A Foundation for Proper Rationalizability from an Incomplete Information Perspective

Andrés Perea and Souvik Roy
Maastricht University

August 6, 2010

Abstract

Proper rationalizability (Schuhmacher (1999), Asheim (2001)) is a concept in epistemic game theory that is based on two assumptions: (1) every player is *cautious,* i.e., does not exclude any opponent's choice from consideration, and (2) every player *respects* the opponent's preferences, i.e., deems one opponent's choice to be infinitely more likely than another whenever he believes the opponent to prefer the one to the other. In this paper we provide a new foundation for proper rationalizability, by assuming that players have incomplete information about the opponent's utilities. We show that, if the uncertainty of each player about the opponent's utilities vanishes gradually in some regular manner, then the choices he can rationally make under common belief in rationality are all properly rationalizable in the original game with no uncertainty about the opponent's utilities.

JEL classification: C72

Keywords and phrases: Epistemic game theory, incomplete information, proper rationalizability.

Contact information:

Department of Quantitative Economics
Maastricht University
P.O. Box 616, 6200 MD Maastricht
The Netherlands.
E-mail: a.perea@maastrichtuniversity.nl, s.roy@maastrichtuniversity.nl.

	d	e	f
a	$0,2$	$1,1$	$1,0$
b	$1,2$	$0,1$	$1,0$
c	$1,2$	$1,1$	$0,0$

Figure 1: An example for proper rationalizability

1 Introduction

Epistemic game theory deals with the ways the players may reason about their opponents before making a decision. More precisely, in epistemic game theory players base their choices on the beliefs about the opponents' behavior, which in turn depend on their beliefs about the opponents' beliefs about others' behavior, and so on. A major goal of epistemic game theory is to study such infinite belief hierarchies, to impose reasonable conditions on these, and to investigate their behavioral implications.

A central idea in epistemic game theory is *common belief in rationality* (Tan and Werlang (1988)), stating that a player believes that his opponents choose rationally, believes that his opponents believe that their opponents choose rationally, and so on. In our view, one of its most natural refinements is the concept of *proper rationalizability* (Schuhmacher (1999) and Asheim (2001)), which is based on Myerson's (1978) notion of proper equilibrium, but without imposing any equilibrium assumption. Proper rationalizability is based on the following two conditions: The first states that players are cautious, meaning that they do not exclude any opponents' choice from consideration. The second condition is an extension of Myerson's ϵ-proper trembling condition, which states that whenever you believe that a choice a is better than another choice b for your opponent, then the probability you assign to b must be at most ϵ times the probability you assign to a. Under ϵ-proper rationalizibility there is common belief in the event that every player is cautious and satisfies the ϵ-proper trembling condition. A choice is called properly rationalizable if it can be chosen under ϵ-proper rationalizability for every $\epsilon > 0$.

We will now explain this concept by means of an example. Consider the game in Figure 1, where player 1 chooses between a, b and c and player 2 chooses between d, e and f. Note that for player 2, choice d is better than choice e, and choice e is better than choice f. Hence, under proper rationalizability player 1 deems d for player 2 much more likely than e, and e much more likely than f. Consequently, only choice c will be optimal for player 1. So, if $\epsilon > 0$ is small enough, then only choices c and d can rationally be made under ϵ-proper rationalizability. As such, only the choices c for player 1 and d for player 2 are properly rationalizable.

The usual interpretation of proper rationalizability is that you assume that your opponent makes mistakes, but that you deem more costly mistakes much less likely than less costly mistakes. In this paper we offer a rather different foundation for proper rationalizability. Instead of assuming that you believe your opponent to make mistakes, we rather suppose that you have uncertainty about his utility function, while believing that he chooses rationally. We thus consider a game with incomplete information. Our main result states that, if we let the uncertainty about the opponent's utility go to zero in some regular manner, then every choice that can rationally be made under common belief in rationality in the game with incomplete information, will be properly rationalizable in the original game, in which there is no uncertainty about the opponent's utilities.

In the game with incomplete information, we impose some regularity conditions on the players' beliefs about the opponent's utility functions which can be summarized as follows: First, for every outcome in the game, the belief that player i has about player j's utility from this outcome, is always normally distributed with its mean at the "original" utility in the original game. As a consequence, player i deems any utility function possible for player j, and hence every choice for player j can be optimal for some utility function deemed possible by i. Together with the condition that i believes in j's rationality, this actually makes sure that player i deems every choice possible for player j, thus mimicking the cautiousness condition described above. Secondly, i's belief about j's utility function should be independent from his belief about j's belief hierarchy. This makes intuitive sense since j's belief hierarchy is an epistemic property of this player, whereas his utility function is not. So there is no obvious reason to expect any correlation between these two characteristics. Thirdly, i's belief about j's utilities from different outcomes in the game should be independent from each other. Possibly some of these conditions can be relaxed for the proof of our main result, but we leave this issue for future research.

Our game with incomplete information is related to the one used in Dekel and Fudenberg (1990). They also consider games with incomplete information where the player's uncertainty about the opponent's utilities goes to zero. An important difference with our approach is that Dekel and Fudenberg apply the concept of *iterated elimination of weakly dominated choices* to the games with incomplete information. They show that if the uncertainty about the opponent's utilities vanishes, then we obtain one round of deletion of weakly dominated strategies, followed by iterated deletion of strongly dominated strategies, in the original game. The latter procedure is also called the Dekel-Fudenberg procedure in the literature. In contrast, we apply common belief in rationality to our games with incomplete information. We then show that if the uncertainty about the opponent's utilities vanishes, we obtain a subselection (that is some, but in general not all) of the properly rationalizable choices in the original game, which is fundamentally different from the Dekel-Fudenberg procedure. Another fundamental difference between our paper and Dekel and Fudenberg lies in the way the uncertainty about the opponent's utilities is modeled. Their model assumes that players only deem possible finitely many utility functions for the opponent, and that a large probability must be assigned to the opponent's "original" utility function. In contrast, we assume that the uncertainty about the opponent's utilities is given by a normal distribution. In particular, players deem *every* utility function possible for the opponent.

The paper is organized as follows: In Section 2 we introduce our epistemic model for games with incomplete information, we formalize the idea of common belief in rationality for these games, and show that common belief in rationality is always possible. In Section 3 we introduce our epistemic model for games with complete information, and present the concept of proper rationalizability for these games. In Section 4 we state our main result, establishing the connection between common belief in rationality in the game with incomplete information (in the presence of small uncertainty about the opponent's utility function), and proper rationalizability in the original game. In Section 5 we provide some concluding remarks. All proofs are collected in Section 6.

2 Rationalizability in Games with Incomplete Information

2.1 Epistemic Model

Throughout this paper we restrict attention to games with two players. Let $\Gamma = (C_i, u_i)_{i \in I}$ be a finite, static game where $I = \{1, 2\}$ is the set of players, C_i is the finite set of choices of player i, and u_i is player i 's utility function. The function u_i assigns to every pair of choices $(c_{1,}c_2) \in C_1 \times C_2$ a utility $u_i\,(c_1, c_2) \in \mathbb{R}$.

In a game with incomplete information players do not only have uncertainty about the opponent's choices, they also have uncertainty about the opponent's utility function. Hence a belief hierarchy should not only specify what the player believes about the opponent's choice but also what he believes about the opponent's utility function. Not only this, it should also specify what the player believes about the opponent's belief about his own choice and utility function, and so on. A possible way of modeling such belief hierarchies is by means of the following definition.

Definition 2.1 (Epistemic model) *An epistemic model for* Γ *with incomplete information is a tuple* $M = (T_i, b_i, v_i)_{i \in I}$ *where (1)* T_i *is the set of types for player* i*, (2)* $b_i : T_i \longrightarrow \triangle\,(C_j \times T_j)$ *is the belief assignment taking only finitely many different probability distributions on* $\triangle\,(C_j \times T_j)$*, and (3)* v_i *is the utility assignment that assigns to every* $t_i \in T_i$ *a utility function* $v_i\,(t_i) : C_1 \times C_2 \longrightarrow \mathbb{R}$.

By $\triangle\,(X)$ we denote the set of probability distributions on X. So, in an epistemic model, each type t_i has a belief about player j's choice-type combinations. And hence, in particular, it has a belief about j's choice. But, as player j's type also specifies his utility function and his belief about player i's choice, player i also has some belief about player j's utility function, and about player j's belief about his own choice, and so on. In this way one can derive a complete belief hierarchy for every given type.

Note that each type t_i can be indentified with a pair $(v_i(t_i), b_i(t_i))$ where $v_i(t_i)$ is its utility function and $b_i(t_i)$ is its belief hierarchy. Since we required the belief assignment to take only finitely many different probability distributions, the epistemic model contains only finitely many different belief hierarchies.

2.2 Restrictions on the Epistemic Model

Our goal will be to model the situation where the players have uncertainty about the opponent's utility function, but where this uncertainty "vanishes in the limit". In order to formalise this we need to impose additional restrictions on the epistemic model.

Recall that every type t_i can be identified with a pair $(v_i(t_i), b_i(t_i))$, where $v_i(t_i)$ is t_i's utility function and $b_i(t_i)$ is its belief hierarchy. Denote by V_i the set of all possible utility functions, and by B_i the set of all belief hierarchies in the epistemic model $M = (T_i, v_i, b_i)_{i \in I}$. The first condition we impose is that $T_i = V_i \times B_i$, that is, for every possible utility function we can think of, and every belief hierarchy in the model, there exists a type in the model with exactly this combination of utility function and belief hierarchy. So in a sense we assume that the type space is rich enough.

Secondly we assume that t_i's belief about j's utility from (c_1, c_2) is statistically independent from its belief about j's utility from (c_1', c_2') whenever $(c_1, c_2) \neq (c_1', c_2')$, and that this belief is also statistically independent from its belief about j's belief hierarchy.

Finally we assume that t_i's beliefs about j's utilities from the various outcomes in the game are all induced by a unique normal distribution. More formally, t_i's belief about j's utility from (c_1, c_2) is given by a normal distribution with its mean at $u_j(c_1, c_2)$ – the "true" utility of player j in the original game. So, all these beliefs are distributed identically around the mean. By collecting all these conditions we arrive at the following definition.

Definition 2.2 (σ-regular epistemic model) *Let P be the normal distribution on $\mathbb{R}$ with mean 0 and variance $\sigma^2 > 0$. Then an epistemic model $M = (T_i, b_i, v_i)_{i \in I}$ is σ-regular if for both players i, (1) $T_i = V_i \times B_i$, (2) for every type $t_i \in T_i$, his belief about j's utility from (c_1, c_2) is statistically independent from his belief about j's utility from (c_1', c_2') whenever $(c_1, c_2) \neq (c_1', c_2')$, and his belief about j's utilities is statistically independent from his belief about j's belief hierarchy, and (3) for every type $t_i \in T_i$, and every choice-pair (c_1, c_2), the belief of t_i about j's utility from (c_1, c_2) is given by P, up to a shift of the mean to $u_j(c_1, c_2)$.*

2.3 σ-Rationalizability

In this subsection we will define common belief in rationality inside an epistemic model with incomplete information. In addition, if we require the epistemic model to be σ-regular for a given normal distribution with mean 0 and variance σ^2, then we obtain the concept of σ-rationalizability.

We first need some more notation. For given type t_i and choice c_i, let $v_i(t_i)(c_i)$ be the expected utility for type t_i from choosing c_i, given his belief $b_i(t_i)$ about the opponent's choice, and given his utility function $v_i(t_i)$.

Definition 2.3 (Rational choice) *A choice c_i is rational for t_i if $v_i(t_i)(c_i) \geq v_i(t_i)(c_i')$ for all $c_i' \in C_i$.*

We will now define common belief in rationality. In words it says that a player believes that his opponent makes rational choices, and believes that his opponent believes that he makes rational choices, and so on.

Formally, for every $\widetilde{T}_i \subseteq T_i$, let

$$\left(C_i \times \widetilde{T}_i\right)^{rat} = \{(c_i, t_i) \in C_i \times \widetilde{T}_i : c_i \text{ is rational for } t_i\}.$$

Definition 2.4 (Common belief in rationality) *For both players i we define subsets of types $T_i^1, T_i^2, \ldots$ in a recursive way as follows:*

$$\begin{aligned} T_i^1 &: = \{t_i \in T_i : b_i(t_i)[(C_j \times T_j)^{rat}] = 1\}, \\ T_i^2 &: = \{t_i \in T_i : b_i(t_i)[(C_j \times T_j^1)^{rat}] = 1\}, \\ &\vdots \\ T_i^l &: = \{t_i \in T_i : b_i(t_i)[(C_j \times T_j^{l-1})^{rat}] = 1\}, \\ &\vdots \end{aligned}$$

Type t_i expresses common belief in rationality if $t_i \in \cap_{l \in \mathbb{N}} T_i^l$.

A type t_i is σ-rationalizable if it expresses common belief in rationality within a σ-regular epistemic model.

Definition 2.5 (σ-rationalizable type) *Let $M = (T_i, b_i, v_i)_{i \in I}$ be a σ-regular epistemic model. Every type $t_i \in T_i$ that expresses common belief in rationality is called σ-rationalizable.*

Now we show that σ-rationalizable types always exist.

Theorem 2.1 (σ-rationalizable types always exist) *Consider a finite static game $\Gamma = (C_i, u_i)_{i \in I}$, and some $\sigma > 0$. Then there is a σ-regular epistemic model $M = (T_i, b_i, v_i)_{i \in I}$ for Γ where all types are σ-rationalizable.*

The proof can be found in Section 6.

2.4 Limit Rationalizability

In this subsection we focus on those choices which can rationally be made under common belief in rationality when the uncertainty about the opponent's utility vanishes. This will lead to the concept of limit rationalizability. We first need an additional definition.

Definition 2.6 (Constant type spaces and utility assignments) *A sequence of epistemic models* $((T_i^n, b_i^n, v_i^n)_{i\in I})_{n\in\mathbb{N}}$ *has constant type spaces and utility assignments if* $T_i^n = T_i^m$ *and* $v_i^n = v_i^m$ *for all* n *and* m*, and for both players* i*.*

We are now ready to define the concept of limit rationalizable choice.

Definition 2.7 (Limit rationalizable choice) *Consider a finite static game* $\Gamma = (C_i, u_i)_{i\in I}$ *with two players. A choice* c_i *is limit rationalizable if there is a sequence* $(\sigma_n)_{n\in\mathbb{N}} \to 0$*, and a sequence* $(M^n)_{n\in\mathbb{N}}$ *of* σ_n*-regular epistemic models with constant type spaces and utility assignments, such that in every* M^n *there is a* σ_n*-rationalizable type* t_i^n *with utility function* u_i*, for which choice* c_i *is optimal.*

3 Proper Rationalizability in Games with Complete Information

3.1 Epistemic Model

Let $\Gamma = (C_i, u_i)_{i\in I}$ be a finite, static game with two players. In a game with complete information players do not have uncertainty about the opponent's utility function. Therefore a belief hierarchy only needs to specify what a player believes about the opponent's choice, what he believes about the opponent's belief about his own choice, and so on. Therefore the epistemic model will be simpler compared to the case of incomplete information.

Definition 3.1 (Epistemic model) *An epistemic model for* Γ *with complete information is a tuple* $M = (\Theta_i, \beta_i)_{i\in I}$ *where (1)* Θ_i *is the finite set of types for player* i*, and (2)* $\beta_i : \Theta_i \longrightarrow \triangle(C_j \times \Theta_j)$ *is the belief assignment.*

So, in an epistemic model, each type θ_i has a belief about player j's choice-type combinations. And hence, in particular, it has a belief about j's choice. But, as player j's type also specifies his belief about player i's choice, player i also has some belief about player j's belief about his own choice, and so on. In this way one can derive a complete belief hierarchy for every given type.

For given type θ_i and choice c_i we define $u_i(c_i, \theta_i)$ as the expected utility for type θ_i from choosing c_i given his belief $\beta_i(\theta_i)$ about his opponent's choice (and given his "fixed" utility function u_i). Type θ_i is said to *prefer* choice c_i to choice c_i' when $u_i(c_i, \theta_i) > u_i(c_i', \theta_i)$. We say that a type θ_i *considers possible* some opponent's type θ_j if $\beta_i(\theta_i)(c_j, \theta_j) > 0$ for some $c_j \in C_j$. Now we introduce the key condition in proper rationalizability, which is the ϵ-proper trembling condition. Intuitively it says that (1) a player should deem possible all opponent's choices, and (2) if a player believes choice a is better than choice b for the other player, then he should deem choice a much more likely than choice b.

Definition 3.2 (ϵ-proper trembling condition) *Let $\epsilon > 0$. A type θ_i satifies the ϵ-proper trembling condition if*
(1) for each θ_j that θ_i deems possible, $\beta_i(\theta_i)(c_j, \theta_j) > 0$ for all $c_j \in C_j$, and
(2) for every θ_j that θ_i deems possible, whenever θ_j prefers c_j to c'_j, then $\beta_i(\theta_i)\left(c'_j, \theta_j\right) \leq \epsilon \cdot \beta_i(\theta_i)(c_j, \theta_j)$.

So, the first condition says that whenever θ_i deems some type θ_j possible, θ_i also assumes every choice is possible for θ_j.

Proper rationalizability is based on the event that the types should not only satisfy the ϵ-proper trembling condition themselves, but also express common belief in the event that types satisfy the ϵ-proper trembling condition.

Definition 3.3 (ϵ-properly rationalizable type) *A type θ_i is ϵ-properly rationalizable if:*
θ_i satisfies the ϵ-proper trembling condition,
θ_i only deems possible opponent's types θ_j which satisfy the ϵ-proper trembling condition,
θ_i only deems possible opponent's types θ_j which only deem possible player i's types θ'_i which satisfy the ϵ-proper trembling condition, and so on.

Properly rationalizable choices are those choices which can rationally be made by ϵ-properly rationalizable types, for all ϵ.

Definition 3.4 (Properly rationalizable choice) *A choice c_i is ϵ-properly rationalizable if there is an epistemic model and an ϵ-properly rationalizable type θ_i within it for which c_i is optimal. A choice c_i is properly rationalizable if it is ϵ-properly rationalizable for all $\epsilon > 0$.*

3.2 Example

Consider again the game in Figure 1. Let the type sets of player 1 and player 2 be $\Theta_1 = \{\theta_1, \theta'_1\}$ and $\Theta_2 = \{\theta_2, \theta'_2\}$. For $\epsilon > 0$ (small), let the beliefs for the types be given by

$$\begin{aligned}
\beta_1(\theta_1) &= \left(1 - \epsilon^2 - \epsilon^3\right)(d, \theta_2) + \epsilon^2 (e, \theta_2) + \epsilon^3 (f, \theta_2), \\
\beta_1(\theta'_1) &= \frac{1}{6}(d, \theta_2) + \frac{1}{6}(e, \theta_2) + \frac{1}{6}(f, \theta_2) + \frac{1}{6}\left(d, \theta'_2\right) + \frac{1}{6}\left(e, \theta'_2\right) + \frac{1}{6}\left(f, \theta'_2\right), \\
\beta_2(\theta_2) &= \left(1 - \epsilon^2 - \epsilon^3\right)(c, \theta_1) + \epsilon^2 (b, \theta_1) + \epsilon^3 (a, \theta_1), \text{ and} \\
\beta_2(\theta'_2) &= \frac{1}{6}(a, \theta_1) + \frac{1}{6}(b, \theta_1) + \frac{1}{6}(c, \theta_1) + \frac{1}{6}\left(a, \theta'_1\right) + \frac{1}{6}\left(b, \theta'_1\right) + \frac{1}{6}\left(c, \theta'_1\right).
\end{aligned}$$

It may be verified that the types θ_1 and θ_2 both satisfy the ϵ-proper trembling condition. Also, type θ_1 only deems possible the opponent's type θ_2, and θ_2 only deems possible the opponent's type θ_1. This implies that both θ_1 and θ_2 are ϵ-properly rationalizable. So, choice c

for player 1, and d for player 2 are ϵ-properly rationalizable for any $\epsilon > 0$ small enough. Hence, choice c for player 1, and d for player 2 are properly rationalizable.

On the other hand, we see that the type θ_1' of player 1 believes that the choices d, e and f are equally likely to be taken by type θ_2 of player 2 while for type θ_2, d is better than e, and e is better than f. So, type θ_1' of player 1 does not satisfy the ϵ-proper trembling condition. Similarly, type θ_2' also does not satisfies the ϵ-proper trembling condition.

4 Main Result

4.1 Statement of the Main Result

For a static game we analysed two contexts, one with incomplete information and another with complete information. In the context with incomplete information, where players have uncertainty about the opponent's utility, we introduced the concept of a limit rationalizable choice. In the context with complete information, where players have no uncertainty about the opponent's utility, we discussed the concept of a properly rationalizable choice. In our main result we connect these two concepts.

Theorem 4.1 (Limit rationalizability implies proper rationalizability) *Consider a finite static game with two players. Every limit rationalizable choice for the context with incomplete information is a properly rationalizable choice for the context with complete information.*

4.2 Illustration of the Main Result

By means of an example we provide some intuition for our main result. More precisely we show how a σ-rationalizable type in the context of incomplete information can be transformed into an ϵ-properly rationalizable type in the context of complete information. Also we show that when σ goes to zero then ϵ goes to zero as well.

Consider again the game from Figure 1. Let us start with the context of incomplete information. Let P be the normal distribution with mean 0 and variance σ^2. From the proof of Theorem 2.1 we know that there exists a σ-regular epistemic model $M = (T_i, b_i, v_i)_{i\in I}$ where every type is σ-rationalizable and all the types have the same belief hierarchy. So, types only differ by their utility function. For each of the types t_1 of player 1 we denote by β_1 the belief about player 2's choice, and for each type t_2 let β_2 be the belief about player 1's choice. As we assume that all the types have the same belief hierarchy, β_1 and β_2 are unique.

For both players i let Q_i be the probability distribution on player i's utility functions generated by P. Since the epistemic model is σ-regular every type t_j has the belief Q_i about i's utility function. Let $V_i(c_i, \beta_i)$ be the set of utility functions for player i such that choice c_i is optimal under the belief β_i about the opponent's choice. Since every type t_i expresses common belief in rationality, the probability it assigns to an opponent's choice c_j is exactly the probability it assigns to the event that j's utility function is in $V_j\left(c_j, \beta_j\right)$, which is $Q_j\left(V_j\left(c_j, \beta_j\right)\right)$. So, we can derive the following six equations:

$$\begin{aligned}
\beta_1(d) &= Q_2(V_2(d, \beta_2)) \\
\beta_1(e) &= Q_2(V_2(e, \beta_2)) \\
\beta_1(f) &= Q_2(V_2(f, \beta_2)) \\
\beta_2(a) &= Q_1(V_1(a, \beta_1)) \\
\beta_2(b) &= Q_1(V_1(b, \beta_1)) \\
\beta_2(c) &= Q_1(V_1(c, \beta_1)).
\end{aligned}$$

Since P has full support on $\mathbb{R}$, it follows that all these probabilities are positive.

Now we turn to the context of complete information. We construct an epistemic model with a single type θ_1 for player 1 and a single type θ_2 for player 2. Let the belief of θ_1 about player 2's choice be given by the β_1 constructed above, and similarly for the belief of θ_2. So, the belief about the opponent's choice has not changed by moving from the context with incomplete information to the context with complete information.

Since in the original game d is better than e and e is better than f for player 2, for small σ we will have that $Q_2(V_2(d, \beta_2))$ is much bigger than $Q_2(V_2(e, \beta_2))$, and $Q_2(V_2(e, \beta_2))$ is much bigger than $Q_2(V_2(f, \beta_2))$. So, by our equations above we have that $\beta_1(d)$ is much bigger than $\beta_1(e)$, and $\beta_1(e)$ is much bigger than $\beta_1(f)$. Given such a β_1, in the original game c will be better than b and b will be better than a. So, similarly, for small σ we will have that $Q_1(V_1(c, \beta_1))$ is much bigger than $Q_1(V_1(b, \beta_1))$, and $Q_1(V_1(b, \beta_1))$ is much bigger than $Q_1(V_1(a, \beta_1))$. And hence, from the equations above, we have that $\beta_2(c)$ is much bigger than $\beta_2(b)$, and $\beta_2(b)$ is much bigger than $\beta_2(a)$. Now define

$$\epsilon = \max\{\frac{\beta_2(a)}{\beta_2(b)}, \frac{\beta_2(b)}{\beta_2(c)}, \frac{\beta_1(e)}{\beta_1(d)}, \frac{\beta_1(f)}{\beta_1(e)}\}.$$

Then, by construction, θ_1 and θ_2 are ϵ-properly rationalizable. Moreover, if σ goes to zero then the associated ϵ would go to zero as well.

If the variance of P is small then choice c is optimal for the σ-rationalizable type t_1 in the model with incomplete information that has the original utility function. Similarly, d is optimal for the σ-rationalizable type t_2 that has the original utility function in the model with incomplete information. As a consequence, c and d are limit rationalizable in the context with incomplete information. On the other hand, in the associated epistemic model with complete information c is optimal for the ϵ-properly rationalizable type θ_1 and d is optimal for the ϵ-properly rationalizable type θ_2. As ϵ goes to zero when σ goes to zero, we conclude that c and d are properly rationalizable. So, in this example the limit rationalizable choices are also properly rationalizable.

5 Concluding remarks

We believe that proper rationalizability is a very natural concept in game theory, but it has not yet received the attention it deserves. In this paper we have established a new foundation for proper rationalizability from the viewpoint of games with incomplete information. In games with incomplete information we define a choice as limit rationalizable if it can rationally be made under common belief of rationality when the uncertainty vanishes gradually in some regular way. We show the existence of such choices. We then prove that each limit rationalizable choice in the game with incomplete information is properly rationalizable for the context with complete information.

Throughout this paper it is assumed that the players' uncertainty about the opponent's utilities are described by a normal distribution. We have used the normal distribution as it is a very natural candidate to describe the uncertainty. We believe, however, that we can extend our framework to wider classes of probability distributions here, as long as this class is closed under taking convex combinations, and Lemma 6.4 is satisfied.

In this paper we restricted our attention to two players for the sake of simplicity. However, we believe our result can be extended to more than two players in a natural way.

6 Proofs

6.1 Existence of σ-Rationalizable Types

We prove Theorem 2.1, which guarantees the existence of σ-rationalizable types. Consider a finite static game $\Gamma = (C_i, u_i)_{i \in I}$, and some $\sigma > 0$. Let P be the normal distribution with mean 0 and variance σ^2. In fact we will construct a σ-regular epistemic model where all types of player 1 have the same belief β_2 about player 2's choice and all types of player 2 have the same belief β_1 about player 1's choice. We construct β_1 and β_2 by means of the fixed point of some correspondence.

For every belief $\beta_j \in \Delta(C_j)$ and every utility function w_i, we define

$$C_i\left(\beta_j, w_i\right) := \{c_i \in C_i : w_i(c_i, \beta_j) \geq w_i(c_i', \beta_j) \text{ for all } c_i'\}.$$

We also define Q_i as the probability distribution on the set of utility functions of player i induced by P. For every $\beta_j \in \Delta(C_j)$ we define

$$\begin{aligned} F_i(\beta_j) \quad : \quad &= \{\beta_i \in \Delta(C_i) : \beta_i = \int_{w_i \in V_i} \gamma_i\left(w_i\right) dQ_i \ , \\ \text{where } \gamma_i\left(w_i\right) \quad \in \quad &\Delta(C_i\left(\beta_j, w_i\right)) \text{ for every } w_i \in V_i\}. \end{aligned}$$

Here V_i denotes the set of all possible utility functions for player i. So every $\beta_i \in F_i(\beta_j)$ is obtained by taking for every utility function w_i a randomization over optimal choices against β_j

and then taking the expected randomization with respect to Q_i. Now we define a correspondence F from $\Delta(C_1) \times \Delta(C_2)$ to $\Delta(C_1) \times \Delta(C_2)$ by

$$F(\beta_1, \beta_2) := F_1(\beta_2) \times F_2(\beta_1).$$

Now we use Kakutani's fixed point theorem to prove that F has a fixed point. Clearly F is upper hemi-continuous and compact valued. We show that F is convex valued. For this it is sufficient to show that F_1 and F_2 are convex valued. For a given β_2, take β_1', β_1'' in $F_1(\beta_2)$. We show that $\lambda\beta_1' + (1-\lambda)\beta_1''$ is also in $F_1(\beta_2)$. By definition

$$\beta_1' = \int_{w_1} \gamma_1'(w_1)\, dQ_1 \text{ and } \beta_1'' = \int_{w_1} \gamma_1''(w_1)\, dQ_1$$

where $\gamma_1'(w_1), \gamma_1''(w_1) \in \Delta(C_1(\beta_2, w_1))$ for every w_1. So we have

$$\lambda\beta_1' + (1-\lambda)\beta_1'' = \int_{w_1} (\lambda\gamma_1'(w_1) + (1-\lambda)\gamma_1''(w_1)) dQ_1$$

where $\lambda\gamma_1'(w_1) + (1-\lambda)\gamma_1''(w_1) \in \Delta(C_1(\beta_2, w_1))$ for every w_1. Hence by definition $\lambda\beta_1' + (1-\lambda)\beta_1'' \in F_1(\beta_2)$. This implies that F_1 is convex valued. The same applies to F_2 and hence we can conclude that F is convex valued. Now using Kakutani's fixed point theorem F has a fixed point $\left(\beta_1^*, \beta_2^*\right)$.

Since $\beta_1^* \in F_1\left(\beta_2^*\right)$ it follows that

$$\beta_1^* = \int_{w_1} \gamma_1^*(w_1)\, dQ_1$$

where $\gamma_1^*(w_1) \in \Delta(C_1(\beta_2^*, w_1))$ for every w_1. Similarly

$$\beta_2^* = \int_{w_2} \gamma_2^*(w_2)\, dQ_2$$

where $\gamma_2^*(w_2) \in \Delta(C_2(\beta_1^*, w_2))$ for every w_2.

We will now construct an epistemic model $M = (T_i, b_i, v_i)_{i \in I}$. For both players i, define

$$T_i = \{t_i^{w_i} \ : \ w_i \in V_i\}.$$

Let the utility assignment v_i be given by

$$v_i(t_i^{w_i}) = w_i$$

for every $t_i^{w_i} \in T_i$. In order to define the belief assignment b_i we first define for every type $t_i^{w_i}$ a density function $\tilde{b}_i(t_i^{w_i})$ on $C_j \times T_j$ as follows:

$$\tilde{b}_i(t_i^{w_i})\left(c_j, t_j^{w_j}\right) := \gamma_j^*(w_j)(c_j),$$

where $\gamma_j^*(w_j)(c_j)$ is the probability that probability distribution $\gamma_j^*(w_j)$ assigns to c_j. For every type $t_i^{w_i}$ let $b_i(t_i^{w_i}) \in \Delta(C_j \times T_j)$ be the probability distribution induced by density function $\tilde{b}_i(t_i^{w_i})\left(c_j, t_j^{w_j}\right)$ and the probability distribution Q_j on V_j. That is, for every set of types $E \subseteq T_j$ given by

$$E := \{t_j^{w_j} : w_j \in F\}$$

we have that

$$b_i(t_i^{w_i})(\{c_j\} \times E) := \int_{w_j \in F} \tilde{b}_i(t_i^{w_i})\left(c_j, t_j^{w_j}\right) dQ_j.$$

It follows that the belief of type $t_i^{w_i}$ about player j's choice is given by β_j^*. Namely, the probability that type $t_i^{w_i}$ assigns to choice c_j is equal to

$$\begin{aligned} b_i(t_i^{w_i})(\{c_j\} \times V_j) &= \int_{w_j \in V_j} \tilde{b}_i(t_i^{w_i})\left(c_j, t_j^{w_j}\right) dQ_j \\ &= \int_{w_j \in V_j} \gamma_j^*(w_j)(c_j)\, dQ_j \\ &= \beta_j^*(c_j). \end{aligned}$$

So all types of player i have the same belief β_j^* about player j's choice. This completes the construction of the epistemic model. It follows directly from the construction that the epistemic model is σ-regular.

We now show that every type in this model expresses common belief in rationality. For this it is sufficient to show that every type $t_i^{w_i}$ believes in the opponent's rationality. So, we must show for both players i and every $t_i^{w_i} \in T_i$ that $b_i(t_i^{w_i})[(C_j \times T_j)^{rat}] = 1$. In order to prove so we show that $\tilde{b}_i(t_i^{w_i})\left(c_j, t_j^{w_j}\right) > 0$ only if c_j is rational for $t_j^{w_j}$.

Suppose that $\tilde{b}_i(t_i^{w_i})\left(c_j, t_j^{w_j}\right) > 0$. Since $\tilde{b}_i(t_i^{w_i})\left(c_j, t_j^{w_j}\right) := \gamma_j^*(w_j)(c_j)$, it follows that $\gamma_j^*(w_j)(c_j) > 0$. As by definition $\gamma_j^*(w_j) \in \Delta(C_j(\beta_i^*, w_j))$ it follows that $c_j \in C_j(\beta_i^*, w_j)$. Remember that the belief of type $t_j^{w_j}$ about player i's choice is exactly β_i^*. Since $c_j \in C_j(\beta_i^*, w_j)$ it follows that c_j is rational for type $t_j^{w_j}$. So we have shown that $\tilde{b}_i(t_i^{w_i})\left(c_j, t_j^{w_j}\right) > 0$ only if c_j is rational for $t_j^{w_j}$. This implies that type $t_i^{w_i}$ believes in the opponent's rationality. Since this holds for every type in the model it follows that every type in the epistemic model expresses common belief in rationality. So every type in the model is σ-rationalizable because the model is σ-regular. This completes the proof. ■

6.2 Some Technical Lemmas

In this subsection we state some technical lemmas which we need for the proof of the main result.

Lemma 6.1 *If X, Y and Z are real valued, independent random variables then* $\Pr(X \geq \max\{Y, Z\}) \geq \Pr(X \geq Y) \cdot \Pr(X \geq Z)$.

Proof. Let f_Y and f_Z be the probability density functions of the random variables Y and Z. Now,

$$\begin{aligned}
& \Pr(X \geq \max\{Y, Z\}) \\
= & \int_y \int_z \Pr(X \geq \max\{y, z\})\, df_Y(y)\, df_Z(z) \\
\geq & \int_y \int_z \Pr(X \geq \max\{y, z\}) \cdot \Pr(X \geq \min\{y, z\}) df_Y(y)\, df_Z(z) \\
= & \int_y \int_z \Pr(X \geq y) \cdot \Pr(X \geq z)\, df_Y(y)\, df_Z(z) \\
= & \int_y \Pr(X \geq y)\, df_Y(y) \cdot \int_z \Pr(X \geq z)\, df_Z(z) \\
= & \Pr(X \geq Y) \cdot \Pr(X \geq Z).
\end{aligned}$$

Note that the first and third equality follow from the fact that Y and Z are independent, and the inequality holds because $\Pr(X \geq \min\{y, z\}) \leq 1$. ■

We now state the well-known Chebyshev's inequality, which we use in the proof of Lemma 6.3.

Lemma 6.2 (Chebyshev's inequality) *Let X be a random variable with $E(X) = \mu$. Then for any number $k > 0$,*

$$\Pr(|X - \mu| \geq k) \leq \frac{Var(X)}{k^2}.$$

Lemma 6.3 *For every $n \in \mathbb{N}$, let $X_n^1, X_n^2, ..., X_n^m$ be independent random variables with $E(X_n^i) = \mu^i$ for all n and i, $\mu^1 > \mu^2 > ... > \mu^m$, and $\lim_{n\to\infty} Var(X_n^i) = 0$ for all i. Then,*

$$\lim_{n\to\infty} \Pr\left(X_n^1 \geq X_n^2 \geq ... \geq X_n^m\right) = 1.$$

Proof. For a given n,

$$\Pr\left(X_n^1 \geq X_n^2 \geq ... \geq X_n^m\right) \geq 1 - \Pr\left(X_n^i < X_n^j \text{ for some } i < j\right).$$

For fixed $i < j$ we have,

$$\begin{aligned}
\Pr\left(X_n^i < X_n^j\right) &= \Pr\left(X_n^j - X_n^i > 0\right) = \Pr\left(\left(X_n^j - X_n^i\right) - \left(\mu^j - \mu^i\right) > \mu^i - \mu^j\right) \\
&\leq \Pr\left(\left|\left(X_n^j - X_n^i\right) - \left(\mu^j - \mu^i\right)\right| > \mu^i - \mu^j\right) \\
&\leq \frac{Var\left(X_n^j - X_n^i\right)}{\left(\mu^i - \mu^j\right)^2} \\
&= \frac{Var\left(X_n^j\right) + Var\left(X_n^i\right)}{\left(\mu^i - \mu^j\right)^2}.
\end{aligned}$$

Here, the inequality comes from Chebyshev's inequality and the last equality follows from the fact that X_n^j and X_n^i are independent. Now, note that $\lim_{n\to\infty} Var\left(X_n^i\right) = 0$ and $\lim_{n\to\infty} Var\left(X_n^j\right) = 0$, which implies $\lim_{n\to\infty} \Pr\left(X_n^i < X_n^j\right) = 0$. Then, from above it follows that

$$\lim_{n\to\infty} \Pr\left(X_n^1 \geq X_n^2 \geq ... \geq X_n^m\right) = 1. \blacksquare$$

Consider a sequence $(P_n)_{n\in\mathbb{N}}$ of normal distributions with mean 0 and variance σ_n^2 such that $\sigma_n \to 0$ as $n \to \infty$. The density function f_n of P_n is given by

$$f_n(x) = \frac{1}{\sigma_n\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma_n^2}} \text{ for all } x.$$

We show that for large n the right tail of P_n becomes arbitrarily steep everywhere.

Lemma 6.4 *Consider a sequence $(P_n)_{n\in\mathbb{N}}$ of normal distributions with mean 0 and variance σ_n^2, such that $\sigma_n \to 0$ as $n \to \infty$. Let f_n be the density functions of these distributions. Then for all $c > 0$ and $\epsilon > 0$ there is $N \in \mathbb{N}$ such that $\frac{f_n(x+c)}{f_n(x)} \leq \epsilon$ for all $n \geq N$ and all $x > 0$.*

Proof. Take $c > 0$ and $\epsilon > 0$. Then

$$\frac{f_n(x+c)}{f_n(x)} = \frac{e^{-\frac{(x+c)^2}{2\sigma_n^2}}}{e^{-\frac{x^2}{2\sigma_n^2}}} = e^{-\frac{1}{2\sigma_n^2}\left((x+c)^2 - x^2\right)} = e^{-\frac{1}{2\sigma_n^2}\left(2cx + c^2\right)} \leq e^{-\frac{c^2}{2\sigma_n^2}}.$$

Now as $c > 0$ is fixed and $\sigma_n \to 0$ as $n \to \infty$, we can find N large enough such that $e^{-\frac{c^2}{2\sigma_n^2}} \leq \epsilon$ for $n \geq N$. $\blacksquare$

Lemma 6.5 *Consider a sequence $(X_n)_{n\in\mathbb{N}}$ of normally distributed random variables such that $E(X_n) = 0$ for all n, and $var(X_n) \to 0$ as $n \to \infty$. Let f_n be the density functions of these random variables. Then, for every $0 < x < y$ it holds that*

$$\lim_{n\to\infty} \frac{\Pr(X_n \geq y)}{\Pr(X_n \geq x)} = 0.$$

Proof. Fix $0 < x < y$, and fix an $\epsilon > 0$. Then, by Lemma 6.4 there is an N such that $\frac{f_n(z+(y-x))}{f_n(z)} \leq \epsilon$ for all $n \geq N$ and all $z > 0$. Take some $n \geq N$. Then,

$$\begin{aligned} \Pr(X_n \geq y) &= \int_y^\infty f_n(z)\, dz = \int_x^\infty f_n(z + (y-x))\, dz \\ &\leq \epsilon \cdot \int_x^\infty f_n(z)\, dz = \epsilon \cdot \Pr(X_n \geq x). \end{aligned}$$

This implies that $\lim_{n\to\infty} \frac{\Pr(X_n \geq y)}{\Pr(X_n \geq x)} = 0$. ■

6.3 Proof of the Main Result

We finally prove or main theorem, which is Theorem 4.1. We proceed by three steps.

In step 1, we show how a σ-regular epistemic model M with incomplete information can be transformed into an epistemic model $\hat{M}$ with complete information. More precisely, we transform every type t_i in M into a type $\theta_i(t_i)$ in $\hat{M}$ which has the same belief about the opponent's choice as t_i.

In step 2, we take a choice c_i^* that is limit rationalizable. So we can find a sequence $(P_n)_{n\in\mathbb{N}}$ of normal distributions with mean 0 and variance σ_n^2, with $\sigma_n^2 \to 0$ as $n \to \infty$, and a sequence $(M^n)_{n\in\mathbb{N}}$ of σ_n-regular epistemic models with constant type spaces and utility assignments, such that in every M^n there is a σ_n-rationalizable type t_i^n with utility function u_i for which choice c_i^* is optimal. We show that the type t_i^n is transformed into a type $\theta_i(t_i^n)$ which is ϵ_n-properly rationalizable for some ϵ_n. Since, for all n, c_i^* is rational for t_i^n, and $\theta_i(t_i^n)$ has the same belief about the opponent's choice and the same utility function as t_i^n, it follows that c_i^* is rational for $\theta_i(t_i^n)$ for all n. As $\theta_i(t_i^n)$ is ϵ_n-properly rationalizable for every n, it follows that c_i^* is ϵ_n-properly rationalizable for all n.

In step 3, we prove that $\lim_{n\to\infty} \epsilon_n = 0$. Hence, c_i^* is ϵ-properly rationalizable for every $\epsilon > 0$ and therefore properly rationalizable.

Step 1. Take some $\sigma > 0$. Let $M = (T_i, b_i, v_i)_{i\in I}$ be a σ-regular epistemic model for Γ with incomplete information. Now we transform this epistemic model M into an epistemic model $\hat{M} = (\Theta_i, \beta_i)_{i\in I}$ with complete information. Using the fact that M is σ-regular we can write

$$T_i = V_i \times B_i,$$

where V_i is the set of all possible utility functions and B_i is the finite set of belief hierarchies in T_i. Then, for $t_i \in T_i$,

$$b_i(t_i) \in \triangle(C_j \times V_j \times B_j).$$

Now take $\Theta_i = B_i$ and $\Theta_j = B_j$. Clearly, Θ_i and Θ_j are finite sets as B_i and B_j are finite. For every $t_i \in T_i$ define the type $\theta_i(t_i) \in \Theta_i$ by

$$\beta_i(\theta_i(t_i)) := \text{marg}_{C_j \times B_j} b_i(t_i).$$

So,

$$\beta_i(\theta_i(t_i))(c_j, b_j) = b_i(t_i)(V_j \times \{(c_j, b_j)\})$$

for all $(c_j, b_j) \in C_j \times B_j$. Hence,

$$\beta_i(\theta_i(t_i)) \in \triangle(C_j \times B_j) = \triangle(C_j \times \Theta_j).$$

By construction $\theta_i(t_i)$ has the same belief about j's choice as t_i. This completes the construction of the epistemic model $\hat{M} = (\Theta_i, \beta_i)_{i \in I}$.

Step 2. Take a choice c_i^* that is limit rationalizable. Hence, there exists a sequence $(P_n)_{n \in \mathbb{N}}$ of normal distributions with mean 0 and variance σ_n^2, with $\sigma_n^2 \to 0$ as $n \to \infty$, and a sequence $(M^n)_{n \in \mathbb{N}}$ of σ_n-regular epistemic models with constant type spaces and utility assignments, such that in every M^n there is a σ_n-rationalizable type t_i^n with utility function u_i for which choice c_i^* is optimal. Let the constant type spaces in the sequence $(M^n)_{n \in \mathbb{N}}$ of epistemic models be T_i and T_j, and the constant utility assignments be v_i and v_j.

Fix an n. Then, within the epistemic model $M^n = (T_i, b_i^n, v_i)_{i \in I}$ there is a σ_n-rationalizable type $t_i^n \in T_i$ with utility function u_i for which c_i^* is optimal. Since type t_i^n only deems possible j's types which are σ_n-rationalizable, and only deems possible j's types which only deem possible i's types which are σ_n-rationalizable, and so on, we may assume without loss of generality that all the types in M^n are σ_n-rationalizable. Let $\hat{M}^n = (\Theta_i^n, \beta_i^n)_{i \in I}$ be the corresponding epistemic model with complete information, as constructed in step 1.

For every $\theta_i \in \Theta_i^n$, we define a number $\epsilon_n(\theta_i)$ as follows: Let Poss(θ_i) be the set of types in Θ_j that θ_i deems possible. For a given type $\theta_j \in$ Poss(θ_i), suppose that θ_j prefers choice c_j^1 to c_j^2, c_j^2 to c_j^3, and so on. So, we obtain an ordering $\left(c_j^1, c_j^2, c_j^3, ..., c_j^m\right)$ of j's choices. Then define

$$\epsilon_n(\theta_i, \theta_j) = \max_{k \in \{2,3,...,m\}} \frac{\beta_i^n(\theta_i)\left(c_j^k, \theta_j\right)}{\beta_i^n(\theta_i)\left(c_j^{k-1}, \theta_j\right)}.$$

Next we define

$$\epsilon_{i,n} = \max_{\theta_i \in \Theta_i^n, \theta_j \in \text{ Poss}(\theta_i)} \epsilon_n(\theta_i, \theta_j).$$

Finally let

$$\epsilon_n = \max\{\epsilon_{i,n},\ \epsilon_{j,n}\}.$$

Note that by construction every type in $\hat{M}^n$ satisfies the ϵ_n-proper trembling condition, hence every type in $\hat{M}^n$ is ϵ_n-properly rationalizable. In particular $\theta_i\left(t_i^n\right)$ is ϵ_n-properly rationalizable.

Step 3. Now we show that $\lim_{n\to\infty}\epsilon_n = 0$. It is sufficient to show that

$$\lim_{n\to\infty}\frac{\beta_i^n\left(\theta_i\right)\left(c_j^k,\theta_j\right)}{\beta_i^n\left(\theta_i\right)\left(c_j^{k-1},\theta_j\right)} = 0 \tag{1}$$

for every $\theta_i\in\Theta_i^n$, and every $\theta_j\in\mathrm{Poss}(\theta_i)$ and every k. As before, player j's choices are ordered $c_j^1,...,c_j^m$ such that θ_j prefers choice c_j^1 to c_j^2, c_j^2 to c_j^3, and so on. We assume, without loss of generality, that all preferences are strict.

Fix some $\theta_i\in\Theta_i^n$ and $\theta_j\in\mathrm{Poss}(\theta_i)$. Suppose that $\theta_i=\theta_i(t_i)$ for some $t_i\in T_i$, and that $\theta_j=\theta_j(t_j)$ for some $t_j\in T_j$. Let $\gamma_j\in\Delta(C_i)$ be θ_j's belief about i's choice. As before, let V_j be the set of utility functions for player j. For every $k\in\{1,...,m\}$, let $X^k:V_j\to\mathbb{R}$ be given by

$$X^k(v_j) := v_j(c_j^k,\gamma_j) = \sum_{c_i\in C_i}\gamma_j(c_i)\cdot v_j(c_j^k,c_i)$$

for every $v_j\in V_j$. So, $X^k(v_j)$ denotes the expected utility for player j induced by choice c_j^k, under the belief γ_j and the utility function v_j. Note that X^k is a random variable, as player i holds a probability distribution on V_j, induced by P_n. The probability distribution of X^k depends on n, and is denoted by $\varphi^{nk}(X^k)$. Note that X^k has a normal distribution with mean

$$E(X^k) = u_j(c_j^k,\gamma_j),$$

and variance

$$Var^n(X^k) = \sum_{c_i\in C_i}(\gamma_j(c_i))^2\cdot\sigma_n^2. \tag{2}$$

In particular, it follows that $\lim_{n\to\infty}Var^n(X^k)=0$, as $\lim_{n\to\infty}\sigma_n^2=0$. Since, by assumption, θ_j strictly prefers c_j^1 to c_j^2, strictly prefers c_j^2 to c_j^3, and so on, we have that $E(X^1)>E(X^2)>...>E(X^m)$.

Let φ^n be the probability distribution of the random vector $(X^1,...,X^m)$. Recall that all types in M^n are σ_n-rationalizable, which implies that all types in M^n express common belief in rationality. As such, type $t_i\in T_i$ (which generates θ_i) expresses common belief in rationality. In particular, t_i only assigns positive probability to those choice-type combinations (c_j,t_j) where

c_j is optimal for t_j. Now, as $\theta_i = \theta_i(t_i)$ and $\theta_j = \theta_j(t_j)$, we have that $\beta_i^n(\theta_i)\left(c_j^k, \theta_j\right)$ is the probability that c_j^k is optimal for t_j, and that is $\varphi^n(X^k \geq X^l \text{ for all } l)$. Then,

$$\frac{\beta_i^n(\theta_i)\left(c_j^k, \theta_j\right)}{\beta_i^n(\theta_i)\left(c_j^{k-1}, \theta_j\right)} = \frac{\varphi^n(X^k \geq X^l \text{ for all } l)}{\varphi^n(X^{k-1} \geq X^l \text{ for all } l)}. \tag{3}$$

Hence, in order to prove (1), we must show that

$$\lim_{n\to\infty} \frac{\varphi^n(X^k \geq X^l \text{ for all } l)}{\varphi^n(X^{k-1} \geq X^l \text{ for all } l)} = 0$$

for all $k \in \{2, ..., m\}$. We distinguish two cases.

Case 1. First we consider the case where $k = 2$. Then we have,

$$\frac{\varphi^n(X^k \geq X^l \text{ for all } l)}{\varphi^n(X^{k-1} \geq X^l \text{ for all } l)} \leq \frac{\varphi^n(X^2 \geq X^1)}{\varphi^n\left(X^1 \geq X^2 \geq X^3 \geq ... \geq X^m\right)}.$$

Recall that $E(X^1) > E(X^2) > ... > E(X^m)$. But then, by Lemma 6.3, $\varphi^n(X^2 \geq X^1) \to 0$, and $\varphi^n\left(X^1 \geq X^2 \geq X^3 \geq ... \geq X^m\right) \to 1$, and hence

$$\frac{\varphi^n(X^2 \geq X^1)}{\varphi^n\left(X^1 \geq X^2 \geq X^3 \geq ... \geq X^m\right)} \to 0,$$

which implies that

$$\frac{\varphi^n(X^k \geq X^l \text{ for all } l)}{\varphi^n(X^{k-1} \geq X^l \text{ for all } l)} \to 0$$

as $n \to \infty$.

Case 2. Now we consider the case where $k > 2$. Let $X^{\max}$ be the random variable given by $X^{\max} := \max_{j \neq k, k-1} X_j$. We have

$$\begin{aligned}
& \frac{\varphi^n(X^k \geq X^l \text{ for all } l)}{\varphi^n(X^{k-1} \geq X^l \text{ for all } l)} \\
= \;& \frac{\varphi^n(\left(X^k \geq X^{k-1}\right) \text{ and } \left(X^k \geq X^{\max}\right))}{\varphi^n((X^{k-1} \geq X^k) \text{ and } (X^{k-1} \geq X^{\max}))} \\
\leq \;& \frac{\varphi^n\left(X^k \geq X^{\max}\right)}{\varphi^n((X^{k-1} \geq X^k) \text{ and } (X^{k-1} \geq X^{\max}))} \\
\leq \;& \text{(by Lemma 6.1)} \frac{\varphi^n\left(X^k \geq X^{\max}\right)}{\varphi^n\left(X^{k-1} \geq X^k\right) \cdot \varphi^n\left(X^{k-1} \geq X^{\max}\right)} \\
= \;& \frac{\varphi^n\left(X^k \geq X^{\max}\right)}{\varphi^n\left(X^{k-1} \geq X^{\max}\right)} \cdot \frac{1}{\varphi^n\left(X^{k-1} \geq X^k\right)} \\
= \;& \frac{\varphi^n\left(X^k \geq X^{\max}\right)}{\varphi^n\left(X^k \geq X^{\max} - (E(X^{k-1}) - E(X^k)\right)} \cdot \frac{1}{\varphi^n\left(X^{k-1} \geq X^k\right)},
\end{aligned}$$

where the last equality follows from the observation that $X^{k-1} - E(X^{k-1})$ and $X^k - E(X^k)$ have the same distribution.

Now, from Lemma 6.3 it follows that $\varphi^n\left(X^{k-1} \geq X^k\right) \to 1$ as $n \to \infty$. We show that

$$\frac{\varphi^n\left(X^k \geq X^{\max}\right)}{\varphi^n\left(X^k \geq X^{\max} - (E(X^{k-1}) - E(X^k)\right)} \to 0$$

as $n \to \infty$.

Let us define $c := E(X^{k-1}) - E(X^k)$. So, we have to show that

$$\frac{\varphi^n\left(X^k \geq X^{\max}\right)}{\varphi^n\left(X^k \geq X^{\max} - c\right)} \to 0 \tag{4}$$

as $n \to \infty$. Note that $\varphi^n\left(X^k \geq X^{\max}\right) \leq \varphi^n\left(X^k \geq X^1\right)$. We first show that there exists $N \in \mathbb{N}$ such that for all $n \geq N$,

$$\varphi^n\left(X^k \geq X^{\max} - c\right) \geq \varphi^n\left(X^k \geq X^1 - c/2\right). \tag{5}$$

Now,

$$\begin{aligned}
& \varphi^n\left(X^k \geq X^{\max} - c\right) \\
= \; & \varphi^n\left(X^k \geq X^{\max} - c \mid X^{\max} = X^1\right) \cdot \varphi^n\left(X^{\max} = X^1\right) \\
& + \varphi^n\left(X^k \geq X^{\max} - c \mid X^{\max} \neq X^1\right) \cdot \varphi^n\left(X^{\max} \neq X^1\right) \\
\geq \; & \varphi^n\left(X^k \geq X^{\max} - c \mid X^{\max} = X^1\right) \cdot \varphi^n\left(X^{\max} = X^1\right) \\
= \; & \varphi^n\left(X^k \geq X^1 - c\right) \cdot \varphi^n\left(X^{\max} = X^1\right).
\end{aligned}$$

So, to show (5) it is sufficient to show that there exists $N \in \mathbb{N}$ such that for all $n \geq N$,

$$\varphi^n\left(X^k \geq X^1 - c\right) \cdot \varphi^n\left(X^{\max} = X^1\right) \geq \varphi^n\left(X^k \geq X^1 - c/2\right). \tag{6}$$

Using Lemma 6.3, $\varphi^n\left(X^{\max} = X^1\right) \to 1$ as $n \to \infty$. We have,

$$\begin{aligned}
& \frac{\varphi^n\left(X^k \geq X^1 - c/2\right)}{\varphi^n\left(X^k \geq X^1 - c\right)} \\
= \; & \frac{\varphi^n\left(\left(X^k - X^1\right) - \left(E\left(X^k\right) - E\left(X^1\right)\right) \geq -c/2 - \left(E\left(X^k\right) - E\left(X^1\right)\right)\right)}{\varphi^n\left(\left(X^k - X^1\right) - \left(E\left(X^k\right) - E\left(X^1\right)\right) \geq -c - \left(E\left(X^k\right) - E\left(X^1\right)\right)\right)}.
\end{aligned}$$

Note that $\varphi^n \left(\left(X^k - X^1 \right) - \left(E \left(X^k \right) - E \left(X^1 \right) \right) \right)$ has a normal distribution with mean 0, and where the variance of $\varphi^n \left(X^k - X^1 \right)$ tends to 0 as $n \to \infty$. Moreover, $-c - \left(E \left(X^k \right) - E \left(X^1 \right) \right) > 0$ as $E \left(X^k \right) - E \left(X^1 \right) < E \left(X^k \right) - E \left(X^{k-1} \right) = -c$. Hence, using Lemma 6.5,

$$\frac{\varphi^n \left(\left(X^k - X^1 \right) - \left(E \left(X^k \right) - E \left(X^1 \right) \right) \geq -c/2 - \left(E \left(X^k \right) - E \left(X^1 \right) \right) \right)}{\varphi^n \left(\left(X^k - X^1 \right) - \left(E \left(X^k \right) - E \left(X^1 \right) \right) \geq -c - \left(E \left(X^k \right) - E \left(X^1 \right) \right) \right)} \to 0$$

as $n \to \infty$. Then, we have,

$$\frac{\varphi^n \left(X^k \geq X^1 - c/2 \right)}{\varphi^n \left(X^k \geq X^1 - c \right)} \to 0.$$

So, there exists $N \in \mathbb{N}$ such that for all $n \geq N$,

$$\varphi^n \left(X^{\max} = X^1 \right) \geq \frac{\varphi^n \left(X^k \geq X^1 - c/2 \right)}{\varphi^n \left(X^k \geq X^1 - c \right)}.$$

This proves (6), which, as we have shown, implies (5).

Now, by (5) we have

$$\begin{aligned}
& \frac{\varphi^n \left(X^k \geq X^{\max} \right)}{\varphi^n \left(X^k \geq X^{\max} - c \right)} \\
\leq & \frac{\varphi^n \left(X^k \geq X^1 \right)}{\varphi^n \left(X^k \geq X^1 - c/2 \right)} \\
= & \frac{\varphi^n \left(\left(X^k - X^1 \right) - \left(E \left(X^k \right) - E \left(X^1 \right) \right) \geq - \left(E \left(X^k \right) - E \left(X^1 \right) \right) \right)}{\varphi^n \left(\left(X^k - X^1 \right) - \left(E \left(X^k \right) - E \left(X^1 \right) \right) \geq -c/2 - \left(E \left(X^k \right) - E \left(X^1 \right) \right) \right)} \\
= & \frac{\varphi^n \left(\left(X^k - X^1 \right) - \left(E \left(X^k \right) - E \left(X^1 \right) \right) \geq \left(E \left(X^1 \right) - E \left(X^k \right) \right) \right)}{\varphi^n \left(\left(X^k - X^1 \right) - \left(E \left(X^k \right) - E \left(X^1 \right) \right) \geq \left(E \left(X^1 \right) - E \left(X^k \right) \right) - c/2 \right)} \\
\to & \ 0
\end{aligned}$$

as n goes to infinity. Here the convergence follows from Lemma 6.5 as $\left(E \left(X^1 \right) - E \left(X^k \right) \right) - c/2 > 0$. So, we have shown (4), which completes case 2. Hence, we have shown that (1) holds for all k. Therefore, $\lim_{n \to \infty} \epsilon_n = 0$ and hence the proof is complete. ■

References

[1] Asheim, G.B. (2001), Proper rationalizability in lexicographic beliefs. Int. J. Game Theory 30, 453–478.

[2] Dekel, E., Fudenberg, D. (1990), Rational behavior with payoff uncertainty. J. Econ. Theory 52, 243–267.

[3] Myerson, R. (1978), Refinement of the Nash equilibrium concept. Int. J. Game Theory 7, 73–80.

[4] Schuhmacher, F. (1999), Proper rationalizability and backward induction. Int. J. Game Theory 28, 599–615.

[5] Tan, T. and S.R.C. Werlang (1988), The bayesian foundations of solution concepts of games, J. Econ. Theory 45, 370-391.