Review of Economic Design
https://doi.org/10.1007/s10058-022-00304-8

ORIGINAL PAPER

Robust implementation in sequential information design under supermodular payoffs and objective

Hiroto Sato[1]

Received: 26 October 2021 / Accepted: 25 June 2022

Abstract
This paper studies *sequential information design* (Doval and Ely in Econometrica 88:2575–2608, 2020) in which a designer can construct the extensive form along with the information structure. In this framework, I investigate robust implementations against adversarial equilibrium selection, when players and the designer have a supermodular payoff function with dominant states and an outside option. The main results show that the optimal partially implementable outcome is fully implementable in sequential information design, which essentially coincides with the optimal partially implementable outcome in static information design. For economic applications such as global game of regime change, this paper proposes a way to robustly achieve the desired outcome in static information design by providing the extensive form and the information structure.

Keywords Information design · Sequential information design · Supermodular game

JEL Classifiction C72 · D82 · D83

1 Introduction

Information design explores how the designer can implement an outcome through manipulation of players' beliefs about the state and their higher-order beliefs, which influences their behavior.[1] This paper studies a more general framework of *sequential*

[1] Information design is a multiple-player generalization of Bayesian persuasion pioneered by Kamenica and Gentzkow (2011). Bergemann and Morris (2019) and Kamenica (2019) survey literature on Bayesian persuasion and information design.

✉ Hiroto Sato
hirotosato@g.ecc.u-tokyo.ac.jp

1 Graduate School of Economics, University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Published online: 12 July 2022

information design (introduced by Doval and Ely (2020)[2]) in which the designer can construct the extensive form along with the information structure. In other words, the designer can release information not only about the state of the world, but also about past players' moves.

In static information design, several notions of implementation are studied such as *partial implementation*, *smallest equilibrium implementation*, and *full implementation*. Partial implementation requires that the target outcome is induced by some preferred Bayes Nash equilibrium (BNE) under the information structure. Bergemann and Morris (2016) and Taneva (2019) characterize the partially implementable outcomes by using the concept of *obedience* condition. Recently, a growing body of literature focuses on "robust" implementation against *adversarial equilibrium selection*. In supermodular game, Mathevet et al. (2020); Hoshino (2021); Sandmann (2020), and Morris et al. (2020) characterize the smallest equilibrium implementable outcomes. Under the monotone designer's payoff, the smallest equilibrium is equal to the worst equilibrium under adversarial selection.[3] In addition, they analyze the more demanding notion, full implementation, which requires the target outcome to be induced by all equilibria. Since the outcome is uniquely implemented, full implementation is also robust against adversarial equilibrium selection.

To the best of my knowledge, this is the first study on "robust-approach" for implementation in sequential information design. It shows that the optimal partially implementable outcome is fully implementable in sequential information design in binary action supermodular games with a monotone and supermodular designer's objective function under dominant state assumption and outside option. This result indicates that the designer can uniquely, and hence robustly, implement the desired outcome without sacrificing payoff by using sequential information design. Notice that, it is not generally possible in static information design as in Morris et al. (2020). Moreover, in this setting, this optimal outcome is approximately equivalent to the optimal partially implementable outcome in static information design. Thus, even though the designer has no gain under partial implementation, under full implementation she does better when sequential information design is feasible instead of only static information design.

These results provide new insight into the benefit of applying sequential information design in line with Doval and Ely (2020). They characterize the set of partially implementable outcomes in sequential information design, which is a superset of the partially implementable outcomes in static information design. In contrast, this paper shows how the designer can fully implement her optimal outcome using sequential information design. The construction behind the main result manipulates strategic uncertainty using the extensive form of the game.

Additionally, the main results have economic applications in the global game literature (Morris and Shin 2003), which includes many settings such as currency attacks (Morris and Shin 1998), bank runs (Goldstein and Pauzner 2005), debt pricing (Mor-

[2] Doval and Ely (2020) is a generalization of Salcedo (2017) to incomplete information games. Furthermore, Makris and Renou (2021) characterize the (partially) implementable outcomes on the multi-stage base games.

[3] For economic application, Goldstein and Huang (2016); Inostroza and Pavan (2022), and Li et al. (2019) study information design with adversarial equilibrium selection in a regime change game.

ris and Shin 2004), investment game (Carlsson and van Damme 1993), team work production (Moriya and Yamashita 2020), and policy changes (De Mesquita 2010; Edmond 2013; Chen and Suen 2017), among others. In these settings, the designer often wants to maximize the probability of players taking a particular action, e.g. not attacking the currency, not withdrawing their deposits, investing in a project, and exerting effort. The main result of this paper proposes a way to achieve this goal robustly by providing the extensive form and the information structure that fully implements the optimal outcome distribution under partial implementation.

1.1 Motivating example

To illustrate the ideas in this paper, I use an example of the investment game similar to Mathevet et al. (2020) and Morris et al. (2020). There are two players 1 and 2, where each player chooses action 1 and 0. Moreover, there are two states θ_1 and θ_0 which occur with the same probability. The payoffs are summarized in Fig. 1 where player 1 (resp. 2) is the row (resp. column):
In this base game, action 1 (resp. 0) is dominant at state θ_1 (resp. θ_0) for each player. Under the prior, action 0 is strictly dominant for each player. Payoffs are supermodular: a player's payoff to action 1 is 1 larger if the other player takes action 1 at each state. Additionally, payoffs are asymmetric: player 1 obtains a payoff 1 larger than player 2 from action 1 at state θ_1. In these setting, the designer wants to maximize the expected number of action 1 taken by players regardless of state.

In static information design (where the designer can only choose the information structure), by the characterization in Bergemann and Morris (2016), the optimal equilibrium distribution over actions and states (call outcome) under partial implementation is shown in Fig. 2:

However, by the characterization of Morris et al. (2020), the above outcome is not fully implementable, and thus the designer faces a trade-off between the robustness of implementation and the payoff in static information design. The optimal fully

Fig. 1 Base game

θ_1	1	0
1	$3, 2$	$2, 0$
0	$0, 1$	$0, 0$

θ_0	1	0
1	$-4, -4$	$-5, 0$
0	$0, -5$	$0, 0$

Fig. 2 Optimal partially implementable outcome in static and sequential information design

θ_1	1	0
1	$\frac{1}{2}$	0
0	0	0

θ_0	1	0
1	$\frac{1}{4}$	$\frac{1}{10}$
0	0	$\frac{3}{20}$

Fig. 3 Optimal fully implementable outcome in static information design

θ_1	1	0
1	$\frac{1}{2}$	0
0	0	0

θ_0	1	0
1	$\frac{2}{9}$	0
0	0	$\frac{5}{18}$

θ_1 $t_1\backslash t_2$	1	2	3	$\cdots$	∞
1	ε	$\eta\left(\frac{1}{2}-\varepsilon\right)$			
2			$\eta(1-\eta)\left(\frac{1}{2}-\varepsilon\right)$		
3				$\ddots$	
$\vdots$					
∞					

θ_0 $t_1\backslash t_2$	1	2	3	$\cdots$	∞
1		$\eta\left(\frac{1}{9}-\delta\right)$			
2	$\eta\frac{1}{9}$		$\eta(1-\eta)\left(\frac{1}{9}-\delta\right)$		
3		$\eta(1-\eta)\frac{1}{9}$		$\ddots$	
$\vdots$			$\ddots$		
∞					$\eta\left(\frac{5}{18}+\delta\right)$

Fig. 4 Construction of the information structure by Morris et al. (2020)

implementable outcome[4] in static information design is essentially equivalent to the outcome in Fig. 3:
To implement this, Morris et al. (2020) construct the following information structure: where ε, δ, and η are sufficiently small positive numbers such that $\varepsilon \gg \eta$, and $t_1, t_2 \in \{1, 2, \ldots\}$ are players' types. An intuition behind this information structure is as follows: first, type $t_i = 1$ believes that the true state is θ_1 with probability close to 1 as $\eta \approx 0$, and chooses action 1 since it is dominant at state θ_1. Then, given type $t_i = 1$ playing action 1, the optimal action of type $t_j := t_{3-i} = 2$ is action 1 regardless of the actions by other types $t_i \geq 3$.[5] Additionally, given type $t_j = 2$ playing action 1, the optimal action of type $t_i = 3$ is action 1 regardless of the actions by other types $t_j \geq 4$. By this logic, both players of types $t_i, t_j < \infty$ choose action 1 in the unique equilibrium.

In sequential information design (where the designer can construct the extensive form and the information structure), the outcome in Fig. 2 is also the optimal partially implementable outcome as shown by Lemma 1 in Sect. 2.3. Thus, the designer has no gain from sequential information design under partial implementation in this example. Note that this outcome can be induced by the extensive form representing the incomplete information game associated with the optimal information structure in static information design. Since there are no off-path information sets in this extensive

[4] By the result of Morris et al. (2020) (Proposition 6 and Corollary 2), under this designer's objective function, the optimal fully implementable outcome (which is equivalent to the optimal smallest equilibrium implementable outcome) can be obtained by maximizing the objective subject to consistency and (weak) sequential obedience. Moreover, since this base game admits potential satisfying convexity and the designer's objective also satisfies restricted convexity, Theorem 2 of Morris et al. (2020) indicates that the optimal outcome can be derived by maximizing the objective with respect to perfectly coordinated outcomes subject to consistency and weak grand coalitional obedience.

[5] For player $j = 1$ (resp. $j = 2$), action 1 induces the expected payoffs (with positive multiplication) greater than

$$\left(\frac{1}{2}-\varepsilon\right)\times 2+\left(\frac{1}{9}-\delta\right)(-5)+\frac{1}{9}(-4) \text{ and } \left(\frac{1}{2}-\varepsilon\right)\times 2+\left(\frac{1}{9}-\delta\right)(-4)+\frac{1}{9}(-5)$$

respectively, which are strictly positive.

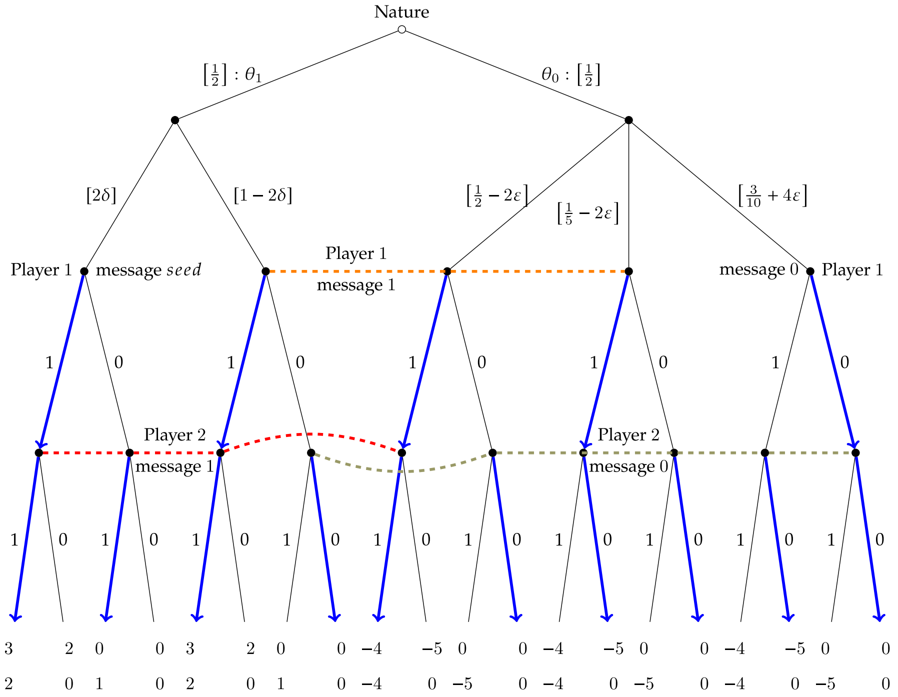

Fig. 5 Construction of the extensive form

form, this outcome is also induced by a self-contained BNE. Thus, there is no loss from restricting what the designer could achieve off the equilibrium path.[6]

My main results will establish that the designer can approximately induce this outcome as full implementation in sequential information design by using the extensive form as shown in Fig. 5 with $\varepsilon > \delta > 0$. Thus, in contrast to static information design, the designer can fully implement the optimal partially implementable outcome in sequential information design. Even though the designer has no gain under partial implementation, under full implementation she does better when sequential information design is feasible instead of only static information design.

In Fig. 5, dashed lines represent players' information sets, and blue arrows represent players' behavior at the equilibrium. Furthermore, the designer's messages and players' actions are labeled with associated branches and information sets where the conditional probability is in square brackets.

The intuition behind the construction of this extensive form is a finite variant of Rubinstein's (1989) contagion through players' information sets: (i) When player 1 receives message *seed*, he knows that the true state is θ_1, and hence optimally chooses action 1. This corresponds to the crazy type in Rubinstein (1989) contagion argument. (ii) When player 2 receives message 1, her optimal choice is action 1 regardless of player 1's move in the remaining information sets. This is because if player 1 who receives message 1 chooses action 1, then the designer tells player 2

[6] I would like to express gratitude to the anonymous referee for pointing out this relationship.

that player 1 does so with probability 1 and the true state is θ_1 with (conditional) probability $\frac{2}{3-4\varepsilon} > \frac{2}{3}$.[7] Otherwise, the designer tells player 2 that the true state is θ_1 with (conditional) probability 1. Thus, in either case, player 2's optimal choice is action 1 when she receives message 1. (iii) When player 1 receives message 1, he believes that the true state is θ_1 with probability $\frac{10-20\delta}{17-20\delta-40\varepsilon} > \frac{10}{17}$ and action 1 is chosen by player 2 who receives message 1. Then, player 1's optimal choice is action 1 regardless of the action of player 2 who receives message 0.[8] Note that, to obfuscate information about player 2's move for player 1, the designer makes the more persuadable player (player 1) move before the less persuadable player (player 2) in this extensive form. (iv) When each player receives message 0, they know the true state is θ_0, and hence optimally choose action 0. Thus, the iterated elimination of strictly dominated strategies leads to the unique (self-contained) BNE.

The above intuition highlights the reason why the designer can fully implement this outcome in sequential information design even though it is not fully implementable in static information design. Precisely, this comes from the difference of the constructions between this extensive form in Fig. 5 and the information structure of Morris et al. (2020) in Fig. 4. In their construction of the information structure (Fig. 4), the designer should provide probability such that, given player 1's type $t_1 = 1$ playing action 1, player 2 with type $t_2 = 2$ has incentive of obeying to choose action 1 regardless of action by player 1's type $t_1 = 3$. However, in the extensive form (Fig. 5), player 2 receives message 1 (corresponding to $t_2 = 2$) if player 1 with message 1 (corresponding to $t_1 = 3$) obeys, and receives message 0 otherwise. Therefore, this variant of contagion in the extensive form involves the true sequentiality of the recommendations. A recommendation depends on information about the players' past moves in addition to payoff-relevant states. Consequently, the designer can fully implement this outcome in sequential information design even though it is not fully implementable in static information design.

Moreover, in contrast to the construction behind Morris et al. (2020) with infinitely many types, the extensive form consists of action recommendations for all players except for the first player, which is similar to the usual information design framework. Additionally, despite the asymmetry of payoffs, the optimal fully implementable outcome does not necessarily satisfy the perfect coordination property which is often discussed in the literature of robust implementation in static information design such as Inostroza and Pavan (2022) and Morris et al. (2020). The rest of this paper is organized as follows. A model and the main results are provided in Sect. 2. In Sect. 3, I discuss the assumptions behind the main results. Finally, the concluding remarks are summarized in Sect. 4. The proof of Lemma 1 and Theorem 1 is presented in Appendix A.

[7] Player 2's expected payoff is: $\frac{2}{3-4\varepsilon} \times 2 + \frac{1-4\varepsilon}{3-4\varepsilon} \times (-4) = \frac{4\varepsilon}{3-4\varepsilon} > 0$.

[8] Note that player 1's expected payoff from action 1 when player 2 who receives message 0 chooses action 0 is: $\frac{10-20\delta}{17-20\delta-40\varepsilon} \times 3 + \frac{5-20\varepsilon}{17-20\delta-40\varepsilon} \times (-4) + \frac{2-20\varepsilon}{17-20\delta-40\varepsilon} \times (-5) = \frac{180\varepsilon-60\delta}{17-20\delta-40\varepsilon} > 0$. By supermodularity, player 1's expected payoff from action 1 is larger than this when player 2 who receives message 0 chooses action 1.

2 Model and results

2.1 Base game

In the *base game* $G = (I, (A_i, u_i)_{i \in I}, \Theta, \mu)$, there are n-players $I = \{1, \dots, n\}$, where each player $i \in I$ takes binary action $a_i \in A_i = \{1, 0\}$ and denotes $A = \prod_i A_i$. As usual, fix player i, A_{-i} represents the set of other players' actions $A_{-i} = \prod_{j \neq i} A_j$. There are two payoff-relevant states $\Theta = \{\theta_1, \theta_0\}$, which are drawn from a probability distribution $\mu \in \Delta(\Theta)$, where $\mu(\theta_1) \in (0, 1)$. Given the action profile $a \in A$ and the state $\theta \in \Theta$, the player i obtains the payoff according to $u_i : A \times \Theta \to \mathbb{R}$.

Throughout this paper, I impose the following assumptions:

Assumption 1 (Supermodular payoffs) For each $i \in I$ and $\theta \in \Theta$,

$$u_i((1, a_{-i}), \theta) - u_i((0, a_{-i}), \theta)$$

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

Assumption 2 (Symmetric dominant state assumption) For each $i \in I$ and $a_{-i} \in A_{-i}$,

$$u_i((1, a_{-i}), \theta_1) > u_i((0, a_{-i}), \theta_1), \quad u_i((1, a_{-i}), \theta_0) < u_i((0, a_{-i}), \theta_0).$$

Assumptions 1 and 2 are typically assumed in the literature of full implementation in information design, such as in Morris et al. (2020) and Mathevet et al. (2020).

Assumption 3 (Dominance of action 0 under the prior) For each $i \in I$ and $a_{-i} \in A_{-i}$,

$$\mu(\theta_1)u_i((1, a_{-i}), \theta_1) + (1 - \mu(\theta_1))u_i((1, a_{-i}), \theta_0) < \mu(\theta_1)u_i((0, a_{-i}), \theta_1) \\ + (1 - \mu(\theta_1))u_i((0, a_{-i}), \theta_0).$$

Assumption 3 implies that action 0 is strictly dominant without any information about the state $\theta \in \Theta$. This is the situation in which the designer's task is the most difficult as seen in Sect. 2.3.

Assumption 4 (Outside option) For each $i \in I$, $u_i((0, a_{-i}), \theta)$ is independent of a_{-i} for each $a_{-i} \in A_{-i}$ and $\theta \in \Theta$.

Although Assumption 4 loses generality, it is acceptable as in Mathevet and Taneva (2020), as most applications including the regime change game of Morris and Shin (2003) and investment game of Carlsson and van Damme (1993) impose this condition. In the static information design, this assumption is vacuous since any players cannot react to the others players' actions. However, in sequential information design, this assumption matters. I discuss this point in Sect. 3.

2.2 Implementation in sequential information design

Following Doval and Ely (2020), this paper examines sequential information design. In this framework, the designer can construct the extensive form and the information structure. Formally, the designer can commit to the canonical extensive forms which are characterized by $\Gamma = (G, \xi, X, T, \pi, (u_i^\Gamma)_{i \in I})$. G is the base game. Nature moves once at the beginning and draws the payoff-relevant states $\theta \in \Theta$ following to the distribution $\mu \in \Delta(\Theta)$. A one-to-one mapping $\xi : I \to I$ is a rule for the players' order of move, where $\xi(i) = j$ represents player i as the j-th mover.[9] $X = Z \cup (\bigcup_j X_j)$ is the union of disjoint sets where Z is the finite set of terminal nodes, and $X_j = X_j^{\theta_1} \cup X_j^{\theta_0}$ is the union of disjoint sets of j-th mover's decision nodes after the realization of state θ_1 and θ_0 for each $j \in I$ (where $X_j^{\theta_1}$ and $X_j^{\theta_0}$ are finite sets). For each j's decision node $x_j \in X_j$, there are two immediate successor branches labeled as 1 and 0. $T = \bigcup_j T_j$ is the union of finite sets where T_j partitions X_j into information sets. A mapping $\pi^\theta \in \Delta(X_1^\theta)$ is the signal structure for the first mover at state $\theta \in \Theta$.[10] Player i's preference over terminal nodes $u_i^\Gamma : Z \to \mathbb{R}$ is consistent with payoff function u_i.[11] Player j's behavioral strategy is defined as a function $\sigma_j : T_j \to \Delta(\{1, 0\})$ for each $j \in I$. $\nu \in \Delta(A \times \Theta)$ is denoted as the outcome induced by the profile of players' strategy $\sigma = (\sigma_j)_{j \in I}$ in the extensive form Γ.[12]

Then, I introduce two notions of implementation: *partial implementation* and *full implementation*.

Definition 1 Outcome $\nu \in \Delta(A \times \Theta)$ is a *partially implementable BNE outcome* if there exists an extensive form Γ such that some BNE σ induces ν in Γ.

Let $a_{\prec i}$ and $a_{\succ i}$ be action profiles of i's leaders and followers. By Salcedo (2017) and Doval and Ely (2020), BNE-partially implementable outcomes in this setting are characterized as follows:

Proposition 1 *Outcome $\nu \in \Delta(A \times \Theta)$ is a partially implementable BNE outcome if and only if it satisfies: for all i-th movers,*

$$\sum_{a_{\succ i}, a_{\prec i}, \theta} \nu((a_i, a_{-i}), \theta)[u_i((a_i, a_{-i}), \theta) - \min_{a'_{\succ i}}\{u_i((a'_i, a'_{\succ i}, a_{\prec i}), \theta)\}] \geq 0 \tag{1}$$

for all $a_i, a'_i \in A_i$.[13]

[9] Therefore, the designer does not mix the player's order of moves. Under Assumption 4, it is without loss to assume it. This is because the designer cannot make a threat of punishment since action 0 works as the outside option for each player.

[10] This construction captures the case in which the designer sends a message to the first mover and to the j-th mover after the $j-1$-th mover's move as this case can be reduced to the choice of one simple lottery corresponding to the compound lottery before the first mover's move.

[11] Precisely, for every terminal node z, there is some profile $(a^z, \theta^z) \in A \times \Theta$ such that $u_i^\Gamma(z) = u_i(a^z, \theta^z)$, where (a^z, θ^z) is the profile of actions and state associated with the terminal node z.

[12] Formally, let $\zeta(x \mid \mu, \pi^\theta, \sigma)$ be the probability that node $x \in X$ is reached given Nature's move μ, the signal structure π^θ, and the profile of players' strategies σ. The outcome ν is induced by σ in Γ if it satisfies $\nu(a, \theta) = \sum_{z \in Z(a,\theta)} \zeta(z \mid \mu, \pi^\theta, \sigma)$, where $Z(a, \theta) = \{z \in Z \mid a^z = a, \theta^z = \theta\}$.

[13] To prove that inequality (1) implies the Doval and Ely 's (2020) condition, the plan is constructed as follows: (i) the obedient path is attached to the equilibrium path and (ii) all followers take punishment action off the path of equilibrium. Then, inequality (1) indicates that this plan satisfies the Doval and Ely 's(2020) condition. For the other direction, take any plan that satisfies the condition of Doval and Ely (2020). Then, incentive conditions are more severe under the plan in which all followers take punishment action after deviation, since it induces the lowest payoffs. Thus, inequality (1) is also satisfied.

In contrast to Doval and Ely (2020), this paper focuses on full implementation.[14]

Definition 2 Outcome $\nu \in \Delta(A \times \Theta)$ is *fully implementable* if for any sufficiently small $\varepsilon > 0$, there exists an extensive form Γ in which all BNE σ induce ν' such that $\|\nu - \nu'\| < \varepsilon$.[15]

In the setting of this paper, the main result holds if "all BNE" is replaced by either "the unique BNE" or "the unique self-contained BNE" in the above definition of full implementation. Note that the latter restricts the designer's ability at off-path information sets since this equilibrium concept requires that the designer (essentially) implements the target outcome without off-path information sets.

2.3 Main results

Let $V : A \times \Theta \to \mathbb{R}$ be the designer's objective function .

Assumption 5 (Monotone and supermodular objective) The designer's objective function V is

1. monotone increasing on A: $V(a, \theta)$ is weakly increasing in $a \in A$ for each $\theta \in \Theta$, and
2. supermodular on A: $V(a \vee a', \theta) + V(a \wedge a', \theta) \geq V(a, \theta) + V(a', \theta)$ for all $a, a' \in A$ and $\theta \in \Theta$.[16]

As in the previous example, Assumption 5 is satisfied when the designer wants to maximize the expected number of action 1 taken by players regardless of the state as in most applications, such as currency attacks, bank runs, liquidity crises, and policy changes.

To investigate the optimal outcome under full implementation in sequential information design, I derive the optimal partially implementable outcome in sequential information design as the upper bound.

Lemma 1 *Under Assumptions 1–5, the optimal partially implementable BNE outcome in sequential information design is equivalent to the optimal Bayes Correlated equilibrium (BCE), that is, the optimal outcome $\nu \in \Delta(A \times \Theta)$ satisfies the obedience conditions;*

$$\sum_{a_{-i},\theta} \nu((a_i, a_{-i}), \theta)[u_i((a_i, a_{-i}), \theta) - u_i((a_i', a_{-i}), \theta)] \geq 0 \tag{2}$$

for all $i \in I$ and $a_i, a_i' \in A_i$.

Proof is in Appendix A.

[14] Under the full implementation, the induced outcome is uniquely implemented, and thus it is robust against adversarial equilibrium selection as in Mathevet et al. (2020), Hoshino (2021), Sandmann (2020), and Morris et al. (2020).

[15] As usual, the sup-norm is considered, that is, $\|\nu - \nu'\| := \sup_{a,\theta} |\nu(a, \theta) - \nu'(a, \theta)|$.

[16] $\vee$ and $\wedge$ represent join and meet, respectively.

Fig. 6 Base game violating assumption 4

θ_1	1	0
1	$3, 2$	$2, 0$
0	$0, 1$	$-1, -1$

θ_0	1	0
1	$-4, -4$	$-5, 0$
0	$0, -5$	$-1, -1$

Note that Bergemann and Morris (2016) show that an outcome is partially implementable if and only if it is a BCE in static information design. By combining this, Lemma 1 shows that under Assumptions 1–5, sequential information design is not helpful when working with partial implementation. This is because, unlike static information design, sequential information design with partial implementation enables the designer to incentivize a punishment by players, but under Assumption 4, action 0 is not influenced by such punishment.[17]

The main theorem of this paper is stated as follows:

Theorem 1 *Under Assumptions 1–5, the optimal partially implementable BNE outcome ν is fully implementable under sequential information design.*

Proof is in Appendix A.

An insight from Theorem 1 is as follows: in general, the designer cannot fully implement the optimal partially implementable outcome in static information design as discussed in the example in Sect. 1.1. By contrast, Theorem 1 shows that in the current setting the designer can fully implement the desired partially implementable outcome without sacrifice in terms of payoff under sequential information design. Moreover, by combining with Lemma 1, even though the designer has no gain under partial implementation, under full implementation she does better when sequential information design is feasible instead of only static information design.

Note that there are two important aspects in the construction of the extensive form. First, in contrast to the construction behind Morris et al. (2020) with infinitely many types, the extensive form consists of action recommendations for all players except for the first player, which is similar to the usual information design framework. Additionally, despite the asymmetry of payoffs, the optimal fully implementable outcome does not necessarily satisfy the perfect coordination property which is often discussed in the literature of robust implementation in static information design such as Inostroza and Pavan (2022) and Morris et al. (2020).

3 Discussion about assumption 4 (outside option)

In this Section, I will discuss the importance of Assumption 4. Without Assumption 4, Lemma 1 does not hold in the following example which slightly modifies the example in Sect. 1.1. Precisely, the payoff of $(0, 0, \theta)$ decreases from 0 to -1 for each player and state θ (Fig. 6).

Suppose that the designer wants to maximize the expected number of action 1 taken by players regardless of state. Then, the optimal partially implementable outcome in static information design is shown in Fig. 7:

[17] This is not true without Assumption 4 as discussed in Sect. 3.

Fig. 7 Optimal partially implementable outcome in static information design

θ_1	1	0
1	$\frac{1}{2}$	0
0	0	0

θ_0	1	0
1	$\frac{1}{4}$	$\frac{1}{10}$
0	0	$\frac{3}{20}$

Fig. 8 Better partially implementable outcome in sequential information design

θ_1	1	0
1	$\frac{1}{2}$	0
0	0	0

θ_0	1	0
1	$\frac{1}{4}$	$\frac{1}{8}$
0	0	$\frac{1}{8}$

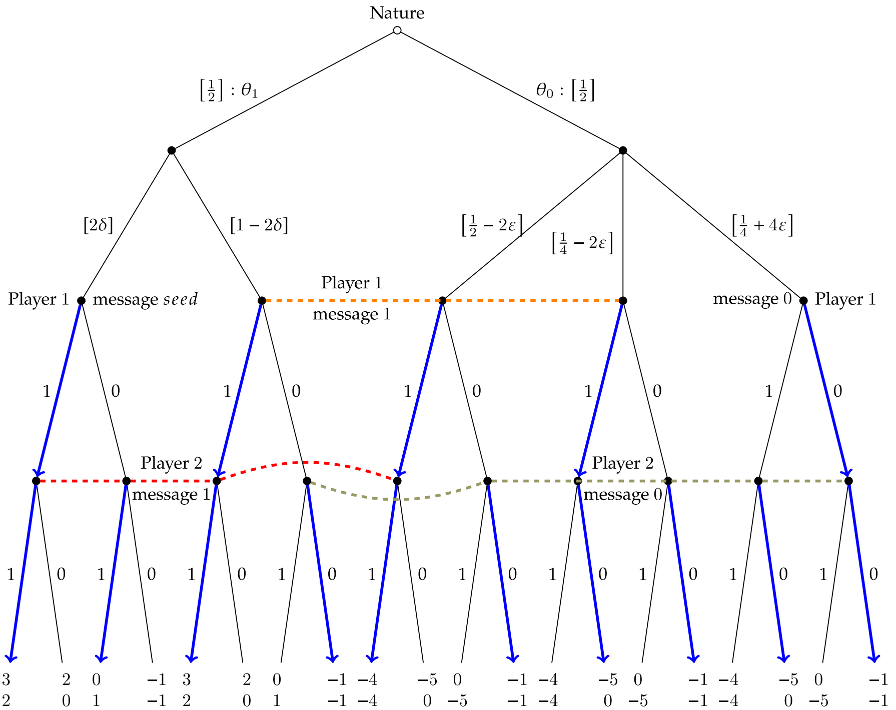

Fig. 9 Construction of the extensive form

However, the designer can fully implement the better outcome in sequential information design as shown in Fig. 8 and construction of the extensive form is shown in Fig. 9 with $\varepsilon > \delta > 0$:

An intuition of the difference between the optimal outcomes in static and sequential information design is the existence of punishments by followers. In this example, if player 1 who received message 1 deviates from action 1 to 0, then player 2 chooses action 0 (punishment) which leads to payoff -1 rather than 0 for player 1. In static information design, it is assumed that the opponent players do not change their actions when the other players deviate from the designer's recommendation. In contrast, in sequential information design, the designer can send a signal that depends on the past movers' actions. Thus, without Assumption 4, the equivalence of optimal outcomes in static and sequential information design does not hold in general.

Moreover, without Assumption 4, it may lose the generality to assume that the order of moves is deterministic in the extensive forms. Thus, it becomes complicated to derive the optimal partially implementable outcome in sequential information design, although it is simplified under Assumption 4 since the result of Arieli and Babichenko (2019) can be used.

4 Conclusion and remarks

This paper contributes to the study of robust implementation against adversarial equilibrium selection in the framework of sequential information design. In this framework, I showed that the optimal partially implementable outcome is fully implementable under binary action supermodular games with a monotone and supermodular designer under dominant state assumption and outside option. Moreover, this outcome is approximately equivalent to the optimal partially implementable outcome in static information design. The characterization of the set of fully implementable outcomes in sequential information design requires further research.

Acknowledgements I would like to express my gratitude to the editor, Moritz Meyer-ter-Vehn, and the two anonymous referees for their constructive suggestions, which helped to significantly improve the paper. Special thanks to Michihiro Kandori and Daisuke Oyama for their advice and helpful comments. I am also thankful to Yuma Noritomo, Masanori Kobayashi, and Ryo Shirakawa. The author acknowledges financial support by JSPS KAKENHI Grant Number 22J10043

Appendix A: Proof of Theorem 1

Step 1: Derivation of the optimal BCE

Under Assumption 4, it is without loss to normalize that for each i, $u_i((0, a_{-i}), \theta) = 0$ for each $a_{-i} \in A_{-i}$ and each $\theta \in \Theta$. The following notations are introduced as in Arieli and Babichenko (2019): Denote $S_{-i} = S \backslash \{i\}$ for a subset $S \subset I$. For simplicity, I denote $u_i(S_{-i}, \theta)$ as player i's payoff from taking action 1 at state θ when only players in S_{-i} choose action 1.

Lemma 2 *(Theorem 5 in* Arieli and Babichenko (2019)) *Under Assumptions 1–5, the optimal BCE can be derived by the following maximum correlation policy.*

Maximum correlation policy

Let $f_i^0(p){:=}p \cdot u_i(I_{-i}, \theta_0) + \frac{\mu(\theta_1)}{1-\mu(\theta_1)}[u_i(I_{-i}, \theta_1)]$ where $p \in [0, 1]$. Define p_0 as follows:

$$p_0{:=}\max \left\{ 0 \leq p \leq 1 \mid f_i^0(p) \geq 0 \ \forall i \in I \right\}$$

Assign probability p_0 to the grand coalition $P^0 = I$ at state θ_0.[18] Then, $P^1 = I_{-j}$ is defined, where $j \in I$ is such that $f_j^0(p_0) = 0$ and $(f_j^0)' < 0$. Let

$$f_i^1(p) := p_0 u_i(I_{-i}, \theta_0) + p u_i(P^1_{-i}, \theta_0) + \frac{\mu(\theta_1)}{1-\mu(\theta_1)}[u_i(I_{-i}, \theta_1)]$$
$$p_1 := \max\left\{0 \le p \le 1 - p_0 \mid f_i^1(p) \ge 0 \ \forall i \in P^1\right\}$$

Assign probability p_1 to the coalition P^1 at state θ_0. If $p_0 + p_1 = 1$, this procedure is stopped. Otherwise, let $P^2 = P^1_{-j}$ where $f_j^1(p_1) = 0$ and $(f_j^1)' < 0$. Recursively, I define

$$f_i^k(p) := p u_i(P^k_{-i}, \theta_0) + \sum_{j=0}^{k-1} p_j u_i(P^j_{-i}, \theta_0) + \frac{\mu(\theta_1)}{1-\mu(\theta_1)}[u_i(I_{-i}, \theta_1)]$$
$$p_k := \max\left\{0 \le p \le 1 - \sum_{j=0}^{k-1} p_j \mid f_i^k(p) \ge 0 \ \forall i \in P^k\right\}$$

Define $P^{k+1} = P^k_{-j}$ where $j \in P^k$ such that $f_j^k(p_k) = 0$ and $(f_j^k)' < 0$.

Let players be indexed according to the order in which they drop out in the maximal correlation policy, i.e. $P_i = \{1, 2, \dots, n-i\}$. Thus, players drop out from the least persuadable player to the most persuadable. This maximum correlation procedure induces the optimal BCE ν as follows:

$$\nu(\mathbf{1}, \theta_1) = \mu(\theta_1)$$
$$\nu((\mathbf{1},\mathbf{0})_{n-k}, \theta_0) = (1-\mu(\theta_1)) p_k \text{ for } 0 \le k \le n$$

where $(\mathbf{1},\mathbf{0})_{n-k}$ represents that each player i such that $i \le n-k$ takes action 1 and the remaining players take action 0 for $0 \le k \le n$, by convention I write $\mathbf{1}$ for $(\mathbf{1},\mathbf{0})_n$ and $\mathbf{0}$ for $(\mathbf{1},\mathbf{0})_0$.

Step 2: Proof of Lemma 1

Proof By Proposition 1, the partially implementable BNE outcome is characterized by the following incentive conditions: for the i-th mover,

$$1: \sum_{a_{\succ i}, a_{\prec i}, \theta} \nu((1, a_{-i}), \theta)[u_i((1, a_{-i}), \theta) - \min_{a'_{\succ i}}\{u_i((0, a'_{\succ i}, a_{\prec i}), \theta)\}] \ge 0 \quad \text{(A.3)}$$
$$0: \sum_{a_{\succ i}, a_{\prec i}, \theta} \nu((0, a_{-i}), \theta)[u_i((0, a_{-i}), \theta) - \min_{a'_{\succ i}}\{u_i((1, a'_{\succ i}, a_{\prec i}), \theta)\}] \ge 0 \quad \text{(A.4)}$$

[18] Note that $p_0 < 1$ by Assumption 3.

First, by Assumption 4, (A.3) is equivalent to

$$\sum_{a_{-i},\theta} \nu((1, a_{-i}), \theta)[u_i((1, a_{-i}), \theta) - u_i((0, a_{-i}), \theta)] \geq 0$$

which is equivalent to the obedience condition (2) for action 1. Thus, the incentive condition for action 1 is equivalent between static and sequential information design under partial implementation. Moreover, as in the optimal BCE derived in Step 1, $\nu(\mathbf{1}, \theta_1) = \mu(\theta_1)$ also holds under the optimal partially implementable BNE outcome in sequential information design under Assumptions 1–5.

Now, (A.4) can be rewritten as

$$\begin{aligned}
&\sum_{a_{\succ i}, a_{\prec i}, \theta} \nu((0, a_{-i}), \theta)[u_i((0, a_{-i}), \theta) - \min_{a'_{\succ i}}\{u_i((1, a'_{\succ i}, a_{\prec i}), \theta)\}] \\
&= \sum_{a_{\succ i}, a_{\prec i}} \nu((0, a_{-i}), \theta_0)[u_i((0, a_{-i}), \theta_0) - \min_{a'_{\succ i}}\{u_i((1, a'_{\succ i}, a_{\prec i}), \theta_0)\}] \\
&\geq 0
\end{aligned}$$

where the equality holds by $\nu(\mathbf{1}, \theta_1) = \mu(\theta_1)$ and the inequality holds since action 0 is dominant in state θ_0 by Assumption 2. Hence, (A.4) is always satisfied under $\nu(\mathbf{1}, \theta_1) = \mu(\theta_1)$.

Therefore, the optimal partially implementable BNE outcome in sequential information design is equivalent to the optimal BCE under Assumptions 1–5. □

Step 3: Construction of the extensive form

In this step, I construct an extensive form that fully implements ν. Taking any sufficiently small ε and δ such that $\varepsilon > \delta > 0$, the desired extensive form can be constructed with the following steps:

1. Nature moves at stage 0, and draws the state θ according to the probability distribution μ.
2. Set player i as the ith mover, that is, $\xi(i) = i$ for all $i \in I$. Thus, the order of moves is from most persuadable player to least persuadable.
3. The designer constructs player 1's decision nodes $X_1 = X_1^{\theta_1} \cup X_1^{\theta_0}$ identified with label,[19]

$$\begin{aligned}
X_1^{\theta_1} &= \{seed, (\mathbf{1}, \theta_1)\} \\
X_1^{\theta_0} &= \{(\mathbf{1}, \theta_0), \ldots, ((\mathbf{1}, \mathbf{0})_{n-k}, \theta_0), \ldots, (\mathbf{0}, \theta_0)\}
\end{aligned}$$

[19] Thus, $|X_1| \leq n + 2$ where equality holds when all nodes are realized with positive probability.

where these nodes realize with (conditional) probability π^{θ} as follows;

$$\pi^{\theta_1}(x_1) = \begin{cases} \frac{\delta}{\mu(\theta_1)} & \text{if } x_1 = seed \\ \frac{\nu(\mathbf{1},\theta_1)-\delta}{\mu(\theta_1)} & \text{if } x_1 = (\mathbf{1}, \theta_1) \end{cases}$$

$$\pi^{\theta_0}(x_1) = \begin{cases} \frac{\nu((\mathbf{1},\mathbf{0})_{n-k},\theta_0)-\varepsilon}{1-\mu(\theta_1)} & \text{if } x_1 = ((\mathbf{1},\mathbf{0})_{n-k}, \theta_0) \text{ and } \nu((\mathbf{1},\mathbf{0})_{n-k}, \theta_0) > 0 \\ & \quad \text{for each } k = 0, 1, 2, \ldots, n-1 \\ \frac{\nu(\mathbf{0},\theta_0)+m\varepsilon}{1-\mu(\theta_1)} & \text{if } x_1 = (\mathbf{0}, \theta_0) \\ 0 & \text{otherwise} \end{cases}$$

where m is the number of nodes in $X_1^{\theta_0} \backslash \{(\mathbf{0}, \theta_0)\}$ with positive probability under π^{θ_0}.

4. The designer divides player 1's decision nodes into three information sets as $T_1 = \{seed, 0, 1\}$ by sending message $seed$ to the node with label $seed$, and message 0 to the node $(\mathbf{0}, \theta_0)$. Otherwise, player 1 receives message 1. Thus, information sets $seed$ and 0 are singleton, and the information set 1 includes the following nodes: $(\mathbf{1}, \theta_1)$, $(\mathbf{1}, \theta_0) \ldots ((\mathbf{1}, \mathbf{0})_{n-k}, \theta_0)$, $\ldots ((\mathbf{1}, \mathbf{0})_{n-1}, \theta_0)$.
5. For player $j = 2, \ldots, n$, the designer divides decision nodes into two information sets as $T_j = \{1, 0\}$ by sending message 1 and 0 conditional on the past movers' actions as follows: player j receives message 1 if (i) nodes are successor of $seed$ and (ii) all past movers choose action 1 when they received message 1 and nodes are successor of $(\mathbf{1}, \theta_1)$, $(\mathbf{1}, \theta_0)$, $\ldots$, $((\mathbf{1}, \mathbf{0})_j, \theta_0)$. Otherwise, the designer sends a message 0.

Step 4: Incentive compatibility and uniqueness of equilibrium

Players' incentives are checked as follows:

1. When player 1 received message $seed$, he believes that the true state is θ_1 with (conditional) probability 1, and hence it is optimal to choose action 1 since it is dominant at state θ_1 by Assumption 2. Thus, player 1's strategies consisting of action 0 at the information set in which he receives message $seed$ are eliminated in the first round.
2. It is seen that player n's optimal choice is action 1 when she received message 1. Now, there are two cases: (i) all players obey and choose action 1 at the information set in which they receive message 1, or (ii) some players deviate from action 1 to action 0 when they receive message 1. In case (i), by obedience condition (2) and $\varepsilon > \delta$, player n's optimal choice is action 1 since she believes that the true state is θ_1 and all past movers choose action 1 with enough probability. In case (ii), player n receives message 1 only if player 1 received message $seed$ which is only at state θ_1 with (conditional) probability 1, and therefore, by Assumption 2, her optimal choice is also action 1 by dominance of action 1 at state θ_1. Therefore, player n's strategies consisting of action 0 at the information set 1 are eliminated in the second round.

3. Given that player 1 and n choose action 1 at the information set *seed* and 1, respectively, I show that player $(n-1)$ optimally chooses action 1 at the information in which he receives message 1. By similar logic in the above, there are two cases: (i) all past movers, players who move before player $(n-1)$, obey and choose action 1 at the information set in which they receive message 1, or (ii) some players deviate from action 1 to action 0 when they receive message 1. In case (i), I show that player $(n-1)$'s optimal choice is action 1 regardless of the actions of player n who receives message 0. If player n who receives message 0 chooses action 0, then player $(n-1)$'s optimal choice is action 1 by obedience condition (2) and $\varepsilon > \delta$. Otherwise, player $(n-1)$'s incentive to choose action 1 is higher than the former case by supermodularity of Assumption 1, and therefore the optimal choice is also action 1. In case (ii), player $(n-1)$ receives message 1 only if player 1 received message *seed* which is only at state θ_1 with (conditional) probability 1, and therefore, by Assumption 2, his optimal choice is also action 1 by dominance of action 1 at state θ_1. Hence, player $(n-1)$'s strategies consisting of action 0 at the information set 1 are eliminated in the third round.
4. Given the above step, by similar logic, player $j = 1, \ldots, n-2$ optimally chooses action 1 when player j receives message 1 regardless of the actions by followers who receive message 0 by supermodularity of Assumption 1, the obedience condition (2), and $\varepsilon > \delta$. Therefore, player j's strategies consisting of action 0 at the information set receiving message 1 are strictly dominated, and hence such strategies are eliminated from the 4th round to the $(n+1)$th round.
5. Given the above argument, all players believe that the true state is θ_0 with (conditional) probability 1 when they receive message 0 which realizes strictly positive probability by $\varepsilon > 0$ and dominance of action 0 under prior μ, and thus they optimally take action 0 since it is dominant at state θ_0 by Assumption 2. Therefore, strategies including action 1 at the information set receiving message 0 are eliminated in the remaining rounds.

Therefore, iterated eliminations of strictly dominated strategies leads to the unique (self-contained) BNE outcome ν' such that $\|\nu - \nu'\| < \varepsilon$ for any sufficiently small $\varepsilon > 0$. □

References

Arieli I, Babichenko Y (2019) Private bayesian persuasion. J Econ Theory 182:185–217

Bergemann D, Morris S (2016) Bayes correlated equilibrium and the comparison of information structures in games. Theor Econ 11:487–522

Bergemann D, Morris S (2019) Information design: a unified perspective. J Econ Lit 57:44–95

Carlsson H, van Eric D (1993) Global games and equilibrium selection. Econ J Econ Soc 61:989–1018

Chen H, Suen W (2017) Aspiring for change: a theory of middle class activism. Econ J 127:1318–1347

Doval L, Ely JC (2020) Sequential information design. Econometrica 88:2575–2608

Edmond C (2013) Information manipulation, coordination, and regime change. Rev Econ Stud 80:1422–1458

Goldstein I, Huang C (2016) Bayesian persuasion in coordination games. Am Econ Rev 106:592–96

Goldstein I, Pauzner A (2005) Demand-deposit contracts and the probability of bank runs. J Financ 60:1293–1327

Hoshino T (2021) Multi-agent persuasion: leveraging strategic uncertainty. Int Econ Rev 63:755

Inostroza N, Pavan A (2022) Adversarial coordination and public information design
Kamenica E (2019) Bayesian persuasion and information design. Ann Rev Econ 11:249–272
Kamenica E, Gentzkow M (2011) Bayesian persuasion. Am Econ Rev 101:2590–2615
Li F, Yangbo S, Mofei Z (2019) Global manipulation by local obfuscation: information design in coordination games. Available at SSRN
Makris M, Ludovic R (2021) Information design in multi-stage games. arXiv preprint arXiv:2102.13482
Mathevet L, Perego J, Taneva I (2020) On information design in games. J Polit Econ 128:1370–1404
Mathevet L, Ina T (2020) Organized information transmission. Available at SSRN 3656555
Mesquita D, Bueno E (2010) Regime change and revolutionary entrepreneurs. Am Political Sci Rev 104:446–466
Moriya F, Yamashita T (2020) Asymmetric-information allocation to avoid coordination failure. J Econ Manag Strateg 29:173–186
Morris, Stephen, Daisuke Oyama, Satoru Takahashi (2020), "Implementation via information design in binary-action supermodular games. Technical report
Morris, Stephen, Hyun Song Shin (1998), "Unique equilibrium in a model of self-fulfilling currency attacks." *American Economic Review*, 587–597
Morris S, Hyun SS (2003) Global games: theory and applications. In: Advances in economics and econometrics: theory and applications, eighth world congress. 1: 56–114, Cambridge University Press
Morris S, Shin HS (2004) Coordination risk and the price of debt. Eur Econ Rev 48:133–153
Rubinstein A (1989) The electronic mail game: strategic behavior under almost common knowledge. Am Econ Rev 79:385–391
Salcedo B (2017) Interdependent choices. Technical report, University of Western Ontario
Sandmann C (2020) Recursive information design. Technical report, Toulouse School of Economics
Taneva I (2019) Information design. Am Econ J: Microecon 11:151–85