpha^\ell_{i,j}$, $k+1 \leqslant i,j$, have been defined. The inductive procedure has three steps that are analogous to Steps 1–3.

Step 4: Defining the measure on the truncated column $\mathrm{Col}(\overrightarrow{k+1},k)$**.** For each $i \geqslant k+1$, apply the $(p^1_k; r^\ell_{i,k+1}, \rho^\ell_{i,k+1})$-scheme on $C_{i,k}$.

Step 5: Defining the measure the truncated row $\mathrm{Row}(k,\overrightarrow{k+1})$**.** For each $j \geqslant k+1$, apply the $(p^2_k; d^\ell_{k+1,j}, \gamma^\ell_{k+1,j})$-scheme on $C_{k,j}$.

Step 6: Defining the measure on the diagonal cell $C_{k,k}$**.** We iterate the construction in Step 3. The weights $\mu^\ell(C_{k,k})$ and $\mu^\ell(E_{k,k})$ are defined in such a way that (a) $\rho^\ell_{i,k+1}$ is close, but not equal to p^1_k; and (b) $\gamma^\ell_{k,k}$ is different from p^2_k. That is, (13) holds for k. Moreover, we choose the measures so that $\rho^\ell_{k,k}$ and $\gamma^\ell_{k,k}$ are bounded away across different ℓ's, respectively, from p^1_k and p^2_k. This will avoid a coincidence as ℓ goes to infinity.

We choose $\tilde{q}^\ell_k$ in the interval $[p^1_k + \varepsilon_k, p^1_k + 2\varepsilon_k]$. Its precise value will be determined shortly. Note that since $\varepsilon_k < \xi/2$, we have $0 < \tilde{q}^\ell_k < 1$. Let

$$\mu^\ell(C_{k,k}) = \max\left\{2\varphi(\tilde{q}^\ell_k; r^\ell_{k,k+1}, \rho^\ell_{k,k+1}), \varepsilon_k\right\} \cdot r^\ell_{k,k+1}. \tag{15}$$

We now use Lemma 4 with $A = \mathrm{Row}(k,\overrightarrow{k+1})$, $y = \rho^\ell_{k,k+1}$ and $B = C_{k,k}$. The lemma states that there is z such that if $\mu^\ell(E_{k,k} \mid C_{k,k}) = z$, then $\rho^\ell_{k,k} = \tilde{q}^\ell_k$.

When setting $\mu^\ell(C_{k,k})$, the induced conditional probability of E given $\mathrm{Col}(\overrightarrow{k},k)$ is determined as well. Our goal is to have it different from p^2_k by at least ε^2_k. Since, $\mu^\ell(C_{k,k})$ is at least $\varepsilon_k \cdot r^\ell_{k,k+1}$ when we choose $\tilde{q}^\ell_\ell$ in the interval $[p^1_k + \varepsilon_k, p^1_k + 2\varepsilon_k]$ and move from one end of the interval to the other, the probability of E given $\mathrm{Col}(\overrightarrow{k},k)$ (recall, it is denoted $\gamma^\ell_{k,k}$) is changing by at least ε^2_ℓ (as is Step 3, one ε_k for the size of $\mu^\ell(C_{k,k})$ and the other because the size of the interval is ε_k). We obtain that by a proper choice of $\tilde{q}_k^\ell$ we have

$$|\rho_{k,k}^\ell - p_k^1| \geqslant \varepsilon_k, \ \text{ and, } \ |\gamma_{k,k}^\ell - p_k^2| \geqslant \varepsilon_k^2/2. \tag{16}$$

It is important to note that the bounds in (16) are independent of ℓ. This implies that when we take the limits as $\ell \to \infty$, these bounds stay untouched.

Due to Steps 4 and 6, the conditional probability of E given $H_{k,k}$ is a convex combination of p_k^1 and $\tilde{q}_k^\ell$ and we denote it by q_k^ℓ. Since, $|\tilde{q}_k^\ell - p_k^1| \leqslant 2\varepsilon_k$, we obtain that $|q_k^\ell - p_k^1| \leqslant 2\varepsilon_k$,

By how much the total weight has increased? We start by estimating the weight increase due to Step 4, which takes care of $\mathrm{Col}(\overrightarrow{k+2},k)$ and of $C_{k+1,k}$. We start with the estimation of the weight of $\mathrm{Col}(\overrightarrow{k+2},k)$. For this purpose we use Lemma 5 with $D = H_{k+1,k+1}$, $D_1 = H_{k+2,k+1}$ and $D_2 = \mathrm{Row}(k+1,\overrightarrow{k+1})$. Using the notation of this lemma, we have $p = p_{k+1}^1, z = \rho_{k+1,k+1}^\ell$ and $q = q_{k+1}^\ell$. Letting $p' = p_k^1$, we obtain

$$\begin{aligned} &\mu^\ell(D_1 \mid D)|p_{k+1}^1 - p_k^1| + \mu^\ell(D_2 \mid D)|z - p_k^1| \\ &\qquad\qquad \leqslant |p_{k+1}^1 - p_k^1| + |q_{k+1}^\ell - p_{k+1}^1|. \end{aligned} \tag{17}$$

In the inductive construction, $\rho_{i,k+1} = p_{k+1}^1$ when $i \geqslant k+2$. Step 4 states that the cell $C_{i,k}$ should have the size $2\varphi(p_k^1, p_{k+1}^1) r_{i,k+1}^\ell$. Therefore the total size of $\mathrm{Col}(\overrightarrow{k+2},k)$ is $\mu^\ell(D_1) \cdot 2\varphi(p_k^1, p_{k+1}^1)$. The cell $C_{k+1,k}$ is also defined in Step 4. Its size, $\mu^\ell(C_{k+1,k})$, is $\mu^\ell(D_2) \cdot 2\varphi(p_k^1, z)$.

We conclude that the total weight added in Step 4 is

$$\begin{aligned} & \mu^\ell(D_1) \cdot 2\varphi(p_k^1, p_{k+1}^1) + \mu^\ell(D_2) \cdot 2\varphi(p_k^1, z) \\ \leqslant\ & 2\mu^\ell(D) \left[\mu^\ell(D_1 \mid D)|p_{k+1}^1 - p_k^1| + \mu^\ell(D_2 \mid D)|z - p_k^1| \right] / \xi \\ \leqslant\ & 2\mu^\ell(D) \left[|p_{k+1}^1 - p_k^1| + |q_{k+1}^\ell - p_{k+1}^1| \right] / \xi \\ =\ & 2\mu^\ell(H_{k+1,k+1}) \left[|p_{k+1}^1 - p_k^1| + |q_{k+1}^\ell - p_{k+1}^1| \right] / \xi. \end{aligned} \tag{18}$$

The first inequality holds by (10) and the second inequality is by (17).

We apply a similar calculation to determine the growth of μ^ℓ due to Step 5. We obtain that this growth is bounded by

$$2\mu^\ell(H_{k+1,k+1}) \left[|p_{k+1}^2 - p_k^2| + |q_{k+1}^\ell - p_{k+1}^2| \right] / \xi. \tag{19}$$

Recall that we defined the weight of the cell $C_{k,k}$ in (15). In case $2\varphi(\tilde{q}^{\ell}_k; r^{\ell}_{k,k+1}, \rho^{\ell}_{k,k+1}) \geqslant \varepsilon_k$, we define $D = H_{k,k+1}$, $D_1 = H_{k+1,k+1}$ and $D_2 = \mathrm{Row}(k, \overrightarrow{k+1})$. We get $p = q^{\ell}_{k+1}, z = \gamma^{\ell}_{k,k+1}$ and $q = p^2_k$. Thus, by Lemma 5

$$\begin{aligned}
&\mu^{\ell}(D_2)\cdot 2\varphi(\widetilde{q}^{\ell}_k, z) \leqslant 2\mu^{\ell}(D)\mu^{\ell}(D_2|D)|z-\widetilde{q}^{\ell}_k|/\xi \\
\leqslant\ & 2\mu^{\ell}(D)\left[\mu^{\ell}(D_1|D)|q^{\ell}_{k+1}-\widetilde{q}^{\ell}_k| + \mu^{\ell}(D_2|D)|z-\widetilde{q}^{\ell}_k|\right]/\xi \\
\leqslant\ & 2\mu^{\ell}(D)\left[|q^{\ell}_{k+1}-\widetilde{q}^{\ell}_k| + |q^{\ell}_{k+1}-p^2_k|\right]/\xi \\
=\ & 2\mu^{\ell}(H_{k+1,k+1})\left[|q^{\ell}_{k+1}-\widetilde{q}^{\ell}_k| + |q^{\ell}_{k+1}-p^2_k|\right]/\xi.
\end{aligned} \tag{20}$$

In the other case where $2\varphi(\tilde{q}^{\ell}_k; r^{\ell}_{k,k+1}, \rho^{\ell}_{k,k+1}) < \varepsilon_k$, we employ the estimation related to Step 5. This bound is given by (19). We see that the added weight in step 6 is bounded by

$$2\mu^{\ell}(H_{k+1,k+1})\max\left\{\left[|q^{\ell}_{k+1}-\widetilde{q}^{\ell}_k| + |q^{\ell}_{k+1}-p^2_k|\right], \left[|p^2_{k+1}-p^2_k| + |q^{\ell}_{k+1}-p^2_{k+1}|\right]\right\}/\xi. \tag{21}$$

To summarize, the total weight defined in Steps 4–6 is bounded from above by total weights added in (18), (19) and (21), which is

$$\begin{aligned}
2\mu^{\ell}(H_{k+1,k+1})\Big(&\left[|p^1_{k+1}-p^1_k| + |q^{\ell}_{k+1}-p^1_{k+1}|\right] + \left[|p^2_{k+1}-p^2_k| + |q^{\ell}_{k+1}-p^2_{k+1}|\right] \\
&+\max\left\{\left[|q^{\ell}_{k+1}-\widetilde{q}^{\ell}_k| + |q^{\ell}_{k+1}-p^2_k|\right], \left[|p^2_{k+1}-p^2_k| + |q^{\ell}_{k+1}-p^2_{k+1}|\right]\right\}\Big)/\xi.
\end{aligned} \tag{22}$$

When we start with Steps 0–3 and add up the their added weights, we obtain that the total weight added in all Steps 0-6 is bounded from above by

$$\begin{aligned}
M_1\Pi^{\ell}_{k=1}\Big(1+\Big(&\left[|p^1_{k+1}-p^1_k| + |q^{\ell}_{k+1}-p^1_{k+1}|\right] + \left[|p^2_{k+1}-p^2_k| + |q^{\ell}_{k+1}-p^2_{k+1}|\right] \\
&+\max\left\{\left[|q^{\ell}_{k+1}-\widetilde{q}^{\ell}_k| + |q^{\ell}_{k+1}-p^2_k|\right], \left[|p^2_{k+1}-p^2_k| + |q^{\ell}_{k+1}-p^2_{k+1}|\right]\right\}\Big)/\xi\Big).
\end{aligned} \tag{23}$$

We show that these products have a uniform bound. In other words, there is a constant W such that the quantity in (23) is bounded by W, for every ℓ. We do it by using Lemma 1 and showing that each of the sequences involved in (23) has bounded variation which is independent of ℓ.

In order to verify it, we have to check only the total variation of the following sequences:

$$|q^{\ell}_{k+1}-p^1_{k+1}| \leqslant 2\varepsilon_{k+1}.$$

$$|q^{\ell}_{k+1}-p^2_{k+1}| \leqslant |q^{\ell}_{k+1}-p^1_{k+1}| + |p^1_{k+1}-p^2_{k+1}| \leqslant 2\varepsilon_{k+1} + |p^1_{k+1}-p^2_{k+1}|.$$

$$|q^{\ell}_{k+1} - \widetilde{q}^{\ell}_{k}| \leqslant |q^{\ell}_{k+1} - p^{1}_{k+1}| + |p^{1}_{k+1} - p^{1}_{k}| + |p^{1}_{k} - \widetilde{q}^{\ell}_{k}| \leqslant 4\varepsilon_{k+1} + |p^{1}_{k+1} - p^{1}_{k}|.$$

$$|q^{\ell}_{k+1} - p^{2}_{k}| \leqslant |q^{\ell}_{k+1} - p^{2}_{k+1}| + |p^{2}_{k+1} - p^{2}_{k}| \leqslant 2\varepsilon_{k+1} + |p^{1}_{k+1} - p^{2}_{k+1}| + |p^{2}_{k+1} - p^{2}_{k}|.$$

Note that the right-hand sides of all these inequalities do not depend on ℓ. Moreover, by the assumptions of (ii) of Theorem 2, the sequences on the right-hand sides have bounded variation. Thus, there is a universal constant W, which bounds from above the quantity in (23) for every ℓ. One can therefore normalize the measure μ^{ℓ} to obtain a probability measure on Ω that satisfies (12), (13) and (16) for any $k \leqslant \ell$. We denote the normalized measure by $\overline{\mu}^{\ell}$.

Stage 2: Taking a limit of $\overline{\mu}^{\ell}$. In order to define the measure μ we take a converging subsequence of [8] $\overline{\mu}^{\ell}$ and denote it by μ. It is important to note that over the margins the measure μ^{ℓ} does not depend on ℓ: it was defined in Step 0 once and for all. Moreover, the μ^{ℓ}-measure of every cell in the margins is positive. After normalization, since the normalizing factors across ℓ are smaller than W, the $\overline{\mu}^{\ell}$-probability of every cell in the margins is bounded away from zero. Hence, the μ-probability of every cell in the margins is strictly positive. Since the μ-measures of the cells in the margins are all positive, the conditional probabilities discussed above are all well defined. In particular, all the conditional probabilities in (12) are well defined and moreover, the equalities in (12) are satisfied by μ. Furthermore, due to (16), the conditional probabilities on the diagonal cells are kept bounded away from their respective probabilities. Therefore, the inequalities in (13) are also satisfied by μ.

It may be the case that the limit measure μ is not a probability measure. However, since the margin cells have a positive measure, μ is not zero. We can now normalize μ, if needed, in order to obtain a probability measure. This completes the proof. □

Proof of Proposition 2. In order to prove part (i), let (p^{1}_{k}) and (p^{2}_{k}) be two internal sequences. We use the basic joint learning process and construct the CP μ such that the two sequences form a positive dialogue at ω.

In this construction we use the term *center* to denote the set of diagonal cells $C_{i,i}$ and the cells adjacent to the diagonal. Thus, the *off center* cells are those $C_{i,j}$ with i, j in $[0,\infty]$ such that either $i > j+1$ or $j > i+1$. The construction of μ is carried out in four steps.

In Step 1, $\mu(C_{i,j})$ will be defined for all cells $C_{i,j}$. In Step 2 we define $\mu(E_{i,j})$ for all cells off center. In Step 3, we fix $n < \infty$ and define a probability μ^{n} that agrees with μ off center

[8] Each of the probability measures μ^{ℓ} is defined on a countable space, Ω, and so the sequence has a converging subsequence.

and generates a dialogue that agrees with the given sequences in the first n stages. In Step 4 we define μ to be limit of the probabilities μ^n.

Step 1: **Defining μ on all cells.** For each $k < \infty$ let

$$e_k = \left(\max\{1/p_k^1, 1/\bar{p}_k^1, 1/p_k^2, 1/\bar{p}_k^2\}\right)^{-1},$$

and $\varepsilon_k = \min\{e_0, \ldots, e_k\}/3$. Define for each $i, j < \infty$, $\mu(C_{i,j}) = W\varepsilon_j^j \varepsilon_i^i$, $\mu(C_{\infty,j}) = W\varepsilon_j^{2j}$, $\mu(C_{i,\infty}) = W\varepsilon_i^{2i}$, and finally $\mu(C_{\infty,\infty}) = 0$. The constant W is chosen to normalize the sum of these numbers.

The only property of μ needed in the following steps is described in the next claim.

Claim 4. *Let $a_k^i = \mu(\mathrm{Row}(i,\vec{k}))$ and $b_k^j = \mu(\mathrm{Col}(j,\vec{k}))$. Then, for all i, j in $[1,\infty]$ and $k < \infty$ such that $i, j \geqslant k+1$,*

$$\frac{a_k^i}{a_{k+1}^i} \geqslant \max\left\{\frac{1}{p_k^1}, \frac{1}{\bar{p}_k^1}\right\} \quad \textit{and} \quad \frac{b_k^j}{b_{k+1}^j} \geqslant \max\left\{\frac{1}{p_k^2}, \frac{1}{\bar{p}_k^2}\right\}. \tag{24}$$

Indeed, consider first a pair (i,k) where $k+1 \leqslant i < \infty$. The ratio a_k^i/a_{k+1}^i is $(\sum_{j\geqslant k} \varepsilon_i^i \varepsilon_j^j + \varepsilon_i^{2i})/(\sum_{j\geqslant k+1} \varepsilon_i^i \varepsilon_j^j + \varepsilon_i^{2i})$. After cancelling ε_i^i this ratio becomes $(\sum_{j\geqslant k} \varepsilon_j^j + \varepsilon_i^i)/(\sum_{j\geqslant k+1} \varepsilon_j^j + \varepsilon_i^i)$. The numerator exceeds ε_k^k. We increase the denominator by replacing each ε_j^j in the infinite sum by ε_k^j. This results in a geometric series whose sum is smaller than $2\varepsilon_k^{k+1}$. Now, we further increase the denominator by replacing ε_i^i with ε_k^{k+1} ($\varepsilon_i^i > \varepsilon_k^{k+1}$ because $1 > \varepsilon_i > \varepsilon_k$ and $i < k$). Thus, the ratio is bigger than

$$\varepsilon_k^k/(3\varepsilon_k^{k+1}) = 1/e_k \geqslant \max\{1/p_k^1, 1/\bar{p}_k^1\}.$$

This proves the first part of (24) for $i < \infty$.

Next, consider the pair (∞, k). The ratio $a_\infty^k/a_\infty^{k+1}$ is equal to $\sum_{j\geqslant k} \varepsilon_j^{2j}/\sum_{j\geqslant k+1} \varepsilon_j^{2j}$. The numerator exceeds ε_k^{2k}. We increase the denominator by replacing each ε_j^{2j} by ε_k^{2j}, getting $\sum_{j=k+1}^{\infty} \varepsilon_k^{2j} = \varepsilon_k^{2k+2}(1-\varepsilon_k^2) \leqslant 2\varepsilon_k^{2k+2}$. Thus, the ratio is greater than $1/(2\varepsilon_k^2) \geqslant 1/(2\varepsilon_k) \geqslant 1/((2/3)e_k) \geqslant \max\{1/p_k^1, 1/\bar{p}_k^1\}$. This proves the first part of (24) for $i = \infty$. The proof for individual 2 is similar.

Step 2: **Defining $\mu(E_{i,j})$ off center.** For $k < \infty$ and $i \geqslant k+2$ we let $\mu(E_{i,k}) = p_k^1 a_k^i - p_{k+1}^1 a_{k+1}^i$. To justify this definition, we need to show that this difference falls between 0 and

$\mu(C_{i,k})$. Note that $\max\{1/p_k^1, 1/\bar{p}_k^1\} \geqslant \max\{p_{k+1}^1/p_k^1, \bar{p}_{k+1}^1/\bar{p}_k^1\}$. Thus, (24) and the equivalence between (6) and (7) imply that $\mu(E_{i,j})$ falls in the required range. Similarly, for $k < \infty$ and $j \geqslant k+2$ we let $\mu(E_{k,j}) = p_k^2 b_k^i - p_{k+1}^2 b_{k+1}^i$.

Observe that if for $i > k$, $\mu(E \mid \mathrm{Row}(i, \vec{k})) = p_k$, then by the definition of $\mu(E_{i,k-1})$ it follows that $\mu(E \mid \mathrm{Row}(i, \overrightarrow{k-1})) = p_{k-1}^1$. Thus, in order to show that (12) holds in row i it is enough to show that $\mu(E \mid \mathrm{Row}(i, \overrightarrow{i-1})) = p_{i-1}^1$, and similarly for the second agent. This is done in the next step.

Step 3: **Constructing probabilities in the center.** It remains to define $\mu(E_{i,j})$ for the cells in the center. This is done as follows. For a fixed $n > 1$ we define $\mu(E_{i,j})$ for center cells such that (12) and (13) hold for all $i, j \leqslant n$. We denote the resulting probability by μ^n. Obviously, all measures μ^n coincide off center. The construction of μ^n is carried out by induction on $k = n+1, \ldots, 1$.

For $k = n+1$ we define arbitrarily $\mu^n(E_{n+1,n+1})$, $\mu^n(E_{n+2,n+1})$, and $\mu^n(E_{n+1,n+2})$. Suppose the construction was carried out for $k+1$. We construct $\mu^n(E_{k,k})$, $\mu^n(E_{k+1,k})$, and $\mu^n(E_{k,k+1})$.

We start with $\mu(E_{k+1,k})$. Denote $p = \mu(E \mid \mathrm{Row}(k+1, \overrightarrow{k+1}))$ and define $\mu(E_{k+1,k}) = p_k^1 a_k^{k+1} - p a_{k+1}^{k+1}$. As $\max\{1/p_k^1, 1/\bar{p}_k^1\} \geqslant \max\{p/p_k^1, \bar{p}/\bar{p}_k^1\}$, it follows from (24) and the equivalence between (6) and (7) that $\mu(E_{i,j})$ falls between 0 and $\mu(E_{k+1,k})$, and thus the definition is valid. Moreover, $\mu(E \mid \mathrm{Row}(k+1, \vec{k}) = p_k^1$. Thus, (12) holds in row $k+1$. We similarly define $\mu(E_{k,k+1})$.

We need to define $\mu^n(E_{k,k})$ so that (13) is satisfied. Since we want to keep the inequality in the limit of μ^n, we need $\hat{p}_k^1$ and $\hat{p}_k^1$ to be bounded away from p_k^1 and p_k^2, respectively, uniformly for all n. Let $M_1 = \mu^n(E \cap \mathrm{Row}(k, \overrightarrow{k+1}))$ and $K_1 = \mu^n(\mathrm{Row}(k, \vec{k}))$. We similarly define M_2 and K_2 for agent 2.

If we set $\mu^n(E_{k,k}) = 0$, then $\mu^n(E \mid \mathrm{Row}(k, \vec{k})) = M_1/K_1$ and $\mu^n(E \mid \mathrm{Col}(\vec{k}, k)) = M_2/K_2$. If we set $\mu^n(E_{k,k}) = \mu^n(C_{k,k}) = \varepsilon_k^{2k}$, then $\mu^n(E \mid \mathrm{Row}(k, \overrightarrow{k+1})) = (\varepsilon_k^{2(k)} + M_1)/K_1$ and $\mu^n(E \mid \mathrm{Col}(\overrightarrow{k+1}, k)) = (\varepsilon_k^{2k} + M_2)/K_2$. Thus, we can choose the pair $(\hat{p}_k^1, \hat{p}_k^2)$ in such a way that $\hat{p}_k^1$ is in the interval $(M_1/K_1, (\varepsilon_k^{2k} + M_1)/K_1)$ and $\hat{p}_k^2 \in (M_2/K_2, (\varepsilon_k^{2k} + M_2)/K_2)$. Note that the lengths of these intervals are ε_k^{2k}/K_1 and ε_k^{2k}/K_2, which depend only on the definition of $\mu(C_{i,j})$ (that is, neither on the definition of $\mu(E_{i,j})$ nor on n)

Thus, we can find sufficiently small $\rho_k > 0$ and a pair $(\hat{p}_k^1, \hat{p}_k^2)$ (in the respective intervals), such that $|\hat{p}_k^1 - p_k^1| \geqslant \rho_k$ and $|\hat{p}_k^2 - p_k^2| \geqslant \rho_k$. The choice of the pair $\hat{p}_k^1, \hat{p}_k^2$ may depend on n, but ρ_k will be the same for all n.

Step 4: **Taking the limit.** Let I be the set of all the center cells indices. The set $[0,1]^I$ with the product topology is compact. For each n, $(\mu^n(E_{i,j}))_{(i,j)\in I}$ is an element of this set. Thus, there exists a limit point $x_{i,j}$ of this sequence. Obviously, for each $(i,j)\in I$, $0\leqslant x_{i,j}\leqslant \mu(C_{i,j})$. We can therefore extend μ to $E_{i,j}$ in the center by defining $\mu(E_{i,j})=x_{i,j}$.

We need to show that μ satisfies (12) and (13). Each equation for i and k in (12) is a linear equation in the three numbers $\mu^n(E_{i,i-1})$, $\mu^n(E_{i,i})$, and $\mu^n(E_{i,i+1})$, where the coefficients are the same for all n. Thus, the equation holds also in the limit, that is, for μ.

For (13), $\mu^n(\text{Row}(k,\vec{k}))$ is a linear expression in $\mu^n(E_{i,i})$, and $\mu^n(E_{i,i+1})$ where the coefficients are independent of n. Thus, $\mu^n(\text{Row}(k,\vec{k}))\to_n \mu(\text{Row}(k,\vec{k}))$. Since, $|\mu^n(\text{Row}(k,\vec{k}))-p^1_k|\geqslant\rho_k>0$ for all n, it follows that $\mu(\text{Row}(k,\vec{k}))\neq p^1_k$. The argument for agent 2 is the same, which completes the proof of part (i).

We now proceed to proving part (ii) of Proposition 2. For this purpose we use twice part (ii) of the theorem. In the construction we use the same space twice, but each time assign different probabilities to the same cells.

In the first step we construct a dialogue in which both agents have the sequence p^1_k. That is, we construct a dialogue generating p^1_k and $\widetilde{p}^2_k$, where $\widetilde{p}^2_k=p^1_k$. Here, the two sequences coincide. Since p^1_k is a positive monologue, these two sequences satisfy the conditions of Theorem 2(ii). The construction of Theorem 2(ii) yields a space, a sequence of partitions and a measure μ_1, with respect to which p^1_k and $\widetilde{p}^2_k$ is a positive dialogue. The measure μ_1 is defined to be the prior of agent 1.

In the second step we construct a dialogue generating the sequences $\widetilde{p}^1_k=p^2_k$ and p^2_k. The space and the sequence of partitions are the same as in the first step. The measure μ_2 might differ from μ_1. The measure μ_2 is defined to be the prior of individual 2. □

References

R. J. Aumann. Agreeing to disagree. *Annals of Statistics*, 4:1236–1239, 1976.

M. Bacharach. Normal bayesian dialogues. *Journal of the American Statistical Association*, 74:837–846, 1979.

D. L. Burkholder. Martingale transforms. *Annals of Mathematical Statistics*, 37:1494–1504, 1966.

Y. Feinberg. Characterizing common priors in the form of posteriors. *Journal of Economic Theory*, 91:127–179, 2000.

J. D. Geanakoplos and H. Polemarchakis. We cannot disagree forever. *Journal of Economic Theory*, 28:192–200, 1982.

F. Gul. A comment on aumann's bayesian view. *Econometrica*, 66:923–927, 1998.

S. Hart and Y. Taumann. Market crashes without external shocks. *Journal of Business*, 77:1–8, 2004.

A. Heifetz. The positive foundation of the common prior assumption. *Games and Economic Behavior*, 56:105–120, 2006.

E. Lehrer and D. Samet. Belief consistency and trade consistency. *Games and Economic Behavior*, 83:165–177, 2014.

B. L. Lipman. Finite order implications of common priors. *Econometrica*, 71:1255–1267, 2003.

S. Morris. Trade with heterogenous prior beliefs and asymmetric information. *Econometrica*, 62:1327–1347, 1994.

S. Morris. The common prior assumption in economic theory. *Economics and Philosophy*, 11: 227–253, 1995.

L. T. Nielsen. Common knowledge, communication, and convergence of beliefs. *Mathematical Social Sciences*, 8:1–14, 1984.

H. Polemarchakis. Rational dialogs. Warwick mimeo, 2016.

D. Samet. Common priors and the separation of convex sets. *Games and Economic Behavior*, 24:172–174, 1998a.

D. Samet. Iterated expectations and common priors. *Games and Economic Behavior*, 24: 131–141, 1998b.

D. Shaiderman. Variation of the conditional means of a martingale. Unpublished manyscript, 2018.

T. Tsuchikura and M. Yamasaki. Martingale sequence of bounded variation 123-127. *Tôhoku Mathematics Journal*, pages 123–127, 1976.