ARTICLE

Strict robustness to incomplete information

Stephen Morris[1] · Daisuke Oyama[2] · Satoru Takahashi[2,3]

Abstract

We study a strict version of the notion of equilibrium robustness by Kajii and Morris (Econometrica 65:1283–1309, 1997) that allows for a larger class of incomplete information perturbations of a given complete information game, where with high probability, players believe that their payoffs are close to (but may be different from) those of the complete information game. We show that a strict monotone potential maximizer of a complete information game is strictly robust if either the game or the associated strict monotone potential is supermodular, and that the converse also holds in all binary-action supermodular games.

Keywords Incomplete information · Robustness · Monotone potential · Sequential obedience · Information design

✉ Satoru Takahashi
satorut@e.u-tokyo.ac.jp

Stephen Morris
semorris@mit.edu

Daisuke Oyama
oyama@e.u-tokyo.ac.jp

[1] Department of Economics, Massachusetts Institute of Technology, Cambridge, MA, USA

[2] Faculty of Economics, University of Tokyo, Tokyo, Japan

[3] Department of Economics, National University of Singapore, Singapore, Singapore

 Springer

1 Introduction

Equilibrium may be fragile to seemingly small departure from common knowledge (Rubinstein, 1989; Carlsson and van Damme, 1993). Then how can an analyst predict strategic behavior when he is almost certain about players' payoff functions? Kajii and Morris (1997) formalized this question and proposed the notion of robust equilibrium as an equilibrium of a complete information game such that every "nearby" incomplete information game has at least one equilibrium that induces an action distribution close to the original equilibrium. In this definition, a "nearby" incomplete information game refers to a game where with high probability, all players know that their own payoff functions are given by those of the complete information game. Kajii and Morris (1997) showed that if a complete information game has a unique correlated equilibrium or a **p**-dominant equilibrium with $\sum_{i \in I} p_i < 1$, then it is robust. Subsequent papers provided sufficient conditions for robustness in terms of potential maximizer (Ui, 2001) and generalized/monotone potential maximizer (Morris and Ui, 2005); Oyama and Takahashi (2020) showed the converse of the latter result in generic binary-action supermodular games.[1]

In this paper, we propose the notion of strict robustness by adopting a larger class of "nearby" incomplete information games than Kajii and Morris (1997), where with high probability, all players believe that their payoffs are close to, but may be different from, those of the complete information game.[2] This class of incomplete information games arises naturally if the analyst assigns high probability on a given profile of payoff functions, but he allows for the possibility that players may not know their own payoff functions exactly. We then establish strict counterparts of the existing results on robustness (with appropriate modifications whenever necessary). More specifically, we show that if the complete information game has a unique correlated equilibrium, it is the unique strictly robust equilibrium; a strict monotone potential maximizer is strictly robust if either the game or the associated strict monotone potential is supermodular; and finally, a strictly robust equilibrium must be a strict monotone potential maximizer in all binary-action supermodular games.

Our results for strict robustness are closely analogous to those for robustness in the sense of Kajii and Morris (1997). We think that strict robustness may be more natural than Kajii-Morris' robustness. However, our main reason for pursuing the results in this paper is that the strict robustness notion connects more cleanly with other work. Indeed, the present study stemmed from our previous study (Morris et al., 2022) on (smallest equilibrium or full) implementation by information design in games, which built in part on the literature on higher order beliefs in general and incomplete information robustness in particular. The robustness question has a metaphorical interpretation, that behind the scene, there is an "adversarial" information designer that tries to design an information structure such that all the equilibrium outcomes thereof are bounded away from the analyst's prediction about the players' behavior; loosely speaking, the prediction is strictly robust only if the information

---

[1] See Kajii and Morris (2020) for a brief survey of the literature.

[2] Kajii and Morris (1997, Remark 4) discussed a similar strengthening.

designer is unable to design such an information structure. In the Appendix, we formalize this argument and, more generally, document the implications of the results in Morris et al. (2022) to strict robustness. In the context of information design, it is natural to allow the information designer to construct information structures in which players do not necessarily know their own payoffs; thus, our definitions of "nearby" incomplete information games and hence of strict robustness are compatible with information design problems.

2 Framework

2.1 Complete information games

A complete information game consists of a finite set I of players, a finite set A_i of actions for each $i \in I$, and a payoff function $g_i : A \to \mathbb{R}$ for each $i \in I$, where we denote $A = \prod_{i \in I} A_i$ and $A_{-i} = \prod_{j \neq i} A_j$ as usual. For brevity, we suppress I and A, and identify a complete information game with the profile $\mathbf{g} = (g_i)_{i \in I}$ of its payoff functions.

An action distribution $\mu \in \Delta(A)$ is a correlated equilibrium of $\mathbf{g}$ if it satisfies the obedience condition:

$$\sum_{a_{-i} \in A_{-i}} \mu(a_i, a_{-i})(g_i(a_i, a_{-i}) - g_i(a'_i, a_{-i})) \geq 0$$

for all $i \in I$ and $a_i, a'_i \in A_i$.[3]

2.2 Incomplete information games

With the player set I and the action sets $(A_i)_{i \in I}$ fixed, an incomplete information game consists of a countable set Θ of payoff-relevant states, a bounded payoff function $u_i : A \times \Theta \to \mathbb{R}$ for each $i \in I$, a countable set T_i of types for each $i \in I$, and a common prior $P \in \Delta(\Theta \times T)$, where we denote $T = \prod_{i \in I} T_i$ and $T_{-i} = \prod_{j \neq i} T_j$. We refer to an incomplete information game as $(\Theta, \mathbf{u}, T, P)$, where $\mathbf{u} = (u_i)_{i \in I}$. For each $i \in I$ and $t_i \in T_i$, we assume without loss of generality that $P(\Theta \times \{t_i\} \times T_{-i}) > 0$, and define $P(\cdot|t_i) \in \Delta(\Theta \times T_{-i})$ by $P(\theta, t_{-i}|t_i) = P(\theta, t_i, t_{-i})/P(\Theta \times \{t_i\} \times T_{-i})$ for $\theta \in \Theta$ and $t_{-i} \in T_{-i}$.

A (behavioral) strategy for player $i \in I$ is a function $\sigma_i : T_i \to \Delta(A_i)$. We denote by Σ_i the set of strategies for $i \in I$, and write $\Sigma = \prod_{i \in I} \Sigma_i$ and $\Sigma_{-i} = \prod_{j \neq i} \Sigma_j$. A strategy profile $\sigma = (\sigma_i)_{i \in I} \in \Sigma$ induces an action distribution $\sigma_P \in \Delta(A)$ by

$$\sigma_P(a) = \sum_{t \in T} P(\Theta \times \{t\}) \prod_{i \in I} \sigma_i(t_i)(a_i)$$

for $a \in A$.

---

[3] For a finite or countably infinite set X, we write $\Delta(X)$ for the set of probability distributions on X.

A strategy profile $\sigma \in \Sigma$ is a Bayes-Nash equilibrium of $(\Theta, \mathbf{u}, T, P)$ if

$$\sigma_i(t_i)(a_i) > 0 \Rightarrow a_i \in \arg\max_{a_i' \in A_i} \sum_{\theta \in \Theta, t_{-i} \in T_{-i}} P(\theta, t_{-i}|t_i) u_i(a_i', \sigma_{-i}(t_{-i}), \theta)$$

for all $i \in I$, $t_i \in T_i$, and $a_i \in A_i$, where $\sigma_{-i}(t_{-i}) = (\sigma_j(t_j))_{j \neq i}$, and the domain of u_i is extended to mixed action profiles in the usual way. By Kakutani's fixed point theorem (with, e.g., the product topology on Σ), a Bayes-Nash equilibrium always exists.

2.3 Strict robustness

Now we introduce our class of incomplete information perturbations. Given a complete information game $\mathbf{g}$ and an incomplete information game $(\Theta, \mathbf{u}, T, P)$, for $i \in I$ and $\eta \geq 0$, let $T_i^{g_i, \eta}$ denote the set of all types of player i for which the payoff function differs from g_i by at most η in expectation, i.e.,

$$T_i^{g_i, \eta} = \left\{ t_i \in T_i \ \middle| \ \sum_{\theta \in \Theta} P(\{\theta\} \times T_{-i}|t_i) \max_{a \in A} |u_i(a, \theta) - g_i(a)| \leq \eta \right\}.$$

We denote $T^{\mathbf{g}, \eta} = \prod_{i \in I} T_i^{g_i, \eta}$.

Definition 1 For $\epsilon \geq 0$ and $\eta \geq 0$, an incomplete information game $(\Theta, \mathbf{u}, T, P)$ is an (ϵ, η)-*elaboration of* $\mathbf{g}$ if $P(\Theta \times T^{\mathbf{g}, \eta}) \geq 1 - \epsilon$.

Our notion of strict robustness requires an action distribution to be close to some equilibrium behavior in all (ϵ, η)-elaborations for sufficiently small $\epsilon > 0$ and $\eta > 0$.

Definition 2 An action distribution $\mu \in \Delta(A)$ is *strictly robust in* $\mathbf{g}$ if for every $\delta > 0$, there exist $\epsilon > 0$ and $\eta > 0$ such that every (ϵ, η)-elaboration $(\Theta, \mathbf{u}, T, P)$ of $\mathbf{g}$ has a Bayes-Nash equilibrium σ such that $\max_{a \in A} |\sigma_P(a) - \mu(a)| \leq \delta$.

An immediate implication of the definition is that if μ is strictly robust in $\mathbf{g}$, then it must be the action distribution of an essential equilibrium of $\mathbf{g}$ in the sense of Wu and Jiang ([1962](#)) (hence a Nash equilibrium). Indeed, by the definition of strict robustness, it is necessary that for every $\delta > 0$, there exists $\eta > 0$ such that every $(0, \eta)$-elaboration with $|\Theta| = 1$ and $|T_i| = 1$ for all $i \in I$, which is a complete information game in the η-neighborhood of $\mathbf{g}$, has an equilibrium $\alpha \in \prod_{i \in I} \Delta(A_i)$ such that $\max_{a \in A} |\prod_{i \in I} \alpha_i(a_i) - \mu(a)| \leq \delta$, which is precisely the definition of essential equilibrium.

Strict robustness strengthens robustness of Kajii and Morris ([1997](#)) by requiring robustness against a larger class of perturbations than the latter. In our language, Kajii and Morris ([1997](#)) only considered $(\epsilon, 0)$-elaborations, where with probability at least $1 - \epsilon$, every player knows that his own payoffs are given by g_i, whereas this "known own payoffs" constraint is relaxed by η for strict robustness. Formally, we say that an action distribution μ is *KM-robust in* $\mathbf{g}$ if for every $\delta > 0$, there exists $\epsilon > 0$ such that every $(\epsilon, 0)$-elaboration of $\mathbf{g}$ has a Bayes-Nash equilibrium σ such that

$\max_{a \in A} |\sigma_P(a) - \mu(a)| \leq \delta$. A strictly robust equilibrium is KM-robust while the converse does not hold.

Example 1 Consider the trivial game **g** with $g_i(a) = 0$ for all $i \in I$ and $a \in A$ with $|A| \geq 2$. Then all product action distributions are KM-robust, but none of them is strictly robust.

3 Results

3.1 Games with unique correlated equilibria

Kajii and Morris (1997, Proposition 3.2) showed that if a complete information game has a unique correlated equilibrium, then it is the unique KM-robust equilibrium in this game. We show that the unique correlated equilibrium is indeed strictly robust.

Proposition 1 *If* **g** *has a unique correlated equilibrium, then it is the unique strictly robust equilibrium of* **g**.

Proof For any Bayes-Nash equilibrium σ of any (ε, η)-elaboration $(\Theta, \mathbf{u}, T, P)$ of **g**, the induced action distribution σ_P satisfies

$$\sum_{a_{-i} \in A_{-i}} \sigma_P(a_i, a_{-i})(g_i(a_i, a_{-i}) - g_i(a'_i, a_{-i}))$$

$$= \sum_{t \in T} P(\Theta \times \{t\})\sigma_i(t_i)(a_i)(g_i(a_i, \sigma_{-i}(t_{-i})) - g_i(a'_i, \sigma_{-i}(t_{-i})))$$

$$= \sum_{t_i \in T_i^{g_i, \eta}} P(\Theta \times \{t_i\} \times T_{-i})\sigma_i(t_i)(a_i)$$

$$\times \sum_{\theta \in \Theta, t_{-i} \in T_{-i}} P(\theta, t_{-i}|t_i) \begin{pmatrix} (u_i(a_i, \sigma_{-i}(t_{-i}), \theta) - u_i(a'_i, \sigma_{-i}(t_{-i}), \theta)) \\ - (u_i(a_i, \sigma_{-i}(t_{-i}), \theta) - g_i(a_i, \sigma_{-i}(t_{-i}))) \\ + (u_i(a'_i, \sigma_{-i}(t_{-i}), \theta) - g_i(a'_i, \sigma_{-i}(t_{-i}))) \end{pmatrix}$$

$$+ \sum_{t \notin T_i^{g_i, \eta} \times T_{-i}} P(\Theta \times \{t\})\sigma_i(t_i)(a_i)(g_i(a_i, \sigma_{-i}(t_{-i})) - g_i(a'_i, \sigma_{-i}(t_{-i})))$$

$$\geq -2\eta - 2\varepsilon \max_{a \in A} |g_i(a)|$$

for all $i \in I$ and $a_i, a'_i \in A_i$. By passing to a subsequence, σ_P converges to a correlated equilibrium of **g** as $\varepsilon, \eta \to 0$. Thus, if **g** has a unique correlated equilibrium, then it must be strictly robust in **g**. $\square$

3.2 Strict monotone potential games

Morris and Ui (2005, Proposition 2) provided a sufficient condition for KM-robustness in terms of *monotone potential maximizer* (MP maximizer). In this section, we show that a strict version of their condition implies strict robustness.

In what follows, we equip each action set A_i with a linear order $\leq$. For $i \in I$ and $a_i, a_i' \in A_i$ with $a_i \leq a_i'$, let $[a_i, a_i']$ denote the interval $\{a_i'' \in A_i \mid a_i \leq a_i'' \leq a_i'\}$. For $h : A \to \mathbb{R}$, $i \in I$, $\pi_i \in \Delta(A_{-i})$, and $B_i \subset A_i$, let $\mathrm{br}_i^h(\pi_i | B_i)$ denote the set of player i's best responses against π_i when he is restricted to play within B_i and his payoff function is given by h, i.e.,

$$\mathrm{br}_i^h(\pi_i | B_i) = \underset{a_i \in B_i}{\arg\max} \sum_{a_{-i} \in A_{-i}} \pi(a_{-i}) h(a_i, a_{-i}).$$

The definition of MP maximizer of Morris and Ui (2005) (of a simpler form) is as follows[4]:

Definition 3 Action profile $a^* \in A$ is an *MP maximizer of* **g** if there exists a function $v : A \to \mathbb{R}$ with $v(a^*) > v(a)$ for all $a \neq a^*$ such that

$$\max \mathrm{br}_i^v(\pi_i | [a_i^*, \max A_i]) \geq \min \mathrm{br}_i^{g_i}(\pi_i | [a_i^*, \max A_i]),$$
$$\min \mathrm{br}_i^v(\pi_i | [\min A_i, a_i^*]) \leq \max \mathrm{br}_i^{g_i}(\pi_i | [\min A_i, a_i^*])$$

for all $i \in I$ and $\pi_i \in \Delta(A_{-i})$. Such a function v is called a *monotone potential* for a^* in **g**.

We employ the strict version of MP maximizer due to Oyama et al. (2008):

Definition 4 Action profile $a^* \in A$ is a *strict MP maximizer of* **g** if there exists a function $v : A \to \mathbb{R}$ with $v(a^*) > v(a)$ for all $a \neq a^*$ such that

$$\max \mathrm{br}_i^v(\pi_i | [a_i^*, \max A_i]) \geq \max \mathrm{br}_i^{g_i}(\pi_i | [a_i^*, \max A_i]),$$
$$\min \mathrm{br}_i^v(\pi_i | [\min A_i, a_i^*]) \leq \min \mathrm{br}_i^{g_i}(\pi_i | [\min A_i, a_i^*])$$

for all $i \in I$ and $\pi_i \in \Delta(A_{-i})$. Such a function v is called a *strict monotone potential* for a^* in **g**.

In words, a^* is an MP maximizer of **g** if there exists a common payoff function v strictly maximized at a^* such that every player i, if restricted to play above (resp. below) a_i^*, has at least one best response under v that is larger (resp. smaller) than or equal to *some* best response under g_i. In contrast, strict MP maximizer requires that every player i, if restricted to play above (resp. below) a_i^*, have at least one best response under v that bounds *all* best responses under g_i from above (resp. below).

By definition, a strict MP maximizer is an MP maximizer. The converse holds in generic games, but not in general. For example, in the nongeneric game in Example 1, all action profiles are MP maximizers, but none of them is a strict MP maximizer.

A complete information game **g** is supermodular if for any $i \in I$ and $a_i, a_i' \in A_i$ with $a_i > a_i'$, $g_i(a_i, a_{-i}) - g_i(a_i', a_{-i})$ is weakly increasing in a_{-i}; a function $v : A \to \mathbb{R}$ is supermodular if the game with common payoff v is supermodular. Morris and Ui

---

[4] Morris and Ui (2005) introduced a set-valued version of MP maximizer.

(2005, Proposition 2) showed that an MP maximizer is KM-robust if either the complete information game or the monotone potential is supermodular. The following theorem shows that a strict MP maximizer is strictly robust under the same supermodularity assumption as in Morris and Ui (2005).

Theorem 1 *If $a^* \in A$ is a strict MP maximizer of* **g** *with strict monotone potential v, and either* **g** *or v is supermodular, then the degenerate distribution on a^* is strictly robust in* **g**.

The theorem immediately implies that a unique potential maximizer of a (weighted) potential game (Monderer and Shapley, 1996) is strictly robust if the game is supermodular.[5]

In the statement of Theorem 1, "strict MP maximizer" cannot be weakened to "MP maximizer". Indeed, in Example 1, the game is supermodular and all action profiles are MP maximizers, but none of them is strictly robust.

The proof of Theorem 1, which is given in Appendix A.1, proceeds almost in the same way as in Morris and Ui (2005). But in one step of the proof, we use the property of strict MP maximizer that the monotone relationship on best responses extends to η-best responses (Lemma A.1) and absorb uncertainty about own payoffs in $(\varepsilon, \eta/2)$-elaborations.

For $\mathbf{p} = (p_i)_{i \in I} \in [0, 1)^I$, an action profile $a^* \in A$ is a *strictly* **p**-*dominant equilibrium in* **g** if

$$\pi_i(a_i^*) > p_i \Rightarrow \mathrm{br}_i^{g_i}(\pi_i | A_i) = \{a_i\}$$

for all $i \in I$ and $\pi_i \in \Delta(A_{-i})$. Strict **p**-dominance is a natural generalization of the classical condition of risk dominance for 2×2 coordination games to many-player many-action games. If a^* is strictly **p**-dominant in **g** with $\sum_{i \in I} p_i < 1$, then the function $v : A \to \mathbb{R}$ given by

$$v(a) = \begin{cases} 1 - \sum_{i \in I} p_i & \text{if } a = a^*, \\ - \sum_{i \in I : a_i = a_i^*} p_i & \text{otherwise} \end{cases}$$

is a strict monotone potential for a^* in **g**. Moreover, v is supermodular if we reorder actions so that $a_i^* = \max A_i$ for all $i \in I$. Thus, together with Lemma 5.5 in Kajii and Morris (1997), Theorem 1 implies the following, which parallels the corresponding result for KM-robustness by Kajii and Morris (1997, Corollary 5.6).

Corollary 2 *If a^* is strictly* **p**-*dominant in* **g** *with $\sum_{i \in I} p_i < 1$, then the degenerate distribution on a^* is the unique strictly robust equilibrium of* **g**.

---

[5] It is not known whether the supermodularity assumption is indispensable. Takahashi (2020) presented an example where a non-singleton potential maximizer set of a non-supermodular game is not KM-robust, hence not strictly robust.

3.3 Binary-action supermodular games

In this subsection, we restrict our attention to binary-action supermodular (BAS) games: all players have binary action sets $A_i = \{0, 1\}$, and the payoff functions are supermodular, i.e., for each $i \in I$, the payoff difference

$$f_i(a_{-i}) := g_i(1, a_{-i}) - g(0, a_{-i})$$

is weakly increasing in $a_{-i} \in A_{-i}$.

Oyama and Takahashi (2020, Theorem 3) showed that a KM-robust equilibrium is a strict MP maximizer in *generic* BAS games. Here, we show that a strictly robust equilibrium is a strict MP maximizer in *all* BAS games. Thus, together with Theorem 1, we establish the exact equivalence between strict robustness and strict MP maximization in this class of games.

The main tool in the proof is a certain obedience condition: it was implicitly introduced in Oyama and Takahashi (2020) and extended to an incomplete information environment and termed *sequential obedience* in Morris et al. (2022). Let Γ denote the set of all sequences of distinct players (including the null sequence), and for each $i \in I$, let Γ_i denote the set of all sequences in Γ where player i is listed. For each $\gamma \in \Gamma$ and $k = 0, 1$, let $a^k(\gamma) \in A$ denote the action profile such that player i plays action k if and only if i is listed in γ; for each $i \in I$, $\gamma \in \Gamma_i$, and $k = 0, 1$, let $a^k_{-i}(\gamma) \in A_{-i}$ denote the action profile of player i's opponents such that player $j \neq i$ plays action k if and only if j is listed in γ before i (therefore, player j plays action $1 - k$ if and only if either j is not listed in γ or j is listed in γ after i).

Definition 5 Let **g** be a BAS game. Action distribution $\mu \in \Delta(A)$ satisfies *sequential obedience in* **g** if there exists $\rho \in \Delta(\Gamma)$ such that

$$\sum_{\gamma \in \Gamma_i} \rho(\gamma) f_i(a^1_{-i}(\gamma)) \geq 0 \tag{1}$$

for all $i \in I$ and $\mu(a) = \rho(\{\gamma \in \Gamma \mid a^1(\gamma) = a\})$ for all $a \in A$; μ satisfies *reverse sequential obedience in* **g** if there exists $\rho' \in \Delta(\Gamma)$ such that

$$\sum_{\gamma \in \Gamma_i} \rho'(\gamma) f_i(a^0_{-i}(\gamma)) \leq 0 \tag{2}$$

for all $i \in I$ and $\mu(a) = \rho'(\{\gamma \in \Gamma \mid a^0(\gamma) = a\})$ for all $a \in A$.

In any BAS game, there exists at least one (in fact, degenerate) action distribution that satisfies both sequential obedience and reverse sequential obedience (Morris et al., 2022, Proposition B.2). Uniqueness of such an action distribution characterizes existence of a strict MP maximizer.

Proposition 3 *In any BAS game* **g**, a^* *is a strict MP maximizer if and only if the degenerate action distribution on* a^* *is the unique action distribution that satisfies sequential obedience and reverse sequential obedience.*

This follows from Lemmas 2 and A.1 in Oyama and Takahashi (2020) and Proposition B.2 in Morris et al. (2022); see Appendix A.2.

With this characterization, we can show the converse of Theorem 1 for all BAS games.

Theorem 2 *In any BAS game* **g**, *an action distribution is strictly robust if and only if it is degenerate on a strict MP maximizer.*

By Proposition 3, proving the "only if" direction in Theorem 2 reduces to proving that if an action distribution is strictly robust, then it must be the unique action distribution that satisfies sequential obedience and reverse sequential obedience. In Appendix A.3, we prove the latter statement by showing that for any action distribution μ that satisfies sequential obedience and reverse sequential obedience, there exists an (ε, η)-elaboration where a unique Bayes-Nash equilibrium induces an action distribution in a neighborhood of μ: the existence of an action distribution that satisfies sequential obedience and reverse sequential obedience is guaranteed (see Appendix A.2), and if there are multiple such action distributions, then no action distribution is strictly robust.

The arguments in Appendix A.3 are an application of those developed by Morris et al. (2022) in the framework of information design. There, a (finite) state space Θ and a profile $\mathbf{u} = (u_i)_{i \in I}$ of payoff functions $u_i : A \times \Theta \to \mathbb{R}$ are exogenously given and a prior on Θ is fixed. A version of the question studied in Morris et al. (2022) concerns what outcomes can be implemented as a unique Bayes-Nash equilibrium in some information structure (full implementability). Under a dominance state assumption that each action is a dominant action at some state, Morris et al. (2022) showed that an outcome is fully implementable if and almost only if it satisfies strict and incomplete information versions of sequential obedience and reverse sequential obedience. We apply this characterization to the limit case as the prior on Θ becomes degenerate on some state θ^* (limit full implementability), noticing that the implementing information structures are indeed (ε, η)-elaborations of $\mathbf{u}(\cdot, \theta^*)$ in our sense with $\varepsilon, \eta \to 0$.

Oyama and Takahashi (2020) earlier obtained a generic characterization of KM-robustness in BAS games, where they appealed to a genericity assumption and assumed strict inequalities in (1) or (2) in Definition 5 to guarantee strict incentives in the desired $(\varepsilon, 0)$-elaborations. In contrast, we exploit payoff uncertainty allowed in (ε, η)-elaborations to provide strict incentives.

4 Concluding remarks

In this paper, we introduced the notion of strict robustness, which tests the robustness of an action distribution against all (ε, η)-elaborations, where, unlike in $(\varepsilon, 0)$-elaborations used for KM-robustness, the players are only required to believe with high probability that their payoff functions are close in expectation to those of the original complete information game **g**. We showed that a strict MP maximizer of **g** is

strictly robust if either **g** or the associated strict monotone potential is supermodular, and that the converse also holds if **g** is a BAS game, thus establishing the equivalence between strict robustness and strict MP maximization in all BAS games.

Oyama and Takahashi (2023) refined Oyama-Takahashi's (2020) generic characterization of KM-robustness and showed that the equivalence between KM-robustness and MP maximization holds in all BAS games for extreme action profiles **0** and **1**. This equivalence, however, fails for nonextreme action profiles, as the following example shows:

Example 2 Consider the following (nongeneric) three-player BAS game from Oyama and Takahashi (2019, Section A.6):

	0	1			0	1
0	1, 1, 0	0, 0, 0		0	0, 0, 0	0, 0, 0
1	0, 0, 0	1, 1, 0		1	0, 0, 0	2, 2, 0
		0				1

As shown in Oyama and Takahashi (2019, Proposition A.2) the degenerate distribution on action profile (1, 1, 0) is KM-robust. On the other hand, one can verify that (1, 1, 0) is not an MP maximizer.

It is known that in *generic* supermodular games, a KM-robust equilibrium, if any, is a noise-independent global game selection (Basteck et al., 2013; Oury and Tercieux, 2007). Given that our perturbations allow for uncertainty about own payoffs as in global games (modulo the apparent difference between the formulations with discrete and continuous states), one may expect that strict robustness implies noise-independent selection in *all* supermodular games. Oyama and Takahashi (2023) show that in the class of all BAS games, an action profile is a noise-independent global game selection if and only if it is a strict MP maximizer; combined with the result of the present paper, it implies that noise-independent global game selection is equivalent to strict robustness in these games.

Finally, it is left as an open problem to characterize (strict) robustness in many-action games beyond BAS games. Note that it is known that the equivalence between robustness and noise-independent global game selection fails in generic symmetric two-player three-action supermodular games (Basteck and Daniëls, 2011; Oyama and Takahashi, 2011); it has not been proved or disproved whether the equivalence between (strict) robustness and (strict) MP maximization holds even in this class of games.

Appendix

A.1. Proof of Theorem 1

For $i \in I$, $\pi_i \in \Delta(A_{-i})$, $\eta \geq 0$, and $B_i \subset A_i$, let $\mathrm{br}_i^{g_i, \eta}(\pi_i | B_i)$ denote the set of player i's η-best responses in B_i against π_i:

$$\mathrm{br}_i^{g_i,\eta}(\pi_i|B_i) = \left\{ a_i \in B_i \mid \sum_{a_{-i} \in A_{-i}} \pi_i(a_{-i})(g_i(a_i, a_{-i}) - g_i(a_i', a_{-i})) \geq -\eta \text{ for all } a_i' \in B_i \right\}.$$

The following lemma extends the monotone relationship on best responses in the definition of strict MP maximizer to η-best responses.[6]

Lemma A.1 *Action profile $a^* \in A$ is a strict MP maximizer of $\mathbf{g}$ with strict monotone potential v if and only if there exist $\eta > 0$ and a function $\tilde{v} : A \to \mathbb{R}$ with $\tilde{v}(a^*) > \tilde{v}(a)$ for all $a \neq a^*$ such that*

$$\min \mathrm{br}_i^{\tilde{v}}(\pi_i|[a_i^*, \max A_i]) \geq \max \mathrm{br}_i^{g_i,\eta}(\pi_i|[a_i^*, \max A_i]),$$
$$\max \mathrm{br}_i^{\tilde{v}}(\pi_i|[\min A_i, a_i^*]) \leq \min \mathrm{br}_i^{g_i,\eta}(\pi_i|[\min A_i, a_i^*])$$

for all $i \in I$ and $\pi_i \in \Delta(A_{-i})$. Moreover, if v is supermodular, then $\tilde{v}$ can be taken to be supermodular.

Proof The "if" direction is obvious. For the "only if" direction, without loss of generality, we assume $A_i \subset \mathbb{R}$ for any $i \in I$. Let $\tilde{v} : A \to \mathbb{R}$ be given by

$$\tilde{v}(a) = v(a) + \zeta \sum_{i \in I} |a_i - a_i^*|$$

for each $a \in A$, where $0 < \zeta < (v(a^*) - \max_{a \neq a^*} v(a)) / \sum_{i \in I}(\max A_i - \min A_i)$. Then we have $\tilde{v}(a^*) > \tilde{v}(a)$ for all $a \neq a^*$, and

$$\min \mathrm{br}_i^{\tilde{v}}(\pi_i|[a_i^*, \max A_i]) \geq \max \mathrm{br}_i^{v}(\pi_i|[a_i^*, \max A_i]), \tag{A.1}$$

$$\max \mathrm{br}_i^{\tilde{v}}(\pi_i|[\min A_i, a_i^*]) \leq \min \mathrm{br}_i^{v}(\pi_i|[\min A_i, a_i^*]) \tag{A.2}$$

for all $i \in I$ and $\pi_i \in \Delta(A_{-i})$. Moreover, if v is supermodular, then $\tilde{v}$ is also supermodular.

For each $i \in I$ and $\pi_i \in \Delta(A_{-i})$, since A_i is finite, there exist a neighborhood $O(\pi_i)$ of π_i and $\eta_i(\pi_i) > 0$ such that

$$\mathrm{br}_i^{\tilde{v}}(\pi_i|B_i) \supset \mathrm{br}_i^{\tilde{v}}(\pi_i'|B_i), \tag{A.3}$$

$$\mathrm{br}_i^{g_i}(\pi_i|B_i) \supset \mathrm{br}_i^{g_i,\eta_i(\pi_i)}(\pi_i'|B_i) \tag{A.4}$$

for all $\pi_i' \in O(\pi_i)$ and $B_i \subset A_i$. Since $\{O(\pi_i)\}_{\pi_i \in \Delta(A_{-i})}$ is an open cover of compact set $\Delta(A_{-i})$, it has a finite subcover $\{O(\pi_i)\}_{\pi_i \in \Pi_i}$.

Let $\eta = \min_{i \in I} \min_{\pi_i \in \Pi_i} \eta_i(\pi_i) > 0$. Then for any $i \in I$ and $\pi_i' \in \Delta(A_{-i})$, since there exists $\pi_i \in \Pi_i$ such that $O(\pi_i) \ni \pi_i'$, we have

---

[6] Oyama and Tercieux (2009, Lemma 2.2) showed a similar result for a set-valued version of strict MP maximizer.

$$\min \operatorname{br}_i^{\tilde{v}}(\pi_i'|[a_i^*, \max A_i]) \geq \min \operatorname{br}_i^{\tilde{v}}(\pi_i|[a_i^*, \max A_i])$$
$$\geq \max \operatorname{br}_i^{v}(\pi_i|[a_i^*, \max A_i])$$
$$\geq \max \operatorname{br}_i^{g_i}(\pi_i|[a_i^*, \max A_i])$$
$$\geq \max \operatorname{br}_i^{g_i,\eta}(\pi_i'|[a_i^*, \max A_i]),$$

where the first inequality follows from (A.3), the second from (A.1), the third since v is a strict monotone potential for a^*, and the fourth from (A.4), and

$$\max \operatorname{br}_i^{\tilde{v}}(\pi_i'|[\min A_i, a_i^*]) \leq \max \operatorname{br}_i^{\tilde{v}}(\pi_i|[\min A_i, a_i^*])$$
$$\leq \min \operatorname{br}_i^{v}(\pi_i|[\min A_i, a_i^*])$$
$$\leq \min \operatorname{br}_i^{g_i}(\pi_i|[\min A_i, a_i^*])$$
$$\leq \min \operatorname{br}_i^{g_i,\eta}(\pi_i'|[\min A_i, a_i^*]),$$

where the first inequality follows from (A.3), the second from (A.2), the third since v is a strict monotone potential for a^*, and the fourth from (A.4). $\qquad\square$

Now we prove Theorem 1. Let $\tilde{v} : A \to \mathbb{R}$ and $\eta > 0$ be as in Lemma A.1. Let $v^* = \tilde{v}(a^*)$, $\bar{v} = \max_{a \neq a^*} \tilde{v}(a)$, and $\underline{v} = \min_a \tilde{v}(a)$.

For each $\delta > 0$, we let

$$\varepsilon = \frac{v^* - \bar{v}}{v^* - \underline{v}} \frac{\delta}{2} > 0, \tag{A.5}$$

and consider any $(\varepsilon, \eta/2)$-elaboration $(\Theta, \mathbf{u}, T, P)$ of $\mathbf{g}$.

For each $i \in I$, let $S_i = A_i^{T_i}$ denote the set of all pure strategies $s_i : T_i \to A_i$ endowed with the product topology, and

$$S_i^+ = \{s_i \in S_i \mid s_i(t_i) \geq a_i^* \text{ for all } t_i \in T_i^{g_i,\eta/2}, s_i(t_i) = \max A_i \text{ for all } t_i \notin T_i^{g_i,\eta/2}\},$$
$$S_i^- = \{s_i \in S_i \mid s_i(t_i) \leq a_i^* \text{ for all } t_i \in T_i^{g_i,\eta/2}, s_i(t_i) = \min A_i \text{ for all } t_i \notin T_i^{g_i,\eta/2}\}.$$

We write $S = \prod_i S_i$, $S^+ = \prod_i S_i^+$, and $S^- = \prod_i S_i^-$.

Define $V : S \to \mathbb{R}$ by

$$V(s) = \sum_{t \in T} P(\Theta \times \{t\}) \tilde{v}(s(t))$$

for $s \in S$. Since S^+ and S^- are nonempty and compact subsets of S, and V is continuous, $\arg\max_{s \in S^+} V(s)$ and $\arg\max_{s \in S^-} V(s)$ are nonempty. Let s^+ and s^- denote arbitrary maximizers of V in S^+ and in S^-, respectively. Note that s^+ and s^- are Bayes-Nash equilibria when players have a common payoff function $\tilde{v}$ and are restricted to play strategies in S^+ and S^-, respectively, i.e.,

$$s_i^+(t_i) \in \operatorname{br}_i^{\tilde{v}}(s_{-i,P(\cdot|t_i)}^+|[a_i^*, \max A_i]), \tag{A.6}$$

$$s_i^-(t_i) \in \mathrm{br}_i^{\bar{v}}(s_{-i,P(\cdot|t_i)}^-|[\min A_i, a_i^*]) \tag{A.7}$$

for all $i \in I$ and $t_i \in T_i^{g_i,\eta/2}$, where $\sigma_{-i,P(\cdot|t_i)}$ denotes the distribution over A_{-i} induced by σ_{-i} under $P(\cdot|t_i)$, i.e., $\sigma_{-i,P(\cdot|t_i)}(a_{-i}) = \sum_{t_{-i} \in T_{-i}} P(\Theta \times \{t_{-i}\}|t_i) \prod_{j \neq i} \sigma_j(t_j)(a_j)$ for $a_{-i} \in A_{-i}$.

Let $\hat{s}^+$ denote the strategy profile given by $\hat{s}_i^+(t_i) = a_i^*$ for $i \in I$ and $t_i \in T_i^{g_i,\eta/2}$ and $\hat{s}_i^+(t_i) = \max A_i$ otherwise. Since $\hat{s}^+ \in S^+$ and $P(\Theta \times T^{\mathbf{g},\eta/2}) \geq 1 - \varepsilon$, we have

$$s_P^+(a^*)v^* + (1 - s_P^+(a^*))\bar{v} \geq V(s^+) \geq V(\hat{s}^+) \geq (1 - \varepsilon)v^* + \varepsilon\underline{v},$$

and hence by (A.5),

$$s_P^+(a^*) \geq 1 - \frac{v^* - \underline{v}}{v^* - \bar{v}}\varepsilon = 1 - \frac{\delta}{2}.$$

Similarly, we have $s_P^-(a^*) \geq 1 - \delta/2$.

Let

$$\Sigma_i^0 = \{\sigma_i \in \Sigma_i \mid \sigma_i(t_i) \in \Delta([s_i^-(t_i), s_i^+(t_i)]) \text{ for all } t_i \in T_i^{g_i,\eta/2}\}.$$

We write $\Sigma^0 = \prod_i \Sigma_i^0$ and $\Sigma_{-i}^0 = \prod_{j \neq i} \Sigma_j^0$.

In what follows, we will show that for any $i \in I$ and $\sigma_{-i} \in \Sigma_{-i}^0$, player i's best responses to σ_{-i} in game $(\Theta, \mathbf{u}, T, P)$ belong to Σ_i^0. Then, it follows from Kakutani's fixed point theorem that $(\Theta, \mathbf{u}, T, P)$ has a Bayes-Nash equilibrium in Σ^0, which induces a^* with probability at least $1 - \delta$.

Take any $i \in I$, $t_i \in T_i^{g_i,\eta/2}$, and $\sigma_{-i} \in \Sigma_{-i}^0$. Since

$$\left|\sum_{\theta,t_{-i}} P(\theta, t_{-i}|t_i)u_i(a_i, \sigma_{-i}(t_{-i}), \theta) - g_i(a_i, \sigma_{-i,P(\cdot|t_i)})\right|$$
$$\leq \sum_{\theta,t_{-i}} P(\theta, t_{-i}|t_i)|u_i(a_i, \sigma_{-i}(t_{-i}), \theta) - g_i(a_i, \sigma_{-i}(t_{-i}))|$$
$$\leq \eta/2$$

for each $a_i \in A_i$, we have

$$\arg\max_{a_i \in B_i} \sum_{\theta,t_{-i}} P(\theta, t_{-i}|t_i)u_i(a_i, \sigma_{-i}(t_{-i}), \theta) \subset \mathrm{br}_i^{g_i,\eta}(\sigma_{-i,P(\cdot|t_i)}|B_i) \tag{A.8}$$

for all $B_i \subset A_i$.

First, suppose that $\mathbf{g}$ is supermodular. Then by (A.8), the supermodularity of g_i, Lemma A.1, and (A.6) or (A.7), we have

$$\max \operatorname*{arg\,max}_{a_i \in [a_i^*, \max A_i]} \sum_{\theta, t_{-i}} P(\theta, t_{-i}|t_i) u_i(a_i, \sigma_{-i}(t_{-i}), \theta)$$

$$\leq \max \operatorname{br}_i^{g_i, \eta}(\sigma_{-i, P(\cdot|t_i)}|[a_i^*, \max A_i])$$

$$\leq \max \operatorname{br}_i^{g_i, \eta}(s_{-i, P(\cdot|t_i)}^+|[a_i^*, \max A_i])$$

$$\leq \min \operatorname{br}_i^{\tilde{v}}(s_{-i, P(\cdot|t_i)}^+|[a_i^*, \max A_i])$$

$$\leq s_i^+(t_i)$$

and

$$\min \operatorname*{arg\,max}_{a_i \in [\min A_i, a_i^*]} \sum_{\theta, t_{-i}} P(\theta, t_{-i}|t_i) u_i(a_i, \sigma_{-i}(t_{-i}), \theta)$$

$$\geq \min \operatorname{br}_i^{g_i, \eta}(\sigma_{-i, P(\cdot|t_i)}|[\min A_i, a_i^*])$$

$$\geq \min \operatorname{br}_i^{g_i, \eta}(s_{-i, P(\cdot|t_i)}^-|[\min A_i, a_i^*])$$

$$\geq \max \operatorname{br}_i^{\tilde{v}}(s_{-i, P(\cdot|t_i)}^-|[\min A_i, a_i^*])$$

$$\geq s_i^-(t_i).$$

Thus type t_i's best responses belong to $\Delta([s_i^-(t_i), s_i^+(t_i)])$.

Second, suppose that v is supermodular. Then by (A.8), Lemma A.1, the super-modularity of $\tilde{v}$, and (A.6) or (A.7), we have

$$\max \operatorname*{arg\,max}_{a_i \in [a_i^*, \max A_i]} \sum_{\theta, t_{-i}} P(\theta, t_{-i}|t_i) u_i(a_i, \sigma_{-i}(t_{-i}), \theta)$$

$$\leq \max \operatorname{br}_i^{g_i, \eta}(\sigma_{-i, P(\cdot|t_i)}|[a_i^*, \max A_i])$$

$$\leq \min \operatorname{br}_i^{\tilde{v}}(\sigma_{-i, P(\cdot|t_i)}|[a_i^*, \max A_i])$$

$$\leq \min \operatorname{br}_i^{\tilde{v}}(s_{-i, P(\cdot|t_i)}^+|[a_i^*, \max A_i])$$

$$\leq s_i^+(t_i)$$

and

$$\min \operatorname*{arg\,max}_{a_i \in [\min A_i, a_i^*]} \sum_{\theta, t_{-i}} P(\theta, t_{-i}|t_i) u_i(a_i, \sigma_{-i}(t_{-i}), \theta)$$

$$\geq \min \operatorname{br}_i^{g_i, \eta}(\sigma_{-i, P(\cdot|t_i)}|[\min A_i, a_i^*])$$

$$\geq \max \operatorname{br}_i^{\tilde{v}}(\sigma_{-i, P(\cdot|t_i)}|[\min A_i, a_i^*])$$

$$\geq \max \operatorname{br}_i^{\tilde{v}}(s_{-i, P(\cdot|t_i)}^-|[\min A_i, a_i^*])$$

$$\geq s_i^-(t_i).$$

Thus type t_i's best responses belong to $\Delta([s_i^-(t_i), s_i^+(t_i)])$.

A.2. Strict monotone potential and (reverse) sequential obedience in BAS games

Fix a BAS game **g**. The following lemma characterizes strict MP maximizer by the existence of an inequality form of weighted potential.[7]

Lemma A.2 *Action profile a^* is a strict MP maximizer of* **g** *if and only if there exist $v : A \to \mathbb{R}$ with $v(a^*) > v(a)$ for all $a \neq a^*$ and $(\lambda_i)_{i \in I}$ with $\lambda_i > 0$ for all $i \in I$ such that*

$$
\begin{aligned}
a_i^* = 0 &\implies v(1, a_{-i}) - v(0, a_{-i}) \geq \lambda_i f_i(a_{-i}), \\
a_i^* = 1 &\implies v(1, a_{-i}) - v(0, a_{-i}) \leq \lambda_i f_i(a_{-i})
\end{aligned}
$$

for all $i \in I$ and $a_{-i} \in A_{-i}$.

Proof The "if" direction is straightforward. For the "only if" direction, suppose that v is a strict monotone potential for a^* in **g**. Pick any $i \in I$, and assume $a_i^* = 0$ without loss of generality. Then there exists no $(\pi_i, \delta_i) \in \mathbb{R}_+^{|A_{-i}|+1}$ such that

$$
\sum_{a_{-i} \in A_{-i}} \pi_i(a_{-i}) f_i(a_{-i}) \geq 0,
$$

$$
- \sum_{a_{-i} \in A_{-i}} \pi_i(a_{-i})(v(1, a_{-i}) - v(0, a_{-i})) - \delta_i \geq 0,
$$

$$
- \delta_i < 0.
$$

By Farkas' lemma, there exists $(\lambda_{i,1}, \lambda_{i,2}) \in \mathbb{R}_+^2$ such that

$$
\begin{aligned}
&\lambda_{i,1} f_i(a_{-i}) - \lambda_{i,2}(v(1, a_{-i}) - v(0, a_{-i})) \leq 0 \text{ for all } a_{-i} \in A_{-i}, \\
&- \lambda_{i,2} \leq -1.
\end{aligned}
$$

By $v(a^*) > v(1, a_{-i}^*)$, we have $\lambda_{i,1} f_i(a_{-i}^*) < \lambda_{i,1} f_i(a_{-i}^*) - \lambda_{i,2}(v(1, a_{-i}^*) - v(a^*)) \leq 0$, and hence $\lambda_{i,1} > 0$. Then $\lambda_i = \lambda_{i,1}/\lambda_{i,2} > 0$ satisfies the desired property. □

For an action profile $a^* \in A$, let $S = \{i \in I \mid a_i^* = 1\}$, let $\mathbf{0}^+$ denote the profile of action 0 among the players in $I \setminus S$, and let $\mathbf{1}^-$ denote the profile of action 1 among the players in S. We consider two games $\mathbf{g}^+ = (g_i^+)_{i \in I \setminus S}$ and $\mathbf{g}^- = (g_i^-)_{i \in S}$, where the former is the restricted game among the players in $I \setminus S$ given by $g_i^+(\cdot) = g_i(\cdot, \mathbf{1}^-)$, and the latter is the restricted game among the players in S given by $g_i^-(\cdot) = g_i(\cdot, \mathbf{0}^+)$. Denote by $X^+ \subset \Delta(A^+)$ the set of action distributions that satisfy sequential obedience in $\mathbf{g}^+$ and by $X^- \subset \Delta(A^-)$ the set of action distributions that satisfy reverse sequential obedience in $\mathbf{g}^-$, where $A^+ = \prod_{i \in I \setminus S} A_i$ and $A^- = \prod_{i \in S} A_i$.

Given Lemma A.2 above, the duality result of Oyama and Takahashi (2020, Lemma 2) together with Lemma A.1 in Oyama and Takahashi (2020), establishes a characterization of strict MP maximizer in terms of (reverse) sequential obedience.

---

[7] Morris and Ui (2005, Lemma 9) obtained a similar characterization for MP maximizer.

Proposition A.1 *Action profile a^* is a strict MP maximizer of* **g** *if and only if* $X^+ = \{\mathbf{0}^+\}$ *and* $X^- = \{\mathbf{1}^-\}$.

By the proof of Lemma 2(2) in Oyama and Takahashi (2019), the set of action distributions that satisfy sequential obedience in **g** has the largest element (in the sense of first-order stochastic dominance), which is degenerate on some pure action profile, which we denote by $\overline{a}$, and symmetrically, the set of action distributions that satisfy reverse sequential obedience in **g** has the smallest element, which is degenerate on some pure action profile, which we denote by $\underline{a}$. By the maximality and minimality of $\overline{a}$ and $\underline{a}$, we have the following:

Proposition A.2

(1) *If* $a^* = \overline{a}$, *then* $X^+ = \{\mathbf{0}^+\}$; *if* $X^+ = \{\mathbf{0}^+\}$, *then* $a^* \geq \overline{a}$.
(2) *If* $a^* = \underline{a}$, *then* $X^- = \{\mathbf{1}^-\}$; *if* $X^- = \{\mathbf{1}^-\}$, *then* $a^* \leq \underline{a}$.

Finally, by Morris et al. (2022, Proposition B.2) we have the following:

Proposition A.3 *The degenerate distribution on* $\overline{a}$ *satisfies reverse sequential obedience, and the degenerate distribution on* $\underline{a}$ *satisfies sequential obedience. In particular,* $\underline{a} \leq \overline{a}$.

Proposition 3 in Sect. 3.3 follows from Propositions A.1–A.3.

A.3. Relation to information design

In this section, we let $A_i = \{0, 1\}$ for all $i \in I$.

A.3.1. Full implementation

This section summarizes the relevant results on implementation by information design from Morris et al. (2022).

Fix a payoff structure $(\Theta, \mathbf{u})$, where Θ is a finite set of states, and $\mathbf{u} = (u_i)_{i \in I}$ is a profile of payoff functions $u_i : A \times \Theta \to \mathbb{R}$ for $i \in I$. We assume that $\mathbf{u}(\cdot, \theta)$ is supermodular for all $\theta \in \Theta$. We also impose the *two-sided dominance state assumption*: there exist $\overline{\theta}, \underline{\theta} \in \Theta$ such that $u_i(1, a_{-i}, \overline{\theta}) > u_i(0, a_{-i}, \overline{\theta})$ and $u_i(1, a_{-i}, \underline{\theta}) < u_i(0, a_{-i}, \underline{\theta})$ for all $i \in I$ and $a_{-i} \in A_{-i}$.

An outcome $\nu \in \Delta(A \times \Theta)$ is *fully implementable in* $(\Theta, \mathbf{u})$ if there exists a type space (T, P) over Θ such that every Bayes-Nash equilibrium of $(\Theta, \mathbf{u}, T, P)$ induces ν.[8]

An outcome $\nu \in \Delta(A \times \Theta)$ satisfies *strict sequential obedience in* $(\Theta, \mathbf{u})$ if there exists $\nu_\Gamma \in \Delta(\Gamma \times \Theta)$ such that

---

[8] Under supermodularity, full implementation in fact requires the uniqueness of Bayes-Nash equilibrium.

$$\sum_{\gamma \in \Gamma_i, \theta \in \Theta} \nu_\Gamma(\gamma, \theta)(u_i(1, a^1_{-i}(\gamma), \theta) - u_i(0, a^1_{-i}(\gamma), \theta)) > 0$$

for all $i \in I$ with $\nu_\Gamma(\Gamma_i \times \Theta) > 0$ and $\nu(a, \theta) = \nu_\Gamma(\{(\gamma, \theta) \in \Gamma \times \Theta \mid a^1(\gamma) = a\})$ for all $a \in A$ and $\theta \in \Theta$; ν satisfies *strict reverse sequential obedience in* $(\Theta, \mathbf{u})$ if there exists $\nu'_\Gamma \in \Delta(\Gamma \times \Theta)$ such that

$$\sum_{\gamma \in \Gamma_i, \theta \in \Theta} \nu'_\Gamma(\gamma, \theta)(u_i(1, a^0_{-i}(\gamma), \theta) - u_i(0, a^0_{-i}(\gamma), \theta)) < 0$$

for all $i \in I$ with $\nu'_\Gamma(\Gamma_i \times \Theta) > 0$ and $\nu(a, \theta) = \nu'_\Gamma(\{(\gamma, \theta) \in \Gamma \times \Theta \mid a^0(\gamma) = a\})$ for all $a \in A$ and $\theta \in \Theta$.

Denote $\mathbf{0} = (0, \ldots, 0), \mathbf{1} = (1, \ldots, 1) \in A$. Morris et al. (2022, Theorem B.1) showed that strict sequential obedience and strict reverse sequential obedience are necessary and jointly almost sufficient for full implementation.

Proposition A.4 *If an outcome* $\nu \in \Delta(A \times \Theta)$ *is fully implementable in* $(\Theta, \mathbf{u})$, *then it satisfies strict sequential obedience and strict reverse sequential obedience in* $(\Theta, \mathbf{u})$. *Conversely, if an outcome* ν *with* $\nu(\mathbf{1}, \overline{\theta}), \nu(\mathbf{0}, \underline{\theta}) > 0$ *satisfies strict sequential obedience and strict reverse sequential obedience in* $(\Theta, \mathbf{u})$, *then it is fully implementable in* $(\Theta, \mathbf{u})$.

An action distribution $\mu \in \Delta(A)$ is *limit fully implementable in* $(\Theta, \mathbf{u})$ *at* θ^* if there exists a sequence $\{\nu^k\}$ of fully implementable outcomes in $(\Theta, \mathbf{u})$ such that $\sum_{a \in A} \nu^k(a, \theta^*) \to 1$ and $\sum_{\theta \in \Theta} \nu^k(\cdot, \theta) \to \mu$ (hence $\nu^k(\cdot, \theta^*) \to \mu$) as $k \to \infty$.

Proposition A.5 *Action distribution* $\mu \in \Delta(A)$ *is limit fully implementable in* $(\Theta, \mathbf{u})$ *at* θ^* *if and only if it satisfies sequential obedience and reverse sequential obedience in* $\mathbf{u}(\cdot, \theta^*)$.

Proof The "only if" direction is immediate from Proposition A.4. For the "if" direction, suppose that μ satisfies sequential obedience and reverse sequential obedience in $\mathbf{u}(\cdot, \theta^*)$. We assume $\theta^* \neq \overline{\theta}, \underline{\theta}$ without loss of generality. For any $\delta \in (0, 1]$, define $\nu_\delta \in \Delta(A \times \Theta)$ by

$$\nu_\delta(a, \theta) = \begin{cases} (1 - \delta)\mu(a) & \text{if } \theta = \theta^*, \\ \frac{\delta}{2} & \text{if } (a, \theta) = (\mathbf{0}, \underline{\theta}) \text{ or } (\mathbf{1}, \overline{\theta}), \\ 0 & \text{otherwise.} \end{cases}$$

Then we have $\sum_{a \in A} \nu_\delta(a, \theta^*) \to 1$ and $\sum_{\theta \in \Theta} \nu_\delta(\cdot, \theta) \to \mu$ as $\delta \to 0$. Also ν_δ satisfies strict sequential obedience and strict reverse sequential obedience in $(\Theta, \mathbf{u})$. To see this, let $\rho, \rho' \in \Delta(\Gamma)$ be such that

$$\sum_{\gamma \in \Gamma_i} \rho(\gamma)(u_i(1, a^1_{-i}(\gamma), \theta^*) - u_i(0, a^1_{-i}(\gamma), \theta^*)) \geq 0,$$

$$\sum_{\gamma \in \Gamma_i} \rho'(\gamma)(u_i(1, a^0_{-i}(\gamma), \theta^*) - u_i(0, a^0_{-i}(\gamma), \theta^*)) \leq 0$$

for all $i \in I$ and $\mu(a) = \rho(\{\gamma \mid a^1(\gamma) = a\}) = \rho'(\{\gamma \mid a^0(\gamma) = a\})$ for all $a \in A$, and define $v_{\delta,\Gamma}, v'_{\delta,\Gamma} \in \Delta(\Gamma \times \Theta)$ by

$$v_{\delta,\Gamma}(\gamma, \theta) = \begin{cases} (1 - \delta)\rho(\gamma) & \text{if } \theta = \theta^*, \\ \frac{\delta}{2} & \text{if } (\gamma, \theta) = (\emptyset, \underline{\theta}) \text{ or } (\pi, \overline{\theta}), \\ 0 & \text{otherwise,} \end{cases}$$

$$v'_{\delta,\Gamma}(\gamma, \theta) = \begin{cases} (1 - \delta)\rho'(\gamma) & \text{if } \theta = \theta^*, \\ \frac{\delta}{2} & \text{if } (\gamma, \theta) = (\pi, \underline{\theta}) \text{ or } (\emptyset, \overline{\theta}), \\ 0 & \text{otherwise,} \end{cases}$$

where $\emptyset$ denotes the null sequence and π is an arbitrarily fixed permutation of I. Then we have

$$\sum_{\gamma \in \Gamma_i, \theta \in \Theta} v_{\delta,\Gamma}(\gamma, \theta)(u_i(1, a^1_{-i}(\gamma), \theta) - u_i(0, a^1_{-i}(\gamma), \theta))$$

$$= (1 - \delta) \sum_{\gamma \in \Gamma_i} \rho(\gamma)(u_i(1, a^1_{-i}(\gamma), \theta^*) - u_i(0, a^1_{-i}(\gamma), \theta^*)) + \frac{\delta}{2}(u_i(\mathbf{1}, \overline{\theta}) - u_i(0, \mathbf{1}_{-i}, \overline{\theta})) > 0,$$

$$\sum_{\gamma \in \Gamma_i, \theta \in \Theta} v'_{\delta,\Gamma}(\gamma, \theta)(u_i(1, a^0_{-i}(\gamma), \theta) - u_i(0, a^0_{-i}(\gamma), \theta))$$

$$= (1 - \delta) \sum_{\gamma \in \Gamma_i} \rho'(\gamma)(u_i(1, a^0_{-i}(\gamma), \theta^*) - u_i(0, a^0_{-i}(\gamma), \theta^*)) + \frac{\delta}{2}(u_i(1, \mathbf{0}_{-i}, \underline{\theta}) - u_i(\mathbf{0}, \underline{\theta})) < 0$$

for all $i \in I$ and $v_\delta(a, \theta) = v_{\delta,\Gamma}(\{(\gamma, \theta) \mid a^1(\gamma) = a\}) = v'_{\delta,\Gamma}(\{(\gamma, \theta) \mid a^0(\gamma) = a\})$ for all $a \in A$ and $\theta \in \Theta$, as claimed. Thus, by Proposition A.4, v_δ is fully implementable in $(\Theta, \mathbf{u})$, and therefore, μ is limit fully implementable in $(\Theta, \mathbf{u})$ at θ^*. $\square$

In particular, as long as payoff structure $(\Theta, \mathbf{u})$ is supermodular and satisfies the two-sided dominance state assumption, limit full implementation at θ^* depends only on $\mathbf{u}(\cdot, \theta^*)$ and independent of the payoff structure that embeds $\mathbf{u}(\cdot, \theta^*)$, and due to Proposition A.3, there exists at least one (in fact, degenerate) limit fully implementable action distribution at θ^*.

A.3.2. Strict robustness and limit full implementation

The formal relationship between strict robustness and limit full implementation is as follows. First, if a type space (T, P) assigns high probability on θ^*, then $(\Theta, \mathbf{u}, T, P)$ is an (ε, η)-elaboration of $\mathbf{u}(\cdot, \theta^*)$ with small $\varepsilon, \eta > 0$.

Lemma A.3 *Let* $(\Theta, \mathbf{u})$ *be a finite-state payoff structure with* $\theta^* \in \Theta$ *and* (T, P) *a type space over* Θ. *Then incomplete information game* $(\Theta, \mathbf{u}, T, P)$ *is an* (ε, η)-*elaboration of* $\mathbf{u}(\cdot, \theta^*)$ *with* $\varepsilon = |I|\sqrt{1 - p^*}$ *and* $\eta = 2\sqrt{1 - p^*} \max_{i \in I, a \in A, \theta \neq \theta^*} |u_i(a, \theta)|$, *where* $p^* = P(\{\theta^*\} \times T)$.

Proof For each $i \in I$, let $T_i^* = \{t_i \in T_i \mid P(\{\theta^*\} \times T_{-i}|t_i) \geq 1 - \sqrt{1 - p^*}\}$ and $T^* = \prod_{i \in I} T_i^*$. Then for each $t_i \in T_i^*$, we have

$$\sum_{\theta \in \Theta} P(\{\theta\} \times T_{-i}|t_i) \max_{a \in A} |u_i(a, \theta) - u_i(a, \theta^*)|$$

$$= \sum_{\theta \neq \theta^*} P(\{\theta\} \times T_{-i}|t_i) \max_{a \in A} |u_i(a, \theta) - u_i(a, \theta^*)|$$

$$\leq 2\sqrt{1 - p^*} \max_{a \in A, \theta \neq \theta^*} |u_i(a, \theta)| \leq \eta.$$

Also, since $P(\Theta \times T_i^* \times T_{-i}) \geq 1 - \sqrt{1 - p^*}$, we have $P(\Theta \times T^*) \geq 1 - |I|\sqrt{1 - p^*} = 1 - \varepsilon$. $\qquad\square$

Thus, if μ is strictly robust in $\mathbf{g}$ and $(\Theta, \mathbf{u})$ embeds $\mathbf{g}$ at $\theta^* \in \Theta$, then no action distribution other than μ is limit fully implementable in $(\Theta, \mathbf{u})$ at θ^*. Indeed, we have:

Proposition A.6 *Let* $\mathbf{g}$ *be a BAS game, and let* $(\Theta, \mathbf{u})$ *be a finite-state binary-action supermodular payoff structure that satisfies the two-sided dominance state assumption and* $\mathbf{u}(\cdot, \theta^*) = \mathbf{g}$ *with* $\theta^* \in \Theta$. *Then an action distribution is strictly robust in* $\mathbf{g}$ *if and only if it is the unique action distribution that is limit fully implementable in* $(\Theta, \mathbf{u})$ *at* θ^*.

Proof The "only if" direction follows from Propositions A.3 and A.5 and Lemma A.3. The "if" direction follows from Theorem 1 and Propositions 3 and A.5. $\qquad\square$

Finally, the "only if" direction of Theorem 2 follows from Propositions 3, A.5, and A.6.

References

Basteck, C., & Daniëls, T. R. (2011). Every symmetric 3×3 global game of strategic complementarities has noise-independent selection. *Journal of Mathematical Economics, 47*, 749–754.

Basteck, C., Daniëls, T. R., & Heinemann, F. (2013). Characterising equilibrium selection in global games with strategic complementarities. *Journal of Economic Theory, 148*, 2620–2637.

Carlsson, H., & van Damme, E. (1993). Global games and equilibrium selection. *Econometrica, 61*, 989–1018.

Kajii, A., & Morris, S. (1997). The robustness of equilibria to incomplete information. *Econometrica, 65*, 1283–1309.

Kajii, A., & Morris, S. (2020). Notes on "refinements and higher order beliefs". *Japanese Economic Review, 71*, 35–41.

Monderer, D., & Shapley, L. S. (1996). Potential games. *Games and Economic Behavior, 14*, 124–143.

Morris, S., Oyama, D., & Takahashi, S. (2022). Implementation via information design in binary-action supermodular games, SSRN 3697335.

Morris, S., Oyama, D., & Takahashi, S. (2023). Global game selections in binary-action supermodular games, SSRN 4392531.

Morris, S., & Ui, T. (2005). Generalized potentials and robust sets of equilibria. *Journal of Economic Theory, 124*, 45–78.

Oyama, D., & Takahashi, S. (2011). On the relationship between robustness to incomplete information and noise-independent selection in global games. *Journal of Mathematical Economics, 47*, 683–688.

Oyama, D., & Takahashi, S. (2019). Generalized belief operator and the impact of small probability events on higher order beliefs, SSRN 3375777.

Oyama, D., & Takahashi, S. (2020). Generalized belief operator and robustness in binary-action supermodular games. *Econometrica, 88*, 693–726.

Oyama, D., & Takahashi, S. (2023). Robustness in binary-action supermodular games revisited, SSRN 4392540.

Oyama, D., Takahashi, S., & Hofbauer, J. (2008). Monotone methods for equilibrium selection under perfect foresight dynamics. *Theoretical Economics, 3*, 155–192.

Oury, M., & Tercieux, O. (2007). Contagion in games with strategic complementarities, Mimeo.

Oyama, D., & Tercieux, O. (2009). Iterated potential and robustness of equilibria. *Journal of Economic Theory, 144*, 1726–1769.

Rubinstein, A. (1989). The electronic mail game: strategic behavior under "almost common knowledge". *American Economic Review, 79*, 385–391.

Takahashi, S. (2020). Non-equivalence between all and canonical elaborations. *Japanese Economic Review, 71*, 43–57.

Ui, T. (2001). Robust equilibria of potential games. *Econometrica, 69*, 1373–1380.

Wu, W.-T., & Jiang, J.-H. (1962). Essential equilibrium points of *n*-person non-cooperative games. *Scientia Sinica, 11*, 1307–1322.