Bayesian coalitional rationality in games*

Xiao Luo[a]†, Yongchuan Qiao[a], Chih-Chun Yang[b]

[a] *Department of Economics, National University of Singapore, Singapore 117570*

[b] *Institute of Economics, Academia Sinica, Taipei 115, Taiwan*

August 2019

Abstract

The rational behavior for a coalition of players in a noncooperative situation requires recourse to an epistemic precondition that each coalition member believes about the coalition members' behavior and beliefs. In the context of games where players are with the Bayesian view of the world, we define the notion of "Bayesian coalitional rationality" (i.e., Bayesian c-rationality) as a distinctive mode of collective behavior that no coalition of players wishes to change, in light of coalition-wise common belief that mutually beneficial gains exist by a tacit coalitional agreement. To demonstrate its usefulness, we provide an epistemic analysis of the solution concepts such as (strong) Nash equilibrium and Bayesian coalitional rationalizability within a framework where players are allowed to hold introspective beliefs about their own behavior and mental states. *JEL Classification: C70, C72, D81.*

Keywords: Coalition-wise common belief, Bayesian coalitional rationality, Bayesian coalitional rationalizability, strong Nash equilibrium, conditional probability system

---

*This paper is based on part of an earlier manuscript "Bayesian Coalitional Rationalizability." We thank Geir Asheim, Adam Brandenburger, Yi-Chun Chen, Yossi Greenberg, Takashi Kunimoto, Shravan Luckraz, Andrés Perea, Xuewen Qian, Chen Qu, Yang Sun, Satoru Takahashi, Licun Xue, Junjian Yi, Shmuel Zamir, Shenghao Zhu, and seminar participants at National University of Singapore and Academia Sinica. The earlier version of the paper was presented at International Conference on Game Theory, SAET Conference on Current Trends in Economics, and China Meeting of the Econometric Society. Financial support from the National University of Singapore and the National Science Council of Taiwan is gratefully acknowledged. The usual disclaimer applies.

†Corresponding author. Fax:+65 6775 2646. E-mail: ecslx@nus.edu.sg (X. Luo); qiaoyongchuan@gmail.com (Y. Qiao); cyang16@econ.sinica.edu.tw (C.C. Yang).

1 Introduction

The conventional concept of rationality used in economics is linked to the optimal mode of behavior –that is, people act in their own best interests, given their information. From the Bayesian viewpoint, subjective probabilities should be assigned to every prospect (e.g., the prospect of a player choosing a certain strategy in a certain game). Based on Harsanyi's [39] seminal work on games with incomplete information, players in games can resolve the exhausted uncertainty that exists in games in the same way as nonstrategic uncertainty being resolved in standard decision theory; thus, each player in a game has a subjective probabilistic belief about the exhausted uncertainty facing the player. In the context of a game, the classic notion of "Bayesian rationality" requires a player to choose an action that maximizes his expected payoff, under his probabilistic belief about the players' epistemic types and the actual play of the game (see, e.g., Aumann [10] and Dekel and Siniscalchi [29]). To put it in a slightly different way,

> *a player is not Bayesian rational* if, on self-reflection and introspection, the player knows/believes that he can choose a replacement of his action to attain a higher expected payoff.

That is, a rational player should be aware of the consequences by choosing differently, and decide not to make a different choice of strategy. This definition of Bayesian rationality manifests the epistemic precondition for the player's rational behavior; the definition is harmonious with the standard one under the usual assumption that the player knows his own belief and using strategy, although it calls for self-awareness and introspection on purpose (see Proposition 1(b)).[1]

A salient feature of the notion of Bayesian rationality is that an individual player makes his own optimal decision, in the absence of coalitional considerations. To see this point, consider the coordination game (where the first player picks a row and the second player picks a column):

	a	b
a	$1, 1$	$0, 0$
b	$0, 0$	$2, 2$

[1]In an influential paper "Backward Induction and Common Knowledge of Rationality," Aumann [11] adopted a definition of rationality on the basis of self-knowledge: a player is "rational" if it is not the case that he knows that he would be able to do better; see also Samet's [71, Section 4] belief-based definition of doxastic rationality and Heller's [43, Definition 3] knowledge-based definition of "profitable joint deviation." In the philosophy of mind and epistemology, no rational agent could be incapable of self-knowledge (see, e.g., Shoemaker [74], Korsgaard [49] and Gilboa [34] for more discussions).

Intuitively, a Bayesian rational player can choose to play the strategy a if he believes that the opponent player plays the strategy a; alternatively, it is not the case that this player, on self-reflection and introspection, believes he can attain a higher expected payoff by unilaterally changing his using strategy a. In this noncooperative game, the notion of Bayesian rationality does not rule out playing the strategy a; indeed, both players jointly playing the strategy a constitutes a Nash equilibrium. However, it is clear that the players should coordinate to play the strategy b, but avoid playing the strategy a, from the coalition's viewpoint.

The coalitional behavior has been recognized as an important issue in economics and social sciences; real-life examples include cartels, trade blocs, political party formation, special interest groups, public goods provision, social networks and matchings, and so on. It is important and fundamental to explore what is a "rational" behavior for a coalition in noncooperative games, e.g., the coordination game.[2] In this paper, we offer an epistemological definition of "Bayesian coalitional rationality" (henceforth, Bayesian c-rationality), in the same way as the individual version of Bayesian rationality. Just like rational individuals, groups should be rational. We define Bayesian c-rationality as a mode of behavior that no group of players wishes to change –that is, every coalition is "rational" in a game. More specifically,

> *a coalition is not Bayesian rational* if members of the coalition commonly believe that they can change their actions, through an (implicit) coalitional agreement, to attain mutually beneficial expected payoffs.

The notion defines, from the inside, a kind of the rational behavior that is compatible with the access to information possessed by coalitions in a society.[3] Interactive beliefs among a coalition are crucial for the concept of coalitional rationality. To facilitate their behavior, the coalition members typically need to have common understandings of mutual behavior

[2] Liu et al. [52] made use of the idea of "coalitional rationality" to define a stable matching with incomplete information, which requires to attain common knowledge among a blocking pair that no profitable pairwise deviation exists. Their analysis, however, refrains from considering deviations by larger groups of agents. Kobayashi [47] made use of the similar idea to study equilibrium contracts for syndicates with differential information, by assuming "an agent declares his intention to join a blocking coalition only when it is common knowledge for the coalition members that he would intend to join." See also Acemoglu et al. [1], Ambrus and Argenziano [5], and Jullien [46] for more applications.

[3] Human beings have a distinct form of self-consciousness that enables us to be aware of the motivations or potential reasons on which we are tempted to act, to evaluate those potential reasons, and to be moved to act accordingly. According to a traditional philosophical view, the concept of rationality is naturally tied to a form of introspection, one that makes us aware of, and capable of evaluating, our reasons themselves. The formalism is also related to Simon's [75, p. 278] jargon of "subjective" rationality (i.e., a mode of behavior that is rational, given the perceptual and evaluational premises of the subject). The rational behavior for a coalition in the context of complex interactions is innately subjective, and is unlikely to be "objective" (as viewed by the experimenter).

and interpersonal reasoning (see, e.g., Schelling [73], Lewis [51] and Chwe [28]). To illustrate this point, again consider the coordination game. In this noncooperative game, the players are coalitionally "irrational" to play the strategy a, because it is a "common belief" among the two players that, as a coalition, they can attain mutually beneficial expected payoffs by changing the *status quo* choice of the strategy a.[4] The key idea here is based on a coalitional reasoning argument that players can look for a "coalitional agreement" in the sense of Hume [44], by which the coalition members confine play to a subset of their strategies, to avoid certain "inferior" strategies in noncooperative games.[5] This is different from an "equilibrium" argument, because playing a is a Nash equilibrium behavior given that each player is Bayesian rational and has a correct belief about the opponent's strategy choice (see Aumann and Brandenburger [13]). The main purpose of this paper is to offer an epistemological definition of coalitional rationality in the context of a game where players are with the Bayesian view of the world, being aware of the effects of coalitional reasoning.

From an epistemic perspective, Bayesian rationality for an individual player is referred to a state of affairs: it is not the case that the player is attentively aware that an alternative choice of strategy can attain a higher expected payoff. The coalitional version of Bayesian rationality captures the similar idea for coalitional reasoning: it is not the case that coalition members share a common belief that they could improve coalition members' expected payoffs by coordinating their moves; that is, the coalition does not wish to change its behavior, in light of coalition-wise common belief that mutually beneficial gains exist by an (implicit) coalitional agreement.[6] The notion of coalitional rationality needs to call for the coalition-wise common-belief presumption, because, with the Bayesian view of the world, each coalition member's behavior and beliefs will be typically affected if an agreement is made by a coalition. In fact, actions undertaken deliberately by a group of players are related to some particular epistemic state, which often interferes with the other states under consideration of players in the group. That is, each player in the group has to think about what would happen in

[4] An event is common belief/knowledge among a group of players, if it is believed/known to all players in the group that it is believed/known to all players in the group, and so on *ad infinitum* (see Aumann [9, 12]). As usual, we use the term "knowledge" to mean true belief with no possibility for error, and use the shorthand term "common knowledge of rationality" to stand for "rationality and common belief of rationality."

[5] As Hume [44] put it in *A Treatise of Human Nature*, "... this may properly enough be called a convention or agreement betwixt us, though without the interposition of a promise ... Two men, who pull the oars of a boat, do it by an agreement or convention, though they have never given promises to each other."

[6] In the spirit of Bernheim et al.'s [20] notion of "coalition-proof Nash equilibrium," we impose a mild "credibility" condition that requires any meaningful change of strategies in a coalitional agreement be justified at least by the principle of individual Bayesian rationality. Technically, the requirement purports to overcome the notorious problem of emptiness, typified by Condorcet's paradox, under the core-like locking arrangements; e.g., in the Prisoner's Dilemma game, the noncooperative Nash strategy can be sustained by our definition of Bayesian c-rationality, because the agreement on the Pareto-optimal cooperative strategy profile is not "credible."

hypothetical circumstances, in one of which every player takes into account what the group of players would do and would contemplate in the hypothetical circumstances.[7] Thus, it is necessary to consider what a coalition member believes about not only the contemplation of joint movements, but also what the other coalition members believe. One important ingredient of our approach is how to formally express that it is a common belief among members of a coalition that mutually beneficial gains exist by a coalitional agreement.[8] To deal with the hypothetical coalitional reasoning in a Bayesian paradigm, we carry out our analysis in an epistemic framework in which each player has a "conditional probability system" (CPS) belief at a type (Definition 1). A CPS belief of a player specifies the family of probabilistic beliefs about the players' strategies and types in all contingencies; even in the case of an unexpected or hypothetical event, a player must have a probabilistic assessment of opponents' strategies contingent on that event (e.g., the coalition member must have a probabilistic belief about the opponents' behavior contingent on a coalitional agreement).

Our work is closely related to Ambrus [4], and provides complementary support to his approach to coalitional rationality. Ambrus [4] offered a wide range of epistemic definitions of coalitional γ-rationality, and obtained the concept of coalitional γ-rationalizability by "rationality and common certainty of coalitional γ-rationality" (where γ is a coalitional best response operator). Despite its great advantages in terms of generality and simplicity, his formalism is less comprehensive from an epistemic perspective, because it does not provide a full epistemic expression of coalitional rationality. In other words, the notion of γ-rationality, although defines the scope of "conceivable" γ-rational strategic behavior, does not explicate what the γ-rational epistemic state of affairs is. The major issue is that the shortcut operator γ is not based on assumptions about players' behavior and beliefs in the possible worlds associated with coalitions; for instance, the conjectures used in the "sensible" operator γ are not linked to epistemic states. Nevertheless, the missing part of the expression is critical in the epistemic analysis. Examples such as backward induction and iterated elimination of weakly dominated strategies show the hidden assumptions in the causal arguments about players' beliefs and rationality at epistemic states are crucial (see, e.g., Brandenburger [21],

[7] To coordinate its interrelated actions, a group of players needs to have recourse to "common knowledge" (see, e.g., Schelling [73], Lewis [51], Chwe [28], and Rubinstein [70]). Each of group members is willing to do so only if he knows that other group members are willing to do so. Members need to have knowledge of each other, knowledge of that knowledge, knowledge of the knowledge of that knowledge, and so on. See also Halpern and Moses [40] and Fagin et al. [31, Chapter 6] for extensive discussions on reasoning about the states of knowledge of the components of a distributed system. In particular, they showed common knowledge plays a critical role in reaching an agreement and coordinating actions in a distributed environment.

[8] In the classic article "The Use of Knowledge in Society," Hayek (1945, p.530) pointed out that "we must show how a solution is produced by the interactions of people each of whom possesses only partial knowledge. To assume all the knowledge to be given to a single mind in the same manner in which we assume it to be given to us as the explaining economists is to assume the problem away and to disregard everything that it is important and significant in the real world."

Dekel and Siniscalchi [29], and Perea [64] for extensive discussions). Moreover, there could be the "self-referential" problem in Ambrus's [4] framework, if the operator γ is viewed as part of the description of a state of the world. One particular troubling issue is how the states can be used to capture knowledge/belief about the model itself (see, e.g., Aumann [9, 10, 12] and Brandenburger and Dekel [23]), such as the expression that the operator γ is common knowledge/belief among a coalition.

To overcome the aforementioned shortcomings in coalitional γ-rationality, we offer a more delicate definition of coalitional rationality in a type-structure framework, which explicates how a player interacts with other players within a coalition if a coalitional agreement is made in game situations (see Definition 2). Our analytical framework is immunized from the "self-referential" problem, because the framework accommodates what the players believe about the information, knowledge and the behavior of coalition members at states of the world.

In the special case of singleton coalitions, our definition of Bayesian c-rationality is harmonious with the definition of Bayesian rationality in terms of the behavioral implication (Proposition 1(a)) and, if the player knows his own "primary" belief and using strategy, our definition of Bayesian c-rationality is equivalent to the definition of Bayesian rationality (Proposition 1(b)). The definition of Bayesian c-rationality for the case of singleton coalitions gives rise to a novel notion of individual rationality under the behavior and type uncertainty (Definition $2'$).

For the individual case, common knowledge of Bayesian rationality is closely related to Bernheim [19] and Pearce [63] notion of "rationalizability": the notion of (correlated) rationalizability can be characterized by "common knowledge of Bayesian rationality" –that is, "Bayesian rationality and common belief of Bayesian rationality" (cf. Tan and Werlang [77] and Brandenburger and Dekel [22]). For coalitional rationality, we obtain a similar result: common knowledge of Bayesian c-rationality provides an epistemic characterization for the solution concept of "Bayesian coalitional rationalizability" in Luo and Yang [53] (Proposition 4). We also discuss the solution concepts of Nash equilibrium and strong Nash equilibrium in a fairly flexible framework where players are allowed to have introspective beliefs about their own strategies and types (Propositions 2 and 3).

It is worth noting that, according to the notion of Bayesian c-rationality, we need to check $2^n - 1$ feasible coalitions (i.e., all subsets of n players except the empty set) and, for a coalition J, $\Pi_{j \in J} \left(2^{|S_j|} - 1\right)$ possible coalitional deviations (i.e., all restrictions of strategies in S_j for each coalition member $j \in J$ except the empty set). Due to the enormous complexity of coalitional reasoning, one may wonder whether a Bayesian c-rational state exists or not; one might be curious about how we express epistemic states about collective beliefs and intention for "rational" coalitions in complex interactions, given that coalitions do not have minds or brains. Proposition 4(b) asserts that, for any finite game, we can find the desirable state(s) by constructing a finite CPS type structure. Our study contributes to a better understanding of the coalition's behavior in game situations, by illuminating the informational preconditions how interactive beliefs are distributed among coalition members.

The rest of this paper is organized as follows. Section 2 provides an example to illustrate the main idea and results in this paper. Section 3 introduces the preliminary notation and definitions. Section 4 defines the concept of Bayesian c-rationality. Section 5 offers an epistemic analysis of Nash equilibrium and strong Nash equilibrium in our analytical framework. Section 6 studies the behavioral implication of common knowledge of Bayesian c-rationality. Section 7 is concluding remarks. To facilitate reading, all the proofs are relegated to the Appendix.

2 An illustrative example

Example 1. Consider a two-person symmetric game (where the first player picks a row and the second player picks a column):

	a_2	b_2	c_2
a_1	2, 2	3, 2	0, 0
b_1	2, 3	3, 3	0, 0
c_1	0, 0	0, 0	1, 1

.

Intuitively, confining the players' play to a subset of strategies $\{a_1, b_1\} \times \{a_2, b_2\}$ is in their mutual interest. The solution concepts of "(Bayesian) coalitional rationalizability" defined in Ambrus [3] and Luo and Yang [53] yield the same outcome set of (Bayesian) coalitional rationalizable strategy profiles: $\{a_1, b_1\} \times \{a_2, b_2\}$.

Next, we explain the concept of "Bayesian coalitional rationality" in an epistemic framework. For simplicity, we consider a type-structure model for this game, which specifies a set of "types" for each player, and for each type, a (probabilistic) belief over the opponents' strategies and types. That is, each player $i = 1, 2$ has two types t_i and t_i' (i.e., $T_i = \{t_i, t_i'\}$); for $i, j = 1, 2$ and $i \neq j$, player i's belief at t_i is $\beta_i(t_i) = 1 \circ (a_j; t_j)$ and player i's belief at t_i' is $\beta_i(t_i') = 1 \circ \left(b_j; t_j'\right)$. A state –a specification of the players' strategies and the players' types– is said to be *Bayesian c-rational* if no group of players commonly believes that there exists a mutually beneficial agreement that excludes the group's using strategy profile at the state. For example, the state $(b_1, b_2; t_1', t_2')$ is Bayesian c-rational, because at this state both players commonly believe that they are playing (b_1, b_2) and each receives the highest expected payoff of 3. However, the state $(a_1, a_2; t_1, t_2)$ is not Bayesian c-rational, because at this state the two players commonly believe that they would be willing to confine their play to the mutually beneficial agreement $\{b_1\} \times \{b_2\}$, instead of playing (a_1, a_2). Moreover, at state $(b_1, b_2; t_1', t_2')$, the players commonly believe in $(b_1, b_2; t_1', t_2')$, and hence Bayesian c-rationality is commonly known at $(b_1, b_2; t'_1, t'_2)$. In Proposition 4(a), we show a player plays a Bayesian c-rationalizable strategy in Luo and Yang [53] under common knowledge of Bayesian c-rationality in a "belief-rich" type structure (cf. Example 4 in Section 6.2).

The Bayesian c-rationalizable strategy profile (a_1, a_2) cannot be attained under common knowledge of Bayesian c-rationality in the above type-structure model. We can, however, find another type-structure model such that common knowledge of Bayesian c-rationality yields all the Bayesian c-rationalizable strategy profiles, including (a_1, a_2). For example, consider a type structure: $T_i = \{t_i, t'_i\}$, $\beta_1(t_1) = 1 \circ (a_2; t_2)$, $\beta_1(t'_1) = 1 \circ (b_2; t'_2)$, $\beta_2(t_2) = 1 \circ (a_1; t'_1)$ and $\beta_2(t'_2) = 1 \circ (b_1; t_1)$. In this type structure, the states $(a_1, a_2; t'_1, t_2)$, $(a_1, b_2; t'_1, t'_2)$, $(b_1, a_2; t_1, t_2)$ and $(b_1, b_2; t_1, t'_2)$ are Bayesian c-rational. (For instance, the set $\{(a_1, a_2; t'_1, t_2), (a_1, b_2; t'_1, t'_2), (b_1, a_2; t_1, t_2), (b_1, b_2; t_1, t'_2)\}$ is commonly believed at the state $(a_1, a_2; t'_1, t_2)$. For each player $i = 1, 2$, there exists a state in this set such that player i plays a_i and gets the highest expected payoff of 3. Therefore, the two players commonly believe there is no mutually beneficial agreement excluding the profile (a_1, a_2) at state $(a_1, a_2; t'_1, t_2)$.) Thus, (a_1, a_2), (a_1, b_2), (b_1, a_2) and (b_1, b_2) are attained under common knowledge of Bayesian c-rationality. In Proposition 4(b), we show the set of Bayesian c-rationalizable strategy profiles can be attained by common knowledge of Bayesian c-rationality in a (finite) type-structure model.

3 Preliminaries

Consider a finite game:

$$\mathcal{G} = (I, \{S_i\}_{i \in I}, \{u_i\}_{i \in I}),$$

where I is a finite set of players, S_i is a finite set of i's pure strategies (as usual, $S = \times_{i \in I} S_i$), and $u_i : S \to \mathbb{R}$ is i's payoff function. We say J is a coalition if J is a nonempty subset of I. For each coalition J, let $S_J = \times_{j \in J} S_j$ and $S_{-J} = \times_{i \in I \setminus J} S_i$. For any $s_i \in S_i$ and any probability distribution μ over $S_{-i} = \times_{j \in I \setminus \{i\}} S_j$, player i's expected payoff function is defined by $U_i(s_i, \mu) = \sum_{s_{-i} \in S_{-i}} \mu(s_{-i}) u_i(s_i, s_{-i})$.

3.1 Type structure with conditional probability systems

Let X be a topological space endowed with the Borel sigma-algebra Σ_X. The set of Borel probability measures on X with the weak convergence topology is denoted by $\Delta(X)$. Cartesian product sets are endowed with the product topology and the product sigma-algebra.

Definition 1. A *conditional-probability-systems type structure (or simply a type structure)* for the game $\mathcal{G} = (I, (S_i, u_i)_{i \in I})$ is a tuple $\mathcal{T}(\mathcal{G}) = \left(I, (S_i, u_i)_{i \in I}, (T_i, \beta_i)_{i \in I}\right)$, where each T_i

is a (nonempty) compact metric space and each $\beta_i : T_i \to \Delta^* (S \times T)$ is continuous.[9]

That is, a type structure specifies, for each player $i \in I$ and each type $t_i \in T_i$, a "conditional probability system (CPS)" on the strategies that the players can choose and the types that the players are endowed with –that is, through β_i, each type t_i is associated with a CPS belief concerning states of the world: players' strategies and types.

A type t_i for player i induces a CPS, denoted by $f_i(t_i)$, on the opponent players' strategies (that is, for each nonempty subset $B \subseteq S$, $f_i(t_i)|_{B_{-i}} \in \Delta(S_{-i})$ is the marginal distribution of $\beta_i(t_i)|_{B\times T} \in \Delta(S \times T)$, denoted by $f_i(t_i)|_{B_{-i}} = \mathrm{marg}_{S_{-i}}\beta_i(t_i)|_{B\times T}$). Because a type's primary probabilistic beliefs on the opponents' strategies and on states play a particularly important role, it is convenient to introduce the specific notation for them: $f_i^0(t_i) = f_i(t_i)|_{S_{-i}}$ and $\beta_i^0(t_i) = \beta_i(t_i)|_{S\times T}$.

Remark 1. The analytical framework of type structure is by now commonly used in the epistemic game theory. The notion of types is due to Harsanyi [39]. When types are associated with ordinary probability measure, Mertens and Zamir [56], Brandenburger and Dekel [23], and among others constructed a "universal" type structure that embeds all type structures under the assumption that the set of states of nature is a compact topological space. Battigalli and Siniscalchi [15] constructed the CPS counterpart of "universal" type space. Like Aumann [10], Aumann and Brandenburger [13] and Brandenburger and Dekel [22], our analytical framework is fairly flexible and allows players to have beliefs about their own strategies and types (see, e.g., Dekel and Siniscalchi [29, Section 12.2.6.3]). In particular, we do not impose the assumption that each player knows his own type and using strategy (cf., e.g., psychological games (Geanakoplos et al. [33]), the absent-minded driver (Piccione and Rubinstein [65]), and the Kahaneman-Tversky man (Lambert-Mogiliansky et al. [50])).[10]

In the type structure $\mathcal{T}(\mathcal{G})$, $S \times T$ is the *state space*. A strategy-type pair $(s, t) \in S \times T$ is called a state (of the world). An *event* E is a measurable subset of $S \times T$. Player i *believes*

[9] The notation $\Delta^*(S \times T)$ denotes the class of "conditional probability systems" on $S \times T$ (with conditioning the collection of nonempty subsets of finite strategy profles), in which a *conditional probability system (CPS)* in $\Delta^*(S \times T)$ is defined as a conditional-probability function $\mu|$ on $S \times T$ that satisfies Bayes' rule, i.e., if $\mu| \in \Delta^*(S \times T)$, then for every nonempty subset $B \subseteq S$ (i) $\mu|_{B\times T} \in \Delta(S \times T)$ and $\mu|_{B\times T}(B \times T) = 1$; (ii) for every set $A \supseteq B$ and $E \in \Sigma_{S\times T}$ satisfying $E \subseteq B \times T$, $\mu|_{A\times T}(E) = \mu|_{B\times T}(E)\mu|_{A\times T}(B \times T)$. See, e.g., Renyi [69] and Myerson [57, 58].

[10] This kind of introspective beliefs are also related to the concepts of self (e.g., Jungian archetypes of the human psyche) and self-knowledge (e.g., the ancient Greek maxim "know thyself" and the ancient Chinese philosopher Lao Tzu's saying "Knowing others is wisdom, knowing yourself is Enlightenment"). Applications abound in the economics literature, e.g., identity (Akerlof and Kranton [2]), self-confidence (Benabou and Tirole [17, 18]), self-control (Gul and Pesendorfer [38]), multi-self (Fudenberg and Levine [32]), rational addiction (Becker and Murphy [16]), and nudge (Thaler and Sunstein [78]).

E *at state* (s,t) if $\beta_i^0(t_i)(E) = 1$, that is, i attributes probability 1 (in terms of the primary probabilistic belief) to the event E. Let

$$B_i(E) = \{(s,t) \in S \times T : i \text{ believes } E \text{ at } (s,t)\}.$$

Let

$$B_J(E) = \bigcap_{j \in J} B_j(E)$$

denote the event "E is mutually believed among coalition J," and let

$$CB_J(E) = B_J(E) \cap B_J(B_J(E)) \cap \cdots$$

denote the event "E is commonly believed among coalition J." (For the grand coalition I, we use a shorthand notation: $CBE = CB_I(E)$.) For a type profile $t_J \in T_J = \times_{j \in J} T_j$, let $\boxed{t_J} \subseteq S \times T$ denote the finest event commonly believed by t_J; that is, the set $\boxed{t_J}$ is the "minimal state subspace" which is belief-closed by all the members in coalition J at the state of mind t_J.[11] It is well known that $CB_J(E) = \left\{(s,t) \in S \times T : \boxed{t_J} \subseteq E\right\}$ (see, e.g., Zamir and Vassilakis [82]). For the special case of singleton coalition $J = \{j\}$, we also write $\boxed{t_j}$ for the minimal state subspace $\boxed{t_J}$ at t_j. Obviously, if player j knows his primary belief, then j knows event E if and only if j "introspectively" knows E; that is, for $t_j \in T_j$, if $\beta_j^0(t_j)\left(\left\{(s',t') \in S \times T : \beta_j^0\left(t'_j\right) = \beta_j^0(t_j)\right\}\right) = 1$, then $\beta_j^0(t_j)(E) = 1$ iff $\boxed{t_j} \subseteq E$.

A player is said to be Bayesian rational if he chooses an action that maximizes his expected payoff given his information (see, e.g., Aumann [10] and Dekel and Siniscalchi [29]). More specifically, *player* i *is Bayesian rational at state* (s,t) if s_i is a best response to $f_i^0(t_i)$ –that is, $s_i \in \mathbf{BR}_i(t_i)$ where

$$\mathbf{BR}_i(t_i) = \left\{\widehat{s}_i \in S_i : U_i(\widehat{s}_i, f_i^0(t_i)) \geq U_i(s'_i, f_i^0(t_i)) \ \forall s'_i \in S_i\right\}.$$

Let

$$R_i = \{(s,t) \in S \times T : i \text{ is Bayesian rational at } (s,t)\}$$

denote the event "player i is Bayesian rational."

For event E and player i, let $\mathbf{s}_i(E) = \{s_i \in S_i : (s,t) \in E\}$ and $\mathbf{t}_i(E) = \{t_i \in T_i : (s,t) \in E\}$. For coalition J, define $\mathbf{s}_J(E) = \times_{j \in J} \mathbf{s}_j(E)$ and $\mathbf{s}_{-J}(E) = \times_{i \in I \setminus J} \mathbf{s}_i(E)$. For brevity, we write $\mathbf{s}(E) = \mathbf{s}_I(E)$.

---

[11] That is, the subset $\boxed{t_J} \subseteq S \times T$ is the smallest state subspace (with respect to set inclusion) such that $\beta_j^0(t_j)\left(\boxed{t_J}\right) = 1$, $\forall j \in J$, $\forall (s,t) \in \boxed{t_J}$ (cf., e.g., Zamir [81] and Maschler, Solan and Zamir [55, Chapter 10.4] for extensive discussions). The minimal state subspace $\boxed{t_J}$ is an analog notion of the "finest common coarsening" of partitional information structures (in Aumann [9]) for the coalition J.

4 Bayesian coalitional rationality: a definition

We formulate a notion of "Bayesian c-rationality" to prescribe, in the context of a game where players are with the Bayesian view of the world, a kind of rational behavior for coalitions that is compatible with the access to information. The notion of Bayesian c-rationality is defined as a mode of behavior that no group of players wishes to change: it is impossible for members of a coalition to have a common belief that they could attain higher expected payoffs by confining to a set of strategies, in a "credible" way.

Definition 2. In the type structure $\mathcal{T}(\mathcal{G})$, *coalition J is Bayesian rational at state* (s,t) if $s_J \in A_J$ whenever the nonempty product subset $A_J \subseteq S_J$ satisfies the following two conditions:

(C1) [κ-**Profitability**] for all $j \in J$ and $(s',t') \in \boxed{t_J}$, if $s'_j \notin A_j$ or $f_j\left(t'_j\right)|_{A_{J\setminus\{j\}}\times \mathbf{s}_{-J}\left(\boxed{t_J}\right)} \neq f_j\left(t'_j\right)|_{\mathbf{s}_{-j}\left(\boxed{t_J}\right)}$,[12] then there exists $\widehat{s}_j \in S_j$ such that

$$U_j\left(s'_j, f_j\left(t'_j\right)|_{\mathbf{s}_{-j}\left(\boxed{t_J}\right)}\right) < U_j\left(\widehat{s}_j, f_j\left(t'_j\right)|_{A_{J\setminus\{j\}}\times \mathbf{s}_{-J}\left(\boxed{t_J}\right)}\right);$$

(C2) [κ-**Credibility**] for all $j \in J$ and $a_j \in A_j \setminus \mathbf{s}_j\left(\boxed{t_J}\right)$, there exists $t'_j \in \mathbf{t}_j\left(\boxed{t_J}\right)$ such that

$$U_j\left(a_j, f_j\left(t'_j\right)|_{A_{J\setminus\{j\}}\times \mathbf{s}_{-J}\left(\boxed{t_J}\right)}\right) \geq U_j\left(s'_j, f_j\left(t'_j\right)|_{A_{J\setminus\{j\}}\times \mathbf{s}_{-J}\left(\boxed{t_J}\right)}\right) \quad \forall s'_j \in S_j.$$

Let $\Re_J = \{(s,t) \in S \times T : J \text{ is Bayesian rational at } (s,t)\}$.

In Definition 2, "J is Bayesian rational at (s,t)" says that coalition members in J share a common belief that no credible profitable coalitional deviation, from the initial set $\mathbf{s}\left(\boxed{t_J}\right)$, precludes jointly playing strategies s_J of coalition members at the state (s,t). In other words, a coalition is Bayesian rational at a state if coalition members commonly believe that they are playing strategies within each vital coalitional agreement A_J of the coalition, to which

[12]For the singleton coalition $J=\{j\}$, decree $f_j\left(t'_j\right)|_{A_{J\setminus\{j\}}\times \mathbf{s}_{-J}\left(\boxed{t_J}\right)} = f_j\left(t'_j\right)|_{\mathbf{s}_{-j}\left(\boxed{t_J}\right)}$.

players in the coalition would be willing to confine their play.[13] In particular, κ-Profitability entails that at every state in the minimal state subspace $\boxed{t_J}$, every coalition member expects a higher payoff if a coalitional agreement is made; κ-Credibility ensures each incremental strategy $a_j \in A_j \setminus \mathbf{s}_j\left(\boxed{t_J}\right)$ satisfies individual rationality. Our formalism of Bayesian c-rationality is genuinely intersubjective in nature; it is explicitly and entirely based on the assumptions about coalition members' behavior and (common) beliefs at epistemic states, in terms of what the players in coalitions believe about each other's strategies, information, and beliefs.

Although the set A_J in Definition 2 can be interpreted as a variant of the "coalitional best response" set in Ambrus [4], the notions of Bayesian c-rationality and γ-rationality differ in some important aspects.[14] Because the operator γ contains only restricted subsets of strategies, the notion of γ-rationality entails an excessive amount of attention to a wide range of "initial agreements" –all product supersets of $\mathbf{s}\left(\boxed{t_J}\right)$ that satisfy "closed under rational behavior (curb)" in the sense of Basu and Weibull [14] (see also Grandjean et al. [36] for this requirement for coalitions); that is, it requires players in J to be logically omniscient with respect to a special class of "curb" agreements hinging on a "behavior-richness" framework. By contrast, our notion of Bayesian c-rationality merely requires the initial agreement under consideration to be the "exact" product set $\mathbf{s}\left(\boxed{t_J}\right)$ (i.e., the "finest" agreement commonly believed by coalition J at its epistemic state t_J), without the requirement of "closed under rational behavior."

One important feature of Definition 2 is that the concept entails coalition-common-belief preconditions: (i) coalition members in J share a common belief that play is in the initial agreement $\mathbf{s}\left(\boxed{t_J}\right)$, and (ii) coalition members in J share a common belief that the contemplating joint movement from the agreement $\mathbf{s}\left(\boxed{t_J}\right)$ to a new coalitional agreement A_J is a vital deviation. Interactive belief plays a crucial role in all accounts of joint actions by rational players in coalitions; it allows players to explore the mutually beneficial opportunity by joint movements (see, e.g., Schelling [73], Lewis [51], Chwe [28] and Sugden [76] for more discussions). The first coalition-common-belief condition (i) is harmonious with the prereq-

[13] Coalition members are not allowed to pool their private information in our noncooperative framework. The epistemic prerequisite for a vital coalitional deviation is that the coalition members share a common knowledge that the deviation is mutually profitable. This idea has the same spirit of Wilson's [79] concept of coarse core in an exchange economy with asymmetric information.

[14] To see this, let $\mathbf{s}\left(\boxed{t_J}\right) \Rrightarrow A_J$ denote "coalition J's κ-credible and κ-profitable movement from $\mathbf{s}\left(\boxed{t_J}\right)$ to a coalitional agreement A_J in Definition 2". Coalition J is Bayesian rational at state (s,t) iff $s_J \in \gamma\left(\mathbf{s}\left(\boxed{t_J}\right), J\right) \equiv \cap_{A_J \subseteq S_J;\, \mathbf{s}\left(\boxed{t_J}\right) \Rrightarrow A_J} A_J$.

uisite for Ambrus's [4] γ-rationality: there is common certainty among J that play is in the initial agreement. The second coalition-common-belief condition (ii) is an implicit precondition for the changing process from the initial agreement to a new coalitional agreement in the concept of coalition rationality.

The classic definition of individual rationality requires that an agent ought to choose an optimal action with respect to his probabilistic belief about the exhaustive uncertainty, irrespective of whether the agent is capable of knowing precisely his epistemic type (which is part of the structure of this agent is uncertain about in a model of decision making under uncertainty). For the special case of singleton coalitions, Proposition 1 below establishes a relationship between the definitions of Bayesian c-rationality and Bayesian rationality; under the commonly used assumption that the player knows his own primary belief and using strategy, the two notions indeed coincide.

Proposition 1. *Consider a type structure $\mathcal{T}(\mathcal{G})$. Suppose $J = \{j\}$.*

(a) $(s,t) \in \Re_J$ *iff* $s_j \in \mathbf{BR}_j\left(\boxed{t_j}\right)$ *where*

$$\mathbf{BR}_j\left(\boxed{t_j}\right) = \left\{s'_j \in S_j : \exists\,(s',t') \in \boxed{t_j} \text{ s.t. } s'_j \in \mathbf{BR}_j\left(t'_j\right)\right\};$$

thus, $\mathbf{s}_j(\Re_J) \subseteq \mathbf{s}_j(R_j)$.

(b) *If j knows his own using strategy s_j and primary belief $\beta^0_j(t_j)$ at state (s,t) –i.e., $\beta^0_j(t_j)\left(\left\{(s',t') \in S \times T : s'_j = s_j \text{ and } \beta^0_j\left(t'_j\right) = \beta^0_j(t_j)\right\}\right) = 1$, then $\Re_J = R_j$.*

Proposition 1(a) brings out a novel way to define individually rational behavior in environments where players hold introspective beliefs about their own behavior and epistemic types.

Definition 2′. In the type structure $\mathcal{T}(\mathcal{G})$, *player i is Bayesian c-rational at state* (s,t) if $s_i \in \mathbf{BR}_i\left(\boxed{t_i}\right)$. Let $\Re_i = \{(s,t) \in S \times T : i \text{ is Bayesian c-rational at } (s,t)\}$.

That is, a singleton player is Bayesian c-rational if it is not the case that the player is self-aware he could do better by a replacement of his playing strategy. The alternative definition of Bayesian rationality expounds the precondition for the player's rational behavior, in terms of inner perception and introspection about how the player's behavior and mental states. Proposition 1(a) asserts that the behavioral implication of Bayesian c-rationality in Definition 2′ must fall within the scope of Bayesian rationality; notably, $\Re_i = \Re_{\{i\}}$.

Remark 2. Definition $2'$ entails that, on self-reflection and introspection, a rational player (as a singleton coalition) cannot refute that he is doing the optimal action given his imperfect information about his own type; it is in the same spirit of Gilboa et al.'s [35] "subjective" definition of rationality in a multiple prior model: the decision maker cannot be convinced that he is wrong in making a decision.[15] Our formalism shares a common feature of coalitional γ-rationality in Ambrus [4]; that is, the classic definition of individual rationality is distinct from the notion of coalitional rationality under the restriction of singleton coalition, because the former one is silent on the epistemic precondition for rational behavior. As Example 2 below demonstrates, an individually rational state might not be rational in the sense of Definition $2'$, because the individual player (unlike an omniscient analyst) fails to know his type at that state.

Example 2. Consider a two-person game $\mathcal{G}$ (where the first player picks a row and the second player picks a column):

	a_2	b_2
a_1	$1,0$	$0,0$
b_1	$0,0$	$1,0$

Consider a type structure $\mathcal{T}(\mathcal{G})$ as follows:

$T_1 = \{t_1, t_1'\}$ and $T_2 = \{t_2\}$;

$\beta_i : T_i \to \Delta^* (S \times T)$ (for $i = 1, 2$) satisfying $\beta_1^0 (t_1') = 1 \circ (b_1, a_2; t_1, t_2)$ and $\beta_1^0 (t_1) = \beta_2^0 (t_2) = 1 \circ (b_1, b_2; t_1, t_2)$.

In this example, $(a_1, b_2; t_1', t_2) \in R_1$, but $(a_1, b_2; t_1', t_2) \notin \Re_1$.

Intuitively, because player 1 does not know his own type t_1', he wrongly believes in t_1; player 1 has the state subspace $\boxed{t_1'} = \{(b_1, a_2; t_1, t_2), (b_1, b_2; t_1, t_2)\}$. Thus, the singleton player 1 is not Bayesian c-rational to play a_1 at t_1', because $a_1 \notin \mathbf{BR}_i \left(\boxed{t_i'}\right)$ –i.e., player 1 cannot find a reason to justify playing a_1. Nevertheless, player 1 is Bayesian rational to play a_1 at t_1', simply because a_1 is a best response to $f_1^0 (t_1')$, even though the player wrongly believes in t_1 and hence should refrain from playing a_1.

[15] This kind of definition is in the spirit of Gilboa's [34] ascriptive theory: it describes a decision maker's behavior, but can also be ascribed to the decision maker without refuting itself.

5 (Strong) Nash equilibrium and Bayesian coalitional rationality

In this section, we discuss the solution concepts of Nash equilibrium and strong Nash equilibrium in our framework where players are allowed to have beliefs about their own strategies and types, through the lens of the notion of Bayesian coalitional rationality.

5.1 Nash Equilibrium

Aumann and Brandenburger [13] observed, under the assumption that each player knows his own strategy choice and (probabilistic) belief, that if all the players are Bayesian rational and mutually know the strategy choices of the others, then the players' choices constitute a Nash equilibrium in the game being played. We extend this result to environments where the players have introspective beliefs about their strategy choices and beliefs.

Proposition 2. *Consider a type structure $\mathcal{T}(\mathcal{G})$.*

(a) *If every player is Bayesian c-rational and "introspectively knows" that they are playing s at state (s,t), then s is a (pure-strategy) Nash equilibrium in $\mathcal{G}$. That is, if $(s,t) \in \Re_i$ and $\mathbf{s}\left(\boxed{t_i}\right) = \{s\}$ $\forall i \in I$, then $u_i(s_i, s_{-i}) \geqslant u_i(s_i', s_{-i})$ $\forall s_i' \in S_i$ $\forall i \in I$.*

(b) *Suppose each player knows his own using strategy and primary belief at state (s,t). If each player is Bayesian rational and knows the opponents are playing s_{-i}, then s is a (pure-strategy) Nash equilibrium in $\mathcal{G}$. That is, if $(s,t) \in R_i$ and $f_i^0(t_i)(s_{-i}) = 1$ $\forall i \in I$, then $u_i(s_i, s_{-i}) \geqslant u_i(s_i', s_{-i})$ $\forall s_i' \in S_i$ $\forall i \in I$.*

5.2 Strong Nash Equilibrium

Aumann [7] introduced the notion of "strong Nash equilibrium" by checking deviations by every conceivable coalition, instead of checking deviations by every individual player only. A strategy profile $s^* \in S$ is a *(pure-strategy) strong Nash equilibrium* if, for every coalition $J \subseteq I$ and for every profile $s_J \in S_J$ there exists a coalition member $j \in J$ such that $u_j(s^*) \geqslant u_j\left(s_J, s_{-J}^*\right)$. That is, an equilibrium is strong if no coalition, taking the actions of its complement as given, can jointly deviate in a way that benefits all of its members. For the notion of strong Nash equilibrium, we need a stronger version of Bayesian coalitional rationality by removing the Credibility condition in Definition 2.

Definition 3. In a type structure $\mathcal{T}(\mathcal{G})$, *coalition J is strong Bayesian rational at state* (s,t) if $s_J \in A_J$ whenever the nonempty product subset $A_J \subseteq S_J$ satisfies the κ-Profitability condition: for all $j \in J$ and $(s',t') \in \boxed{t_J}$, if $s'_j \notin A_j$ or $f_j\left(t'_j\right)|_{A_{J\setminus\{j\}}\times\mathbf{s}_{-J}\left(\boxed{t_J}\right)} \neq f_j\left(t'_j\right)|_{\mathbf{s}_{-j}\left(\boxed{t_J}\right)}$, then there exists $\widehat{s}_j \in S_j$ such that

$$U_j\left(s'_j, f_j\left(t'_j\right)|_{\mathbf{s}_{-j}\left(\boxed{t_J}\right)}\right) < U_j\left(\widehat{s}_j, f_j\left(t'_j\right)|_{A_{J\setminus\{j\}}\times\mathbf{s}_{-J}\left(\boxed{t_J}\right)}\right).$$

Let $\Re_J^{\text{strong}} = \{(s,t) \in S \times T : J \text{ is strong Bayesian rational at } (s,t)\}$.

Proposition 3. *Consider a type structure $\mathcal{T}(\mathcal{G})$. If each coalition is strong Bayesian rational and "commonly knows" players are playing s at state (s,t), then s is a (pure-strategy) strong Nash equilibrium. That is, if $(s,t) \in \Re_J^{\text{strong}}$ and $\mathbf{s}\left(\boxed{t_J}\right) = \{s\}$ for all coalitions $J \subseteq I$, then for every coalition $J \subseteq I$ and for every profile $s'_J \in S_J$ there exists a coalition member $j \in J$ such that $u_j(s) \geqslant u_j(s'_J, s_{-J})$.*

6 Bayesian coalitional rationalizability and common knowledge of Bayesian coalitional rationality

In this section, we study the behavioral implication of "common knowledge of Bayesian c-rationality" –that is, the strategic implication of "Bayesian c-rationality and common belief of Bayesian c-rationality."

6.1 Bayesian coalitional rationalizability

In noncooperative games, Ambrus [3] first offered a game-theoretic solution concept of "coalitional rationalizability" to capture the idea of common knowledge of coalitional rationality. This concept is defined through an "internal coalitional reasoning":

> ... players look for agreements to avoid certain strategies, without specifying play within the set of non-excluded strategies ... A restriction is supported if every group member always (for every possible expectation) expects a higher payoff if the agreement is made than if he instead chooses to play a strategy outside the agreement. (Ambrus [3, p.904])

Luo and Yang [53] proposed an alternative solution concept of "Bayesian coalitional rationalizability" (henceforth, Bayesian c-rationalizability) for situations in which, in pursuit of mutually beneficial interests, the players in a coalition (i) evaluate their payoff expectations by Bayesian updating, if a coalitional agreement is made, and (ii) contemplate various plausible deviations –that is, the validity of deviation is checked not only against restricted subsets of strategies as in Ambrus [3], but also against arbitrary sets of strategies. The following example shows the main feature of Bayesian c-rationalizability.

Example 3. Consider the two-person symmetric game (where the first player picks a row and the second player picks a column):

	a_2	b_2	c_2
a_1	$3,0$	$0,3$	$0,2$
b_1	$0,3$	$3,0$	$0,0$
c_1	$2,0$	$0,0$	$1,1$

From the Bayesian viewpoint, the two players would like to confine their play to a subset of strategies $\{a_1, b_1\} \times \{a_2, b_2\}$ (which is also called a "coalitional agreement"). Let $i, j \in \{1, 2\}$ and $j \neq i$. Intuitively, if player i assigns a prior probability of less than 0.5 to the j's strategy a_j, then the player achieves an expected payoff of less than 1.5 by playing c_i. It is beneficial for the players to reach the coalitional agreement because each can guarantee a higher expected payoff of 1.5. If player i assigns a prior probability of more than or equal to 0.5 to a_j, then it is also beneficial for the players to reach the coalitional agreement instead of playing c_i because in this case each player can achieve a higher expected payoff under the updated belief. (Without appealing for Bayes' rule, player i could achieve an expected payoff of 2 by playing c_i, higher than some expected payoff that might be obtained by playing a_i or b_i after the coalitional agreement is made. Due to this difference, Ambrus's [3] notion of coalitional rationalizability does not rule out strategy c_i.) The notion of Bayesian c-rationalizability yields the outcome set $\{a_1, b_1\} \times \{a_2, b_2\}$. This notion reflects the idea that a group of players can coordinate their play to achieve mutually beneficial outcomes, because of common knowledge of the fact that the players in group are Bayesian rational and aware of mutually beneficial arrangement of strategies.

Formally, consider a game $\mathcal{G}$. For player j, let $\Delta^*(S_{-j})$ denote the set of all conditional probability systems (CPSs) on finite set S_{-j} of the opponents' strategy profiles. For nonempty subset $A \subseteq S$, let

$$\Delta^*_A(S_{-i}) = \left\{ \mu| \in \Delta^*(S_{-i}) : \mu|_{S_{-i}}(A_{-i}) = 1 \right\}.$$

For nonempty product subsets $A, B \subseteq S$, we say a coalition $\mathcal{J}_{AB}$ is a "feasible coalition from A to B" if $B = B_{\mathcal{J}_{AB}} \times A_{-\mathcal{J}_{AB}}$.

Definition 4 (Luo and Yang [53]). A nonempty product subset $\mathcal{R} \subseteq S$ is a *coalitional rationalizable set (CRS)* if $\mathcal{R} \Rrightarrow \mathcal{R}'$ only for $\mathcal{R}' = \mathcal{R}$, where for $\mathcal{R} \neq \emptyset$, we define $\mathcal{R} \Rrightarrow \mathcal{R}'$ as: $\exists$ a feasible coalition $\mathcal{J}_{\mathcal{R}\mathcal{R}'}$, $\forall j \in \mathcal{J}_{\mathcal{R}\mathcal{R}'}$ such that

(1) **[Profitability]** $\forall r_j \in \mathcal{R}_j$, $\forall \mu| \in \Delta^*_{\mathcal{R}}(S_{-j})$, if $r_j \notin \mathcal{R}'_j$ or $\mu|_{\mathcal{R}_{-j}} \neq \mu|_{\mathcal{R}'_{-j}}$, then $U_j(r_j, \mu|_{\mathcal{R}_{-j}}) < U_j(s_j, \mu|_{\mathcal{R}'_{-j}})$ for some $s_j \in S_j$, and

(2) **[Credibility]** $\forall r'_j \in \mathcal{R}'_j \backslash \mathcal{R}_j$, there exists $\mu| \in \Delta^*_{\mathcal{R}}(S_{-j})$ such that $U_j(r'_j, \mu|_{\mathcal{R}'_{-j}}) \geq U_j(s_j, \mu|_{\mathcal{R}'_{-j}})$ $\forall s_j \in S_j$.

A strategy $r_i \in \mathcal{R}_i$ is said to be a *Bayesian c-rationalizable strategy* for player i; let $\mathcal{R}^*$ denote the set of Bayesian c-rationalizable strategy profiles in game $\mathcal{G}$.

That is, a CRS $\mathcal{R}$ is a (nonempty) product set of pure strategies from which no group of players would like to make a "profitable" and "credible" deviation. With the restriction of $|\mathcal{J}_{\mathcal{R}\mathcal{R}'}| = 1$, Definition 4 yields a correlated version of rationalizability (see Luo and Yang [53]).

6.2 Common knowledge of Bayesian coalitional rationality

Consider a type structure $\mathcal{T}(\mathcal{G})$. Let

$$\mathcal{A}_J\left(\boxed{t_J}\right) = \left\{A_J \subseteq S_J : \text{deviation } A_J \times \mathbf{s}_{-J}\left(\boxed{t_J}\right) \text{ from } \mathbf{s}\left(\boxed{t_J}\right) \text{ satisfies Credibility in Definition 4}\right\}$$

denotes the collection of coalition J's credible agreements at epistemic state t_J. We say *coalition J is belief-rich-for-credible-deviations (BRCD) at $t \in T$*, if for each $A_J \in \mathcal{A}_J\left(\boxed{t_J}\right)$, $j \in J$ and $a_j \in A_j \setminus \mathbf{s}_j\left(\boxed{t_J}\right)$, there exists $t'_j \in \mathbf{t}_j\left(\boxed{t_J}\right)$ such that

$$U_j\left(a_j, f_j\left(t'_j\right)|_{A_{J\setminus\{j\}} \times \mathbf{s}_{-J}\left(\boxed{t_J}\right)}\right) \geq U_j\left(s'_j, f_j\left(t'_j\right)|_{A_{J\setminus\{j\}} \times \mathbf{s}_{-J}\left(\boxed{t_J}\right)}\right) \ \forall s'_j \in S_j.$$

In words, the BRCD property requires any credible deviation from $s\left(\boxed{t_J}\right)$ via J in game $\mathcal{G}$ satisfies κ-Credibility in the type structure $\mathcal{T}(\mathcal{G})$. A *type structure $\mathcal{T}(\mathcal{G})$ is BRCD* if every coalition is BRCD at every type profile $t \in T$. Let

$$\mathring{\Re} = \bigcap\nolimits_{J \subseteq I} \Re_J,$$

denote the event "every coalition is Bayesian rational in Definition 2." For subset $E \subseteq S \times T$, let $\mathrm{proj}_S(E)$ denote the projection of E on S; that is, $\mathrm{proj}_S(E) = \{s \in S : (s,t) \in E\}$. We are now in a position to present a main result of this paper.

Proposition 4. *(a) For any BRCD* $\mathcal{T}(\mathcal{G})$, $\text{proj}_S\left(\overset{\circ}{\Re}\cap CB\overset{\circ}{\Re}\right)\subseteq\mathcal{R}^*$. *(b) There exists a finite BRCD* $\mathcal{T}(\mathcal{G})$ *such that* $\text{proj}_S\left(\overset{\circ}{\Re}\cap CB\overset{\circ}{\Re}\right)=\mathcal{R}^*$.

In Proposition 4, we need to consider a type structure model, in which each player is rich in beliefs such that no credible deviation in Definition 4 is excluded from consideration. This sort of belief-richness condition in the analytical framework is commonly used for the epistemic analysis of game-theoretic solution concepts, for example, extensive-form rationalizability in a complete CPS type structure model in Battigalli and Siniscalchi [15] and iterated weak dominance in a complete "lexicographic conditional probability system (LCPS)" type structure model in Brandenburger et al. [24].[16] Remarkably, our belief-richness condition accommodates the finite model of type structure; Proposition 4(b) shows the existence of a finite BRCD type structure that is compatible with the assumption "common knowledge of Bayesian coalitional rationality."

The following example shows that, without imposing the belief-richness condition, common knowledge of Bayesian c-rationality may generate a strategy profile that is not Bayesian c-rationalizable.

Example 4. Consider again the two-person game in Example 3:

	a_2	b_2	c_2
a_1	$3,0$	$0,3$	$0,2$
b_1	$0,3$	$3,0$	$0,0$
c_1	$2,0$	$0,0$	$1,1$

Consider a type structure $\mathcal{T}(\mathcal{G})$: for $i,j=1,2$ and $i\neq j$, $T_i=\{t_i\}$, $\beta_i^0(t_i)=1\circ(c_1,c_2;t_1,t_2)$ and $f_i(t_i)|_{\{a_j,b_j\}}=1\circ a_j$. Because no coalitional deviation from $\mathbf{s}\left(\boxed{t}\right)=\{c_1\}\times\{c_2\}$ is κ-credible in $\mathcal{T}(\mathcal{G})$, $(c_1,c_2;t_1,t_2)$ is a Bayesian c-rational state; hence, Bayesian c-rationality is commonly known at $(c_1,c_2;t_1,t_2)$. But, (c_1,c_2) is not a Bayesian c-rationalizable strategy profile, because confining the players' play to a subset of strategies $\{a_1,b_1\}\times\{a_2,b_2\}$ is in their mutual interest (cf. Example 3). The main issue here is that, in this type structure $\mathcal{T}(\mathcal{G})$, the CPS beliefs of each player are rather sparse; no rich beliefs support the desirable deviation $\{a_1,b_1\}\times\{a_2,b_2\}$ from $\{c_1\}\times\{c_2\}$ to be κ-credible in $\mathcal{T}(\mathcal{G})$. That is, the type structure $\mathcal{T}(\mathcal{G})$ fails to satisfy the BRCD condition.

[16] Although this kind of "belief richness" is crucial to our epistemic analysis, an analogous concept does not appear in Ambrus's [4] framework, because of making use of the shortcut operator γ. In Ambrus [4], the prominent operator γ^* implicitly requires a sort of belief-richness condition: a player must consider all possible conjectures in the game, each of which supports a best response strategy (excluded from the restricted set of strategies).

Proposition 4(a) asserts this Bayesian c-rationalizable profile (c_1, c_2) cannot be attained by common knowledge of Bayesian c-rationality in a BRCD type structure. To see this point, consider a type structure $\widehat{\mathcal{T}}(\mathcal{G})$ such that $\widehat{T}_1 = \{t_1, t_1'\}$, $\widehat{T}_2 = \{t_2, t_2'\}$ and

$$\beta_1^0(t_1) = 1 \circ (a_1, a_2; t_1, t_2);$$

$$\beta_1^0(t_1') = 1 \circ (b_1, b_2; t_1', t_2');$$

$$\beta_2^0(t_2) = 1 \circ (b_1, a_2; t_1', t_2);$$

$$\beta_2^0(t_2') = 1 \circ (a_1, b_2; t_1, t_2');$$

$$f_1(t_1)|_{\{b_2, c_2\}} = f_1(t_1')|_{\{a_2, c_2\}} = 1 \circ c_2;$$

$$f_2(t_2)|_{\{a_1, c_1\}} = f_2(t_2')|_{\{b_1, c_1\}} = 1 \circ c_1.$$

It is easy to verify that $\widehat{\mathcal{T}}(\mathcal{G})$ is a BRCD type structure. For example, for the grand coalition $\{1, 2\}$, $\boxed{t} = \{(a_1, a_2; t_1, t_2), (b_1, b_2; t_1', t_2'), (b_1, a_2; t_1', t_2), (a_1, b_2; t_1, t_2')\}$ $\forall t \in T$. Consider a credible deviation $\{b_1, c_1\} \times \{c_2\}$ from $s(\boxed{t}) = \{a_1, b_1\} \times \{a_2, b_2\}$, player 1 can use $f_1(t_1)|_{\{c_2\}} = 1 \circ c_2$ to justify playing c_1 and player 2 can use $f_2(t_2')|_{\{b_1, c_1\}} = 1 \circ c_1$ to justify playing c_2. Hence, $\{b_1, c_1\} \times \{c_2\}$ is κ-credible in $\widehat{\mathcal{T}}(\mathcal{G})$.

Because every player attains the highest expected payoff of 3 at every state in $\boxed{t}$, there is no κ-profitable deviation by changing the play of player(s). Therefore, Bayesian c-rationality is commonly known across $\boxed{t}$. Consequently, in the BRCD type structure $\widehat{\mathcal{T}}(\mathcal{G})$, common knowledge of Bayesian c-rationality yields the set of Bayesian c-rationalizable strategy profiles: $\{a_1, b_1\} \times \{a_2, b_2\}$.

With the restriction of singleton coalitions, Definition 4 yields the definition of (correlated) rationalizability. As an immediate corollary of Theorem 1, we obtain a characterization for rationalizability without appealing to the belief-richness condition. Let $\mathcal{R}$ denote the set of (correlated) rationalizable strategy profiles in $\mathcal{G}$, and let

$$\Re = \cap_{i \in I} \Re_i$$

denote the event "every player is Bayesian c-rational in Definition $2'$."

Corollary 1. *(a) In any type structure* $\mathcal{T}(\mathcal{G})$, $\text{proj}_S(\Re \cap CB\Re) \subseteq \mathcal{R}$. *(b) There is a type structure* $\mathcal{T}(\mathcal{G})$ *such that* $\text{proj}_S(\Re \cap CB\Re) = \mathcal{R}$.

Remark 3. The BRCD condition in Proposition 4 ensures the credibility requirement for coalitional deviations in Definition 2 (cf. Bernheim et al. [20]). The BRCD condition is no longer needed if we remove the credibility requirement in Definition 2. However, like Aumann's [7] notion of strong Nash equilibrium, no "strong Bayesian c-rational" state (Definition 3) exists in any type structure for a certain game (e.g., the Prisoner's Dilemma).

7 Concluding remarks

The study of how groups of players act in their mutually beneficial interest in social environments is of great importance in economics and social sciences; for example, Paul Samuelson's [72] "The Pure Theory of Public Expenditure," Garrett Hardin's [42] "The Tragedy of the Commons" and Mancur Olson's [61] "The Logic of Collective Action" have provided examples to demonstrate the difference between the individual rationality and collective rationality.

In a broader sense, there is a variety of theories about the collectively rational behavior in game situations. The pattern of Nash equilibrium behavior prevails, in the absence of coalitional considerations, if each player is Bayesian rational and has a correct belief about the opponents' strategy choices (see Aumann and Brandenburger [13]). In an equilibrium approach, Aumann [7] and Bernheim et al. [20] offered the solution concepts of strong Nash equilibrium and coalition-proof Nash equilibrium for the collectively rational behavior, which takes into account the interests of coalitions in games. In a non-equilibrium paradigm, Bernheim [19] and Pearce [63] demonstrated the rationalizable strategic behavior is obtained if each player is Bayesian rational and Bayesian rationality is common belief among the players (see also Tan and Werlang [77]); the notion of rationalizable rational behavior is defined in a purely noncooperative environment, without taking into consideration the coalition's behavior in strategic interactions.

In this paper, we have offered a definition of coalitional rationality –i.e., Bayesian c-rationality– in the context of noncooperative games where players are with the Bayesian view of the world, being aware of the effects of coalitional reasoning. The notion of Bayesian c-rationality prescribes a mode of behavior that no group of players wishes to change –that is, it is not the case that the coalition members commonly believe that mutually beneficial gains exist by a (credible) coalitional agreement. Our approach adheres to the conventional point of view that the game model under consideration fully describes any aspect of coalitional bargaining and coalitional negotiation. Accordingly, although players cannot make any explicit/binding form of coalitional agreements, they can undertake a deductive reasoning by means of tacit/implicit coalitional agreements existing in a noncooperative environment.[17]

From an epistemic perspective, exploring how to make "rational" states possible involving the distinctive mode of coalitional reasoning is theoretically important and profound in the context of games (see also Chant and Ernst [25] more discussions). Such an epistemic analysis can help in better understanding when a particular solution concept is applicable in practical

[17] There are other approaches to coalitional behavior in the literature (see, e.g., Ray and Vohra [68]). For instance, Greenberg [37, Chapter 5], Chwe [27], Mariotti [54], Ray and Vohra [66, 67], and Xue [80] studied models for coalitional negotiation/bargaining in which coalitions act publicly (and thus consequences are publicly observed) or coalitions make binding "point" agreements (rather than the nonbinding coalitional agreements used in this paper). See also Newton [59, 60] for an in-depth study of coalitional behavior from the point of view of evolutionary game theory.

circumstances. Ambrus [4] made an attempt to present a definition of coalitional γ-rationality by highlighting the precondition for the coalitionally rational behavior. Along the lines of the epistemic game theory program (see Aumann and Brandenburger [13], Aumann [10, 11], Brandenburger [21], Dekel and Siniscalchi [29], and Perea [64]), we have offered a more comprehensive definition for coalitional deliberations in noncooperative games (Definition 2). Like Aumann [9, 8, 10, 11] and Aumann and Brandenburger [13], we carry out our analysis within an arbitrary model, including finite and infinite models.

Our approach sheds light on the preconditions for coalitionally rational behavior in the context of a noncooperative game where members of coalitions are with the Bayesian view of the world. With restricting the size of coalition to one, our definition of Bayesian c-rationality is harmonious with the individual Bayesian rationality (Proposition 1). By using the notion of Bayesian c-rationality, we have provided an epistemic analysis of the solution concepts such as (strong) Nash equilibrium (Aumann [7]) and Bayesian coalitional rationalizability (Luo and Yang [53]). In particular, we have established epistemic conditions for the notions of Nash equilibrium and strong Nash equilibrium in situations where players are allowed to have introspective beliefs about their own strategies and types (Propositions 2 and 3); we have offered an epistemic characterization of the notion of Bayesian coalitional rationalizability in terms of "common knowledge of Bayesian c-rationality" (Proposition 4).

In closing, we mention some possible extensions. In this paper, we define Bayesian c-rationality by assuming players are subjective expected utility maximizers and coordinate their play to achieve mutual gains through nonbinding and tacit agreements on joint actions. Alternatively, we can consider the coalitional preferences as the aggregation of the preferences of coalition members; see, for example, Hara et al. [41] for a coalitional expected multi-utility theory. The extension of this paper to games with different modes of coalitional behavior is an intriguing subject for further research; cf. Asheim [6], Chen et al. [26], and Epstein [30] for related work on rationalizability under general preferences. The exploration of the notion of coalitional rationality in dynamic settings and incomplete information settings is also an important research topic for further study.

8 Appendix: Proofs

Proof of Proposition 1. (a) The "only if" part: Suppose $(s,t) \in \Re_J$ where $J = \{j\}$. It suffices to show that the set $\mathbf{BR}_j\left(\boxed{t_j}\right)$ is nonempty and satisfies κ-Profitability and κ-Credibility in Definition 2.

Step 1. Assume, in negation, that $\mathbf{BR}_j\left(\boxed{t_j}\right) = \emptyset$. Because $f_j^0\left(t_j'\right)\left(\mathbf{s}_{-j}\left(\boxed{t_j}\right)\right) = 1$ for any $(s',t') \in \boxed{t_j}$, s_j' is not a best response to $f_j^0\left(t_j'\right) = f_j\left(t_j'\right)|_{\mathbf{s}_{-j}\left(\boxed{t_j}\right)}$; that is,

$$U_j\left(s_j', f_j\left(t_j'\right)|_{\mathbf{s}_{-j}\left(\boxed{t_j}\right)}\right) < U_j\left(\widehat{s}_j, f_j\left(t_j'\right)|_{\mathbf{s}_{-j}\left(\boxed{t_j}\right)}\right) \text{ for some } \widehat{s}_j \in S_j.$$

Now, define

$$\mathbf{BR}_j\left(\mathbf{t}_j\left(\boxed{t_j}\right)\right) = \left\{\widehat{s}_j \in S_j : \exists t_j' \in \mathbf{t}_j\left(\boxed{t_j}\right) \text{ s.t. } \widehat{s}_j \in \mathbf{BR}_j\left(t_j'\right)\right\}.$$

Observe that every nonempty subset $A_j \subseteq \mathbf{BR}_j\left(\mathbf{t}_j\left(\boxed{t_j}\right)\right)$ satisfies κ-Profitability and κ-Credibility. By Definition 2, $s_j \in A_j \subseteq \mathbf{BR}_j\left(\mathbf{t}_j\left(\boxed{t_j}\right)\right)$ for all nonempty subsets $A_j \subseteq \mathbf{BR}_j\left(\mathbf{t}_j\left(\boxed{t_j}\right)\right)$ because $(s,t) \in \Re_J$. Therefore, $\mathbf{BR}_j\left(\mathbf{t}_j\left(\boxed{t_j}\right)\right) = \{s_j\}$; thus, $s_j \in \mathbf{BR}_j\left(t_j'\right)$ for all $t' \in \mathbf{t}_j\left(\boxed{t_j}\right)$. But, by Definition 2, $s_j \in \mathbf{s}_j\left(\boxed{t_j}\right)$ because $\mathbf{s}_j\left(\boxed{t_j}\right) \neq \emptyset$ also satisfies κ-Profitability and κ-Credibility. Hence, $s_j \in \mathbf{BR}_j\left(\boxed{t_j}\right)$. A contradiction.

Step 2. Since $\mathbf{BR}_j\left(\boxed{t_j}\right) \subseteq \mathbf{s}_j\left(\boxed{t_j}\right)$, $\mathbf{BR}_j\left(\boxed{t_j}\right)$ satisfies κ-Credibility. Assume, in negation, that $\mathbf{BR}_j\left(\boxed{t_j}\right) \neq \emptyset$ fails to satisfy κ-Profitability. Since $f_j\left(t_j'\right)|_{A_{J\setminus\{j\}}\times\mathbf{s}_{-J}\left(\boxed{t_j}\right)} = f_j\left(t_j'\right)|_{\mathbf{s}_{-j}\left(\boxed{t_j}\right)} = f_j^0\left(t_j'\right)$ $\forall t_j' \in T_j$ when $J = \{j\}$, there exists $(s',t') \in \boxed{t_j}$ such that $s_j' \notin \mathbf{BR}_j\left(\boxed{t_j}\right)$ and

$$U_j\left(s_j', f_j^0\left(t_j'\right)\right) \geq U_j\left(s_j'', f_j^0\left(t_j'\right)\right) \ \forall s_j'' \in S_j.$$

Therefore, $s_j' \in \mathbf{BR}_j\left(\boxed{t_j}\right)$. A contradiction.

The "if" part: Suppose $s_j \in \mathbf{BR}_j\left(\boxed{t_j}\right)$. It suffices to show that $\mathbf{BR}_j\left(\boxed{t_j}\right) \subseteq A_j$ whenever $A_j \neq \emptyset$ satisfies κ-Profitability and κ-Credibility in Definition 2. Assume, in negation, that there exists $s'_j \in \mathbf{BR}_j\left(\boxed{t_j}\right)$ such that $s'_j \notin A_j$ and $A_j \neq \emptyset$ satisfies κ-Profitability and κ-Credibility. Since $s'_j \in \mathbf{BR}_j\left(\boxed{t_j}\right)$, there exists $(s', t') \in \boxed{t_j}$ such that $s'_j \in \mathbf{BR}_j\left(t'_j\right)$. Therefore, $s'_j \in \mathbf{s}_j\left(\boxed{t_j}\right)$ and $s'_j \notin A_j$. Because $A_j \neq \emptyset$ satisfies κ-Profitability,

$$U_j\left(s'_j, f_j\left(t'_j\right)|_{\mathbf{s}_{-j}\left(\boxed{t_j}\right)}\right) < U_j\left(\widehat{s}_j, f_j\left(t'_j\right)|_{\mathbf{s}_{-j}\left(\boxed{t_j}\right)}\right) \text{ for some } \widehat{s}_j \in S_j.$$

But, since $f_j^0\left(t'_j\right) = f_j\left(t'_j\right)|_{\mathbf{s}_{-j}\left(\boxed{t_j}\right)}$, $s'_j \notin \mathbf{BR}_j\left(t'_j\right)$. A contradiction.

(b) Let $(s,t) \in S \times T$. Since $\beta_j^0(t_j)\left(\left\{(s',t') \in S \times T : s'_j = s_j \text{ and } \beta_j^0\left(t'_j\right) = \beta_j^0(t_j)\right\}\right) = 1$, $\boxed{t_j} \subseteq \left\{(s',t') \in S \times T : s'_j = s_j \text{ and } \beta_j^0\left(t'_j\right) = \beta_j^0(t_j)\right\}$. Therefore, $f_j^0\left(t'_j\right) = f_j^0(t_j)\ \forall t'_j \in \mathbf{t}_j\left(\boxed{t_j}\right)$ and $\mathbf{s}_j\left(\boxed{t_j}\right) = \{s_j\}$. Thus, $\mathbf{BR}_j\left(\boxed{t_j}\right) = \mathbf{BR}_j(t_j) \cap \{s_j\}$. By Proposition 1(a), $(s,t) \in \Re_J$ iff $s_j \in \mathbf{BR}_j(t_j)$. ■

Proof of Proposition 2. (a) Let $i \in I$ and $(s,t) \in \Re_i$. By Proposition 1(a), $s_i \in \mathbf{BR}_i\left(\boxed{t_i}\right)$. Since $\mathbf{s}\left(\boxed{t_i}\right) = \{s\}$, there exists $(s,t') \in \boxed{t_i}$ such that $s_i \in \mathbf{BR}_i(t'_i)$. Again, by $\mathbf{s}\left(\boxed{t_i}\right) = \{s\}$, $f_i^0(t'_i) = 1 \circ s_{-i}\ \forall t' \in \mathbf{t}_j\left(\boxed{t_i}\right)$. Therefore, $u_i(s_i, s_{-i}) \geqslant u_i(s'_i, s_{-i})\ \forall s'_i \in S_i\ \forall i \in I$.

(b) Let $i \in I$ and $(s,t) \in R_i$. Because i knows his own using strategy, primary belief and the opponents are playing s_{-i} at state (s,t), $\boxed{t_j} \subseteq \{s\} \times T$, i.e., $\mathbf{s}\left(\boxed{t_i}\right) = \{s\}$. By Proposition 1(b), $\Re_i = R_i$. By Proposition 2(a), s is a Nash equilibrium in $\mathcal{G}$. ■

Proof of Proposition 3. Suppose that $(s,t) \in \Re_J^{\text{strong}}$ and $\mathbf{s}\left(\boxed{t_J}\right) = \{s\}$ for all coalitions $J \subseteq I$. Assume, in negation, that s is not a strong Nash equilibrium, then there exist a coalition J and $\widehat{s}_J \in S_J$ such that $u_j(s) < u_j(\widehat{s}_J, s_{-J})$ for all $j \in J$. We show $A_J = \{\widehat{s}_J\}$ satisfies the κ-Profitability. For all $j \in J$ and $(s',t') \in \boxed{t_J}$, $s'_j = s_j$ and $f_j\left(t'_j\right)|_{\mathbf{s}_{-j}\left(\boxed{t_J}\right)} = 1 \circ s_{-j}$ because J commonly believes that players play s at (s,t). Therefore, for all $j \in J$ and $(s',t') \in \boxed{t_J}$,

$$U_j\left(s'_j, f_j\left(t'_j\right)|_{\mathbf{s}_{-j}\left(\boxed{t_J}\right)}\right) = u_j(s) < u_j(\widehat{s}_J, s_{-J}) = U_j\left(\widehat{s}_j, f_j\left(t'_j\right)|_{A_{J\setminus\{j\}} \times \mathbf{s}_{-J}\left(\boxed{t_J}\right)}\right).$$

Since $(s,t) \in \Re_J^{\text{strong}}$, $s_J \in A_J$ and hence $s_J = \widehat{s}_J$. A contradiction. ■

Observe that, in a type structure $\mathcal{T}(\mathcal{G})$, player i believes event E if and only if the support of i's primary belief is a subset of event E; that is, for $t_i \in T_i$, $\beta_i^0(t_i)(E) = 1$ iff supp $\beta_i^0(t_i) \subseteq E$ where supp $\beta_i^0(t_i)$ denotes the support of distribution $\beta_i^0(t_i)$ (cf., e.g., Zamir and Vassilakis [82]). For $\mu \in \mathbf{\Delta}(S_{-i})$ define

$$BR_i(\mu) = \{s_i \in S_i : U_i(s_i, \mu) \geq U_i(s_i', \mu) \ \forall s_i' \in S_i\};$$

for $Z_{-i} \subseteq S_{-i}$ define

$$BR_i(Z_{-i}) = \{s_i \in S_i : \exists \mu \in \mathbf{\Delta}(S_{-i}) \text{ s.t. } \mu(Z_{-i}) = 1 \text{ and } s_i \in BR_i(\mu)\}.$$

Proof of Proposition 4. (a) Assume, in negation, that $\text{proj}_S\left(\mathring{\Re} \cap CB\mathring{\Re}\right) \nsubseteq \mathcal{R}^*$. By Proposition 1 in Luo and Yang [53], there exists a reduction (product-set) sequence $\{\mathcal{D}^\tau\}$ such that $\mathcal{R}^* = \bigcap_{\tau=0}^{\infty}\mathcal{D}^\tau$ with $\mathcal{D}^0 = S$ and $\mathcal{D}^\tau \Rrightarrow \mathcal{D}^{\tau+1}$ for all $\tau \geq 0$. Hence, there exists k such that $\text{proj}_S\left(\mathring{\Re} \cap CB\mathring{\Re}\right) \subseteq \mathcal{D}^k$ and $\text{proj}_S\left(\mathring{\Re} \cap CB\mathring{\Re}\right) \nsubseteq \mathcal{D}^{k+1}$, where $\mathcal{D}^{k+1} \subseteq \mathcal{D}^k$ and $\mathcal{D}^k \Rrightarrow \mathcal{D}^{k+1}$ via J. That is, $s_J \notin \mathcal{D}_J^{k+1}$ for some $(s,t) \in \mathring{\Re} \cap CB\mathring{\Re}$ because $\mathcal{D}_{-J}^k = \mathcal{D}_{-J}^{k+1}$. Clearly, $(s,t) \in CB\left(\mathring{\Re} \cap CB\mathring{\Re}\right)$ implies $\boxed{t_J} \subseteq \boxed{t} \subseteq \mathring{\Re} \cap CB\mathring{\Re}$, because $\boxed{t}$ is the finest event commonly believed by t. Therefore, $\mathbf{s}\left(\boxed{t_J}\right) \subseteq \mathcal{D}^k$. Let $J^0 = \left\{j \in J : \mathbf{s}_j\left(\boxed{t_J}\right) \cap \mathcal{D}_j^{k+1} = \emptyset\right\}$. We distinguish three cases.

1. $|J^0| = 0$: Define $A_J = \left(\mathbf{s}_j\left(\boxed{t_J}\right) \cap \mathcal{D}_j^{k+1}\right)_{j \in J}$. Obviously, A_J satisfies κ-Credibility at (s,t) because $A_j \subseteq \mathbf{s}_j(\boxed{t_J})\ \forall j \in J$. We proceed to show A_J satisfies κ-Profitability at (s,t). Apparently, for all $j \in J$ and $(s',t') \in \boxed{t_J}$, $f_j\left(t_j'\right)|_{\mathbf{s}_{-j}\left(\boxed{t_J}\right)} = f_j\left(t_j'\right)|_{\mathcal{D}_{-j}^k}$ because $f_j^0\left(t_j'\right)\left(\mathbf{s}_{-j}\left(\boxed{t_J}\right)\right) = 1$ and $\mathbf{s}_{-j}\left(\boxed{t_J}\right) \subseteq \mathcal{D}_{-j}^k$. Note that, supp $f_j^0\left(t_j'\right) \cap \mathcal{D}_{-j}^{k+1} =$ supp $f_j^0\left(t_j'\right) \cap \left(A_{J\setminus\{j\}} \times \mathbf{s}_{-J}\left(\boxed{t_J}\right)\right)$ because supp $f_j^0\left(t_j'\right) \subseteq \mathbf{s}_{-j}\left(\boxed{t_J}\right)$ and $\mathbf{s}_{-J}\left(\boxed{t_J}\right) \subseteq \mathcal{D}_{-J}^{k+1}$. (i) If supp $f_j^0\left(t_j'\right) \cap \mathcal{D}_{-j}^{k+1} \neq \emptyset$ then $f_j\left(t_j'\right)|_{A_{J\setminus\{j\}} \times \mathbf{s}_{-J}\left(\boxed{t_J}\right)} = f_j\left(t_j'\right)|_{\mathcal{D}_{-j}^{k+1}}$. By Profitability in Definition 4, it follows that if $s_j' \notin A_j$ or $f_j\left(t_j'\right)|_{\mathbf{s}_{-j}\left(\boxed{t_J}\right)} \neq f_j\left(t_j'\right)|_{A_{J\setminus\{j\}} \times \mathbf{s}_{-J}\left(\boxed{t_J}\right)}$, then $u_j(s_j', f_j\left(t_j'\right)|_{\mathbf{s}_{-j}\left(\boxed{t_J}\right)}) < u_j(s_j^*, f_j\left(t_j'\right)|_{A_{J\setminus\{j\}} \times \mathbf{s}_{-J}\left(\boxed{t_J}\right)})$ for some $s_j^* \in S_j$. (ii) If supp $f_j^0\left(t_j'\right) \cap \mathcal{D}_{-j}^{k+1} = \emptyset$ then $f_j\left(t_j'\right)|_{\mathbf{s}_{-j}\left(\boxed{t_J}\right)} \neq f_j\left(t_j'\right)|_{A_{J\setminus\{j\}} \times \mathbf{s}_{-J}\left(\boxed{t_J}\right)}$.

Since $\mathbf{s}_{-j}\left(\boxed{t_J}\right) \subseteq \mathcal{D}^k_{-j}$ and $\left(A_{J\backslash\{j\}} \times \mathbf{s}_{-J}\left(\boxed{t_J}\right)\right) \subseteq \mathcal{D}^{k+1}_{-j}$, there exists a CPS $\mu| \in \Delta^*_{\mathcal{D}^k}(S_{-i})$ such that $\mu|_{\mathcal{D}^k_{-j}} = f_j(t'_j)|_{\mathbf{s}_{-j}\left(\boxed{t_J}\right)}$ and $\mu|_{\mathcal{D}^{k+1}_{-j}} = f_j(t'_j)|_{A_{J\backslash\{j\}}\times\mathbf{s}_{-J}\left(\boxed{t_J}\right)}$.
By Profitability in Definition 4, it follows that $U_j(s'_j, f_j(t'_j)|_{\mathbf{s}_{-j}\left(\boxed{t_J}\right)}) < U_j(s^*_j, f_j(t'_j)|_{A_{J\backslash\{j\}}\times\mathbf{s}_{-J}\left(\boxed{t_J}\right)})$ for some $s^*_j \in S_j$.

2. $|J^0| > 1$: Define $\widetilde{A}_J = \mathcal{D}^{k+1}_{J^0} \times \left(\mathbf{s}_j\left(\boxed{t_J}\right) \cap \mathcal{D}^{k+1}_j\right)_{j\in J\backslash J^0}$. By Lemma 2 in Luo and Yang [53], for all $\tau \geq 0$, $BR_i\left(\mathcal{D}^\tau_{-i}\right) \subseteq \mathcal{D}^\tau_i$ $\forall i \in I$. Hence, for all $j \in J^0$, $\widetilde{A}_j \supseteq BR_j\left(\widetilde{A}_{J\backslash\{j\}} \times \mathbf{s}_{-J}\left(\boxed{t_J}\right)\right)$ because $\left(\widetilde{A}_{J\backslash\{j\}} \times \mathbf{s}_{-J}\left(\boxed{t_J}\right)\right) \subseteq \mathcal{D}^{k+1}_{-j}$. Let $A_J \times \mathbf{s}_{-J}\left(\boxed{t_J}\right)$ be the (nonempty) set of strategies surviving iterated elimination of never-best responses for all players in J^0 starting from $\widetilde{A}_J \times \mathbf{s}_{-J}\left(\boxed{t_J}\right)$. Hence, $A_j = BR_j\left(A_{J\backslash\{j\}} \times \mathbf{s}_{-J}\left(\boxed{t_J}\right)\right)$ $\forall j \in J^0$ and $A_j = \widetilde{A}_j$ $\forall j \in J\backslash J^0$. We proceed to show A_J satisfies κ-Profitability and κ-Credibility at (s,t). Note that, $\mathbf{s}_{-j}\left(\boxed{t_J}\right) \cap \mathcal{D}^{k+1}_{-j} = \emptyset$ $\forall j \in J$ because $|J^0| > 1$.
κ*-Profitability:* Since $\mathbf{s}\left(\boxed{t_J}\right) \subseteq \mathcal{D}^k \Rrightarrow \mathcal{D}^{k+1}$ via J and $\mathbf{s}_{-j}\left(\boxed{t_J}\right) \cap \mathcal{D}^{k+1}_{-j} = \emptyset$ $\forall j \in J$, by Profitability in Definition 4, $\forall j \in J$ $\forall s'_j \in \mathbf{s}_j\left(\boxed{t_J}\right)$ and $\forall \mu| \in \Delta^*_{\mathbf{s}\left(\boxed{t_J}\right)}(S_{-j})$, we have $u_j(s'_j, \mu|_{\mathbf{s}_{-j}\left(\boxed{t_J}\right)}) < u_j(s^*_j, \mu|_{\mathcal{D}^{k+1}_{-j}})$ for some $s^*_j \in S_j$. By $\left(A_{J\backslash\{j\}} \times \mathbf{s}_{-J}\left(\boxed{t_J}\right)\right) \subseteq \mathcal{D}^{k+1}_{-j}$, for all $j \in J$ and $(s', t') \in \boxed{t_J}$, $U_j\left(s', f_j(t'_j)|_{\mathbf{s}_{-j}\left(\boxed{t_J}\right)}\right) < U_j\left(s^*_j, f_j(t'_j)|_{A_{J\backslash\{j\}}\times\mathbf{s}_{-J}\left(\boxed{t_J}\right)}\right)$ for some $s^*_j \in S_j$.
κ*-Credibility:* For $j \in J^0$, $\mathbf{s}_j\left(\boxed{t_J}\right) \cap A_j = \emptyset$ and hence $\operatorname{supp} f^0_j(t_j) \cap \left(A_{J\backslash\{j\}} \times \mathbf{s}_{-J}\left(\boxed{t_J}\right)\right) = \emptyset$ because $|J^0| > 1$ and $\operatorname{supp} f^0_j(t_j) \subseteq \mathbf{s}_{-j}\left(\boxed{t_J}\right)$. Note that, $A_j = BR_j\left(A_{J\backslash\{j\}} \times \mathbf{s}_{-J}\left(\boxed{t_J}\right)\right)$ $\forall j \in J^0$ and $A_j \subseteq \mathbf{s}_j\left(\boxed{t_J}\right)$ $\forall j \in J\backslash J^0$. Thus, $A_J \times \mathbf{s}_{-J}\left(\boxed{t_J}\right)$ is credible from $\mathbf{s}\left(\boxed{t_J}\right)$ via J in Definition 4. Since coalition J is BRCD at t, κ-Credibility is satisfied.

3. $|J^0| = 1$ (with $J^0 = \{j^0\}$): Let $d_{j^0} \in BR_{j^0}\left(f_{j^0}\left(t'_{j^0}\right)|_{\left(\mathbf{s}_j\left(\boxed{t_J}\right)\cap\mathcal{D}^{k+1}_j\right)_{j\in J\backslash J^0}\times\mathbf{s}_{-J}\left(\boxed{t_J}\right)}\right)$

for some $t'_{j^0} \in \mathbf{t}_{j^0}\left(\boxed{t_J}\right)$. Similarly, by Lemma 2 in Luo and Yang [53], $d_{j^0} \in \mathcal{D}_{j^0}^{k+1}$. Define $A_J = \{d_{j^0}\} \times \left(\mathbf{s}_j\left(\boxed{t_J}\right) \cap \mathcal{D}_j^{k+1}\right)_{j \in J \setminus J^0}$. We proceed to show A_J satisfies κ-Profitability and κ-Credibility at (s,t).

κ-Profitability: For all $j \in J$, if $j \neq j^0$ then $\mathbf{s}_{-j}\left(\boxed{t_J}\right) \cap \mathcal{D}_{-j}^{k+1} = \emptyset$ and hence it is similar to κ-Profitability in Case 2; if $j = j^0$, then $\mathbf{s}_{-j}\left(\boxed{t_J}\right) \cap \mathcal{D}_{-j}^{k+1} \neq \emptyset$ and hence it is similar to κ-Profitability in Case 1.

κ-Credibility: By construction, $d_{j^0} \in BR_{j^0}\left(f_{j^0}\left(t'_{j^0}\right)|_{A_{J \setminus J^0} \times \mathbf{s}_{-J}\left(\boxed{t_J}\right)}\right)$ for some $t'_{j^0} \in \mathbf{t}_{j^0}\left(\boxed{t_J}\right)$, and hence κ-Credibility is satisfied.

That is, we can always find a $A_J \subseteq \mathcal{D}_J^{k+1}$ satisfying κ-Profitability and κ-Credibility in Definition 2 at (s,t). Thus, $(s,t) \in \mathring{\Re}$ implies $s_J \in A_J \subseteq \mathcal{D}_J^{k+1}$. A contradiction.

The road map for the proof of Proposition 4(b). We construct a type structure $\mathcal{T}(\mathcal{G}) = \left(I, (S_i, u_i)_{i \in I}, (T_i, \beta_i)_{i \in I}\right)$ such that, for every type profile $t \in T$ and coalition J, the minimal state subspace $\boxed{t_J} = \mathcal{R}^* \times T$, where $\mathcal{R}^*$ is the set of the c-rationalizable strategy profiles in game $\mathcal{G}$. More specifically, we construct, in two steps, a finite type space T_i by identifying each $t_i \in T_i$ with a "coherent" infinite hierarchies of CPS beliefs (where an arbitrary order belief in the belief hierarchy specifies i's CPS belief about not only the opponents' lower-order beliefs and using strategies, but also about his own lower-order beliefs and using strategy).

First, we construct a finite set of i's first order CPS beliefs T_i^1, which is comprised of three kinds of first order CPS beliefs in $\Delta^*(S)$: (1) the first order "anchoring" belief $\widehat{\nu}^1|$ such that the primary probabilistic belief has a uniform distribution on $\mathcal{R}^*$, (2) i's first order blocking beliefs for the purpose of the "profitability" requirement, and (3) i's first order BRCD beliefs for the purpose of the "credibility" requirement. The third kind of beliefs purports to accommodate the "belief-richness" condition in an analytical framework. In the construction of T_i^1, $r_i \in \mathcal{R}_i^*$ can be supported as a best response by some $\nu^1|$ in T_i^1.

Second, we construct the desirable types by specifying higher order beliefs in T_i (where T_i has exactly the same cardinality of T_i^1). As usual, the "coherency" criterion is respected for the infinite hierarchies of CPS beliefs. The novelty is that, we require (i) the second order "anchoring" belief $\widehat{\nu}^2| \in \Delta^*(S \times T^1)$ to assign a uniform distribution on $\mathcal{R}^* \times T^1$ (in terms of the primary primary probabilistic belief), and (ii) i's other second order belief $\nu^2| \neq \widehat{\nu}^2|$ to assign (marginal) probability one to the "anchoring" belief $\widehat{\nu}^1|$. Higher order beliefs are constructed in a similar way.

We finally verify $\mathcal{R}^* = \text{proj}_S \left(\overset{\circ}{\Re} \cap CB\overset{\circ}{\Re} \right)$ in this constructed type structure $\mathcal{T}(\mathcal{G})$.

Proof of Proposition 4(b). We construct a finite type structure $\mathcal{T}(\mathcal{G}) = \left(I, (S_i, u_i)_{i\in I}, (T_i, \beta_i)_{i\in I}\right)$ such that, for every type profile $t \in T$ and every coalition $J \subseteq I$, $\boxed{t_J} = \mathcal{R}^* \times T$. For any given player $i \in I$, we construct a finite type space T_i by identifying each $t_i \in T_i$ with a "coherent" infinite hierarchies of CPS beliefs in two steps.

Step 1. Construct the set of i's 1-order belief T_i^1, which is comprised of three kinds of first order CPS beliefs in $\Delta^*(S)$ as follows.

(1) i's first order "anchoring" belief. Define i's first order "anchoring" belief $\widehat{\nu}^1| \in \Delta^*(S)$ such that $\widehat{\nu}^1|_S$ is a uniform distribution on $\mathcal{R}^*$.

(2) i's first order blocking beliefs. Let $A = A_J \times \mathcal{R}^*_{-J}$ be a deviation from $\mathcal{R}^*$ via coalition J. If $r_i \in \mathcal{R}_i^*$ and there exists $\mu| \in \Delta^*_{\mathcal{R}^*}(S_{-i})$ such that (i) $r_i \notin A_i$ or $\mu|_{\mathcal{R}^*_{-i}} \neq \mu|_{A_{-i}}$ and (ii) $U_i(r_i, \mu|_{\mathcal{R}^*_{-i}}) \geq U_i(s_i, \mu|_{A_{-i}})$ $\forall s_i \in S_i$, then we pick one such CPS $\mu|$ for r_i and A_J. For this pair (r_i, A_J), define i's first order blocking belief $\nu^1| \in \Delta^*_{\mathcal{R}^*}(S)$ such that $\text{marg}_{S_{-i}} \nu^1| = \mu|$. (In the case of $J = \{i\}$, if $r_i \in \mathcal{R}_i^* \backslash A_i$, there is i's first order blocking belief $\nu^1|$ such that $r_i \in BR_i\left(\text{marg}_{S_{-i}} \nu^1|_s\right)$, because $\mathcal{R}_i^* \subseteq BR_i\left(\mathcal{R}^*_{-i}\right)$.)

(3) i's first order BRCD beliefs. Let $A = A_J \times \mathcal{R}^*_{-J}$ be a deviation from $\mathcal{R}^*$ via coalition J. If $a_i \in A_i \backslash \mathcal{R}_i^*$ and there exists $\mu| \in \Delta^*_{\mathcal{R}^*}(S_{-i})$ such that $a_i \in BR_i\left(\mu|_{A_{-i}}\right)$, then we pick one such CPS $\mu|$ for a_i and A_J. For this pair (a_i, A_J), define i's first order BRCD belief $\nu^1| \in \Delta^*_{\mathcal{R}^*}(S)$ such that $\text{marg}_{S_{-i}} \nu^1| = \mu|$.

Let

$$T^1 \equiv \times_{i\in I} T_i^1 \text{ and } \Omega^1 \equiv S \times T^1.$$

Step 2. Construct the set of i's higher-order "coherent" CPS belief hierarchies T_i^{k+1} for all $k \geqslant 1$. Assume inductively that T_i^k, T^k and Ω^k are defined. (We decree $\Omega^0 \equiv S$.) For $\widehat{t}_i^k = \left(\widehat{\nu}^\ell|\right)_{\ell=1}^k$ in T_i^k, define i's $(k+1)$-th order belief $\varphi\left(\widehat{t}_i^k\right) = \widehat{\nu}^{k+1}| \in \Delta^*\left(\Omega^k\right)$ such that

1. $\widehat{\nu}^{k+1}| \in \Delta^*_{\mathcal{R}^* \times T^k}\left(\Omega^k\right)$;
2. [Coherence] $\text{marg}_{\Omega^{k-1}} \widehat{\nu}^{k+1}| = \widehat{\nu}^k|$;
3. $\widehat{\nu}^{k+1}|_{\Omega^k}$ is a uniform distribution on $\mathcal{R}^* \times T^k$.

For $t_i^k = \left(\nu^\ell|\right)_{\ell=1}^k$ in $T_i^k \backslash \left\{\widehat{t}_i^k\right\}$, define i's $(k+1)$-th order belief $\varphi\left(t_i^k\right) = \nu^{k+1}| \in \Delta^*\left(\Omega^k\right)$ such that

1. $\nu^{k+1}| \in \Delta^*_{\mathcal{R}^* \times T^k}\left(\Omega^k\right)$;

2. [Coherence] $\operatorname{marg}_{\Omega^{k-1}} \nu^{k+1}| = \nu^k|$;

3. $\operatorname{marg}_{T_i^k} \nu^{k+1}|_{\Omega^k} \left(\widehat{t}_i^k\right) = 1$.

Let

$$T_i^{k+1} \equiv \left\{ t_i^{k+1} = \left(\nu^\ell|\right)_{\ell=1}^{k+1} : t_i^k = \left(\nu^\ell|\right)_{\ell=1}^{k} \in T_i^k \text{ and } \nu^{k+1}| = \varphi\left(t_i^k\right) \right\}$$

denote the set of "i's $(k+1)$-order "coherent" CPS belief hierarchies." Let

$$T^{k+1} \equiv \times_{i \in I} T_i^{k+1} \text{ and } \Omega^{k+1} \equiv S \times T^{k+1}.$$

Finally, let

$$T_i \equiv \left\{ t_i = \left(\nu^\ell|\right)_{\ell=1}^{\infty} : \forall k \geq 1,\ t_i^k = \left(\nu^\ell|\right)_{\ell=1}^{k} \in T_i^k \text{ and } \nu^{k+1}| = \varphi\left(t_i^k\right) \right\}$$

be the set of infinite hierarchies of i's "coherent" CPS beliefs. By our construction, T_i is finite and has the same cardinality of T_i^1. For each $t_i = \left(\nu^\ell|\right)_{\ell=1}^{\infty} \in T_i$, there is a CPS belief $\nu| \in \Delta^*(S \times T)$ such that $\operatorname{marg}_{\Omega^{k-1}} \nu| = \nu^k|$ for all $k \geq 1$ (cf. Battigalli and Siniscalchi [15, Proposition 1]). Define a (continuous) mapping β_i from T_i to $\Delta^*(S \times T)$ by letting $\beta_i(t_i) \equiv \nu|$.

To accomplish our proof, it remains to verify that the constructed type structure $\mathcal{T}(\mathcal{G})$ satisfies BRCD and $\mathcal{R}^* = \operatorname{proj}_S \left(\mathring{\Re} \cap CB\mathring{\Re} \right)$. We proceed to show the minimal state subspace $\boxed{t_J} = \mathcal{R}^* \times T$ for all $t \in T$ and coalition J. Let $j \in J$. If $t_j = \widehat{t}_j$ is j's "anchoring type" in T_j, then by the construction of $\widehat{t}_j$, $\beta_j^0(t_j)$ is a uniform distribution on $\mathcal{R}^* \times T$ and hence, $\boxed{t_J} = \mathcal{R}^* \times T$. If $t_j \neq \widehat{t}_j$, then by the construction of t_j, $\operatorname{marg}_{T_j} \beta_j^0(t_j)$ assigns probability one to $\widehat{t}_j$. Therefore, $\boxed{t_J} = \mathcal{R}^* \times T \ \forall t \in T$.

To see BRCD, we need to show all credible deviations from $\mathcal{R}^*$ are κ-credible in $\mathcal{T}(\mathcal{G})$. Consider a type profile $t \in T$ and a coalition $J \subseteq I$. Let $A = A_J \times \mathcal{R}_{-J}^*$ be a credible deviation from $\mathcal{R}^* = \mathbf{s}\left(\boxed{t_J}\right)$ via J. For each $j \in J$ and $a_j \in A_j \backslash \mathcal{R}_j^*$, by our construction in Step 1(3), there exists j's first order BRCD belief $\nu^1| \in T_j^1$ such that a_j is a best response to $\operatorname{marg}_{S_{-j}} \nu^1|_A = \operatorname{marg}_{S_{-j}} \nu^1|_{A_J \times \mathbf{s}_{-J}\left(\boxed{t_J}\right)}$. By the construction of T_j, there is $t_j' \in T_j = \mathbf{t}_j\left(\boxed{t_J}\right)$ such that $f_j\left(t_j'\right) = \operatorname{marg}_{S_{-j}} \nu^1|$. That is, BRCD is satisfied.

To see $\mathcal{R}^* = \operatorname{proj}_S \left(\mathring{\Re} \cap CB\mathring{\Re} \right)$, we show $\mathring{\Re} = \mathcal{R}^* \times T$; i.e., every coalition J is Bayesian rational at every state $(s,t) \in \mathcal{R}^* \times T$. Suppose, in negation, that some coalition J is not Bayesian rational at state (s,t). Then, there is a (nonempty) product subset $A_J \subseteq S_J$

such that $s_J \notin A_J$ and, moreover, κ-Profitability and κ-Credibility hold in Definition 2. Clearly, $A_J \times \mathcal{R}^*_{-J}$ is a credible deviation from $\mathcal{R}^*$ because κ-Credibility implies Credibility in Definition 4. By our construction, $A_J \times \mathcal{R}^*_{-J}$ is also a profitable deviation from $\mathcal{R}^*$. (Otherwise, there exist $j \in J$, $r_j \in \mathcal{R}^*_j$ and $\mu| \in \Delta^*_{\mathcal{R}^*}(S_{-j})$ such that (i) $r_j \notin A_j$ or $\mu|_{\mathcal{R}^*_{-j}} \neq \mu|_{A_{-j}}$ and (ii) $U_j(r_j, \mu|_{\mathcal{R}^*_{-j}}) \geq U_j(s'_j, \mu|_{A_{-j}})$ $\forall s'_j \in S_j$. By our construction, there exists $(r, t') \in \boxed{t_J} = \mathcal{R}^* \times T$ such that (i) $r_j \notin A_j$ or $f_j\left(t'_j\right)|_{\mathcal{R}^*_{-j}} \neq f_j\left(t'_j\right)|_{A_{-j}}$ and (ii) $U_j(r_j, f_j\left(t'_j\right)|_{\mathcal{R}^*_{-j}}) \geq U_j(s'_j, f_j\left(t'_j\right)|_{A_{-j}})$ $\forall s'_j \in S_j$. That is, κ-Profitability fails to hold for A_J.) Therefore, $\mathcal{R}^* \not\Rrightarrow A_J \times \mathcal{R}^*_{-J} \neq \mathcal{R}^*$, contradicting the fact that $\mathcal{R}^*$ is a CRS. Thus, $\mathring{\Re} = \mathcal{R}^* \times T$.

By our construction, for every $i \in I$ and $(s,t) \in \mathcal{R}^* \times T$, t_i assigns probability one to $\mathcal{R}^* \times T$ and thus, $(s,t) \in \mathring{\Re} \cap B\left(\mathring{\Re}\right)$. Therefore, $\mathcal{R}^* \times T \subseteq \mathring{\Re} \cap B\left(\mathring{\Re}\right) \subseteq \mathring{\Re} = \mathcal{R}^* \times T$; i.e., $\mathring{\Re} \cap B\left(\mathring{\Re}\right) = \mathcal{R}^* \times T$. Repeat this argument, $\mathring{\Re} \cap B^n\left(\mathring{\Re}\right) = \mathcal{R}^* \times T$ for all $n \geq 1$. Hence, $\left(\mathring{\Re} \cap CB\mathring{\Re}\right) = \mathcal{R}^* \times T$.■

Proof of Corollary 1. (a) Let $S^\infty = \cap_{k \geq 0} S^k$ (with $S^0 = S$) denote the outcome of iterated elimination of strictly dominated strategies (IESDS) in $\mathcal{G}$. Then, $\mathcal{R} = S^\infty$. Apparently, $\mathrm{proj}_S(\Re \cap CB\Re) \subseteq S^0$. We assume $\mathrm{proj}_S(\Re \cap CB\Re) \subseteq S^k$ for some $k \geqslant 0$. We inductively show $\mathrm{proj}_S(\Re \cap CB\Re) \subseteq S^{k+1}$. Let $(s,t) \in \Re \cap CB\Re$ and $i \in I$. By Proposition 1(a), $s_i \in \mathbf{BR}_i(t'_i)$ for some $(s', t') \in \boxed{t_i}$. Because $(s', t') \in \boxed{t_i} \subseteq \boxed{t} \subseteq CB\Re$, $(s', t') \in B_I(\Re \cap CB\Re) \subseteq B_i(\Re \cap CB\Re)$; i.e., $\beta^0_i(t'_i)(\Re \cap CB\Re) = 1$. By the induction hypothesis, $f^0_i(t'_i) \in \triangle S^k_{-i}$. Since $s_i \in \mathbf{BR}_i(t'_i)$, $s_i \in S^{k+1}_i$. Therefore, $\mathrm{proj}_S(\Re \cap CB\Re) \subseteq \times_{i \in I} \mathrm{proj}_{S_i}(\Re \cap CB\Re) \subseteq S^{k+1}$. Thus, $\mathrm{proj}_S(\Re \cap CB\Re) \subseteq \mathcal{R}$.

(b) With the restriction of singleton coalitions, the notion of Bayesian c-rationalizablility is equivalent to the notion of rationalizablility; that is, $\mathcal{R} = \mathcal{R}^*$. By the construction in the proof of Proposition 4(b), we can find a (finite) type structure such that $\mathrm{proj}_S(\Re \cap CB\Re) = \mathcal{R}$. ■

References

[1] D. Acemoglu, G. Egorov, K. Sonin, Coalition formation in non-democracies, RES 75(2008), 987-1009.

[2] G.A. Akerlof, R.E. Kranton, Economics and identity, QJE 115(2000), 715-775.

[3] A. Ambrus, Coalitional rationalizability, QJE 121(2006), 903-929.

[4] A. Ambrus, Theories of coalitional rationality, JET 144(2009), 676-695.

[5] A. Ambrus, R. Argenziano, Asymmetric networks in two-sided markets, AEJ: Microeconomics 1(2009), 17-52.

[6] G.B. Asheim, The Consistent Preferences Approach to Deductive Reasoning in Games, Springer, 2006.

[7] R. Aumann, Acceptable points in general cooperative n-person games, in: A.W. Tucker and R.D. Luce (eds.), Contributions to the Theory of Games IV, Princeton Univ. Press (1959), pp.287-324.

[8] R.J. Aumann, Subjectivity and correlation in randomized strategies, JME 1(1974), 67-96.

[9] R.J. Aumann, Agreeing to disagree, Ann. Stat. 4(1976), 1236-1239.

[10] R.J. Aumann, Correlated equilibrium as an expression of Bayesian rationality, Econometrica 55(1987), 1-18.

[11] R.J. Aumann, Backward induction and common knowledge of rationality, GEB 8(1995), 6-19.

[12] R.J. Aumann, Interactive epistemology I: knowledge, IJGT 28(1999), 263-300.

[13] R.J. Aumann, A. Brandenburger, Epistemic conditions for Nash equilibrium, Econometrica 63(1995),1161-1180.

[14] K. Basu, J.W. Weibull, Strategy subsets closed under rational behavior, Econ. Letters 36(1991), 141-146.

[15] P. Battigalli, M. Siniscalchi, Hierarchies of conditional beliefs and interactive epistemology in dynamic games, JET 88(1999),188-230.

[16] G.S. Becker, K.M. Murphy, A theory of rational addiction, JPE 96(1988), 675-700.

[17] R. Benabou, J. Tirole, Self-confidence and personal motivation, QJE 117(2002), 871-915.

[18] R. Benabou, J. Tirole, Intrinsic and extrinsic motivation, RES 70(2003), 489-520.

[19] B.D. Bernheim, Rationalizable strategic behavior, Econometrica 52(1984), 1007-1028.

[20] B.D. Bernheim, B. Peleg, D.Whinston, Coalition-proof Nash equilibria I. Concepts, JET 42(1987), 1-12.

[21] A. Brandenburger, The power of paradox: some recent developments in interactive epistemology, IJGT 35(2007), 465-492.

[22] A. Brandenburger, D. Dekel, Rationalizability and correlated equilibria, Econometrica 55(1987), 1391-1402.

[23] A. Brandenburger, D. Dekel, Hierarchies of beliefs and common knowledge, JET 59(1993), 189-198.

[24] A. Brandenburger, A. Friedenberg, H.J. Keisler, Admissibility in games, Econometrica 76(2008), 307-352.

[25] S.R. Chant, Z. Ernst, Epistemic conditions for collective action, Mind 117(2008), 549-573.

[26] Y.C. Chen, X. Luo, C. Qu, Rationalizability in general situations, ET 61(2016), 147-167.

[27] M.S.-Y. Chwe, Farsighted coalitional stability, JET 63(1994), 299-325.

[28] M.S.-Y. Chwe, Rational Ritual: Culture, Coordination, and Common Knowledge, Princeton Univ. Press, 2001.

[29] E. Dekel, M. Siniscalchi, Epistemic game theory. In: H.P. Young and S. Zamir (eds.), Handbook of Game Theory with Economic Applications, Volume 4. Elsevier (2015), pp. 619-702.

[30] L. Epstein, Preference, rationalizability and equilibrium, JET 73(1997), 1-29.

[31] R. Fagin, J.Y. Halpern, Y. Moses, M.Y. Vardi, Reasoning About Knowledge, The MIT Press, 1995.

[32] D. Fudenberg, D.K. Levine, A dual-self model of impulse control, AER 96(2006), 1449-1476.

[33] J. Geanakoplos, D. Pearce, E. Stacchetti, Psychological games and sequential rationality, GEB 1(1989) 60-79.

[34] I. Gilboa, Rationality and ascriptive science, Mimeo (1991).

[35] I. Gilboa, F. Maccheroni, M. Marinacci, D. Schmeidler, Objective and subjective rationality in a multiple prior model, Econometrica 78(2010), 755-770.

[36] G. Grandjean, A. Mauleon, V. Vannetelbosch, Strongly rational sets for normal-form games, ET Bulletin 5(2017), 35-46.

[37] J. Greenberg, The Theory of Social Situations, Cambridge Univ. Press, 1990.

[38] F. Gul, W. Pesendorfer, Temptation and self-control, Econometrica 69(2001), 1403-1435.

[39] J. Harsanyi, Games with incomplete information played by 'Bayesian' players, I-III, Manage Sci 14(1967-68),159-182, 320-334, 486-502.

[40] J.Y. Halpern, Y. Moses, Knowledge and common knowledge in a distributed environment, J. ACM 37(1990), 549-587.

[41] K. Hara, E. Ok, G. Riella, Coalitional expected multi-utility theory, Econometrica 87(2019), 933-980.

[42] G. Hardin, The Tragedy of the Commons, Science 162(1968), 1243-1248.

[43] Y. Heller, All-stage strong correlated equilibrium, GEB 69(2010), 184-188.

[44] D. Hume, A Treatise of Human Nature, 1738.

[45] F.A. Hayek, The use of knowledge in society, AER 35(1945), 519-530.

[46] B. Jullien, Competition in multi-sided markets: divide and conquer, AEJ: Microeconomics 4(2011), 186-219.

[47] T. Kobayashi, Equilibrium contracts for syndicates with differential information, Econometrica 48(1980), 1635-1665.

[48] E. Kohlberg, J.-F. Mertens, On the strategic stability of equilibria, Econometrica 54(1986), 1003-1037.

[49] C. Korsgaard, Self-Constitution: Agency, Identity, and Integrity, Oxford Univ. Press, 2009.

[50] A. Lambert-Mogiliansky, S. Zamir, H. Zwirn, Type indeterminacy: a model for the KT (Kahaneman Tversky) man, J. Math. Psychology 53(2009), 349-361.

[51] D. Lewis, Convention: A Philosophical Study, Harvard Univ. Press, 1969.

[52] Q. Liu, G.J. Mailath, A. Postlewaite, L. Samuelson, Stable matching with incomplete information, Econometrica 82(2014), 541-587.

[53] X. Luo, C.C. Yang, Bayesian coalitional rationalizability, JET 144(2009), 248-263.

[54] M. Mariotti, A model of agreements in strategic form games, JET 74(1997), 196-217.

[55] M. Maschler, E. Solan, S. Zamir, Game Theory, Cambridge Univ. Press, 2013.

[56] J.-F. Mertens, S. Zamir, Formulation of Bayesian analysis for games with incomplete information, IJGT 14(1985), 1-29.

[57] R. Myerson, Multistage games with communication, Econometrica 54(1986), 323-358.

[58] R. Myerson, Game Theory: Analysis of Conflict, Harvard Univ. Press, 1991.

[59] J. Newton, Coalitional stochastic stability, GEB 75(2012), 842-854.

[60] J. Newton, Shared intentions: The evolution of collaboration, GEB 104(2017), 517-534.

[61] M. Olson, The Logic of Collective Action: Public Goods and the Theory of Groups, Harvard Univ. Press, 1965.

[62] E. Ostrom, Governing the Commons: The Evolution of Institutions for Collective Action, Cambridge Univ. Press, 1991.

[63] D. Pearce, Rationalizable strategic behavior and the problem of perfection, Econometrica 52(1984), 1029-1051.

[64] A. Perea, Epistemic Game Theory: Reasoning and Choice, Cambridge Univ. Press, 2012.

[65] M. Piccione, A. Rubinstein, On the interpretation of decision problems with imperfect recall, GEB 20(1997), 3-24.

[66] D. Ray, R. Vohra, Equilibrium binding agreements, JET 73(1997), 30-78.

[67] D. Ray, R. Vohra, A theory of endogenous coalition structures, GEB 26(1999), 286-336.

[68] D. Ray, R. Vohra, Coalition formation, in: H.P. Young, S. Zamir (eds), Handbook of Game Theory Vol 4, Elsevier, 2014.

[69] A. Renyi, On a new axiomatic theory of probability, Acta Math. Acad. Sci. Hungaricae 6(1955), 285-335.

[70] A. Rubinstein, The electronic mail game: strategic behavior under “almost common knowledge”, AER 79(1989), 385-391.

[71] D. Samet, Common belief of rationality in games of perfect information, GEB 79(2013) 192-200.

[72] P.A. Samuelson, The pure theory of public expenditure, Review of Economics and Statistics 36(1954), 387-389.

[73] T.C. Schelling, The Strategy of Conflict, Harvard Univ. Press, 1960.

[74] S. Shoemaker, Self-knowledge and ‘inner-sense’, Philosophy and Phenomenological Research 68(1994), 249-314.

[75] H.A. Simon, Models of Man: Social and Rational, Wiley, 1957.

[76] R. Sugden, The logic of team reasoning, Philosophical Explorations 6(2003), 165-181.

[77] T. Tan, S. Werlang, The Bayesian foundations of solution concepts of games, JET 45(1988), 370-391.

[78] R.H. Thaler, C.R. Sunstein, Nudge: Improving Decisions about Health, Wealth, and Happiness, Yale Univ. Press, 2008.

[79] R. Wilson, Information, efficiency and the core of an economy, Econometrica 46(1978), 807-816.

[80] L. Xue, Coalitional stability under perfect foresight, ET 11(1998), 603-627.

[81] S. Zamir, Bayesian Games: Games with Incomplete Information, in: Robert A. Meyers (ed.), Encyclopedia of Complexity and Systems Science, Springer (2009), pp.426-441.

[82] S. Zamir, S. Vassilakis, Common beliefs and common knowledge, JME 22, 495-505.