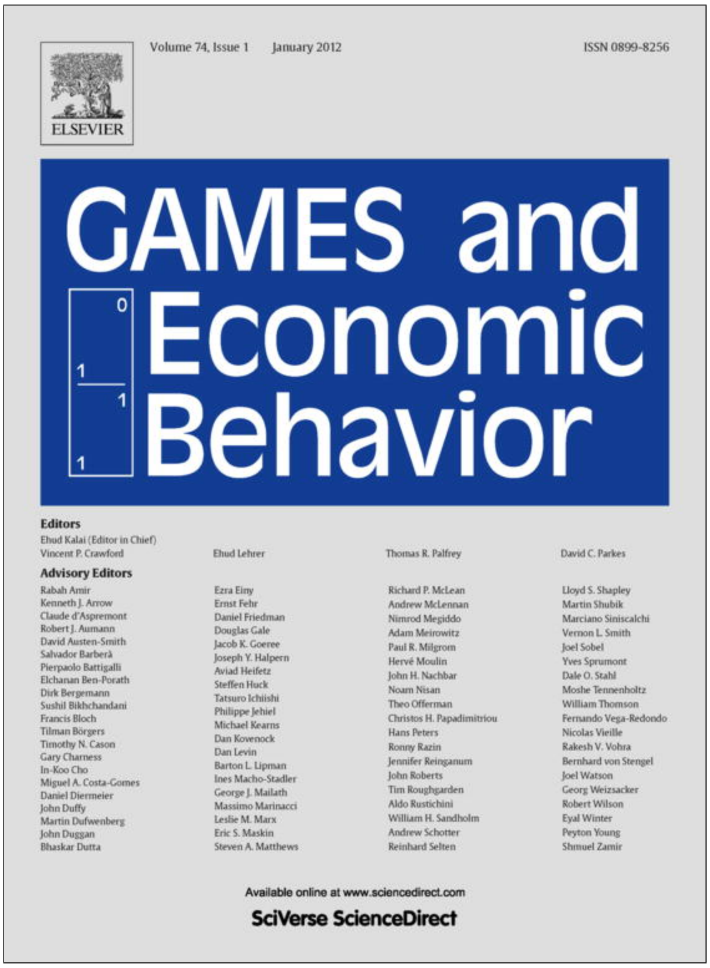
Volume 74, Issue 1 January 2012
ISSN 0899-8256
ELSEVIER
GAMES and Economic Behavior
Editors
Ehud Kalai (Editor in Chief)
Vincent P. Crawford
Ehud Lehrer
Thomas R. Palfrey
David C. Parkes
Advisory Editors
Rabah Amir
Kenneth J. Arrow
Claude d'Aspremont
Robert J. Aumann
David Austen-Smith
Salvador Barberà
Pierpaolo Battigalli
Elchanan Ben-Porath
Dirk Bergemann
Sushil Bikhchandani
Francis Bloch
Tilman Börgers
Timothy N. Cason
Gary Charness
In-Koo Cho
Miguel A. Costa-Gomes
Daniel Diermeier
John Duffy
Martin Dufwenberg
John Duggan
Bhaskar Dutta
Ezra Einy
Ernst Fehr
Daniel Friedman
Douglas Gale
Jacob K. Goeree
Joseph Y. Halpern
Aviad Heifetz
Steffen Huck
Tatsuro Ichiishi
Philippe Jehiel
Michael Kearns
Dan Kovenock
Dan Levin
Barton L. Lipman
Ines Macho-Stadler
George J. Mailath
Massimo Marinacci
Leslie M. Marx
Eric S. Maskin
Steven A. Matthews
Richard P. McLean
Andrew McLennan
Nimrod Megiddo
Adam Meirowitz
Paul R. Milgrom
Hervé Moulin
John H. Nachbar
Noam Nisan
Theo Offerman
Christos H. Papadimitriou
Hans Peters
Ronny Razin
Jennifer Reinganum
John Roberts
Tim Roughgarden
Aldo Rustichini
William H. Sandholm
Andrew Schotter
Reinhard Selten
Lloyd S. Shapley
Martin Shubik
Marciano Siniscalchi
Vernon L. Smith
Joel Sobel
Yves Sprumont
Dale O. Stahl
Moshe Tennenholtz
William Thomson
Fernando Vega-Redondo
Nicolas Vieille
Rakesh V. Vohra
Bernhard von Stengel
Joel Watson
Georg Weizsacker
Robert Wilson
Eyal Winter
Peyton Young
Shmuel Zamir
Available online at www.sciencedirect.com
SciVerse ScienceDirect

Games and Economic Behavior 74 (2012) 321–331

Contents lists available at ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

On the strategic impact of an event under non-common priors ☆

Daisuke Oyama [a,*], Olivier Tercieux [b]

[a] *Faculty of Economics, University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*
[b] *Paris School of Economics and CNRS, 48 Boulevard Jourdan, 75014 Paris, France*

ARTICLE INFO

Article history:
Received 3 September 2010
Available online 21 June 2011

JEL classification:
C72
D82

Keywords:
Common prior assumption
Higher order belief
Rationalizability
Contagion
Belief potential

ABSTRACT

This paper studies the impact of a small probability event on strategic behavior in incomplete information games with non-common priors. It is shown that the global impact of a small probability event (i.e., its propensity to affect strategic behavior at all states in the state space) has an upper bound that is an increasing function of a measure of discrepancy from the common prior assumption. In particular, its global impact can be arbitrarily large under non-common priors, but is bounded from above under common priors. These results quantify the different implications common prior and non-common prior models have on the (infinite) hierarchies of beliefs.

1. Introduction

While controversial, the common prior assumption (hereafter, CPA) is used in most models of incomplete information in game theory and economics. This assumption says that the beliefs of all players are generated from a single prior, updated by Bayes' rule, so that differences in their beliefs are due solely to differences in information that they receive. It is well known that the CPA is crucial for many results in incomplete information games (e.g., Aumann's, 1976 result on agreeing to disagree and no trade theorems by Milgrom and Stokey, 1982). The purpose of this paper is to clarify the restrictions that we implicitly impose on strategic behavior in game theoretic models when we accept the CPA. Specifically, we focus on the strategic impact that a small amount of payoff uncertainty has through "contagion" effects over players' beliefs about payoffs, their beliefs about others' beliefs, and so on, i.e., hierarchies of beliefs. We introduce a measure of discrepancy from the CPA and give a tight upper bound of the strategic impact as an increasing function of this discrepancy measure. In particular, it is shown that the global impact of a small probability event (i.e., its propensity to affect strategic behavior at all states in the state space) can be arbitrarily large under non-common priors, while it is bounded from above under common

☆ Earlier versions of this paper were circulated under the title "Strategic Implications of (Non-)Common Priors I: Belief Potential". We would like to thank Akihiko Matsui, Stephen Morris, Dov Samet, Takashi Ui, Muhamet Yildiz, anonymous referees, and especially, Atsushi Kajii and Michihiro Kandori for helpful comments and discussions. This paper has been presented at the 2006 Kyoto Game Theory Workshop, Hitotsubashi University, University of Tokyo, Yokohama National University, GAMES 2008, Ben-Gurion University, University of Edinburgh, Max Planck Institute for Research on Collective Goods, National University of Singapore, and the 10th SAET Conference. D. Oyama acknowledges Grant-in-Aid for JSPS Fellows and Scientific Research (S) and financial support from the Zengin Foundation for Studies on Economics and Finance. O. Tercieux acknowledges financial support from the French Ministry of Research (Action Concertée Incitative). This research was initiated while O. Tercieux was visiting the Institute of Economic Research, Kyoto University, whose hospitality is gratefully acknowledged.

* Corresponding author.
E-mail addresses: oyama@e.u-tokyo.ac.jp (D. Oyama), tercieux@pse.ens.fr (O. Tercieux).
URLs: http://www.oyama.e.u-tokyo.ac.jp (D. Oyama), http://www.pse.ens.fr/tercieux (O. Tercieux).

0899-8256/$ – see front matter
doi:10.1016/j.geb.2011.06.003

priors. From these results, we obtain new insights which contribute to our deeper understanding of the recent literature on higher order beliefs and (non-)common priors, in particular, studies of Lipman (2003) and Weinstein and Yildiz (2007).

It has been known that once we depart from common knowledge of payoffs by introducing a small amount of incomplete information, strategic behavior may change dramatically through higher order beliefs. Rubinstein (1989), Carlsson and van Damme (1993), and Morris et al. (1995), among others, show how a small probability event can have a large impact on strategic behavior (under common prior). To see the logic behind, suppose that player 1 is known to take a certain action at some event E which has a very small ex ante probability. If player 2 puts high conditional probability on the event E at his information sets at which E is thought possible, this knowledge might imply a unique best response by player 2 at these information sets. This, in turn, implies how player 1 responds to that knowledge at larger information sets, and so on. If this iterative argument results in a unique action profile played everywhere on the state space, then we have a *contagion* of this action profile from event E. Then the question is: when is it the case that a certain action profile being chosen at some event (which, again, may have a very small probability) implies that this action profile is chosen everywhere on the state space, in other words, when is an action profile contagious?

In Section 2, we first discuss an example of a 2×2 coordination game embedded in an information system of the type of Rubinstein's (1989) email game, but with heterogeneous priors. In the example, we demonstrate that, from a small probability event, any strict Nash equilibrium however "weak" it is can be contagious under sufficiently heterogeneous priors, while only a risk-dominant equilibrium can be contagious under common prior. We may thus say for this case that this small probability event has an arbitrarily large impact under heterogeneous priors, while its impact is bounded from above under common prior.

The strategic impact of an event in the above sense is formally measured by the notion of *belief potential* proposed by Morris et al. (1995). First, say that an event E has an *impact* p on a state ω (which we also refer to as the "local impact" of event E) if the statement "player 1 believes with probability at least p that 2 believes with probability at least p that 1 believes ... event E" is true at state ω. Then, the belief potential (or "global impact") of event E is the largest probability p such that E has impact p on *all* states in the state space. As Morris et al. (1995) demonstrate, the belief potential of an event E characterizes, through the notion of p-dominance, the necessary "strength" of an action profile to be contagious from E.

In Section 3, we find the measure of discrepancy from the CPA that provides an upper bound on the global impact of small probability events as an increasing function of this measure. The measure of discrepancy is the supremum of the ratios between the players' prior probabilities over the states in the state space. This result implies, first, that for a small probability event to have a maximum global impact, this measure of discrepancy from the CPA has to be large. Second, it implies that the global impact of a small probability event is bounded from above under the CPA. The latter result also quantifies the implications of the CPA on infinite hierarchies of beliefs.

Given these formal results, we explore the relationship between common and non-common prior models in terms of the global and the local impact of an event, which helps to clarify recent important results in the literature. We first point out that, in contrast with the global impact (i.e., on all states in the state space), the local impact (i.e., on a given state) of a small probability event can be arbitrarily large even under the CPA. Indeed, we show that for any integer $N > 0$, we can find an information system with a common prior, and a small probability event E such that at some state ω players mutually know up to order N that E did not occur but where still this small probability event has an arbitrarily large impact on state ω. This is to translate into our setting a result by Weinstein and Yildiz (2007) (reported also in Yildiz, 2004) that for any strict Nash equilibrium a^* of a complete information type,[1] there exist nearby types (with respect to product topology in the universal type space; see Mertens and Zamir, 1985 and Brandenburger and Dekel, 1993) such that a^* is the unique rationalizable outcome for these types, and moreover, these types can be chosen from models *with common prior*.[2] The latter result crucially relies on that of Lipman (2003) who considers the implications of the CPA for finite hierarchies of beliefs. Lipman (2003) shows that for any state in a partition model where players may have heterogeneous priors (but with common support), there is a corresponding state in another partition model with a common prior that is close to the original state with respect to product topology. That is, the CPA does not have any implication on finite order beliefs, if one is interested only in local properties of the beliefs (i.e., properties at a given state). Note, however, that Lipman's (2003) result says nothing about the restrictions imposed on global properties of the whole state space. In fact, we show that under the CPA, the set of states on which a vanishingly small probability event has a large local impact has an arbitrarily small probability with respect to the common prior. This is the sense in which we say that the global impact of a small probability event cannot be arbitrarily large under the CPA. In relation to Weinstein and Yildiz' (2007) result, this implies that for some complete information type and for some strict Nash equilibria a^* (e.g., the risk-dominated equilibrium in a 2×2 coordination game), the ex ante probability of the set of types (in models from which nearby types are extracted) for which a^* is uniquely rationalizable vanishes, if we require the set of types such that the complete information game is not played to be vanishingly small with respect to the common prior.

[1] A complete information type is a degenerate type in the universal type space where it is common knowledge that a given complete information game is played.

[2] The result of Weinstein and Yildiz (2007) is much more general: first they consider types that need not be complete information types, and second they deals not only with strict Nash equilibria but also with rationalizable outcomes that need not be strict Nash equilibria. For our purpose, it is not necessary to consider the most general version of their result.

The example in Section 2 shows that in 2×2 coordination games, where there are two strict Nash equilibria, the risk-dominated equilibrium can be contagious from a vanishingly small probability event if the players are allowed to have heterogeneous priors. On the other hand, Kajii and Morris (1997) have shown that under the CPA, this is not possible: in their terminology, the risk-dominant equilibrium is robust to incomplete information. Our example shows that even the risk-dominant equilibrium is not robust if we allow for non-common prior perturbations. In a companion paper (Oyama and Tercieux, 2010), we show that in generic games, a Nash equilibrium is robust to incomplete information under non-common priors if and only if it is a unique action profile that survives iterative elimination of strictly dominated actions. That paper also shows that the risk-dominant equilibrium (a p-dominant equilibrium with low p, more generally) continues to be robust to non-common prior perturbations that are close to common prior perturbations with respect to the discrepancy measure developed in the present paper.

The remainder of the paper is organized as follows. Section 2 provides an example that summarizes our formal analysis. It illustrates why without the CPA, any strict Nash equilibrium can be contagious and how it is related to the discrepancy from the CPA. Section 3 introduces the concept of belief potential and states our results linking the measure of discrepancy from the CPA with the belief potential of small probability events. It also compares the local and the global impacts of small probability events, relating our results to those by Lipman (2003) and Weinstein and Yildiz (2007). Section 4 discusses alternative distance measures from the CPA. as well as an extension to the many player case.

2. Example

In this section, we illustrate the analyses in the next section with a simple example. Consider the following 2×2 coordination game with complete information which we denote by $\mathbf{g}$. There are two players, 1 and 2, each of whom has two actions L and R. Throughout the paper, for $i = 1, 2$ we write $-i$ for player $j \neq i$. The payoffs are given by

	L	R
L	p, p	$0, 0$
R	$0, 0$	$1-p, 1-p$

where $p \in (1/2, 1)$, so that (L, L) is (both Pareto-dominant and) risk-dominant. We say that (L, L) is a strict $(1-p)$-dominant equilibrium while (R, R) is a strict p-dominant equilibrium; formally, an action profile (a_1^*, a_2^*) is a strict p-dominant equilibrium if a_i^* is a unique best response for player i to any conjecture that places a probability larger than p on player $-i$ playing a_{-i}^* (see Morris et al., 1995 and Kajii and Morris, 1997). As p becomes close to one, the strict Nash equilibrium (R, R) becomes "weaker".

Now, we ask the following question: For each strict Nash equilibrium $a^* = (L, L), (R, R)$ of $\mathbf{g}$, are there "perturbations" arbitrarily "close" to $\mathbf{g}$ in which a^* is played as a *unique* rationalizable strategy outcome? The question, of course, is not well defined unless what we mean by "perturbations" being "close" to $\mathbf{g}$ is specified.

2.1. Perturbations via incomplete information games

Here, as perturbations of $\mathbf{g}$ we consider incomplete information games with an information partition structure as well as the same sets of players and actions as in $\mathbf{g}$, where we allow the players to have different prior beliefs. The complete information game $\mathbf{g}$ is considered as a degenerate incomplete information game. We regard a perturbed incomplete information game to be close to $\mathbf{g}$ if the event that both players know that their payoffs are given by $\mathbf{g}$ has probability close to one with respect to both players' prior distributions.

To address the question, we consider the following class of perturbations of $\mathbf{g}$. The state space Ω is given by $\{1, 2\} \times \mathbb{Z}_+$. Player $i = 1, 2$ has information partition $\mathcal{Q}_i$ which consists of (i) the event $\{(-i, 0)\}$ and (ii) all the events of the form $\{(i, k-1), (-i, k)\}$ for $k \geqslant 1$. Observe that this partition structure is of the same type as that in the electronic mail game of Rubinstein (1989).

The players may have different prior beliefs. For $r \in [1, \infty)$ and $\varepsilon \in (0, 1)$, let player i's prior P_i be defined by

$$P_i(i, k) = \frac{r}{r+1} \cdot \varepsilon(1-\varepsilon)^k,$$
$$P_i(-i, k) = \frac{1}{r+1} \cdot \varepsilon(1-\varepsilon)^k.$$

The players have a common prior if and only if $r = 1$. Observe that for all $\omega \in \Omega$, $P_i(\omega)/P_{-i}(\omega) = r$ if $\omega = (i, k)$, while $P_{-i}(\omega)/P_i(\omega) = r$ if $\omega = (-i, k)$. We will use the parameter r to measure the degree of discrepancy from the CPA.

Finally, let $E_i = \{(-i, 0)\}$ and $E = E_1 \cup E_2$. The payoffs of each player i are given by $\mathbf{g}$ at all states in $\Omega \setminus E_i$, while a_i^* is a strictly dominant action for player i on event E_i, where $a^* = (a_1^*, a_2^*)$ will be (L, L) or (R, R). Verify that $P_i(E) = \varepsilon$ for each i. Let us denote this incomplete information game by $\mathcal{U}(r, \varepsilon; a^*)$.

(1) Common prior case ($r = 1$): As demonstrated by Morris et al. (1995),[3] if L is a dominant action for each player i at state $(-i, 0)$, then however small $\varepsilon > 0$ is, the incomplete information game $\mathcal{U}(1, \varepsilon; (L, L))$ has a unique rationalizable

[3] Kajii and Morris (1997) extend their argument to the countable state space case.

strategy profile, where (L, L) is played at all $\omega \in \Omega$: that is, we have a "contagion" of the risk-dominant action L. On the other hand, as established by Kajii and Morris (1997), even if R is a dominant action for each player i at state $(-i, 0)$, the incomplete information game $\mathcal{U}(1, \varepsilon; (R, R))$ has a Bayesian Nash equilibrium in which (L, L) is played with high (ex ante) probability whenever ε is sufficiently small. We may say that under a common prior, the event E, however small its (ex ante) probability is, has an impact large enough to make the risk-dominant action contagious, but not large enough to make the risk-dominated one contagious.

(2) Non-common prior case ($r > 1$): We show that for r sufficiently large, each action is contagious: for each equilibrium a^*, if a_i^* is a dominant action at state $(-i, 0)$ for each player i, then there exists $\bar{r}$ such that for all $r > \bar{r}$ and all $\varepsilon \in (0, 1)$, the incomplete information game $\mathcal{U}(r, \varepsilon; a^*)$ has a unique rationalizable strategy profile, where a^* is played at all $\omega \in \Omega$. To see this, suppose that for each player i, R is a dominant action at $(-i, 0)$. Observe that

$$P_i\bigl(\bigl\{(i, k-1)\bigr\} \bigm| \bigl\{(i, k-1), (-i, k)\bigr\}\bigr) = \frac{r}{r+1-\varepsilon} \tag{2.1}$$

for all $k \geqslant 1$. Now, given $p \in (1/2, 1)$, let $\bar{r} = p/(1-p)$ (> 1), and take any $r \geqslant \bar{r}$ and $\varepsilon \in (0, 1)$. Then, if player $-i$ plays R at $(i, k-1)$ in any rationalizable strategy, then it implies that player i plays R at $(-i, k)$ in any rationalizable strategy, since i assigns a probability $r/(r+1-\varepsilon) > p$ to the event $-i$ plays R, which makes R the unique best response. We may hence say that under non-common priors, the event E, however small its (ex ante) probability is, may have an impact large enough that any strict Nash equilibrium is contagious. The key to this result is that by increasing the value of r, we can have the relevant conditional probabilities, $P_i(\{(i, k-1)\} \mid \{(i, k-1), (-i, k)\})$, be as close to one as possible. The supremum of such conditional probabilities relevant to the contagion argument will be called the *belief potential* of the event E (see Definition 3.1 for the precise definition). In this particular information system with given r and ε, the belief potential of E is $r/(r+1-\varepsilon)$, as given by (2.1). But it will turn out that this is the "best case", in which a small probability event has the largest impact. We will show that given values of discrepancy measure, r, and small probability, ε, the value $r/(r+1-\varepsilon)$ is the maximum of the belief potential of a small probability event *over information systems* (see Theorem 3.2 for the precise statement). This implies that an event can have a larger impact on higher order beliefs under non-common prior than under common prior.

2.2. Perturbations via types

The second notion uses, as perturbations, states instead of incomplete information games. A state is considered to be close to the complete information game $\mathbf{g}$ if at this state (together with the associated partition model) players know up to high level that payoffs are given by $\mathbf{g}$. The corresponding notion in the universal type space (Mertens and Zamir, 1985; Brandenburger and Dekel, 1993) is known as the product topology and has been studied by Weinstein and Yildiz (2007), among others. They identify the complete information game $\mathbf{g}$ with a point $t_{\mathbf{g}}$ in the universal type space, i.e., the hierarchy of degenerate beliefs, and considers as "perturbations" being "close" to $\mathbf{g}$, types in the universal type space that are close to $t_{\mathbf{g}}$ with respect to product topology. Their results imply, in particular, that for any strict Nash equilibrium a^* of complete information game $\mathbf{g}$, there exists a sequence of types converging to $t_{\mathbf{g}}$ each of which plays a^* as a unique rationalizable strategy outcome. Moreover, by appealing to Lipman's (2003) result, they show that those converging types can be taken from models (i.e., belief-closed subspaces) with common prior. We will discuss in Section 4 the relationship between Lipman's (2003) and Weinstein and Yildiz' (2007) results and ours.

Weinstein and Yildiz' (2007) result can be stated in our framework as follows: for any strict Nash equilibrium a^* of complete information game $\mathbf{g}$, there exists a sequence of perturbed incomplete information games $\mathcal{U}^k$ with common prior and states ω^k such that any rationalizable strategy profile of $\mathcal{U}^k$ plays a^* at ω^k, where in each $\mathcal{U}^k$, a_i^* is a strictly dominant action for player i on an event E_i^k, and at each ω^k, players mutually know up to kth order that the payoffs are given by $\mathbf{g}$. To see this in our example, let $a^* = (R, R)$. By modifying the incomplete information game in the previous subsection with given $p \in (1/2, 1)$, $\mathcal{U}^k$ can be constructed as follows (common for all k). Let the state space be $\bar{\Omega} = \Omega \cup \{\infty\}$, and the information partition for each player i be $\bar{\mathcal{Q}}_i = \mathcal{Q}_i \cup \{\{\infty\}\}$, where Ω and $\mathcal{Q}_i$ are the state space and the information partitions defined in the previous subsection. Define the common prior $\bar{P}$ by

$$\bar{P}(1, k) = \bar{P}(2, k) = \frac{1}{2}\varepsilon\left(\frac{1-\varepsilon}{r}\right)^k$$

for $k \geqslant 0$ and

$$\bar{P}(\infty) = 1 - \frac{r}{r-(1-\varepsilon)}\varepsilon,$$

where r is such that $r \geqslant p/(1-p)$. Note that we need to add a state, denoted ∞, in order for $\bar{P}$ to sum up to one. The payoffs of each player i are given by $\mathbf{g}$ at all states in $\bar{\Omega} \setminus E_i$, while a_i^* is a strictly dominant action for player i on event E_i. Then, since the relevant posteriors are given by

$$\bar{P}\bigl(\bigl\{(i, k-1)\bigr\} \bigm| \bigl\{(i, k-1), (-i, k)\bigr\}\bigr) = \frac{r}{r+1-\varepsilon}$$

for all $k \geqslant 1$, the same argument in the previous subsection shows that any rationalizable strategy plays R in every state in $\bar{\Omega} \setminus \{\infty\}$. In addition, it is easy to check that for each k, at $(i, k+1)$ $(i = 1, 2)$ it is mutually known up to order k that payoffs are given by $\mathbf{g}$.

Now, if we require that $\bar{P}(E_1 \cup E_2)$ $(= \varepsilon)$ vanish along the sequence, then $\bar{P}(\bar{\Omega} \setminus \{\infty\})$ must vanish accordingly, and so the ex ante probability of the event that R is played as a unique rationalizable strategy action converges to 0. In fact, as we will argue in Section 3.3 (see Proposition 3.4), this is the case not only in this particular construction of incomplete information games, but also in any such construction. This is to be contrasted with the non-common prior case in the previous subsection, where any strict Nash equilibrium can be contagious over the state space. In this sense, if one is interested in strategic behavior on the whole state space, rather than local behavior (i.e., behavior at a particular state as in Weinstein and Yildiz, 2007), then models with common priors may be significantly different from those with non-common priors.

Remark 2.1. In the construction above, we could have assumed that the payoffs of each player i are given by $\mathbf{g}$ at all states in $\Omega \setminus E_i$, while a_i^* is a strictly dominant action for player i on event E_i and at state ∞ some action b_i^* (possibly equal to a_i^*) is a strictly dominant action for each player i. Clearly, we would have obtained a dominance solvable game with a common prior (note, however, that the prior probability of the event "payoffs are given by $\mathbf{g}$" converges to 0 as ε vanishes). This implies that for any strict Nash equilibrium a^* of complete information game $\mathbf{g}$, there exists a sequence of types in the universal type space converging to $t_{\mathbf{g}}$ (in the product topology) each of which plays a^* as a unique rationalizable strategy outcome. Moreover, those converging types can be taken from models (i.e., belief-closed subspaces) that satisfy *both* dominance solvability and the CPA.

3. Belief potential

3.1. Information systems and belief potential

An *information system* is the structure $IS = (\Omega, (P_i)_{i=1,2}, (\mathcal{Q}_i)_{i=1,2})$, where Ω is a countable set of states, P_i is the prior distribution on Ω for player $i = 1, 2$, and $\mathcal{Q}_i$ is the partition of Ω representing the information of player i. We write $Q_i(\omega)$ for the element of $\mathcal{Q}_i$ containing ω. Given an information system, we write $\mathcal{F}_i$ for the sigma algebra generated by $\mathcal{Q}_i$. We assume that $P_i(Q_i(\omega)) > 0$ for all $i = 1, 2$ and $\omega \in \Omega$. Under this assumption, the conditional probability of ω' given $Q_i(\omega)$, $P_i(\omega' | Q_i(\omega))$, is well defined by $P_i(\omega' | Q_i(\omega)) = P_i(\omega')/P_i(Q_i(\omega))$ whenever $\omega' \in Q_i(\omega)$. For an event $E \subset \Omega$, we let E^c denote its complement, i.e., $E^c = \{\omega \in \Omega \mid \omega \notin E\}$. Given an information system *IS*, define the following measure of discrepancy from the common prior case, ρ, by

$$\rho(P_1, P_2) = \max_{i,j} \sup_{\omega \in \Omega} \frac{P_i(\omega)}{P_j(\omega)} \tag{3.1}$$

with a convention that $q/0 = \infty$ for $q > 0$, and $0/0 = 1$. Note that $\rho(P_1, P_2) < \infty$ only if (P_1, P_2) has common support. The information system satisfies the CPA if and only if $\rho(P_1, P_2) = 1$.

We use the notion of p-belief as defined by Monderer and Samet (1989). For $p \in (0, 1]$, the p-belief operator for player $i = 1, 2$, $B_i^p : 2^\Omega \to 2^\Omega$, is defined by

$$B_i^p(E) = \{\omega \in \Omega \mid P_i(E | Q_i(\omega)) \geqslant p\}.$$

That is, $B_i^p(E)$ is the set of states where player i believes E with probability at least p (with respect to his own prior P_i). We will also use the knowledge operator for player i, $K_i : 2^\Omega \to 2^\Omega$, defined by

$$K_i(E) = \{\omega \in \Omega \mid Q_i(\omega) \subset E\}.$$

That is, $K_i(E)$ is the set of states where player i knows that event E is true. Let $K_*(E) = K_1(E) \cap K_2(E)$ be the set of states where it is mutual knowledge that event E is true, i.e., where both players know that event E is true. At a state ω, an event E is said to be mutual knowledge at order N if $\omega \in \bigcap_{n=1}^{N} [K_*]^n(E)$. Finally, at state ω, an event E is said to be common knowledge if $\omega \in \bigcap_{n=1}^{\infty} [K_*]^n(E)$.

We define the *contagion operator* $H_i^p : 2^\Omega \to 2^\Omega$ by

$$H_i^p(E) = B_i^p(B_{-i}^p(E)) \cup E.$$

We denote $(H_i^p)^0(E) = E$ and for $k \geqslant 1$, $(H_i^p)^k(E) = H_i^p((H_i^p)^{k-1}(E))$.

Denote $(H_i^p)^\infty(E) = \bigcup_{k=1}^{\infty} (H_i^p)^k(E)$. We follow Morris et al. (1995) to measure the *global* impact of an event by the notion of *belief potential*. The belief potential of an event E is the largest probability p such that a statement of the form "player i believes with probability at least p that player $-i$ believes with probability at least p that i believes … that the true state is in E" is true at every state in Ω.

Definition 3.1. The *belief potential of event* E, $\sigma(E)$, is

$$\sigma(E) = \max_{i=1,2} \sigma_i(E),$$

where

$$\sigma_i(E) = \sup\{p \in [0,1] \mid (H_i^p)^\infty(E) = \Omega\}.$$

To illustrate these concepts, consider the information system and the event $E = \{(1,0),(2,0)\}$ in Section 2.1. Note that this information system satisfies $\rho(P_1,P_2) = r$. Observe first that for each $i = 1,2$, $B_i^p(E) = \{(-i,0)\} \cup \{(i,0),(-i,1)\}$ and $H_i^p(E) = \{(-i,0)\} \cup \{(i,0),(-i,1)\} \cup \{(i,1),(-i,2)\}$ if $p \leqslant r/(r+1-\varepsilon)$, and $B_i^p(E) = \{(-i,0)\}$ and $H_i^p(E) = E$ otherwise. Thus,

$$(H_i^p)^K(E) = \{(-i,0)\} \cup \bigcup_{k=1}^{K+1} \{(i,k-1),(-i,k)\}$$

and therefore $(H_i^p)^\infty(E) = \Omega$ if $p \leqslant r/(r+1-\varepsilon)$, and $(H_i^p)^\infty(E) = E$ otherwise. This implies that for this information system,

$$\sigma(E) = \frac{r}{r+1-\varepsilon}.$$

In Section 3.2, we will show that, given $r \geqslant 1$ and $\varepsilon > 0$, this is the maximum value of the belief potential of an event with probability ε over the information systems such that $\rho(P_1,P_2) = r$.

As demonstrated in Section 2, the impact of a small probability event as measured by the belief potential characterizes the contagion of Nash equilibria played at that event (which motivates our study of the belief potential) through the notion of p-dominance. We describe the connection briefly here; see Morris et al. (1995, Theorem 5.1) or the working paper version of the present paper Oyama and Tercieux (2005) for the formal presentation. An action profile (a_1^*,a_2^*) of a complete information game $\mathbf{g} = (g_1,g_2)$ is a *strict* p*-dominant equilibrium* if a_i^* is a unique best response for player i to any conjecture that places a probability larger than p on player $-i$ playing a_{-i}^*. Given an information system IS and an event $E = E_1 \cup E_2 \subset \Omega$ with $E_i \in \mathcal{F}_i$, $i = 1,2$, suppose that for each player i, a_i^* is a strictly dominant action at every $\omega \in E_i$ and the payoffs at every $\omega \notin E_i$ are given by g_i. Then, if the event E has belief potential $\sigma > 0$ and the action profile (a_1^*,a_2^*) is a strict p-dominant equilibrium for some $p < \sigma$, then for each player i, a_i^* is the unique rationalizable action at all $\omega \in \Omega$.

3.2. Upper bound of belief potential

Our main goal is to characterize the upper bound of the belief potential of small probability events over information systems with a given value of the discrepancy from the CPA (i.e., $\rho(P_1,P_2)$). Given an information system $IS = (\Omega,(P_i)_{i=1,2},(\mathcal{Q}_i)_{i=1,2})$, we denote

$$\mathcal{F}_1 \oplus \mathcal{F}_2 = \{E \subset \Omega \mid E = E_1 \cup E_2 \text{ for some } E_i \in \mathcal{F}_i \text{ for each } i = 1,2\}$$

(recall that $\mathcal{F}_i$ is the sigma algebra generated by $\mathcal{Q}_i$). For $p \in (0,1]$ and $E \in \mathcal{F}_1 \oplus \mathcal{F}_2$, we define

$$H_*^p(E) = B_1^p(E) \cup B_2^p(E).$$

We denote $(H_*^p)^k(E) = H_*^p((H_*^p)^{k-1}(E))$ for $k \geqslant 1$, where $(H_*^p)^0(E) = E$, and $(H_*^p)^\infty(E) = \bigcup_{k=1}^\infty (H_*^p)^k(E)$. Verify that $(H_1^p)^\infty(E) \cup (H_2^p)^\infty(E) = (H_*^p)^\infty(E)$, so that if $(H_i^p)^\infty(E) = \Omega$, then $(H_*^p)^\infty(E) = (H_i^p)^\infty(E)$. It is thus sufficient to characterize the (ex ante) probability of $(H_*^p)^\infty(E)$.

Again, in this subsection, we aim at providing an upper bound on the belief potential of an event over information systems. To do so, by definition of belief potential, what matters is the "propensity" of $(H_*^p)^\infty(E)$ to cover the whole state space Ω for a given $p \in (0,1]$. Intuitively, given a certain measure of discrepancy from the common prior assumption (i.e., given $\rho(P_1,P_2)$), it is harder for $(H_*^p)^\infty(E)$ to cover Ω as p becomes large. For instance, in the information system in Section 2.1, for a given r (i.e., $\rho(P_1,P_2) = r$), if p is large enough ($p > r/(r+1-\varepsilon)$), then $(H_*^p)^\infty(E) = E$ ($\neq \Omega$) and hence $P_i(H_*^p)^\infty(E) = P_i(E)$ for each player $i = 1,2$. The following lemma formalizes this idea in general. More specifically, for any information system such that $\rho(P_1,P_2) = r$, whenever p is above a certain threshold, the lemma provides an upper bound on $P_i(H_*^p)^\infty(E)$ as a function of the prior probabilities assigned to E. The two important facts here are that (as in the example) the required threshold for p is an increasing function of r and that the upper bound on $P_i(H_*^p)^\infty(E)$ goes to zero as the probability of E tends to zero.

This result is the "conjugate" of Proposition 5.5 in Oyama and Tercieux (2010), where the upper bound for $P_i([(H_*^p)^\infty(E^c)]^c)$ is obtained for the many-player case.

Lemma 3.1. *For any $r \geqslant 1$, if $p > r/(1+r)$, then in any information system with $\rho(P_1, P_2) = r$, any event $E \in \mathcal{F}_1 \oplus \mathcal{F}_2$ satisfies*

$$P_i\big(\big(H_*^p\big)^\infty(E)\big) \leqslant \frac{p}{(1+r)p - r} \max\big\{P_1(E), P_2(E)\big\}$$

for all $i = 1, 2$.

Proof. See Appendix A. □

The following is the main result of this subsection, which shows that the belief potential of small probability events has an upper bound that is an increasing function of the discrepancy from the CPA.

Theorem 3.2. *For any $r \geqslant 1$ and any information system with $\rho(P_1, P_2) = r$, if $E \in \mathcal{F}_1 \oplus \mathcal{F}_2$ and $\max\{P_1(E), P_2(E)\} \leqslant \varepsilon$, then*

$$\sigma(E) \leqslant \frac{r}{1 + r - \varepsilon}.$$

Proof. Take any $q > r/(1+r-\varepsilon)$ $(> r/(1+r))$. If $\max\{P_1(E), P_2(E)\} \leqslant \varepsilon$, then by Lemma 3.1, for each $i = 1, 2$,

$$P_i\big(\big(H_*^q\big)^\infty(E)\big) \leqslant \frac{q}{(1+r)q - r}\varepsilon < 1,$$

meaning that $(H_*^q)^\infty(E) \neq \Omega$, and hence $(H_i^q)^\infty(E) \neq \Omega$. This implies that $\sigma(E) \leqslant r/(1+r-\varepsilon)$, as claimed. □

Note that the upper bound given above is tight: it is attained by the event E in the information system considered in Section 2.1.

Theorem 3.2 proves that for small probability events to have a large global impact, the value of the measure $\rho(P_1, P_2)$ must be large; otherwise it is bounded from above. Note that under the CPA where this measure is by definition minimal, the global impact of a small probability event is bounded from above. Following the construction in Section 2.1, one can readily show that for any strict Nash equilibrium a^* of any complete information game **g**, provided that measure $\rho(P_1, P_2)$ is sufficiently large, we can construct an incomplete information game in which the event "**g** is played" has a prior probability arbitrarily close to 1 (for each agent) but a^* is the unique rationalizable action at any state. Otherwise stated, any strict Nash equilibrium can be made contagious provided that we allow for the measure of discrepancy from the CPA to be large enough.

Moreover, given Proposition 5.7 in Oyama and Tercieux (2010), for any strict Nash equilibrium to be contagious, the discrepancy from the CPA (measured by $\rho(P_1, P_2)$) must also be (arbitrarily) large. Otherwise stated, whenever this measure is assumed to be bounded, we can find a complete information game and a strict Nash equilibrium that cannot be contagious. For instance, given $r < \infty$, consider the example in Section 2.1, where $\rho(P_1, P_2) = r$, and let $p > r/(1+r-\varepsilon)$ (recall (R, R) is p-dominant). Then, R is not contagious from the event E; i.e., even when R is a dominant action on E, there is a Bayesian Nash equilibrium such that L is played at every $\omega \notin E$. In fact, (L, L) being $(1-p)$-dominant and $1 - p < 1/(1+r)$, Proposition 5.7 in Oyama and Tercieux (2010) says that L is not eliminated and hence R cannot be contagious in *any* nearby incomplete information perturbation.

3.3. Local versus global impact of an event

Lipman (2003, 2010) shows that given any partition model *IS* with common support (and tail consistency in the case of infinite state space) and any state ω in the model, for any finite $N > 0$ there is a partition model with a common prior $\overline{IS}$ and a state $\bar{\omega}$ in that model at which all the same facts about the world are true and all the same statements about beliefs and knowledge of order less than N are true. That is, the common prior assumption does not impose any restriction on *finite* order beliefs. A similar argument allowed us to show in the example of Section 2.2 that the small probability event E can have an arbitrarily large *local* impact even if the common prior assumption holds.

On the other hand, the *global* properties of the state space in $\overline{IS}$ may be very different from the one in *IS* (as illustrated in the leading example). When one is interested in global properties of the whole state space, models with non-common priors may be quite far from any model with a common prior. In this subsection, we formalize this observation with the notions of global and local impact.

We measure the local impact of an event at a given state in the following way. Event E is said to have impact p on a state ω if a statement of the form "player i believes with probability a least p that player $-i$ believes with probability at least p that i believes ... that the true state is in E" is true at ω.

Definition 3.2. Event E is said to have *impact p at state ω* if $\omega \in (H_1^p)^\infty(E) \cap (H_2^p)^\infty(E)$. The *belief potential of event E at state ω*, $\sigma(\omega|E)$, is

$$\sigma(\omega|E) = \sup\big\{p \in [0,1] \bigm| E \text{ has impact } p \text{ at } \omega\big\}.$$

As we already explained, the impact of a small probability event E as measured by the belief potential characterizes the global contagion of Nash equilibria played at E. Similarly, the local impact of event E on a state ω characterizes the contagion to the state ω of Nash equilibria played at E. This connection between the belief potential at a state and the contagion of a Nash equilibrium to that state is described as follows.[4] Given an information system *IS* and an event $E = E_1 \cup E_2 \subset \Omega$ with $E_i \in \mathcal{F}_i$, $i = 1, 2$, suppose that for each player i, a_i^* is a strictly dominant action at every $\omega \in E_i$ and the payoffs at every $\omega \notin E_i$ are given by g_i. Then, if the event E has belief potential $\sigma > 0$ at state ω and the action profile (a_1^*, a_2^*) is a strict p-dominant equilibrium for some $p < \sigma$, then for each player i, a_i^* is the unique rationalizable action at ω.

Consider the information system $(\bar{\Omega}, (\bar{P}_i)_{i=1,2}, (\bar{\mathcal{Q}}_i)_{i=1,2})$ and the event $E = \{(1, 0), (2, 0)\}$ in Section 2.2. Recall that this information system satisfies the CPA. Note that, for any $p \in [0, 1)$, if $r \geqslant p/(1-p)$, we have $(H_i^p)^\infty(E) = \Omega$ $(\neq \bar{\Omega})$. This implies that, for any $p \in [0, 1)$ and $N > 0$ (provided that r is large enough), there exists ω such that it is mutually known up to order N that payoffs are given by the complete information game and where still E has a large impact on ω, more formally, $\omega \in \bigcap_{n=1}^N (K_*)^n(E^c)$ and $\sigma(\omega|E) \geqslant p$. This is the sense in which we say that a small probability event can have an arbitrarily large local impact. This is true irrespective of whether one assumes the CPA. Hence, in relation to contagion of Nash equilibrium, we thus have the following. For any strict Nash equilibrium (a_1^*, a_2^*) of a complete information game $\mathbf{g}$, we can find an incomplete information game satisfying the CPA and a state ω such that (1) the probability of the event "$\mathbf{g}$ is played" is arbitrarily close to 1, (2) it is mutually known up to arbitrarily high finite levels at ω that $\mathbf{g}$ is played, and (3) for each player i, a_i^* is the unique rationalizable action at ω.

This result is tightly connected to a result of Weinstein and Yildiz (2007) which shows, under the richness assumption, that for any type in the universal type space, there exists an arbitrarily close type (with respect to product topology in the universal type space) where rationalizability yields a unique action profile, and moreover, due to Lipman (2003) such a type can always be taken from a model with a common prior. In our language, the richness assumption asserts that for each action a_i, there is a set $Q_i \in \mathcal{Q}_i$ at which a_i is a strictly dominant action for player i (recall that $E_i = \{(-i, 0)\}$ in the example in Section 2 is such a set). Hence, the above points (1)–(3) describe Weinstein and Yildiz' (2007) result (restricted to complete information types), where the complement of the event "$\mathbf{g}$ is played" is a dominant action types set ("crazy types set") playing the role of the richness assumption.

While a small probability event can have an arbitrarily large local impact under the CPA, we show, in contrast, that the global impact of a small probability event E cannot be arbitrarily large under the CPA. The main point here is that under common prior, the set of states on which a given small probability event has a "large" impact is small with respect to prior probabilities. The following lemma formalizes this point.

Lemma 3.3. *Let $r \geqslant 1$ and $p > r/(r+1)$. For any $\delta > 0$, there exists $\varepsilon > 0$ such that for any information system with $\rho(P_1, P_2) \leqslant r$ and any event $E \in \mathcal{F}_1 \oplus \mathcal{F}_2$ such that $\max\{P_1(E), P_2(E)\} \leqslant \varepsilon$, we have*

$$P_i\big(\big\{\omega \in \Omega \bigm| \sigma(\omega|E) \geqslant p\big\}\big) \leqslant \delta$$

for all $i = 1, 2$.

Proof. Given $p > r/(r+1)$ and $\delta > 0$, set $\varepsilon = \delta\{(1+r)p - r\}/p$. Then by Lemma 3.1, we have for each $i = 1, 2$,

$$P_i\big(\big(H_i^p\big)^\infty(E)\big) \leqslant P_i\big(\big(H_*^p\big)^\infty(E)\big) \leqslant \frac{p}{(1+r)p - r}\varepsilon \leqslant \delta,$$

as claimed. □

As a corollary of the previous lemma, by setting $r = 1$ we have the following main result of this subsection (see also Section 2.2).

Proposition 3.4. *For any $p > 1/2$ and any $\delta > 0$, there exists $\varepsilon > 0$ such that for any information system $\overline{IS} = (\bar{\Omega}, (\bar{P})_{i=1,2}, (\bar{\mathcal{Q}}_i)_{i=1,2})$ that satisfies the CPA and any event $\bar{E} \in \bar{\mathcal{F}}_1 \oplus \bar{\mathcal{F}}_2$ such that $\bar{P}(\bar{E}) \leqslant \varepsilon$, we have*

$$\bar{P}\big(\big\{\omega \in \bar{\Omega} \bigm| \sigma(\omega|\bar{E}) \geqslant p\big\}\big) \leqslant \delta.$$

To summarize, under heterogeneous priors, both the local and the global impact of any small probability event can be arbitrarily large, whereas under common prior, only the local impact can be arbitrarily large.

In terms of contagion of Nash equilibria, while any strict Nash equilibrium at a small probability event can spread in some partition model with non-common priors, it may not be the case for partition models with a common prior. Indeed, in 2×2 coordination games, the risk-dominated equilibrium cannot spread from a small probability event when we assume the existence of a common prior, as shown by Kajii and Morris (1997). In light of the result of Weinstein and Yildiz

[4] See the working paper version Oyama and Tercieux (2005) for details.

(2007) described above in common prior models the set in which the risk-dominated equilibrium is played as a unique rationalizable outcome is assigned probability close to zero by the common prior whenever this prior assigns a vanishingly small probability to the crazy types set. That is, *one cannot simultaneously have a common prior, make any given rationalizable action uniquely rationalizable on a large probability set, and assign small probability to the crazy types set*.

4. Discussion

4.1. Alternative discrepancy measures

Here we discuss our discrepancy measure ρ in relation to other measures in the literature.

The measure ρ is equivalent to the *separation distance*:

$$d_S(P_1, P_2) = \sup_{\omega \in \Omega}\left(1 - \frac{P_1(\omega)}{P_2(\omega)}\right)$$

(Aldous and Diaconis, 1987; see also Gibbs and Su, 2002). One can observe that

$$\max\{d_S(P_1, P_2), d_S(P_2, P_1)\} = 1 - \frac{1}{\rho(P_1, P_2)}.$$

An analogous notion to ρ, "δ-closeness", is used in Monderer and Samet (1995): translated into our setting, P_1 and P_2 are said to be δ*-close* if

$$e^{-\delta} < \frac{P_1(\omega)}{P_2(\omega)} < e^{\delta}$$

for all $\omega \in \Omega$. If $\rho(P_1, P_2) = r$, then P_1 and P_2 are $(\log r)$-close.

Another measure is the one defined by

$$d_0(P_1, P_2) = \sup_{E \subset \Omega}\left|P_1(E) - P_2(E)\right|,$$

which equals $(1/2)\sum_{\omega \in \Omega}|P_1(\omega) - P_2(\omega)|$. Note that the information system satisfies the CPA if and only if $d_0(P_1, P_2) = 0$. One can show that $d_0(P_1, P_2) \leqslant 1 - \rho(P_1, P_2)^{-1}$, while the reverse inequality does not hold in general. While this measure gives an upper bound of the belief potential analogously in Theorem 3.2 (although it is tight only in the limit as $P_i(E) \to 0$), it does not allow to prove the analogue of Lemma 3.3; see Oyama and Tercieux (2005, Section 4.2). The point is that there exists no direct connection between proximity of priors using d_0 and proximity of conditional beliefs. This has to be contrasted with the measure ρ used in this paper. To understand this point, let us consider (without loss of generality) a fixed Ω. For the measure ρ, we have that for any sequence $\{(P_i^m)_{i=1,2}\}_{m=0}^{\infty}$, $\rho(P_1^m, P_2^m) \to 1$ if and only if

$$\max_{i,j}\sup_{E,F \subset \Omega}\frac{P_i^m(E|F)}{P_j^m(E|F)} \to 1,$$

where $P_i^m(E|F)$ is the conditional probability of event E given F in terms of P_i.

4.2. Many players

A straightforward way to consider the case with many (say, I) players would be to extend the definition of H_*^p (used in Lemma 3.1) by $H_*^p(E) = \bigcup_{i=1}^{I} B_i^p(E)$ and employ a version of belief potential defined by $\sigma'(E) = \sup\{p \in [0,1] \mid (H_*^p)^{\infty}(E) = \Omega\}$. Then, by Proposition 5.5 in Oyama and Tercieux (2010) we have $\sigma'(E) \leqslant r(I-1)/\{1 + r(I-1) - \varepsilon\}$ when $\max_i P_i(E) \leqslant \varepsilon$. In particular, $\sigma'(E)$ is bounded by $(I-1)/(I-\varepsilon)$ under the CPA.

With this definition, however, the tight connection to the contagion effect will be lost, i.e., for $I > 2$, the p-dominance of a Nash equilibrium smaller than the number $\sigma'(E)$ so defined, while necessary, is not sufficient for contagion of that equilibrium. If we instead use an appropriate definition of belief potential that allows a contagion result, then the upper bound given by H_*^p will not be tight. Still, our qualitative point will hold true that any strict Nash equilibrium is contagious from a small probability event in some information system with sufficiently heterogeneous priors and under the CPA this is not the case for some strict Nash equilibria. Obtaining a quantitative result (i.e., deriving a tight upper bound of an appropriately defined belief potential) is left for future research.

Appendix A. Proof of Lemma 3.1

We first note the following, which is essentially equivalent to Lemma A in Kajii and Morris (1997).

Lemma A.1. *Let $p > 0$. For any event E and player i, if $F_i \in \mathcal{F}_i$ and $F_i \subset B_i^p(E)$, then $P_i(F_i \setminus E) \leqslant ((1-p)/p)P_i(F_i \cap E)$.*

Fix $r \geqslant 1$, and consider any information system with $\rho(P_1, P_2) = r$ and any event $E = E_1 \cup E_2$, each $E_i \in \mathcal{F}_i$. In the following, we want to obtain an upper bound for $P_j((H_*^p)^K(E))$. The argument proceeds as follows. We consider families of events ($\{D_1^\ell \setminus E^{\ell-1}\}_{\ell=1}^K$ and $\{D_2^\ell \setminus E^{\ell-1}\}_{\ell=1}^K$ to be defined below) that (together with E) cover the event $(H_*^p)^K(E)$. To bound the probabilities of these events, we first use Lemma A.1 as an intermediate step in Lemma A.2. We then derive a recursive formula of upper bounds in Lemma A.3, which gives an upper bound of $P_j((H_*^p)^K(E))$ in Lemma A.4. The key is to obtain tight enough bounds in Lemma A.3 so that the limit of the bound in Lemma A.4 as $K \to \infty$ remains finite whenever $p > r/(1+r)$.

Let $E_i^0 = E_i$ and $E^0 = E_1^0 \cup E_2^0$. Given $K \geqslant 1$ and $p \in (0, 1]$, define $\{E_1^k, E_2^k, E^k\}_{k=1}^K$ recursively by

$$E_i^k = B_i^p\big(E^{k-1}\big), \quad E^k = E_1^k \cup E_2^k.$$

Then, $(H_*^p)^K(E) = E^K$. Let $D_i^0 = E_i^0$ and $D_i^k = E_i^k \setminus E_i^{k-1}$ for $k = 1, \dots, K$. Observe that $D_i^k \in \mathcal{F}_i$ for all k.

For $i, j = 1, 2$, let $x_i(j, 0) = 0$, and

$$x_i(j,k) = \sum_{\ell=1}^{k} P_i\big(D_j^\ell \setminus E^{\ell-1}\big) \tag{A.1}$$

and

$$x_i(k) = x_i(1,k) + x_i(2,k)$$

for $k = 1, \dots, K$. Let also

$$z_i(j,k) = \sum_{\ell=1}^{k} P_i\big(D_j^\ell \cap E^{\ell-1}\big) \tag{A.2}$$

for $k = 1, \dots, K$. Note that

$$P_i\big(\big(H_*^p\big)^K(E)\big) \leqslant P_i(E) + x_i(K) \tag{A.3}$$

and that

$$z_i(j,k) \leqslant x_i(-j,k-1) + P_i\big(E_{-j}^0 \setminus E_j^0\big). \tag{A.4}$$

By using Lemma A.1, we have the following.

Lemma A.2. *For all $k = 1, \dots, K$ and $i = 1, 2$,*

$$x_i(i,k) \leqslant \frac{1-p}{p} z_i(i,k).$$

Now, $\rho(P_1, P_2) = r$ implies that $x_i(j,k) \leqslant r x_{-i}(j,k)$ and $z_i(j,k) \leqslant r z_{-i}(j,k)$. Thus by Lemma A.2, we have the following.

Lemma A.3. *For all $k = 1, \dots, K$ and $i = 1, 2$,*

$$x_i(k) \leqslant \frac{r(1-p)}{p} x_{-i}(k-1) + \frac{r(1-p)}{p} P_{-i}\big(E^0\big).$$

Proof. By Lemma A.2 and (A.4),

$$\begin{aligned}
x_i(k) &= x_i(i,k) + x_i(-i,k) \\
&\leqslant x_i(i,k) + r x_{-i}(-i,k) \\
&\leqslant \frac{1-p}{p} z_i(i,k) + \frac{r(1-p)}{p} z_{-i}(-i,k) \\
&\leqslant \frac{r(1-p)}{p}\big(z_{-i}(i,k) + z_{-i}(-i,k)\big) \\
&\leqslant \frac{r(1-p)}{p} x_{-i}(k-1) + \frac{r(1-p)}{p} P_{-i}\big(E^0\big),
\end{aligned}$$

as claimed. □

By recursively applying Lemma A.3 to (A.3), we obtain the upper bound of $P_i((H_*^p)^K(E))$.

Lemma A.4. *In any information system with $\rho(P_1, P_2) = r$, any event $E \in \mathcal{F}_1 \oplus \mathcal{F}_2$ satisfies*

$$P_i\big(\big(H_*^p\big)^K(E)\big) \leqslant \max\big\{P_1(E), P_2(E)\big\} \sum_{k=0}^{K} \left\{\frac{r(1-p)}{p}\right\}^k \tag{A.5}$$

for all $i = 1, 2$.

We are now in a position to prove Lemma 3.1. It remains to consider the limit of the right hand side of (A.5) as $K \to \infty$. This is where the assumption that $p > r/(1+r)$ is used.

Proof of Lemma 3.1. If $p > r/(1+r)$, or $r(1-p)/p < 1$, then the right hand side of (A.5), $\sum_{k=0}^{K}\{r(1-p)/p\}^k$, converges to

$$\frac{1}{1 - \frac{r(1-p)}{p}} = \frac{p}{(1+r)p - r}$$

as $K \to \infty$. Hence, by Lemma A.4 we have the desired inequality. $\square$

References

Aldous, D., Diaconis, P., 1987. Strong uniform times and finite random walks. Adv. Appl. Math. 8, 69–97.
Aumann, R.J., 1976. Agreeing to disagree. Ann. Statist. 4, 1236–1239.
Brandenburger, A., Dekel, E., 1993. Hierarchies of beliefs and common knowledge. J. Econ. Theory 59, 189–198.
Carlsson, H., van Damme, E., 1993. Global games and equilibrium selection. Econometrica 61, 989–1018.
Gibbs, A.L., Su, F.E., 2002. On choosing and bounding probability metrics. Int. Statist. Rev. 70, 419–435.
Kajii, A., Morris, S., 1997. The robustness of equilibria to incomplete information. Econometrica 65, 1283–1309.
Lipman, B.L., 2003. Finite order implications of common priors. Econometrica 71, 1255–1267.
Lipman, B.L., 2010. Finite order implications of common priors in infinite models. J. Math. Econ. 46, 56–70.
Mertens, J.F., Zamir, S., 1985. Formulation of Bayesian analysis for games with incomplete information. Int. J. Game Theory 14, 1–29.
Milgrom, P., Stokey, N., 1982. Information, trade, and common knowledge. J. Econ. Theory 26, 17–27.
Monderer, D., Samet, D., 1989. Approximating common knowledge with common beliefs. Games Econ. Behav. 1, 170–190.
Monderer, D., Samet, D., 1995. Stochastic common learning. Games Econ. Behav. 9, 161–171.
Morris, S., Rob, R., Shin, H.S., 1995. p-Dominance and belief potential. Econometrica 63, 145–157.
Oyama, D., Tercieux, O., 2005. On the strategic impact of an event under non-common priors, ssrn.com/abstract=998022.
Oyama, D., Tercieux, O., 2010. Robust equilibria under non-common priors. J. Econ. Theory 145, 752–784.
Rubinstein, A., 1989. The electronic mail game: strategic behavior under 'almost common knowledge'. Amer. Econ. Rev. 79, 385–391.
Weinstein, J., Yildiz, M., 2007. A structure theorem for rationalizability with application to robust predictions of refinements. Econometrica 75, 365–400.
Yildiz, M., 2004. Generic uniqueness of rationalizable actions. Mimeo.