Continuous virtual implementation: complete information*

Ritesh Jain† Michele Lombardi‡

November 11, 2021

Abstract

A social choice rule (SCR) is a mapping from preference profiles to lotteries over outcomes. When preference profiles are close to being common knowledge among players, an SCR is continuously virtually fully implementable if there exists a mechanism such that all its equilibrium outcomes are arbitrarily close to the outcomes recommended by the SCR. When there are at least three players and a domain condition is satisfied, we obtain the following result: any SCR is continuously virtually fully implementable in Bayesian Nash equilibria, as well as in interim correlated rationalizable strategies, by a finite mechanism.

Keywords: Continuous implementation, virtual implementation, social choice rules, uniform-weak topology, common knowledge.
JEL Classifications: C72; D71; D82; D83.

*Special thanks go to Siyang Xiong, Arunava Sen and a referee, whose comments and suggestions have led to substantial improvements in the article. We are grateful to Debasis Mishra, Hannu Vartiainen, Kirby Neilsen, Olivier Bochet, Paul Healy, Yaron Azrieli, Ville Korpela, John Levy, and seminar participants at the Ohio State University, Indian Statistical Institute and National Taiwan University. Ritesh Jain greatly acknowledges financial support from the Ministry of Science and Technology Grant, Taiwan under the grant: 107-2410-H-001-113-MY2. All remaining errors are ours.

†Institute of Economics, Academia Sinica, e-mail: ritesh@econ.sinica.edu.tw.

‡University of Liverpool Management School, Liverpool, UK. E-mail: michele.lombardi@liverpool.ac.uk. Department of Economics and Statistics, University of Napoli Federico II.

1 Introduction

A traditional assumption of implementation theory is that of complete information, meaning that the state (of nature) is common knowledge among players (but unknown of course to the designer). However, a fundamental contribution of game theory shows that predictions of strategic situations are sensitive to common knowledge assumptions (see, for example, Rubinstein, 1989 and Weinstein and Yildiz, 2007). This paper considers the question of when an SCR is virtually implementable not only when the state is common knowledge among players but also when it is *close* to common knowledge.[1] Thus, the paper studies virtual continuous implementation, and it provides a characterization of continuously virtually implementable SCRs in the preference environment introduced by Abreu and Matsushima (1992a).[2]

Instead of exact implementation, in which every outcome recommended by the SCR is achieved with probability one, our designer wishes to achieve outcomes specified by an SCR with probability arbitrarily close to one: he requires *virtual* implementation.

By following Oury and Tercieux (2012), we use the model of incomplete information introduced in Harsanyi (1967) and developed in Mertens and Zamir (1985). Our notion of continuous virtual implementation requires that in any "nearby" model in the universal type space that embeds our baseline complete information model, all Bayesian Nash equilibrium outcomes of the devised mechanism are ε-close to the recommendations provided by a given SCR, at all types close to the initial common knowledge types, and for every $\varepsilon > 0$. If such a mechanism exists, we say that the SCR is continuously virtually implementable.

We consider continuity with respect to the uniform–weak topology (see, for example, Monderer and Samet, 1989; and Chen et al., 2010). Roughly speaking, this topology preserves common knowledge. That is, types close to the common knowledge types in this topology have approximately common knowledge of the state in the following sense: for some number p close to one, they assign probability of at least p to the state, assign probability of at least p to the event that the state occurs in and the other players assign probability of at least p to the state, and so forth, ad infinitum.

Under a well-known domain restriction due to Abreu and Matsushima (1992a), we show that any SCR is continuously virtually implementable in Bayesian Nash equilibria when there are at least three players. Moreover, we also achieve continuous virtual implementation in interim correlated rationalizable (ICR) strategies, which assures that the outcomes generated by a given SCR are achieved despite the presence of strategic uncertainty. These results are obtained by devising a *finite* mechanism. As a consequence, every player's best response correspondence is always well-defined, and it does not rely on any tail-chasing procedure to

[1] In this paper, "implementation" without qualification always refers to full-implementation, which means that every equilibrium outcome of the devised mechanism is socially optimal. This contrasts with partial-implementation where only some equilibria need to be desirable.

[2] They impose a weak domain restriction ruling out identical preferences.

eliminate unwanted best replies, such as integer games. Although we use a module game in our construction, our characterization result accounts for mixed-strategy equilibria, unlike existing results that rely on modulo games. We also show that for the devised mechanism, virtual implementation in strict Nash equilibria, and in rationalizable strategies, suffices for continuous virtual implementation when the state is common knowledge among players. Similar to Chung and Ely (2003), strict incentives are used in order to attain continuous virtual implementation in uniform–weak topology for any "nearby" model.

In a seminal paper, Abreu and Sen (1991) characterize the class of correspondences that are virtually implementable in Nash equilibria. This result relies on a tail-chasing construction, which is typical in the classical literature on implementation theory. Following Abreu and Matsushima (1991), we characterize the class of correspondences that are virtually implementable by a finite mechanism.

There are several reasons why focussing only on social choice functions can be considered unsatisfactory (see, for instance, Thomson (1996)). Firstly, multi-valued SCRs typically represent many social decisions. Prominent examples include the Pareto, the Walrasian, the Condorcet, and the no-envy correspondences. Secondly and foremost, since F represents the social objectives that the society or its representative want to achieve, its full implementation is the correct objective of the society. It would be unacceptable to partially implement F by implementing a social choice function f which systematically picks, for each θ, a socially optimal outcome $f(\theta) \in F(\theta)$.[3] The reason is that the implementation of this subselection $f \in F$ may violate some of the normative properties that led the society or its representatives to choose F. As Thomson (1986, p. 135) aptly noted "it most certainly will be unacceptable when the correspondence embodies some minimal concerns about fairness distribution."

An argument made in favor of a partial implementation of F is based on the interpretation that the mechanism designer views the outcomes in $F(\theta)$ as equally good (Abreu and Sen, 1991; Mezzetti and Renou, 2012). A shortcoming of this interpretation is that in some situations we do not know whether the mechanism designer is indifferent between socially optimal outcomes or not. Moreover, the classical interpretation of implementation of F requires that each outcome in $F(\theta)$ must be supported by a distinct equilibrium (Maskin, 1999; Abreu and Sen, 1991) and it does not assume planner's indifference. This paper shows that this classical interpretation is not restrictive when the objective is to implement F virtually. This is in sharp contrast to the case of "exact" implementation where implementation in the classical sense is more restrictive than implementation under the assumption of planner's indifference (Mezzetti and Renou, 2012).

The remainder of the paper is organized as follows. Section 2 defines the implementation model. Section 3 presents the characterization result and provides an informal discussion of the implementing mechanism, with the proof offered in Section 4. This result builds on

[3] Here, partial implementation means that the set of equilibrium outcomes is a non-empty subset of the socially optimal outcomes.

the existing literature on implementation theory, which is discussed in Section 5. Section 6 concludes by highlighting possible extensions.

2 Model

Preliminaries

Let $N = \{1, ..., n\}$ be a finite set of players with $n \geq 3$. The finite set of (pure) outcomes is denoted by X. The set of all lotteries over X is denoted by Y. Player i's utility function is indexed by a parameter θ_i. We refer to θ_i as player i's payoff type. The set of admissible payoff types for player i is assumed to be finite, and it is denoted by Θ_i. Player i's preferences over lotteries is described by a continuous and bounded utility function $u_i : Y \times \Theta_i \to \mathbb{R}$, where $u_i(y, \theta_i)$ is player i's utility of the lottery y when he is of payoff type θ_i. For each $\theta_i \in \Theta_i$, $u_i(\cdot, \theta_i)$ satisfies the expected utility hypothesis. A payoff type profile is described by an n-tuple of types $\theta \in \prod_{i \in N} \Theta_i = \Theta$. For any $\theta \in \Theta$, θ_{-i} denotes the $n-1$-tuple $(\theta_1, ..., \theta_{i-1}, \theta_{i+1}, ..., \theta_n)$.

Following Abreu and Matsushima (1992a), we introduce a domain restriction that will play a key role in our analysis.

Assumption 1 (Abreu and Matsushima, 1992a) For every $i \in N$ and $\theta \in \Theta$, lotteries $\bar{a}(i, \theta), \underline{a}(i, \theta) \in Y$ exist such that

$$u_i(\bar{a}(i, \theta), \theta_i) > u_i(\underline{a}(i, \theta), \theta_i) \text{ and } u_j(\underline{a}(i, \theta), \theta_j) \geq u_j(\bar{a}(i, \theta), \theta_j) \text{ for each } j \in N \setminus \{i\}.$$

Assumption 1 requires that for each payoff type profile θ and each player i, there exist two lotteries that are strictly ranked by player i, but every other player has just the opposite ranking. This assumption is satisfied in environments with transferable private goods that are positively valued by players. In the implementation literature, this kind of assumptions is often made in studies relating to well-behaved implementing mechanisms (Jackson et al., 1994; Kartik et al., 2014). Kunimoto and Serrano (2011) even argue that Assumption 1 is indispensable.

The goal of the designer is to implement a (stochastic) SCR F, which is a mapping $F : \Theta \twoheadrightarrow Y$ from Θ to a nonempty compact set of Y.[4] The common interpretation is that F represents the social objectives that the planner wants to achieve. If $x \in F(\theta)$, we say that x is socially optimal at θ. A social choice function (SCF) $f : \Theta \to Y$ is a single-valued SCR.

The implementation problem arises from the fact that the planner's goal depends on the true payoff type profile and he does not know it. To elicit it, the planner designs a

[4] The compactness of $F(\theta)$ is needed to make sure that the players' strategy space is compact (see below). This assumption also appears in Kunimoto and Serrano (2019), and it is consistent with the implementation model of Mezzetti and Renou (2012). The latter authors study Nash implementation in terms of the support of the equilibrium, with a finite set of outcomes and deterministic SCRs.

(stochastic) mechanism. A mechanism is a game form $\Gamma = (M_1, ..., M_n, g)$, where M_i is player i's compact set of pure strategies, and $g : \prod_{i \in N} M_i \to Y$ is the outcome function. We denote a pure strategy of player i by $m_i \in M_i$ and a profile of pure strategies is denoted by $m = (m_1, ..., m_n) \in M = \prod_{i \in N} M_i$. A mechanism is finite if M_i is a finite set for each player $i \in N$.

Universal type space and topologies

Our goal is to study virtual implementation when the designer entertains some doubts about whether the true model is of complete information. To model this, we embed the complete information model in the general model of incomplete information introduced in Harsanyi(1967) and developed in Mertens and Zamir (1985). According to Oury and Tercieux (2012), a model $\mathcal{T}$ is a pair (T, κ), where $T = \prod_{i \in N} T_i$ is a countable type space and $\kappa[t_i] \in \triangle(\Theta \times T_{-i})$ is the associated belief for each type $t_i \in T_i$. An incomplete information model $\mathcal{T}$ and a mechanism Γ induce an incomplete information game, which is denoted by $\mathcal{U}(\Gamma, \mathcal{T})$. Let $\bar{\mathcal{T}} = (\bar{T}, \bar{\kappa})$ denote the complete information model, that is, $\bar{T}_i = \{t_i^\theta | \theta \in \Theta\}$ and $\bar{\kappa}[t_i^\theta](\theta, t_{-i}^\theta) = 1$ for each $\theta \in \Theta$.

For any two models $\mathcal{T}$ and $\mathcal{T}'$, we write $\mathcal{T}' \subseteq \mathcal{T}$ if $T' \subseteq T$, and for each $t_i \in T_i'$, $\kappa[t_i](E) = \kappa[t_i]\left(\left(\Theta \times T_{-i}'\right) \cap E\right)$ for any measurable set $E \subseteq \Theta \times T_{-i}$.

For any type t_i of the model $\mathcal{T}$, we can compute the first-order belief of t_i (that is, her belief about Θ), denoted by $h_i^1[t_i]$, by setting $h_i^1[t_i]$ equal to the marginal distribution of $\kappa[t_i]$ on Θ. Also, her second-order belief (that is, her belief about θ and the others' first-order beliefs) can be computed by setting

$$h_i^2[t_i](E) = \kappa[t_i]\left(\left\{(\theta, t_{-i}) \mid \left(\theta, h_1^1[t_1], ..., h_n^1[t_n]\right) \in E\right\}\right)$$

for each measurable $E \subseteq \Theta \times (\triangle(\Theta))^n$. An entire hierarchy of beliefs

$$h_i[t_i] = \left(h_i^1[t_i], h_i^2[t_i], ..., h_i^k[t_i], ...\right)$$

of type t_i can be computed by proceeding in this way. Note that $h_i^1[t_i] \in \triangle(\Theta)$, $h_i^2[t_i] \in \triangle(\Theta \times (\triangle(\Theta))^n)$, and so on.

Let us now introduce the notion of distance used in this paper. As in Chen et al. (2018a), let $Z^0 = \Theta$ and let $Z^k = \left[\triangle\left(Z^{k-1}\right)\right]^n \times Z^{k-1}$ for each $k \geq 1$. Note that $h_i^k[t_i] \in \triangle\left(Z^{k-1}\right)$ for every $k \geq 1$. Let d^0 denote the discrete metric on Θ and let d^1 denote the Prohorov distance on the space of first-order beliefs.[5] Recursively, for any $k \geq 2$, we endow $\triangle\left(Z^{k-1}\right)$ with the

[5] Given a metric space (X, d), the Prohorov distance between any two $\mu, \mu' \in \triangle(X)$ is

$$\inf\{\gamma > 0 | \mu'(E) \leq \mu(E^\gamma) + \gamma\}$$

for every Borel set $E \subseteq X$, where $E^\gamma = \{x \in X | \inf_{y \in E} d(x, y) < \gamma\}$.

Prohorov distance d^k, where Z^{k-1} is endowed with the sup-metric induced by $d^0, d^1, ..., d^{k-1}$. Let $T_i^* \subseteq \prod_{k=0}^{\infty} \triangle\left(Z^k\right)$ be the player i's universal type space constructed by Mertens and Zamir (1985). This space has the property that $h_i[t_i] \in T_i^*$ if there exists some type t_i' in some model such that t_i and t_i' have the same k-th-order belief for every k. Each T_i^* is endowed with a product topology.

A sequence of types $\{t_{i,n}\}_{n=1}^{\infty}$ in T_i^* converges uniform–weakly to a type t_i if

$$d_i^{\text{uw}}(t_{i,n}, t_i) \equiv \sup_{k \geq 1} d_i^k\left(h_i^k[t_{i,n}], h_i^k[t_i]\right) \to 0.$$

We write that $d^{\text{uw}}(t_n, t) \to 0$ if $d_i^{\text{uw}}(t_{i,n}, t_i) \to 0$ for each i.[6]

A sequence of types $\{t_{i,n}\}_{n=1}^{\infty}$ in T_i^* converges in product topology to a type t_i if

$$d_i^{\text{p}}(t_{i,n}, t_i) \equiv \sum_{k=1}^{\infty} \frac{1}{2^k} d_i^k\left(h_i^k[t_{i,n}], h_i^k[t_i]\right) \to 0.$$

Again, we write that $d^{\text{p}}(t_n, t) \to 0$ if $d_i^{\text{p}}(t_{i,n}, t_i) \to 0$ for each i.

Solution concepts

Fix any game with incomplete information $\mathcal{U}(\Gamma, \mathcal{T})$. In this game, player i's (mixed) strategy is any measurable function $\sigma_i : T_i \to \triangle(M_i)$. We write $\sigma_i(m_i|t_i)$ for the probability that strategy σ_i assigns to message m_i when player i is of type t_i. For each player $i \in N$, player i's best response correspondence $BR_i : \triangle(\Theta \times M_{-i}) \twoheadrightarrow M_i$ is defined by

$$BR_i(\pi_i|\mathcal{U}(\Gamma, \mathcal{T})) = \arg\max_{m_i \in M_i} \int_{(\theta, m_{-i}) \in \Theta \times M_{-i}} u_i(g(m_i, m_{-i}), \theta) \, \mathrm{d}\pi_i(\theta, m_{-i}),$$

for each belief $\pi_i \in \triangle(\Theta \times M_{-i})$, where $\arg\max$ is the set of maximizers.

A strategy profile $\sigma = (\sigma_1, ..., \sigma_n)$ is a Bayes-Nash equilibrium of $\mathcal{U}(\Gamma, \mathcal{T})$ if, for each $i \in N$ and each $t_i \in T_i$,

$$m_i \in \text{Supp}(\sigma_i(t_i)) \implies m_i \in BR_i(\pi_i(\cdot|t_i, \sigma_{-i})|\mathcal{U}(\Gamma, \mathcal{T})),$$

where $\pi_i(\cdot|t_i, \sigma_{-i}) \in \triangle(\Theta \times M_{-i})$ denotes the joint distribution on the underlying uncertainty and the messages of other players induced by type t_i and strategy profile σ_{-i}. We denote by $\text{BNE}(\mathcal{U}(\Gamma, \mathcal{T}))$ the set of Bayes-Nash equilibria of the game of incomplete information $\mathcal{U}(\Gamma, \mathcal{T})$. The set of Nash equilibria of the game of complete information $\mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)$ at type profile $t^\theta \in \bar{T}$ is denote by $\text{NE}\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$. Given a model of incomplete information $\mathcal{U}(\Gamma, \mathcal{T})$ such that $\mathcal{T} \supseteq \bar{\mathcal{T}}$, a strategy $\sigma \in \text{BNE}(\mathcal{U}(\Gamma, \mathcal{T}))$ and a profile $t^\theta \in \bar{T}$, we write $\sigma_{|t^\theta}$ for the strategy restricted to the profile t^θ. It can easily be checked that $\sigma_{|t^\theta} \in \text{NE}\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$.

[6] We refer the reader to Chen et al. (2010) for further details about this topology.

The second solution concept that we study is that of interim correlated rationalizability (ICR), which was introduced by Dekel et al. (2007). As with rationalizability in complete-information games, ICR is defined by an iterative deletion procedure. In line with Weinstein and Yildiz (2017), fix any game with incomplete information $\mathcal{U}(\Gamma,\mathcal{T})$. For each player $i \in N$ and integer $k \in \mathbb{Z}_+$, define the family of correspondences $\mathrm{R}_i^k : T_i \twoheadrightarrow M_i$ iteratively, by setting $\mathrm{R}_i^0(t_i,\mathcal{U}(\Gamma,\mathcal{T})) = M_i$, and for each $k > 0$,[7]

$$\mathrm{R}_i^k(t_i,\mathcal{U}(\Gamma,\mathcal{T})) = \left\{ m_i \in M_i \;\middle|\; \begin{array}{l} \text{there exists a belief } \lambda_i \in \triangle(T_{-i}\times\Theta\times M_{-i}) \text{ such that:} \\ (1)\ m_i \in BR_i(\pi_i|\mathcal{U}(\Gamma,\mathcal{T})) \text{ such that } \pi_i = \mathrm{marg}_{\Theta\times M_{-i}}\lambda_i \\ (2)\ \mathrm{marg}_{\Theta\times T_{-i}}\lambda_i = \kappa[t_i] \\ (3)\ \lambda_i\left(\left\{(\theta,t_{-i},m_{-i})\,|m_{-i}\in \mathrm{R}_{-i}^{k-1}(t_{-i},\mathcal{U}(\Gamma,\mathcal{T}))\right\}\right) = 1 \end{array} \right\}.$$

Property 1 requires that m_i is the best response to belief λ_i; property 2 requires consistency of λ_i with type t_i's beliefs about $T_{-i}\times\Theta$; and the last property requires that the other players play according to $\mathrm{R}_{-i}^{k-1}(t_{-i},\mathcal{U}(\Gamma,\mathcal{T}))$ under λ_i. The limiting correspondence $\mathrm{R}_i : T_i \twoheadrightarrow M_i$ is defined by

$$\mathrm{R}_i(t_i,\mathcal{U}(\Gamma,\mathcal{T})) = \bigcap_{k\geq 0} \mathrm{R}_i^k(t_i,\mathcal{U}(\Gamma,\mathcal{T})).$$

The set of ICR strategy profiles of the game of incomplete information $\mathcal{U}(\Gamma,\mathcal{T})$ at type profile $t \in T$ is denoted by $\mathrm{R}(t,\mathcal{U}(\Gamma,\mathcal{T}))$. The set of rationalizable strategy profiles of the game of complete information $\mathcal{U}(\Gamma,\bar{\mathcal{T}})$ at type profile $t^\theta \in \bar{T}$ is denoted by $\mathrm{R}(t^\theta,\mathcal{U}(\Gamma,\bar{\mathcal{T}}))$.

Implementation

An SCR F is virtually implementable if there exists a "nearby" nonempty correspondence H that is "exactly" implementable in a solution concept. Formally, let $d(x,y)$ be the Euclidean distance between any pair of lotteries. The SCR $F:\Theta\twoheadrightarrow Y$ is ε-close to a nonempty correspondence $H:\Theta\twoheadrightarrow Y$ if for each $\theta\in\Theta$, $d_h(F(\theta),H(\theta)) \leq \varepsilon$, where d_h is the the Hausdorff distance.[8] Our notion of closeness is inspired by Abreu and Sen (1991), but is weaker than their definition (p. 1005).[9] Our definition of continuous virtual (full) implementation can be stated as follows.

Definition 1 A finite mechanism Γ continuously virtually implements $F:\Theta\twoheadrightarrow Y$ in Bayesian Nash equilibria (resp., in ICR strategies) with respect to (w.r.t.) d^{uw} if for each $\varepsilon > 0$, a nonempty correspondence $H:\Theta\twoheadrightarrow Y$ exists which is ε-close to F such that for any $\theta\in\Theta$ and any model $\mathcal{T}\supseteq\bar{\mathcal{T}}$:[10]

[7] marg_A takes the marginal with respect to the set A.

[8] For a definition of Hausdorff distance see, for instance, Taylor (1986; p. 127).

[9] It is weaker because Abreu and Sen's (1991) notion of closeness requires the existence of a bijection $\tau_\theta : F(\theta)\to H(\theta)$ such that the Euclidean distance $\rho(x,\tau_\theta(x)) \leq \varepsilon$ for all $x\in F(\theta)$.

[10] Let $\mathrm{Supp}(\sigma(t_n)) \equiv \prod_{i\in N}\mathrm{Supp}(\sigma_i(t_{i,n}))$.

(a) for each $x \in F(\theta)$, there exists $\sigma \in \mathrm{BNE}(\mathcal{U}(\Gamma, \mathcal{T}))$ such that (i) $g\left(\sigma\left(t^{\theta}\right)\right) \in \rho_H^{\theta}(x)$ and $\sigma_{|t^{\theta}} \in \mathrm{NE}\left(t^{\theta}, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$ [resp., $\sigma_{|t^{\theta}} \in \mathrm{R}\left(t^{\theta}, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$] and (ii) for any sequence of types $\{t_n\}$ in T with $d^{\mathrm{uw}}\left(t_n, t^{\theta}\right) \to 0$, there exists $\underline{n}$ large enough such that for any $n \geq \underline{n}$: $g(\sigma(t_n)) \in \rho_H^{\theta}(x)$ [resp., $\sigma(t_n) \in \mathrm{R}(t_n, \mathcal{U}(\Gamma, \mathcal{T}))$].

(b) for each $\sigma \in \mathrm{BNE}(\mathcal{U}(\Gamma, \mathcal{T}))$ [resp., $\sigma(t_n) \in \mathrm{R}(t_n, \mathcal{U}(\Gamma, \mathcal{T}))$], $\bigcup_{m \in \mathrm{Supp}(\sigma(t_n))} g(m) \subseteq H(\theta)$.

When there exists a finite mechanism that continuously virtually implements F in Bayesian Nash equilibria (resp., in ICR strategies) w.r.t. d^{uw}, we say that F is continuously virtually implementable in Bayesian Nash equilibria (resp., in ICR strategies) w.r.t. d^{uw} by a finite mechanism.

Only part (a) of Definition 1 is directly comparable with Definition 2 of Oury and Tercieux (2012) when F is an SCF. Part (a) requires that if x is socially optimal at θ, then a Bayes-Nash equilibrium σ exists such that for some large enough $\underline{n}$ and each $n \geq \underline{n}$, it holds that the outcome corresponding to $g(\sigma(t_n))$ is ε-close to x, and such that σ is a Nash equilibrium strategy profile for type profile $\theta \in \Theta$ when it restricted to t_{θ}. Part (b) requires that for every Bayes-Nash equilibrium strategy profile σ and every "nearby" type t_n, it holds that any outcome of any pure strategy profile in the support of the equilibrium profile $\sigma(t_n)$ must be ε-close to some outcome in F at θ.

Definition 2 above is closely related to the definition of continuous implementation of Oury and Tercieux (2012). However, there are important differences. First, our continuity notion is based on the uniform weak topology, whereas Oury and Tercieux (2012)'s notion of continuity relies on the product topology. Second, we require full implementation, wheares Oury and and Tercieux (2021)'s notion of implementation is that of partial implementation. Specifically, if we replace the uniform weak topology with product topology and drop part (b), Definition 2 is exactly the definition of continuous implementation of Oury and Tercieux (2012).

The following lemma, due to Abreu and Matsushima (1992a; p. 999), will be used throughout the paper. It requires the existence of a set of lotteries for player i such that each of her type has a distinct maximal element within this set. Recall that different payoff types induce different preferences over lotteries, like in Abreu and Matsushima (1992).

Lemma 1 (Abreu and Matsushima, 1992a) Let Assumption 1 hold. Let $i \in N$. Then, a function $f_i : \Theta_i \to Y$ exists such that for each $\theta_i \in \Theta_i$, it holds that

$$u_i(f_i(\theta_i), \theta_i) > u_i(f_i(\theta_i'), \theta_i)$$

for all $\theta_i' \in \Theta_i \backslash \{\theta_i\}$.

3 Result

The characterization result is stated below. Its proof is constructive and can be found in the next section. The devised mechanism is a variant of the Abreu-Matsushima mechanism (Abreu and Matsushima, 1992a), henceforth, AM-mechanism, which virtually implements any SCF in iteratively undominated strategies when there are at least three players and Assumption 1 is satisfied. It is worth mentioning that the result does not depend on the private-value assumption because the original benchmark model is the complete information model.

Theorem 1 Let Assumption 1 hold and let $n \geq 3$. Then, any SCR $F : \Theta \twoheadrightarrow Y$ is continuously virtually implementable in Bayesian Nash equilibria (resp., in ICR strategies) w.r.t. d^{uw} by a finite mechanism.

Remark 1 In Theorem 1, continuity is with respect to the uniform–weak topology. For continuity with respect to the coarser product topology, that is, with respect to the topology of weak convergence of k–order beliefs, for each $k \geq 1$, we can show that part (b) of Definition 1 holds. The reason is that we achieved full implementation in rationalizable strategies on the complete information model $\bar{\mathcal{T}}$, and in every game induced by our finite mechanism the ICR correspondence is upper hemicontinuous in the product topology. However, it is not clear whether part (a) of Definition 1 holds. The reason is that it is generally impossible to obtain multiple equilibria for "nearby" types in product topology (Rubinstein, 1989; Weinstein and Yildiz, 2007). Note that one of the equilibria is continuous with respect to the product topology.

In what follows, we intuitively discuss the basic arguments in the proof of the complete information model. In the devised mechanism, each player makes $(K+2)$ simultaneous announcements. A typical announcement is indexed by $k \in \{-1, 0, 1, ..., K\}$, where K is an integer that is yet to be specified.

Fix any $\delta > 0$. For each $\theta \in \Theta$, $F(\theta)$ is δ-closed to a finite set $A(\delta, F(\theta))$ in the Hausdorff distance such that the union of the delta open balls of the elements of $A(\delta, F(\theta))$ covers $F(\theta)$.[11] Let $\mathcal{F}_{F,\delta} = \{f : \Theta \to Y | \text{for every } \theta \in \Theta,\ f(\theta) \in A(\delta, F(\theta))\}$ be a collection of SCFs, each of which assigns, to each type profile θ, an element $f(\theta) \in A(\delta, F(\theta))$. Since Θ is finite and the SCR is compact-valued, the collection $\mathcal{F}_{F,\delta}$ is finite.[12]

[11] Look at the collection of open balls $B_\delta(x)$, where x runs over all elements of $F(\theta)$. Since $F(\theta)$ is compact, this open covering has a finite sub-covering–i.e., using finitely many of these open balls. The centres of the balls is the finite set $A(\delta, F(\theta))$ we were looking for. We are grateful to a referee for having drawn our attention to this point, which has allowed us to devise a finite implementing mechanism. Jain and Lombardi (2019) provide a characterization result via a bounded (not necessarily finite) mechanism.

[12] To understand it, suppose that the cardinality of $|\Theta| = J$. Then, $\mathcal{F}_{F,\delta} \equiv A(\delta, F(\theta_1)) \times ... \times A(\delta, F(\theta_J))$ and $A(\delta, F(\theta_j))$ is finite for each state $\theta_j \in \Theta$, with $j = 1, ..., J$.

Each player i reports an ε-approximated SCF in the $k = -1$ announcement, her type in the $k = 0$ announcement, and an entire type profile in each of the remaining announcements. That is, player i's message space is

$$M_i = \mathcal{F}_{F,\delta} \times \Theta_i \times \Theta \times ... \times \Theta = M_i^{-1} \times M_i^0 \times M_i^1 \times ... \times M_i^K.$$

By construction, player i's strategy space M_i is finite. Therefore, Γ is a finite mechanism.

The devised mechanism is, roughly speaking, an augmented AM-mechanism with a voting scheme over the elements of $\mathcal{F}_{F,\delta}$, which happens in stage $k = -1$. The voting scheme can be described as follows. Suppose that the designer has designated $f^* \in \mathcal{F}_{F,\delta}$ as the default SCF to be implemented. Players can change f^* into $f \in \mathcal{F}_{F,\delta}$ if all players agree on this change. The selected SCF is used to determine the outcome of the *decision rule* of the mechanism in each stage $k \geq 1$. The mechanism is augmented without loss of the attractive properties of the AM-mechanism. Moreover, as the AM-mechanism, in our implementing mechanism there is no tail–chasing and there are not integer games. Finally, it satisfies the best response property: optimal strategies always exists.

Though the constructed mechanism is a simultaneous mechanism, it can be useful to think of it as a sequential mechanism with $K + 2$ stages, where players make simultaneous announcements in each stage.

Suppose that the default SCF f^* is to be virtually implemented by an arbitrarily small $\varepsilon > 0$. Then, the outcome function selects a lottery over the following three components:

Dictator rule: With probability $\frac{\varepsilon}{n}$, player i is selected as a *dictator*. Based on her announcement at the stage $k = 0$, her best outcome from a predetermined set of outcomes is selected.

Audit rule: With probability $\frac{\varepsilon^2}{n}$, player i is *audited* for consistency. To conduct this audit, the designer considers all announcements made by the players from stage $k = 1$ to stage K, and compares them with the message profile, m^0, reported by the players at stage $k = 0$. Player i is punished by selecting $\underline{a}(i, m^0)$ if she is the first one to announce a type profile different from m^0. Otherwise, she is rewarded by selecting $\bar{a}(i, m^0)$.

Decision rule: With probability $\frac{1-\varepsilon-\varepsilon^2}{K}$, at each stage $k \geq 1$, the outcome is determined as follows:

- If all players make exactly the same announcement, θ', then the selected lottery is $\hat{f}(\theta', f^*)$, which is arbitrarily close to $f^*(\theta')$.
- If all but player i make exactly the same announcement, θ', then the outcome function selects the lottery $\hat{f}_i(\theta', f^*)$, which is arbitrarily close to $f^*(\theta')$, where,

for a small number $\alpha > 0$, $\hat{f}$ and $\hat{f}_i$ are defined by:

$$\begin{aligned}\hat{f}(\theta', f^*) &= (1 - n\alpha) f^*(\theta') + \alpha \sum_{j \in N} \bar{a}(j, \theta') \\ \hat{f}_i(\theta', f^*) &= (1 - n\alpha) f^*(\theta') + \alpha \sum_{j \in N \setminus \{i\}} \bar{a}(j, \theta') + \alpha \underline{a}(i, \theta').\end{aligned}$$

- In all other cases, an arbitrary lottery y is selected by the mechanism.

An important feature of the AM-mechanism is that if every player reports her true type θ_i in her $k = 0$ announcement, and everyone reports the true type profile θ in each stage $k \geq 1$, then $f^*(\theta)$ is implemented with probability $1 - \varepsilon - \varepsilon^2$, where $\varepsilon > 0$ is an arbitrarily small parameter chosen by the designer. Another important feature is that truthful reporting is the uniquely rationalizable strategy for each player i. This feature is due to the following two main insights.

First, each player i's strictly dominant strategy is to truthfully report her type θ_i in stage $k = 0$. The possibility that each player is nominated as a dictator is key to an understanding of this insight. To understand this, suppose that player i plays any strategy $\hat{m}_i$ such that $\hat{m}_i^0 = \hat{\theta}_i \neq \theta_i$. By changing $\hat{m}_i$ into m_i, where $m_i^0 = \theta_i$ and $m_i^k = \hat{m}_i^k$ for each $k \geq 1$, player i has a utility gain of $u_i(f_i(\theta_i), \theta_i) - u_i\left(f_i\left(\hat{\theta}_i\right), \theta_i\right) > 0$, by Lemma 1, when she is chosen as the dictator. To provide player i with incentives to truthfully report in stage $k = 0$, this utility gain must be greater than the maximal utility gain from lying. Since the gain from lying comes only from the auditing component of the mechanism, with a probability that depends on ε, the designer provides incentives to truthfully report to player i by choosing ε appropriately.[13]

The second insight is that the audit component of the mechanism, as well as the appropriate choice of K, provides players with incentives to be truthful in each stage $k \geq 1$. To understand this, recall that by the above discussion, everyone is truthful in stage $k = 0$. Let θ be the true profile, so that $m^0 = \theta$. Fix $k = 1$ and any player i. Suppose that player i plays the strategy $\hat{m}_i$ such that $\hat{m}_i^1 \neq \theta = m^0$ and that every other player j plays m_j.

Let us suppose that player i is not the only player who makes a $k = 1$ announcement that is inconsistent with $m^0 = \theta$. By changing $\hat{m}_i$ into m_i such that $m_i^1 = \theta$ and $m_i^k = \hat{m}_i^k$ for each $k > 1$, player i has a utility gain of $u_i(\bar{a}(i, (\theta_i, \theta_{-i})), \theta_i) - u_i(\underline{a}(i, (\theta_i, \theta_{-i})), \theta_i) > 0$ when she is audited—by the domain assumption. When some other player is audited, truthtelling by player i does not affect the outcome of the mechanism. However, a truthful report by player i may cause herself a utility loss in the decision component of the mechanism when stage $k = 1$ is selected by the designer. Given that this loss can happen with probability $\frac{1-\varepsilon-\varepsilon^2}{K}$, the designer can make this loss arbitrarily small by choosing K appropriately.

[13] To understand why lying can be profitable, let us consider a case where everyone else is truthful in all stages. In this case, a lie of player i induces punishments for other players in the auditing component, which may be beneficial to her.

Let us suppose that player i is the only player who makes a $k=1$ announcement that is inconsistent with $m^0=\theta$. By changing $\hat{m}_i$ into m_i such that $m_i^1=\theta$ and $m_i^k=\hat{m}_i^k$ for each $k>1$, player i does not suffer any utility loss when she is audited as $u_i(\bar{a}(i,(\theta_i,\theta_{-i})),\theta_i)-u_i(\underline{a}(i,(\theta_i,\theta_{-i})),\theta_i)>0$. When some other player is audited, truthtelling by player i can not harm her. The reason is that player i can only harm other players in the auditing phase by truthtelling—player i may only have a utility gain by the domain assumption. Player i has incentives to change $\hat{m}_i$ into m_i because when stage $k=1$ is selected by the designer, given that all players but player i make the same $k=1$ announcement, player i's utility is $u_i(f_i^*(\theta),\theta_i)$, which, by the domain assumption, is strictly lower than the utility she obtains under truthtelling; that is, $u_i(f_i^*(\theta),\theta_i)<u_i(f^*(\theta),\theta_i)$.

Since our goal is to virtually implement F, by implementing $f^*\in\mathcal{F}_{F,\delta}$ we have achieved our goal partially. To virtually implement F, as mentioned earlier, we augment the AM-mechanism with a voting rule over $\mathcal{F}_{F,\delta}$, which happens in stage $k=-1$.

Recall that in our augmented mechanism, players can coordinate on any $f\in\mathcal{F}_{F,\delta}$ by reaching an unanimous consensus on f. If they fail to do so, then f^* is implemented. Note that the "elected" SCF is used when stage $k\geq 1$ is chosen to determine the outcome of the decision component of the mechanism.

An attractive feature of the voting game is that any unanimous agreement on $f\in\mathcal{F}_{F,\delta}$ forms a *strict* Nash equilibrium. This feature allows us to create multiple strict Nash equilibria in the augmented mechanism. Indeed, we show that the strategy profile in which every player i plays $m_i=(f,\theta_i,\theta,...,\theta)$ forms a strict Nash equilibrium. Moreover, we also show that player i's rationalizable strategies are of the form $m_i=(\cdot,\theta_i,\theta,...,\theta)$ (see Lemma 4 below). Thus, even though the players fail to coordinate on one strict Nash equilibrium (that is, one SCF), or even though they are playing some mixed equilibrium, the realized outcome will be ε-close to an $f\in\mathcal{F}_{F,\delta}$.

It is worth emphasizing that our mechanism does not rely on any tail chasing construction. The reason is that the constructed mechanism is a finite mechanism.

4 Proof of the result

Suppose that $n\geq 3$ and that Assumption 1 holds. We proceed by breaking the proof in two cases. Case 1 provides the proof for the complete information model $\bar{\mathcal{T}}$. Case 2 extends the implementation result for $\bar{\mathcal{T}}$ to any "nearby" model of incomplete information $\mathcal{T}$ such that $\mathcal{T}\supseteq\bar{\mathcal{T}}$.

Case 1: The complete information model $\bar{\mathcal{T}}$

Let us define $\Gamma = (M, g)$ as follows.

$$\begin{aligned} M &= \prod_{i \in N} M_i, \\ M_i &= M_i^{-1} \times M_i^0 \times M_i^1 \times ... \times M_i^K, \end{aligned}$$

where the integer K is yet to be specified, and where[14]

$$M_i^{-1} = \mathcal{F}_{F,\delta},\ M_i^0 = \Theta_i \text{ and } M_i^k = \Theta \text{ for all } k \in \{1, ..., K\}.$$

Since Lemma 1 holds and since Θ_i is finite, it follows that a real number $\eta > 0$ exists such that for each $\theta_i \in \Theta_i$, it holds that

$$u_i(f_i(\theta_i), \theta_i) - u_i(f_i(\theta_i'), \theta_i) > \eta$$

for each $\theta_i' \in \Theta_i \backslash \{\theta_i\}$.

By using the lotteries specified by Assumption 1, let us define the function $\xi : N \times M \to Y$ by:

$$\xi(i, m) = \begin{cases} \underline{a}(i, m^0) & \text{if for some } k \in \{1, ..., K\},\ m_j^h = m^0 \text{ for all } h = 1, ..., k-1 \\ & \text{and all } j \in N \backslash \{i\}, \text{ and } m_i^k \neq m^0; \\ \bar{a}(i, m^0) & \text{otherwise.} \end{cases}$$

Since Assumption 1 holds, for every $f \in \mathcal{F}_{F,\delta}$, a nearby SCF $\hat{f} : \Theta \to Y$ exists and, for each $i \in N$, a nonempty single-valued function $\hat{f}_i : \Theta \to Y$ exists such that

$$u_i\left(\hat{f}(\theta_i), \theta_i\right) - u_i\left(\hat{f}_i(\theta), \theta_i\right) > 0 \tag{1}$$

for all $\theta \in \Theta$. To see it, for a small number $\alpha > 0$, let us define $\hat{f}$ and $\hat{f}_i$ by:

$$\begin{aligned} \hat{f}(\theta, f) &= (1 - n\alpha) f(\theta) + \alpha \sum_{j \in N} \bar{a}(j, \theta) \\ \hat{f}_i(\theta, f) &= (1 - n\alpha) f(\theta) + \alpha \sum_{j \in N \backslash \{i\}} \bar{a}(j, \theta) + \alpha \underline{a}(i, \theta). \end{aligned}$$

By definition of $\hat{f}$ and $\hat{f}_i$, it can be checked that (1) holds. Moreover, by definition, it also follows that f, $\hat{f}$ and $\hat{f}_i$ are all ε-close to each other.

For every $k \in \{1, ..., K\}$, define the function $\rho^k : M^k \times M^{-1} \to Y$ as follows.

Rule 1 (Universal agreement): If $\left|\left\{i \in N | m_i^{-1} = f\right\}\right| = n$ for some $f \in \mathcal{F}_{F,\delta}$, then:

(a) If $m_i^k = \theta$ for all $i \in N$, then $\rho^k\left(m^k, m^{-1}\right) = \hat{f}(\theta, f)$.

[14] $\mathcal{F}_{F,\delta}$ has been defined in the previous section.

(b) For all $i \in N$, if $m_j^k = \theta$ for all $j \in N \setminus \{i\}$ and $m_i^k \neq \theta$, then $\rho^k\left(m^k, m^{-1}\right) = \hat{f}_i(\theta, f)$.

(c) Otherwise, $\rho^k\left(m^k, m^{-1}\right) = y$ for some $y \in Y$.

Rule 2 ($n-1$ agreement): If $\left|\left\{i \in N | m_i^{-1} = f\right\}\right| = n-1$ for some $f \in \mathcal{F}_{F,\delta}$ and $m_\ell \neq f$ for some $\ell \in N$, then:

(a) If $m_i^k = \theta$ for all $i \in N$, then $\rho^k\left(m^k, m^{-1}\right) = \hat{f}_\ell(\theta, f)$.

(b) For all $i \in N$, if $m_j^k = \theta$ for all $j \in N \setminus \{i\}$ and $m_i^k \neq \theta$, then $\rho^k\left(m^k, m^{-1}\right) = \hat{f}_i(\theta, f)$.

(c) Otherwise, $\rho^k\left(m^k, m^{-1}\right) = y$ for some $y \in Y$.

Rule 3 (Disagreement): Otherwise, for some $f^* \in \mathcal{F}_{F,\delta}$,

(a) If $m_i^k = \theta$ for all $i \in N$, then $\rho^k\left(m^k, m^{-1}\right) = \hat{f}(\theta, f^*)$.

(b) For all $i \in N$, if $m_j^k = \theta$ for all $j \in N \setminus \{i\}$ and $m_i^k \neq \theta$, then $\rho^k\left(m^k, m^{-1}\right) = \hat{f}_i(\theta, f^*)$.

(c) Otherwise, $\rho^k\left(m^k, m^{-1}\right) = y$ for some $y \in Y$.

Let $\varepsilon > 0$ be an arbitrary small number such that $1 - \varepsilon - \varepsilon^2 > 0$. The outcome function $g : M \to Y$ is defined, for all $m \in M$, by:

$$g(m) = \frac{\varepsilon}{n}\sum_{i \in N} f_i\left(m_i^0\right) + \frac{\varepsilon^2}{n}\sum_{i \in N}\xi(i, m) + \frac{1 - \varepsilon - \varepsilon^2}{K}\sum_{k=1}^{K}\rho^k\left(m^k, m^{-1}\right). \tag{2}$$

For each player i's type θ_i, let

$$E_i(\theta_i) = \max_{m \in M}\left(\sum_{j \in N}\left|u_i\left(\xi(j, m), \theta_i\right)\right|\right).$$

We fix $\varepsilon > 0$ such that for all $i \in N$,

$$\eta > 2\varepsilon E_i(\theta_i) \tag{3}$$

for all $\theta_i \in \Theta_i$.

For each $i \in N$ and each $\theta \in \Theta$, define

$$B_i(\theta) = u_i\left(\bar{a}(i, \theta), \theta_i\right) - u_i\left(\underline{a}(i, \theta), \theta_i\right)$$

and for each $k = 1, ..., K$, define

$$D_i(\theta) = \max_{\left(m^k, m^{-1}\right) \in M^k \times M^{-1}}\left[u_i\left(\rho^k\left(m^k, m^{-1}\right), \theta_i\right) - u_i\left(\rho^k\left(\left(m_{-i}^k, \bar{m}_i^k\right), m^{-1}\right), \theta_i\right)\right],$$

where $\bar{m}_i^k = \theta$.

By Assumption 1, it follows that for all $i \in N$, $B_i(\theta) > 0$ for all $\theta \in \Theta$. Thus, an integer $K > 0$ exists such that for all $i \in N$,

$$K \frac{\varepsilon^2}{n} B_i(\theta) > \left(1 - \varepsilon - \varepsilon^2\right) D_i(\theta) \tag{4}$$

for all $\theta \in \Theta$.

Fix any SCR F. Clealy, Γ is a finite mechanism. To prove that Γ virtually implements F in Nash equilibria (resp., in rationalizable strategies), we need the following lemmata for any $t^\theta \in \bar{\mathcal{T}}$.

Lemma 2 For all $m \in M$, $m \in \mathrm{R}\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right) \implies m_i = (\cdot, \theta_i, \theta, ..., \theta)$ for all $i \in N$.

Proof. The proof of this statement is based on the proof of Abreu and Matsushima (1992a). We report it for the sake of completeness. Take any $m \in \mathrm{R}\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$. We proceed by cases.

Case 0: $m_i^0 \neq \theta_i$ for some $i \in N$.

Suppose that $m_i^0 \neq \theta_i$ for some $i \in N$. Let $\bar{m}_i \in M_i$ be such that $\bar{m}_i^0 = \theta_i$ and $\bar{m}_i^k = m_i^k$ for each $k \in \{-1, 1, ..., K\}$. Fix any $m_{-i} \in M_{-i}$. To save space, let $\bar{m} = (\bar{m}_i, m_{-i})$ and $m = (m_i, m_{-i})$. Note that, by construction, $\bar{m}^k$ and m^k fall into the same rule for each $k \in \{1, ..., K\}$. By definition of g, Lemma 1 and the fact that $\bar{m}_i^{-1} = m_i^{-1}$, we have that

$$\begin{aligned} u_i(g(\bar{m}), \theta_i) - u_i(g(m), \theta_i) &= \frac{\varepsilon}{n}\left[u_i\left(f_i\left(\bar{m}_i^0\right), \theta_i\right) - u_i\left(f_i\left(m_i^0\right), \theta_i\right)\right] \\ &\quad + \frac{\varepsilon^2}{n} \sum_{j \in N} \left[u_i(\xi(j, \bar{m}), \theta_i) - u_i(\xi(j, m), \theta_i)\right] \\ &> \frac{\varepsilon}{n}\left(\eta - 2\varepsilon E_i(\theta_i)\right) \\ &> 0, \end{aligned} \tag{5}$$

where the last inequality uses (3). This means that $\bar{m}_i$ strictly dominates m_i, and so $m_i \notin \mathrm{R}_i\left(t_i^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$, which is a contradiction. Then, $m_i \in \mathrm{R}_i\left(t_i^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$ is such that $m_i^0 = \theta_i$. It follows that if $m \in \mathrm{R}\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$, then $m_i^0 = \theta_i$ for all $i \in N$.

For each $h \in \{0, 1, ..., K\}$, let $P(h)$ be the statement "If $m \in \mathrm{R}^h\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$, then for all $i \in N$, it holds that

$$m_i^0 = \theta_i \text{ and } m_i^\ell = \theta \text{ for each } \ell = 1, ..., h."$$

By the above arguments, we know that $P(0)$ holds. Assume that $P(h-1)$ holds for $0 \leq h - 1 < K$. We show that $P(h)$ holds.

Assume, to the contrary, that $P(h)$ is false, that is, $m_i^h \neq \theta$ for some $i \in N$. Recall that $m \in \mathrm{R}\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$. Let $\bar{m}_i \in M_i$ be such that $\bar{m}_i^h = \theta$ and $\bar{m}_i^k = m_i^k$ for all $k \in \{-1, 0, ...K\} \setminus \{h\}$.

Take any $m_{-i} \in \mathrm{R}_{-i}^{h-1}\left(t_{-i}^{\theta}, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$. To save space, let $\bar{m} = (\bar{m}_i, m_{-i})$ and $m = (m_i, m_{-i})$. Again, note that, by construction, $\bar{m}^k$ and m^k fall into the same rule for each $k \in \{1, ..., K\} \setminus \{h\}$. We proceed according to the following two cases.

Case 1: There exists an player $j \neq i$ such that $m_j^h \neq \theta$

By definition of g, Assumption 1 and the fact that $\bar{m}_i^{-1} = m_i^{-1}$, we have that

$$\begin{aligned} u_i\left(g\left(\bar{m}\right), \theta_i\right) - u_i\left(g\left(m\right), \theta_i\right) &= \frac{\varepsilon^2}{n}\left[u_i\left(\bar{a}\left(i, \bar{m}^0\right), \theta_i\right) - u_i\left(\underline{a}\left(i, m^0\right), \theta_i\right)\right] \\ &+ \frac{1-\varepsilon-\varepsilon^2}{K}\left[u_i\left(\rho^h\left(\bar{m}^h, \bar{m}^{-1}\right), \theta_i\right) - u_i\left(\rho^h\left(m^h, m^{-1}\right), \theta_i\right)\right] \\ &= \frac{\varepsilon^2}{n} B_i\left(\theta\right) - \frac{1-\varepsilon-\varepsilon^2}{K}\left[u_i\left(\rho\left(m^h, m^{-1}\right), \theta_i\right) - u_i\left(\rho^h\left(\bar{m}^h, \bar{m}^{-1}\right), \theta_i\right)\right] \\ &\geq \frac{\varepsilon^2}{n} B_i\left(\theta\right) - \frac{1-\varepsilon-\varepsilon^2}{K} D_i\left(\theta\right) \\ &> 0, \end{aligned} \quad (6)$$

where the last inequality uses (4). Since the choice of $m_{-i} \in \mathrm{R}_{-i}^{h-1}\left(t_{-i}^{\theta}, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$ is arbitrary, it follows that $m_i \notin \mathrm{R}_i^h\left(t_i^{\theta}, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$, which is a contradiction.

Case 2: For all $j \neq i$, $m_j^h = \theta$

We proceed according to whether $\xi\left(i, \bar{m}\right) = \xi\left(i, m\right)$ or not.

Suppose that $\xi\left(i, \bar{m}\right) = \xi\left(i, m\right)$. It simplifies the argument, and causes no loss of generality, to assume that $m_j^{-1} = f = \bar{m}_i^{-1}$ for all $j \neq i$. Then, $\bar{m}$ and m fall into Rule 1. Then, by definition of g and the fact that $m_j^h = \theta$ for all $j \neq i$, we have that

$$\begin{aligned} u_i\left(g\left(\bar{m}\right), \theta_i\right) - u_i\left(g\left(m\right), \theta_i\right) &= \frac{1-\varepsilon-\varepsilon^2}{K}\left[u_i\left(\rho^h\left(\bar{m}^h, \bar{m}^{-1}\right), \theta_i\right) - u_i\left(\rho^h\left(m^h, m^{-1}\right), \theta_i\right)\right] \\ &= \frac{1-\varepsilon-\varepsilon^2}{K}\left[u_i\left(\hat{f}\left(\theta, f\right), \theta_i\right) - u_i\left(\hat{f}_i\left(\theta, f\right), \theta_i\right)\right] \\ &> 0, \end{aligned} \quad (7)$$

where the last inequality uses (1). Suppose that $\xi\left(i, \bar{m}\right) \neq \xi\left(i, m\right)$. Then, by applying the same reasoning used in Case 1, we have that $u_i\left(g\left(\bar{m}\right), \theta_i\right) - u_i\left(g\left(m\right), \theta_i\right) > 0$. In either case, since the choice of $m_{-i} \in \mathrm{R}_{-i}^{h-1}\left(t_{-i}^{\theta}, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$ is arbitrary, it follows that $m_i \notin \mathrm{R}_i^h\left(t_i^{\theta}, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$, which is a contradiction.

By the principle of mathematical induction, it follows that $P\left(h\right)$ holds for each $h \in \{1, ..., K\}$. Since $\mathrm{R}\left(t^{\theta}, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right) = \cap_{k=1}^{K} \mathrm{R}^k\left(t^{\theta}, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$, it follows that if $m \in \mathrm{R}\left(t^{\theta}, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$, then $m_i = (\cdot, \theta_i, \theta, ..., \theta)$ for all $i \in N$. ■

Lemma 3 For all $f \in \mathcal{F}_{F,\delta}$ and all $m \in M$, if $m_i = (f, \theta_i, \theta, ..., \theta)$ for all $i \in N$, then $m \in \text{NE}\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$ such that for all $i \in N$, $u_i\left(g\left(m\right), \theta_i\right) > u_i\left(g\left(m'_i, m_{-i}\right), \theta_i\right)$ for all $m'_i \in M_i \backslash \{m_i\}$.

Proof. Take any $m \in M$ such that $m_i = (f, \theta_i, \theta, ..., \theta)$ for all $i \in N$. We show that m is a (pure) strict Nash equilibrium of $\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$. To this end, and without loss of generality, we can just focus on the set of rationalizable strategies. Fix any $i \in N$ and any $\bar{m}_i \in \text{R}_i\left(t_i^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right) \backslash \{m_i\}$. Lemma 2 implies that $\bar{m}_i = (\cdot, \theta_i, \theta, ..., \theta)$. Since $\bar{m}_i \neq m_i$, it must be the case that $\bar{m}_i^{-1} = \bar{f} \neq f$ for some $\bar{f} \in \mathcal{F}_{F,\delta}$. It follows from the definition of g that

$$\begin{aligned} u_i\left(g\left(m_i, m_{-i}\right), \theta_i\right) - u_i\left(g\left(\bar{m}_i, m_{-i}\right), \theta_i\right) &= \left(1 - \varepsilon - \varepsilon^2\right)\left[u_i\left(\hat{f}\left(\theta, f\right), \theta_i\right) - u_i\left(\hat{f}_i\left(\theta, f\right), \theta_i\right)\right] \\ &> 0, \end{aligned}$$

where the last inequality follows from (1). Since the choice of both player i and $\bar{m}_i \in \text{R}_i\left(t_i^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right) \backslash \{m_i\}$ is arbitrary, it follows that for each $i \in N$, $u_i\left(g\left(m\right), \theta_i\right) > u_i\left(g\left(\bar{m}_i, m_{-i}\right), \theta_i\right)$ for all $\bar{m}_i \in \text{R}_i\left(t_i^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right) \backslash \{m_i\}$. This implies that m is a strict Nash equilibrium of $\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$. ■

Lemma 4 For all $m \in M$, $m \in \text{R}\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right) \iff m_i = (\cdot, \theta_i, \theta, ..., \theta)$ for all $i \in N$.

Proof. Lemma 2 implies the "only if" of the statement. Thus, let us show the "if" part. Take any $m \in M$ such that $m_i = (\cdot, \theta_i, \theta, ..., \theta)$ for all $i \in N$. We show that $m_i \in \text{R}_i\left(t_i^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$ for each $i \in N$. Fix any $i \in N$ and suppose that $m_i^{-1} = f$. Let $\bar{m}_{-j} \in M_{-j}$ be such that $m_j = (f, \theta_j, \theta, ..., \theta)$ for all $j \in N \backslash \{i\}$. Lemma 3 implies that $(m_i, \bar{m}_{-j}) \in \text{NE}\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$. It follows that $m_i \in \text{R}_i\left(t_i^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$. Since the choice of i is arbitrary, we conclude that $m \in \text{R}\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$. ■

To show that Γ virtually implements F in Nash equilibria (resp., in rationalizable strategies) we need first to define the correspondence $H : \Theta \twoheadrightarrow Y$ that is ε-close to F. To this end, for each $x \in A\left(\delta, F\left(\theta\right)\right)$ and each $i \in N$, define $\gamma\left(x, \theta\right)$ and $\gamma_i\left(x, \theta\right)$ as follows.

$$\begin{aligned} \gamma\left(x, \theta\right) &= \frac{\varepsilon}{n}\sum_{i \in N} f_i\left(\theta_i\right) + \frac{\varepsilon^2}{n}\sum_{i \in N}\bar{a}\left(i, \theta\right) + \frac{1 - \varepsilon - \varepsilon^2}{K}\left[\left(1 - n\alpha\right)x + \alpha\sum_{j \in N}\bar{a}\left(j, \theta\right)\right] \\ \gamma_i\left(x, \theta\right) &= \frac{\varepsilon}{n}\sum_{i \in N} f_i\left(\theta_i\right) + \frac{\varepsilon^2}{n}\sum_{i \in N}\bar{a}\left(i, \theta\right) + \frac{1 - \varepsilon - \varepsilon^2}{K}\left[\left(1 - n\alpha\right)x + \alpha\sum_{j \in N \backslash \{i\}}\bar{a}\left(j, \theta\right) + \alpha\underline{a}\left(i, \theta\right)\right]. \end{aligned}$$

By definition and the fact that $\alpha > 0$ is a small number, it follows that for all $x \in A\left(\delta, F\left(\theta\right)\right)$, $d\left(x, \gamma\left(x, \theta\right)\right) \leq \varepsilon$ and that $d\left(x, \gamma_i\left(x, \theta\right)\right) \leq \varepsilon$ for all $i \in N$. Thus, H can be defined as follows. For all $\theta \in \Theta$,

$$H\left(\theta\right) = \left\{\left\{\gamma\left(x, \theta\right), \left\{\gamma_i\left(x, \theta\right)\right\}_{i \in N}\right\} | x \in A\left(\delta, F\left(\theta\right)\right)\right\}. \tag{8}$$

For all $\theta \in \Theta$, consider the surjection $\rho_H^\theta : F(\theta) \twoheadrightarrow H(\theta)$ defined by $\rho_H^\theta(x) = \left\{\gamma(x,\theta), \{\gamma_i(x,\theta)\}_{i\in N}\right\}$, for each $x \in A(\delta, F(\theta))$. By using the mapping ρ_H^θ and by choosing $\delta > 0$ small enough, it can be checked that H is ε-close to F. To complete the proof, we need the following useful result.

Lemma 5 For all $m \in \mathrm{R}\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$, $g(m) \in \rho_H^\theta(x)$ for some $x \in F(\theta)$.

Proof. Take any $m \in \mathrm{R}\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$. Lemma 2 implies that $m_i = (\cdot, \theta_i, \theta, ..., \theta)$ for all $i \in N$. This means that either Rule 1(a), Rule 2(a) or Rule 3(a) applies to m. If Rule 1(a) applies, then $g(m) = \gamma(f(\theta), \theta) \in \rho_H^\theta(f(\theta))$ and $m_i^{-1} = f$ for all $i \in N$. If Rule 3(a) applies, then $g(m) = \gamma(f^*(\theta), \theta) \in \rho_H^\theta(f^*(\theta))$. Finally, suppose that Rule 2(a) applies. Then, $m_j^{-1} = f$ for all $j \in N \setminus \{i\}$ and $m_i^{-1} \neq f$, for some $i \in N$. By definition, $g(m) = \gamma_i(f(\theta), \theta) \in \rho_H^\theta(f(\theta))$. ■

We now show that, for each $t^\theta \in \bar{T}$, it holds that:

(a) for all $x \in F(\theta)$, there exists $m \in \mathrm{NE}\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$ (resp., $m \in \mathrm{R}\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$) such that $g(m) \in \rho_H^\theta(x)$.

(b) for each $\sigma \in \mathrm{NE}\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$ (resp., $\sigma \in \mathrm{R}\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$), $\bigcup_{m \in \mathrm{Supp}(\sigma)} g(m) \subseteq H(\theta)$.

To show part (a), suppose that $x \in F(\theta)$. Let m be such that $m_i = (f, \theta_i, \theta, ..., \theta) \in M_i$ for each $i \in N$ and $f(\theta) = x$. By Lemma 3, $m \in \mathrm{NE}\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$ (and so $m \in \mathrm{R}\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$). Moreover, by definition of g, $g(m) = \gamma(x, \theta)$. Since the choice of $x \in F(\theta)$ is arbitrary, it follows that part (a) is satisfied.

Let us now show part (b). For every $m \in \mathrm{R}\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$, Lemma 5 implies that $g(m) \in \rho_H^\theta(x)$ for some $x \in F(\theta)$. By definition of H, it follows that $g(m) \in H(\theta)$. This completes the proof of part (b) for the case of rationalizable strategies. Finally, take any $\sigma \in \mathrm{NE}\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$ and any $m \in$Supp(σ). By definition of $\mathrm{R}\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$, we have that Supp$(\sigma) \subseteq \mathrm{R}\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$, and so $m \in \mathrm{R}\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$. Again, Lemma 5 implies that $g(m) \in \rho_H^\theta(x)$ for some $x \in F(\theta)$, and so $g(m) \in H(\theta)$, by definition of H.

Case 2: Extension to any "nearby" model

Fix any model $\mathcal{T} \supseteq \bar{\mathcal{T}}$, any $\theta \in \Theta$ and any $\varepsilon > 0$ such that Γ virtually implements F in strict Nash equilibria (resp., in rationalizable strategies). Let us show that Definition 1 is met.

Let us first show that part (a) of Definition 1 is met. Fix any $x \in F(\theta)$. Lemma 3 shows that the strategy profile $m^*(\theta)$, where $m_i^*(\theta) = (f, \theta_i, \theta, ..., \theta)$ for all $i \in N$ and where $f(\theta) = x$, is a strict Nash equilibrium of $\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$. From the above, we know that $g(m^*(\theta)) \in \rho_H^\theta(x)$. We will use the existence of $m^*(\theta)$ to show the existence of strategy profile satisfying part (a) of Definition 1. We distinguish two cases.

Case A: Bayesian Nash equilibria

The fact that $m^*(\theta)$ is a strict Nash equilibrium at $\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$ implies that there exists $\underline{n}$ such that for each $n \geq \underline{n}$ and each $m_i \neq m_i^*$, it holds that

$$(1-\epsilon)\left[u_i\left(g\left(m_i^*(\theta), m_{-i}^*(\theta)\right), \theta_i\right) - u_i\left(g\left(m_i, m_{-i}^*(\theta)\right), \theta_i\right)\right] > \epsilon A \tag{9}$$

where $\epsilon > 0$ and

$$A \equiv \max_{i,m,m',\hat{\theta}} \left| u_i\left(g(m), \hat{\theta}_i\right) - u_i\left(g(m'), \hat{\theta}_i\right)\right|.$$

The following two conditions can be satisfied by further decreasing the real number ϵ:

C1 For all $i \in N$ and all $\theta' \neq \theta$, the $(d_i^{\text{uw}}, \epsilon)$-ball around $\left(\theta, t_i^\theta\right)$, denoted by $\left\{\left(\theta, t_i^\theta\right)\right\}^\epsilon$, is disjoined from $\left\{\left(\theta', t_i^{\theta'}\right)\right\}^\epsilon$.

C2 For all $t_i \in T_i$, $d_i^{\text{uw}}\left(t_i, t_i^\theta\right) < \epsilon \implies \kappa[t_i]\left(\left\{\left(\theta, t_{-i}^\theta\right)\right\}^\epsilon\right) > 1-\epsilon$,

where $\left\{\left(\theta, t_{-i}^\theta\right)\right\}^\epsilon$ denotes the $\left(d_{-i}^{\text{uw}}, \epsilon\right)$-ball around $\left(\theta, t_{-i}^\theta\right)$.

Let us consider the agent normal form of the game $\mathcal{U}(\Gamma, \mathcal{T})$ with the restriction that every type t_i of player i in $(d_i^{\text{uw}}, \epsilon)$-ball around $\left(\theta, t_i^\theta\right)$ plays $m_i^*(\theta)$. Let us denote this game by $\widetilde{\mathcal{U}}(\Gamma, \mathcal{T})$. Since $\mathcal{T}$ is countable and M is finite, a standard fixed-point argument implies that $\text{BNE}\left(\widetilde{\mathcal{U}}(\Gamma, \mathcal{T})\right) \neq \emptyset$. Let $\sigma \in \text{BNE}\left(\widetilde{\mathcal{U}}(\Gamma, \mathcal{T})\right)$. For any sequence $\{t_n\}$ in T with $d^{\text{uw}}\left(t_n, t^\theta\right) \to 0$, there exists $\underline{n}$ such that $\sigma(t_n) = m^*(\theta)$ for all $n \geq \underline{n}$, by construction of $\widetilde{\mathcal{U}}(\Gamma, \mathcal{T})$ and by C2 and (9).

Let us now consider the original game $\mathcal{U}(\Gamma, \mathcal{T})$. Observe that for every type t_i of player i in $(d_i^{\text{uw}}, \epsilon)$-ball around $\left(\theta, t_i^\theta\right)$, the unique best response for t_i is to play $\sigma(t_i) = m_i^*(\theta)$. This is due to C2 and (9). Moreover, for every t_i of player i that is not in $(d_i^{\text{uw}}, \epsilon)$-ball around $\left(\theta, t_i^\theta\right)$, $\sigma(t_i)$ is a best response to σ_{-i}. This is because $\sigma \in \text{BNE}\left(\widetilde{\mathcal{U}}(\Gamma, \mathcal{T})\right)$. Therefore, $\sigma \in \text{BNE}(\mathcal{U}(\Gamma, \mathcal{T}))$. Finally, observe that, by construction, $\sigma_{|t^\theta} \in \text{NE}\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$. This shows that part (a) of Definition 1 holds for BNE.

Case B: Interim correlated rationalizability

Let us now show that part (a) of Definition 1 holds for ICR. To this end, we need to show that there exists a large enough $\underline{n}$ such that for all $n \geq \underline{n}$,

$$\text{R}(t_n, \mathcal{U}(\Gamma, \mathcal{T})) \subseteq \text{R}\left(t^\theta, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right). \tag{10}$$

Since $\mathcal{F}_{F,\delta}$ is finite, (10) follows directly from Lemma 1 of Dekel et al. (2006).

To complete the proof of Theorem 1, we are left to show part (b) of Definition 1.

Let us first show it for the set of ICR strategies. Take any $\sigma(t_n) \in \text{R}(t_n, \mathcal{U}(\Gamma, \mathcal{T}))$. Note that $\text{Supp}(\sigma(t_n)) \subseteq \text{R}(t_n, \mathcal{U}(\Gamma, \mathcal{T}))$. Recall that $\varepsilon > 0$ is such that Γ virtually implements F in strict Nash equilibria (resp., in rationalizable strategies). Since there exists $\underline{n}$ such that for all $n \geq \underline{n}$, $m_i = (\cdot, \theta_i, \theta, ..., \theta)$ for all $i \in N$ and all $m \in \text{Supp}(\sigma(t_n))$. Lemma 4 implies that Supp$(\sigma(t_n)) \subseteq \mathrm{R}\left(t^{\theta}, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$. Fix any $m \in$Supp$(\sigma(t_n))$. Lemma 5 implies that $g(m) \in \rho_H^{\theta}(x)$ for some $x \in F(\theta)$. By definition of H, it follows that $g(m) \in H(\theta)$.

Finally, take any $\sigma \in \mathrm{BNE}(\mathcal{U}(\Gamma, \mathcal{T}))$ and any $m \in$Supp$(\sigma(t_n))$. By the arguments in the preceding paragraph, we have that $m \in \mathrm{R}\left(t^{\theta}, \mathcal{U}\left(\Gamma, \bar{\mathcal{T}}\right)\right)$. Thus, $g(m) \in H(\theta)$, by Lemma 5 and definition of H.

5 Related literature

This paper contributes to two new strands on the literature on mechanism design and implementation theory.

Robust mechanism design

The first strand is that of robust mechanism design, pioneered by Bergemann and Morris (2005). In this literature, many works, such as Bergemann and Morris (2005, 2009a,b, 2011a) and Chung and Ely (2007), adopt a "global" approach by studying settings in which the designer does not have any idea of the information structure prevailing among players. The designer's objective is to implement an SCF on all models (that is, all information structures) that he deems possible. By contrast, Chung and Ely (2003), Oury and Tercieux (2012), Jehiel et al. (2012), Aghion et al. (2012), Chen et al. (2018a) follow a "local" approach. They consider settings in which the designer knows the initial model under study, but not perfectly (for example, our designer knows that the state is common knowledge among players). The designer wishes to implement an SCF not only at all types of the initial model but also at all types "close" to initial types. An intermediate approach is followed by Ollár and Penta (2017), who study settings in which the designer rules out some possible beliefs among players. Given that we follow a local approach to robust mechanism design, in what follows, we discuss only papers on local robustness.

Oury and Tercieux (2012) relate partial implementation of an SCF on the neighborhood of a type space to its full implementation. More precisely, under a domain assumption (of costly messages), they show that partial continuous implementation in product topology is equivalent to full implementation in rationalizable strategies. In a recent paper, Chen et al. (2018a) study partial continuous implementation of SCFs under the uniform–weak topology. Specifically, by focussing on direct revelation mechanisms, they show that partial continuous implementation is tightly connected to partial implementation in strict Nash equilibrium in the initial model. In contrast to these contributions, we focus on continuous full implementation in rationalizable strategies under the uniform–weak topology. Moreover, we allow for multi-valued SCRs.[15]

[15] Jehiel et al. (2012) also study partial implementation of SCFs via mechanisms that are incentive compatible only for beliefs that lie in a neighborhood of some benchmark beliefs (which may be derived from some common prior as usually assumed in the mechanism design literature).

Indeed, this paper belongs more to the literature on the robustness of full implementation under complete information, pioneered by Chung and Ely (2003). They show, by assuming an arbitrarily small uncertainty about the state and strict preferences among players, that only Maskin monotonic SCFs can be implemented in undominated Nash equilibria, by requiring that any Bayesian Nash equilibria of the implementing mechanism must be arbitrarily close to the SCF.[16] This is in contrast with the permissive result of Palfrey and Srivastava (1991), who show that any SCF can be fully implemented in undominated Nash equilibria. Aghion et al. (2012) study full implementation in subgame-perfect equilibria under similar perturbations and uncover a similar negative result, in the sense that whenever an extensive form mechanism implements a non-Maskin monotonic SCF, there exists an undesirable equilibrium in some nearby environment. In contrast to these authors, we focus on implementation via finite mechanisms which are robust to strategic uncertainty.

Robustness to strategic uncertainty

A game theoretic solution that is weaker than Nash equilibrium is that of rationalizability, pioneered by Bernheim (1984) and Pearce (1984). This solution, which builds solely on the assumption of common knowledge of rationality, asks "What might a rational player do?".[17] By allowing players' beliefs to be correlated, Brandenburger and Dekel (1987) propose a weaker version of rationalizability, which is fully characterized by the set of strategies that survive the thought process of iterative deletion of never best responses. This is the definition of rationalizable strategies used in the literature on implementation in rationalizable strategies. We adopt this definition as well.

Bergemann et al. (2011b) study the implementation of SCFs under complete information in rationalizable strategies. They show that a necessary and almost sufficient condition for implementation is strict Maskin monotonicity*, which is stronger than Maskin monotonicity (1999). In recent studies, Kunimoto and Serrano (2019) and Jain (2019) examined implementation of correspondences in rationalizable strategies. All characterization results on implementation in rationalizable strategies are far from complete and, moreover, are derived by devising implementing mechanisms that rely on questionable tail-chasing procedures to eliminate unwanted best responses, such as integer or modulo games (Jackson, 1992). Our result show that any SCR is virtually implementable in rationalizable strategies by a finite mechanism.

Chen et al. (2018b) have recently studied the implementation of SCFs under complete information in rationalizable strategies by finite mechanisms in an environment with transfers

[16] Maskin monotonicity is a remarkably strong invariance condition, which is necessary for the full implementation in Nash equilibria (Maskin, 1999). For exaple, Maskin monotonicity is precisely the property that the SCFs typically studied in contract theory do not satisfy.

[17] The type of rationality captured by the notion of Nash equilibrium can be described by two rationality assumptions: (1) common knowledge of rationality, that is, the best responses of players to their beliefs and (2) rational expectations, that is, players' beliefs are correct.

and lotteries, as in Abreu and Matsushima (1992a, 1994). Similar to our mechanism, their mechanism borrows ideas from the mechanism devised by Abreu and Matsushima (1992a, 1994). Chen et al.'s (2018b) contribution is an attempt to unify the classical approach of exact implementation, which usually relies on mechanisms having questionable features, with the approach of virtual implementation, which relies on well-behaved mechanisms.

6 Concluding remarks

We characterized the class of SCRs that are virtually implementable not only when the state is common knowledge among players but also when it is "close" to common knowledge. The closeness between types is measured in terms of the uniform–weak topology. This result is achieved by devising a finite mechanism. We also show that virtual implementation in strict Nash equilibria, and in rationalizable strategies, imply continuous virtual implementation when the state is common knowledge among players. As in Chung and Ely (2003), in the true model, strict incentives are used to attain continuous implementation in uniform–weak topology.

The results of the paper provide a theoretical benchmark for continuous virtual implementation of SCRs. We obtain our results under the assumption that the baseline model is that of complete information. This assumption may not be satisfied in certain situations. Abreu and Matsushima (1992b) have generalized the AM–mechanism to Bayesian environments. They show that any SCF that can be virtually Bayesian Nash implemented in these environments must satisfy a measurability condition, namely AM-measurability. To characterize the class of social choice sets which are virtually implementable in both Bayesian Nash and ICR strategies, we believe that a construction like the one presented in this paper would be useful. We believe that such a construction will hinge on the identification of an appropriate variant of the AM-measurability condition. We leave this subject for future research.

Finally, let us remark that the AM-mechanism has been the focus of attention in several recent strands of implementation theory, such as robust virtual implementation (Bergemann and Morris, 2009a; Muller, 2016), level-k implementation (Serrano et al., 2018), and implementation with verification (Matsushima, 2019). Indeed, these papers provide constructive proofs that rely, directly or indirectly, on the AM-mechanism. We believe that our construction may play an important role in these strands of the literature when the objective of the designer is represented by an SCR.

References

[1] Aghion P, Fudenberg D, Holden R, Kunimoto T, Tercieux O. Subgame–Perfect Implementation Under Information Perturbations. The Quarterly Journal of Economics 127

(2012) 1843–1881.

[2] Abreu D, Matsushima H. Virtual implementation in iteratively undominated strategies: complete information. Econometrica 60 (1992a) 993–1008.

[3] Abreu D, Matsushima H. Virtual implementation in iteratively undominated strategies II: Incomplete information. Princeton University working paper (1992b).

[4] Abreu D, Matsushima H. Exact Implementation. Journal of Economic Theory 64 (1994) 1–19.

[5] Abreu D, Sen A. Virtual implementation in Nash equilibrium. Econometrica 59 (1991) 997–1021.

[6] Bergemann D, Morris S. Robust Mechanism Design. Econometrica 73 (2005) 1521–1534.

[7] Bergemann D, Morris S. Robust virtual implementation. Theoretical Economics 4 (2009a) 45–88.

[8] Bergemann D, Morris S. Robust Implementation in Direct Mechanisms. Review of Economic Studies 76 (2009b) 1175–1204.

[9] Bergemann D, Morris S. Robust Implementation in General Mechanisms. Games and Economic Behavior 71 (2011a) 261–281.

[10] Bergemann D, Morris S, Tercieux O. Rationalizable implementation. Journal of Economic Theory 146 (2011b) 1253–1274.

[11] Bernheim D. Rationalizable Strategic Behavior. Econometrica 52 (1984) 1007–1028.

[12] Bertsekas DP, Shreve SE. Stochastic Optimal Control: The Discrete-Time Case. Academic Press, 1978.

[13] Brandenburger A, Dekel E. Rationalizability and Correlated Equilibria. Econometrica 55 (1987) 1391–1402.

[14] Chen Y-C, Di Tillio A, Faingold E, Xiong S. Uniform topologies on types. Theoretical Economics, 5 (2010) 445–478.

[15] Chen Y-C, Frank MM, Pai MM. Truthful continuous implementation. Working Paper (2018a).

[16] Chen YC, Kunimoto T, Sun Y, Xiong S. Rationalizable Implementation in Finite Mechanisms. Working Paper (2018b).

[17] Chung K-S, Ely JC. Implementation with Near-Complete Information. Econometrica 71 (2003) 857–871.

[18] Chung K-S, Ely JC. Foundations of dominant-strategy mechanisms. Review of Economic Studies 74 (2007) 447–476.

[19] Dekel E, Fudenberg D, Morris S. Topologies on types. Theoretical Economics 1 (2006) 275–309

[20] Dekel E, Fudenberg D, Morris S. Interim correlated rationalizability. Theoretical Economics, 2 (2007) 15–40.

[21] Harsanyi JC. Games with Incomplete Information Played by "Bayesian" Players, I-III. Part I. The Basic Model. Management Science 14 (1967) 159–182.

[22] Jackson MO. Implementation in undominated strategies: A look at bounded mechanisms. The Review of Economic Studies, 59 (1992) 757–775.

[23] Jackson MO, Palfrey TR, Srivastava S. Undominated Nash implementation in bounded mechanisms. Games and Economic Behavior 6 (1994) 474–501.

[24] Jain R. Rationalizable implementation of social choice correspondences. Games and Economic Behavior 127 (2021) 47-66

[25] Jain R, Lombardi M.Continuous Virtual Implementation: Complete Information (April 30, 2019). Available at SSRN: https://ssrn.com/abstract=3369470.

[26] Jehiel P, Meyer-ter-Vehn M, Moldovanu B. Locally robust implementation and its limits. Journal of Economic Theory 147 (2012) 2439–2452.

[27] Kartik N, Tercieux O, Holden R. Simple mechanisms and preferences forhonesty. Games and Economic Behavior 83 (2014) 284–290.

[28] Kunimoto T, Serrano R. A New Necessary Condition for Implementation in Iteratively Undominated Strategies. Journal of Economic Theory 146 (2011) 2583–2595.

[29] Kunimoto T, Serrano R. T. Rationalizable Implementation of Correspondences. Mathematics of Operations Research 44 (2019) 1326–1344.

[30] Mertens J-F, Zamir. Formulation of Bayesian analysis for games with incomplete information. International Journal of Game Theory 14 (1985) 1–29.

[31] Maskin E. Nash equilibrium and welfare optimality. Rev Econ Studies 66 (1999) 23–38.

[32] Matsushima H. Implementation without expected utility: Ex-post verifiability. Social Choice and Welfare 53 (2019) 575–585.

[33] Mezzetti C, Renou L. Implementation in Mixed Nash Equilibrium. J Econom Theory 147 (2012) 2357–2375.

[34] Monderer D, Samet D. Approximating Common Knowledge with Common Beliefs, Games and Economic Behavior 1 (1989) 170–190.

[35] Muller C. Robust virtual implementation under common strong belief in rationality. Journal of Economic Theory 162 (2016) 407–450.

[36] Ollár M, Penta A, Full Implementation and Belief Restrictions, Am Econ Rev 107 (2017) 2243–2277.

[37] Oury M, Tercieux O. Continuous implementation. Econometrica 80 (2012) 1605–1637.

[38] Palfrey T, Srivastava S, Nash-implementation using undominated strategies, Econometrica 59 (1991) 479–501.

[39] Pearce D. Rationalizable Strategic Behavior and the Problem of Perfection. Econometrica 52 (1984) 1029–1050

[40] Rubinstein A. The Electronic Mail Game: Strategic Behavior Under "Almost Common Knowledge", American Economic Review 79 (1989) 385–391.

[41] Serrano R, de Clippel G, Saran R. Level-k Mechanism Design. The Review of Economic Studies 86 (2018) 1207–1227.

[42] Taylor AE. General Theory of Functions and Integration. Dover Publications, New York, 1986.

[43] Thomson, W. Concepts of implementation. The Japanese Economic Review 47(2) (1996) 133-143.

[44] Weinstein J, Yildiz M. A structure theorem for rationalizability with application to robust predictions of refinements. Econometrica 75 (2007) 365–400.

[45] Weinstein J, Yildiz M. Sensitivity of equilibrium behavior to higher-order beliefs in nice games. Games Econ Behav 72 (2011) 288–300

[46] Weinstein J, Yildiz M. Interim correlated rationalizability in infinite games. Journal of Mathematical Economics 72 (2017) 82–87.